

UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS

Facultad 6



*Algoritmo para la detección de
Publicidad en Materiales Audiovisuales.*

**TRABAJO DE DIPLOMA PARA OPTAR POR EL TÍTULO DE INGENIERO EN CIENCIAS
INFORMÁTICAS.**

AUTOR: Ivelin Ibarra Pérez

TUTOR: MSc. Yanio Hernández Heredia

Co-tutor: Ing. Yoandri Quintana Rondón.

La Habana, 1 de julio, 2011.


Año 53 de la Revolución.



“Si no existe la organización, las ideas, después del primer momento de impulso, van perdiendo eficacia.”

Ernesto Guevara de la Serna

Avales


Universidad de las Ciencias Informáticas
Federación Estudiantil Universitaria

Otorga el presente:

Reconocimiento


A. Algoritmo para la detección de publicidad en materiales audiovisuales

Por haber obtenido:

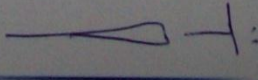
Relevante

"El conocimiento nos hace responsable.."
Ernesto Ché Guevara

Dado a los 27 días del mes de mayo de 2011
"Año 53 de la Revolución"



Lester González López
Presidente FEU-UCI



Ángel Vega García
Vicerrector de Investigación y Postgrado



Jornada
Científica
Estudiantil



9na
jornada
científica
estudiantil



Se le otorga
el presente

Reconocimiento

*A. Algoritmo para la selección de publicidad
en materiales audiovisuales.*

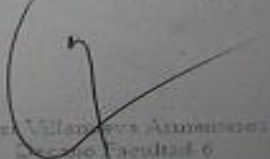
Por haber obtenido la categoría de

Relevante

En la 9na edición de la Jornada Científica Estudiantil


Alex. Casasa Hernández
Presidente FEUC


Luis Manuel Vidal Pich
Presidente Consejo Científico
FEUC


Yara Velázquez
Asesora Asesoría
Docente Facultad 6

Dedicatoria

A mis queridos padres Idalmis y Carlos.

A mi abuela Nena, mi hermanito Ernestico y demás familiares...

Agradecimientos

Al culminar el presente trabajo de diploma agradezco en primer lugar a mis padres Idalmis y Carlos, por confiar en mí y apoyarme en cualquier decisión que tomase. Los quiero mucho y son muy importantes en mi vida; por ustedes es que he logrado las metas que me he propuesto.

Agradezco a mi hermanito Ernestico por alegrarme los días con sus ocurrencias y sus bellos sentimientos. Gracias por creer siempre que soy la mejor, te quiero con la vida, eres muy importante para mí.

Agradezco a mis abuelos maternos Nena y Gerardo que aunque ya mi abuelo no esté siempre estará en mi corazón. Mi abuelita por estar pendiente de todo a pesar de su avanzada edad, ojalá dures muchos años para poder compartir contigo muchas experiencias. Te quiero mucho.

Agradezco a mi tía Martha y mi tío Rubén por todo el apoyo que me han dado en el transcurso de la carrera y durante toda mi vida, porque se han comportado como unos verdaderos padres, preocupándose a cada momento por todos mis problemas. A mis primos Ridel y Naviel por estar siempre ahí cuando los he necesitado y brindarme tanto cariño.

Agradezco a mi familia por parte de padre a mis abuelos María y Segundo, mi tía Ileana, mi tío Piti, Mi tío Alexis, mi tía Yamilé, mi primo Alexey, a todos por preocuparse por mí siempre.

Agradezco a mi novio Ismael por apoyarme aunque lejos. Por ser conmigo la bella persona que es y estar siempre ahí cuando lo he necesitado. Te amo.

Agradezco a todos mis profesores por formarme como la profesional que hoy soy, a mi Tutor Yanio por encaminarme y guiarme en toda la investigación, a mi cotutor Yoandri por luchar día a día con la investigación , apoyarme incondicionalmente y asumir el papel de co-tutor cuando más falta me hacía con la responsabilidad que lo hizo.

Agradezco a todas mis amistades de la Universidad que de una forma u otra me apoyaron y ayudaron. A mi amiga de 5 años Lisandra que siempre fue mi gran apoyo, por ser tan preocupada con los estudios y encaminarme con su ejemplo y sacrificio; a las amigas del grupo de CDI a Adayli, Yamilé, Glenda, Yisell, Krysna, Saily, Cecilia, Ariadna, Elizabeth que aunque no fue mucho el tiempo que compartimos siempre nos esmeramos en sentirnos como una pequeña familia; a las amigas de grupos anteriores a Yudalis, Aliuska y a Yanisley y Jose por nunca olvidarse de mí incluso cuando estaba en Venezuela.

Agradezco al tribunal por guiarme en la investigación y recomendar siempre de una forma constructiva, en especial a la profe Yanisley por revisar siempre el documento exhaustivamente cada vez que se lo envié.

Declaración de Autoría

Declaro que soy el único autor de este trabajo y autorizo a la Universidad de las Ciencias Informáticas a hacer uso del mismo en su beneficio.

Para que así conste firmo la presente a los ____ días del mes de julio del año 2011.

Ivelin Ibarra Pérez

Yoandri Quintana Rondón.

Datos de Contacto

Tutor: Msc. Yanio Hernandez Heredia.

E-mail: yhernandezh@uci.cu

- Ingeniero en Ciencias Informáticas, Universidad de Ciencias Informáticas, 2007.
- Vinculado al Proyecto Productivo UCITeVe durante los 5 años de estudios universitarios.
- Líder del departamento de Señales Digitales.
- Profesor del Departamento de Técnicas de Programación, Facultad 6.

Co-Tutor: Ing. Yoandri Quintana Rondón.

E-mail: yqrondon@uci.cu

- Ingeniero en Ciencias Informáticas, Universidad de Ciencias Informáticas, 2009.
- Vinculado al Proyecto Productivo UCITeVe durante 3 años de sus estudios universitarios.
- Ocupa el rol de arquitecto en el proyecto al cual está vinculado, Sistema de captura y catalogación de medias (SCCM).
- Profesor del Departamento de Señales Digitales, Facultad 6.

Resumen

El intercambio de información por parte del hombre desde su surgimiento ha repercutido directamente en el desarrollo de la sociedad. En la actualidad son los medios de difusión masiva los encargados de llevar el mismo mensaje a un mayor número de personas, con el objetivo de mantenerlos informados del acontecer. Mientras mayor sea la veracidad del contenido de los anuncios que se transmiten mayor será la aceptación por parte de la población. Mediante la presente investigación se propone un algoritmo capaz de detectar anuncios televisivos automáticamente por ciertas características que los identifiquen, de esta forma podrá ser analizado el contenido de los mismos. La investigación se encuentra dividida en tres capítulos de los cuales el primero contiene los conceptos que facilitan el entendimiento del dominio del problema, así como un estudio del estado actual de las técnicas existentes para la detección, en el segundo capítulo se propone un nuevo algoritmo que da solución a la problemática detectada en un inicio y por último en el capítulo tres se analizan los resultados obtenidos por la propuesta de solución. Como resultado de la investigación se obtiene el algoritmo para la detección de publicidad en materiales audiovisuales.

Palabras Claves.

Detección de publicidad, fotogramas, histograma, píxeles.

Contenido

Introducción.	1
CAPÍTULO 1. Fundamentos teóricos del algoritmo para la detección de publicidad en materiales audiovisuales.	6
1.1 Conceptos asociados al dominio de la investigación.	7
1.1.1 Publicidad.	7
1.1.2 Fotograma.	7
1.1.3 Histograma.	7
1.1.4 Shots.	8
1.2 ¿En qué consisten las técnicas de detección de publicidad en materiales audiovisuales? 8	
1.3 Técnicas de detección de publicidad en materiales audiovisuales.	9
1.3.1 Detección por fotogramas negros y disminución del volumen.	9
1.3.2 Alta tasa de corte y acción.	11
1.3.3 Algoritmo basado en la detección de logotipos.	11
1.3.4 Descriptores de Audio.	13
1.3.5 Reconocimiento de texto en video.	14
1.4 Conclusiones Parciales.	15
CAPÍTULO 2. Solución Propuesta.	16
2.1 Especificación de requisitos del algoritmo propuesto.	16
2.1.1 Requisitos funcionales.	16
2.3 Detección de fotogramas claves.	18
2.3.1 Cálculo de la intensidad media de la imagen.	18

2.3.2	Calculo del umbral.	22
2.4	Detección de picos elevados de volumen.	23
2.4.1	Extracción de Características del audio.	23
2.5	Detección de publicidad por texto en video.	23
2.5.1	Binarización de la Imagen.	24
2.5.2	Fragmentación o segmentación de la imagen.	24
2.5.3	Adelgazamiento de las componentes.	25
2.5.4	Comparar con patrones.	25
2.6	Conclusiones Parciales.	26
Capítulo 3.	Resultados Obtenidos.	27
3.1	Herramientas utilizadas.	27
3.1.1	MATLAB.	27
3.1.2	C++.	28
3.1.4	Framework de desarrollo QT.	30
3.2	Detección de fotogramas claves.	31
3.3	Detección de aumento de audio.	32
3.4	Detección de texto en video.	34
3.4	Realización de pruebas al algoritmo propuesto.	39
3.5	Conclusiones parciales.	43
Conclusiones Generales	44
Recomendaciones	45
Referencias Bibliográficas.	46

Bibliografía consultada.....	49
Anexos.....	52
Glosario de términos.....	61

Índice de figuras.

Figura 1: Bloque comercial separado por fotogramas negros (Sadlier, 2002)	10
Figura 2: Ejemplos de Logotipos que identifican a diferentes televisoras	12
Figura 3: Ejemplo de texto publicitario en video (Cade, 2007).....	15
Figura 4: Diagrama de procesos del algoritmo propuesto.	18
Figura 5: Número de píxeles por fila (N) y columna (M) de la imagen (Bueno, 2009).	19
Figura 6: Diagrama para la detección de cambios de toma.....	32
Figura 7: Representación de audio publicitario.....	34
Figura 12: Función para separar línea a línea el texto de la imagen.	37
Figura 13: Representación de la base de datos.	38
Figura 14: Resultados de las técnicas para la detección de Spots publicitarios.	43

Índice de tablas.

Tabla 1: Análisis de la técnica de detección por cambios de toma(Spots5).	39
Tabla 2: Análisis de la técnica de detección por cambios de toma(Spots4).	40
Tabla 3: Análisis de la técnica de detección por cambios de toma (Nov1, Nov2).....	40
Tabla 4: Análisis de la técnica de detección por cambios de toma (Peli1, Peli2, Peli3).....	41
Tabla 5: Análisis de la técnica de detección por aumento de volumen (Spots4, Spots5).	42

Introducción.

Desde sus inicios el hombre sintió la necesidad de comunicarse e intercambiar información, por lo que buscó diferentes vías para lograrlo, se comunicaba mediante señales de humo, pintando en las paredes de las cuevas, mediante las aves, entre otras técnicas. Luego surge la invención de la escritura, permitiendo la comunicación entre puntos distantes a través de mensajeros. Esto constituía un verdadero problema, ya que la información no llegaba en tiempo, debido a las grandes distancias y en muchas ocasiones el mensaje llegaba muy distinto a como lo había emitido la fuente de origen. A medida que la mente humana se fue desarrollando se buscaron nuevas alternativas con el fin de facilitar las actividades sociales. Los primeros pasos hacia una sociedad de la información estuvieron marcados por el telégrafo eléctrico, después del teléfono y la radiotelefonía.

A finales del siglo XIX surgen las tecnologías de la informática y las comunicaciones (TIC), las cuales han aportado al hombre una serie de herramientas capaces de facilitar la comunicación, constituyendo el uso y el acceso a la información su objetivo principal. El manejo de la información es cada vez más dependiente de la tecnología, lo que obliga a un tratamiento con medios cada vez más sofisticados.

Los sistemas de difusión audiovisual cobran cada vez más repercusión en la población, influyendo de forma notable en la vida de los ciudadanos. En la actualidad, en la mayoría de los hogares en países desarrollados, se encuentran varios dispositivos receptores de radio y televisión. La difusión alcanzada por estos medios alrededor del mundo así como la cantidad de telespectadores o radioyentes hacen de la televisión y la radio los medios publicitarios por excelencia. Los anuncios o spots comerciales son pequeñas secuencias de audio o video que permiten darle publicidad a un producto o servicio determinado **(Bueno, 2009)** .

La publicidad, como forma de comunicación, siempre se ha mostrado ágil para adaptarse a los cambios sociales, a las tendencias y a las modas, resultando ser, en incontables ocasiones, parte del sistema difusor e impulsor que ha favorecido la implantación de esos cambios.

También ha liderado innovaciones creativas en la forma de contar y comunicar. En ese continuo estado de adaptación, la publicidad se ha preocupado constantemente por encontrar nuevas maneras de relacionarse con el medio de comunicación que la acoge.

En la TVi¹ se pueden encontrar actualmente tanto mensajes publicitarios convencionales como interactivos. Se entiende por formatos convencionales aquellos que utilizan exclusivamente el lenguaje y las posibilidades tecnológicas propias de la televisión de flujo tradicionalmente entendida.

Los formatos convencionales no son más que los mensajes publicitarios audiovisuales diseñados para un consumo lineal, con inicio y final predeterminados por el emisor, en los que el receptor no tiene capacidad de alterar la linealidad, ni puede ejercer ningún otro tipo de interacción sobre el mensaje **(Quijada, 2007)** .

Los mensajes publicitarios convencionales no aprovechan las ventajas de la interactividad, aunque la naturaleza de la publicidad sea precisamente ésta: una respuesta del público en forma de compra del producto o servicio publicitado.

Se entiende por publicidad interactiva todo el conjunto de aplicaciones concebidas en cualquier formato publicitario que permita una acción mutua entre el usuario y la aplicación publicitaria.

La sociedad cubana no se ha mantenido ajena a estos avances, el perfeccionamiento constante ha permitido cada vez más, emplear los medios de difusión con el objetivo de elevar la educación y la cultura del pueblo. A diferencia de los grandes países consumistas la televisión cubana ofrece algunos espacios habituales sobre diferentes problemáticas por las que atraviesa el país, como son el ahorro económico, situaciones ambientales, temas relacionados con la medicina, la educación, el deporte, entre otras. Pero solo con estos espacios ambientales, muchos de ellos en horarios nocturnos no es suficiente el trabajo ideológico con la población. Es necesario además utilizar otras vías de difusión tan o más importantes que ellos y precisamente uno de estos son los spots que salen al aire entre los diferentes programas o dentro de espacios habituales **(Martínez., 2005)** .

Estos spots deben reflejar los principales problemas de cada territorio y a su vez los del país. Deben diseñarse para utilizar al máximo las potencialidades del canal visual ya que precisamente en este momento entre programa y programa es cuando más personas abandonan por instantes la TV².

La publicidad puede ser usada con fines maliciosos, puede llevar a la población un mensaje que carezca de veracidad, con el objetivo de confundir y crear malentendidos, esto es muy común en países donde hay

¹TVi: televisión interactiva.

²TV: televisión.

gran número de televisoras privadas, por lo que se hace necesario el uso de herramientas eficientes que permitan extraer información del material publicitario, para poder clasificarlo sin intervención humana.

El problema de detección de anuncios (o de contenido audiovisual en general) ha sido abordado de formas distintas por diferentes autores, **(Bueno, 2009) (Henneberger, 2008) (Albiol, 2004) (CVISION Technologies, 1998)** donde el objetivo fundamental es la discriminación entre contenidos y anuncios. Es por esto que se hacen necesarios sistemas capaces de monitorizar y reconocer de forma automática la emisión de un gran número de anuncios de TV.

La capacidad de detectar automáticamente los anuncios en sistemas de video posibilita, que las instituciones públicas verifiquen que las emisiones cumplen con la legalidad vigente en lo referente a tiempo de emisión de anuncios (Ley 25/1994 y sus posteriores modificaciones (ley 22/1999)) **(Bueno, 2009)** .

Para los anunciantes que compran tiempo en televisión o radio, la hora del día y la programación que se emite son aspectos de su interés. El horario y los programas afectan la audiencia probable y como resultado, la tarifa que se cobra por el tiempo de emisión. Estas tarifas son bastante altas, por lo que constituye una necesidad para los anunciantes comprobar el cumplimiento de sus contratos con las emisoras de TV o radio. Del mismo modo, es importante conocer el contexto en el que se produce la emisión: qué artículo o producto se han publicitado inmediatamente antes o después, si es o no un competidor directo. La detección y el reconocimiento precisos de anuncios es actualmente, una necesidad tanto para los anunciantes como para las televisoras.

