

# ViRe-YouTube: Visualizando los resultados de búsquedas en YouTube

## ViRe-YouTube: Visualizing search results in YouTube

Carlos Mera<sup>1</sup>, MSc y Roberto Therón<sup>2</sup>, PhD

1. Escuela de Ingeniería de Sistemas y Computación, Universidad del Valle, Cali - Colombia
2. Departamento de Informática y Automática, Universidad de Salamanca, Salamanca - España

carlosm@univalle.edu.co, theron@usal.es

Recibido para revisión 28 de enero de 2009, aceptado 20 de mayo de 2009, versión final 1 de junio de 2009

**Resumen-** En este trabajo se aborda el problema de la visualización de los resultados de búsqueda de videos en YouTube [36] desde tres perspectivas: (1) La adaptación de una técnica de visualización de información para dar una vista general de los videos 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 videos utilizando las relaciones implícitas entre estos, etiquetas y otros videos relacionados; (3) La implementación de técnicas de previsualización de videos, para brindar información de su contenido visual. Como propuesta de investigación se implementó ViRe-YouTube un sistema para la recuperación de videos 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 videos recuperados y por ende a hacer una selección más adecuada de los videos a reproducir.

**Palabras Clave-** Visualización de Información, Previsualización de Vídeos, Visualización de resultados de Búsqueda de Vídeos

**Abstract-** In This paper we discuss the problem of visualizing video search results in YouTube [36] from three approaches: (1) The adaptation of a visualization technique to give an overview of recovered videos and help users to recognize the relationships, patterns and trends among videos; (2) Use a graph-based technique for exploratory video search using the relations among labels and other related videos; (3) The implementation of video preview techniques, which provides information about the visual content. We implemented ViRe-YouTube, a system to video search that replace the traditional interaction with an interaction style similar to exploration. In this kind of interaction, the organization and presentation of results help users to analyze retrieved videos and make a better choice for the video they want to playback.

**Keywords-** Information Visualization, Video Summarization, Visualization of video Search Results

### I. INTRODUCCIÓN

El creciente número de videos de YouTube ha convertido este repositorio en uno de los lugares más populares y visitados de Internet. Se estima que hay más de 45'000.000 de videos y que el repositorio crece a un ritmo asombroso, en el que cada minuto los usuarios suben cerca de 7 horas de video [3]. YouTube tiene el potencial de contener información de interés para muchos tipos usuarios. Sin embargo, el disponer de tanta información genera inconvenientes en la búsqueda y exploración de videos, lo que conlleva a que este proceso sea lento, tedioso, y algunas veces, frustrante.

La búsqueda de videos en YouTube, al igual que en otros repositorios de información multimedia, se realiza a través de una interfaz en la que el usuario ingresa los términos que describen los videos a buscar. Cuando se obtienen los resultados, el sistema muestra en una lista una imagen representativa y la descripción general de los videos 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 videos 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 suficientes [19] ya que requieren un gran esfuerzo cognitivo para la exploración; Y (c) una imagen representativa del video no da la suficiente información de su contenido para realizar una selección cuidadosa de los videos que son relevantes para el usuario.

Existen diferentes propuestas en la literatura que intentan dar solución a estos y otros inconvenientes de los sistemas de recuperación de videos. Algunos autores, por ejemplo, se centran en el proceso de generación de previsualizaciones del video

[14], [17], [19], [22], [29], [34], otros proponen mecanismos de búsqueda que involucran conceptos semánticos [21], y otros exploran la posibilidad de realizar búsquedas basadas en contenido [11]. 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 [10].

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 brindan a los usuarios la posibilidad de 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 brindan al usuario información del contenido visual, y 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 descripción del estado del arte. En la Sección 3 se presentan algunos trabajos relacionados. En la sección 4 se detalla ViRe-YouTube, la herramienta de visualización propuesta. En la Sección 5 se presentan las conclusiones. Finalmente, la Sección 6 señala algunos trabajos futuros.

## II. ESTADO DEL ARTE

El uso de las técnicas de visualización de información en los sistemas de recuperación de vídeos, ofrece la posibilidad de desarrollar nuevas formas para presentar la información que es recuperada, haciendo más sencilla la experiencia de búsqueda de los usuarios. A continuación se presentan los conceptos relacionados con el área de visualización de información y se describen, brevemente, algunas técnicas de previsualización de vídeos.

