

Avances en Informática y Automática Segundo Workshop

Máster Oficial en Sistemas Inteligentes
Departamento de Informática y Automática
Universidad de Salamanca

Ángel Luis Sánchez Lázaro
Carlos García-Figuerola Paniagua
Vivian López Batista
Eladio Sanz García
Sara Rodríguez González
Raquel Sánchez Díaz (Eds.)

Universidad de Salamanca
Diciembre 2008

Avances en Informática y Automática Segundo Workshop

Departamento de Informática y Automática
Facultad de Ciencias
Universidad de Salamanca
Plaza de la Merced S/N,
37008, Salamanca, Spain

Tel.: + 34 923294653

Fax: + 34 923294514

Web: <http://mastersi.usal.es>

Email: secdia@usal.es

Universidad de Salamanca, Salamanca, España

Ninguna parte de esta obra puede ser reproducida, almacenada en ningún soporte o transmitida en diversos medios sin la autorización previa de los autores.

© Ángel Luis Sánchez Lázaro, Carlos García-Figuerola Paniagua, Vivian López Batista, Eladio Sanz García, Sara Rodríguez González, Raquel Sánchez Díaz.
Todos los derechos reservados. Impreso en Salamanca.

Prólogo

El congreso que se celebra dentro del Máster en Sistemas Inteligentes de la Universidad de Salamanca proporciona la oportunidad ideal para que sus estudiantes presenten las principales aportaciones del trabajo de Tesis de Máster y obtengan una realimentación del interés de sus trabajos de investigación. Esta segunda edición, correspondiente al bienio 2008 - 2009, será un encuentro interdisciplinar, supervisado por un grupo de investigadores de reconocido prestigio, pertenecientes a la Universidad de Salamanca.

Entre los principales objetivos del congreso se encuentran:

- Ofrecer a los estudiantes un marco donde exponer sus primeros trabajos de investigación.
- Proporcionar a los participantes un foro donde discutir ideas y encontrar nuevas sugerencias de compañeros, investigadores y otros asistentes a la reunión.
- Permitir a cada estudiante una realimentación de los participantes sobre su trabajo y una orientación sobre las futuras direcciones de investigación.
- Contribuir al desarrollo del espíritu de colaboración en la investigación.

Índice general

Segundo Workshop - Máster en Sistemas Inteligentes

Redes Neuronales aplicadas al Reconocimiento de Huellas Dactilares.	1
<i>Griselda Cobos Estrada y Angélica González Arrieta</i>	
1. Introducción	1
2. Conceptos Teóricos	2
2.1. Huellas Dactilares	2
2.2. Redes Neuronales	3
3. Redes Neuronales aplicadas a la detección de Huellas Dactilares.	3
3.1. Diseño de los datos	4
4. Pruebas y Resultados.	6
5. Conclusiones	8
Recuperación de la Información Y Técnicas de Social Bookmarking	11
<i>Beatriz Fortínez and Carlos G. Figuerola</i>	
1. Introducción	11
2. Esquema Clásico de la Recuperación de la Información	12
3. Tendencias y Paradigmas en la Recuperación de la información	14
4. Social Bookmarking	15
5. Papel de las Folksonomías	16
6. Caso de Estudio, Delicious	17
7. Aportaciones del Documento	18
Modeling, Estimation and Optimal Control Design of a Biological Wastewater Treatment Process	22
<i>Hicham El BAHJA and Pastora Vega</i>	
1. Introduction	22
2. Process modeling	24
3. Estimation	26
3.1. Extended Kalman filter	26
3.2. Asymptotic estimator	28
4. Control	29
4.1. One step predictive control method	30
4.2. Predictive control	30
5. Simulation and results	31
6. Conclusion	32
MAIM: Multiagent Architecture Incidents Management	35
<i>Rosa María Martín Moreno Juan Manuel Corchado Rodríguez</i>	
1. Introducción	35
2. Gestión de Incidencias	37

2.1. Mejores prácticas para la gestión de incidencias	37
3. Metodologías de sistemas multiagentes	38
3.1. GORMAS	38
4. MAIM: Multiagent architecture for Incident Management	39
4.1. Sistema de penalizaciones/recompensas	43
5. Conclusiones	44
Sistema Bioinspirado de Reconocimiento Fonético en Entornos Ruidosos	46
<i>Javier Martínez Elicegui</i>	
1. Introducción	46
2. Fisiología del Habla	47
3. Fisiología del oído y su simulación con un modelo oído computacional	48
3.1. Oído externo	48
3.2. Oído Medio	49
3.3. Oído Interno: la cóclea	50
3.4. Membrana basilar	50
3.5. Órgano de Corti	51
3.6. Modelo Oído Artificial	52
4. Fases de un Sistema de Reconocimiento de Habla	52
4.1. Extracción características	53
4.2. Modelos Acústicos	54
5. Conclusiones	54
<i>ViRe-YouTube: Visualizando los Resultados de Búsquedas de YouTube</i>	56
<i>Carlos Andrés Mera Banguero</i>	
1. Introducción	56
2. Estado del arte	57
2.1. Visualización de Información	57
2.2. Previsualización de Vídeos	59
3. Trabajos Relacionados	60
4. <i>ViRe-YouTube</i>	61
4.1. Un Gráfico de Dispersión para analizar los Vídeos Recuperados	62
4.2. Búsqueda Exploratoria de Vídeos a través de Grafos	63
4.3. Filtros para las Visualizaciones de <i>ViRe-YouTube</i>	65
4.4. Métodos de Previsualización de Vídeos	65
5. Conclusiones	67
6. Trabajos Futuros	68
Métodos del Análisis de Redes Sociales Aplicados a la Recolección de un Crawler	71
<i>Loretto Peirano Cordano y José Luis Alfonso Berrocal</i>	
1. Introducción	71
2. Análisis de Redes Sociales	72
2.1. Teoría de Redes	72
2.2. Trabajando con Redes	76
3. Crawler	77

4. Análisis de Datos	78
5. Conclusiones	80
Stereo-MAS: Multi-Agent System for Image Stereo Processing	84
<i>Sara Rodríguez and Juan M. Corchado</i>	
1. Introducción	84
2. Estado del arte	86
3. Análisis de imágenes: Fases y Técnicas	88
4. Propuesta	93
5. Resultados y Conclusiones	96
Arquitectura MVOSMA: Mundos Virtuales a través de Sistemas Multiagente basados en Organizaciones	99
<i>Jesús Ángel Román Gallego, Juan Manuel Corchado Rodríguez</i>	
1. Introducción	99
2. Conceptos Previos	100
2.1. Mundos Virtuales y Web 3D	100
2.2. Agentes y Sistemas Multiagente	101
2.3. Las Organizaciones	102
3. Arquitectura MVOSMA	103
4. Conclusiones y Trabajo Futuro	107
Sistema de Inteligencia ambiental mediante visión estereoscópica y representación 3D en tiempo real para la localización de personas en entornos de dependencia	109
<i>Manuel P. Rubio, Juan M. Corchado</i>	
1. Introducción	109
2. Revisión del problema	110
3. Estereoscopia	111
3.1. Principios básicos	111
3.2. Condicionantes técnicos del sistema estereoscópico elegido	112
3.3. Justificación de la tecnología adoptada	113
4. Reconstrucción estereoscópica: fases y técnicas	114
4.1. Reconocimiento de postura	115
4.2. Representación 3D	117
5. Propuesta	118
6. Conclusiones y trabajo futuro	119
Librería Simulink de Sistemas HVAC	122
<i>Héctor Salazar y Eladio Sanz</i>	
1. Introducción	122
2. Sistema HVAC	123
3. Operación del Sistema	123
4. Modelación del Sistema	123
4.1. Recinto	124
4.2. Unidad de Intercambio aire-agua (UTA)	125
4.3. Unidad de Producción (Caldera)	125

5. Variables del Sistema	125
6. Sistema de Ecuaciones	126
7. Librería Simulink <i>ISTemperature</i>	127
8. Resultados	128
9. Conclusiones	129
Simulaciones con EJS y Matlab/Simulink	133
<i>Raquel Sánchez Díaz, Eladio Sanz, and Belén Pérez-Lancho</i>	
1. Introducción	133
2. Objetivos	134
3. Herramientas utilizadas	135
4. Situación actual de la investigación en este campo.....	136
4.1. Sistemas de simulación basados en <i>web</i>	137
4.2. Laboratorios virtuales y remotos	137
5. Resultados	138
5.1. Sistema real	139
5.2. Simulaciones para su ejecución en local.....	139
5.3. SBW utilizando un Matlab remoto	142
6. Conclusiones	143
6.1. Líneas de trabajo futuro.....	144
Análisis Visual de Redes de Co-autoría basado en Mapas	
Auto-Organizados: aplicación a publicaciones Biomédicas	146
<i>Mary Carmen Trejo Ávila</i>	
1. Introducción	147
1.1. Problema.....	148
1.2. Objetivo	148
2. Análisis Inteligente de Datos	150
3. Red Neuronal de Kohonen (SOM).....	151
4. Análisis de Dominios	152
5. Caso de Estudio: Red Social de premios Nobel en Medicina de los	
últimos 10 aos	153
5.1. Recuperación de Datos (acquire)	153
5.2. Preprocesamiento de Datos (parse)	154
5.3. Procesamiento de Datos (filter)	154
5.4. Minería de Datos (mine)	154
5.5. Visualización (represent, refine, interact)	154
6. Conclusiones y Trabajo Futuro	158
6.1. Recomendación para trabajo a futuro	159
Open, online & collaborative knowledge visualization? Architecture	
proposal to achieve it	161
<i>Marcela Villalobos</i>	
1. Introduction.....	161
1.1. Information visualization	162
1.2. Objectives.....	163

2. State of the art	163
2.1. The concept	164
3. A new architecture	165
3.1. Back-End	166
3.2. Front-End	168
3.3. The whole picture	170
4. First proof of concept	171
5. Conclusions	172
6. Future Work	173

Redes Neuronales aplicadas al Reconocimiento de Huellas Dactilares

Griselda Cobos Estrada y Angélica González Arrieta

Departamento de Informática y automática. Universidad de Salamanca.
Patio de Escuelas 1, 37008 Salamanca
gris_wwjed@usal.es angelica@usal.es

Resumen En el presente trabajo se hace uso de una red neuronal Perceptrón Multicapa con el algoritmo de aprendizaje de retropropagación de errores, para su aplicación en el reconocimiento de huellas dactilares, teniendo como objetivo el medir la eficiencia de la red neuronal variando los datos de prueba, observando el comportamiento de la red en el caso especial de obtener solamente parte de una huella dactilar, donde no existan las características que dependen del centro de la huella.

Una vez diseñada la estructura general de la red después de varias pruebas y debidamente entrenada, se procedió a realizar las pruebas variando los puntos característicos y las características de dichos puntos. Los resultados en general, muestran un mejor reconocimiento al tener todos los puntos característicos de la huella y se va reduciendo el porcentaje de reconocimiento al disminuir el número de puntos característicos a 12, mas sin embargo recupera el nivel de porcentaje cuando el número de puntos son de 10, 8 y 5. En cuanto al número de características disminuidas, se tiene en general un menor nivel de porcentaje de reconocimiento, al quitar 2 características. Se obtuvo un buen porcentaje de aciertos al quitar las características que dependían del centro de la huella, así como el código de la huella, llegando de esta manera al objetivo deseado.

1. Introducción

Existen numerosas técnicas para la identificación de las personas que van desde claves que son identificadas por procesos informáticos, tarjetas con códigos de barras o chips integrados, hasta métodos biométricos. La selección de alguna de estas técnicas dependerá de la seguridad requerida en el sistema, así como de los recursos con los que cuente la empresa.

Se debe de tomar en cuenta que los sistemas tradicionales de identificación como los de clave y tarjetas, entre otros, tienen ciertas deficiencias que hacen más vulnerables a las personas y a las empresas, frente a daños de diversas índoles, como robos financieros o de información, etc. Esto ha provocado un gran desarrollo en la identificación con sistemas biométricos[1][2][3].

Los sistemas de identificación biométrica, basan sus características de detección en rasgos personales físicos que son medibles. Estos rasgos son captados por un sistema que automáticamente realiza el reconocimiento de la persona y deben

de cumplir con las siguientes características: Universalidad.- significa que todo individuo debe de poseer dicha característica. Unicidad.- la característica debe de ser única para cada individuo. Permanencia.- debe de ser invariable a través del tiempo. Cuantificable.- debe de ser medible de una manera cuantitativa.

El uso de la huella dactilar como técnica de identificación biométrica, es aceptada ampliamente debido a que cumple con las características antes mencionadas, además de contar con altos valores de fiabilidad, facilidad de uso, contra prevención de ataques, aceptación en general, estabilidad, autenticación, precio y estándares[4].

En el presente trabajo, se hace uso de una red neuronal artificial para la identificación de individuos mediante la huella dactilar. Se hará uso de un conjunto de datos que incluyen los puntos característicos que son extraídos de las huellas dactilares de varias personas mediante un sistema especializado en ello.

Se pretende valorar la efectividad de la red neuronal artificial en la detección de la huella dactilar variando el número de características de los puntos de la huella dactilar a identificar. Estas características incluyen: la distancia de crestas entre los diferentes puntos, el ángulo de acuerdo al centro de la huella, la posición del punto respecto al centro de la huella, el tipo de punto característico (abrupta, fusión), entre otros. El hecho de ir disminuyendo las características de las pruebas realizadas, se debe a que en la realidad se pueden presentar casos especiales en los que se tendrá que identificar alguna huella incompleta debido a múltiples razones.

El buen funcionamiento de la red neuronal depende, en gran manera de un adecuado entrenamiento, para lo cual se realizaron varias pruebas de donde se seleccionó el entrenamiento que obtuvo un rendimiento del 86.8% tomando en cuenta todas las características y todos los puntos característicos de una huella. Y en base a este entrenamiento, se hicieron otras pruebas variando el número de puntos característicos y de características de la huella.

2. Conceptos Teóricos

2.1. Huellas Dactilares

Biometría Identificativa Se llama biometría al estudio mensurativo o estadístico de los fenómenos o procesos biológicos [RaE]. Si hablamos de biometría identificativa queremos estudiar como fenómenos o procesos biológicos nos pueden ayudar a la identificación de individuos por medio de medidas.

La biometría identificativa busca un rasgo físico, dactilograma, iris, etc. que perdure en el tiempo, ya que un atributo que dependa del comportamiento de la persona puede cambiar con mayor facilidad.

Para que una característica física o anatómica sea utilizada por un sistema biométrico debe cumplir varias propiedades:

1. Universalidad: todo individuo debe poseer dicha característica.

2. Unicidad: la característica en única para una persona dada.
3. Permanencia: debe ser perdurable en el tiempo.
4. Cuantificable: debe ser posible su medición de forma cuantitativa.

Dactiloscopia Es el estudio y observación de las crestas y dibujos papilares que se encuentran en la cara interna de las manos y las plantas de los pies. Un dactilograma se diferencia de otro en las crestas, líneas de las yemas de los dedos, y los surcos, espacio entre dichas líneas.

Morfología general de las crestas capilares Para determinar si una huella es igual a otra se tiene que observar los puntos característicos de ambas, si en las dos huellas existen los mismos puntos en los mismos lugares podremos asegurar que las dos huellas son de la misma persona. Se puede definir punto característico como: la circunstancia particular o individual que presenta las crestas papilares en un lugar exacto, otros autores recurren a las palabras variedades o anomalías. Lo normal, estadísticamente hablando, es que se encuentre un número mínimo de treinta puntos característicos por cada huella de un dedo. Dependiendo del país se utiliza un número u otro para acordar la identidad, en España se estima como suficiente de diez a doce puntos característicos.

En los sistemas de identificación digitales, se suele buscar por dos clases de puntos, las abruptas y las bifurcadas, también llamadas fusiones[5][6][7].

2.2. Redes Neuronales

Una red neuronal artificial (RNA) es un modelo computacional o un sistema de procesamiento de información diseñado en base a las redes neuronales biológicas, que es capaz de aprender mediante ejemplos de adaptabilidad y que tiene características de robustez, manejo de fallos y capacidad de generalización[8].

Las redes neuronales artificiales son estructuras distribuidas, de procesamiento paralelo y que se encuentran formadas por neuronas o elementos de procesamiento, los cuales se encuentran interconectados entre sí mediante conexiones o sinapsis, en los cuales se almacenará conocimiento.

Las RNA generalmente son usadas en procesos de clasificación, identificación de patrones, predicciones y modelación matemática [9][10][11][12]. Entre las ventajas existentes de una RNA está su capacidad de manejar cambios no importantes en la información de entrada, como señales con ruido.

Tomando en cuenta las ventajas de las RNA, en el presente trabajo se utilizan para implementar una clasificación y reconocimiento de huellas dactilares.

3. Redes Neuronales aplicadas a la detección de Huellas Dactilares

A continuación se explicará la manera en que se aplicó la RNA en la detección de las huellas dactilares, los diferentes experimentos realizados y los resultados obtenidos.

3.1. Diseño de los datos

Existen métodos de detección de huellas dactilares por medio de redes neuronales artificiales[13][14][15][16], pero en este estudio se enfocará la atención a la detección de las huellas teniendo como base cantidades cada vez más pequeñas de puntos característicos. Hay que recordar que los puntos característicos son ciertas características que tienen las huellas dactilares que las hacen únicas entre todas las demás[17]. Por lo general existen como mínimo, 30 de ellos en las huellas dactilares. Entre ellos se pueden mencionar a los siguientes: abruptas y bifurcadas.

Es importante mencionar que se toma en cuenta un aproximado de 18 puntos característicos que coincidan con otra imagen de huella dactilar para poder determinar si pertenece a la misma persona.

En el presente estudio se tomó en cuenta un conjunto de datos de puntos característicos de diversas huellas de individuos. Para esto, cabe hacer mención de que el procesamiento de la imagen de la huella ya se llevó a cabo y ya se obtuvieron los datos importantes de dicha imagen.

Para el apropiado uso de la información contenida en la base de datos en la red neuronal artificial, se realizó primeramente un análisis para determinar los campos que serían importantes. Se tomaron en cuenta dos tablas: Malla y Punto_característico. La tabla de Malla es aquella donde se tiene el detalle de todos los puntos característicos que hacen referencia a otros cercanos, así como el número de crestas que existen entre los dos puntos. En esta tabla, los puntos característicos pueden repetirse pues el mismo punto puede apuntar a 1 ó más puntos cercanos a él. La tabla Punto_característico es aquella donde se enlistan los puntos característicos con todos sus detalles, es decir, su ángulo, posición, etc.

De esta manera se realizó una relación entre las dos tablas, dejando solamente un archivo csv, el cual se tomó de referencia para realizar el entrenamiento de la red neuronal. Este archivo de entrenamiento, contiene todos los datos de todas las personas, con una o todas sus huellas.

Para una mejor comprensión de los resultados obtenidos es necesario resaltar los dos términos principales que se usarán, los cuales comprenden los siguientes:

Puntos característicos.- su significado ya ha sido explicado, mas sin embargo, físicamente en el archivo csv se refiere a los registros de la tabla.

Características.- son los campos propiamente dichos de la tabla o archivo csv, y se enlistaron en los párrafos anteriores (dis, ang_direccion, etc).

Posteriormente, para determinar los datos de entrenamiento, se tomó en cuenta el hacer pruebas para identificar una huella con todos sus puntos y con todas sus características y una huella con menos puntos y con todas sus características, así como otras pruebas para identificar huellas con todos sus puntos característicos y con 5 características, además de pruebas con menos puntos característicos y las mismas 5 características.

Las características que se dejaron para las pruebas donde se tomaron en cuenta menos datos, son las siguientes: Dis, Posicion_x, Posicion_y, Tipo_punto, Codigo_huella.

Matlab requiere que los datos sean introducidos a manera de matrices. En el caso de las tablas, las características se encuentran en forma de vectores columna, sin embargo, Matlab las requiere en forma de vectores fila, por lo que hay que aplicarles una transpuesta. Lo que significa que las columnas se convierten en filas.

subsectionDiseño Red Neuronal Se utilizará una arquitectura de Perceptrón Multicapa[18][19]. El algoritmo de entrenamiento usado es el de retropropagación de errores[20].

La función de transferencia o de activación utilizada en las capas ocultas es la sigmoidea. Esta función proporciona el estado de activación actual partir del potencial sináptico o peso y del propio estado de activación anterior. Dado que la función de salida es la función identidad (es decir $f(x)=x$), por tanto el estado de activación es igual a la salida de la neurona.

Las redes multicapa se forman por un conjunto de redes de capa simple en cascada unidas por pesos, donde la salida de una capa es la entrada de la siguiente capa. Generalmente son capaces de aprender funciones que una red de capa simple no puede aprender, por lo que ofrecen mejores capacidades computacionales. Para que este incremento en poder computacional sea tal, tiene que existir una función de activación no lineal entre las capas, por lo que generalmente se utilizará una función de activación sigmoidea en detrimento de la lineal o umbral.

La función usada para la capa de salida es la lineal. La elección de las funciones de las capas ocultas y de salida, es el resultado de diversas pruebas realizadas. En los resultados obtenidos se pudo observar que las pruebas hechas con la función sigmoidea en las capas ocultas y la función lineal en la capa de salida tienen un mayor porcentaje de aciertos.

La estructura general de la red implementada se conforma de la siguiente manera: 10 neuronas en la capa de entrada, dos capas ocultas de 11 neuronas cada una y una capa de salida de 11 neuronas.

Capa de entrada.- está conformada por 10 neuronas que representan las características de las huellas dactilares:

1. Dis
2. ang_direccion
3. Posicion_x
4. Posicion_y
5. Tipo_punto
- 6.Codigo_huella
7. N_crestas
8. Cuad_x
9. Cuad_y

La capa de salida está compuesta por 11 neuronas, que son las personas que se encuentran en la base de datos. Cada neurona de la capa de salida, entregará el valor de 0 ó 1, dependiendo del usuario elegido como el dueño de la huella.

Capas ocultas.- la red cuenta con dos capas ocultas de 11 neuronas cada una. Para determinar el número de capas y de neuronas asignadas, se realizaron 22 pruebas diferentes desde 1 hasta 3 capas y con combinaciones desde 5 a 20 neuronas. Se observó que los resultados perdían efectividad al aumentar a 3 las capas

ocultas y al asignar a más de 15 las neuronas en dichas capas. Se hicieron pruebas con dos capas y con 11 neuronas en cada una y se observó que mejoraban los resultados.

Los pesos usados en la red neuronal, son los que inicializa automáticamente Matlab. Una matriz de pesos de entrada (es decir, entre la capa de entrada y la primer capa oculta) tiene las siguientes dimensiones: $M \times N$, donde M corresponde al tamaño de la matriz oculta (11 en este caso) y N se obtiene como el producto entre el tamaño de la entrada (10 en este caso) y el número de retardos asociados a los pesos.

La dimensión de la matriz de pesos de las capas ocultas es $M \times N$, donde M corresponde al tamaño de la segunda capa oculta y N al producto entre el tamaño de la primer capa (11 en este caso) y el número de retardos asociados a los pesos.

Finalmente se presenta la matriz de pesos de salida, es decir, entre la última capa oculta y la capa de salida, donde el tamaño de la matriz es $M \times N$, donde M corresponde al tamaño de la capa de salida y N al producto entre el tamaño de la segunda capa (11 en este caso) y el número de retardos asociados a los pesos.

4. Pruebas y Resultados

Primeramente se entrenó la red neuronal con el conjunto de datos completo de los puntos característicos de las huellas de los usuarios.

El primer parámetro que se consideró para entrenar la red neuronal fue el valor del error cuadrático medio, para no realizar un sub-entrenamiento, ni un sobre-entrenamiento en detrimento del desempeño de la red neuronal.

Para lo anterior, se realizaron diferentes pruebas modificando el parámetro MSE. La red que tiene un mejor desempeño es a la que se le ha asignado un valor MSE de cero. a partir de este resultado las pruebas posteriores se implementaron usando un error MSE igual a cero.

Como se pudo observar en los resultados anteriores, el error cuadrático medio no llega a ser 0, pues el entrenamiento se detiene a los 81 epochs. Se alcanza un error cuadrático medio de 0.0163885. Sin embargo, el error de generalización nos indica lo inadecuado de tener errores muy pequeños, por lo que este error se considera aceptable para la realización de las pruebas.

Posteriormente se realizaron otros entrenamientos sobre el mismo conjunto de datos, con la finalidad de obtener el mejor comportamiento de la red neuronal. Para ello se generaron cuatro entrenamientos, net, net1, net2 y net3. De esta forma se obtuvieron cuatro redes neuronales entrenadas de manera diferente, sobre el mismo conjunto de datos.

Posteriormente se pusieron a prueba las cuatro diferentes redes neuronales, variando el número de puntos.

antes de empezar a realizar la mayor parte de las pruebas y en base a varias pruebas, se planteó una hipótesis, basada en el hecho de que algunas características estaban introduciendo ruido a los resultados. Debido a ello se realizaron diversas

pruebas que permitieran identificar la característica que introduce dicho ruido. El experimento se realizó con el mismo número de puntos, siendo estos 121, y las características que se tomaron en cuenta para todas las pruebas fueron: dis, posicion_x, posicion_y, tipo_punto y codigo_huella; ya que el desempeño con dichas características fue del 100 %. Para cada experimento se añadieron cuatro de las cinco características restantes (ang_direccion, n_crestas, cuad_x, cuad_y y contabilizarComo) dejando una fuera para identificar la característica productora del ruido.

Se notó que la característica “contabilizarComo” era un dato de entrada “no útil”, pues era exactamente igual para todos los casos, con valor igual a uno. Por tal motivo se decidió eliminar dicha característica.

a partir de la eliminación de la característica anteriormente mencionada, se realizó nuevamente el entrenamiento de la red neuronal, siendo en este caso tres diferentes entrenamientos.

Como se pudo observar en los resultados, el desempeño de la red neuronal sin la característica “contabilizarComo” es mucho mejor que cuando es considerada. Teniendo en cuenta este hecho, se procedió a realizar las pruebas siguientes sin considerar la característica “contabilizar como”. así mismo, la red neuronal que se consideró para continuar las pruebas fue Net5, debido a que presentó el mejor porcentaje de desempeño, 79.21 %. De esta forma, las pruebas se realizaron con una red neuronal entrenada que presenta un desempeño del 79.21 %. Tomando como base este entrenamiento se empezó a variar el número de puntos y posteriormente el número de características.

Dados los resultados anteriores podemos decir que, a medida que se reducen los

Cuadro 1. Detalle de pruebas (porcentaje de desempeño)

Características	121	20	18	16	14	12	10	8	5	%
Todas	86.77	75.00	72.22	75.00	71.43	75.00	90.00	87.50	80.00	79.21
Sin ang_direccion	86.77	75.00	72.22	75.00	71.43	75.00	90.00	87.50	80.00	79.21
Sin ang_direccion y sin N_crestas	86.77	75.00	72.22	100	71.43	75.00	90.00	87.50	80.00	81.99
Sin ang_direccion, sin N_crestas, sin Cuad_x, sin Cuad_y	77.68	70.00	66.67	68.75	64.29	58.33	80.00	62.50	80.00	69.80
Sin ang_direccion, sin N_crestas, sin Cuad_x, sin Cuad_y, sin codigo_huella	86.77	80.00	77.78	81.25	78.57	75.00	90.00	75.00	80.00	80.48

puntos característicos de 20 a 5, el desempeño se ve afectado. Sin embargo, eso

no es del todo exacto pues, podemos apreciar como en los resultados con puntos característicos 10, 8 y 5, el porcentaje de rendimiento es sumamente elevado. No obstante son muy parecidos al desempeño cuando se tienen todos los puntos característicos totales.

Por otro lado, también se puede ver que al ir disminuyendo el número de características el comportamiento de la red neuronal se mantiene constante, alrededor del 80 %, salvo en el caso en que se suprime las características `cuad_x` y `cuad_y`. Sin embargo, cuando se suprime la característica `codigo_huella`, se observa una mejora en el rendimiento de la red. Concluyendo de esta manera que quitando el conjunto de dichas características en la prueba final, mejora el rendimiento de la red, pues no las está considerando como un aspecto de suma importancia para la clasificación de las clases.

Podemos concluir que la red neuronal quedó entrenada bajo un desempeño genérico del 79.21 % y que al ir disminuyendo sus puntos característicos (caso en que la captura de la huella no fuera correcto) el desempeño también disminuye. También podemos decir que al ir eliminando características, la red neuronal se mantiene sin alteraciones importantes, salvo que se eliminen características fundamentales, como el caso de `cuad_x` y `cuad_y`, y en el que el desempeño de la red neuronal se decreta de manera sustancial. Finalmente, podemos mencionar también la existencia de características que no son tan fundamentales y que incluso se puede prescindir de ellas, como el caso de `contabilizarComo`, evidenciando de esta forma la hipótesis declarada al inicio de las pruebas.

5. Conclusiones

Las redes neuronales artificiales en el campo de la detección de las huellas dactilares son de amplio uso y en el presente trabajo se ha implementado al caso específico de tener solamente unos puntos característicos cuyas especificaciones o características especiales no dependan del centro de la huella, pues se está suponiendo que la red neuronal vaya a detectar una huella incompleta. Los resultados obtenidos, tomando en cuenta lo anteriormente mencionado, fueron satisfactorios, ya que se logró un porcentaje general de acierto del 80.48 % en todas las pruebas donde se retiraron las cuatro características que dependían del centro de la huella y donde se varió el número de puntos característicos, ignorando también la característica del código de la huella, con lo que se logró mejorar la eficiencia de la red en comparación con las pruebas donde solamente se retiraron las 4 características dependientes del centro de la huella.

Particularizando el punto anterior, se pudo observar el comportamiento que tuvo la red neuronal artificial en su entrenamiento y posteriormente en las pruebas realizadas en cuanto al patrón de aprendizaje generado con y sin ciertas características, como fue el caso especial de: “`contabilizarComo`”. Característica que tuvo que retirarse del entrenamiento para poder generar mejores resultados. Y después el caso especial de la característica: `codigo_huella`, a la cual la red no necesitó para dar un mejor rendimiento en la detección de la huella, aún toman-

do en cuenta que ya se habían ignorado las cuatro características relacionadas con el centro de la huella. De esta manera se tiene que la red neuronal fue capaz de generar un patrón de aprendizaje lo suficientemente estable tomando en cuenta ciertas características que determinó ser las más importantes de manera automática. Siendo éstas, realmente, las características que se requerían para la detección de la huella.

La red neuronal requirió de varias pruebas para concretizar su estructura general. Para determinar las capas ocultas, se realizaron pruebas con diferente número de capas y neuronas. Y se vio que con dos capas ocultas de once neuronas cada una, era suficiente para el buen funcionamiento de la red con nueve entradas. Con esto se confirma el hecho de que un perceptrón con dos capas ocultas es suficiente para la representación de regiones de decisión complejas.

Se determinó el uso de una red perceptrón multicapa con el algoritmo de aprendizaje de retropropagación. El uso de esta red fue debido al extensivo uso que tiene en tareas similares donde se requiere el reconocimiento de clases en base a patrones de aprendizaje que es capaz de generar con los datos de entrenamiento que se le asignen.

Los pesos iniciales deben de ser aleatorios, por eso, se tomaron los que asigna por default la red dentro de Matlab. También para determinar el error cuadrático medio, se hicieron diferentes pruebas, donde se determinó dejarlo en valor de cero, pues proporcionó el mejor porcentaje en el entrenamiento y sobre esa red, se realizaron las pruebas necesarias.

Es necesario recordar que para la selección de la arquitectura de la red, se requiere de un proceso de aprendizaje con experimentación, es decir, no hay una regla que determine con exactitud la mejor opción, y debido a esto se puede dejar como objeto de estudio posterior el realizar más combinaciones en las pruebas para llegar a mejores resultados e incluso hacer uso de otro tipo de red que dé una mejor opción para el caso de estudio planteado en el presente trabajo.

Referencias

1. Mislav Grgic Kresimir Delac. a survey of biometric recognition methods. In *46th International SyrnPoSium Electronics in Marine*, 2004.
2. Ming Zhang and John Fulcher. Face recognition using artificial neural network group-based adaptive tolerance trees. *IEEE Transactions on Neural Networks*, 7:555–568, 1996.
3. abdeljaoued Y. Mayoraz E. Ben-Yacoub, S. Fusion of face and speech data for person identity verification. *IEEE Transactions on Neural Networks*, 10:1065–1074, 1999.
4. Salil Prabhakar Sharath Pankanti and anil K. Jain. On the individuality of fingerprints. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 24:16, 2002.
5. Josef Strom Bartunek. Minutiae extraction from fingerprint with neural network and minutiae based fingerprint verification. Master's thesis, Blekinge Tekniska Hogskola, 2005.

6. F.H. Herrera Santana and L. Uribe Kaffure. Detección de detalles en huellas dactilares usando redes neuronales. Master's thesis, Pontificia Universidad Javeriana, 2005.
7. Gilberto Emilio Hernández Sorroza. *Obtención de características deltas, cores y encierros en una huella dactilar*. PhD thesis, Universidad Tecnológica de la Mixteca, 2002.
8. Pablo García Estévez. Las redes neuronales supervisadas. Technical report, Universidad Complutense de Madrid, 2005.
9. González arrieta angélica. *Identificación del arma y la Munición utilizadas en un disparo con técnicas conexionistas*. PhD thesis, Universidad de Salamanca; Departamento de Informática y automática, 2000.
10. Y. Q. Shi E. You and P. Engler. Fingerprint pattern recognition for medical uses a frequency domain approach. Technical report, New Jersey institute Technology.
11. Carlos Mallo González. Predicción de la demanda eléctrica horaria mediante redes neuronales artificiales. Technical report, Universidad de Oviedo.
12. V.H. García Ortega. Sistema de reconocimiento de huellas dactilares para el control de acceso a recintos. Technical report, Instituto Tecnológico de Chihuahua, 2001.
13. Airam Carlos Pais Barreto Marques and antonio Carlos Gay Thomé. Neural network fingerprint segmentation method. Technical report, Universidade Federal do Rio de Janeiro.
14. Pedro Mayorga Ortiz. *Reconocimiento de Huellas Digitales Utilizando Redes Neuronales*. PhD thesis, Instituto Politécnico Nacional, 1998.
15. Hugo Vieira Neto. Fingerprint classification with neural networks. *IEEE*, 1997.
16. E. anbalagan. E. Mohan B. Jayaraman, C. Puttarnadappa and Srinivasarao Madane. Fingerprint authentication using back-propagation algorithm. *Medwell International Journal of Soft Computing*, 3:6, 2008.
17. García Díaz Pablo. Reconocimiento óptico de los puntos característicos de dibujos lofoscópicos. Master's thesis, Universidad de Salamanca, 2008.
18. K. I. Funahashi. On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 2:183–192, 1989.
19. Stichcombe M. White H. Hornik, K. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.
20. Williams R.J. Rumelhart D. E., Hinton G.E. Learning representations by back-propagation errors. *Nature*, 323:533–6, 1986.

Recuperación de la Información Y Técnicas de Social Bookmarking

Beatriz Fortínez and Carlos G. Figuerola

Departamento de Informática y automática
Universidad de Salamanca, Salamanca, España.
{beatriz_fortinez, figue}@usal.es

Resumen El campo de la recuperación de la información en la actualidad se orienta hacia esquemas colaborativos de recuperación, teniendo como foco de interés al usuario en su dimensión de consumidor de información y generador de la misma. Los sistemas de social bookmarking emergentes de la Web 2.0, Delicious(<http://delicious.com/>) como caso de estudio, brinda al usuario de búsquedas específicas y avanzadas una gamma de herramientas que faciliten la búsqueda, organización y categorización de la información. Se desarrollan las técnicas de bookmarking, que por medio de tags o etiquetas identifican recursos y enlaces de interés individual y colectivo. Surgen las folksonomías que clasifican la información de manera flexible puesto que se dejan de lado las estructuras jerárquicas de clasificación antes usadas. Esta nueva clase de esquema le da al usuario poder en el manejo de información, selecciona sus fuentes y filtra sus resultados con criterios humanos apoyados en el proceso algorítmicos previos, sustentando la base social de Internet. El artículo consta de dos partes, la descripción del Esquema Clásico RI y el Esquema RI emergente, cuya base es la web social con los servicios interactivos de la Web 2.0.

Palabras Claves: Recuperación de la Información, Web 2.0, Social Bookmarking, Folksonomías.

1. Introducción

La actividad de la recuperación de la información desde sus comienzos ha ido en continuo desarrollo de técnicas y métodos efectivos de búsqueda y de obtención de resultados satisfactorios para el usuario. La recuperación de la información, en su concepción actual, es un sistema de extracción de documentos en formato electrónico relevantes a una temática específica. En un sistema clásico RI el proceso es a partir de una necesidad informativa que se convierte en una consulta compuesta de palabras claves y consultas en lenguaje natural, la cual es utilizada como entrada en un Sistema de recuperación tradicional (figura 1)[1] y procesadas por los motores de búsqueda que despliegan una lista con los resultados más relevantes de la búsqueda, son dos las medidas consideradas en

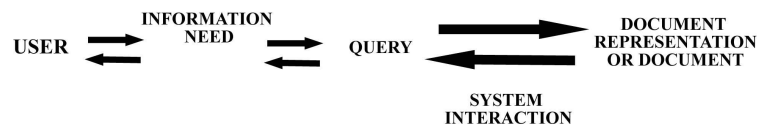


Figura 1. Esquema Clásico RI.

este esquema: exhaustividad y precisión. El ideal es que ambas características sean alcanzadas por los algoritmos encargados de este proceso.

Hoy en día, los sistemas de recuperación se apoyan en la contribución humana, agregando criterios más complejos:

- Intuición
- Razonamiento
- Conocimientos y Dominio de temas
- Experiencia

A partir de los criterios anteriores, la RI se dirige hacia el desarrollo de algoritmos que apliquen los comportamientos del usuario en el proceso de búsqueda, siendo la fundamentación y filosofía de los servicios colaborativos que se van introduciendo gradualmente en el actual world wide web.

De lo anterior surge el término de inteligencia colectiva. La inteligencia colectiva descansa en el intercambio de opiniones y juicios del público miembros de la comunidad de internautas que posee Internet en el punto dinámico de la Web 2.0. actualmente la Web ofrece servicios de orden social, promueven la formación de redes sociales, enciclopedias de conocimiento abierto y portales para compartir multimedia y enlaces. Figura 2.

Son los servicios de Social Bookmarking a los cuales dirigimos nuestra atención en este documento, definimos su función social y mecanismos de almacenamiento, recuperación y búsqueda de información, como ejemplo de estos servicios tomaremos a Delicious uno de los pioneros en las anotaciones sociales.

2. Esquema Clásico de la Recuperación de la Información

El objetivo de un Sistema RI, es proveer al usuario de la existencia de documentos relacionados a su consulta. Sin embargo dicha tarea se lleva a cabo teniendo en consideración la búsqueda, organización y selección de material relevante además del tiempo de búsqueda. De estos se derivan 3 áreas en RI:

- Análisis de Contenido
- Estructura de la información
- Evaluación



Figura 2. Servicios Web 2.0.

El análisis de contenido y la estructura de la información están estrechamente vinculados, puesto que las técnicas de crawler se encargan de rastrear información en la web identificando los contenidos de las páginas web a través de las etiquetas de código HTML (HyperText Markup Language), a partir de este análisis se deriva la necesidad de una estructura de la información, lo cual dicho sea de paso es el principal problema de la RI, la pobre estructura de la información. De ahí que el estudio de la organización de los contenidos de las páginas web, aplicando tecnologías de descripción de contenidos tales como: RDF (Resource Description Framework), XML (Extensible Markup Language); las cuales contribuyen a convertir el contenido de la web en algo entendible y procesable para los agentes de software involucrados en el proceso de clasificación de información.

La semántica de la web representa un punto clave en el desarrollo y evolución de mejores sistemas RI. Se necesita cubrir la organización lógica de la información y manejar variables descriptivas de los contenidos de las páginas web que faciliten el proceso de localización, categorización y selección de información. Sin embargo hay variables que van más allá del control sistemático, exigen el factor humano.

Las herramientas de búsqueda en el esquema clásico están basadas en palabras claves e índices, localizados por crawlers o robots que navegan la web en busca de contenidos para una consulta dada. Se distinguen tres modelos clásicos de búsqueda: El modelo Booleano, Probabilístico y el Vectorial[2]. Este último es ampliamente utilizado por su desempeño en la recuperación de información.

Tanto la taxonomía y la semántica de la información son dos de los elementos de los cuales depende que el proceso de extracción y recuperación de información sea más fácil de realizar. Estos modelos son aplicados a las herramientas de búsqueda convencionales: Motores de búsqueda, Directorios y Meta buscadores, los cuales a nivel de usuario no representan ninguna complejidad de uso, sin embargo la evaluación de resultados reposa en el usuario que es al final el que decide si su necesidad informativa se ha cumplido.

Las nuevas tendencias en la web orientan el paradigma de búsqueda en dirección al usuario, dándole un valor agregado a los sistemas RI actuales, que amplían sus fronteras a la contribución colectiva de la red de usuarios de Internet.

3. Tendencias y Paradigmas en la Recuperación de la información

Es a partir de la Web 2.0 que los servicios de connotación social emergen de los paradigmas tradicionales de Internet. La valoración de las colecciones documentales recuperadas por el usuario, adquieren relevancia gracias a la votación popular de la comunidad virtual, estableciendo funciones intermediarias de los usuarios en los Sistemas RI. Nacen así las folksonomías derivadas de las taxonomías, solo que en este caso son los usuarios los que establecen la categorización y estructuración no formal de la información. Internet se encuentra en el punto de evolución adquiriendo un rol dinámico y participativo del público. La idea básica es la conexión entre las diferentes comunidades web. Este nuevo enfoque promueve la inteligencia colectiva por medio del intercambio de recursos, opiniones y portales de conocimiento abierto. De pronto el usuario se convierte en un agente activo en la producción de la información, adquiriendo al mismo tiempo la administración de la categorización y clasificación de los recursos. Lo que busca este esquema es la sinergia entre hombre - máquina, desarrollando el potencial para obtener información con diversidad de perspectivas. Figura 3.

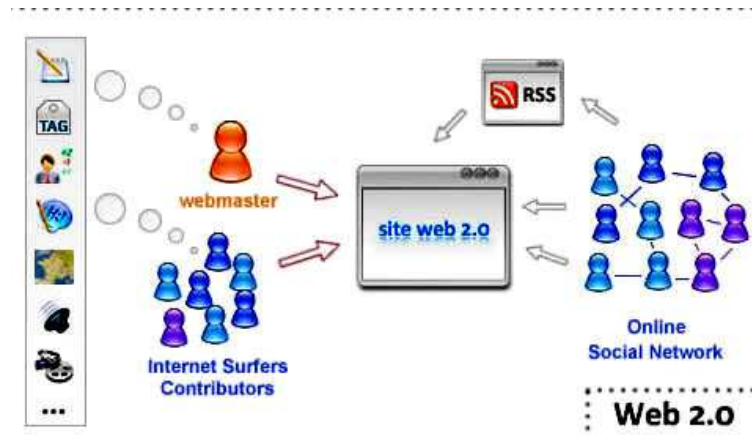


Figura 3. Web 2.0

En materia RI contamos con servicios que permiten la organización y clasificación de material en formato electrónico por medio de los siguientes métodos:

- Bookmarking de URL's(Etiquetado de enlaces).
- Catergorización de multimedia y datos (portales como: Flickr y Technorati).

Sin embargo dirigimos nuestra discusión a los servicios de Bookmarking, para explorar los mecanismos colaborativos de categorización y clasificación de información a través del etiquetado de enlaces bajo el criterio y juicio de los usuarios.

4. Social Bookmarking

Al hablar de social bookmarking nos referimos a las anotaciones sociales realizadas en portales de intercambio de recursos y enlaces. Se trata del almacenamiento en línea de enlaces de interés, en primera instancia a nivel individual, el cual trasciende a lo social y público una vez que son consultados y compartidos por el resto de usuarios del servicio. Un Sistema Social Bookmarking trata de un trabajo de cooperación, la suma del esfuerzo individual para aumentar la efectividad de la RI.

La vinculación de los servicios de Social bookmarking y los motores de búsqueda así como tambien los directorios tradicionales, es en ambas direcciones puesto que es a partir de los sistemas previos que extraemos los enlaces para almacenarlos y administrarlos dentro de los servicios colaborativos. Figura 4.



Figura 4. Esquema Social Bookmarking.

Un Sistema de Social Bookmarking trata de un trabajo de cooperación al sumar el esfuerzo individual al de otros, para el intercambio de información.

Mencionamos algunas de las características de los Servicios de Social Bookmarking:

- Presencia de un buen volumen de usuarios, que hacen relevante el aspecto social.
- El servicio posee una extensiva documentación del mismo.
- Posibilidades de importación y exportación de bookmarks.
- Un buen servicio de bookmarking deberá tener feeds, agregadores tal como: RSS(Really Simple Syndication), nueva tecnología que permite actualización de la información suscrita.

El mecanismo de estos servicios esta basado en etiquetas, palabras que organizan la información, categorizando y clasificando un enlace. La tarea de categorización y clasificación de material es realizada por la comunidad de usuarios que comparten una jerga común en un tema determinado. Dentro de estos servicios se simplifica la distribución de material, listas bibliográficas, enlaces, referencias, para un grupo de usuarios con intereses comunes. El valor agregado de estos servicios radica en el juicio humano, de ahí que la red de usuarios es para un grupo de usuarios avanzados y grupos académicos e investigadores que tienen definido temas de interés y que buscan una constante actualización del material. La contribución participativa del usuario es base de la web social, la cual pretende ser flexible y enriquecedora a nivel de contenidos y de oferta informativa.

5. Papel de las Folksonomías

El poder descriptivo de las Folksonomías representa una ventaja en la adopción del esquema participativo de la recuperación de la información. Las folksonomías describen estructuras parecidas a las taxonomías, definen estructuras conceptuales de bajo peso las cuales emergen de comunidades de usuarios que colectan recursos, reflejando la visión de las comunidades, de atributos asociados supliendo la categorización formal de los recursos[3,4]. Enriquecen el conocimiento base con descripciones y categorización de intereses, recursos y enlaces compartidos a través de los Sistemas de social Bookmarking.

Las ventajas se sintetizan en dos: la conexión estrecha entre los términos y los recursos descritos; son menos costosos en cuanto desarrollo, dado que surgen de los mismos usuarios[5]. La flexibilidad estructural de estas Folksonomías demanda un control de vocabulario y el establecer determinados parámetros de etiquetado, que son acordados en consenso por el grupo o comunidad de usuarios.

Los Sistemas RI basados en el etiquetado de enlaces, sugieren etiquetas populares y frecuentes en la comunidad, hay una relación entre etiquetas marcadas con anterioridad que facilitan el manejo y asignación de las mismas, que permiten al usuario añadir un significado explícito y propio entendimiento, expresando opiniones y perspectivas de los bookmarks almacenados.

El objetivo es disminuir la ambigüedad de términos fuera del contexto así como también la inconsistencia entre los contenidos y etiquetas, nos planteamos pues la interrogante ¿Cómo controlar o estandarizar el vocabulario colectivo?, son

estructuras no formales que necesitan una parametrización de uso sin dejar la flexibilidad de comprensión cognitiva de lado, al parecer trataríamos de crear patrones de asignación y clasificación más o menos comunes que se presten a la comprensión hombre- máquina y que en un futuro se extiendan a entradas de sistemas de recomendación y sugerencias de contenidos, por el momento, de este mecanismo destacamos sus bondades, el valor descriptivo, la posibilidad de un sumario de contenidos, intercambio de recursos y características de uso como la flexibilidad y comprensión de parte de la comunidad de usuarios.

Cuadro 1. Taxonomías vs Folksonomías

Taxonomías	Ontologías formales y clasificación jerárquica de la información	Proceso algorítmico.	Relevancia de información a partir de hiperenlaces.
Folksonomías	Clasificaciones no jerárquicas de la información.	Proceso realizado por la comunidad de usuarios.	Sistema de ranking por popularidad de etiquetas

6. Caso de Estudio, Delicious

Delicious[6] es considerado dentro de la categoría de software social, el que facilita la interacción social, intercambio de colaboraciones y de información, promueven las comunidades, basadas en actividades de grupos de usuarios. Social bookmarking permite tecnológicamente compartir recursos a través de enlaces almacenados o de hiperlinks para conectar personas con una base común de intereses. Delicious es un sitio para compartir enlaces y recursos con un grupo de la comunidad de usuarios de Internet que realizan búsquedas más avanzadas y específicas según el interés y la necesidad informativa.

Delicious provee de tags que identifican los recursos, las cuales pueden ser marcadas mediante etiquetas populares, sugeridas o ciegas (asignadas por los usuarios tomando criterios individuales de categorización). Las etiquetas más populares, utilizadas por la comunidad web, muestra el número de personas que han etiquetado el enlace así como también muestra las etiquetas que le identifican para el grupo. Creando así un tesoro, que facilitará la búsqueda de temas relacionados con el enlace de interés. a partir de estos parámetros de popularidad el usuario ha de realizar el filtro en la búsqueda de información y explorar enlaces indirectos. La búsqueda de información esta basada en el uso de etiquetas, asignadas por los diferentes miembros de la comunidad de usuarios del servicio. Delicious facilita la exploración de las etiquetas más populares y las del interés del usuario. La mayor de las ventajas de Delicious es la accesibilidad de los recursos, tanto para los usuarios como para la comunidad web. El filtro humano representa uno de los mayores capitales de la base de conocimiento

creada con la participación colaborativa de la comunidad. La connotación social de Delicious, lo incluye en el top de la lista de herramientas de aprendizaje para el 2008 (<http://www.c4lpt.co.uk/recommended/top100.html>, consultado en septiembre 2008). Por sus características de eficiencia y productividad en la organización de etiquetas proporcionando soluciones a los que los utilizan en el ámbito académico.



Figura 5. Delicious

7. Aportaciones del Documento

La revisión bibliográfica realizada sobre los Sistemas de Recuperación de Información con base social, en específico los Servicios de Social Bookmarking, representan un cambio en el esquema tradicional de almacenamiento y recuperación de información. a nivel de usuario se tiene a disposición herramientas y recursos para realizar búsquedas efectivas y satisfactorias[7]. El nuevo paradigma descansa en la etapa social de Internet con el advenimiento de servicios interactivos derivados de la Web 2.0[8], actualmente observamos en Internet el intercambio y generación de información a partir de las comunidades de usuarios y público en general que hace uso de la red de redes. El crecimiento de Internet se justifica con el desarrollo de tecnologías y el dinamismo de los servicios ofrecidos. Podemos comparar las características de cada esquema con la tabla 2.

Cuadro 2. Comparación de Esquemas

Esquema	Objetivo	Paradigma	Estructura de la información	Ventaja
Esquema Clásico RI	Esquema orientado a la funcionalidad de los procesos algorítmicos.	automatización de selección y clasificación de material.	Indexación de la información.	Tienen la características de ser exhaustivos en la búsqueda de material.
Tendencias Colaborativas	Orientado al usuario.	Contribuciones colectivas de información (consumo y producción).	Etiquetado de la información.	Personalización de búsquedas.

Ahora con los nuevos esquemas y paradigmas cognitivos en los Sistemas RI, Delicious se nos presenta como una herramienta, en principio de propósito individual, que permite marcar y etiquetar enlaces y tenerlos disponibles desde cualquier ordenador. Sin embargo esta idea alcanza la dimensión social de compartir e intercambiar información desde este portal de marcado. Cambiamos métodos de almacenamiento, administración y recuperación de información a partir de categorizaciones y clasificaciones manuales con el valor agregado del juicio humano, desde este punto podemos considerar el proceso de filtro y categorización de recursos como un proceso de connotación cognitiva y orientado al usuario. Utilizamos patrones de búsquedas flexibles y entendibles tanto para el humano como para la máquina.

La organización y clasificación previa de la información realizada por los Sistemas RI tradicionales realizan procesos automáticos de búsqueda y selección de documentos que satisfagan más o menos la consulta del usuario, se establecen paradigmas de efectividad y medidas de precisión que evalúan el desempeño de estos sistemas, sin embargo la naturaleza social de los servicios interactivos de la web 2.0 atraen a más usuarios y por ende más contribuciones que facilitan el rastreo de información convirtiendo la Web en algo más dinámico y personalizado. El valor agregado de la web social es la cobertura de perspectivas individuales y colectivas, las cuales generan una independencia de organizaciones y de entidades influyentes en el World Wide Web, es el poder del público el que produce y demanda conocimiento en blogs, Wikis y servicios bookmarking.

El individuo se especializa en temas puntuales, descentralizando las fuentes generadoras de información tradicionales. No hemos de desvincular el uso de los motores de búsqueda tradicionales, pues son a partir de ellos que localizamos la información y posteriormente la marcamos con etiquetas sugeridas por la comunidad de usuarios y por el contenido de la página web seleccionada.

Trabajos futuros en esta línea de investigación apuntan a la colaboración conjunta de Sistemas RI tradicionales y alternos, dada la interrelación que guardan

entre sí. Sistemas de recomendación que combinan lo mejor de ambos esquemas, de los Sistemas RI tradicionales conservamos la base científica para la búsqueda algorítmica de colecciones de documentos que son seleccionados y clasificados de acuerdo a consultas lanzadas por los usuarios y de los sistemas de marcado colaborativo las ventajas de una clasificación previa del material realizado por humanos, sin costo computacional adicional y sin límites en cuanto cantidad de información; desventajas presentes en los Directorios que son más o menos el modelo previo al esquema de categorización de información en la que participa el humano.

aunque la iniciativa de estos Sistemas de Recomendación basados en el filtro colaborativo datan de la años noventa, investigaciones y trabajos en desarrollo maduran en concepto y técnica; incorporan matrices de similitud y algoritmos de solución derivados de la inteligencia artificial que permitirán la actualización en tiempo real de perfiles y sugerencias[9].

Referencias

1. Carol A. Hert. *Understanding information retrieval interactions: theoretical and practical implications*. Ablex Publishing Corporation, Greenwich, Connecticut Londo, England, 1997.
2. F. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley, Harlow, England., 2000.
3. Andreas Hotho, Robert Jaschke, Christoph Schmitz, and Gerd Stumme. FolkRank: a ranking algorithm for folksonomies. *in Proc. FGIR 2006*, 2006.
4. Thomas Vander Wal. Folksonomy, february 2007.
5. Pierpaolo Basile, Domenico Gendarmi, Filippo Lanubile, and Giovanni Semeraro. Recommending smart tags in a social bookmarking system. *University of Bari*,, pages 22–29, 2007.
6. Matt Biddulph. Introducing del.icio.us what is del.icio.us?, november 2004.
7. Don Turnbull. Interacting with recommender systems. acm sigchi (computer-human interface). The University of Texas, School of information, 1999.
8. Tim O'Reilly. What is web 2.0 design patterns and business models for the next generation of software. O'Reilly online magazine, September 2005.
9. Anne Yun-an Chen and Dennis McLeod. Collaborative filtering for information recommendation systems. *anne Yun-an Chen and Dennis McLeod Department of Computer Science and Integrated Media System Center University of Southern California, Los angeles, California, USA*, 2005.
10. Delicious. Servicio de social bookmarking.
11. Google. Motor de búsqueda.
12. Yahoo. Directorio de yahoo.
13. unknow. *Delicious Handout, consultado en septiembre 2008*.
14. Inc. singhal@google.com amit Singhal Google. Modern information retrieval: a brief overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 2001.
15. J alonso arévalo. Recuperación de la información: La búsqueda bibliográfica. E-LIS Library and Information Science (LIS), 2004.
16. Eneko astigarraga. Social bookmarking - compartiendo enlaces de internet. online magazine, febrero 2005.

