

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA**

### **ANÁLISIS DE DOS ALGORITMOS PARA DETECCIÓN DE ROSTROS E IMPLEMENTACIÓN DE UNO DE ELLOS UTILIZADO EN ANALÍTICA DE VIDEO PARA SISTEMAS DE VIDEOVIGILANCIA**

#### **PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO EN ELECTRÓNICA Y TELECOMUNICACIONES**

**SÁNCHEZ SALAZAR CRISTIAN RENÉ**  
cristian4248@hotmail.com

**DIRECTOR: ING. DIANA MARTÍNEZ, MSc.**  
diana.martinez@epn.edu.ec

**CODIRECTOR: ING. JORGE CARVAJAL, MSc.**  
jorge.carvajal@epn.edu.ec

**Quito, junio 2017**

## **DECLARACIÓN**

Yo, Cristian René Sánchez Salazar, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

---

Cristian Sánchez

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por Cristian René Sánchez Salazar, bajo mi supervisión.

---

**Ing. Diana Martínez, MSc**  
**DIRECTOR DEL PROYECTO**

---

**Ing. Jorge Carvajal, MSc**  
**CODIRECTOR DEL PROYECTO**

## **AGRADECIMIENTOS**

A mi Directora de Tesis Ing. Diana Martínez, por su comprensión, paciencia y apoyo; por ser maestra y guía en cada momento durante la realización de mi trabajo de titulación.

A mi Codirector de Tesis Ing. Jorge Carvajal, por su orientación y enseñanza durante el desarrollo de mi trabajo de titulación.

A la Escuela Politécnica Nacional, por el conocimiento impartido a través de sus profesores durante mi trayectoria universitaria.

A mis compañeros, por brindarme su amistad y apoyo en el recorrido de mi formación profesional.

**Cristian**

## **DEDICATORIA**

A mi Dios, por darme su bendición y por ser parte de cada decisión tomada en mi vida, con alegrías y tristezas siempre me alentó a ser cada vez más fuerte y perseverante para lograr lo que me propongo.

A mi esposa y compañera de vida Mayra Sandoval, por cruzarte en mi camino y brindarme todo tu amor, por permitirme ser tu esposo y compartir una vida. Por ser mi pilar principal para terminar mi formación profesional. Por bendecirme con un hijo hermoso, Cristian Sebastián que eres mi alegría y mi todo. Por ser parte de esta familia hermosa de la cual doy gracias cada instante de mi vida.

A mi padre formativo Manuel Onofa por abrimme las puertas de su corazón y hogar, permitiéndome crecer y formarme como un hijo verdadero brindándome su confianza, alegría, respeto, firmeza, cariño y sobre todo muchos cuidados.

A mi madre Flora por estar ahí en los momentos buenos y malos de mi niñez y adolescencia, cada aliento y regaño fueron bien aplicados para guiarme en el camino correcto de mi profesión.

A mi madre Guille por ser esa Figura de rectitud, responsabilidad, honestidad, nobleza y trabajo que es el complemento necesario para determinar en mí una persona de carácter formado y seguro de las decisiones que en la vida debes tomar. Además de ser el ejemplo de mujer abnegada para sus hijos, esposo y hogar.

A mis hermanos y a mis primos que los considero como hermanos por permitirme ser parte de su crecimiento y por dejarme compartir su espacio, por ser mi fuerza de superación y ejemplo para ellos para que sus vidas sean tan bendecidas como la mía por contar con una familia hermosa.

Finalmente, a todas las personas que formaron parte de mi vida, unas para ayudarme y otras para obligarme a ser más fuerte; para tomar mis decisiones, para aprender con las que se puede contar y en quienes se puede confiar incondicionalmente.

**Cristian**

## CONTENIDO

CONTENIDO.....	I
RESUMEN .....	XIII
PRESENTACIÓN.....	XIV
CAPÍTULO 1 FUNDAMENTOS TEÓRICOS DE LOS SISTEMAS DE VIDEOVIGILANCIA.....	1
1.1 INTRODUCCIÓN .....	1
1.2 GENERALIDADES .....	1
1.3 VIDEOVIGILANCIA.....	2
1.4 EVOLUCIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA .....	3
1.4.1 SISTEMAS DE VIDEOVIGILANCIA ANALÓGICOS.....	4
1.4.1.1 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN VCR.....	5
1.4.1.2 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN DVR.....	5
1.4.1.3 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN DVR DE RED .....	6
1.4.2 SISTEMAS DE VIDEOVIGILANCIA IP.....	7
1.5 COMPONENTES DE LOS SISTEMAS DE VIDEOVIGILANCIA.....	8
1.5.1 MONITOR.....	9
1.5.2 VIDEO GRABADOR NVR.....	10
1.5.3 SERVIDOR DE VIDEO.....	10
1.5.4 CONMUTADOR DE RED.....	11

1.5.5 CONJUNTO DE CÁMARAS.....	11
1.5.5.1 DEFINICIÓN DE CÁMARA DE VIDEO IP .....	11
1.5.5.2 FUNCIONAMIENTO DE UNA CÁMARA DE VIDEOVIGILANCIA IP .....	12
1.5.5.3 COMPONENTES DE UNA CÁMARA IP .....	12
1.5.5.4 TIPOS DE CÁMARAS .....	18
1.5.6 CONEXIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA.....	20
1.5.6.1 MEDIOS GUIADOS.....	20
1.5.6.2 MEDIOS NO GUIADOS .....	20
1.6 AMBIENTES DE INSTALACIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA ..	20
1.7 ANALÍTICA DE VIDEO.....	21
1.7.1 INICIOS DE LA ANALÍTICA DE VIDEO.....	22
1.7.2 COMO FUNCIONA LA VISIÓN ARTIFICIAL.....	24
1.7.2.1 ADQUISICIÓN DE LA IMAGEN .....	25
1.7.2.2 PRE-PROCESADO .....	25
1.7.2.3 SEGMENTACIÓN .....	25
1.7.2.4 REPRESENTACIÓN Y DESCRIPCIÓN (EXTRACCIÓN DE CARACTERÍSTICAS) .....	26
1.7.2.5 RECONOCIMIENTO E INTERPRETACIÓN.....	26
1.7.3 TIPOS DE SISTEMAS PARA PROCESAMIENTO DE DATOS EN ANALÍTICA DE VIDEO.....	27

1.7.4 TIPOS DE FILTROS EN ANALÍTICA DE VIDEO.....	29
1.7.5 INSTALACIÓN DE UN SISTEMA DE VIDEOVIGILANCIA.....	32
1.7.5.1 FACTORES DE INSTALACIÓN .....	32
CAPÍTULO 2 ANÁLISIS DE DOS ALGORITMOS PARA DETECCIÓN DE ROSTROS, VIOLA&JONES Y HAUSDORFF.....	34
2.1 INTRODUCCIÓN .....	34
2.2 DETECCIÓN DE ROSTROS EN IMÁGENES DE VIDEO .....	34
2.3 ANTECEDENTES.....	35
2.4 CLASIFICACIÓN DE MÉTODOS PARA DETECCIÓN DE ROSTROS .....	36
2.5 TIPOS DE ALGORITMOS PARA DETECCIÓN DE ROSTROS.....	39
2.5.1 DETECCIÓN DE ROSTROS EN IMÁGENES CON FONDO CONTROLADO	39
2.5.2 DETECCIÓN DE ROSTROS POR COLOR.....	39
2.5.3 DETECCIÓN DE ROSTROS POR EL MOVIMIENTO.....	39
2.5.4 DETECCIÓN DE ROSTROS USANDO VARIOS ALGORITMOS.....	40
2.5.5 DETECCIÓN DE ROSTROS EN ESCENAS SIN RESTRICCIONES.....	40
2.6 ALGORITMO DE VIOLA&JONES .....	41
2.6.1 DESCRIPCIÓN,.....	41
2.6.2 INDICADORES.....	41
2.6.3 METODOLOGÍA.....	42
2.6.4 ETAPAS.....	43



2.6.4.1 ETAPA DE CARACTERÍSTICAS .....	44
2.6.4.2 ETAPA DE LA IMAGEN INTEGRAL .....	46
2.6.4.3 ETAPA DE ENTRENAMIENTO .....	49
2.6.4.4 ETAPA CLASIFICADORES EN CASCADA .....	52
2.7 DISTANCIA DE HAUSDORFF .....	58
2.7.1 DEFINICIÓN.....	58
2.7.2 DETECCIÓN DE ROSTROS BASADO EN LA DISTANCIA DE HAUSSDORFF.....	62
2.7.3 DETECCIÓN BASADO EN MODELOS GEOMÉTRICOS.....	64
2.7.4 ETAPAS DEL ALGORITMO PARA LA DETECCIÓN [21].....	66
2.7.4.1 DETECCIÓN GRUESA O GENERAL.....	66
2.7.4.2 REFINAMIENTO .....	70
2.8 TABLA COMPARATIVA DE LOS ALGORITMOS .....	70
CAPÍTULO 3 IMPLEMENTACIÓN DEL ALGORITMO Y ANÁLISIS DE RESULTADOS.....	73
3.1 INTRODUCCIÓN .....	73
3.2 SOFTWARE DE SIMULACIÓN.....	73
3.3 AMBIENTE DE SIMULACIÓN.....	75
3.4 DESARROLLO DEL CÓDIGO .....	77
3.5 EJECUCIÓN DEL CÓDIGO EN MATLAB .....	85

3.5.1 INICIALIZACIÓN Y PRESENTACIÓN DEL CÓDIGO.....	85
3.5.2 EJECUCIÓN DEL CÓDIGO PARA DETECCIÓN DE ROSTROS.....	87
3.5.2.1 PRUEBAS EN AMBIENTES CON ILUMINACIÓN LIMITADA .....	87
3.5.2.2 PRUEBAS EN AMBIENTES CON ILUMINACIÓN NORMAL.....	90
3.5.2.3 PRUEBAS CON OTROS OBJETOS PARA DETECCIÓN DE ROSTROS ....	93
3.6 CÓDIGO COMERCIAL DE MATHWORKS .....	96
3.7 EJECUCIÓN DEL CÓDIGO DE MATHWORKS .....	99
3.7.1 EJECUCIÓN DE LA DETECCIÓN DE ROSTROS.....	99
3.7.1.1 PRUEBAS EN AMBIENTE CON ILUMINACIÓN LIMITADA.....	99
3.7.1.2 PRUEBAS EN AMBIENTES CON ILUMINACIÓN ADECUADA .....	102
3.7.1.3 PRUEBAS CON OTROS OBJETOS PARA DETECCIÓN DE ROSTROS ..	104
3.8 ANÁLISIS DE LAS PRUEBAS REALIZADAS .....	108
CAPÍTULO 4 CONCLUSIONES Y RECOMENDACIONES.....	112
4.1 CONCLUSIONES .....	112
4.2 RECOMENDACIONES .....	115
REFERENCIAS BIBLIOGRÁFICAS.....	118

## ÍNDICE DE FIGURAS

### CAPÍTULO 1

Figura 1.1: Sistema de CCTV tradicional con VCR. [50] .....	5
Figura 1.2: Sistema de CCTV con DVR [50] .....	6
Figura 1.3: Sistema de CCTV conectado a la Red [50] .....	6
Figura 1.4: Sistema de videovigilancia IP. a) Red LAN. b) Red WAN [48].....	7
Figura 1.5: Esquema de sistema de videovigilancia [80] .....	9
Figura 1.6: Dispositivos de monitoreo [86] .....	10
Figura 1.7: Funcionamiento básico interno de una cámara de video IP [39].....	12
Figura 1.8: Esquema básico de una cámara IP [52] .....	13
Figura 1.9: Distancia focal de un lente IP [52] .....	14
Figura 1.10: Localización en el espectro radioeléctrico de la parte visible por el hombre (luz) [52] .....	15
Figura 1.11: Apertura del iris [52] .....	15
Figura 1.12: Etapas de una aplicación de visión artificial [66].....	24
Figura 1.13: Sistema embebido [67] .....	28
Figura 1.14: Sistema basado en un servidor [67] .....	28
Figura 1.15: Sistema híbrido [67] .....	29

### CAPÍTULO 2

Figura 2.1: Clasificación de la detección de rostros basado en las características e imágenes [5]. .....	38
Figura 2.2: Característica de Haar [11].....	45
Figura 2.3: Imagen integral en el punto (x, y) [95] .....	47
Figura 2.4: Representación de píxeles [11] .....	48
Figura 2.5: Escalado de características [11].....	49
Figura 2.6: Curva ROC para un clasificador compuesto por 200 características rectángulo [11]. .....	54
Figura 2.7: Representación esquemática de una cascada de detección [11]. .....	54
Figura 2.8: Características utilizadas en Viola&Jones [11]. .....	56

Figura 2.9: Conjunto de puntos de A y B [16].....	60
Figura 2.10: Distancias más cortas desde A hacia B [16].....	60
Figura 2.11: $h = (A, B)$ Distancia de A hacia B [16].....	60
Figura 2.12: $h(A, B) = d(a_1, b_1)$ Distancia más lejana [16].....	61
Figura 2.13: Procedimiento para buscar rostros [79].....	63
Figura 2.14: Modelo de ajuste por escala, traslado y rotación [20]. ....	65
Figura 2.15: Etapa de detección gruesa o general [20] .....	66
Figura 2.16: Etapa de detección de refinamiento [20] .....	66
Figura 2.17: Detección de segmentos .....	67
Figura 2.18: Esquema del funcionamiento de la fase de localización facial [22]. ....	68
Figura 2.19: Problema de escala. a) Detección de rostros sin factor de normalización. b) Detección de rostros usando factor de normalización [22]. ....	69

