

**Universidad de las Ciencias Informáticas**

**“Facultad 6”**



**Procedimiento para el reconocimiento de rostro en  
sistemas de catalogación de audiovisuales.**

Trabajo de Diploma para optar por el título de  
Ingeniero en Ciencias Informáticas

Autor: Nilo Tomás Díaz Alés

Tutor: Ing. Zorilin Alonso Guerrero

Cotutor: Ing. Jean Michael Suárez Pérez

Junio 2012

## DECLARACIÓN DE AUTORÍA

Declaramos ser autores de la presente tesis y reconocemos a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los \_\_\_\_ días del mes de \_\_\_\_\_ del año \_\_\_\_\_

Nilo Tomás Díaz Alés

Ing. Zorilin Alonso Guerrero

\_\_\_\_\_

Firma del Autor

\_\_\_\_\_

Firma del Tutor

## AGRADECIMIENTOS

*Con esta tesis termina una etapa en mi vida de esfuerzo y estudio, por eso quiero mostrar mi agradecimiento a todos los que han compartido conmigo, a todos esos que me han ayudado a superar las dificultades, que me han dado su apoyo y sobre todo que han tenido tanta paciencia. Pero en especial quiero agradecer a un grupo que siempre ha estado a mi lado:*

*A mis padres por ser mi razón de ser, por haber creído en mi en momentos en que yo mismo no lo hacía, por todo su amor, y por ser los mejores papas del mundo, gracias.*

*A mi tata por ser mi alegría y mi inspiración, te quiero.*

*A Eleidita y Tony por ser tan hospitalarios y buenos conmigo, su apoyo en estos años ha sido incondicional, gracias.*

*A mis amigos Alain Peña, Yidian, Frank, Liuver, Yoan, Betancourt, Humbe que han estado presente para apoyarme en todo momento, y hacerme tan feliz.*

*A la FEU por darme tantas herramientas para ser mejor persona, capaz de mirar más allá de mi mismo.*

*A mi tutora que está conmigo en una segunda tesis, fiel a la causa, una chica increíble, que siempre está ahí cuando se necesita.*

*A toda mi familia en general, abuelos, primos, tíos les agradezco por brindarme su apoyo y amor incondicional.*

*A los amigos que compartieron años de su vida conmigo en el trabajo, el estudio, y también en las fiestas.*

*A los muchachos del proyecto, Jean Michael, Yoendry, Adnan, Yoandry, Luis, que siempre han pensado en grande. A todos, gracias.*

## DEDICATORIA

*Esta tesis quiero dedicársela a tres personas muy especiales en mi vida:*

*A mis padres y mi hermana por estar presentes apoyándome en todo, por ser mi ejemplo a seguir en todos los aspectos, sin ellos, nada en mi vida hubiese sido posible, hoy les debo todo lo que soy, los quiero mucho.*

*Quiero dedicársela además a la Revolución y a Fidel por haberme dado la posibilidad de cumplir mi sueño, estudiar en la mejor Universidad de Cuba, la UCI que tanto me ha enseñado.*



## RESUMEN

El procesamiento de imágenes brinda en la actualidad innumerables ventajas para el almacenamiento, localización y tratamiento de materiales audiovisuales. Particularmente el reconocimiento de rostros puede aplicarse a diferentes áreas como son: la autenticación en sistemas, control de inmigración, sistemas de vigilancia, entre otros. En el presente trabajo se define un procedimiento para la utilización del reconocimiento de rostro en la catalogación de audiovisuales, se precisan las etapas que requiere su implementación: detección del rostro, extracción de características y confrontación con base de datos. A partir del estudio y la valoración de los algoritmos y técnicas empleadas en los sistemas de detección de rostro se proponen las que presentan mejor rendimiento: *Viola and Jones* para la detección, Jets de Gabor para la extracción de características y distancia euclidiana para la confrontación. Se analizan además las principales problemáticas a las que se enfrenta hoy esta área del conocimiento como son el cambio de iluminación de las imágenes, el cambio de pose y orientación del rostro y otros factores que la hacen un lugar donde aún se debe crear y encontrar soluciones. Experimentalmente se ha implementado un prototipo para probar el procedimiento utilizando OpenCV como biblioteca fundamental para el procesamiento de los datos.

**Palabras claves:** catalogación, procesamiento de imágenes, reconocimiento de rostro.

## **ABSTRACT**

*Image processing offers many advantages nowadays for the storage, location and treatment of audiovisual materials. Particularly face recognition can be applied to different areas such as: authentication systems, immigration control, surveillance systems, among others. In this work is defined a procedure for the use of face recognition in audiovisual cataloging, are described the necessary steps needed to implement it: face detection, feature extraction and comparison with database. After the study and evaluation of algorithms and techniques used in face detection systems are proposed which have better performance: Viola and Jones for the detection, Gabor Jets for the extraction of features and Euclidean distance for confrontation. It also analyzes the main problems that facing today this area of knowledge such as lighting change images, change of pose and orientation of the face and other factors that make the face recognition a flied where is possible to create and find solutions. Experimentally it has implemented a prototype to test the procedure using OpenCV library to the processing of data and show the recognition results.*

**Key Words:** cataloging, image processing, face recognition

## ÍNDICE DE TABLAS

Tabla 1. Resultados del proceso de detección.....	43
Tabla 2. Resultados del reconocimiento de rostro.....	45

## ÍNDICE DE FIGURAS

Figura 1. Problemas a los que se enfrenta la detección de rostros. ....	5
Figura 2: Procedimiento general de reconocimiento de rostro.....	14
Figura 3: Métodos de detección de rostros .....	16
Figura 4: Métodos de reconocimiento de rostro .....	21
Figura 5: El operador LPB básico. ....	26
Figura 6: Procedimiento propuesto para el reconocimiento de rostro.....	29
Figura 7: Resultados reportados con los métodos más representativos .....	30
Figura 8: Resultados experimentales reportados con los métodos .....	33
Figura 9: Filtros de Wavelets de Gabor aplicados en un punto de la imagen .....	34
Figura 10. Método de detección.....	41
Figura 11. Método de extracción de características.....	41
Figura 12. Método de distancia euclidiana .....	42



# ÍNDICE

<b>INTRODUCCIÓN</b> .....	<b>1</b>
<b>CAPÍTULO 1: Fundamento teórico</b> .....	<b>3</b>
1.1. Conceptos asociados al dominio del problema.....	3
1.2. Objeto de estudio .....	4
1.2.1. Descripción general.....	4
1.2.2. Descripción actual del dominio del problema .....	9
1.2.3. Situación Problemática .....	10
<b>CAPÍTULO 2: Solución Propuesta</b> .....	<b>14</b>
2.1. Diseño del procedimiento propuesto para el reconocimiento de rostro.....	14
2.1.1. Adquisición de la imagen.....	15
2.1.2. Detección de rostro.....	15
2.1.3. Procesamiento y creación del vector .....	18
2.1.4. Conformación de la plantilla o normalización .....	18
2.1.5. Extracción de características.....	19
2.1.6. Creación del vector de características.....	27
2.2. Procedimiento propuesto .....	28
2.2.1. Selección manual del rostro .....	29
2.2.2. Validación del rostro, métodos de detección .....	30
2.2.3. Conformación de la plantilla .....	32
2.2.4. Extracción de características.....	32
2.2.5. Confrontación de los resultados .....	36
2.3. Aplicación del procedimiento.....	36
2.4. Conclusiones parciales .....	38
<b>CAPÍTULO 3: Análisis de los resultados</b> .....	<b>39</b>
3.1. Análisis de los resultados del procedimiento .....	39
3.1.1. Creación del proyecto.....	40

3.1.2. Primera etapa: detección de rostro.....	40
3.1.3. Segunda etapa: extracción de características.....	41
3.1.4. Tercera etapa: confrontación.....	42
3.2. Validación del procedimiento .....	42
3.2.1. Detección de rostro.....	43
3.3. Conclusiones parciales .....	46
<b>CONCLUSIONES .....</b>	<b>47</b>
<b>RECOMENDACIONES.....</b>	<b>48</b>
<b>Referencias Bibliográficas .....</b>	<b>49</b>
<b>Bibliografía .....</b>	<b>52</b>
<b>Glosario de Términos .....</b>	<b>56</b>

## INTRODUCCIÓN

El reconocimiento de rostros es en la actualidad una de las problemáticas con mayor grado de dificultad a las que se enfrentan los sistemas computacionales, muy necesario en la creación de aplicaciones que favorezcan el trabajo del hombre. La variedad de software que se nutren de los resultados que ofrecen estos métodos abarcan casi todos los campos, haciéndose más notables en el ámbito de la Seguridad (1).

Por otra parte, el audiovisual está presente en todos los escenarios por los que transita la humanidad, en la última noticia, en la película del momento, en el video más pegado; se utiliza a diario para plasmar recuerdos, vivencias y los hechos más relevantes de la cotidianidad. El cúmulo de materiales audiovisuales es elevado por lo que se necesitan mecanismos automáticos o semiautomáticos para lograr clasificar la información y así favorecer la recuperación de estos de forma rápida y sencilla (2). El conocimiento de las personas que aparecen en un material es un dato relevante a almacenar sobre los mismos, muy útil para su búsqueda, por lo que resulta de gran utilidad el reconocimiento de rostro en ese contexto.

En muchas instituciones donde se trabaja con este tipo de contenidos: televisoras, productoras de cine, videotecas; la catalogación de personas se hace indispensable. En estudios realizados en instituciones donde se realiza la catalogación durante el levantamiento de requisitos de software para el sistema de catalogación de audiovisuales del Departamento Señales Digitales cuyo objetivo es la informatización de este proceso, se ha observado que el personal pocas veces cuenta con la ayuda de herramientas de apoyo para la definición de los personajes que aparecen en los materiales. Muchas de estas etiquetas se completan utilizando para ello la apreciación propia del documentalista al visualizar el material. En el caso de los actores o personas significativas que participan en determinada escena, se dificulta el completamiento de los descriptores, debido a que los métodos mencionados son deficientes y requieren de mucha experiencia en la identificación de rostros. Es por ello que determina en la realización de la actividad de catalogación, el conocimiento del ámbito del material, una amplia cultura general y una buena preparación del individuo que se encarga del completamiento de la ficha de descripción del material audiovisual, si se tiene en cuenta además la variedad de personas a encontrar o similitud entre ellos, se puede comprender que la tarea en particular se hace engorrosa y compleja, desacertándose en muchos de los casos con el nombre correcto de la persona identificada. La producción de nuevo contenido audiovisual se basa en la adquisición de nuevas imágenes y en el uso de las ya existentes en archivo, de ahí la importancia de tenerlo ampliamente descrito para

garantizar que según diversos criterios de búsqueda se responda con celeridad y efectividad al proceso productivo de medias.

En el presente trabajo se realiza el estudio y puesta en práctica de algunos de los métodos y algoritmos más significativos del reconocimiento de rostro para su aplicación en la catalogación de materiales audiovisuales, proceso que se dificulta por lo novedoso de la temática que se aborda donde numerosas universidades de prestigio internacional se encuentran en estos instantes realizando pequeños avances para encontrar soluciones efectivas a los problemas que se exponen en esta área. Su aplicación en un proceso semiautomático que sirva de apoyo al personal que realiza la tarea de catalogar audiovisuales, tributa a la descripción más completa del archivo audiovisual de las instituciones y facilita que la recuperación de escenas, cortos u otro elemento significativo se haga en un menor tiempo y de forma sencilla.

Por tanto se presenta como **problema de investigación científica** que no se obtienen adecuados resultados de búsqueda ante el parámetro de aparición de determinado personaje durante la recuperación de videos catalogados con el sistema de catalogación del departamento Señales Digitales interfiriendo en el proceso de producción y gestión de audiovisuales de la institución donde se aplique.

El **objeto de estudio** de esta investigación está dado por los procesos y métodos involucrados en el reconocimiento de rostros en materiales audiovisuales teniendo como **campo de acción** el reconocimiento de rostros en sistemas de catalogación del Departamento Señales Digitales.

Siendo el **objetivo general** de la investigación el determinar un procedimiento de reconocimiento de rostro para sistemas de catalogación de audiovisuales en el departamento Señales Digitales. Para lograr su realización se plantean las siguientes tareas:

1. Caracterización de los métodos y descriptores visuales factibles en imágenes de rostro.
2. Análisis y resumen de los métodos y técnicas existentes para la identificación de rostros en audiovisuales.
3. Propuesta de un procedimiento a seguir en la obtención de descriptores de rostro, a partir de la utilización de métodos de reconocimiento.
4. Implementación y prueba del procedimiento propuesto.
5. Validación de la fiabilidad del procedimiento propuesto a través del rendimiento de los descriptores utilizados.

Se presenta como **idea a defender** que la determinación de un procedimiento de reconocimiento de rostro para sistemas de catalogación de audiovisuales, contribuirá en el proceder adecuado para la implementación de la identificación de personajes en los sistemas de catalogación del departamento Señales Digitales.

### **Métodos de Investigación Científica**

**Histórico-lógico:** Se utiliza en la investigación realizada para conocer el comportamiento histórico y actual de los diversos procesos de reconocimiento y comparación de rostro, así como de las técnicas y normas empleadas para la construcción de futuros sistemas.

**Analítico-sintético:** Se emplea para el procesado de toda la información y técnicas referentes al reconocimiento de rostro, las cuales son sintetizadas seleccionando las que son viables y necesarias para la propuesta final del procedimiento.

**Comparativo:** Se utilizó al comparar el rendimiento de los algoritmos y técnicas empleados en el reconocimiento de rostros a partir del análisis de las variables que retornan cada uno. Es utilizado para analizar los resultados obtenidos en cada una de las etapas del procedimiento que se propone, desde la detección como primer elemento, hasta el control de los resultados del reconocimiento. Se emplea para comparar la efectividad de los descriptores propuestos en el reconocimiento de rostro.

## **CAPÍTULO 1: Fundamento teórico**

En este capítulo se ofrecerá una descripción detallada de los conceptos principales utilizados en la investigación, se proveerá una visión general de la temática abordada para comprender el contexto teórico sobre el que se trabaja. Se analizará explícitamente la situación problemática con la intención de comprender el dominio del problema.

### **1.1. Conceptos asociados al dominio del problema**

Los conceptos utilizados en la investigación asociados al campo de la informática y específicamente a los medios audiovisuales son diversos. Para el mejor entendimiento del trabajo en desarrollo se han descrito los términos más representativos, un ejemplo claro es multimedia, que se refiere a objetos que usan de forma simultánea varios tipos de contenidos, como pueden ser: video, audio, texto, imágenes, animación y otros (3). Oficialmente se le denomina multimedia, pero en numerosos medios de comunicación y empresas de este sector se le nombra también media, siendo un apócope del término anteriormente mencionado (2).

Los procesos de catalogación están dados por un grupo de actividades enfocadas al registro ordenado de libros, documentos o material semejante, de acuerdo con normas o “control bibliográfico” (4), por lo que resulta importante entenderla basado en ese contexto.

Las entidades que trabajan con materiales audiovisuales se han planteado progresivamente la digitalización de los archivos de medias, proceso que permite la mejor conservación del acervo audiovisual y que da la posibilidad de una mejor organización y gestión del archivo a través de un sistema informático. Como parte de las funcionalidades que se presentan en un sistema informático que trabaja con audiovisuales, se encuentra el procesamiento de imágenes, que es útil para analizar secuencias de imágenes de los materiales, con variadas funciones que pueden estar enfocadas a la calidad del material, identificar los objetos presentes en las secuencias o en este caso como herramienta de apoyo para la catalogación de personajes.

En el procesamiento de imágenes está definido como elemento significativo la presencia del descriptor, que según la Real Academia Española (RAE) es un símbolo válido y formalizado que se emplea para representar inequívocamente los conceptos de un documento o de una búsqueda. Una definición que complementa la anterior y que se ajusta aún más al contexto que se aborda es que los descriptores: son el nexo entre el objeto real (la imagen de rostro en este caso) y la representación que el sistema tiene del mismo.

### 1.2. Objeto de estudio

A continuación se describen los procesos involucrados en el reconocimiento de rostros en materiales audiovisuales. Se presentan las principales actividades asociadas al proceso.

#### 1.2.1. Descripción general

Pocos son los años del nacimiento de los primeros archivos audiovisuales digitales, desde entonces se han venido desarrollando sistemas tecnológicos cuyos objetivos no sólo son digitalizar y gestionar diversas fuentes para su posterior catalogación y almacenamiento, sino la búsqueda de modelos y soluciones que permitan acceder a ellos de manera rápida y eficiente.