El cumplimiento de las normativas de publicidad en la transmisión de contenidos audiovisuales en canales de televisión es inadecuado cuando no se monitorea o cuando se realiza manualmente. Con vista a darle solución al **problema** planteado queda definido como **objeto de estudio** el análisis de las técnicas de reconocimiento de texto publicitario en videos, la detección de audio por sus características y la variación de shots³ según cambios de tomas, enmarcado al **campo de acción** de la detección de publicidad para el área de gestión y procesamiento de medias en las aplicaciones de Monitoreo y Catalogación de canales de Televisión en el Departamento de Señales Digitales.

³ Shots: secuencia de cámara ininterrumpida.

El **objetivo general** de la investigación consiste en proponer un algoritmo para el proceso de detección de publicidad en transmisiones televisivas, utilizando técnicas de procesamiento de audio y video para el Departamento de Señales Digitales, para darle cumplimiento se definieron las siguientes tareas:

1. Identificar las variantes de técnicas y algoritmos existentes de mejores resultados, así como las tendencias utilizadas en la solución del problema planteado, a partir de un estudio del estado del arte.
2. Seleccionar las técnicas eficientes de búsqueda y detección de texto en imágenes y video a utilizar en el algoritmo a desarrollar.
3. Seleccionar las técnicas eficientes de detección de audio según el volumen, para utilizar en el algoritmo a desarrollar.
4. Seleccionar las técnicas eficientes para la detección de shots en videos.
5. Diseñar un algoritmo para detectar segmentos de publicidad en materiales audiovisuales utilizando las técnicas seleccionadas.
6. Validar el algoritmo propuesto con el desarrollo de una herramienta que implemente las técnicas seleccionadas.

Para obtener los conocimientos necesarios que hagan posible el cumplimiento del **objetivo** planteado en el trabajo y cumplir las **tareas** propuestas, se lleva a cabo una investigación científica en la que se utilizan los siguientes **métodos científicos**:

Analítico -sintético: ha permitido definir la propuesta del algoritmo de detección de publicidad, a partir del conocimiento, las teorías y documentos que ocupan el centro de la investigación, extrayendo los aspectos más significativos relacionados con el objeto de estudio de la misma.

Inductivo-deductivo: mediante un estudio y análisis de las técnicas generales para la detección de publicidad en materiales audiovisuales, se pudo llegar a deducir un algoritmo de menor nivel de generalidad, con mayor grado de especificidad.

Histórico-Lógico: sirvió para el estudio y análisis de las técnicas y algoritmos de detección de publicidad en la evolución del desarrollo computacional.

Métodos Empíricos: el experimento permitió validar la información del algoritmo propuesto.

La investigación consta de **3 capítulos** donde se realiza un estudio crítico y descriptivo de las técnicas de reconocimiento de publicidad.

El capítulo 1 trata los temas relacionados con la fundamentación teórica que justifica la investigación, se analiza el estado actual del tema a nivel nacional e internacional.

En el capítulo 2 se expone todo lo referente a la solución propuesta, así como las nuevas tendencias, tecnologías empleadas en la solución y otros temas de interés.

El capítulo 3 muestra los resultados obtenidos después de la aplicación de la solución propuesta, aportando conclusiones respecto al uso de la misma así como recomendaciones para trabajos futuros.

Los Anexos incorporados al final del documento contienen elementos que apoyan la fundamentación teórica de la investigación así como una muestra de los resultados obtenidos y se definen en el **Glosario de Términos** los conceptos relacionados con el tema de la tesis u otro tema que sea de difícil comprensión.

CAPÍTULO 1. Fundamentos teóricos del algoritmo para la detección de publicidad en materiales audiovisuales.

Las técnicas para la detección de anuncios televisivos es un tema que ha sido abordado por varios autores, **(Bueno, 2009) (Henneberger, 2008) (Albiol, 2004) (CVISION Technologies, 1998)** por lo que han surgido diferentes definiciones, interrogantes y soluciones. En el presente capítulo se abordan los conceptos y temas relacionados con el proceso de detección de publicidad en materiales audiovisuales mediante técnicas de detección de texto publicitario, de sonido con picos elevados de volumen y shots con gran variabilidad de efectos y transiciones. Con el objetivo de ubicar el contexto en el que se enmarca la investigación, se hace un estudio detallado de las técnicas y algoritmos eficientes existentes en el mundo para dar solución a problemas similares a los de la investigación.

Una secuencia de video es una rica fuente de información multimodal, con voz, audio, texto (si los subtítulos están disponibles), los patrones de color, forma de los objetos reflejados y el movimiento de estos objetos (cambios en los marcos). Aunque el ser humano puede interpretar con rapidez el contenido semántico de la fusión de la información de las diferentes modalidades. Con el auge de la Internet y varios tipos de recursos multimedia, es de interés la obtención de herramientas eficaces que permitan facilitar la difusión de la información audiovisual por el ser humano. Los recursos multimedia deben ser indexados, almacenados y recuperados de una manera similar a la forma en que un cerebro humano los procesa. Las aplicaciones que requieren la comprensión de la escena incluyen la localización y rastreo de eventos especiales en el video, el seguimiento activo de objetos especiales, edición de video y la composición entre otros. La clave para entender el contenido de una secuencia de video es la segmentación y clasificación de la escena **(Liu, 1998)** .

En el área de gestión y procesamiento de medias del departamento de señales digitales es de utilidad extraer información del material publicitario, para así analizar el contenido del mismo y velar porque cumpla con la calidad y una serie de normativas establecidas como son: cumplimiento por parte de las emisoras de TV o radio de los contratos con los anunciantes, la legalidad vigente en cuanto a tiempo de emisión de anuncios, conocer el contexto en el que se produce la emisión, qué artículo o producto se han publicitado, tiempo máximo entre la salida de un comercial al aire y el siguiente, entre otras.

1.1 Conceptos asociados al dominio de la investigación.

Para un mejor entendimiento durante el desarrollo de la investigación a continuación se explican algunos conceptos que son utilizados con frecuencia en cada uno de los capítulos del trabajo de diploma.

1.1.1 Publicidad.

La publicidad es una poderosa herramienta de la promoción que puede ser utilizada por empresas, organizaciones no lucrativas, instituciones del estado y personas individuales, para dar a conocer un determinado mensaje relacionado con sus productos, servicios, ideas u otros, a un determinado grupo. Es la acción de transmitir un mensaje utilizando medios de comunicación con la intención de influir sobre el comportamiento del consumidor. El objetivo es el dar a conocer el producto, para así aumentar su nivel de ventas (**Chust, 2010**).

La publicidad es una forma de comunicación impersonal de gran alcance, ya que utiliza medios masivos de comunicación como la televisión, la radio, el internet entre otros. Sin esta información suministrada por los anunciantes al consumidor, el producto no puede competir en igualdad de condiciones con los productos de la competencia.

1.1.2 Fotograma.

Se conoce como fotograma a cada una de las imágenes que se suceden en una película cinematográfica y que están consideradas de manera aislada (**DefiniciónABC, 2010**).

Un espectador siente la sensación de movimiento en una producción audiovisual cuando son visualizados los fotogramas con una determinada frecuencia, la velocidad con que se sucedan las imágenes es la responsable de que la escena tenga mayor similitud a escenas del mundo real.

1.1.3 Histograma.

El histograma de una imagen f es una gráfica que representa los niveles de intensidad del color de f con respecto al número de píxeles presentes en f con cada intensidad de color (**Serra, 2005**).

Mediante el histograma se obtiene la intensidad de los píxeles por cada nivel de gris. Esta información será mostrada en una gráfica.

1.1.4 Shots.

Se define como shots a la secuencia de cámara ininterrumpida que no es más que la unión de cada una de las tomas que forman una secuencia de video.

1.2 ¿En qué consisten las técnicas de detección de publicidad en materiales audiovisuales?

Los segmentos comerciales son materiales audiovisuales con muy escasa duración. Por este motivo suelen tratarse de secuencias con un elevado número de efectos, generalmente ausentes en otros tipos de producciones audiovisuales **(Bueno, 2009)** .

Los segmentos comerciales tienen ciertas características que los hacen más fáciles de identificar que los segmentos de video en general. Los comerciales son casi siempre agrupados en bloques, por lo general poseen de 4 a 10 anuncios cada uno. La duración de los distintos anuncios es corto, casi siempre menos de noventa segundos y por lo general es un múltiplo entero de quince segundos. Además estos tienden a tener un alto grado de acción, para captar la atención de los espectadores en la pequeña cantidad de tiempo disponible para transmitir un mensaje, caracterizados por un número significativo de cortes **(Henneberger, 2008)**.

Señala Lienhart que la tasa de cortes por minuto en una muestra de 200 anuncios de la televisión alemana fue de 20.9, mientras que en materiales de transmisión normal fue de sólo 3,7 **(Marques, 2004)**. Cuando un comercial sale al aire hay una tendencia de aumentar el volumen de las pistas de audio, además delimitar al principio y al final con fotogramas⁴ de color negro o blanco, acompañados de silencio **(Tie-Yan Liu, 2004)**. Muchas estaciones de televisión exhiben una red de logos en la esquina inferior derecha de la pantalla durante la programación regular y en el momento que sale el comercial este logo se elimina **(Henneberger, 2008)** .

⁴ Fotograma: cada una de las imágenes que se suceden.

El uso de los métodos tradicionales para hacer búsquedas y obtener información relevante en los spots comerciales se muestran ineficientes e inadecuados, ya que requieren de mucho tiempo y recae toda la responsabilidad en el ojo humano. La obtención de herramientas capaces de detectar automáticamente comerciales en materiales audiovisuales, agiliza y optimiza este proceso. Las características antes expuestas presentadas por los spots hacen posible el uso de algoritmos de detección aptos para la extracción. La mayoría de los algoritmos se basan en un gran número de estas pistas, combinando las mismas con el fin de lograr mejorar la tasa de detección.

1.3 Técnicas de detección de publicidad en materiales audiovisuales.

Como parte del proceso de desarrollo de las TIC han surgido una serie de algoritmos y técnicas capaces de facilitar en gran medida el intercambio de la información, formando parte de estas las técnicas de detección de publicidad en materiales audiovisuales, las cuales se han ido perfeccionando y adaptando a las necesidades inmediatas de la sociedad, por lo que se hace necesario hacer un estudio detallado de algunos algoritmos eficientes existentes en el mundo relacionado con la temática antes expuesta.

1.3.1 Detección por fotogramas negros y disminución del volumen.

Se determina que un fotograma dado es negro si la intensidad media del valor de los píxeles de la imagen está por debajo de un determinado valor de umbral, la obtención de la intensidad se calcula fácilmente en el dominio analógico según Hurs. Para una mejor detección se debe garantizar que la desviación estándar de los valores de intensidad también sea inferior a un umbral de acuerdo Lienhart (**Henneberger, 2008**) .

Para determinar si un marco negro está acompañado de silencio se toma la suma del valor absoluto de todas las muestras de audio que corresponde a un fotograma de video, se define un valor de umbral para indicar silencio y se compara el valor calculado con el valor del umbral. En base a este resultado, se puede tomar la decisión de que se está en presencia o no de silencio (**Sadlier, 2001**) .

Se establece un máximo tiempo entre un fotograma negro y otro. Si el tiempo entre las secuencias de un fotograma negro, es mayor que este, el segmento se considera ser parte del programa. El algoritmo propuesto por Sadlier establece que la longitud máxima de un comercial es de noventa segundos (2250 imágenes a 25 cuadros por segundo). Este algoritmo también mira cuántos segmentos consecutivos de comerciales se producen, para determinar si el comercial candidato es en realidad parte de un bloque de comerciales. Si un bloque no contiene por lo menos tres comerciales individuales, debe ser parte del programa. Si por lo menos cuatro secuencias de fotogramas negros se producen con una separación máxima entre cada uno de noventa segundos, se clasifica como un corte comercial (Marques, 2004).

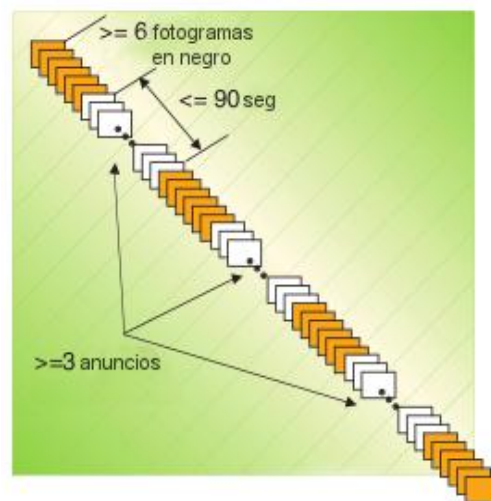


Figura 1: Bloque comercial separado por fotogramas negros (Sadlier, 2002).

Se concluye diciendo que mediante este algoritmo, un fotograma negro se detecta solo si es acompañado de un silencio. El algoritmo analizado se rige por una serie de características mostradas por los comerciales y se establecen ciertas condiciones para tomar la decisión de si se está o no en presencia de un corte comercial, como por ejemplo, para reducir aún más la posibilidad de una detección errónea el algoritmo requiere que sean detectados cierto número de los sucesivos fotogramas negros acompañados de silencio en conjunto. Se combinan tanto técnicas de procesamiento de audio como de video, dando una mayor precisión en cuanto al resultado.

1.3.2 *Alta tasa de corte y acción.*

En (**Television Commercial Detection Techniques., 2008**) se propone un algoritmo para la detección de comerciales mediante tasas de cortes que se rige por: una secuencia candidata debe tener una tasa de corte superior a cinco por minuto para su totalidad y como mínimo en un minuto más de treinta cambios de toma.

Lienhart utilizó un grupo de dos indicadores adicionales relacionados con el alto nivel de acción en los comerciales. En primer lugar, el "índice de cambio del borde", describe el número de píxeles del borde (según lo encontrado por un algoritmo de detección de bordes). El segundo indicador es el vector de movimiento, el cual describe el movimiento de objetos en la imagen. Los métodos de detección sobre la base de estos dos parámetros tienen un índice de repetición en torno al 96% cuando se usa en bases de datos de prueba.

Se concluye diciendo que es una técnica que se basa en una característica muy significativa y común de los comerciales, por lo que pudiese ser aplicada a cualquier televisora, no dependiendo de tendencias de las mismas. Este algoritmo tiene un índice de repetición de 93,43% y una tasa de detección de falsos 0,09%, lo que confirma la idoneidad del mismo.

1.3.3 *Algoritmo basado en la detección de logotipos.*

Los sistemas de detección de logotipos⁵ pretenden reconocer la existencia de los mismos sobre el flujo de video del programa. Estos suelen existir en cualquiera de las 4 esquinas de la pantalla, lo que suele suceder con mayor probabilidad en la parte inferior derecha de la pantalla, por tanto estas técnicas de detección permiten distinguir los programas que tienen la presencia de logotipos. A continuación se analizarán dos técnicas utilizadas para la detección de logos.

En la técnica de **detección por contornos estables** analizada se emplearán solo la detección de logotipos opacos o transparentes.

En (**Albiol, 2004**) se plantea que los logotipos pueden ser clasificados en tres categorías.

⁵ Logotipo: coloquialmente conocido como **logo**, es un elemento gráfico, que identifica a una persona, empresa, institución o producto.

- Logotipos con animación.
- Logotipos opacos.
- Logotipos transparentes.

Los logotipos opacos o transparentes son el tipo menos difícil de distinguir según Albiol. El valor de los píxeles se mantiene relativamente constante y una vez detectados suelen ser fáciles de rastrear, sin embargo, logotipos que parecen ser opacos a veces son realmente transparentes. En este caso, se propone un método en el que los fotogramas de video se buscan en un área que contiene contornos estables. Si un área general tiene un contorno estable durante un largo periodo de tiempo y un transcurso de muchos fotogramas de video, se puede suponer que existe un logotipo (**Henneberger, 2008**) (**Albiol, 2004**).

La **búsqueda de patrones** es una técnica de análisis que se puede aplicar en detección de objetos, reconocimiento, seguimiento y correspondencia, dada una imagen (un patrón o modelo) encontrar sus apariciones dentro de otra imagen mayor. No se buscan sólo las apariciones “exactas”, sino permitiendo cierto grado de variación respecto al patrón (**Albiol, 2004**).

