



UNIVERSIDAD
POLITECNICA
DE VALENCIA



Máster Universitario
en Tecnologías, Sistemas y
Redes de Comunicaciones

Desarrollo de plataforma web para reproducción de vídeos de YouTube con aromas sincronizados y evaluación subjetiva.

Autor: Yherson Gabriel Ortiz Villafañe

Directores:

Dr. Fernando Boronat Seguí

Dr. Mario Alberto Montagut Climent

Fecha de inicio: Marzo 2014

Lugar de trabajo: Departamento de Comunicaciones de la Escuela Politécnica Superior de
Gandía

Objetivos

El objetivo principal de este Trabajo Final de Máster (TFM) ha sido desarrollar una plataforma web para la visualización de contenidos multimedia, agregando efectos multi-sensoriales de tipo olfativo, controlados mediante etiquetas de sincronización de aromas con respecto al vídeo, y la realización de evaluaciones subjetivas con usuarios para comprobar los beneficios aportados. Este objetivo principal, a su vez, se puede dividir en los siguientes sub-objetivos asociados:

- Diseño y desarrollo de la plataforma web, utilizando HTML, JavaScript, PHP y la integración de APIs.
- Diseñar un software de acople y control para el dispositivo hardware dispensador de aromas.
- Estudiar y clasificar los diferentes tipos de contenidos, complementarios a los de audio y vídeo, susceptibles de ser incluidos en sistemas multimedia, como datos de tipo textual (p.ej., subtítulos), datos de aplicaciones de Social Media (ej. twitter), datos con contenido de tipo sensorial (p.ej., aromas), así como objetos obtenidos vía web (p.ej., widgets, scripts, ficheros HTML...).
- Incluir en un prototipo una experiencia multi-sensorial, incluyendo aromas. Se definirá un protocolo para insertar, codificar, etiquetar, transmitir, interpretar y generar información (p.ej., identificación, instante inicial, duración de emisión, intensidad, persistencia, combinaciones posibles...) sobre los aromas, de manera sincronizada con los flujos audiovisuales.
- Diseño de varios clips (en concreto 7) multimedia que contenga la información de aromas deseada.
- Definir los parámetros que se evaluarán de forma subjetiva, realización de pruebas a usuarios, recopilación de respuestas y posterior análisis estadístico.
- Analizar los resultados para comprobar el incremento de inmersividad y realismo debido a la inclusión de efectos olfativos cuando se consumen contenidos audiovisuales.

Metodología

- Estudio y clasificación de los diferentes tipos de contenidos, complementarios a los de audio y vídeo, susceptibles de ser incluidos en sistemas multimedia.
- Revisión del estado del arte, incluyendo estándares y estudios científicos relacionados.
- Diseño de un prototipo de sistema multimedia que incluya efectos olfativos.- Determinación del hardware a usar y desarrollo de software basado en su propio SDK.
- Selección de los componentes web y APIs más adecuadas para la implementación de la plataforma.
- Diseño de clips multimedia, incluyendo los elementos visuales correspondientes a los aromas que se pretenden incluir.
- Diseño y planteamiento del tipo de pruebas de evaluación a realizar. Delimitación y alcances de las mismas.
- Documentación y análisis de los resultados obtenidos de usuarios extrayendo conclusiones de los mismos.

Desarrollos teóricos realizados

Para este trabajo y su respectivo desarrollo se analizó todo lo referente a estudios e investigaciones llevados a cabo previamente, para analizar las bases, los parámetros de evaluación empleados en los mismos, así como para determinar qué parámetros complementarios se podrían añadir con el objetivo de obtener nuevos resultados.



Se ha especificado un protocolo para insertar, codificar, etiquetar, transmitir, interpretar y generar información (p.ej., identificación, instante inicial, duración de emisión, intensidad, persistencia, combinaciones posibles...) sobre los aromas, de manera sincronizada con los flujos audiovisuales.

Desarrollo de prototipos y trabajo de laboratorio

Se ha diseñado una plataforma web (mediante HTML5, JavaScript y PHP), que integra el control del dispositivo hardware de emisión de aromas (modelo SBix de la empresa Exhalia) y permite el control de la visualización de vídeos y de la emisión de aromas de los mismos de forma sincronizada en los instantes en los que se reproduce información audiovisual relacionada.

Resultados

Se ha obtenido, por un lado, una plataforma para la visualización de vídeos de YouTube con efectos olfativos, codificados mediante etiquetas (metadatos). Cualquier usuario que disponga del dispensador de aromas, puede utilizar este prototipo mulsemmedia, tan solo accediendo a la URL de la plataforma web. Asimismo, mediante esta plataforma se puede evaluar el incremento de inmersividad de los usuarios debido a la inclusión de efectos olfativos en los vídeos y, además, los efectos que provocan en los usuarios diferentes grados de asincronía en el momento del inicio real de la emisión de aromas con respecto al momento en el que debería iniciarse, así como cambios en la duración de las emisiones, separación temporal adecuada entre dos emisiones consecutivas, etc. Con todo ellos se pueden evaluar parámetros subjetivos, como grado de inmersividad, de molestia, sensación de realidad, distracción generada por los efectos multi-sensoriales, etc. En este TFMse presentan y analizan los resultados de las evaluaciones subjetivas realizadas a 44 usuarios.

Líneas futuras

- Analizar más patrones y variables en las pruebas para lograr resultados complementarios.
- Mejorar la interfaz de usuario de la plataforma con el fin de proporcionar una mayor inmersividad a los usuarios.
- Incluir más efectos mulsemmedia, además de los efectos olfativos, como, por ejemplo, táctiles (haptics), gusto, sensitivas a nivel físico (frío, calor), vibración, etc.
- Extender el protocolo de etiquetado de aromas, incluyendo parámetros como la intensidad, persistencia, la combinación de aromas, para intentar ofrecer una mejor calidad de experiencia (QoE), así como incrementar las variables de análisis.

Publicaciones

Los resultados del TFM se intentarán publicar más adelante en conferencias y/o revista de esta área de investigación.

Abstract

The concept of Mulsemmedia (multi-sensory media) covering new multi-sensory experiences, such as taste, touch or smell, is gaining momentum in multimedia communications. Mulsemmedia effects including generates immersivity greater in users of multimedia applications. The measure to achieve quality of experience (QoE) of users are added when these effects is still a challenge.

This TFM focuses on synchronization (flavorings) in olfactory sensory effects audiovisual content. For this we have developed a web platform for playing YouTube videos, including metadata about olfactory multisensory effects, according to a mechanism defined in TFM. In addition, it has been subjectively evaluated by tests of users (including 44 participants) increased its immersivity and the highest degree of realism due to the inclusion of such olfactory effects in videos. Also, they have also evaluated the effects they have on users varying degrees of asynchrony at the time of the actual start of the emission of aromas with respect to the time when it should start and forcing changes in the length of emissions, separation adequate time between two consecutive releases, etc.

Resumen

El concepto de Mulsemmedia (multi-sensorial media), que abarca nuevas experiencias multisensoriales, como el gusto, tacto o el olfato, está cobrando fuerza en las comunicaciones multimedia. La inclusión de efectos mulsemmedia genera un mayor grado de inmersividad en los usuarios de aplicaciones multimedia. El poder lograr medir la calidad de experiencia (QoE) de los usuarios cuando se añaden estos efectos todavía es un reto.

El presente TFM se centra en la sincronización de efectos sensoriales olfativos (aromas) en contenidos audiovisuales. Para ello se ha desarrollado una plataforma web para la reproducción de vídeos de YouTube, incluyendo metadatos sobre efectos multisensoriales olfativos, siguiendo un mecanismo definido en el TFM. Además, se ha evaluado de forma subjetiva mediante tests de usuarios (incluyendo 44 participantes) el incremento de su inmersividad y el mayor grado de realismo debido a la inclusión de dichos efectos olfativos en los vídeos. Asimismo, también se han evaluado los efectos que provocan en los usuarios diferentes grados de asincronía en el momento del inicio real de la emisión de aromas con respecto al momento en el que debería iniciarse, así como forzando cambios en la duración de las emisiones, separación temporal adecuada entre dos emisiones consecutivas, etc.

Autor: Ortiz Villafañe Yherson Gabriel, email: yheorvil@teleco.upv.es

Directores: Dr. Boronat seguí Fernando, email: fboronat@dcom.upv.es

Dr. Montagut Climent Mario, email: mamontor@posgrado.upv.es

Fecha de entrega: 7-09-2015

Agradecimientos

Primero que todo, quiero darle muy sinceras gracias a mis Directores, Fernando Boronat Seguí y Mario Montagut Climent, por todo el aporte y confianza depositada en mí, para poder realizar este proyecto. También quiero agradecer a mis compañeros del grupo de investigación, Immersive Interactive Media R&D de la EPSG, toda la ayuda brindada en este trabajo, mención especial al profesor Francisco Javier Pastor Castillo y a mi compañero Jordi Belda Valls por sus aportes fundamentales en el desarrollo del mismo.

A todos ellos, muchas gracias.

TABLA DE CONTENIDO

CAPÍTULO 1. INTRODUCCIÓN Y OBJETIVOS	6
1.1. Introducción	6
1.2. Objetivos y sub-objetivos	6
1.3. Estructura de la memoria	7
CAPÍTULO 2. ESTADO DEL ARTE Y ESTÁNDARES RELACIONADOS	7
2.1 Estado del arte.....	7
2.2 Estándares relacionados	17
2.2.1 MPEG- V	18
2.2.2 MPEG-7.....	19
CAPÍTULO 3. DESARROLLO DEL PROYECTO	20
3.1. Hardware difusor de aromas	20
3.2. Protocolo de etiquetado y codificación de los metadatos sobre aromas.....	21
3.3. Vídeos en YouTube.....	22
3.3.1 Lectura de las etiquetas con metadatos en YouTube	23
3.3.2. YouTube Data API v3	24
3.4 Desarrollo e integración de software.....	24
3.4.1 Funcionalidades mediante el uso de de JavaScript	24
3.4.2 Funcionalidades mediante el uso de PHP	24
CAPÍTULO 4. EVALUACIÓN SUBJETIVA	26
4.1 Escenario de realización de las pruebas.....	26
4.2 Metodología y Descripción de las pruebas.....	27
4.3 Prototipo desarrollado para la realización de las pruebas.....	28
4.4. Objetivos de la evaluación.....	30
4.5. Resultados.....	31
CAPÍTULO 5. CONCLUSIONES Y TRABAJO FUTURO	37
5.1 Conclusiones	37
5.2 Líneas futuras.....	37
REFERENCIAS	38

CAPÍTULO 1. INTRODUCCIÓN Y OBJETIVOS

1.1. Introducción

En la actualidad, encontramos muchas opciones tecnológicas basadas principalmente en mejorar la experiencia de los usuarios cuando consumen contenidos multimedia, ofreciéndoles un mayor grado de realidad y/o inmersividad. En muchas de ellas van implícitas sensaciones multi-sensoriales (*mulsemedia* o *multi-sensorial media*). Algunos casos puntuales son, por ejemplo, el cine interactivo en 4D o los videojuegos con plataformas para generación de realidad virtual, entre otros. Todas ellas han logrado despertar el interés de la comunidad investigadora para explorar y comprender la influencia de todo tipo de estímulos sensoriales, como, por ejemplo, el olfativo, el gusto, el tacto, o también las sensaciones de humedad y temperatura, entre otros. Se intenta determinar su impacto sobre la calidad de la experiencia (Quality of Experience, QoE) percibida por los usuarios, las sensaciones que generan, así como el grado de inmersividad y realismo que se consigue.

Este TFM se centra en la inclusión de efectos olfativos (aromas) en aplicaciones multimedia audiovisuales. Cabe destacar la importancia del sentido del olfato, el cual, comparado con otros sentidos, es 10.000 veces más sensible.

En el TFM se pretende analizar, mediante evaluaciones subjetivas, el impacto sobre los usuarios cuando se incluyen nuevas experiencias multi-sensoriales de tipo olfativas en los contenidos multimedia tradicionales. Para lograrlo, en primer lugar, se debe desarrollar una plataforma web que proporcione los mecanismos de visualización y sincronización de contenidos audiovisuales, en este caso vídeos de Youtube, con información sobre aromas asociados a los elementos visuales que aparecen en los vídeos.