### A. 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 incluyen 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..." [17]. Esta definición indica que las técnicas de visualización de información buscan generar representaciones visuales de los datos, tales que éstas maximicen la posibilidad de comprensión e interiorización del conocimiento inherente en ellos.

Se han propuesto diferentes modelos que permiten llevar un conjunto de datos en crudo a una representación visual para su interpretación y análisis [5], [7], [13]. Ben Fry propone un modelo

que consta de siete etapas que van desde la adquisición de los datos hasta su presentación visual, e interacción con el usuario [15]. 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, se hace una breve descripción de algunos 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.

1) Psicología de la Percepción: Cómo menciona Shneiderman [29], 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 [32]. Por ejemplo, en la Figura 1 se muestran dos conjuntos de puntos. Uno de ellos contiene un punto rojo dispuesto aleatoriamente de tal manera que la atención no se encuentra prefijada sobre ningún objeto [24]. Antes de que el observador pueda enfocar la atención sobre algún objeto y comience una tarea consciente de búsqueda, la tarea de detectar el punto rojo ya ha sido realizada

Parte de la importancia de los sistemas de visualización de información recae en la capacidad para ayudar a "ver" la información que se oculta en los datos. La utilización de propiedades preatentativas ayuda a los sistemas de visualización en esa tarea de percepción. Algunas propiedades preatentativas ampliamente utilizadas son:



Figura 1. Detección preatentativa basada en color

- 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.

Estas propiedades visuales producen un impacto sobre cuatro niveles de percepción del usuario [4]:

- 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.

2) Metáforas de Representación de Datos: Una de las decisiones más importantes en el proceso de visualización de información, es determinar de qué manera se van a presentar los datos al usuario. Ben Fry [15] señala que esta selección debe hacerse de tal forma que la representación sea lo más simple posible pero que trasmite los aspectos más relevantes del conjunto de datos. Estás 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 2 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 de rectángulos. En este tipo de representación, cada nivel de anidación se corresponde con un nivel de descomposición de la jerárquica.

3) Interacción con el Usuario: El mantra de búsqueda visual de información de Shneiderman [29] "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 cómo 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 sí lo son.

- Detalles Bajo 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): poder extraer un conjunto de objetos de la visualización para su análisis posterior.

## B. Previsualización de Videos

El objetivo de un usuario al previsualizar un video es evaluar si ese video es de su interés, o no, sin tener que descargarlo y reproducirlo en su totalidad. La previsualización de un video se realiza disminuyendo la información redundante y eliminando el contenido que se puede predecir viendo una parte del mismo [34].

Comúnmente, la previsualización de videos se clasifica como previsualización estática ó previsualización dinámica. La previsualización estática se genera utilizando un conjunto de frames del video [10], [11], [12], [20], [29], [32], [34]. La previsualización dinámica es una versión corta del video original generada a partir de algunos planos del video [23], [19], [14], [6].

En el caso de las previsualizaciones estáticas, un tema de investigación es cómo 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 [10] (Figura 2a). Smoliar et al. [30] proponen otra manera de organización llamada micon, que es un sólido rectangular en el que los frames se disponen en orden secuencial (Figura 2b). Kim et al. [18] crean mosaicos utilizando frames con características de color similar (Figura 2c) y Goeau et al. [16] proponen una disposición llamada Table Of Video Contents que se ilustra en la Figura 2d.

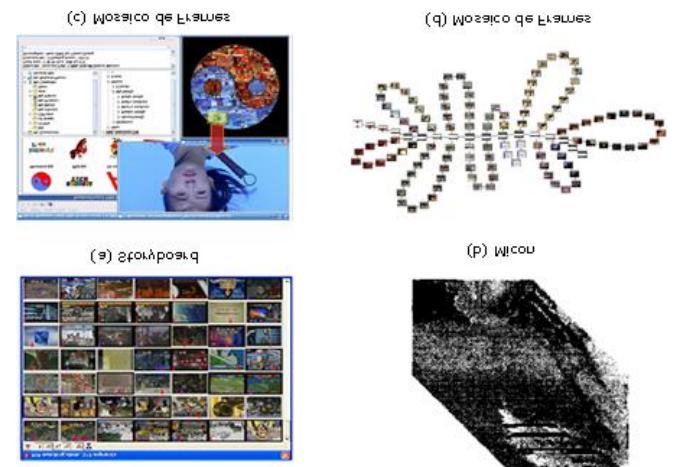


Figura 2. Métodos de organización especial de frames

### III. 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. Sin embargo, en los sistemas de recuperación de vídeos estos elementos descriptivos suelen ser insuficientes para que el usuario seleccione rápidamente los vídeos que son relevantes para él.

