

# Sistema de Telecomunicaciones: ITVNET

Juan Diego López Vargas, Liliana Gómez Meza, Carlos Díaz  
Universidad Santo Tomás - Colombia  
Universidad Politécnica de Madrid - España  
diegolopez@usantotomas.edu.co

## Resumen

El presente artículo tiene como objetivo fundamental presentar un proyecto en desarrollo por un consorcio Colombo-Español, que busca transformar la plataforma iiTV, de difusión y sindicación multiformato orientada hacia los contenidos audiovisuales digitales en un sistema inteligente que clasifique los contenidos a partir de un proceso de filtrado de imagen y audio y de los metadatos extraídos de un proceso de segmentación automática.

## Palabras clave:

IPTV, OCW, MPEG, Contenidos

## Abstract

This article aims to present a major development project by a consortium Colombo-Spanish, where it seeks to transform the platform iiTV, broadcast and syndication-oriented multiformat digital audiovisual content in an intelligent system that classifies the content based a process of filtering image and audio and metadata drawn from an automatic segmentation process.

## Keywords:

IPTV, OCW, MPEG, Contents

---

## Introducción

El objetivo principal del proyecto iTVNet (Interactive TV Net) es fortalecer y mejorar el acceso y la gestión de los contenidos que, en estos tiempos de crisis generalizada, son más demandados por los usuarios, como son los contenidos formativos y educativos.

Facilitar y optimizar el acceso a los contenidos educativos permitirá un profundo desarrollo de las habilidades de los usuarios del sistema. Esto, que es algo que solicita actualmente el mercado laboral, ayudará a obtener una plena inclusión profesional. Este punto es un claro valor añadido para las empresas que hagan uso del sistema.

Para lograr el objetivo previsto, se plantea obtener una plataforma web de contenidos audiovisuales formativos y educativos completa. Con "completa" queremos decir que disponga de todas las facilidades y funcionalidades que permitan al usuario del sistema optimizar las búsquedas de los contenidos que necesita o solicita, minimizar los tiempos de las mismas, asegurar que los resultados devueltos por estas búsquedas están alineados con las preferencias de los usuarios, entrenar con el propio sistema para que mejore los resultados a medida que se utiliza. En definitiva, lograr un sistema de información multimedia cuyos contenidos estén plenamente indexados.

Optiva Media, USTA, UMB y sus socios quieren estar presentes en el nicho de mercado donde se sitúa iTVNet. Para ello se articula este proyecto de colaboración, para que cada socio aporte su experiencia y sus recursos y así conseguir ser pioneros en un mercado tan específico.

## EL CONTENIDO ES LA CLAVE

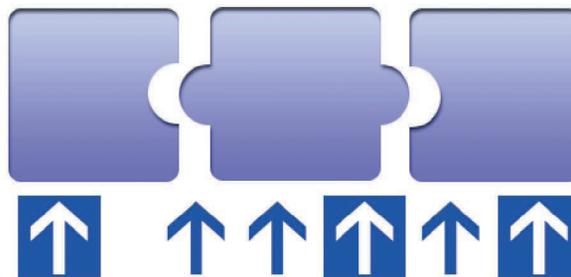


Figura 1: El contenido es la clave

Fuente: Autores

Para obtener la plataforma web de contenidos audiovisuales formativos y educativos, se desarrollarán tres subsistemas:

**Subsistema de segmentación automática:** será capaz de filtrar audio y vídeo en los contenidos audiovisuales digitales que los usuarios hayan ingresado en el sistema para extraer metainformación de los mismos y poner a disposición de los usuarios únicamente los segmentos que más se ajusten al contenido que solicitan. Este subsistema irá de la mano del módulo de segmentación manual que ya se encuentra desarrollado y que permitirá optimizar los resultados devueltos por el segmentador automático.

**Subsistema de perfilado de usuarios:** se encargará de recoger información de los usuarios del sistema y del contexto que les rodea con el fin de disponer de un perfil detallado de los mismos.

**Subsistema de recomendación:** hará uso de los resultados obtenidos por los dos subsistemas anteriores para ofrecer a los usuarios los contenidos que mejor se adapten a sus necesidades o preferencias de la forma más concreta posible (segmentando el contenido original).