1.2. Objetivos y sub-objetivos

El objetivo principal es desarrollar una plataforma web para visualización de contenidos multimedia, agregando efectos multisensoriales de tipo olfativo, controlados mediante metadatos de generación y sincronización de aromas con respecto al vídeo. Estos metadatos se corresponden con etiquetas, incluyendo información de inicio y duración de las difusiones que, a su vez, también servirán para generar las asincronías entre los aromas y el vídeo en los procesos de evaluación.

Dicha plataforma se utilizará para la realización de evaluaciones subjetivas en usuarios.

Para conseguir este objetivo, se planificaron varios sub-objetivos asociados:

- Programación de la plataforma mediante HTML, JavaScript, PHP y el manejo e integración de APIs.
- Diseño un software de acople y control para el dispositivo hardware dispensador de aromas.

- Se estudió y clasificó los diferentes tipos de contenidos, complementarios a los de audio y vídeo, susceptibles de ser incluidos en sistemas multimedia, como datos de tipo textual (p.ej., subtítulos), datos de aplicaciones de Social Media (ej. twitter), datos con contenido de tipo sensorial (p.ej., aromas), así como objetos obtenidos vía web (p.ej., widgets, scripts, ficheros HTML...).
- Diseño de protocolo para insertar, codificar, etiquetar, transmitir, interpretar y generar información (p.ej., identificación, instante inicial, duración de emisión, intensidad, persistencia, combinaciones posibles...) sobre los aromas, de manera sincronizada con los flujos audiovisuales. Estas etiquetas se insertan como metadatos de los vídeos de Youtube
- Selección de los parámetros de evaluación subjetiva más relevantes, para la realización de pruebas a usuarios, y la recopilación de respuestas para su posterior análisis.
- Análisis de los resultados obtenidos.

1.3. Estructura de la memoria

En el Capítulo 2, se presenta todo lo relacionado con el estudio de estándares y de trabajos científicos existentes y relacionados con la temática del TFM. A continuación, en el Capítulo 3 se describe el proceso de desarrollo del TFM, abarcando todo lo referente a la integración del hardware usando al sistema multi-sensorial, la implementación de la plataforma web, los mecanismos de codificación, sincronización de datos olfativos con contenidos multimedia, el uso de la API de Youtube, etc. En el Capítulo 4 se incluyen la metodología y los resultados de las evaluaciones subjetivas. El TFM finaliza con las conclusiones y líneas de trabajo futuras en el Capítulo 5.

CAPÍTULO 2. ESTADO DEL ARTE Y ESTÁNDARES RELACIONADOS

2.1 Estado del arte

Cuando se habla del concepto multimedia, se hace referencia al uso e integración de 2 o más elementos de comunicación siendo al menos uno de ellos dependiente del tiempo. Tradicionalmente los sistemas multimedia han sido audiovisuales, es decir, incluyen principalmente audio y vídeo. Esta situación está en contradicción con el hecho de que el 60% de la comunicación humana no se hace de manera bi-sensorial, sino que la mayoría de nosotros percibimos el mundo a través de una combinación de cinco sentidos (la vista, el oído, el tacto, el gusto y olfato). En este aspecto, se pretende incluir las sensaciones multi-sensoriales, en las aplicaciones multimedia, resultando en el concepto *mulsemedia*. Aunque suene innovador, ya se trataron de introducir los efectos mulsemedia décadas atrás. Un caso en particular del que se tiene conocimiento, sucedió en 1943, cuando Hans Laube, perfeccionó la técnica de concentración de aromas y la aplicó en un sala de cine, utilizando el sistema de aire acondicionado, para expulsar 35

olores artificiales, en momentos específicos de una película. Entre los años 1959 y 1960 otras personas siguieron intentando hacer este tipo de cine. Fue el caso de Michael Todd Jr, que creó un sistema de difusión llamado *Smell-O- Vision*-. Estas innovaciones tuvieron buena acogida inicialmente, pero fueron quedando en el olvido.

En la actualidad existen numerosos trabajos y artículos relacionados con el tema principal de este TFM (mulsemmedia), en los que se añaden efectos multisensoriales a los contenidos multimedia y se evalúa de forma subjetiva la respuesta de los usuarios. A continuación, se exponen, de manera resumida, varios de los trabajos que se han llevado a cabo en el pasado y que han servido como base para el TFM.

Un claro ejemplo fue el trabajo *“Reproduction of scent and video at remote site using odor sensing system and olfactory display together with camera”* [1], en el que los participantes podían seleccionar y reproducir una serie de vídeos, almacenados en un servidor remoto, y sincronizarlos con un difusor de aromas ubicado en el lugar del usuario.

Entre las características de esta prueba vídeo-olfativa están el poder manejar el porcentaje de concentración de aroma en la prueba, logrando comprobar el efecto de mejora de la percepción de realismo ante este variable.

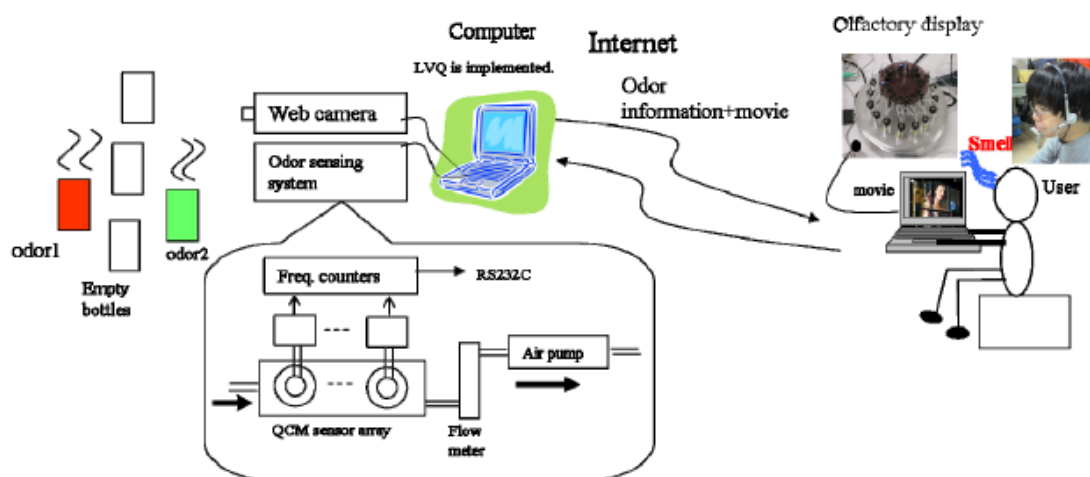


Figura 1. Estructura del Sistema tele-olfativo [1]

Esta prueba se realizó con dos esencias de frutas, manzanas y naranjas, representadas en la Figura 1 mediante odor 1 y odor 2, las cuales se emitían automáticamente cuando aparecían en pantalla las imágenes de dichas frutas. Esta técnica se realizó con microbalanzas de cristal de cuarzo, las cuales se conectaban a circuitos osciladores y su función era medir los cambios de frecuencia, para obtener respuestas del sensor y así obtener la orden de expulsión de las esencias.

En general este trabajo buscaba 3 objetivos: la visualización remota de contenidos multimedia, la sincronización audiovisual y olfativa y el control de concentración del olor al final de la prueba realizada en 41 personas, se deseaba conocer el promedio de personas que identificaran los dos

olores transmitidos, realizando varias pruebas continuas variando su concentración. Del total, 33 identificaron los dos olores en las pruebas realizadas, mientras que 6 personas solo identificaron un olor y 2 no identificaron ninguno de los dos.

Otro trabajo, más acorde a los objetivos de este TFM es descrito en la publicación *“Synchronization of Olfaction-Enhanced Multimedia”* [2]. Se basa en presentar los resultados de QoE realizados por un estudio experimental que explora, desde una perspectiva del usuario final, los límites temporales dentro de los cuales los datos olfativos se pueden presentar para mejorar la QoE. Entre sus resultados, se resalta la presencia de regiones temporales para el correcto funcionamiento de la sincronización y también cómo incide el adelanto y el retraso de los aromas con respecto a los contenidos audiovisuales. Una continuación de este trabajo, por los mismos investigadores es *“Perceived Synchronization of Olfactory Multimedia”* [3], donde se realiza un análisis más exhaustivo, añadiendo más variables. Uno de los factores analizados fue la incidencia de la sincronización labial con la auditiva. Sin embargo, el factor más relacionado con este TFM fue el análisis de la percepción de los usuarios sobre los niveles de sincronización olfativa con imágenes de un vídeo.

Asimismo, a diferencia de en [2], también se consideran otros factores que inciden en las pruebas, como es el caso de parámetros de QoS como, por ejemplo, el jitter, entre otros. Para lograr obtener mejores resultados se necesitó hacer un estudio experimental y considerar los límites temporales entre los cuales los aromas se podían sincronizar con éxito con otros objetos multimedia (en este caso, un vídeo). En concreto, se concluye que los efectos de adelantamiento olfativo en presencia de contenidos audiovisuales son más tolerables que el presentar un retraso olfativo con respecto al contenido audiovisual. Otro componente importante que se tuvo en cuenta fue la volatilidad de los aromas, así como otros factores externos que inciden en la vida cotidiana y no están presentes en las pruebas, como, por ejemplo, el viento. Otro factor que incidió en este estudio fue el concepto de multi-variedad de edad, género, profesión y clase social en los 42 usuarios que realizaron la prueba. A nivel de software, se implementó un programa de visualización donde se podía reproducir el contenido audiovisual. El hardware utilizado para la difusión de aromas es la marca Vortex¹, mostrado en la Figura 2, el cual presenta conectividad USB.

¹ Difusor de aromas, de la empresa Dale Air con sede en Reino Unido.



Figura 2. Difusor Vortex Activ.²

La prueba de este estudio se llevó a cabo con 6 olores ampliamente reconocibles e identificables, sincronizándolos cada uno con un videoclip independientemente, para que no se afectase la prueba por mala interpretación de algún aroma. Además, la prueba presentaba una relación temporal síncrona, con la idea de descubrir los límites aceptables en los cuales la presentación olfativa en relación a los contenidos audiovisuales se percibía por los usuarios.

El enfoque de este estudio se basaba en los estudios de Steinmetz [16]-[18] que trataban de encontrar los umbrales de asincronía tolerables para los usuarios, para diferentes combinación de componentes multimedia, como audio, video, texto, punteros, etc. La idea es analizar el impacto de diferentes niveles de asincronía sobre la percepción del usuario. Para lograr esto, en [3] se introducían asincronías de hasta 10 segundos, en intervalos de tiempo de 60 segundos. Las razones esgrimidas en el estudio fueron las siguientes:

- El rango de tiempo medio de un humano para percibir un olor está entre los 0,5s hasta los 2,5s después de ser generado, teniendo en cuenta la distancia y posición del difusor requerida [4], pero si se está a la espera del mismo este efecto se puede prolongar.
- El rango máximo aceptable para crear asincronías se limita a 90s para acomodar la adaptación olfativa con respecto a los contenidos audiovisuales. En [5], se demuestra que es mejor limitar la exposición a términos cortos de tiempo, como intervalos entre los 20s y 30s.

En la Figura 3, se observa el modelo OCPN (*Object Composition Petri-Net*), con el que se mide el lapso de tiempo antes de la presentación del contenido de los datos olfativos (P_o) en relación con la presentación del contenido audiovisual (P_{AV}), denominado $P_{\delta 1}$. Con esto se obtiene una variable representando el desfase temporal inicial, si existe. Del mismo modo, se puede obtener el desfase posterior denominado $P_{\delta 2}$.

² <http://www.daleair.com/dispensers/machines/vortex-activ>

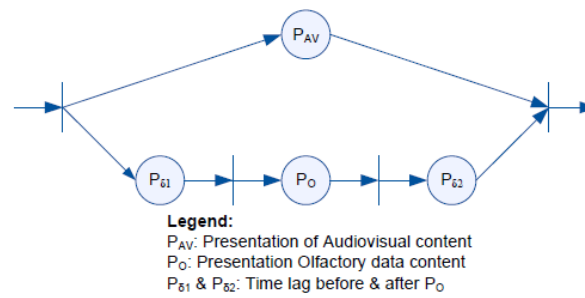


Figura 3. Modelo del escenario temporal usado en los experimentos [3]

En el proceso del experimento, los participantes visualizaron 6 vídeos acompañados con datos olfativos asociados, cada vídeo con una duración aproximada de 90 segundos. Las pruebas tuvieron una duración promedio de unos 20 minutos. Se realizó la primera difusión con una duración de 30 segundos, se llevaron a cabo 7 casos por cada videoclip, el primero se llevó a cabo en correcta sincronía, después se empezaron a lanzar las respectivas difusiones con cierta asincronía. Primero, se generaron con 30 segundos de adelanto ($P_{\delta 1} - 30$) y después se iban reduciendo hasta los 10 segundos ($P_{\delta 1} - 10$), con intervalos de 10s. Del mismo modo se hicieron difusiones iniciando con 30 segundos de retraso ($P_{\delta 2} + 30$), reduciendo este valor hasta 10 segundos ($P_{\delta 2} + 10$), tal como se muestra en la Tabla 1.