En el proyecto Informedia [8], [9], [26] se desarrolló 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 diferentes segmentos de vídeos que se recuperan en una consulta. No obstante, los métodos de representación que se proponen en Informedia están orientados a colecciones de vídeos estructuradas (noticias y entrevistas), que además, pueden ser complementados con otros elementos de visualización de información.

YouTube propone una herramienta visual para la exploración de vídeos. Sin embargo, ésta no cuenta con las iteraciones de las que Shneiderman se refiere en su mantra de búsqueda visual de información [29].

SearchMe [27] y DiscoverGalaxy [13] son dos aplicaciones que proponen componentes visuales para la búsqueda de vídeos. SearchMe utiliza una metáfora basada en listas llamada Cover Flow y DiscoverGalaxy proporciona dos vistas basadas en tablas y una más en la que se organizan 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 oSkope [25] y TimeTube [31], los cuales utilizan visualizaciones basadas en gráficos de dispersión. No obstante, éstas se limitan a organizar los vídeos a partir de dos de sus atributos.

Flookon [14] 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, Blinkx [5] y VideoSurf [33] brindan información del contenido visual de los vídeos recuperados. A pesar de ello, su representación visual está basada en listas.

### IV. VIRE-YOUTUBE

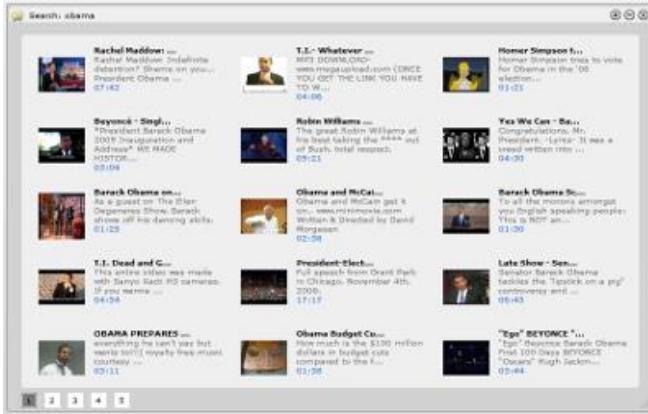
ViRe-YouTube es una herramienta de visualización de resultados de búsqueda de vídeos de YouTube, que cuenta con múltiples vistas para analizar los vídeos recuperados, permite búsquedas exploratorias y la previsualización del contenido de los vídeos, brindando al usuario la posibilidad de hacer una selección 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 cómo esos componentes implementan las tareas principales del mantra de búsqueda de información visual de Shneiderman.

#### A *Un Gráfico de Dispersión para Analizar los Vídeos Recuperados*

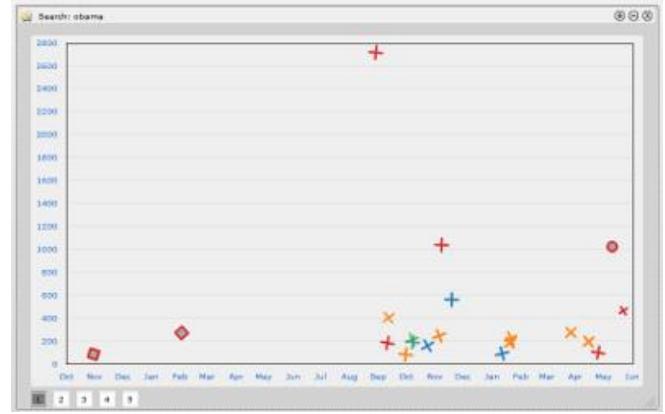
Baeza y Ribeiro [2] mencionan que una vista general del conjunto de datos recuperados puede dar al usuario pistas de su organización, y más aún, le puede ayudar a seleccionar o eliminar fuentes de información. Una de las limitaciones de las listas de documentos (y en este caso particular, de las listas de vídeos) es que no le dan al usuario esa vista general a la que se refieren Baeza y Ribeiro en [2].