El procesamiento de imágenes constituye el punto de partida para la realización del análisis del audiovisual. Su implementación ofrece diversas posibilidades de uso, puede utilizarse de manera automática, es decir, en componentes de software que ejecutan el procesamiento por lotes de materiales o mediante la orden de un operador para asistir el trabajo que realiza, todas las variantes con el objetivo de lograr registrar la mayor cantidad de datos asociados a los documentos audiovisuales que se almacenan en una institución. De manera general la catalogación de personajes en video se puede beneficiar por la utilización de métodos de reconocimiento de rostro que se presentan como una buena herramienta de apoyo a los documentalistas y demás entes que intervienen en el proceso.

El reconocimiento de rostro es un proceso que se realiza a través de tres etapas fundamentales. En la primera se realiza la detección del rostro en la imagen, en la segunda se efectúa la extracción de características y en la tercera se confrontan estas características con la información almacenada y se devuelve un resultado de reconocimiento del rostro (5).

**La detección del rostro en imágenes** se presenta como una etapa de menos dificultad que las posteriores obteniéndose resultados en la actualidad aceptables, pero aún se enfrenta a diferentes problemáticas como son las escalas, iluminación, oclusiones, ángulos de cámara, expresiones variadas o accesorios presentes en la persona. Numerosos autores han hecho referencia a la forma de ocurrencia y repercusión de estas problemáticas (6) (7) (8).



**Figura 1. Problemas a los que se enfrenta la detección de rostros.**

La adquisición de una imagen que esté afectada por las agravantes mencionadas, condiciona el resultado final del procesamiento, es por ello que son múltiples los algoritmos que van enfocados a la normalización de la imagen del rostro y a su detección en entornos variados (9) (10). En el caso de la presente investigación se asumirá la adquisición de la imagen en situaciones controladas donde debe garantizarse que el ruido que afecta a la detección sea el mínimo, ya que el documentalista debe seleccionar una imagen nítida y que presente al rostro con el mínimo de accesorios o elementos que atenten contra el proceso de reconocimiento.

Atendiendo a la forma de proceder en la detección de rostros se han definido dos grupos de métodos:

**Métodos basados en rasgos faciales:** Buscan encontrar aquellas características presentes en cualquier rostro: ojos, cejas, labios, boca, mentón, líneas de contorno, entre otros.

**Métodos basados en la imagen:** Aplican herramientas generales de reconocimiento de patrones para sintetizar un modelo a partir de un conjunto de imágenes de entrenamiento. Trabajan con la imagen completa o una región de esta sin buscar rasgos faciales de forma localizada (5).

**La extracción de características** constituye la etapa más compleja del proceso, es aquí donde se definen los principales descriptores que representarán cada una de las características importantes del rostro para la comparación con los datos almacenados de las personas. Los algoritmos de reconocimiento de rostro, durante la extracción de características se basan en los descriptores para la representación de objetos de interés, se puede plantear que los resultados del sistema mejoran en la medida que mejor sea su representación.

Es conocido que los algoritmos de reconocimiento de rostro han venido evolucionando con el transcurso de los años y la variedad de ellos en la actualidad es elevada, por lo que se han clasificado en dos grupos según los tipos de rasgos que utilizan para el procesamiento: métodos basados en características geométricas y métodos basados en la apariencia (11).



Los métodos basados en características geométricas fueron los más populares en los inicios del procesado automático de rostro (12). Este tipo de métodos se basa en los mismos descriptores que utilizan las personas, explícitamente del conocimiento sobre el rostro: los rasgos faciales (nariz, boca, ojos) y las propiedades y relaciones entre estos (áreas, distancias, ángulos). Estos métodos son eficientes y efectivos cuando alcanzan reducción de la información e insensibilidad a las variaciones de iluminación y pose; además requieren detectar los rasgos faciales con una alta probabilidad y contar con imágenes de buena calidad (13). Estas características geométricas por sí solas son inadecuadas para el reconocimiento de rostros, ya que existe información como: la textura o la apariencia facial, por solo citar algunas, que tienen alto grado de detalles y son obviadas.

Por otra parte los métodos basados en la apariencia, han sido los más dominantes en los últimos años y sobre los que se han puesto los mayores esfuerzos por parte de los investigadores (14), debido a que se enfocan directamente en las intensidades de los píxeles u otras representaciones basadas en la imagen que logran una mejor descripción de la misma, constituyendo un avance significativo para la eficiencia y efectividad de los sistemas de reconocimiento de rostros (15).

Los métodos basados en apariencia se clasifican en dos tipos a partir de las técnicas y algoritmos que utilizan:

**Métodos holísticos:** utilizan toda la imagen de la cara como un vector de entrada al sistema de reconocimiento, siendo esta la unidad básica de procesamiento.

**Métodos basados en características locales:** se representa a partir de vectores de las diferentes regiones del rostro, como pueden ser características locales: ojos, nariz, boca. Sus posiciones y estadísticas locales constituyen la entrada al sistema de reconocimiento (11). Apoyan el criterio de que es mejor representar un rostro mediante un conjunto de vectores de características locales de pocas dimensiones, en vez de un solo vector de grandes dimensiones.

Existen métodos híbridos que combinan técnicas holísticas y locales. Los métodos basados en características locales, comparados con los globales, son más apropiados para el reconocimiento efectivo de rostro con una sola imagen de muestra (5). Los métodos basados en la apariencia o características locales involucran cuatro pasos generales:

1. Partición en regiones locales.
2. Extracción de rasgos locales.
3. Selección de rasgos.

### 4. Clasificación.

Se reporta que la forma más simple y comúnmente utilizada para extraer la información del rostro es dividir la imagen en ventanas rectangulares las cuales pueden estar solapadas o no. Sin embargo, se han utilizado también otras formas como elipses y franjas (16). El tamaño y la forma de las regiones varía en dependencia del método utilizado, es esta una cuestión que tiene influencia directa en la robustez de los métodos, por lo que aún se investiga para llegar a estandarizar una propuesta de uso en este sentido. Cuando las regiones locales han sido definidas, es necesario decidir cómo representar la información. La meta es representar la mayor cantidad de información posible y para esto se utilizan los descriptores.

Una vez que se han extraído las características del rostro representadas en los descriptores se procede a la clasificación de los mismos, cuyo objetivo consiste en retener solamente un subconjunto del vector de rasgos original, evitando la pérdida de información discriminativa. Este paso será necesario en dependencia de la forma de representar la información, algunos métodos de extracción de rasgos incluyen ellos mismos un paso de selección. Para la clasificación se utiliza comúnmente la combinación de clasificadores, esta puede ser llevada a cabo de dos formas diferentes: concatenando en un solo vector los rasgos de las diferentes regiones o realizando la clasificación por separado en cada región y luego tomar la decisión final combinando el resultado.

Atendiendo a que la información que recogen estos métodos básicamente está definida por los descriptores de imagen, a continuación se detallará la función de los mismos y se agruparán según sus características para lograr una mejor comprensión del proceso.

**Los descriptores de imagen** constituyen “el lenguaje” en que el sistema reconocerá las características de la imagen que se procesa. Idealmente, un descriptor visual debería poseer las siguientes propiedades:

- **Simplicidad:** El descriptor debería representar las características extraídas de la imagen de manera clara y sencilla para permitir una fácil interpretación de su contenido.
- **Repetibilidad:** El descriptor generado a partir de una imagen debe ser independiente del momento en el que se genere.
- **Diferenciabilidad:** Dada una imagen, el descriptor generado debe poseer alto grado de discriminación respecto a otras imágenes y al mismo tiempo contener información que permita establecer una relación entre imágenes similares.

- **Invarianza:** Cuando existen deformaciones en la representación de dos imágenes, es deseable que los descriptores que las representan aporten la robustez necesaria para poder relacionarlas aún bajo diferentes transformaciones.
- **Eficiencia:** Es importante que los recursos consumidos para generar el descriptor sean aceptables para poder ser utilizados en aplicaciones con restricciones críticas de espacio y tiempo (17).

De acuerdo al grado de profundidad de la representación del contenido llevada a cabo por los descriptores y dependiendo del nivel de abstracción al que se refieran es posible clasificarlos en dos grupos:

**Descriptores de información general:** engloban los descriptores también llamados de bajo nivel, que proporcionan una descripción respecto del color, formas, regiones, texturas y movimientos presentes en la imagen.

**Descriptores de información de dominio específico:** también llamados descriptores semánticos, proporcionan información acerca de los objetos y eventos que constituyen la escena. Se apoyan en los descriptores visuales para llevar a cabo la difícil tarea de realizar una descripción semántica de las imágenes. Un ejemplo podría ser el reconocimiento de objetos dentro de una imagen o la extracción de características relacionadas con sentimientos o sensaciones (17).

Se analizarán los descriptores de información general, pues muestran resultados en cuanto al área de reconocimiento de rostro. Se clasifican los descriptores de información general atendiendo a qué parte de la imagen procesada representan:

- **Descriptores globales:** son aquellos que toman la imagen como un todo y no utilizan información a priori o sobre el objeto a representar. Resumen el contenido de la imagen en un único vector o matriz de características. Poseen la ventaja de encapsular una gran cantidad de información de la imagen requiriendo una pequeña cantidad de datos para describirla. A pesar de su simplicidad, este tipo de descriptores han resultado ser ampliamente utilizados para diferentes tareas, debido entre otras cosas a su bajo coste computacional unido a prestaciones relativamente buenas.
- **Descriptores locales:** son los que extraen información de puntos o regiones particulares de la imagen, considerando el hecho de que se trabaja con un objeto dado. Actúan sobre regiones de interés, previamente calculadas o identificadas, construyendo un vector de características de esa región que tiene en cuenta la información contenida tanto en el punto de interés como en la

región adyacente al mismo o vecindario. Son aplicables al área de reconocimiento de rostros los descriptores locales, ya que el rostro será tomado como el objeto a representar (5).

Durante la **confrontación de la información** se comparan las características extraídas con las almacenadas de cada uno de los rostros se presenta como la etapa final del proceso y se basa en métodos de comparación de características que buscan encontrar esas diferencias o similitudes que existen entre los rostros que se analizan, en muchos de los casos se utilizan algoritmos de similitud como puede ser similitud entre vectores de Wavelets de Gabor (Jet), ponderados o por magnitud o algoritmos que buscan definir la distancia entre los vectores o grafos de características, como es el caso de la distancia del coseno. Dependiendo del modo de funcionamiento del sistema se compara contra el registro del mismo usuario, modo verificación, o contra los registros de todos los usuarios, modo identificación.

### 1.2.2. Descripción actual del dominio del problema

El trabajo con archivos multimedia en una entidad que maneja audiovisuales es tarea diaria, la generación de materiales es constante, siendo el proceso de catalogación uno de los más importantes dentro de la empresa. El proceso de catalogación de audiovisuales en el ámbito digital incluye un flujo de actividades que se describen en este acápite para una mejor comprensión del entorno en el que se aplicarán.

El documentalista audiovisual, que es la persona con competencias para realizar esta labor, realiza la búsqueda del material audiovisual que debe catalogar atendiendo a los datos primarios que se introducen del material durante su digitalización e incorporación al servidor de almacenamiento. Una vez localizado el material se procede a su descripción, la que se basa en una ficha de catalogación que contiene datos según el tipo de audiovisual que se esté catalogando. Finalmente el audiovisual es almacenado en el archivo digital de conjunto con su ficha de catalogación para su conservación y futura utilización.

Es precisamente durante la descripción del audiovisual que se deben introducir los nombres de los personajes significativos que hacen aparición en él, pero puede darse el caso de que el documentalista no sea capaz de identificar a una personalidad determinada por varias razones: que por error no lo identifique como alguien importante o que no conozca su nombre o papel social. Los dos casos están muy ligados, pero se observa con frecuencia que los documentalistas pueden identificar que se trata de una persona importante sin embargo no conocen su nombre u origen. Para el primer caso aislado la solución no está a la mano del software con el que se interactúa y deberá estar encaminada a

fortalecer la preparación integral del individuo, pero para atenuar el efecto del segundo caso se puede incluir en el sistema informático una herramienta que le permita al operador hacer más fácil y efectiva su tarea, será allí donde intervendrá en el sistema el procedimiento de reconocimiento de rostros.

### 1.2.3. Situación Problemática

La catalogación de medias es una tarea pendiente y de constante realización en las empresas que manejan materiales audiovisuales, este proceso en muchas televisoras, videotecas, empresas que trabajan con medias, no está automatizado en todas sus etapas y no se ajusta a las exigencias actuales para el trabajo con archivos multimedia.

Uno de los problemas que afecta la adecuada catalogación de los materiales es la falta de información que pueden tener los documentalistas sobre los personajes que aparecen en los mismos, dato que es oportuno conocer para disímiles usos. Se da con frecuencia que en la etapa de recuperación de los audiovisuales archivados en la entidad, se define con prioridad la etiqueta de personajes presentes en el material, representando para periodistas, editores y demás especialistas que trabajan con este tipo de elementos, un atributo de inestimable valía. La descripción en este caso es incompleta o deficiente, a partir de que durante la realización del proceso de catalogación el documentalista o personal que trabaja en el archivo no contó con todas las herramientas necesarias para identificar correctamente a los personajes que aparecen en un material audiovisual e insertarlos en su ficha de catalogación.

No contar con la información completa de los personajes que aparecen en los materiales influirá directamente en la completitud del catálogo y por tanto en la efectividad de las búsquedas que realicen los usuarios, que puede ocurrir cuando no se tienen etiquetados a todos los personajes, afectándose desde los resultados de la búsqueda hasta la producción de nuevos materiales.

Estos elementos dan como resultado una inadecuada gestión del patrimonio audiovisual que conserva la institución, que se ejemplifica en una recuperación pobre del contenido y en tiempos altos en el proceder de cada búsqueda de materiales audiovisuales. Por ello en la medida que se tenga automatizado el proceso, mediante la utilización de herramientas que presenten las opciones necesarias para el adecuado funcionamiento, la calidad del servicio crecerá exponencialmente.

### 1.3. Análisis de otras soluciones existentes

En la bibliografía consultada, en Cuba se presentan un grupo de trabajos del Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV) que se especializa en investigaciones sobre diversas temáticas, entre ellos la detección de patrones y el reconocimiento de rostro como rama de esta área

del conocimiento. Esta institución hasta la fecha ha realizado estudios del estado del arte de variadas temáticas sobre reconocimiento de rostro, como son: el estado actual de los algoritmos de reconocimiento de rostro usando tecnología GPU, el estado actual de la indexación en espacios métricos aplicado al reconocimiento de rostros, estado del arte de los métodos de extracción de rasgos invariantes a la iluminación de rostros, así como otros temas de reconocimiento enfocados en el seguimiento de rostros en video. Se ha analizado como relevante para la presente investigación la recientemente publicada tesis de doctorado: Algoritmo de Reconocimiento de Rostros basado en la Apariencia Local para Aplicaciones Reales en Condiciones Variables de Iluminación, de Heydi Méndez Vázquez, que se enfoca en la creación de un algoritmo para el reconocimiento del rostro.

El algoritmo propuesto por Méndez se enfoca principalmente en el análisis de la influencia de la iluminación en el rostro, determinando como parte de su funcionamiento si se encuentra la imagen del rostro comprometida por esta variable. El algoritmo desempeña una etapa de pre-procesamiento que es más eficaz que los métodos existentes en esta área bajo condiciones de iluminación variables.

De forma general se propone en el referido trabajo un algoritmo de reconocimiento de rostro que incluye las etapas clásicas del proceso, que incorpora una etapa extra para el análisis de la incidencia de la iluminación en el rostro, que en caso de estar presente, logra corregirla, pasándose posteriormente a realizar la extracción de características la cual alcanza mejores resultados a partir del pre-procesado. En detección se emplea *Viola and Jones*, en extracción se utiliza Patrones Binarios Locales Multiresolución (MLBP) con la clasificación basada en LDA y en la etapa de confrontación utilizan una función de similitud que no se define en la investigación.

Los inconvenientes que presenta este trabajo para la tesis que se desarrolla es que la propuesta va enfocada a la solución de la problemática de la iluminación, aspecto en los que obtienen una alta efectividad, y que es válido sobre todo en cámaras de seguridad en locaciones como la aduana donde las fotos son frontales y los ángulos de la pose son de menos de quince grados, elementos que son útiles para el trabajo que se presenta pero no definitivos debido a que el catalogador en este caso deberá seleccionar imágenes de rostros poco afectados por la iluminación y con un buen grado de orientación del mismo.

