

# *Mecanismos de Interactividad en un Ambiente de Televisión Digital y Propuesta de Interactividad Basada en Voz<sup>1</sup>*

## *Interactive Television Mechanisms in a Digital Television Environment and a Proposal of Interactivity Controlled by Voice<sup>1</sup>*

**Oscar J. Calderón C.**

*Especialista en Redes y Servicios Telemáticos  
Ingeniero en Electrónica y Telecomunicaciones  
Docente Investigador Facultad de Ingeniería Electrónica y Telecomunicaciones  
Universidad del Cauca  
Miembro IEEE  
Grupo Nuevas Tecnologías en Telecomunicaciones - GNTT  
oscaro@unicauca.edu.co*

**Laura Gómez Figueroa**

*Estudiante de Pregrado de Ingeniería en Electrónica y Telecomunicaciones  
Universidad del Cauca  
Departamento de Telecomunicaciones  
Grupo Nuevas Tecnologías en Telecomunicaciones - GNTT  
lfigueroa@unicauca.edu.co*

**Jesús Armando Sandoval León**

*Estudiante de Pregrado de Ingeniería en Electrónica y Telecomunicaciones  
Universidad del Cauca  
Departamento de Telecomunicaciones  
Grupo Nuevas Tecnologías en Telecomunicaciones - GNTT  
jarsandoval@unicauca.edu.co*

Recibido Agosto 20 de 2010 – Aceptado Diciembre 13 de 2010

---

<sup>1</sup> Documento derivado del proyecto de investigación: “Análisis del Uso de Interactividad Basada en Voz para el Manejo del Servicio de Video Bajo Demanda”, realizado por los estudiantes Laura Gómez y Armando Sandoval para optar al título de ingenieros en Electrónica y Telecomunicaciones, bajo la dirección del ingeniero Oscar J. Calderón C. El proyecto se encuentra inscrito en el Departamento de Telecomunicaciones de la Facultad de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca.

## SÍNTESIS

*Con la llegada de los nuevos servicios asociados a la televisión digital, se hace necesaria la inclusión de nuevos mecanismos de interactividad diferentes al tradicional control remoto, los cuales facilitarán el acceso a la información de manera más sencilla y eficiente.*

*Este artículo propone un análisis de las ventajas y desventajas de los principales mecanismos de interactividad empleados en un ambiente de televisión digital y a su vez plantea una solución de interactividad basada en Voz sobre IP para el manejo del servicio de video bajo demanda.*

**Descriptor:** *interactividad, video bajo demanda, televisión interactiva, control remoto, voz, interfaz.*

## ABSTRACT

Interactive television brings new services which need interaction mechanisms different from remote control for allowing effective and simple access to the information.

Therefore, this document presents the analysis about the advantages and disadvantages the main interactivity devices used in a digital television environment have. Furthermore, this paper proposes a solution of Video on Demand service controlled through Voice over IP.

**Descriptors:** *interactivity, video on demand, interactive television, remote control, voice, interface.*

## 1. INTRODUCCIÓN

La televisión es uno de los medios de comunicación y entretenimiento de mayor uso por parte de los seres humanos; a su vez, se encuentra en constante evolución incorporando nuevos conceptos y requerimientos.

Uno de los nuevos conceptos asociados a la televisión digital es la interactividad gracias a la cual, el usuario cambia su rol dejando de ser un actor pasivo para convertirse en uno activo, que influye en la forma y tiempo en que va a recibir sus contenidos de televisión.

A pesar de la innovación en los servicios de la televisión, aún se conserva el control remoto como mecanismo de interacción para el usuario final y es la base para la ejecución de mandos a distancia. Este mecanismo tradicional genera inconvenientes cuando se desea obtener contenidos de entretenimiento de alta carga interactiva, pues no es muy eficiente debido al elevado número de botones que posee, a la difícil recordación de los mismos y a que muchos de ellos no son en realidad utilizados. Es necesario por tanto, indagar acerca de nuevas formas de interacción con la televisión que permitan el acceso a los nuevos servicios de manera oportuna y simple. En este artículo se analizan los actuales mecanismos de interacción con la televisión y se describe una propuesta de interactividad basada en Voz sobre IP (VoIP, Voice over IP) para el manejo del servicio de Video bajo Demanda (VoD, Video on Demand).

## 2. TELEVISIÓN INTERACTIVA

La Corporación de Radiodifusión Británica (BBC, British Broadcasting Corporation) (Rees, 2004), define la televisión interactiva así: "...is the content and services (in addition to linear TV and radio channels) which are available for digital viewers to navigate through on their TV screen. In practice, at the moment, this means giving the viewer control over some video, audio, graphical and text elements or allowing them to use simple games and quizzes or send simple communications back to the broadcaster".