Para realizar la búsqueda se utiliza el método **template matching** el cual consiste en la comparación de plantillas, una de las principales aplicaciones del método es la detección de objetos.



Figura 2: Ejemplos de Logotipos que identifican a diferentes televisoras (Disey Publicidad, 2011).

Se concluye diciendo que la ventaja de la técnica antes planteada es que una vez que un logotipo de una estación de radiodifusión es estable durante un programa, el algoritmo simplemente debe detectar la ausencia de la presencia del mismo para determinar qué secuencia de video es parte de un segmento

comercial. Este algoritmo solo sería eficiente frente a estaciones que mantienen una red de logotipos durante sus transmisiones, en caso contrario la técnica sería un error.

1.3.4 *Descriptores de Audio*

Recientemente varios investigadores han comenzado a estudiar el potencial del análisis de la señal de acompañamiento de audio, para la clasificación de video de la escena **(Leonardi, 1997) (S. Pfeiffer, 1998) (Tewfik, 1997) (Y. Wang, 1997)**. Esto es factible ya que el audio de un spot comercial es muy diferente al de un partido de fútbol o al de un informe de prensa. La información de audio por sí sola no puede ser suficiente para comprender el contenido de la escena, debe ser analizada en conjunto con la información visual.

Los descriptores de audio extraen de los materiales audiovisuales una serie de características que los identifican en forma de un vector de características. Este conjunto de características ocupa menos espacio que el objeto, ya que se almacenan de forma numérica.

El análisis de audio basado en el cálculo requiere mucho menos espacio de almacenamiento y puede ser utilizado en una etapa del proceso previo antes del análisis global de la información visual.

La comprensión de audio se puede basar en las características de tres capas: las características acústicas de bajo nivel, las firmas de audio de nivel intermedio asociado con diferentes objetos de resonancia y los modelos de alto nivel semántico de audio en las clases de escena diferentes.

En la capa de características acústicas, se analizan las de bajo nivel de género, tales como volumen y ancho de banda de una señal de audio. Esto constituye la etapa previa al procesamiento que se requiere en cualquier sistema de procesamiento de audio. En la capa de la firma acústica, se determina el objeto que produce un sonido particular.

Los sonidos producidos por objetos diferentes tienen diferentes firmas. Por ejemplo, cada instrumento musical tiene su propia respuesta al impulso al ser golpeados. En la capa de alto nivel se hacen uso de algunas reglas antes una conocida semántica sobre la estructura de audio en diferentes tipos de escenas, normalmente existe sólo el habla en el informe de noticias, pronóstico del tiempo, pero en un comercial, por lo general siempre hay una música de fondo. **(Liu, 1998)**.

1.3.5 Reconocimiento de texto en video.

Un OCR⁶ permite capturar e identificar texto incluido dentro de imágenes, reconocer ópticamente caracteres y es capaz de hacer ficheros de texto con capacidad de búsqueda mediante el reconocimiento de las palabras individuales dentro de un archivo imagen. Actualmente estos programas son capaces de reconocer no solo el texto en sí, sino también el estilo y formato de este (**CVISION Technologies, 1998**) .

El sistema OCR separa la imagen correspondiente a un caracter y compara el mismo con una base de caracteres para determinar su correspondiente código ASCII⁷ o bien Unicode⁸. OCR funciona utilizando un tipo de inteligencia artificial, conocido como reconocimiento de formas y estructuras, que identifica los caracteres individuales de texto, incluidos signos de puntuación, espacios y finales de línea.

Para su correcto funcionamiento es necesario que la imagen de donde provenga dicho texto esté en las mejores condiciones posibles para que pueda reconocer correctamente los caracteres. Hay que tener en cuenta que factores tales como un texto borroso (aunque sea solo ligeramente), falta de una parte de una letra, entre otros, dificultan el correcto reconocimiento de este texto. El promedio de efectividad de los programas OCR, en perfectas condiciones, ronda el 90%, disminuyendo al disminuir la calidad de la imagen (**CVISION Technologies, 1998**) .

OCR posee diferentes opciones como son:

1. Auto girar: Si el papel se ha digitalizado de lado o al revés, al seleccionar Auto girar se vuelve a colocar la página en la posición correcta.
2. Idioma de OCR: La selección de un idioma diferente en esta lista, permite digitalizar documentos en un idioma distinto del predeterminado en el equipo.

⁶ OCR: algoritmo de reconocimiento de texto.

⁷ ASCII: código de caracteres basado en el alfabeto latino.

⁸ Unicode: estándar de codificación de caracteres diseñado para facilitar el tratamiento informático, transmisión y visualización de textos de múltiples lenguajes y disciplinas técnicas.



Figura 3: Ejemplo de texto publicitario en video (Cade, 2007).

Se concluye diciendo que los OCR pueden aportar valiosas pistas para la detección de productos y servicios anunciados por los comerciales, ya que permite extraer palabras claves. Por lo general cuando se hace publicidad mediante texto publicitario, el mismo posee un tamaño grande, para que llame la atención al ojo humano ya sea mediante los colores, la forma, entre otras alternativas tomadas por los anunciantes. El texto puede ser detectado por los programas actuales de OCR, que reconocen el estilo y el formato del mismo, además permiten leer texto en cualquier dirección, con la opción de auto girar y hacer búsquedas en cualquier idioma. Se debe tener en cuenta que uno de los requisitos básicos para que un programa OCR funcione correctamente es la presencia de una imagen de gran calidad.

1.4 Conclusiones Parciales.

Los métodos analizados en este capítulo para la detección de cortes comerciales en transmisiones televisivas se basan en una serie de características mostradas por los spots publicitarios. Los algoritmos tienen un por ciento muy bajo en cuanto a la detección de falsos positivos, los mismos pueden ser muy rápidos y precisos según el grado de calidad de la imagen que se esté analizando, sin embargo existen otros algoritmos que su resultado depende solo de tendencias mostradas por las televisoras, siendo esto un factor que no siempre permite lograr un resultado satisfactorio. Mientras mayor sea la combinación de las técnicas de detección mayor será la precisión y rapidez en tomar una decisión con respecto al resultado, actualmente no existe una técnica de detección que se base en un número significativo de características.

CAPÍTULO 2. Solución Propuesta.

En el presente capítulo se define la solución a la problemática planteada en un inicio, a partir de un estudio detallado del estado del arte de las técnicas eficientes para detección de publicidad en materiales audiovisuales existente en el mundo. Se caracterizan cada una de las fases del proceso de detección de anuncios del algoritmo propuesto, el cual combina técnicas tanto de procesamiento de audio como de video, con el fin de lograr un resultado que gane en precisión.

2.1 Especificación de requisitos del algoritmo propuesto.

Luego de un estudio detallado del problema de la investigación y a partir de los objetivos planteados para resolver el mismo, es necesario que la solución propuesta cumpla con una serie de condiciones y capacidades de interés para el usuario.

2.1.1 Requisitos funcionales.

RF1. Detectar material publicitario.

El algoritmo propuesto debe ser capaz de determinar los cortes comerciales existentes en un material audiovisual.

1.1. Detectar audio fuera de rango.

El algoritmo propuesto debe detectar el aumento considerado del volumen dentro de una media.

1.2. Detectar shots variables.

El algoritmo propuesto debe detectar los cambios de tomas que se produzcan dentro del material audiovisual y llevar el control de los mismos para así determinar si se está o no en presencia de secuencias candidatas.

1.3. Detectar texto publicitario.

El algoritmo propuesto debe determinar cuándo se está en presencia de algún texto dentro del material audiovisual que represente promoción a un determinado servicio o producto.

2.2 Proceso de detección de anuncios.

El algoritmo propuesto combina tres técnicas para la detección de materiales publicitarios, la primera (T-0) es la encargada de detectar los cambios de toma existentes dentro de la secuencia de video. Se prioriza debido a que está basada en una característica muy significativa de los comerciales, la cual tiene un porcentaje elevado de efectividad. Una vez comprobada si el material audiovisual no cumple con dicha técnica se descarta la posibilidad de estar en presencia de un spots publicitario, si ocurre lo contrario se comprobará la técnica 1(T-1) la cual detecta aumentos considerables en el volumen. Será la segunda en el análisis debido a que la detección puede ser rápida, pues la tendencia de este aumento es solamente al principio de cada anuncio dentro del bloque. Si la técnica arroja un resultado positivo el material pasa a formar parte de una secuencia candidata, de lo contrario se comprueba la técnica 2(T-2), la cual detecta todos los textos existentes en la media y verifica que estén dentro de la misma con el fin de promocionar algún servicio o producto. Si el resultado arrojado es negativo la producción no es clasificada como un spot, si ocurre lo contrario el algoritmo identificará un spot publicitario.

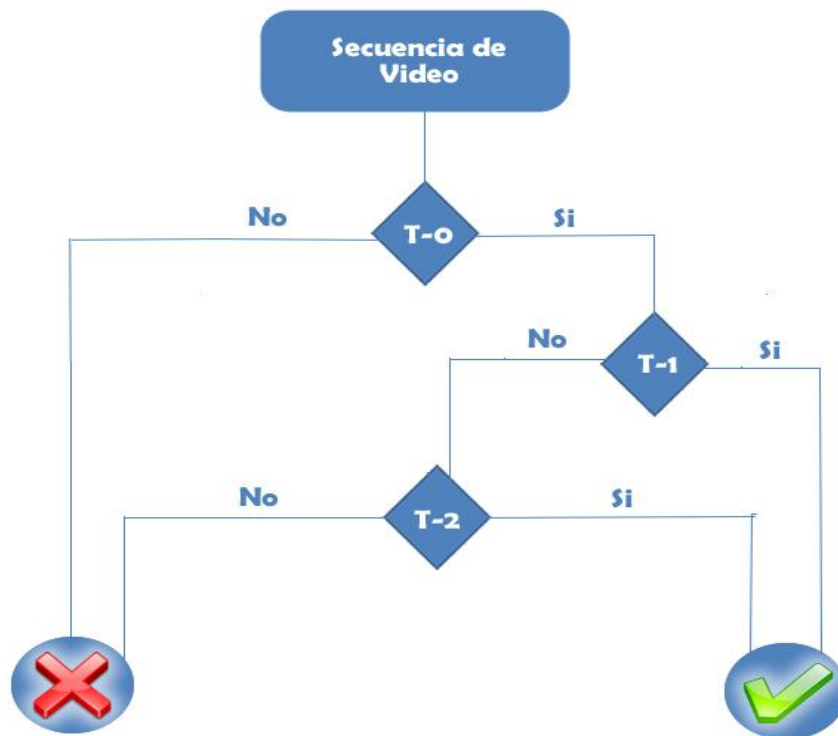


Figura 4: Diagrama de procesos del algoritmo propuesto.

2.3 Detección de fotogramas claves.

El procedimiento empleado parte de definir una métrica que permite establecer algún tipo de distancia entre dos fotogramas consecutivos, i e $i+1$. Se usan las métricas basadas en el histograma de la imagen $H_i(j)$ o $H_{i+1}(j)$, (j varía típicamente de 0 hasta un cierto nivel máximo) ya que presentan un bajo coste computacional, además de ofrecer un procedimiento de detección bastante robusto (Bueno, 2009).

2.3.1 Cálculo de la intensidad media de la imagen.

Una vez calculado el histograma se calcula la **diferencia de intensidad media de la imagen**. Para ello se calcula la **intensidad media**, de la forma siguiente:

Intensidad media de la imagen.

$$AI_{i=\frac{1}{NM}} \sum_{j=0}^{j_{max}} j \cdot H_i(j)$$

Donde N y M son respectivamente el número de píxeles por fila y columna de la imagen, como se ilustra en la figura 1

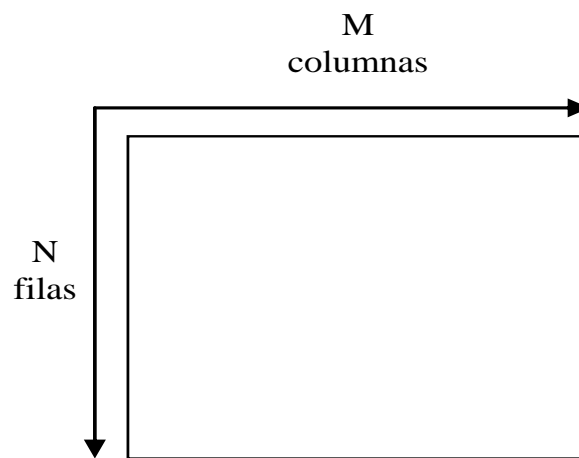


Figura 5: Número de píxeles por fila (N) y columna (M) de la imagen (Bueno, 2009).

Para calcular la intensidad se utiliza un indicador adicional relacionado con el alto nivel de acción en los comerciales, con el objetivo de lograr una mayor precisión en cuanto a un resultado, este está relacionado con el **índice de cambio del borde**, el cual describe el número de píxeles del borde, los materiales que tienen un alto grado de acción se caracterizan por tener mayor nivel de intensidad.

El algoritmo de detección de bordes Canny es usado para detectar todos los bordes existentes en una imagen. Este algoritmo está considerado como uno de los mejores métodos de detección de contornos mediante el empleo de máscaras de convolución y basado en la primera derivada. Los puntos de contorno son zonas de píxeles en las que existe un cambio brusco de nivel de gris.

Algoritmo de Canny.

Uno de los métodos relacionados con la detección de bordes es el uso de la primera derivada, la misma toma valor cero en todas las regiones donde no varía la intensidad y tiene un valor constante en toda la transición de intensidad. Por tanto un cambio de intensidad se manifiesta como un cambio brusco en la

primera derivada (**Cruz, 2001**), característica que es usada para detectar un borde, y en la que se basa el algoritmo de Canny.

El algoritmo de Canny consiste en tres pasos:

- Obtención del gradiente: en este paso se calcula la magnitud y orientación del vector gradiente en cada píxel.
- Supresión no máxima: en este paso se logra el adelgazamiento del ancho de los bordes obtenidos con el gradiente, hasta lograr bordes de un píxel de ancho.
- Histéresis de umbral: en este paso se aplica una función de histéresis basada en dos umbrales, con este proceso se pretende reducir la posibilidad de aparición de contornos falsos.

Obtención del gradiente

Para la obtención del gradiente, lo primero que se realiza es la aplicación de un filtro gaussiano a la imagen original, con el objetivo de suavizar la imagen y tratar de eliminar el posible ruido existente. Sin embargo, se debe tener cuidado de no realizar un suavizado excesivo, pues se podrían perder detalles de la imagen y provocar un pésimo resultado final. Este suavizado se obtiene promediando los valores de intensidad de los píxeles en el entorno de vecindad con una máscara de convolución de media cero y desviación estándar σ . Una vez que se suaviza la imagen, para cada píxel se obtiene la magnitud y módulo (orientación) del gradiente, obteniendo así dos imágenes. (**Rebaza, 2010**).

Pasos a seguir para la detección del gradiente:

1. Suavizar la imagen I con H mediante un filtro gaussiano y obtener J como imagen de salida.
2. Para cada píxel (i, j) en J , obtener la magnitud y orientación del gradiente basándose en las siguientes expresiones:

El gradiente de una imagen $f(x, y)$ en un punto (x, y) se define como un vector bidimensional dado por la ecuación:

$$G[f(y, x)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix}$$

Siendo un vector perpendicular al borde, donde el vector G apunta en la dirección de variación máxima de f en el punto (x, y) por unidad de distancia, con la magnitud y dirección dadas por:

$$|G| = \sqrt{G_x^2 + G_y^2} = |G_x| + |G_y|$$

$$\phi(x, y) = \tan^{-1} \frac{G_y}{G_x}$$

Supresión no máxima al resultado del gradiente.

Las dos imágenes generadas en el paso anterior sirven de entrada para generar una imagen con los bordes adelgazados. El procedimiento es el siguiente: se consideran cuatro direcciones identificadas por las orientaciones de 0° , 45° , 90° y 135° con respecto al eje horizontal. Para cada píxel se encuentra la dirección que mejor se aproxime a la dirección del ángulo de gradiente

Se obtiene E_m a partir de la magnitud de gradiente y E_0 a partir de la orientación, de acuerdo a las expresiones anteriores. Para ello se observa si el valor de la magnitud de gradiente es más pequeño que al menos uno de sus dos vecinos en la dirección del ángulo obtenida en el paso anterior. De ser así se asigna el valor 0 a dicho píxel, en caso contrario se asigna el valor que tenga la magnitud del gradiente.

