



HACÍA LA INDIZACIÓN AUTOMÁTICA DE DOCUMENTOS AUDIOVISUALES TELEVISIVOS

Jorge Caldera-Serrano

Doutor en Documentación por la Universidad de Salamanca, España.

Profesor de la Universidad de Extremadura, España.

E-mail: jcalser@alcazaba.unex.es

Resumen

Se especifican y analizan las principales herramientas con las que se cuenta en la actualidad para llevar a cabo indización automática de información audiovisual especialmente válido para las cadenas de televisión. Para ello, se describe la viabilidad de la indización textual sobre la descripción de imagen en movimiento realizada desde los servicios de documentación, así como se muestra la potencialidad que plantean el reconocimiento de voz y el reconocimiento de imagen en el ámbito de la indización automática.

Palabras claves: Indización automática. Información audiovisual. Reconocimiento de audio. Reconocimiento de vídeo. Televisión. Sistema de Información Multimedia.

TOWARDS THE AUTOMATIC INDEXING OF AUDIOVISUAL TELEVISION INFORMATION

Abstract

You specify and analyze the main tools which now has to perform automatic indexing visual information especially true for television. This describes the feasibility of indexing textual description on motion picture made from documentation services and shown the potential to raise the voice recognition and image recognition in the area of automatic indexing.

Key words:

Keywords: *Automatic indexing. Audiovisual information. Audio recognition. Video Attribution. TV. Media Information System*

1 OBJETIVO Y ESTRUCTURA

El carácter multidisciplinar y transversal de nuestra disciplina hace necesario contar con diferentes áreas del conocimiento para poder realizar la labor investigadora. Ciencias de la Comunicación, la Teoría de Sistemas, Ciencias de la Computación y Ciencias de la Documentación e Información son elementos a tener en cuenta para desarrollar los procesos de indización automática en documentos creados y procesados en los medios de comunicación televisivos.

El objetivo del trabajo es determinar posibles vías para indizar de forma automática documentos de carácter informativo para los medios televisivos, especialmente para información audiovisual, identificando problemas y posibles líneas de solución.

Para ello, se llevará a cabo un breve análisis de las estructuras discursivas de los elementos analizados y se estudiará la indización automática tradicional. El carácter acumulativo de la ciencia nos obliga a plantearnos formas operativas de trabajo, valorando como válidos trabajos anteriores y experiencias demostradas, por lo que el tratar de profundizar sobre indización automática sería un error.

La investigación desarrollada desea facilitar los métodos y modos de indización sobre documentos especialmente delicados como son los audiovisuales, sobre los cuales aún no se han probado las herramientas de indización, aportando para ello la visión clásica de la indización sobre información textual pero también aportando elementos relacionados con el reconocimiento de imagen y de voz que pueden potenciar, implementándolo de forma conjunta, la indización y la posterior recuperación de imagen.

2 LAS UNIDADES SEMÁNTICAS DE DOCUMENTOS AUDIOVISUALES EN LOS MASS MEDIA

El estudio, valoración y análisis de la documentación audiovisual ha sufrido un impulso en el ámbito de las *Ciencias de la Documentación* motivado por el aumento de los tipos documentales y la cantidad de documentos generados de dicha naturaleza. Este desarrollo viene avalado por el aumento en el número de publicaciones que en nuestra disciplina se realizan sobre este tipo documental, sobre todo por un grupo de investigadores que centran sus trabajos sobre documentación en medios de comunicación y documentación cinematográfica desde los puntos de vista más dispares (gestión, bases de datos, organización, lenguajes documentales, conceptualización, etc.) (CALDERA-SERRANO, 2005)

Esta parcela del conocimiento, integrado de pleno derecho en las *Ciencias de la Documentación* y con intereses transversales y horizontales con las *Ciencias de la Comunicación*, absorbe conocimientos de ambas disciplinas para desarrollar técnicas, herramientas e implementación conceptual para una cuestión de muy reciente calado en la comunidad investigadora y/o universitaria. Ello se traduce no sólo en la escasez de trabajos sino en la adaptación de teorías y de trabajos anteriores a una nueva realidad, especialmente a la de los Sistemas de Información Documental en las televisiones (a partir de ahora SID-TV).