De la definición de Televisión Interactiva (iTV, Interactive Television), se destaca el control que se brinda al usuario sobre los contenidos y servicios que recibe, así como la comunicación que se puede establecer con el proveedor de contenido (Picard, 2005),

(Jensen, 1998). Además, es importante que la definición de iTV no se encuentre ligada a un prototipo de servicio, un dispositivo multimedia en particular, o a ciertas características a cumplir, evitando el riesgo de incluir servicios que no son interactivos y excluir algunos que sí lo son (Jensen, 1998).

Los beneficios de la interactividad se pueden resumir en control, personalización y selección de contenidos, a través de pantallas de navegación, intercambio de información, retroalimentación hacia el proveedor de contenido, por medio de un canal de retorno. También, se encuentra el enriquecimiento y participación en los programas y/o servicios, y crecimiento del nicho de mercado para los proveedores de servicios y contenidos (Picard, 2005), (Buchanan-Oliver & Chan, 2004).

iTV trae consigo un cambio en la forma como los contenidos se entregan (Lu, 2005), pasando de un modelo Push, en el cual el contenido entregado es estático y lineal a un modelo Pull, en el cual el televidente selecciona los servicios a los cuales desea acceder, influenciando de alguna manera el contenido que observa y puede llegar a ser el conductor del proceso que ocurre entre él y el proveedor de contenido (Picard, 2005).

Con iTV, el televidente puede participar de programas y servicios, producir y compartir contenidos propios, incluso comunicarse con otros televidentes, como si lo hiciera a través de otros medios de comunicación como Internet (Estebanell, 2002), (Jensen, 1998). Ahora bien, con la evolución de la televisión y la llegada de nuevos servicios y las nuevas características ya explicadas, es necesario el uso de mecanismos de interactividad adecuados a las nuevas necesidades de intercambio de información. A continuación se realiza un análisis de los principales mecanismos de interactividad empleados en un ambiente de televisión digital.

### **3. PRINCIPALES MECANISMOS DE INTERACTIVIDAD**

Luego de analizar la literatura existente acerca del tema, se

ha identificado el control remoto, la voz, las interfaces táctiles, multimodales y gestuales, como las principales formas de interacción con la televisión. Así mismo, se concluye que la realimentación, la eficiencia, la naturalidad, la carga cognitiva, son algunas propiedades importantes de las mismas.

La realimentación corresponde a la entrega de respuestas al usuario cuando éste ha entregado una orden (O'Modhain & Oakley, 2003), (Eriksson & Filip, 2007). La carga cognitiva es la relación entre el conocimiento y nivel de concentración necesario para realizar una tarea (de Oliveira Neto & Martins, 2008). La eficiencia es la relación entre la carga cognitiva, el número de errores y el tiempo requerido para el desempeño de una tarea (Ibrahim, Lundberg, & Johansson, 2001). La naturalidad hace referencia al uso intuitivo de una interfaz en la cual el televidente intuye el comando a entregar, sin que sea necesario recurrir al manual de usuario (Freeman & Weissman, 1994).

### **3.1 El Control Remoto**

El control remoto es la forma tradicional para la entrega de comandos a distancia empleado en la interacción con la televisión y otros electrodomésticos (Freeman & Weissman, 1994), (de Oliveira Neto & Martins, 2008). El usuario lo acepta como una extensión del cuerpo y opina que no desea aprender acerca del uso de otras interfaces, ya que se encuentra familiarizado con su uso (de Oliveira Neto & Martins, 2008). Así mismo, el televidente recibe realimentación tangible por medio de la presión ejercida sobre los botones, incrementando el grado de inmersión en el ambiente de interacción (O'Modhain & Oakley, 2003).

En relación con iTV, se considera que el control remoto no es adecuado (de Miranda, hayashi, & Baranauskas, 2009), (de Miranda, Piccolo, & Baranauskas, 2007), (Berglund, 2003), (de Oliveira Neto & Martins, 2008), debido a que son necesarios demasiados pasos para lograr acceder a un servicio (Ibrahim, Lundberg, & Johansson, 2001). Además, presenta mayor carga cognitiva como consecuencia del número de etiquetas asociadas

a los botones, las cuales se encuentran en diferentes idiomas o cambian de posición entre el diseño de un control remoto y otro (Eriksson & Filip, 2007), (de Miranda, hayashi, & Baranauskas, 2009). Finalmente, la entrada de datos para los servicios de comunicación y el acceso a diferentes puntos de la pantalla son otros problemas reportados para el control remoto (de Miranda, Piccolo, & Baranauskas, 2007), (Enns & MacKenzie, 1998).