### CAPÍTULO 3

Figura 3.1: Diagrama de flujo del código implementado .....	78
Figura 3.2: Registro de temperatura en la sala de pruebas .....	85
Figura 3.3: Ventana de comandos corriendo código implementado .....	86
Figura 3.4: Ventana de presentación del trabajo realizado.....	86
Figura 3. 5: Presentación completa del trabajo realizado .....	86
Figura 3.6: Detección frontal con poca iluminación .....	88
Figura 3.7: Cámara web con enfoque a 45° con poca iluminación .....	88
Figura 3.8: Detección de rostros girados a 45° con poca iluminación.....	89
Figura 3.9: Detección de rostros girados a 90° con poca iluminación.....	90
Figura 3.10: Detección frontal con iluminación normal .....	91
Figura 3.11: Cámara web con enfoque a 45° con iluminación normal .....	91
Figura 3.12: Detección de rostros girados a 45° con iluminación normal .....	92
Figura 3.13: Detección de rostros girados a 90° con iluminación normal .....	92
Figura 3. 14: Detección de rostro con facciones completas .....	94
Figura 3.15: Detección de facciones de un rostro sin facciones completas de rostro	94

Figura 3.16: Detección de rostro blanco y negro .....	95
Figura 3.17: Detección de rostro casi completo.....	95
Figura 3.18: Detección de rostro confuso.....	96
Figura 3. 19: Detección frontal con poca Iluminación .....	99
Figura 3.20: Cámara con enfoque a 45° con poca iluminación.....	100
Figura 3.21: Detección de rostros girados a 45° con poca iluminación.....	101
Figura 3.22: Detección de rostros girados a 90° con poca iluminación.....	101
Figura 3.23: Detección frontal con iluminación normal .....	102
Figura 3. 24: Cámara con enfoque a 45° con iluminación normal .....	103
Figura 3.25: Detección de rostros girados a 45° con iluminación normal .....	103
Figura 3.26: Detección de rostros girados a 90° con iluminación normal .....	104
Figura 3.27: Detección de rostro con facciones completas .....	105
Figura 3.28: Detección de facciones de un rostro sin facciones completas de rostro .....	106
Figura 3.29: Detección de rostro blanco y negro .....	106
Figura 3.30: Detección de rostro casi completo.....	107
Figura 3.31: Detección de rostro confuso.....	107
Figura 3.32: Análisis de resultados del código desarrollado .....	109
Figura 3.33: Análisis de resultados del código de Mathworks .....	110

## ÍNDICE DE TABLAS

### CAPÍTULO 1

Tabla 1.1: Tipos de Filtros en analítica de video en videovigilancia [78] .....	30
---	----

### CAPÍTULO 2

Tabla 2.1: Comparación entre los algoritmos Viola&Jones y la distancia de Hausdorff .....	70
---	----

### CAPÍTULO 3

Tabla 3.1: Ejemplos de Intensidad Luminosa [83] .....	76
Tabla 3.2: Detección frontal con poca Iluminación .....	88
Tabla 3.3: Resultados con enfoque a 45° con poca iluminación.....	89
Tabla 3.4: Detección de rostros girados a 45° con poca iluminación.....	89
Tabla 3.5: Detección de rostros girados a 90° con poca iluminación .....	90
Tabla 3.6: Detección frontal con iluminación normal .....	91
Tabla 3.7: Resultados con enfoque a 45° con iluminación normal .....	91
Tabla 3.8: Detección de rostros girados a 45° con iluminación normal.....	92
Tabla 3.9: Detección de rostros girados a 90° con iluminación normal.....	93
Tabla 3.10: Detección de rostro con facciones completas.....	94
Tabla 3.11: Detección de facciones de un rostro sin facciones completas de rostro .	94
Tabla 3.12: Detección de rostro blanco y negro .....	95
Tabla 3.13: Detección de rostro casi completo .....	95
Tabla 3.14: Detección de rostro confuso .....	96
Tabla 3.15: Detección frontal con poca Iluminación .....	100
Tabla 3.16: Resultados con enfoque a 45° con poca iluminación.....	100
Tabla 3.17: Detección de rostros girados a 45° con poca iluminación .....	101
Tabla 3.18: Detección de rostros girados a 90° con poca iluminación .....	101
Tabla 3.19: Detección frontal con iluminación normal .....	102
Tabla 3.20: Resultados con enfoque a 45° con iluminación normal .....	103
Tabla 3.21: Detección de rostros girados a 45° con iluminación normal.....	104
Tabla 3.22: Detección de rostros girados a 90° con iluminación normal.....	104

Tabla 3.23: Detección de rostro con facciones completas .....	105
Tabla 3.24: Detección de facciones de un rostro sin facciones completas de rostro .....	106
Tabla 3.25: Detección de rostro blanco y negro .....	106
Tabla 3.26: Detección de rostro casi completo .....	107
Tabla 3.27: Detección de rostro confuso .....	108
Tabla 3.28: Tabla de resultados 1 código desarrollado .....	109
Tabla 3.29: Tabla de resultados 2 código de Mathworks.....	110

## ÍNDICE DE ECUACIONES

Ecuación 1: Ecuación para calcular el valor de una característica .....	45
Ecuación 2: Imagen Integral (suma de píxeles).....	46
Ecuación 3: Suma Acumulativa de píxeles .....	46
Ecuación 4: Peso modificado .....	48
Ecuación 5: Clasificador débil .....	51
Ecuación 6: Clasificador débil con respecto al error ponderado .....	51
Ecuación 7: Clasificador Fuerte .....	52
Ecuación 8: Cálculo de la tasa de falsos positivos .....	56
Ecuación 9: Cálculo de la tasa de detección .....	56
Ecuación 10: Número de características .....	57
Ecuación 11: Distancia de Hausdorff .....	59
Ecuación 12: Distancia de Hausdorff general.....	61
Ecuación 13: Distancia de Hausdorff dirigida .....	64
Ecuación 14: Distancia de Hausdorff dirigida modificada .....	64
Ecuación 15: Formula detección utilizando la distancia de Hausdorff .....	65



## ÍNDICE DE CODIGOS

Código 1: Código para detección de rostros en tiempo real implementado .....	84
Código 2: Código para detección de rostros de la Herramienta comercial [5] .....	98

## RESUMEN

La implementación de sistemas de videovigilancia<sup>1</sup> con análisis de video proporciona mejores características de registros, almacenamiento y monitoreo de manera automática en muchos ambientes de implementación, permiten la supervisión de objetos y personas para prevenir y garantizar la seguridad. En la actualidad los sistemas de videovigilancia más usados son los que cuentan con tecnología IP.

La detección de rostros es una aplicación de los sistemas de videovigilancia y es el primer paso para otras aplicaciones posteriores como seguimiento y reconocimiento. Para la detección facial existen una serie de algoritmos que permiten extraer el rostro en una imagen de video, entre los cuales se encuentran el método de cascada de Viola&Jones y el método por modelos geométricos usando la distancia de Hausdorff.

El algoritmo de Viola&Jones divide la imagen integral en subregiones de tamaños diferentes y utiliza una serie de clasificadores mientras que el método de Hausdorff está basado en bordes y utiliza la distancia como medida de similitud entre un modelo de rostro general. Por su fiabilidad y eficiencia el algoritmo más utilizado es el desarrollado por Viola&Jones.

El programa Matlab y su conjunto de herramientas para cálculos matemáticos y visualización gráfica permiten realizar una simulación del código del algoritmo en interiores en un ambiente controlado para la detección de rostros en tiempo real por el método de Viola&Jones y así obtener una serie de datos para comprobar la efectividad del sistema.

---

<sup>1</sup> Vigilancia por medio de un sistema de cámaras, fijas o móviles.

## PRESENTACIÓN

En el presente trabajo se realiza el análisis de dos algoritmos para detección de rostros en tiempo real y se realiza la implementación de uno de ellos utilizado en analítica de video para sistemas de videovigilancia.

En el capítulo 1 se describe de manera detallada la importancia de los sistemas de videovigilancia, su evolución a través de los años hasta la tecnología IP actualmente utilizada. Se detallan los componentes que lo conforman, aplicaciones y ambientes de operación. Se describen los fundamentos de analítica de video, visión artificial, segmentación y formatos de procesamiento de la imagen.

En el capítulo 2 se analizan dos algoritmos utilizados para la detección de rostros. El primer algoritmo que se analiza está basado en modelos geométricos usando la distancia de Hausdorff de Jesorsky, Kirchberg, Frischholz. El segundo algoritmo que se analiza está basado en clasificadores en cascada de Viola&Jones. Se estudian los principios que utilizan, así como su estructura y funcionamiento para determinar el algoritmo de mejores características.

En el capítulo 3, se realiza la implementación del código de mejores características encontrado en el capítulo anterior, en un ambiente de simulación en interiores controlado para detección de rostros en tiempo real en el programa Matlab y se realiza un análisis del mismo. Además, se realiza un análisis comparativo del algoritmo implementado con una herramienta comercial disponible ya desarrollada en Matlab.

Finalmente, en el capítulo 4 se presentan las conclusiones y recomendaciones del trabajo realizado.

# **CAPÍTULO 1**

## **FUNDAMENTOS TEÓRICOS DE LOS SISTEMAS DE VIDEOVIGILANCIA**

### **1.1 INTRODUCCIÓN**

Este capítulo tiene como objetivo describir de manera detallada la importancia de los sistemas de videovigilancia, su evolución a través de los años, desde los sistemas analógicos hasta la tecnología IP.

Adicionalmente se describen los componentes que lo conforman y las funciones de cada uno de los dispositivos.

Finalmente, el capítulo concluye con una breve descripción de los ambientes de aplicación y operación de los sistemas de videovigilancia en la actualidad.

### **1.2 GENERALIDADES**

Se puede decir que la tecnología avanza a pasos gigantes en *hardware*, *software*, aplicaciones y funcionalidades, así como en la cantidad de datos y metadatos que se generan diariamente. Muchos de estos datos son imperceptibles al humano que no sabe que se han generado o que están ahí por sus actividades o acciones, la automatización de los sistemas con los que se convive hacen indispensable contar con tecnologías y aplicaciones que detecten, organicen y muestren estos datos en un lenguaje más amigable para el entendimiento humano.

Gran parte de la información que se utiliza en la actualidad está en forma digital y es presentada en videos e imágenes, convirtiendo esta práctica en una de las mejores técnicas de aprendizaje, conocimiento, detección e incluso decisión de eventos o acciones.

Los videos e imágenes son los principales productos considerados información o fuente de datos en videovigilancia que deben ser revisados, administrados y almacenados de forma adecuada con el objetivo de analizar eventos en tiempo real. Los sistemas de videovigilancia en la actualidad permiten precautelar la seguridad de personas y objetos en sus actividades diarias y en varios entornos como centros comerciales, bancos, estadios deportivos, tiendas, etc. [4].

El contar con videos no ha garantizado una vigilancia completa en tiempo real por los errores o vulnerabilidades del sistema al depender de un recurso humano que identifique de forma continua los eventos, las 24 horas del día sujeto a cansancio y fatiga. Por esta problemática las investigaciones se enfocaron en sistemas de videovigilancia más eficientes que almacenen automáticamente la información de video y que generen reportes de eventualidades en tiempo real [2]. Además, un sistema de videovigilancia automatizado permite la administración local y/o remota del sistema [68].

Por esta razón, existen en el mercado sistemas de videovigilancia con la capacidad de realizar la generación de alarmas inteligentes, reporte de intrusos, movimiento de objetos, ente otros. También existen aplicaciones que permiten al usuario detectar movimientos sospechosos, detectar intrusos en áreas de acceso restringido, seguimiento de comportamientos sospechosos o anómalos, identificación a diferentes distancias de rostros u objetos, [4].

Estas aplicaciones han ganado el mercado de los sistemas de videovigilancia por lo que es importante la revisión de los algoritmos que se utilizan para conocer cómo se realiza la detección de rostros.

### **1.3 VIDEOVIGILANCIA**

La videovigilancia es la actividad que consiste en la colocación de cámaras de grabación para la captación de imágenes y/o sonidos con la finalidad de garantizar la

seguridad de una instalación o de las personas [87] y prevenir, averiguar o esclarecer hechos delictivos [23].

## 1.4 EVOLUCIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA

Brevemente se revisará una reseña de los sistemas de videovigilancia. Las cámaras de video comenzaron a ser utilizadas como un elemento de seguridad de la preparación militar. El primer uso de un circuito cerrado de televisión CCTV por sus siglas en inglés *Closed Circuit Television* se produjo en el año 1942. Los primeros sistemas se componían únicamente de cámaras en blanco y negro conectadas a monitores que eran importantes para la observación de ensayos de misiles [48].

Posteriormente, aparecieron los generadores de cuadrantes con el fin de poder visualizar varias cámaras a la vez en el mismo monitor. Los primeros servidores de grabación utilizaban cintas de video para luego ser reemplazadas por los discos digitales. Paralelamente las cámaras también evolucionaron, dejando de lado las monocromáticas<sup>2</sup> para utilizar gamas de color, mejorando las resoluciones [48].

Los sistemas de videovigilancia actuales distan mucho de los primeros sistemas instalados. Al principio los sistemas de cámaras CCTV, se trataban de circuitos cerrados de televisión sin la posibilidad de acceder remotamente. Con los avances en telecomunicaciones, aparecen los primeros grabadores de video con tarjeta de red, los cuales permiten visualizar las cámaras desde varios dispositivos externos a la instalación del circuito.

La tecnología analógica fue pionera en popularizar la videovigilancia para diversos entornos como tiendas, bancos, centros comerciales, entre otros. Sin embargo,

---

<sup>2</sup> Cámaras que en su momento solo tenían la capacidad de mostrar un solo color

desde los años 90 se introdujo la transmisión digital, dando origen a los sistemas de videovigilancia IP (*Internet Protocol*).