No se pretende hacer directamente la extrapolación de las unidades semánticas de los textos escritos a los documentos audiovisuales sino que cuenta con la finalidad de identificar y clarificar las posibles unidades de tratamiento documental, cuya determinación para documentos audiovisuales televisivos parece complicarse en caso de compararlo con los textos escritos.

2.1 Unidades semánticas del discurso

Para el estudio de las unidades semánticas tomamos como punto de inicio el trabajo de Moreiro (2004) que analiza las unidades semánticas discursiva textuales para así determinar su posterior indización automática.

Las unidades semánticas determinadas por Moreiro (2004) son las microestructuras (estructura superficial), macroestructura como estructura semántica parcial y la superestructura como disposición de las partes del discurso (estructuras intermedias); y la macroestructura global con elementos sintácticos, semánticos y pragmáticos. Por lo tanto, las macroestructuras organizan jerárquicamente el texto de tal manera que en él existe una información contenida en el discurso o en parte del mismo, definiendo el significado del discurso, el tema y materias. A su vez puede estar integrado en una macroestructura global que ofrece el sentido al discurso en su totalidad, y que es desde esta macroestructura global desde la que debe entenderse el texto, mientras que las microestructuras serían los elementos que ofrecen significado local.

Por lo tanto, se observan estructuras de significación específicas y globales, siendo las primeras las que dan sentido a las siguientes y el conjunto de significados parciales dan sentido global al discurso, todo ello estructurado por la superestructura que organiza y presenta el texto formalmente.

2.2 Estructura de la documentación audiovisual en televisión

Las superestructuras de los documentos científicos presentados a revistas cuentan con una estructura conocida por todos y estudiados entre otros por Drop (1987); los documentos audiovisuales también cuenta con una estructura fija en su formato de informativos que ayuda su conservación y también su posterior recuperación, aunque no asegure la recuperación documental.

La superestructura de un noticiario estaría conformado por:

- La cabecera. Se desarrolla un avance de las noticias destacadas. Suele incluir la locución en estudio acompañado de las imágenes, pudiendo contar con enviados especiales o corresponsales desplazados al lugar de los hechos. No suele contar con más de tres noticias.
- El sumario. Se especifican las noticias que se desarrollan en el informativo, por lo que es una pieza de escasa duración en la cual se observan, junto a un rótulo explicativo, las noticias del informativo.
- Noticias por bloques. Las noticias suelen agruparse por bloques temáticos, aunque las noticias más importantes del día suelen ubicarse al comienzo del noticiario. El orden habitual suele ser el siguiente: sección *nacional*, *internacional*, *economía*, *deportes* y en último lugar *cultura y sociedad*.
- Cierre. Último bloque que suele repetir las noticias de la cabecera y las últimas informaciones llegadas a la redacción.

La superestructura de una noticia se estructura de la siguiente forma:

- Locución en estudio. El periodista en estudio (plano medio) introduce al telespectador en la noticia, realizando a veces un resumen del contenido.
- Entradilla. En la crónica o reportaje se lleva a cabo una introducción realizada por el periodista (corresponsal/desplazado/enviado especial) en el cual vuelve a darse las pautas sobre la noticia o se contesta a elementos planteados por el periodista en estudio (en el caso de los directos)
- Recursos/Locución. El cuerpo de la noticia está integrado por declaraciones e imágenes que ayudan a entender la noticia, junto con la locución del periodista que nos facilita la información.
- Postproducción y/o grafismo. Utilización de elementos infográficos que ayudan al entendimiento de la información: gráficos, mapas, rótulos explicativos, etc.
- Cierre locución. Resumen general de la información que suele estar integrada por otra nueva valoración del periodista desplazado. No siempre se realiza esta labor en los noticiarios.

En los SID-TV el material emitido no es sino un reflejo de horas y horas de material que deben ser igualmente analizados como los brutos, material de agencia, corresponsalías, centros territoriales, free lance, otras cadenas, entre otras, por lo que es complicado la extrapolación de la arquitectura textual.

El discurso televisivo es complejo, derivado de que toda la parrilla de emisión de la cadena conforma una macroestructura global que define la ideología que el medio desea ofrecer. Los informativos son programa en los que se difunden dichos discursos, que a su vez contarán con otros discursos menores que son las noticias. Son, desde las macroestructura, desde las que se puede entender una información, aunque éstas a su vez estén conformadas por estructuras inferiores.