17. Simone Braun, andreas Schmidt, and andreas Walter. Ontology maturing: a collaborative web 2.0 approach to ontology engineering. 2007.
18. Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Science Department, Stanford University, Stanford, Ca 94305, USA*, 1997.
19. Javier Cremades. *MICROPODER La fuerza del ciudadano en la era digital*. Editorial Espasa Calpe S.a., 2007.
20. Delicious. Sección de ayuda de delicious.
21. Thomas Gruber. Collective knowledge systems: Where the social web meets the semantic web. *Journal of Web Semantics*, julio, 2007.
22. Susan Gunelius. Overview of del.icio.us. web blog, , about.com.
23. Maria Hernandez de Juan. Modelos de recuperacion de la informacion: Recuperacion y acceso a la informacion. pagina web actualizada, mayo 2008.
24. Educause Learning Initiative. 7 things you should know about... social bookmarking. Educause Learning Initiative, mayo 2005.
25. alex Iskold. The social bookmarking faceoff. Read Write Web online magazine, september 2006.
26. Jinbeom Kang and Joongmin Choi. Detecting informativeweb page blocks for efficient information extraction using visual block segmentation. *IEEE 2007 International Symposium on Information Technology Convergence*, 2007.
27. kevin Kelly. We are the web. *Revista online Wired*, 2005.
28. Jon M. Kleinberg. authoritative sources in a hyperlinked environment. *Journal of the ACM*, Vol. 46, No. 5, 1999.
29. Jane Klobas. *Wikis: Tools for Information Work and Collaboration*. Chandos Publishing (Oxford), 2006.
30. Frederick W. Lancaster. *El control del vocabulario en la recuperacion de informacion*. Universidad de Valencia, 1995.
31. Mary Madden and Susannah Fox. Riding the waves of web 2.0 more than a buzzword, but still not easily defined. *Pew Internet Project*, October, 2006.
32. David Muir. Simply del.icio.us online social bookmarking, or: Tagging for teaching. 2005.
33. Eva Maria Méndez Rodríguez and José antonio Merlo Vega. Localización, identificación y descripción de documentos web: tentativas hacia la normalización.
34. Lee Rainie. 28dewey and his decimals, internet users are revolutionizing the way we classify information and make sense of it. *P E W I N T E R N E T & a M E R I C A N L I F E P R O J E C T*, january 2007. January 31, 2007.
35. C. J. van Rijsbergen. *INFORMaTION RETRIEVAL*. Dept. of Computer Science, University of Glasgow, 2da. Edición 1979.
36. S.E. Robertson and K. Spark Jones. Relevance weighting of search terms. *Journal of america Society for information science*, 27:129-146, 1976.
37. Sebastian Ryszard Kruk and Stefan Decker. Jeromedl and foafrealm - taking advantage of semantic social collaborative filtering in digital libraries. *Digital Enterprise Research Institute, NUIG, Ireland* <http://www.deri.org>, September, 2005.
38. Gerard Salton, amit Singhal, Mandar Mitra, and Chris Buckley. automatic text structuring and summarization. *Information Processing & Management*, Vol. 33, No. 2:pp. 193-207, 1997.
39. M. Szomszor, C. Cattuto, H. alani, K. O'Hara, a. Baldassarri, V. Loreto, and V. D. P. Servedio. Folksonomies, the semantic web, and movie recommendation. 2007.
40. Daniel G Teivelis. Semantic web: information and services integration. Technical report, CONSIST Business information technology, 2005.

Modeling, Estimation and Optimal Control Design of a Biological Wastewater Treatment Process

Hicham El BAHJA and Pastora Vega

University of Salamanca, Department of Computer Science and Automatic,
Faculty of Sciences, Plaza de los Cados s/n, 37008 Salamanca, Spain
<http://informatica.usal.es>, {hicham,pvega}@usal.es

Resumen The aim of this work is to give a mathematical model, estimate and control a biological wastewater treatment process. The paper illustrates the use of a complex nonlinear model in the design of a software sensor and two predictive control techniques. Process modelling describes the complete dynamics of autotrophic and heterotrophic biomasses, biodegradable organic matters. The control approach structure is combined with the estimation algorithm, for the on-line reconstruction of unmeasured biological states and unknown parameters of the bioprocess. The efficiency of both the control and estimation schemes are demonstrated via computer simulations.

Keywords: Predictive control; Activated sludge process; Estimation; Optimization; Filtering; Dissolved oxygen control.

1. Introduction

Wastewater treatment is just one component in the urban water cycle; however, it is an important component since it ensures that the environmental impact of human usage of water is significantly reduced. It consists of several processes: biological, chemical and physical processes. Wastewater treatment aims to reduce: nitrogen, phosphorous, organic matter and suspended solids. To reduce the amount of these substances, wastewater treatment plants (WWTP) consisting of (in general) four treatment steps, have been designed. The steps are: a primarily mechanical pre-treatment step, a biological treatment step, a chemical treatment step and a sludge treatment step.

The purpose of the mechanical pre-treatment step is to remove various types of suspended solids from the incoming wastewater. To a large extent this step is meant to protect the following steps from various types of grits and larger particles. Typically, the step consists of grids that remove larger objects in the wastewater, an aerated sand filter that removes sand and a primary sedimentation unit that reduces the content of suspended solids in the wastewater by

means of sedimentation. The primary sedimentation may also remove considerable amounts of organic matter in the particulate form and, hence, reduce the need for aeration later in the process.

The aim of the biological treatment step has originally been solely to remove organic matter. However, today many wastewater treatment plants are also designed for the biological removal of nitrogen and phosphorous. The most common type of biological treatment step is based on the activated sludge process. The biological reactor contains a mixture of microorganisms suspended in wastewater; called activated sludge. The microorganisms degrade the content of organic matter in the wastewater aerobically, i.e. when air is supplied to the biological reactor. To retain the sludge in the system, the biological reactor is followed by a sedimentation unit that separates the clean effluent wastewater from the sludge. The sludge is then recycled into the biological reactor. Due to the growth of the microorganisms, sludge has to be removed from the system continuously via the sludge outtake. In this simple system, the main control handles are: aeration, sludge outtake and sludge recirculation. These variables should be controlled to ensure a suitable treatment efficiency of the process, which includes maintaining a correct amount of sludge in the system.

In the biological treatment of wastewater, the sedimentation process enables to separate the treated wastewater from the biomass sludge and produces a clear treated effluent. In addition to clarification, secondary settler tanks or clarifiers have the function of thickening the activated sludge for returning to the bioreactor and even, as storage tank. The settling process can take place in the same reactor or in a secondary settler. By all these reasons, secondary settling tanks have been considered essential and often they can be limiting factors for good removal efficiencies of the activated sludge system. Tkacs and assistants developed the first model of settling process that can consider or not possible biological reactions in the clarifier [1].

For the chemical precipitation process to function, two reactors are needed: a flocculation chamber where the chemicals are added and the flocs are formed and a sedimentation unit, which separates the flocs from the water. The precipitation process may take place at several locations in the wastewater treatment plant. In pre-precipitation plants, the process is carried out in the mechanical pre-treatment step. In simultaneous precipitation the precipitation is performed in the biological step and in post-precipitation plants the process is carried out in a separate chemical step following the biological step. These are the basic options, but others exist. Often a combination of two of these structures is used.

The purpose of the sludge treatment step is to prepare the sludge for end disposal. Anaerobic digestion is probably one of the most used processes for reducing the amount of sludge. At the same time, the digestion process produces gas, providing a significant source of energy, which is usually used at the WWTP. Sludge treatment also includes various dewatering processes, which reduce weight and volume of the sludge. Sludge treatment is gaining in importance as it becomes increasingly difficult to dispose of the sludge. Sludge disposal is in many countries becoming one of the large costs of wastewater treatment.

In this paper, we are concerned with the problem of regulating the residual pollutant substrate and the dissolved oxygen concentrations around prespecified levels inside an activated sludge for wastewater treatment process. The purpose is two-fold field of the first control action is to provide an acceptable pollutant level in the effluent while the second is to provide the biomass with the necessary and sufficient amount of energy to carry on the oxidation. As a fact, proper aeration is crucial to process efficiency, since an insufficient dissolved oxygen level would impair the oxidation process and eventually lead to biomass death, whereas a too high dissolved oxygen may cause the sludge to settle poorly. Excessive aeration is also undesirable from an economic point of view, since the oxygen in excess is simply lost to the atmosphere.

The control law is based on one step ahead prediction of the controlled variables and optimization of an appropriate quadratic cost function with *fmincon* in matlab, and is associated to an estimator for on-line tracking of unavailable states and unknown parameters of the bioprocess. The estimator uses process physics in a more direct manner to develop a nonlinear observer applicable to the estimation problem of activated sludge process.

The wastewater treatment process is modeled by differential equations derived from mass-balance considerations, which involve nonlinear terms. The most important parameter is the specific growth rate which is a complex nonlinear function of plant states and several uncertain biological parameters. In the estimation scheme the specific growth rate is considered as an unknown time varying quantity and is estimated without a priori knowledge about its kinetic model structure. In [3] the specific growth rate was reconstructed on the basis of its analytical expression via the estimation of the kinetic parameters appearing in this expression.

The proposed control scheme presents several basic features. It is based on the well-known nonlinearities of the process, it is adaptive in order to deal with the parameter uncertainty and it allows reliable information about the current state of the biochemical reaction by simultaneous states and parameter estimation.

The paper is organized as follows. The modeling of the continuous wastewater treatment is detailed in Section 2. Section 3 is dedicated to the extended Kalman filter and the asymptotic estimator presentations. The predictive control laws are then described in Section 4. In Section 5 the efficiency of the control scheme is illustrated via simulation studies. Finally Section 6 ends the paper.

2. Process modeling

The activated sludge process [5] is usually constituted by a bioreactor and a settler. The aerator is taken to be a well-stirred tank in which suspended microorganisms biochemically degrade the dissolved substrate. The suspended microorganisms are separated completely in the settler. A portion of the concentrated biomass is recycled to the bioreactor. The remainder is wasted to maintain a bounded organism level in the system. The energy required is supplied by dissolved oxygen and carbon dioxide is released in turn. We assume that no bio-

reaction takes place in the settler and the aerator is considered to be perfectly mixed so that the concentration of each component is spatially homogeneous.

The mass-balance around the aerator and the settler give the following equations:

$$\dot{X} = \mu(\cdot) X - (1 + q_r(t)D(t)) X + q_r(t)D(t)X_r \quad (1)$$

$$\dot{S} = -\frac{\mu(\cdot)}{Y_s} X - (1 + q_r(t)D(t)) S + D(t)S_{in}(t) \quad (2)$$

$$\dot{C} = -\frac{\mu(\cdot)}{Y_C} x - (1 + q_r(t)D(t)) C + K_{La}(C_s - C) \quad (3)$$

$$\dot{X}_r = (1 + q_r(t)D(t)) X - (1 + q_r(t)D(t)) X_r \quad (4)$$

Where:

- $X(t)$, $S(t)$, $X_r(t)$ and $C(t)$ are, respectively, the biomass, the substrate, the recycled biomass and the dissolved oxygen concentrations.
- $\mu(t)$ corresponds to the biomass specific growth rate. It is assumed to follow the olsson model[6]:

$$\mu(t) = \mu_{max} \frac{S(t)}{K_S + S(t)} \frac{C(t)}{K_C + C(t)} \quad (5)$$

where μ_{max} is the maximum specific growth rate, K_S is the affinity constant and K_C is saturation constant.

- $D(t)$ is the dilution rate.
- q_r and q_p represent the ratio of recycled flow to influent flow and the ratio of waste flow to influent flow.
- K_{La} represents the oxygen mass transfer coefficient. We assume that this quantity is in a linear relationship with the air flow rate $W(t)$

$$K_{La} = \alpha W(t), \alpha > 0. \quad (6)$$

- S_{in} corresponds to the influent substrate concentration.
- C_S is the maximum dissolved oxygen concentration.
- Y_S and Y_C are constants yields coefficients.

The objective in this work, we have investigated the problem of regulating the process states (S, C) around a specified setpoints (S^*, C^*) under the five assumptions:

A1: The dilution rate $D(t)$ and the aeration flow $W(t)$ are the two control variables.

A2: The yields coefficients Y_S and Y_C are knowns.

A3: The biomass, substrate and recycled biomass concentrations $(X(t), S(t), X_r)$ are unavailable on-line.

A4: A noisy measure of the dissolved oxygen concentration C is available.

A5: The control variables are bounded.

3. Estimation

The implementation of efficient modern control strategies in bioprocesses [7][9][10], highly depends on the availability of on-line information about the key biological process components like biomass and substrate. But due to lack or prohibitive cost, in many instances, of on-line sensors for these components and due to expense and duration (several days or hours) of laboratory analyses, there is a need to develop and implement algorithms which are capable of reconstructing the time evolution of the unmeasured state variables on the base of the available on-line data. However, because of the nonlinear feature of the biological processes dynamics and the usually large uncertainty of some process parameters, mainly the process kinetics, the implementation of extended versions of classical observers proves to be difficult in practical applications, and the design of new methods is undoubtedly an important research matter nowadays. In that context, Extended Kalman Filter (EKF) and the Asymptotic Estimator (AE) procedures are presented in this work.

The aim of the estimation procedure is to compute estimated values of the unavailable state variables of the process ($X(t), S(t), X_r(t)$) and the specific growth rate $\mu(t)$ using the dissolved oxygen concentration $C(t)$ as a unique measurable variable.

3.1. Extended Kalman filter

The EKF estimator uses a non-linear mathematical model of the process and a number of measures for estimating the states and parameters not measurable. The estimation is realised in three stages: prediction, observation and registration.

Let a dynamic non-linear system be characterised by a model in the state space form as following:

$$\frac{dX(t)}{dt} = f(X(t), u(t), t) + v(t) \quad (7)$$

where:

- $X(t)$ represents the state vector.
- $f(\cdot)$ non-linear function of $X(t), u(t)$.
- $u(t)$ represents the input vector.
- $v(t)$ Vector of noise on the state equation of dimension n , assumed Gaussian white noise, medium null and covariance matrix known $q(t) = Var(v(t))$.

The state of the system is observed by m discrete measures related to the state $X(t)$ by the following equation of observation:

$$Z(t_k) = h(X(t_k), t_k) + \omega(t_k) \quad (8)$$

where:

- $h(\cdot)$ the observation matrix.
- t_k the observation instant.
- $\omega(t_k)$ Vector of noise on the measure, of dimension m , independent of $v(t)$, assumed Gaussian white noise, medium null and covariance matrix known $r(t) = Var(\omega(t))$.

The EKF algorithm corresponding to the continuous process in discrete observation, where the measurements are acquired at regular intervals, is given by [4]:

1. Initialisation filtre $t = t_0$:

$$X(t_0) = E(X(t_0)) \quad (9)$$

$$L(t_0) = Var(X(t_0)) \quad (10)$$

2. Between two instant of observation:

The estimated state $\hat{X}(t)$ and its associated covariance matrix $L(t)$ are integrated by the equations:

$$\frac{d\hat{X}(t)}{dt} = f(\hat{X}(t), u(t), t) \quad (11)$$

$$\frac{dL(t)}{dt} = F(t)L(t) + L(t)F^T(t) + q(t) \quad (12)$$

$$F(t) = \frac{\partial f(\hat{X}(t), u(t), t)}{\partial \hat{X}} \quad (13)$$

Then we have, before the observation at $t = t_{k-}$, an estimated of $\hat{X}(t_{k-})$ and its covariance matrix $L(t_{k-})$.

3. Updating the gain:

$$K(t_k) = L(t_{k-})H^T(\hat{X}(t_{k-}), t_k) \left[H(\hat{X}(t_{k-}), t_k)L(t_{k-})H^T(\hat{X}(t_{k-}), t_k) + r(t_k) \right]^{-1} \quad (14)$$

4. Update of the estimated state:

$$\hat{X}(t_k) = \hat{X}(t_{k-}) + K(t_k) \left[Z(t_k) - h(\hat{X}(t_{k-}), t_k) \right] \quad (15)$$

5. Update of the covariance matrix:

$$L(t_{k+}) = L(t_{k-}) - K(t_k)H(\hat{X}(t_{k-}), t_k)L(t_{k-}) \quad (16)$$

$$H(\hat{X}(t_{k-})) = \frac{\partial h(\hat{X}(t_{k-}), t_k)}{\partial \hat{X}(t_{k-})} \quad (17)$$

3.2. Asymptotic estimator

The estimator proposed by Bastin and Dochain [11] is based on reasoning stability first. It was developed for a particular class of nonlinear systems and especially suitable for biotechnological processes. It allows partial or complete estimation of the bioprocess state.

The asymptotic estimator (AE) [11] deals with state and parameter estimation in activated sludge process. It consists of an adaptive observer which combines a priori knowledge of the physical bioprocess with the experimental data.

The dynamics of the bioprocess (Eqs. (1)-(4)) are rewritten in the following manner:

$$\frac{d\xi(t)}{dt} = K\varphi(X) - D^*(t)X(t) + U \quad (18)$$

where:

$$\xi = \begin{pmatrix} X \\ S \\ X_r \\ C \end{pmatrix}; K = \begin{pmatrix} 1 \\ \frac{-1}{Y_S} \\ 0 \\ \frac{-1}{Y_C} \end{pmatrix}; U = \begin{pmatrix} 0 \\ D(t)S_{in} \\ 0 \\ K_{La}C_s \end{pmatrix}$$

$$D^* = \begin{pmatrix} (1 + q_r(t))D(t) & 0 & -q_r D(t) & 0 \\ 0 & (1 + q_r(t))D(t) & 0 & 0 \\ -(1 + q_r(t))D(t) & 0 & 0 & (q_p + q_r(t))D(t) \\ 0 & 0 & 0 & K_{La} + (1 + q_r)D(t) \end{pmatrix} \xi(t) \text{ re-}$$

presents the state vector, $\varphi(t)$ is the vector of reaction kinetics, K corresponds to yield coefficients matrix and $U(t)$ is the feed rate vector.

Let us consider the following partition $\xi = \begin{bmatrix} \xi_a \\ \xi_b \end{bmatrix}$ with the corresponding matrices and vectors (D_a^*, K_a, U_a) and (D_b^*, K_b, U_b) and such that K_a is full rank.

$$\xi_a = S, K_a = [\frac{-1}{Y_S}], U_a = DS_{in}, D_a^* = D = D(1 + q_r)$$

$$\xi_b = \begin{bmatrix} X \\ X_r \\ C \end{bmatrix}, K_b = \begin{bmatrix} 1 \\ 0 \\ \frac{-1}{Y_C} \end{bmatrix},$$

$$D_b^* = \begin{bmatrix} D(1 + q_r) & -q_r D & 0 \\ -D(1 + q_r) & D(q_p + q_r) & 0 \\ 0 & 0 & D(1 + q_r) + K_{La} \end{bmatrix}.$$

Eq. (18) can then rewritten as follows:

$$\frac{d\xi_a}{dt} = K_a\varphi - D_a^*\xi_a + U_a \quad (19)$$

$$\frac{d\xi_b}{dt} = K_b\varphi - D_b^*\xi_b + U_b \quad (20)$$

We can introduce the following state transformation:

$$Z(t) = A_0 \xi_a(t) + \xi_b(t) \quad (21)$$

Where A_0 is the solution of the matrix equation:

$$A_0 K_a + K_b = 0 \quad (22)$$

The auxiliary state vector obtained is:

$$Z = [Z_1 Z_2 Z_3]^T = [Y S + X X_r - K_0 S + C]^T$$

The dynamics of which are governed by the following differential equation:

$$\frac{dZ}{dt} = -A_0 D_a^* \xi_a + A_0 U_a - D_b^* (Z - A_0 \xi_a) + U_b \quad (23)$$

The bioprocess dynamics can then be equivalently represented by Equations (19) and (23).

The vector Z can explicitly be rewritten as a linear combination of the vector $\xi_m(t) = C(t)$ and $\xi_e = [X(t)S(t)X_r(t)]$ of respectively measured and no measured state variables in the following manner:

$$Z(t) = A_1 \xi_m(t) + A_2 \xi_e(t) \quad (24)$$

Then an asymptotic observer for the reconstruction of the unmeasured state variables is given by:

$$\frac{d\widehat{Z}}{dt} = -(A_0 D_a^* + D_b^* A_0) \widehat{\xi}_a - D_b^* \widehat{Z} + A_0 U_a + U_b \quad (25)$$

$$\frac{d\widehat{\xi}_a}{dt} = \frac{-1}{Y_S} \mu(\widehat{\xi}_a) \widehat{Z}_1 + \mu(\widehat{\xi}_a) \widehat{\xi}_a - D_a^* \widehat{\xi}_a + D S_{in} \quad (26)$$

$$\widehat{\xi}_e = A_2^+ (\widehat{Z} - A_1 \xi_m) \quad (27)$$

4. Control

The control objective is to make the residual substrate concentration and the dissolved oxygen concentration to track the states of a given reference model despite the disturbances and system parameter uncertainties by acting on the dilution rate and the aeration flow rate.

4.1. One step predictive control method

Using the Euler formula with a sampling time period T_s , the model equation (2) (3) are discretized to yield a one-step-ahead predictor as follows:

$$S^p(k+1) = \widehat{S}(k)[1 - T_s D(k)(1+r)] + (-\mu(k)/Y_S)\widehat{X}(k) + T_s D(k)S_{in} \quad (28)$$

$$C^p(k+1) = C(k)[1 - T_s D(k)(1+r)] + T_s(-\mu(k)/Y_C)\widehat{X}(k) + T_s(D(k)C_{in} - \alpha W(k)(C_S - C^m(k))) \quad (29)$$

where $\widehat{S}(k)$, $\widehat{X}(k)$ and $\widehat{\mu}(k)$ are estimated values of S , X , μ . They are updated by the estimator algorithm and provided to the controller according to the certainty equivalence principle. $C^m(k)$ is the measurement value of the dissolved oxygen concentration.

let the reference model squaring with asymptotic attenuation of the regulation error at a rate defined by the control gain parameters g_s and g_c be:

$$S^{mdr}(k+1) = S^* + (1-g_s)[\widehat{S}(k) - S^*], C^{mdr}(k+1) = C^* + (1-g_c)[C(k) - C^*] \quad (30)$$

with $0 < g_s < 1$ and $0 < g_c < 1$

The control is therefore computed by minimizing the following quadratic cost function:

$$Min_{D(k), W(k)} \left\{ J = \left\| \begin{bmatrix} S^p(k+1) \\ S^p(k+1) \end{bmatrix} - \begin{bmatrix} S^{mdr}(k+1) \\ S^{mdr}(k+1) \end{bmatrix} \right\|_{\Pi}^2 + \left\| \begin{bmatrix} D(k) - D(k-1) \\ W(k) - W(k-1) \end{bmatrix} \right\|_{\Lambda}^2 \right\} \quad (31)$$

where Π and Λ are weighting diagonal matrices.

4.2. Predictive control

As in the paragraph below [12], the control objective is to make the dissolved oxygen and the residual substrate concentration to track the states of a given set point. Using the Euler formula with a sampling period T_s , all the model equations following are discretized :

$$X(k+H_p) = T_s \mu(\cdot)X(k) + (1 - T_s(1+q_r(k))D(k))X(k) + T_s q_r D(k)X_r(k) \quad (32)$$

$$S(k+H_p) = -\frac{T_s \mu(\cdot)X(k)}{Y_S} + (1 - T_s(1+q_r(k))D(k))S(k) + T_s D(k)S_{in}(k) \quad (33)$$

$$C(k+H_p) = -\frac{T_s \mu(\cdot)X(k)}{Y_C} + (1 - T_s(1+q_r(k))D(k))C(k) + T_s K_{La}(C_S - C(k)) \quad (34)$$

$$X_r(k+H_p) = T_s(1+q_r(k))D(k)X(k) + T_s(q_p(k) + q_r(k))D(k)X_r(k) \quad (35)$$

Where: H_p Is the horizon of prediction.

Let S_{sp} and C_{sp} are the reference concentrations of the substrate and the dissolved oxygen respectively. Where:

$$S_{sp} = [r_s(k+1), \dots, r_s(k+H_p)]^T, C_{sp} = [r_c(k+1), \dots, r_c(k+H_p)]^T \quad (36)$$

The control objective is to minimize the squared deviation between the predictor output (32-35) and the reference model output (36).

The control is therefore computed by minimizing the following quadratic cost function:

$$\begin{aligned} Min_{D(k), W(k)} J = & \left\| \begin{bmatrix} S^p \\ S^p \end{bmatrix} - \begin{bmatrix} S_{sp} \\ S_{sp} \end{bmatrix} \right\|_{\Pi}^2 + \\ & \left\| \begin{bmatrix} D(k)D(k+1)\dots D(k+H_c) \\ W(k)W(k+1)\dots W(k+H_c) \end{bmatrix} - \begin{bmatrix} D(k-1)D(k)\dots D(k+H_c-1) \\ W(k-1)W(k)\dots W(k+H_c-1) \end{bmatrix} \right\|_{\Lambda}^2 \end{aligned} \quad (37)$$

where:

- H_c is the control horizon.
- $D = [D(k)D(k+1)\dots D(k+H_c)]$ input (dilution rate) increment vector.
- $W = [W(k)W(k+1)\dots W(k+H_c)]$ input (Air flow rate) increment vector.
- $S^p = [S(k)S(k+1)\dots S(k+H_p)]$ output (substrate) increment vector.
- $C^p = [C(k)C(k+1)\dots C(k+H_p)]$ output (dissolved oxygen) increment vector.
- Π and Λ are weighting diagonal matrices.

5. Simulation and results

The simulation results are obtained by using a fourth order Runge-Kutta with the same typical values of process and kinetic parameters $Y_S = 0,65$, $Y_C = 1,3$, $q_r = 0,6$, $q_p = 0,2$, $\alpha = 0,018m^{-3}$, $C_s = 10mgl^{-1}$, $S_{in} = 200mgl^{-1}$. The design parameters of the control law parameters are: $D_{min} = 0,02h^{-1}$, $D_{max} = 0,15h^{-1}$, $W_{max} = 300m^3h^{-1}$, $W_{min} = 0m^3h^{-1}$, $T_s = 1min$, $g_s = 0,7$, $g_c = 0,85$, $\Pi = I$, $\Lambda = diag(0,8e^{-6})$. Simulation of the first algorithm without estimation (figure 1) but the (figure 2) with the EKF estimator. Simulation of the second algorithm without estimation (figure 3) but the (figure 4) with the EKF estimator. Simulation with the EKF estimator.

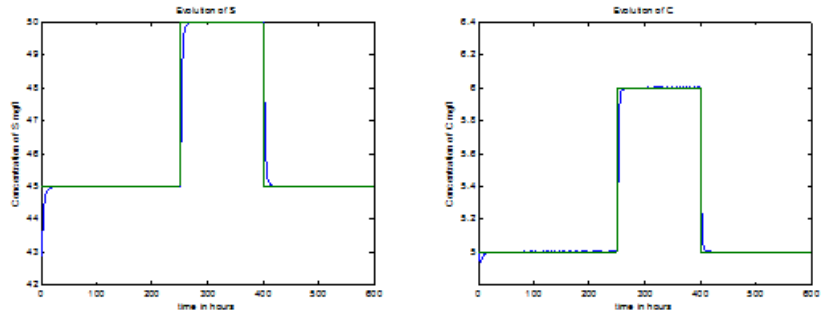


Figure 1. evolution of the substrate and the dissolved oxygen concentrations with a variables set point.

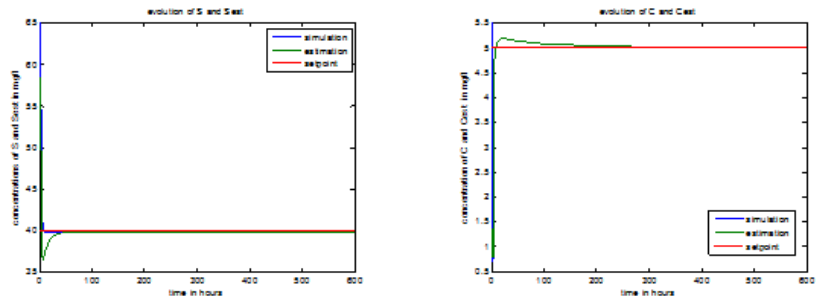


Figure 2. evolution of the substrate and the dissolved oxygen concentrations.

6. Conclusion

This study is concerned with the design of evaluating a suitable controller and obtaining real time information about the biological activity in an activated sludge wastewater treatment process. The controls laws are based on direct exploitation of the nonlinear model of the wastewater treatment process and are coupled with an EKF estimator for on-line tracking of unavailable states and time varying parameters. The predictive controllers are used to maintain the residual substrate concentration and the dissolved oxygen concentration by acting on the dilution rate and the air flow rate. The performances of the control and estimation strategies are evaluated by simulation studies.

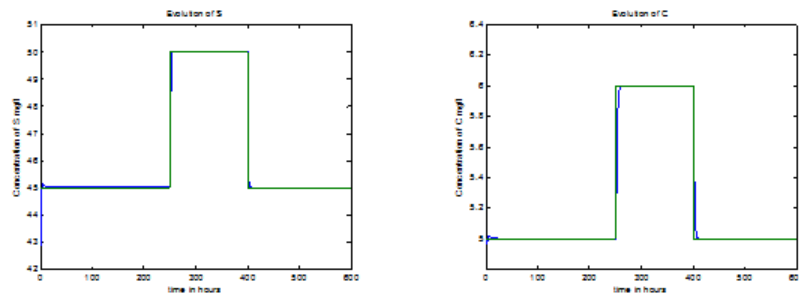


Figure 3. evolution of the dissolved oxygen and the substrate concentrations with a variable references.

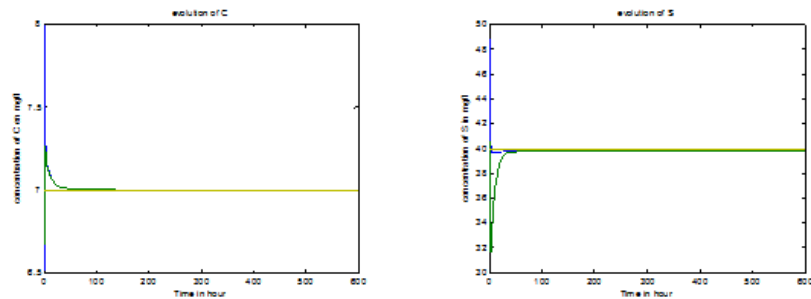


Figure 4. evolution of the dissolved oxygen with the reference 7 and the substrate with reference 40 concentrations.

Referencias

1. Takacs I, G. G. Patry and D. Nolasco (1991) A dynamic model of the clarification-thickening process. *Water Research* 25(10), pp. 1263-127
2. Brandt J, Hein W (2001) Polymer materials in joint surgery. In: Grellmann W, Seidler S (eds) *Deformation and fracture behavior of polymers*. Engineering materials. Springer, Berlin Heidelberg New York
3. F.Nejjari, *Modlisation, Estimation et commande d'un bioproc de traitement des eaux uses*, Thesis Report, Faculty of Sciences, Marrakesh, Morocco, June 1997.
4. Labarrere, M., Krief, J.P. ET Gimonet, B. (1982). *Le filtrage et ses applications*. Cepadues Edition, Toulouse.
5. J.F. Andrews Ross DW (1977) *Lysosomes and storage diseases*. MA Thesis, Columbia University, New York
6. Olsson, G and Piani, G. (1992). *Computer systems for automation and control*. Prentice Hall International Ltd., UK, ISBN 0-13-45781-4.
7. B. Dahhou, G. Roux, G. Chamiltoris, *Modelling and adaptive predictive control of a continuous fermentation process*, *Appl. Math. Modelling* 16 (1992) 545-552.
8. Ross DW (1977) *Lysosomes and storage diseases*. MA Thesis, Columbia University, New York

9. D. Dochain, Design of adaptive controllers for nonlinear stirred tank bioreactors: extension to the MIMO situation, *J. Proc. Cont.* 1 (1991) 41-48.
10. F. Nejjari, A. Benhammou, B. Dahhou, G. Roux, Nonlinear multivariable control of a biological wastewater treatment process, in: *Proceedings of ECC 97*, Brussels, Belgium, 1-4 July 1997.
11. BASTIN, G., AND DOCHAIN, D. *On-line Estimation and Adaptive Control of Bioreactors*. Elsevier, Amsterdam, 1990.
12. Ashley RM, Wotherspoon DJJ, Petrie M, Jack A (1997) *Modelling to Assist with the Evaluation of the Holistic Performance of Sewers and Wastewater Treatment Plants*. CIWEM Conference on Integrated Modelling 97. Ballyonnell, Irland, October 1997.

MAIM: Multiagent Architecture Incidents Management

Rosa María Martín Moreno¹
Juan Manuel Corchado Rodríguez²

Dpto. Informática y automática
Plaza de los Caídos S/N 37008, Salamanca, España

¹rosa.martin@red.es

²corchado@usal.es

Abstract This report presents an architecture of agents for the process of incidents management. This architecture allows the automatic classification of incidents taking into account its category and priority. It also provides a meeting point for the different actors involved in resolving incidents. This way we intend to maximize the number of incidents handled by each operator and minimize the resolution time, increasing customer satisfaction. The architectural design has been based upon the methodology oriented by GORMA organizations.

- Keywords: Incidents Management, arquitectura, multiagent, GORMA

Resumen En este artículo se presenta una arquitectura de agentes para el proceso de gestión de incidencias microinformáticas. Esta arquitectura permite la clasificación automática de las incidencias teniendo en cuenta su categoría y prioridad, y proporciona un punto de encuentro entre los diferentes actores que intervienen en la resolución de las incidencias. De esta forma se pretende maximizar el número de incidencias resueltas por cada operador y minimizar el tiempo de resolución, aumentando la satisfacción de los clientes. Para el diseño de la arquitectura se ha tomado como base la metodología orientada a organizaciones GORMA.

- Palabras claves: Gestión de incidencias, arquitectura, multiagentes, GORMA

1. Introducción

MAIM se presenta como una arquitectura multiagente para la Gestión de Incidencias microinformáticas, que permita mejorar los tiempos de resolución de las mismas y por tanto, la satisfacción de los clientes.

Tanto los individuos como las organizaciones dependen cada vez más de la Informática para alcanzar sus objetivos. Esta dependencia en aumento ha dado como resultado una necesidad creciente de servicios informáticos de calidad, que se correspondan con los objetivos del negocio y que satisfagan los requisitos y

las expectativas del cliente. a través de los años, el énfasis ha pasado de estar sobre el desarrollo de las aplicaciones a la gestión de servicios. Las aplicaciones y elementos hardware de comunicación contribuyen a que sus usuarios alcancen sus objetivos, y en caso de fallos, éstos son soportados por los procesos de mantenimiento y operaciones.

En este contexto, nace la necesidad de gestionar los fallos o incidencias de una forma eficiente, para que se minimice el tiempo en el que los elementos que dieron fallo se encuentren sin dar servicio y por tanto, aumente la satisfacción de los usuarios.

Normalmente, cuando se produce una incidencia se acaba resolviendo, de una u otra forma, pero este proceso suele ser largo y tedioso, sin ser un proceso formal, ni documentado y no suele resolverse siempre de la misma manera. Los usuarios cada vez con más frecuencia demandan un mantenimiento de alta calidad, eficiente y continuo e independiente de su localización geográfica.

La Gestión de Incidencias es responsable de tratar la incidencia, registrarla, clasificarla, identificar los roles necesarios para resolverla e informar al cliente / usuario sobre el progreso de la misma. actualmente, para este proceso se utilizan herramientas de gestión de incidencias que requieren de la intervención manual de los operadores, la incidencia se registra, clasifica y se asigna al operador correspondiente de forma manual, interviniendo la subjetividad, lo que puede producir errores que alargan la vida de la incidencia.

La resolución de las incidencias normalmente se lleva a cabo a través de empresas privadas que realizan el mantenimiento remoto, in-situ, de la conectividad y la sustitución de hardware cuando este se estropea. Por tanto, para cada centro están implicados una gran variedad de operarios y de empresas de mantenimiento.

La investigación en este tema surge a raíz de la relación del autor con el proceso de Gestión de Incidencias de los programas de fomento de la sociedad de la información de Red.es. Teniendo en cuenta la experiencia del autor junto con un análisis inicial de los indicadores de resolución de incidencias por parte de los operadores, se deduce que existe un retardo considerable en su resolución debido a errores provocados por una gestión manual, lo que provoca una insatisfacción de los clientes.

Si una incidencia se clasifica mal y se escala de forma errónea, provoca que un operador que no tiene los medios suficientes para solucionarla, dedique tiempo a esa incidencia. En este caso se está retrasando su resolución y ocupando un recurso que podría ser utilizado para resolver otra incidencia, por tanto, se está perdiendo el dinero asociado a las horas de ese operador.

además de los errores manuales, existen otros dos elementos que provocan el retraso en la resolución de las incidencias. El primero de ellos es la deficiencia de la planificación de los operarios. a la hora de planificar los recursos hay que tener en cuenta el número de operadores, la prioridad de las incidencias y la localización de las mismas en el caso de que se tengan que desplazar.

El segundo elemento que influye en el retraso es el tener que contactar con el usuario después de escalada la incidencia al operador. El centro de atención a

usuarios recoge la incidencia y es el operador el que se pone en contacto con el usuario para concertar una cita en la que se solucionará la incidencia.

Por otra parte, la insatisfacción de los clientes puede aumentar por la falta de información sobre el tiempo que se tardará en resolver la incidencia y el estado de la misma en cada momento.

Teniendo en cuenta esta problemática se decide centrar la investigación en el estudio de una arquitectura de agentes que permita automatizar el proceso de gestión de incidencias, reduciendo el tiempo de resolución y por tanto aumentando la satisfacción de los clientes

2. Gestión de Incidencias

Se puede definir una incidencia como cualquier suceso que no forma parte del funcionamiento estándar de un servicio y que causa o puede causar una interrupción o degradación de la calidad de ese servicio [1]

Gestión de incidencias es el proceso responsable de la gestión del ciclo de vida de todos los incidentes. El objetivo de este proceso es recuperar el servicio de TI (Tecnologías de la Información) para los usuarios afectados lo antes posible. [4]

Las actividades principales de la gestión de incidencias son las siguientes:

- Detección y registro de las incidencias. Todas las incidencias deben de registrarse, siendo la automatización el método idóneo
- Clasificación. Proceso de identificación de la razón de la incidencia y la correspondiente acción de resolución
- Comparación. Búsqueda de la incidencia en la base de datos de conocimiento (KDB) para determinar la solución definitiva o temporal
- Escalado. Una vez clasificada la incidencia se determina cómo resolverla y si es necesario se escala al operador adecuado
- Resolución y recuperación. Se resuelve la incidencia y se recupera el servicio.
- Cierre de la incidencia. Confirmación de la resolución de la incidencia y cierre de la misma.

2.1. Mejores prácticas para la gestión de incidencias

Tanto ITIL (Information Technology Infrastructure Library) como COBIT (Control Objectives for Information and related Technology) se definen como las mejores prácticas para la gestión de Servicios Informáticos, dentro de esta gestión se encuentra la Gestión de Incidencias.

ITIL fue desarrollada a finales de 1980 como guía para el gobierno de Reino Unido y se ha convertido en un estándar de facto en la Gestión de Servicios Informáticos. ITIL aborda la gestión de todos los procesos relacionados con los servicios de Tecnologías de la información, entre los que se encuentra la gestión de incidencias.

COBIT es una guía para la gestión de los Sistemas de Información en las organizaciones, desarrollado por la ISaCa (Information Systems audit and Control

association) y el ITGI (IT Governance Institute).

COBIT es un conjunto de mejores prácticas para la seguridad, la calidad, la eficacia y la eficiencia en TI que permiten alinear TI con el negocio, identificar riesgos, entregar valor al negocio, gestionar recursos y medir el desempeño, el cumplimiento de metas y el nivel de madurez de los procesos de la organización.

3. Metodologías de sistemas multiagentes

El desarrollo de los sistemas multiagentes tradicionalmente se ha realizado a partir de un diseño centrado en el estudio de cada agente de forma aislada e independiente, orientado hacia su estructura y comportamiento. La integración con otros agentes se llevaba a cabo mediante plataformas de localización de agentes a través de servicios de direccionamiento y descripción de los mismos, tal y como propone el estándar FIPA . Otro sistema utilizado es el que propone la plataforma RETINSA [6] mediante agentes intermedios que buscan los mejores proveedores de servicios a los clientes.

Recientemente este enfoque ha cambiado y el diseño se está dirigiendo hacia los aspectos organizativos de los agentes donde empiezan a aparecer conceptos como grupo, roles, funciones, etc. estableciéndose dos niveles descriptivos: el de la organización y el del agente [5].

Debido a este nuevo enfoque se habla de dos metodologías de desarrollo diferentes: orientada a la organización y orientada al agente.

3.1. GORMAS

GORMAS (Guidelines for Organization-based Multiagent Systems) es una guía metodológica para el diseño de sistemas multiagentes abiertos desde la perspectiva de las organizaciones humanas [2].

Las organizaciones resultan ser un medio efectivo para la coordinación de las actividades, no sólo de los humanos sino también de los agentes.

El concepto de organización ha adquirido una gran importancia en el área de los sistemas multiagente, al facilitar el análisis y diseño de mecanismos de coordinación y colaboración en sistemas abiertos.

Esta guía se basa en la Teoría de Organización y en el Modelo de Organización. Esta teoría aborda las propiedades de las organizaciones humanas y sus aspectos de diseño, mientras que el Modelo de Organización describe los principales aspectos de las organizaciones: estructura, funcionalidad, normalización, dinamicidad y entorno. Este modelo consta de un conjunto de meta-modelos que extienden las propuestas de INGENIAS y ANEMONA, empleando fundamentalmente los conceptos de unidad organizativa, servicio, norma y entorno. además, contiene un conjunto de patrones de diseño, con los que facilitar el modelado de la estructura de la organización.

GORMAS consta de un conjunto de fases que cubren el análisis, el diseño de la

estructura organizativa y el diseño de la dinámica de la organización. Con estas fases se especifica cuáles son los servicios que ofrece la organización, cuál es su estructura interna y qué normas rigen su comportamiento.

4. MAIM: Multiagent architecture for Incident Management

En este artículo se propone una arquitectura de agentes para el proceso de Gestión de Incidencias. Esta arquitectura se centra en las organizaciones, basándose en la metodología de GORMAS. Teniendo en cuenta esta arquitectura, el sistema permitiría registrar una incidencia, clasificarla de forma automática y buscar el operador que mejor la resolviera. El usuario de forma inmediata podría elegir la fecha en la que desea que se resolviera y conocería el operador que le iría a atender.

Por su parte, la unión de las diferentes empresas al sistema les permitiría una mejor planificación de sus recursos, lo que da lugar a un mayor número de incidencias resueltas, un aumento en la productividad y por tanto, en los beneficios. Inicialmente se determinan cuales son los servicios y productos que debe ofrecer el sistema, los objetivos globales que persigue (misión), los grupos de interés y la información que requiere, como se muestra en la figura 1.

CGIS representa la unidad organizativa principal del sistema, sus objetivos principales son maximizar la cantidad de incidencias resueltas y minimizar el tiempo que se tarda en resolverlas. Para ello, se ofrecen cuatro servicios principales: clasificar la incidencia, buscar un operador que la resuelva, reservarlo y facturar la intervención. Estos servicios requieren de información que les proporciona cada uno de los proveedores. El producto obtenido es la reserva de los operadores necesarios para resolver la incidencia en una fecha determinada(cita).

Para determinar el tipo de organización de agentes más adecuada se realiza un análisis de las dimensiones organizativas que definen las características intrínsecas de la organización. Estas dimensiones son las siguientes:

- **Departamentalización: Divisional.**
La agrupación se realiza principalmente en base a los operadores que intervienen en la resolución de la incidencia (Helpdesk, operadores in-situ, operadores de conectividad, compañías de hardware), teniendo en cuenta que inicialmente hay que realizar una clasificación de incidencias.
- **Especialización: ampliación horizontal y vertical.**
Los roles proporcionan servicios poco relacionados entre sí, con interdependencia secuencial y recursos compartidos.
Los operadores deciden cómo implementar sus propios servicios.

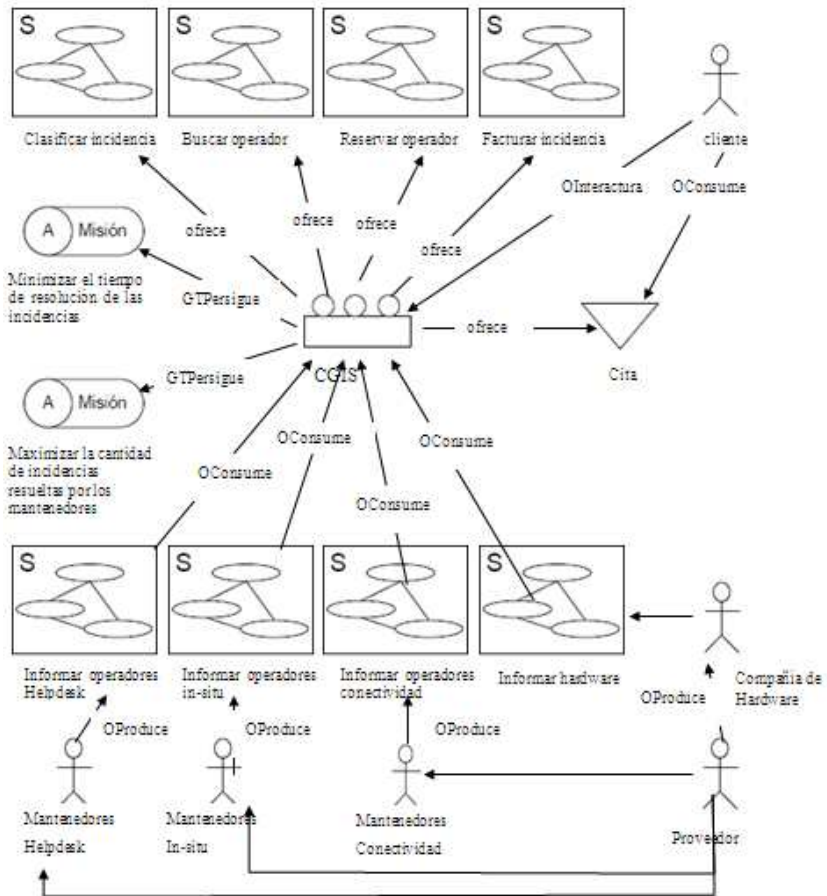


Figura 1. Diagrama del modelo de organización

- **Coordinación y Formalización:** adaptación mutación y normalización de tareas. Se necesitan mecanismos de negociación para concretar las características específicas de la ejecución de las tareas, ya que existe ampliación vertical y la tecnología de producción es de lotes.
 La normalización es necesaria para garantizar un orden preciso en la invocación de los servicios y un orden en la ejecución de las tareas internas.

Teniendo en cuenta estas dimensiones, la congregación se considera el tipo de organización más adecuada y sus características se muestran en la figura 2 de la vista estructural.

Las congregaciones se forman entre agentes interesados que presentan características similares o bien complementarias sobre un área de interés común. Los agentes buscan maximizar su utilidad local a largo plazo, en función de su interacción con los agentes de la congregación. Los agentes no tienen el concepto de grupo como tal y no les importan los objetivos que persigan los otros agentes, aunque sí dependen de otros para satisfacer sus propias necesidades.[2]

La unidad CGIS contiene tantas unidades como tipos de operadores, los roles cliente y proveedor se refinan dentro de cada nueva unidad. Se identifica un nuevo servicio Registrar Unidad. asociado a la unidad principal CGIS, este servicio permite crear dinámicamente unidades nuevas.

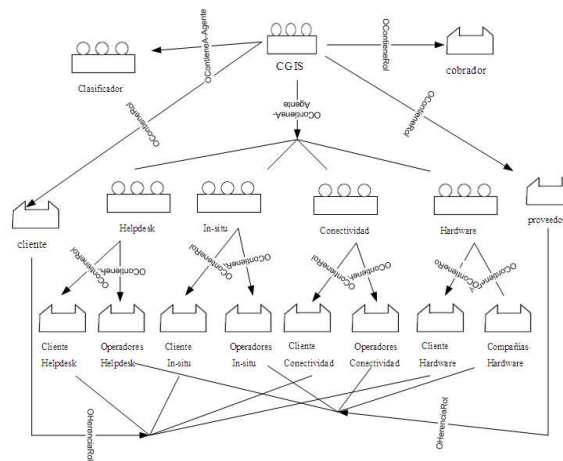


Figura 2. Vista estructural

En cada una de las unidades organizativas relacionadas con los operadores se incluye un rol propio de GestorUnidad, que proporciona los servicios de adquirir Rol, Dejar Rol e Informar Unidad.

En la figura 3 se muestra el modelado interno de la unidad Helpdesk, su relación con la unidad principal CGIS y el resto de unidades.

Para realizar la apertura del sistema hay que determinar que funcionalidad se realiza por agentes internos y cuál se publica para que sea proporcionada por agentes externos.

Los agentes internos se desarrollan por el diseñador del sistema y tiene roles no accesibles, por tanto, su funcionalidad es exactamente la requerida. Por otro lado, los agentes externos son agentes que participan en la organización pero que su procedencia es heterogénea, en estos casos es necesario que adopten roles accesibles a las unidades y que su comportamiento sea controlador a través de normas.

La funcionalidad a publicar es la de los servicios buscar Operador Helpdesk, buscar Operador Conectividad, buscar Operador In-situ, buscar Componente Hardware, reservar Operador Helpdesk, reservar Operador Conectividad, reservar Operador In-situ, reservar Componente Hardware.

Para representar la publicación de los servicios es necesario definir la unidad Directorio de Servicios, del tipo Puerto Servicio, relacionada con la unidad CGIS. a través de ella se accede a los perfiles de los servicios mediante la relación Iaccede. De esta forma los agentes externos pueden obtener información sobre qué realiza el servicio para hacer uso del servicio (cliente) o implementar su funcionalidad (proveedor).

Los roles clientes y proveedores son accesibles por agentes externos (incluidos los roles relacionados mediante la relación OHerenciaRol), mediante la adquisición del rol adecuado. Por otro lado, los roles Gestor Unidad y Cobrador se asocian a agentes internos, por tanto, no son accesibles de forma externa.

En la figura 4 se muestra el modelo de organización indicando los agentes externos (a-agentes) y los internos (agentes), junto con los roles que juegan.

4.1. Sistema de penalizaciones/recompensas

Para potenciar los objetivos del sistema se utilizan penalizaciones o recompensas, interesa que los proveedores atiendan las incidencias lo más rápido posible y que esa atención esté asociada con la prioridad de la incidencia. Normalmente con los proveedores de mantenimiento se llegan a acuerdos de nivel de servicio, estos acuerdos indican el máximo tiempo que se tiene que tardar en atender un determinado tipo de incidencia. Si este tiempo se supera se debe penalizar al proveedor.

Por tanto, el sistema debe tener en cuenta el tiempo que se tarda en resolver una incidencia por cada proveedor (empresa a la que se le asocia la incidencia) y establecer los SLas (Service Level agreement) para cada uno de ellos. Evidentemente los tiempos en resolver una incidencia por un operador remoto deben ser bastante inferiores a los tiempos de un operador in-situ, ya que en el segundo caso se tiene que tener en cuenta el desplazamiento.

La forma de penalizar a los proveedores cuando no cumplan los SLas es disminuyendo el número de incidencias asignadas, por tanto, disminuyendo los beneficios obtenidos. Según recuperen los niveles de servicio se les irá asignando un mayor número de incidencias, aumentando el beneficio. Por tanto, aquellos proveedores

estructura definida y unas normas sociales.

La arquitectura propuesta se basa en la metodología de organizaciones GORMAS, lo que permite definir los objetivos de la organización, su estructura, su dinámica y el entorno. Permite diseñar un sistema abierto donde las empresas que ofrecen servicios de mantenimiento microinformático se pueden adherir para obtener clientes que requieran la resolución de incidencias. Por su parte, los clientes se benefician de este sistema viendo aumentada su satisfacción al disminuir el tiempo de resolución de sus incidencias y al conocer de forma inmediata cuándo se van a solucionar.

Si se hace un análisis del gran volumen de incidencias registradas en centros de acceso a internet durante un año, como se muestra en la figura 5, nos podemos hacer una idea de la importancia de la investigación en este campo.

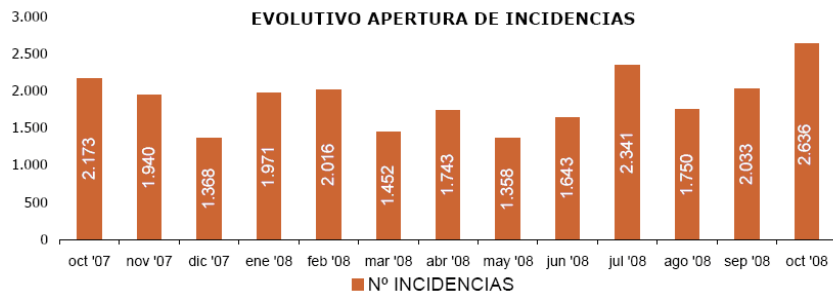


Figura 5. Número de incidencias registradas

Referencias Bibliográficas

1. *Fundamentos de Gestion de Servicios TI basado en ITIL*. Van Haren Publishing, 2007.
2. E. argente Villaplana. Gormas: Guías para el desarrollo de sistemas abiertos basados en organizaciones. Technical report, Universidad Politécnica de Valencia, 2008.
3. F. Dignum, V.; Dignum. a logic for agent organizations. In *Formal approaches to Multi-agent Systems (FaMaS)*. Multi-agent Logics, Languages and Organizations Federated Workshops (MaLLOW), 2007.
4. ITIL. *Glosario de términos ITIL, definiciones y acrónimos*, Mayo 2006.
5. J. S. O. Boissier, J.F. Hubner. Organization oriented programming, from closed to open organizations. In *Engineering Societies in the agents World VI, Sixth International Workshop, ESaW06*, 2007.
6. K. Sycara. Multi-agent infrastructure, agent discovery, middle agents for web services and interoperation. In *9th ECCAI advanced Course, EaSS01*, pages 17–49. Springer, 2001.