La salida de este segundo paso es la imagen I_n con los bordes adelgazados, es decir, $E_m(i, j)$ después de la supresión no máxima de puntos de borde (Rebaza, 2010).

Histéresis de umbral a la supresión no máxima.

La imagen obtenida en el paso anterior suele contener máximos locales creados por el ruido. Una solución para eliminar el ruido es la histéresis del umbral. El proceso consiste en tomar la imagen obtenida del paso anterior, tomar la orientación de los puntos de borde de la imagen y tomar dos umbrales, el primero más pequeño que el segundo. Para cada punto de la imagen se debe localizar el siguiente punto de borde

no explorado que sea mayor al segundo umbral. A partir de este punto seguir las cadenas de máximos locales conectados en ambas direcciones perpendiculares a la normal del borde siempre que sean mayores al primer umbral. Así se marcan todos los puntos explorados y se almacena la lista de todos los puntos en el contorno conectado. Es así como en este paso se logra eliminar las uniones en forma de Y de los segmentos que confluyen en un punto (**Rebaza, 2010**) .

La distancia entre dos fotogramas se mide como la diferencia de intensidad media entre ambos:

Diferencia entre dos fotogramas.

$$AID_i = |AI_i - AI_{i-1}|$$

Se considera que se produce un cambio de escena si la magnitud anterior supera un umbral previamente establecido.

2.3.2 Cálculo del umbral.

Se calcula el umbral mediante el método Otsu, el cual elige el umbral óptimo, maximizando la varianza⁹ entre clases realizando una búsqueda exhaustiva. Otsu fue uno de los mejores métodos de selección de umbral para imágenes del mundo real. La importancia del método radica en que es automático, es decir, no necesita supervisión humana ni información previa de la imagen antes de su procesamiento (**Otsu, 1979**) .

Si la tasa de cambios de escena de un material es superior a 15 cambios por minuto para su totalidad y en un minuto del análisis supera los 30 pasa a formar parte de una secuencia candidata. Esta decisión se toma luego de realizar varias pruebas y percibir que el promedio de cambios para una muestra de 10 materiales que no representan publicidad es de 14 por minuto y para materiales publicitarios sobre la misma muestra es de 32.5.

⁹ Varianza: muestran la variabilidad de una distribución, indicando por medio de un número, si las diferentes puntuaciones de una variable están muy alejadas de la media de dicha variable aleatoria.

2.4 Detección de picos elevados de volumen.

Existen numerosos descriptores que pueden ser utilizados para la clasificación de señales de audio. Generalmente son clasificados en dos categorías: descriptores de dominio de tiempo y de dominio de frecuencia. Estos descriptores pueden ser extraídos en dos niveles diferentes: a nivel de ventana o a nivel de clip (**Liu, 1998**) .

2.4.1 Extracción de Características del audio.

Para la extracción de características del audio el mismo se divide en frames o ventanas entre 1 y 60 segundos de duración aproximadamente. Si estas ventanas son muy grandes, la precisión con respecto a la caracterización de los cambios de la señal de audio no será la mejor, mientras que si son muy pequeñas no se podrán extraer descriptores válidos. Luego de obtener el valor de la muestra de cada pista de audio se procede a graficar cada señal, con el objetivo de mostrar la variación temporal de la magnitud.

Se debe tener en cuenta que el volumen de una señal de audio depende de la ganancia del sistema de grabación y digitalización (**Liu, 1998**) .

2.5 Detección de publicidad por texto en video.

El texto publicitario tiene la responsabilidad de brindar a los telespectadores mayor claridad acerca del producto o servicio que se está promocionando. El texto en un video es capaz de transmitir mensajes al mismo tiempo del audio, ampliando la información a ser divulgada.

La mayor parte de los algoritmos para reconocer escritura están escritos a partir de imágenes binarias, por lo que se hace conveniente el paso de una imagen a color a una binaria, además esto permite reducir el volumen de los datos a tratar (**Nagy, 1990**) .

2.5.1 Binarización de la Imagen.

La binarización de una imagen digital consiste en convertir la imagen digital en una imagen en blanco y negro, de tal manera que se preserven las propiedades esenciales de la imagen. Uno de los métodos para poder binarizar una imagen digital es mediante el histograma de dicha imagen. A través del histograma se obtiene una gráfica donde se muestran el número de píxeles por cada nivel de gris que aparecen en la imagen. Para binarizar la imagen, se deberá elegir un valor adecuado dentro de los niveles de grises (umbral), de tal forma que el histograma forme un valle en ese nivel. Todos los niveles de grises menores al umbral calculado se convertirán en negro y todos los mayores en blanco. Se usa la binarización con selección del nivel, mediante el método de Otsu (**Otsu, 1979**), que persigue minimizar la varianza intraclase de la luminosidad de los píxeles blancos y negros (**Elliman, 1982**).

2.5.2 Fragmentación o segmentación de la imagen.

La generación de metodologías para la extracción de información a partir de regiones u objetos de interés de la imagen posibilita una mejor comprensión del procesamiento de la misma. La segmentación es una importante tarea hacia un sistema de análisis de la imagen a alto nivel, esta permite la cuantificación y visualización de las regiones u objetos de interés y de su calidad depende en gran medida el resultado final de la interpretación automática de una secuencia de imágenes. No existe un método universal para realizar la segmentación de una imagen, este proceso está íntimamente ligado a la tarea que se desee resolver que satisfaga las expectativas del observador.

Una vez obtenida la imagen binaria se deberá fragmentar o segmentar en las diferentes componentes conexas que la componen. La fragmentación o segmentación es la operación que permite la descomposición de un texto en diferentes entidades lógicas. Estas entidades lógicas deben ser lo suficientemente invariables, para ser independientes del escritor y lo suficientemente significativas para su reconocimiento. Este proceso constituye una de las mayores dificultades del reconocimiento, y se hace necesario para poder reconocer cada uno de los caracteres de la imagen binaria.

Se propone utilizar el método de detección de bordes Algoritmo de Canny antes explicado en el epígrafe 2.4.1 para el cálculo de la intensidad de una imagen.

2.5.3 Adelgazamiento de las componentes.

Una vez obtenida las componentes conexas de la imagen, se deberá realizar un proceso de adelgazamiento de cada uno de ellas. El procedimiento de adelgazamiento consiste en ir borrando sucesivamente los puntos del borde de cada componente conexa, de forma que se preserve su topología.

Las condiciones exactas que determinan si un punto se puede borrar están relacionadas con el concepto de punto simple (un píxel negro P del borde de la imagen se considera simple si el conjunto de los vecinos en negro de P tienen exactamente una componente conexa que es adyacente a P) y punto final (un punto es final si tiene exactamente un vecino negro; un punto final no es más que un punto extremo de la imagen). Un punto del borde de cada componente se puede eliminar si es simple y no es final.

El borrado de puntos debe seguir un esquema de barridos sucesivos para que la imagen siga teniendo las mismas proporciones que la original y conseguir así que no quede deformada. El borrado en cada rastreo debe hacerse en paralelo, es decir, señalar los píxeles borrables para eliminarlos todos a la vez.

2.5.4 Comparar con patrones.

Una vez realizado el adelgazamiento de la imagen, se realiza la comparación con patrones o plantillas. En esta etapa se debe comparar los caracteres obtenidos con unos caracteres teóricos almacenados en una base de datos. En esta base de datos se almacena, por tanto, todos los patrones o plantillas (**Elliman, 1982**).

La base de datos consta de dos tablas, la **tabla de valores predeterminados** que contiene al **campo Palabras claves** y la **tabla de Valores dinámicos** que contiene al **campo Texto extraído**. Estos campos almacenarán respectivamente textos que representan publicidad (nombres de productos, direcciones electrónicas, la palabra publicidad) que son muy frecuentes encontrar en medio de un corte comercial y el texto encontrado en las producciones audiovisuales. Posteriormente se establecen comparaciones entre cada uno de los atributos de las tablas.

2.6 Conclusiones Parciales.

En el capítulo anterior se ha mostrado la propuesta de solución al problema identificado en un comienzo de la investigación, cumpliendo con los requisitos funcionales especificados. Se ha desglosado el análisis de las medias en tres aspectos fundamentales, los cuales aportan en gran medida al resultado final, por lo que se está en presencia de un algoritmo híbrido que utiliza un árbol de decisión para la detección de publicidad. El algoritmo propuesto se basa fundamentalmente en la detección por cambios de toma, debido a que es la técnica de mayor por ciento de efectividad de las fusionadas con un 96%.

Capítulo 3. Resultados Obtenidos.

En el presente capítulo se abordan los temas relacionados con los resultados obtenidos luego de la aplicación del algoritmo propuesto con anterioridad. Las técnicas que combina la propuesta de solución son probadas con distintas tecnologías, debido a las ventajas mostradas por estas según las particularidades de cada técnica. Se describe el hardware utilizado para la realización de las pruebas y se arriba a conclusiones con respecto a los resultados del por ciento de efectividad del algoritmo para la detección de publicidad en materiales audiovisuales, así como el por ciento de fallo.

3.1 Herramientas utilizadas.

Para el desarrollo de la implementación de las técnicas fusionadas por el algoritmo, se hace necesario estudiar a profundidad una serie de herramientas que facilitan en gran medida cada uno de los pasos a seguir en el logro del resultado. Para la realización de las pruebas prácticas del algoritmo se utiliza el software matemático MATLAB R2008a, el framework Qt bajo código C++ y su entorno de desarrollo integrado Qt Creator, así como la librería OpenCV sobre sistema operativo GNU/Linux Ubuntu. Para la utilización de estas herramientas se hace necesario un Hardware con las siguientes características: una computadora de escritorio con procesador Intel Core Duo a 1.86 Ghz y 1 Gb de memoria RAM o superior. A continuación se mencionan una serie de aspectos relacionados con las tecnologías utilizadas, los cuales justifican el uso de las mismas.

3.1.1 MATLAB.

MATLAB¹⁰ es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio, permite la representación de datos, funciones e imágenes mediante la manipulación de matrices, la implementación de algoritmos, la creación de interfaces de usuario y la comunicación con programas en otros lenguajes .

El nombre MATLAB surge dado que el elemento básico de dato es una matriz, la cual no requiere de dimensionamiento. Tiene un ambiente de cómputo de alta ejecución numérica y de visualización, que

¹⁰ MATLAB: matrix laboratory

integra el análisis numérico, cálculo de matrices, procesamiento de imágenes y diseño gráfico, en un ambiente sencillo de utilizar.

La Herramienta tiene implementado varios algoritmos para el procesamiento de imágenes dentro de los cuales se encuentran: mejora de propiedades de la imagen como contraste y color, transformaciones geométricas (rotación, traslación, escalado), filtros (suavizado de bordes y morfología), transformación binaria, detección de bordes y contornos, posición y coincidencias de formas, reconocimiento de modelos geométricos (líneas, círculos) y formas entre otros (**Jalón, 2005**).

Luego de profundizar en varias de las funciones que posee la herramienta se llega a la conclusión que la misma aporta ciertas facilidades en cuanto al procesamiento de videos en general, por lo que dos de las técnicas combinadas por el algoritmo propuesto son probadas en MATLAB, la detección de texto en video y picos elevados de audio. Se decide que sean implementadas en esta herramienta ya que ofrece muchas facilidades graficas y funciones que tributan en la obtención del resultado. No es un problema realizar las en el hardware con que se cuenta, ya que el texto no tiene una frecuencia de aparición muy alta por fotogramas y el audio una vez extraído del video disminuye su tamaño en disco.

Según el algoritmo propuesto para la implementación de la técnica de detección por fotogramas claves, se requiere de un hardware con ciertas características con el que no se cuenta actualmente en el proyecto Captura y Catalogación de Medias (SCCM) del departamento de Señales Digitales de la Universidad de Ciencias Informáticas. Las pruebas tendrán que efectuarse a materiales audiovisuales de un tiempo de duración no menor a 15 minutos, anteriormente explicado. El análisis en MATLAB para cada una de las medias se realizará a nivel de fotograma y un video de esta duración posee una cantidad considerable de los mismos y por consiguiente de cambios de tomas, por lo que se hace necesario el uso de otro lenguaje de programación que solucione el problema con los medios disponibles (C++).

3.1.2 C++.

El lenguaje de programación C++ fue diseñado a mediados de los años 1980, con la intención de extender al exitoso lenguaje de programación C, con mecanismos que permitan la manipulación de objetos. Posteriormente se añadieron facilidades de programación genérica, con la particularidad de

redefinir los operadores (sobrecarga de operadores) y de poder crear nuevos tipos de datos que se comporten como tipos fundamentales **(Jiménez, 2008.)** .

C++ es un lenguaje versátil y potente que ha llegado a ocupar el primer puesto como herramienta en el desarrollo de aplicaciones entre los programadores. Permite el control de memoria y una capacidad de programación de alto y bajo nivel. Algunas de las principales ventajas que posee este lenguaje es la versatilidad, la portabilidad pues está estandarizado y un mismo código fuente se puede compilar en diversas plataformas, es uno de los lenguajes más rápidos en cuanto a ejecución, los programas que genera son considerados compactos, rápidos y consumen menor cantidad de recursos, ya que permite la creación de punteros a objetos entre otras. Muchos sistemas Operativos, compiladores e intérpretes han sido escritos en C++ (el propio Windows y Java). El código es portable, un programa ANSI en C o C++ podrá ejecutarse en cualquier máquina y bajo cualquier sistema operativo. Y si es necesario, proporcionan un acceso a bajo nivel de hardware sólo igualado por el ensamblador. Este lenguaje de programación ofrece la librería OpenCV, la misma facilita el trabajo con el procesamiento de imágenes **(González, 2008)**.

3.1.3 OpenCV.

OpenCV es una biblioteca open source para C/C++ para procesamiento de imágenes y visión computarizada, desarrollada inicialmente por Intel. Su primera versión estable fue liberada en 2006. En Octubre de 2009, se liberó el segundo release mayor OpenCV v2.

Disponible en Linux, Mac, y Windows. Tiene estructuras básicas de datos para operaciones con matrices y procesamiento de imágenes. Permite visualizar datos muy sencillamente y extraer información de imágenes y videos. Tiene funciones de captura y presentación de imágenes **(Furfaro, 2010)** .

OpenCV es una librería que contiene muchas estructuras de datos, permite el reconocimiento de patrones, procesamiento 2D-3D en tiempo real, calibración de cámaras, entre otras. Además permite trabajar con múltiples formatos de video como AVI y contiene una interfaz gráfica de usuario que permite una fácil y rápida forma de interactuar y visualizar las imágenes. Con el uso de la librería se hace más fácil realizar la implementación y las pruebas a la técnica de detección de publicidad por fotogramas claves.

3.1.4 Framework de desarrollo QT.

Qt es un framework multiplataforma, fue creado en 1991 por la compañía Quasar Technologies, el mismo está escrito en el lenguaje de programación C++, pero es posible utilizarlo con otros lenguajes. En un principio sólo ofrecía bibliotecas de código para la creación de interfaces gráficas de usuario. En la actualidad el Framework está compuesto por una serie de módulos que proveen funcionalidad específica a través de una biblioteca de clases multiplataforma como: Bases de datos, XML, multimedia, comunicación en red, OpenGL, entre otras.

El propósito del framework es permitir a los desarrolladores construir aplicaciones multiplataforma a partir de una misma base de código de manera rápida y sencilla. El mismo provee poderosas herramientas de desarrollo, entre ellas destaca un completo entorno de desarrollo (IDE) multiplataforma muy completo, llamado Qt Creator.

Este entorno de desarrollo está integrado con el Qt Designer para ayudarle a diseñar formas de la interfaz de usuario, como lo haría con la versión independiente. La integración de Qt Designer incluye también la gestión y finalización del proyecto de código. Qt Creator abastece no solo a los desarrolladores que están acostumbrados a utilizar el ratón, sino también a los desarrolladores que se sienten más cómodos con el teclado. Una amplia gama de métodos abreviados de teclado y navegación están disponibles para ayudar a acelerar el proceso de desarrollo de su aplicación.

Como características tiene el reconocimiento de métodos, la facilidad de creación de formularios y amplia documentación On-line, así como opciones que facilitan el desarrollo de aplicaciones. Qt Creator utiliza la biblioteca de software Qt para crear interfaces gráficas de usuario, incluye un editor de texto con autocompletado, sistemas de depuración, integración con sistemas de control de versiones.