Evidenciar que en un SID-TV no sólo se analizan y conservan el material que ha sido emitido, sino que se analiza (para su potencial conservación) el material con el que ha sido confeccionado. Todas y cada una de las fuentes han de ser analizadas contando con criterios tanto periodísticos como documentales, seleccionando las imágenes e informaciones contenidas en dichos documentos, los cuales cuenta con una superestructura propia, diferente al de la noticia del informativo, y a macroestructuras globales y parciales (así como microestructuras) que son partes integrantes de la noticia que posteriormente se emite.

En este caso no sólo deben analizar las macroestructuras globales y parciales para así poder identificar temas y conceptos pudiendo realizar el análisis documental, sino que es necesario la identificación de elementos menos significativos del discurso como son las microestructuras que son en si misma parte que pueden ser solicitadas.

Partiendo de la simplificación discursiva en la que se traduce sólo el análisis del material emitido, es notoria la necesidad de analizar documentalmente tanto las microestructuras como las macroestructuras de los documentos audiovisuales.

Cuadro 1 - Registro documental de documento audiovisual televisivo

Registro documental de documento audiovisual televisivo		
Título	Cumbre de Bill Clinton y Boris Yeltsin con motivo de la celebración del 50 aniversario de la ONU. Ataque de risa de Clinton en la conferencia de prensa conjunta	
Análisis cronológico		
21:20:03	Entradilla Corresponsal en Estados Unidos	
21:20:15	Clinton y Yeltsin sentados en exterior de antigua casa del presidente Roosevelt en Manhattan dialogando	
21:20:23	P.G. Ambos se levantan y caminan hacia el interior de la casa, rodeados de escolta	
21:20:40	P.A. Clinton y Yeltsin posando en interior de mansión	
21:20:51	Estudio	
21:20:56	R.P. de Yeltsin y Clinton. Carcajadas de Clinton por declaraciones de Yeltsin sobre los periodistas	
21:21:32	Fin	
Puntos de acceso		
	Visionado	Referenciado
Onomásticos	Clinton, Hill Yeltsin, Boris	Roosevelt
Temáticos	Dirigentes políticos Anécdotas personales Actos oficiales – cumbre	Periodismo – Periodista Política internacional

Fonte: Autoria propria

Este registro documental no es completo sino que reseñamos los más relevantes desde el punto de vista de acceso al contenido. En el campo *título documental* se identifica la macroestructura del documento. En análisis cronológico se realiza la descripción detallada de las secuencias o conjuntos de secuencias que conforman el clip de vídeo (nivel de microestructuras discursivas). También se indican los puntos de acceso temáticos y onomásticos (nivel de macroestructuras globales y parciales), realizando la extracción de los puntos de acceso de información visualizada y referenciada.

3 INDIZACIÓN AUTOMÁTICA

El aumento de la cantidad de la información disponible es una realidad que a nadie escapa, lo que plantea una serie de retos derivados de su conservación y más importante de su recuperación. Los motores de búsqueda cuentan cada vez con mayor relevancia ya que son estas herramientas las que nos facilitan el acceso a la información existente tanto en redes

como en Sistema de Gestión de Bases de Datos. La gestión del conocimiento está destinada a utilizar los datos semánticos subyacentes en una colección de documentos para ayudar a comprender y utilizarla.

Importante para el desarrollo de la Organización del Conocimiento además de una apertura mental derivada del crecimiento exponencial de la información (y no sólo en Internet) son los desarrollos tecnológicos en Inteligencia Artificial. El conjunto de tecnologías del lenguaje humano facilita la organización del conocimiento, mejorando la recuperación y la extracción automática de resúmenes, presentaciones e identificación del contenido temático.

Las principales aplicaciones de la minería de textos son la extracción de información de un documento; agregar y comparar información automáticamente; clasificar y organizar documentos; clasificar textos e indizarlos.

Las principales ventajas de la tecnología de minería frente a la tradicional actividad documental son la capacidad de procesar rápidamente grandes cantidades de datos textuales, lo que sería muy complicado por lectores humanos; objetividad en el proceso de extracción; y dejar las tareas más exigentes para lectores humanos.

3.1 Los *keyphrases* y palabras claves

La extracción de palabras claves suele ser habitual en una variada tipología documental, siendo especialmente útil, sistematizado y estandarizado en los documentos científicos. Estas palabras claves y *keyphrases* son incluidos normalmente por el autor y suelen proporcionar una información sumaria del documento, ofreciendo una descripción breve y exacta de su contenido. Cuenta también con la opción de clasificar documentos agrupándolos por temas facilitando así la elaboración automática de tesauros.