Estos tres elementos se basarán en un sistema ontológico que permitirá establecer un esquema conceptual claro y exhaustivo que defina el dominio concreto de los contenidos formativos y educativos. Esta ontología hará posible, entre otras cosas, la comunicación entre diferentes agentes software sin necesidad de intervención humana.

Además, se plantean como objetivos la creación de un módulo de comunicaciones móviles que permita ofrecer los contenidos difundidos por Internet TV a casi todos los entornos de comunicaciones móviles y, del mismo modo, otro módulo para integrar el sistema en una plataforma de TV digital (satélite, cable o terrestre).

Por otro lado, se aplicarán los principios del Diseño para Todos que permitirán que el sistema sea accesible y usable por cualquier persona independientemente de sus limitaciones, ya sean éstas de carácter físico, psicológico, motriz, por razones culturales, demográficas, etc.

## Conclusiones

### a. Innovaciones Tecnológicas y Ventajas para la Empresa

El carácter innovador del proyecto iTVNet viene dado, principalmente, por dos aspectos:

- La segmentación automática de los contenidos, que va a permitir al usuario escoger la parte que realmente es interesante para él del contenido completo que se encuentra en el sistema.
- La indexación semántica de esos fragmentos de los contenidos, lo cual va a facilitar la búsqueda y la recuperación por parte del usuario.

Estos dos aspectos van a obligar a ahondar en trabajos que, aunque ya han supuesto estudios y desarrollo previo, se encuentran, hoy muy por detrás de lo que precisa el consumo de contenido audiovisual. Como sabemos, en estos días de convergencia digital y de hibridación de dispositivos, este tipo de consumo está sufriendo grandes cambios. El panorama audiovisual ha experimentado una revolución radical en los últimos años en que las premisas del sector han cambiado de forma absoluta.

Hoy, podemos descartar la idea de que los contenidos audiovisuales sólo se consumen en casa –a través de la televisión o el vídeo– o en el cine. Aunque la televisión sigue siendo la forma de visionado más universal y el cine sobrevive renqueante como escaparate de las principales novedades del sector audiovisual, han aparecido nuevos competidores. Ordenadores y teléfonos móviles se han ido transformando desde sus funciones originales hasta convertirse en nuevos soportes multimedia provocando que los cimientos del sector se tambaleen. A éstos se unen otros dispositivos a través de los que también se puede acceder al mercado audiovisual, como agendas electrónicas, televisores de bolsillo o incluso consolas originalmente diseñadas para juegos.