En el ámbito internacional, se encontraron diversos trabajos enfocados a la solución de problemas en el área del reconocimiento de rostro, en muchos casos solo se planteaban temáticas específicas como la detección o solo la extracción de características. También son muchos los estudios que se presentan en determinada técnica de una de las etapas que comprenden el proceso, solo dos trabajos de los analizados modelan el proceso de manera general y es el caso de un proyecto desarrollado en

Uruguay denominado Proyecto Aguará, que se perfila en la creación de un sistema biométrico para una aplicación de control de acceso.

Aguará en su exposición presenta un estudio bastante completo de la utilización de un algoritmo orientado al reconocimiento de rostro y de manera práctica exponen resultados de su trabajo en la creación de un sistema biométrico. Para ello se basan en Facel, un proyecto desarrollado en la Universidad Estatal de Colorado que realiza la detección y reconocimiento de rostro a partir de una cámara web. Ambos proyectos fueron útiles para el desarrollo de la presente investigación, pero de forma general resultaron incompletos dado que solo se exponen en el primer caso, elementos teóricos que se ponen en práctica en una aplicación privativa a la cual no se pudo acceder y en el segundo caso se estudiaron los trabajos que fueron publicados por la Universidad Estatal de Colorado, pero solo brindan demos desarrollados en python y ejemplos de prueba que impiden su implementación en un sistemas propio.

Por otra parte se puede apreciar en esta área bibliotecas de algoritmos disponibles para la creación de sistemas de biometría o en este caso para el trabajo con audiovisuales, ejemplo de ello es: VeriLook Extended SDK, Morpheus SDK o FaceVACS SDK, que gozan de buen prestigio internacional pero que presentan licencias muy costosas para Cuba, mostrándose como una solución inviable.

El procedimiento que se expone en el presente trabajo va encaminado a implementar las diferentes funcionalidades con la utilización de bibliotecas de código abierto que permitan su integración en aplicaciones comerciales, elemento que realza su utilidad. Es por ello que el procedimiento que se propone debe definir las etapas clásicas del reconocimiento de rostros, pero que se presente como único y diferente, que combine las técnicas relevantes en cada etapa y que además utilice tecnologías propias o herramientas que permitan su integración en sistemas comerciales del Departamento Señales Digitales, los cuales deben perfeccionarse para lograr altos niveles de calidad.

### **1.4. Conclusiones**

Una vez realizado un bosquejo de las diferentes técnicas que se utilizan en la temática tratada y haberse entendido el dominio del problema así como su situación problemática se puede concluir que: el área del conocimiento en la que se desenvuelve la investigación es amplia y constituye un reto, a partir de la dificultad que representa como proceso para la comunidad internacional. El tema está sujeto a numerosas problemáticas, entre ellas: la iluminación, la pose de la persona, presencia de lentes u otros accesorios, que pueden atenuar el resultado esperado. Se puede apreciar además que el campo donde mejores resultados se han obtenido es en el de la seguridad, a partir de que se tiene

una mayor cantidad de imágenes de rostro de una misma persona, lo que contribuye que la realización de la confrontación tenga más posibilidades de acierto, dado que se cuenta con un espectro más amplio de comparación. Se presenta además que para la problemática de este trabajo existen soluciones que se acercan a lo esperado, teniendo en cuenta que el resultado final constituye una herramienta de apoyo al documentalista o catalogador, donde lo que se necesita es similitud entre rostros, lo que posibilita un margen de error un poco más abierto. Es esencial tener en cuenta que las técnicas que se propongan deben ser rápidas y precisas, debido a que estos aspectos definirán la efectividad del proceso de la catalogación.



## CAPÍTULO 2: Solución Propuesta.

Las tecnologías informáticas se han desarrollado y perfeccionado vertiginosamente con el objetivo de crear software más robustos y funcionales, que contribuyan al propio desarrollo del hombre; su estudio y uso adecuado garantiza, que las fronteras del conocimiento cada día se extiendan un poco más. En este capítulo se realiza la descripción del proceso en general, viéndose en cada etapa los diferentes métodos y técnica que se emplean, decantando en una propuesta final para el presente trabajo.

### 2.1. Diseño del procedimiento propuesto para el reconocimiento de rostro

Según la Real Academia Española un procedimiento es el método de ejecutar algunas cosas o la acción de proceder de forma adecuada frente a determinadas actividades (18).

Para un entendimiento del proceso y que a la vez permita exponer los diferentes métodos y técnicas a emplear en cada etapa, se han representado gráficamente en la Figura 2 los subprocesos que intervienen desde que una imagen es capturada hasta que se compara con las almacenadas.

El procedimiento se ha separado en varios bloques o actividades que se ejecutan una vez terminado el anterior. En el primer bloque se realiza la adquisición de la imagen, la que puede ser de forma automática o semiautomática. Luego se ejecuta el procesamiento y creación del vector de la imagen, para finalmente efectuar la confrontación y mostrar el resultado. Este diagrama representa el flujo de un sistema de reconocimiento de forma general, los cuales se tendrán en cuenta para la definición del procedimiento.

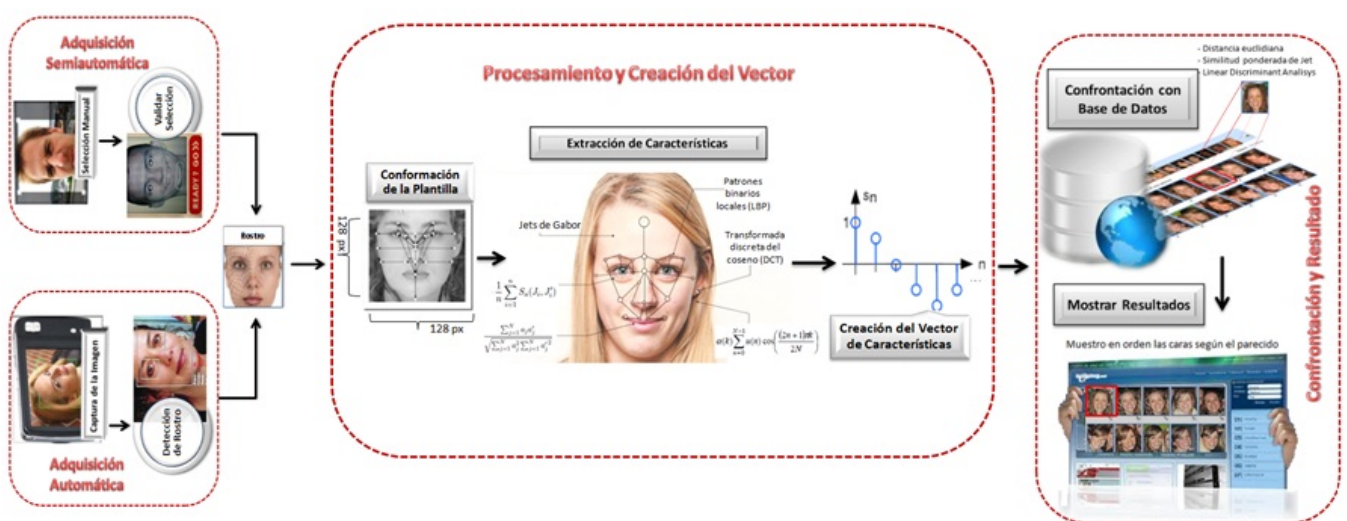


Figura 2. Procedimiento general de reconocimiento de rostro.

La creación de un sistema de reconocimiento de rostro, debe tener en cuenta el campo de acción donde se va a efectuar finalmente, debido a que esto puede conllevar a la utilización explícita de determinadas técnicas y métodos que conformarán los pasos a seguir. Es por ello que se presenta a continuación de forma general las diferentes técnicas y métodos que resalta la literatura en cada etapa, para definir finalmente una propuesta ajustada a aspectos muy vinculados propiamente a este campo de acción.

### **2.1.1. Adquisición de la imagen**

La adquisición de la imagen inicial para la realización del proceso constituye un primer paso significativo, como se presenta en el gráfico de la Figura 2, puede ser realizado de manera automática o semiautomática, en ambos casos se debe efectuar una validación de la imagen que se presenta para definir la existencia de un rostro en la escena. La técnica utilizada en este paso deberá tratar de atenuar las variables que atentan contra la detección como puede ser la iluminación, utilización de accesorios, oclusión, fondo complejo entre otros factores que fueron mencionados en el capítulo 1.

### **2.1.2. Detección de rostro**

En esta etapa se hace indispensable la utilización de métodos de detección de rostros, presentándose a continuación una selección de las más representativas, agrupadas en diferentes categorías a partir de las características comunes que poseen, definiéndose su robustez y utilización según el entorno donde se vayan a emplear.

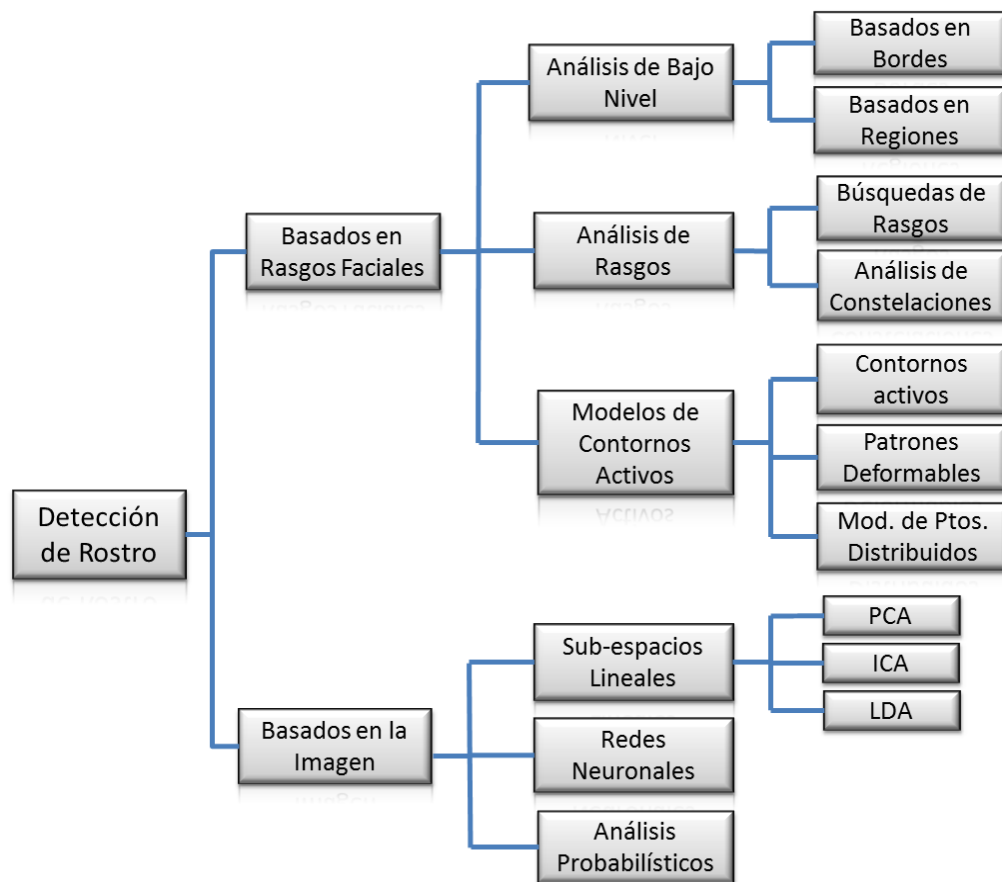


Figura 3. Métodos de detección de rostros

Los métodos de detección de rostros se clasifican según la forma en que realizan la detección, más específicamente lo que buscan en la imagen para determinar la presencia del rostro. La primera clasificación de los métodos de detección está basada en la forma en que se analiza la imagen para detectar un rostro, que fue explicada anteriormente; dentro de cada grupo de métodos se puede distinguir lo siguiente:

Los métodos **basados en rasgos faciales** se agrupan en tres subgrupos a partir de características y formas de proceder que tienen en común:

- **Análisis de bajo nivel:** se enfocan en técnicas que trabajan directamente con los píxeles, principalmente existen dos:
  - Basados en bordes, que buscan bordes, los afinan, etiquetan y finalmente buscan estructuras similares a las de una cara.
  - Basados en regiones, que aprovechan el hecho de que hay zonas más oscuras que el resto de la cara (cejas, pupilas, etc), umbralizan y separan la imagen en regiones.

Localizan la cara comparando la distribución de las regiones presentes con la distribución de regiones tipo de una cara.

- **Análisis de rasgos faciales:** son una familia de métodos que buscan encontrar implícitamente los rasgos faciales. Se basan fuertemente en las relaciones geométricas que cumplen los diferentes rasgos presentes en una cara (16), presentándose dos grandes aproximaciones al respecto:
  - Búsqueda de rasgos, que intentan realizar una búsqueda ordenada de los rasgos característicos de una cara. Por ejemplo, primero buscan la frente, luego los ojos, continúan con la nariz, etc. Se basan en hipótesis sobre la pose y orientación de la cara y utilizan heurística.
  - Análisis de constelaciones, que buscan levantar algunas de las hipótesis de los métodos anteriores sobre la pose y orientación de la cara. Se basan en un análisis probabilístico de la ubicación de ciertos puntos característicos (constelación) de una cara.
- **Análisis mediante modelos de contornos activos:** son métodos basados en modelos de contornos activos que buscan adaptar un modelo genérico de un rasgo (ojo, boca, contorno de la cara) a la imagen o porción de imagen en cuestión (5). Para esto, buscan iterar deformando el modelo hasta adaptarlo al rasgo buscado. Se basan fuertemente en la información local de la imagen (bordes, nivel de gris) encontrando tres grandes técnicas:
  - Snakes o contornos activos, que son comúnmente utilizados para encontrar el contorno de la cara. Se basan en la minimización de una función de energía para adaptar el modelo.
  - Patrones deformables, buscan adaptar modelos paramétricos de cada rasgo. Al igual que los Snakes se basan en minimizar una función de energía para la adaptación.
  - Modelos de puntos distribuidos (PDM), son una manera compacta de representar de manera paramétrica las formas buscadas. El ajuste de estos modelos se basa en discretizar el contorno PDM en diferentes puntos etiquetados. Las variaciones posibles de estos puntos se guardan en un modelo estadístico realizado cuidadosamente a partir de un conjunto de entrenamiento.

En los métodos **basados en la imagen** se pueden apreciar variadas técnicas, entre ellas:

- **Métodos de sub-espacios lineales:** se basan en representar las imágenes de caras en espacios lineales. Para esto buscan encontrar el espacio lineal al que pertenecen mediante un análisis estadístico. Existen diferentes aproximaciones entre ellas: PCA (Análisis de

Componentes Principales), ICA (Análisis de Componentes Independientes), LDA (Análisis de Discriminante Lineal). Cada uno de estos métodos se diferencia del otro en la forma de hallar el subespacio de caras. Hasta el día de hoy son los métodos más explorados, posiblemente debido a su generalidad para aplicarlos a toda clase de problema de reconocimiento de patrones (5).

- Las **Redes neuronales**: se presenta como otro método de este grupo, que necesitan de una etapa inicial de aprendizaje con conjuntos representativos de entrenamiento, para que luego el funcionamiento sea el deseado. Existen aproximaciones locales de redes neuronales (no holísticas), cuyo objetivo es encontrar diferentes estructuras locales y luego en base a ellas reconocer el objeto en cuestión (5).
- **Análisis probabilísticos**: se basan en principios fundacionales del reconocimiento de patrones como pueden ser el principio de máxima verosimilitud. Esta clase de métodos busca estimar distribuciones de probabilidad mediante histogramas y luego comparar los mismos frente a histogramas medios aprendidos estadísticamente de imágenes con caras (5).

### 2.1.3. Procesamiento y creación del vector

El siguiente subproceso define los pasos fundamentales en la extracción de características del rostro procesado, constituye el hito más significativo dentro del proceso, el cual va a determinar los descriptores adecuados, los métodos de reconocimiento a utilizar y la creación del vector de características que se empleará para la correcta realización de la etapa de confrontación y devolución de resultado.

### 2.1.4. Conformación de la plantilla o normalización

La imagen que se tiene luego de la detección del rostro, carece de un grupo de elementos necesarios para realizar una adecuada selección de sus características, es por ello que se le realiza un proceso de normalización, donde se estandarizan un grupo de elementos importantes que se explicarán a continuación.