Actualmente, los dispositivos de conexión remota se incrementaron con la aparición de *laptops*, *smartphones*, *tablets* y demás dispositivos móviles y el acceso a Internet ha permitido el acceso remoto a los sistemas de videovigilancia desde cualquier lugar del mundo.

A continuación, se describen los dos tipos de sistemas de videovigilancia que existen [50]:

- 1) Sistemas de videovigilancia analógicos.
- 2) Sistemas de videovigilancia digitales sobre IP.

#### **1.4.1 SISTEMAS DE VIDEOVIGILANCIA ANALÓGICOS**

Los sistemas de circuito cerrado de televisión CCTV, fueron los más utilizados en la última década. La imagen obtenida por estos sistemas es analógica, es decir, una señal con amplitud y período variable en el tiempo. La interfaz de salida está conformada por un conector BNC (*Bayonet Neill-Concelman*) para cable coaxial con impedancia de 75 ohm, por lo que utilizan el mismo cable coaxial de 75 ohm que un sistema de televisión [49].

Estos sistemas en la actualidad han sido desplazados por los sistemas digitales debido a que el usuario prefiere las funcionalidades de un sistema IP, por ejemplo, el monitoreo remoto a través de Internet. Sin embargo, los sistemas de videovigilancia analógicos también pueden ser monitoreados remotamente a través de Internet con sus componentes de red o DVR (*Digital Video Recorder*) [49].

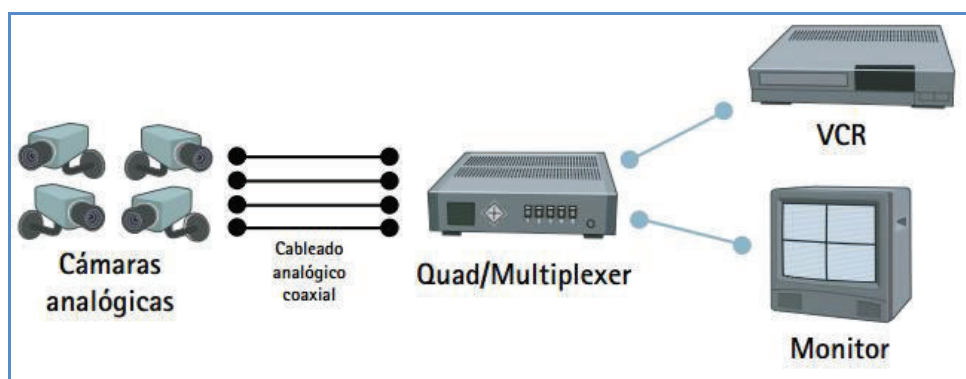
Cuando se habla de un sistema de videovigilancia analógico con DVR, no se accede directamente a la cámara de seguridad desde Internet, sino que se accede al DVR y éste muestra las imágenes de las cámaras de seguridad [49].

Los sistemas analógicos pueden estar conformados por diferentes dispositivos que les permitan ofrecer nuevas aplicaciones y funciones conforme se describe a continuación.

#### 1.4.1.1 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN VCR

Es un sistema de videovigilancia analógico tradicional formado por cámaras analógicas con salida de cable coaxial conectada al VCR (*Video Cassette Recorder*) para grabar videos.

En sistemas mayores, se puede conectar un *quad* o un multiplexor<sup>3</sup> entre la cámara y el VCR. El quad/multiplexor permite grabar el video procedente de varias cámaras en un solo grabador<sup>4</sup>. Para monitorizar el video es necesario un monitor analógico como se muestra en la Figura 1.1.



**Figura 1.1:** Sistema de CCTV tradicional con VCR. [50]

#### 1.4.1.2 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN DVR

Es un sistema de vigilancia con cámaras analógicas conectadas a un DVR, el cual incluye un multiplexor y suministra grabación digital como se muestra en la Figura 1.2.

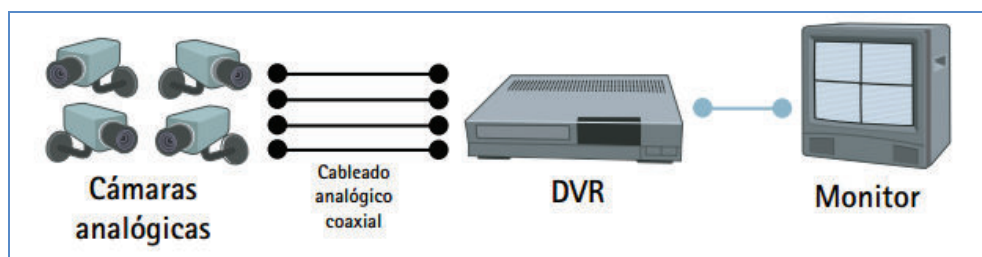
<sup>3</sup> Equipo que cuenta con varias entradas y una única salida de datos.

<sup>4</sup> Dispositivo de grabación de vídeo en formato digital.



La introducción del sistema DVR ofrece ventajas como:

- Funcionamiento sin cintas ni cambios de cinta.
- Calidad de grabación constante.
- Capacidad de buscar información con rapidez en el video grabado.



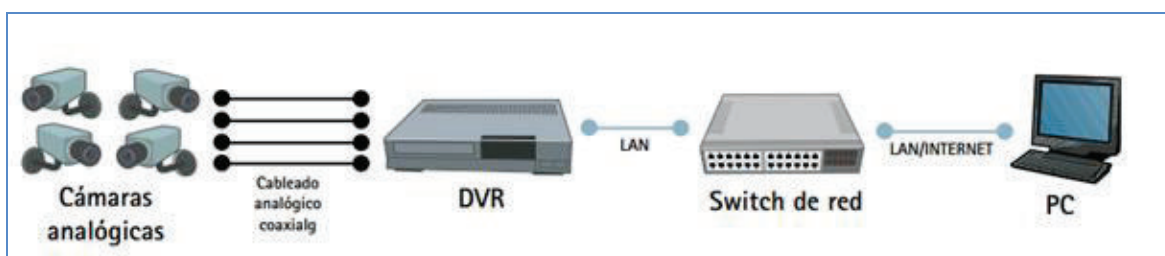
**Figura 1.2:** Sistema de CCTV con DVR [50]

#### 1.4.1.3 SISTEMAS ANALÓGICOS DE CCTV BASADOS EN DVR DE RED

Este sistema permite conectar en red las cámaras analógicas mediante un DVR de red para monitorizar de forma remota el video grabado o en directo como se muestra en la Figura 1.3.

El sistema DVR de red ofrece las siguientes ventajas:

- Monitorización remota del video a través de un computador.
- Funcionamiento remoto del sistema.



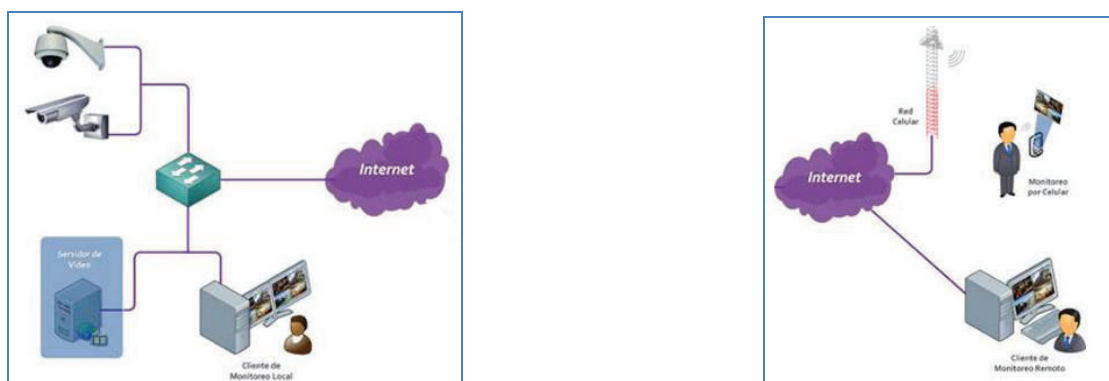
**Figura 1.3:** Sistema de CCTV conectado a la Red [50]

A continuación, se describen los sistemas de videovigilancia IP de igual forma como se abordaron los sistemas de videovigilancia analógicos.

#### 1.4.2 SISTEMAS DE VIDEOVIGILANCIA IP

Los sistemas de videovigilancia IP aprovechan la red informática sin necesidad de desplegar una infraestructura de cableado coaxial específica para la red de videovigilancia como la analógica, es decir, se utiliza el mismo cableado que se emplea para la comunicación de datos. Las instalaciones modernas están optando por una red IP convergente que brinde voz, datos y video en una sola infraestructura, con esto están abandonando la tecnología analógica en favor de la videovigilancia IP [51].

Un sistema de videovigilancia IP puede estar instalado en una red de área local LAN (*Local Area Network*) donde el sistema se administra de forma interna, así como también en una red de área extendida WAN (*Wide Area Network*) permitiendo el acceso remoto desde redes externas como se muestra en las Figuras 1.4a) y 1.4b).



**Figura 1.4:** Sistema de videovigilancia IP. a) Red LAN. b) Red WAN [48]

En la actualidad trabajar con redes digitales en los sistemas de videovigilancia tiene beneficios los cuales se describen a continuación [50]:

- Integración del video y audio del sistema de videovigilancia a la red de datos.

- Flexibilidad del sistema de videovigilancia para incrementar el número de cámaras.
- Escalabilidad del sistema a versiones superiores que soporten mayor número de licencias para administrar las cámaras.
- Compartición de recursos de almacenamiento como servidores de almacenamiento y *firewall* de una red.
- Monitoreo local y remoto desde cualquier punto de acceso que se proporcione al sistema de videovigilancia.
- Reducción de costos por compartición de infraestructura de red.

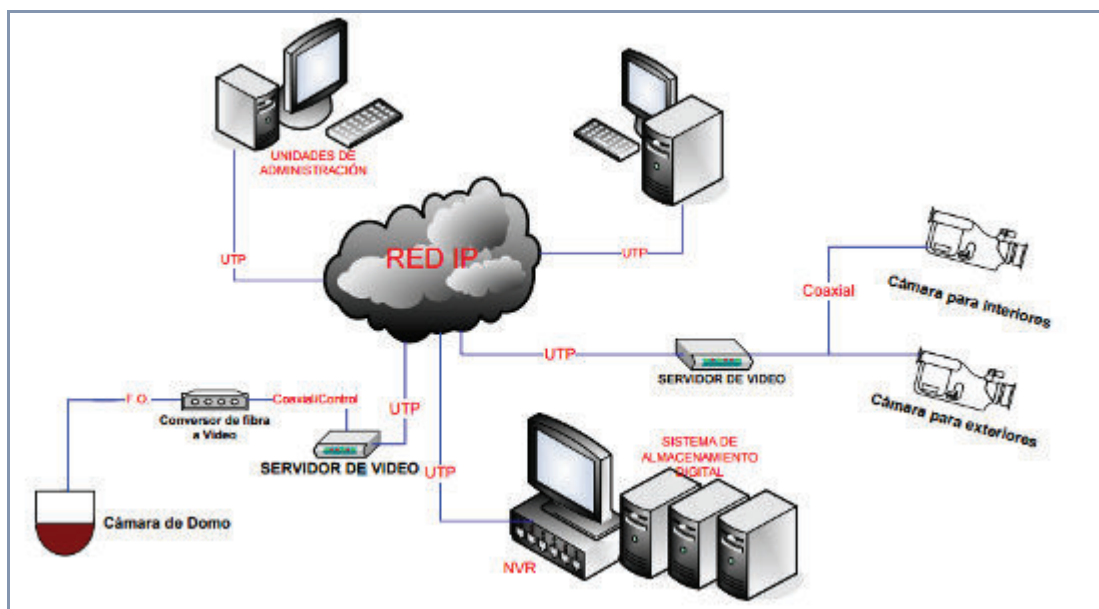
Todos los componentes de un sistema de videovigilancia tienen gran importancia y cada uno de ellos desempeña funciones específicas dentro de su estructura como receptar, procesar, comunicar, almacenar las imágenes. A continuación, se describen cada uno de ellos.

## **1.5 COMPONENTES DE LOS SISTEMAS DE VIDEOVIGILANCIA**

Un sistema de videovigilancia IP consta de los siguientes componentes [68]:

- Monitor.
- Video Grabador NVR (*Network Video Recorder*).
- Servidor de Video.
- Conmutador de Red.
- Conjunto de Cámaras.

En la Figura 1.5 se puede observar la ubicación de los elementos que conforman el sistema de videovigilancia.



**Figura 1.5:** Esquema de sistema de videovigilancia [80]

Los elementos del sistema de videovigilancia permiten el procesamiento de la información y se describen con mayor detalle a continuación.

### 1.5.1 MONITOR

El monitor es un dispositivo de salida que reproduce en su pantalla las imágenes capturadas por las cámaras. Un monitor dispone de varios aspectos que deben ser considerados, según la aplicación en la que se vaya a utilizar, como resolución máxima, tamaño de la pantalla, tipo de monitor, alimentación eléctrica, entre otros.

En la Figura 1.6 se muestran otros periféricos de salida que pueden ser utilizados para monitorizar un sistema de videovigilancia como son las *laptops*, celulares, *tablets*, entre otros.



**Figura 1.6:** Dispositivos de monitoreo [86]

### 1.5.2 VIDEO GRABADOR NVR

El grabador de cámaras IP NVR permite la grabación simultánea y continua de video de las cámaras IP, liberando automáticamente el espacio disponible. Como soporte de grabación, admite un disco duro interno, disco externo o memorias flash USB.