Las *Keyphrases* proporcionan un breve resumen del contenido de un documento, lo que sin duda aporta un valor añadido cuando se trabaja con grandes colecciones digitales de documentos, al ser muy complicado realizar esta labor de forma eficiente y rápida. Por medio de estas palabras claves se puede interpretar el contenido del documento de forma individual pudiendo ser utilizadas como elemento posterior de recuperación.

Son variados los métodos que se han propuesto para generar o extraer la información sumaria del texto. En el dominio específico de *keyphrases*, hay dos acercamientos fundamentales: desde la asignación del *keyphrase* y extracción del *keyphrase*. Ambos utilizan métodos de aprendizaje de la máquina por medio del entrenamiento. La asignación de *Keyphrase* intenta seleccionar las frases de un vocabulario controlado que describen lo mejor posible un documento. Los datos del entrenamiento asocian un sistema de documentos a cada frase en el vocabulario, y construyen un clasificador para cada frase. Un nuevo documento es procesado por cada clasificador, y se asigna el *keyphrase* de cualquier modelo que lo clasifique positivamente (DUMAIS, 1998). La extracción de *Keyphrase* no utiliza un vocabulario controlado sino que por el contrario elige los *keyphrases* del texto mismo.

En la asignación de *keyphrases* (también denominada clasificación del texto) un análisis de un documento conduce a la selección de los *keyphrases* para ese documento de un vocabulario controlado. Tiene dos ventajas principales: el vocabulario controlado se asegura de que los documentos similares estén clasificados constantemente, y los documentos se pueden asociar a los conceptos que no se mencionan explícitamente en su texto.

4 INDIZANDO DOCUMENTOS AUDIOVISUALES SOBRE TEXTO

La realización de la extracción de contenido en el ámbito audiovisual es una labor ardua y delicada de las tareas encomendadas a los gestores de la información, de ahí los intentos de buscar nuevas vías y el hallazgo de nuevos métodos para ir automatizando total o

parcialmente estas operaciones.

La indización automática se basa en la idea de poder de resolución de un término, método basados en frecuencia de aparición, lo cual lo habilita como término de indización. La ya conocida Ley de Zipf y fundamenta que ordenadas las palabras de un texto por su frecuencia de uso, el producto de su frecuencia de uso por su posición en el ordenamiento es constante.

Sobre este material se podrá hacer una indización integral ya que por medio de los métodos estadísticos demostrados como válidos se podrá extraer los términos más representativos desde el punto de vista temático.

La indización se realizará sobre los campos de texto libre aportados tradicionalmente por el documentalista: resumen, título y descripción de imágenes.

Una vez realizado catas de pruebas con colecciones documentales de documentación audiovisual de programas informativos observamos el principal inconveniente: el número de términos es escaso, por lo que la frecuencia de aparición de los términos no siempre es adecuada (la campana de repetición de términos es muy plana).

Por lo tanto es necesario contar con elementos que potencien y aumenten el contenido textual de estos documentos, y para ello nos vamos a referir rápidamente a la Identificación Biométrica para posteriormente señalar la potencialidad que nos ofrece.

4.1 Identificación Biométrica

La Identificación Biométrica consiste en verificar la identidad de seres humanos, aunque puede servir igualmente para identificar otros parámetros distintos. La biometría (aplicación de métodos matemáticos y estadísticos a la investigación biológica) ofrece una identificación de personas por como son, no de elementos externos que lo identifican (tarjetas, credenciales, etc.). Estas técnicas proveen al cuerpo de una serie de elementos que los individualizan y distinguen del resto de individuos. Las técnicas biométricas se basan en medir al usuario para reconocerlo automáticamente aplicando técnicas estadísticas y de Inteligencia Artificial (lógica borrosa, redes neuronales, etc.) (CALDERA, 2009).

4.1.1 Reconocimiento de audio

Uno de los tipos de identificación biométrica es el Automatic Speech Recognition (ASR). Estos sistemas permiten el acceso a la información por medio de una forma tan natural y habitual como el habla, siendo especialmente relevante las posibilidades en la telefonía móvil, la computación y la televisión (HATZIS et. al., 2003).