En este trabajo se propone la utilización de un gráfico de dispersión enriquecido con propiedades visuales preatentivas para organizar y presentar los vídeos recuperados en 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, dando la vista general que propone Shneiderman y que mencionan Baeza y Ribeiro en [2].

1) Traducción de Atributos Gráficos: Cada asignación de características visuales a los objetos resuelve con diferente grado de efectividad los cuatro niveles de percepción propuestos por Bertin en [4], siendo importante determinar qué atributos de los vídeos se van a representar con cada característica visual:



(a) Lista 2D de los videos recuperados



(b) Gráfico de dispersión de los videos recuperados

**Figura 3.** Dos tipos de vistas en los que se visualizan los vídeos recuperados para la consulta “obama”

- Posición: la posición de los objetos le permite al usuario asociar y agrupar los videos por proximidad; además, 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 videos son su duración (en X) y el número de visitas (en Y).
- Forma: es una característica visual ideal para el nivel asociativo [4]. Por esta razón, se propone utilizar la forma para representar la antigüedad de los videos.
- Color: en términos de los niveles de percepción de Bertín [4], el color es ideal para impactar en el nivel asociativo y selectivo, por esta razón se utiliza para representar las categorías de videos 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 eficaz, en la que en un espacio limitado se presenta un conjunto amplio de atributos informativos de la colección de videos recuperados. La Figura 3 ilustra un ejemplo para la consulta "Obama" realizada a finales de mayo de 2008. En este ejemplo, el usuario puede rápidamente observar en el gráfico de dispersión (Figura 3b) 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 cuatro categorías diferentes y que uno de ellos 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 (Figura 3a).

- 2) Controles e Interacción con el Usuario: Esta visualización cuenta con un mecanismo que le permite al usuario

ver los detalles de los videos, evitando de esta manera, la perdida del contexto de la consulta, a diferencia de los que sucede en los sistemas presentados en la sección III. Otro control propuesto e implementado en ViRe-YouTube permite a los usuarios cambiar los atributos que se representan en los ejes del gráfico facilitando el análisis de los videos desde diferentes puntos de vista.

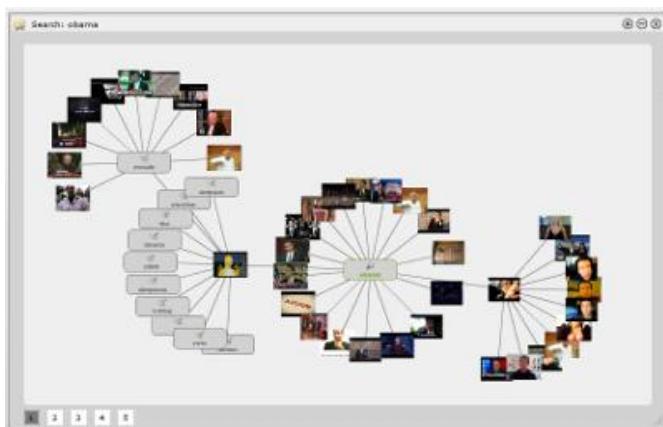
#### B. Búsqueda Exploratoria de Videos Usando Grafos

ViRe-YouTube usa un mecanismo de exploración de videos del tipo Query searching and Browsing [37], similar al utilizado por Flokoon, en el que se emplean las etiquetas y las listas de videos relacionados para explorar otros videos que pueden estar interconectados con los términos de la consulta inicial.

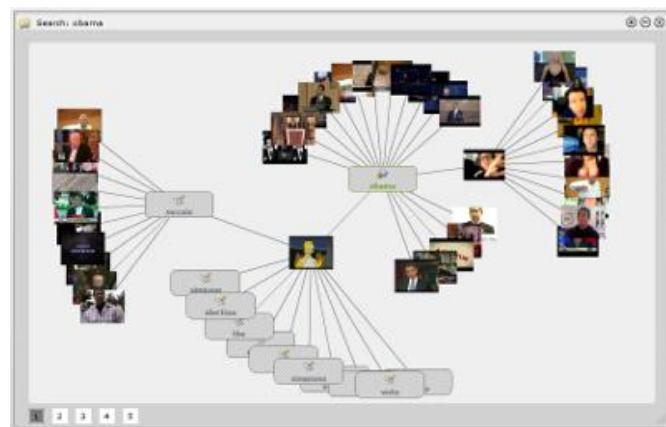
Los grafos, a diferencia de las visualizaciones basadas en listas, ofrecen un mecanismo que le permite al usuario explorar el espacio de búsqueda sin perder el contexto de la consulta inicial. Además, esta metáfora de representación visual, disminuye la carga cognitiva y las tareas memorísticas de exploración,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

ViRe-YouTube utiliza dos tipos de grafos: grafos dirigidos por fuerzas y grafos radiales. La Figura 4 ilustra una búsqueda exploratoria iniciada con los videos recuperados para la consulta "obama" utilizando los dos tipos de grafos, los cuales se explican a continuación.