Tabla 1. Intervalo de valores asignados a los participantes de la prueba

Case	Clip 1 Skew	Clip 2 Skew	Clip 3 Skew	Clip 4 Skew	Clip 5 Skew	Clip 6 Skew
1	0s	-30s	-20s	-10s	10s	20s
2	30s	0s	-30s	-20s	-10s	10s
3	20s	30s	0s	-30s	-20s	-10s
4	10s	20s	30s	0s	-30s	-20s
5	-10s	10s	20s	30s	0s	-30s
6	-20s	-10s	10s	20s	30s	0s
7	-30s	-20s	-10s	10s	20s	30s

Dichos valores de desfase temporal se aplicaron para cada vídeo, y se mostraron aleatoriamente a los participantes para poder tener una mayor variación en los resultados y que no afectase a la información entre ellos. Entre las condiciones del lugar, se buscaba no crear persistencia entre los olores de cada prueba, se utilizó un salón con buena ventilación de aire. Cuanto a los parámetros de evaluación, se escogieron 5 conceptos: “demasiado pronto”, “pronto”, “tiempo adecuado”, “tarde” y “muy tarde”, para definir las percepciones de los participantes. Se obtuvo como media de sincronización entre las difusiones y el contenido audiovisual los valores con un rango entre ($P_{\delta 1} - 20$) y ($P_{\delta 2} + 20$), es decir, entre estos intervalos la media de participantes notó una sincronización mayor, mostrando un cambio de tipo exponencial en valores específicos como de ($P_{\delta 2} + 30$) de “tarde” y en ($P_{\delta 1} - 30$) de “pronto”, pero no considerando error de sincronizado el valor de ($P_{\delta 1} - 30$), lo cual quiere decir, que este es tolerable.

Los resultados obtenidos en este trabajo, nos servirán para definir nuestros límites de medida, entre los cuales se van a forzar situaciones de asincronías. En concreto, para este TFM se han considerado valores de desfase temporal de las difusiones de aromas con respecto al contenido audiovisual entre -30 segundos y + 25 segundos.

En el trabajo “*QoE assessment in olfactory and haptic media transmission: Influence of inter-stream synchronization error*” [7], también se abarca el tema de la integración de contenidos mulsemmedia con vídeos pero, además, combinando efectos táctiles (*haptics*) junto con los aromas. Las pruebas realizadas consistían en crear escenarios, por ejemplo, simulando el efecto de recogida de una fruta virtual 3D de un árbol mediante y percibir el olor de dicha fruta recolectada mediante el uso de un dispositivo olfativo. Para simular la recolección de las frutas se utiliza un dispositivo, con una interfaz de control táctil que proporciona fuerzas de retorno y presión, denominado PHANToM Omni, mostrado en la Figura 4. A partir de la recolección mediante el uso de este dispositivo, se generan las respectivas difusiones de los aromas correspondientes sincronizados con los objetos 3D representando las frutas.

En la Figura 5, se pueden observar, además del PC, los dos equipos hardware usados en este trabajo: el difusor de aromas SyP@D2³ y el dispositivo de interfaz táctil, el PHANToM.



Figura 4. Configuración de la prueba en [7].

Asimismo, en dicha prueba se generaban diferentes niveles de asincronías entre la presentación de los elementos visuales y olfativos, analizando su impacto sobre la QoE.

Este trabajo incluye factores adicionales no considerados en otros estudios [2, 3], como:

- La fluctuación o retardo de red.
- Las distancias entre el difusor y los usuarios (su nariz).
- La sensación del ambiente después de una difusión, que puede afectar o confundir al usuario con nuevos olores.
- Los tiempos de duración de las difusiones, ya que pueden generar efectos distorsionadores en la prueba.

³ Difusor de 6 cartuchos, de la compañía Exhalia.

Asimismo, en la prueba se controlaron las variables intensidad y duración en el difusor. La distancia entre éste y el usuario se fijó en 0,30 metros y se verificó que se necesitaba 2 segundos aproximadamente para que el usuario percibiera el olor. La velocidad del viento era aproximadamente 0,15 m/s.

La información táctil es interpretada como una variable de fuerza y se actualizan los parámetros definidos de variables de tiempo y posición, se almacenan en los diferentes buffers, y se genera una orden directa al difusor con los valores adecuados de intensidad y duración de la respectiva esencia. Después de recogida la fruta esta desaparece y la difusión se detiene.

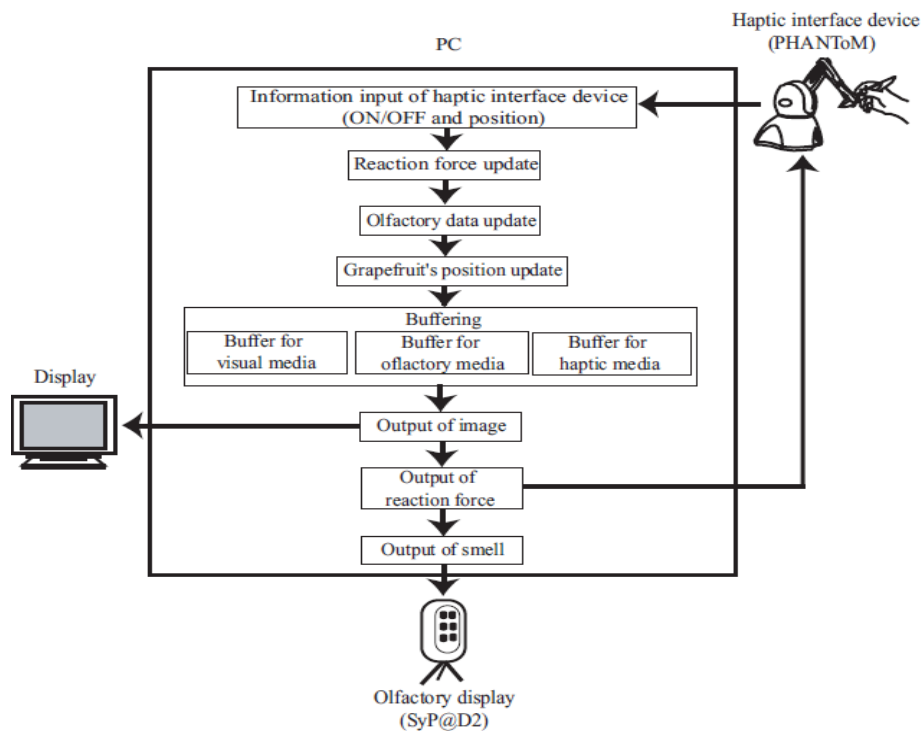


Figura 5 Funcionamiento del sistema propuesto en [7].

Para el análisis del QoE se llevaron a cabo 2 evaluaciones por usuario, cambiando la variable de recogida, forzando retardos y adelantos en la presentación de la información olfativa.

La prueba consiste en simular la recolección de una fruta con el dispositivo táctil y en el momento de obtener dicha fruta, emitir una difusión de olor que debe estar sincronizado con el sistema táctil. El mismo procedimiento se realiza nuevamente, pero en este caso, se presenta una asincronía con respecto a la difusión de olor, ya sea de adelanto o retraso. Se comparan ambos valores para obtener la calidad de experiencia QoE.

Para la evaluación subjetiva, se usó una escala Mean Opinion Score (MOS) [10] de 5 valores, definidos tal y como se indica en la siguiente tabla.

Tabla 2. Escala MOS

Descripción	Excelente	Bueno	Justo	Pobre	malo
Valor	5	4	3	2	1

La prueba se hizo con 20 personas de ambos sexos, con edades entre los 21 y 30 años, a cada uno de los cuales se le pidió valorar la calidad de sincronización utilizando dicha escala MOS.

En la evaluación, los errores de sincronización entre 1380ms y 250ms, de adelantamiento táctil, se corresponden con un valor medio por encima de 4. Esto se debe principalmente al tiempo aproximado de 2 segundos que tarda para percibir el olor, desde el instante de su emisión. Por otro lado, los valores más bajos de MOS, con un nivel mínimo de 2, se obtienen cuando se adelanta la emisión de olores por encima de los 500 ms. Por tanto, los asincronías son más molestos cuando los aromas están retrasados que cuando están adelantados.

Cabe destacar que existen muchos otros factores que inciden en la percepción de eventos en las personas, entre los que están el estado de ánimo, la cultura, la edad o el género, entre otros. El estudio en *“Age and gender influence on perceived olfactory & visual media synchronization”* [9], además de ser uno de los más recientes en este tema, analizalos resultados de los trabajos anteriores [1][2], considerando además el impacto de otros factores, como la edad y el género de los usuarios, haciendo ciertas combinaciones para obtener nuevos parámetros de análisis diferente a los trabajos anteriores. Se descartó el sonido de los vídeos como información relevante, dándole mayor importancia a la relación visual y olfativa.

Tal como se puede apreciar en la figura 7, para la prueba se utilizó un PC para la visualización del vídeo, un equipo difusor modelo SBi4, de la compañía Exhalia, y un ventilador adicional para ampliar la difusión hacia el usuario y para eliminar rápidamente los aromas de los difusores entre las secuencias de ensayo.



Figura 6. Sistema de prueba para integración olfativa y de vídeo. [9]

A nivel de software se desarrolló un programa con el Software Development Kit (SDK) del hardware, basado en JAVA y se integró en el reproductor VLC Media Player para sincronizar la reproducción del vídeo con la difusión de aromas.

La prueba se realizó con 43 usuarios, entre los 20 y 41 años, clasificados del modo indicado en la Tabla 3.

Tabla 3. Organizaciones por edades de cada género para la prueba.

Genero\Edad	Entre los 20 y 30 años	31 – 40 años	Más de 41 años
23 Hombres	11	10	2
20 Mujeres	9	8	3

La prueba consta de 6 vídeos de diferentes géneros, como documentales, programas de cocina, noticias, entre otros, cada uno con una duración aproximada de 90 segundos y divididos en bloques de 30 segundos. En el segundo de bloque entre los 31 y los 60 segundos se incorpora la difusión del aroma relacionado con el vídeo.

Los olores podían ser agradables o, en algunos casos, desagradables. Consistían en esencias de frutas, resinas, aromas picantes, etc. Dichos aromas ya fueron usados de forma similar en [2].

Para la evaluación primero se enseñó a los participantes un vídeo sincronizado y después el mismo vídeo con asincronías en la emisión de aromas. Se les pidió a los usuarios que respondieran ciertas preguntas sobre su experiencia. Esta metodología sigue lo definido en la recomendación ITU-T P.910 [10].

Se les hicieron 5 preguntas a los usuarios, a las que tenían que contestar seleccionando valores del 1 al 5, cuyo significado se incluye la Tabla 4. La primera pregunta busca identificar el error de sincronización y, por consiguiente, los límites de sincronización, la segunda pregunta busca determinar cómo se asimilaban o qué tolerantes eran con respecto a la asincronía, y las tres finales analizan el impacto de la asincronía sobre la QoE o (disfrute y nivel de realismo).

Tabla 4. Escala de respuesta MOS para las preguntas

Puntaje	Pregunta 1	Pregunta 2	Preguntas 3, 4 y 5.
5	Muy tarde	Imperceptible	Fuertemente agradable
4	Tarde	Perceptible pero no molesto	Agradable
3	Ni rápido, ni tarde	Ligeramente molesto	Ni agradable, ni desagradable
2	Rápido	Molesto	Desagradable
1	Muy rápido	Muy molesto	Fuertemente Desagradable

Los valores de asincronía en cada vídeo se representan como negativos cuando la difusión de aromas está por delante del vídeo y van a variar desde un rango de -30 segundos hasta +30 segundos, con sesgos cada 5 segundos. Es decir, se tomaran valores de -30s, -25s, -20s, -15s, -10s, y para el caso contrario donde la difusión de aromas esté después del contenido de vídeo estará representada por los sesgos 5s, 10s, 15s, 20s, 25s y 30s.