Los sistemas ASR van encaminados tanto a identificar la voz de manera automática como para pasar a texto aquello que se escucha (dictadores).

Dentro de la tecnología ASR se pueden distinguir tres importantes campos de aplicación en los que el uso de reconocedores de voz, contando actualmente con diferentes productos comerciales que proporcionan unas prestaciones bastante aceptables. Tales sistemas son (ESPINOSA, 1998): dictado automático, Interactive Voice Response (comandos de control) y los sistemas dedicados a la segmentación y etiquetado automático de voz (órdenes sonoras por comandos).

La banda sonora es muy importante en el material de informativos, ya que en esta banda cuenta con el principal aporte informativo derivado del periodismo declarativo que se hace en estos momentos. Al igual que la locución (voz en off) aporta información relevante para la comprensión de la información.

Por medio del reconocimiento de la voz se logrará pasar a texto la locución y las declaraciones, aumentando así el número de términos para una correcta indización automática.

Otra opción que aporta los sistemas de detección de voz es la opción de poder

reconocer a los personajes que se encuentran en un banco de voces. Si estimamos que lo habitual es que en información exista una clara redundancia de los personajes involucrados en la información, esto se traduce en la identificación rápida de los personajes habituales. Se podrá por tanto indizar a los personajes escuchados -y por tanto visualizados-.

4.1.2 Reconocimiento de vídeo

El principal desarrollo para el reconocimiento de vídeo viene dado por el reconocimiento facial estudiado y desarrollado a partir de estudios biométricos. La disciplina biométrica utiliza información biológica para verificar las identidades. La idea básica que fundamenta la biométrica es que nuestros cuerpos contienen propiedades e información única que pueden ser utilizadas para distinguirnos e individualizarlos del resto de individuos del colectivo (CALDERA; ZAPICO, 2009).

Los puntos nodales de nuestro rostro nos individualizan. Estos software mapean nuestra cara e identificando la distancia entre los puntos nodales descritos como válidos (puede llegarse a contabilizar más de 80 sólo para el rostro). Las distancias entre estos puntos nodales confeccionan nuestro faceprint que será integrado en un banco de imágenes, con la cual se comparará las nuevas imágenes. Evidentemente el factor humano es clave para incluir la información en el banco de imágenes, e igualmente es fundamental para verificar los resultados que aún hoy no están garantizados.

Otra opción de reconocimiento de imágenes de sobra conocidas por todos son los Reconocedores Ópticos de Caracteres, que posibilitarán sin ningún problema el identificar cualquier elemento de producción que se incluya en el material audiovisual y que ofrecen valor añadido a las imágenes ofrecidas (KIMURA; SHRIDAR, 1991).

4.2 Utilizando la identificación biométrica

Lo que se pretende en estos momentos es solucionar el problema del número de términos que componen un documento audiovisual de informativos utilizando los elementos textuales realizados por el documentalista en texto libre: título, resumen y descripción de imágenes.

Una de las maneras sería la utilización de los reconocedores de audio con el fin de poder transcribir a texto aquello que se dicen en estudio y se indica en la pieza informativa, tanto en la voz en off (locución) como en las declaraciones de las personas involucradas. La transcripción del material no sólo aporta información textual sino pistas importantes, debido a que el periodista en estudio suele realizar un resumen general del contenido del documento que posteriormente se podrá visualizar.

Otro elemento que podría ser especialmente útil sería el reconocimiento de los caracteres de la rotulación utilizada en la postproducción de la noticia. Esta información nos dará pistas claves para la indización de cierto tipo de elementos, especialmente los onomásticos y los geográficos.

Por lo tanto, con esta fácil técnica, aumenta la potencialidad de la indización automática debido al número de texto representativo. No obstante, existe otro elemento que debiera tenerse en cuenta para poder ayudar a la descripción de términos relevantes. En el texto de los informativos existen partes más importantes para la descripción temática de la información. A saber: en la parte realizada por el documentalista sería el título de la noticia, en el cual se deberá aportar al menos información que responda a cuestiones tales como quién y qué, y la entrada realizada en estudio por parte del periodista. Estas partes son las más relevantes desde el punto de vista del contenido desde el momento que el título documental aporta lo más significativo del documento y el periodista en estudio realiza un breve resumen

de contenido.

Podría realizarse una ponderación del texto, valorando los términos aparecidos en estos textos con un peso mayor por medio de algoritmos correctores, lo que facilitaría la descripción textual.