El procesamiento de imágenes conlleva en la mayoría de los casos un alto coste computacional estando determinado por el volumen y tamaño de la representación con que se trabaje. Teniendo en cuenta este elemento existen técnicas que se han enfocado en la reducción del tamaño como es el caso del análisis de componentes principales (PCA) que se enfoca en representar las medidas numéricas de un conjunto de variables en un espacio de dimensión mucho menor. Esta representación

debe ser tal que al eliminar dimensiones superiores la pérdida de información sea mínima. Por lo tanto, la idea es representar cada imagen en un espacio de baja dimensión (19).

Según estudios realizados por algunos autores el procesamiento está determinado por la cantidad de píxeles de una imagen, el cual puede demorar pocos segundos para imágenes de 32x32 píxeles, y puede demorar minutos para imágenes de más de 128x128 píxeles. La reducción del espacio siempre está condicionada a la pérdida de información, por lo que se debe determinar un tamaño adecuado a las exigencias de los métodos de reconocimiento que se utilizarán.

Otro elemento a tener en cuenta en la etapa de normalización es la apariencia del color del rostro, que está determinada por las condiciones de iluminación existentes en la escena. Para la solución de este problema algunos autores recomiendan aplicar la técnica de compensación de la iluminación como puede ser definir un blanco de referencia considerando los píxeles que pertenecen al 5% superior de valores de luminancia. Luego, los canales R, G y B de la imagen se ajustan de manera que el valor medio de los píxeles pertenecientes al blanco de referencia sea escalado linealmente a 255 (23).

La normalización de los bordes de la imagen es algo más a considerar, estos pueden provocar falsos positivos, por lo que se realiza un suavizado de los mismos. Esto hace con el objetivo de evitar efectos de borde introducidos durante el proceso de selección o escalado del área del rostro (20).

Es importante destacar como otro caso de normalización, la selección de zonas significativas dentro de la imagen, como pueden ser: los ojos, cejas, boca, nariz u otra de interés, las cuales son segmentadas para su procesado por separado, sumándose finalmente los resultados que arrojen cada una de las áreas. Es empleado por muchos autores como un mecanismo de extracción de características en zonas representativas, siendo avalado por sus buenos resultados.

Los elementos de normalización que se exponen, son los más significativos atendiendo a las técnicas y descriptores que se van a utilizar en este trabajo para la extracción de características, su uso está condicionado a la finalidad propia de la etapa y a los parámetros de que exigen las etapas posteriores del procedimiento.

### **2.1.5. Extracción de características**

Esta etapa del procedimiento representa como se ha mencionado anteriormente un paso crítico. De la fiabilidad de los descriptores seleccionados, depende que en la etapa de confrontación, se obtengan los mejores resultados. Luego de haber concluido el proceso de normalización es necesario decidir

cómo representar esa información y es cuando se definen los descriptores, que van a estar asociados a los métodos o técnicas que se empleen.

Se utilizan normalmente vectores formados directamente con las intensidades de los píxeles (21) (22) o rasgos derivados de estos como: la transformada discreta del coseno (DCT) (23), Wavelets de Gabor, Wavelets de Harr (24) (25), entre otros.

Estos descriptores van desde los más complejos hasta los más simples, estando determinados por la intensidad, las texturas o rasgos propios del descriptor, que tienen en cuenta en algunos casos las variaciones de iluminación del área donde se enfoca.

Los métodos de reconocimiento de rostro basados en una sola imagen de muestra se pueden clasificar en dos grupos de acuerdo al tipo de rasgos que utilizan: métodos basados en características geométricas y métodos basados en la apariencia, los cuales se detallan en el capítulo 1. Los métodos basados en la apariencia se han mostrado más efectivos y también pueden clasificarse de acuerdo a la forma en que se utilizan: holísticos (global) y basados en características locales. Dentro de estas clasificaciones se puede destacar lo siguiente:

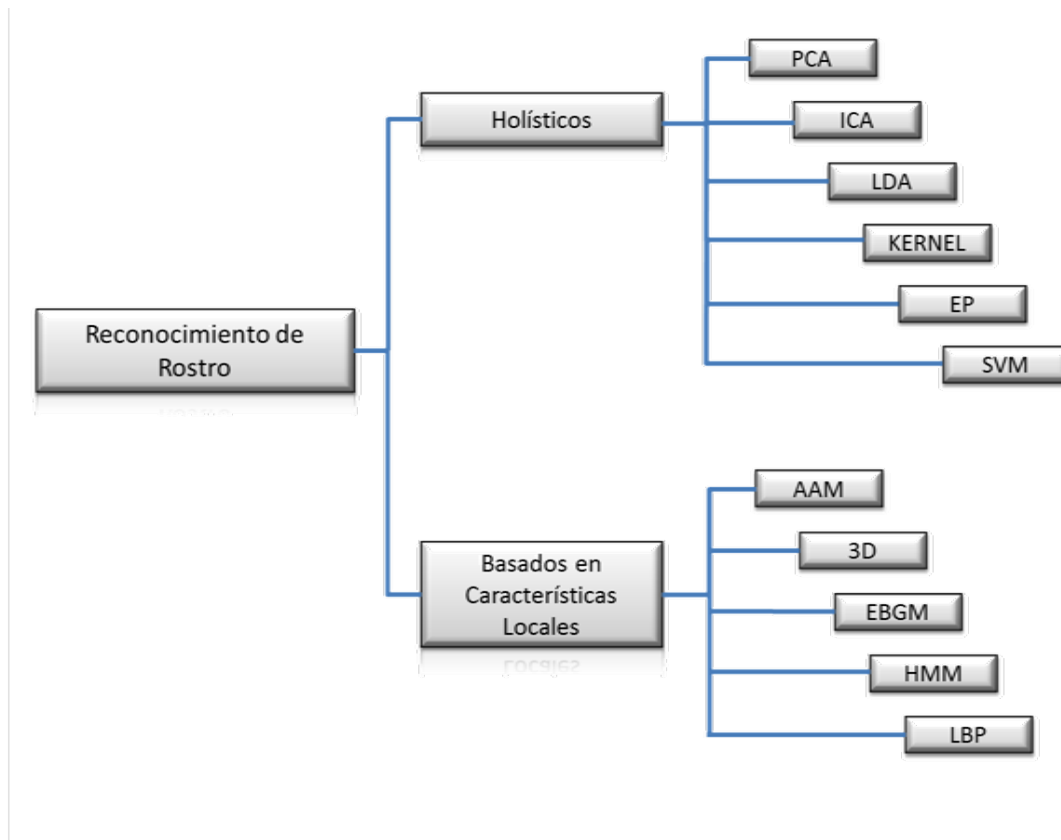


Figura 4. Métodos de reconocimiento de rostro

Los métodos holísticos agrupan una serie de algoritmos que presentan características comunes en cuanto a cómo realizan el procesamiento del rostro para la extracción de las principales características, como es el caso de:

- **Análisis de Componentes Principales (PCA):** que teniendo en cuenta la distribución de imágenes que contienen caras, intenta capturar la variabilidad de estas imágenes, buscando la independencia de cualquier cara o característica particular. Esto se logra aplicando la transformada de Karhunen-Loeve al espacio de caras. Se denomina PCA dado que se busca extraer, de un conjunto de imágenes de entrenamiento, un subespacio cuya base maximice la varianza del espacio original. A estos vectores se les denomina Eigenfaces dado que son los vectores propios correspondientes a los valores propios más grandes de la matriz de covarianza. Se logra así reducir considerablemente la dimensión del problema pasando a trabajar con las proyecciones de las caras originales en el subespacio generado. Fijando luego una métrica, se implementa el reconocimiento encontrando la distancia entre el vector de características de la imagen de entrada y los vectores almacenados en la base de datos (26).



- **Análisis de Componentes Independientes (ICA):** es otro método de la misma familia que intenta representar el espacio de caras en un subespacio que minimice la dependencia de segundo y de mayor orden entre sus componentes. Se asume que las señales de entrada son combinaciones de fuentes no observables estadísticamente independientes. Si la combinación es lineal, se puede definir una matriz de combinación cuyos coeficientes son los que definen la combinación lineal. Entonces ICA estima la matriz inversa de la matriz de combinación. Vale acotar que para aplicar ICA se utiliza previamente PCA para reducir la dimensionalidad del espacio original de caras para disminuir el costo computacional (27).
- **Análisis de Discriminante Lineal (LDA):** este algoritmo intenta llevar el espacio de caras a un subespacio de baja dimensionalidad que aumente la separabilidad de las clases presentes. La idea del algoritmo es encontrar la base de vectores en un subespacio que mejor discrimine entre las diferentes clases, en el caso del reconocimiento las identidades. Se utilizan todas las muestras de todas las clases y se calcula la matriz de dispersión entre clases distintas (inter-clase) y la matriz de dispersión en la misma clase (intra-clase). Se busca maximizar la relación entre el determinante de la matriz inter-clase y el determinante de la matriz intra-clase. Los elementos de la base que maximiza la relación anterior, se denominan Fisherfaces (28).

Los tres métodos antes mencionados son definidos como de baja dimensionalidad, teniendo este aspecto en común, según estudio realizados. Representan además la base para el resto de los algoritmos que se enfocan en esta área.

- **Métodos basados en Kernels:** este grupo de métodos que se describen a continuación son una generalización a los métodos de análisis de componentes (PCA, IDA, LDA). En los métodos de componentes, se construye un subespacio que cumpla determinadas restricciones y luego se elige una base que lo genere de alguna manera particular, componentes de mayor potencia, mejor discriminación. La novedad introducida por los métodos de Kernels es poder tener en cuenta momentos de mayor orden sin tener un costo computacional excesivamente grande. Buscan llevar el problema de clasificación a un espacio de mayor dimensión donde las clases sean linealmente separables.

Dentro de las ventajas de estos métodos se puede destacar que obtienen un mejor desempeño que los métodos de Análisis de Componentes. Además intentan capturar la información presente en los momentos estadísticos de mayor orden. Como contrapartida requieren un mayor costo computacional que los métodos de subespacios lineales y además presentan los mismos problemas frente a iluminación no uniforme y variación de pose (29).

- **Evolutionary Pursuit (EP):** este método, al igual que PCA, IDA y LDA se basa en el análisis de componentes. EP plantea una manera novedosa de obtener una base de vectores eficiente para la representación de las imágenes de caras. Para encontrar la base, se realiza una búsqueda de manera que se trate de maximizar una función *fitness* que mide al mismo tiempo la precisión de la clasificación y la habilidad de generalización del sistema. Como el problema de buscar la base óptima es una dificultad de dimensión alta se utiliza un tipo de algoritmos genéticos al cual se le llama Evolutionary Pursuit.

Esta técnica novedosa obtiene una mejor performance que PCA, ICA o LDA debido a que es más general, es decir aprenden mejor que los anteriores con las mismas imágenes de entrenamiento. La desventaja principal es que se debe resolver un problema de búsqueda en un espacio de muy alta dimensión (30).

- **Support Vector Machine (SVM):** constituye un método genérico para resolver problemas de reconocimiento de patrones. Dado un conjunto de puntos en un determinado espacio que pertenecen a dos clases distintas, SVM encuentra el hiperplano que separa la mayor cantidad de puntos de la misma clase del mismo lado. Esto se realiza maximizando la distancia de cada clase al hiperplano de decisión denominado OSH (*Optimum Separating Hyperplane*). Los puntos más cercanos al hiperplano, de cada conjunto en cuestión, son los llamados vectores soporte (*supportvectors*). Para aplicar SVM es necesario reducir la dimensionalidad de los datos para lo que se utilizan herramientas como PCA o LDA. En general se obtienen resultados superiores a los métodos clásicos de análisis de componentes y también se logra cierta robustez frente a cambios en la iluminación (31).

En el otro grupo, se encuentran los métodos basados en características locales, que se agrupan a partir de que se enfocan propiamente en elementos más específicos dentro del rostro y son el caso de:

- **Active Appearance Model (AAM):** que es un modelo estadístico de la forma y la apariencia en niveles de gris del objeto de interés que se puede generalizar a casi cualquier ejemplo válido de dicho objeto (32).

Ajustar dicho modelo a una imagen implica encontrar los parámetros del modelo que minimizan la diferencia entre la imagen y una síntesis del modelo proyectado en la imagen. El mismo se genera combinando modelos de variación de la forma y la apariencia de la cara. Estos se construyen a partir de imágenes de prueba donde se marcan puntos de interés (puede ser de forma manual o automática).

Dentro de las ventajas se destaca poder obtener representación de vistas no conocidas de la cara a partir de la interpolación de poses conocidas. Como aspecto negativo se menciona la complejidad en la implementación. Además los resultados son fuertemente sensibles a la correcta selección de los puntos principales.

- **Modelado 3D:** cuya idea central de esta familia de técnicas es ajustar un modelo genérico en tres dimensiones (3D) a cada imagen en cuestión. Existen distintas técnicas de adquisición de imágenes (o reconstrucción) 3D, entre ellas: cámaras estéreo, escáner láser, SLS (*Structure Light System*), secuencia de imágenes 2D (33) (34).

Primero mediante una serie de imágenes 3D de unos pocos individuos se construye un modelo de cara genérico 3D (por ejemplo un modelo de malla triangular). Luego dicho modelo es ajustado de manera iterativa, para cada individuo de la base completa. Finalmente se registran las variaciones respecto al modelo genérico y a su vez alguna información de textura.

Los métodos basados en modelos 3D son menos sensibles a variaciones en la pose o en cambios en la iluminación que otros métodos. Su principal contrapartida es que necesitan imágenes de adquisición compleja (Adquisición 3D), además de ser extremadamente costos computacionalmente.

- **Elastic Bunch Graph Matching (EBGM):** este método consiste en extraer una representación de la cara en forma de grafo y el reconocimiento se realiza comparando los grafos correspondientes a las distintas imágenes. Se define un conjunto de puntos principales, como por ejemplo las pupilas, esquinas de la boca, etc. Un grafo etiquetado representando una cara consiste en lo siguiente: nodos ubicados en los puntos principales y las aristas que se forman entre parejas de nodos. Cada nodo es etiquetado con los denominados jets y cada arista es etiquetada con la distancia entre los nodos correspondientes (35). Los jets se basan en una transformada wavelet definida como la convulsión de la imagen con una familia de núcleos de Gabor con distintas frecuencias y orientaciones.

Luego de extraído el grafo de la imagen de prueba, se utiliza una función de similitud para evaluar la similitud de dicho grafo con los grafos de la base de datos. Un criterio de decisión puede ser tomar el máximo valor de similitud como cara encontrada.

En este proceso es necesaria una etapa de marcado manual de puntos para la generación del modelo estadístico, el algoritmo logra buenos resultados con una única imagen por persona. El

problema principal de este algoritmo es el costo computacional de convulsionar la familia de núcleos de Gabor, es decir de machear o sobreponer los filtros sobre determinados puntos de la imagen para generar datos estadísticos que se emplean en la creación del vector.

- **Modelos de Markov escondidos (HMM):** es un conjunto de modelos estadísticos utilizados para caracterizar las propiedades estadísticas de una señal. Para el caso de imágenes, las dependencias temporales se transforman en dependencias espaciales, por lo que es necesario encontrar una forma de recorrer una señal en dos dimensiones (36).

HMM consiste en dos procesos interrelacionados:

1. Un proceso subyacente, modelado por una cadena de Markov no observable con un número finito de estados, una matriz de transición de estados y una distribución de probabilidad de estados inicial.
2. Un proceso observable dado por los vectores de observación y relacionado por un conjunto de densidades de probabilidad asociadas a cada estado.

La dificultad en este tipo de algoritmos se centra en el entrenamiento de las cadenas. Además es costoso desde el punto de vista computacional. Para aplicar HMM al reconocimiento de caras, lo que se hace es dividir la imagen en bloques y se aplica la DCT. Luego ciertos coeficientes de la transformada del bloque se utilizan como los vectores de observación.

- **Local Binary Patterns (LBP):** constituye un operador que fue diseñado originalmente para la descripción de textura, la idea de usar el LBP para la descripción de rostros está motivada por el hecho de que los rostros pueden ser vistos como una composición de micro-patrones los cuales son bien descritos por este operador. Asigna una etiqueta a cada píxel de una imagen usando en una vecindad de 3x3 píxeles el valor del píxel central como umbral y considerando el resultado como un número binario, luego el histograma de las etiquetas es utilizado como descriptor de la textura. Este método muestra también ser bastante invariante a cambios de expresión facial y de iluminación. Otra ventaja es la eficiencia computacional del operador LBP y no se necesita ninguna normalización previa antes de aplicar el operador a la imagen del rostro (16).