De los resultados de las respuestas a las preguntas realizadas se pueden extraer varias conclusiones importantes, aunque no se encontró una diferencia significativa en los patrones de percepción de asincronías en los usuarios cuando se analizó por género y edad. Es decir, los valores medidos en cada una de las pruebas mostraron una tendencia muy similar, salvo algunas medidas que se detallan a continuación:

- En la evaluación por edad, se notó una calificación mayor en los jóvenes, independientemente del género, al evaluar la sensación de realidad en la pruebas.
- En los valores de asincronía máxima, cuando se evalúa la sensación de disfrute, las personas de mayor edad calificaron por debajo de la media.
- Los rangos o zonas de sincronización obtenidos, respaldan los resultados de trabajos anteriores [1][2], comprobando niveles de percepción similares ante valores de asincronía establecidos.

En la publicación “*Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery*” [13], se propone un sistema adaptivo de transmisión de datos mulsemmedia, bautizado como ADAMS (*ADaptive MulSemedia delivery solution*), para la entrega de vídeo escalable y de los datos sensoriales a los usuarios (véase la Figura 8). ADAMS proporciona una transmisión adaptativa, en base a la propia información del vídeo, al contenido multisensorial y estado de la red.

ADAMS recomienda el uso de un esquema de descripción MPEG-7 (descrito a continuación), para la multiplexación e integración de múltiples efectos sensoriales (como, el olfato, movimiento de aire, táctil, etc.). Además, incluye módulos de adaptación en el lado del servidor: adaptación de flujo mulsemmedia y programación de prioridad de paquetes. A nivel de resultados, la evaluación subjetiva dedujo “niveles de disfrute mayores” al introducir secuencias mulsemmedia a contenidos multimedia y se estudió la preferencia de la inclusión de algunos efectos sensoriales ante variaciones del nivel de calidad de los vídeos. Los resultados del estudio demuestran las directrices de una estrategia adaptativa que selecciona la combinación óptima para los vídeos.

Las pruebas de percepción de los usuarios mostraron que ADAMS superaba las soluciones de entrega multimedia existentes, en términos de percepción del usuario y la calidad de disfrute, al añadir sistemas sensoriales adaptativos.

Las pruebas están basadas en los parámetros de medida y vídeos utilizados en trabajos anteriores [3], que se toman como punto de referencia para el posterior análisis de los resultados obtenidos y para demostrar que las nuevas variables pueden mejorar la inmersividad. Las condiciones de la prueba se llevaron a cabo basadas en las condiciones sugeridas en las recomendaciones ITU-T P.910 [10] y la ITU ITU-T P.911 [14] .

Los resultados principales están basados en medir la experiencia y el nivel de percepción de los usuarios, variando la escala de calidad de los vídeos usados y comparando los valores sin y con el

método ADAMS, demostrando una mayor satisfacción del sistema al incluirlo con vídeos de mayor calidad.

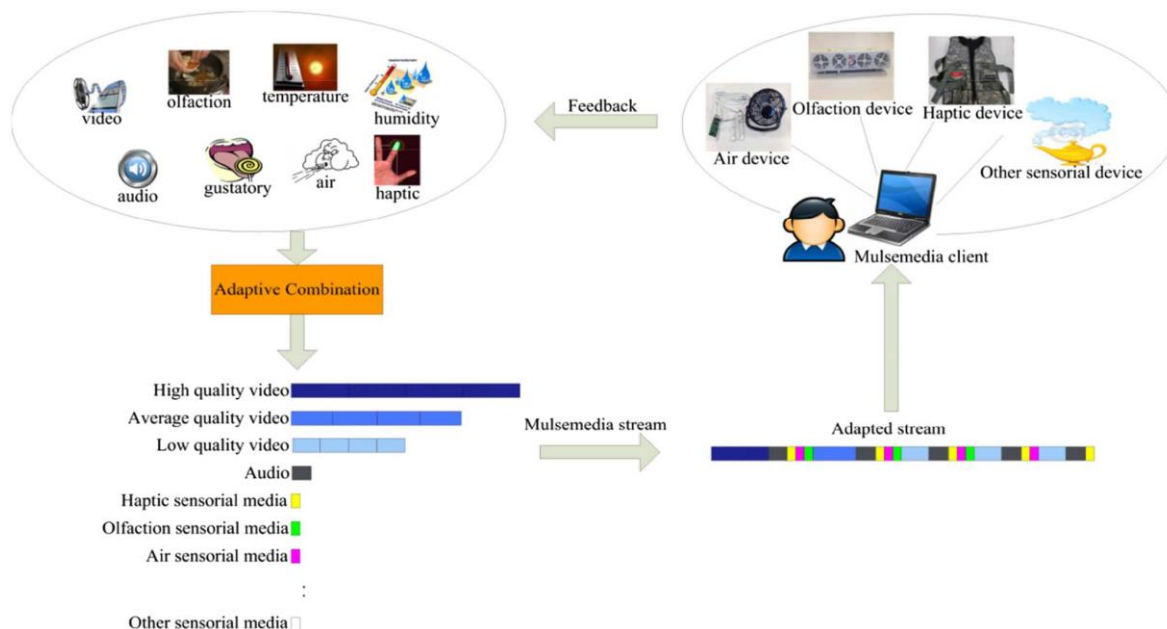


Figura 7. Ilustración del sistema ADAMS [13].

Como base fundamental de análisis para este TFM, por su alcance en los resultados y metodología de las pruebas, se resalta, la publicación “*Perceived Synchronization of Multimedia Services*” [15], que, al igual que [1] y [2], se basa en evaluar la inclusión de efectos sensoriales en los contenidos multimedia adicionando. Entre las variaciones están los valores de los límites temporales entre los cuales se llevarán las respectivas asincronías. Los nuevos valores se toman después de la realización de pruebas piloto para definir las regiones ideales de evaluación.

La prueba se realizó con 48 participantes, y se utilizaron 12 videoclips de 20 segundos seleccionados de dos películas.

Los resultados muestran los valores medios en los cuales los contenidos audiovisuales y los sistemas multisensoriales podrían ser aceptable para los usuarios.

También se analizan los umbrales tolerables en cuanto a errores de sincronización.

Los vídeos y metodología de las pruebas están basadas en los estudios previos [3] y [13], con tal de disponer de referencias de comparación.

2.2 Estándares relacionados

En este TFM, no podía faltar la consideración de los trabajos que el grupo MPEG está realizando en esta área de investigación. Hay dos estándares que de manera independiente están desarrollando

técnicas de transmisión de datos multimedia con contenidos adicionales. Dicha información adicional se transmite en forma de metadatos o etiquetas, según el caso, y sirven como mecanismo de control de efectos multisensoriales. A continuación, se detallan los estándares MPEG que incluyen estos aspectos y que se sirven como herramienta de referencia para este TFM.

2.2.1 MPEG- V

El estándar MPEG-V [18] proporciona herramientas para el control y renderizado de efectos sensoriales. Consta de 7 partes, de las cuales las siguientes están relacionadas con este TFM:

- La parte 1, que comprende la descripción de la arquitectura de MPEG-V.
- La parte 3, que define un lenguaje de descripción de efectos sensoriales (SEDL, Sensory Effect Description Language) y el método para ser transmitidos como metadatos, junto con los contenidos multimedia adicionales.
- La parte 6, que especifica la sintaxis y la semántica de los tipos de datos y herramientas comunes a las definidas en las demás partes de la norma ISO / IEC 23005.[22]

La integración de los sistemas Mulsemmedia, definiendo parámetros que pueden ser usados para una mejor experiencia multimedia, está incluida en la parte 3, donde se encuentra el segmento de información sensorial. En esta parte, se especifica la descripción del lenguaje de efectos sensoriales, SEDL, que viene siendo la representación digital de la información para los sentidos de la vista, el oído y el olfato.

El lenguaje SEDL incluye las etiquetas SEM (Sensory Effect Metadata), que se encargan de identificar atributos de los efectos sensoriales, como duración, intensidad, etc. Estos atributos pueden estar asociados a cualquier tipo de contenido multimedia, como videoclips, música, películas, juegos, etc.

Mediante las etiquetas SEM se puede tener el control e interacción mulsemmedia con contenidos audiovisuales, controlar dispositivos hardware que simulen efectos sensoriales, como, por ejemplo, el uso de difusores de aromas, control de flujo de aire, sillas con efectos de vibración, lámparas, etc.

En la Figura 9 se puede observar como los contenidos multimedia incluyen etiquetas, que definen los parámetros de los efectos multi-sensoriales, y que son procesadas para luego ser presentadas en los dispositivos apropiados en el instante adecuado.

Un ejemplo de pruebas mulsemmedia utilizando MPEG-V se puede consultar en [11].

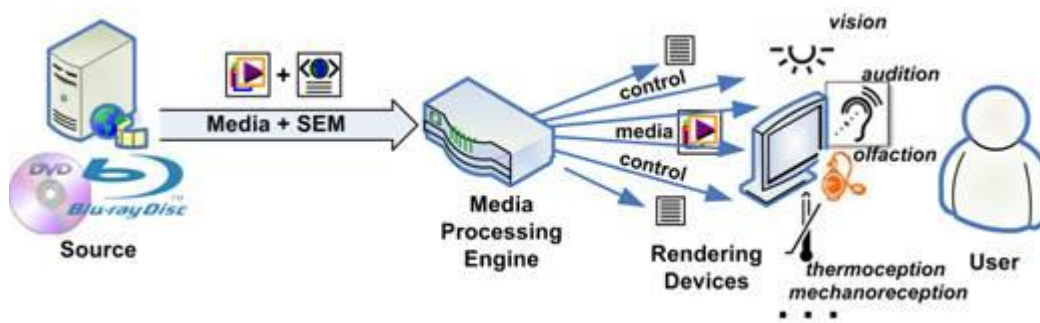


Figura 8. Funcionamiento del envío de etiquetas en MPEG-V [19]

2.2.2 MPEG-7

MPEG-7 [21] nace con la necesidad de disponer de mecanismos para la gestión, filtrado, organización y descripción de contenidos multimedia heterogéneos.

MPEG-7 ofrece un mecanismo para representar información audiovisual, de manera que sea posible desarrollar sistemas capaces de indexar grandes bases de datos incluyendo material multimedia (este puede incluir: gráficos, imágenes estáticas, audio, modelos 3D, vídeo y escenarios de cómo estos elementos se combinan) y su posterior filtrado o selección.

Se trata de un estándar relevante para este TFM, ya que se requiere manejar información de descripción de contenidos multimedia, insertando información de control de efectos multi-sensoriales y enviándola de manera paralela a los contenidos audiovisuales.

CAPÍTULO 3. DESARROLLO DEL PROYECTO

Para la consecución de los objetivos de este TFM, se necesitaba el desarrollo de una plataforma web, integrando diferentes componentes software y hardware (difusores de aromas). En este capítulo se describe el proceso de desarrollo del proyecto, así como las decisiones que se han ido tomando durante el mismo.

3.1. Hardware difusor de aromas

Para la difusión de aromas, se ha utilizado el dispositivo hardware SBix v2.0, de la empresa francesa Exhalia, que se suministran con un kit de desarrollo (SDK). El SDK lo conforman varias librerías, plugins, y ejemplo de uso, para diferentes lenguajes de programación, como Java, HTML, C++ y Visual Basic.

La plataforma multimedia de este TFM se ha desarrollado mediante tecnologías web estándar, como HTML y Javascript, para no evitar la instalación de ninguna aplicación específica en los equipos de prueba. El poder utilizar un navegador proporciona más versatilidad a los usuarios.

El control de los difusores se realiza a través de JavaScript, y funciona correctamente en el navegador Microsoft Internet Explorer de 32 bits, a partir de la versión 9 (restricción impuesta por el propio SDK).

Cada dispensador de aromas contiene 4 difusores de aromas, que son ventiladores que se activan y desactivan de manera controlada (véase Figura 9). Cada difusor presenta una cavidad donde se pueden insertar cartuchos del mismo fabricante, con las esencias de aromas elegidas.