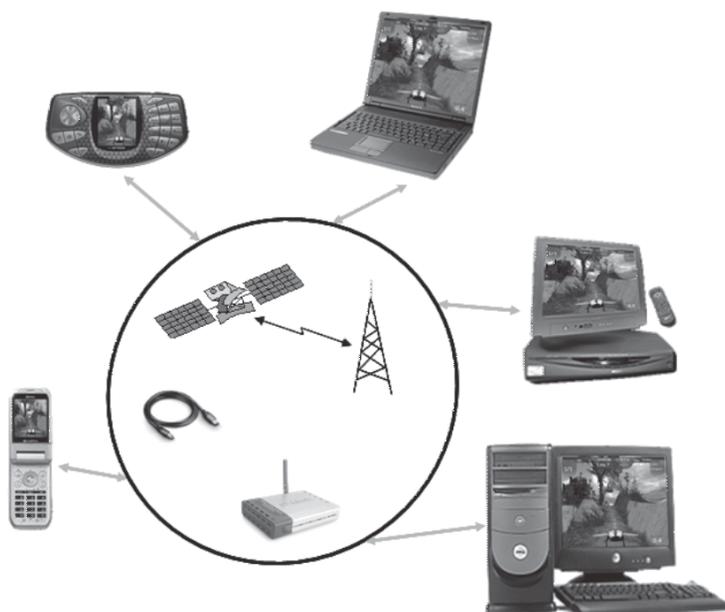


Figura 2: Disfrute de Contenidos en Entornos Multiplataforma  
Fuente: Autores

El reciente Barómetro Motorola de Consumo Multimedia en España durante 2010 [9] es revelador en cuanto al calibre de los cambios que estamos viviendo en el consumo de productos audiovisuales. De acuerdo con los datos de este estudio, aunque todavía un 72% de los españoles consume televisión en directo todas las semanas, un creciente porcentaje de consumidores de contenidos audiovisuales los obtiene por otros medios. Así, prácticamente un 50% de los espectadores consume al menos una vez por semana vídeos por streaming a través de Internet, un 28% se descarga algún vídeo y un 18% ve televisión a la carta. Las cifras absolutas de consumo en streaming son realmente impresionantes, pues el portal Youtube, que es el máximo exponente de este tipo de consumo audiovisual, alcanzó, durante 2009, los mil millones de visitas diarias (La Vanguardia, 2011).

Los paradigmas han cambiado no sólo respecto al soporte del visionado, sino también respecto al momento y el lugar del mismo. La clásica emisión lineal de contenidos, en la que la voluntad del espectador se limitaba a escoger si veía o no la emisión broadcast que tenía disponible en su región, no ha desaparecido, es evidente,

pero ha nacido otra línea de oferta audiovisual atomizada que pivota en torno a la voluntad del espectador. El espectador escoge el momento y el lugar en el que va a disfrutar del contenido audiovisual. Tal y como decía uno de los eslóganes de una conocida compañía de telecomunicaciones, "la vida es móvil". En nuestros días ya no es necesario estar sentado ante nuestra mesa de trabajo para consumir contenidos audiovisuales a través de un ordenador. Podemos hacerlo desde un portátil, desde un teléfono móvil de última generación, desde una tableta, desde un pequeño televisor de bolsillo o desde una PDA.

Para ver nuestra serie favorita, no es necesario esperar el día y hora de su emisión por la cadena de televisión. Ahora podemos buscarlo en la web de la cadena, que probablemente lo ponga a disposición del espectador por este medio. También podemos buscarlo en páginas especializadas donde podemos bajarnos capítulos de cientos de series distintas o incluso podemos encontrar partes del episodio o escenas concretas en portales de Video on Demand como Youtube o Dailymotion. En conclusión, es más fácil ver nuestros contenidos favoritos.

Sin embargo, aún queda mucho trabajo por hacer en cuanto al filtrado o fragmentación (segmentación) de los muchos contenidos que se están generando (anteriormente el usuario era un mero espectador, pero ahora se ha convertido en otro generador de contenidos, un prosumer) o a la categorización de los mismos (indexación). En estos campos es donde el consorcio iTVNet pone todas las fuerzas.

### b. Tecnologías más significativas

A continuación se detalla, desde un punto de vista tecnológico, cómo se utilizarán las herramientas de las que dispone el proyecto para la ejecución de las tareas.

En uno de los puntos más importantes de iTVNet, como es el de la segmentación automática de contenidos, se seguirá un esquema similar al que se muestra a continuación en referencia al reconocimiento de vídeo:

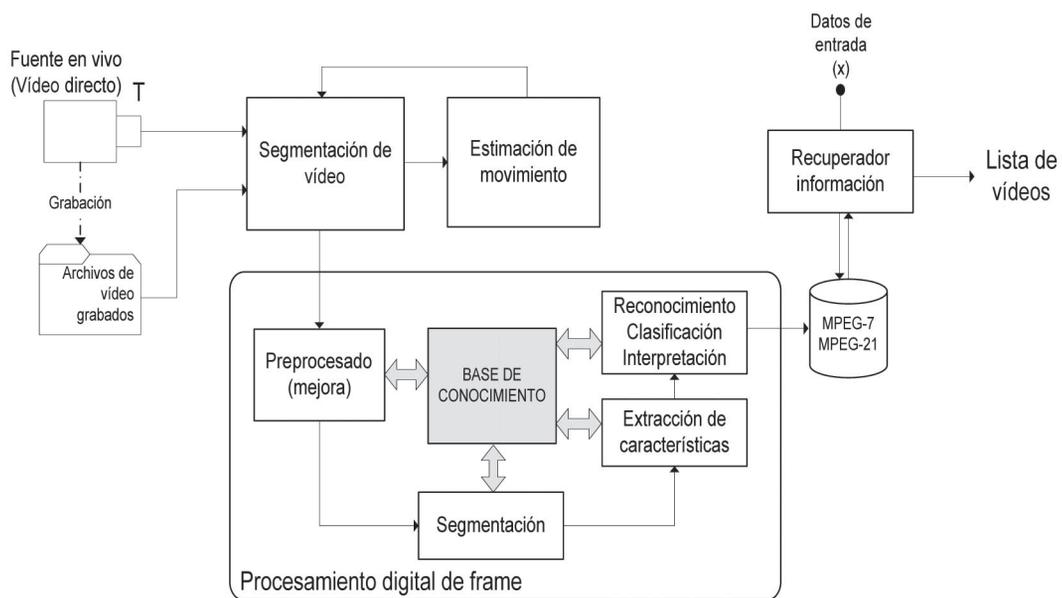


Figura 3: Esquema reconocimiento de vídeo

Fuente: Autores

A partir de una fuente de vídeo, bien de una fuente en vivo o bien previamente grabada, se realiza la adquisición de frames mediante segmentación del vídeo. La segmentación de vídeo consiste en obtener por un lado, los objetos relevantes del primer plano (foreground) y, por otro, los objetos pertenecientes al fondo (background) donde habrá que identificar los objetos que se mueven en el frame mediante técnicas de estimación de movimiento. Para ello, utilizaremos como punto de partida dos frames consecutivos. El primer frame captado se corresponderá con el fondo.

La estimación de movimiento indica el factor de cambio entre frames en el tiempo. La técnica de detección de movimiento que se empleará es la estimación de flujo óptico. El flujo óptico mide el movimiento de los objetos en una secuencia. Los objetos cercanos generan mucho flujo, mientras que los lejanos apenas se mueven. El flujo óptico indica la dirección, el sentido y la intensidad del movimiento, describiendo adecuadamente los desplazamientos que se producen en la imagen. Se utilizará la variante del método basado en bloques del frame.

Una vez segmentado el vídeo y detectado el flujo óptico, se pasará al procesamiento digital de cada plano del frame, foreground y background, pasando por los siguientes bloques:

- Preprocesado: en este bloque se realizará una mejora del frame (a partir de ahora imagen) para conseguir un plano más limpio para que así aumenten las probabilidades de éxito en las etapas posteriores. Se elimina el ruido, sombras, cambios de iluminación, etc. que se pueden confundir con regiones. Las técnicas que actualmente existen son:
  - Aumento del contraste / luminosidad. Se utilizan técnicas de umbralización de histograma de la imagen. El umbral se calcula en función de la información que se pretenda modificar.
  - Realce de bordes o contornos. El realce de bordes o contornos es muy interesante para resaltar objetos que queremos detectar o clasificar. Se consigue mediante el uso de filtros paso alto.
  - Reducción de ruido. Mediante operaciones de filtrado paso bajo se consigue reducir el ruido de la imagen.
- Segmentación: éste es uno de los bloques principales del sistema ya que es el punto donde, en función de unas características, se divide la imagen en áreas con significado.

La segmentación es la operación de bajo nivel que determina regiones homogéneas y disjuntas o, equivalentemente, que encuentra los bordes de dichas regiones. Estas regiones homogéneas, o los bordes, corresponden a los objetos, o partes de ellos, dentro de la imagen. Por lo tanto, la segmentación juega un papel muy importante como el primer paso antes de aplicar operaciones de alto nivel, como reconocimiento, interpretación semántica y representación.

Las técnicas de segmentación son muy dependientes del propósito de la aplicación y del tipo de imagen a analizar. Antes de segmentar, es preciso definir qué objetos/características interesa determinar.

Los tipos de segmentación de imágenes se reparten en cinco grandes grupos:

- Basadas en características. Se asigna cada píxel a una región en función de características locales de la imagen en el píxel y en su vecindad.
- Basadas en transiciones. Los métodos de segmentación basada en transiciones aplican sobre la imagen un detector de bordes. Las regiones se definen a partir de las fronteras delimitadas por los bordes detectados. La detección de bordes se realiza mediante la medición del gradiente de la imagen en dirección determinada a lo largo de una recta.
- Basadas en modelos. Estas técnicas se presuponen conocidas algunas características de los objetos o regiones en la imagen, tales como, rectas, objetos circulares, etc. La transformada de Hough es la más utilizada permitiendo localizar objetos dentro de una imagen.
- Basadas en homogeneidad. Son técnicas de crecimiento de regiones, métodos de abajo-arriba. Parte de unos puntos o semillas, incluye en la misma región píxeles vecinos con características (niveles de gris, textura, color) similares hasta que todo los píxeles se han asignado a uno de los puntos de partida.
- Basadas en morfología matemática (Transformada watershed). La transformada watershed es una técnica morfológica de segmentación de imágenes de niveles de gris. Es un método basado en regiones, que

divide todo el dominio de la imagen en conjuntos conexos. Esta técnica no va a ser implementada ya que es un método difícil de automatizar.

- Extracción de características. Una vez dividida la imagen en zonas con características de más alto nivel, se pasará a su extracción de características de tipo morfológico, como el área, perímetro o excentricidad o características basadas en la textura o en el color.
- Reconocimiento / clasificación / interpretación. Con las características analizadas de cada región, se clasificarán las regiones analizadas, bien en aspectos de color y textura o bien en objetos pertenecientes a clases, como personas, vehículos, etc.

A partir de este punto, se guardará la información obtenida de cada plano del frame etiquetando los contenidos mediante metadatos siguiendo el estándar MPEG-7.

La información obtenida seguirá un esquema conceptual definido en una ontología. Las ontologías tienen los siguientes componentes que sirven para representar el conocimiento de un dominio:

- Conceptos: son las ideas básicas que se intentan formalizar. Los conceptos pueden ser clases de objetos, métodos, planes, estrategias, procesos de razonamiento, etc.
- Relaciones: representan la interacción y enlace entre los conceptos del dominio. Suelen formar la taxonomía del dominio. Por ejemplo: subclase-de, parte-de, parte-exhaustiva-de, conectado-a, etc.
- Funciones: son un tipo concreto de relación donde se identifica un elemento mediante el cálculo de una función que considera varios elementos de la ontología. Por ejemplo, pueden parecer funciones como categorizar-clase, asignarFecha, etc.
- Instancias: se utilizan para representar objetos determinados de un concepto.
- Axiomas: son teoremas que se declaran sobre relaciones que deben cumplir los elementos de la ontología. Por ejemplo: "Si A y B son de la clase C, entonces A no es subclase de B", "Para todo A que cumpla la condición C, A es B", etc.

Para poder realizar la definición de ontologías, se necesitan lenguajes de marcado apropiados que representen el conocimiento de estas ontologías. Uno de los lenguajes con gran capacidad expresiva que se ha convertido en estándar para realizar anotaciones de ontologías es OWL (Ontology Web Language).

## Referencias

- [1] R. Ll. Roberto. "Especificación owl de una ontología para teleeducación en la web semántica" Tesis doctoral, Valencia, 2007.
- [2] R. P. Keily, R. L Rodrigo, Web semántica: un nuevo enfoque para la organización y recuperación de información en el web. ISSN 1024-9435 ACIMED v.13 n.6 Ciudad de La Habana nov.-dic. 2005
- [3] Optiva Media [Online]. Available: [http://www.optivamedia.com/productos\\_broadbandtv.html](http://www.optivamedia.com/productos_broadbandtv.html).
- [4] V.G. Rafael. Un Entorno para la Extracción Incremental de Conocimiento desde Texto en Lenguaje Natural, Departamento de Ingeniería de la Información y las Comunicaciones. Tesis doctoral Murcia, 2005.
- [5] S. M. Aurelio, Desarrollo de una aplicación Web 2.0 para la captura y análisis de la Valoración del usuario Proyecto Final De Carrera , Valencia, 2010.

- 
- [6] C, L, Angel. Estudio Técnico de factibilidad para la implementación de un canal de Televisión digital, Proyecto Final de Carrera, Quito 2008.
  - [7] Aragall, Presidente de la Coordinadora del Dpt en España
  - [8] G. Vanderheiden. M. Cooper. Comprender las WCAG 2.0 .W3C, University of Wisconsin-Madison Ben Caldwell, Trace R&D Center, Universidad de Wisconsin-Madison.
  - [9] Mediacenter Motorola, El reciente Barómetro Motorola de Consumo Multimedia en España durante 2010.