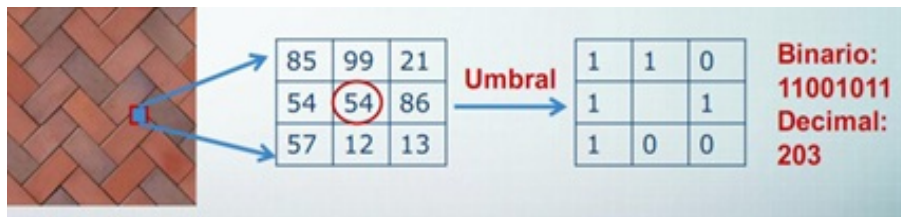


Figura 5. El operador LPB básico (16).

En cualquiera de sus variantes, los histogramas LBP contienen información acerca de la distribución de los micro-patronos locales como bordes, manchas y áreas lisas. Para una representación eficiente del rostro se debe retener también información espacial. Para este propósito la imagen se divide en regiones y los descriptores son extraídos de cada una de estas regiones independientemente, los descriptores son entonces concatenados para formar el descriptor global de la imagen obteniendo un histograma mejorado espacialmente. En el histograma mejorado espacialmente se tiene efectivamente la descripción del rostro en tres niveles diferentes de localización: las etiquetas LBP para los histogramas contienen información acerca de los patrones a niveles de píxeles, las etiquetas se suman en una región pequeña para producir información a un nivel regional y luego los histogramas son concatenados para construir una información global del rostro.

Desde el punto de vista de la clasificación, un problema usual es tener muchas clases y solo muy pocas, probablemente sólo una muestra de ejemplo por clase. Muchas medidas de similitud han sido propuestas para tratar este problema con histogramas: intersección de histogramas, estadística de probabilidad logarítmica, estadística Chi cuadrado (37).

Cuando la imagen es dividida en regiones, se espera que algunas regiones contengan más información útil que otras, en términos de distinguir un individuo de otros, por ejemplo, los ojos y la boca. Para tomar ventaja de esto, un peso puede ser asignado para cada región basándose en la importancia de la información que contiene. La tasa de reconocimiento utilizando este operador es bastante alta, un 93% si no se utilizan pesos y 97% si se utilizan, mostrando ser superior a otros métodos (38).

Durante el estudio de los métodos de reconocimiento de rostro, se ha verificado que los métodos basados en características locales son los más extendidos y que con más frecuencia se utilizan en sistemas de reconocimiento. Estos métodos vienen muy unidos a los descriptores y a las técnicas que se utilizan en la propuesta final, es frecuente además que se utilicen la conjugación de varios de ellos.

### 2.1.6. Creación del vector de características

La creación del vector de características es lo que se obtiene como resultado de aplicar un descriptor combinado intrínsecamente con un método de extracción de características, básicamente es lo que generan como salida de ese proceso, que se empleará para la realización de la confrontación, a partir de que este elemento debe ser único para cada rostro que se analice. Un ejemplo de ello es el reconocimiento de rostros con el método EBGM. Este método, logra conformar un vector a partir de los puntos característicos obtenidos en la imagen del rostro, pero cada punto característico (nodo del grafo) va acompañado de la información de apariencia local en su vecindad representada por los jets de Gabor (39). Muchos autores referencian a este método entre los de mejores resultados y que ha alcanzado popularidad en los últimos tiempos.

### 2.1.7. Confrontación y resultados

La última etapa del proceso es la que se determina como confrontación y muestra de los resultados del reconocimiento del rostro. La misma se desarrolla teniendo en cuenta la finalidad del modelo o aplicación que se está implementando, dado que ello determinará los métodos o formas que se utilicen para alcanzar los resultados.

Para la realización de este proceso, se parte de la existencia de información ya almacenada sobre los rostros a analizar, como son las características o descriptores de cada persona, datos asociados como: nombre, edad, sexo u otros metadatos necesarios para la comparación, en su mayoría extraídos en etapas anteriores. Es importante determinar el modo en que se va a realizar la confrontación de los datos, porque existen dos formas de determinar los resultados, definidos por: modo verificación o modo identificación. La determinación del modo en que se esté efectuando esta etapa agiliza y hace más efectivo el proceso, en el primer caso solo se verifican o comparan características entre dos entes, el entrado por el usuario y su supuesta identidad almacenada en el sistema, dando como devolución una variable *booleana*, positivo o negativo. El modo identificación realiza una confrontación más general, entre el rostro introducido por el usuario y todos o una parte de los almacenados por el sistema, siendo más ágil y eficaz el resultado según la combinación efectiva de los métodos de comparación que se utilicen, teniéndose en cuenta la cantidad de rostros con que se realice la confrontación, que pueden seccionarse por grupos, según algunos datos que se tengan como precondition, como pudiera ser el sexo o grupo generacional de la persona.

Los métodos de comparación que se utilizan tienen un mismo objetivo y están determinados por comprobar el grado de similitud entre las características de los rostros, que ya en esta etapa están

contenidos en un vector de características según el método de reconocimiento que se haya utilizado. En algunos trabajos se pone en práctica como métodos de comparación la clasificación por el vecino más cercano, específicamente las reglas 1-NN y 3-NN. Las distancias consideradas varían según los descriptores utilizados.

Otra de las técnicas utilizadas son las que buscan definir la distancia entre los vectores de características, uno de ellos es la distancia del coseno. Pero sin duda los algoritmos más generalizados son los algoritmos de similitud, que pueden clasificarse en los de Similitud Promedio, Similitud por Magnitud o Similitud Ponderada, se puede utilizar la combinación de los tres.

### **2.2. Procedimiento propuesto**

La propuesta final a partir del estudio de las diferentes técnicas y métodos de reconocimiento de rostro, se basa fundamentalmente en elementos avalados internacionalmente y extendidos en diferentes sistemas que tienen una finalidad cercana a la que se busca con el que se presenta en este trabajo.

De forma general se puede plantear que la utilización de las técnicas más avanzadas en cada una de las etapas da como combinación un sistema que define lo mejor de cada una de ellas con sus particularidades, que será en el caso que se presenta la solución al problema planteado. Es por ello que se define para emplear en la detección de rostros la utilización de varios métodos basados en rasgos faciales que en conjunto definen elementos distintivos de una persona, que lo hacen inequívoco frente a otros objetos.

En la etapa de extracción de características se expone a los métodos basados en la apariencia, destacando como los de mejores resultados, los que se enfocan en varios puntos de la imagen, así como la utilización de descriptores de información general, sin embargo cuando se quiera lograr el resumen semántico de materiales audiovisuales deberán utilizarse los descriptores de dominio específico, factibles para esos casos.

Finalmente es importante resaltar en este punto que la confrontación deberá realizarse de forma rápida y efectiva, la utilización de un método de similitud enfocado en la efectividad y agilidad del proceso garantizan que el procedimiento propuesto sea conveniente para su implementación en las aplicaciones del departamento Señales Digitales.

La creación del procedimiento comprende las tres etapas fundamentales definidas para este proceso, se muestran a continuación las actividades que se proponen para la implementación final del mismo, explicando en cada caso la utilización de los algoritmos a desarrollar. Cada actividad será la

consecución de la anterior, se trata de una concatenación donde la salida de una constituye la entrada de la siguiente.

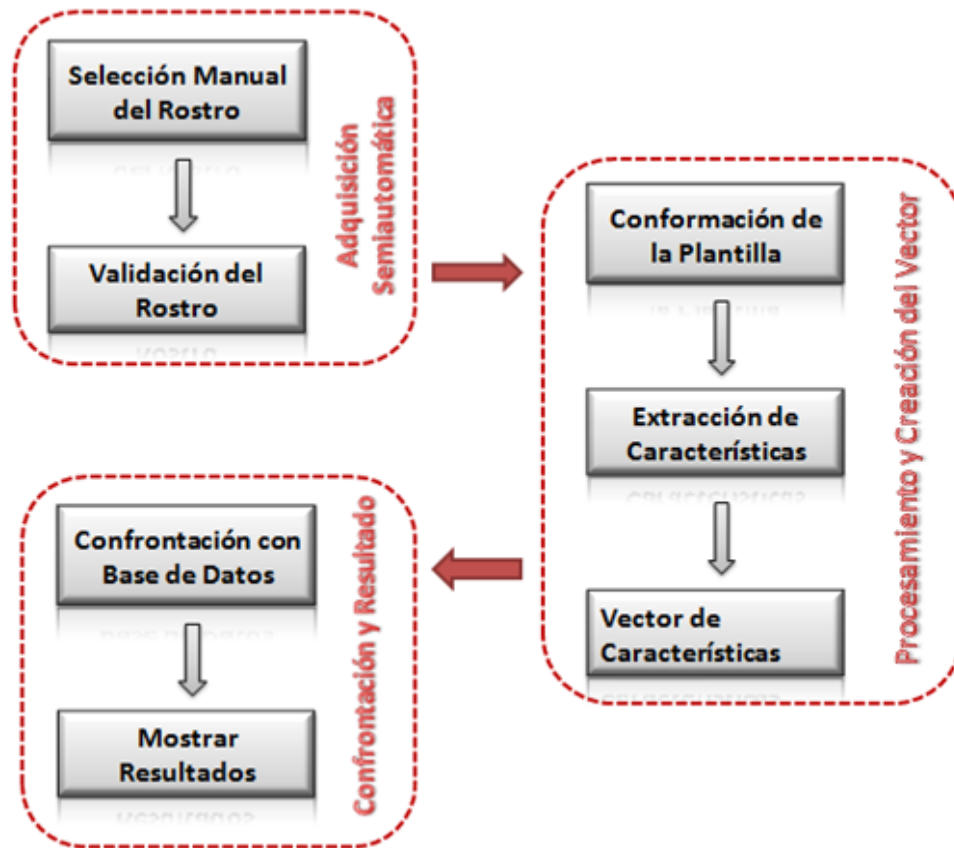


Figura 6. Procedimiento propuesto para el reconocimiento de rostro

### 2.2.1. Selección manual del rostro

La adquisición de la imagen, como se ha mostrado al inicio del capítulo, puede realizarse de dos formas, siendo la semiautomática la que se propone debido a que es la que se ajusta al proceso de catalogación pues el documentalista seleccionará la imagen que será procesada mediante una herramienta de selección manual, lo que además genera un menor índice de error en su selección explícita, luego se dará paso a la actividad de validación de rostro.



2.2.2. Validación del rostro, métodos de detección

		Ventajas	%	Desventajas
Técnicas basadas en rasgos	Análisis de bajo nivel en el color	Minimizan los efectos que tienen las variaciones de iluminación o pose.	90%	Dificultad al localizar ciertos rasgos faciales por oclusión.
	Red Neuronal	Alto porcentaje de aciertos que produce, superior al 95%.	95%	Se caracteriza por ser lento.
Técnicas basadas en la imagen	Métodos Estadísticos: Viola y Jones	Procesa imágenes extremadamente rápido. Eficaz y rápido en término de detección.	99%	Genera falsos positivos, al reconocer objetos en la imagen que no son rostros.

Figura 7. Resultados reportados con los métodos más representativos (14).

Según los autores Li y Jain (14), en sus comparaciones de los diversos métodos de detección del rostro, el método de *Viola and Jones* basado en Adaboost ha sido el más eficaz de todos los trabajos desarrollados. Es por ello que para la realización de la validación se ha definido como método de detección de rostro, los implementados por *Viola and Jones*, que gozan de un alto prestigio en el campo que se estudia y se encuentran contenidos de forma nativa en la biblioteca OpenCV desarrollada por Intel.

Este algoritmo fue propuesto en el año 2001 por Paul Viola y Michael Jones, basado en la detección de objetos en tiempo real. A pesar de que puede ser entrenado para detectar una variedad de clases de objetos, fue motivado principalmente por el problema de la detección de rostros. Como método de detección logra de forma efectiva resolver el problema de validación de rostro presente en esta etapa y su comportamiento es propio de métodos basados en rasgos faciales. Es válido resaltar que goza de gran prestigio internacional y es utilizado en muchos de los sistemas que se desarrollan sobre esta temática.

Para la realización de este proceso se tendrá en cuenta tres elementos fundamentales:

### Imagen Integral

La estructura de Imagen Integral introducida por *Viola and Jones* permite un cálculo rápido de las diferentes características de una imagen. Para esto se suma el valor en escala de grises (o intensidad) de cada píxel desde la esquina superior derecha hasta el punto objetivo (40).

### Adaboost en la selección de clasificadores

Los clasificadores utilizados en el método de detección de rostro propuesto por *Viola and Jones* son las características. En una cuadrícula de 24x24 se pueden encontrar aproximadamente 160000 características candidatas (41). Adaboost permite la selección de una colección de clasificadores débiles. La combinación lineal de estos clasificadores débiles dará como resultado el clasificador fuerte. Los pesos de cada clasificador serán inversamente proporcionales a sus errores en el entrenamiento.

### Cascada de clasificadores fuertes

El método de *Viola and Jones*, crea una estructura en cascada con varios clasificadores fuertes de forma que se rechacen rápidamente las zonas de *background* (todo lo que es diferente al rostro), y se aumente la precisión de las regiones donde hay rostros, además de que el tiempo de ejecución es mucho menor. Para este acercamiento se partió de la idea, de que es más fácil para el sistema inteligente rechazar una región como *background*, que confirmar que realmente es un rostro (42). Además las regiones en una imagen que constituyen *background* tienden a ser mucho mayores que las regiones que contienen caras. En esto es en lo que radica la eficiencia del método al descartar rápidamente la mayor cantidad de regiones de una imagen, los clasificadores concentrarán su capacidad de cómputo en aquellas regiones más sospechosas de ser rostros.

La utilización de estos tres pasos de forma combinada ante la presencia de cada análisis resultan en un alto grado de efectividad y bajos niveles de falsos positivos (rostros que marcó como tal y que no lo son) y falsos negativos (rostros que no reconoció y que están presente). Es importante resaltar que en el presente trabajo, se utilizará el algoritmo **cvHaarDetectObjects** presente en la biblioteca OpenCV, así también una cascada de clasificadores ya entrenada y disponible en la misma biblioteca de código abierto.

### 2.2.3. Conformación de la plantilla

Una vez obtenido un rostro acorde con las especificaciones del sistema y del documentalista, se pasa a la segunda etapa del proceso, enfocada en varias actividades importantes, comenzando con la normalización y creación de la plantilla a partir del rostro entrado como parámetro. Atendiendo a los métodos y descriptores que se va a utilizar se han definido para el presente trabajo un grupo de acciones de normalización que son convenientes llevar a cabo antes de pasar a la etapa posterior:

1. Se implementará una transformación geométrica que ubica las coordenadas de los ojos en un valor predeterminado y se realiza un escalado de la imagen para llevarla a un tamaño dado (128 x 128 píxeles), logrando con esta proporción un rango en el que no se pierde información necesaria para la extracción de características y reduce el tiempo de procesamiento.
2. Se creará con la nueva dimensión de la imagen, una máscara en escala de grises del rostro para resaltar en una dimensión más estrecha de colores los diferentes elementos resaltantes de la cara, simplifica el análisis del descriptor para la definición de las diferentes características.
3. Los bordes de la imagen serán suavizados en un margen de 10 píxeles. Esto se aplica para evitar posibles efectos de borde introducidos durante la transformación geométrica.

### 2.2.4. Extracción de características

Luego de haber conformado la plantilla, se pasa a la extracción de las características, para lo cual se presenta a continuación una tabla con los resultados que reporta la literatura con respecto a la utilización de los métodos de reconocimiento de rostro basados en apariencia local, siendo estos los más representativos para el presente trabajo, se muestra una comparación realizada en base a los mismos.