Sistema Bioinspirado de Reconocimiento Fonético en Entornos Ruidosos

Javier Martínez Elicegui

Departamento de Informática y Automática
Universidad de Salamanca
<http://diaweb.usal.es>

Resumen Este trabajo de investigación está enfocado hacia un tratamiento bioinspirado de la señal acústica y de esta forma, obtener una mejor discriminación frente al ruido en los sistemas automáticos de reconocimiento del habla.

En este trabajo ponemos en contacto dos áreas de investigación que han seguido habitualmente caminos separados. Por un lado, las investigaciones relacionadas con el desarrollo de modelos computacionales de la fisiología auditiva aplicados en los implantes cocleares, y por otro lado, las investigaciones para el reconocimiento automático del habla.

actualmente los reconocedores de voz se aplican directamente a la señal percibida por micrófonos, sin aplicarles las transformaciones no lineales de la señal que se producen en el oído humano. Basándonos en el hecho de que personas con lesiones leves en el oído reconocen el habla relativamente bien en entornos limpios de ruido pero tienen problemas auditivos muy superiores al resto de la población en entornos con ruido, queremos demostrar que aplicando modelos del comportamiento del oído humano en una fase previa al reconocimiento del habla, podemos mejorar significativamente su calidad en entornos ruidosos.

Palabras Clave: ASR: Automatic Speech Recognition, Oído Computacional, Entornos con Ruido.

1. Introducción

Los reconocedores de habla se han extendido extraordinariamente en los últimos años para la atención telefónica automática. Sin embargo, estos sistemas de reconocimiento incrementan de forma alarmante la tasa de errores y resultan ineficaces cuando la relación señal-ruido (SNR) es pobre.

A medida que se ha ido generalizando el uso habitual del teléfono móvil, cada vez es más normal hacer llamadas a sistemas de reconocimiento automático desde entornos con ruido (cafetería, coche, calle, etc.), situaciones muy diferentes al entorno de laboratorio en el que se han diseñado y configurado los algoritmos de reconocimiento. Debido a esta circunstancia, se han ido enriqueciendo los algoritmos de reconocimiento con toda una serie de métodos para filtrar las señales de ruido, pero con resultados a día de hoy bastante limitados.

La propuesta que presentamos en este documento se basa en un nuevo enfoque para resolver este problema. Se ha comprobado que la inteligibilidad del

habla por parte de personas con dao en el oído interno se aproxima a lo normal en ambientes silenciosos, pero sin embargo, esta percepción se deteriora más rápidamente que para las personas con audición normal a medida que el ruido de fondo aumenta [8]. Del mismo modo, los índices de inteligibilidad del habla de los usuarios de implantes cocleares son casi normales para SNR altas pero se reducen considerablemente al disminuirla [13].

Esto sugiere que las características del habla sufren importantes transformaciones en cómo son preprocesadas por el oído y cómo son trasladadas al nervio auditivo. Asimismo, sugiere que la eficacia de los sistemas que utilizan procesadores de voz mejorarían considerablemente, sobre todo en ambientes ruidosos, si sus procesadores incorporasen características de procesamiento del sistema auditivo periférico. Sobre esta hipótesis, el objetivo de esta investigación es producir un sistema de reconocimiento de vocabularios que incorpore un preprocesado basado en los modelos computacionales del sistema auditivo [4][5][6].

Para la preparación de este trabajo de Tesis de Máster se ha analizado abundante información de áreas tan complejas y extensas como son la física, la psicoacústica y los modelos computacionales de reconocimiento del habla. Ha sido especialmente complejo adquirir los conocimientos base que nos permiten comprender la abundante formulación matemática que a menudo aparece en las publicaciones.

Quiero dedicar especial mención al libro *"Signal, Sound and Sensations (Modern Acoustics and Signal Processing)"* de William M. Hartmann [1], por su extensa y formidable recopilación de conceptos, también ha sido de gran ayuda el libro *"An Introduction to the Psychology of Hearing"* de Brian C.J. Moore [7], y como no mencionar *"Speech and Language Processing"* de Daniel Jurafsky y James H. Martin [2], o el clásico mejorado ao a ao *"The HTK Book"*.

Una vez adquiridos los conocimientos base a través de la documentación anterior, han sido clave para el planteamiento de este trabajo, las publicaciones de Enrique López Poveda y Ray Meddis sobre modelos computacionales del oído.

2. Fisiología del Habla

El estudio de la pronunciación de las palabras en los distintos lenguajes es analizado por la fonética. La fonética modela la pronunciación de las palabras en una secuencia de símbolos que los representan. Existen principalmente dos alfabetos para describir los fonemas. El *International Phonetic Alphabet (IPA)* y *ARPabet* que usa símbolos ASCII para su representación.

El aire al ser expulsado por los pulmones y pasar por la laringe hace vibrar un par de estructuras elásticas llamadas cuerdas vocales, de 15 mm en hombres y 13 mm en las mujeres. Este sonido generado es posteriormente modificado al pasar a través del tracto vocal y tracto nasal en función de la posición que adopten los órganos vocales.

Para el análisis de los fonemas una herramienta fundamental es el espectrograma, que nos representa la distribución en frecuencias de la energía emitida

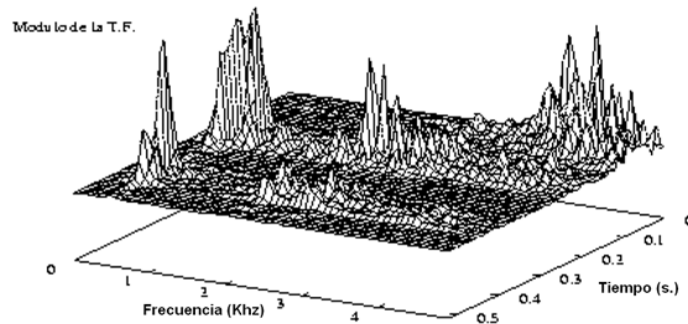


Figura 1. Representación Espectral pronunciación palabra [Siete]

en cada instante del tiempo. Cada pico de energía es llamado formante y corresponde a una frecuencia que es amplificada por la forma del tracto vocal/nasal. Cada fonema tiene una secuencia de formantes y valles que lo caracterizan.

A estos fonemas iniciales hay que unir toda una serie de fenómenos que complican en gran medida la clasificación y reconocimiento automático del habla. Entre estos fenómenos están las sílabas acentuadas, los alófonos de un fonema (diferentes pronunciaciones debido a la reducción/hiperarticulación o preparación para articular el siguiente fonema), la supresión de fonemas, la velocidad de pronunciación, la prosodia, los dialectos, etc.

Existen diccionarios fonéticos entre los que cabe destacar CELEX, CMUdict y PRONLEX para la lengua inglesa, así como otra serie de ellos para otras lenguas (Español, Alemán, Japonés, Árabe, ...) accesibles en el *Linguistic Data Consortium (LDC)*.

3. Fisiología del oído y su simulación con un modelo oído computacional

El oído humano puede percibir una extensa serie de frecuencias comprendidas entre 16 Hz y 20,000 Hz. Por regla general, la percepción de las altas frecuencias es mejor en la infancia y disminuye gradualmente con el tiempo, de manera que a un adulto normal le es difícil oír frecuencias que pasan de 12,000 Hz.

La intensidad sonora se expresa en decibeles (dB SPL). Una intensidad sonora de 0 dB es apenas perceptible, 20 dB es equivalente a un susurro a 1m de distancia y es 100 veces más intenso, 60 dB es equivalente a una conversación normal y es un millón de veces más intenso, y 100 dB es equivalente a un martillo neumático a 10m de distancia, siendo 10.000 millones de veces más intenso.

3.1. Oído externo

El oído externo además de protegernos de que no se nos introduzcan cuerpos extraños en el sistema auditivo, tiene la misión de concentrar las ondas sonoras, especialmente las de alta frecuencia y llevarlas al canal auditivo (Figura 2).

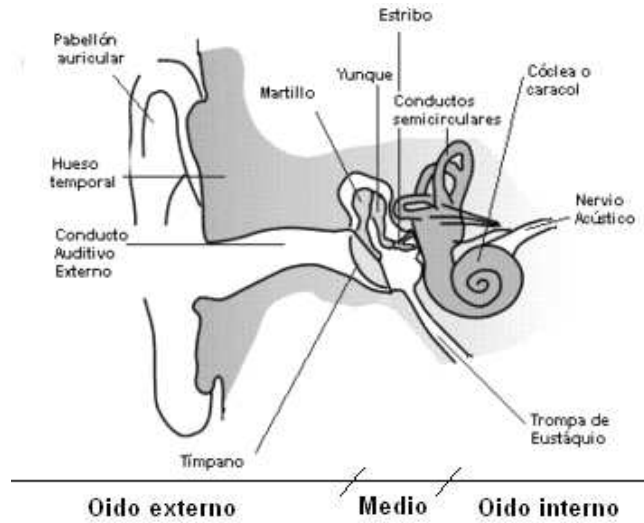


Figura 2. Representación secciones del Oído

Del mismo modo que los ojos nos proporcionan visión estereoscópica permitiéndonos calcular distancias, los oídos nos proporcionan audición estereofónica para determinar la dirección del sonido. Resulta destacable que el oído externo nos permite determinar la posición en altura de la fuente del sonido, para ello el cerebro es capaz de determinar las transformaciones que se producen en la señal acústica según el ángulo con que nos llega la señal al oído externo.

El pabellón auricular se comunica con el conducto auditivo que se extiende hasta el tímpano. En el tímpano los cambios de presión de las ondas sonoras se transforman en vibraciones mecánicas de longitud de onda sumamente pequeña. En una conversación normal el desplazamiento de la membrana es del orden del diámetro de una molécula de hidrógeno. El oído externo en su conjunto actúa como un elemento resonante que resalta las frecuencias entre 2 KHz y 5.5 KHz.

3.2. Oído Medio

El oído medio está compuesto por la cadena de huesecillos: martillo, yunque y estribo (Figura 2), que toman las vibraciones proyectadas sobre el tímpano y las conducen a la ventana oval (oído interno). Valga como curiosidad indicar que el estribo es el hueso más pequeño del cuerpo humano.

La principal función del oído medio es la adaptación de impedancias entre aire y el fluido que existe en la cóclea, ya que de no existir se provocaría un rebote de la energía que se pretende transmitir al oído interno.

Otra importante función de la cadena de huesecillos es tener controlada por músculos y ligamentos la movilidad máxima necesaria para la transmisión sonora. Los músculos timpánicos se combinan de tal manera que se contraen al mismo

tiempo formando una unidad de defensa ante los ruidos intensos, amortiguando el sonido a altas intensidades. Su eficiencia en esta tarea depende de la frecuencia de las vibraciones, transmitiendo óptimamente frecuencias medias (1 a 4 KHz), pero ofreciendo resistencia a las oscilaciones de otras frecuencias. Este fenómeno determina en buena medida la forma baera invertida de los audiogramas [6].

3.3. Oído Interno: la cóclea

En la cóclea es donde ocurre la transformación de energía mecánica en eléctrica mediante un fenómeno mecánico-químico-eléctrico. Esta energía bioeléctrica es conducida por los centros nerviosos hasta la corteza cerebral que es donde se integran las seales y tomamos conciencia de la imagen acústica.

La cóclea es un órgano en forma de caracol que descompone la seal acústica en las diferentes frecuencias, ya que el cerebro funciona en base a un mapa tonotópico, es decir que diferentes áreas del cerebro tratan en primera instancia la interpretación de distintos rangos de frecuencias, para posteriormente otras áreas del cerebro ir integrando información y abstrayendo sensaciones conjuntas.

La cóclea está formada por un sistema de tres rampas paralelas enrolladas: rampa vestibular, rampa media y rampa timpánica. La rampa vestibular y media se hallan separadas por la membrana de Reissner y la rampa timpánica y la rampa media se hallan separadas por la membrana basilar. Sin entrar en mayores detalles, la presión generada por el estribo en la rampa vestibular se traslada a movimientos en la membrana basilar.

3.4. Membrana basilar

Tal como demostró George von Békésy, Nobel Medicina 1961, cada zona de la membrana basilar es más sensible a un estrecho rango de frecuencias. Este fenómeno nos permite discriminar entre sonidos compuestos por frecuencias muy parecidas pero con diferente contenido armónico.

En el modelo utilizado se simula este comportamiento mediante un banco de filtros denominados *DRNL: Dual-Resonance NonLinear*, que representa cada región de la membrana basilar [4]. Los filtros DRNL reproducen toda una serie de efectos no lineales que se han comprobado vía experimental como son el hecho de que la frecuencia central, la anchura y la forma de la función del filtro varían con la intensidad del sonido.

Este comportamiento no lineal es el responsable del efecto de *supresión* por el cual la percepción de un sonido puede disminuir en presencia de seal en bandas adyacentes.

La Figura 3 muestra de forma simplificada los efectos de supresión ante un estímulo de seal una vez aplicado el banco de filtros que representan las sucesivas zonas de la membrana basilar. El resultado de este efecto de supresión es que se resaltan la amplitud entre picos y valles de frecuencias formantes, lo cual debería contribuir a aumentar la SNR en ambientes ruidosos [10].

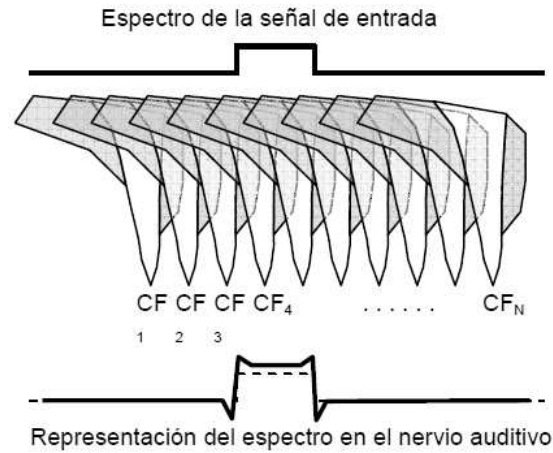


Figura 3. Efecto del fenómeno de *supresión* ante un estímulo de señal de entrada

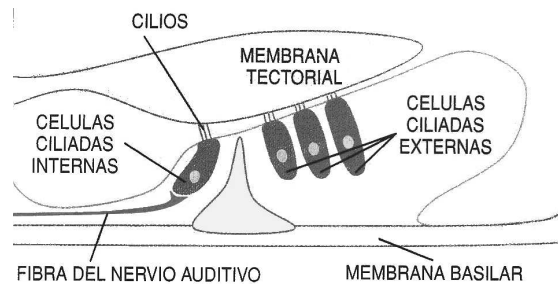


Figura 4. Órgano de Corti

3.5. Órgano de Corti

A todo lo largo de la superficie de la membrana basilar se halla una estructura denominada órgano de Corti (Figura 4), que contiene en los humanos unas 16.000 células ciliadas, y que son las encargadas de transformar los movimientos mecánicos en señales eléctricas que facilitan la descargas de neurotransmisores sobre el nervio acústico. En el modelo desarrollado se ha generado un algoritmo que simula la amplitud de oscilación de los cilios, genera el valor del potencial intracelular, y a partir de ahí se simulan el resto de fenómenos hasta la excitación del nervio auditivo. Para mayor detalle de todos estos mecanismos recomendamos la lectura (López-Poveda 2006 [3]).

Estos comportamientos de las células ciliadas son los responsables del efecto de *adaptación*, por el cual la actividad de las fibras del nervio auditivo es mayor al producirse un cambio brusco en un estímulo sonoro que una vez transcurrido un cierto tiempo [11]. Este fenómeno de *adaptación* se produce por los mecanismos de intercambio entre las células ciliadas y el nervio auditivo, en los que

disminuye la respuesta ante un mismo estímulo pasados unos milisegundos. Esta característica facilita la percepción de algunos sonidos consonánticos como la $[t]$ o la $[k]$ a pesar de su brevedad. Asimismo, tras el cese de un estímulo sonoro se reduce la actividad del nervio auditivo lo cual dificulta la percepción de sonidos breves que siguen a un sonido prolongado.

3.6. Modelo Oído Artificial

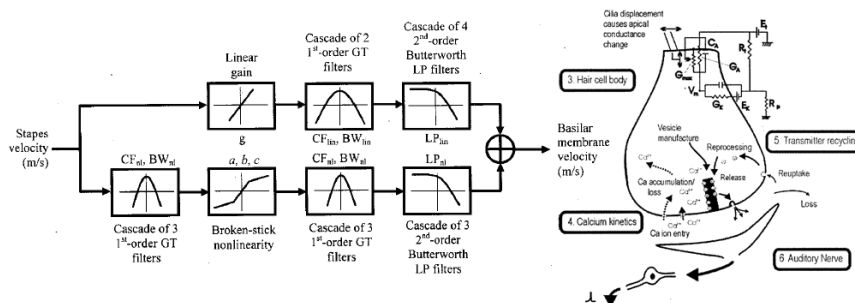


Figura 5. Modelo de Oído Artificial [6]

Simular todos estos comportamientos es una tarea compleja que ha llevado muchos años de investigación y cuyos resultados se conocen principalmente a través de sucesivas publicaciones de Enrique López Poveda y Ray Meddis.

Aislar y obtener las curvas y parámetros que caracterizan cada una de las fases del oído humano nos permiten tener una extraordinaria base de experimentación para simular diferentes patologías, o como en el objeto de este trabajo de investigación, para analizar los reconocedores automáticos del habla y comprender en profundidad qué características influyen de forma más significativa.

Para finalizar en este apartado, la Figura 5 muestra un esquema de bloques de las diferentes fases que simulan el comportamiento del oído humano. Este modelo está implementado en rutinas Matlab y en librerías de código C, disponibles bien a través del laboratorio de investigación liderado por Enrique López Poveda, o bien desde el proyecto DSAM: Development System for auditory Modelling, accesible en (<http://www.pdn.cam.ac.uk/groups/dsam/>).

4. Fases de un Sistema de Reconocimiento de Habla

Para la realización de nuestro reconocedor del habla se está usando HTK (Hidden Markov Model Toolkit). HTK se ha convertido en una herramienta fundamental en la investigación de las técnicas de reconocimiento del habla.

Dentro de la arquitectura de un reconocedor del habla, podemos distinguir las siguientes funciones: *extracción de características*, *algoritmos de búsqueda*, *modelos acústicos*, *diccionarios* y *modelos del lenguaje*.

4.1. Extracción características

Durante la *extracción de características* se obtiene información relevante de la seal analógica, de forma que nos permita distinguir de forma eficiente los fonemas que forman las palabras. Los fonemas están caracterizados por su envolvente espectral y la energía de la seal.

Un primer paso es digitalizar la seal analógica. Para ello debemos determinar la frecuencia de muestreo de la seal analógica (frecuencia Nyquist), y la granularidad o detalle con que medimos cada muestra. Valores típicos pueden ser un muestreo de 16 Khz y una granularidad de 8-bits (valores entre ± 127).

A continuación se realiza un pre-énfasis de la seal para amplificar las frecuencias altas y compensar el efecto de los pulsos glotales y la diferente impedancia de radiación según la frecuencia. Un siguiente paso es determinar los intervalos de seal que tomamos para extraer las características (ej: 25ms), y el desplazamientos entre intervalos (ej: 10ms). Para la obtención de estos intervalos se aplica una función coseno (ventana Hamming) que evita ondas laterales en el espectro de frecuencias.

Un paso final es obtener las características de la seal de cada ventana por medio del análisis espectral. Aunque existen distintos métodos (LPC, MFSC, PLP, VTLN, MFCC, ...), *MFCC: Mel Frequency Cepstral Coefficients* es el más usado por su mejor equilibrio entre complejidad y resultados obtenidos.

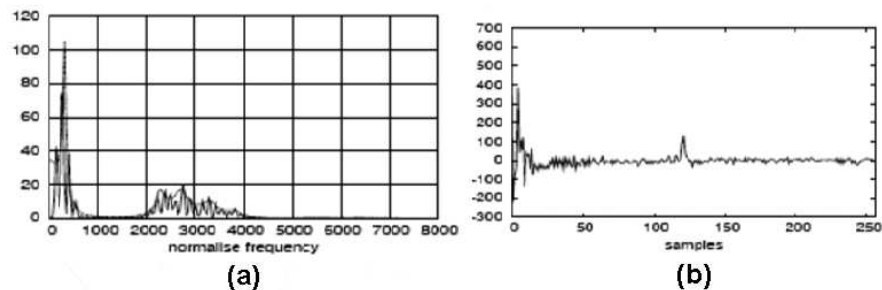


Figura 6. Cálculo del Cepstrum

El nombre *Cepstral* surge de invertir el orden de las 4 primeras de la palabra spectral, dado que pasamos del dominio de tiempo al dominio de frecuencias y luego nuevamente al dominio del tiempo.

El *cepstrum* es el *spectrum del logaritmo del spectrum* [2]. En la Figura 5-a se aprecia el espectro de frecuencias de un fonema y en la Figura 5-b se observa la seal del cepstrum, en el que se distingue el pulso 120 correspondiente a la seal

generada por las cuerdas vocales y los pulsos de valores bajos que se corresponden a las transformaciones del tracto vocal, que es de donde obtendremos los coeficientes cepstrales según una escala de percepción humana o escala Mel.

4.2. Modelos Acústicos

El modelado de reconocimiento del habla en base a Modelos Ocultos de Markov (HMM: Hide Markov Model), está siendo aplicado con éxito en numerosas soluciones comerciales y académicas. En este modelado del habla cada estado del HMM representa un fragmento de fonema y las observaciones se corresponden con los coeficientes cepstrales obtenidos anteriormente [9].

Una vez determinada la topología del HMM, es siguiente paso es el entrenamiento (algoritmo Baum-Welch) para obtener los coeficientes de transición entre estados del HMM y las funciones de densidad probabilísticas de los coeficientes cepstrales para cada uno de los estados. Estas funciones de densidad se modelan como mezcla de curvas gaussianas.

Existe el problema de que un mismo fonema puede variar mucho debido a la coarticulación o preparación de los órganos vocales para la articulación del siguiente fonema. Para considerar esta circunstancia se crean HMM's que reconocen un fonema dependiendo del contexto (trifonema). Esta evolución nos hace pasar de 50 fonemas a 100.000 trifonemas en el idioma inglés, lo cual obliga a buscar simplificaciones como es crear para cada fonema, clusters de fonemas precedentes, que condicionan de forma semejante la coarticulación del mismo.

Una vez entrenado el HMM, se le somete al reconocimiento de fonemas (algoritmo Viterbi). Para mejorar la tasa de aciertos, se complementa el resultado del modelo acústico con HMM, con diccionarios de palabras válidas y modelos del lenguaje como son las gramáticas y los modelos *N-gram* que representan las probabilidades en secuencias de palabras de longitud "N".

5. Conclusiones

La realización de este trabajo de Tesis de Máster nos ha permitido un primer acercamiento al estado del arte de estas tecnologías, y servir como punto de arranque para la posterior Tesis Doctoral donde se demuestre formalmente respuestas a las siguientes preguntas:

1. ¿Mejoran los reconocedores automáticos del habla, si se les antepone un modelo computacional del oído?
 - ¿En entornos sin ruido?. ¿En entornos de ruido blanco, ruido cafetería (bubble speech) y ruido de calle?. ¿Tiene un comportamiento lineal para diferentes niveles señal-ruido?. ¿Tiene un comportamiento semejante para diferentes niveles sonoros?
2. ¿Mejoran/empeoran los experimentos anteriores, sustituyendo la estructura fina de las bandas del modelo computacional del oído por su frecuencia característica, pero manteniendo la envolvente de la señal? [12]

agradecimientos

Quiero agradecer a Angel Luis Sánchez Lázaro y a Enrique López Poveda por el apoyo y ayuda prestada durante la preparación de este Trabajo de Master.

Referencias Bibliográficas

1. WH. Hartmann. *Signals, Sound, and Sensation (Modern acoustics and Signal Processing)*. ISBN 1-56396-283-7, fifth printing 2005.
2. D. Jurafsky and J.H. Martin. *Speech and Language Processing: an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall, 2008.
3. E.A. Lopez and a. Eustaquio-Martín. a biophysical model of the inner hair cell: The contribution of potassium current to peripheral compression. *JaRO-J assoc. Res. Otolaryngol* 7(3), pages 218–235, 2006.
4. EA. Lopez-Poveda and R. Meddis. a human nonlinear cochlear filterbank. *Journal of the acoustical Society of america* 110 (6), pages 3107–3118, 2001.
5. EA. Lopez-Poveda, LPO. O'Mard, and R. Meddis. a revised computational inner hair cell mode. *Psychophysical and Physiological advances in Hearing (Whurr Publishers Ltd, London, UK)*, pages 112–119, 1998.
6. EA. López-Poveda and R. Meddis. Oídos artificiales. *Mente y Cerebro* 10, pages 24–32, 2005.
7. Brian C.J. Moore. *an Introduction to the Psychology of Hearing*. aCaDEMIC PRESS INC., Third Printing 1992.
8. RW. Peters, BCJ. Moore, and T. Baer. Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people. *J. acoust. Soc. am.*, pages no 103, 577–587, 1998.
9. LR. Rabiner. a tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77, No 2:257–286, 1989.
10. M.B. Sachs and E.D. Young. Encoding of steady-state vowels in the auditory nerve: Representation in terms of discharge rate. *J. acoust. Soc. am.* 66, pages 470–479, 1979.
11. LA. Westerman and RL. Smith. Rapid and short-term adaptation in auditory nerve responses. *Hear. Res*, pages no 15, 249–60., 1984.
12. B.S. Wilson, R. Schatzer, E.a. Lopez-Poveda, X. Sun, D. Lawson, and R.D. Woford. Two new direction in speech processor design for cochlear implants. *EaR aND HEaRING*, pages VOL 26, NUMB 4; 735–815, 2005.
13. BW. Wilson, DT. Lawson, JM. Mller, RS. Tyler, and J. Kiefer. Cochlear implants: Some likely next steps. *ann. Rev. Biomed. Eng.* 5, 2003.

ViRe-YouTube: Visualizando los Resultados de Búsquedas de YouTube

Carlos Andrés Mera Banguero

Universidad de Salamanca, Departamento de Informática y automática
Salamanca - España
carlosmera20@gmail.com

Resumen En este trabajo se aborda el problema de la visualización de los resultados de búsqueda de vídeos en YouTube¹ desde tres perspectivas. (1) La adaptación de una técnica de visualización de información para dar una vista general de los vídeos recuperados y facilitar el reconocimiento de relaciones, patrones y tendencias entre ellos. (2) La utilización de una técnica basada en grafos para motivar la búsqueda exploratoria de vídeos utilizando las relaciones implícitas entre vídeos, etiquetas y otros vídeos relacionados. (3) La implementación de técnicas de previsualización de vídeos, que dan información de su contenido visual. Como resultado de este trabajo se implementó *ViRe-YouTube* un sistema para la recuperación de vídeos de YouTube que reemplaza la interacción tradicional (en la cual el usuario envía una consulta y el sistema le muestra una lista de resultados) con un estilo de interacción similar al de navegación, en el que la organización y presentación de los resultados ayudan al usuario a comprender mejor el conjunto de vídeos recuperados y por ende a hacer una selección más adecuada de los vídeos a reproducir.

1. Introducción

El creciente número de vídeos de YouTube² ha hecho de este repositorio uno de los lugares más populares y visitados de Internet. YouTube tiene el potencial de contener información de interés para muchos tipos usuarios. No obstante, el disponer de tanta información genera inconvenientes en el proceso de búsqueda y exploración de vídeos, lo que conlleva a que éste sea lento, tedioso, y algunas veces, frustrante.

La búsqueda de vídeos en YouTube, al igual que en otros repositorios de su tipo, se realiza a través de una interfaz en la que el usuario ingresa los términos que describen los vídeos a buscar. Cuando se obtienen los resultados, el sistema muestra en una lista una imagen representativa y la descripción general de los vídeos recuperados. Este esquema de presentación, que es ampliamente utilizado, tiene algunas limitaciones: (a) No es fácil determinar si existe, o no, una relación entre los vídeos recuperados; (b) Cuando se hacen búsquedas exploratorias, en las que el usuario tiene una vaga, o ninguna, idea de lo que busca, las listas no son

¹ <http://www.youtube.com/>

² Se estima que hay más de 45'000.000 de vídeos y que el repositorio crece a un ritmo asombroso, en el que cada minuto los usuarios suben cerca de 7 horas de vídeo [1]

suficientes, [2] ya que requieren un gran esfuerzo cognitivo para la exploración; Y (c) una imagen representativa del vídeo no da la suficiente información de su contenido para realizar una selección cuidadosa de los vídeos que son relevantes para el usuario.

En la literatura hay diferentes trabajos que intentan dar solución a estos y otros inconvenientes de los sistemas de recuperación de vídeos. algunos autores, por ejemplo, se centran en el proceso de generación de previsualizaciones del vídeo [3,4,5,6,7,8], otros proponen mecanismos de búsqueda que involucran conceptos semánticos [9], y otros exploran la posibilidad de realizar búsquedas basadas en contenido [10]. Sin embargo, con respecto a la forma en que se visualizan los resultados de dichas búsquedas, son pocos los trabajos que se encuentran en la literatura [11].

En este artículo se propone y se describe una herramienta que utiliza diferentes técnicas de visualización de información para presentar y organizar los resultados de las búsquedas de vídeos en YouTube. Las representaciones propuestas le permiten al usuario analizar la colección de vídeos recuperados, detectar patrones y encontrar relaciones, que en una lista no son evidentes. además, con el fin de proporcionar elementos de juicio más intuitivos, se propone la utilización de diferentes mecanismos de previsualización de vídeos, los cuales dan al usuario información del contenido visual, y que por tanto, le permiten hacer una selección más cuidadosa de los vídeos que son de su interés.

El resto del artículo se organiza como sigue. En la Sección 2, se hace una breve introducción del estado del arte. En la Sección 3 se describen algunos trabajos relacionados. En la sección 4 se detalla *ViRe-YouTube*, la herramienta que se propone en este trabajo. En la Sección 5 se presentan las conclusiones. Y en la Sección 6 se describen algunos trabajos futuros.

2. Estado del arte

La aplicación de las técnicas de visualización de información a los sistemas de recuperación de vídeos, abre la posibilidad de desarrollar nuevos sistemas de recuperación que mejoren la experiencia de búsqueda del usuario. a continuación se presentan conceptos relacionados con el área de visualización de información y se describen, brevemente, algunas técnicas de previsualización de vídeos.

2.1. Visualización de Información

*“...los sistemas informáticos de Visualización de Información, se enfocan en brindar ayudas para la exploración o la explicación de datos a través de sistemas que incluyan representaciones visuales, contemplando intervenciones interactivas del observador. El principal desafío en el diseño de visualizaciones, es el de representar el conjunto de datos abstractos, no necesariamente espaciales, reforzando los aspectos cognitivos...”*³

³ <http://www.infovis.org/infovis/2005/>

Como se describe en la definición anterior, lo que se busca al aplicar técnicas de visualización de información es generar una representación visual de los datos, tal que ésta optimice la posibilidad de comprensión e interiorización del conocimiento inherente en ellos.

Para ello, se han propuesto diferentes modelos que permiten llevar un conjunto de datos en crudo a una representación visual para su interpretación y análisis [12,13,14]. Ben Fry [12], por ejemplo, propone un modelo que consta de siete etapas que van desde la adquisición de los datos hasta su presentación visual, refinamiento e interacción con el usuario. Cualquiera que sea el modelo seleccionado para crear visualizaciones de los datos, estos involucran la aplicación y estudio de diferentes aspectos de minería de datos, psicología cognitiva, y principios de diseño de interfaces gráficas de usuario, entre otros.

Dada la importancia de estos aspectos, a continuación se hace una breve descripción de ciertos elementos de la psicología de la percepción, metáforas de representación de datos y se presentan algunos elementos de interacción, en los sistemas de visualización de información.

Psicología de la Percepción. Cómo menciona Shneiderman [15], una representación visual es más fácil de usar y comprender que una descripción textual. El sistema visual humano es un poderoso sistema de procesamiento capaz de detectar cambios de forma, color, textura, tamaño, orientación y posición de los objetos, lo que favorece la transmisión de conocimiento a través de esas características visuales.

En términos de percepción, el sistema visual está especialmente adaptado para procesar, más rápidamente y con mayor precisión, ciertas propiedades visuales. a este conjunto de propiedades se les llama preatentivas, ya que su detección parece preceder al momento de la atención [16]. Parte de la importancia de los sistemas de visualización de información recae en su capacidad para ayudar a “ver” la información que se oculta en los datos. La utilización de propiedades preatentivas ayuda a los sistemas de visualización en esa tarea de percepción. algunas propiedades preatentivas ampliamente utilizadas son:

- *La Posición:* que determina el lugar en el espacio del objeto.
- *El Tamaño:* que está relacionado con el área del objeto.
- *El Color:* que especifica la longitud de onda de la luz con que se pinta el objeto.
- *La Forma:* que define el símbolo utilizado para representar el objeto.

Jacques Bertin [17] menciona que estas propiedades visuales producen un impacto sobre cuatro niveles de percepción del usuario:

- *Asociativo:* que permite formar grupos de objetos.
- *Selectivo:* que responde a la percepción de objetos dentro de un grupo que poseen una particularidad.
- *Ordinal:* que se expresa con la especificación de una variable particular cuantitativa y ordenada.
- *Cuantitativo:* que responde a la percepción de la razón entre dos o más objetos, permitiendo la comparación con respecto a una o varias características.

Métáforas de Representación de Datos. Una de las decisiones más importantes en el proceso de visualización de información, es determinar de que manera se van a presentar los datos al usuario. Fry [12] señala que esta selección debe hacerse de tal forma que la representación sea lo más simple posible pero que transmita los aspectos más relevantes del conjunto de datos. Estas son algunas de las representaciones más comunes:

- *Gráficos de Dispersión*: los datos se representan como una nube de puntos cuya posición horizontal y vertical se determina por los valores en los datos. Estos gráficos pueden ser extendidos a 3 o más dimensiones.
- *Mapas de Calor*: son una representación en la que los datos se codifican como colores en un mapa de dos dimensiones.
- *Árboles*: son utilizados para representar datos ordenados con estructuras jerárquicas.
- *Grafos*: emplean una colección de nodos y aristas para representar las relaciones binarias entre los objetos de un conjunto.
- *TreeMaps*: son un tipo de visualización de jerarquías que se representan por el anidamiento rectángulos. En este tipo de representación, cada nivel de anidación se corresponde con un nivel de descomposición de la jerárquica

Interacción con el Usuario. El mantra de búsqueda visual de información de Shneiderman [15] “*Overview first, zoom and filter, then details-on-demand*” plantea el siguiente conjunto de tareas, como las más relevantes a tener en cuenta en el diseño de sistemas de visualización interactivos:

- *Vista General (Overview)*: proporcionar una vista general de la colección de datos que permita, al usuario, ver como estos están organizados.
- *Escala (Zoom)*: permitir al usuario escalar los datos que son de su interés, por ejemplo, variando la distancia focal de la visualización.
- *Filtrado (Filter)*: proveer un mecanismo que permita eliminar de la visualización los objetos que no son de interés, dándole la posibilidad al usuario de centrarse en los objetos que si lo son.
- *Detalles a Demanda (Details-on-Demand)*: la idea de este concepto es brindar al usuario la posibilidad de solicitar información detallada sobre un elemento, o un grupo de elementos, cuando él lo requiera.
- *Relacionar (Relate)*: permitir al observador reconocer las relaciones entre los objetos.
- *Historial (History)*: mantener la historia de acciones realizadas por los operadores para permitir volver a un estado anterior o repetir una acción.
- *Extracción (Extract)*: para poder extraer un conjunto de objetos de la visualización para su análisis posterior.

2.2. Previsualización de Vídeos

El interés de un usuario al previsualizar un vídeo es evaluar si ese vídeo es de su interés, o no, sin tener que descargarlo y reproducirlo en su totalidad. La

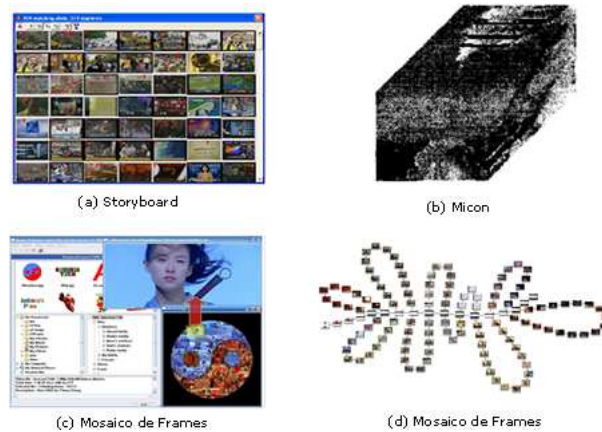


Figura 1. Métodos de Organización Espacial de Frames

previsualización de un vídeo se realiza disminuyendo la información redundante y eliminando el contenido que se puede predecir viendo una parte del mismo [18].

La mayoría de los autores clasifican la previsualización en uno de dos tipos: previsualización estática, la cual se genera utilizando un conjunto de frames del vídeo [19,4,20,5,21,22,23], y la previsualización dinámica que es una versión corta del vídeo original generada a partir de algunos planos del vídeo [24,7,8,25].

En el caso de las previsualizaciones estáticas, un tema de investigación es como organizar los frames que hacen parte de la previsualización. La forma más común de hacerlo es utilizar una tabla a la que se le llama *storyboard* [11] (1a). Smoliar *et al.* proponen en [26] otra manera de organización llamada *micon*, que es un sólido rectangular en el que los frames se disponen en orden secuencial (figura 1b). Kim *et al.* [27] crean mosaicos utilizando frames con características de color similar (figura 1c) y Goeau *et al.* [28] proponen una disposición llamada *Table Of Video Contents* que se ilustra en la figura 1d.

3. Trabajos Relacionados

Quizá la manera más común de organizar y visualizar los resultados de una búsqueda es a través de una lista textual en la que se presenta el título del documento y otros elementos descriptivos como el autor, el URL y algunas veces una imagen o *screenshot* del documento. No obstante, en los sistemas de recuperación de vídeos estos elementos descriptivos suelen ser insuficientes para que el usuario seleccione rápidamente los vídeos son relevantes para él.

En el proyecto Informedia [11,29,30] se ha desarrollado una interfaz de consulta y presentación de resultados que a través de diferentes técnicas de visualización de información (líneas de tiempo, gráficos de dispersión, etc), proporcionan al usuario una herramienta para analizar y explorar las relaciones entre los dife-

rentes segmentos de vídeos que se recuperan en una consulta. No obstante, los métodos de representación que se proponen en Informadia están orientados a colecciones de vídeos estructurados (noticias y entrevistas), que además, pueden ser complementadas con otros elementos de visualización de información.

Con respecto a trabajos que se centran en la visualización de resultados de búsqueda de vídeos en YouTube, el propio YouTube propone una herramienta visual para la exploración de vídeos, sin embargo, ésta no cuenta con ninguna de las iteraciones de las que Shneiderman se refiere en su mantra de búsqueda visual de información [15].

SearchMe⁴ y DiscoverGalaxi⁵ son dos aplicaciones que proponen componentes visuales para la búsqueda de vídeos. SearchMe utiliza una metáfora basada en listas llamada *cover flow* y DiscoverGalaxi proporciona dos vistas basadas en tablas y una adicional que organiza los vídeos en un espacio tridimensional. Sin embargo, estas propuestas cuentan con los mismos problemas de las clásicas listas de resultados. Otros dos sistemas de búsqueda de vídeos en YouTube son oSkope⁶ y TimeTube⁷, los cuales utilizan visualizaciones basadas en gráficos de dispersión. No obstante, éstas se limitan a organizar los vídeos en términos de dos de sus atributos.

Flookon⁸ es una aplicación de búsqueda y exploración que propone una visualización basada en grafos para representar las relaciones entre los vídeos recuperados, las etiquetas y otros vídeos relacionados de YouTube. El problema de esta visualización, al igual que las anteriores, es que no proporciona detalles de los vídeos, ni dan información sobre su contenido visual.

Por otro lado, Blinkx⁹ y VideoSurf¹⁰ sí dan información del contenido visual de los vídeos recuperados. a pesar de ello, su representación visual está basada en listas.

4. *ViRe-YouTube*

La herramienta que se propone en este trabajo, es una primera aproximación de un sistema visual de recuperación de vídeos de YouTube que cuenta con múltiples vistas para analizar los vídeos recuperados, realizar búsquedas exploratorias y previsualizar el contenido de los vídeos, esto con el fin de permitir al usuario hacer una selección más cuidadosa de los vídeos que son de su interés. En las subsecciones siguientes se describen los componentes principales de *ViRe-YouTube* y se explica como esos componentes implementan las tareas principales del mantra de búsqueda de información visual de Shneiderman.

⁴ <http://www.searchme.com/>

⁵ <http://www.tkaap.com/galaxy/>

⁶ <http://oskope.com/>

⁷ <http://www.dipity.com/mashups/timetube>

⁸ <http://www.flookoon.com/#youtube>

⁹ <http://www.blinkx.com/>

¹⁰ <http://www.videosurf.com/>

4.1. Un Gráfico de Dispersión para analizar los Vídeos Recuperados

Una vista general del conjunto de datos recuperados puede dar al usuario pistas de su organización, y más aún, le puede ayudar a seleccionar o eliminar fuentes de información [31]. Una de las limitaciones de las listas de documentos (y en este caso particular, de las listas de vídeos) es que no le dan al usuario esa vista general que mencionan Baeza y Ribeiro en su libro “*Modern Information Retrieval*” [31].

En este trabajo se propone la utilización de un gráfico de dispersión enriquecido con propiedades visuales preatentivas para organizar y presentar los vídeos recuperados de una consulta. La utilización de este tipo de visualización abre la posibilidad al usuario de identificar, rápidamente, con respecto a los diferentes atributos de los vídeos, grupos, tendencias, comportamientos y *outliers*. además, por la naturaleza lineal de su algoritmo de representación, el número de vídeos que se pueden representar es casi ilimitado, lo que favorece la vista general que propone Shneiderman en su mantra de búsqueda de información visual; y que por otro lado mencionan Baeza y Ribeiro en [31].

Traducción de atributos Gráficos. Cada asignación de características visuales a los objetos resuelve con diferente grado de efectividad los cuatro niveles de percepción propuestos Bertín en [17]. Por eso es importante determinar que atributos de los vídeos se van a representar con cada característica visual:

- *Posición*: la posición de los objetos le permite al usuario asociar y agrupar los videos por proximidad, a demás que permite encontrar videos fuera de rango o que no siguen las tendencias de la mayoría de los videos recuperados. Los atributos utilizados para determinar la posición de los vídeos son su duración (eje X) y el número de visitas (eje Y).
- *Forma*: es una característica visual ideal para el nivel asociativo [17]. Por esta razón, se propone utilizar la forma para representar la antigüedad de los vídeos.
- *Color*: en términos de los niveles de percepción de Bertín [17], el color es ideal para impactar en el nivel asociativo y selectivo, por esta razón se utiliza para representar las categorías de vídeos de YouTube.
- *Tamaño*: como el tamaño es adecuado para los niveles ordinal y cuantitativo, este se utiliza para representar la calificación del video, que es un atributo de naturaleza ordinal cuyo rango no permite que se desborden los objetos por su tamaño.

Esta asignación de características logra una estructura visual efectiva, en la que en un espacio limitado se presenta un conjunto amplio de atributos informativos de la colección de videos recuperados. La figura 2 muestra un ejemplo para la consulta “Obama”. En este ejemplo, el usuario puede rápidamente observar en el gráfico de dispersión que todos los videos recuperados tienen una duración corta, y que la mayoría fueron publicados a finales del 2008, además, se evidencia que los videos se agrupan en tres categorías diferentes y que uno de ellos

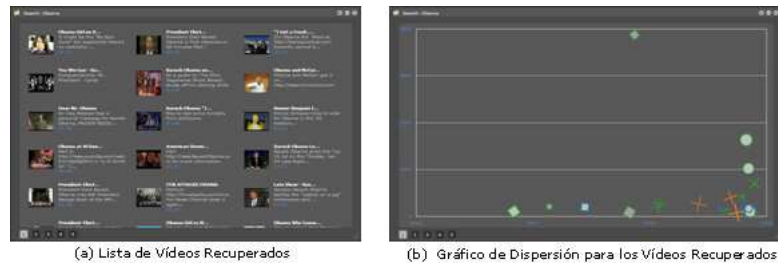


Figura 2. Dos Vistas en la que se Visualizan los Vídeos resultantes de la consulta “Obama”

está fuera de rango, con respecto a la duración. abstraer esta información del gráfico requiere menor tiempo, procesamiento y esfuerzo cognitivo que cuando se abstrae de una lista.

Controles e Interacción con el Usuario. Esta visualización cuenta con un mecanismo que le permite al usuario ver en la propia visualización los detalles de los vídeos, evitando de esta manera, la pérdida del contexto de la consulta, a diferencia de los que sucede en los sistemas presentados en la sección 3. Otro control prouesto e implementado en *ViRe-YouTube* permite a los los usuarios cambiar los atributos que se representan en los ejes del gráfico, con lo que facilita el análisis de los vídeos desde diferntes puntos de vista.

4.2. Búsqueda Exploratoria de Vídeos a través de Grafos

En este trabajo, se propone un mecanismo de exploración de videos del tipo *Query searching and Browsing*[32], similar al utilizado por Flokoon, en el que se emplean las etiquetas y las listas de vídeos relacionados para explorar otros vídeos que pueden estar interconectados con los términos de la consulta inicial.

Los grafos, a diferencia de las visualizaciones basadas en listas, ofrecen un mecanismo de recuperación visual en el que se permite al usuario explorar el espacio de búsqueda sin perder el contexto de la consulta inicial. además, esta metáfora de representación visual, disminuye la carga cognitiva y las tareas memorísticas de la exploración y hace visibles las relaciones, que en las representaciones basadas en listas, no se perciben.

En *ViRe-YouTube* se propone la utilización de dos tipos de grafos: grafos dirigidos por fuerzas y grafos radiales. La figura 3 muestra una búsqueda exploratoria iniciada con los vídeos recuperados de la figura 2 utilizando los dos tipos de grafos, los cuales se explican a continuación.

Grafos Dirigidos por Fuerzas. Se modelan como sistemas físicos, en los que se definen fuerzas que actúan sobre los nodos, tal que su posición se establece encontrando el equilibrio en el sistema físico [33].

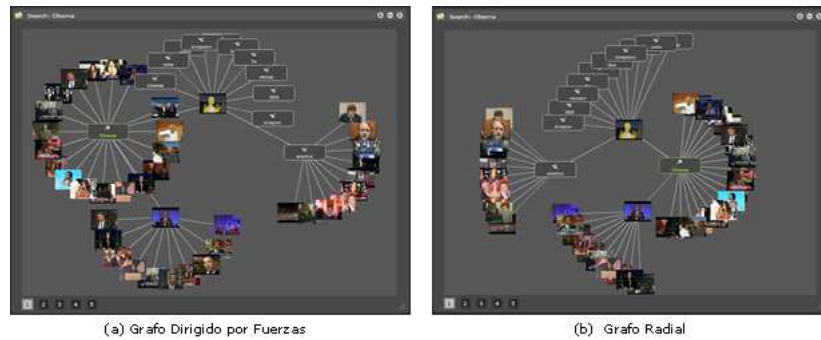


Figura 3. Búsqueda Exploratoria Iniciada con la Consulta “Obama”

Una de las motivaciones para utilizar grafos dirigidos por fuerzas, en este trabajo, son los buenos resultados estéticos que se consiguen con ellos. Este tipo de grafos logra un equilibrio visual que se refleja en aristas de igual tamaño y nodos simétricamente distribuidos, características de organización, que disminuyen la carga en el proceso cognitivo del usuario y que facilitan la exploración de la colección de vídeos. además, como se basan en analogías físicas de objetos comunes, su comportamiento es relativamente fácil de predecir y entender.

Grafos Radiales. Un grafo organizado en una disposición radial, o “grafo radial”, es tratado como un árbol cuya raíz es definida por un nodo foco que se posiciona en el centro de la visualización. Los otros vértices se organizan en anillos concéntricos alrededor del nodo foco, dependiendo de su distancia hasta dicho nodo foco [34].

La propuesta de utilizar dos tipos de grafos en la búsqueda exploratoria se sustenta en una anotación de Yee *et al.* [34] que sugiere que la habilidad para ver un grafo, interactivamente, desde diferentes perspectivas puede producir nuevo conocimiento sobre los datos.

Controles e Interacción con el Usuario. Basados en el mantra de búsqueda de información visual de Shneiderman [15], los grafos, al igual que los gráficos de dispersión, dan al usuario una vista general de los vídeos recuperados, para ello cuentan con controles que le permiten al usuario variar la distancia focal de la visualización y obtener detalles de los vídeos bajo demanda. Estos grafos también cuentan con un control que les permite explorar el espacio de búsqueda de los vídeos (explosión de nodos) y eliminar de la visualización aquellos que no son de su interés (eliminación de nodos). Igualmente, tienen mecanismos que le permite al usuario desplazarse por la visualización (paneo de cámara) y reacomodar los nodos de la misma (arrastrar y soltar).

4.3. Filtros para las Visualizaciones de *ViRe-YouTube*

El filtrado, es una de las técnicas básicas de interacción que hace parte de las tareas que propone Shneiderman [15] en su mantra. Los filtros permiten al usuario controlar el contenido de la visualización, de tal manera que estos le ayudan a centrar su atención eliminando los elementos que no son de su interés.

En este trabajo se proponen dos tipos de filtros que, además de dar información sobre la agrupación de los vídeos en categorías y autores, permiten recuperar los vídeos que pertenecen a un determinado autor o categoría. Estos criterios son útiles para discernir y seleccionar los vídeos de interés que pertenecen a una clase determinada.

Filtro Basado en un Árbol Expansible . Un árbol expansible es una representación visual de un árbol en la que los nodos padres se pueden expandir o contraer con el fin de mostrar u ocultar la rama del árbol que se deriva de ellos. El árbol expansible, que se propone en este trabajo, organiza los nodos como títulos de una tabla de contenido en la que el espacio de la sangría, con la que se separan los nodos del margen izquierdo, depende de la profundidad del nodo en el árbol, como se ilustra en la figura 4a.

Este filtro tiene implementados algunos controles que le permiten al usuario expandir y contraer las ramas del árbol (interacción utilizada para filtrar los vídeos en las otras visualizaciones); y modificar la estructura del árbol a una de cuatro disponibles (sólo categorías, sólo autores, categorías-autores o autores-categorías).

Filtro Basado en un *TreeMap* . Un *treemap* [35] es una técnica de visualización en la que el árbol se representa con un conjunto de rectángulos anidados que simbolizan los nodos del árbol.

Algunas de las motivaciones por las que se propone un *treemap* como filtro son: su idea de representación es relativamente sencilla y por tanto su curva de aprendizaje es baja; proporcionan al usuario una visión de conjunto, de árboles complejos, más rápido que una estructura de representación basada en nodos y aristas; y además, su uso eficiente de espacio disminuye la necesidad de usar controles de desplazamiento y *zoom*. La figura 4b pone en evidencia estas ventajas, con respecto al árbol expansible.

Con los controles de interacción implementados en el *treemap*, el usuario puede seleccionar un nodo tipo categoría y filtrar los datos en la visualización que pertenecen a esa categoría, además puede cambiar el atributo que determina el tamaño de los rectángulos, que por defecto es el número de calificaciones del vídeo.

4.4. Métodos de Previsualización de Vídeos

Como se ha mencionado antes, uno de los problemas de la mayoría de sistemas de recuperación de vídeos de YouTube, es que sólo muestran al usuario una

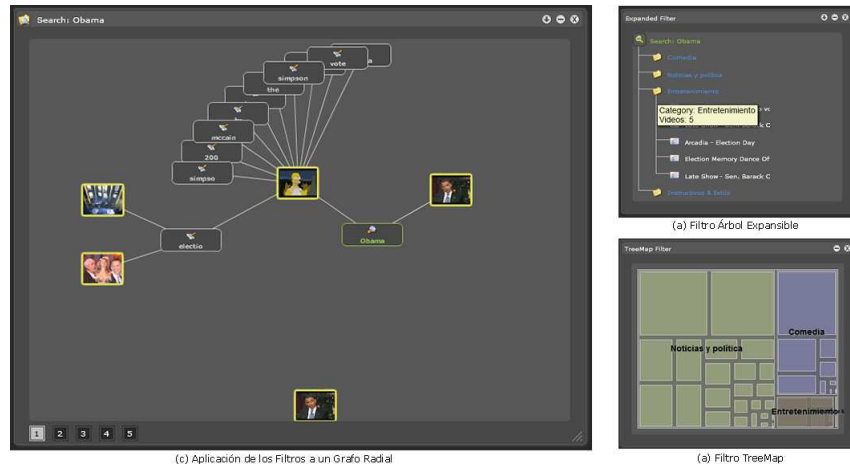


Figura 4. Filtros y su aplicación en la Búsqueda Exploratoria de la 3

imagen representativa del contenido del vídeo, lo que dificulta y retrasa la tarea de encontrar los vídeos relevantes para el usuario. Para cubrir esa desventaja, en este trabajo se propone la integración de cuatro métodos de previsualización que ayudan al usuario a identificar de forma más rápida los vídeos que son de su interés.

Los métodos de previsualización que se explican a continuación utilizan un conjunto de frames uniformemente seleccionados, por cada segundo de duración del vídeo. Este tipo de selección de frames se escogió con base en la discusión de que no hay diferencia entre la comprensión adquirida al ver un conjunto de frames cuidadosamente seleccionados, o al ver un conjunto de frames seleccionados a través de un muestreo uniforme en el tiempo, la cual fue presentada por Simone Santini en [36].

Previsualización Dinámica. La previsualización dinámica que integra *ViRe-YouTube* consiste en un GIF animado creado a partir de un conjunto de frames del vídeo obtenidos de una selección a intervalos uniformemente distribuidos de tiempo. La utilización de este tipo de previsualización da al usuario una idea del contenido del vídeo a través de una vista rápida del mismo. La figura 5a muestra la previsualización de uno de los vídeos obtenidos con la consulta “Obama”.

Previsualización en Espiral. Se propone una previsualización en la que los frames del vídeo se organizan en una espiral logarítmica, tal que el tamaño de los frames aumenta a medida que el punto en el que se ubican en la espiral se aleja del centro del gráfico. La ecuación polar de la espiral logarítmica está determinada por la ecuación 1.

$$r = ae^{b\theta} \quad (1)$$

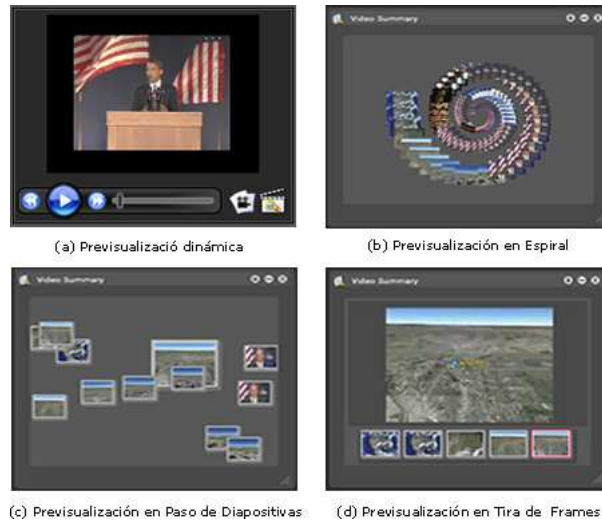


Figura 5. Tipos de Previsualización en *ViRe-YouTube*

donde, r es la distancia del punto al origen, θ es el ángulo formado con el eje X , y a y b son constantes arbitrarias.