### **3.2 La Voz**

Una interfaz por voz entrega las ordenes al servicio de televisión por medio de palabras y se reconoce como natural e intuitiva debido a que es la forma de comunicación entre los seres humanos (Freeman & Weissman, 1994), (de Oliveira Neto & Martins, 2008).

Para la interacción con iTV, la voz facilita la entrada de datos, brinda acceso a un mayor número de usuarios y es más eficiente en comparación con el control remoto (Ibrahim, Lundberg, & Johansson, 2001), (Berglund, 2003). Además, ésta requiere menos esfuerzo para la entrega de comandos, debido a que el televidente no debe conocer etiquetas en otros idiomas o hacer uso de demasiados botones, y no desvía la atención de la pantalla del televisor para la entrega de comandos (Eriksson & Filip, 2007).

Para una interfaz por voz, es recomendable emplear formas de activación/desactivación (Bhuiyan & Picking, 2009), que disminuyan la probabilidad de error en el reconocimiento de las palabras y la entrega de comandos por error que pueda causar el ambiente ruidoso en el cual se observa televisión (Berglund, 2003), (Ibrahim, 2000). También se recomiendan sistemas de ayuda que guíen al usuario, prevengan la confusión y su percepción negativa frente al servicio (Eriksson & Filip, 2007), (Ibrahim, Lundberg, & Johansson, 2001), (Telefónica I+D, 2005), (de Miranda, Hotnung, & Baranauskas, 2009).

La realimentación se debe entregar por medio de mensajes en

pantalla, ya que las instrucciones a través de comandos de voz perturban la visualización de los contenidos emitidos (de Oliveira Neto & Martins, 2008), (Ibrahim & Johansson, 2002).

Una de las restricciones de las interfaces por voz es el uso de un número bajo de palabras preestablecidas en la gramática del sistema, las que podrían incrementar la probabilidad de error de reconocimiento y limitar la interacción del usuario. Por tanto, se deben evitar las palabras homófonas y tener en cuenta que la gramática sea de uso común (de Oliveira Neto & Martins, 2008). La carga cognitiva de una interfaz por voz está condicionada al número de comandos y al tiempo de interacción, debido a que para largas sesiones el usuario puede sentir cansancio (Eriksson & Filip, 2007), (Ibrahim, 2000).

### 3.3 Interfaz Táctil

La interfaz táctil permite el acceso a la información por medio de pantallas sensibles al tacto. Éstas se emplean generalmente en computadores y dispositivos móviles sin embargo, han sido consideradas como una opción de interacción con la televisión (Telefónica I+D, 2005).

Este tipo de interfaces se asocian al control remoto, Figura 1, entregando realimentación tangible a través del tacto (O'Modhrain & Oakley, 2003). Dicha integración, reduce el número de botones en el control (Eriksson & Filip, 2007), facilita la aceptación de la interfaz (Eriksson & Filip, 2007) y evita la inserción en el ambiente de interacción de un nuevo dispositivo que deba ser manipulado (O'Modhrain & Oakley, 2003), (Enns & MacKenzie, 1998).



Figura 1. Interfaz táctil asociada al control remoto

En la interacción con iTV, la interfaz táctil facilita el alcance a cualquier punto en la pantalla, disminuyendo el número de pasos necesarios para completar una tarea (Malik & Laszlo, 2004).

Con una interfaz táctil la carga cognitiva se reduce en comparación con los otros mecanismos de interacción, debido a que el usuario no debe aprender comandos y de forma natural deduce la parte de la pantalla que debe presionar (de Miranda, Hotnung, & Baranauskas, 2009). No obstante, es posible que en algunos casos este tipo de interfaces requieran la lectura de manuales para configurarlas y utilizarlas correctamente (de Miranda, Piccolo, & Baranauskas, 2007).

Dada la distancia a la cual se encuentra el televidente del televisor, no se recomienda integrar la interfaz táctil a la pantalla del televisor, ya que puede ocasionarse cansancio en el mismo durante la entrega de instrucciones (de Miranda, Hotnung, & Baranauskas, 2009).

### **3.4 Interfaz Gestual**

Los gestos son un tipo de comunicación no verbal realizada con partes del cuerpo. Éstos han sido pensados en la interacción con la televisión (Eriksson & Filip, 2007), porque el reconocimiento se realiza a través de cámaras y no es necesaria la adición de un dispositivo externo que deba ser manipulado (Bhuiyan & Picking, 2009).