Las características de los dispositivos difusores de aromas son los siguientes:

- Dimensiones: 40*40*10 mm
- Tasa voltaje por conexión USB: 5V DC
- Tasa de corriente: $0,14 \pm 10\%$ A
- Tasas de velocidad de giro: $5500 \pm 10\%$ RPM
- Noise: $26 \pm 10\%$ dBA

La comunicación con el PC es mediante una conexión serial USB, desarrollada principalmente para trabajar sobre la versión de Windows XP y 7 (32 bits). Para las versiones de sistemas operativos de 64 bits, el SDK presenta problemas de incompatibilidad y funcionamiento irregular. Para mayor fiabilidad y estabilidad, se decidió trabajar en sistemas de 32 bits, aunque se ha solicitado a la empresa fabricante que mejore y proporcione los drivers para sistemas de 64 bits.



Figura 9. Dispositivo hardware para la difusión de aromas, Exhalia SBix⁴

3.2. Protocolo de etiquetado y codificación de los metadatos sobre aromas

En este apartado se describe el protocolo de etiqueta y codificación de los aromas, insertados como metadatos de los vídeos de Youtube, que se ha especificado en este TFM. La codificación consiste en representar en la descripción de un vídeo de Youtube, información sobre ciertos parámetros para la presentación de los aromas. Para poder entender mejor este protocolo, en primer lugar se explicará cómo funciona un poco el SDK de los dispositivos dispensadores de aroma.

El SDK permite activar uno de sus difusores⁵ del dispositivo mediante una llamada a una función que admite 4 parámetros. El número del difusor, el instante de inicio de la difusión, la duración de cada difusión y la intensidad (velocidad) del ventilador.

Es por ello que se han definidos los metadatos sobre los aromas a través de etiquetas textuales que contienen información sobre dichos parámetros, siguiendo el siguiente formato:

OlfactoryD/Difusor/Retardo/Duración/Intensidad.

En concreto, los 5 campos de estas etiquetas (en formato de cadenas de texto), separados por el carácter “/”, tienen el siguiente significado:

- El primer valor es la cadena fija “OlfactoryD” que indica que se trata de una etiqueta con datos de efectos olfativos.
- El segundo valor, “Difusor”, indica el número de difusor que se debe activar. Cada dispositivo de difusión hardware consta de 4 difusores independientes, y se permite instalar varios dispositivos en un mismo PC, cada uno conectado a un puerto USB. Por tanto, se pueden conseguir números de aromas múltiplos de 4. El SDK contiene un fichero de configuración en el que a cada difusor se le puede asignar una etiqueta de texto de la forma ODEUR_X, donde X indica el número del difusor.

⁴ Tomado de: <http://perceptionofpoverty.org>

⁵ Definimos como difusión el proceso de activación de un ventilador, con las respectivas características de inicio, duración y velocidad.

- El tercer valor, “Retardo”, indica el instante de tiempo de inicio de la difusión. Esta variable está configurada en segundos relativos a la posición inicial del vídeo.
- El cuarto valor, “Duración”, define el intervalo de tiempo durante el cual el difusor deberá estar activado (ventilador en marcha).
- El último valor, “Intensidad”, indica, en una escala de 1 a 10, las revoluciones de giro del ventilador de cada difusor.

Un ejemplo de aplicación sería la siguiente etiqueta:

OlfactoryD/ODEUR_2/15/10/5,

En este caso, la etiqueta indica que se deberá activar el difusor número 2, en el segundo de 15 del vídeo, durante 10 segundos y con una intensidad media (nivel 5) del ventilador.

En adelante, a cada una de estas etiquetas se la denominará “Etiqueta de Aroma”.

3.3. Vídeos en YouTube

Al tener que trabajar con vídeos de YouTube, uno de los problemas encontrados fue la necesidad de integración y uso de la YouTube API⁶, necesaria para poder obtener la información relevante (metadatos) sobre efectos olfativos, como es el caso de la “descripción” del vídeo, que es dónde se han colocado dichos metadatos para este TFM, tal y como se indica posteriormente.

Para el TFM, se visualizan vídeos de YouTube, de una cuenta propia, a los cuáles se les han colocado etiquetas con información sobre aromas en la descripción del mismo, en la parte de la ‘descripción’. Ello se ha realizado al subir el vídeo a la plataforma de YouTube. Aquí se ha seguido la filosofía del estándar MPEG-7, planteado en [13]. Dicha información va a ser la referencia ya que se escribirá codificada en el formato de etiquetas descrito que serán leídas e interpretadas por la aplicación web desarrollada.

En la figura 10, se muestra cómo en la parte de descripción del vídeo aparecen las etiquetas introducidas para que sean interpretadas por la interfaz web desarrollada en el TFM.

⁶ Se trabajará con la YouTube Data API V3.

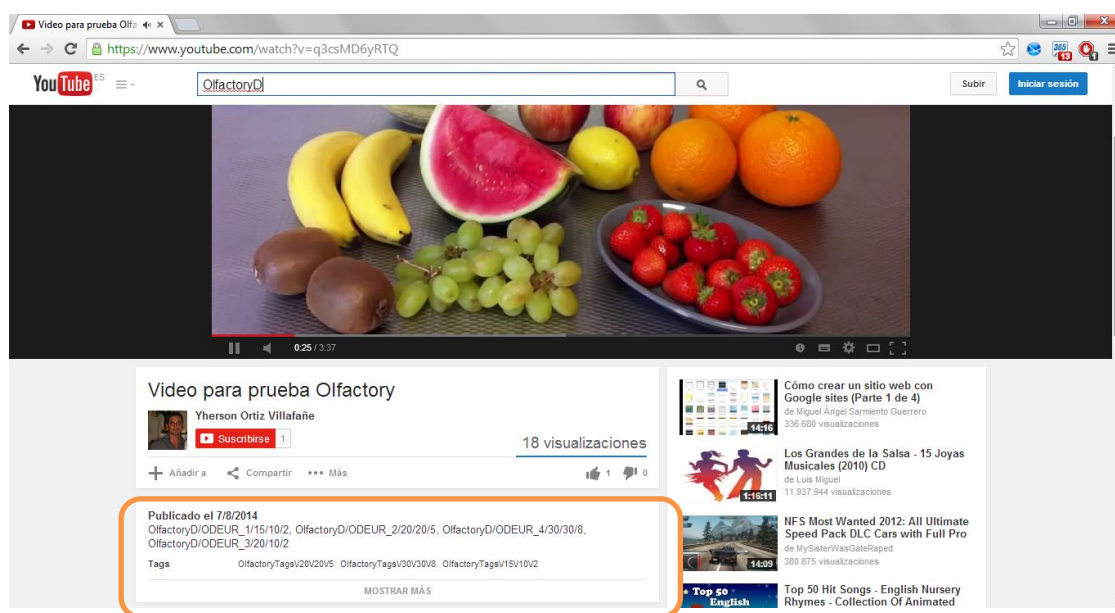


Figura 10. Etiquetas incluidas en la descripción de los vídeos.

3.3.1 Lectura de las etiquetas con metadatos en YouTube

En este TFM se decidió utilizar Youtube como servidor de vídeo. Esto evita tener que configurar un propio servidor, garantiza escalabilidad y ubicuidad, así como facilita que los usuarios puedan contribuir con más vídeos con datos sobre aromas y que puedan disfrutar del consumo de todos los vídeos disponibles con dicha información en Youtube, su disponen del dispositivo generador de aromas.

Para ello, se ha utilizado la API de YouTube, que permite capturar la información complementaria, en este caso la descripción del vídeo que contendrá las etiquetas sobre aromas, de los vídeos subidos a dicha plataforma.

En general, en la descripción de un vídeo de Youtube se pueden incluir cualquier información textual complementaria sobre el vídeo. En nuestro caso, se añade una lista de etiquetas, separadas por comas, conteniendo los metadatos necesarios para la presentación de los aromas, respetando la norma sobre la generación de aromas para la pruebas en recintos cerrados [12].

Por ejemplo, para un caso específico, podemos tener en la descripción de un vídeo de YouTube, el siguiente texto.

En el ejemplo mostrado en la Figura 10, se han incluido 4 “Etiquetas de Aroma” separadas por comas, en la sección de descripción del vídeo:

- “Etiqueta de Aroma” 1 —————> ODEUR_1/15/10/2
- “Etiqueta de Aroma” 2 —————> ODEUR_2/20/20/5
- “Etiqueta de Aroma” 3 —————> ODEUR_4/30/30/8
- “Etiqueta de Aroma” 4 —————> ODEUR_3/20/10/2

3.3.2. *YouTube Data API v3*

La API de datos de YouTube es una herramienta de gran ayuda para este TFM ya que mediante la misma se puede obtener, modificar e insertar información sobre los vídeos. Además, permite realizar búsquedas de vídeos o de listas de reproducción.

Para poder utilizarla, se necesita una credencial de autenticación o clave de API, y para ello se deben seguir los pasos detallados en el anexo 5.

3.4 *Desarrollo e integración de software*

La interfaz web desarrollada en el TFM consta de 4 archivos, cada uno de los cuales cumple una función determinada. Se trata de un fichero principal escrito en HTML, uno escrito en Javascript (index.js), otro en PHP (extensión .php) y un fichero de estilos CSS (styles .css).

3.4.1 *Funcionalidades mediante el uso de de JavaScript*

El núcleo principal del proyecto está basado en JavaScript. Mediante este lenguaje se controlan varios aspectos, entre los que se destaca:

- El módulo de activación y desactivación de los difusores durante los intervalos adecuados.
- La lectura y uso de la información decodificada proveniente de las descripciones de los vídeos. Es el encargado de decodificar las variables obtenidas del vídeo seleccionado en YouTube. Este procedimiento se lleva a cabo mediante el uso de arrays descriptivos, que organizan la información, y la almacenan en variables definidas para su posterior lectura por el complemento desarrollado en HTML.
- La sincronización del vídeo con las variables de control multi-sensorial, definidas en los metadatos de los vídeos, mediante las etiquetas previamente mencionadas. Aplicación de desfases temporales a los valores especificados en dichas etiquetas para evaluar su impacto sobre la QoE percibida por los usuarios.

3.4.2 *Funcionalidades mediante el uso de PHP*

Asimismo, también se ha utilizado PHP para poder utilizar la YouTube Data API, resaltando las siguientes funciones:

- El uso o llamado de la API. Esta característica es una de las más importantes ya que de ella se derivan los siguientes procesos de la aplicación.
- La búsqueda selectiva por palabra clave de vídeos y la respectiva obtención e interpretación de los metadatos relacionados, como son el título y la descripción, entre otros. Estos metadatos son muy relevantes ya que servirán para la configuración y activación de los difusores.
- Autenticación mediante el uso de la KEY API generada para poder utilizar la API.

- Almacenamiento de la información proporcionada por los usuarios durante la fase de evaluaciones subjetivas del TFM.

En el próximo Capítulo se proporcionan varias capturas de pantalla que muestran las funcionalidades de la plataforma web desarrollada.



CAPÍTULO 4. EVALUACIÓN SUBJETIVA

En este Capítulo se presenta el prototipo desarrollado para realizar las evaluaciones subjetivas, la metodología seguida en las mismas, así como los resultados obtenidos.

En dichas pruebas participaron un total de 44 usuarios, 30 hombres y 14 mujeres, de edades comprendidas entre 18 y 53 años, con una media de edad de 27 años, todos ellos estudiantes o trabajadores del Campus de Gandia de la UPV.

Los participantes debían acceder a una página web alojada en un servidor del mismo Campus, a través de la URL: <http://158.42.144.236/olfactory/index.html>.

4.1 Escenario de realización de las pruebas

Todos los usuarios realizaron la prueba en el mismo lugar, que reúne las condiciones básicas que exige la normativa relacionada [12]. Entre ellas, destacan el uso máximo de 10 aromas para que se obtenga un resultado fiable de la prueba, que el recinto tenga una temperatura y la ventilación apta para la circulación de los aromas y que no favorezca su persistencia, pues podría afectar a los resultados de la prueba.

La distancia entre los difusores y los usuarios fue de 50 centímetros, ubicados en posición horizontalmente al rostro y la altura de la nariz.