- 1) Grafos Dirigidos por Fuerzas: La disposición de los elementos del grafo se modela 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



(a) Grafo Dirigido por Fuerzas



(b) Grafo Radial

**Figura 4.** Búsqueda exploratoria usando grafos para la consulta iniciada con el término clave "obama"

fuerzas 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 videos. De forma natural se forman e identifican fácilmente grupos entre los elementos representados mediante el grafo. Además, como se basan en analogías físicas de objetos comunes, su comportamiento es relativamente fácil de predecir y entender.

2) Grafos Radiales: Un grafo organizado en una disposición radial, o "grafo radial", es tratado como un árbol cuya raíz es definida por un nodo raíz (que es en el que se centra el usuario) 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 [35].

ViRe-YouTube utiliza los dos tipos de grafos en la búsqueda exploratoria teniendo en cuenta una anotación de Yee et al. [35] donde sugiere que la habilidad para ver un grafo, interactivamente, desde diferentes perspectivas puede producir nuevo conocimiento sobre los datos.

3) Controles e Interacción con el Usuario: Basados en el mantra de búsqueda de información visual de Shneiderman [29], los grafos, al igual que los gráficos de dispersión, dan al usuario una vista general de los videos recuperados, para ello cuentan con controles que le permiten al usuario variar la distancia focal de la visualización y obtener detalles de los videos bajo demanda. Estos grafos también cuentan con un control que les permite explorar el espacio de búsqueda de los videos (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).

### C. 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 [29] 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 videos en categorías y autores, permiten recuperar los videos que pertenecen a un determinado autor o categoría, como se ilustra en la Figura 5. Estos criterios son útiles para discernir y seleccionar los videos de interés que pertenecen a una clase determinada.

1) 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 5b.

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

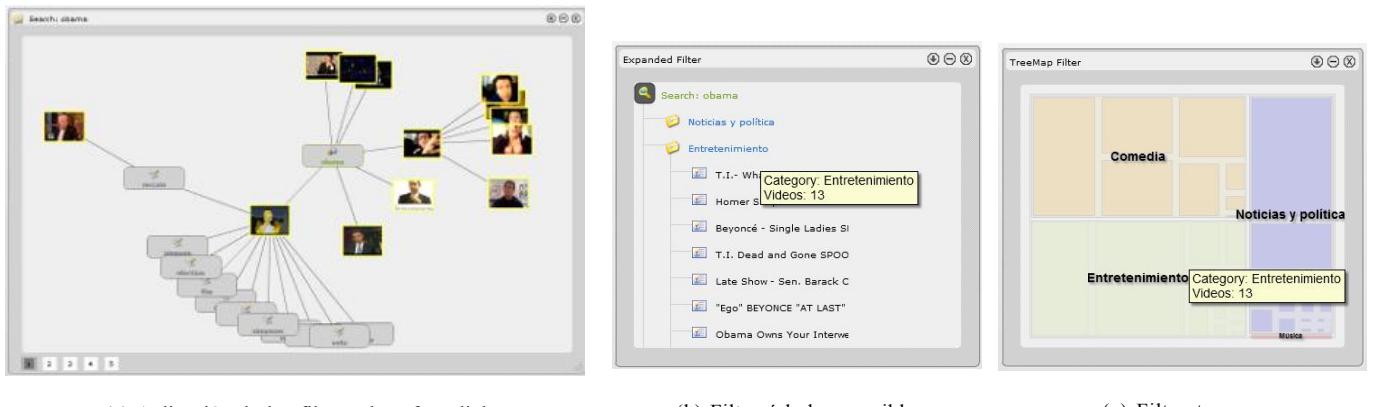
2) Filtro Basado en un TreeMap: Un treemap [28] es una técnica de visualización en la que el árbol se representa con un conjunto de rectángulos anidados que simbolizan los nodos del árbol.