Como consecuencia del uso de cámaras, es posible que el televidente sienta pérdida de privacidad, de esta manera algunos sistemas centran el reconocimiento en algunas partes del cuerpo o del ambiente de entretenimiento, como las manos o las mesas de centro (Vatavu & Pentiuc, 2008).

Este tipo de interfaces posibilitan la manipulación directa de los objetos en pantalla, permitiendo la entrega de realimentación visual al televidente durante la interacción (Eriksson & Filip, 2007), (Ahn, Lee, Kim, & Kwon, 2004).



Para hacer uso de una interfaz gestual, los usuarios deben aprender la forma de los comandos a entregar, convirtiéndola en un mecanismo de interacción poco natural e intuitivo, generando confusión, desorientación y carga cognitiva (Freeman & Weissman, 1994), (de Miranda, Hotnung, & Baranauskas, 2009), (Bhuiyan & Picking, 2009). Se considera poco eficiente, a causa del uso de un gran número de comandos necesarios para completar una tarea (Eriksson & Filip, 2007).

En la interacción con la televisión a través de los gestos, se deben evitar largas sesiones para evitar causar cansancio muscular (Freeman & Weissman, 1994), (Vatavu & Pentiu, 2008). Tampoco es recomendable un ambiente de entretenimiento cargado de elementos que genere alta carga de procesamiento y problemas de reconocimiento (Eriksson & Filip, 2007), (Vatavu & Pentiu, 2008), (Chen, Mummert, Pillai, Hauptmann, & Sukthankar, 2010). La interfaz táctil es el único mecanismo de interacción que permite la entrega de comandos desde varios usuarios en el mismo instante de tiempo (Vatavu & Pentiu, 2008).

### **3.5 Interfaces Multimodales**

Por último, se encuentran las interfaces multimodales que emplean diversas maneras de suministrar y obtener información de un servicio, brindándole al usuario la posibilidad de elegir la forma de completar una tarea (Telefónica I+D, 2005), (Ibrahim & Johansson, 2002). Éstas incrementan la eficiencia y reducen la probabilidad de error (Ibrahim, 2000), porque las deficiencias de una interfaz son compensadas con las ventajas de otras (Telefónica I+D, 2005).

Este tipo de mecanismos de interacción, requieren altos niveles de procesamiento, sincronismo y coordinación, como consecuencia del incremento en el número de comandos a reconocer desde diferentes interfaces de entrada (Telefónica I+D, 2005).

En relación con iTV, las interfaces multimodales permiten la interacción con los nuevos servicios de manera exitosa

(Stegmann, Henke, & Kirchherr, 2008), Así mismo, permiten que todos los usuarios puedan acceder a la información (Instituto Nacional de Tecnologías de la Comunicación, 2009).

Si el televidente no ha tenido experiencia previa con las interfaces multimodales, puede ser necesario, el aprendizaje por medio de ensayo y error (Stegmann, Henke, & Kirchherr, 2008), generando quizás una experiencia poco satisfactoria (de Oliveira Neto & Martins, 2008).

#### **4. PROPUESTA DE UNA SOLUCIÓN DE INTERACTIVIDAD BASADA EN VOZ**

Con el objetivo de analizar y evaluar desde la perspectiva objetiva y subjetiva la viabilidad de una interfaz por voz en un ambiente de televisión interactiva, específicamente para el manejo de un servicio de video bajo demanda y determinar si ésta puede suplir los problemas reportados para el control remoto, se propone una solución de interactividad basada en Voz sobre IP, la cual permitirá obtener la opinión de los usuarios acerca de este tipo de interfaces, y el comportamiento de la red frente a un sistema de este tipo.

Con la construcción de una interfaz por voz, se ofrecerá a los televidentes otra alternativa para la entrega de mandos a distancia que, posiblemente, permitirá el acceso a la información por parte de algunos usuarios que actualmente no pueden hacerlo a través del control remoto. Así mismo, se considera que por ser una solución novedosa, la voz contribuirá a atraer la atención de aquellos televidentes que se encuentran reacios a los nuevos servicios interactivos.

La solución que se propone en este artículo presenta un diseño general de una interfaz basada en voz para el manejo de un servicio de video bajo demanda, sin embargo la integración de otros módulos es posible, permitiendo el acceso a otros servicios interactivos. El diseño modular que presenta la propuesta posibilita la elección y cambio de herramientas de cada módulo de manera independiente, rápida y sencilla.

La Figura 2, muestra el esquema general de la solución, que cuenta con 5 componentes, denominados, módulo de servidor de video streaming, de servidor web, de reconocimiento del habla, cliente multimedia y de VoIP.