Método	Base de Datos de Rostros	Cantidad de Imágenes Probadas	Taza de Reconocimiento	Taza estándar de comparación	Principales variaciones en las imágenes
Método local probabilístico de sub-espacios	AR	600	82.3 %	70.2 %	expresión, tiempo
		400	71.0%	33.0%	oclusiones
Método local utilizando mapas auto-organizados	AR	600	93.7%	70.2 %	expresión, tiempo
		400	76.0%	33.0%	oclusiones
Modelos ocultos de Markov	AR	1440	89.8%	67.2%	expresión, tiempo, iluminación
Jets de Gabor	FERET	1196	95.0%	79.7%	expresión
Patrones binarios locales	FERET	1196	97.0%	67.2%	expresión, tiempo, iluminación

**Figura 8. Resultados experimentales reportados con los métodos (16).**

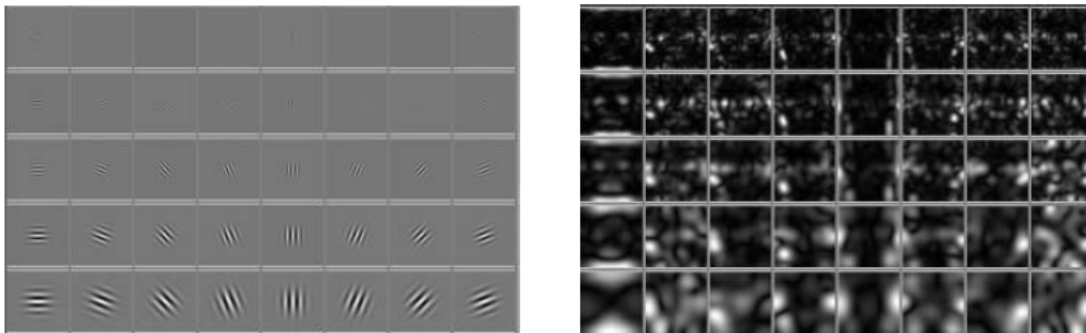
La tabla muestra que cada método tiene sus fortalezas y debilidades en el reconocimiento de rostro (16), sin embargo LBP es el que mejor tasa de reconocimiento muestra, pero se ha considerado que tanto él como Jets de Gabor resuelven la problemática presentada ya que los problemas de iluminación se atenúan debido a que la imagen proporcionada por el documentalista deberá poseer el mínimo de factores que atenten con el proceso de reconocimiento. Además es referenciado el caso de Gabor por el proyecto Aguará y por los trabajos de la Universidad Estatal de Colorado, como uno de los mejores métodos a emplear en esta etapa por su utilización ante la presencia de solo una imagen por rostro. Es válido señalar que se reporta como una limitante en cuanto al uso de Wavelets de Gabor el costo computacional del algoritmo, pero si se tiene en cuenta que los vectores de las imágenes contra las que se comparará la imagen de entrada son previamente calculados y almacenados, este factor no influirá demasiado en la celeridad del proceso de reconocimiento, pues solo se requerirá extraer el vector de características de la imagen de entrada para compararlo con los ya almacenados.

Habiéndose estudiado los métodos de reconocimiento presentes en la literatura y algunos de los descriptores más frecuentes que se utilizan en ellos, se definirá con más detalles la propuesta final para esta etapa, que está avalada por sus altos porcentajes de efectividad y por numerosos autores en la bibliografía. La misma se presenta a partir de la utilización del descriptor de Wavelets de Gabor y el método EBGM, denominándose esta técnica por algunos autores como Jets de Gabor.

### Wavelets de Gabor

Wavelets de Gabor es uno de los descriptores que actúa de forma local, muy popular actualmente dada su particularidad de poseer un alto potencial discriminador en el reconocimiento de caras. Con él se ve la posibilidad de analizar el espectro de una imagen de la misma forma que lo realiza el análisis tradicional de Fourier, permitiendo conocer su contenido espectral. Tiene como ventaja que permite el estudio localizado del espectro, centrando la atención en una región particular.

Los Wavelets de Gabor se basan en una familia de filtros que permiten ser sintonizados en determinada orientación, en determinada frecuencia y además pueden limitar su rango de acción a determinada región de la imagen (38).



**Figura 9. Filtros de Wavelets de Gabor aplicados en un punto de la imagen (39)**

Los Wavelets de Gabor están formados por un banco de filtros lineales e invariante temporales, por lo que se los caracteriza por su respuesta al impulso. Los elementos descriptores son aquí los denominados jets, resultado de aplicar los Filtros de Gabor a la imagen normalizada y evaluar el resultado en el punto principal deseado.

La configuración estándar de los Jets de Gabor está basada en 40 wavelets (8 orientaciones, 5 frecuencias) complejas, donde cada una de ellas tiene una componente real y una imaginaria. Un jet para la posición  $(x,y)$  de una imagen, se produce macheando ese punto con cada uno de los wavelets obtenidos mediante la ecuación (5):

$$W(x, y, \theta, f, \phi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}} \cos(2\pi f x' + \phi)$$

Donde:

$$x' = x \cos\theta + y \sin\theta$$

$$y' = -x \sin\theta + y \cos\theta$$

$\theta$  especifica la orientación de la wavelet (8 orientaciones:  $0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8, 5\pi/8, 6\pi/8, 7\pi/8$ )

$\lambda$  especifica la frecuencia (5 frecuencias:  $4, 4^2, 8, 8^2, 16$ )

$\phi$  especifica la fase (2 fases:  $0, \pi/2$ )

$\sigma$  especifica el radio de la gaussiana ( $\sigma=\lambda$ )

$\gamma$  especifica la razón del aspecto de la gaussiana ( $\lambda=1$ )

La ecuación descrita se le denomina también Banco de Filtros Gabor Base, para un mejor entendimiento se desglosan los parámetros de esta clase de filtros.

**Frecuencia:** Es el inverso de la distancia en píxeles entre máximos locales contiguos. En la ecuación es representado por  $f$ , frecuencia de la senoide. También se puede caracterizar por su período  $\lambda$ , siendo este  $\lambda = 1/f$ .

**Selectividad espacial:** Representa qué tan local es el filtro en el dominio espacial. Es representado por el radio de la gaussiana  $\sigma$ . Para valores mayores de  $\sigma$  menor es la selectividad espacial. Con el objetivo de que todos los wavelets sean escalas y/o rotaciones de un par de filtros madre, se deberá cumplir que  $\sigma = k\lambda$ .

**Selectividad frecuencial:** Representa qué tan local es el filtro en el dominio de la frecuencia. En la ecuación es representado indirectamente por  $1/\sigma$ . Para valores mayores de sigma mayor es la selectividad frecuencial.

**Relación de aspecto:** La relación de aspecto de la gaussiana es representada por  $\gamma$ .

**Orientación:** Es el ángulo que forma el centro del filtro en el dominio frecuencial con el eje horizontal. En la ecuación es representado por  $\theta$ .

**Fase:** Es la fase de la senoide que modula la gaussiana. Siendo representada por  $\phi$ . Por lo general se utiliza dos fases que distan  $\pi/2$  entre sí. La dupla resultante de aplicar esta pareja de filtros puede ser vista como un número complejo (5).

Al aplicar a cada punto de la selección los distintos filtros de Gabor se obtienen los denominados jets los cuales van a contener las coordenadas polares almacenadas en un arreglo. Entonces la información de cada punto analizado será la combinación de fase, orientación y frecuencia, por tanto,

en la práctica se obtiene una descripción a través de múltiples frecuencias y orientaciones, de una región de la imagen.

El método que implementa este resultado recogido en esta etapa es EBGM, que será utilizado a partir de algoritmos implementados con la ayuda de OpenCV.

El algoritmo EBGM es una herramienta de reconocimiento de objetos, aplicada a caras en 1997 por Wiskott (39), que implementa tres etapas básicas:

- Localización de puntos principales en el objeto de interés.
- Extracción de características en dichos puntos.
- Confrontación de estructuras de características para la validación del objeto.

Se utiliza aquí este algoritmo como herramienta de localización de puntos característicos para el reconocimiento de caras.

### **2.2.5. Confrontación de los resultados**

Con el vector de características generado se da paso a la tercera y última etapa que es donde se confrontan los resultados. Para ello se estudiaron tres métodos de similitud: distancia del coseno, comparación de histograma y el método euclidiano.

De las pruebas realizadas con los tres métodos el que mejor resultados ofreció para el presente trabajo fue el método euclidiano, pues los datos estadísticos que se generaron o más específicamente la cercanía a cero en caso de muy parecido o lejanía en caso que la imagen fuera muy diferente, fueron mejores que los generados por los otros métodos de similitud analizados y probados.

Este método va a realizar la comparación del vector del rostro entrado inicialmente con los vectores almacenados con anterioridad. Este proceso a partir de las distancias más cercanas entre estos datos estadísticos, va a dar como resultado una muestra de los parecidos al rostro que el documentalista preseleccionó inicialmente, con la intención de corroborar o conocer su ficha, siendo así una herramienta de consulta muy útil para la realización de la catalogación de personajes.

### **2.3. Aplicación del procedimiento**

El sistema de catalogación del departamento de Señales Digitales es una aplicación web desde la cual el documentalista mientras se reproduce el video deberá seleccionar la imagen que contiene el rostro que desea identificar, esta imagen será enviada al núcleo de procesamiento del sistema que está

implementado con tecnologías de escritorio, específicamente C++ con Qt, donde se validará que existe rostro, en caso positivo se normalizará y serán extraídas sus características, las que se compararán con el resto de las imágenes que se tienen almacenadas para determinar las imágenes candidatas que se enviarán a la interfaz web para ser mostradas al documentalista de más parecido a menos parecido. Teniendo esta herramienta de apoyo el documentalista podrá definir con más claridad el etiquetado de personajes que aparece en el material. El núcleo de procesamiento del sistema está basado en una arquitectura de *plugin*<sup>1</sup>, en la cual el reconocimiento de rostro será una nueva funcionalidad a agregar en un *plugin*. Las funcionalidades del componente de reconocimiento de rostros se pueden especificar de la siguiente manera:

- Permitir seleccionar la imagen que será procesada: mientras se reproduce un material el usuario podrá seleccionar la imagen en la que aparece la personalidad que desea reconocer. En caso de que conozca su identidad podrá introducir directamente sus datos en la ficha de catalogación, la que deberá poseer auto completamiento en caso de que el sujeto esté presente en la base de datos, en caso negativo se le deberá sugerir al usuario la inserción del nuevo personaje en la base de datos.
- Detectar el rostro en la imagen: de la imagen seleccionada por el usuario se deberá detectar la región de interés para su representación y procesamiento, que en este caso constituye el rostro de la persona, esta información será enviada para su comparación.
- Comparar el rostro con la información contenida en la base de datos: la información del rostro detectado deberá ser comparada con la información presente en la base de datos para detectar las imágenes más similares y que son candidatas a ser el rostro especificado por el usuario.
- Enviar las posibles respuestas de acuerdo a la similitud del rostro: el sistema deberá enviar el conjunto de posibles identidades con el nombre y la imagen de muestra en la base de datos, la respuesta deberá estar ordenada del más al menos similar.
- Insertar un nuevo sujeto en la base de datos de personas: en caso de que no coincida el rostro indicado por el usuario con ninguno de los de la base de datos o el usuario indique que las respuestas son erróneas se deberá permitir insertar al nuevo personaje en la base de datos con la imagen de muestra indicada por el usuario, para insertar al personaje deberá introducir además otros datos de identidad como el nombre y apellidos, nacionalidad y papel social, cargo político u otro dato que explique su relevancia.
- Permitir seleccionar la respuesta adecuada: el usuario podrá determinar cuál de las respuestas es la correcta e indicarlo.

---

<sup>1</sup> Es un complemento que se relaciona a una aplicación para aportarle nueva funcionalidad.



- Enviar los datos de identidad de la persona seleccionada por el usuario, los que serán almacenados en la ficha de catalogación del material audiovisual: una vez que se haya indicado la respuesta por el usuario los datos de identidad del personaje almacenados en la base de datos deberán almacenarse en la ficha de catalogación del material (43).

### 2.4. Conclusiones parciales

Después de realizado el estudio de las diferentes etapas que comprende el reconocimiento de rostro y de haberse evaluado los diversos métodos y técnicas que existen en cada una de ellas, se puede concluir que: los métodos aplicables al reconocimiento de rostro son los basados en características geométricas y los basados en la apariencia, estos últimos se pueden clasificar según la forma de analizar la información, pueden ser locales y globales, los locales han demostrado ser más efectivos en entornos con una sola imagen de muestra como es el caso de Wavelets de Gabor. Se muestra como un algoritmo que resuelve de forma efectiva la validación de los rostros entrados a *Viola and Jones* que además se facilita su implementación a partir de la utilización de la biblioteca OpenCV. Finalmente se expone la creación de métodos basados en la apariencia local capaces de lidiar con los problemas de iluminación, cantidad de imágenes de muestra, variación de la postura; Elastic Bunch Graph Matching se presenta como uno de los más efectivos y que se ajusta a la presente investigación ya que está ligado a la utilización del descriptor propuesto Wavelets de Gabor.

## **CAPÍTULO 3: Análisis de los resultados**

El presente capítulo presenta los resultados de la implementación del procedimiento en un entorno de prueba, para corroborar que el mismo muestra resultados utilizables en la catalogación de audiovisuales. Se detalla su implementación, así como se describen las principales pruebas aplicadas a cada etapa, con el objetivo de comprobar el funcionamiento individual de cada técnica utilizada y finalmente realizar la valoración de la efectividad del procedimiento de forma general a partir de los resultados obtenidos.

### **3.1. Análisis de los resultados del procedimiento**

Para la realización del prototipo de prueba se han tenido en cuenta varios elementos: lo primero es que se trató de que el proyecto generado implementara los elementos de programación más cercanos a la línea de desarrollo del departamento donde se enmarca el problema. Se utilizó una base de datos de rostros representativos, que refleje adecuadamente un proceso real en una entidad que trabaje con audiovisuales. Por último tema estuvo que se debe cumplir a cabalidad cada una de las etapas definidas en los capítulos anteriores e implementarse usando los algoritmos propuestos para cada una de ellas. Es importante destacar que en investigaciones de este corte, las validaciones de las propuestas se hacen en un entorno que resulte cómodo para el investigador, un ejemplo de ello es MatLab o en el presente caso Visual Studio, debido a que provee numerosas bibliotecas que facilitan el trabajo con las matrices de datos y con los métodos matemáticos a comprobar.

Teniendo en cuenta los elementos planteados anteriormente se estableció lo siguiente: el lenguaje y las bibliotecas a utilizar, debían estar acorde con la política del departamento. El prototipo de prueba se implementó en el lenguaje C++ debido a que es muy utilizado internacionalmente en este tipo de aplicaciones y presenta una alta integración con la biblioteca OpenCV, de código abierto, libre del pago de licencias y muy representativo en el reconocimiento de patrones.

Para dar solución a la segunda premisa, se investigó sobre cual base de datos de rostros es la más adecuada para probar el procedimiento. Se valoraron dos casos: el primero es la base de datos de rostros FERET creada entre 1993 y 1997, su misión principal es la de servir de pruebas para proyectos de tecnologías de reconocimiento automático de caras para aplicaciones diversas de inteligencia y seguridad. Consta de un conjunto de 14051 imágenes en escala de gris de 8 bits de caras humanas con distintas poses, desde frontales a perfiles derechos e izquierdos, así como también distinta iluminación, expresión y tomas de diferentes momentos de la vida de la persona (44).

Es actualmente una de las base de datos de rostros más utilizadas a nivel internacional, en aplicaciones de esta rama. Para el presente trabajo presenta una serie de inconvenientes, entre ellos que las imágenes están en una escala de 8 bit, escala de grises, con una definición marcada de su pose e iluminación, un poco alejada de los intereses del trabajo debido a que las imágenes reales, son recortes de un material audiovisual con elementos comunes en su pose e incidencia de la iluminación. Otro aspecto importante es que contienen rostros de personas poco relevantes, es decir, personas no populares o conocidas elemento a tener en consideración debido a que el principal interés durante la catalogación es el etiquetado de personalidades relevantes. Como último elemento está que para acceder a FERET se necesitan una serie de requisitos descritos en una carta de permiso elaborada por la entidad que los necesita, donde se deja claro que su uso es académico, de lo contrario se deberá pagar la licencia y como es conocido el objetivo principal de la UCI es insertarse en el mercado de software por lo que el objetivo de la inclusión de este valor agregado en el producto de catalogación del Departamento Señales Digitales es brindarlo a clientes que lo requieran.

La segunda base de rostros analizada fue **Labeled Faces in the Wild Home** una base de datos de fotografías faciales diseñada para estudiar el problema de reconocimiento de rostros sin restricciones. El conjunto de datos contiene más de 13000 imágenes de rostros de personalidades importantes en diferentes ámbitos de la vida recogidos de la web, en total 1680 personas representadas tienen dos o más fotos distintas en el conjunto de datos. Representa una base de rostros con gran reconocimiento internacional, utilizada en numerosas investigaciones, se puede descargar gratuitamente, por lo que se ha empleado para la realización de las pruebas.