Este tipo de previsualización le permite al usuario formarse una idea rápida de la estructura del vídeo, e identificar fácilmente los planos del mismo. La previsualización en espiral le da al usuario información del contenido del vídeo y le muestra cómo cambian los frames de principio a fin o de fin a inicio, como se ilustra en la figura 5b.

Previsualización como Paso de Diapositivas. Otra previsualización propuesta e implementada es la de paso de diapositivas. En este tipo de previsualización el usuario puede ver como cambia el vídeo frame a frame, como se ilustra en figura 5c. Esta previsualización tiene la ventaja de que el usuario puede ver todos los frames extraídos del vídeo a una velocidad constante de tal manera que se facilita la tarea de encontrar detalles que no pueden ser vistos ni en la previsualización dinámica, ni en la previsualización en espiral.

Previsualización en Tira de Frames Esta es una previsualización similar a la que utiliza VideoSurf, en la que los frames se organizan en una lista horizontal y en la que a demanda del usuario se amplía uno de los frames en la parte central de la visualización, como se ilustra en la figura 5d.

5. Conclusiones

El desarrollo de este trabajo muestra cómo la aplicación de técnicas de visualización de información, en sistemas de recuperación de vídeos, abre la posi-

bilidad de explorar la colección de vídeos, a la vez que provee un mecanismo de análisis que le facilita al usuario la tarea de descubrir patrones, tendencias y las relaciones implícitas entre los vídeos.

Las visualizaciones propuestas e implementadas en este trabajo cuentan con la habilidad de presentar y organizar un conjunto de vídeos, en un espacio que facilita la comparación y favorece el análisis y la exploración de vídeos, a diferentes niveles de detalle. así mismo, toman provecho de las capacidades de procesamiento del sistema visual humano para detectar patrones y tendencias de los vídeos recuperados a través de su representación con propiedades preatentivas.

Finalmente, el uso en conjunto de las diferentes visualizaciones, filtros y mecanismos de previsualización proporciona al usuario una poderosa herramienta que reduce la carga cognitiva y las tareas memorísticas involucradas en el proceso de búsqueda y exploración de vídeos.

6. Trabajos Futuros

Con el desarrollo de este trabajo se abre la posibilidad de abordar algunos problemas no contemplados en él, algunos de ellos son:

- Realizar estudios de impacto, aceptación y efectividad de los diferentes tipos de visualización utilizados.
- Proponer visualizaciones que soporten la búsqueda basada en recuperación semántica y/o características de bajo nivel de los vídeos.
- analizar y mejorar algunos de los algoritmos implementados, con el fin de disminuir la carga de la memoria y su tiempo de respuesta.

Referencias Bibliográficas

1. Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D., aly, M.: Video suggestion and discovery for youtube: taking random walks through the view graph. In: Proceeding of the 17th international conference on World Wide Web, NY, USa, aCM (2008) 895–904
2. Kules, W.M.: Supporting exploratory web search with meaningful and stable categorized overviews. PhD thesis, University of Maryland, College Park (2006)
3. Nagasaka, a., Tanaka, Y.: automatic video indexing and full-video search for object appearances. 2nd IFIP Working Conference on Visual Database Systems (1991) 119–133
4. Wolf, W.: Key frame selection by motion analysis. IEEE International Conference on acoustics, Speech, and Signal Processing **2** (May 1996) 1228–1231
5. Zhuang, Y., Rui, Y., Huang, T., Mehrotra, S.: adaptive key frame extraction using unsupervised clustering. Proceedings of International Conference on Image Processing **1** (October 1998) 866–870
6. Li, Y., Zhang, T., Tretter, D.: an overview of video abstraction techniques. Technical report, HP Laboratories - Imaging Systems Laboratory (July 2001)
7. Ma, Y.F., Lu, L., Zhang, H.J., Li, M.: a user attention model for video summarization. In: Proceedings of the tenth aCM international conference on Multimedia, NY, USa, aCM (2002) 533–542

8. Huet, B., Merialdo, B.: automatic Video Summarization. *Signals and Communication Technology*. In: *Interactive Video*. Springer (2007) 27–42
9. Luo, H.: *Concept-Based Large-Scale Video Database Browsing and Retrieval via Visualization*. PhD thesis, University of North Carolina at Charlotte (2007)
10. Declair, C., Hacid, M.S., Kouloumdjian, J.: a genetic model for video content based retrieval. In: *SaC '98: Proceedings of the 1998 aCM symposium on applied Computing*, NY, USA, aCM (1998) 458–459
11. Christel, M.G.: Supporting video library exploratory search: when storyboards are not enough. In: *Proceedings of the 2008 international conference on Content-based image and video retrieval*, NY, USA, aCM (2008) 447–456
12. Fry, B.J.: *Computational information design*. PhD thesis, School of architecture and Planning, Massachusetts Institute of Technology (2004)
13. Card, S.K., Mackinlay, J.: The structure of the information visualization design space. *Proceedings of IEEE Symposium on Information Visualization* (October 1997) 92–99
14. Chi, E.H.H., Riedl, J.: an operator interaction framework for visualization systems. *Proceedings of the IEEE Symposium on Information Visualization* (October 1998) 63–70
15. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. *Proceedings of IEEE Symposium on Visual Languages* (September 1996) 336–343
16. Triesman, a.: Preattentive processing in vision. In: *Computer Vision, Graphics, and Image Processing*. (1985) 156–177
17. Bertin, J.: *Semiology of graphics*. University of Wisconsin Press (1983)
18. Wang, F., Ngo, C.W.: Rushes video summarization by object and event understanding. In: *Proceedings of the international workshop on TRECVID video summarization*, NY, USA, aCM (2007) 25–29
19. Mills, M., Cohen, J., Wong, Y.Y.: a magnifier tool for video data. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, NY, USA, aCM (1992) 93–98
20. Zhang, H.J., Wu, J., Zhong, D., Smoliar, S.W.: an integrated system for content-based video retrieval and browsing. *Pattern Recognition* **30**(4) (april 1997) 643–658
21. DeMenthon, D., Kobla, V., Doermann, D.: Video summarization by curve simplification. In: *Proceedings of the sixth aCM international conference on Multimedia*, NY, USA, aCM (1998) 211–218
22. Divakaran, a., Radhakrishnan, R., Peker, K.a.: Motion activity-based extraction of key-frames from video shots. *Proceedings of International Conference on Image Processing* **1** (2002) 932–935
23. de avila, S.E.F., da Luz Jr., a., de a. Araújo, a.: Vsumm: a simple and efficient approach for automatic video summarization. *15th International Conference on Systems, Signals and Image Processing* (June 2008) 449–452
24. Nam, J., Tewfik, a.H.: Video abstract of video. *IEEE 3rd Workshop on Multimedia Signal Processing* (1999) 117–122
25. Chen, F., Cooper, M., adcock, J.: Video summarization preserving dynamic content. In: *Proceedings of the international workshop on TRECVID video summarization*, NY, USA, aCM (2007) 40–44
26. Smoliar, S.W., Zhang, H., Koh, S.L., Lu, G.J.: Interacting with digital video. *Proceedings of IEEE Region 10's Ninth annual International Conference* **2** (august 1994) 852–856

27. Kim, K., Essa, I., abowd, G.D.: Interactive mosaic generation for video navigation. In: Proceedings of the 14th annual aCM international conference on Multimedia, NY, USA, aCM (2006) 655–658
28. Goeau, H., Thievre, J., Viaud, M.L., Pellerin, D.: Interactive visualization tool with graphic table of video contents. IEEE International Conference on Multimedia and Expo (July 2007) 807–810
29. Christel, M.G.: Visual digests for news video libraries. In: Proceedings of the aCM Multimedia Conference. (1999) 303–311
30. Stevens, S., Christen, M., Wactlar, H.: Informedia: improving access to digital video. Interactions **1**(4) (1994) 67–71
31. Yates, R.B., Neto, B.R.: User Interfaces and Visualization. In: Modern Information Retrieval. Addison Wesley (1999) 257–324
32. Zhang, J.: Visualization for Information Retrieval. Volume 23 of The Information Retrieval Series. Springer (2008)
33. Aiello, A., Silveira, R.I.: Trazado de grafos mediante métodos dirigidos por fuerzas: revisión del estado del arte y presentación de algoritmos para grafos donde los vértices son regiones geográficas. Technical report, Departamento de Computación, Universidad de Buenos Aires (2004)
34. Yee, K.P., Fisher, D., Dhamija, R., Hearst, M.: animated exploration of dynamic graphs with radial layout. In: Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01), Washington, DC, USA, IEEE Computer Society (2001) 43
35. Shneiderman, B.: Tree visualization with tree-maps: 2-d space-filling approach. ACM Transactions on Graphics **11**(1) (1992) 92–99
36. Santini, S.: Who needs video summarization anyway? In: Proceedings of the International Conference on Semantic Computing. (September 2007) 177–184

Métodos del Análisis de Redes Sociales Aplicados a la Recolección de un Crawler

Loretto Peirano Cordano y José Luis Alfonso Berrocal

Departamento de Informática y Automática
Universidad de Salamanca

Resumen La presente investigación tiene como objetivo principal comprobar la factibilidad de aplicar metodologías asociadas al análisis de redes sociales sobre datos que han sido recolectados a través de un Crawler, como así mismo, obtener resultados concretos sobre dichos datos. Los objetivos específicos apuntan a conocer y comprender las metodologías que componen el análisis de redes sociales para poder ser aplicadas. Para lograr estos objetivos, es necesario conocer las herramientas informáticas que apoyan el análisis de redes sociales, sobre todo de grandes redes las cuales sin el apoyo de la velocidad de cálculo que ofrecen los computadores no sería posible estudiar.

1. Introducción

Para realizar un análisis de las redes es primordial en primera instancia recopilar la información necesaria. Para generar una red del tipo “social” es necesario hacerlo a partir de los datos reticulares, los cuales están representados a través de los actores y sus relaciones más que de sus atributos.

Las redes resultantes requieren herramientas que apoyen el proceso de obtención de resultados, digo grandes redes ya que difícilmente en el área de red social se trabaja con muestras, en general, es necesario trabajar con un censo completo de la población.

Como herramienta de recopilación de información se utilizó un Crawler para generar una red, las semillas iniciales corresponden a 110 páginas web relacionadas con la alfabetización digital y la resultante es una red conformada por 30.779 páginas y sus hiperenlaces entre éstas, el Crawler rastrea los hiperenlaces buscando nuevas páginas las cuales analiza buscando otros enlaces existentes y así sucesivamente, de esta manera es como se conforma la red.

El análisis de los datos fue realizado mediante una herramienta de software llamada Pajek, la cual está orientada al análisis de diversos tipos de redes utilizando algoritmos basados en los métodos del análisis de las redes sociales, esta herramienta dentro de sus utilidades se caracteriza por ofrecer algoritmos enfocados a distintos tipos de redes, ya sean del tipo dirigidas, no dirigidas, temporales, genealógicas, de pequeño, mediano o gran tamaño, etcétera, Esta investigación se centra principalmente en el estudio de las opciones vinculadas con el análisis de grandes redes del tipo dirigido que es la característica de la red en estudio.

2. Análisis de Redes Sociales

El análisis de redes sociales (ARS) se centra en las relaciones entre los actores y no en los actores individuales y sus atributos, entonces los actores no son muestreados de forma independiente si no que se deben considerar las relaciones que tiene con otras personas y considerarlos a ellos también dentro de la obtención de datos. [1]

El ARS combina las áreas de sociología, psicología, antropología, estadística y matemáticas, entre éstas empezaron a moldear y respaldar las ideas iniciales en el estudio de las relaciones y los individuos, los grafos comenzaron a jugar un papel importante y a dar forma a las ideas, posteriormente fueron incorporadas las matrices las cuales permiten el análisis de grupos mas grandes.

Aún así, fue necesaria la incorporación de las herramientas tecnológicas para completar el ciclo, con esto comienzan a desarrollarse programas computacionales para realizar el análisis de grandes volúmenes de datos el cual no era posible realizar manualmente.

2.1. Teoría de Redes

Algunos de los conceptos asociados a la teoría de redes son los siguientes: [2]
Visión Relacional: Análisis construido por medio de los lazos o vínculos entre los actores he intenta explicar como las relaciones crean diferencias de poder y prestigio entre las personas.

Formalización: La teoría de los grafos no es el único formalismo utilizado para modelar las redes aunque si fue uno de los primeros, Otro mecanismo es a través de la teoría matricial la cual es de gran ayuda para análisis de grandes redes de datos.

Niveles de Análisis: Existen diversas formas de analizar redes, sin embargo, una buena estrategia es hacer un análisis de forma conjunta sobre éstas.

Nivel egocéntrico: Se intenta explicar las diferencias entre actores según las diversas posiciones sociales, mostrando a los individuos y a las redes que los rodean. Se destaca que este método NO conduce a una descripción completa de la estructura social.

Nivel Intermedio: Se analizan las unidades entre si, se hace un análisis entre las unidades y la red parcial o total. Generalmente el estudio esta basado en díadas pero puede extenderse a subconjuntos de la red.

Nivel de Estructura Completa: Consiste en comparar entre si estructuras sociales completas, por lo general se requiere disponer de todos los datos concernientes a los lazos que vinculan una estructura con otra, la enumeración completa de la población es a su vez esencial.

Perpectiva en el ARS. Se considera no solo un elemento en la red, si no que se observa como éste se coloca dentro del sistema, también se tienen en cuenta otros factores como los roles y grupos, estructuralismos v/s individualismo, por ejemplo el desempeño en el grupo, la posición que indica las oportunidades y limitaciones, además considera la hipótesis de que el capital social prevalece sobre

el capital humano. [3]

Relacional: Se basa en conexiones directas e indirectas que existen entre unidades de una red y están más focalizados sobre el análisis micro con énfasis en díadas y triadas

Posicional: Se basa sobre la similaridad en las pautas de relaciones de unas unidades con otras. Esta centrado en el análisis macro, describiendo estructuras sociales globales.

Mundo Pequeño: Es el tipo de red donde el ARS ha tenido resultados mas fructíferos, comenzando por los experimentos de Milgram, donde se indica que dos personas desconocidas están a una distancia de 5 enlaces. Este fenómeno se ha difundido a otras redes tales como: redes de artistas de cines, redes tróficas marinas, redes de computadores, redes neuronales, etcétera, estas redes no pequeñas estarían a una distancia media de 4 pasos.

Las propiedades de una red de mundo pequeño es de un diámetro pequeño y una alta fragmentación al igual que los grafos aleatorios, otra característica de las redes mundo pequeño es que cuando se agregan mas nodos su diámetro aumenta lentamente, entendiéndose por lento un avance del diámetro de la red de forma logarítmica con respecto a los nodos agregados. [4]

Formas de Crecimiento de las Redes. *Enlace Aleatorio:* Consiste en que cuando se anexa un nuevo nodo, éste se enlaza aleatoriamente a otro nodo de la red. Existen muy pocas redes que tengan realmente este comportamiento.

Enlace Preferencial: Consiste en enlazar un nuevo nodo a un mejor nodo, un mejor nodo se determina por la cantidad de enlaces que ya tiene y corresponde a los nodos con un mayor número de enlaces, por lo que resultaría un modelo poco equitativo.

Componente Gigante: es una característica de la red en la que un grupo de nodos agrupa a la mayoría de los nodos en la red, este fenómeno se da en la mayoría de las redes sociales.

Ley de Potencias: una frase representativa sería que el “20% de la población tiene el 80% de las riquezas”, esto se da en las redes sociales donde los nodos que disponen de una mayor cantidad de enlaces son la minoría.

Ley de Zipf: Se aplica sobre las conexiones de nodos, donde el nodo mas enlazado es un múltiplo fijo del segundo mas enlazado, éste a su vez del tercero mas enlazado y así sucesivamente, por ejemplo, la cantidad de visitas que tienen las páginas web.

En las redes sociales es posible aplicar leyes de potencias donde habrá algunos nodos con muchos enlaces y muchos nodos con ningún enlace. Por otro lado se considera que las redes exponenciales son más vulnerables ya que basta atacar los nodos centrales para eliminar la conectividad de toda la red. Algunos consideran que para mantener una red limpia se puede enfocar la desinfección sobre los nodos centrales quienes tienen la mayor cantidad de enlaces.

Algo a considerar es que los nodos con más enlaces entrantes y salientes podrían no ser los más importantes, por ejemplo, en una red donde se necesite ir de un punto a otro serán más importantes aquellos nodos que no se pueden “evitar”, a

esta inevitabilidad se le denomina centralidad o *betweenness* o la traducción al castellano como enmediedad, por tanto, es interesante considerar la enmediedad o *betweenness* en la red ya que cualquier ataque a un nodo inevitable afectaría las conexiones con gran parte de la red. [5]

Propiedades de las Redes. La estructura y el comportamiento de las redes está basado y activado por las interacciones locales entre los nodos.

Las propiedades básicas de las redes sociales tienen importantes consecuencias. Normalmente, algunos nodos tienen muchas conexiones mientras que otros tienen pocas, especialmente cuando las poblaciones se convierten en grandes y no todas las conexiones posibles están presentes existen “agujeros estructurales”. entonces surge el cuestionamiento de ¿Hasta qué punto los individuos están conectados unos con otros y hasta qué punto la red en su conjunto está integrada?. Los nodos individuales pueden tener pocos o muchos lazos. Los individuos pueden ser “fuentes” de relaciones, “agujeros” (nodos que reciben pero no emiten) o ambos. Esas clases de diferencias básicas entre las conexiones inmediatas de los nodos pueden ser críticas explicando cómo ellos ven el mundo y cómo el mundo los ve a ellos. El número y clase de lazos que los nodos tienen son claves para determinar hasta qué punto su inclusión en la red limita su conducta y el rango de oportunidades, influencia y poder que tienen. Si no es posible que un grupo de nodos alcance a otro, querrá decir que la red esta formada por mas de un grupo.

La regla de tamaño existente en redes es: para una red de tamaño N donde cada nodo esta conectado a cada uno de los otros nodos incluyendo la conexión consigo mismo, entonces se dice que existen N^2 conexiones posibles.

Si cada unidad está conectada a cada una de las otras unidades excepto a si misma entonces el número de conexiones posibles es $N*(N-1)$.

Algunas características elementales existentes en las redes serían: [5], [6]

Distancia: es la longitud entre un punto y otro, podría por ejemplo expresarse en kilómetros.

Enmediedad o Betweenness Cantidad de nodos por los que hay que pasar para llegar a un destino.

Es interesante considerar la distancia si lo que se pretende es rapidez para entregar una información, en ese caso se escoge el camino mas corto o geodesia, pero si lo que se pretende es propagar información por la red, valdrá mas tener un mayor alcance y que se propague por múltiples caminos a que tan rápido pueda llegar la información.

Conexión Débil: Representa la cantidad de caminos que se disponen para transmitir un mensaje, si existiesen pocas opciones, entonces se dice que la conexión es débil.

Hubs y Autoridades: Un Hub es aquel nodo desde donde salen las conexiones y una autoridad es el nodo que recibe las conexiones.

Media: Indica la cantidad de lazos que están presentes en relación al máximo posible.

Desviación Estándar: Representa la cantidad de elementos que son diferentes, si

todos los elementos fueran 1 entonces la desviación sería de 0.

Accesibilidad: Un nodo es “accesible” por otro si existe un conjunto de conexiones mediante las cuales se pueda trazar un camino.

Centralidad de Grado: Tiene relación con la cantidad de enlaces que posee un nodo, cuando éste tiene muchos lazos de entrada se dice que es un nodo de prestigio y cuando posee muchos enlaces de salida se dice que es un nodo influyente.

Centralidad de Cercanía: Es la distancia de un nodo con los demás.

Centralidad por grado de intermediación: son nodos que se encuentran en el camino geodésico entre otros pares de nodos, es decir, sirve de intermediario.

Centralidad de Flujo: Se refiere a la cantidad de caminos que se pueden tomar para llegar a otro nodo sin tener que optar únicamente por el camino geodésico. ¿Qué define el poder?: la cantidad de conexiones, la cercanía, y el grado de intermediación, a este conjunto lo denominan centralidad (ha existido debates de cómo denominarlos pero varios coinciden en que una posición central es una posición de poder).

Índice de poder de Bonacich: Hace una modificación a la centralidad de grado, la cual tiene relación con la cantidad de enlaces que se poseen. Bonacich sostiene que el tener mas enlaces no necesariamente hace a un nodo mas poderoso, si no que además deben ser consideradas las conexiones de sus conexiones.

Proximidad: Se refiere a la posibilidad de que un nodo se relacione con otro nodo pero a través de un número de pasos dentro de la red. Entre mas central es el nodo mayor es su grado de cercanía, esto significará que necesitará en promedio un menor número de pasos para llegar a los otros nodos.

El índice de centralidad-proximidad se puede representar así, donde D es la distancia es $D_{i+} = \sum_{j=1}^N D_{ij}$ De este modo el índice es mayor cuando aumenta la proximidad

La centralidad vista como mediación se define como el nivel en que otros actores deben pasar a través de un actor focal para comunicarse con el resto de los actores. El valor de la mediación para un actor mide la proporción de las geodésicas que pasan por él y se representa como:

$$C_B(K) = \frac{2 \sum_{i=1}^n \sum_{j=1}^n (g_{ik}(k)/g_{ij})}{n^2 - 3n + 2}$$

Para todos los puntos no ordenados, i,j,k , donde $i < j$, n es el número de nodos de la red y $g_{ij}(k)$ es el número de geodésicas (caminos más cortos) entre i y j , que pasan por k . Por tanto si k está en el camino más corto del par (i,j) , K tiene alta centralidad-mediación.

El grado representa el nivel de la actividad comunicativa (la capacidad de comunicar directamente con otros); la proximidad (Closeness) representa la independencia (la capacidad de llegar a muchos de los otros miembros de la red directamente, esto es sin apoyarse en intermediarios), mientras que la mediación (Betweenness) representa el control de la comunicación de otros y su capacidad de restringirla.

2.2. Trabajando con Redes

“Para los analistas de redes los actores se describen a través de sus relaciones, no de sus atributos. Y las relaciones en sí mismas son tan fundamentales como los actores que se conectan a través de ellas”. [7]

Cuando se hacen estudios de redes sociales en general se trabaja con un censo completo y por tanto es muy importante tener en cuenta los límites de la población a estudiar y la forma que se seleccionaran dentro de la población, además cuando un nodo es considerado se debe incluir a todos los nodos que tienen relación con él.

Muestreo de Relaciones. Existen diversos métodos para hacer una recogida referente a las relaciones que vinculan a los actores, algunos de estos métodos son los siguientes: [7]

Método de Redes completas: Consiste en recoger la información de los lazos que tiene un actor con los demás, esto significa realizar un censo de los lazos que existen entre los actores, para esto se recolecta la información sobre los lazos entre pares o díadas, estos datos aportan una fotografía completa de los lazos existentes en una población determinada.

Método Bola de Nieve: El nombre es bastante representativo, este método consiste en seleccionar un actor o un conjunto de actores y se les consulta sobre los lazos que tienen con otros actores, posteriormente se le realiza la misma consulta a los actores que fueron mencionados (pero que no estén en la lista anterior) y se les realiza la misma consulta, de esta manera se “hecha a rodar la bola de nieve”, este proceso se detiene cuando se decida hacerlo o cuando ya no se han encontrado nuevos actores.

Escala de Medidas. Consiste en asignarle puntuaciones a las relaciones, los niveles de medición consiste de tres tipos: nominal, ordinal y de intervalo, la medida nominal puede ser dividida en binario o de categoría múltiple. [8]

Medida Binaria: Consiste en colocar 0 ó 1, se pone un 1 cuando existe la relación y un 0 cuando no existe, es uno de los métodos mas usados, la desventaja es que no indica que tan fuerte es la relación.

Medida Nominal de Categoría Múltiple: Consiste en una selección múltiple de relaciones donde dado un conjunto de relaciones, estas son enumeradas y cada actor deberá indicar cual es la relación mas fuerte que posee con los otros actores, desde luego dentro de las opciones existirá la alternativa “ninguna relación”. Es posible después crear conjuntos separados por lazos. La desventaja de este tipo de escala es que solo se permite seleccionar una relación, por ejemplo, amistad, parentesco, esto significa que las densidades pudiesen bajar artificialmente.

Medida Ordinal Agrupada por Relaciones: Mide el índice de agrado, desagrado o indiferencia entre actores, usualmente se usan los valores +1, -1 y 0 para representar esta medida. Este método ayuda a medir la fuerza de los lazos y la reciprocidad de las relaciones, comúnmente la reciprocidad y la fuerza están ligadas, es decir, si la relación es recíproca es posible que el lazo existente sea

fuerte.

Medida Ordinal de Ranking de Relaciones: El actor debe indicar para cada tipo de relación la posición que ocupa cada actor en la fuerza de la relación, es decir, en primer lugar se coloca quién agrada más y así sucesivamente, esto reflejaría la posición de los nodos y la diferencia de intensidad en las relaciones.

Medida de Intervalo de Relaciones: Es la forma más avanzada y permite discriminar lazos, por ejemplo, indicar que el lazo X es dos veces más fuerte que el lazo Y. Estos lazos son valorados dentro de escalas, donde la diferencia entre 1 y 2 es la misma diferencia que 15 y 16, caso que no ocurriría con el método anterior.

Cómo Utilizar los Métodos del ARS. Las herramientas utilizadas para trabajar con redes sociales son los grafos y las matrices.

Los grafos se pueden clasificar en orientados, binarios, ponderados o de relaciones múltiples.

Las matrices aventajan a los grafos cuando se trabaja con grandes volúmenes de información ya que con muchos datos los grafos se vuelven más difíciles de visualizar.

La matriz de adyacencia es la más común para la representación en redes sociales y está compuesta tanto en sus filas como en sus columnas por la cantidad de actores existente para el análisis y donde los elementos representan los vínculos entre los actores. La más simple de todas es la matriz binaria donde se representa con un 1 la existencia de la relación y con un 0 la no existencia. Cabe señalar además que el actor origen del vínculo va en la fila y el actor objeto en la columna. Se observa además que si todas las relaciones existentes en la matriz fuesen recíprocas entonces la matriz resultante sería una matriz simétrica.

Al igual que en los grafos, las relaciones positivas negativas o nulas se representan con un +1, -1 ó 0. Y cuando se miden las relaciones otorgándoles un valor ordinal entonces ese número es el que se coloca en los elementos de la matriz.

3. Crawler

El Crawler fue la herramienta utilizada para realizar la recolección de información, el resultado final entregó una red compuesta de 30.779 nodos o páginas. Los puntos iniciales utilizados sobre el Crawler corresponden a 110 páginas web relacionadas con la alfabetización digital, estos puntos de entrada no fueron seleccionados al azar, si no que corresponden a un estudio realizado por terceros y que actualmente se encuentra en proceso, este estudio ha determinado cuáles son las 110 páginas de mejor calidad existentes en el universo web relacionadas con el tema de alfabetización digital.

Un Crawler denominado también como robot o araña es un componente de un motor de búsqueda (o un programa independiente) que trabaja bajo el puerto 80, su función es inspeccionar la web creando copias de las páginas visitadas las que posteriormente serán indexadas por un motor de búsqueda logrando así un sistema de búsquedas más rápido.

El trabajo del Crawler consiste en visitar un conjunto de páginas e identificar

todos los hiperenlaces existentes en éstas, tanto, dentro del mismo dominio como fuera de él y las añade a una lista de páginas a visitar para posteriormente analizar los hiperenlaces de éstas y así sucesivamente hasta que ya no aparezcan nuevas páginas web ó haya llegado al término indicado como máximo de páginas a analizar.

Hay tres características importantes de la Web que hacen dificultoso el rastreo:

1. su gran volumen
2. su acelerado ritmo de cambio
3. La generación de páginas dinámicas

Por tanto, dado los problemas descritos anteriormente y sumando a esto el hecho de que el ancho de banda no es ilimitado ni exclusivo, el comportamiento de un Crawler es el resultado de una combinación de políticas tales como:

1. Una política de selección que indique que páginas descargar
2. Una política sobre los estados de las páginas para comprobar la existencia de cambios
3. Una política para no recargar los sitios web
4. Una política para distribuir coordinadamente los Crawler en la red

Dado que los Crawler pueden recuperar los datos más rápidamente y en mayor profundidad que las personas, el uso y abuso de ellos podría provocar un “efecto paralizante” sobre el desempeño de un sitio. El tiempo que un servidor invierte en atender las múltiples peticiones que un solo Crawler puede realizar por segundo es considerable, más aún, si debe atender las peticiones de múltiples Crawler a la vez.

El uso de los rastreadores web es útil para una serie de tareas, pero está sujeto a un precio que debe pagar la comunidad en general. Algunos de los costos a considerar serían: [9]

1. Consumo de los recursos de la red, dado el considerable ancho de banda que requieren los Crawler y el alto grado de paralelismo con que operan
2. Pueden sobrecargar un servidor si la frecuencia de consultas es muy alta
3. Los Crawler mal diseñados podrían llegar a “colgar” un servidor o un Router
4. Los Crawler personales utilizados por muchas personas simultáneamente podrían llegar a interrumpir los servidores web

4. Análisis de Datos

Como ya se ha descrito anteriormente los datos recolectados a través de un Crawler corresponden a páginas web relacionadas con la alfabetización digital, el resultado obtenido corresponde a una red del tipo dirigida, esta red además está conformada por relaciones asimétricas, es decir, donde la relaciones no son necesariamente correspondidas entre los vértices, en este caso en particular desde una página web hacia otra. La herramienta de software seleccionada para realizar el análisis de esta red fue Pajek.

Para analizar una red hay que saber desde luego con que tipo de red estamos tratando, las redes con 10 vértices son llamadas pequeñas redes, aquellas redes con mas de 1000 nodos y conexiones son llamadas grandes redes, en nuestro caso existe un total de 30.779 nodos por lo tanto se esta hablando de una gran red. Una red es llamada “Sparse” cuando el número de líneas es igual al número de vértices, y cuando el número de líneas es mayor al número de vértices entonces la red se denomina “Dens”, para este caso concreto se trata de una red del tipo “Dense”.

La información básica que se puede obtener de una red es:[10]

1. Cantidad de vértices
2. Cantidad de líneas dirigida
3. Cantidad de líneas no dirigidas
4. Cantidad de loops dirigidos
5. Cantidad de loops no dirigidos
6. Densidad con y sin loop, en pajek se representan como densidades 1 y 2

La red a analizar contiene 30.779 vértices y 471.212 conexiones entre vértices, como esta red corresponde al tipo de red dirigida la totalidad de líneas son de este tipo. El tiempo de carga de la red fue de 5 segundos, la densidad 1 y 2 es de 0,0004974, ambas son iguales ya que no existen loops en la red.

Con 30.779 nodos existentes en la red, el número máximo de líneas posibles existentes hubiese sido de 947.316.062, la densidad la puedo determinar dividiendo el número real de líneas por el número posible de líneas, esto sería $472.212/947.316.062$ lo que me da un total de 0,000497417 correspondiente al resultado indicado por Pajek.

También se observa analizando los componentes strong/weak que existen 36 grupos dentro de la red de los cuales 10 de ellos pertenecen a grupos aislados y donde uno de ellos posee mas de la mitad de los nodos integrantes de la red con 28.144 páginas vinculadas a este grupo y las cuales están relacionadas de alguna manera.

Esta red fue sometida a múltiples observaciones las cuales entregaron información bastante interesante, se destaca dentro de éstas las siguientes:

1. Existen 28 nodos que no pueden ser accesados desde ningún otro nodo, en el sentido práctico desde ninguna otra página web.
2. El mayor número de influencia detectado es de 21.210, la cantidad de nodos que posee esa influencia es de 140
3. Se observa que la siguiente mayor influencia existente es de 17.716
4. Se observan grandes saltos en el valor de influencia, donde desde una influencia de 17.347 salta a una influencia de 4.558 desde ahí disminuye paulatina-mente hasta 3.720 para posteriormente dar un brinco hasta una influencia de 654 desde ahí existe una disminución paulatina con algunos pequeños saltos en los valores de la influencia.
5. Las páginas web que poseen una influencia menor de 12 corresponden a un total de 89 páginas.

6. La mayor proporción de influencia corresponde a 0,689 considerando a 1 como el máximo posible, es decir, más de la mitad de la red puede acceder de algún modo a estas páginas.
7. La mínima distancia entre una página y otra es desde luego 1 la cantidad de páginas que tienen un promedio de distancia de 1 es de 82 páginas.
8. La mayor distancia promedio detectada considera una longitud de 41 y corresponde al vértice 7497
9. El promedio total de distancia entre páginas que existe en la red es de 7 enlaces.
10. La página con mayor proximidad por prestigio corresponde al vértice 3.381 con una proximidad de 0,35
11. La siguiente página posee un prestigio de 0,304 y corresponde al vértice 3.199
12. En general observando las 10 páginas con mayores proximidad por prestigio, éstas corresponderían a la web de <http://portal.unesco.org/ci/en/>.
13. La menor proximidad por prestigio detectada después de 0 corresponde a 0,0000325 y la poseen 9 vértices

Con respecto a los hubs y autoridades, se determinan que los vértices que cumplen estos roles son los siguientes: 4161, 4128, 4121, 4117, 4111, 4104, 4100, 4095, 4081, 3381. Todos estos vértices son los mejores en ser hub y autoridad, estos corresponden al dominio de: portal.unesco.org/ci/en/

5. Conclusiones

Al inicio de esta investigación se describió como uno de los objetivos principales determinar si era posible aplicar los métodos del Análisis de Redes Sociales (ARS) a la recolección de datos que realiza un Crawler, como también obtener resultados concretos al aplicar las metodologías correspondientes sobre dichos datos.

Después de haber estudiado y comprendido las técnicas del ARS se puede decir con seguridad que no tan sólo es posible aplicar estas metodologías sobre los datos recolectados a través de un Crawler que era el objetivo en cuestión, si no que además es una herramienta muy útil para análisis de grandes redes, se debe considerar además, que el realizar un análisis manual sobre estas redes sería prácticamente imposible si se quisiesen obtener resultados en un corto periodo de tiempo, por lo tanto, las intervenciones de las herramientas computacionales del ARS resultan tremendamente benéficas.

En general, conociendo bien la estructura del archivo con el cual trabajan las herramientas de redes sociales, es posible capturar los datos desde un Crawler y transformarlos a un formato legible por el programa de redes sociales a utilizar y de este modo generar la red a estudio, una vez que se obtiene el formato requerido, la tarea de obtener respuestas se vuelve mas simple, o por decirlo de otra manera, cualquier red que pueda ser representada por una matriz podría aplicar técnicas del ARS. Desde luego, la utilidad de aplicar esta metodología estará vinculada a la conformación de la red, a los objetivos que se persiguen y a las preguntas que se quieren responder.

Así como los sociólogos utilizan como herramientas de recopilación de datos los cuestionarios, entrevistas, observaciones, rastreos de llamadas telefónicas, etcétera, las cuales son aplicadas a un conjunto inicial de actores para posteriormente aplicarlas en cadena sobre los actores señalados por los primeros hasta llegar a un límite previamente establecido o hasta que no aparezcan nuevos actores. De este mismo modo, un Crawler recolecta la información, trabajando con un conjunto inicial de datos previamente definido y estudiando sus enlaces con otras páginas, el equivalente a las relaciones en sociología, para posteriormente estudiar los enlaces de las páginas que fueron señaladas por los primeros y así sucesivamente hasta llegar al límite previamente establecido o hasta que no aparezcan nuevas páginas, si se realiza una comparación se puede decir que un Crawler tiene un comportamiento similar a los sociólogos al momento de recopilar la información para conformar una red. Por lo tanto, su forma de recopilación lo hace una herramienta capaz de generar una red que pueda ser sometida a análisis a través de las metodologías del ARS.

Referente al estudio de la red generada a partir de las 110 páginas web de alfabetización digital cuya resultante después de aplicar el Crawler corresponde a una red conformada por 30.779 páginas y 471.212 hiperenlaces, se indica que del análisis de ésta se desprende una longitud media de 7 hiperenlaces para llegar desde una página web hasta casi cualquiera de las 30.779 páginas que componen esta red, digo casi por que existen algunos grupos aislados a los que no es posible llegar. Esta longitud media es bastante similar a algunos experimentos que se han realizado a lo largo de la evolución del análisis de redde sociales, de hecho, sigue la tendencia de éstos.

Otro aspecto destacable es que dentro del estudio se observa que quiénes obtuvieron mayores puntuaciones como nodos (páginas) hubs, autoridades, de prestigio y en general aquellas que son más centrales dentro de la red, son las páginas pertenecientes al dominio de <http://portal.unesco.org/ci/en>

Referencias Bibliográficas

1. Molina, J.L.: El análisis de redes sociales Una introducción. Edicions Bellaterra (2001) ISBN: 84-7290-160-2.
2. Lozares, C.: La teoría de redes sociales. Technical report (1999)
3. Borgatti, S.: Conceptos básicos de redes sociales, Boston College (Febrero 2003) Presentación.
4. Molina, J.L.: La ciencia de las redes. Apuntes 11, Departamento de Antropología Social, Universidad Autónoma de Barcelona, joseluis.molina@uab.es (Junio 2004) Asociación para el Avance de la Ciencia y la Tecnología en España (AACTE).
5. Guervós, J.J.M.: Redes sociales: Introducción. Depto. Arquitectura y Tecnología de Computadores (2005) jj@merelo.net.
6. Menéndez, L.S.: Análisis de redes sociales: o como representar las estructuras sociales subyacentes. Apuntes de Ciencia y Tecnología **7** (2003)
7. Hanneman, R.A.: Introducción A Los Métodos Del Análisis De Redes Sociales. (2000)
8. Hanneman, L.R.I.R.A.: Introduction to the formal analysis of social networks using mathematica. Technical report (2006)

9. Crawler, W.: http://es.wikipedia.org/wiki/web_crawler (2008)
10. Batagelj, V., Mrvar, A.: Pajek program for large network analysis. Technical report, University of Ljubljana (1999)
11. Alejandro Velázquez, N.A.: Manual introductorio al análisis de redes sociales. (Junio 2005)
12. Anuska Ferligoj, Patrick Doreian, V.B.: Optimizational approach to blockmodeling. Technical report
13. Batagelj V., M.A.: <http://pajek.imfm.si/doku.php> (Mayo 2008)
14. Batagelj, V.: Networks / pajek program for large network analysis (2008)
15. Batagelj, V.: Analysys of large network with pajek. (2007)
16. Batagelj, V.: Notes en blockmodeling. (1994)
17. Borgatti: Concepto de teoría de grafos. (2003)
18. Borgatti, S.P.: 2-mode concepts in social network analysis. Technical report, University of Kentucky (2007)
19. Carvajal, A.S.G.M.C.R.: Las redes sociales como herramienta de análisis estructural input-output. Technical report, Universidad de Oviedo (2003)
20. Christian Kohlschutter, P.A.C., Nejd, W.: Efficient parallel computation of pagerank. (2006) 241–252
21. Dirson, G.: Page rank (2008)
22. Doreian, P.: Event sequences as generators of social network evolution. *Social Network* **24** (2002) 93–119
23. Espinoza, J.F.N.: Manual para operaciones básicas con visone en el análisis de redes sociales para el desarrollo rural, un acercamiento a una herramienta de apoyo para la evaluación de proyectos de desarrollo rural. (2008)
24. Everton, S.F.: A guide for the visually perplexed: Visually representing Social networks. Stanford University. (2004)
25. Figaredo, L.A.D.D.D.: Análisis de redes sociales (ars) aplicado a foros didácticos virtuales. Technical report (2004)
26. Guimarães, P.: Tutorial for Inputting an Interaction Matrix into Pajek. Universidad Estadual de Campinas, Brasil. (2007)
27. Josep Rodriguez, F.M.: Guia practica de analisis de redes. (2004)
28. José Luis Molina (ed., UAB), g.Q.U.J.M.U.I.M.J.U.A.d.F.U.: Talleres de autoformación con programas informáticos de análisis de redes sociales. Universitat Autnoma de Barcelona (2006)
29. José Luis Molina, Juan Manuel Muñoz, M.D.: Redes de publicaciones científicas: un análisis de la estructura de coautorías. *Revista Redes* **3** (2000)
30. Junghoo Cho, H.G.M.: The evolution of the web and implications for an incremental crawler. (1999) 18
31. León, A.A.R.: Representación de series temporales con pajek. Technical report, Laboratorio de Redes, IIMAS-UNAM. (2002)
32. León, A.R.: Redes: teoría y práctica. Technical report, Unam, México (Agosto 2007)
33. Lozares, C.: Valores, campos y capitales sociales. *REDES- Revista hispana para el análisis de redes sociales* **2**(2) (Junio 2003) 30
34. Manz: Robot.txt todo lo que debería saber (Agosto 2007)
35. Miceli, J.E.: Los problemas de validez en el ars. (2007)
36. Molina, J.: Operaciones Básicas con Ucinet 6. (2005)
37. Molina, J.L.: Medidas de centralidad. (2002)
38. Molina, J.L.: Definiciones elementales: densidad, distancias. centralidad. cohesión. posiciones y estructuras. In: Presentación. (2000)

39. Molina, J.L.: Introducción al análisis de redes sociales. (1999)
40. Mrvar, A.: Network analysis using pajek. Technical report (2004)
41. Quijada, D.F.: El análisis de redes sociales aplicado al estudio de la estructura de las industrias culturales. (2008) 16
42. águeda Quiroga: Introducción al análisis de datos reticulares prácticas con ucinet6 y netdraw1 versión 2. Technical report, Universidad Pompeu Fabra (Septiembre 2003)
43. Ricardo Baeza, C.C.: Wire - web information retrieval environment (Abril 2004)
44. Robles, G.: Introducción al análisis de redes (sociales). (2004)
45. Roldán, C.L.C.J.M.V.P.J.M.O.P.L.: Relaciones, redes y discurso: revisión y propuestas en torno al análisis reticular de datos textuales. Technical report, Universidad Autonoma de Barcelona (2001)
46. Santos, F.R.: Análisis de Redes Sociales. Centro de Investigaciones Sociologicas, CIS (2003)
47. SocSciBot: Socscibot 4 and socscibot 3 (2008)
48. Vladimir Batagelj, A.M.: Program for Analysis and Visualization of Large Networks. (Abril 2008)
49. Wikipedia: Web crawler (Agosto 2008)
50. Wolfe, A.W.: The rise of network thinking in antropology. *Social Network* 1(1) (1978) 53–64 University of South Florida.

Stereo-MAS: Multi-Agent System for Image Stereo Processing

Sara Rodríguez and Juan M. Corchado

Departamento de Informática y automática
Universidad de Salamanca
{srg, corchado}@usal.es

Resumen El artículo presenta una arquitectura distribuida basada en agentes que permite procesar la información visual obtenida por cámaras estereoscópicas. El sistema está inmerso dentro de un proyecto global que tiene como objetivo el desarrollo de un entorno inteligente para la localización e identificación en entornos de dependencia donde confluirán otra serie de tecnologías. Los algoritmos de visión son muy costosos y toman tiempos de respuesta muy altos, lo que en nuestro caso es un gran inconveniente si consideramos que muchas de las aplicaciones pueden requerir operación en tiempo real. La arquitectura de agentes pretende automatizar el proceso de análisis de las imágenes obtenida por las cámaras y optimizar su procesamiento.

Palabras clave: estereoscopia, cámaras estereoscópicas, visión artificial, SMA, agentes, análisis correspondencia, identificación, entornos de dependencia.

1. Introducción

Uno de los mayores retos para Europa y su comunidad científica es encontrar vías más efectivas para proveer cuidados al creciente número de personas que forman los sectores de discapacitados y ancianos [6]. Los sistemas multi-agente y las arquitecturas basadas en dispositivos inteligentes han sido recientemente explorados como sistemas de supervisión de cuidado médico para ancianos y personas dependientes, por lo que estos sistemas podrían proveer apoyo constante en su vida diaria [1] [7].

El problema de la visión artificial, en concreto de la visión estereoscópica, ha recibido una considerable atención en los últimos años en la comunidad científica. Las aplicaciones del procesado de imagen son muy variadas, e incluyen aspectos como la medición remota, el análisis de imágenes biomédicas, la simulación de cirugía guiada remota, el reconocimiento de caracteres, aplicaciones de realidad virtual y realidad aumentada en sistemas colaborativos, entre otros.

El tema del presente trabajo se engloba dentro de un proyecto de investigación en el que se viene trabajando dentro del Grupo de Biomedicina, Sistemas Informáticos Inteligentes y Tecnología Educativa (BISITE) de la Universidad de Salamanca y que tiene como objetivo desarrollar un sistema de inteligencia ambiental [6] para el cuidado y supervisión de pacientes en entornos de dependencia, proporcionando un entorno capaz de realizar tareas de localización, identificación y monitorización de pacientes de forma automática. Dicho entorno permite

además, la supervisión de los pacientes y la simulación de situaciones, por parte del personal médico, a través de un entorno virtual. Para alcanzar este objetivo se utilizan técnicas de inteligencia artificial, agentes inteligentes y tecnologías inalámbricas. La infraestructura tecnológica consiste en la utilización de dispositivos ZigBee [28] para identificar y localizar a los usuarios, cámaras de vigilancia y tecnología Wi-Fi. Todos estos dispositivos serán distribuidos estratégicamente por el entorno de dependencia. Los agentes son capaces de reaccionar de acuerdo a las características de una situación determinada del entorno, creando nuevos comportamientos y fijándose nuevas metas.



Figura 1. Entorno Inteligente para la Localización e Identificación en Entornos de Dependencia.

En colaboración con las tecnologías inalámbricas empleadas, el sistema cuenta con un módulo de video vigilancia que permite extraer una representación automática en 3D de un entorno a partir de la información obtenida de cámaras estereoscópicas. asumiendo que se desconoce la situación de un paciente en un entorno en cada momento, sus movimientos, postura o estado, se plantea la posibilidad de unión de ambas tecnologías; complementando así la información de localización obtenida a través de los dispositivos ZigBee con la información de las cámaras. Esta cooperación permite reducir los costes técnicos empleados en la localización de personas dependientes, así como dar un mayor nivel de detalle de la misma.

Dentro del proyecto global, el sistema de localización por radiofrecuencia se encuentra ya en fase de finalización de su desarrollo y son los módulos de video-vigilancia, análisis y representación 3D los que centran actualmente la atención del grupo. En concreto, el presente trabajo se centra en el desarrollo de una arquitectura distribuida basada en agentes que permita procesar la información visual obtenida por las cámaras estereoscópicas. La arquitectura de agentes pretende automatizar el proceso de análisis de las imágenes obtenidas por las cámaras y optimizar su procesamiento. Paralelamente a este trabajo se está desarrollando la parte del proyecto que permite, una vez analizada la información recogida por las cámaras estereoscópicas, representarla en formato tridimensional así como interpretarla de manera adecuada para el reconocimiento de objetos y posturas de los personajes de la escena.

La descripción de este proyecto global junto con un resumen de las principales tecnologías empleadas son abordados en las secciones 1 y 2 de este artículo. En la sección 3 se hace hincapié en las técnicas utilizadas para el análisis de las imágenes. Una vez mostradas todas las herramientas, en las secciones 4 y 5 se mostrará la propuesta del sistema en sí y los resultados y conclusiones obtenidos.

2. Estado del arte

Sin duda alguna la visión es el más completo y a la vez complejo de los sentidos y el que permite al hombre interactuar con su medio ambiente y desenvolverse en él con un alto grado de eficiencia gracias a su adaptabilidad y robustez como sistema de percepción. Considerando la importancia de la información visual como fuente de datos del mundo real, resulta interesante plantear la posibilidad de proveer a un ordenador de dicho "sentido", lo que aunado a otros mecanismos tales como razonamiento y aprendizaje hagan de éste una herramienta capaz de interactuar eficientemente con el medio ambiente dinámico en el que se desenvuelve. En este apartado haremos hincapié en los mecanismos principales utilizados en nuestro sistema, la estereoscopia y los sistemas multi-agente (SMA).

Tradicionalmente, la técnica de recuperación de forma a partir de estereoscopia se ha dividido en dos problemas diferentes. Partiendo de dos imágenes bidimensionales en el espacio de coordenadas (u, v) , la imagen (I) izquierda y (D) derecha de la figura 2a, el *problema de la correspondencia* trata de buscar qué dos puntos $m_I(u_I, v_I)$ de la imagen izquierda y $m_D(u_D, v_D)$ de la imagen derecha corresponden a un mismo punto M del espacio tridimensional (X, Y, Z) . Una vez encontrados estos dos puntos, el *problema de la reconstrucción* trata de encontrar las coordenadas del punto M . En la figura 2a se observa el eje óptico que es la línea imaginaria ortogonal al plano imagen y que contiene el centro óptico del objetivo.

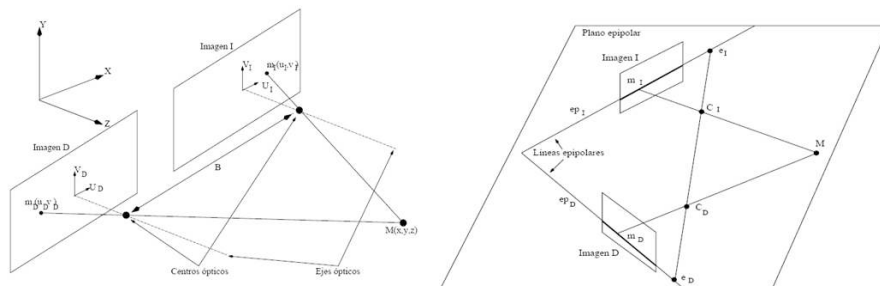


Figura 2. (a) Geometría estéreo: Marco de referencia y sistemas de referencia de las cámaras. (b) Restricción epipolar. Plano epipolar del punto P de la escena y líneas epipolares.

La tarea más difícil es sin duda responder al problema de la correspondencia. Como en general hay varias posibilidades para escoger el elemento correspondiente en la imagen D de un elemento de la imagen I , el problema de la correspondencia estéreo se dice que es ambiguo. Debido a esta ambigüedad, se hace necesario averiguar qué elementos, qué características, qué restricciones y qué consideraciones se pueden aplicar para reducirla al máximo. a ello se dedi-

cará una parte importante de este artículo, pero en primer lugar se explicará en este apartado el problema de la reconstrucción, que es un problema más cerrado y determinista.

En la figura 2b, consideramos plano epipolar aquel que forman los dos centros ópticos C_I y C_D de los objetivos de las cámaras con cualquier punto M del espacio objeto. Por otra parte, el plano epipolar ($C_I M C_D$) corta a las dos superficies imagen I e imagen D en sendas líneas rectas ep_I y ep_D , que se denominan líneas epipolares. Finalmente, la proyección del centro óptico de cada cámara sobre la otra cámara define el llamado *epipolo* e . Los epipolos de cada una de las cámaras (ep_I y ep_D) van a ser los puntos por los que van a pasar todas las líneas epipolares. Utilizando cualquier plano epipolar como base, todos los puntos de la escena que pertenecen a este plano tendrán su imagen en cada una de las líneas epipolares de las dos imágenes. Si consideramos que los planos retinales son coincidentes entre sí, y paralelos a la línea base (CI, CD) que une los dos centros ópticos de las cámaras, observaremos que los epipolos de ambas cámaras se situarán en el infinito, y por lo tanto, las líneas epipolares serán todas paralelas entre sí, y paralelas a su vez a la línea base. Esta configuración especial se denomina *configuración de cámaras paralelas*. Con esta configuración, y un adecuado posicionamiento de los planos retinales se puede conseguir que las líneas epipolares coincidan con las filas de las imágenes digitales tomadas [21].

El objetivo final de la reconstrucción es hallar las coordenadas del punto M (x, y, z) partiendo de las coordenadas de las proyecciones de dicho punto sobre los planos de imagen (u_I, v_I) y (u_D, v_D). Para ello se necesita la distancia focal f de las cámaras y la distancia entre sus dos centros ópticos o línea base B .

La reconstrucción tridimensional puede resolverse, bien mediante geometría proyectiva, o bien mediante geometría euclídea. En el caso de la geometría euclídea, que es la forma más sencilla y ofrece unos resultados suficientemente aproximados tenemos, para las imágenes I y D , por semejanza de triángulos:

$$\frac{u_I}{f} = \frac{x}{z}; x = \frac{u_I}{f} \cdot z; x = \frac{u_D}{f} \cdot z - B; y = \frac{v_D}{f} \cdot z; \frac{v_I}{f} = \frac{y}{z}; y = \frac{v_I}{f} \cdot z$$

Desarrollando estas ecuaciones podemos llegar a las expresiones siguientes:

$$x = \frac{u_I \cdot B}{d}; y = \frac{v_I \cdot B}{d}; z = \frac{f \cdot B}{d}; d = u_D - u_I$$

En estas expresiones el valor d es la denominada disparidad que hace referencia a la diferencia entre las coordenadas u_I y u_D respecto del centro de sus imágenes. Para la obtención de dichas coordenadas u_I y u_D se ha tomado como origen de coordenadas de cada imagen (U, V) el punto de corte del eje óptico y el plano retinal.

Con estas expresiones podemos concluir que, una vez conocidas la distancia focal de las cámaras, la línea base y la disparidad entre los píxeles correspondientes, es sencillo calcular las coordenadas (x, y, z) del punto M del espacio para la configuración de cámaras paralelas.

La cámara elegida para llevar a cabo el proceso de estereoscopia es la Bumblebee2, modelo BB2-COL-ICX424 (640x480 Color 3.8mm) del fabricante Point

Grey [16]. El sistema estereoscópico incluye dos minicámaras que generan dos secuencias de vídeo sincronizadas y permiten obtener información tridimensional del entorno mediante el estudio de las disparidades entre ellas.

Por otro lado, la utilización de agentes [4] deliberativos BDI (*Belief, Desire, Intention*) es una pieza fundamental en el desarrollo de la plataforma propuesta. aparentemente el sistema visual humano maneja un alto grado de especialización al clasificar y procesar la información visual que recibe, tal es el caso de recuperación de forma por textura, sombreado, profundidad etc. Computacionalmente es muy difícil competir con esta especialización y discriminar en una imagen sólo aquella información relevante para un propósito en particular. En respuesta a esta problemática se ha propuesto implementar estos algoritmos sobre una arquitectura distribuida basada en agentes que permita procesar en tiempo real la información visual contenida en una imagen.