Qt es utilizado por gran cantidad de dispositivos, algunos de ellos son: computadoras de escritorio, teléfonos celulares, lectores electrónicos, impresoras, computadoras de automóvil, entre otras. Algunas compañías de renombre que han utilizado Qt son: Google, HP, Samsung y Asus. Aplicaciones como: Google Earth, Skype, Adobe Photoshop Album, el reproductor multimedia VLC, la herramienta de modelado Maya, VirtualBox entre muchas otras hacen uso del mismo.

3.2 Detección de fotogramas claves.

Como primer paso del proceso de detección de fotogramas claves se carga el video mediante la función CvCapture* capture = cvCreateFileCapture("/home/lvelin/Escritorio/Spots1.avi"), especificando la dirección donde se encuentra el video, la cual crea un objeto donde quedará guardado el mismo. Una vez almacenado, se procede a extraer fotograma por fotograma, este proceso se realiza mediante la función frame2 = cvQueryFrame(capture). Luego se calcula la diferencia (distancia) de histogramas entre cada una de las imágenes y se almacenan estos valores en una lista. Posteriormente se establece un umbral calculado a partir de la media entre cada una de las diferencias almacenadas. Los valores de histograma que estén por encima de este representarán cambios de toma, los que estén por debajo no. Luego de este proceso se almacenan en una lista todos los cambios existentes. Como resultado final se obtiene la cantidad de cambios existentes en el material analizado, lo que permite identificar si la técnica es válida o no.

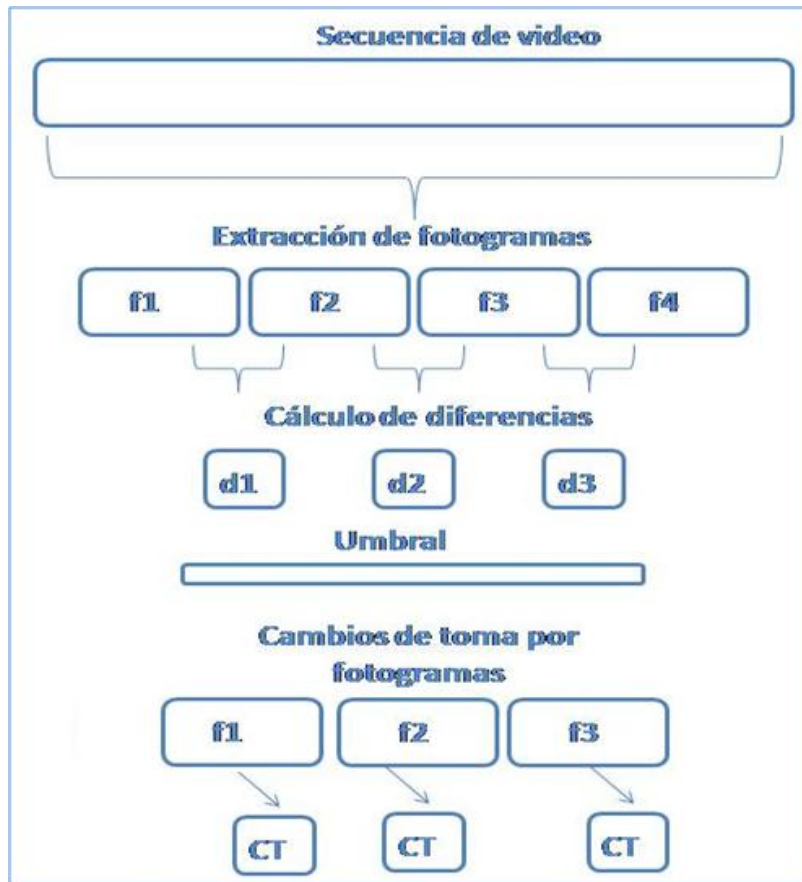


Figura 6: Diagrama para la detección de cambios de toma.

3.3 Detección de aumento de audio.

La herramienta propuesta para aplicar esta técnica hasta la versión que se tiene para el sistema operativo Linux es R2008a y la misma para el tratamiento del audio permite interactuar solo con los formatos wav y au. Las extensiones .wav son tratadas mediante la librería wavread, $[y,fs,nbits] = wavread(filename)$ la cual permite cargar el audio y obtener información como número de bits que lo compone, longitud del mismo, frecuencia de muestreo. Con respecto a esta extensión también se puede crear una señal de audio, los datos de entrada para la construcción son los mismos a los datos de salida de la librería wavread, la sintaxis de la misma es $wavwrite(y,fs,nbits,filename)$. Las extensiones .au permiten crear y cargar audio de forma similar a las librerías antes mencionadas, solo diferenciándose en el tipo de formato.

Las bibliotecas antes mencionadas solo soportan audio con los formatos especificados, por lo que representa una limitación si el audio con el que se necesita interactuar forma parte de un video, es por esto que se utiliza la librería `mmread`.

Para detectar si ocurre o no aumento dentro de las pistas de audio de un video se captura el mismo mediante la librería `[video, audio] = mmread('video.avi')`, la cual logra separar el audio del video, devolviendo cada uno en una variable independiente. Solo se tomará el audio puesto que el hardware con que se cuenta no permite cargar videos de larga duración. Una vez extraída la señal se obtiene la duración total de la misma en segundos mediante la función `duracionS=audio.totalDuration`, la cantidad total de frames de audio es almacenada en `totalFrame=audio.nrFramesTotal`, el valor de cada una de las muestras que la componen se observa en `audio.data`, la frecuencia de muestro mediante `fs=audio.rate` y el tiempo en que ocurre cada frames queda guardado en la variable `time=audio.times`.

Luego se grafica el audio por intervalos no mayores de 60 segundos con la función `plot(audio)`, para obtener el efecto visual del mismo en cada instantes donde existe o no publicidad, demostrando que la amplitud de la señal es mayor en casos donde la producción forma parte de un spots y los valores de los picos de audio son mucho mayor ver figura 7 y figura 8.

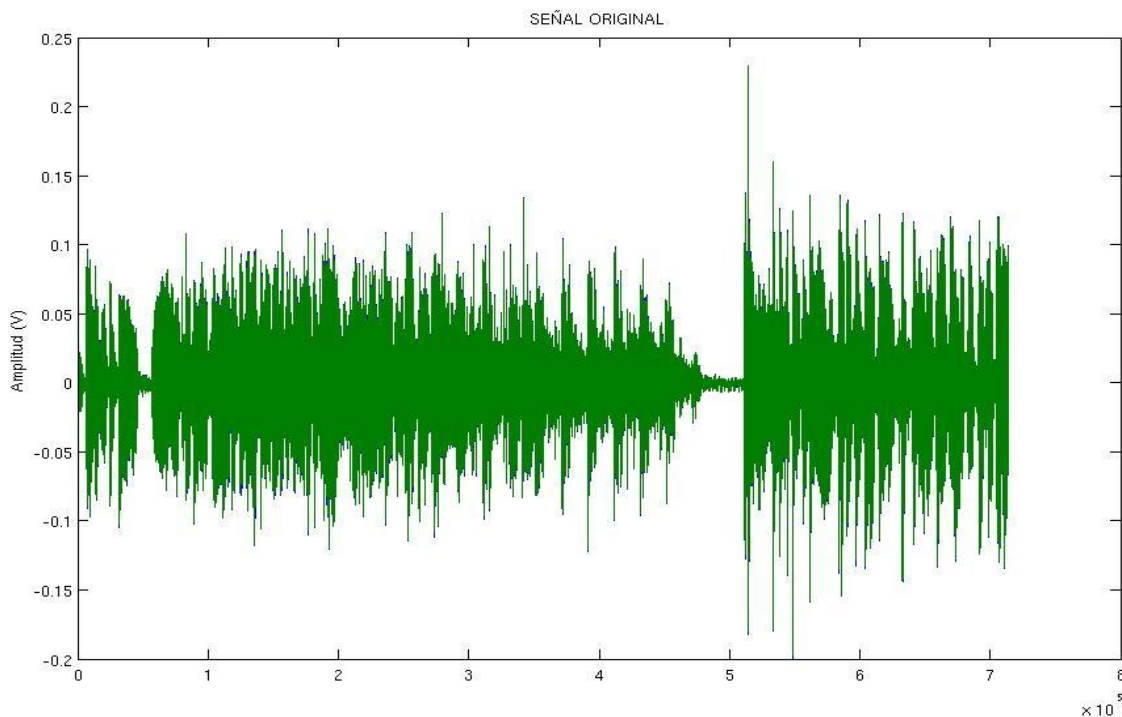


Figura 7: Representación de audio publicitario.

3.4 Detección de texto en video.

El experimento se realiza sobre imágenes de formatos (JPEG extensión .jpg y BPM extensión .bmp) aunque el software matemático puede procesar imágenes en formatos como TIFF, GIF, PNG y XWD.

En la fase de preprocesamiento de la imagen digital se tienen en cuenta los procedimientos a realizar en la solución propuesta. Para ello en la primera etapa se adquiere la imagen mediante la función `imread`, si se desea visualizar la misma en MATLAB, se ejecuta la función `imshow`, como lo muestra la figura 7, esta función puede ser usada siempre que se necesite ver los resultados de alguna transformación, tendrá como parámetro la imagen a ser visualizada con su correspondiente extensión. Luego si la imagen está en el plano RGB lo cual se prueba con la función `length(size(imagen))==3`, se convierte a escala de grises mediante la función `rgb2gray(imagen)` ver resultado de la ejecución de la función figura 8. Esta es una escala empleada en la imagen digital en las que el valor de cada píxel posee un valor equivalente a una

graduación de gris. Las imágenes representadas están compuestas de sombras de grises, que van desde el negro más profundo, variando gradualmente en intensidad de grises hasta llegar al blanco.

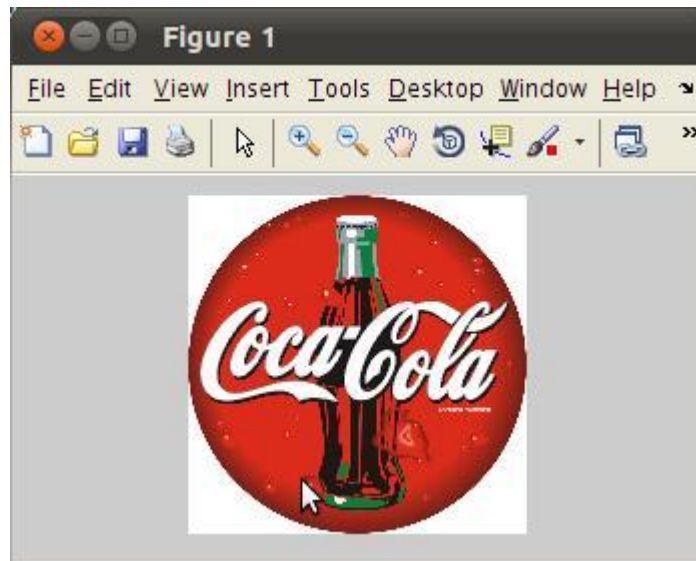


Figura 8: Imagen mostrada al utilizar la función imshow.



Figura 9: Conversión de la imagen a escala de grises.

Cabe aclarar que en estos resultados la descripción RGB proporcionada de un color de la imagen hace referencia a la descomposición del color en términos de la intensidad de los colores con que se forma el rojo, el verde y el azul. Esta representación de la imagen es un modelo de color basado en la síntesis aditiva con el que es posible representar un color mediante la mezcla por adición de los tres colores de luz primarios. Mediante los valores RGB pueden mostrarse colores notablemente diferentes que aunque utilicen un mismo modelo de color, los espacios de colores varían considerablemente ver figura 9.

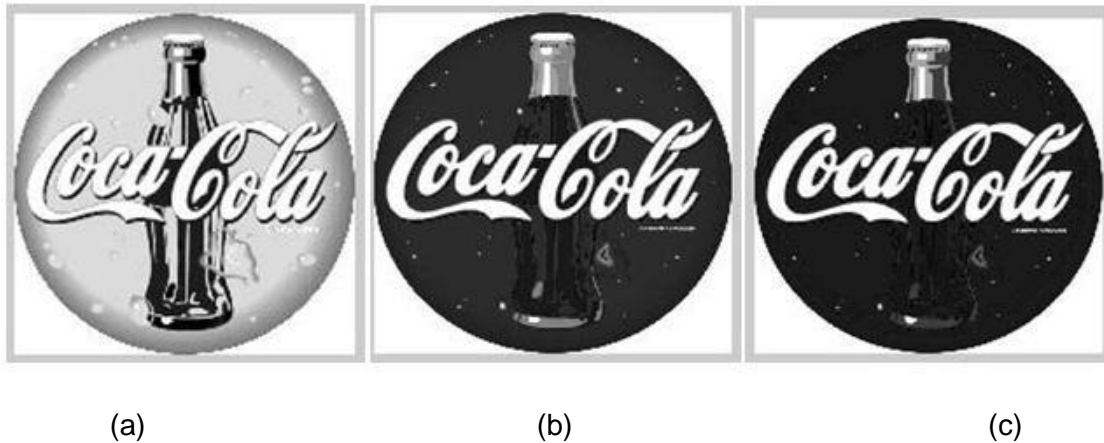


Figura 10: Planos de la imagen (a) rojo, (b) verde, (c) azul.

Luego de tener la imagen en escala de grises mediante la función `medfilt2 (imagen)` se reduce el ruido y se mantienen los bordes, esto se hace con el objetivo de aumentar la calidad de los pixeles de la misma para el posterior procesamiento, el resultado de este proceso se muestra en la figura 10. Se hace necesario obtener las dimensiones por filas y columnas convirtiéndola en una matriz mediante la función `[f c]=size(imagen)` donde `f` guardará la cantidad de filas y `c` la cantidad de columnas.



Figura 11: Imagen luego de aplicar filtrado.

Una vez obtenida la imagen matricial se procede a abrir un fichero con extensión `.txt` mediante la función `fopen('text.txt', 'wt')` con permiso de escritura, ya que el texto detectado en la imagen se plasmará en él. Luego se crea una matriz **word** vacía, para almacenar cada línea del texto que esté dentro de la imagen. Se separa línea por línea mediante la función `lines` mostrada en la figura 11, la cual llama a la función `clip`. Posteriormente se separa letra por letra mediante la función `edge(double(imgn),'canny')`, la cual etiqueta las componentes conexas de la imagen. Este proceso es denominado como fragmentación o

segmentación. La segmentación realizada es el resultado del proceso de binarización en la que se separaron los objetos o regiones de interés del resto de la misma, llegándose a formar un conjunto de estructuras o descriptores de la escena del objeto o región de interés mediante la detección de bordes y contornos de la misma.

```
function [fl re]=lines(aa)
aa=clip();
r=size(aa,1);
for s=1:r
    if sum(aa(s,:))==0
        nm=aa(1:s-1,1:end);|
        rm=aa(s:end,1:end);
        fl=~clip(~nm);
        re=~clip(~rm);
        break
    else
        fl=~aa;
        re=[];
    end
end
```

Figura 12: Función para separar línea a línea el texto de la imagen.

Los elementos que constituyen una imagen en MATLAB tienen el formato entero uint8, que es un tipo de dato que puede variar de 0 a 255, sin poder soportar decimales y valores que salgan fuera de ese rango. Lo anterior resulta una desventaja principalmente en aquellos casos donde se implementan algoritmos que trabajan con estos tipos de datos para realizar operaciones de división o multiplicación por tipo de dato flotante. En estos casos es necesario transformar la imagen de tipo de dato uint8 a double. Es importante tener en cuenta que si se utiliza la función imshow para desplegar las imágenes, esta no tiene la capacidad de poder desplegar imágenes del tipo double por lo que una vez realizado las operaciones de punto flotante es necesario después convertir al tipo de dato uint8.

Luego cada letra es normalizada a un tamaño de 42 x 24 píxeles, que es el tamaño de la plantilla con la que realizará la correlación. Para la normalización de cada fragmento encontrado en la imagen se usó la siguiente función `imresize(imagen_g,[42 24])`.

Luego se establece una comparación mediante la correlación en dos dimensiones mediante la función `corr2`. Esta operación da un valor de la semejanza entre dos matrices (la obtenida luego del proceso de fragmentación y las plantillas almacenadas, cada plantilla es una imagen binaria bmp de 42 x 24 píxeles. Como paso final se escribe el contenido de la imagen en el fichero por la ejecución de la función `fprintf(fid,'%s\n',word)`.