El estándar ONVIF (*Open Network Video Interface Forum*) es un modelo universal de compatibilidad de dispositivos IP para permitir el uso de cámaras IP de diferentes fabricantes.

El NVR no necesita dedicar un computador para la grabación y visualización de sus cámaras de vigilancia IP, pues puede manejarse desde una televisión, monitor o mediante un navegador en un computador (local o remoto) [70].

### 1.5.3 SERVIDOR DE VIDEO

Un servidor de video es un dispositivo que permite la transición tecnológica entre los sistemas analógicos de vigilancia conocidos como CCTV y las nuevas formas de vigilancia conocidas como vigilancia IP [71].

Un servidor para video está conectado a una red de datos y puede ofrecer video de forma automática o bajo petición desde un navegador web a otras aplicaciones profesionales de seguridad [72].

#### 1.5.4 CONMUTADOR DE RED

Es el dispositivo de interconexión de equipos que opera en la capa de enlace de datos del modelo OSI (*Open System Interconnection*). Su función es interconectar dos o más segmentos de red, mejorar el rendimiento y la seguridad de las redes de área local (LAN). Se usan en los sistemas de videovigilancia para agrupar varias cámaras IP en el mismo segmento de red [73].

#### 1.5.5 CONJUNTO DE CÁMARAS

Es un grupo de dispositivos analógicos o digitales de diferentes modelos y tipos de funcionamiento para registrar imágenes y sonido que se encuentran distribuidos en áreas específicas para monitoreo.

A continuación, se detalla de manera más específica la definición, funcionamiento y componentes de las cámaras de video.

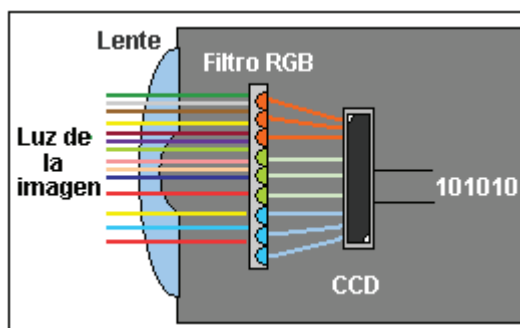
##### 1.5.5.1 Definición de cámara de video IP

Una cámara de video IP es un dispositivo que captura imágenes convirtiéndolas en señales eléctricas denominadas señales de video. Es por esto que se las considera como transductores que captan señales o imágenes ópticas visibles y las convierten en eléctricas para ser transmitidas por los medios guiados o no guiados hasta la unidad de control y monitoreo [38].

En una cámara de video, la luz de la imagen pasa por el lente y se refleja en un filtro RGB (*Red-Green-Blue*) el cual descompone la luz en tres colores básicos; rojo, verde y azul. Esta división de rayos se concentra en un sensor sensible a la luz, el cual asigna valores a cada píxel<sup>5</sup> y envía los datos analógicos o digitales para su codificación [39] como se muestra en la Figura 1.7.

---

<sup>5</sup> Superficie homogénea más pequeña de las que componen una imagen, que se define por su brillo y color.



**Figura 1.7:** Funcionamiento básico interno de una cámara de video IP [39]

Las cámaras de video son utilizadas para ver lo que está pasando en un momento determinado en un sitio indicado. Por tal motivo, es importante señalar que la ubicación de las cámaras dependerá del control y seguimiento que el usuario requiere [40].

#### 1.5.5.2 Funcionamiento de una cámara de videovigilancia IP

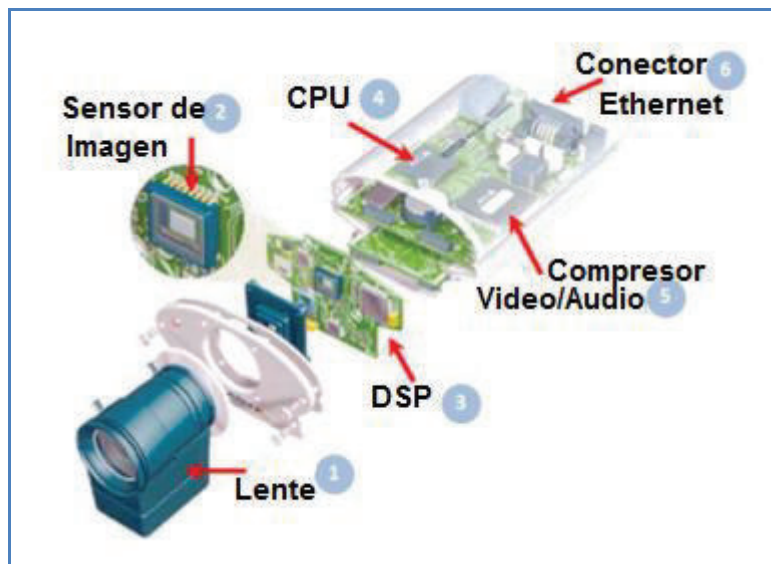
Las cámaras IP poseen un lente que enfoca la imagen en el sensor de imágenes. Antes de llegar al sensor, la imagen pasa a través de un filtro óptico, que elimina cualquier luz infrarroja, para que los colores mostrados sean correctos. El sensor de imagen convierte la imagen, compuesta por información lumínica, en señales eléctricas. Estas señales eléctricas digitales están ya en un formato que puede comprimirse y enviarse a través de la red. La CPU (*Central Processing Unit*) representa las funciones de procesamiento de la cámara y están diseñadas específicamente para aplicaciones de red [73].

El esquema de funcionamiento y componentes de una cámara IP se muestran en la Figura 1.8.

#### 1.5.5.3 Componentes de una cámara IP

Una cámara IP está compuesta por un lente, un sensor de imagen, un procesador de imagen DSP (*Digital Signal Processor*), un CPU, la etapa de compresión y una tarjeta Ethernet como se puede observar en la Figura 1.8.

La mayoría de las cámaras IP actuales incluyen una memoria interna que normalmente es una tarjeta microSD para almacenar los videos [40].



**Figura 1.8:** Esquema básico de una cámara IP [52]

Se describirá en la siguiente sección cada uno de los componentes de una cámara IP.

#### *1.5.5.3.1 Lente*

Los lentes son los “ojos” de un sistema de videovigilancia y su función es determinar la escena que se muestra en el monitor, dependen directamente de la distancia o longitud focal y controlan la cantidad de luz que llega al sensor (iris).

La distancia focal, como se puede ver en la Figura 1.9, es la distancia entre el centro del lente y el sensor de imagen y se mide en milímetros. Los rayos de objetos infinitamente distantes se condensan internamente en el lente en un punto común en el eje óptico. El punto en el que se coloca el sensor de imagen de la cámara se llama punto focal. Las ópticas con distancia focal pequeñas tienen un ángulo de apertura grande lo que permite observar zonas extensas. Un objetivo de 50 mm equivaldría al ángulo de visión humano [50].

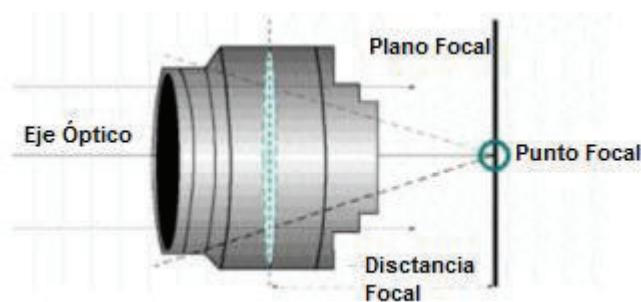


Según la distancia focal los lentes se clasifican en lentes fijos y lentes vari-focales:

- Los lentes fijos son el tipo más simple de lente, y por lo tanto, menos costosos.
- Los lentes de distancia focal variable (vari-focal), aunque son más costosos son los más usados porque se puede conseguir un ajuste más preciso de la escena haciendo que el sistema de videovigilancia sea más flexible.

Para encontrar el valor fijo del lente se requiere un cálculo preciso para seleccionar el lente más adecuado para una escena determinada. Este cálculo se basa en conocer el tamaño deseado del área de visualización y la distancia a la cámara.

Se puede ajustar el valor de la distancia focal de forma remota, esto significa que un solo lente se puede utilizar para ver una amplia zona hasta que se detecta un intruso y en ese mismo momento hacer un zoom para capturar detalles faciales.

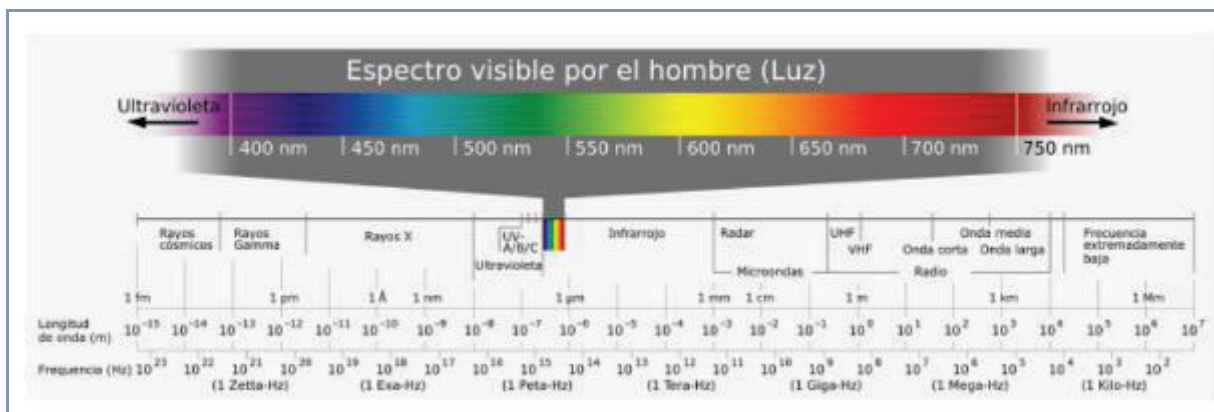


**Figura 1.9:** Distancia focal de un lente IP [52]

Otra característica de los lentes es la corrección infrarroja (IR), el ojo humano es capaz de ver únicamente la parte de "luz visible" del espectro mostrado en la Figura 1.10, más allá de la luz visible está la porción del espectro infrarrojo.

Para compensar la falta de iluminación muchas cámaras llevan incorporados *leds* (*light-emitting diode*) de iluminación infrarroja e incluso focos térmicos. En el caso de que la cámara no lleve incorporada iluminación se pueden utilizar focos de iluminación adicional. Por lo general son de iluminación infrarroja, dependiendo del

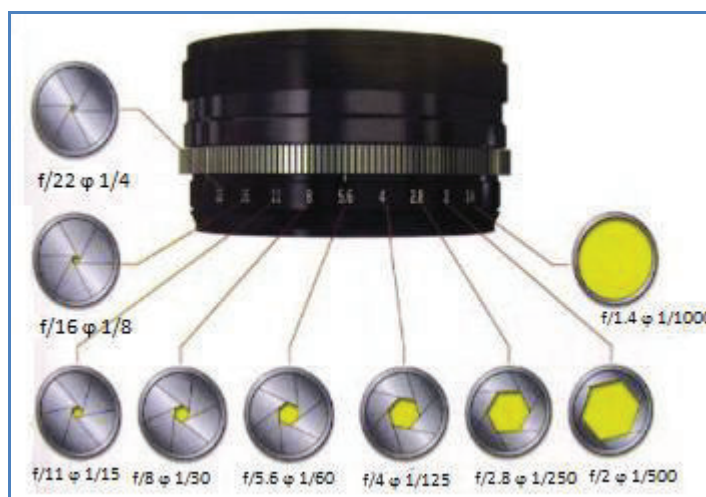
modelo pueden iluminar la escena desde 10 m hasta 350 m y con un ángulo de apertura entre 3° y 120° [52].



**Figura 1.10:** Localización en el espectro radioeléctrico de la parte visible por el hombre (luz) [52]

La luz IR afecta negativamente a la exactitud de la reproducción del color, por esta razón todas las cámaras a color llevan incluido un filtro de corrección IR para minimizar o eliminar la luz IR que llega al sensor de imagen [52].

Otro concepto relacionado con los lentes es el iris que controla la cantidad de luz que incide sobre la cara del sensor de imagen. La medida de la apertura del iris se hace en pasos discretos conocidos como *f-stops* como se puede ver en la Figura 1.11



**Figura 1.11:** Apertura del iris [52]

#### *1.5.5.3.2 Sensor de imagen*

Existen dos tipos de tecnologías utilizadas para la fabricación de sensores para las cámaras. Se trata de los dispositivos de carga acoplada CCD (*Charge-Coupled Device*) o semiconductores complementarios de óxido metálico CMOS (*Complementary Metal-Oxide-Semiconductor*). Ambos tipos de sensores están formados en su esencia por semiconductores de metal-óxido (MOS) y están distribuidos en forma de matriz. Su función es la de acumular una carga eléctrica en cada una de las celdas de esta matriz. Estas celdas son los llamados píxeles. La carga eléctrica almacenada en cada píxel, dependerá en todo momento de la cantidad de luz que incida sobre el mismo. Cuanta más luz incida sobre el píxel, mayor será la carga que éste adquiera [52].

#### *1.5.5.3.3 Procesador de imagen DSP*

La función principal de un procesador es recibir la imagen digitalizada por parte del sensor y después procesarla para enviarla a la etapa de compresión. La calidad de una imagen proporcionada por el sensor se puede mejorar gracias al procesador de imagen que puede ajustar o aplicar diferentes técnicas y parámetros para conseguir mejor control del tiempo de exposición, iris y ganancia, compensación de luz de fondo, rango dinámico, algoritmos de mosaico, reducción de ruido y procesamiento del color [52].