Una arquitectura multi-agente distribuida presenta una mejor capacidad de recuperación ante errores, permitiendo utilizar agentes autónomos, capaces de tomar decisiones y de adaptarse al entorno por separado. Dentro de las arquitecturas deliberativas se encuentran las intencionales, y entre ellas el modelo más utilizado hoy en día es el BDI, en el cual, la estructura interna de los agentes y sus capacidades de elección, se basan en aptitudes mentales: creencias, deseos, e intenciones [19]. Una arquitectura distribuida basada en agentes, los cuales se ejecutan bajo demanda, permite mover el código a lugares donde las acciones son requeridas. Esto permite obtener respuestas en tiempo de ejecución, autonomía, continuidad de los servicios, así como mayores niveles de flexibilidad y escalabilidad que arquitecturas centralizadas [2]. Por otra parte, se reduce el esfuerzo de programar tareas múltiples, ya que sólo es necesario especificar objetivos globales para que los agentes cooperen entre ellos y así lograr los objetivos señalados, siendo el sistema capaz de generar conocimiento y experiencia[15]. La plataforma presentada en detalle en la sección 5 de este artículo, se basa en un modelo de arquitectura multi-agente distribuida orientada a servicios [2].

3. Análisis de imágenes: Fases y Técnicas

Dentro de las técnicas de recuperación tridimensional de la escena se puede realizar una primera clasificación en técnicas pasivas y activas [18]. Las primeras se basan en información captada sin modificar las condiciones de captación de la escena. Por contra, las técnicas activas modifican las condiciones del entorno con el fin de extraer información de su estructura tridimensional. Cada una de ellas presenta características específicas en el proceso de interpretación visual, desde la etapa de formación de la imagen, hasta el análisis e interpretación de la misma, pero todas presentan unas fases que podríamos considerar comunes en el análisis. La idea general es que después de procesar la imagen (reducir ruido, mejorar contraste, realizar análisis de correspondencia, disparidad y otras técnicas que veremos en este apartado), podamos definir la escena como una representación icónica del mundo visible; para que a partir de los datos de esta representación se pueda calcular las acciones correspondientes. Este *modelo*

icónico no necesitará representar todos los detalles de la escena, sino solo los necesarios para desarrollar la tarea que queramos.

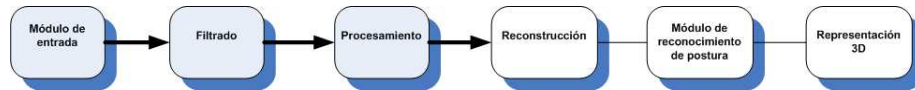


Figura 3. Módulos del sistema de análisis de imágenes.

En figura 3 podemos ver la secuencia de acciones a seguir (clasificadas en diferentes módulos) para llevar a cabo el filtrado, procesado y reconstrucción de las imágenes obtenidas de la escena. Estos módulos están incluidos dentro del *sistema de análisis de imágenes* que podíamos ver en la figura 1.

El primer problema que nos encontramos, antes de llegar a analizar imágenes estereoscópicas, es la *entrada de datos* de la captación de las escenas. El número de cámaras que se van a utilizar (y número de imágenes de que podremos disponer para realizar la reconstrucción), disposición de las cámaras, calibración, y su relación entre sí, y con el sistema de coordenadas del espacio tridimensional y de las imágenes, posibles defectos de las cámaras, etc., son cuestiones que se abordan en este módulo y que nos permitirán saber cómo es el proceso de obtención de las imágenes.

En el *módulo de filtrado* se llevan a cabo acciones más directas (o a un nivel más bajo) sobre las imágenes obtenidas. Estas acciones serán, entre otras, la reducción de ruido, la mejora del contraste, el realce de bordes o la corrección de distorsiones. algunas de estas acciones podrán llevarse a cabo a nivel hardware, es decir, por medio de los recursos proporcionados por las cámaras [16].

Según el dominio donde actúen, los filtros podemos dividirlos en filtros de dominio espacial (convolución) y filtros en el dominio de la frecuencia (multiplicación y transformadas de Fourier). Los primeros operan directamente sobre los píxeles de la imagen y en ellos se utilizan generalmente matrices denominadas máscaras que operan sobre una vecindad de píxeles, centrándose en el pixel de interés. Cada pixel de la nueva imagen se obtiene mediante el sumatorio de la multiplicación de la máscara por la vecindad del pixel.

Hoy en día existe una gran cantidad de algoritmos para la restauración y mejoramiento de imágenes: filtros de suavizamiento (mejoran la detección de bordes y son útiles cuando la imagen presenta por ejemplo ruido causado por efecto de iluminación deficiente: filtros pasa-baja, filtros gaussianos) [9], filtros de realce (destacan detalles: filtros pasa-alta, filtros High-Boost.) y filtros de énfasis de altas frecuencias (acentúan los detalles pero preservan las zonas uniformes). Cada clase de filtro difiere en los valores utilizados en la máscara.

El *módulo de procesamiento* puede considerarse el corazón del sistema ya que será donde se pongan en práctica los algoritmos para el análisis de disparidad y correspondencia de los pares estereoscópicos y donde se obtengan las medidas de distancia a cámara. Medidas que nos ayudarán en la siguiente fase a reconstruir

la escena. Esta reconstrucción consiste en, a partir del modelo icónico obtenido después de las fases anteriores, representar los objetos del entorno. En nuestro caso, tenemos un entorno interior donde lo que nos interesa representar (a partir de los datos obtenidos del análisis) son los pacientes y residentes del centro. Para ello, tendremos un *módulo de reconocimiento de postura* y otro de *representación 3D* que modelarán la escena con los datos que le lleguen.

Según Marr y Poggio [12], existen tres etapas en el proceso de recuperación de la estructura de una escena. Estas son, primero, seleccionar un punto característico de un objeto en una de las imágenes (*preprocesamiento*); segundo, encontrar el mismo punto característico en la otra imagen complementaria (*análisis de correspondencia*); y tercero, medir la diferencia relativa entre la posición de estos dos puntos (*análisis de disparidad y obtención de distancia*). La diferencia de posición de esos puntos en cada una de las imágenes es como ya se ha nombrado con anterioridad, la disparidad.

El *preprocesamiento* tiene como objetivo identificar las características representativas de cada imagen [14]. Estas características tendrán que ser elegidas cuidadosamente porque con ellas se llevarán a cabo los análisis. En visión artificial y procesamiento de imágenes el concepto de característica es usado para definir una pieza de información que es relevante para llevar a cabo una tarea computacional. Una *característica* puede ser un borde, una esquina o un blob. Muchos algoritmos de visión utilizan la detección de características como paso inicial, y como consecuencia, existe una gran cantidad de detectores ya desarrollados (Canny [5], Sobel [24], Prewitt [18], Marr-Hildreth [13], Harris and Stephens/Plessey [10], SUSaN [23], Shi and Tomasi [22], FaST [25], Laplaciano del Gaussiano, Determinante del Hessiano[11], MSER, Grey-level blobs[11]).

En visión artificial es la detección de bordes la técnica más comunmente utilizada y más robusta. Los bordes son causados por una variedad de factores: discontinuidades en la normal a la superficie, en la profundidad, en el color y/o en la iluminación. Lo normal es que sean encontrados en zonas de la imagen donde el nivel de intensidad fluctúa bruscamente. así, un punto de borde puede ser visto como un punto en una imagen donde se produce una discontinuidad en el gradiente. El algoritmo de Canny [5] es usado para detectar todos los bordes existentes en una imagen. Este algoritmo está considerado como uno de los mejores métodos de detección de contornos.

El *problema de la correspondencia* descrito en el apartado 2.1 consiste básicamente, en dadas dos imágenes de partida, encontrar las parejas de puntos de ambas imágenes que se corresponden con un mismo punto de la escena o de la imagen en 3D. La información de profundidad (distancia) se obtiene a partir del hecho de que el mismo punto característico en la escena aparece en una posición ligeramente distinta en los dos planos de las imágenes. El desplazamiento entre las dos imágenes, es lo que conocemos como *disparidad*.

Las estrategias de obtención de correspondencia entre dos puntos pueden clasificarse de varias formas [8]:

- Según las primitivas utilizadas podemos clasificarlas en técnicas basadas en áreas (p.e. Suma de las Diferencias absolutas), técnicas basadas en carac-

terísticas (p.e. Marr-Poggio[12], Pollard-Mayhew-Frisby[17], Segmentos de borde o Características múltiples), y en técnicas basadas en programación dinámica (p.e. Cox, Otha y Kanade o Bircheld y Tomasi [3]).

- Según la geometría empleada están las que utilizan ejes paralelos y las que no.
- Y según el paradigma en el que se incluyen están las estrategias binoculares, trinoculares y multioculares.

Las *técnicas basadas en área* consideran las dos imágenes captadas como una señal bidimensional trasladada. Tratan de obtener, para cada punto de la imagen, dicha traslación minimizando un cierto criterio (correlación). Para cada píxel de una imagen se calcula la correlación entre la distribución de intensidades de una ventana centrada en dicho píxel y una ventana del mismo tamaño centrada en el píxel a analizar de la otra imagen.

Una de las técnicas más sencillas a utilizar e implantar es la Suma de Diferencias absolutas (SDA) ya que únicamente se realizan operaciones con números enteros. Dado un píxel de coordenadas (x, y) en la imagen izquierda, se calcula un índice de correlación $C(x, y, s)$ a cada desplazamiento s de la ventana de correlación en la imagen derecha. Para calcular el índice de correlación se tiene:

$$C(x, y, s) = \sum_{u=-w, v=-w}^{u=w, v=w} |I_l(x+u, y+v) - I_r(x+u+s, y+v)|$$

donde $2w + 1$ es el tamaño de la ventana centrada en el píxel considerado, ubicado en la posición (x, y) e I_l, I_r son los valores de gris de los píxeles en las imágenes izquierda y derecha respectivamente. El cómputo del índice de correlación puede realizarse de forma eficiente evitando las redundancias entre ventanas sucesivas [9]. La disparidad $dl(x, y)$ entre el píxel de la imagen izquierda y aquel de la imagen derecha se define como el desplazamiento s que minimiza el índice de correlación: $dl(x, y) = \operatorname{argmin}_s C(x, y, s)$.

La librería utilizada en el trabajo (Triclops SDK[26]) establece la correspondencia entre las imágenes usando esta técnica. Entre las ventajas de estas técnicas están el obtener buenos resultados sobre imágenes con textura importante, permitir crear mapas densos de disparidad y su facilidad de paralelizar. Como inconvenientes decir que presentan problemas con imágenes con elevadas discontinuidades de superficie, que son muy sensibles a variaciones fotométricas debidas a sombras, que requieren un proceso posterior de eliminación de falsas correspondencias y que tiene problemas con las oclusiones.

Las *técnicas basadas en características* obtienen primitivas de alto nivel (puntos de borde, segmentos, curvas, regiones, etc.) que atesoran un conjunto de propiedades invariantes a la proyección en mayor o menor medida. Tienen las ventajas de que las primitivas de alto nivel proporcionan información más rica que los niveles de intensidad, permiten utilizar restricciones geométricas entre las primitivas, son robustas aunque proporcionan información dispersa.

Entre los métodos basados en características nos encontramos unas de las teorías más famosas e importantes que hemos estudiado dentro de las opciones

barajadas, la teoría computacional de Marr-Poggio[12]. También otras teorías y técnicas entran dentro de los métodos basados en características como la teoría computacional de Pollard-Mayhew-Frisby [17] o las técnicas basadas en segmentos borde.

La teoría computacional de Mayhew-Frisby (PMF) está considerada dentro de la categoría de los algoritmos de correspondencia soportados por vecindad cuya característica es la de establecer la relación de rasgos entre imágenes a partir del análisis de sus posiciones con respecto a los puntos de los rasgos que los rodean [17]. En la figura 4 se esquematiza la relación de vecindad (V_i, V_d) para los puntos P y P' que representan la proyección de un punto en dos perspectivas (imagen izquierda y derecha).

Este algoritmo asume dos restricciones fundamentales, la primera está dada por la arquitectura desde la que fueron adquiridas las imágenes, en la que como recordaremos respetaba una disposición paralela entre los dispositivos receptores imágenes, particularidad que permite asumir la existencia de una relación directa entre la información visual contenida a la misma altura de las imágenes (*raster*); en otras palabras, los rasgos contenidos en la línea n de la imagen izquierda deberán aparecer igualmente (conservando cierta disparidad) en la línea n de la imagen derecha y por lo tanto el proceso de correspondencia entre rasgos se realizará solo entre aquellos que se ubiquen en el mismo raster de ambas imágenes. La segunda restricción está dada por el concepto de gradiente de disparidad (GD). La idea de aplicar el GD a un algoritmo de correspondencia parte de los trabajos sobre percepción visual con los que se estableció que el límite con el que el sistema visual humano logra fusionar puntos pertenecientes a un par estereoscópico de imágenes está dado por un $GD \leq 1$ el cual es calculado como se observa en la figura 4.

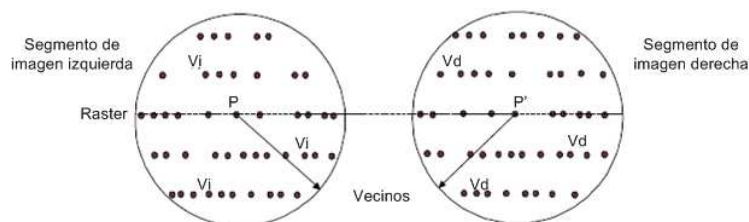


Figura 4. Relación de vecindad (V_i, V_d) para los puntos P y P' y cálculo de GD

Las técnicas de basadas en programación dinámica plantean la búsqueda de correspondencias como un problema de optimización, descomponiendo el problema en otros más sencillos; por ello son considerados programación dinámica. Una de las ventajas de utilizar estas técnicas es la rapidez, debido a la descomposición en subtareas que realiza. Otra de las ventajas es que suelen ser métodos sencillos de implementar. Entre los inconvenientes, que no calculan muy bien la disparidad entre puntos de los objetos pequeños.

Rubio de Lemus [20] aplica una metodología de comparación y evaluación de algoritmos de búsqueda de correspondencias en estereopsis, mostrando los aspectos relevantes de los mismos.

El *cálculo de disparidad* nos permitirá obtener la profundidad de cada uno de los puntos de la escena, obteniendo una única imagen que se denomina *mapa de disparidad*. Puesto que existe una relación directa entre la profundidad relativa de los objetos en una imagen y su disparidad en un par estéreo, podemos coger como valores relativos de la profundidad de los objetos la información extraída del mapa de disparidad, es decir, tomaremos dicha imagen como una aproximación válida del *mapa de profundidad*. Nuestra reconstrucción, puesto que partimos de imágenes que son 2D, tomará como valor aproximado de la profundidad relativa de un punto con respecto al observador, el correspondiente a dicho punto en el mapa de profundidad; para así poder obtener una tercera coordenada para poder representarlo en un espacio 3D.

4. Propuesta

El proceso de visión estereoscópica es implementado sobre una arquitectura distribuida construida con agentes la cual permite paralelizar tareas utilizando cada servicio (distribuido) como una unidad de procesamiento independiente. Cabe mencionar que inicialmente no todos los procesos realizados por el sistema son implementados para operar en forma paralela, sino únicamente el algoritmo de correspondencia por resultar computacionalmente muy costoso. El sistema está compuesto por un conjunto de agentes, con roles bien definidos que comparten información y servicios. El análisis de imágenes implica un proceso complejo donde cada agente ejecuta su tarea con la información disponible en cada momento. En la figura 5a se muestra un esquema de la arquitectura multi-agente propuesta.

Las aplicaciones consisten en todos aquellos programas que pueden ser utilizados para acceder a las funcionalidades del sistema. Las aplicaciones son dinámicas y adaptables al contexto, reaccionando de forma distinta ante determinadas situaciones (por ejemplo, si se encuentra en una habitación o en otra, o si el usuario tiene alguna discapacidad) y el tipo de servicio solicitado. Los servicios representan el grueso de las funcionalidades del sistema a nivel de procesamiento, entrega y adquisición de información. Se pretende que el sistema cuente con un directorio flexible de servicios, locales o remotos, de manera que puedan ser modificados, añadidos o eliminados dinámicamente y bajo demanda.

Como se aprecia en la figura 5a, una de las características es que, al no estar restringidas las aplicaciones por un determinado lenguaje de programación, se amplía el abanico de posibilidades para desarrollar interfaces sobre un mayor número de dispositivos y sistemas. Esto es destacable, ya que por ejemplo, esta característica será vital a la hora de integrar el módulo encargado del análisis de imágenes con el de representación 3D, o con el de localización.

El funcionamiento del sistema multi-agente conlleva la coordinación de agentes que cooperan y comparten entre sí la información necesaria durante el pro-

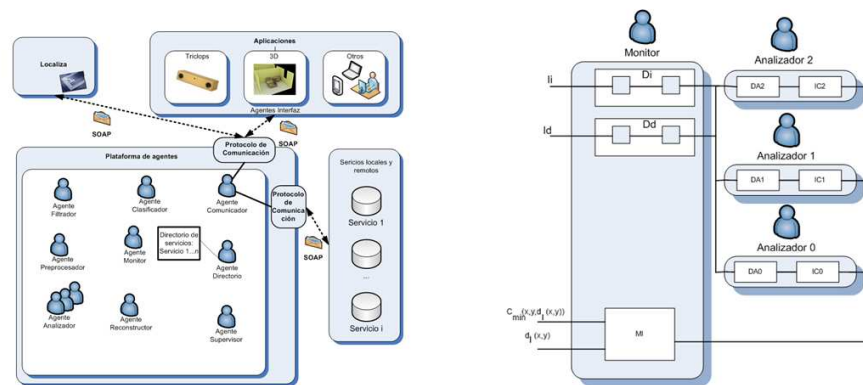


Figura 5. (a) Esquema de la arquitectura. (b) Estructura del algoritmo SDA

ceso de análisis. al descomponer el proceso de análisis descrito en los anteriores apartados, las tareas generadas se distribuyen entre cada uno de los roles del sistema. Se han obtenido los siguientes roles: *Clasificador*: encargado de recibir las imágenes, comprobar su formato y cambiarlo en el caso de que sea necesario. *Filtrador*: encargado de las tareas de filtrado básico. Coordina los filtrados de las primeras fases ya sea desde las aplicaciones incrustadas en las cámaras o desde el propio sistema de análisis. *Preprocesador*: encargado de la extracción de características. Podrá hacer uso del servicio de extracción que crea conveniente (Canny, etc.). *Monitor*: se encarga del control de los procesos. asigna tareas a los demás agentes. Divide en bloques de las imágenes y asigna las tareas de análisis a los agentes analizadores. Los agentes analizadores le devolverán sus resultados y éste enviará la información concentrada al agente Reconstructor que se encargará de hacer que se represente tridimensionalmente. *Interfaz*: este tipo de agente es diseñado para ser embebido en las aplicaciones de los usuarios para la comunicación directa con los agentes de la plataforma. *analizador*: lleva a cabo el análisis de correspondencia. *Reconstructor*: recoge la información analizada para enviarla a la aplicación de representación 3D. *Comunicador*: responsable de las comunicaciones entre las aplicaciones y la plataforma y los servicios y la plataforma. Maneja las peticiones de entrada desde las aplicaciones para ser procesadas por los servicios. *Supervisor*: analiza la estructura y sintaxis de todos los mensajes de entrada y salida del sistema y supervisa el correcto funcionamiento de los otros agentes en el sistema (verifica periódicamente el estado de todos los agentes registrados en la arquitectura enviando mensajes ping). *Directorio*: llevará un listado activo de todos los servicios que ofrece la arquitectura (locales o remotos).

En este caso se ha trabajado un prototipo implementando el componente analizador siguiendo dos de las técnicas estudiadas para el análisis de correspondencia: el algoritmo Mayhew-Frisby (PMF), basado en características y el algoritmo de emparejamiento de la Suma de Diferencias absolutas (SDA) ba-

sado en áreas. En ambos casos se realiza una optimización aprovechando las ventajas que nos proporciona la plataforma de agentes, haciéndose una distribución de tareas y paralelización, para conseguir una disminución en el tiempo de procesado.

PMF: La base que fundamentó la *paralelización del algoritmo de correspondencia PMF* fue la restricción de *correspondencia entre líneas*, es decir, un punto ubicado en la línea n de la imagen izquierda solo puede ser correspondido con algún punto de la imagen derecha que se encuentre en la misma línea. Partiendo de esta premisa se direccionó la ejecución del algoritmo de correspondencia a unidades de procesamiento independientes, pasando como parámetro solo los dos bloques de datos contenidos en el mismo *número de línea* de cada imagen.

El agente Monitor toma el control de los procesos desarrollados por el sistema, asignando las tareas de aplicación de los filtros, detección de características, etc. Después de que las imágenes han sido reducidas a sus rasgos principales estas son segmentadas en 20 bloques de 24 líneas cada uno (por estar trabajando con imágenes de 640×480). La longitud de cada bloque está en función del tamaño en filas de la imagen a procesar. Es importante mencionar que la formación de los bloques se fuerza para que las 24 líneas que integran cada bloque no se encuentren dispuestas secuencialmente en la imagen, es decir, sean seleccionadas de tal forma que existan 20 líneas de separación entre ellas. Un ejemplo del contenido de un bloque sería: $a = \{0, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460\}$ y así sucesivamente para los bloques restantes, esta convención fue adoptada con el fin de uniformizar la distribución de cargas de trabajo entre los agentes ya que generalmente existen zonas en que la imagen concentra mayor información.

Una vez que los bloques han sido formados, éstos son direccionados por el agente Monitor para ser procesados paralelamente en agentes independientes (analizadores). En el momento en que el último agente finaliza su proceso, la información es concentrada por el Monitor para realizar la reconstrucción 3D con el agente correspondiente.

SDA: El cálculo estéreo se realiza con la librería Triclops [26], que por defecto utiliza un algoritmo de emparejamiento basado en la Suma de Diferencias absolutas (SDA). La utilización de esta medida de comparación resulta más rápida que otros algoritmos basados en correlación, aunque es más sensible a los cambios de iluminación. al ser el método utilizado en las librerías proporcionadas por Point Grey [16] se realiza su implementación y también una propuesta de optimización del mismo mediante la paralelización de las tareas llevadas a cabo por el algoritmo.

El algoritmo a implantar debe descomponerse en un conjunto de subfunciones, las cuales actúan secuencialmente en el flujo de datos. Estas subfunciones son asignadas a los agentes analizadores a través del agente Monitor. El agente Monitor será el encargado de controlar y asignar tareas a los demás agentes. Identifica las subfunciones a realizar en el análisis de correspondencia y asigna las tareas de cada uno de los agentes analizadores. La estructura del algoritmo implantado se muestra en la figura 5(b). Para tener acceso a los píxeles vecinos

de la ventana de correlación una vez que han sido escaneados, se utilizan unas estructuras tipo FIFO (*First In, First Out*) en el agente Monitor para las imágenes izquierda y derecha (I_i , I_d), representadas como D_i y D_d . La diferencia absoluta entre los píxeles de la imagen derecha y la imagen izquierda es calculada en los módulos D_{ai} , donde $i = 1, 2, \dots, d_{max}$, y d_{max} es el máximo desplazamiento de la ventana de correlación. El índice de correlación se calcula en los módulos IC_i . Tanto el cálculo de la diferencia absoluta como el índice de correlación se realiza en los agentes analizadores de manera paralela. La disparidad $dl(x, y)$ se determina haciendo una comparación entre todos los índices de correlación que corresponden a los d_{max} desplazamientos. El módulo MI determina la disparidad $dl(x, y)$ así como el mínimo índice de correlación $C_{min}(x, y, dl(x, y))$, el cual puede utilizarse para calcular un índice de confianza del emparejamiento.

5. Resultados y Conclusiones

La arquitectura de agentes propuesta nos permite automatizar el análisis descrito y optimizar su procesamiento. La utilización de agentes nos ofrece grandes niveles de flexibilidad [2] ya que con ellos podemos mover el código a lugares donde las acciones son requeridas consiguiendo respuestas en tiempo de ejecución, autonomía y continuidad de los servicios. además, se reduce el esfuerzo de programar tareas múltiples, ya que sólo es necesario especificar objetivos globales para que los agentes cooperen entre ellos y así lograr los objetivos señalados. Los algoritmos de visión estereoscópica son implementados sobre esta arquitectura, la cual permite paralelizar tareas utilizando cada servicio como una unidad de procesamiento independiente.

Inicialmente, el algoritmo elegido para su optimización fue el algoritmo de correspondencia por ser el más costoso computacionalmente. así, se definen dos propuestas de paralelización de técnicas de correspondencia: el algoritmo Mayhew-Frisby (PMF), basado en características y el algoritmo de la Suma de Diferencias absolutas (SDA) basado en áreas. Para ello se lleva a cabo la implementación de un prototipo *analizador* que sea capaz de llevar a cabo este tipo de análisis.

La primera fila de la imagen 6 muestra el procesamiento de una imagen de muestra. Es una imagen PPM (*Portable Pixmap Format*) formada por tres imágenes estéreo.

Podemos coger como valores relativos de la profundidad de los objetos la información extraída del mapa de disparidad. Los valores más oscuros representan objetos de la escena más alejados y los más claros más cercanos. Con los primeros resultados pudimos observar que la detección de bordes es un paso que, aunque no es necesario, es muy útil ya que nos ayuda en entornos en los que las condiciones de iluminación cambian de manera significativa. Por este motivo, se decidió ampliar el prototipo añadiendo más funcionalidades. En este caso se desarrollaron los algoritmos de detección de bordes de Sobel [24] y Canny[5]. Como se observa, el mapa de disparidad obtenido utilizando bordes, refleja mucho mejor la profundidad de los objetos de la imagen original.

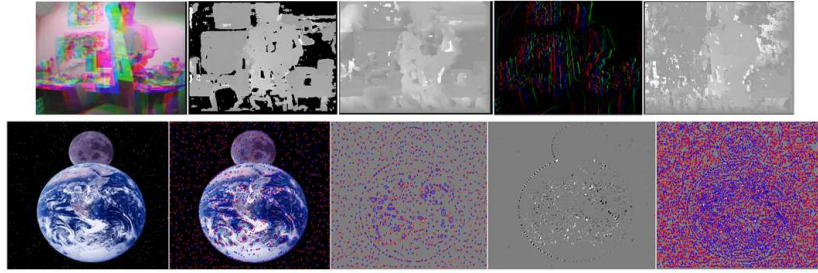


Figura 6. Resultados procesado de imagen. De izquierda a derecha y de arriba a abajo: imagen original, mapa de disparidad utilizando detección de bordes, mapa de disparidad obtenido sin utilizar detección de bordes, utilización del método Sobel sobre la imagen original, mapa de profundidad obtenido utilizando Sobel en lugar de Laplaciana (utilizada por defecto en Triclops); imagen original estéreo derecha, extracción de características por número de líneas n sobre la imagen original, extracción de características por número de líneas n , detección de contorno y extracción de características por número de líneas $m > n$

El siguiente paso fue el desarrollo de la siguiente técnica de análisis de correspondencia basada en características (PMF). La segunda fila de la imagen 6 muestra la secuencia de resultados obtenidos.

En función de todo lo expuesto, y a partir de los resultados obtenidos, el algoritmo SDA es el óptimo. Para realizar una comparación objetiva, los algoritmos se han aplicado sin mejoras. Para poder perfeccionar los resultados, se han desarrollado filtros y máscaras adaptables a las características de la imagen (por ejemplo, un conversor de imágenes estéreo JPG (*Stereo.jpg*) a PPM y a escala de grises) cuyos resultados no se presentan en este trabajo. además, se ha propuesto la mejora de ambos algoritmos mediante su paralelización utilizando la arquitectura de agentes. Como líneas de trabajo futuro están: (i) continuar con el desarrollo de las funcionalidades del análisis, comparando los resultados, eligiendo las mejores técnicas y utilizando el hardware Point Grey [16]; (ii) desarrollar la propuesta de paralelización de los algoritmos de correspondencia para la optimización del sistema; (iii) integrar las funcionalidades dentro de la arquitectura distribuida propuesta (junto con posibles ampliaciones que puedan darse de desarrollos externos); (iv) integrar los módulos del sistema global (módulo de análisis, módulo de localización por radiofrecuencia, módulo de reconstrucción 3D).

Referencias

- [1] Angulo, C. and Tellez, R. (2004). Distributed Intelligence for smart home appliances. Tendencias de la minería de datos en España. Red Española de Minería de Datos.
- [2] Ardissono, L., Petrone, G. and Segnan, M. (2004). a conversational approach to the interaction with Web Services. Computational Intelligence, 20, 693-709.
- [3] Bircheld S. and Tomasi C. (1998) Depth Discontinuities by Pixel-to-Pixel. Stereo. International Conference on Computer Vision, pp. 1073-1080

- [4] Brenner, W., Wittig, H. and Zarnekow, R. (1998). Intelligent software agents: Foundations and applications. Secaucus, NJ, USA: Springer-Verlag New York, Inc
- [5] Canny, J., (1986). a computational approach to edge detection. IEEE Trans Pattern analysis and Machine Intelligence, 8 (6), pp. 679-698. 1986.
- [6] Carretero, N. and Bermejo, a. B.: Inteligencia ambiental. CEDITEC: Centro de Difusión de Tecnologías, Universidad Politécnica de Madrid, España
- [7] Corchado, J. M., Bajo, J., De Paz, Y. and Tapia, D. I. (2006). Intelligent Environment for Monitoring alzheimer Patients, agent Technology for Health Care. Decision Support Systems. Eslevier Science. amsterdam, Netherlands.
- [8] Dhond U.R, aggarwal J.K., (1989) Structure From Stereo - a Review, IEEE Trans. on Systems, Man. and Cybernetics, Vol. 19, No.6, Nov/Dec
- [9] Digital Image Processing. Image analysis and Understanding. accesible en: <http://www.icaen.uiowa.edu/dip/LECTURE/contents.html>
- [10] Harris C. and Stephens M., (1988). a combined corner and edge detector. Cuarta conferencia alvey Vision. Páginas 147-151.
- [11] Lindeberg T., (1998). Feature detection with automatic scale selection. International Journal of Computer Vision 30. Páginas 77-116.
- [12] Marr D. y Poggio T. (1979). a computational theory of human stereo vision. Proc R Soc Lond, pp. 301-328
- [13] Marr D., E.C. Hildreth (1980) Theory of edge detection, Proc. Roy. Soc. London B-207 187-217.
- [14] Pearson, Don. (1991) Image Processing. Great Britain, McGrawHill.
- [15] Pecora, F. and Cesta, a. (2007). Dcop for smart homes: a case study. Computational Intelligence , 23 (4), 395-419
- [16] Point Grey Research Inc. <http://www.ptgrey.com/> (2008)
- [17] Pollard, S.B., Mayhew, J.E.W. and Frisby, J.P. (1985). a stereo correspondence algorithm using a disparity gradient constraint, Perception, vol. 14, 445-470.
- [18] Prewitt J.M.S. (1970) Object enhancement and extraction, in: B.S. Lipkin, a. Rosenfeld (Eds.), Picture analysis and Psychopictorics, academic Press, New York
- [19] Rao, a. S. and Georgeff, M. P. (1995). BDI agents: from theory to practice. In Proceedings of ICMaS'95. San Francisco, Ca, USA
- [20] Rubio De Lemus, P. (1993) aplicacion De Una Metodologia De Evaluacion De Sistemas De Emparejamiento En Vision Tridimensional. Psicothema ISSN edición en papel: 0214-9915 1993 . Vol. 5 , n 1 , pp. 135-159
- [21] Sáez, J.M., Peñalver, a. and Escolano, F., (2003) Compact mapping in plane-parallel environments using stereo Vision, CIARP 2003, 659-666.
- [22] Shi J. and Tomasi C. (1994). Good Features to Track. 9 IEEE Computer Vision and Pattern Recognition.
- [23] Smith S. M. and Brady J. M. (1997). SUSaN - a new approach to low level image processing. International Journal of Computer Vision 23. Páginas 45-78
- [24] Sobel, I. and Feldman,G. (1968). a 3x3 Isotropic Gradient Operator for Image Processing. Conferencia Stanford artificial Project
- [25] Trajkovic M. and Hedley M., (1998). Fast corner detection. Image and Vision Computing. Páginas 75-87
- [26] Triclops StereoVision System Manual Version 3.1. User guide and command reference. (2003) Point Grey Research Inc.
- [27] Vázquez, I. and López de Ipiña, D. (2005). Inteligencia ambiental: la presencia invisible. Revista solo programadores No. 127, Revistas Profesionales 16-19.
- [28] ZigBee. (2006). ZigBee Specification Document 053474r13. ZigBee alliance.

Arquitectura MVOSMA: Mundos Virtuales a través de Sistemas Multiagente basados en Organizaciones

Jesús Ángel Román Gallego, Juan Manuel Corchado Rodríguez

Universidad de Salamanca, Departamento de Informática y Automática
Salamanca - España

Resumen. La realidad virtual supone para las personas la opción de introducirse en un mundo aparentemente real e interactuar con el entorno mediante dispositivos electrónicos, existiendo muchas aplicaciones relativas a la realidad virtual como por ejemplo simuladores de vuelo, simuladores anatómicos para operaciones, etc. Los sistemas multiagentes han sido empleados en esta realidad virtual para la representación de objetos dentro de estos mundos virtuales. El presente artículo describe una arquitectura para proveer estos mundos virtuales, *MVOSMA*, la cual está basada en un diseño organizacional derivado de la teoría de las organizaciones capaz de suministrar mundos virtuales personalizados basándose en una serie de características del usuario que solicita este mundo, introduciendo mecanismos de comunicación entre agentes y servicios Web.

1. Introducción

Se puede definir la realidad virtual como la representación en tiempo real de un entorno simulado e interactivo generado a través de medios informáticos [14]. Esta realidad virtual puede ser inmersiva cuando el usuario se encuentra inmerso en un entorno tridimensional y no inmersiva cuando el usuario interactúa con un entorno representado en su monitor mediante el teclado, el ratón u otros dispositivos. La extensión de esta realidad virtual a la Web se conoce como Web 3D. Se dice que el futuro de Internet es crear espacios en 3D accesibles de forma estándar [11], los cuales son modelados a través de la utilización de la especificación [17] la cual describe *Virtual Reality Modeling Language (VRML)*.

La adaptación de los sistemas distribuidos a través de inteligencia artificial la encontramos a través de la tecnología de agentes, pasándose a denominar estos sistemas distribuidos, Sistemas Multiagente (*MAS*) [18]. La construcción de MAS integra tecnologías multidisciplinares como son técnicas de *ingeniería del software*, *inteligencia artificial*, *programación concurrente y distribuida* [5].

El diseño de MAS, generalmente, en lugar de realizarse como un sistema con componentes que ejecutan métodos, se ha de pensar en los objetivos que los componentes deben alcanzar y en las tareas necesarias para que lo consigan, de este modo se intuye que el proceso es más intuitivo ya que esta forma de modelar y de razonar se halla más cerca del pensamiento humano que los paradigmas de programación tradicionales [12].

La Real academia de la Lengua Española define el término *organización* como la “*asociación de personas, regulada por un conjunto de normas en función de determinados fines. Surgen siempre de una idea previa inspirada por un interés o una necesidad general*”. Esto implica que podemos considerar la organización como un grupo de entidades que realizan una serie de actividades específicas y diferenciadas, el cual, conforma una estructura y cumple una serie de normas para obtener un fin común y determinado [15]. La aplicación de una estructura organizacional en un sistema multiagente facilita su desarrollo y prepara al mismo para su crecimiento y consistencia en cuanto a tolerancia a los fallos que se puedan producir.

El trabajo que aquí se presenta consiste en la definición de una arquitectura capaz de implementar de forma relacionada los tres conceptos vistos anteriormente: *Mundos Virtuales, Sistemas Multiagentes y Organizaciones*. La posibilidad de dotar de inteligencia a un sistema distribuido para proveer mundos virtuales a un usuario de forma personalizada a través de una serie de características del propio usuario permitirá acercarnos más a la Web 3D. además si este sistema presenta una composición basada en una de las organizaciones dentro de la propia “Teoría de Organizaciones”, permitirá que las comunicaciones entre los distintos agentes que componen el sistema multiagente, así como su mantenimiento se realice más fácilmente. Todo esto hace que la arquitectura *MVOSMA (Mundos Virtuales a Través de Sistemas Multiagente basados en Organizaciones)* se presente como una arquitectura consistente y novedosa para la visualización de mundos virtuales personalizados.

2. Conceptos Previos

En este punto se introducirán una serie de conceptos los cuales son de vital importancia en la creación de la arquitectura *MVOSMA* ya que el estudio y comprensión de su estado del arte ha permitido la interrelación de los mismos para dar lugar a dicha arquitectura.

2.1. Mundos Virtuales y Web 3D

Los orígenes de la Realidad Virtual se remontan a mediados de los años 70 en el Departamento de Defensa de los Estados Unidos, donde comienza el desarrollo de este tipo de proyectos con la intención de permitir el adiestramiento de los

pilotos de aviones de combate sin la necesidad del riesgo que conlleva la práctica real y el coste que supone [14]. En [14] se muestran una serie de principios que rigen la realidad virtual como son: **Simulación; Interacción; Percepción.**

Una clasificación de esta realidad virtual se realiza acorde a la base sobre la que se desarrolla cada uno de los soportes físicos que permiten la inmersión en el mundo virtual, así se puede aludir a la diferenciación entre *Realidad Virtual Inmersiva*, y *Realidad Virtual no Inmersiva* según [14][19].

El estándar para modelos 3D en la red es actualmente *VRML (Virtual Reality Modeling Language)* [17]. Este lenguaje tiene como misión la misma que HTML para las páginas Web, pero en este caso para escenarios tridimensionales. En sus comienzos los mundos virtuales VRML eran ambientes monousuarios, es decir, para ser vistos por todos los usuarios conectados a la red, pero navegando en forma solitaria. La “*Computer Support Cooperative Work (CSCW)*” [9] y el proyecto de mundos virtuales 3D del “*Institute for applied Information Technology*” [10] ha enfocado el desarrollo de VRML hacia mecanismos y tecnologías colaborativas para entornos virtuales para Internet. Los principales objetivos son la representación e interacción de múltiples usuarios distribuidos en el mundo, mediante la infraestructura de la red. Con este proyecto ha sido realizado un prototipo de sistema distribuido de Realidad Virtual, mediante el protocolo *DWTP (Distributed Worlds Transfer and Communication Protocol)* [1].

2.2. Agentes y Sistemas Multiagente

Una definición de agente se da en [13], donde se concibe a éste como un sistema que reúne las características de: *autonomía, Reactividad, Pro-actividad y Habilidad social*. además, existen ciertos atributos que los agentes pueden poseer, estos son: *Razonamiento y aprendizaje*, mediante los cuales los agentes son capaces de comportarse de manera inteligente. *Movilidad*, que permite a los agentes desplazarse entre los nodos de una red y ejecutarse en distintas plataformas.

Un modelo de clasificación de agentes característico es el modelo de clasificación llamado de las vocales, propuesto por *Yves Demazeau* [4], donde la *a (de agente)* caracteriza sus rasgos individuales: arquitectura, funcionamiento interno, complejidad, etc. La *E (de entorno)* caracteriza los requisitos computacionales para que el agente funcione correctamente; la *I (de interacción)* tiene en cuenta las capacidades de comunicación del agente; la *O (de organización)* considera el papel del agente en el sistema; y, la *U (de utilidad)* referencia a la aplicación de la que forma parte el agente.

Según las características individuales, los agentes pueden clasificarse en *agentes reactivos y agentes cognitivos* [3]. Los primeros realizan tareas sencillas y su modelo computacional está basado en un ciclo recepción de eventos. Por otra parte, los agentes cognitivos realizan tareas complejas. De hecho, utilizan algún

tipo de representación explícita del conocimiento. Para realizar las tareas necesitan llevar a cabo procesos de razonamiento y otros procesos cognitivos, como por ejemplo la planificación y el aprendizaje. En cuanto al modo de organización [3], podemos considerar agentes individuales, que no tienen capacidad de cooperación, y agentes cooperativos, que pueden realizar tareas solos o colaborando con otros agentes.

Respecto a los tipos de arquitecturas para la construcción de Sistemas Multiagente, las deliberativas utilizan modelos de representación simbólica del conocimiento del conocimiento, y suelen estar basadas en la teoría de planificación [18][2][13][3]. Estos agentes parten de un estado inicial y tienen la capacidad de concebir y generar planes para alcanzar sus objetivos. El uso de una representación simbólica trae consigo una serie de problemas, y es por lo que se han estudiado modelos más efectivos de representación del conocimiento. Es por ello que, las arquitecturas reactivas se caracterizan por no tener como elemento central de razonamiento un modelo simbólico y por no utilizar razonamiento simbólico complejo. En último lugar, contar con la existencia de arquitecturas híbridas que tratan de solventar las limitaciones que ambas arquitecturas presentan por separado.

2.3. Las Organizaciones

Las Organizaciones Humanas. La definición de organización que aparece en el diccionario de la Real academia de la Lengua Española, corresponde con la de *“asociación de personas regulada por un conjunto de normas en función de determinados fines”*. Por lo tanto, una organización consiste en una serie de individuos que realizan unas actividades o funciones específicas y diferenciadas. además, estos individuos se estructuran siguiendo unas pautas y reglas determinadas que les permitan alcanzar los fines u objetivos de la organización. Los fines deben ser objetos de conocimiento común, que cumplan principalmente las funciones de guiar los esfuerzos de los miembros en vistas a su consecución [15]. asimismo, deben proporcionar una fuente de legitimidad que determine las conductas y acciones adecuadas en el contexto organizacional y establezcan los niveles mínimos que han de conseguirse.

Las Organizaciones de agentes. Dentro de los sistemas multiagente, la organización se ha empleado fundamentalmente para describir a un conjunto de agentes que, mediante una serie de roles y reglas de interacción establecidas, se coordinan tratando de alcanzar los objetivos globales del sistema. Una definición de organización en el contexto de los sistemas multiagente es dada por *L. Gasser: “las organizaciones son sistemas estructurados con patrones de actividad, conocimiento, cultura, historia y habilidades distintas de cualquier agente particular”*. Las organizaciones existen a un nivel completamente independiente

de los agentes individuales que los constituyan, los cuales pueden ser reemplazables. además, ocupan alguna región del espacio, bien sea geográfico, temporal, simbólico, etc. [8].

3. Arquitectura MVOSMA

Teniendo en cuenta los conceptos anteriores, aquí se muestra el diseño de una arquitectura que engloba los conceptos vistos. Para ello la arquitectura se centra en dos grupos como contenedores de agentes: *la Organización y los Gremios*. La organización es la encargada de dar respuesta a las peticiones de los clientes y es la que dirige y determina el funcionamiento del sistema. Esta organización está dividida en departamentos, los cuales tienen las siguientes tareas.

- *Departamento de Dirección:* Funcionamiento, coordinación de los demás departamentos y comunicaciones externas.
- *Departamento de Producción:* análisis del perfil del solicitante, generación del mundo virtual y "subcontratación" de los agentes de los gremios.
- *Departamento de Mantenimiento:* Comprobación del estado de los agentes de los diferentes departamentos.

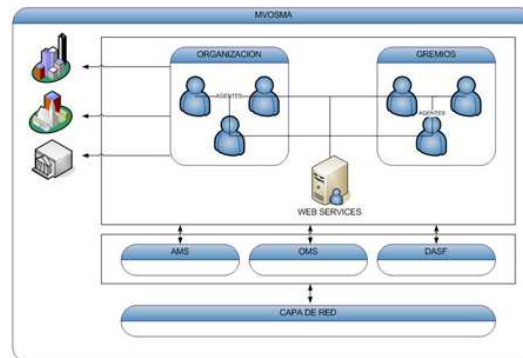


Figura 1. arquitectura MVOSMA

En los *Gremios* se encuentran agentes con alto grado de especialización formando grupos con un fin común, de este modo si desde la organización se requieren varios agentes que provean servicios para lograr una finalidad determinada, se facilita la labor de búsqueda y contratación ya que solamente han de encontrar el gremio que persiga ese fin en vez de buscar a los agentes uno por uno. La "subcontratación" de los agentes pertenecientes a los gremios, llevada a cabo por el departamento de producción, se realiza de acuerdo a las necesidades en un momento determinado. Esta subcontratación se realiza mediante una serie de *Servicios Web* los cuales están a disposición de los agentes encargados de la

misma, de este modo los agentes de la organización se descargan de servicios y se centran en las tareas productivas.

Tipo	Meta-Servicio	Descripción
Laboral	Contractagent	Contrata a un agente existente en un gremio
	Dimissagent	Despide a un agente que finaliza su tarea
	asignTask	asigna una tarea a un agente contratado
	InformTask	Informa acerca de una tarea determinada

Cuadro 1. Meta-Servicios para la Contratación de agentes

Se han propuesto dos grupos como contenedores de agentes, ahora bien, las otras partes de la arquitectura que proveen los servicios que son utilizados por los agentes del sistema para las operaciones internas, tanto de los agentes, como de la organización en si son: el *aMS* (*agent Management System*), *OMS* (*Organización Management System*) y el *DaSF* (*Directory and Service Facilitator*). El *DaSF* ofrece un directorio de agentes y de los servicios que éstos proveen a modo de páginas amarillas. Este *DaSF* actúa de forma que cualquier agente que ofrezca un determinado servicio ha de registrarse, de este modo cuando fueren requeridos por algún otro agente, ya sea para su contratación o simplemente para conocimiento, pueda proveer esta información en el menor tiempo posible y con la máxima exactitud.

Tipo	Meta-Servicio	Descripción
aBMReg	RegisterProfile	Crea la descripción de un nuevo servicio (Profile)
	RegisterProcess	Crea la implementación del servicio (Process)
	RegisterGremio	Crea un Gremio
	ModifyProfile	Modifica un Profile Existente
	MofifyProcess	Modifica un Process Existente
	ModifyGremio	Modifica un Gremio
	DesregProfile	Elimina un Profile. También su Process asociado
	DesregGremio	Elimina un Gremio existente y todos sus servicios
asociacion	addProvider	Enlaza un proveedor con un servicio ya existente
	DelProvider	Elimina un proveedor de un servicio
Descubrimiento	SearchService	Busca un servicio determinado
	GetProfile	Extrae la descripción de un servicio especificado
	GetProcess	Extrae la implementación de un servicio especificado
	GetGremio	Extrae el gremio de un servicio especificado

Cuadro 2: Meta-Servicios del DaSF

Otro elemento es la *OMS* la cual presta servicios que conciernen al mantenimiento y explotación de la organización. así pues, los servicios prestados están orientados al mantenimiento de la misma a través de normas que rigen el funcionamiento de la organización. También se incluyen las políticas de contratación de agentes las cuales determinan los agentes necesarios en cada momento.

Tipo	Meta-Servicio	Descripción
Structural	RegisterRole	Crea un nuevo rol dentro de un departamento
	RegisterNorm	Crea una nueva norma dentro de un departamento
	RegisterDepartment	Crea un departamento dentro de la organización
	RegisterDeptNorm	Crea una norma de departamento en la organización
	DesregRole	Elimina un rol de un departamento
	DesregNorm	Elimina una norma de un departamento
	DesregDepartment	Elimina un departamento dentro de una organización
	DesregDeptNorm	Elimina una norma de departamento
	InformagentRole	Informa de que un agente adopta un rol determinado
	InformMembers	Indica que existen agentes en un departamento
	CountMembers	Devuelve el número de agentes de un departamento
	InformDepartment	Devuelve la descripción de un departamento
	InformDepartRoles	Muestra los roles existentes en un departamento
InformRoleProfiles	Indica los perfiles asociados a un determinado rol	
InformRoleNorms	Provee las normas asociadas a un determinado rol	
Dynamical	RegisteragentRole	Crea una nueva relación $\langle \text{entity}, \text{department}, \text{rol} \rangle$
	DesregagentRole	Elimina una reacción $\langle \text{entity}, \text{department}, \text{rol} \rangle$ dada
	acquireRole	Solicita la adopción de un rol en un departamento
	LeaveRole	Solicita la desvinculación de un rol departamental
	Expulse	Obliga a un agente a desvincularse de un rol
	Duplicateagent	Duplica un agente dentro de un departamento

Cuadro 3: Meta-Servicios del OMS

El último componente corresponde al *aMS*, que se encarga de la creación y destrucción de agentes definiendo el ciclo de vida de los mismos. además realiza diferentes funciones básicas para el mantenimiento de la plataforma. Por último la arquitectura cuenta con la capa de comunicaciones, la cual se encarga del envío y recepción de mensajes entre distintos agentes y con el exterior de la plataforma.

Tipo	Meta-Servicio	Descripción
Registration	Register	Registra un nuevo agente en la plataforma
	Deregister	Elimina el registro de un agente ya existente
	UpdateRegister	Modifica el registro de un agente excepto su nombre
Discovery	Searchagent	Solicita información a cerca de un agente registrado
	GetDescription	Obtiene una descripción de la plataforma

Management	Suspendagent activateagent	Suspende la ejecución de un agente en activo activa la ejecución de un agente suspendido
Communication	Send	Envía un mensajes dentro y fuera de la plataforma

Cuadro 4: Meta-Servicios del AMS

Para el desarrollo de esta arquitectura se ha decidido seguir los estándares que propone la *FIPA (The Foundation for Intelligent Physical agents)* [16]. Estas especificaciones *FIPA* representan un conjunto de estándares que persiguen el objetivo de promover la interacción entre agentes heterogéneos y los servicios que representan los mismos. *FIPA* mantiene un ciclo de vida para sus distintas especificaciones, compuesto por las etapas *Preliminary, Experimental, Standard, Deprecated, y Obsolete*. a cada especificación se le asigna un identificador único, en función del estado en el que se encuentran dentro del ciclo de vida. La elección de *FIPA* como estándar para el desarrollo de la arquitectura *MVOSMA* se basa en que es la especificación más seria en lo que a desarrollo de sistemas multiagente se refiere, además de ofrecer una alta escalabilidad. Uno de los principales documentos es el *FIPA abstract architecture Specification* [6]. Dicho documento, y las especificaciones derivadas del mismo, definen la arquitectura abstracta propuesta por *FIPA*, siendo el propósito principal de este documento el garantizar la interoperabilidad y la reusabilidad. Otro de los documentos que constituye el pilar principal de *FIPA* es el documento *FIPA agent Management Specification* [7]. Dicho documento contiene especificaciones para la gestión de agentes, incluyendo servicios de gestión de agentes, ontologías, y transporte de mensajes dentro de la plataforma de agentes.

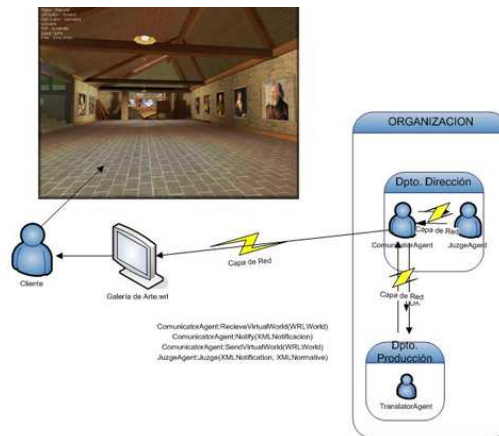


Figura 2. Detalle de Mundo Virtual Suministrado

En la *Figura 2* se muestra el proceso final en la generación de un mundo virtual personalizado, en este caso, una galería de arte. Previamente se ha realizado la petición del cliente del mundo virtual, a partir de este momento se ponen en funcionamiento los mecanismos necesarios de la arquitectura para ajustar el perfil del solicitante, y contratar los agentes necesarios para la correcta visualización del mundo virtual solicitado.

4. Conclusiones y Trabajo Futuro

Se observa en la bibliografía que los agentes se muestran como objetos dentro de los mundos virtuales, para aquellos objetos que tengan cierta capacidad de percibir su entorno y actuar de forma autónoma. En la arquitectura descrita se pretende que sea el propio sistema multiagente el proveedor de mundos virtuales solicitados por el cliente de modo que a través de técnicas de inteligencia artificial el sistema sea capaz de personalizar el mundo solicitado.

La organización que se adopta para el sistema multiagente es sencilla y basada en una jerarquía. Este tipo de organización acelera las búsquedas de servicios que se demanden por los agentes gracias a la creación de los *gremios*, o grupo de agentes que están especializados en la realización de tareas determinadas, es decir, cada gremio de agentes provee una serie de servicios los cuales tienen un mismo fin, de este modo si desde la organización se requieren varios agentes que provean servicios para lograr una finalidad determinada, se facilita la labor de búsqueda y contratación ya que solamente han de encontrar el gremio que persiga ese fin en vez de buscar a los agentes uno por uno.

Una de las pretensiones de este trabajo es la relación del sistema multiagente propuesto con una serie de servicios Web. Estos servicios Web se utilizan para la contratación de los agentes pertenecientes a los gremios. El acceso a estos servicios se realiza a través de mensajes XML estándar, lo cual se ha adoptado como protocolo de comunicación entre las distintas entidades de nuestro sistema haciendo así totalmente estándar el lenguaje de comunicación dentro y fuera del sistema.

La organización del sistema multiagente se compone del departamento de dirección, departamento de producción y departamento de control. Dentro de cada uno de ellos existen varios agentes que se coordinan y cumplen una serie de normas para lograr un fin común. Esta organización viene dada, pero la propia definición de la arquitectura a través de las ontologías y parametrización de las mismas hace que se pueda adoptar cualquier organización y por lo tanto abordar cualquier problema, convirtiéndola en una arquitectura abierta, y cambiante en tiempo real.

Como continuidad del presente trabajo y líneas de investigación futuras, se pretende lograr la definición de un formato determinado para el lenguaje de

comunicaciones preservando la forma de XML estándar, que sea compatible no solamente con agentes sino con cualquier servicio Web con el que se desee interactuar. Otra línea a seguir es la implementación de esta plataforma en un lenguaje concreto, lo cual permitirá probar la misma y medir su efectividad de forma cuantificable. Esta implantación es laboriosa, lo que implica un trabajo a medio, largo plazo.