Una vez obtenido todo el texto de la imagen se procede a adicionar cada palabra en el campo de palabras extraídas, que contiene la tabla de Valores_Predeterminados de la base de datos, cada uno de los atributos de este campo serán comparados con los atributos del campo palabras extraídas de la tabla Valores_Dinámicos, para buscar similitud, mediante una consulta ver figura 12. Luego de este proceso se mostrará una vista con las palabras que coincidan ver figura 13, si se verifica que ocurre una aparición como mínimo de 5 palabras claves en al menos un minuto del análisis, la secuencia de video se considera candidata en la aplicación de esta técnica.

	id [PK] integer	texto_extraido character varying		id [PK] integer	palabras_claves character varying
1	1	cola	1	1	refresco
2	2	la	2	2	colgate
3	3	rexona	3	3	gatorade
4	4	colagate	4	4	rexona
*			*		

Figura 13: Representación de la base de datos.

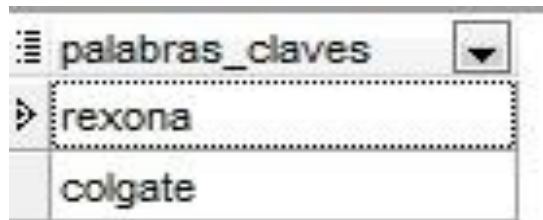


Figura 14: Detección de palabras publicitarias.

3.4 Realización de pruebas al algoritmo propuesto.

Para realizar las pruebas al algoritmo propuesto se toma una muestra de 2 Spots publicitarios, 2 segmentos de una novela y 3 segmentos de una película.

T0: Detección de spots por cambios de toma.

Material audiovisual	Tiempo de duración	Detección automática	Detección Visual	Falsos positivos	Falsos negativos
Spots5.avi	1m	46	49		3
	1m	44	46		2
	1m	33	39		6
	1m	23	20	3	
Total	4m	146	154	3	11
Promedio		36.5	38.5	0.75	3.7
% Detección		94.8%			

Tabla 1: Análisis de la técnica de detección por cambios de toma(Spots5).

Luego de analizar el primer spots se llega a la conclusión que para un promedio de 38 cambios de toma por minuto la técnica implementada detecta 36 cambios de toma por minuto por lo que son introducidos 2 falsos negativos en la detección.

Material	Tiempo de	Detección	Detección	Falsos	Falsos
----------	-----------	-----------	-----------	--------	--------

audiovisual	duración	automática	Visual	positivos	negativos
Spots4.avi	1m	33	30	3	
	1m	36	32	4	
	1m	33	34		1
	1m	21	17	4	
Total	4m	123	113	11	1
Promedio		30.75	28.25	2.75	0.25
% Detección		91.9%			

Tabla 2: Análisis de la técnica de detección por cambios de toma(Spots4).

Luego de analizar el segundo spots se llega a la conclusión que para un promedio de 28 cambios de toma por minuto la técnica implementada detecta 30 cambios de toma por minuto, por lo que son introducidos 2 falsos positivos en la detección. El promedio de detección para los dos Spots analizados representa un porcentaje de 93.35, nótese que la muestra es pequeña y cualquier fallo representa un porcentaje elevado.

Material audiovisual	Tiempo de duración	Detección automática	Detección Visual	Falsos positivos	Falsos negativos
Nov1.avi	1m	11	11	0	0
Nov2.avi	1m	12	12	0	0
Total	2m	23	23	0	0
Promedio		19.3	19.3	0	0
% Detección		100%			

Tabla 3: Análisis de la técnica de detección por cambios de toma (Nov1, Nov2).

Material audiovisual	Tiempo de duración	Detección automática	Detección Visual	Falsos positivos	Falsos negativos
Peli1.avi	1m	24	24	0	0
Peli2.avi	1m	21	21	0	0
Peli3.avi	1m	13	13	0	0
Total		58	58	0	0
Promedio		19.3	19.3	0	0
% Detección		100%			

Tabla 4: Análisis de la técnica de detección por cambios de toma (Peli1, Peli2, Peli3).

Luego del análisis representado en las dos tablas anteriores se llega a la conclusión que son detectados todos los cambios ocurridos para un porcentaje de detección de un 100%, sin ser introducidos falsos positivos o falsos negativos.

Al concluir las pruebas a la T0 son detectados como posible publicidad los materiales: Spots4 y Spots5 debido a que en todos los minutos analizados la cantidad de cambios de toma superaban los 15 y de igual modo como mínimo en un minuto sobrepasaron los 30 cambios. Los materiales: Nov1, Nov2, Peli1, Peli2, Peli3 no representan publicidad debido a que no en todos los minutos del análisis superan los 15 cambios y en ningún punto supera los 30 cambios. El porcentaje de detección de la técnica es de un 96.68%.

T1: Detección de spots por aumento de volumen.

Las pruebas realizadas mediante esta técnica se le aplican a los materiales: Spots4 y Spots5 ya que representan la salida de la T0.

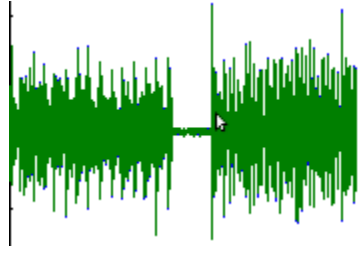
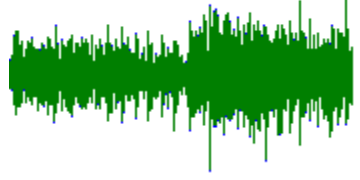
Material audiovisual	Tiempo de duración	Detección Visual	Falsos positivos	Falsos negativos
Spots4.avi	14s		0	0
Spots5.avi	12s		0	0
Total	26s		0	0
% Detección		100%		

Tabla 5: Análisis de la técnica de detección por aumento de volumen (Spots4, Spots5).

Mediante la representación de la frecuencia del audio se percibe que en el momento que el spots publicitario sale al aire la amplitud es mayor (ver anexo 1), además de acercarse a cero en el instante antes del comercial, este efecto se nota de forma más clara cuando el fotograma de color negro que marca el comienzo del comercial es mayor, debido a que este está acompañado de silencio.

Luego de analizar las dos producciones audiovisuales se toma la decisión de estar en presencia de 2 spots publicitarios debido a que arrojan resultados positivos en la detección por cambios de toma y en la detección por aumento de volumen, por lo que no es necesario analizar la T2.

Se está en presencia de un algoritmo altamente fiable debido a que tiene un porcentaje de detección en su conjunto para la muestra seleccionada de un 100%, superando los existentes actualmente en el mundo utilizados con el mismo fin (ver figura 15). El resultado final del algoritmo depende en gran medida de la primera técnica fusionada por el mismo, mientras mayor sea la precisión de esta mayor será la exactitud del resultado final.

Porciento de detección de los algoritmos para la detección de publicidad.

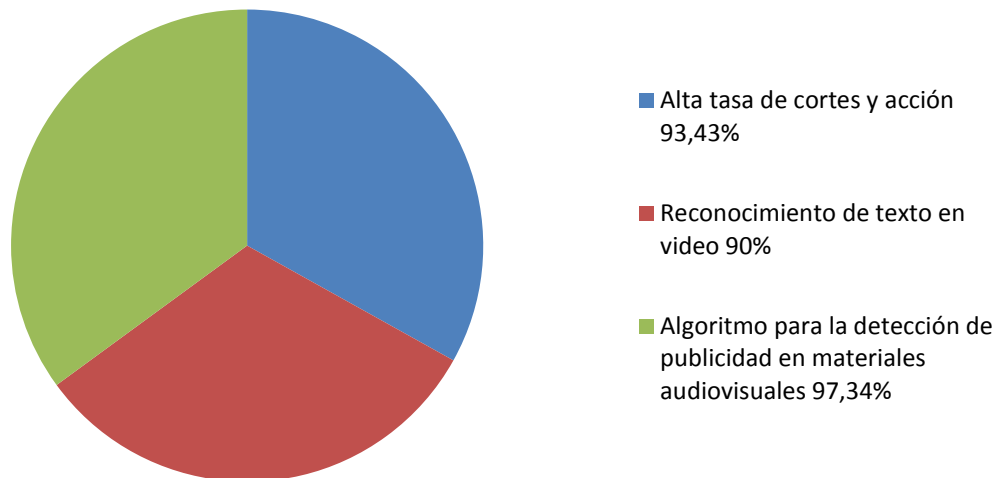


Figura 15: Resultados de las técnicas para la detección de Spots publicitarios.

3.5 Conclusiones parciales.

Luego de la culminación del capítulo Análisis de los resultados, se llega a la conclusión que la tecnología utilizada para la realización de las pruebas es en gran medida útil, pero en ocasiones presenta limitaciones debido al hardware con que se dispone. El algoritmo propuesto tiene un porciento de detección elevado superando a los existentes en la actualidad, esto se debe en gran medida a la integración de las técnicas que lo conforman. En materiales que no representan publicidad la detección es muy precisa debido a que los cambios de toma son abruptos. El número de falsos positivos y falsos negativos introducidos en materiales que representan publicidad no es significativo, por lo que no influye en el resultado final del algoritmo.

Conclusiones Generales

Durante el proceso de desarrollo de la investigación se analizaron distintas técnicas para la detección de publicidad en materiales audiovisuales, llegando a la conclusión que algunas de las estudiadas no brindan un porcentaje alto en la detección y varias de ellas tienen ciertas dependencias de las televisoras. Son seleccionadas dos de las técnicas de mejores resultados e implementada una nueva luego de un análisis detallado del comportamiento de la métrica que mide. Se propone un nuevo algoritmo el cual combina las tres técnicas antes analizadas, puesto que luego de realizar pruebas a la solución propuesta arroja mejores resultados que cada técnica por separada. Se deduce que mientras mayor sea la fusión alcanzada por el algoritmo mayor será la precisión con respecto al resultado. Al concluir la investigación quedan sentadas las bases para el desarrollo de las restantes fases del proceso, como son: el análisis, el diseño y la implementación.

Recomendaciones

A partir de la investigación realizada se recomienda realizar las restantes fases, el análisis, el diseño y la implementación. Desarrollar la implementación de las tres técnicas combinadas por el algoritmo sobre una herramienta libre. Luego de la implementación lograr la integración con las aplicaciones del proyecto de Captura y catalogación de medias del departamento de Señales digitales.

Referencias Bibliográficas.

Albiol. 2004. **Detection of TV Commercials. 2004.**

Bueno, José Ramón Cerquides. 2009. **Sistema de reconocimiento y monitorización de anuncios en TV. 2009.**

Cade, Robert. 2007. <http://www.mopo.ca/2007/11/gatorade-inventor-dies-at-80.html>.

<http://www.mopo.ca/2007/11/gatorade-inventor-dies-at-80.html>. [En línea] 28 de Noviembre de 2007. [Citado el: 14 de Marzo de 2011.]

Chust, José Vicente Mestre. 2010. [http://www.suite101.net/content/el-concepto-de-](http://www.suite101.net/content/el-concepto-de-publicidad-a16347)

[publicidad-a16347](http://www.suite101.net/content/el-concepto-de-publicidad-a16347). <http://www.suite101.net/content/el-concepto-de-publicidad-a16347>. [En línea] 6 de Mayo de 2010. [Citado el: 14 de Marzo de 2011.]

Cruz, Jesús M. de la. 2001. *Visión por Computador. Imágenes digitales y aplicaciones. 2001.*

CVISION Technologis. 1998. [http://www.cvisiontech.com/reference/ocr/ocr-](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp)

[capture.html?lang=esp](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp). [http://www.cvisiontech.com/reference/ocr/ocr-](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp)

[capture.html?lang=esp](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp). [En línea] 1998. <http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp>.

DefiniciónABC. 2010. **Definición ABC. www.definicionabc.com/general/fotograma.php.** [En

línea] 5 de enero de 2010. [Citado el: 14 de Marzo de 2011.]

Disey Publicidad. 2011. <http://www.disey.com.mx/index.php>. 2011.

Elliman, D.G., Lancaster, I.T. 1982. **A Review of segmentation and contextual analysis techniques for text recognition. Munich.Alemania. : s.n., 1982.**

Furfaro, Alejandro. 2010. *Manejo de bibliotecas Opencv. 2010.*

González, Carlos D. 2008. <http://www.usabilidadweb.com.ar/cpp.php>.

<http://www.usabilidadweb.com.ar/cpp.php>. [En línea] 2008. [Citado el: 17 de Marzo de 2011.] <http://www.usabilidadweb.com.ar/cpp.php>.

Henneberger, Bryson. 2008. **Television Commercial Detection Techniques. 2008.**

- Jalón, Javier García de. 2005. ***Aprenda Matlab 7.0 como si estuviera en primero***. Madrid : s.n., 2005.
- Jiménez, Leandro Rodríguez. 2008.. ***Procesamiento de imágenes biológicas***. Bellaterra. : s.n., 2008.
- Latecki, L.J. 1998. **Discrete Representation of Spatial Objects in Computer Vision**. 1998.
- Leonardi, C. Saraceno and R. 1997. **Audio as a Support to Scene Change Detection and**. 1997.
- Liu, Z., Wang, Y. & Chen, T. 1998. **AUDIO FEATURE EXTRACTION AND ANALYSIS FOR**. 1998.
- Marques, B. Satterwhite and O. 2004. ***Automatic Detection of TV Commercials***. 2004.
- Martínez., Carlos Alberto Saez. 2005. **Los spots, una posibilidad para desarrollar la educación ambiental a través de la televisión**. 2005.
- Mateos, Ginés García. **Universiadd de Murcia**. [En línea]
- Nagy, Thomas A. Nartker, Stephen V. Rice. 1990. **Optical Character Recognition**. 1990.
- Otsu, Nobuyuki. 1979. **A threshold selection method from gray-level histogram**. 1979.
- Quijada, Prado Franquet Ribes y Soto Fernández. 2007. **La publicidad televisiva ante el reto de la interactividad**. 2007.
- Rebaza, Jorge Valverde. 2010. **Escuela Académico Profesional de Informática.Escuela Académico Profesional de Informática.Universidad Nacional de Trujillo**. Trujillo : s.n., 2010.
- S. Pfeiffer, S. Fischer and W. Effelsberg. 1998. **Automatic Audio Content Analysis**. 1998.
- Sadlier, D. A., Marlow, S., O'Connor, N. & Murphy, N. 2002. **Automatic TV Advertisement Detection From MPEG Bitstream**. 2002.
- Sadlier, D.A. 2001. **Automatic TV Advertisement Detection from MPEG Bitstream**. 2001.
- Serra, Jesús Angulo y Jean. 2005. **Segmentación de Imágenes en Color utilizando Histogramas Bi-Variables** . Paris : s.n., 2005.
- Tewfik, J. Nam and A. H. 1997. **Combined Audio and Visual Streams Analysis for Video**. 1997.

Tie-Yan Liu, Tao Qin and Hong-Jiang Zhang,. 2004. **Time-constraint boost for TV commercials detection. 2004.**

Y. Wang, J. Huang, Z. Liu, and T. Chen. 1997. **Multimedia Content Classification using Motion. 1997.**

Bibliografía consultada.

Albiol. 2004. **Detection of TV Commercials. 2004.**

Bueno, José Ramón Cerquides. 2009. **Sistema de reconocimiento y monitorización de anuncios en TV. 2009.**

Cade, Robert. 2007. <http://www.mopo.ca/2007/11/gatorade-inventor-dies-at-80.html>.

<http://www.mopo.ca/2007/11/gatorade-inventor-dies-at-80.html>. [En línea] 28 de Noviembre de 2007. [Citado el: 14 de Marzo de 2011.]

Corporation, Nokia. 2010. <http://developer.qt.nokia.com/wiki/QtCreatorWhitepaper.2011>.

Chust, José Vicente Mestre. 2010. <http://www.suite101.net/content/el-concepto-de-publicidad-a16347>.

<http://www.suite101.net/content/el-concepto-de-publicidad-a16347>. [En línea] 6 de Mayo de 2010. [Citado el: 14 de Marzo de 2011.]

Cruz, Jesús M. de la. 2001. *Visión por Computador. Imágenes digitales y aplicaciones. 2001.*

CVISION Technologies. 1998. [http://www.cvisiontech.com/reference/ocr/ocr-](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp)

[capture.html?lang=esp](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp). [http://www.cvisiontech.com/reference/ocr/ocr-](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp)

[capture.html?lang=esp](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp). [En línea] 1998. [http://www.cvisiontech.com/reference/ocr/ocr-](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp)

[capture.html?lang=esp](http://www.cvisiontech.com/reference/ocr/ocr-capture.html?lang=esp).