#### *1.5.5.3.4 CPU*

La unidad de procesamiento central CPU de una cámara IP es una tarjeta electrónica basada en algún sistema operativo o lenguaje de alto nivel que controla y administra todas las funciones de la cámara, como por ejemplo la compresión, envío de imágenes, gestión de alarmas y avisos, entre otros [52].

#### *1.5.5.3.5 Compresión*

La compresión resulta imprescindible para la transmisión de imágenes y video a través de una red IP ya que permite disminuir el tamaño de una imagen digital para una transferencia más eficiente a través de la red [73].

La transmisión masiva de datos de video sin comprimir ocasionaría saturación de la red, por esto desde la aparición de las redes de datos se han creado algoritmos que procesan la señal para eliminar redundancias y para aplicar filtros que reducen el tamaño de la imagen a costa de deteriorar su calidad de forma mínima. Los estándares de compresión más utilizados en las cámaras IP son; MJPEG, MPEG-4, H.264 y H.265 [52].

- MJPEG es un estándar en el que cada fotograma es comprimido como una imagen JPEG [52].
- MPEG-4 es un conjunto de 27 estándares y protocolos usados para codificación y transmisión de flujos de video/audio en entornos de reducido ancho de banda (hasta 1,5 Mbps). Es el primer gran estándar en la transmisión de videos por redes IP y es usado también en dispositivos móviles y en televisión [52].
- H.264 también conocido como MPEG-4 Parte 10 se trata del estándar para la compresión de vídeo digital el cual ofrece una mayor resolución de vídeo que MJPEG o MPEG-4 a la misma velocidad de bits y el mismo ancho de banda o visto desde otra forma la misma calidad de vídeo con una velocidad de bits inferior [52].
- H.265 o HEVC *High Efficiency Video Coding* se trata del estándar para la compresión de vídeo más eficiente que H.264. Aproximadamente es capaz de comprimir un archivo de la misma calidad en la mitad de espacio que H.264 [52].

#### *1.5.5.3.6 Tarjeta Ethernet*

La tarjeta o interfaz Ethernet de la cámara IP es la encargada de ofrecer conectividad de red para poder transmitir las imágenes capturadas a través de la red IP [52].

En la siguiente sección se estudiarán los diferentes tipos de cámaras entre los cuales se tienen interiores, exteriores, entre otros.

#### **1.5.5.4 Tipos de cámaras**

Existen una variedad de cámaras que cumplen con diferentes requerimientos y su clasificación depende de varios factores como; el entorno donde serán instaladas, almacenamiento de las imágenes (analógico o digital), tamaño del sensor o película, grado de automatismo, entre otros.

A continuación, se detallan algunos tipos de cámaras existentes [42]:

- **Cámaras Interiores.**- Las cámaras más sencillas que se pueden encontrar son las destinadas a áreas interiores. No tienen la necesidad de una carcasa o visión nocturna ya que suele haber iluminación permanente durante las horas que se necesita supervisión.
- **Cámaras Exteriores.**- Las cámaras con más complejidad por el ambiente externo donde se desenvuelven, tienen la necesidad de una carcasa y visión nocturna ya que no suele haber iluminación permanente durante las 24 horas de supervisión.
- **Cámaras con Infrarrojos.**- Si la cámara va a estar colocada en un lugar con poca iluminación, o se necesita vigilancia 24 horas, la mejor opción es colocar cámaras con visión nocturna. Estas cámaras graban el video durante el día a color y cuando existe poca iluminación encienden de forma automática sus infrarrojos para seguir grabando en blanco y negro.
- **Cámaras Anti-vandálicas.**- Las zonas transitadas por muchas personas o locales que son especialmente vulnerables a robos y agresiones son los indicados para las cámaras anti-vandálicas. Estas cámaras, de forma obligatoria, cuentan con una carcasa resistente a golpes y se mantienen fijas

para seguir grabando todo lo que ocurre. Perfectas para parqueaderos, almacenes, discotecas, bares o exteriores de tiendas, entre otros.

- Cámaras con Movimiento y Zoom.- Las cámaras con zoom y movimiento permiten ampliar o reducir la imagen y son idóneas para instalaciones de videovigilancia que tienen a una persona monitorizando las cámaras o para grandes superficies que se vigilan siguiendo una ruta de movimiento.
- Cámaras Ocultas.- Si se necesita vigilar con total discreción algún lugar de su casa o negocio se recomienda cámaras espías. Estas cámaras se colocan dentro de algún objeto (detectores de humo, sensores de movimiento, espejos, tornillos, enchufes, entre otros) y pasan 100% desapercibidas a todas las personas que pasen por delante. Sin embargo, se recomienda respetar la privacidad de las personas.
- Cámaras de día/noche.- Están elaboradas para compensar las condiciones variables de luz para que la cámara pueda capturar imágenes. Se utilizan principalmente en aplicaciones al aire libre. En muchos casos las unidades tienen un amplio rango dinámico para funcionar en el resplandor de luz.
- Cámaras varifocales.- Una cámara con un lente de distancia focal variable permite al operador ampliar o reducir el ángulo de visión manteniendo el foco en la imagen.
- Cámaras inalámbricas.- Su principal funcionalidad es la transmisión de datos por el espectro electromagnético, lo cual permite una gran facilidad en la conexión hasta algún punto de control y administración del sistema de videovigilancia.
- Cámaras de alta definición.- Estas cámaras dan a los operadores la posibilidad de ampliar la imagen con extrema claridad.

A continuación, se describirá los tipos de conexiones existentes para los sistemas de videovigilancia.

### **1.5.6 CONEXIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA**

Los sistemas de videovigilancia dependen directamente del medio físico que conecta la cámara con los dispositivos de monitoreo. Los medios de comunicación se clasifican en guiados y no guiados, en el siguiente apartado se describirá cada uno de ellos.

#### **1.5.6.1 Medios guiados**

Los medios de transmisión guiados están constituidos por cables que se encargan de la conducción de las señales desde un extremo al otro [75].

Los medios físicos o cables se utilizan para crear enlaces de comunicaciones, en donde la información se propaga por el medio de transmisión de un dispositivo a otro.

#### **1.5.6.2 Medios no guiados**

En este tipo de medios, la transmisión y la recepción de información se lleva a cabo a través del espectro electromagnético mediante el uso de antenas [75].

Las redes inalámbricas son una de las tecnologías más utilizadas para comunicar dispositivos de forma libre.

En la siguiente sección se estudiarán los ambientes de instalación de los sistemas de videovigilancia.

## **1.6 AMBIENTES DE INSTALACIÓN DE LOS SISTEMAS DE VIDEOVIGILANCIA**

Los sistemas de videovigilancia se implementan en diferentes ambientes y se requiere de una ingeniería de diseño para su óptimo funcionamiento.

Se considera que los sistemas de videovigilancia que no requieren de un sistema muy exigente en seguridad son locales comerciales, universidades, restaurantes,

hogares, en los cuales se puede utilizar soluciones analógicas o IPs, pero en ambientes más rigurosos como puertos, aeropuertos, estaciones de ferrocarril, metro, autobús, polígonos industriales, centros penitenciarios, museos, edificios públicos, hospitales, entre otros, se recomiendan sistemas de videovigilancia IP por sus mejores características [4].

Es importante considerar también otros factores que determinan los entornos de instalación de los sistemas de videovigilancia como son el área a ser cubierta, la infraestructura, los horarios en lo que se requiere la vigilancia, las adversidades climatológicas, protecciones anti-vandálicas, apuntamiento, movilidad, entre otros.

Los sistemas de videovigilancia están directamente relacionados con la analítica de video, ya que ésta permite analizar señales de video, extraer y procesar datos específicos, comportamientos o actitudes para generar una reacción de acuerdo a las alarmas entregadas por el sistema [76]. A continuación, se trata la analítica de video con mayor detenimiento.

## 1.7 ANALÍTICA DE VIDEO

El análisis de video o analítica de vídeo es una tecnología que combina *software* de análisis con un sistema de videovigilancia elevando las prestaciones de un sistema convencional<sup>6</sup> de visualización de cámaras de seguridad [76].

La analítica de vídeo ofrece muchas otras ventajas, por ejemplo, previene la transmisión de gran cantidad de datos de vídeo, que contienen poca información o sin procesar, consumiendo recursos de red y de energía de cómputo además de altos costos en medios de almacenaje y discos duros [78].

---

<sup>6</sup> Dicho de un acto, de una costumbre, de una indumentaria, etc.: Que se atienen a las normas mayoritariamente observadas.



### **1.7.1 INICIOS DE LA ANALÍTICA DE VIDEO**

La analítica de video se remonta a los años 60's cuando se empieza a extraer información útil de las imágenes de video, convirtiendo el video en datos. La tecnología detrás del análisis de video es la visión artificial [77].

La visión artificial es una rama de la inteligencia artificial o ciencia de razonamiento de las máquinas, que tiene por objetivo modelar matemáticamente los procesos de percepción visual en los seres vivos y generar programas que permitan simular estas capacidades visuales mediante el uso de un computador. La visión artificial permite la detección automática de la estructura y propiedades de un posible mundo dinámico en 3 dimensiones a partir una o varias imágenes bidimensionales del mundo. Las imágenes pueden ser monocromáticas o a color; pueden ser capturadas por una o varias cámaras y cada cámara puede ser estacionaria o móvil. La estructura y propiedades del mundo tridimensional que se intentan deducir en la visión artificial incluyen no sólo propiedades geométricas (tamaños, formas, localización de objetos, etc.), sino también propiedades del material (sus colores, sus texturas, la composición, etc.) y la luminosidad u oscuridad de las superficies [66].

Los sistemas de percepción computacional, como también se conoce a la visión artificial, van más allá de medir o detectar, estos sistemas perciben, descifran o reconocen el mensaje sensorial. La información visual es una proyección bidimensional de objetos tridimensionales y por tanto la imagen que capta el ojo humano o una cámara digital tiene infinitas interpretaciones posibles. La percepción es un proceso que se distribuye a lo largo del espacio y del tiempo [66].

La visión artificial pretende capturar la información visual del entorno físico para extraer características relevantes utilizando procedimientos automáticos [66].

Los sistemas de visión artificial pueden hacer cosas de manera mucho más efectivas y adecuadas que la visión humana, como por ejemplo:

- Dentro del espectro electromagnético la visión humana solamente capta un pequeño rango de frecuencias y amplitudes conocido como “rango de luz visible” los sistemas de visión artificial pueden captar todo el espectro electromagnético.
- La velocidad de respuesta de la visión humana es de 0,06 segundos mientras que en las cámaras de estado sólido es de 0,01 milisegundos y este tiempo va disminuyendo según se mejora la electrónica de estos dispositivos.
- Las cámaras de estado sólido están compuestas por sensores de imagen de estado sólido CCD (*Charge Coupled Device*) lo que permite que sean livianas y de tamaño reducido.
- La visión humana se cansa y se ve afectada por las emociones, no sucede así en los sistemas de visión artificial.
- El ser humano puede discernir entre 10 ó 20 niveles de gris mientras que los sistemas de visión artificial tienen una definición muy superior.
- La visión humana tiene muy poca precisión apoyándose en instrumentos de medida, los sistemas de visión artificial tienen gran precisión en la medición dependiendo solamente de la resolución espacial de los componentes del sistema.
- La visión humana es poco consistente por la fatiga y las distracciones, en cambio la visión artificial mantiene su nivel de rendimiento constante a lo largo de su vida útil.
- Los sistemas de visión artificial pueden trabajar en entornos muy peligrosos con riesgos radioactivos, químicos, biológicos, ruido, polución, temperaturas muy altas y muy bajas.

Así también se pueden señalar ventajas de la visión humana frente a la visión artificial como mejor reconocimiento de objetos, mejor adaptación a situaciones imprevistas, mejor procesamiento de tareas de alto nivel y mejor resolución, 576

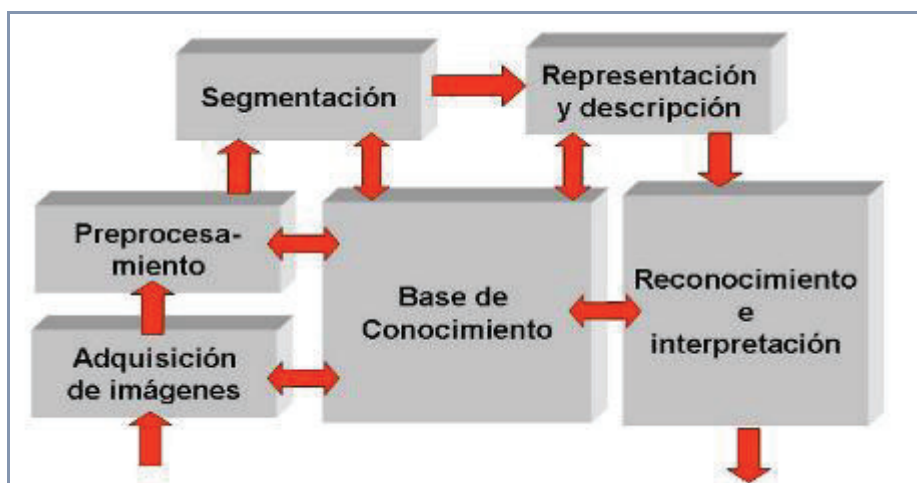
megapíxeles<sup>7</sup> frente a los 50 megapíxeles de una cámara. Por tanto, la disciplina de la visión artificial abarca enfoques sofisticados sobre la representación del mundo a partir de la manera en que lo ve el lente de una cámara de video [64].

A continuación, se detalla el funcionamiento de la visión artificial.

### 1.7.2 COMO FUNCIONA LA VISIÓN ARTIFICIAL

La analítica de video es una tecnología de visión artificial que se utiliza para analizar videos de datos específicos por ejemplo comportamiento, objetos o actitudes.