Referencias Bibliográficas

1. Wolfgang Broll. Dwtip-an internet protocol for shared virtual environments. *Virtual Reality Modeling Language Symposium*, pages 49–56, 1998.
2. R. Corchado, J.M. & Laza. Constructing deliberative agents with case-based reasoning technology. *International Journal of Intelligent Systems*, 18:227–241, 2003.
3. Grupo de Investigación HIDRa Departamento de Ingeniería de Sistemas y Computación Universidad de los andes. agentes y sistemas multiagente: Integración de conceptos básicos.
4. Yves Demazeau. Multi-agent system methodology. In *CNRS Leibniz IMaG*, 2003.
5. V. J. Botti F. Garijo et al M. Belmonte. *agentes Software y Sistemas Multiagente. Conceptos, arquitecturas, y aplicaciones*. ana Mas, 2005.
6. Foundation for Intelligent Physical agents. Fipa00001 fipa abstract architecture specification, 2000.
7. Foundation for Intelligent Physical agents. Fipa00023 fipa agent management specification, 2004.
8. L. Gasser. Perspectives on organizations in multiagente systems. *Lecture Notes in Computer Science*, pages 1–16, 2001.
9. <http://wikipedia.org>. Enciclopedia de internet.
10. http://www.fit.fraunhofer.de/index_e.n.html. Fraunhofer institute for applied information technology.
11. <http://www.web3d.org/>. The ultimate 3d viewer solutions.
12. Plans Intentions and Practical Reason. *Bratman, M. E.* 1987.
13. M. Wooldridge N. R. Jennings. *agent-oriented software engineering*. 2000.
14. García alvarado R. Santelices Malfanti I Parra Márquez, J.C. *Introducción práctica a la realidad virtual*. 2001.
15. J. Peiro. *Psicología de la Organización*. Universidad Nacional de Educación a Distancia, 1991.
16. Paul Grima Raimund Dachselt, Ralf Drner. Web3d 2003-workshop, 2003.
17. Chris Marrin Rikk Carey, Gavin Bell. Iso/iec 14772-1:1997 virtual reality modeling language (vrml97). *The VRML Consortium Incorporated*, 1997.
18. Jorge J. Gómez Sanz. Metodologías para el desarrollo de sistemas multi-agente. *Revista Iberoamericana de Inteligencia artificial*, 18:51–63, 2003.
19. K.W.; Gibson I. Zheng, J.M.; Chan. Virtual reality. potentials. *IEEE Issue-2*, 17:20–23, 1998.

Sistema de Inteligencia ambiental mediante visión estereoscópica y representación 3D en tiempo real para la localización de personas en entornos de dependencia

Manuel P. Rubio, Juan M. Corchado

Departamento de Informática y automática
Plaza de la Merced s/n, 37008, Salamanca, España
(mprc, corchado)@usal.es

Resumen. Este trabajo presenta un estudio de técnicas para el tratamiento de imágenes procedentes de cámaras estereoscópicas; y de los datos obtenidos desarrollar un sistema de presentación de dicha información en un entorno 3D interactivo en tiempo real. El sistema estará incluido dentro de un proyecto global que tiene como objeto el desarrollo de un entorno inteligente para la localización e identificación en entornos de dependencia donde confluirán otra serie de tecnologías. Dentro de este proyecto, se encuentra ya en fase de finalización un sistema de localización por radiofrecuencia y son los módulos de videovigilancia, análisis y representación 3D los que centran actualmente la atención. En concreto, el presente trabajo se centra en el desarrollo de una arquitectura que permita, una vez analizada la información recogida por las cámaras estereoscópicas, representarla en formato tridimensional así como interpretarla de manera adecuada para el reconocimiento de objetos y posturas de una persona en la escena.

1. Introducción

En Europa, durante las últimas tres décadas, el número de personas mayores de 60 años aumentó cerca de un 50 %, por lo que en la actualidad, el 25 % de la población de Europa es mayor de 60 años y se estima que dentro de 20 años este porcentaje aumente a un 30 %, dando como resultado un total de 100 millones de ciudadanos[2].

La importancia de desarrollar nuevos y más rentables métodos para suministrar cuidados médicos y asistencia a los ancianos se acentúa al considerar tales tendencias[2], siendo necesaria para una mejor administración de los servicios, la creación de entornos seguros, discretos y adaptables para la monitorización y fomento de la salud. Las arquitecturas basadas en dispositivos inteligentes están siendo investigadas como sistemas de supervisión de cuidado médico para ancianos y personas dependientes, por lo que podrían proporcionar apoyo constante en su vida diaria[3].

El término “ambient Intelligence” (amI) define un modelo de interacción[11] en el que las personas están rodeadas de un entorno inteligente, consciente de su

presencia, sensible al contexto y capaz de adaptarse a sus necesidades, a través de tecnología incrustada, no invasiva y transparente para los usuarios con el objetivo de facilitar sus actividades diarias[1]. La inteligencia ambiental hace uso de tecnologías que permiten estas posibilidades: redes de comunicaciones, dispositivos de presentación de información, redes de sensores, sistemas de visión artificial, sistemas adaptativos y sensibles al contexto (inteligentes), etc.

El problema de la visión artificial, en general y de la visión estereoscópica en particular, ha recibido una considerable atención en los últimos años, que ha llevado a la aparición de nuevos métodos teóricos y desarrollos prácticos para el diseño de sistemas de procesado de imagen. Sus aplicaciones son muy variadas, e incluyen aspectos como el reconocimiento de caracteres, la detección de blancos, análisis de imágenes biomédicas, sistemas de seguridad y vigilancia, inspección automática o medición remota.

Se presenta en este trabajo un sistema basado en amI diseñado con el objetivo de mejorar la eficiencia de los cuidados médicos, la atención y la seguridad de pacientes dependientes. El sistema interactúa con los usuarios y el ambiente a través de técnicas de visión artificial estereoscópica y proporciona una monitorización en tiempo real de los pacientes mediante un entorno tridimensional virtual.

La descripción de este proyecto global junto con un resumen de la tecnología estereoscópica empleada serán abordados en las secciones 2 y 3 de este artículo. En la sección 4 se hará hincapié en las técnicas de reconstrucción y representación 3D. En las secciones 5 y 6 se mostrará la propuesta del sistema y las conclusiones y líneas futuras.

2. Revisión del problema

El presente trabajo se engloba dentro de un proyecto de investigación en el que se está trabajando dentro del Grupo de Biomedicina, Sistemas Informáticos Inteligentes y Tecnología Educativa (BISITE) de la Universidad de Salamanca. El objetivo del proyecto global es desarrollar un sistema de inteligencia ambiental para el cuidado y supervisión de pacientes en entornos de dependencia[4], proporcionando un entorno capaz de realizar tareas de localización, identificación y monitorización de pacientes de forma automática.

Se supervisará automáticamente la ubicación de los pacientes a través de tecnología ZigBee[12] (sistema de localización por radiofrecuencia) y cámaras de vídeo estereoscópicas. asumiendo que se desconoce la situación de un paciente en un entorno en cada momento, sus movimientos, postura o estado, se plantea la posibilidad de unión de ambas tecnologías; complementando así la información de localización obtenida a través de los dispositivos ZigBee con la información de las cámaras. Una vez analizada la información recogida por las cámaras estereoscópicas, se interpreta de manera adecuada para el reconocimiento de objetos y posturas de los personajes de la escena y se representa en un modelo tridimensional que facilitará un seguimiento personalizado de los pacientes por parte del

personal médico y los propios familiares, permitiendo de este modo controlar de manera más exhaustiva el estado de los enfermos. La innovación de este concepto reside en la representación en tercera dimensión y en el seguimiento del paciente para representar la escena correspondiente según su posición en cada instante.

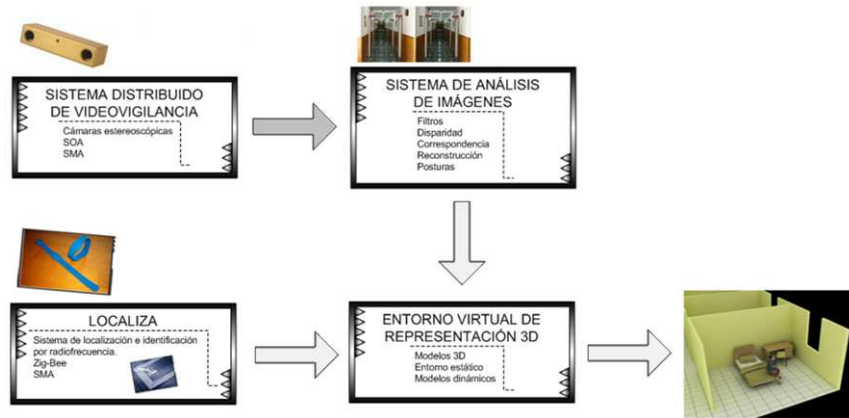


Figura 1. Entorno Inteligente para la Localización en entornos de dependencia

3. Estereoscopia

La visión artificial ha tratado de reproducir las funciones del sistema visual mediante el análisis y procesamiento de imágenes obtenidas desde cámaras de video. Muchas han sido las técnicas propuestas por visión artificial para tratar de alcanzar tan aparentemente simple objetivo y aún cuando cada una propone enfoques diferentes todas pueden considerarse complementarias. El presente trabajo está basado en la técnica de recuperación de forma y posición a partir de la estereoscopia, la cual tiene como modelo funcional la estructura binocular sobre la que opera el sistema visual humano.

3.1. Principios básicos

La visión del hombre es un complejo y sofisticado sistema que abarca procesos fisiológicos, bioquímicos, neurológicos y psicológicos. De manera natural el mecanismo de visión es estereoscópico, es decir, es capaz de apreciar, a través de la visión binocular, las diferentes distancias y volúmenes del entorno, permitiendo la observación tridimensional del mismo. Los ojos, debido a su separación,

obtienen dos imágenes con pequeñas diferencias entre ellas, denominadas disparidades. El cerebro las procesa y las interpreta de forma que se percibe la sensación de profundidad, lejanía o cercanía, formando la imagen tridimensional de los objetos del entorno. Este proceso se denomina estereopsis y se regula en función de 10 señales, cuatro de las señales están dentro del campo fisiológico y seis en el psicológico[6].

La fotografía estereoscópica, también denominada abreviadamente fotografía 3D, intenta reproducir la sensación de profundidad de la visión estereoscópica natural. Si obtenemos dos fotografías con una separación adecuada, correspondientes a la visión que se obtendría con cada ojo, y se observan con un visor apropiado, es posible recrear la sensación de profundidad, de la cual carecen las fotografías convencionales. Si se quiere obtener fotografías 3D en movimiento (vídeo estereoscópico) se puede emplear una cámara estéreo especial, una cámara convencional dotada de un accesorio especial con espejos o bien dos cámaras disparadas sincronizadamente. a este último tipo pertenece el sistema estereoscópico elegido en este trabajo.

3.2. Condicionantes técnicos del sistema estereoscópico elegido

La cámara elegida en el sistema de visión es la Bumblebee2, modelo BB2-COL-ICX424 (640x480 Color 3.8mm) del fabricante canadiense Point Grey[5].

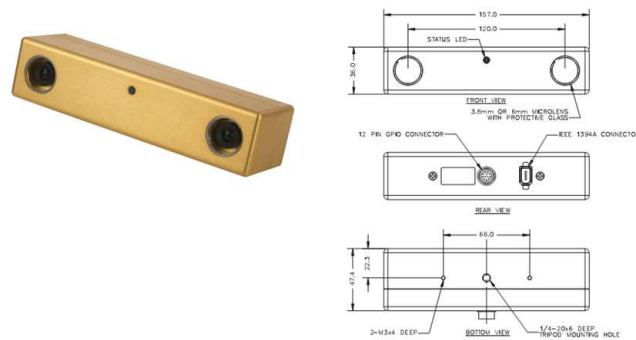


Figura 2. Cámara Bumblebee2 y esquema

Las principales características de esta cámara son:

- Dos CCDs Sony ICX204 de 1/3", Color progressive scan.
- 640x480 píxeles cuadrados hasta 48 fotogramas por segundo.
- Conversor analógico/digital de 12 bit.
- Línea base (distancia entre cámaras) 120mm.
- Distancia focal de la óptica: 3,8mm con 70 de campo de visión horizontal.
- Tamaño: 157mm x 36mm x 47.4mm

- Peso: 342gramos.
- Interface: 6-pin IEEE-1394 (FireWire) para el control de la cámara y transmisión de los datos de vídeo.
- alimentación suministrada a través del puerto IEEE-1394.
- Consumo menor de 3W.

Esta cámara es ideal para aplicaciones tales como el seguimiento de personas, reconocimiento de gestos y posturas, robótica móvil y otras aplicaciones de visión por computador. Está precalibrada para corregir las distorsiones de las lentes y los desajustes temporales y espaciales, por lo que no es necesario hacerlo manualmente. La información de calibración está precargada en la cámara, permitiendo que el software recupere la corrección de la imagen.

Para el software de control y reconocimiento se empleará un sistema dividido en tres categorías relacionadas entre ellas.

Digiclops proporciona el control de la cámara y la transmisión vía firewire al ordenador receptor de las imágenes. También permite configurar la cámara ajustando el balance de blancos, la velocidad de obturación, exposición, ganancia, etc.

Triclops SDK es un conjunto de funciones C++ que permite obtener la información 3D de las imágenes. Proporciona una rápida y precisa generación del mapa de profundidad de la escena. Esto se puede lograr aplicando múltiples algoritmos especificando todas las características del procesado estéreo. Tanto Digiclops como Triclops están incluidos con la cámara estereoscópica Bumblebee2.

Por último, la tercera parte del software aplicación se basa en las otras dos y es generada por el usuario. Con el software de usuario se recogen los resultados y se presentan en el formato adecuado. Se escribe en visual C++ utilizando las funciones del SDK Triclops y Digiclops. Con ello se puede lograr código muy optimizado y con la rapidez de ejecución necesaria para la obtención de datos en tiempo real.

3.3. Justificación de la tecnología adoptada

Dentro de un sistema de vídeo vigilancia, como el que se quiere diseñar, se pueden presentar distintas tecnologías en la parte de hardware y transmisión de datos y distintas técnicas y algoritmos en el software de reconocimiento de imagen.

El parámetro más importante a tener en cuenta al diseñar un sistema de visión artificial es el número de cámaras a emplear. Habitualmente se utilizan configuraciones de una cámara (monocular), dos cámaras (formando un par estereoscópico) y múltiples cámaras (normalmente entre tres y cinco, dependiendo del entorno de trabajo). a mayor número de cámaras los resultados son más precisos pero es mayor la potencia de cálculo necesaria así como el ancho de banda en la transmisión de las imágenes. En nuestro caso se planteó la utilización de una cámara, dos cámaras independientes desde puntos de vista distintos y dos cámaras formando un par estereoscópico.

Empleando una cámara solo se obtiene información de zonas de la secuencia de imágenes y reconocimiento de formas y patrones, estáticos o en movimiento; pero no se logran datos de posición o distancias del entorno tridimensional y de los elementos del mismo. Salvo que la cámara cambie su posición y se desplace durante el proceso de grabación de imágenes. Como en nuestro caso la cámara es estática, una no es suficiente para los resultados buscados.

El uso de dos cámaras desde puntos de vista distintos aporta gran exactitud en los resultados pero requiere procedimientos de configuración y calibración muy complejos y críticos en la precisión obtenida.

Un par estereoscópico desde una posición fija reduce los problemas de calibración y da muy buenos resultados en la determinación de posiciones y distancias espaciales, también permite el reconocimiento de formas y patrones. además existen algoritmos ya muy desarrollados.

En cuanto a la tecnología de las cámaras a emplear son dos principalmente, analógicas y digitales y ambas con dos sistemas de transmisión de las imágenes, cable y radio. Las cámaras analógicas son las utilizadas habitualmente en seguridad, están más desarrolladas y son más baratas pero necesitan una etapa intermedia de digitalización de las imágenes para poder ser tratadas informáticamente y extraer su información. Las cámaras digitales utilizan sensores CMOS o CCD que generan las imágenes directamente en formato digital, normalmente en formato de vídeo mpeg4 (lo que evita el uso de la etapa de digitalización).

Como punto final del hardware se tiene que evaluar y decidir si el ordenador en el que se ejecute el software de tratamiento de las imágenes se encuentra en el propio entorno y se transmiten a la central de seguimiento sólo los resultados (deberá ser bastante potente) o se enviará el “videostreaming” a la central y se procesará allí (se requerirá un gran ancho de banda).

En el caso concreto de este proyecto se emplearán cámaras estereoscópicas digitales para el reconocimiento de situación y postura de los pacientes. Irán conectadas a los ordenadores de procesamiento vía cable con interface IEEE 1394 de alta velocidad y alimentación incorporada. Cada sistema estereoscópico incluye dos minicámaras que generan dos secuencias de vídeo sincronizadas y permiten obtener información tridimensional del entorno mediante el estudio de las disparidades entre ellas. La información obtenida se complementa con la del sistema de radiofrecuencia para evitar errores de posición y mejorar los resultados[12].

4. Reconstrucción estereoscópica: fases y técnicas

En la figura 3 se presenta el proceso completo diseñado para el tratamiento de las secuencias de imágenes estereoscópicas, obtención de los datos de posición, orientación y postura del paciente, y su representación tridimensional en el entorno virtual de presentación. Es necesario un módulo de entrada que adecúe las imágenes al formato deseado para realizar el análisis. Un módulo de filtrado que mejore su calidad, aplicando acciones como la reducción de ruido, la mejora del contraste, el realce de bordes o la corrección de distorsiones[10]. En el módulo

de procesamiento se ponen en práctica las técnicas de análisis estereoscópico. Será aquí donde se implementen los algoritmos para el análisis de disparidad y correspondencia de los pares estereoscópicos y donde se obtengan las medidas de distancia a cámara[7]. Esta reconstrucción consiste en representar los objetos del entorno. En nuestro caso, tenemos un entorno interior donde lo que nos interesa representar son los pacientes y residentes del centro. Para ello, tendremos un módulo de reconocimiento de postura y otro de representación 3D que modelarán la escena con los datos recogidos.

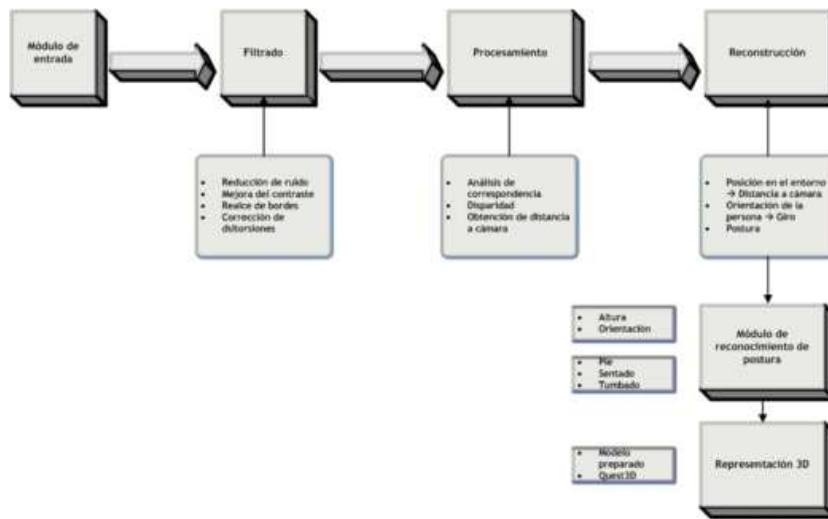


Figura 3. Módulos del sistema

En el resto de este artículo se presentan técnicas del módulo de reconocimiento y representación 3D. Los otros módulos se muestran en el otro trabajo paralelo a este.

4.1. Reconocimiento de postura

Los sistemas de reconocimiento de movimiento humano suelen estar compuestos por cuatro fases bien diferenciadas, inicialización, seguimiento, estimación de la pose y reconocimiento[8]. antes de que un sistema esté listo para procesar datos, necesita ser inicializado, es decir, debe de crearse un modelo adecuado del sujeto. a continuación se realiza un seguimiento del movimiento del individuo. Para ello se realiza un proceso de segmentación de la imagen diferenciando al sujeto del fondo. La pose del sujeto puede ser estimada ya que puede ser la salida del sistema, p.e. para controlar

un avatar (representación gráfica de un humano) en un entorno virtual, o puede ser procesada posteriormente en la fase de reconocimiento. La última fase realiza un análisis de la pose u otros parámetros para poder reconocer las acciones realizadas por el sujeto. Evidentemente, no todos los sistemas abarcan o contienen todos estos subsistemas. En aplicaciones de control, como es nuestro caso, puede ser suficiente con estimar la pose, prescindiendo del último subsistema de reconocimiento.

La **inicialización** cubre los pasos necesarios para asegurarse de que el sistema comienza con una correcta interpretación de la escena. Esto incluye calibración de las cámaras, adaptación a las características de la escena e inicialización del modelo.

En el **seguimiento** se establecen relaciones coherentes del sujeto y/o extremidades entre fotogramas. El seguimiento puede verse como una fase en la que se prepara la información para poder estimar la pose, o una fase donde se preparan los datos para el reconocimiento. En el primer caso, su propósito es extraer información específica de las imágenes, ya sea a bajo nivel, como bordes o bien de alto nivel, como las manos, y la cabeza. En el segundo su tarea normalmente se limita a representar los datos de forma adecuada.

Como regla general, en el seguimiento se pueden identificar tres fases: segmentación y separación del sujeto y el fondo. a continuación, estas imágenes segmentadas son transformadas a otra representación que reduzca la cantidad de información a procesar, o que se adecue más a un algoritmo concreto. Por último, definen como se debería hacer un seguimiento del sujeto entre fotogramas. En lo que respecta a la representación de la información, podemos encontrar sistemas que hacen uso de representaciones basadas en objetos (puntos, cajas, siluetas, blobs), o aquéllos que hacen uso de la propia imagen (Fourier, segmentos y puntos, características de la escena, etc.). En el skeleton se dibuja la figura

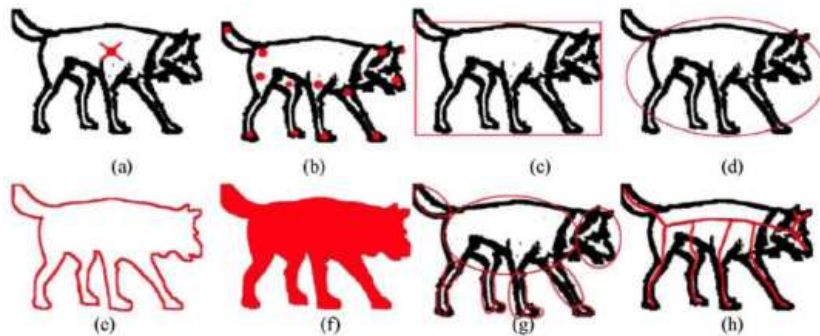


Figura 4. Representación de los objetos. (a) centroide, (b) puntos múltiples, formas básicas (c) rectángulo, (d) elipse, (e) contorno, (f) silueta, (g) formas articuladas, (h) skeleton

como líneas que representan el eje de cada sección de las figuras. Se logra al lle-

var la representación con contornos a su eje mayor o principal. Se utilizan sobre todo en personajes articulados, y pueden ser en 2D o en 3D. Es el equivalente al esqueleto real en el cuerpo humano y ayuda mucho a definir los movimientos y poses del entorno virtual de representación 3D.

La **estimación de la pose** es el proceso de identificar la configuración de un cuerpo humano y/o de los miembros individuales en una determinada escena. Puede ser la salida del sistema, como en ciertas aplicaciones de control en entornos virtuales (el caso de este proyecto), o simplemente servir de entrada al módulo de reconocimiento que seguirá procesando esta información. Un aspecto común en la fase de estimación de la pose es el uso de un modelo que permita incorporar conocimiento sobre el movimiento humano en el proceso de análisis. Normalmente se utiliza un modelo geométrico de un cuerpo humano, que puede ser acompañado de información sobre el movimiento. El modelo humano es representado por un número de articulaciones y “huesos” (segmentos) que los conectan. Cuanto más complejo es el modelo, mejores son los resultados obtenidos, pero se requiere un mayor tiempo de proceso.

La fase de **reconocimiento** puede verse como un tipo de postproceso. El objetivo último de muchos de los sistemas de captura de movimiento suele ser el reconocimiento del movimiento. Este reconocimiento normalmente se lleva a cabo clasificando el movimiento capturado como una acción de entre un conjunto de acciones posibles. Tradicionalmente, se han usado dos paradigmas distintos: reconocimiento mediante reconstrucción, o reconocimiento directo. El primero se basa en el concepto de reconstruir la escena y luego analizarla, el segundo realiza el análisis directamente sobre los datos de bajo nivel.

4.2. Representación 3D

Un aspecto muy importante en este proyecto es la presentación de datos a los usuarios del sistema de inteligencia ambiental que deban tomar decisiones o ser informados. Dichos datos tienen una componente principal relacionada con la posición y orientación de la persona en entornos tridimensionales y se obtienen mediante el sistema de radiofrecuencia y el de vídeo vigilancia estereoscópica inteligente. Pero para obtener una comprensión rápida y precisa es necesario visualizarlos en un entorno tridimensional virtual en tiempo real.

Para la representación del paciente en el entorno 3D se utilizará un modelo articulado (skeleton) construido a partir de componentes que están conectados entre sí mediante articulaciones y conocen la relación que mantienen con los que están conectados. Se trata de componentes rígidos que únicamente admiten movimientos de rotación entre ellos. La unión de las distintas articulaciones da como resultado la estructura interna del esqueleto. Se trata de una estructura jerárquica que puede ser descrita mediante un árbol que también puede interpretarse como un grafo dirigido sin ciclos (figura 5).

El modelo geométrico se construye situando el esqueleto en el interior de una malla poligonal cerrada que se deforma adaptándose a los movimientos del mismo y que incluye el aspecto visual del paciente visualizado.

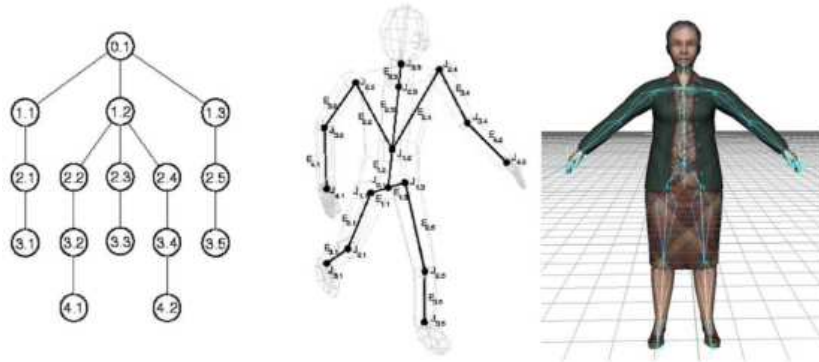


Figura 5. árbol jerárquico de las articulaciones, estructura interna y modelo geométrico elegido

Para la programación del entorno tridimensional virtual en tiempo real se utilizará una aplicación específica de propósito general llamada Quest3D. Es la conjunción de un Motor gráfico con una plataforma de desarrollo y se utiliza en arquitectura, videojuegos, software de entrenamiento y simuladores. Los datos y animaciones son importados de paquetes CaD tales como Maya o 3D Studio Max. Es un producto desarrollado por act-3D B.V. en Holanda y su primera versión fue publicada en septiembre del 2001[9].

Una de las características más importantes de Quest3D es la metodología de programación, que es totalmente diferente a la de los habituales lenguajes de programación. Su entorno de desarrollo es casi por completo visual. Las aplicaciones se desarrollan conectando componentes funcionales, denominadas “Channels” que vinculados componen una estructura de árbol, que representa la del programa que se implementa. El árbol se ejecuta por completo una vez (al menos) por frame, invocando a cada “channel”. Lo que da como resultado una aplicación 3D en tiempo real, permite que el programador pueda modificarla mientras ésta se ejecuta y no exista compilación de código como en los lenguajes habituales.

5. Propuesta

El sistema de visión y reconstrucción 3D que se desarrollará en el proyecto necesita los datos descriptivos de la posición tridimensional del paciente dentro de la habitación, su orientación espacial respecto a la cámara y su postura o pose. La solución tecnológica la da el sistema de estereovisión. además de proporcionar información de la distribución de objetos y personas en la escena, permite calcular también su distancia respecto al par de observación y por lo tanto respecto a cualquier sistema de referencia elegido. así mismo con los datos conseguidos se modelará y posicionará un esqueleto virtual que representará al paciente dentro de un entorno tridimensional interactivo y en tiempo real de la habitación.

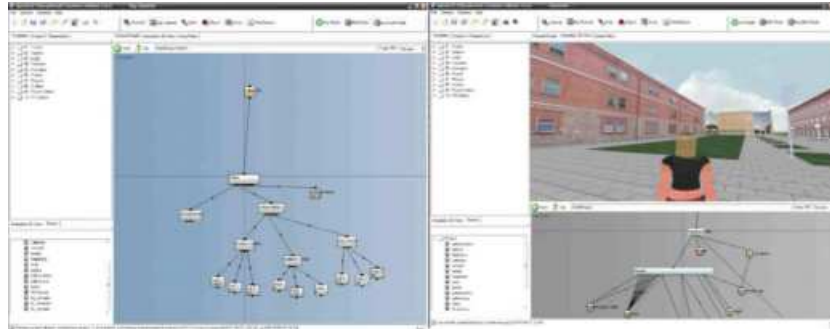


Figura 6. Pantallas de la aplicación Quest3D

Para el desarrollo del proyecto se elegirá un entorno lo más parecido posible a una habitación habitual de una residencia o clínica geriátrica y se situará en ella la cámara con su correspondiente ordenador conectados vía cable con interface IEEE 1394. En él se llevarán a cabo las capturas de secuencias en distintas condiciones de iluminación, distancias, movimientos y vestuario del paciente (que será un voluntario). Las secuencias se almacenarán y se utilizarán como pruebas en la aplicación multiagente generada en el proyecto. Por lo tanto, inicialmente el sistema no trabajará en tiempo real.

Una vez obtenidos los datos de situación, orientación y postura se llevarán a la aplicación 3D interactiva donde, ahora sí, se podrán observar en tiempo real, cambiando el punto de vista y navegando por el entorno. Los datos visuales se complementarán con otro tipo de información aportada por el resto del sistema de inteligencia permitiendo el control completo del paciente por parte de la persona encargada de su cuidado.

Después de comprobado y optimizado el sistema con secuencias grabadas, es el momento de pasar a la utilización del mismo en tiempo real, ello requerirá gran capacidad de cálculo pues los procesos de búsqueda de correspondencias, determinación de profundidades, seguimiento y representación 3D son muy costosos.

6. Conclusiones y trabajo futuro

Dentro del proyecto global, el presente trabajo se ha centrado particularmente en el estudio de las técnicas empleadas en las tareas de seguimiento, estimación de la pose y reconocimiento a partir de la información visual obtenida por cámaras estereoscópicas, así como de la presentación de los datos obtenidos en el entorno tridimensional interactivo. a lo largo del proceso de formación que supone el estudio de las técnicas de análisis visual, se han analizado las diferentes fases necesarias para construir un módulo basado en visión artificial (figura 3). Una conclusión que se obtuvo fue que los algoritmos de visión y representación son muy costosos y toman tiempos de respuesta muy altos, lo que en nuestro

caso es un gran inconveniente ya que muchas de las aplicaciones requieren operación en tiempo real. En este sentido, en el trabajo paralelo se incorpora un nuevo enfoque en el que el análisis de imágenes obtenidas por las cámaras se hace mediante agentes inteligentes. La arquitectura de agentes permite automatizar el análisis y la reconstrucción y optimizar su procesamiento, ofreciendo mayores niveles de flexibilidad y escalabilidad que las arquitecturas centralizadas y consiguiendo respuestas en tiempo de ejecución, autonomía y continuidad de los servicios.

Después del estudio de las técnicas en estereoscopia, es necesario elaborar las distintas aplicaciones correspondientes a las fases involucradas. Podemos diferenciar como líneas de trabajo futuro:

- Continuación con el desarrollo de las funcionalidades del análisis, comparando los resultados y eligiendo las mejores técnicas y utilizando el hardware Point Grey[5].
- Desarrollo de técnicas de paralelización de los algoritmos de seguimiento, estimación de pose, reconocimiento y representación 3D para la optimización del sistema.
- Integración de las funcionalidades dentro de la arquitectura distribuida.
- Integración de los módulos del sistema global (módulo de análisis, módulo de localización por radiofrecuencia y módulo de reconstrucción 3D).

Referencias Bibliográficas

1. Carretero N. Bermejo a. B. *Inteligencia ambiental*. CEDITEC: Centro de Difusión de Tecnologías, Universidad Politécnica de Madrid, 2003.
2. Afsarmanesh H. Camarinha-Matos L. Design of a virtual community infrastructure for elderly care. In *3rd IFIP Working Conference on Infrastructures for Virtual Enterprises*, 2002.
3. Cortellessa G. Grisetti G. Giuliani M. et al. Cesta a., Bahadori S. The robocare project, cognitive systems for the care of the elderly. In *Proceedings of International Conference on aging, Disability and Independence.*, 2003.
4. De Paz Y. Tapia D. I. Corchado J. M., Bajo J. *Intelligent Environment for Monitoring alzheimer Patients, agent Technology for Health Care. Decision Support Systems*. Eslevier Science, 2006.
5. Point Grey Research Inc. <http://www.ptgrey.com/>. última consulta: 07-12-2008, 2008.
6. Lipton L. *Foundations of the Stereoscopic Cinema*. Van Nostrand Reinhold Company, 1982.
7. De la Escalera a. *Visión por computador. Fundamentos y métodos*. Prentice Hall, 2001.
8. Granum E. Moeslund T. B. a survey of computer vision-based human motion capture. *Comput. Vis. Image Underst.*, 2001.
9. Quest3D. <http://www.quest3d.com/>. última consulta: 07-12-2008, 2008.
10. Verry a. Trucco E. *Introductory techniques for 3D Computer Vision*. Prentice Hall, 1998.
11. López de Ipiña D. Vázquez I. *Inteligencia ambiental: la presencia invisible*. Revista solo programadores, 2005.

12. ZigBee. *ZigBee Specification Document*. ZigBee Standards Organization. ZigBee alliance., 2006.

Librería Simulink de Sistemas HVAC

Héctor Salazar y Eladio Sanz

Departamento de Informática y automática, Facultad de Ciencias
Universidad de Salamanca, Salamanca, España
hsalazar@usal.es esanz@usal.es

Resumen. Este escrito presenta una librería Simulink pública para sistemas de Calentamiento, Ventilación y aire acondicionado (HVAC). Está constituida por cuatro bloques principales: un recinto donde se controla la temperatura, una unidad de tratamiento de aire (UTA), una unidad de producción de energía y finalmente un regulador que controla la temperatura y la ventilación del recinto. Estos bloques son el resultado de su modelación matemática, su implementación en subsistemas Simulink, y su integración en una librería, llamada *ISTemperature*. Esta librería pretende ser una herramienta pública que permita construir sistemas de manera confiable, sencilla y modular para la implementación de modelos HVAC. En el presente trabajo se desarrolla un sistema HVAC sencillo, controlado por un regulador de acción proporcional, basándose en componenetes de la librería presentada.

1. Introducción

En la actualidad el uso de herramientas de simulación en sistemas de ingeniería es de suma importancia, pues nos permite desarrollar sistemas complejos sin necesidad de su implementación, además de poder realizar diferentes pruebas con un considerable ahorro de tiempo y de recursos. a pesar de la gran variedad de herramientas de simulación, Matlab/Simulink sigue siendo la más utilizada dentro del ámbito académico, debido, entre otras cosas, a su amplia variedad de librerías que permiten simular diferentes tipos de sistemas. así, el desarrollo de librerías de Matlab/Simulink es cada día mayor, aumentando la gama asociada a diferentes campos de la ingeniería.

Herramientas de simulación que implementan sistemas de calefacción, refrigeración, acondicionamiento de aire y ventilación han sido creadas, basándose en desarrollos propios o en librerías generadas en Simulink [1] [2] [3] [4]. En el presente trabajo se presenta una librería pública de Simulink sobre sistemas HVAC. En ella están desarrollados los bloques de los elementos principales que constituyen este tipo de sistemas. Dichos elementos constitutivos están diseñados a partir de una modelación matemática de diversos subsistemas termodinámicos, representativos de un sistema HVAC. Finalmente, un modelo HVAC es construido a partir de la librería propuesta en la presente investigación. En dicho sistema, se controla la temperatura y la ventilación del recinto mediante cinco diferentes señales de control, que pueden actuar una a la vez o de manera simultánea para llevar a cabo un control óptimo. El desempeño de la librería es

observado mediante la utilización de una acción de control de tipo proporcional. El desempeño tiene que ver con el comportamiento de la variable a controlar (temperatura dentro de un recinto) y con el consumo de energía asociado a dicho comportamiento.

2. Sistema HVAC

Actualmente, hablar de sistemas de climatización es hablar de sistemas que involucren, no sólo un confort térmico, sino además calidad del aire y ahorro energético [5] [6]. Con calidad de aire se consideran diversas cuestiones, desde remoción de partículas patógenas, hasta renovación de aire exahusto, según estándares y exigencias mínimas para el buen desempeño de las personas que se desarrollan en un determinado recinto [7] [8] [9] [10] [11]. Uno de los sistemas más desarrollados en la actualidad es el HVAC, regulado y estandarizado por muchas organizaciones tales como ASHRE, SMACNA, ACCA, Uniform Mechanical Code, Internacional Mechanical Code y AMCA.

El desarrollo de los componentes de un sistema HVAC va de la mano con el desarrollo de procesos térmicos y sistemas industriales, donde se involucran nuevos métodos de modernización, eficiencia energética y nuevos paradigmas de control automático de procesos. Es importante mencionar que últimamente las investigaciones encaminadas a una mejor eficiencia energética han dado lugar a sistemas donde el consumo energético es un factor inherente a estos sistemas [5] [12] [6] [13].

3. Operación del Sistema

Existen diferentes esquemas sobre los que opera un sistema HVAC [14] [15], sin embargo el diagrama básico es como el que se muestra en la Fig.1. La manera en que opera es la siguiente:

El aire que entra a la unidad de intercambio aire-agua (UTA) es calentado por conducción y convección, adquiriendo de esta forma una temperatura T_{as} . Este aire es inyectado al recinto y por convección el recinto adquiere una temperatura T_{ar} , que es la temperatura de confort deseada. El aire del recinto, regresa a la UTA, pero antes es *renovado*. Para este propósito el aire del recinto pasa por una etapa de ventilación, donde una parte del aire es expulsado del sistema hacia el exterior y otra parte, proporcional al expulsado, es inyectado desde el exterior. Una vez renovado (ventilado) el aire, éste pasa nuevamente por la unidad de intercambio aire-agua y el ciclo se repite.

4. Modelación del Sistema

El presente modelo define el comportamiento dinámico del sistema HVAC que se muestra en la Fig.1. El sistema está formado básicamente por un recinto, una unidad de intercambio aire-agua y una unidad de producción (caldera).

4.1. Recinto

El modelo del recinto está caracterizado por dos variables de estado: la temperatura del recinto T_{ar} y la temperatura de las paredes internas T_{pi} . El aire en el recinto se considera totalmente mezclado, de tal forma que la distribución de temperatura del recinto es uniforme y la dinámica de la temperatura de la zona puede ser expresada en un modelo de capacidad térmica. además, el efecto que una pared tiene en la temperatura del recinto es el mismo efecto que tienen el resto de las paredes. Por otro lado, el piso y el techo no tienen efecto sobre la temperatura del recinto. La densidad del aire se asume como constante y no se ve afectada por los cambios en la temperatura y la razón de humedad del recinto. Las pérdidas de presión a través del recinto y en la sección de intercambio aire-agua son despreciables. La gente, las luces y las condiciones extremas de temperatura son las entradas no controlables. Bajo las suposiciones anteriores, las ecuaciones de balance de energía del recinto son las siguientes

$$C_{ar} \frac{dT_{ar}}{dt} = -\rho_a c_a (T_{ar} - T_{as}) u_1 q_{ar} - (UA)_v (T_{ar} - T_{ex}) - (UA)_{pi} (T_{ar} - T_{pi}) + Q_p \quad (1)$$

$$C_p \frac{dT_p}{dt} = -(UA)_{pe} (T_{pe} - T_{ex}) + (UA)_{pi} (T_{ar} - T_{pi}) \quad (2)$$

La Ecuación (1) establece que la velocidad de cambio de energía en el recinto es igual a la diferencia entre la energía transferida al recinto, ya sea por conducción o por convección y la energía removida del recinto. En (2) la velocidad de cambio de energía en las paredes es igual a la energía transferida a través de las paredes, debido a la diferencia de temperaturas entre el aire interior y exterior.

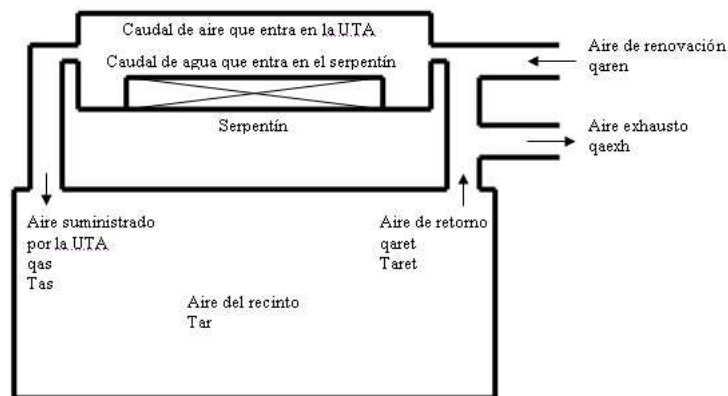


Figura 1. Diagrama esquemático de un sistema HVAC

4.2. Unidad de Intercambio aire-agua (UTA)

Se considera que la velocidad de flujo de la masa del agua dentro del serpentín es variable, por medio de una válvula de tres vías. De esta forma tendremos dos flujos complementarios, regulados por u_2 . Por un lado, el caudal de agua suministrada por la caldera en la válvula de tres vías q_{wc} , a una temperatura de T_{wc} . Y por otro el caudal de retorno en la válvula de tres vías q_{wr} , a una temperatura de T_{wr} . además, se asume que el material que está hecho el serpentín es altamente conductivo, de tal forma que su resistencia térmica se considera despreciable. Por otro lado, se asume que el aire que entra a la zona de intercambio $u_1 q_{as}$ está dado por la superposición del caudal de aire de renovación y el caudal de aire no contaminado, $u_3 q_{as} + (1 - u_3) q_{as}$.

De esta forma, el balance de energía entre el agua caliente y el aire frío puede ser expresado como

$$C_{as} \frac{dT_{as}}{dt} = -\rho_a c_a q_{as} T_{as} + \rho_a c_a (u_3 q_{as} T_{ex} + (1 - u_3) q_{as} T_{ar}) + (UA)_{UTA} (T_{wr} - T_{as}) \quad (3)$$

$$C_{wr} \frac{dT_{wr}}{dt} = -\rho_w c_w q_w T_{wr} + \rho_w c_w (u_2 q_w T_{wc} + (1 - u_2) q_w T_{wr}) - (UA)_{UTA} (T_{wr} - T_{as}) \quad (4)$$

4.3. Unidad de Producción (Caldera)

Considerando la caldera como un recinto a temperatura T_{wc} , podemos modelarla de la siguiente forma

$$C_c \frac{dT_{wc}}{dt} = \rho_w c_w u_2 q_w T_{wr} - \rho_w c_w u_2 q_w T_{wc} - (UA)_c (T_{wc} - T) + \eta P_N u_4 \quad (5)$$

5. Variables del Sistema

Dado el modelo del sistema HVAC, podemos definir las variables que operan en el mismo de la siguiente manera:

Variables de Estado

$x_1 = T_{ar}$	Temperatura del aire del recinto
$x_2 = T_p$	Temperatura de las paredes
$x_3 = T_{as}$	Temperatura del aire suministrado (sale de la unidad UTA)
$x_4 = T_{wr}$	Temperatura del agua de retorno en la válvula de tres vías
$x_5 = T_{wc}$	Temperatura del agua suministrada por la caldera en la válvula de tres vías

Variables de Control

u_1	% caudal nominal aire recinto
u_2	% caudal de agua de caldera en válvula 3 vías
u_3	% caudal de aire de renovación en la sala
u_4	% caudal nominal gas

Perturbaciones

u_5	Temperatura exterior
u_6	aportación calorífica externa: ocupación, radiación, etc.

Por otro lado, es importante mencionar que el valor de las variables anteriores está dado en 1%. además los valores de u_1 y u_3 tienen otros límites inferiores, impuestos por la renovación reglamentaria del aire interior, en base a condiciones ambientales y de ocupación. Dichos valores están en función de u_6 y están dados por (6) y (7).

$$q_R u_1 \geq q_L u_6 \Rightarrow u_1 \geq \frac{q_L u_6}{q_R} \quad (6)$$

$$q_R u_1 u_3 \geq q_L u_6 \Rightarrow u_1 u_3 \geq \frac{q_L u_6}{q_R} \quad (7)$$

6. Sistema de Ecuaciones

Una vez planteadas las ecuaciones y las variables, podemos establecer las ecuaciones en forma de representación de variables de estado, y a su vez en términos de parámetros y coeficientes.

$$\dot{x}_1 = (a_{11})_n x_1 u_1 + a_{11} x_1 + a_{12} x_2 + (a_{13})_n x_3 u_1 + b_{15} u_5 + b_{16} u_6 \quad (8)$$

$$\dot{x}_2 = a_{21} x_1 + a_{22} x_2 + b_{25} u_5 \quad (9)$$

$$\dot{x}_3 = a_{31} x_1 + (a_{31})_n x_1 u_3 + a_{33} x_3 + a_{34} x_4 + (b_{35})_n u_3 u_5 \quad (10)$$

$$\dot{x}_4 = a_{43} x_3 + a_{44} x_4 + (a_{44})_n x_4 u_2 + (a_{45})_n x_5 u_2 \quad (11)$$

$$\dot{x}_5 = a_{51} x_1 + (a_{54})_n x_4 u_2 + (a_{55})_n x_5 u_2 + a_{55} x_5 + b_{54} u_4 \quad (12)$$

Dónde

$$(a_{11})_n = -\frac{q_R}{V_R} \quad a_{11} = -\frac{(UA)_v + (UA)_{pi}}{C_R} \quad a_{12} = \frac{(UA)_{pi}}{C_R} \quad (a_{13})_n = \frac{q_R}{V_R}$$

$$b_{15} = \frac{(UA)_v}{C_R} \quad b_{16} = \frac{1}{C_R}$$

$$\begin{aligned}
 a_{21} &= \frac{(UA)_{pi}}{C_p} & a_{22} &= -\frac{(UA)_{pe}+(UA)_{pi}}{C_p} \\
 b_{25} &= \frac{(UA)_{pe}}{C_p} \\
 a_{31} &= \frac{\rho_a c_a q_a}{C_a} & (a_{31})_n &= -\frac{\rho_a c_a q_a}{C_a} & a_{33} &= -\frac{(UA)_{UTA}}{C_a} - \frac{\rho_a c_a q_a}{C_a} & a_{34} &= \frac{(UA)_{UTA}}{C_a} \\
 (b_{33})_n &= \frac{\rho_a c_a q_a}{C_a} \\
 a_{43} &= \frac{(UA)_{UTA}}{C_w} & a_{44} &= -\frac{(UA)_{UTA}}{C_w} & (a_{44})_n &= -\frac{\rho_w c_w q_w}{C_w} & (a_{45})_n &= \frac{\rho_w c_w q_w}{C_w} \\
 a_{51} &= \frac{(UA)_c}{C_c} & (a_{54})_n &= \frac{\rho_w c_w q_w}{C_c} & (a_{55})_n &= -\frac{\rho_w c_w q_w}{C_c} & a_{55} &= -\frac{(UA)_c}{C_c} \\
 b_{54} &= \frac{\eta P_N}{C_c}
 \end{aligned}$$

7. Librería Simulink *ISTemperature*

La librería que se desarrolla en el presente trabajo tiene que ver con los elementos del sistema HVAC descritos en la Sec.4. Cada elemento desarrollado en la librería, hace uso de máscaras para la implementación de los parámetros de entrada, así como su inicialización y la manera en que estos son visualizados. así mismo, el uso de máscaras permite la implementación del soporte de ayuda “Help” que es muy conocido entre los usuarios de Simulink.

De esta forma, la librería cuenta con un recinto, una unidad de intercambio aire-agua, y una unidad de producción (caldera), como bloques primordiales. además, de existir otros elementos periféricos que le dan a la librería una mayor completitud para poder construir sistemas modulares y sencillos.

La librería Simulink de sistemas HVAC, se halla bajo el nombre de *ISTemperature*. Los elementos disponibles para la librería se encuentran agrupados en diferentes rubros, según su funcionalidad:

- Control de Flujo
 - Válvula de 3 vías (3-Way-Valve)
- Intercambiadores
 - UTA (UTA)
- Producción
 - Caldera (Boiling)
- Recintos
 - Recinto (Room)
- Reguladores
 - Regulador (Regulator)

- Varios
 - Costo (Cost)
 - Parámetros (Parameters)
 - Carga térmica (Termal_Load)

Todos los componentes anteriores también se encuentran agrupados completamente bajo el rubro HVAC.

8. Resultados

El control de un sistema HVAC es un tema muy estudiado y desarrollado en nuestros días. Desde controles clásicos hasta controles inteligentes han sido implementados para lograr un control óptimo que involucre al mismo tiempo, buen desempeño, funcionalidad y ahorro energético [16] [17] [18] [19] [20] [21] [22].

En el presente trabajo, se realizó un control de un sistema HVAC. El sistema es implementado con la librería *ISTemperature*. Se realizó una simulación considerando un tiempo de simulación de 30000 segundos.

El primer paso fué sintonizar el regulador de la señal de control. Para ello se seleccionó la señal u_2 como variable y el resto de las señales como constantes. El valor de estas señales constantes fueron:

- $u_1 = 0.9$
- $u_3 = 0.15$
- $u_4 = 1$

así mismo la acción de control elegida fue de tipo proporcional, con un *set-point* ubicado en los $25^{\circ}C$, con un tolerancia permisible del 2% con respecto al valor final. Una temperatura externa de $7^{\circ}C$ y una ocupación de 10 personas.

Se realizaron diversas pruebas, variando la ganancia proporcional hasta obtener resultados óptimos. Los resultados de estas pruebas son los que se muestran en la Tabla1.

Se puede ver que al incrementar la ganancia, se va incrementando el sobreimpulso, así como el tiempo de estabilización. Con respecto al sobreimpulso, es de alrededor del 20%. Con respecto al consumo energético, es igual en todos los casos, pues la variable que maneja el suministro de combustible, u_4 , se mantiene constante en todo momento.

a partir de los resultados anteriores, se sintonizó de manera manual el regulador que controla la señal u_2 , y el resto de las señales se mantuvieron constantes. De esta forma y considerando una acción de control Proporcional, el valor de la ganancia del regulador y el valor constante de las señales fueron los siguientes:

- Valor constante u_1
 - $u_1 = 0.9$
- Regulador u_2

Tabla 1. Respuesta de salida con u_2 como variable, usando una acción de control Proporcional

Ganancia Proporcional K	Tiempo de estabilización ($^{\circ}C$)	Sobreimpulso (segundos)	Consumo Energético (unidades)
-5	29.5477	9840	3.0e+06
-10	29.8845	10560	3.0e+06
-15	29.9694	10770	3.0e+06
-20	30.0083	10890	3.0e+06
-30	30.0453	11040	3.0e+06
-40	30.0615	11220	3.0e+06
-50	30.0728	11310	3.0e+06

- Ganancia proporcional $K = -2$
- Valor constante u_3
 - $u_3 = 0.15$
- Valor constante u_4
 - $u_4 = 1$

a partir de los valores anteriores se realizaron varias pruebas, variando el *setpoint*, la temperatura externa y la ocupación. En la Fig.2 se muestra la respuesta del sistema, al variar de manera simultánea las siguientes características:

- Un *setpoint* inicial de $25^{\circ}C$
- Una temperatura externa $u_5 = 5^{\circ}C$ con presencia de ruido
- Una ocupación $u_6 = 5$ personas

La respuesta es amortiguada, con un sobreimpulso del 10.3% y no tiene error de offset. así mismo el tiempo de establecimiento es de 6750 seg. Podemos decir que los cambios realizados en el sistema fueron compensados de manera correcta por la señal de control u_2 del sistema HVAC.

9. Conclusiones

La necesidad de utilizar herramientas de simulación de sistemas de calefacción, acondicionamiento de aire y ventilación, para el diseño de edificios sustentables se hace cada vez más imperiosa. La librería presentada en el actual trabajo pretende ser una respuesta a esta necesidad, presentando bloques constitutivos de un sistema HVAC, así como elementos periféricos que le dan mayor completitud, soporte, flexibilidad y modularidad a la librería. a través de la presente librería se muestra la facilidad de parametrización y el uso de modelos para un

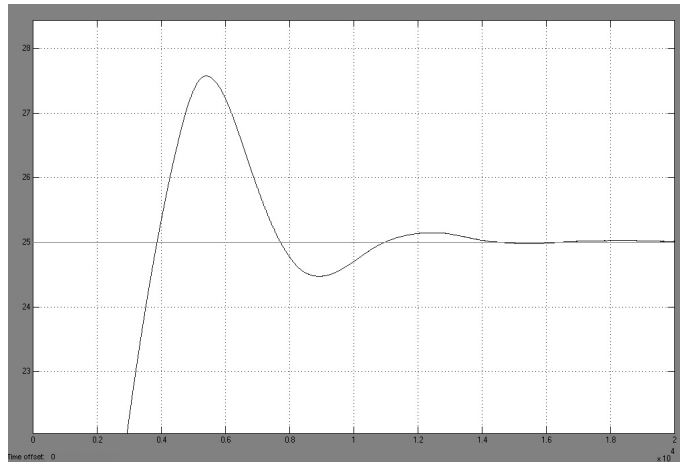


Figura 2. Respuesta de un sistema HVAC, con un *setpoint* de 25°C , una temperatura externa $u_5 = 5^{\circ}\text{C}$ y una ocupación $u_6 = 5$.

sistema HVAC. Mediante la distribución adecuada de entradas y parámetros de entrada, se construyen bloques que simulan de una manera amigable la estructura funcional de un sistema HVAC, así como la implementación de un control óptimo. La construcción modular de cada elemento constituyente del sistema permite generar sistemas según necesidades, mediante la introducción adecuada de los parámetros de cada bloque.

Los resultados obtenidos con la librería nos demuestra la fiabilidad de la misma, pues se genera un control cuya respuesta de salida presenta un comportamiento bastante óptimo. La respuesta generada, en base a las cuatro señales de control, permitió ir regulando de manera satisfactoria el control de la temperatura de un recinto. Por tal motivo, la herramienta presentada en este trabajo puede ser considerada óptima para del diseño de sistemas HVAC.

Pero no sólo la fiabilidad de los resultados del sistema resultó ser óptima, sino que además la manera de construir el sistema resultó ser bastante accesible, simple, rápida y modular. La librería nos proporciona elementos hechos a la medida, sin necesidad de ir divagando por todas las librerías Simulink en busca de elementos adecuados para la construcción de un sistema HVAC.

Como trabajo futuro se considera el aumento de componentes propios del sistema HVAC que le den mayor riqueza a la librería, como por ejemplo, *conductos* donde se simule las pérdidas de energía por rozamiento con el material conductor, diferentes tipos de recinto (modelos de diversos ordenes), así como la inserción de diferentes sistemas de ventilación, como un VaV (Variable air Volume). También queda como trabajo futuro el desarrollo de controles más inteligentes, que permitan desarrollar sistemas más complejos, por ejemplo, sistemas de control predictivo, borroso o neuronal.