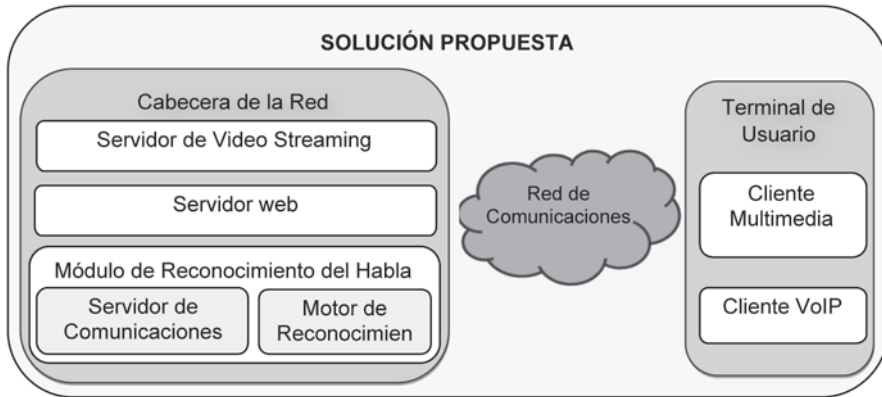


Figura 2. Esquema modular de una solución de interactividad

Para el servicio de video bajo demanda son necesarios el servidor de video, web y el cliente multimedia. Para la interfaz de reconocimiento del habla son necesarios el servidor de comunicaciones IP, el motor de reconocimiento y el cliente VoIP.

El módulo de servidor de video streaming se encarga de enviar los flujos de los contenidos multimedia, siendo parte fundamental del servicio de video bajo demanda. El servidor web quien a su vez tiene la posibilidad de ser reemplazado por un servidor de aplicaciones y cumple con la función el control en el establecimiento de la sesión entre el servidor de video streaming y el reproductor de video; además debe almacenar la guía de contenidos por medio de la cual el usuario accede a los videos disponibles.

El módulo de reconocimiento del habla integra el servidor de comunicaciones IP, que comunica el cliente VoIP con el motor de reconocimiento. Estos elementos en conjunto permiten el reconocimiento de los comandos de voz, y la entrega de los mismos en el formato adecuado para que puedan ser ejecutados.

El motor de reconocimiento hace uso de gramáticas, las cuales contienen las palabras que el sistema está en condición de reconocer, de esta manera durante la construcción de la gramática, se debe diseñar una interfaz gráfica simple que precise de pocos comando para evitar causar carga cognitiva. Gracias al módulo de reconocimiento del habla y el uso de un motor de reconocimiento sin entrenamiento, la solución es independiente del usuario, garantizando el acceso de cualquier televidente.

El cliente VoIP es el encargado de obtener de la voz y enviarla al servidor de comunicaciones de VoIP. Éste se encuentra a cargo de un softphone o un teléfono IP.

El cliente multimedia contiene la Interfaz Gráfica de Usuario (GUI, Graphical User Interface) y el reproductor multimedia. La interfaz gráfica muestra de forma organizada los contenidos multimedia disponibles. El reproductor es necesario para mostrar en pantalla los videos transmitidos bajo streaming desde el servidor de video. Se recomienda que la INTERFAZ GRÁFICA de la solución se implemente por medio de una página web, debido a que es la forma en que los Set Top Box (STB) permiten el acceso a los servicios interactivos.

Con el objetivo de proponer una solución cercana a la realidad, con la opción de ser implementada en un escenario real de servicio de video bajo demanda, en la cabecera de red se encuentran el servidor de video streaming, el servidor web y el bloque de reconocimiento del habla, los cuales realizan procesamiento de manera centralizada atendiendo todos los usuarios que deseen acceder al servicio y disminuyendo los requerimientos necesarios en el terminal de usuario.

En el terminal de usuario reside el cliente VoIP y multimedia, de esta manera es necesario contar con un STB y un teléfono IP o con un computador que permita emularlos.

La red de comunicaciones conecta los servidores con los terminales de usuario por medio del protocolo IP. Es importante

que ésta cuente con una velocidad de transmisión mínima con la cual se disminuya el retardo en la entrega de los contenidos multimedia y no se afecte el transporte de la voz.