El proceso de un sistema de visión artificial se representa de acuerdo a la Figura 1.12 donde se muestra las etapas de adquisición, pre-procesado, segmentación, representación y reconocimiento de una imagen.



**Figura 1.12:** Etapas de una aplicación de visión artificial [66]

A continuación, se detalla cada una de las etapas de la visión artificial.

<sup>7</sup> Superficie homogénea más pequeña de las que componen una imagen, que se define por su brillo y color.

### 1.7.2.1 Adquisición de la imagen

El primer paso es adquirir la imagen digital. Para ello se necesita captar el objeto mediante sensores, para a continuación, digitalizar la señal producida por el sensor y finalmente ser almacenada en la memoria de un computador para su posterior procesamiento [7].

### 1.7.2.2 Pre-procesado

Toda imagen que se adquiere por medios ópticos, electro-ópticos o electrónicos sufre en cierta medida los efectos de la degradación que se manifiestan en forma de ruido, pérdida de definición y de fidelidad de la imagen. La degradación viene provocada por el ruido de los sensores de captura, imprecisiones en el enfoque de la cámara, movimiento de la misma o perturbaciones aleatorias, entre las que tiene relevancia el efecto de la propagación de la radiación en el medio de transmisión. El pre-procesado pretende reparar en la imagen los desperfectos producidos o no eliminados por el *hardware*: deformación de ésta, ruido introducido, poco o mucho contraste o brillo, entre otros. Los algoritmos de pre-procesado permiten modificar la imagen para eliminar ruido, transformarla geométricamente, mejorar la intensidad o el contraste, etc.; procesos que intentan mejorar el resultado final de la imagen capturada [7].

### 1.7.2.3 Segmentación

Segmentar una imagen digital significa dividirla en áreas o zonas separadas e individualizadas. Es decir, consiste en diferenciar del fondo los diversos objetos tomando en cuenta factores como su ubicación.

Al final de la etapa de segmentación, se tienen que conocer los objetos que hay para extraer y las características propias de cada uno de ellos. Además, cada pixel de la imagen tiene que tener una etiqueta que los defina, de forma que simplemente por agrupación de puntos con la misma etiqueta y conectados espacialmente, se pueda determinar la lista de objetos (estos objetos son realmente zonas o regiones

individualizadas dentro de la imagen, ya que un objeto, puede estar repartido en varias regiones diferentes dentro de la imagen obtenida) [7].

#### **1.7.2.4 Representación y descripción (extracción de características)**

Después de la segmentación, la siguiente etapa en el procesamiento de una imagen consiste en la extracción de características como por ejemplo; área, perímetro, excentricidad, momentos de inercia, esqueletos, textura, color de las áreas.

Se puede usar la representación por frontera o región completa. La representación por la frontera es apropiada cuando el objetivo se centra en las características de la forma externa como esquinas o concavidades y convexidades. La representación por regiones es apropiada cuando la atención se centra en propiedades internas como la textura o el esqueleto. Sin embargo, en muchas aplicaciones ambas representaciones coexisten [7].

#### **1.7.2.5 Reconocimiento e interpretación**

El reconocimiento es el proceso que asigna una etiqueta a un objeto basada en la información que proporcionan las características para clasificación. La interpretación lleva a asignar significado al conjunto de objetos reconocidos.

Existen varias técnicas de clasificación, como redes neuronales, sistemas expertos, lógica difusa, clasificadores estadísticos, entre otros.

Una vez que la imagen ha sido satisfactoriamente segmentada y han sido extraídas sus características, el procesador o CPU puede hacer una serie de pruebas en los elementos de interés que aparecen en la escena. Existen varios algoritmos que permiten obtener los resultados deseados. Entre sus funciones principales se encuentran:

- Detección de formas.
- Análisis geométrico.
- Comparación con patrones.
- Medición de objetos.

- Detección de objetos difusos.

Los algoritmos de reconocimiento son usados estratégicamente con diferentes fines, por ejemplo decidir si una pieza se ha fabricado correctamente o no, guiar un robot hacia un objeto concreto, reajustar la maquinaria de producción automáticamente, entre otros.

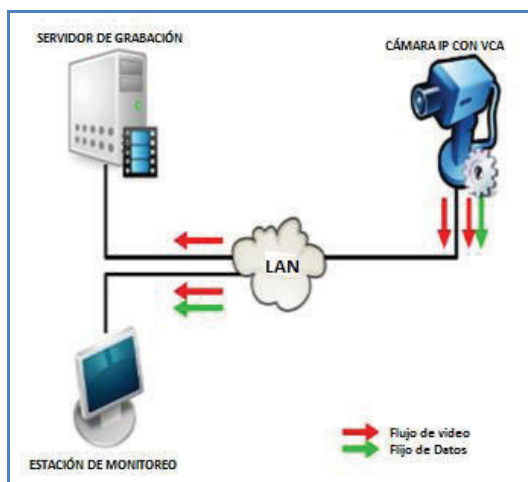
En la siguiente sección se detallan los tipos de sistemas para procesamiento de datos en analítica de video.

### **1.7.3 TIPOS DE SISTEMAS PARA PROCESAMIENTO DE DATOS EN ANALÍTICA DE VIDEO**

Los sistemas para analítica de video pueden procesar los datos de acuerdo a los siguientes esquemas:

VCA es el análisis de contenido de video que se pueden presentar en los sistemas como se muestra a continuación.

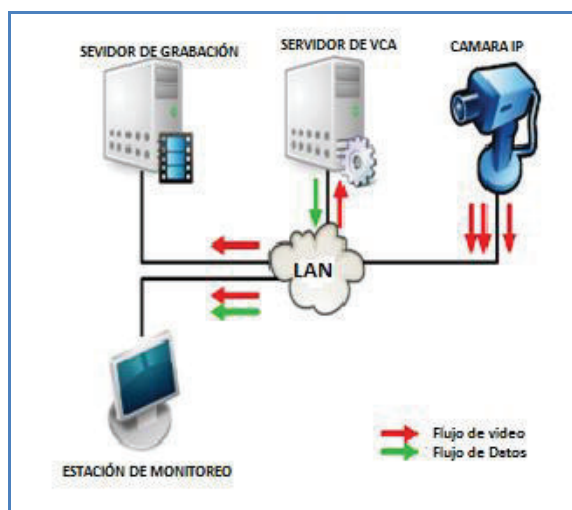
- Sistema Embebido.- La imagen capturada es analizada por el procesador del dispositivo de borde, es decir, en la cámara IP o el codificador.  
Esta solución es la más utilizada en la analítica de video, especialmente cuando se trata de proyectos de cientos de cámaras porque necesitan menos servidores, no consumen tanta memoria del servidor, utilizan un ancho de banda más bajo, y puede realizar varios tipos de alarmas, tales como detección de personas, vehículos y objetos a través de líneas virtuales, objetos que aparecen o desaparecen, entre otros [67]. El esquema de los equipos se muestra en la Figura 1.13.



**Figura 1.13:** Sistema embebido [67]

- Sistema basado en un servidor.- En este caso, además de enviarse al servidor de grabación, y estaciones de monitoreo, se envían imágenes de cámaras IP y codificadores a los servidores de análisis. Estos servidores analizan simultáneamente las imágenes de múltiples cámaras, y se aplicarán diferentes alarmas. Si se detecta una situación, el sistema advierte a los operadores. Estos son los sistemas preferidos para soluciones pequeñas y medianas, de hasta 32 cámaras [67].

El esquema de los equipos se muestra en la Figura 1.14.

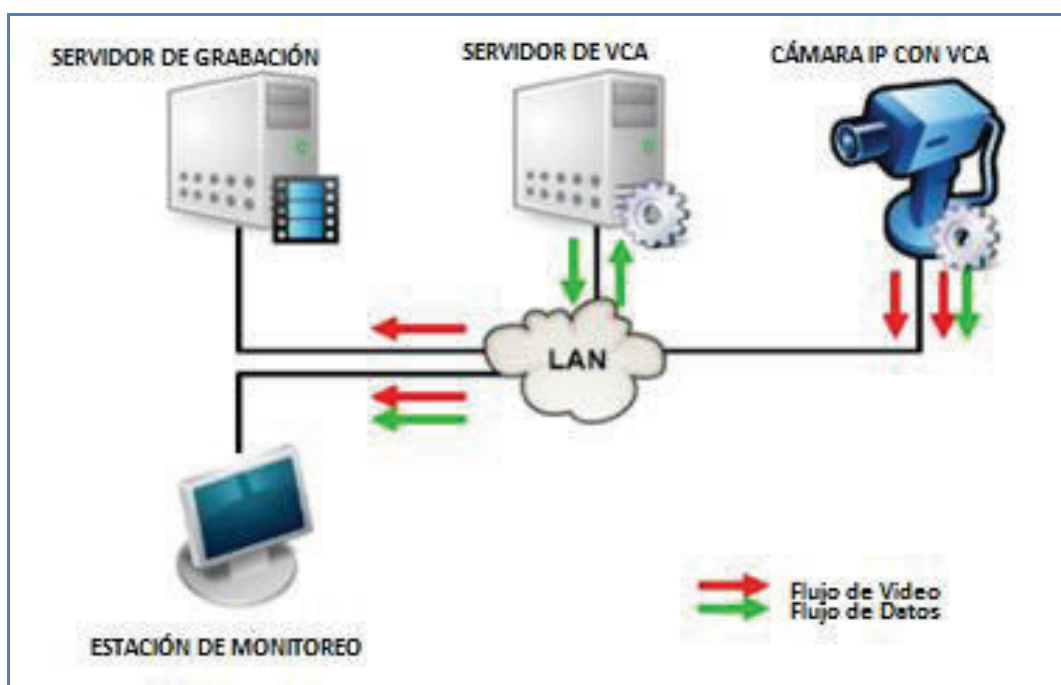


**Figura 1.14:** Sistema basado en un servidor [67]

- Sistema Híbrido.- Es la combinación de un sistema embebido y un sistema basado en un servidor.

Las soluciones híbridas son las más recomendadas para sistemas grandes, los beneficios son la reducción del uso de ancho de banda en los servidores [67].

El esquema de los equipos de un sistema híbrido se muestra en la Figura 1.15.



**Figura 1.15:** Sistema híbrido [67]








A continuación, se describen los tipos de filtros en analítica de video para sistemas de videovigilancia.









#### 1.7.4 TIPOS DE FILTROS EN ANALÍTICA DE VIDEO

En la Tabla 1.1 se detallan los filtros que se utilizan para seleccionar imágenes en analítica de video para un sistema de videovigilancia.



**Tabla 1.1:** Tipos de Filtros en analítica de video en videovigilancia [78]

TIPOS DE FILTRO	ICONO	DESCRIPCIÓN
Clasificación de objetos		La clasificación de objetos es una herramienta que permite diferenciar entre objetos de interés como personas, animales, autos y otros movimientos de la escena; también detecta el tamaño, las dimensiones e incluso la velocidad de los objetos
Conteo de personas		Es una herramienta para recopilar estadísticas de comportamiento de las personas, donde se muestra cuántas personas se encuentran en un determinado horario y en un determinado lugar. Cuenta el flujo de personas sin necesidad de intervención humana, por ejemplo en un banco
Conteo de vehículos		Es una herramienta que permite recopilar estadísticas de comportamiento de los vehículos, ideal para aplicaciones comerciales y de seguridad donde se puede determinar el flujo de vehículos sin necesidad de intervención humana
Entrada y salida		Una alarma es accionada cuando un objeto cruza desde el exterior al interior de la zona de detección. Así mismo, una alarma es activada cuando un objeto cruza desde el interior hacia el exterior de una zona de detección
Aparecer y desaparecer		Este filtro es utilizado para detectar personas apareciendo y desapareciendo dentro de la puerta de entrada, pero ignora personas pasando por delante de la entrada
Parada		Objetos que se encuentran detenidos dentro de la zona durante más tiempo del definido activarán la alarma
Permanencia		Este es un filtro para la detección de objetos cuya velocidad disminuye o aumenta. Objetos que permanecen durante un periodo de tiempo más largo que el especificado dentro de una zona, disparan una alarma (por ejemplo en vigilancia dentro de un túnel o en un sistema de control de tráfico)

TIPOS DE FILTRO	ICONO	DESCRIPCIÓN
Detección de dirección		Objetos que van en la dirección marcada (dentro de los límites de tolerancia de ángulo) de un lado a otro de la zona o encima de la línea, accionan una alarma
Detección de seguimiento		El seguimiento de objetos es rastrear en un periodo de tiempo la ubicación de uno o más objetos móviles
Filtro de color		El filtro de color detecta objetos en base a una intensidad de colores predefinida, por ejemplo para el personal médico en hospitales o en bases militares
Detección de retirada de objetos		Objetos que son retirados de una zona predefinida serán detectados y la alarma se accionará
Detección de objetos abandonados		Este filtro detecta el abandono de un objeto en una zona predeterminada, como por ejemplo un bolso o una maleta abandonados en la calle
Seguimiento de personas		Este filtro es usado para detectar y seguir distintos objetos. Diferencia entre objetos, aunque estén superpuestos. Está provisto de un avanzado sistema de seguimiento de objetos por medio de análisis del color. El filtro es ideal para entornos concurridos y aplicaciones
Estabilizador de video		En entornos donde la estabilidad de la cámara es un problema, esta función compensará cualquier movimiento para proporcionar una visión clara
Detección de manipulación		Con este filtro, cualquier intento de manipulación de la cámara (manosear, cubrir o desenfocar la lente) será detectado y accionará la alarma

### **1.7.5 INSTALACIÓN DE UN SISTEMA DE VIDEOVIGILANCIA**

Los sistemas de videovigilancia por defecto presentan cierto grado de dificultad en su instalación, sin embargo, mientras se tome en consideración los factores que pueden afectar en su instalación, se contará con un mejor desempeño y se podrá aprovechar las ventajas de la analítica de video.