### 3.1.1. Creación del proyecto

Como se expuso en el capítulo anterior el proyecto a desarrollar está compuesto por tres etapas fundamentales, que se han tenido en cuenta para la creación de proyecto para validar el procedimiento.

### 3.1.2. Primera etapa: detección de rostro

Para la creación de la aplicación se tiene como primer paso la adquisición de la imagen, en la cual se realiza una selección manual del rostro dentro de la misma. Posteriormente se pasa a comprobar la existencia o no de la cara, entrando en este punto *Viola and Jones* como recurso para realizar la detección, devolviendo un valor booleano que determina la veracidad de la selección.

```
bool detect_and_draw(IplImage* img, CvHaarClassifierCascade* cascade , ...);
```

### Figura 10. Método de detección

El método **detect\_and\_draw ()** es el encargado de validar la selección manual realizada por el usuario donde se espera la existencia de un rostro, necesita como parámetros la imagen seleccionada y la cascada de clasificadores entre otros que se utilizarán en la realización del proceso. Este método devolverá un valor *booleano* con la certificación de la detección.

La cascada de clasificadores de Haar es una estructura de entrenamiento que se crea con anterioridad donde se predefinen los descriptores más importantes que va a tener en cuenta el método para la definición o no de un rostro dentro de la imagen. En este caso se hace uso de un clasificador genérico presente en la biblioteca de OpenCV, denominado **haarcascade\_frontalface\_alt**, el que ha sido entrenado con un alto número de imágenes de rostros.

### 3.1.3. Segunda etapa: extracción de características

En la segunda etapa del procedimiento se tiene en cuenta la selección del rostro ya validado al cual se le realizará la extracción de características, pero antes se deberá normalizar hasta llevar la imagen a un estándar más adecuado para su procesamiento, que optimizará los resultados finales. Para ello se hace uso de la función **cvResize()** para escalar a 128px por 128px y **cvConvertScale()** para llevar a escalas de grises; también se aplica una reducción de los bordes introducidos.

Con la imagen normalizada se puede proceder a aplicar el método EBGM que utiliza Wavelets de Gabor para la extracción de características, de esta manera se logrará obtener el vector del rostro. El primer paso es generar el banco de 40 filtros, necesario para convolucionarlos sobre determinados puntos del rostro. Para ello se utilizará la función **gabor\_filter ()** presente en la biblioteca **gaborExtract** creada específicamente para la extracción de características a partir de técnicas, está liberada bajo licencia BSD.

```
double* extract_features(const IplImage* img,...)
```

### Figura 11. Método de extracción de características

El método **extract\_features ()** presente en la biblioteca antes mencionada es el encargado de devolver el vector de características, a partir de la generación de varias matrices de comparación y la utilización de funciones presentes en la biblioteca OpenCV, que se encargan de devolver las componentes del

vector de características, el cual modela la identidad propia del rostro analizado y debe ser único para cada uno. Este vector es el que se comparará con el resto de los vectores almacenados para así determinar la similitud entre los mismos.

### 3.1.4. Tercera etapa: confrontación

Luego de la creación del vector de características se pasa a la etapa de confrontación, es en este momento donde se realiza la comparación entre vectores y se identifican los más parecidos. Para la realización de esta etapa se utilizó el método euclidiano.

```
Double euclid (int n, double* data1, double* data2)
```

#### Figura 12. Método de distancia euclidiana

El método euclidiano presenta como parámetros  $n$  que es el tamaño del vector, **data1** que es el vector base entrado por el usuario y **data2** que es el vector almacenado contra el que se está analizando la similitud. El parámetro **data2** puede poseer en otros casos en lugar de un único valor, la referencia a un conjunto de valores que corresponderá a los almacenados en base de datos. El método retorna un valor positivo decimal en los casos donde cero es el más parecido, significa que los vectores comparados corresponden a rostros iguales y mientras más se aleja de él menos será su semejanza. La etapa y el proyecto culminan con la muestra de los rostros ordenados por similitud según sea el caso, mostrándose en orden del más parecido al menos parecido.

### 3.2. Validación del procedimiento

La validación del procedimiento parte de validar cada una de las etapas fundamentales por separado y finalmente hacer un análisis de los resultados de forma general. Un proceso de pruebas que es la forma factible para realizar la validación de este proceso se hace de diferentes maneras, en el caso del reconocimiento de rostros usualmente se realiza analizando los falsos positivos y falsos negativos. Un falso positivo consiste en declarar que una región de una imagen es un rostro cuando no lo es, mientras que un falso negativo es no detectar una región de la imagen en la que sí existe un rostro. A continuación se realiza un análisis de las pruebas realizadas en cada una de las partes del procedimiento.

### 3.2.1. Detección de rostro

Para validar la detección de rostro se han obtenido los resultados necesarios para calcular la tasa de falsos positivos (FMR por sus siglas en inglés) y la tasa de falsos negativos (FNMR por sus siglas en inglés). El proceso de prueba en detección de rostro que se plantea, consiste en ingresar en el sistema imágenes de rostros de varios individuos, las cuales serán procesadas por el algoritmo de detección *Viola and Jones* de una en una y comprobar visualmente si fue correcto o no el resultado. Durante todo este proceso los parámetros que se tomaron fueron los siguientes:

- Número de falsos positivos en el proceso de detección.
- Número de falsos negativos en el proceso de detección.
- Número de intentos o pruebas realizadas.

A continuación se muestra una tabla con los datos obtenidos de falsos positivos y negativos durante el proceso de detección de rostros en la etapa de entrenamiento.

Id_Imagen	Total de intentos	Falsos Positivos	Falsos Negativos
1	10	0	0
2	11	1	0
3	11	0	0
4	10	0	0
5	12	0	0
6	13	0	0
7	12	0	0
8	10	0	0
9	10	0	0
10	10	1	1
11	11	0	0
12	10	0	0
13	10	0	0
14	13	0	0
15	10	0	1
16	10	0	0
17	11	0	1
18	10	0	0
19	12	0	0
20	10	0	2
<b>Total</b>	270	2	4

Tabla 1. Resultados del proceso de detección.

### Análisis de los resultados

Obtenidos los resultados del proceso de las pruebas, se analizan los mismos con el fin de comprobar la efectividad de la etapa de detección. Los resultados obtenidos con el método de detección *Viola and Jones* muestran un FMR de 0.74 % y un FNMR de 1.48 %, un rendimiento total del 97,78%, observándose que a pesar de sus errores, el método se presenta robusto para su utilización.

#### 3.2.2. Extracción de características y confrontación

Para validar los resultados de la etapa de extracción y confrontación se tuvieron en cuenta una serie de elementos por ejemplo: se puede aplicar directamente una técnica para hallar la tasa de falsos positivos y la tasa de falsos negativos debido a que los juegos de datos son diferentes y van enfocados a buscar acercamiento en el parecido de los rostros, lo que requiere la comprobación visual del resultado de comparación, es decir, no se considera un desacierto del procedimiento que la imagen más parecida no sea la primera que proponga el sistema, debido a que los resultados serán analizados por un documentalista que seleccionará la que considere acertada dentro de las respuestas del sistema. La bibliografía estudiada refiere la realización de pruebas y aplicación de técnicas del tipo FMR y FNMR o de *precision* y *recall* en métodos de reconocimiento de rostros aplicados en sistemas de seguridad, donde el objetivo no es el mismo que se persigue en esta investigación sino detectar si un rostro pertenece a una persona.

Para el presente trabajo se realizaron pruebas experimentales donde una imagen de rostro entrada por el usuario se deberá comparar haciendo uso del procedimiento contra cinco imágenes que se tienen almacenadas, existirá en cada caso al menos una de la misma persona pero no obligatoriamente es la imagen idéntica insertada por el usuario. Para el mejor entendimiento de las pruebas, se exponen los resultados de una muestra de cinco de ellas en una tabla donde las tres últimas columnas son los datos que reflejan la efectividad del método planteado. La columna que se nombra **Valor Euclidiano** recoge el dato que retorna el método de similitud entre los vectores comparados, los más cercanos a 0 serán los más semejantes. La columna **Ord. Visual** refleja cuales son los rostros más parecidos por apreciación humana, donde 1 es el más similar. La columna **Ord. Sistem** es el ordenamiento que realiza el sistema al aplicar el procedimiento, lo que se compara con la columna anterior para definir el nivel de efectividad de forma cualitativa.

Juego de Datos					
No.	Imagen de entrada	A comparar con	Valor Euclidiano	Ord. Visual	Ord. Sistem
1	Imagen 1	1 (diferente)	0,005961	2	2

		persona)			
		2 (diferente persona)	0,007511	5	<b>4</b>
		3 (diferente persona)	0,007194	4	<b>3</b>
		4 (diferente persona)	0,010271	3	<b>5</b>
		5 (misma persona)	0,003915	1	<b>1</b>
<b>2</b>	Imagen 2	1 (misma persona)	0,001654	1	<b>2</b>
		2 (diferente persona)	0,001575	2	<b>1</b>
		3 (diferente persona)	0,007863	5	<b>5</b>
		4 (diferente persona)	0,007214	4	<b>4</b>
		5 (diferente persona)	0,005255	3	<b>3</b>
<b>3</b>	Imagen 3	1 (diferente persona)	0,005543	3	<b>3</b>
		2 (misma persona)	0,001794	1	<b>1</b>
		3 (diferente persona)	0,007434	4	<b>5</b>
		4 (misma persona)	0,003115	2	<b>2</b>
		5 (diferente persona)	0,005561	5	<b>4</b>
<b>4</b>	Imagen 4	1 (misma persona)	0,002596	1	<b>2</b>
		2 (diferente persona)	0,002477	2	<b>1</b>
		3 (diferente persona)	0,003416	5	<b>3</b>
		4 (diferente persona)	0,007621	4	<b>4</b>
		5 (diferente persona)	0,007834	3	<b>5</b>
<b>5</b>	Imagen 5	1 (diferente persona)	0,016859	5	<b>5</b>
		2 (diferente persona)	0,007369	2	<b>2</b>
		3 (misma persona)	0,006267	1	<b>1</b>
		4 (diferente persona)	0,009441	3	<b>4</b>
		5 (diferente persona)	0,009102	4	<b>3</b>

**Tabla 2. Resultados del reconocimiento de rostro.**

En la primera prueba se observa como los dos resultados más semejante a los definidos visualmente coinciden con los devueltos por el sistema entre ellos el más parecido coincide exactamente con la



persona entrada por el usuario, sin embargo el tercer caso no coincide, pero se observa que el sistema ha cometido solamente un error en el orden de los 5 elementos con respecto a la percepción visual real de las imágenes. En la segunda prueba se visualiza como se intercambian solamente los dos primeros resultados, debido al parecido real de las imágenes. El resto coincide exactamente con el orden visual propuesto.

En la tercera prueba se pueden observar resultados positivos, dado que se trabajaron con dos imágenes de la misma persona y en ambos casos fue correcta la propuesta del sistema. Se destaca que las dos últimas imágenes ordenadas visualmente no correspondieron con las del sistema. La cuarta prueba arroja resultados interesantes, es una prueba realizada solo con rostros de mujeres, invirtió los dos primeros resultados, pero el caso es que el segundo rostro visualmente tiene un alto parecido a la imagen entrada y se nota una pequeña diferencia con respecto al primero que es la que constituye la misma persona.

En la quinta y última prueba se puede observar que respondió adecuadamente y solo en el cuarto caso intercambió los resultados, cabe destacar que el quinto resultado es un valor bien lejano puesto que corresponde a un rostro bastante diferente, siendo acertada la respuesta.

### 3.3. Conclusiones parciales

El prototipo implementado presenta de manera práctica las potencialidades y debilidades propias del procedimiento, las cuales constituyen recomendaciones para futuros estudios.

La etapa de pruebas que se subdividió en dos momentos, detección de rostros primeramente y después extracción y confrontación, resultó muy útil para determinar la viabilidad del procedimiento, observándose que en el caso de la validación del rostro, con el empleo de *Viola and Jones*, se logró un alto nivel de efectividad, óptimo para su empleo en sistemas de reconocimiento. Se pudo constatar además en un segundo momento al realizar las pruebas a la etapa de extracción de características y confrontación, a partir de la utilización de los Jets de Gabor y distancia euclidiana, que los resultados fueron aceptables, debido a que en la totalidad de los casos se obtuvo el rostros de la persona a encontrar en los primeros mostrados por el sistema, se observaron ciertas fallas pero que eran válidas, porque estas sucedieron en rostros no tan parecidos al entrado por el usuario.

## CONCLUSIONES

### Se puede concluir lo siguiente:

- La catalogación de materiales audiovisuales es una necesidad para garantizar la ágil recuperación de los mismos en empresas que los producen y/o almacenan, en el presente dichas empresas están sumergidas en el proceso de digitalización, el que asegura una mejor gestión de la documentación audiovisual.
- La digitalización ha modificado el proceso de catalogación, si antes se realizaba de forma manual con la presencia de tarjeteros y ficheros físicos con la información relativa a los materiales, ahora se puede realizar a partir de la visualización digital del material y la inserción de datos relativos al mismo también de forma digital.
- Uno de los problemas que afecta la adecuada catalogación de los materiales es la falta de información que pueden tener los documentalistas sobre los personajes que aparecen en los mismos, dato que es oportuno conocer para disímiles usos. Un procedimiento basado en técnicas de reconocimiento de rostro que permita la implementación de sistemas de apoyo al documentalista, es una solución para atenuar el problema mencionado.
- El reconocimiento de rostro presenta una serie de problemáticas en la actualidad que van desde los cambios de iluminación, de pose, presencia de lentes y otros accesorios hasta el hecho de contar con pocas imágenes de una misma persona, dificultando la realización de entrenamientos que reduzcan el margen de error en sistemas encargados de realizarlo.
- Se expone que la etapa de detección presenta actualmente resultados muy favorables para la creación de sistemas enfocados en esta área, dado por la efectividad de los métodos que se presentan en este campo, uno de esos caso es el del método de *Viola and Jones* utilizado en este trabajo. En contraste a esto, no existe el mismo avance en los métodos de extracción de características y definición de descriptores, pues actualmente donde se tienen mejores resultados es en temas de seguridad, y de igual manera no presentan los mismos niveles de efectividad que en el caso de detección.
- Se presentan como métodos destacados ante los problemas de iluminación, presencia de accesorios y cantidad de imágenes de comparación, que son unos de las más frecuentes dentro de las problemáticas planteadas, al descriptor de Wavelets de Gabor y el método EBGMM alcanzándose resultados aceptables para el procedimiento propuesto.
- El procedimiento de detección de rostros constituirá un apoyo al proceso de catalogación y será un valor agregado a los productos de gestión y catalogación de media del departamento Señales Digitales, incrementando sus funcionalidades y potencialidades.

## RECOMENDACIONES

- Se recomienda para el perfeccionamiento de los resultados del trabajo el estudio de técnicas que se puedan combinar con jets de Gabor, como puede ser el caso de Local Binary Patterns y técnicas discriminantes como es el caso de PCA. Se propone además una selección, de los puntos donde se aplican los filtros de Gabor, basado en entrenamientos en rostros, que permita autónomamente la definición de los mismos. Todo ello para lograr mejoras progresivas en los resultados obtenidos hasta el momento.
- Se propone seguir estudiando métodos de similitud que puedan arrojar mejores resultados que los actuales, algunas propuestas pudieran ser, métodos de similitudes por magnitud ponderadas. También pudiera analizarse el caso de métodos radiales todo ello con la intención de extraer de los vectores de características los mejores datos para comparar con los almacenados.
- Otra recomendación válida se enfoca hacia la recuperación de los vectores de características almacenados en base de datos, que puede partir por propuestas de organización según sus características y distancias entre ellos, con el objetivo de acelerar el proceso.
- Se recomienda migrar la implementación del procedimiento a tecnologías libres, acordes con las políticas de soberanía tecnológica del departamento. Se recomienda el uso de C++ con Qt, debido a su potencialidad en el tratamiento de imágenes y su integración con la biblioteca OpenCV.