Referencias Bibliográficas

1. A. Husaunndee, R. Lahrech, H.V.N.J.V.: Simbad: a simulation toolbox for the design and test of HVAC control systems. In: Proceedings of the Fifth International IBPSa Conference, Prague. Czech Republic, 2: 269 276. (2005)
2. Angela Sasic Kalagasidis, Peter Weitzmann, T.R.N.R.P.C.E.H.C.R.: The international building physics toolbox in simulink. ScienceDirect Energy and Buildings **39** (2007) 665 – 674
3. Peter Weitzmann, Angela Sasic Kalagasidis, T.R.N.R.P., Hagentoft, C.E.: Presentation of the international building physics toolbox for simulink. Technical report, Technical University of Denmark
4. Annual Performance Simulation and Benchmarking of an advanced Solar Collector using TRNSYS. Proceedings of Solar 2002 - Australian and New Zealand Solar Energy Society. (2002)
5. Stefano Paolo Corgnati, Enrico Fabrizio, M.F.: The impact of indoor thermal conditions, system controls and building types on the building energy demand. ScienceDirect Energy and Buildings **40** (2008) 627 – 636
6. Ruey-Lung Hwang, Ming-Jen Cheng, T.P.L.M.C.H.: Thermal perceptions, general adaptation methods and occupant's idea about the trade-off between thermal comfort and energy saving in hot humid regions. ScienceDirect Building and Environment **30** (2008) 1 – 7
7. R. Kosonen, F.T.: The effect of perceived indoor air quality on productivity loss. ScienceDirect Energy and Buildings **36** (2004) 981 – 986
8. Liang Zhou, F.H.: Optimization of ventilation systems in office environment, part i: Methodology. ScienceDirect Building and Environment **30** (2008) 1 – 6
9. Liang Zhou, F.H.: Optimization of ventilation systems in office environment, part ii: Results and discussions. ScienceDirect Building and Environment **30** (2008) 1 – 9
10. Hens, H.S.: Thermal comfort in office buildings: Two case studies commented. ScienceDirect Building and Environment (2008) 1 – 10
11. Elena G. Dascalakia, Argyro Lagoudib, C.a.B.a.G.G.: air quality in hospital operating rooms. ScienceDirect Building and Environment (2008) 1945 – 1952
12. S. Ginestet, D. Marchio, O.M.: Evaluation of faults impacts on energy consumption and indoor air quality on an air handling unit. ScienceDirect Energy and Buildings **40** (2008) 51 – 57
13. Ashfaq Ahmed Chowdhury, M.G. Rasul, M.K.: Thermal-comfort analysis and simulation for various low-energy cooling-technologies applied to an office building in a subtropical climate. ScienceDirect Applied Energy **85** (2008) 449 – 462
14. Bourhan Tashtoush, M. Molhim, M.a.R.: Dynamic model of an HVAC system for control analysis. ScienceDirect Energy **30** (2005) 1730 – 1745
15. Shui Yuan, R.P.: Multiple-zone ventilation and temperature control of a single-duct VAV system using model predictive strategy. ScienceDirect Energy and Buildings **38** (2006) 1248 – 1261
16. E. Semsar-Kazerouni, M.J.Y., Lucas, C.: Nonlinear control and disturbance decoupling of HVAC systems using feedback linearization and backstepping with load estimation. IEEE Transactions on Control Systems Technology **16** (2008) 12
17. Chengyi Guo, Q.S., Cai, W.: A neural network assisted cascade control system for air handling unit. IEEE Transactions on Industrial Electronics **54** (2007) 9
18. Michael Anderson, Michael Buehner, P.Y.M.D.H.C.a.J.T., Hodgson, D.: MIMO robust control for HVAC systems. IEEE Transactions on Control Systems Technology **16** (2008) 1 – 9

19. E.Z.E. Conceio, M.M.J.R. Lúcio, a.R.E.C.: Development of a temperature control model used in HVAC systems in school spaces in mediterranean climate. ScienceDirect Building and Environment (2008) 1 – 18
20. Servet Soyguder, Mehmet Karakose, H.a.: Design and simulation of self-tuning pid-type fuzzy adaptive control for an expert HVAC system. ScienceDirect Expert Systems with applications (2008) 1 – 8
21. Jianbo Bai, Shengwei Wang, X.Z.: Development of an adaptive smith predictor-based self-tuning pi controller for an HVAC system in a test room. ScienceDirect Energy and Buildings (2008) 1 – 9
22. Leephakpreeda, T.: Grey prediction on indoor comfort temperature for HVAC systems. Expert Systems with applications **34** (2008) 2284 – 2289

Simulaciones con EJS y Matlab/Simulink

Raquel Sánchez Díaz, Eladio Sanz, and Belén Pérez-Lancho

Departamento de Informática y automática, Universidad de Salamanca
Plaza de la Merced s/n, 37008, Salamanca, Spain
kela@usal.es, esanz@usal.es, lancho@usal.es

Resumen. Las simulaciones se han convertido en herramientas fundamentales tanto en el aprendizaje del control como en la investigación y experimentación con el sistema, ya que proporcionan grandes ventajas. En este trabajo se desarrollarán distintas simulaciones y se probarán diferentes arquitecturas, obteniendo así las bases para el empleo de las herramientas que en un futuro se pretenden utilizar en el desarrollo de laboratorios virtuales remotos y que cada día están más presentes y toman mayor importancia en las investigaciones que se realizan en este campo: Matlab, EJS y JIM.

Key words: Simulación; control; Matlab/Simulink; EJS; JIM; simulaciones basadas en *web*; laboratorio virtual; laboratorio remoto.

1. Introducción

Las nuevas tecnologías de la información y la comunicación (TIC) están transformando muchos aspectos de la sociedad. El paradigma de la educación tradicional no es ajeno a ello, pasando de ser una etapa de preparación a un proceso de constante formación. Un proceso continuo de enseñanza será más exitoso mientras más flexible sea respecto de la hora, el número de personas y el lugar en que se realice la instrucción. Las TIC y sus novedosas características de comunicación (como Internet) y de aprendizaje (como la simulación o la tele-presencia), presentan sin duda un enorme potencial para el desarrollo de la educación a distancia.

Muchas son las áreas de conocimiento e investigación en ingeniería que se han nutrido de los avances de las tecnologías basadas en Internet y la importancia que esta ha tenido en la enseñanza basada en *web*. El control automático es una de las áreas técnicas de la ingeniería en la que el impacto de estas tecnologías ha sido especialmente significativo en todo lo relativo al desarrollo de herramientas para el aprendizaje y experimentación.

Por un lado las simulaciones son herramientas fundamentales tanto en el aprendizaje del control como en la investigación y experimentación con el sistema. Proporcionan grandes ventajas como la interactividad (ver en tiempo de ejecución la evolución de los diferentes aspectos y respuestas del sistema ante cualquier cambio introducido en los parámetros), el poder simular en unos minutos y cuantas veces se quiera lo que en la realidad podría tardar semanas, ahorrar los costes y evitar los riesgos que trabajar con el sistema real supondría,

y además poder operar con el sistema simulado como si fuera el real desde cualquier punto donde se tenga un ordenador y a cualquier hora.

Por otro lado, Internet proporciona los medios necesarios para el intercambio de información entre computadores ubicados en cualquier parte del planeta. Esta característica permite el desarrollo de simulaciones distribuidas, es decir, el cálculo computacional requerido puede ejecutarse en diferentes computadores conectados a Internet, consiguiendo con ello una versión remota de la simulación.

Parece entonces que la Ingeniería de Control debe aprovechar estas capacidades facilitadas por las herramientas de simulación e Internet. En los últimos años se ha observado un aumento significativo en el desarrollo de proyectos y trabajos de investigación orientados a la mejora de la educación en automática. Ya son muchos los ejemplos que se pueden encontrar sobre interactividad, simulación basada en *web*, laboratorios virtuales, remotos, etc.

Las ventajas que tiene el trabajo con simulaciones, tanto para fines de experimentación con el sistema, como para el aprendizaje del control, nos han llevado a comenzar un trabajo de investigación cuyos inicios y primeros resultados presentamos en esta memoria. Principalmente se pretende dotar de interacción e interfaz gráfica a los modelos de una sala de la biblioteca que se tienen para simular cómo varía su temperatura en función de varios parámetros, a la vez que se prueban las distintas arquitecturas y se obtienen las bases para el empleo de las herramientas que en un futuro esperamos utilizar en el desarrollo de laboratorios virtuales remotos y que cada día están más presentes y toman mayor importancia en las investigaciones que se realizan en este campo: Matlab, EJS y JIM.

2. Objetivos

Para este trabajo se han planteado varios objetivos. Unos a corto plazo, que son los que se pretenden cumplir al finalizar el trabajo de tesis de máster y presentaremos ahora, y otros a largo plazo, de los que hablaremos en las líneas de trabajo futuro.

Esta es la primera parte de un trabajo de investigación más amplio en el cual una de las cosas que se pretende es desarrollar laboratorios virtuales remotos e incluirlos en la red automatL@bs, de la que hablaremos después. Por lo tanto, en relación con esto nuestros objetivos son:

- Hacer una revisión de cómo está la investigación en este campo para así ver qué cosas serían interesantes de hacer y hacia dónde se podría investigar.
- adquirir los conocimientos básicos que se necesitan para el desarrollo de estos laboratorios:
 - aprender a utilizar EJS para cosas más complejas y documentar los problemas que vayan surgiendo y no tengan respuesta en los manuales que existen en la actualidad.
 - intentar conectar EJS con Matlab y al igual que antes ampliar la información existente con lo aprendido a lo largo de nuestra experiencia.

- probar las distintas arquitecturas y posibilidades que EJS y Matlab nos ofrecen para ver si alguna nos puede interesar especialmente.
- comprobar que EJS cumple con los requisitos que de él se esperan para decidir si será la herramienta que se utilizará en el desarrollo de estos laboratorios.

También se quiere trabajar con el sistema de climatización de la Biblioteca abraham Zacut. Se disponen de unos modelos en Simulink de este sistema, y además también tenemos las ecuaciones que lo modelan. Por lo tanto, haciendo uso de los modelos y las herramientas de las que disponemos, se pretende:

- Crear una interfaz gráfica que represente la evolución de la temperatura en una sala de la biblioteca para que resulte más visual y cercano al usuario.
- añadir a los modelos interactividad con el usuario para que éste pueda:
 - Variar los parámetros y ver cómo afectan a la temperatura.
 - Probar diferentes tipos de control y sintonizar los controladores.
- Validar los modelos de los que se dispone utilizando datos reales que se deben recoger diariamente.
- Simular en unos minutos lo que en la realidad tardaría semanas.
- ahorrar los costes y evitar los riesgos que trabajar con el sistema real supondría.

Pretendemos que esto nos sirva para dos cosas. Por un lado se puede desarrollar una herramienta para utilizarla con fines educativos, y que así los alumnos aprendan de una forma interactiva, visual y más sencilla las bases de un sistema de control.

Por otro lado podemos desarrollar herramientas para fines de investigación y mejora del sistema de climatización, ya que si logramos obtener un buen modelo del sistema y desarrollar una aplicación que nos dote de interactividad y nos permita utilizar datos reales, podremos realizar simulaciones como si trabajáramos con el sistema real, con las ventajas que ya hemos dicho que esto supone. Haciendo uso de las herramientas que se desarrollen se puede por ejemplo buscar un buen controlador para luego instalarlo en el sistema real.

Por el momento sólo pretendemos centrarnos en la parte final del sistema de calefacción. Trabajaremos con la UTA4, que es una unidad de tratamiento de aire. Por lo tanto, los parámetros que principalmente manejaremos serán la temperatura y el caudal del aire que es impulsado a la habitación.

3. Herramientas utilizadas

Matlab/Simulink [1], es un paquete de software clásico que proporciona facilidades para la construcción de modelos de forma gráfica mediante diagramas de bloques, por lo que se ha convertido en una herramienta ampliamente utilizado en la industria y la enseñanza de gran parte de las ingenierías. Sin embargo, la interactividad proporcionada por los modelos Simulink dista bastante de la

interactividad que se pretende tener en las simulaciones. Es en este aspecto donde EJS (Easy Java Simulations) [2] puede resultar muy útil.

EJS es una herramienta software gratuita diseñada para el desarrollo de simulaciones o laboratorios virtuales interactivos. Proporciona un mecanismo propio para la descripción de sistemas y modelos dinámicos propios de la Ingeniería de Control. También provee de una enorme cantidad de elementos visuales parametrizables e interactivos, que permiten la rápida construcción de la vista de una simulación. además permite integrar modelos Simulink en las simulaciones desarrolladas en EJS, añadiendo así un alto nivel de interactividad a estos modelos.

JIM (Java Internet Matlab) [3], es un paquete escrito en Java orientado a extender las capacidades de EJS para manipular modelos Simulink. La extensión permite al usuario de EJS, ejecutar la simulación sin la necesidad de contar con Matlab/Simulink instalado en su equipo, ya que EJS establece un enlace de red con un equipo remoto, donde está instalado el servidor JIM, que sí posee Matlab/Simulink.

4. Situación actual de la investigación en este campo

La concepción tradicional de laboratorio conlleva algunos problemas como altos costes, espacio limitado, desplazamientos del alumno y un horario de acceso restringido. Debido a las grandes posibilidades que nos ofrece la tecnología actual, en los últimos años se ha trabajado muy intensamente a nivel nacional en la creación de nuevos entornos de experimentación. Estos entornos se pueden clasificar en función de la naturaleza del sistema sobre el que se opera y del acceso al recurso sobre el que se experimenta de la forma que muestra la Fig.1.

Naturaleza del sistema sobre el que se opera		
	REAL	SIMULADA
Acceso al recurso sobre el que se experimenta	LOCAL	Laboratorio tradicional
	REMOTO	Laboratorio virtual multiusuario
		Laboratorio virtual monousuario
		Telelaboratorio o laboratorio remoto

Fig. 1. Clasificación de los nuevos entornos de experimentación.

El Departamento de Informática y automática de la UNED (Universidad Nacional de Educación a Distancia) está trabajando en el desarrollo de nuevos paradigmas de laboratorios para la realización, a través de Internet, de experiencias prácticas de Control automático sobre plantas y sistemas reales o simulados [4]. actualmente los esfuerzos en investigación se dirigen hacia la creación de redes de laboratorios. Hablaremos de ellos más tarde, cuando expliquemos el proyecto automatL@bs.

4.1. Sistemas de simulación basados en web

La Simulación Basada en Web (SBW), se puede definir como el empleo de los recursos y tecnologías ofertados por la *World Wide Web* para la interacción con herramientas de simulación localizadas tanto en el lado del cliente como en el del servidor. El resultado de esto son los laboratorios virtuales basados en la WWW.

Se puede realizar una clasificación de estos sistemas en función de cuatro criterios: la ubicación del motor matemático de cálculo (local o remoto), la naturaleza del núcleo de simulación (herramientas específicas o lenguajes de propósito general), las capacidades de diseño (modificación de los parámetros o la estructura del modelo) y el grado de interactividad con la simulación (*pseudo-batch* u *on-line*).

Independientemente de las posibilidades de diseño y grado de interactividad, las SBW más frecuentes son tres:

- **Configuración monolítica:** la interfaz gráfica y motor de simulación constituyen una aplicación monolítica ejecutándose dentro del navegador WWW y residiendo en el ordenador del cliente.
- **Configuración híbrida:** la interfaz gráfica y el motor son aplicaciones independientes residiendo ambas en el mismo computador, es decir, en el del cliente. La interfaz se localiza dentro del navegador WWW, mientras que el motor habitualmente es un entorno de simulación del tipo Matlab/Simulink.
- **Configuración distribuida:** la interfaz gráfica y el motor son aplicaciones independientes y están físicamente separadas. La interfaz se localiza en el navegador WWW del cliente pero la simulación reside en el servidor remoto, pudiendo estar constituida por una aplicación desarrollada para tal efecto o por herramientas de modelado y simulación.

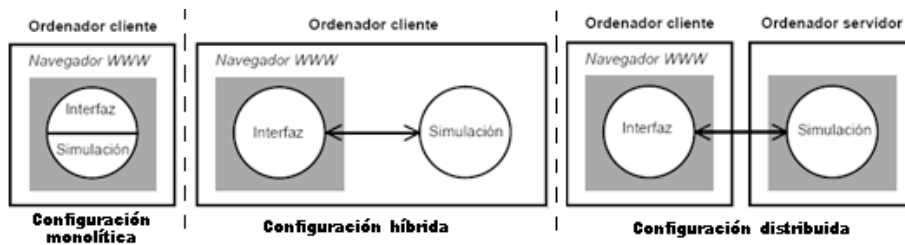


Fig. 2. Configuraciones de SBW.

4.2. Laboratorios virtuales y remotos

Los laboratorios virtuales se basan en el empleo de modelos matemáticos de los objetos físicos para la realización de experiencias prácticas desde un punto de vista docente, formativo o, incluso, con fines de investigación. Su principal ventaja es que, suponiendo que estén bien hechos, se comportan como lo haría

la planta o el experimento real, y por lo tanto se puede trabajar con ellos como si lo hicieras con el real. Esto los convierte en herramientas de gran utilidad tanto para labores de aprendizaje de control como para realizar pruebas y ajustar algoritmos de control o parámetros de controladores sin tener que actuar directamente sobre el sistema real, lo cual conllevaría restricciones de tiempo, espacio, coste, mantenimiento, etc. En el Departamento de Informática y automática de la UNED se han desarrollado varios, por ejemplo [5] y [6].

En un laboratorio de control basado en *web*, los estudiantes o investigadores pueden acceder al laboratorio desde una localización remota y se les permite interactuar con el experimento real. Es posible cambiar los parámetros de control, ejecutar los experimentos, descargar datos y visualizar resultados a través de Internet. Una de las principales ventajas de los laboratorios remotos es la optimización en el aprovechamiento de recursos y la posibilidad de compartirlos. Estos son los objetivos principales del proyecto `automatL@bs` [7].

`automatL@bs` es una red de laboratorios virtuales/remotos para la enseñanza de la automática que se constituye mediante la integración de los recursos que aportan los grupos que participan en el proyecto. Proporciona un sistema de reserva de tiempos para la realización de los experimentos y un entorno de trabajo común que facilita su aprendizaje por parte del alumno. La red de laboratorios remotos en automática `automatL@bs` es algo más que la suma de las partes que la constituyen, ya que debe ser percibida por sus usuarios como un laboratorio con una estructura uniforme independientemente de donde se encuentre la localización física de las plantas. Todo lo que necesita el alumno para conectarse a `automatL@bs` es un navegador y estar dado de alta para la realización de las prácticas. Todos los laboratorios comparten un mismo esquema de trabajo y los materiales que se proporcionan a los alumnos se han cuidado de forma tal que el desarrollo de las prácticas se pueda hacer de manera autónoma. actualmente se ofrecen nueve sistemas: bola y viga, motor de corriente continua, sistema de tres tanques ([8]), sistema de cuatro tanques, planta de cuatro variables, RobUaLab, Roto-ímán, servomotor y sistema `heatflow` ([9]).

El proyecto `automatL@bs` está coordinado por el profesor Sebastián Dormido Bencomo del Departamento de Informática y automática de la UNED. En él participan grupos del área de Ingeniería de Sistemas y automática de las siguientes universidades españolas: Universidad Nacional de Educación a Distancia, Universidad de almería, Universidad de alicante, Universidad Politécnica de Valencia, Universidad Politécnica de Cataluña, Universidad Miguel Hernández y Universidad de León.

5. Resultados

Durante el tiempo dedicado a la realización de la tesis de máster se han dado los primeros pasos de lo que en un futuro se espera que sea la tesis doctoral. Hasta ahora se han ido adquiriendo los conocimientos más básicos que se necesitan, y aprendiendo a manejar las herramientas que se van a utilizar. Se ha hecho una

introducción, pequeña pero necesaria, enfocada al campo por el que se va a tratar de investigar en el futuro.

Para ir adquiriendo el manejo y los conocimientos necesarios se han desarrollado varias simulaciones con EJS y Simulink, probando las diferentes configuraciones explicadas anteriormente (Fig.2), para así poder apreciar mejor cuáles nos resultan interesantes para el futuro. además se han creado dos versiones de cada una de las simulaciones. La primera de ellas es para trabajar en local y mientras desarrollamos. Para que no fuera necesario tener instalado EJS ni Matlab en el ordenador desde el que se ejecutaran las simulaciones, se decidió crear una versión para la *web*, de manera que se trabaja con un Matlab instalado en un servidor (que nosotros hemos configurado) y se hace mediante un *applet* de Java en el navegador, por lo que tampoco sería necesario instalar EJS.

Otra labor muy importante ha sido la de aprender a desarrollar simulaciones avanzadas con EJS, ya que no es una labor trivial. En este aspecto se han encontrado varios obstáculos, debido sobre todo a que hay una falta de documentación sobre el funcionamiento de la herramienta cuando el grado de dificultad de las simulaciones que se quieren construir con ellas aumenta. afortunadamente, al final se ha encontrado solución a todos los problemas que han surgido. Para que a futuros investigadores no les suceda lo mismo, se ha explicado detalladamente en la memoria de tesis ([13]) la forma de solucionar todos esos obstáculos.

5.1. Sistema real

Hemos trabajado con el sistema de climatización de la Biblioteca abraham Zacut. Concretamente tenemos las ecuaciones de un modelo que representa la temperatura de una habitación de la biblioteca, la cual se controla mediante flujo de aire. En lo que se refiere al sistema de calefacción el modelo sólo maneja la parte correspondiente a la UTa4 (unidad de tratamiento de aire), es decir, sólo tiene en cuenta la temperatura a la que se impulsa el aire y el caudal de aire impulsado.

En nuestros ejemplos principalmente vamos a trabajar con el aire impulsado por la UTA, la temperatura que hay en el exterior y las temperaturas real y simulada de la sala. Todos los datos utilizados son datos reales que se han ido almacenando. Debido a que este es un sistema dinámico pero lento, los datos siempre son muestreados cada media hora.

5.2. Simulaciones para su ejecución en local

Se han desarrollado cuatro simulaciones. La versión local es la que se utiliza cuando se está trabajando directamente con EJS y se ejecutan desde allí. El Matlab que utilizan puede ser local o remoto.

La **primera simulación** (Fig.3) ha sido desarrollada por completo con EJS. En él se introducen las ecuaciones del modelo por lo que EJS se encarga tanto de la parte de simulación como de la interfaz. En ningún momento se utiliza Matlab ni Simulink. En esta simulación se pretende probar distintos tipos de control (lazo abierto, On/Off, P, PI y PID) para controlar la temperatura de una sala de la

Biblioteca de abraham Zacut. La variable controlada es la temperatura interior del recinto (el usuario puede fijar su valor de deseado), y la variable de control es el porcentaje de apertura de la válvula que deja entrar el aire caliente.

El usuario puede cambiar los valores de muchos parámetros del modelo: temperatura y caudal del aire que entra, ocupación, altura y acristalamiento de la sala, temperatura exterior (fijada manualmente o leída de un fichero), y ver cómo afectan a la temperatura de la sala, cuyo color variará en función de su valor. además puede variar tipo de control empleado y ajustar los parámetros de los controladores. Todas estas características hace de esta simulación un buen ejemplo para la didáctica del control o la experimentación durante la elección y sintonización de un buen controlador para el sistema real.

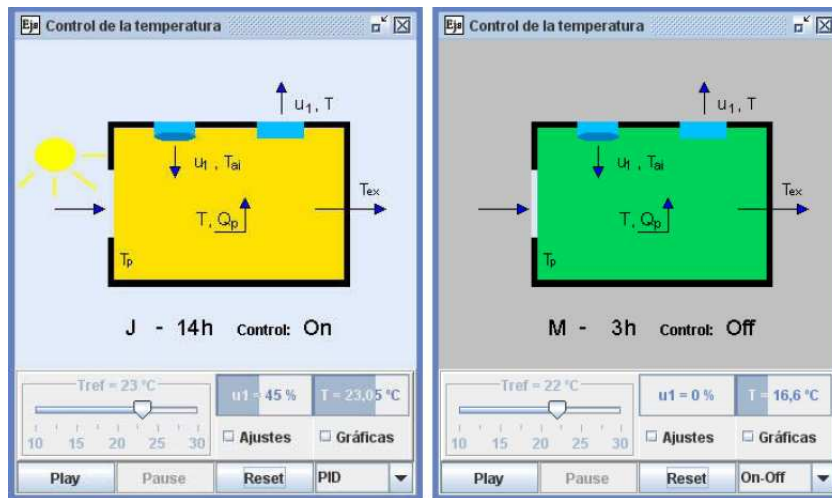


Fig. 3. Primera simulación.

En la **segunda simulación** que construimos pretendíamos aprender a conectar EJS con Simulink, de manera que el modelo del proceso fuera un modelo Simulink y que EJS simplemente se encargara de mostrar la interfaz gráfica y permitiera al usuario modificar algunos de los parámetros. Se disponía ya del modelo Simulink, en lazo abierto, que se encarga de simular la temperatura de la misma sala. En este caso todos los parámetros eran fijos (la temperatura exterior también se generaba en el propio diagrama de bloques como una onda sinusoidal modificada mediante un factor aleatorio). Para añadir esa interacción del usuario con el modelo Simulink decidimos que éste pudiera fijar desde la interfaz de EJS los valores de la temperatura y el caudal de aire impulsado.

Para el funcionamiento de esta simulación es necesario conectar las variables que se utilizan en la interfaz de EJS con sus respectivas del modelo Simulink. En la Fig.4 se muestra cómo se realiza esta conexión. al trabajar en local, con un Matlab instalado en el mismo ordenador, que es el que va a buscar EJS por defecto, sólo es necesario especificar la ruta en la que se encuentra nuestro

modelo Simulink. Más adelante, cuando se utilice un Matlab remoto, veremos que esto cambia.

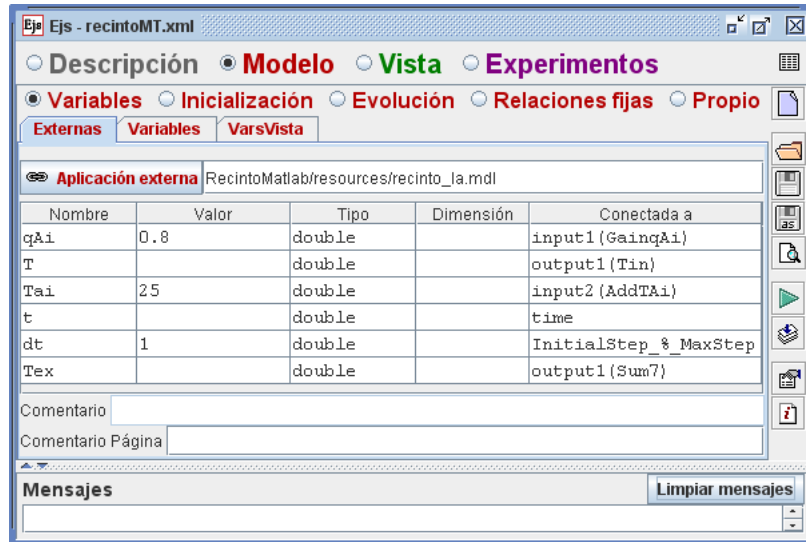


Fig. 4. Conexión de variables entre EJS y Simulink.

Una vez que conseguimos realizar correctamente las conexiones y que todo funcionara como debía, nuestro objetivo principal estaba cumplido, así que decidimos dar el siguiente paso. No nos gustaban los valores que se iban obteniendo para la temperatura exterior porque se generaban de manera aleatoria y eso no nos servía para nada. así que decidimos tomarlos de un fichero *.mat* con datos reales que previamente EJS manda cargar en el *workspace* de Matlab. De esta manera llegamos a la **tercera simulación**.

Una vez solventados todos los problemas que se presentaban en las comunicaciones entre EJS y Matlab decidimos crear una **cuarta simulación** que nos permitiera validar el modelo Simulink. Es decir, ahora utilizamos los datos reales recogidos durante cuatro días para los parámetros del modelo que varían en la realidad, y mientras realizamos la simulación vemos la diferencia entre la temperatura interior real que hay en la sala y la simulada. De esta manera podemos comprobar si nuestro modelo es adecuado y más o menos se ajusta a la realidad.

El control, en la zona que nos interesa que es la de la habitación, es en lazo abierto. No se realiza un control sobre la temperatura del aire impulsado, aunque previamente sí que se ha realizado control sobre la temperatura del agua que calienta ese aire.

En este desarrollo la unidad de tiempo es el minuto. Un paso de la simulación de Matlab se corresponde con un minuto de tiempo transcurrido en la realidad. Es importante tener esto en cuenta a la hora de introducir los datos reales para que la validación sea correcta. Todos los datos son muestreados cada treinta minutos.

Trabajamos con cuatro variables de entrada. Es decir, nuestro fichero de datos *.mat* estará compuesto por cuatro variables que se cargarán en el espacio de trabajo. Las cuatro son datos reales tomados cada media hora del sistema de climatización de la biblioteca: On/Off (encendido y apagado de la caldera/enfriadora), temperatura del aire impulsado por la UTA4 (temperatura del aire que entra en la sala), temperatura exterior y temperatura interior real (los valores se toman de un termopar colocado en la sala de la biblioteca que ha sido modelada). Esta última variable es la que se va a comparar con la simulada para validar el modelo.

La otra variable importante que representamos en esta simulación es la temperatura interior simulada, que es la temperatura de la sala que se va obteniendo según avanza la simulación. Si el modelo es adecuado debería de ser muy parecida a la real. aquí el usuario no puede interactuar con la simulación. Sólo puede observar cómo va evolucionando y con los resultados obtenidos determinar la validez del modelo. Si observamos las gráficas obtenidas (Fig.5) vemos cómo la temperatura simulada se ajusta bastante bien a la real, por lo que podemos concluir que el modelo es suficientemente bueno.

5.3. SBW utilizando un Matlab remoto

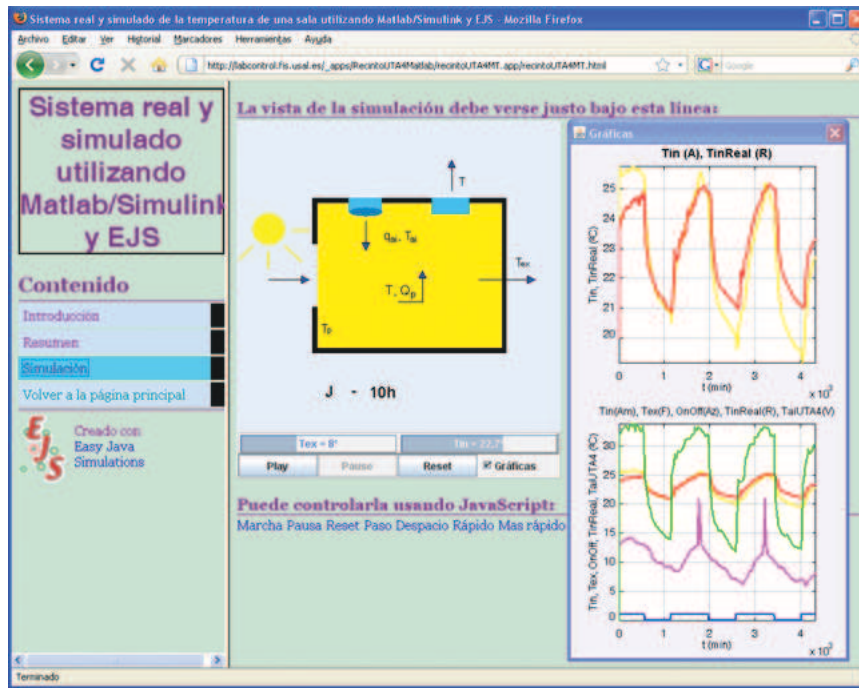


Fig. 5. Versión *web* de la cuarta simulación.

Hemos creado una versión *web* para cada una de nuestras simulaciones (ej. Fig.5). Utilizan un matlab remoto instalado en el servidor. Para ello previamente

ha sido necesario configurar el servidor adecuadamente instalando todo lo que necesita para su funcionamiento (servidor *web*, servidor JIM, Matlab, etc). De esta manera el usuario no necesita tener instalado ni EJS ni Matlab en su ordenador. Simplemente a través del navegador (y teniendo instalado Java) puede ejecutar todas las simulaciones de una manera rápida y sencilla.

La primera simulación no utiliza Matlab así que se correspondería con una configuración monolítica (Fig.2). Las demás sí, por lo que tendrían configuración distribuida. Si en vez de usar un Matlab remoto instalado en el servidor lo utilizaran local, la arquitectura sería híbrida. Para indicar a EJS dónde está instalado el Matlab que se va a utilizar hay que añadir a la ruta del fichero del modelo Simulink (Fig.4) lo siguiente:

```
<matlab:IP:Puerto>RutaDelModeloSimulink
```

donde IP es la dirección IP del servidor donde se tiene instalado Matlab, y Puerto es el número de puerto por donde JIM escucha [10].

6. Conclusiones

Una vez finalizado el trabajo hemos podido comprobar que los objetivos que nos habíamos planteado cumplir a corto plazo han sido alcanzados. Por un lado se han adquirido los conocimientos necesarios para en un futuro realizar desarrollos bastante complejos con EJS y Matlab/Simulink, y se han resuelto y documentado los problemas que han ido surgiendo. además se han aprendido los conceptos clave sobre el control de procesos.

También se ha realizado una revisión del estado de la investigación en este momento que nos ha permitido conocer en qué están trabajando otros grupos de investigación, qué proyectos se están llevando a cabo en la actualidad y si nos interesa entrar a formar parte de alguno, como por ejemplo el de automatL@bs, y además nos ha dado ideas de líneas de investigación en las que podemos trabajar en el futuro y que plantearemos en el apartado siguiente.

Se han desarrollado varias simulaciones que en conjunto cumplen los requisitos que nos habíamos planteado. Las diferentes versiones creadas nos han permitido probar las distintas arquitecturas y ver su rendimiento. La configuración que más nos ha atraído, y la que probablemente se utilizará cuando comencemos a poner las herramientas a disposición del público, es la distribuida. Las razones son dos. Por un lado hemos comprobado que el rendimiento que se obtiene en una red local es suficientemente bueno, por lo que podría utilizarse por ejemplo con los alumnos de asignaturas de control cuando estén en las aulas de la facultad. Si decidiéramos que las aplicaciones fueran accesibles desde Internet tendríamos que hacer más pruebas y tal vez plantearnos un cambio en el tipo de comunicación entre EJS y JIM (utilizar la asíncrona). Eso dependerá también de las características del sistema que pongamos a explotar.

Por otro lado hemos visto que la principal ventaja de esta arquitectura es que los usuarios que ejecuten las simulaciones no necesitan instalar en sus ordenadores ni EJS ni Matlab. Con una conexión a Internet, un navegador y Java

es suficiente. Son requisitos que por defecto cumplen prácticamente todos los usuarios. Esto supone un gran ahorro económico ya que la licencia de Matlab no es gratuita, además de una gran ventaja porque las simulaciones que se distribuyan de esta manera serán fácilmente accesibles para cualquier usuario lo que supondrá que más gente se anime a utilizarlas.

Después de trabajar con EJS podemos concluir que es un entorno muy agradable de trabajo, ya que permite crear simulaciones de una manera bastante rápida, y al utilizar una tecnología tan conocida como Java, toda la problemática de la programación está resuelta si se ha utilizado antes ese lenguaje. además el hecho de que esté preparado para conectarse con Matlab en local y en remoto y que sea capaz de generar *applets* nos ha parecido extremadamente útil. Es algo a lo que creemos que se le puede sacar mucho partido y lo que en principio vamos a utilizar mucho en los desarrollos futuros.

No obstante, también hemos encontrado varias desventajas. La principal y más importante, que sobrecarga demasiado el sistema. Si por ejemplo se quiere ir visualizando las gráficas a la vez que se trabaja con la simulación, la ejecución se ralentiza mucho. En simulaciones sencillas no hay problema, pero cuando se le empieza a exigir rendimiento puede llegar a resultar demasiado pesado.

Por otro lado, la versión con la que hemos trabajado (v3.37) crea varios problemas con las rutas de los ficheros auxiliares. Muchas veces no los encuentra y es necesario realizar muchas pruebas hasta llegar a comprender la estructura de directorios que utiliza cuando se emplean rutas relativas, y lo que considera como el directorio actual en cada modo de ejecución. En este y otros aspectos encontramos que falta bastante documentación.

El hecho de poseer conocimientos de avanzados Java ayuda bastante a la hora de desarrollar, pero también nos ha hecho darnos cuenta de que, al estar EJS preparado para usuarios sin muchos conocimientos sobre programación, no deja mucho lugar a los que sí los tienen para modificar ciertas cosas, meterse un poco más en profundidad y así lograr mejores resultados. En algunos aspectos el entorno está muy cerrado y hay cierto código que EJS genera y que el desarrollador no puede modificar, aunque sepa cómo hacerlo y en muchas ocasiones el realizar estos cambios le simplificaría la labor.

6.1. Líneas de trabajo futuro

En el futuro se pueden seguir varias líneas de trabajo distintas, todas ellas relacionadas. Por un lado sería muy interesante el incluir a nuestra universidad en la red de laboratorios automatL@bs ya que en el laboratorio de automática se dispone de varios experimentos que se podrían preparar para este fin, permitiendo su manejo desde un lugar remoto y con cámaras que permitan visualizar lo que está ocurriendo en la realidad.

además se pretende seguir trabajando con el sistema de climatización de la biblioteca para intentar mejorarlo. Se pueden desarrollar simulaciones más complejas y que manejen las demás partes de este sistema.

También incluso se podría realizar algo parecido con otros modelos de Simulink, en concreto con el de la depuradora con la que se está trabajando en

otro proyecto de investigación que se lleva a cabo en el departamento. Es decir, crear una interfaz gráfica para ver el proceso, crear simulaciones que permitieran variar parámetros, probar diferentes sistemas de control, etc.

Por último, también se podrían desarrollar otros sistemas con EJS destinados a la enseñanza del control para utilizarlos tanto en las clases teóricas como prácticas de las asignaturas ya que, al ser ejemplos interactivos y muy visuales, ayudan a comprender mejor los conceptos y permite su aplicación práctica.

Referencias Bibliográficas

1. The MathWorks: <http://www.mathworks.com>.
2. EJS (Easy Java Simulations) <http://www.um.es/fem/Ejs/>.
3. JIM (Java Internet Matlab) <http://lab.dia.uned.es/rmatlab/>.
4. Dormido, S., Sánchez, J., Morilla, F.: Laboratorios virtuales para la práctica a distancia de la automática. XXI Jornadas de automática, Sevilla (2000).
5. Buccieri, D., Sánchez, J., Dormido, S., Mullhaupt, P., Bonvin, D.: Interactive 3d simulation of flat system: The Spidercrane as a case study. Proceedings of the 44th IEEE Conference on Decision and Control, Sevilla (Diciembre 2005).
6. Dormido, S., Esquembre, F.: The Quadruple-Tank process: an interactive tool for control education. Proc. European Control Conference, Cambridge, UK (2003).
7. automatL@bs: <http://lab.dia.uned.es/automatlab/>.
8. Duro, N., Dormido, R., Vargas, H., Dormido, S., Sánchez, J., Pastor, R.: The Three-Tank system: a remote and virtual control laboratory using easy java simulations. in 44th IEEE CDC/ECC (2005).
9. Vargas, H., Dormido, R., Duro, N., Sánchez, J., Canto, S.D., Farias, G., Dormido, S., Canto, M.a., Esquembre, F.: Heatflow: Un laboratorio basado en web usando Easy Java Simulations y Labview para el entrenamiento en técnicas de automatización. XII Latin-american Congress on automatic Control, CDROM del Congreso, Salvador de Bahía, Brasil (Octubre 2006).
10. Farias, G., Esquembre, F., Sanchez, J., Dormido, S., Vargas, H., Canto, S.D., Dormido, R., Duro, N., Canto, M.a.: Desarrollo de laboratorios virtuales, interactivos y remotos utilizando Easy Java Simulations y modelos Simulink. XII Latin-american Congress on automatic Control, Salvador de Bahía, Brasil (Octubre 2006).
11. Guzmán, J.L., Rodríguez, F., Berenguel, M., Dormido, S.: Laboratorio virtual para la enseñanza de control climático de invernaderos. Revista Iberoamericana de automática e Informática Industrial **2**(2) (abril 2005) 82–92.
12. Dormido, S., Sánchez, J., Esquembre, F., Vargas, H., Dormido-Canto, S., Dormido, R., Duro, N., Farias, G., Canto, M. a.: The development of web-based virtual laboratories using Easy Java Simulations (EJS) Remote Engineering and Virtual Instrumentation Rev07, Oporto, Portugal (Junio 2007).
13. Sánchez, R.: Simulaciones con EJS y Matlab/Simulink. Tesis de Master, Universidad de Salamanca (Diciembre 2008).
14. Dormido, S., Esquembre, F., Farias, G., Sánchez, J.: adding interactivity to existing Simulink models using Easy Java Simulations. Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference 2005 Seville, Spain (Diciembre 2005).
15. Dormido, S., Esquembre, F.: Conectividad de EJS con otras herramientas. Taller de EJS, Desarrollo de simulaciones dinámicas interactivas. (abril 2007).
16. Esquembre, F.: Easy Java Simulations. The Manual (v3.4). (Septiembre 2005).

Análisis Visual de Redes de Co-autoría basado en Mapas Auto-Organizados: aplicación a publicaciones Biomédicas

Mary Carmen Trejo Ávila

Departamento de Informática y automática
Universidad de Salamanca, España
mtrejo@usal.es

Resumen. La visualización de estructuras en el campo científico es un objeto clásico de estudio del área Cienciométrica. Uno de los casos más importantes es la identificación y manipulación de relaciones entre pares de autores determinadas a partir de diversos indicadores. Para ello, métodos tradicionales de Análisis de Redes Sociales complementados con técnicas de Visualización de Grafos son utilizados frecuentemente. Dicha representación no siempre es la mejor solución, ya sea porque muchas veces requiere un tiempo considerable de procesamiento, o simplemente los resultados obtenidos no reflejan con claridad la potencialidad de contribución para la toma de decisiones. Algunos de los factores que afectan a estos resultados son el manejo de grandes cantidades de datos, la calidad de los datos, y en ocasiones la manipulación de datos multidimensionales. En este trabajo se propone un prototipo para construir, representar y analizar redes sociales basadas en relaciones de co-autoría. El conjunto de datos es recuperado de una base de datos digital y gratuita de documentos de literatura Biomédica y de Ciencias de la Salud. Se propone una vista dual la cual permite, por un lado, una fácil identificación de las colaboraciones en comunidades científicas, y por otro, las áreas temáticas de concentración de dichos grupos.

Abstract. The visualization of scientific field structures is a classic object of study in the Scientometric area. One of the most important cases is the retrieval of relationships between author pairs, which is determined by several indicators. Traditional techniques of Social Network Analysis complemented by Graph Visualization are used in most of the cases. Unfortunately this is not always the best solution, mostly because it requires a considerable amount of processing time, or it just doesn't accurately reflect the potential contribution of the obtained results. Some of the factors that affect these results are the handling of great amounts of data, the quality, or even the manipulation of high-dimensional data. In this work we propose a tool to build, represent and analyse social networks based on co-author relationships. The data is gathered from a free digital archive of Biomedical and Life Sciences journal literature. We propose a dual visualization that allows not only an easier recognition of collaborations in the scientific communities but also the identification of concentration areas in each group.

1. Introducción

La tecnología moderna permite la creación de grandes almacenes de datos (crudos) que requieren ser explorados mediante búsqueda de información refinada (conocimiento). Desarrollar agentes que permitan procesar estos grandes volúmenes de datos y convertirlos en conocimiento útil para la toma de decisiones (inteligencia) constituye un reto colosal.

Tradicionalmente, el análisis de datos consiste en la aplicación de métodos matemáticos con la finalidad de obtener información útil para el mejor entendimiento de por ejemplo: un fenómeno natural. En muchas ocasiones es posible conocer de antemano información a priori, que puede ser utilizada en el proceso de análisis de datos. Sin embargo, cuando se cuenta con una gran cantidad de datos y se conoce poco o nada acerca de las estructuras subyacentes en los datos, el problema se ataca de distinta manera; con la finalidad de explorarlos y averiguar: si los datos tienen una estructura particular, si existen datos que no se ajusten al comportamiento de la mayoría, si los datos presentan algún tipo de agrupamiento o si es posible establecer similitudes o diferencias entre dos o más grupos de datos.

Cuando el análisis de datos está dirigido a responder estas cuestiones se conoce como *Análisis Exploratorio de Datos (Exploratory Data Analysis, EDA)* [1]. En la respuesta a estas preguntas, no es suficiente la aplicación de las herramientas analíticas. La presentación de los resultados arrojados por el análisis de datos es de gran importancia, ya que en la mayoría de las ocasiones es necesaria la interpretación de los resultados, por lo que es deseable que éstos sean presentados de manera resumida y ordenada. Con la finalidad de cumplir estos requerimientos se aplican una serie de técnicas que provee el campo de investigación denominado *Visualización de Información (Information Visualization, IV)* [2].

Otra circunstancia de la tecnología actual, es la variedad de paradigmas computacionales que se han desarrollado desde mediados del siglo pasado. Con principios distintos a los utilizados por modelos secuenciales, algunos de estos paradigmas parten de la idea de construir modelos computacionales capaces de adaptarse a distintas situaciones de una manera no predeterminada. En el desarrollo de estos modelos computacionales surge la disciplina denominada *Aprendizaje Automático (Machine Learning, ML)* [3].

Dentro de los distintos paradigmas para el ML destacan las denominadas *Redes Neuronales Artificiales (Artificial Neural Networks, ANN's)* [4]. En este trabajo, la principal herramienta a considerar para el procesamiento y el análisis de datos es el modelo de red neuronal propuesto por Teuvo Kohonen nombrado *Mapas Auto-Organizados (Self-Organizing Maps, SOM)* [5]. El SOM es un algoritmo relativamente nuevo y eficiente para llevar a cabo la visualización de grandes conjuntos de datos multidimensionales.

Es así como el campo llamado *Análisis Inteligente de Datos (Intelligent Data Analysis, IDA)* [6] surge de la combinación de métodos tradicionales para el

análisis de datos, herramientas desarrolladas en el aprendizaje automático y otros métodos más específicos.

Un tipo de aplicación de obtención de estructuras subyacentes en los datos, recae en el área de investigación denominada *Análisis de Redes Sociales (Social Network Analysis, SNA)* [7]. En este caso, las estructuras (redes) son sociales y su metodología atiende dos cuestiones básicas: qué analiza y cómo lo analiza. Para ello centra su estudio en las relaciones (vínculos) existentes entre una serie definida de elementos (individuos, organizaciones). A su vez, se abordan tres tareas principales: 1. identificación de comunidades, 2. identificación de actores principales y 3. análisis de roles y posiciones.

Utilizando las metodologías de SNA, apoyadas por EDA y IV, se pueden analizar diversos tipos de estructuras, como aquellas donde las relaciones o vínculos son científicos entre individuos como lo son los investigadores. En [8] se introduce el problema de visualizar información asociada a la literatura científica, donde se propone la combinación de dos áreas de estudio: gráficos por computadora (IV) con recuperación automática de documentos (*Informetría (Informetric)* y *Recuperación de Información (Information Retrieval, IR)*). La disciplina de especialización dedicada al análisis de ese tipo de estructuras y actividades científicas es llamada *Cienciometría (Scientometric)*.

1.1. Problema

La visualización de redes es un tema muy popular para el campo de IV. En la mayoría de las ocasiones se utiliza la representación de grafo mediante el uso de vértices y aristas. Esta representación se adapta correctamente cuando las redes son dispersas, sin embargo existen aquellas que son globalmente dispersas y localmente densas, tal como lo son las redes sociales; por lo que dicha representación no siempre es la mejor solución. La dificultad es mayor ya que la percepción y representación de la realidad son tareas complicadas, tal como Tufte lo expresa en [9] "... el mundo es complejo, dinámico y multidimensional; el papel es estático y plano..."

La problemática central de este trabajo radica en proponer una visualización de estructuras subyacentes, contenidas en un determinado conjunto de datos, que permita la correcta identificación, manipulación e interpretación de relaciones entre individuos con múltiples características.

1.2. Objetivo

En la actualidad existen diversos y grandes repositorios de documentos científicos. éste es el caso de PubMed, un servicio de la Biblioteca Nacional de Medicina de los Estados Unidos (U.S. National Library of Medicine, NLM) donde se incluyen alrededor de 16 millones de citas de artículos del área Biomédica que datan desde los años 50's. Estas referencias son de MedLine y de otras revistas científicas en el campo de las Ciencias de la Salud. Además provee un tesoro

conocido como **Medical Subject Headings o MeSH**, que es el vocabulario (de más de 33,000 términos) controlado que emplea MedLine y otras bases biomédicas para indizar, catalogar y recuperar la información que se introduce en cada una de ellas. Una extensión de PubMed es GoPubMed [10] [11], un servicio web que permite explorar los resultados de búsquedas en PubMed basadas en la *Ontología Génica (Gene Ontology, GO)*; el cual es un vocabulario controlado y jerárquicamente estructurado para la biología molecular.

Una clase particular de red social es aquella donde los individuos son científicos y las relaciones son colaboraciones científicas. Los resultados recuperados por GoPubMed facilitan la creación de este tipo de redes.

El objetivo de este trabajo es extraer y visualizar de manera oportuna redes sociales de colaboración científica. Primeramente, usando la avanzada metodología de búsqueda semántica que provee GoPubMed se recuperan datos relevantes (autores y términos MeSH). Posteriormente, basados en el modelo del algoritmo SOM se obtiene la distribución de la red social. Finalmente, se aplican diversas técnicas de IV como grafos complementados con zoom semántico y filtros para su manipulación e interpretación.

El algoritmo SOM define una función del espacio de entrada a una red de neuronas en el plano. A su vez, esta función define una proyección del conjunto de datos multidimensionales (invisible, de los que no se conoce a priori ningún tipo de organización) a un espacio visible (normalmente bidimensional). La visualización del conjunto de datos permite que las relaciones de similitud que se presentan entre los datos dentro del espacio multidimensional puedan ser observadas en un despliegue bidimensional denominado *mapa*.

En este caso, los datos de entrada (autores) que comparten características comunes (las colaboraciones y los temas) activan áreas adyacentes en el mapa con topología hexagonal. Distintas tonalidades de color y diversos tamaos se aplican a los elementos en el mapa para representar la estructura de la clasificación. Además, se puede utilizar un *zoom* para explorar partes específicas de la red. En un primer nivel de abstracción se puede construir el grafo correspondiente al *clustering*¹, en segundo la composición de un determinado cluster, y en el último nivel la información contenida en una neurona. Las características (asociadas al mapa) respecto a un determinado autor o un tema pueden ser analizadas mediante la visualización de un mapa de componente. Estadísticas generales de los campos que componen los datos se muestran en un *TreeMap*².

Con esta visualización de vista dual se intenta conseguir una descripción general sobre la colección entera de datos (SOM), ya que identifica clusters en los datos, y al mismo tiempo permite centrarse en nodos particulares (grafo), para analizar datos específicos de la red con mayor detalle. Por lo tanto, la visualización facilita la comprensión de patrones y tendencias contenidas en el

¹ extracción de representantes de un conjunto de datos.

² visualización que permite representar jerarquías de forma que se optimiza el llenado del espacio

conjunto de datos, tales como: identificación de autores que tienden a trabajar en conjunto y autores que tienden a trabajar en áreas de investigación similares.

Este trabajo está compuesto por distintos apartados donde se puntualizan las ideas generales para que en el último de éstos, el cual aborda el caso de estudio, se pueda llevar a cabo una interpretación adecuada de los resultados arrojados por el prototipo. Para un mayor detalle de la perspectiva histórica, nociones básicas, modelos y técnicas de los diferentes campos de investigación revisados para la redacción de este breve documento consúltese [12].

2. Análisis Inteligente de Datos

Un dato representa la medición de distintas características observadas. Al conjunto \mathbb{U} de todas las observaciones se le denomina conjunto universo. En este trabajo, \mathbb{U} es un espacio multidimensional donde cada elemento es representado como un vector de dimensión n ($n > 3$). Así, un conjunto de datos X es un subconjunto de \mathbb{U} . Si $X = \{x_1, x_2, \dots, x_k\}$ el elemento $x_i \in X$ se representará de manera vectorial como: $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$ donde x_{ij} representa el valor de la variable v_j observado en el dato x_i .

El **Análisis Inteligente de Datos** es el estudio interdisciplinar concerniente al análisis de datos efectivo y eficiente, el cual involucra la aplicación iterada de algoritmos matemáticos sobre conjuntos de datos. En IDA las disciplinas que juegan un papel fundamental son: la Estadística y el **Aprendizaje Automático**. En el desarrollo de ML existen distintos paradigmas que surgen o se basan en diversas ramas de las ciencias, uno es el *Conexionista*, también llamadas **Redes Neuronales Artificiales**, que emulan la forma en la cual las neuronas del cerebro procesan información a lo largo de una red de neuronas. Las ANN's pueden ser vistas como un sistema de procesamiento paralelo y distribuido que consiste de un gran número de unidades (neuronas) conectadas entre sí, donde cada una de ellas es un dispositivo simple de cálculo que, a partir de un vector de entrada procedente del exterior o de otras neuronas, proporciona una única respuesta o salida.

Como oposición a la concepción clásica del análisis de datos y por la necesidad de contar con métodos para la exploración de grandes conjuntos multidimensionales, surge la escuela de **Análisis Exploratorio de Datos**. Existen métodos clásicos útiles para ilustrar estructuras o relaciones multivariadas entre los elementos de conjuntos de datos multidimensionales, uno de ellos es la exploración visual de información en espacios complejos, abordada dentro de **Visualización de Información**

“La Visualización de Información puede ser entendida como un proceso asistido por la computadora, en el cual se busca revelar seales de un fenómeno abstracto al transformar datos en formas visuales” (*C. Chen, 2002*) [13]

El proceso para el entendimiento de los datos comienza a partir de un conjunto de números y una pregunta. En [14], Ben Fry, propone las siguientes fases

para revelar una solución a la pregunta correspondiente: *Adquisición (Acquire)*, *Análisis (Parse)*, *Filtro (Filter)*, *Minería de Datos (Mine)*, *Representación (Represent)*, *Refinamiento (Refine)* e *Interacción (Interact)*.

La visualización favorece la adquisición de nuevo conocimiento. Las aptitudes y habilidades humanas que se utilizan para el análisis e interpretación de la información, están directamente relacionadas con la capacidad cognitiva de los individuos, y son consideradas como una más de las operaciones que realiza el cerebro para el procesamiento de información.

La visualización de la información no es el resultado implícito del acto de ver, no es un producto espontáneo del individuo que recibe la información ya visualizada. La visualización es una tarea del comunicador visual, que transforma datos abstractos y fenómenos complejos de la realidad en mensajes visibles, haciendo posible que los individuos vean con sus propios ojos, datos y fenómenos que son directamente incomprensibles, y que por tanto comprendan la información que yace oculta.