DefiniciónABC. 2010. **Definición ABC. www.definicionabc.com/general/fotograma.php.** [En

línea] 5 de enero de 2010. [Citado el: 14 de Marzo de 2011.]

Disey Publicidad. 2011. <http://www.disey.com.mx/index.php>. 2011.

Elliman, D.G., Lancaster, I.T. 1982. **A Review of segmentation and contextual analysis techniques for text recognition. Munich.Alemania. : s.n., 1982.**

Furfaro, Alejandro. 2010. **Manejo de bibliotecas Opencv. 2010.**

- González, Carlos D. 2008. <http://www.usabilidadweb.com.ar/cpp.php>.
<http://www.usabilidadweb.com.ar/cpp.php>. [En línea] 2008. [Citado el: 17 de Marzo de 2011.] <http://www.usabilidadweb.com.ar/cpp.php>.
- Henneberger, Bryson. 2008. **Television Commercial Detection Techniques. 2008.**
- Jalón, Javier García de. 2005. ***Aprenda Matlab 7.0 como si estuviera en primero*. Madrid : s.n., 2005.**
- Jiménez, Leandro Rodríguez. 2008.. ***Procesamiento de imágenes biológicas*. Bellaterra. : s.n., 2008.**
- Latecki, L.J. 1998. **Discrete Representation of Spatial Objects in Computer Vision. 1998.**
- Leonardi, C. Saraceno and R. 1997. **Audio as a Support to Scene Change Detection and. 1997.**
- Lazalde, Cristina. 2010. **EcuRed.**http://www.ecured.cu/index.php/Herramienta_CASE.2010
- Liu, Z., Wang, Y. & Chen, T. 1998. **AUDIO FEATURE EXTRACTION AND ANALYSIS FOR. 1998.**
- Marques, B. Satterwhite and O. 2004. ***Automatic Detection of TV Commercials*. 2004.**
- Martínez., Carlos Alberto Saez. 2005. **Los spots, una posibilidad para desarrollar la educación ambiental a través de la televisión. 2005.**
- Mateos, Ginés García. **Universiadd de Murcia. [En línea]**
- Nagy, Thomas A. Nartker, Stephen V. Rice. 1990. **Optical Character Recognition. 1990.**
- Otsu, Nobuyuki. 1979. **A threshold selection method from gray-level histogram. 1979.**
- Quijada, Prado Franquet Ribes y Soto Fernández. 2007. **La publicidad televisiva ante el reto de la interactividad. 2007.**
- Rebaza, Jorge Valverde. 2010. **Escuela Académico Profesional de Informática.Escuela Académico Profesional de Informática.Universidad Nacional de Trujillo. Trujillo : s.n., 2010.**
- S. Pfeiffer, S. Fischer and W. Effelsberg. 1998. **Automatic Audio Content Analysis. 1998.**
- Sadlier, D. A., Marlow, S., O'Connor, N. & Murphy, N. 2002. **Automatic TV Advertisement Detection From MPEG Bitstream. 2002.**

- Sadlier, D.A. 2001. **Automatic TV Advertisement Detection from MPEG Bitstream. 2001.**
- Serra, Jesús Angulo y Jean. 2005. **Segmentación de Imágenes en Color utilizando Histogramas Bi-Variables . Paris : s.n., 2005.**
- Tewfik, J. Nam and A. H. 1997. **Combined Audio and Visual Streams Analysis for Video. 1997.**
- Tie-Yan Liu, Tao Qin and Hong-Jiang Zhang,. 2004. **Time-constraint boost for TV commercials detection. 2004.**
- Y. Wang, J. Huang, Z. Liu, and T. Chen. 1997. **Multimedia Content Classification using Motion. 1997.**
- Pressman, Roger S. 2005. **Ingeniería del Software.Un enfoque práctico.2005.**

Anexos

Anexo 1: Proceso de ejecución de la T0 para una película. 53

Anexo 2: Resultado de la ejecución de la T0 para un minuto en una película..... 54

Anexo 3: Proceso de ejecución de la T0 para un Spots (4a.avi). 55

Anexo 4: Resultado de la ejecución de la T0 para un minuto en un Spots (4a.avi). 56

Anexo 5: Proceso de ejecución de la T0 para un Spots (5a.avi). 57

Anexo 6: Resultado de la ejecución de la T0 para un minuto en un Spots (5a.avi). 58

Anexo 7: Código en MATLAB para graficar la señal de audio de un video. 59

Anexo 8: Representación en MATLAB de audio publicitario y audio ausente de publicidad luego de la ejecución del código mostrado en el anexo 7. 60

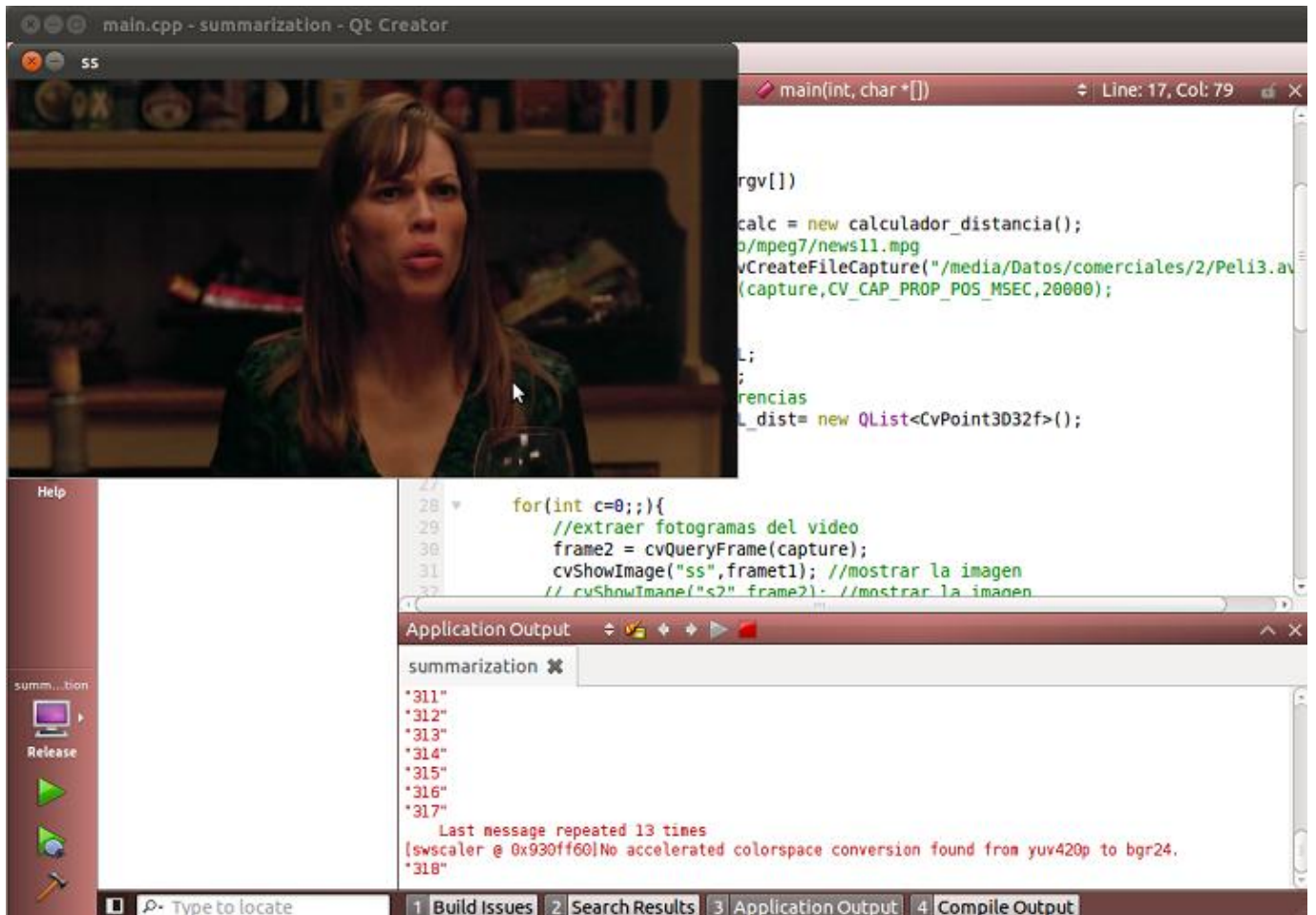
Anexo 1: Proceso de ejecución de la T0 para una película.

Figura 16: Detección de fotogramas claves en un minuto del análisis de una película.

En la figura anterior se muestra el análisis en tiempo de ejecución de la técnica de detección de publicidad por cambios de toma (T0). En el momento que se toma la foto está siendo analizado el fotograma 318.

Anexo 2: Resultado de la ejecución de la T0 para un minuto en una película.

The screenshot displays the Qt Creator IDE with a C++ project named 'summarization'. The main.cpp file contains the following code:

```

10  lude "QVariant"
11  lude <QTime>
12
13  ▼ nain(int argc, char *argv[])
14
15  :calculador_distancia *calc = new calculador_distancia();
16  //home/luis/Escritorio/mpeg7/news11.mpg
17  cvCapture* capture = cvCreateFileCapture("/media/Datos/comerciales/2/Peli3.avi");
18  / cvSetCaptureProperty(capture,CV_CAP_PROP_POS_MSEC,20000);
19  bool is=1;
20  int Ne=2;
21  IplImage* frame2 = NULL;
22  IplImage * framet1 = 0;
23  // crear lista de diferencias
24  QList<CvPoint3D32f> * L_dist= new QList<CvPoint3D32f>();
25  cvPoint3D32f punto;
26
27
28  ▼ for(int c=0;;){
29      //extraer fotogramas del video
30      frame2 = cvQueryFrame(capture);
31      cvShowImage("ss",framet1); //mostrar la imagen
32      // cvShowImage("s2" frame2); //mostrar la imagen

```

The Application Output window shows the following output:

```

summarization ✖
*1456"
[swscaler @ 0x9314c50]No accelerated colorspace conversion found from yuv420p to bgr24.
*1457"
[swscaler @ 0x9313bb0]No accelerated colorspace conversion found from yuv420p to bgr24.
*1458"
[swscaler @ 0x93137c0]No accelerated colorspace conversion found from yuv420p to bgr24.
*1459"
[swscaler @ 0x9314290]No accelerated colorspace conversion found from yuv420p to bgr24.
*1460"
*13"

```

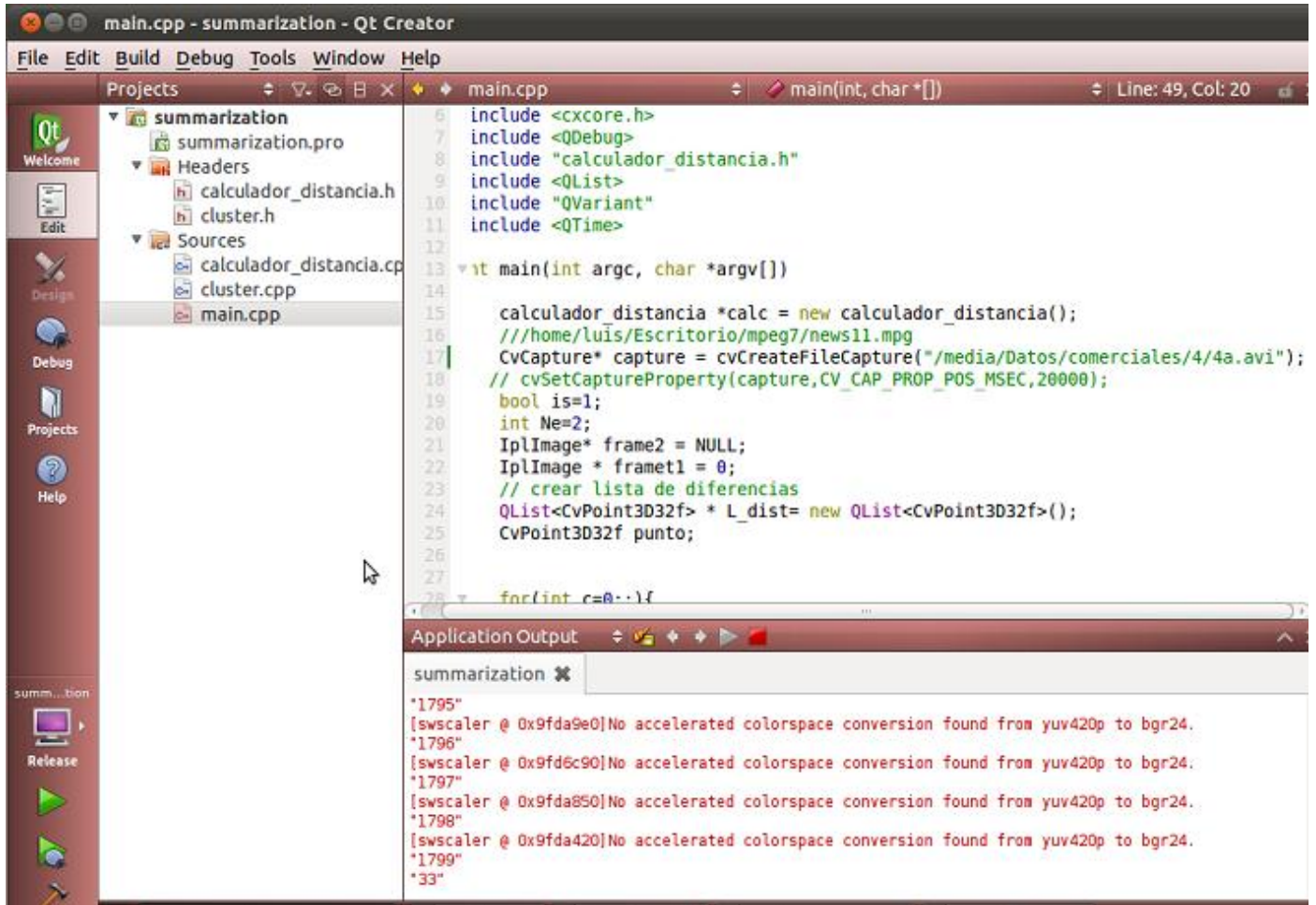
Figura 17: Cantidad de fotogramas claves en un minuto del análisis de una película.

En la figura anterior la técnica detecta 13 cambios de toma para un total de 1460 fotogramas analizados.

Anexo 3: Proceso de ejecución de la T0 para un Spots (4a.avi).

Figura 18: Detección de fotogramas claves en un minuto del análisis de un Spots publicitario (4a.avi).

En la figura anterior se muestra el análisis en tiempo de ejecución de la técnica de detección de publicidad por cambios de toma (T0). En el momento que se toma la foto está siendo analizado el fotograma 1557.

Anexo 4: Resultado de la ejecución de la T0 para un minuto en un Spots (4a.avi).

The screenshot shows the Qt Creator IDE with a C++ project named 'summarization'. The main.cpp file is open, showing code that includes headers like <xcvcore.h>, <QDebug>, and <QList>. The main function initializes a 'calculador distancia' object, sets up video capture for '4/4a.avi', and processes frames. The application output window at the bottom shows the results of the analysis, listing frame numbers from 1795 to 1799 and a total count of 33 keyframes.

```
6 include <xcvcore.h>
7 include <QDebug>
8 include "calculador_distancia.h"
9 include <QList>
10 include "QVariant"
11 include <QTime>
12
13 int main(int argc, char *argv[])
14 {
15     calculador distancia *calc = new calculador_distancia();
16     //home/luis/Escritorio/mpeg7/news11.mpg
17     CvCapture* capture = cvCreateFileCapture("/media/Datos/comerciales/4/4a.avi");
18     // cvSetCaptureProperty(capture,CV_CAP_PROP_POS_MSEC,20000);
19     bool is=1;
20     int Ne=2;
21     IplImage* frame2 = NULL;
22     IplImage * framet1 = 0;
23     // crear lista de diferencias
24     QList<CvPoint3D32f> * L_dist= new QList<CvPoint3D32f>();
25     CvPoint3D32f punto;
26
27
28     for(int c=0;c<1f
```

Application Output

```
summarization ✖
*1795"
[swscaler @ 0x9fda9e0]No accelerated colorspace conversion found from yuv420p to bgr24.
*1796"
[swscaler @ 0x9fd6c90]No accelerated colorspace conversion found from yuv420p to bgr24.
*1797"
[swscaler @ 0x9fda850]No accelerated colorspace conversion found from yuv420p to bgr24.
*1798"
[swscaler @ 0x9fda420]No accelerated colorspace conversion found from yuv420p to bgr24.
*1799"
*33"
```

Figura 19: Cantidad de fotogramas claves en un minuto del análisis de un Spots (4a.avi).