El proceso interactivo general para la solución que se propone será el siguiente:

1. El usuario accederá a la guía de contenidos que se muestra en pantalla y solicitará al servidor web desde el cliente multimedia, para elegir un video y observarlo.
2. El cliente VoIP obtendrá los comandos entregados por el televidente durante la navegación en la guía de contenidos o para el control del video y los enviará al servidor de comunicaciones que a su vez, los entregará al motor de reconocimiento.
3. Si el televidente ha solicitado cambiar de categoría u obtener más información acerca de un contenido en particular, el servidor web responderá a dicha solicitud.
4. Si el usuario ha solicitado un video, el servidor de video streaming responderá enviando el flujo de video.
5. Finalmente, durante la reproducción del contenido multimedia, el usuario podrá pausar, reproducir, y detener el video a través de comandos enviados al servidor de video streaming mediante previo reconocimiento de los mismos.

Con la explicación del proceso interactivo termina la descripción general de la solución que se propone.

## **5. EVALUACIÓN DE LA SOLUCIÓN POR PARTE DE LOS USUARIOS**

Con el objetivo de evaluar la solución propuesta y obtener un acercamiento inicial a la opinión de los usuarios acerca del control por voz para un servicio de VoD, dentro de esta sección se exponen los resultados obtenidos luego de la evaluación de la misma.

## 5.1 Escenario de Evaluación

Para llevar a cabo la evaluación, se utilizó del salón 448 de la Facultad de Ingeniería Electrónica y Telecomunicaciones (FIET) de la Universidad del Cauca, el cual está ubicado dentro en una zona de baja afluencia de personal y es un espacio con bajo ruido, lo que favoreció la evaluación. Así mismo, se realizó el montaje mostrado en la Figura 3. En el equipo servidor se albergó el servidor de video, el servidor web y el módulo de reconocimiento del habla. El cliente multimedia estuvo comprendido por un televisor conectado a un computador portátil, que en conjunto, recibieron los flujos de video, los decodificaron y los mostraron en pantalla. El cliente VoIP estuvo a cargo de un softphone. La comunicación entre el servidor y los elementos del terminal de usuario se realizó por medio de la Red de Área Local (LAN, Local Area Network) de la FIET.

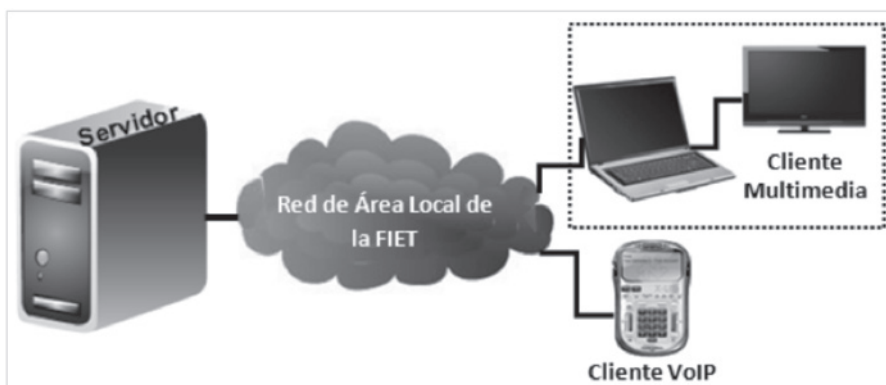


Figura 3. Elementos utilizados en la evaluación

La población que evaluó la solución estuvo comprendida por 7 mujeres y 14 hombres (Nielsen, 2000), quienes respondieron las siguientes preguntas:

1. ¿El sistema le pareció agradable?
2. ¿Disfrutó su sesión con el sistema?
3. ¿El sistema fue fácil de manejar?
4. En general, ¿cree que el control por voz para el servicio de Video bajo Demanda es bueno?

5. ¿Cuál cree que es el mayor problema del sistema?
6. ¿Qué expectativas tenía de la prueba del servicio?

## 5.2 Resultados Obtenidos

Después de aplicada la evaluación, se encontró que el 85.71% de los usuarios consideran agradable el sistema, el 4,76% cree que no lo es y el 9,52% no opinó. Así mismo, el 80.95% de la población disfrutó la sesión con el sistema, porcentaje bastante alto en comparación con el 4,76% que presentó una opinión desfavorable y el 14,29% que se encontraba indeciso.

En relación con la tercera pregunta, el 71,43% de los usuarios opinaron que el sistema es fácil de manejar, el 4,76% manifestó una opinión desfavorable, y el 23,81% estaba indeciso.

Luego de la interacción con el sistema, todos los usuarios consideraron que el sistema de control por voz es bueno; entre las razones más importantes están: lo intuitivo de la interfaz debido a que es posible utilizar diferentes comandos para realizar la misma acción; la facilidad de uso y el poco esfuerzo para lograr encontrar un video y observarlo. Finalmente, el hecho de descartar el control remoto como dispositivo de entrega de mandos a distancia, lo que implica tener que buscarlo cada vez que se deseaba interactuar con el servicio.