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 rápidamente que la adquirida mediante 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 5c 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 video.



**Figura 5.** Aplicación de los filtros en una búsqueda exploratoria iniciada con el término clave "obama"

#### D. Métodos de Previsualización de Vídeos

Como ya se mencionó una de las falencias de algunos sistemas de recuperación de vídeos de YouTube es que sólo muestran al usuario una única imagen (un frame) extraída del contenido del video, lo que dificulta y retraza la tarea de encontrar los videos relevantes para el usuario. Para cubrir esa desventaja, ViRe-YouTube integra cuatro métodos de previsualización que ayudan al usuario a identificar de forma más rápida los videos 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 video. 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, idea presentada por Simone Santini en [26].

1) Previsualización Dinámica: La previsualización dinámica que integra ViRe-YouTube consiste en un GIF animado creado a partir de un conjunto de frames del video 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 video a través de una vista rápida del mismo.

2) Previsualización en Espiral: Se propone una previsualización en la que los frames del video 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 grafico. La ecuación polar de la espiral logarítmica está determinada por la ecuación (1).

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

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

Este tipo previsualización le permite al usuario formarse una idea rápida de la estructura del video, e identificar fácilmente los planos del mismo. La previsualización en espiral le da al usuario información del contenido del video y le muestra cómo cambian los frames de principio a fin o de fin a inicio, como se ilustra en la Figura 6a.

3) Previsualización como Paso de Diapositivas: Otra previsualización implementada en ViRe-YouTube es la de paso de diapositivas. En este tipo de previsualización el usuario puede ver cómo cambia el video frame a frame, como se ilustra en la Figura 6b. Esta previsualización tiene la ventaja de que el usuario puede ver todos los frames extraídos del video 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.

4) Previsualización en Tira de Frames: Ésta es un 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 6c.

#### V. CONCLUSIONES

El uso de técnicas de visualización de información, en sistemas de recuperación de videos, abre la posibilidad de explorar colecciones de videos, 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.

ViRe-YouTube presenta y organiza 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. Asimismo, toma provecho de las capacidades de



(a) Previsualización en espiral

(b) Previsualización en paso de diapositivas

(c) Previsualización en tira de frames

**Figura 6.** Tipos de previsualización estática en ViRe-YouTube para uno de los vídeos recuperados con la consulta "obama"

procesamiento del sistema visual humano para detectar patrones y tendencias de los vídeos recuperados a través de una representación con propiedades preattentivas.

Finalmente, ViRe-YouTube usa en conjunto diferentes visualizaciones, filtros y mecanismos de previsualización para 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.

## VI. TRABAJOS FUTUROS

Las técnicas de visualización de información abren la posibilidad de abordar algunos problemas no contemplados en ViRe-YouTube, algunos de ellos son:

- Estudios de impacto, aceptación y efectividad de los diferentes tipos de visualización utilizados.
- Visualizaciones que soporten la búsqueda basada en recuperación semántica y/o características de bajo nivel de los vídeos.
- Algoritmos que disminuyan la carga de la memoria y su tiempo de respuesta.

## REFERENCIAS

- [1]Aiello, A., and 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", Departamento de Computación, Universidad de Buenos Aires, 2004  
[2]Baeza R., and Ribeiro B.: "Modern Information Retrieval, Chapter: User Interfaces and Visualization", Addison Wesley, 1999, pp. 257-324.  
[3]Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S.,

Ravichandran, D., and 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, New York, NY, ACM, 2008, pp. 895-904