A continuación, se detallan los factores que se deben considerar en la instalación de un sistema de videovigilancia.

#### **1.7.5.1 Factores de instalación**

Para mejorar la eficiencia de un sistema de videovigilancia se deben considerar las siguientes recomendaciones [78]:

- **Ubicación de la cámara.-** La instalación de la cámara en lugares bajos o en sitios donde se cruzan objetos demasiado cerca del rango de detección debe ser evitada, al igual que en sitios donde la iluminación es extrema puesto que la luz que resulta demasiado fuerte puede dañar la calidad del video, inundando el plano de la imagen de la cámara de vigilancia.  
Un enfoque correcto se logrará cuando la cámara es ubicada en un lugar elevado en dirección descendente.
- **Cambios de luz y cantidad de luz.-** El análisis de video puede funcionar en condiciones de luz escasa, sin embargo, el resultado presentará menor rendimiento que con la luz abundante. Por lo tanto, se debe medir con un luxómetro el nivel de luz antes de utilizar el sistema.
- **Movimientos excesivos.-** Los movimientos del ambiente donde se instala la cámara como por ejemplo ondulación en el agua, balanceo de árboles, sombras, condiciones de luz extrema, entre otros pueden ocasionar problemas en la eficacia de los sistemas de videovigilancia. Para ello se debe utilizar cámaras adecuadas para dicho ambiente y las condiciones

perturbadoras deben ser limitadas en la configuración de parámetros de la cámara, si una falsa alarma por movimientos excesivos no puede ser evitada, las áreas más sensibles deben ser marcadas como ignoradas.

- Selección de lentes y ángulo de observación.- Para un sistema de videovigilancia es importante elegir lentes que ofrecen un amplio ángulo de visión como por ejemplo lo lentes ultra gran angular, también conocidas como ojos de pez, que cubren hasta  $180^\circ$ .
- Profundidad de detección.- El campo de visión varía con la distancia focal del lente y con la distancia a los objetos de la escena. Es importante determinar la distancia máxima que se desea cubrir con las cámaras y en función de esto seleccionar la cámara que se adapte a esta aplicación.

Existen una serie de algoritmos que se utiliza en analítica de video para la detección y seguimiento de personas y objetos. Los algoritmos como Viola&Jones y Hausdorff permiten extraer información de un conjunto de imágenes receptadas por un sistema de videovigilancia en una escena.

En el capítulo 2 se detallará de manera profunda el algoritmo por el método cascada de Viola&Jones y el método por modelos geométricos usando la distancia de Hausdorff para la detección de rostros.

## **CAPÍTULO 2**

### **ANÁLISIS DE DOS ALGORITMOS PARA DETECCIÓN DE ROSTROS, VIOLA&JONES Y HAUSDORFF**

#### **2.1 INTRODUCCIÓN**

Este capítulo tiene como objetivo analizar dos algoritmos utilizados para la detección de rostros en un sistema de videovigilancia. El primer algoritmo que se analizará está basado en modelos geométricos usando la distancia de Hausdorff de Jesorsky, Kirchberg, Frischholz. El segundo algoritmo que se analizará está basado en clasificadores en cascada de Viola&Jones. Se estudiarán los principios que utilizan, así como su estructura y funcionamiento. Finalmente se realizará una Tabla comparativa de los dos tipos de algoritmos utilizados.

#### **2.2 DETECCIÓN DE ROSTROS EN IMÁGENES DE VIDEO**

La ciencia y tecnología buscan de manera continua la creación y desarrollo de equipos y sistemas autónomos que trabajen en base a algoritmos<sup>8</sup> y logren extraer la información requerida de manera automática y eficaz.

La detección de rostros es un paso fundamental en los sistemas de seguimiento y reconocimiento de rostros, con el fin de localizar y extraer la región del rostro desde un fondo. También tiene varias aplicaciones en áreas como recuperación de imágenes basadas en contenido, codificación de video, videoconferencia, vigilancia de multitudes e interfaces inteligentes entre humanos y computadoras [10].

Debido a la amplia área de aplicación, la detección de rostros se ha convertido en un proceso importante en el mundo del procesamiento digital de imágenes, y aunque

---

<sup>8</sup> Conjunto de operaciones sistemáticas que permite hacer un cálculo y hallar la solución de un problema.

existen factores que pueden perturbar la detección de un rostro como iluminación, oclusión ambiental, entre otros, existen nuevos y mejorados algoritmos para este fin.

## **2.3 ANTECEDENTES**

Los primeros esfuerzos en la detección de rostros empiezan en los años sesenta donde se utilizaron técnicas heurísticas y antropométricas como el método desarrollado por Toshiyuki Sakai, Makoto Nagao y Takeo Kanade [1]. Estas técnicas son en gran parte básicas debido a varias suposiciones tales como fondo plano y cara frontal. Para estos sistemas cualquier cambio en las condiciones de la imagen significaba un ajuste fino o un rediseño completo. A pesar de estos problemas, el crecimiento de los intereses de investigación permaneció estancado hasta la década de 1990 cuando el desarrollo de algoritmos que detectaban rostros inició su auge, dado que la tecnología empezaba a permitir la reproducción de videos, videoconferencias, los procesadores eran cada vez más rápidos [90].

La detección de rostros tiene muchas aplicaciones pero fue en el campo de la fotografía digital donde más se acerca al público. Fujifilm fue la primera empresa anunciando la tecnología de detección de rostros en el año 2004, pero fue Nikon, en 2005, la primera empresa que sacó al mercado una cámara utilizando este sistema [91].

Durante la última década se han presentado esquemas de segmentación más robustos particularmente aquellos que usan el movimiento, el color y la información generalizada. El uso de estadísticas y redes neuronales también ha permitido que los rostros sean detectados en escenas desordenadas a diferentes distancias de la cámara. Las redes neuronales usan un conjunto de imágenes que representan rostros de todo tipo y otro conjunto de imágenes que no representan rostros, de forma que la red neuronal pueda establecer el criterio adecuado acerca de lo que es un rostro y lo que no lo es. El método estadístico consiste en calcular la probabilidad de que la imagen sea un rostro, y a la probabilidad de que no lo sea [90].

Los métodos de detección de rostros requieren información de las características del rostro y pueden organizarse de manera efectiva en dos grandes categorías que se distinguen por su diferente enfoque en la detección de un rostro.

Las propiedades aparentes del rostro tales como el color de la piel y la geometría del rostro se extraen con los diferentes métodos que se describen a continuación.

## **2.4 CLASIFICACIÓN DE MÉTODOS PARA DETECCIÓN DE ROSTROS**

Los métodos para la detección de rostros en una imagen se dividen en dos enfoques principales; basadas en las características o rasgos faciales y basadas en la imagen.

En la Figura 2.1 se muestran los dos enfoques y sus respectivas subdivisiones que se detallan a continuación [90].

- 1) Enfoque basado en características faciales.- Este enfoque se basa en buscar determinados elementos que componen un rostro, como pueden ser los ojos, líneas de contorno del rostro, entre otros [90].

Este enfoque se divide en tres métodos que se explican a continuación:

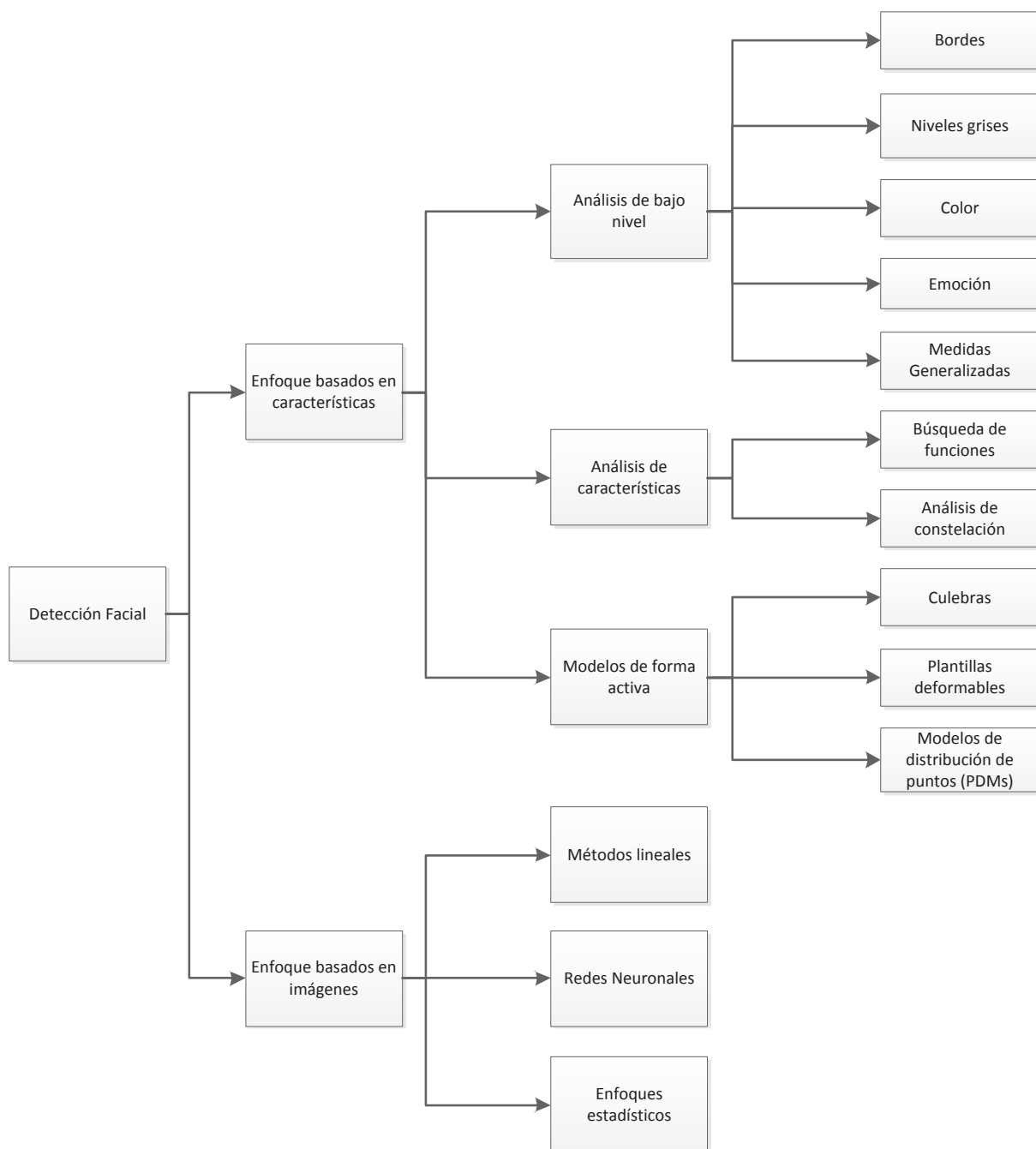
- Análisis de bajo nivel.- Son técnicas que trabajan con la información proporcionada por los píxeles.
- Análisis de características.- Se basa en la geometría del rostro para identificar rasgos.
- Modelos de forma activa.- Usan la información de los bordes y el brillo de la imagen, comparando con los rasgos estándar de forma que poco a poco se van cambiando y tomando la forma del rasgo patrón.

- 2) Enfoques basados en la imagen.- En este enfoque los métodos trabajan con la imagen completa o zonas concretas de la misma, efectuando cálculos matemáticos que determinan si hay un rostro o no, sin buscar rasgos concretos.

Este enfoque se divide en tres métodos que se describen a continuación:

- Métodos lineales.- Partiendo de un conjunto de imágenes que representan rostros, cada fracción de la imagen produce un error llamado “distancia desde el espacio de los rostros”, comparando los valores mínimos de estas distancias se determina la posición de los rostros en la imagen.
- Redes neuronales.- Se usan para la clasificación de imágenes según patrones establecidos, por ejemplo, la red neuronal usa un conjunto de imágenes que representan rostros de todo tipo (de varias razas, tonos de piel, con y sin gafas, posiciones de los labios y ojos, ligeras rotaciones, entre otros) y otro conjunto de imágenes que no representan rostros, de forma que la red neuronal pueda establecer el criterio adecuado acerca de lo que es un rostro y lo que no lo es. La respuesta de la red neuronal ante una imagen de entrada es la de decidir si dicha imagen corresponde o no a un rostro, es decir, una respuesta binaria.
- Análisis estadísticos.- Usa un método estadístico y lo que más se parezca estadísticamente a lo que se conoce como un rostro, tendrá más probabilidades de ser un rostro.





**Figura 2.1:** Clasificación de la detección de rostros basado en las características e imágenes [5].

A continuación, se detallan los tipos de algoritmos para la detección de rostros.

## **2.5 TIPOS DE ALGORITMOS PARA DETECCIÓN DE ROSTROS**

Existen muchos algoritmos para localizar un rostro humano en una escena con diferentes niveles de dificultad. A continuación, se describen algunos de los más utilizados.

### **2.5.1 DETECCIÓN DE ROSTROS EN IMÁGENES CON FONDO CONTROLADO**

Estos algoritmos presentan un nivel bajo de dificultad. Se utilizan imágenes con un fondo monocromático plano con un fondo estático predefinido donde al remover el fondo siempre quedará los límites del rostro deseado [93].

### **2.5.2 DETECCIÓN DE ROSTROS POR COLOR**

Si se tiene acceso a imágenes a color, se puede usar el color de piel típico para encontrar segmentos del rostro. La desventaja es que no funciona con todo tipo de colores de piel y no es muy eficiente bajo condiciones de iluminación variables.