Los principales problemas del sistema, de acuerdo con los usuarios fueron la lenta ejecución de los comandos posterior a la entrega de una orden y algunos errores en el reconocimiento del habla. Además, sería importante incrementar el número de comandos disponibles para una misma acción y que se tengan en cuenta los vocablos propios de cada región para la inclusión de las palabras que el sistema está en capacidad de reconocer. Por último, el 85,71% de la población opinó que el sistema cumplió con sus expectativas, ya que es fácil de manejar, novedoso y original, respondió a los comandos existentes de forma adecuada y porque se podían realizar otras acciones mientras se interactuaba con el sistema sin necesidad de hacer

uso de las manos y del control remoto. El 14% de los usuarios consideró que aún existen factores que se deben mejorar, tal es el caso del número de comandos disponibles actualmente y la ayuda en pantalla acerca de las opciones que se pueden ejecutar para el manejo del servicio.

## **6. CONCLUSIONES**

Son variados los esfuerzos realizados para generar interfaces acordes a las necesidades de intercambio de información en los nuevos servicios interactivos, ya que se desea ofrecer al televidente una interfaz amigable, intuitiva y de baja carga cognitiva. No obstante, todavía se perciben deficiencias y hay mucho por mejorar en el diseño de las mismas.

Las interfaces multimodales se presentan como la mejor alternativa de interacción con un servicio de entretenimiento, debido a que brindan un acceso universal y mejoran la interacción del televidente con el servicio, pues las deficiencias de una interfaz se compensan con las ventajas de otras.

Gracias a que la solución cuenta con un diseño centralizado y divide las funciones entre los servidores y los terminales de usuario, ésta será de fácil administración e integrable de manera sencilla a un escenario real.

Cuando se diseñe un mecanismo de interacción como el propuesto en este artículo, se recomienda hacer un análisis de las opciones disponibles actualmente, con el fin de identificar las falencias de los mismos y evitar generar una solución con los mismos problemas.

A partir de la solución presentada, es posible generar diversas soluciones particulares adaptadas a un servicio o herramienta en particular gracias al diseño modular de la misma, el cual posibilita el intercambio de diferentes elementos sin afectar la solución en su totalidad.



Considerando las ventajas y desventajas para una interfaz por voz, se recomienda que durante la implementación de la solución se ofrezcan mecanismos de activación y desactivación de la interfaz de reconocimiento de voz.

Con base en la solución de interactividad por Voz sobre IP, será posible realizar procesos de evaluación que permitan determinar si la voz puede llegar a ser una opción adecuada para el manejo del servicio de video bajo demanda.

Durante la construcción de un nuevo mecanismo de interacción en un ambiente de televisión, es importante la inclusión de los usuarios durante la etapa de diseño del mismo, puesto que a partir de la realimentación se pueden encontrar requisitos no considerados previamente y algunos problemas durante la etapa de interacción.

Una solución de interactividad basada en voz para el servicio de VoD es apropiada y podría tener buena aceptación por parte de los usuarios finales, siempre y cuando se mejore la velocidad de respuesta y se disminuya la ocurrencia de errores repentinos durante el reconocimiento del habla.

Uno de los principales factores que influyó de manera positiva la opinión de los usuarios acerca de la solución propuesta fue lo innovador de un servicio controlado por voz y la comodidad que ofrece una interfaz de este tipo.

## BIBLIOGRAFÍA

- Ahn, S. C., Lee, T. S., Kim, I. J., & Kwon, Y. M. (2004). ***Large Display Interaction Using Video Avatar and Hand Gesture Recognition***. International Conference on Image Analysis and Recognition (ICIAR), (págs. 261-268). Porto.
- Berglund, A. (2003). ***Future TV Interaction: Augmenting Existing Interaction Techniques***. Linköping University, Department of Computer and Information Science, Linköping.