Otra de las principales técnicas para EDA es Clustering, que tiene como objetivo reducir la cantidad de datos mediante el agrupamiento de datos con características similares. Se define:

“Los algoritmos de clustering son métodos para dividir un conjunto X de k observaciones en K grupos (clusters) de tal manera que miembros del mismo grupo son más parecidos que miembros de distintos grupos”
(*B.D. Ripley, 1996*) [4]

3. Red Neuronal de Kohonen (SOM)

“Las Redes Neuronales Artificiales son redes de elementos simples interconectadas masivamente en paralelo (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico”
(*T. Kohonen, 1998*) [5]

Kohonen presentó en 1982 su modelo de ANN llamado SOM. El algoritmo de la red SOM está basado en el aprendizaje no supervisado y entrenamiento competitivo, lo cual quiere decir que no se necesita intervención humana durante el mismo y que se necesita saber muy poco sobre las características de la información de entrada. La idea de entrenamiento competitivo consiste en determinar cuál de las neuronas es la que mejor representa a un estímulo de entrada dado. A esta neurona se le considera neurona ganadora y tiene la capacidad de inhibir a las otras neuronas; es decir, el vector de pesos asociado a estas neuronas no serán ajustados de igual forma que el vector de la neurona ganadora. Existen dos etapas básicas del proceso durante el entrenamiento:

1. **Ordenamiento Global.** Durante aproximadamente las primeras 1000 competencias se lleva a cabo el ordenamiento de los datos a lo largo y ancho del

mapa. Este ordenamiento consiste en establecer los pesos de cada neurona para que éstas sean capaces de identificar cierto subconjunto característico dentro del conjunto de datos X y para que las relaciones de cercanía entre las distintas neuronas del mapa reflejen cercanía de los datos correspondientes en el espacio multidimensional del cual provienen.

2. **Refinamiento.** Dado que el aprendizaje es un proceso estocástico, la precisión final del mapa dependerá del número de pasos en esta etapa final de la convergencia.

La representación *U-Matrix (Unified distance Matrix)* de la red SOM visualiza la distancia entre neuronas adyacentes. La misma se calcula y se presenta con diferentes colores entre los nodos adyacentes. Un color oscuro entre neuronas corresponde a una distancia grande, que representa un espacio importante entre los valores de los patrones en el espacio de entrada. Un color claro, significa que las neuronas están cerca unas de otras. Las áreas claras pueden pensarse como *clases* y las oscuras como *separadores*.

4. Análisis de Dominios

El **Análisis de Dominios (Domain-Analysis, D-A)** se basa en el paradigma dominio-analítico que establece que la mejor forma de comprender la información, consiste en estudiar los dominios de conocimiento como parte del discurso de las comunidades de las que proceden. En el campo de **Ciencias de la Información (Information Science, IS)** el D-A puede llevarse a cabo mediante estudios bibliométricos. Enfocados en el campo de IS, las colaboraciones científicas pueden ser perfectamente utilizadas para: detectar los nuevos frentes de investigación científica, obtener información etnográfica referente a la presencia y naturaleza de las relaciones sociales, saber cómo se relaciona un determinado autor con el resto de la comunidad científica en una determinada área de trabajo, etc.

Las redes son, paradójicamente, la forma más antigua y más moderna de organización social de la actividad humana. Para Barnes, una **Red Social (Social Network, SN)** consiste en un conjunto de lazos que vinculan a los miembros del sistema social a través y más allá, de las categorías sociales y los grupos cerrados [15]. Originalmente el medio utilizado para representar las conexiones entre los actores y en qué aplicar las medidas con las cuales obtener patrones de comportamiento, fueron las matrices de coincidencia de datos. La representación gráfica de la información para el análisis de este tipo de patrones fue propuesta por primera vez por Moreno mediante el sociograma:

“... el Sociograma es más que un mero sistema de representación... es un método que hace posible la exploración de hechos sociométricos, donde el emplazamiento propio de cada individuo y sus interrelaciones con otros individuos, puede ser mostrado. Hasta ahora, es el único esquema posible para realizar el análisis estructural de una comunidad” (*J.L. Moreno, 1934*) [16]

La aparición de las SN's, el sociograma y la aparición de la Teoría de Grafos no tenían otro objetivo que el realizar de una forma profunda y adecuada el análisis estructural de las redes sociales.

“El Análisis de Redes Sociales se ocupa del estudio de las relaciones entre una serie definida de elementos (personas, grupos, organizaciones, países e incluso acontecimientos). A diferencia de los análisis tradicionales que explican, por ejemplo, la conducta en función de la clase social y la profesión, el Análisis de Redes Sociales se centra en las relaciones y no en los atributos de los elementos” (*J.L. Molina, 2001*) [17]

La **Cienciometría** es el conjunto de métodos cuantitativos y algoritmos cuyo objetivo es evaluar la calidad científica de un investigador, una revista periódica, un país o región. Los indicadores cienciométricos constituyen la herramienta más utilizada para medir el producto de la investigación científica, el manejo particular de estos indicadores permite determinar cuánto (volumen) producen los investigadores, en qué áreas adelantan su trabajo científico (distribución de temáticas y distribución de autores en cada una de esas áreas). Para fines de este trabajo la unidad de análisis será el artículo científico.

Las **Redes de Co-Autoría (Co-Authorship Networks)** son una clase importante de redes sociales y se han utilizado extensivamente tanto para determinar la estructura de colaboraciones científicas, así como el estado individual de los investigadores. Se utiliza el término co-autoría para hacer referencia a la firma conjunta de un trabajo científico por dos autores. Aunque sea algo similar a las redes de co-citación en la literatura científica [18], la relación de coautor implica un vínculo social mucho más fuerte que la relación de cita.

La IV no es una práctica nueva en el campo de la IS, incluso se le considera como una de las pioneras en estas problemáticas. Sugerida hace más de 60 años y puesta en práctica hace ahora poco más de cuarenta, la IV ha sido utilizada para poner al descubierto y divulgar la esencia y estructura de la ciencia.

5. Caso de Estudio: Red Social de premios Nobel en Medicina de los últimos 10 años

El foco de análisis, son destacados grupos de investigación en el campo de la Biomedicina de la última década, vinculados a los investigadores reconocidos con el premio Nobel en Medicina en el intervalo de años [1999-2008]. Se realizó la implementación de un prototipo, el cual da soporte a cada una de las fases de la metodología, constituido por diversos módulos utilizando distintas tecnologías.

5.1. Recuperación de Datos (acquire)

La estrategia de búsqueda puede llevarse a cabo usando PubMed o GoPubMed, e independientemente de la elección, se preserva el procedimiento de ingreso de la consulta para cada motor. Se obtuvieron un total de 12,965 registros, esta

cantidad se logró de la unión de resultados de búsquedas individuales de parte de cada investigador acreedor al premio Nobel. Los campos recuperados son: afiliación, autores, fecha de creación y de publicación, idioma, número de referencias, páginas, país, pmid, tipo de publicación, resumen, revista, nombre de sustancias, términos MeSH, título y volumen.

5.2. Preprocesamiento de Datos (parse)

Se realizaron una serie de acciones manuales (normalización), con la finalidad de contar con el menor número posible de sesgo en el nombre de los investigadores del conjunto objetivo. Para ello, se emplearon distintas sentencias de Language de Consulta Estructurado (Structured Query Language, SQL)³.

5.3. Procesamiento de Datos (filter)

Se obtuvo la matriz de co-ocurrencia a partir de conteos e índices de actividad entre términos MeSH y autores. Si la co-ocurrencia es elevada significa que hay una proximidad; es decir, que dos elementos tienen relación. En cambio si es baja quiere decir que no hay relación; es decir, que hay distancia. Cada 50 autores relacionados directamente con alguno de los 24 principales investigadores, y cada término MeSH sobresaliente y estrechamente asociado con las áreas de investigación de éstos, son nuestras variables. Por lo que, el conjunto de datos está compuesto por 528 variables (468 autores y 60 términos MeSH), que a su vez son de dimensión 528.

5.4. Minería de Datos (mine)

La ejecución del entrenamiento de la red depende de la determinación de los parámetros básicos del mapa, éstos esencialmente consisten en la determinación de su arquitectura y en la forma de las funciones que serán utilizadas para la actualización de los pesos de los vectores de referencia. La configuración es una retícula cuadrada de 225 neuronas, inicializadas de manera aleatoria, con distancia Euclidiana y vecindad Lineal con un total de 75 iteraciones.

5.5. Visualización (represent, refine, interact)

Con el SOM se visualiza la preservación de la topología y la distribución de los datos en un despliegue ordenado; basándose en ellas es posible el establecimiento de relaciones entre variables, la visualización de clusters y la inspección de relaciones de vecindad entre los nodos en el mapa.

Mapa de Entrenamiento. Revela la evolución del aprendizaje durante el proceso de entrenamiento por cada iteración (presentación del conjunto de datos) sobre el mapa. Se le asocian dos atributos del algoritmo, el primero son las

³ lenguaje declarativo de acceso a bases de datos relacionales que permite especificar diversos tipos de operaciones sobre las mismas.

relaciones de proximidad entre los vectores de referencia por medio del método U-Matrix. Para cada nodo se calculan los promedios de las distancias entre éste y sus vecinos inmediatos en la retícula. A estos valores se le asocia una escala cromática, para posteriormente asignar un color a cada nodo. La gama de colores se obtiene de la mezcla de un determinado color con blanco, relacionando valores decrecientes a tonalidades claras y valores crecientes a tonalidades oscuras. Por lo tanto, un valor pequeño indica cercanía (similitud) mientras que un valor grande indica lejanía (disimilitud). El segundo atributo es la cardinalidad del conjunto de Voronoi asociada a cada nodo; es decir, el número de veces que una neurona fue ganadora para un dato. El conjunto de estos valores se representa con el tamaño de cada nodo, por lo que un tamaño reducido indica que allí se concentran pocas variables mientras que en un tamaño extendido indica un aumento de variables sedimentadas. Otro elemento útil para la interpretación es la asignación de etiquetas sobre los nodos, en este caso un autor o un término MeSH.

Mapa de Clustering. Visualiza una división del mapa en clusters respecto a los vectores de referencia de las neuronas. Se implementó el algoritmo de Clustering denominado *SOM-Ward*. éste combina la información local del ordenamiento en el mapa con el algoritmo *Ward* de Clustering jerárquico, y de esta manera se garantiza que las regiones definidas sobre el mapa resulten conexas. Como en el algoritmo Ward clásico, en el punto de partida del algoritmo *SOM-Ward*, cada nodo es considerado un cluster. En cada paso dos clusters son unidos en uno solo. Estos 2 clusters seleccionados son aquellos que tienen la distancia mínima entre todas las distancias entre clusters, en este caso la distancia se calcula entre los centroides de las neuronas adyacentes (véase la Figura 1).

Mapas de Componentes. Representa el promedio de los valores de cada atributo o componente correspondiente a los datos asociados a cada nodo. La distribución se visualiza por medio de una escala de color, que se asocia al rango de valores que los datos toman en el componente correspondiente. Los valores mínimos están, nuevamente, representados por tonalidad clara y valores máximos por tonalidad oscura. La exploración de los mapas de componentes puede ayudar a establecer relaciones entre las distintas variables (véase la Figura 2).

Grafo de Relaciones. Da soporte al mapa generado por el SOM y/o a la identificación de otro tipo de indicadores. Específicamente, en una red social relacionada a través de co-autorías, se muestran los vínculos explícitos entre los autores. En la Figura 3 destacan 2 grupos, de 6 y 8 autores respectivamente, el resto no presentan relación directa. Evans M es quien tiene el mayor número de publicaciones (tamaño del nodo), y Barré-Sinoussi F junto con Montagnier L tienen el mayor número de artículos (27) en conjunto.

TreeMap. Muestra las estadísticas de los campos recuperados. En la Figura 4 se indican: términos MeSH, autores, países, idiomas, años y revistas. Rápidamente se aprecia que los valores predominantes son: humans, Evans M, Estados Unidos, inglés, 2006 respectivamente. También se observa que la revista "Proceedings of the National Academy of Sciences of the United States of America"

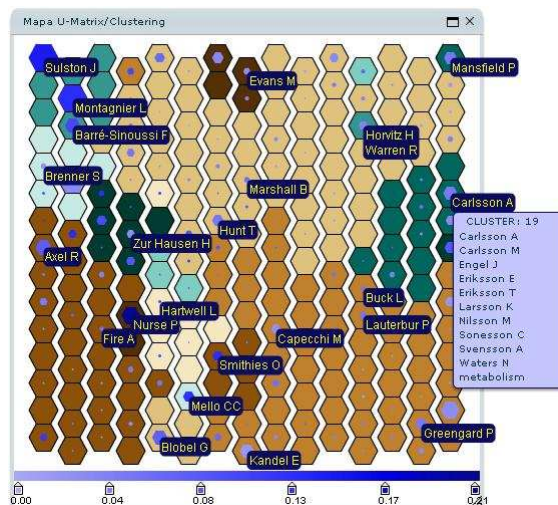


Fig. 1. Sealamiento del cluster 19 (de 24) donde se ubica el investigador Carlsson A, premio Nobel en el 2000, al que se le asocian 724 publicaciones. El resto de los 23 autores se encuentran distribuidos en distintos clusters (gama de colores café-verde), donde un color azul oscuro indica lejanía entre los mismos, así como una región grande azul indica muchas variables (autores o términos MeSH) en dicha neurona.

tiene una ventaja de 226 publicaciones sobre la revista *Nature*, ésta con 321 publicaciones colocada en segundo lugar.

A pesar de que la visualización provista por los mapas puede ser útil para lograr un entendimiento intuitivo, no está claro para que tipo de aplicaciones esta representación es válida. Dado que el SOM trata sobre todo de preservar estructuras locales, la interpretación de los mapas debe hacerse localmente; es decir, basándose en las relaciones entre los vectores de referencia dentro de los conjuntos vecindad de cada nodo. Aunque la interpretación de la estructura global puede ser útil, debe tenerse especial cuidado cuando se inferen propiedades globales del conjunto de datos a partir de la representación bidimensional hecha por el SOM. En general, la calidad de un mapa debe ser evaluada por un experto en el área de aplicación [19]. Existen formas analíticas de medir el desempeño del mapa; entre ellas figuran: error de cuantización y distorsión normalizada.

En cualquier proyección de conjuntos complejos de datos, se produce una pérdida de información respecto a la realidad estructural de los datos originales; el SOM intenta preservar su topología o relaciones de vecindad. La propuesta de realizar una vista dual es proveer un mecanismo dinámico e interactivo que facilite la visualización del SOM hacia el usuario, a su vez permite exhibir datos cuantitativos o cualitativos básicos para este tipo de estudios. Cabe señalar que la elección de la representación en forma de grafo, como vista secundaria, es debido a que este tipo de estructuras semánticas (redes sociales) resultan de mayor familiaridad, para el usuario, asociadas en forma de vértice y arista.

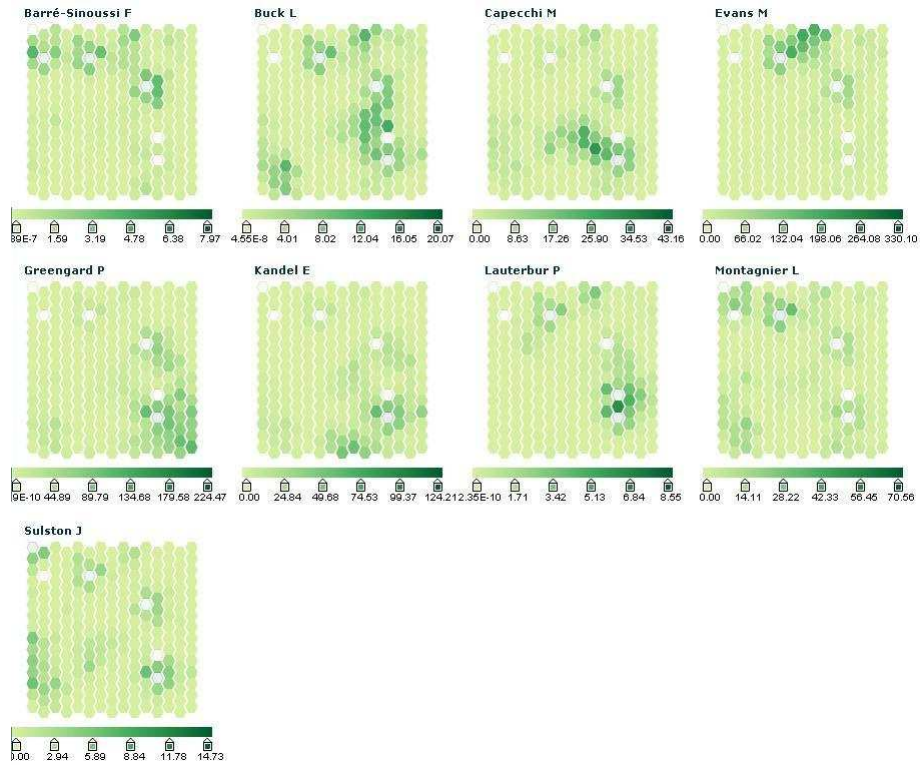


Fig. 2. Comparación de promedios asociados a 9 autores (Barré-Sinoussi F, Buck L, Capecchi M, Evans M, Greengard P, Kandel E, Lauterbur P, Montagnier L, Sulston J) donde los temas contenidos en las neuronas seleccionadas (blancas) son mayormente abarcados, así como los autores con los que más similitud tienen se localizan en las neuronas con verde intenso.

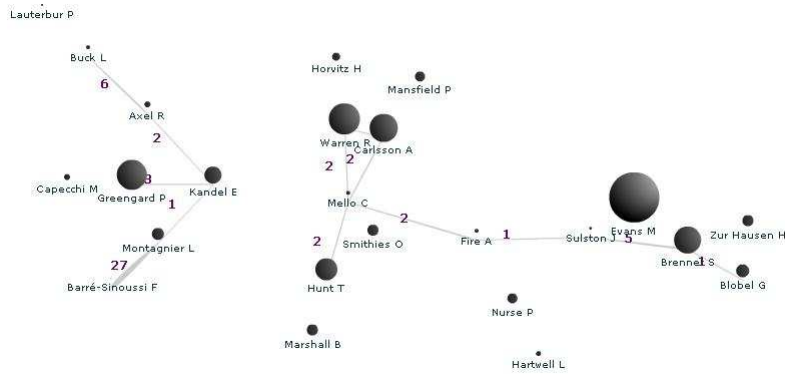


Fig. 3. Red de co-autoría entre los 24 investigadores objetivo

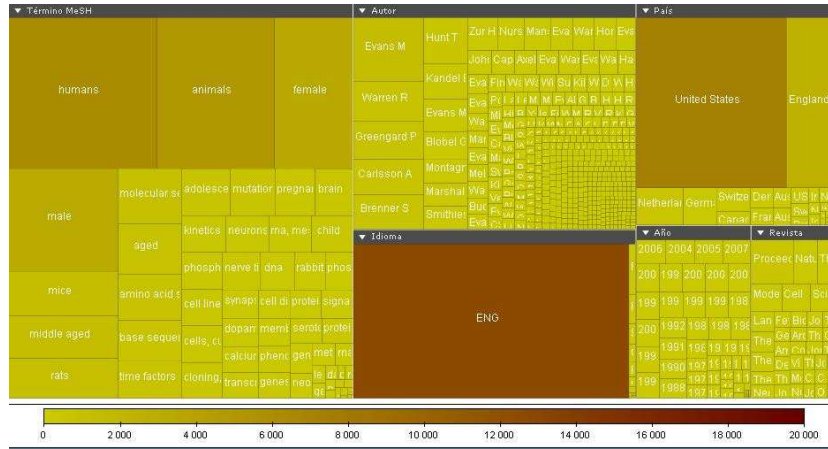


Fig. 4. Estadísticas generales del conjunto de registros

6. Conclusiones y Trabajo Futuro

Las técnicas clásicas de EDA tienen la gran limitante de poseer una alta complejidad computacional y por lo tanto su aplicación en el análisis de grandes conjuntos de datos es poco viable.

La alternativa que se implementó para resolver esta problemática es la utilización de ANN's, en particular el algoritmo SOM. Dadas las propiedades expuestas de este modelo, se puede concluir que efectivamente representa una herramienta de gran utilidad en el análisis de grande conjuntos de datos multidimensionales. Una de las propiedades más útiles del SOM es que brinda una forma de visualizar las relaciones de similitud subyacentes en el conjunto de datos. Además, se observó que el SOM es una alternativa factible para la realización de las tareas de Clustering.

En la propuesta de aplicación al análisis de información científica, a partir de redes de co-autoría, se mostró cómo el SOM puede ser utilizado con fines de investigación académica y como una herramienta para el análisis de la actividad científica y tecnológica.

Para poder llevar a cabo una exploración exitosa no basta con la aplicación del algoritmo, se necesita de una visualización que respalde un eficiente entendimiento de las relaciones en los datos. La animación propuesta para la simulación del aprendizaje de la red neuronal favorecen la comprensión, análisis, e interpretación de la estructura del dominio que se represente, tanto por usuarios no iniciados como por expertos en el uso del SOM.

La combinación de técnicas *Detail + Overview* y SOM no es por definición inadecuada, aunque su grado de adecuación depende directamente de la forma del mapa obtenido y de lo reconocible que resulte su estructura global. Para este caso específico se cumplen los requerimientos de manera positiva y conveniente.

Los mapas hexagonales generados a partir del SOM aprovecha el espacio visual, ya que en su representación gráfica no existen zonas *en blanco*. El grafo despliega características básicas y vitales las cuales se pierden disuelven en el mapa SOM. Con lo anterior se cubre la necesidad de ofrecer una visión global de la representación para orientar al usuario, y por otro una vista para su exploración visual en detalle.

6.1. Recomendación para trabajo a futuro

Un objetivo altamente necesario es concluir en su totalidad la funcionalidad y la correcta integración de las técnicas de visualización empleadas para lograr con totalidad el objetivo planteado. Sería de gran aporte poder descender escalonadamente a niveles inferiores de tipo geográfico, institucional, de autor, e incluso si se desea, llegar hasta el artículo. Y por otro lado, verificar que cada una de las operaciones válidas desde el enfoque de SNA se puedan realizar de manera sencilla y apropiadas.

Una propuesta inicial es la implementación de un algoritmo matemático que indique el número de clusters que mejor represente a un conjunto de datos en grupos.

Otro objetivo inmediato sería integrar la lectura de diversos formatos para los datos de entrada. Debido a la naturaleza de las visualizaciones y algoritmos planteados, dichos métodos son aplicables a una amplia gama de análisis con opción de beneficiarse usuarios con diversos enfoques y en distintos ámbitos.

Cada una de las fases mostradas en el caso de uso puede ser ampliamente estudiada. Por ejemplo, para la etapa de preprocesamiento deseable sería que la intervención manual se redujera para que fuese mínima. Esta es una tarea engorrosa para la cual se necesita emplear demasiado tiempo; áreas de investigación como lo son Web Semántica, Algoritmos Genéticos, y ANN's con Lógica Difusa pueden ser un acercamiento de solución. La etapa de procesamiento de datos puede ser complementada rápidamente con la implementación de distintas funciones matemáticas para normalizar datos. Es importante que los datos estén lo mejor representados, para cada conjunto variará el tipo de transformación necesaria o que mejor se adapte ya sea a los datos objetivo o a los algoritmos por aplicar en posterioridad y de esto depende la obtención de buenos resultados.

Para este prototipo se implementó el algoritmo SOM para la etapa de minería de datos y se observa que tiene buenos resultados. Sin embargo, la asignación de cada parámetro es vital para la obtención de un resultado claro y positivo. El estudio de este tipo de ANN aún se encuentra en apogeo existiendo diversas incertidumbres relacionadas a la misma. Es evidente que un análisis matemático de las propiedades del SOM, requiere la utilización de diversas ramas de las Matemáticas y por lo tanto representa un reto muy interesante de investigación.

Sería oportuno que usuarios dedicados a la elaboración de estudios bibliométricos probaran la herramienta con la finalidad de retroalimentar los indicadores

usados, así como las técnicas de visualización y formas de interacción. Este primer prototipo es tan sólo un bosquejo de una herramienta con amplias posibilidades de extensión y mejoramiento. La integración de un grupo de trabajo de personas tanto del área de visualización de información como de parte del área de ciencias de la información sería idóneo.

Referencias Bibliográficas

1. J.W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1997.
2. M.R. Berthold and D.J. Hand. *Intelligent Data Analysis: An Introduction*. Springer Verlag, 2003.
3. T. Mitchell. *Machine Learning*. McGraw-Hill Education (ISE Editions), 1997.
4. B.D. Ripley. *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.
5. T. Kohonen. *Self-Organizing Maps*. Neurocomputing, 1998.
6. E.R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 1986.
7. J.P. Scott. *Social Network Analysis: A Handbook*. SAGE Publications, 2000.
8. H.D. White and K.W. McCain. Visualizing a discipline: an author co-citation analysis of information science 1972–1995. *J. Am. Soc. Inf. Sci.*, 49:327–355, 1998.
9. E.R. Tufte. *Envisioning Information*. Graphics Press, 1990.
10. R. Delfs and A. Doms and A. Kozlenkov and M. Schroeder. GoPubMed: ontology-based literature search applied to Gene Ontology and PubMed. In *Proceedings of German Bioinformatics Conference*. LNBI., pages 169–178. Springer, 2004.
11. A. Doms and Schroeder M. Gopubmed: exploring pubmed with the gene ontology. *Nucleic Acid Research*, 33(Web Server Issue):W783W786, 2005.
12. M.C. Trejo-Avila. *Anlisis Visual de Redes de Co-Autora basado en Mapas Auto-Organizados: Aplicacin a publicaciones Biomdicas*. 2008.
13. C. Chen. *Information visualization*. *Information Visualization*, 1, 2002.
14. B. Fry. *Visualizing Data*. O'Reilly Media, Inc., 2007.
15. J.A. Barnes. Class and committees in a norwegian island parish. *Human Relations*, (7):3958, 1954.
16. Moreno J.L. *Who shall survive? : a new approach to the problem of Human Inter-relations*. Nervous and Mental Disease Publ., 1934.
17. J.L. Molina. *Anlisis de Redes Sociales. Una Introduccin*. Bellaterra, Ediciones S.A., 2001.
18. E. Garfield. *Citation Indexing: Its Theory and Application in Science, Technology, and Humanities*. John Wiley & Sons, 1979.
19. S. Kaski. Data exploration using self-organizing maps. *Acta Polytechnica Scandinavica, Mathematics, Computing and Management in Engineering Series No. 82*, 1997.

Open, online & collaborative knowledge visualization? Architecture proposal to achieve it

Marcela Villalobos

Departamento de Informática y automática
Facultad de Ciencias - Universidad de Salamanca
Plaza de la Merced s/n 37008 Salamanca, España
{marcela.villalobos}

Abstract In this work we combine the human’s ability to interpret and transmit knowledge through visual images, today’s massive amount of data available, the need of real day to day solutions for researches and developers to share and test their proposals in the visualization field, and the fact that we are moving into a collaborative-information era. We identified the need of a common open solution to be created in this field. a solution that would not only provide the users and researchers with a collaborative environment, but that would also efficiently solve the architectural and logical challenges that such an ambitious work presents. We take previous researchers contributions and experiences, and reinterpret them into a new architectural design based on best software design practices, the well-proven and powerful Entity-Relation data model, current open source technologies available and the existence of several potential partners, willing to share their ideas and artifacts. The transformation of all these into a collaborative information visualization online environment is the main contribution of this work.

1. Introduction

“A picture is worth ten thousand words”. Visualization through visual imagery has been an effective way to communicate both abstract and concrete ideas since the dawn of man, by excellence the way human knowledge is generated and shared. as explained by Spence in [15], visualization is the formation in the mind of the image of an abstract concept. Being a mental phenomenon, vision doesn’t intervene at that moment, although it could have been involved in the formation of the symbols and elements that end up composing the mental image.

The image and hence the visualization, is a mental construction that goes beyond sensorial perception. This mental construction, is directly attached to the concept of knowledge, the intellectual apprehension of things. To understand means to surround, to include something, to interiorise it. We can say that the process of learning is the capacity that we have to create mental images from something real, abstract or invisible, visualizing and contextualizing the idea of that “something” in our minds.

It’s important to point here that as Wurman¹ explains in [20]: “What differentiates knowledge from information is the complexity of the experiences that

¹ Considered a pioneer in the practice of making information understandable

you need to reach it". This means, knowledge cannot be transferred from one person to another, it has to be built by the person him/herself and this represents one of the biggest motivations and objectives of the present work: how to create visual environments that will not only represent stored data, but that will facilitate the creation of new information. It will need to provide the opportunity for the users to have "experiences" with this information leading them to generate new knowledge. Going beyond, we need to provide the means on which this new knowledge can be shared and re-built constantly by a community, creating the so called "collective intelligence".

1.1. Information visualization

As a subject in computer science, data visualization is the use of interactive, sensory representations of abstract data to reinforce cognition, hypothesis building and reasoning. Today most of the data is stored and shared through the use of computers, automating the management and access to thousands of resources specially over the internet. Such a combination of data storage and visualization techniques can be integrated into building a larger process of sensemaking and knowledge creation in today's collaborative global society, potentiated by technological tendencies like the semantic web.

Even though there are much more areas where information visualization contributes (scientific visualizations, software visualization, etc.) we will here focus on displaying quantitative, multivariate information, and there are numerous different types of graphical elements available for displaying quantitative data, such as line graphs, bar graphs, sparklines, etc.(Tuft in [17]).

Interactive exploration of multivariate data sets is challenging both from the user experience and visualization implementation points of view (refer to [11] and [6]). From a user perspective it's challenging, because it is difficult to comprehend patterns in more than three dimensions. From a system design point of view, visualization systems often are a set of graphical and statistical methods, usually selected by the programmer who has a high level idea of how the system will be used, and of what the users goals are. This leaves many researchers uncertain about how to explore their data in an orderly manner, and more important, how their users will require to explore it as well.

The discussion on which visualization techniques are best to apply to different data sets and types has been going on within the visualization community since the beginning. Problem-oriented classifying solutions [19,12], automated generation of graphic representations [5], best mapping techniques based on experience [4] and even the dashboard approach, all of them present with pros and cons to the challenge of choosing the best metaphor to interact with the user.

As well, as the information era keeps developing, data to analyze is not "focused" anymore. On a daily basis we find struggling engineers, statisticians, marketing professionals and lab researchers with the same problem: fighting against the overwhelming wave of multivariate data to obtain a small pattern, hidden secrets from that data sea. Too often, as [1] stresses out, these analyses and software systems built to "help", translate into decision making and analysis

tools that rely more on coincidences and user experience than in the appropriate design for the human tasks they try to support.

1.2. Objectives

We are facing two problems: in one side, how to potentiate the human capacity to abstract data and images and create knowledge and mental models within the boundaries of a software visualization system? and on the other side, how can we create such an environment, overcoming the theoretical and technological issues that have arisen in the past efforts?

This work proposes a flexible approach, in which we will provide the users with an automated environment, where she/he can navigate through this data sea and creating an “open” environment where users can “play” with different options and visualization resources improving their experiences to gain insights that could lead to significant discoveries. To achieve this we designed a robust and scalable architecture design, that will provide the means to bring all these experiences together. Specifically, after an overview of the state of the art, the Snap [7] project system will be used as the base for our proposal.

Our architecture will enable the system to provide the following features:

1. Plug & Play for data: the user will be able to plug in almost any multivariate dataset into the system.
2. Plug & Play for visualizations: the system will provide the mechanism to use different visualization components applied to a selected dataset or subsets.
3. Interaction between components: each component will be a visualization type, and each of them will have a communication channel with other components, allowing them to interact and update based on user interaction.
4. Open free environment: the framework will be an “open” web based system. This will not only facilitate the opportunities for developers to test their visualizations but also for the user to build a community of shared knowledge.

We want this to be a “laboratory” environment for researchers and developers as well, to generate concept proofs or just try new approaches, before getting into the details of building specific visualization applications. as well we will create an initial “proof of concept” exercise that will reinforce our initial theory or point out potential issues in both the technology and architecture proposed. For this case we will implement a component part of our architecture, and will focus it in one of the most common ways for analyzing multivariate data: the parallel coordinates [2].

2. State of the art

The project hypothesis, based on the current literature on information visualization, is that multiple coordinated visualizations will enable users to explore complex multivariate information in a faster and easier way than single, custom-made visualizations. This, combined with the fact that new technology

and resources are being created around the information visualization field (as development frameworks and toolkits) inspired the proposal of exploring the benefits and new features that current tools could offer to a collaborative component based approach.

Multiple view systems are highly challenging to design. While designing and deploying these systems, design decisions are to be made, independently of the visualization area or goals to pursue. Based on Tang's work [16] we identify four of the most common issues to deal with while making design decision for the system to build:

- Data Model: this decision can either simplify the user's experience, as we can provide an abstraction of the data to be analyzed, or in the other side, can really complicate things by creating unnecessary layers or bad abstractions.
- Data access: one of the most complex and time-consuming problems of the visualization systems design, and ironically it belongs to an area that has little relevance to the final objectives of the visualization process, at least from a user point of view. The final goal is that the visualization tool should adapt to the user needs, and make data access not only easy but almost "invisible" to the user, focusing the efforts on what's our end goal: providing insightful visualization tools.
- Meta-data generation: in some cases (and specifically in this work) the access to meta-data and though to additional semantic information of the data set to be analyzed, can enhance not only the performance of the system, but the capabilities that it can offer to the user.
- Modularization: making the right choices in this area is always a tradeoff. However often this choices end on hard coded interactions between the modules chosen, minimizing opportunities for reuse, or limiting the scalability and extensibility of the system. In the case of a more "open" environment, the design decision is a major one that needs to ensure that reuse, scalability, extensibility and interaction mechanisms are provided independently of the visualization technique chosen.

Addressing these issues requires a balance between performance, flexibility and the amount of expertise required from both users and developers. In this work we focus our design proposal to overcome these issues.

2.1. The concept

A similar approach as a collaborative environment was initiated back in 1996 when Shneiderman developed a taxonomy of data types by different domain-specific tasks types [14]. In 2000 they presented the first proposal Snap [7], formalizing the conceptual model of visualization coordination based on the fact that the entity-relation model establishes a direct correspondence between relational data concepts and user-interface concepts, inspired also by previous success story of RRM [3].

Following North's [8] design and theoretical proposal we designed an architecture that provides a solution to the main issue of the original approach: the

automation and encapsulation of multiple diverse data sources. at the same time, we transported the idea back into our current "future", which is an open, diverse and collaborative knowledge sharing environment: open source projects, collaborative code programming, social web environments for knowledge sharing, etc.

North's concept establish a direct correspondence between relational data concepts and user-interface concepts. This correspondence is shown in Table 1:

<i>Relational Data Model</i>	<i>User Interface</i>
Relation	Visualization
Tuple	Item in a visualization
Primary key	Item Id
Join	Coordination

Table 1. Snap Conceptual Model: relational model concepts mapping

Some important details on the implications of this correspondence are:

- A visualization displays a relation (a table or query result) from the database.
- Coordination between two visualizations is based on the join relationship between their relations.
- Each tuple in a database table is depicted as an individual item in the visualization. This requires also the assumption that the table or relation, must have a primary key attribute to uniquely identify the tuple.

This way of modeling the system behaviors is the key to enable a more scalable and extensible solution to the collaborative environment proposal. as extensible and powerful the Entity-Relational model is, as well our system will inherit the strengths (and manage the well-know weaknesses) of the model. This is an important part of how can the software architecture provide such flexibility, and enable the use of independent visualization tools developed by others.

3. A new architecture

After the analyses performed to Snap and it's later Fusion [9] implementation (refer to Section 4 from [18]), we identified four major contributions that will be used as baseline for our proposal:

- Conceptual model: a formal model of visualization coordination based on the relational data model that provides a strong underlying theory and a language for specifying coordinations.
- User interface: a user interface for constructing coordinated-visualization interfaces without programming.
- Empirical evaluation: an evaluation of users ability to construct and operate their own coordinated-visualization interfaces.
- Implementation: an implemented system that is giving us important learnings and experiences.

Probably due to the fact that the project itself was somehow pioneer in the field, several areas for improvement and opportunities were left behind. Table 2 is a summary of the areas for improvement that were identified as part of the research.

<i>Negative design decision</i>	<i>Why?</i>
User oriented	Since Snap places significant design capability in the hands of users, guidelines are needed to help them design appropriate coordinated visualization interfaces for their data
Data sources	Database constraints: issues with integrating different databases technologies, or even other common data sources
Complexity	Visualizations are not “simple” to build. Managing frames, javascript compability concerns, etc. places a roadblock for the system: distribution to users
Independent visualizations or user centered approach	Potentially disparate visualizations can be created

Tabla 2. Snap Conceptual Model areas for improvement

Specially the areas for improvement will take our attention, so that these will be overcome, and in particular we are to give a solution to the second and third issues. Next we present a proposal for an architecture to integrate and inter operate different visualization components; there are two main features that are going to be addressed on this proposal: the first one is to allow different types of data sources and to integrate them in a single data model, the other big feature will be to make the different visualization components to inter operate based on an event-listener model.

3.1. Back-End

The back-end or data-layer of an information visualization system has proven to be one of the most challenging and important features that will determine the usability and effectiveness of a system of this nature.

Today there are many different possibilities and open access for data storage, just to mention some of them we can identify: relational databases in all their different flavors (Oracle, MySQL, Postgres, etc.), hierarchical databases (as JCR or SVN/CVS repositories), spreadsheets, file systems, XML databases and Web Services. almost all of these possibilities might come with their own mechanism for data querying, connectivity and compatibility with other technologies and platforms (operating systems, virtual machines, programming languages).

Our proposal to overcome these issues is be to have a centralized, single data source, which can be accesible from different platforms (operating systems, virtual machines and programming languages) and to have a single query language to access it. Plus, given the fact that it may handle different data sets coming from several data sources (much more options than a relational database) it has

to be largely scalable and well performing on data retrieval operations for high volumes of data. We want to create an “universal” data wrapper that will encapsulate all the data related concerns from the user to allow him/her focus on data visualization and analysis and to provide developers with a robust environment for further visualization development.

In order to allow access to such a diverse range of data stores a single data access point is provided in the form of a search index (powered by apache Lucene), a place where all that different data can be integrated in a scalable way following an Entity-Relationship model and where data can be flexibly accessed through smart fuzzy-logic-enabled queries.

Using a search index implies that an Indexer component (see Figure 1) has to be written for each different type of data source. The Indexer will take the data from the original data source and store it in the global index. Once an Indexer exists for a given data source type it can be reused for any specific instance of the data source.

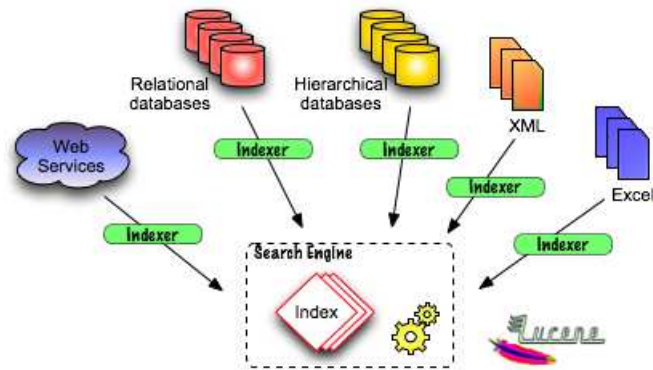


Fig. 1. Data sources exposition

It is the main objective of the Indexer component to provide a generic way of mapping the original data source to an ER model when it's not the case that the data source is already modeled that way.

A service-oriented architecture is provided to encapsulate the access to this centralized data source. Following the current software trends of ubiquity and collaboration, this implies that the data source has to be cross-platform interoperable and web available. This is accomplished by using a combination of well-known software engineering practices and design patterns (see Figure 2).

The component to interact directly with the search index will be a Data access Object (DaO)², it will encapsulate all the technical details of accessing a Lucene search index, thus leaving the architecture in a flexible state enough to

² <http://java.sun.com/blueprints/corej2eepatterns/Patterns/DataaccessObject.html>

change the data source implementation if necessary without affecting the rest of the components involved.

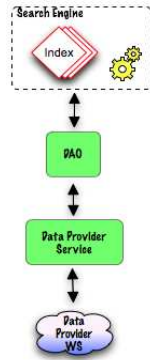


Fig. 2. Data Source Index

different platforms a WebService component will wrap the logic in the Service and will work as a transport layer making the Service available through the Internet.

Finally, the Schema Discover Service is responsible of handling the data schema management, necessary to implement Snap visualization model, based on the relational model approach. This component performs the meta data analysis of the chosen data source, and determines all possibilities for data joins, or in this case, visualization interactions. This component as well implements its corresponding web service.

3.2. Front-End

The front-end design is based on Snap project conceptual design. Being this layer in charge of displaying the visualizations through the visualization components, and of interoperating them.

To increase its potential though, we selected adobe Flex⁵ as the implementation technology. In this case, this decision is specially important even in the design phase, as it will allow us to easily implement the required functionalities but also to complete the open environment and community concept of this project, since Flex is becoming a strong and big community supported tool.

From the front-end point of view, and implementing [7] visualization controller approach, the architecture works based on an Event Manager layer, which handles data sources and visualization collaborative interactions, following the

³ Core J2EE Patterns, Session Facade <http://java.sun.com/blueprints/corej2eepatterns/Patterns/SessionFacade.html>

⁴ Core J2EE Patterns. Transfer Object <http://java.sun.com/blueprints/corej2eepatterns/Patterns/TransferObject.html>

⁵ <http://www.adobe.com/products/flex/>

The DaO will be exposed to external consumers by a Service component using the Session Façade design pattern³. The Service component will provide any logic that goes beyond the simple data access operations provided by the DaO. One of the main contributions of the Service is to serve data using the Data Transfer Object design pattern (DTO)⁴. It will provide data in a way that transcends the technical boundaries of the host language (Java in this case) by converting its data types to XML. This is a major requirement to allow cross-platform interoperability as almost all of the current technology platforms understand XML as a standard language. In order to make the logic available across dif-

entity relation model (applied here to visualizations handling and communication processes) in such a way that visualizations won't even know they are part of such an environment (making it easier to include new and different visualization components as needed -hence empowering tech-savy users to collaborate to the architecture-), and together potentiate the power of visualization techniques.

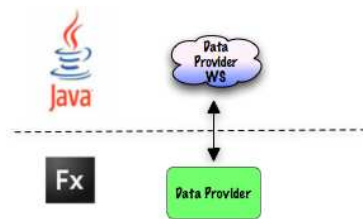


Fig. 3. Java back-end layer inter-operating with Flex front-end layer

The DataProvider Service component will be the consumer of the WebService component defined as part of the back-end and act as an abstract data source for the rest of components on the front-end side, as seen in Figure 3.

The DataProvider component will be the Flex counter-part of the Data Provider Web Service, encapsulating all the WebService-related access details to the rest of the architecture. The data exchanged between Java and Flex will follow the DTO pattern⁶ and will be formatted using XML.

The Visualization Initializer component, which communicates with the back-end through the Schema Discover web Service, will start its work every time a new data source is requested. It will get the results of the meta data analyses (possible joins = possible visualization interactions), initialize the required or available visualization components and register them within the EventManager Component. It will also store information of the possible relations from the data set, (following Snap's proposal of a Visualization schema model [10]).

The communication between the visualization components will be coordinated by the EventManager component. This will follow the Observer design pattern⁷ having the visualization components posting events to the EventManager and this one in turn, broadcasting these events to all the other visualization components who will have a listener implemented for that particular event (see Figure 4).

The EventManager component is in charge of the joins coordination translated as events between visualizations. This implies that each visualization component will have a subcomponent of EventListener that implements listeners for the particular events that are of interest to the requirements of the visualization component. The EventListener component that each visualization should implement simplifies the Snap's aPI approach taken, facilitating the "users distribution" weakness from its design. Handling events while maintaining the required data structures will be handled in the EventManager component.

⁶ Core J2EE Patterns - Transfer Object <http://java.sun.com/blueprints/corej2eepatterns/Patterns/TransferObject.html>

⁷ Observer Pattern definition http://en.wikipedia.org/w/index.php?title=Observer_pattern&oldid=24969593

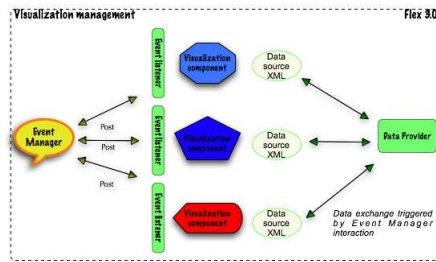


Fig. 4. Event Manager component proposed architecture

3.3. The whole picture

Next we describe a typical flow of conversation in the architecture. Before starting to visualize a data source the user will have to provide his/her data set to the back-end, this could be achieved either by providing the user with a web application where the data can be uploaded, connect to a web service or accessing a database server directly. Once the data set is provided it will get indexed by the corresponding Indexer component according to the dataset's container technology (meaning whether the data was provided using a spreadsheet, a data base, xml, etc.). The process of indexing will only happen the first time a data set is provided, subsequent uses of that particular data set can be achieved without reindexing.

Once the data is in place, the user can start working on the visualization interface and specify which data set is going to be used. This will result on an invocation to the Visualization Initializer component, who will access the Schema Discover Service using the corresponding web service, the Schema Discover Service will query the Data Provider Service to obtain the data set's metadata and determine the possible join relationships and coordinations between visualizations, reusing the design and algorithm proposed by the Snap project ([10]).

For this, the Data Provider will access the index using the DaO component, performing a query to obtain the metadata of the data set. The Visualization Initializer will init the visualization components and store the possible joins or coordinations between visualizations suggested by the Schema Discover Service. It will also start the Event Manager and register the Visualizations' listeners on it.

All requests from the visualization components to the Data Provider will access the Data Provider Service in the back-end ending up on requests to the underlying index through the DaO component.

When the user starts to interact with a visualization component it may start posting events to the Event Manager which will provoke the manager to broadcast the event to the other visualization components suggesting them changes on their behaviour (according to how they are programmed to deal with that particular event); this changes may imply reloading data, zooming, filtering, etc. This

will provide the user with the experience of having integrated and interoperable visualization components.

The unified diagram of the architecture components and their relationships with one another is shown in Figure 5.

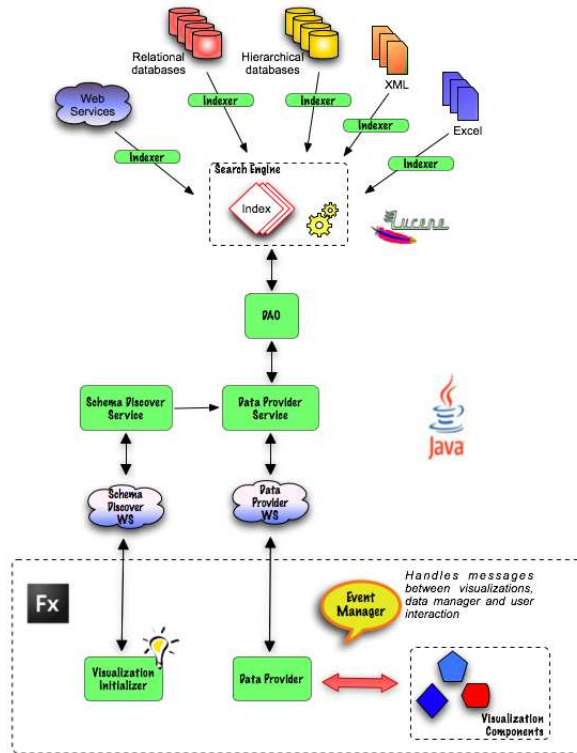


Fig. 5. architecture proposal

4. First proof of concept

To serve as a proof of concept that Flex was one of the best implementation technologies options to bring this work to reality, we designed and implemented one visualization component, that would fit within our architecture proposal. also, an “Integration layer” was built as well to simulate the external data load that the components from our architecture need to implement. In this case, a web service interfaces with the component data manager to load data into the visualization (for details on implementation and final conclusions please refer to [18]).

Due to the fact that parallel coordinates are both: a common way of visualizing multivariate data but a not that popular open source visualization component, we chose this implementation to be our initial practical proof of

concept for Flex and actionScript 3 features to easy this type of development. as well, our implementation includes another visualization technique that intends to enhance the power of the parallel coordinate analysis. We included a Heatmap simple component embedded into the coordinates plot.

Figure 6 shows the parallel coordinates component, including the visualization of a HeatMap⁸. For more details refer to [18].

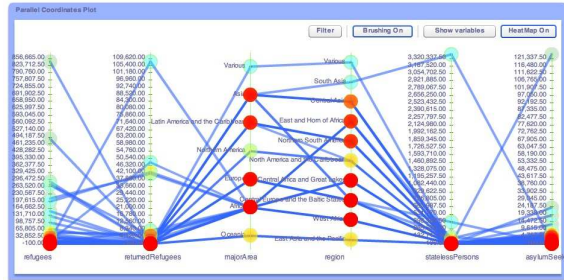


Fig. 6. Parallel Coordinates with Heatmap functionality turned on, presenting data occlusion issues.

The results of this implementation where successful, not only did we verify the fact that Flex is a good option for the creation of visual software pieces, but also it's own architecture and philosophy supports our goal of creating individual component-oriented information visualization pieces. as a first experience and experiment with Flex technology and the design proof of concept we believe we achieved our main goals of getting the required knowledge on both the technology and design to be used in the final implementation of our work.

5. Conclusions

This work started as an individual idea of how to provide an easier access to the information visualization field to common users. after starting the state of the art overview, we found the idea of this collaborative component based environment was not only shared among an important part of the information visualization community, but that previous implementation work, proofs of concept and mostly theoretical researches and proposals were being discussed. However these were never completely implemented, or if so, where not published and shared effectively beyond the scientific community.

From the Snap work and from the review of other authors previous experiences, we detailed and enhanced our conceptual idea, incorporating the positive conceptual aspects(as the Entity Relation model metaphor, visualization schemas and collaborative coordination algorithms) into our initial design. as well

⁸ Graphical representation of data where the values taken by a variable in a two-dimensional map are represented as colors http://www.askdavetaylor.com/whats_a_google_web_page_heat_map.html.

we identified the weaknesses and aspects to be improved from these experiences, and proposed solutions to overcome them.

With the new technology available (both enabled for building visualizations and handling data access), good architecture practices and a simple, clean design, this architecture overcomes the issues that previous researchers encounter in the past 10 years. a collaborative multi-view visualization environment may help users, researchers, designers, and developers in rapidly obtaining results and making decisions from diverse data, transforming it into information. This research and proposal signifies a first step made towards applied information visualization into the “real world” daily problems our users face. It intends to bring the visualizations more close to the user, giving him/her the power and the simplicity to create new knowledge.

6. Future Work

Based both on the analyses of some of the areas for improvement of the tools and concepts utilized to build our proposal, and on the previous experience of the Snap creators and other researches reviewed as part of this work, we set a challenging area for future work to be done. To start, this is a conceptual, paper based proposal. For this to become as powerful as we want it to be, we need to take this ideas and put them in reality, implementing the architecture is obviously our next step. This also implies reviewing and describing all the details and nuances of the design that our proposal requires to become a reality.

In this version of the proposal we are not getting into the details of the “user-profile” components of the design. Meaning the implementation details and structures to support the user preferences to be saved, as well as to be able to save his/her visualization schemas and sub data sets created. For this implementation to be successful, a separate design of the front-end user interface should be carefully evaluated and defined. The user interface design needs not only to comply with information visualization techniques but as well needs to be light and well integrated to maintain the performance and easy-to-use goals that our work included as main objectives.

Two of the more challenging features that would be included in the final implementation include data mining algorithms integrated into the Schema Discover Service or as an independent component of our back-end. as well as the implementation (inspired in Seo and Shneiderman’s work [13]) of a “recommendation” feature, where the system will automate the selection of the best visualization types or components, based on the analysis of the data to be visualized.

This work generated the motivation and investigation field for a much more major, profound and interesting research and work to be done.

Referencias Bibliográficas

1. Robert A. Amar. Knowledge precepts for design and evaluation of information visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11:432–442, 2005.

2. Alfred Inselberg and Bernard Dimsdale. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *VIS '90: Proceedings of the 1st conference on Visualization '90*, pages 361–378, Los Alamitos, Ca, USA, 1990. IEEE Computer Society Press.
3. Tomás Isakowitz, Edward A. Stohr, and P. Balasubramanian. Rmm: a methodology for structured hypermedia design. *Commun. ACM*, 38:34–44, 1995.
4. Andrea Lau and Andrew Vande Moere. Towards a model of information aesthetics in information visualization. *Proceedings of the 11th International Conference Information Visualization*, pages 87–92, 2007.
5. Jock Mackinlay. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics (TOG)*, pages 110 – 141, 1986.
6. A.I. McLeod and S.P. Provost. Multivariate data visualization. In *Encyclopedia of Environmetrics*, pages 1333–1344. Wiley, 2001.
7. Chris North and B. Shneiderman. Snap-together visualization: can users construct and operate coordinated visualizations? *International Journal of Human-Computer Studies*, 53:715–739, 2000.
8. Chris North, Nathan Conklin, Kiran Indukuri, and Varun Saini. Visualization schemas and a web-based architecture for custom multiple-view visualization of multiple-table databases. In *Information Visualization, Palgrave-Macmillan*, pages 211–228, 2002.
9. Chris North, Nathan Conklin, Kiran Indukuri, Varun Saini, and Qiang Yu. Fusion: Interactive coordination of diverse data, visualizations, and mining algorithms. *ACM CHI*, 2003.
10. Varun Saini. *Visualization schemas a user interface extending relational data schemas for flexible, multiple-view visualization of diverse databases*. PhD thesis, Blacksburg, Va., 2003.
11. Ken Brodlie Selan dos Santos. Gaining understanding of multivariate and multidimensional data through visualization. *Computers and Graphics*, 28:311–325, 2004.
12. Hikmet Senay and Eve Ignatius. A knowledge-based system for visualization design. *Computer Graphics and Applications, IEEE*, 14:36–47, 1994.
13. Jinwook Seo and Ben Shneiderman. A rank-by-feature framework for interactive exploration of multidimensional data. *Information Visualization*, 4:96–113, 2005.
14. B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343, 1996.
15. Robert Spence. *Information Visualization: Design for Interaction (2nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2007.
16. Diane Tang. Design choices when architecting visualizations. In *in 'Proceedings of InfoVis 2003, IEEE Symposium on Information Visualization', IEEE Computer Society*, pages 41–48, 2003.
17. Edward R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, second edition edition, 2001.
18. Marcela Villalobos. Proposal: Multivariate collaborative visualization environment, 2008.
19. Stephen Wehrend and Clayton Lewis. A problem-oriented classification of visualization techniques. *Visualization, 1990. Visualization '90., Proceedings of the First IEEE Conference on*, pages 139–143, 469, 1990.
20. Richard S. Wurman, David Sume, and Loring Leifer. *Information Anxiety 2*. Que, 2000.