Las medidas del habitáculo fueron de 4.50m de largo, por 2,75m de ancho, por 2.80 m de alto. Dicho habitáculo presenta una puerta de 0.80 x 2m, y tres ventanas de 50cm de ancho y 75cm de alto, las cuales permitieron la evacuación correcta de los aromas, al terminar cada una de las pruebas. Además, se colocó un ventilador que se accionaba tras cada prueba y que favorecía dicha evacuación hacia las ventanas. La distribución del habitáculo se muestra en la Figura 13, y una imagen de un usuario participando en la prueba se puede ver en la Figura 14. En dicha figura se puede apreciar la distribución y distancia entre los principales elementos de la prueba: usuario, pantalla y dispensadores de aromas.

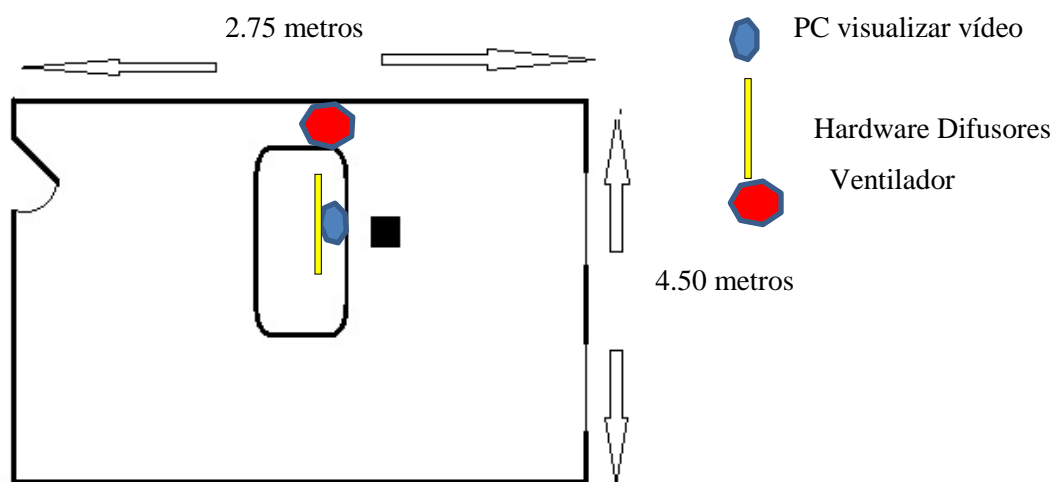


Figura 11. Plano y esquema del lugar donde se realizaron las pruebas



Figura 12. Usuario realizando la pruebas de evaluación

4.2 Metodología y Descripción de las pruebas

La prueba consta de 7 vídeos: un vídeo inicial de referencia, con una sincronización perfecta entre los efectos olfativos y el contenido audiovisual, y 6 vídeos adicionales en los que se forzaron diferentes casos de asincronía para su evaluación por los usuarios. Dichos vídeos se han diseñado en este TFM mediante el software Windows Movie Maker, combinando imágenes en movimiento y secuencias de vídeo en las que aparecen elementos visuales de los que se disponen aromas. Cada uno de las vídeos tiene una duración aproximada de 2 minutos y se pueden visualizar a través de la cuenta de Youtube del grupo de investigación Immersive Interactive Media R&D Group: <https://goo.gl/xS9HVX>

En dichos vídeos el rango de asincronía de la difusión de aromas con respecto al contenido audiovisual oscila entre +25 segundos y -30 segundos, ya que se corresponde con la región de sincronización de trabajos anteriores [2], y los sesgos son de 5 segundos.

En concreto, se han definido 11 casos para los 6 vídeos, tal y como se muestra en la Tabla 5. Par cada vídeo, la columna de la derecha que indica el valor 0 representa que no hay asincronía, mientras que la columna de la derecha indica el valor de asincronía, a excepción del Vídeo 2, en el que presenta la variación de la duración nominal durante la cual se presenta el aroma.

Tabla 5. Descripción de las asincronías de aromas.

Casos	Video Referencia	Vídeo 1		Vídeo 2		Vídeo 3		Vídeo 4		Vídeo 5		Vídeo 6	
Caso 1	0	0	+25	0	40%	0	-5	0	-30	0	C1	0	+5
Caso 2	0	0	+20	0	40%	0	-10	0	-25	0	C2	0	+10
Caso 3	0	0	+15	0	80%	0	-15	0	-20	0	C3	0	+15
Caso 4	0	0	+10	0	80%	0	-20	0	-15	0	C4	0	+20
Caso 5	0	0	+5	0	100%	0	-25	0	-10	0	C2	0	+25
Caso 6	0	0	-5	0	120%	0	-30	0	-5	0	C1	0	-30
Caso 7	0	0	-10	0	120%	0	+25	0	+5	0	C2	0	-25
Caso 8	0	0	-15	0	160%	0	+20	0	+10	0	C3	0	-20
Caso 9	0	0	-20	0	160%	0	+15	0	+15	0	C4	0	-15
Caso 10	0	0	-25	0	200%	0	+10	0	+20	0	C3	0	-10
Caso 11	0	0	-30	0	200%	0	+5	0	+25	0	C1	0	-5

Cada usuario evaluó uno de los 11 casos, que equivale al vídeo de referencia, más 2 condiciones para cada uno de los vídeos, del 1 al 6, primero con sincronización perfecta y después con el respectivo valor de asincronía. Así pues, en total, cada usuario visualizó 13 secuencias. El tiempo medio teórico de duración de la prueba por usuario son 45 minutos.

4.3 Prototipo desarrollado para la realización de las pruebas

Aunque bajo la supervisión del autor del TFM, a cada usuario se le dio el control de la prueba, que le fue guiada en 4 pasos a través de la interfaz web del prototipo desarrollado para las pruebas.

El Paso 1 (Figura 13) consistía en la realización de una encuesta inicial para verificar si los usuarios eran aptos o no para la prueba (anosmia y/o daltonismo). En caso afirmativo se registraba su información personal.



DEPARTAMENTO DE COMUNICACIONES

UNIVERSITAT POLITÈCNICA DE VALÈNCIA
CAMPUS DE GANDIA

immersive interactive media

Prototipo Para Aplicaciones MulSeMedia

PASO 1:
REALIZA LA EVALUACIÓN INICIAL EN EL ENLACE QUE APARECE A CONTINUACIÓN.

[PRESIONA AQUÍ](#)

ESTE PROCEDIMIENTO SE LLEVA A CABO UNA SOLA VEZ

PASO 2:
TODOS LOS VÍDEOS EXCEPTO EL INICIAL, SE VERÁN DOS VECES, ENTONCES SIEMPRE SE DEBE SELECCIONAR EL VÍDEO A REALIZAR Y EL CASO 0 PARA LA PRIMERA VISUALIZACIÓN Y SE PRESIONA EL BOTÓN "SELECCIÓN DE LA PRUEBA". DESPUÉS PARA LA SEGUNDA VISUALIZACIÓN SE MARCA NUEVAMENTE EL MISMO VÍDEO, PERO AHORA CON EL CASO QUE SE ASIGNADO Y SE PRESIONA EL BOTÓN "SELECCIÓN DE LA PRUEBA" NUEVAMENTE

Selecciona el vídeo:

Selecciona el caso:

[Selección de la Prueba](#)

Figura 13. Interfaz web del prototipo, pasos 1 y 2.

En el Paso 2 (Figura 13) se seleccionaba, por un lado, uno de los 7 vídeos y, por otro lado, el caso que le era asignado por el evaluador. Al pulsar el botón de *Selección de Prueba*, automáticamente se utilizaba la API de YouTube para cargar el vídeo seleccionado y obtener los valores correspondientes al control de los difusores de aromas (a través de las etiquetas del vídeo seleccionado).

En este paso, cada usuario visualizó el Vídeo 1, y se le indicó que pulsara el botón rojo (Figura 14) en el instante que percibiese el inicio y el final del aroma que se le indicó (que fue el aroma de curry). Con esto se perseguía medir el retardo de percepción de los aromas, es decir, el instante de tiempo que transcurre desde que se activa el generador de aromas hasta que el usuario percibe el aroma en cuestión. Este vídeo lo visualizó cada usuario dos veces, una vez sin asincronía alguna y otra con un valor de asincronía entre el rango establecido de -30s hasta +25s, para todos los usuarios.



Figura 14. Prueba de retardo de percepción de inicio y final del aroma.

El Paso 3 se corresponde con la evaluación subjetiva de cada uno de los 6 vídeos por parte de los usuarios. Al finalizar la visualización de cada vídeo, incluyendo los dos casos, los usuarios tenían que rellenar un cuestionario on-line preparado para tal fin.

Por último, en el Paso 4, se debía rellenar un cuestionario final on-line (Figura 15), en el que se evaluaba de manera general la experiencia multi-sensorial experimentada.



Figura 15. Interfaz web del prototipo, pasos 3 y 4.

4.4. Objetivos de la evaluación

El objetivo principal de la evaluación subjetiva es determinar el impacto sobre la QoE percibida de los usuarios cuando consumen contenidos audiovisuales, combinados con efectos olfativos (aromas), mediante la plataforma web desarrollada en este TFM. Por un lado, se pretende determinar el impacto sobre el nivel de inmersividad, disfrute y grado de realismo, cuando se incluyen aromas en comparación con cuando no se incluyen. Por otro lado, siguiendo la filosofía de

trabajos relacionados, se pretende evaluar el impacto de varios factores relacionados con la presentación de dichos efectos olfativos, como son los niveles de asincronía tolerables, pero también otros factores relevantes que no han sido considerados en ningún estudio previo, como son la variación porcentual de la duración nominal de los aromas (es decir, se modifica el intervalo durante el cual se emiten los aromas en comparación con el intervalo durante el cual los elementos asociado son visibles en el vídeo), y la variación de la separación temporal entre los aromas, para medir el grado de persistencia entre aroma.

En las pruebas realizadas, se utilizaron las métricas de evaluación y valores de puntuación que se presentan en la Tabla 6, con una escala de 5 valores siguiendo la filosofía MOS. Las variantes de la métrica MOS utilizadas (A, B y C en la Tabla 6) se entenderán mejor cuando, más adelante, se presenten los resultados asociados.

Tabla 6. Métricas de evaluación utilizadas.

Escala	5	4	3	2	1
A	Demasiado pronto (DP)	Pronto (P)	En el instante adecuado (EIA)	Tarde (T)	Demasiado tarde (DT)
B	Totalmente de acuerdo (TDA)	Más bien de acuerdo (MBA)	De acuerdo (DA)	Más bien en desacuerdo (MBD)	Totalmente en Desacuerdo (TED)
C	Imperceptible (I)	Perceptible pero no molesto (PNM)	Ligeramente molesto (LM)	Molesto (M)	Muy molesto (MM)

4.5. Resultados

En primer lugar, se analizan los resultados sobre la percepción de los usuarios sobre el instante de inicio de la difusión de los aromas con respecto al contenido audiovisual, empleándose la escala “A” de la Tabla 6 para la evaluación de esta prueba.

Se han detectado las regiones de sincronía y asincronía (de manera similar a [2]). La región de sincronía se define como el intervalo temporal de asincronía en el que los usuarios NO perciben error de sincronización entre la emisión del aroma y el contenido audiovisual. Se observa que valores entre -15 y 15 segundos de asincronía son aceptables para los usuarios, sin llegar a ser calificados como pronto o tarde. Ésta será, por tanto, la región de sincronía obtenida.

Los valores muestran una correlación en el comportamiento gráfico, con respecto a los resultados de [2]; pero en nuestro caso la región de sincronía es más estrecha. Cuando la emisión del aroma está adelantada o retrasada con respecto al contenido audiovisual, en un valor igual o superior a 20s, los usuarios ya perciben dicho adelanto o retraso.