- [4]Bertin, J. "Semiology of graphics", University of Wisconsin Press, 1983  
[5]Blinkx: <http://www.blinkx.com/> [Última visita: Abril 22 de 2009]  
[6]Card, S.K., and Mackinlay, J.: "The structure of the information visualization design space", in Proceedings of IEEE Symposium on Information Visualization, October 1997, pp. 92-99  
[7]Chen, F., Cooper, M., and Adcock, J.: "Video summarization preserving dynamic content", in Proceedings of the international workshop on TRECVID video summarization, New York, NY, ACM, 2007, pp. 40-44  
[8]Chi, E.H.H., and Riedl, J.: "An operator interaction framework for visualization systems", in Proceedings of the IEEE Symposium on Information Visualization, October 1998, pp: 63-70  
[9]Christel, M.G.: "Visual digests for news video libraries", in Proceedings of the ACM Multimedia Conference, 1999, pp. 303-311  
[10]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, New York, NY, ACM, 2008, pp. 447-456  
[11]Decler, C., Hacid, M.S., and Kouloudjian, J.: "A genetic model for video content based retrieval", in Proceedings of the 1998 ACM symposium on Applied Computing, New York, NY, ACM, 1998, pp. 458-459  
[12]DeMenthon, D., Kobla, V., and Doermann, D.: "Video summarization by curve simplification", in Proceedings of the sixth ACM international conference on Multimedia, New York, NY, ACM, 1998, pp. 211-218  
[13]DiscoverGalaxy: <http://www.tkaap.com/galaxy/> [Última visita: Abril 22 de 2009]  
[14]Folkoon: <http://www.flokoon.com/#youtube> [Última visita: Abril 22 de 2009]  
[15]Fry, B. J.: "Computational information design", PhD thesis, School of Architecture and Planning, Massachusetts Institute of Technology, 2004  
[16]Goeau, H., Thievre, J., Viaud, M.L., and Pellerin, D.: "Interactive visualization tool with graphic table of video contents", IEEE International Conference on Multimedia and Expo, July 2007, pp. 807-810  
[17]Infovis: <http://www.infovis.org/infovis/2005/> [Última visita: Noviembre de 2008]

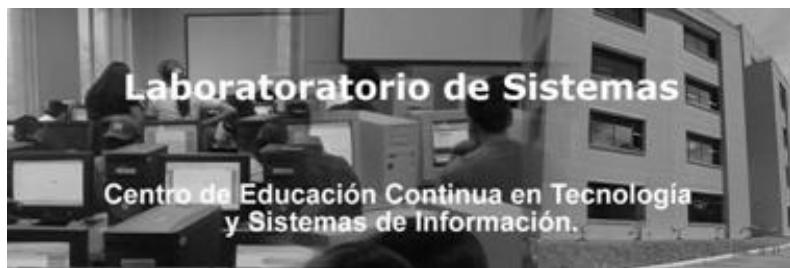
- [18]Kim, K., Essa, I., and Abowd, G.D.: "Interactive mosaic generation for video navigation", in Proceedings of the 14th annual ACM international conference on Multimedia, New York, NY, ACM, 2006, pp. 655-658
- [19]Kules, W.M.: "Supporting exploratory web search with meaningful and stable categorized overviews", PhD thesis, University of Maryland at College Park.
- [20]Li, Y., Zhang, T., and Tretter, D.: "An overview of video abstraction techniques", Technical report, HP Laboratories - Imaging Systems Laboratory, July 2001
- [21]Luo, H.: "Concept-Based Large-Scale Video Database Browsing and Retrieval via Visualization", PhD thesis, University of North Carolina at Charlotte, 2007
- [22]Ma, Y.F., Lu, L., Zhang, H.J., and Li, M.: "A user attention model for video summarization", in Proceedings of the tenth ACM international conference on Multimedia, New York, NY, ACM, 2002, pp. 533-542
- [23]Mills, M., Cohen, J., and Wong, Y.Y.: "A magni<sup>-er</sup> tool for video data.", in Proceedings of the SIGCHI conference on Human factors in computing systems, New York, NY, ACM , 1992, pp. 93-98
- [24]Mislej E.M.: "COCOVAS: Un sistema de visualización de Información documental para sistemas de Recuperación de Información de Internet", Departamento de Computación, Universidad de Buenos Aires, Noviembre de 2006.
- [25]Oskope: <http://oskope.com/> [Última visita: Abril 22 de 2009]
- [26]Santini, S.: "Who needs video summarization anyway?", in Proceedings of the International Conference on Semantic Computing, September 2007, pp. 177-184
- [27]Searchme: <http://www.searchme.com/> [Última visita: Abril 22 de 2009]
- [28]Shneiderman, B.: "Tree visualization with tree-maps: 2-d space filling approach", ACM Transactions on Graphics, 11(1), 1992, pp. 92-99
- [29]Shneiderman, B.: "The eyes have it: a task by data type taxonomy for information visualizations", in Proceedings of IEEE Symposium on Visual Languages, September 1996, pp. 336-343
- [30]Smoliar, S.W., Zhang, H., Koh, S.L., and Lu, G. J.: "Interacting with digital video", in Proceedings of IEEE Region 10's Ninth Annual International Conference, 2, August 1994, pp. 852-856
- [31]TimeTube: <http://www.dipity.com/mashups/timetube> [Última visita: Abril 22 de 2009]
- [32]Triesman, A.: "Preattentive processing in vision", in Computer Vision, Graphics, and Image Processing., 1985, pp. 156-177
- [33]VideoSurf: <http://www.videosurf.com/> [Última visita: Abril 22 de 2009]
- [34]Wang, F., and Ngo, C.W.: "Rushes video summarization by object and event understanding", in Proceedings of the international workshop on TRECVID video summarization, New York, NY, ACM, 2007, pp. 25-29
- [35]Yee, K.P., Fisher, D., Dhamija, R., and 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, 43, 2001
- [36]YouTube: <http://www.youtube.com> [Última visita: Abril 22 de 2009]
- [37]Zhang, J.: "Visualization for Information Retrieval", Vol. 23 of the Information Retrieval Series. Springer 2008.