## Referencias Bibliográficas

1. Biometric Watch. Journal for the Biometric Industry . [Online] Septiembre 20, 2006. [Cited: Diciembre 11, 2011.] <http://www.biometricwatch.com/index.html>.
2. **González, Enrique Almeida Maldonado y Ronald Baby.** Sistema de Catalogación de Medias. Ciudad Habana : s.n., 2009.
3. Dictionary of Video and Television Technology. USA : Elsevier Science, 2002. ISBN 1-878707-99-X..
4. **wordreference.** <http://www.wordreference.com/definicion/catalogacion>. [Online] [Cited: Enero 2, 2012.]
5. **M. Delbracio C. Aguerrebere, G. Capdehourat and M. Mateu.** Proyecto aguara, reconocimiento automatico de caras. Uruguay : s.n., 2006.
6. **AntitzaDantcheva, Jean-Luc Dugelay, Petros Elia and Sophia Antipolis.** Person Recognition using a bag of facial soft biometrics (BoFSB) and Multimedia Communications Department. France : EURECOM.
7. **Chellappa, Kamran Etemad and Rama.** Discriminant analysis for recognition of human face images. In Proceedings of the 1st International Conference on Audio- and Video-Based Biometric Person Authentication. Londres : AVBPA , 1997.
8. **Ael Adini, Yael Moses and Shimon Ullman.** Face Recognition: The Problem of Compensating for Changes in Illumination Direction. s.l. : IEEE , 1997.
9. **Lossurveys de M-H. Yang, D.J Kriegman, and N. Ahuja.** Detecting faces in images: A survey. s.l. : IEEE , January 2002.
10. **Low, Erik Hjelmas and Boon Kee.** Face detection: A survey. Computer Vision and Image Understanding. 2001.
11. **W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld.** Face recognition: A literature survey. s.l. : ACM, 2003.
12. **R. Chellappa, C. Wilson, and S. Sirohey.** Human and machine recognition of faces: a survey. s.l. : IEEE, 1995.
13. **Marcel, Sebastien.** A tutorial on face recognition. Switzerland : IDIAP , June 2007.
14. **Jain, Stan Z. Li and Anil K.** Handbook of face recognition. s.l. : Springer, 2005.
15. **Xiaoyang Tan, Songcan Chen , Zhi-Hua Zhou , Fuyan Zhang.** Face recognition from a single image per person: A survey. 2006.

16. **Reyes, Heidi Méndez Vázquez y Edel García.** Estado actual de los métodos de reconocimiento automático de rostros basados en la apariencia local. Ciudad de La Habana : CENATAV, 2008. 2072-6887.
17. **García, Oscar Boullosa.** Estudio comparativo de descriptores visuales para la detección de escenas cuasiduplicadas. Madrid : s.n., 2011.
18. RAE. [Online] [Cited: Febrero 03, 2012.]  
[http://buscon.rae.es/draeI/SrvltConsulta?TIPO\\_BUS=3&LEMA=procedimiento](http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=procedimiento).
19. **Fuentes, Edwin Omar Ortiz y Henry Argüello.** Reconocimiento de rostros utilizando PCA en dispositivo DSPIng. 2008.
20. **Capdehourat, Cecilia Aguerrebere y German.** Reconocimiento de caras con Características locales. marzo 2006.
21. **Tan X., Chen S.C., Zhou Z.-H., and Zhang F.** Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft kNN ensemble. s.l. : IEEE, 2005.
22. **Martinez, A.M.** Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. s.l. : IEEE , 2002.
23. **Stiefelhagen, Hazim Kemal Ekenel y Rainer.** Local Appearance Based Face Recognition using Discrete Cosine Transform. s.l. : EUSIPCO, 2005.
24. **Kepenekci B., Tek F. B., G. Bozdagi Akar.** Occluded Face Recognition based on Gabor Wavelets. Rochester NY : ICIP , Sept 2002.
25. **Li, Hung-Son Le y Haibo.** Recognizing frontal face images using Hidden Markov models with one training image per person. s.l. : ICPR04, 2004.
26. **Pentland, Matthew Turk and Alex Paul.** Eigenfaces for recognition. Journal of Cognitive Neuroscience. 1991.
27. **C. Havran, L. Hupet, J. Czyz, J. Lee, L. Vandendorpe, and M. Verleysen.** Independent component analysis for face authentication. In Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information and Engineering Systems. s.l. : KES , September 2002.
28. **W. Zhao, R. Chellappa, and A. Krishnaswamy.** Discriminant analysis of principal components for face recognition. Nara, Japan : s.n., 1998.
29. **Yang, Ming-Hour.** Face recognition using kernel methods . s.l. : MIT Press, December 2001.
30. **Wechsler, Chengjun Liu and Harry.** Evolutionary pursuit and its application to face recognition. s.l. : IEEE , 2000.
31. **G. Guo, S. Li, and K. Chan.** Face recognition by support vector machines. 2000.

32. **Gareth J. Edwards, Timothy F. Cootes, and Christopher J. Taylor.** Face recognition using active appearance models. Freiburg, Germany : Springer-Verlag, 1998.
33. **Jain, R. Hsu and A.** Face modeling for recognition. s.l. : IEEE , February 2001.
34. **Vetter, Volker Blanz and Thomas.** Face recognition based on fitting a 3D morphable model. s.l. : IEEE , 2003.
35. **LaurenzWiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg.** Face recognition by elastic bunch graph matching. s.l. : CRC Press, 1999.
36. **Markov, A. Nefian. A hidden.** model-based approach for face detection and recognition. Atlanta : s.n., 1999.
37. **TimoAhonen, AbdenourHadid y MattiPietikäinen.** Face Recognition with Local Binary Patterns. 2004.
38. **James Short, Josef Kittler and Kieron Messer.** A Comparison of Photometric Normalisation Algorithms for Face Verification. s.l. : IEEE , 2004.
39. **Bolme, David S.** Elastic bunch graph matching. s.l. : Colorado State University, 2003.
40. **Ariel, M. A.** Detección y reconocimiento de caras. 2011.
41. **FERNANDO, C. S. M.** Tutorial Viola-Jones. 2009.
42. **CASTRILLÓN M, DÉNIZ O.** A comparison of face and facial feature detectors based on the Viola-Jones general object detection framework. Machine Vision & Applications. 2011.
43. **Zorilin Alonso Guerrero, Nilo Tomás Díaz Ales , Angel Dayán Marín Abreu.** Componente de detección de rostros para la catalogación de audiovisuales: conceptualización. Ciudad Habana : s.n., 2012.
44. National Institute of Standards & Technology (nist). Ferret Evaluation. [Online] [Cited: 04 25, 2012.] <http://www.itl.nist.gov/iad/humanid/feret/perf/eval.html..>

## Bibliografía

1. Biometric Watch. Journal for the Biometric Industry . [Online] Septiembre 20, 2006. [Cited: Diciembre 11, 2011.] <http://www.biometricwatch.com/index.html>..
2. **González, Enrique Almeida Maldonado y Ronald Baby.** *Sistema de Catalogación de Medias.* Ciudad Habana : s.n., 2009.
3. *Dictionary of Video and Television Technology.* USA : Elsevier Science, 2002. ISBN 1-878707-99-X..
4. **wordreference.** <http://www.wordreference.com/definicion/catalogacion>. [Online] [Cited: Enero 2, 2012.]
5. **M. Delbracio C. Aguerrebere, G. Capdehourat and M. Mateu.** *Proyecto aguara, reconocimiento automatico de caras.* Uruguay : s.n., 2006.
6. **AntitzaDantcheva, Jean-Luc Dugelay, Petros Elia and Sophia Antipolis.** *Person Recognition using a bag of facial soft biometrics (BoFSB) and Multimedia Communications Department.* France : EURECOM.
7. **Chellappa, Kamran Etemad and Rama.** *Discriminant analysis for recognition of human face images.* In *Proceedings of the 1st International Conference on Audio- and Video-Based Biometric Person Authentication.* Londres : AVBPA , 1997.
8. **Ael Adini, Yael Moses and Shimon Ullman.** *Face Recognition: The Problem of Compensating for Changes in Illumination Direction.* s.l. : IEEE , 1997.
9. **lossurveys de M-H. Yang, D.J Kriegman, and N. Ahuja.** *Detecting faces in images: A survey.* s.l. : IEEE , January 2002.
10. **Low, Erik Hjelmas and Boon Kee.** *Facetedetection: A survey.* *Computer Vision and Image Understanding.* 2001.
11. **W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld.** *Face recognition: A literature survey.* s.l. : ACM, 2003.
12. **R. Chellappa, C. Wilson, and S. Sirohey.** *Human and machine recognition of faces: a survey.* s.l. : IEEE, 1995.
13. **Marcel, Sebastien.** *A tutorial on facerecognition.* Switzerland : IDIAP , June 2007.
14. **Jain, Stan Z. Li and Anil K.** *Handbook of facerecognition.* s.l. : Springer, 2005.
15. **Xiaoyang Tan, SongcanChen , Zhi-HuaZhou , Fuyan Zhang.** *Facerecognition from a single image per person: A survey.* 2006.

16. **Reyes, Heidi Méndez Vázquez y Edel García.** *Estado actual de los métodos de reconocimiento automático de rostros basados en la apariencia local.* Ciudad de La Habana : CENATAV, 2008. 2072-6887.
17. **García, Oscar Boulosa.** *Estudio comparativo de descriptores visuales para la detección de escenas cuasiduplicadas.* Madrid : s.n., 2011.
18. RAE. [Online] [Cited: Febrero 03, 2012.]  
[http://buscon.rae.es/draeI/SrvltConsulta?TIPO\\_BUS=3&LEMA=procedimiento](http://buscon.rae.es/draeI/SrvltConsulta?TIPO_BUS=3&LEMA=procedimiento).
19. **Fuentes, Edwin Omar Ortiz y Henry Argüello.** *Reconocimiento de rostros utilizando PCA en dispositivo DSPIng.* 2008.
20. **Capdehourat, Cecilia Aguerrebere y German.** *Reconocimiento de caras con Características locales.* marzo 2006.
21. **Tan X., Chen S.C., Zhou Z.-H., and Zhang F.** *Recognizing partially occluded, expression variant faces from single training image per person with SOM and soft kNN ensemble.* s.l. : IEEE, 2005.
22. **Martinez, A.M.** *Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class.* s.l. : IEEE , 2002.
23. **Stiefelhagen, Hazim Kemal Ekenel y Rainer.** *Local Appearance Based Face Recognition using Discrete Cosine Transform.* s.l. : EUSIPCO, 2005.
24. **Kepenekci B., Tek F. B., G. BozdagiAkar.** *Occluded Face Recognition based on Gabor Wavelets.* Rochester NY : ICIP , Sept 2002.
25. **Li, Hung-Son Le y Haibo.** *Recognizing frontal face images using Hidden Markov models with one training image per person.* s.l. : ICPR04, 2004.
26. **Pentland, Matthew Turk and Alex Paul.** *Eigenfaces for recognition.* *Journal of Cognitive Neuroscience.* 1991.
27. **C. Havran, L. Hupet, J. Czyz, J. Lee, L. Vandendorpe, and M. Verleysen.** *Independent component analysis for face authentication.* *In Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information and Engineering Systems.* s.l. : KES , September 2002.
28. **W. Zhao, R. Chellappa, and A. Krishnaswamy.** *Discriminant analysis of principal components for face recognition.* Nara, Japan : s.n., 1998.
29. **Yang, Ming-Hour.** *Face recognition using kernel methods .* s.l. : MIT Press, December 2001.
30. **Wechsler, Chengjun Liu and Harry.** *Evolutionary pursuit and its application to face recognition.* s.l. : IEEE , 2000.
31. **G. Guo, S. Li, and K. Chan.** *Face recognition by support vector machines.* 2000.



32. **Gareth J. Edwards, Timothy F. Cootes, and Christopher J. Taylor.** *Face recognition using active appearance models.* Freiburg, Germany : Springer-Verlag, 1998.
33. **Jain, R. Hsu and A.** *Face modeling for recognition.* s.l. : IEEE , February 2001.
34. **Vetter, Volker Blanz and Thomas.** *Face recognition based on fitting a 3D morphable model.* s.l. : IEEE , 2003.
35. **Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg.** *Face recognition by elastic bunch graph matching.* s.l. : CRC Press, 1999.
36. **Markov, A. Nefian.** *A hidden. model-based approach for face detection and recognition.* Atlanta : s.n., 1999.
37. **Timo Ahonen, Abdenour Hadid y Matti Pietikäinen.** *Face Recognition with Local Binary Patterns.* 2004.
38. **James Short, Josef Kittler and Kieron Messer.** *A Comparison of Photometric Normalisation Algorithms for Face Verification.* s.l. : IEEE , 2004.
39. **Bolme, David S.** *Elastic bunch graph matching.* s.l. : Colorado State University, 2003.
40. **Ariel, M. A.** *Detección y reconocimiento de caras.* 2011.
41. **FERNANDO, C. S. M.** *Tutorial Viola-Jones.* 2009.
42. **CASTRILLÓN M, DÉNIZ O.** *A comparison of face and facial feature detectors based on the Viola-Jones general object detection framework.* *Machine Vision & Applications.* 2011.
43. **Zorilin Alonso Guerrero, Nilo Tomás Díaz Ales , Angel Dayán Marín Abreu.** *Componente de detección de rostros para la catalogación de audiovisuales: conceptualización.* Ciudad Habana : s.n., 2012.
44. National Institute of Standards & Technology (nist). Ferret Evaluation. [Online] [Cited: 04 25, 2012.] <http://www.itl.nist.gov/iad/humanid/feret/perf/eval.html>.
45. **Laurenz Wiskott, Jean-Marc Fellous, Norbert Krüger, and Christoph von der Malsburg.** *Face recognition by elastic bunch graph matching.* Heidelberg, Germany : s.n., 1997.
46. **Kepenekci B., Tek F. B., G. Bozdagi Akar.** *Occluded Face Recognition based on Gabor Wavelets.* New York : Rochester, Sept 2002.
47. Wordreference. [Online] [Cited: 05 01, 2012.] <http://vis-www.cs.umass.edu/lfw>.
48. **Kothari, R. Lotlikar and R.** *Fractional-step dimensionality reduction.* s.l. : IEEE , 2000.
49. **Ruud M. Bolle, Sharath Pankanti, and Nalini K. Ratha.** *Evaluation techniques for biometrics-based authentication systems (frr).* . Barcelona, Spain : s.n., September 2000.

50. **P. Jonathon Phillips, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss.** *The feret evaluation methodology for face-recognition algorithms.* . s.l. : IEEE , 2000.
51. **S. Rizvi, P. Phillips, and H. Moon.** *The FERET verification testing protocol for face recognition algorithms.* s.l. : Image and Vision Computing Journal, 1999.
52. **Phillips, H. Moon and P.J.** *Analysis of pca-based face recognition algorithms.* Santa Barbara, California : s.n., June 1998.
53. **Teixeira, Marcio Luis.** *The bayesian intrapersonal/extrapersonalclassifier.* Colorado, USA : s.n., 2003.

## **Glosario de Términos**

### **Algoritmo**

Conjunto ordenado y finito de operaciones que permite hallar la solución de un problema. En este trabajo se refiere al conjunto de procedimientos que se ejecutan de manera ordenada para lograr el reconocimiento automático de los individuos a partir de imágenes del rostro.

### **Método**

Modo de obrar o proceder. En el contexto del trabajo se refiere a la forma de ejecutar alguno de los pasos necesarios en el algoritmo de reconocimiento de rostros.

### **Sistema**

Conjunto de reglas o principios sobre una materia, estructurados y enlazados entre sí.

### **Software**

Es el conjunto de los programas de cómputo, procedimientos, reglas, documentación y datos asociados que forman parte de las operaciones de un sistema de computación.

### **Convolución**

Se denomina convolución a una función, que de forma lineal y continua, transforma una señal de entrada en una nueva señal de salida.

### **Kernel**

En procesamiento de imágenes, normalmente se le denomina *kernel* a la función de impulso utilizada en una convolución.

### **Eigenfaces**

Término en inglés por el cual se conoce al algoritmo de reconocimiento de rostros basado en el análisis de componentes principales (PCA).

### **Histograma**

El histograma de una imagen es una distribución que representa el número de píxeles por cada valor de intensidad presente en la imagen.

### **Wavelets**

Una *wavelet* es una oscilación de onda, con una amplitud que comienza en cero, aumenta y luego disminuye de nuevo a cero. Una serie de *wavelets* es una representación de una función cuadrática integrable por una cierta serie ortonormal generada por una *wavelet*.