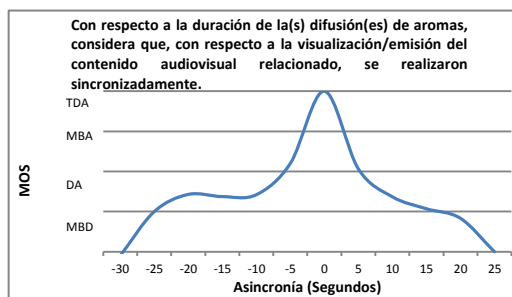
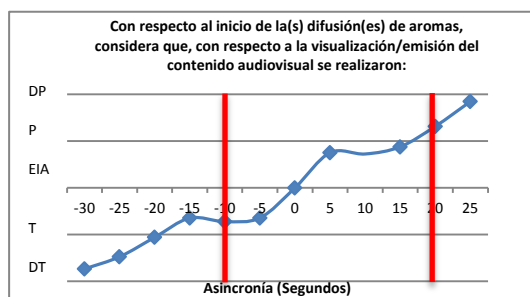


Figura 16. Resultados de la prueba subjetiva sobre el inicio de las difusiones de aromas. Figura 17. Resultados de la prueba subjetiva sobre, de la duración de las difusiones de aromas

A continuación, se analiza la percepción de los usuarios sobre la duración de los aromas con respecto al contenido audiovisual asociado, empleándose la escala “B” de la Tabla 6 para la evaluación de esta prueba. Se debe tener en cuenta que la duración de la emisión de los aromas no se varió, coincidía con la duración exacta del intervalo de tiempo entre la aparición y desaparición del objeto al que se corresponde el aroma emitido. Sin embargo, a medida que se aumentan las asincronías introducidas de retraso y adelanto de inicio de los aromas, también varía el instante de finalización de la emisión de los aromas.

En la Figura 17 se puede apreciar que a medida que aumenta el valor de asincronía, los usuarios percibieron en mayor medida este desfase.

En la figura 18 se muestran los resultados con respecto a la percepción de asincronía entre el aroma y el contenido audiovisual asociado para un vídeo en los casos evaluados. En este caso se utilizó la escala de evaluación “C” de la Tabla 6.

En la gráfica 18 además, se pueden diferenciar dos regiones. Una región con valores de asincronía comprendidos entre -5s y 5s, en la que las asincronías son imperceptibles, y la otra con valores de asincronía mayores a ± 5 segundos, en la que las asincronías ya son perceptibles, e incluso resultan molestas a los usuarios.

Tal y como se puede observar en la tendencia de dicha gráfica, cuanto mayor es la asincronía de la emisión del aroma con respecto al contenido audiovisual, mayor es la molestia ocasionada al usuario.

También se ha analizado la influencia en los usuarios de la separación temporal entre la emisión de dos aromas consecutivos y tratar de determinar la relación de este parámetro con la persistencia entre aromas. Para este caso, se utilizó un vídeo de referencia con dos aromas separados 35 segundos (se corresponde con el vídeo 5 de la prueba), y para la evaluación se varió el tiempo de separación del inicio de su difusión, quedando una separación temporal entre los inicios de difusión de 5s, 20s, 50s y 65s. En este caso se utilizó la escala “B” de la Tabla 6, y los resultados se muestran en la Figura 19.

Dichos resultados marcan una tendencia lineal, aunque suavizada, mostrando una mayor relevancia al reducir la separación temporal entre la emisión de dos aromas consecutivos. Un motivo puede ser que, al estar tan cerca, se genera una sensación de persistencia entre aromas, con lo que el primer aroma se puede mezclar con el segundo produciendo un efecto adverso. Con respecto a la evaluación de vídeos con mayor separación entre aromas, los usuarios no consideraron relevante dicha separación.

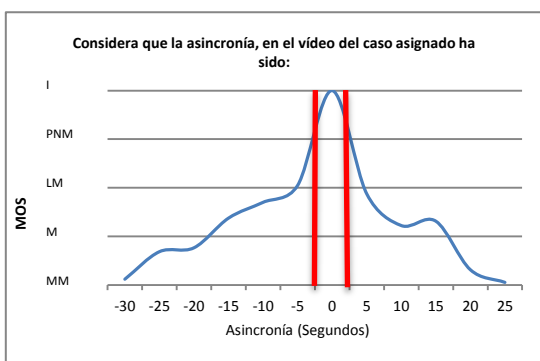


Figura 18. Resultados de la prueba subjetiva sobre la asincronía del aroma.

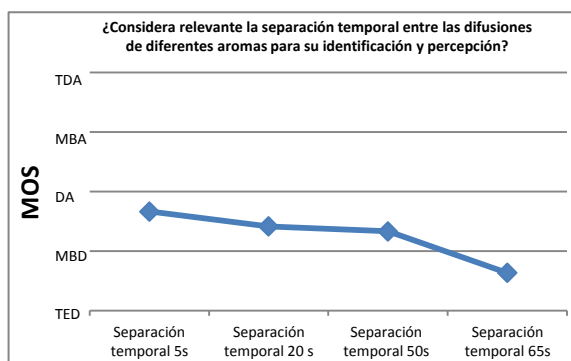


Figura 19. Resultados de la prueba subjetiva sobre relevancia de la separación temporal entre aromas.

La figura 20, muestra los resultados de una prueba en la que se utilizó un vídeo con aromas perfectamente sincronizados, pero al que se le varía porcentualmente la duración del aroma en los 5 valores específicos mostrados en el eje horizontal (el primer número se corresponde con el porcentaje respecto a la duración adecuada o correcta que se utilizó para la duración de la difusión empleada en cada prueba, y el segundo con la duración específica de emisión). El objetivo era medir el grado de percepción de los usuarios sobre la adecuación o no de la duración de la difusión, empleando como métrica la escala “B” de la Tabla 6.

Un efecto que se puede inferir de la gráfica es que al aumentar o reducir la duración de la difusión del aroma un 20% de la duración correcta (es decir, entre un 80 y un 120%), los usuarios todavía continúan percibiéndola como adecuada. Sin embargo, cuando la variación en la duración es superior al 20%, los usuarios ya empiezan a percibir que la duración no es correcta.

La figura 21, representa el tiempo en el que los usuarios percibieron los momentos de inicio y finalización de los aromas, para diferentes valores de asincronía entre el rango de -30s y 25s.

La tendencia de las muestras indican que, a mayor valor de asincronía del inicio de aroma en retraso con respecto al vídeo, el tiempo de percepción también aumenta, tanto para determinar el inicio como el final del olor. Para el caso de asincronías del aroma en adelante con respecto al vídeo, se observan diferentes tendencias para determinar el inicio y final del olor. A medida que aumenta la asincronía, es decir, a mayor adelanto del aroma, más tiempo tarda el evaluador en percibir el inicio de aroma. Con respecto a la percepción del final de aroma, se obtienen valores muy variables, donde poco se puede ver una tendencia o relación con respecto a las asincronías.

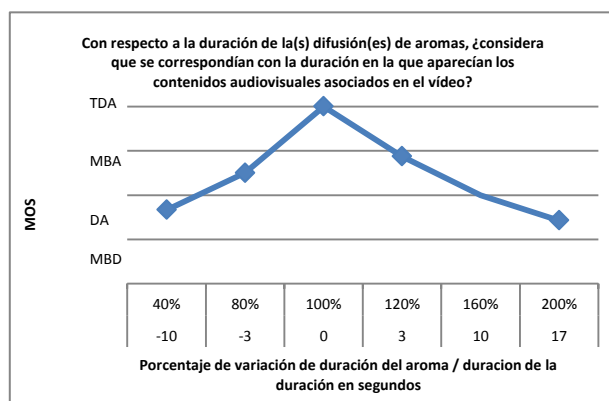


Figura 20. Resultados de la prueba subjetiva sobre variación temporal de aromas.

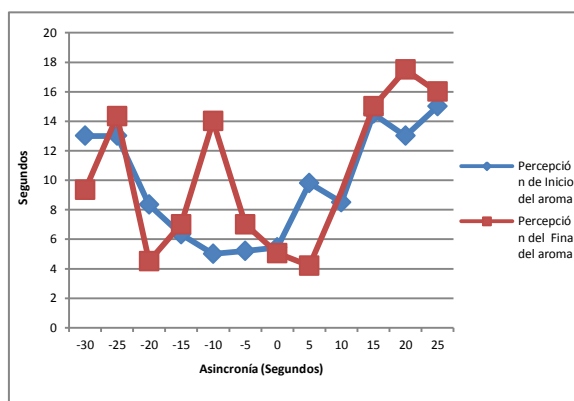


Figura 21. Resultados de la prueba subjetiva sobre la percepción del inicio y final del aroma, con diferentes asincronías.

En uno de los vídeos se incluyó la difusión de un aroma de una fruta que no aparecía en el vídeo de forma visual hasta pasados unos segundos. Se deseaba verificar si el aroma, al estar adelantado con respecto al vídeo, generaba expectativas o información al usuario sobre la imagen que espera ver. En la Figura 22 se muestra el resultado del comportamiento de los evaluadores, al percibir dicho aroma con poca o nula información audiovisual. Casi la mitad de los evaluadores, un 46%, afirmaron que sí les había proporcionado información y les había ayudado a identificar el objeto, mientras que un 17%, aproximadamente, indicaron que no fue así para ellos. Estos resultados indican que la inclusión de aromas pueden ser muy relevante para ayudar a identificar elementos audiovisuales cuando se consumen contenidos multimedia para un porcentaje significativo de usuarios.

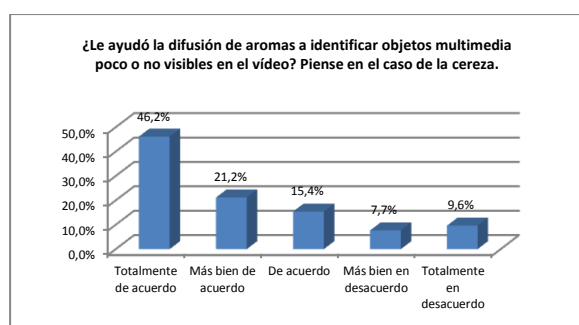


Figura 22. Evaluación del grado de identificación de efectos multi-sensoriales poco o no visibles

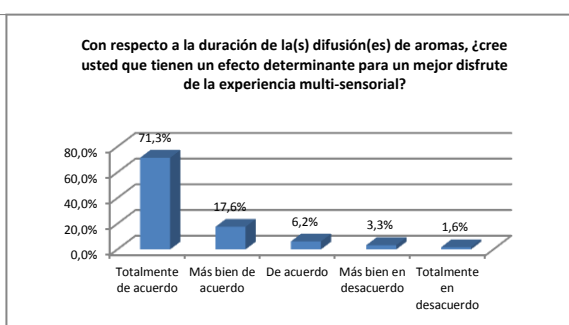


Figura 23. Evaluación de la importancia de la duración de los aromas

En las siguientes figuras se representa la importancia, ventajas y beneficios que se pueden obtener al añadir efectos con aromas a contenido audiovisual, así como el grado de aceptación por parte de los usuarios, valorando el nivel de la molestia percibida y la distracción provocada. Dichos resultados servirán como punto de partida para futuras investigaciones y aplicaciones en el campo de las comunicaciones multimedia.

La duración de los aromas emitidos en las pruebas osciló entre 10s y 18s, con una media de 13s. Con estos valores se aseguraron dos aspectos: i) evitar que fueran demasiado cortos y no percibidos; y ii) evitar que fueran demasiado largos, para evitar el solape entre aromas (por el efecto de persistencia en el ambiente). En la Figura 23 se muestra que los usuarios consideraron dichos valores de duración como correctos en un porcentaje, aproximado, del 72%.

Otro punto que se analizó fue el efecto de incluir varios aromas en un mismo vídeo. Tal y como se puede observar en los resultados de la Figura 24, el grado de persistencia olfativa incide al incluir mayor cantidad de aromas.

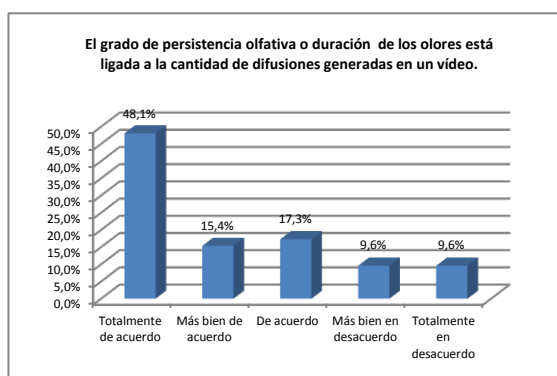


Figura 24. Relación de la persistencia de los aromas con respecto a las difusiones generadas

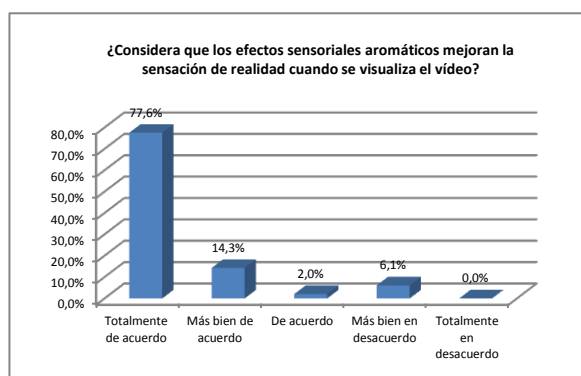


Figura 25. Relación de los efectos multi-sensoriales con la sensación de la realidad

Al preguntar a los usuarios si la inclusión de aromas mejora la sensación de realidad experimentada, los resultados fueron muy positivos, ya que más de un 90% de los usuarios contestó afirmativamente, tal y como se puede ver en la Figura 25. Asimismo, se preguntó a los usuarios sobre la relación entre el grado de realismo e inmersividad y la cantidad de aromas emitidos en un vídeo (Figura 26). En este caso, aunque casi la mitad de usuarios estuvo totalmente de acuerdo, no se obtuvo una tendencia clara o proporcional. Esto puede ser debido a que existen factores más relevantes para proporcionar mayor inmersividad como es el grado de afinidad con los vídeos observados por los usuarios, o la duración de las difusiones, entre otros.