Esta ponderación es potencialmente factible ya que estos elementos son constantes en la información ofrecidas por los medios de comunicación.

5 INDIZACIÓN AUTOMÁTICA POR OTROS PARÁMETROS

Ahora bien, junto a los elementos anteriormente señalados, sería importante que la labor de indización automática se encardinará en el proceso con otras herramientas útiles para la identificación de otros elementos, tantos onomásticos como geográficos.

Por medio del reconocimiento de audio se podrá identificar a personas que hablan en las imágenes y que por lo tanto serán identificadas en la indización. Es cierto que existen algunos casos en los que se cuenta con información sonora únicamente de dicho personaje (conexiones telefónicas), lo cual es de fácil solución desde el momento en que se integrará en el software un elemento de reconocimiento de audio que elimine al personaje de los visualizados si existe el término “telefónico” “conexión telefónica” o “foto archivo” (los términos deberán adaptarse a las pautas y modas de cada una de las cadenas).

Los bancos de datos de voces deberán estar implementados por los gestores de la cadena, no obstante volvemos a señalar que el periodismo actual maneja relativamente pocos personajes en lo que se refiere a la información de actualidad.

Igualmente por medio del reconocimiento de rostros se podrá llegar a identificar a personajes relevantes que forman parte de la noticia. Los reconocedores son cada vez más factibles y menos falibles, por lo que es una herramienta potente para el reconocimiento e indización de estos personajes. Actualmente parece que el mantenimiento de estos bancos visuales también deberá ser mantenidos por el personal de la cadena, no obstante cada vez es mayor el número de empresas que van orientando sus modelos de negocios en el ámbito de la documentación a facilitar este tipo de recursos. Sin olvidar que por medio de la transcripción de texto siempre se podrá identificar a un personaje, no siempre está presente esta información en formato texto (especialmente en información bruta)

Igualmente para información geográfica se podrá contar con el reconocer de caracteres así como con otras herramientas que pudieran reconocer lugares. Esta potencialidad aún es lejana derivado de la inexistencia de bancos de imágenes de lugares, edificios, etc. La teoría es idéntica a la identificación de rostros, y a todas luces más asumible y factible derivado de que zonas geográficas y lugares concretos sufren mucho menos el paso del tiempo que las personas, no obstante estas herramientas no están aun claramente implementadas.

También útil será el reconocimiento de audio para determinar posibles lugares identificados que deberán ser contrastados con un tesoro con los distintas ubicaciones de interés (los tesoros geográficos son de los más fácilmente implementados y desarrollados).

Otra tecnología que se está desarrollando y pudiera ser utilizada con el fin de que ayudara a la indización automática de material audiovisual son las relacionadas con la lectura de patrones, también denominada reconocimiento de formas, la cual consiste en relacionar conceptos especialmente físicos aunque también pudieran ser abstractos que contaran con patrones similares. En este caso el sistema agrupa y clasifica objetos que visualmente cuenta con características comunes por medio del análisis de píxeles, incluso esta tecnología tiene su potencial para la clasificación de ruidos por medio de la interpretación de frecuencias.

Por lo tanto, las imágenes aparecidas serán captadas y analizadas atendiendo a un banco de recursos anteriores, y cuando se detecta características similares lo identificará dentro de dicha clasificación. Sin lugar a dudas el potencial es amplio y con una

implementación nada complicada. Sería necesario realizar la identificación de los elementos diversos de forma pormenorizada en una primera fase con el fin de describir por medio de descriptores a cada clase, siendo posteriormente su extracción de forma automática. El sistema seguirá siendo siempre imperfecto por lo que deberá de seguir implementándose posteriormente, pero nunca como en la primera fase de asignación de descripción a imágenes o ruidos (píxeles o frecuencias).

Estos sistemas aprenden automáticamente, tal y como hace nuestros niños, asignando a las nuevas imágenes a diferentes grupos de clasificación por medio de patrones visuales, auditivos y sensitivos -olvidemos, de momento, este último para la computación- (CARRASCO; MARTÍNEZ, 2011). No obstante, y para ayudar al sistema a mejorar los resultados, el factor humano sigue siendo clave. Esta tecnología es útil para describir animales, edificios, y cualquier tipo de recurso audiovisual.