**Carlos Mera** cursó sus estudios de Ingeniero de Sistemas en la Universidad del Valle (Colombia). Posteriormente obtuvo el título de Magister en Ingeniería de Sistemas y Computación en Universidad del Valle (Colombia) y luego realizó estudios en la Universidad de Salamanca (España), donde recibió el título de Máster en Sistemas Inteligentes. Ahora, trabaja con el grupo de laboratorio de investigación de Multimedia y Visión de la Universidad del Valle (Colombia) y labora como docente de la misma universidad en la Escuela de Ingeniería de Sistemas y Computación.

**Roberto Therón** cursó sus estudios de Informática en la Universidad de Salamanca - España (Diplomatura) y la Universidad de la Coruña - España (Licenciatura). Posteriormente ha obtenido los títulos de Licenciado en Comunicación Audiovisual (Universidad de Salamanca) y Licenciado en Humanidades (Universidad de Salamanca). En la misma Universidad de Salamanca continúa realizando su trabajo de investigador, como encargado del grupo VisUsal que se centra en la combinación de enfoques procedentes de la Informática, Estadística, Diseño Gráfico y Visualización de Información, para obtener una adecuada comprensión de conjuntos de datos complejos. En los últimos años, se ha dedicado al desarrollo de herramientas de visualización avanzada para datos multidimensionales, como por ejemplo datos genéticos o paleoclimáticos. Es autor de más de 60 artículos en revistas y congresos internacionales.

# **Universidad Nacional de Colombia Sede Medellín**

## **Facultad de Minas**



### **Misión**

Ofrecer servicios de apoyo a la docencia en cuanto a la operación de los computadores y del software adecuado con miras al desarrollo integral de los futuros ingenieros.

### **Visión**

Avanzamos en la búsqueda de convertir el Laboratorio de Sistemas e Informática en una dependencia ágil, moderna, facilitadora de procesos y cambios, atenta a las necesidades de otras dependencias de la Universidad, cuya labor apoyamos y articulamos. Serviremos con dinamismo, amabilidad y efectividad a todos los integrantes de la comunidad universitaria y a la sociedad en general. Uniremos esfuerzos para construir un ambiente de trabajo cada vez más positivo que propicie la participación, la creatividad y el desarrollo profesional de los integrantes del equipo de trabajo. Propendemos por un Laboratorio como instrumento gestor y generador de proyectos de investigación sustentado en un equipo interdisciplinario de trabajo en torno a la informática aplicada a la ingeniería.

### **Cursos**

- Lenguajes de Programación: Diseño de Páginas Web en ASP.NET con VB.NET, Programación Web PHP y MYSQL, MS-Visual Basic Básico y Avanzado, Java
- Generales: Latex, ARCGIS, MS-Office (Word, Excel y PowerPoint), Excel Financiero, Excel Avanzado, Mantenimiento de Hardware y Software Niveles I y II, MS-Access Básico, MS-Project Básico (Programación y Gerencia de Proyectos), AUTOCAD 2D Básico y 3D, Matlab, Moodle de Apoyo a los Procesos de Enseñanza y Aprendizaje.

Para mayor información, por favor comunicarse a los teléfonos: 4255313, 4255312, 4255355. Calle 59A No. 63 – 020 Medellín (Colombia), Bloque M7, quinto piso.

Email: [labsis@unalmed.edu.co](mailto:labsis@unalmed.edu.co)

<http://xue.unalmed.edu.co/cursos>