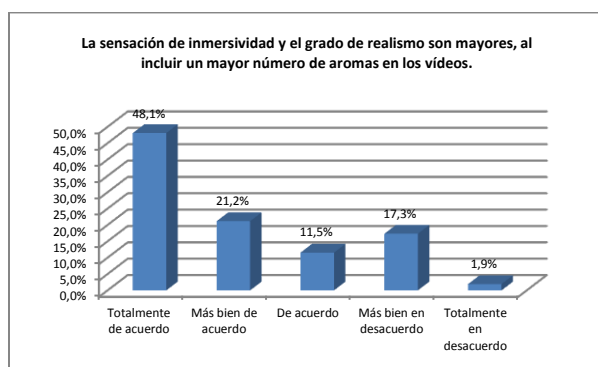


Figura 26. relación de la sensación de inmersividad con el número de aromas en vídeo

Otro factor a tener en cuenta en las pruebas, es que, para el usuario, el uso de elementos externos para crear efectos sensoriales no debería afectar, ni captar su principal atención, dejando de lado el contenido audiovisual. La idea es que constituyan un complemento que ayude a obtener un mayor realismo, pero en ningún momento puede ser distrayente o molesto. Es por ello que a los usuarios también se les consultó sobre las molestias ocasionadas y la distracción ocasionada por la inclusión de los aromas. Los resultados se resumen en las Figuras 27 y 28. En la figura 27 se puede apreciar que a la mayoría de usuarios les resultó poco molesto. En la Figura m se observan valores de opinión más dispersos, aunque la mayoría coinciden en que la inclusión de efectos sensoriales es poco distrayente.

También se les preguntó a los usuarios sobre el grado de disfrute durante las pruebas realizadas. Los resultados se muestran en la Figura 29. Se observa un grado de aceptación muy alto con respecto a las pruebas realizadas.

Otra característica que corrobora dichos resultados fue la sensación de duración de la prueba realizada por cada usuario. En promedio, las 44 pruebas llevadas a cabo tuvieron una duración media de 47 minutos. Sin embargo, al finalizar las mismas se les preguntó a los usuarios sobre cuál creían que había sido el tiempo de duración de las mismas. Se obtuvo un valor en promedio inferior de 32 minutos, es decir un 35% menos de duración con respecto a la duración real. Se puede justificar dicha impresión de los usuarios debida a una sensación de aceptación e inmersividad, logrando un mayor agrado y disfrute del contenido mulsemidia.

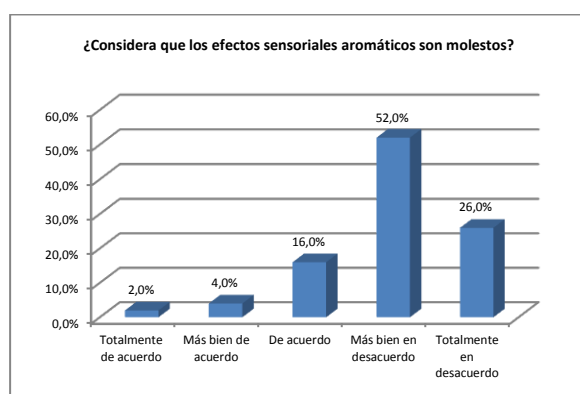


Figura 27. Evaluación del grado de molestia al usar efectos aromáticos

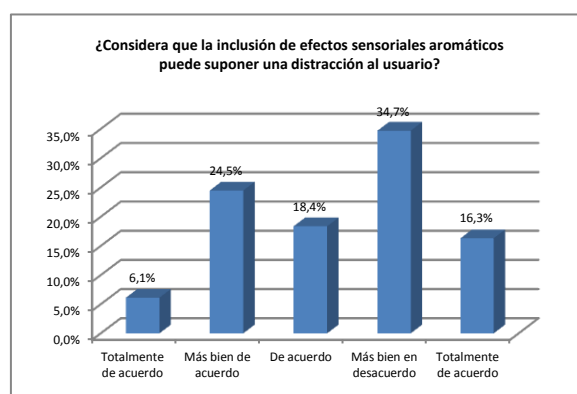


Figura 28. Evaluación del grado de distracción al incluir efectos aromáticos

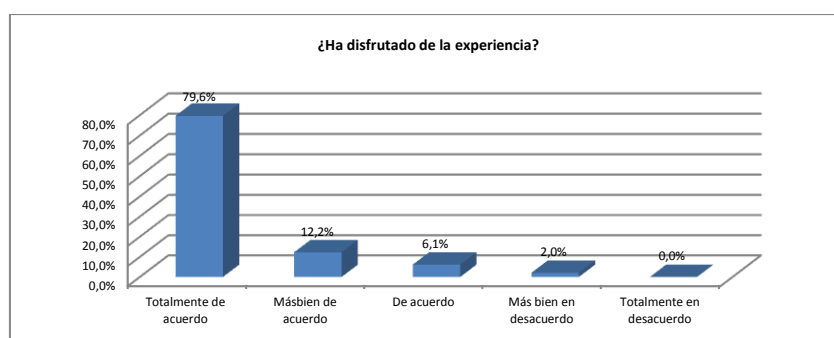


Figura 29. Evaluación del disfrute de la prueba

CAPÍTULO 5. CONCLUSIONES Y TRABAJO FUTURO

5.1 Conclusiones

En este TFM se ha presentado una plataforma web para la reproducción de vídeos de Youtube con la inclusión de efectos olfativos, mediante la integración de unos dispositivos hardware dispensadores de aromas. La información y sincronización de los aromas con el contenido audiovisual asociado se ha conseguido mediante la especificación de un protocolo para el etiquetado e interpretación de la información de los aromas, que se incluyen como metadatos en la descripción de los vídeos almacenados en Youtube. Un prototipo de dicha plataforma se ha utilizado para realizar evaluaciones subjetivas, en las que participaron 44 usuarios, para tratar de determinar la influencia de varios factores cuando se incluyen aromas, como los rangos de asincronía permisibles, la variación de la duración de la emisión de los aromas permisible, el retardo de percepción, la separación temporal entre aromas o la persistencia de los mismos. Asimismo, los resultados de la evaluación subjetiva reflejan que la inclusión de aromas es beneficioso para ayudar a la identificación de elementos visuales poco o no visibles, e incrementan el grado de realismo, disfruta e inmersividad cuando se consumen contenidos multimedia.

Así pues, se puede concluir que los objetivos planificados se han conseguido plenamente, y que los resultados obtenidos han sido satisfactorios y pueden ser útiles para la comunidad científica.

5.2 Líneas futuras

Durante el desarrollo del TFM se han identificado diferentes aspectos a mejorar y que se pueden abordar en una fase de investigación futura, como son los siguientes:

- Definición de más variables para el control del hardware y en el proceso de de codificación empleado.
- Uso de otro tipo de información adicional a la descripción y el título del vídeo, como es el caso de las etiquetas de YouTube, ya que estas ofrecen unas características más estándar que se pueden utilizar para una mejor comprensión e integración.
- Extender el desarrollo para que la plataforma web sea compatible con los diferentes navegadores disponibles.

REFERENCIAS

- [1] T. Nakamoto; N. Nimsuk; B. Wyszynski; H. Takushima; M. Kinoshita; N Cho, “*Reproduction of scent and video at remote site using odor sensing system and olfactory display together with camera*”, published in: Sensors IEEE, pp. 799-802, Lecce (Italia), October 2008.
- [2] G. Ghinea; O. A. Ademoye, “*Synchronization of Olfaction-Enhanced Multimedia*”, Published in: Multimedia, IEEE Transactions, (Volume:11 , Issue: 3), pp 561 – 565, April 2009.
- [3] G. Ghinea; O. A. Ademoye, “*Perceived Synchronization of Olfactory Multimedia*”, Published in: Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on (Volume:40 , Issue: 4), pp 657 – 663, July 2010.
- [4] Laing, D.G. 1983, “*Natural sniffing gives optimum odour perception for humans*”, *Perception*, vol. 12, no. 2, pp. 99-117.
- [5] Washburn, D.A., Jones, L.M., Satya, R.V., Bowers, C.A. & Cortes, A. 2003, “*Olfactory Use in Virtual Environment Training*”, Modelling & Simulation, San Diego, CA
- [6] SensAble Technologies Inc., “3D Touch SDK OpenHaptics Toolkit programmer’s guide version 1.0,” 2004
- [7] S. Hoshino, Y. Ishibashi, N. Fukushima, and S. Sugawara, “*QoE Assessment in Olfactory and Haptic Media Transmission: Influence of Inter-Stream Synchronization Error*”, Published in: Communications Quality and Reliability (CQR), International Workshop Technical Committee on, pp 1-6, IEEE, May 2011.
- [8] ITU-R BT. 500-11, “Methodology for the subjective assessment of the quality of television pictures,” 2002.
- [9] N. murray, Y. Qiao, B. Lee, G. Muntean, Karunakar A. k. “*Age and gender influence on perceived olfactory & visual media synchronization*”, Published in: Multimedia and Expo (ICME), IEEE International Conference on. pp 1 – 6, IEEE, july 2013.
- [10] Subjective Audiovisual Quality Assessment Methods for Multimedia Applications, ITU-T Recommendation P.911, December1998.
- [11] M. Walzl, B. Rainer, C. Timmerer, H. Hellwagner. “*An end-to-end tool chain for Sensory Experience based on MPEG-V*”, Signal Processing: Image Communication, (Volume 28, Issue 2), pp 136–150, February 2013,
- [12] ISO. 2006. ISO 5496:2006 Sensory analysis—Methodology—Initiation and training of assessors in the detection and recognition of odours. International Standards Organization (ISO).
- [13] Z. Yuan, G. Ghinea, and G. Muntean “*Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery*”, IEEE transactions on multimedia, (vol. 17, no. 1), January 2015.
- [14] Subjective Video Quality AssessmentMethods forMultimedia Applications, ITU-T Recommendation P.910, April 2008.
- [15] Z. Yuan, T. Bi, G. Ghinea, and G. Muntean “*Perceived Synchronization of Mulsemedia Services*”, IEEE Transactions on multimedia, (vol 17, no 7), July 2015.
- [16]Steinmetz, R. & Nahrstedt, K. 1995, “*Multimedia: Computing, Communications and Applications*”, Prentice Hall, New Jersey, USA

- [17] Steinmetz, R., "Human Perception of Jitter and Media Synchronization", *IEEE Journal on Selected Areas in Communications*, (vol. 14, no. 1), pp. 61-72, 1996
- [18] Blakowski, G. & Steinmetz, R. "A Media Synchronization Survey: Reference Model, Specification, and Case Studies", *IEEE Journal on Selected Areas in Communications*, (vol. 14, no. 1), pp. 5-3, 1996,.
- [19] ISO/IEC FDIS 23005-1 - Information technology -- Media context and control - Part 1: Architecture
- [20] M. Walzl, C. Timmerer, H. Hellwagner, "A Test-Bed for Quality of Multimedia Experience Evaluation of Sensory Effects", July, 2009.
- [21] official reference definition standard MPEG-7: <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>
- [22] ISO/IEC CD 23005-6 3rd Edition Common types and tools:
<http://mpeg.chiariglione.org/standards/mpeg-v/common-types-and-tools/isoiec-cd-23005-6-3rd-edition-common-types-and-tools>