En la figura anterior la técnica detecta 33 cambios de toma para un total de 1799 fotogramas analizados.

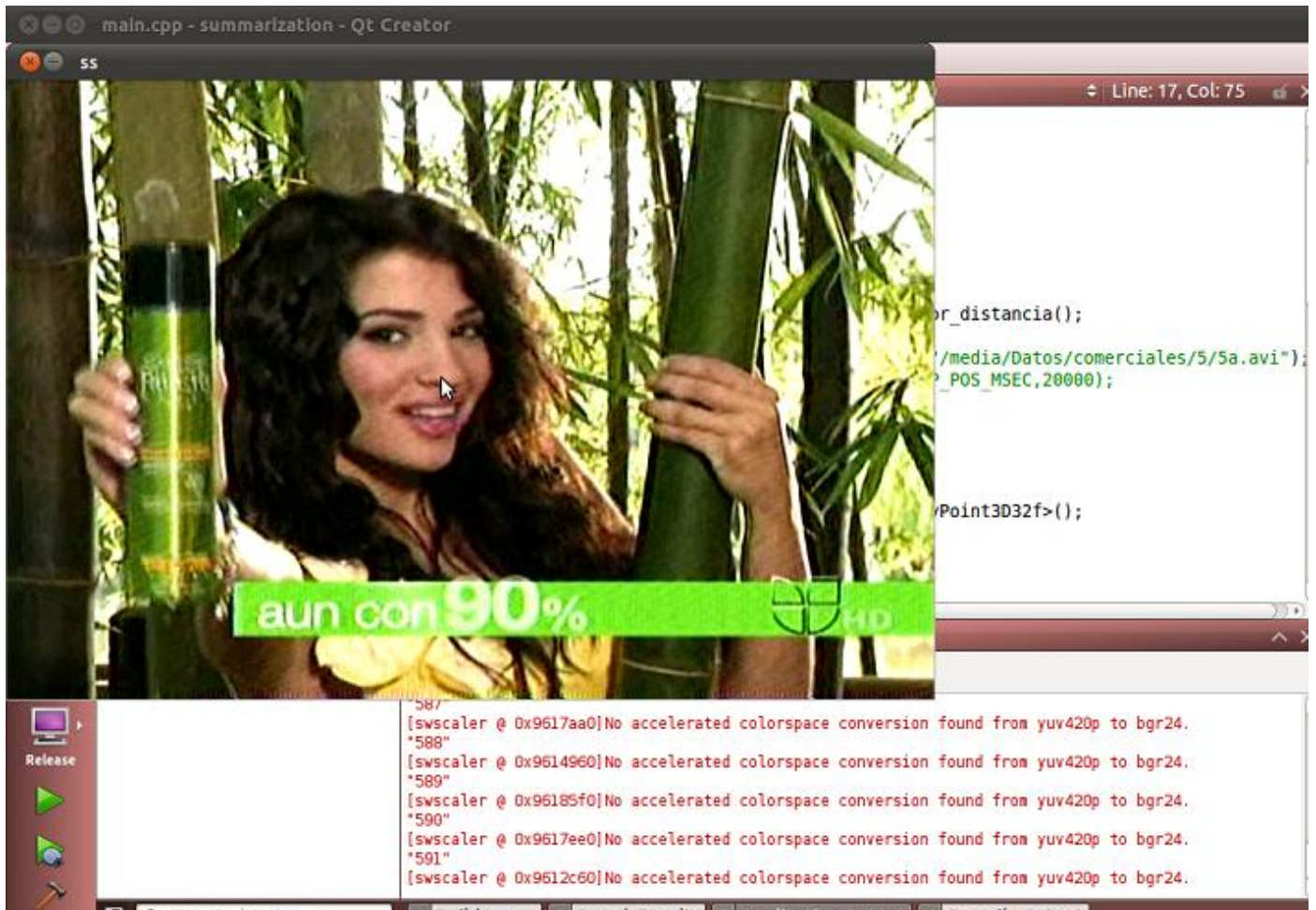
Anexo 5: Proceso de ejecución de la T0 para un Spots (5a.avi).

Figura 20: Detección de fotogramas claves en un minuto del análisis de un Spots publicitario (5a.avi).

En la figura anterior se muestra el análisis en tiempo de ejecución de la técnica de detección de publicidad por cambios de toma (T0). En el momento que se toma la foto está siendo analizado el fotograma 591.

Anexo 6: Resultado de la ejecución de la T0 para un minuto en un Spots (5a.avi).

The screenshot shows the Qt Creator IDE with a C++ project named 'summarization'. The main.cpp file is open, showing code that includes headers like <cxcore.h>, <QDebug>, and <QList>. The main function initializes a 'calculador_distancia' object, creates a video capture for '5a.avi', and sets properties for frame rate and position. It then enters a loop to process frames.

```

6 #include <cxcore.h>
7 #include <QDebug>
8 #include "calculador_distancia.h"
9 #include <QList>
10 #include <QVariant>
11 #include <QTime>
12
13 int main(int argc, char *argv[])
14 {
15     calculador_distancia *calc = new calculador_distancia();
16     //home/luis/Escritorio/mpeg7/news11.mpg
17     CvCapture* capture = cvCreateFileCapture("/media/Datos/comerciales/5/5a.avi")
18     // cvSetCaptureProperty(capture,CV_CAP_PROP_POS_MSEC,20000);
19     bool is=1;
20     int Ne=2;
21     IplImage* frame2 = NULL;
22     IplImage * framet1 = 0;
23     // crear lista de diferencias
24     QList<CvPoint3D32f> * L_dist= new QList<CvPoint3D32f>();
25     CvPoint3D32f punto;
26
27
28     for(int c=0..1{

```

The Application Output window shows the following output:

```

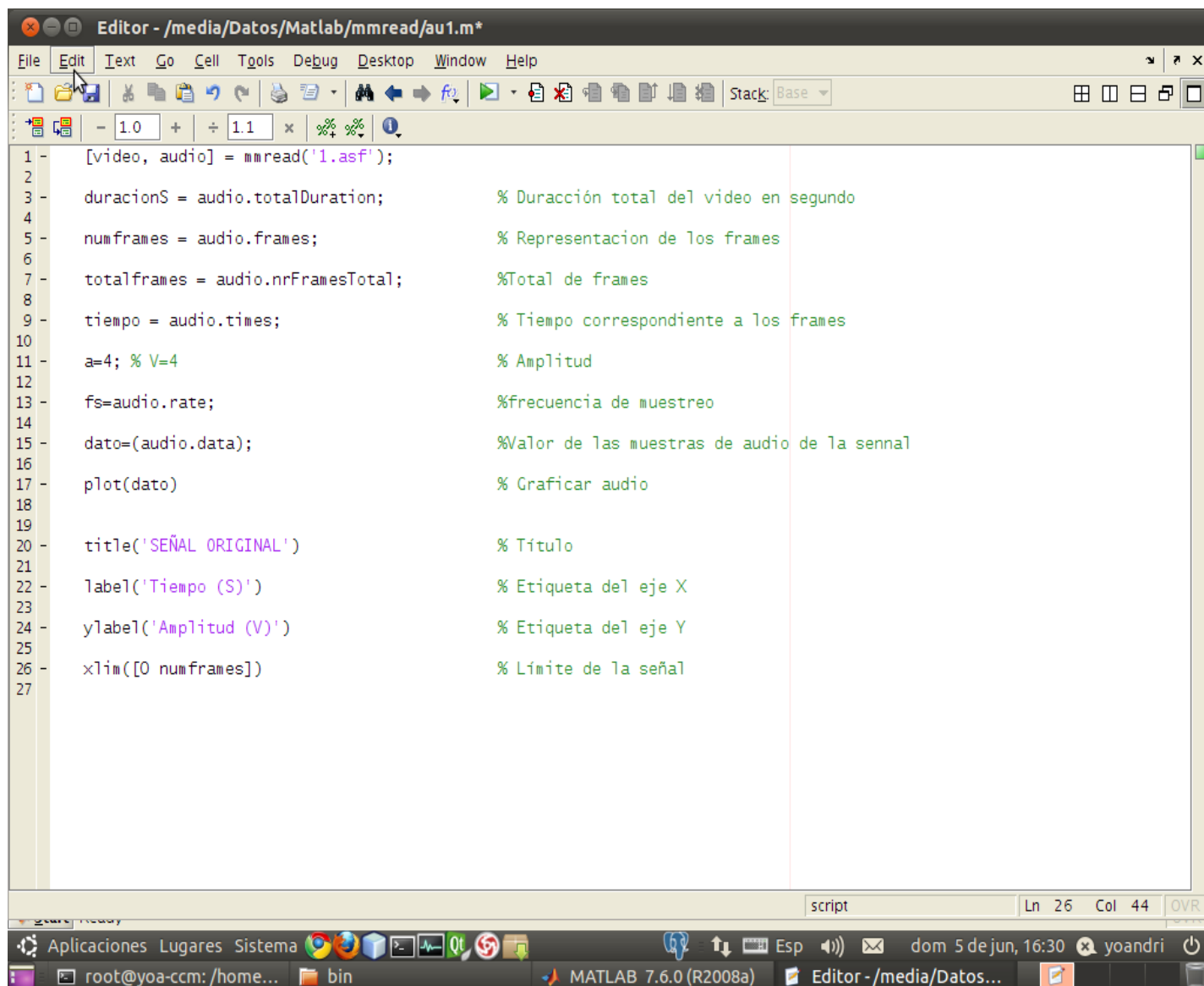
summarization ✖
*1795"
[swscaler @ 0x961f210]No accelerated colorspace conversion found from yuv420p to bgr24.
*1796"
[swscaler @ 0x961f6c0]No accelerated colorspace conversion found from yuv420p to bgr24.
*1797"
[swscaler @ 0x961ae30]No accelerated colorspace conversion found from yuv420p to bgr24.
*1798"
[swscaler @ 0x961fca0]No accelerated colorspace conversion found from yuv420p to bgr24.
*1799"
*46"

```

Figura 21: Cantidad de fotogramas claves en un minuto del análisis de un Spots (5a.avi).

En la figura anterior la técnica detecta 46 cambios de toma para un total de 1799 fotogramas analizados.

Anexo 7: Código en MATLAB para graficar la señal de audio de un video.

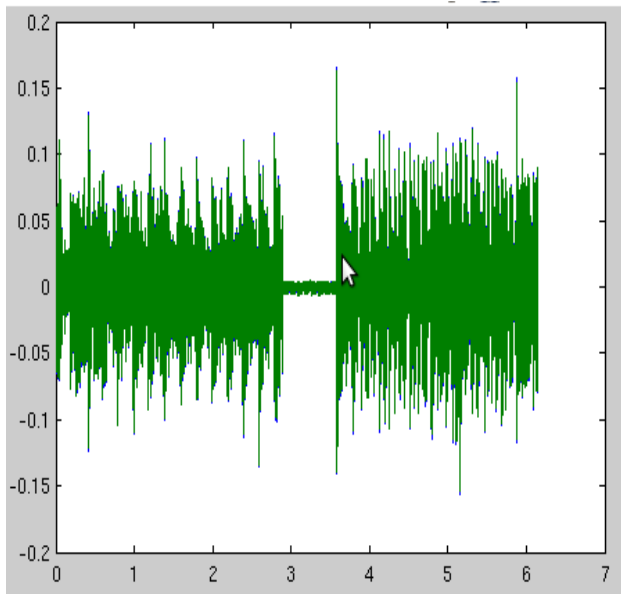
A screenshot of the MATLAB Editor interface. The window title is 'Editor - /media/Datos/Matlab/mmread/au1.m*'. The menu bar includes File, Edit, Text, Go, Cell, Tools, Debug, Desktop, Window, and Help. The toolbar contains various icons for file operations, editing, and execution. The code editor shows the following MATLAB code:

```
1 - [video, audio] = mmread('1.asf');  
2  
3 - duracionS = audio.totalDuration;           % Duración total del video en segundo  
4  
5 - numframes = audio.frames;                 % Representacion de los frames  
6  
7 - totalframes = audio.nrFramesTotal;       %Total de frames  
8  
9 - tiempo = audio.times;                    % Tiempo correspondiente a los frames  
10  
11 - a=4; % V=4                               % Amplitud  
12  
13 - fs=audio.rate;                           %frecuencia de muestreo  
14  
15 - dato=(audio.data);                       %Valor de las muestras de audio de la sennal  
16  
17 - plot(dato)                                % Graficar audio  
18  
19  
20 - title('SEÑAL ORIGINAL')                  % Titulo  
21  
22 - label('Tiempo (S)')                      % Etiqueta del eje X  
23  
24 - ylabel('Amplitud (V)')                  % Etiqueta del eje Y  
25  
26 - xlim([0 numframes])                     % Limite de la señal  
27
```

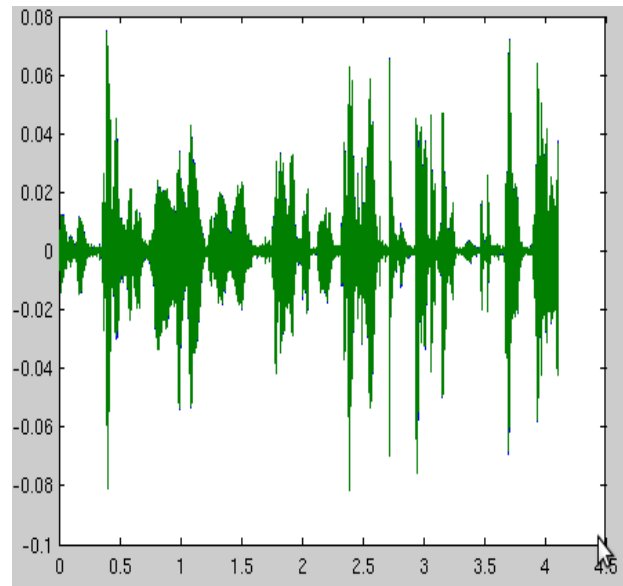
The status bar at the bottom of the editor shows 'script', 'Ln 26 Col 44', and 'OVR'. The system tray at the bottom of the screen shows the taskbar with icons for applications, system status, and network, along with the date and time 'dom 5 de jun, 16:30' and the user name 'yoandri'. The terminal window below the editor shows the prompt 'root@yoa-ccm: /home...' and the directory 'bin'.

Figura 22: Código para graficar señales de audio en MATLAB.

Anexo 8: Representación en MATLAB de audio publicitario y audio ausente de publicidad luego de la ejecución del código mostrado en el anexo 7.



Gráfica 1: Audio con presencia de publicidad.



Gráfica 2: Audio ausente de publicidad.

Glosario de términos.

Secuencia candidata: Secuencia de video que cumple con cierta o ciertas características que puede representar en un futuro publicidad.

Umbral: De acuerdo al contexto en el cual se le emplee, el término **umbral** podrá referir diversas cuestiones, en el transcurso de la investigación se emplea para la toma de una decisión con respecto al mismo, por ejemplo el umbral del audio para determinar si hay aparición de silencio o no, todos los valores que se encuentren por encima del mismo arrojarán un resultado y los que estén por debajo arrojarán otro.

Filtrado: Es el proceso para dar a una imagen un efecto especial para la eliminación de ruido (**Rebaza, 2010**).

Falso positivo: Detección de un segmento de anuncio que no lo es.

Falso negativo: No detección de un segmento que representa publicidad.

Componente conexa: Región que se encuentra dentro de un fotograma donde todos los pixeles son adyacentes entre sí, formando un objeto dentro de la imagen (**Elliman, 1982**).

Topología: La Topología es un área de las Matemáticas que privilegia fundamentalmente la noción de vecindad, adyacencia o relación entre objetos (**Latecki, 1998**).

Gradiente: Es el valor dado por la dirección de más rápido cambio de intensidad de la imagen. Es un vector que se encuentra normal a una superficie o curva en el espacio a la cual se le está estudiando, en un punto cualquiera. Mide la tasa y la dirección del cambio en un campo escalar; el gradiente de un campo escalar es un campo vectorial. (**Elliman, 1982**).

Filtro gaussiano: Es la técnica diseñada para mejorar la imagen donde se presentan bordes (**Rebaza, 2010**).

Primera derivada: La "primera derivada" de una función mide su crecimiento, no su valor en un momento dado (**Cruz, 2001**).