CONCLUSIONES

La labor documental en los SID-TV es una ardua tarea, complicada y con una cantidad ingente de material. Por ello, todos los servicios buscan soluciones para agilizar el análisis documental del material audiovisual que pasan por simplificarlo (perdiendo información básica), realizando una mayor selección sobre el material a analizar (con los peligros de pérdida de información vital) y optimizando los SID-TV sobre todo por medio de la automatización de procesos. No obstante, las labores de indización siguen sin estar automatizadas en la mayor parte de las cadenas de televisión, seguramente por lo complejo del asunto.

En este trabajo se ha planteado la problemática de realizar indización automática partiendo de los textos descritos por el documentalista sobre información audiovisual de actualidad: título, resumen y descripción de secuencias. Igualmente se ha planteado los fallidos intentos que deben ser solucionados incorporando texto por medio de: induir el texto rotulado en las noticias, el de las declaraciones de los personajes y del periodista en estudio, además de ponderando el texto que de forma prioritaria deberá ser tenido en cuenta para describir la información.

Además parece que estas técnicas deberán utilizarse de manera sincronizada con el reconocimiento de rostros y voz, de tal forma que podremos identificar claramente a las personas visionadas y escuchadas en el documento.

No obstante, además de estas soluciones se plantea la mejora de la indización por medio del novedoso sistema de reconocimiento de patrones así como la implementación de bancos de imágenes para el reconocimiento de otras características visuales.

En definitiva, estimamos que la indización automática tal y como se realiza con información textual es absolutamente inviable, siendo necesario la utilización de diferentes tecnologías para la discriminación de contenidos e identificación de material visualizado y referenciado en la información. Estas tecnologías no son utopías sino que todas existen en la actualidad, cuestión diferente es la necesidad de adaptarlas y hacerlas converger de manera conjunta en los Sistemas de Información Documentación de las cadenas televisivas.

Reconocimientos

Este trabajo ha sido financiado por la Junta de Extremadura (Consejería de Educación, Ciencia y Tecnología) y el Fondo Social Europeo dentro del plan de apoyo a las actuaciones de los Grupos de Investigación inscritos en el catálogo de la Junta de Extremadura. (GR10019).

REFERENCIAS

CALDERA-SERRANO, J. Unidades semánticas discursivas en la información audiovisual televisiva. **Ciencias de la Información**, v. 36, n. 3, p. 39-48, 2005.

_____. Changes in the management of information in audio-visual archives following digitization: current and future outlook. **Journal of Librarianship and Information Science**, v. 40, n. 1, p. 13-20, 2008.

_____; ZAPICO-ALONSO, F. El reconocimiento facial biométrico como método de identificación de personas en los archivos de televisión. **El Profesional de la Información**, v. 18, n. 4, p. 427-431, 2009.

CARRASCO OCHOA, J. A.; MARTÍNEZ TRINIDAD, J. F. Reconocimiento de patrones. **Komputer Sapiens: Revista de divulgación de la Sociedad Mexicana de Inteligencia Artificial**, año III, v. 2, p. 5-9, 2011.

DROP, W. Planificación de textos con ayuda de modelos textuales. In: E. Bernárdez (Comp.) **Lingüística del texto**. Madrid: Arco, p. 293-316, 1987.

DUMAIS, S. T. et al. Inductive learning algorithms and representations for text categorization. In: INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, **Proceedings...** Philadelphia: ACM PRESS, p. 148-155, 1998.

ESPINOSA, M. **Comparación entre un sistema de reconocimiento de voz con el enfoque de redes neuronales y un sistema basado en modelos ocultos de Harkov utilizando el CSLU Toolkit**. México: Tesis UDLAP, 1998.

HATZIS, A. et al. An integrated Toolkit deploying speech technology for computer based speech training with applications to dysarthric speakers. In: EUROPEAN CONFERENCE ON SPEECH COMMUNICATION TECHNOLOGY, 8th, **Proceedings...**, Genova: Eurospeech, p. 2213-2216, 2003.

KIMURA, F.; SHRIDAR, M. Handwritten Numerical Recognition Based on Multiple Algorithms. **Pattern Recognition**, v. 24, n. 10, p. 963-969, 1991.

MOREIRO GONZÁLEZ, José Antonio. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Ediciones Trea, 2004.

Artigo recebido em 15/02/2013 e aceito para publicação em 25/05/2013