- Bhuiyan, M., & Picking, R. (2009). **Gesture-controlled User Interfaces, What Have We Done and What's Next?** 5th Collaborative Research Symposium on Security, E-Learning, Internet and Networking (SEIN), (págs. 59-60). Darmstadt.
- Buchanan-Oliver, M., & Chan, N. (2004). **Defining Interactivity: Building Relationships Online.** 20th Industrial Marketing and Purchasing Conference. Copenague.
- Chen, M. y., Mummert, L., Pillai, P., Hauptmann, A., & Sukthankar, R. (2010). **Controlling your TV with Gestures.** 11th ACM SIGMM International Conference on Multimedia Information Retrieval, (págs. 405-408). Philadelphia.
- de Miranda, L. C., hayashi, E. C., & Baranauskas, M. C. (2009). **An Experimental Scenario to Investigate the Remote Control as Artifact for Interaction with Television.** Reporte Técnico, Instituto de Computação, Campinas.
- de Miranda, L. C., Hotnung, H. H., & Baranauskas, M. C. (2009). **MULTIS: A Gesture Based Interaction Model for iDTV.** Reporte Técnico, Universidade Estadual de Campinas, Instituto de Computação, Campinas.
- de Miranda, L. C., Piccolo, L. S., & Baranauskas, M. C. (2007). **Uma Proposta de Taxonomia e Recomendação de Utilização de Artefatos Físicos de Interação com a TVDI.** Workshop Perspectives, Challenges and Opportunities Human-Computer Interaction Latin America (CLIHC). Rio de Janeiro.
- de Oliveira Neto, J. S., & Martins, V. F. (2008). **Non Functional Requeriment to Voice User Interface on Interactive Television> An initial Study.** IADIS International Conference Interfaces and Human Computer Interaction. Amsterdam.
- Enns, N., & MacKenzie, S. (1998). **Touchpad-Based Remote Control Devices.** ACM SIGCHI Conference on Human Factors in Computing Systems "Making the Impossible Possible". Los Angeles.
- Eriksson, R., & Filip, S. (2007). **Enhancing the User Experience with New Interaction Techniques for**

**Interactive Television.** Tesis de Maestría, Umeå University, Umeå.

- Estebanell, M. (2002). **Interactividad e Interacción.** Revista Latinoamericana de Tecnología Educativa , I (1), 11-18.
- Freeman, W., & Weissman, C. (1994). **Television Control by Hand Gestures.** Reporte Técnico, Mitsubishi Electric Research Laboratories, Cambridge, Cambridge.
- Ibrahim, A. (2000). **Multimodal Interaction with Interactive TV.** Linköping University, Department of Computer and Information Science, Linköping.
- Ibrahim, A., & Johansson, P. (2002). **Multimodal Dialogue Systems for Interactive TV Applications.** 4th IEEE International Conference on Multimodal Interfaces. Pittsburgh.
- Ibrahim, A., Lundberg, J., & Johansson, J. (2001). **Speech Enhanced Remote Control for Media Terminal.** EUROSPEECH: 7th European Conference on Speech Communication and Technology.
- Instituto Nacional de Tecnologías de la Comunicación. (2009). **Guía de Recomendaciones de Accesibilidad en TDT.** Instituto Nacional de Tecnologías de la Comunicación (INTECO).
- Jensen, J. (1998). **'Interactivity': Tracking a New Concept in Media and Communication Studies.** Nordicom Review , XIX (1), 185-204.
- Lu, K. (2005). **Interaction Design Principles for Interactive Television.** Tesis de Maestría, Georgia Institute of Technology, Atlanta.
- Malik, S., & Laszlo, J. (2004). **Visual Touchpad: A Two-handed Gestural Input Device.** 6th International Conference on Multimodal interfaces (ICMI), (págs. 289-296.). State College.
- Nielsen, J. (19 de Marzo de 2000). **Why You Only Need to Test with 5 Users.** Recuperado el 15 de Junio de 2010, de <http://www.useit.com/alertbox/20000319.html>
- O'Modhrain, S., & Oakley, I. (2003). **Touch TV: Adding**

***Feeling to Broadcast Media.*** 1st European Conference in Interactive Television: Viewers Actors. Brighton.

- Picard, R. (2005). ***A Consumer Perspective on Digital Terrestrial and Interactive Television.*** En Digital Terrestrial Television in Europe (págs. 135-149). New Jersey: Lawrence Erlbaum Associates.
- Rees, C. (2004). ***BBC Enhanced TV Formats.*** Londres: British Broadcasting Corporation.
- Stegmann, J., Henke, K., & Kirchherr, R. (2008). ***Multimodal Interaction for Access to Media Content.*** 12th International Conference ICIN. Bordeaux.
- Telefónica I+D. (2005). ***Interfaces multimodales. En Las Telecomunicaciones y la Movilidad en la Sociedad de la Información (pág. 289/304).*** Fundación Telefónica.
- Vatavu, R. D., & Pentiu, S. G. (2008). ***Interactive Coffee Tables: Interfacing TV within an Intuitive, Fun and Shared Experience.*** 6th European Conf. Changing Television Environments, (págs. 183–187). Salzburg.