La detección de rostros por color se puede realizar con los métodos que se enlistan a continuación [93]:

- 1) Por extracción básica de color.
- 2) Con la detección de rostros en imágenes en color.
- 3) Utilizando PCA (Análisis de Componentes Principales).
- 4) Con la detección de color de piel bajo condiciones de iluminación cambiantes.
- 5) Con la detección de rostros y reconocimiento en imágenes en color con un fondo complejo.
- 6) Con la visión por ordenador y color de la piel humana.

### **2.5.3 DETECCIÓN DE ROSTROS POR EL MOVIMIENTO**

En un video en tiempo real, se puede utilizar el rostro que siempre se encuentra en movimiento en este escenario. Basta con calcular el área en movimiento y se obtiene

el rostro. La principal dificultad se presenta si existen otros objetos moviéndose en un segundo plano.

La detección de rostros por movimiento se puede realizar de las formas que se describen a continuación:

- 1) Explicación de la detección básica de movimientos para encontrar el rostro.
- 2) Detección de parpadeo, los ojos humanos parpadean simultáneamente; esto puede usarse para encontrar y normalizar rostros.

#### **2.5.4 DETECCIÓN DE ROSTROS USANDO VARIOS ALGORITMOS**

La combinación de varios algoritmos normalmente produce un resultado mejor. Los métodos para detección de rostros que se usan en conjunto son:

- 1) Una mezcla de color y 3D.
- 2) Una mezcla de color y eliminación de fondo.

#### **2.5.5 DETECCIÓN DE ROSTROS EN ESCENAS SIN RESTRICCIONES**

Este es el algoritmo con mayor nivel de dificultad en la detección de rostros. Dada una imagen fija en blanco y negro los seres humanos pueden diferenciar el rostro, pero en la visión artificial se requiere de alguno de los algoritmos que se describen a continuación.

- 1) Seguimiento de rostro basado en modelos.- Consiste en la detección de bordes utilizando modelos geométricos. A continuación, se mencionan dos métodos de alto rendimiento:
  - Detección de rostros en tiempo real mediante coincidencia de orientación de bordes.
  - Detección robusta del rostro usando la distancia de Hausdorff.
- 2) Clasificadores en Cascadas.- El avance en la detección de rostros ocurrió con el algoritmo desarrollado por Viola&Jones. Utiliza una serie de clasificadores

(clasificadores en cascada), cada uno con un conjunto de características visuales. Este algoritmo es el más utilizado para la detección de rostros.

A continuación, se describe de manera detallada el algoritmo por el método de cascada de Viola&Jones.

## **2.6 ALGORITMO DE VIOLA&JONES**

### **2.6.1 DESCRIPCIÓN**

El algoritmo de P. Viola y M.J. Jones es capaz de realizar el procesamiento de imágenes de forma extremadamente rápida, logrando una alta tasa de detección. Es utilizado en un rango importante de aplicaciones debido a que su alta velocidad posibilita la detección en tiempo real (aplicable en codificación de video, teleconferencias, seguimiento de objetos en movimiento, entre otros.), y por otro lado se utiliza en diversos dispositivos de baja potencia de cálculo, *handhelds*<sup>9</sup> y sistemas embebidos (por ejemplo cámaras de fotos).

### **2.6.2 INDICADORES**

Para medir el desempeño de un algoritmo de detección de rostros se utilizan los siguientes indicadores fundamentales:

- Tasa de detección.- Es el porcentaje de rostros correctamente detectados respecto del total de rostros de la imagen.
- Tasa de falsos positivos.- Indica el número de regiones que se han marcado como rostros donde realmente no hay un rostro.
- Tasa de falsos negativos.- Indica el número de rostros que no han sido correctamente detectados. Es decir que han sido clasificados negativamente.

---

<sup>9</sup> Dispositivo de mano

### 2.6.3 METODOLOGÍA

El principio básico del algoritmo de Viola&Jones es explorar cada sub-ventana capaz de detectar rostros a través de una imagen de entrada dada.

Una metodología estándar realiza el procesamiento de una imagen cambiando la escala de la imagen de entrada a diferentes tamaños, para luego ejecutar el detector de tamaño fijo a través de estas imágenes. Esto toma bastante tiempo debido al cálculo de tamaño de las diferentes imágenes y por tanto al tiempo de detección de rostros.

Contrariamente a la metodología estándar anteriormente mencionada, Viola&Jones ajusta el detector en lugar de la imagen de entrada y ejecuta el detector muchas veces a través de la imagen cada vez con un tamaño diferente. En un principio podría parecer que ambos enfoques son igualmente largos o demorosos pero Viola&Jones ha ideado un detector de escala invariable que requiere el mismo número de cálculos cualquiera que sea el tamaño. Este detector se construye utilizando una imagen llamada integral y algunas características rectangulares simples que recuerdan a las *wavelets*<sup>10</sup> de Haar, por lo que es posible implementar el algoritmo utilizando únicamente aritmética entera lo que proporciona también una rapidez en la detección [93].

La versión final del algoritmo de Viola&Jones está compuesta por 32 clasificadores puestos en cascada. En el proceso, una imagen que es reconocida como un rostro necesita alrededor de 80.000 operaciones, sin embargo, gracias a esta arquitectura en cascada el promedio de operaciones para una sub-ventana es muy bajo [93].

Mediante el uso de clasificadores en cascada muy sencillos que se ejecutan uno detrás de otro, cada clasificador de la cascada selecciona características y se

---

<sup>10</sup> Secuencia de funciones matemáticas para la representación de datos o funciones

entrena mediante el algoritmo de AdaBoost<sup>11</sup>. Los primeros clasificadores son muy sencillos y permiten rechazar una gran cantidad de no rostros mientras que aceptan un porcentaje muy alto de rostros

La cascada permite desechar gran parte de las regiones de la imagen y solo concentrarse en las zonas en las que es más probable que haya un rostro por lo que la detección es bastante rápida.

#### 2.6.4 ETAPAS

El algoritmo de Viola&Jones tiene cuatro pasos descritos a continuación [94]:

- 1) Etapa de Características.
- 2) Etapa de Creación de una Imagen Integral.
- 3) Etapa de Entrenamiento.
- 4) Etapa de Clasificadores en Cascada.

El primer y segundo paso son conocidos como la "Imagen Integral", que permite que las características tipo "*Haar*" utilizadas por este detector se puedan procesar muy rápido.

La tercera es un proceso de aprendizaje automático, "AdaBoost", que selecciona sólo las características importantes de todo el conjunto.

La cuarta etapa es la creación de una estructura en "cascada", la combinación de clasificadores complejos, que rechaza el fondo de la imagen de entrada pasando más tiempo de cálculo en las áreas que puedan contener el objeto de interés.

A continuación se describe el primer paso del algoritmo de Viola&Jones.

---

<sup>11</sup> abreviatura de "Adaptive Boosting ", es un algoritmo de aprendizaje automático formulado por Yoav Freund y Robert Schapire que ganó el Premio Gödel en 2003 por su trabajo.

#### 2.6.4.1 Etapa de características

El procedimiento de detección de objetos del algoritmo de Viola&Jones clasifica las imágenes basándose en la evaluación de características. Existen buenas razones para usar las características en lugar de los píxeles:

- La razón más común es que las características pueden actuar como un codificador del dominio del conocimiento *ad-hoc* que es difícil de aprender usando una cantidad finita de datos de entrenamiento.
- Una segunda buena razón crítica para las características es que funciona mucho más rápido que un sistema basado en píxeles.

Las características utilizadas por Viola&Jones son del estilo de las características *Haar Basis*<sup>12</sup> y específicamente se utilizan tres tipos:

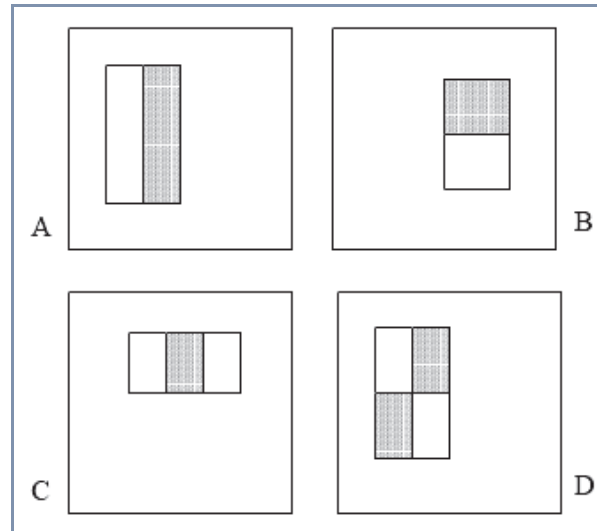
- 1) El valor de una característica de dos rectángulos que se muestran en la Figura 2.2 (A) y la Figura 2.2 (B) es la diferencia entre la suma de los píxeles dentro de dos regiones rectangulares. Las regiones tienen el mismo tamaño y forma y están horizontalmente o verticalmente adyacentes como se muestra en la Figura 2.2.
- 2) El valor de una característica de tres rectángulos que se muestran en la Figura 2.2 (C) calcula la suma dentro de dos rectángulos exteriores sustraídos de la suma en un rectángulo central.
- 3) El valor de una característica de cuatro rectángulos que se muestran en la Figura 2.2 (D) calcula la diferencia entre pares diagonales de rectángulos.

La resolución base con la que trabaja el detector es de 24x24 píxeles donde el conjunto total de características para este tamaño de imagen es extremadamente

---

<sup>12</sup> Función matemática para la representación de datos

alto llegando actualmente a los 180000 píxeles, mucho mayor que la cantidad de píxeles contenidos en la imagen de 24x24 píxeles (576 píxeles), necesarios para describir completamente la imagen. En este sentido es que se utilizará AdaBoost para seleccionar las más importantes [11].



**Figura 2.2:** Característica de Haar [11]

La Figura 2.2 muestra un ejemplo de características de rectángulo mostradas en relación con la ventana de detección. En resumen, la suma de los píxeles que se encuentran dentro de los rectángulos blancos se sustrae de la suma de píxeles en los rectángulos grises.

Una característica está formada por un número finito de rectángulos y su valor escalar consistirá en la suma de los píxeles de cada rectángulo, aplicando un cierto factor de peso.

La Ecuación 1 permite calcular el valor de una característica.

$$caracteristica = \sum_{1 \leq i \leq N} w_i * suma\_rectangulo(r_i)$$

**Ecuación 1:** Ecuación para calcular el valor de una característica



Donde  $\{r_1, \dots, r_N\}$  son los rectángulos que forman la característica y  $w_i$  el peso<sup>13</sup> de cada uno.

A continuación, se describe el segundo paso del algoritmo de Viola&Jones.

#### 2.6.4.2 Etapa de la Imagen Integral

Las características de rectángulo se pueden calcular muy rápidamente usando una representación intermedia para la imagen denominada la imagen integral.

Un valor de la imagen integral  $ii(x, y)$  será igual al valor del pixel  $i(x, y)$  de la imagen sumado a todos los pixeles de la imagen que estén a la izquierda y arriba de la posición  $(x, y)$  expresada en la Ecuación 2 [95].

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

**Ecuación 2:** Imagen Integral (suma de pixeles)

La imagen integral se puede calcular iterativamente comenzando por la posición (0,0). En la Figura 2.3 se puede ver una representación de la imagen integral de una fotografía. El pixel de abajo a la derecha contiene la suma de todas las intensidades de los pixeles de la imagen.

$$s(x, y) = s(x, y - 1) + i(x, y)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y)$$

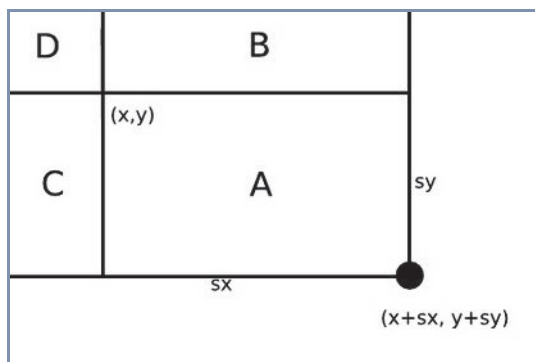
**Ecuación 3:** Suma Acumulativa de pixeles

Gracias a esta representación como se muestra en la Ecuación 3, se puede calcular la suma de los pixeles de un rectángulo de cualquier tamaño utilizando únicamente 4 accesos a memoria. Para calcular el valor de un rectángulo A cuya esquina superior

---

<sup>13</sup> Espacio en bytes que ocupa una imagen en el dispositivo de almacenamiento

izquierda está en  $(x, y)$  y cuyo tamaño es  $(sx, sy)$  únicamente se necesita acceder al valor de la imagen integral de 4 esquinas.



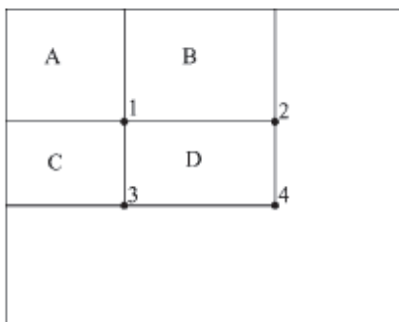
**Figura 2.3:** Imagen integral en el punto  $(x, y)$  [95]

Utilizando la imagen integral cualquier suma rectangular se puede calcular en cuatro referencias de matrices como se muestra en la Figura 2.4. Claramente la diferencia entre dos sumas rectangulares se puede calcular en ocho referencias. Dado que las características de dos rectángulos definidas anteriormente implican sumas rectangulares adyacentes, pueden calcularse en seis referencias de arreglo, ocho en el caso de las características de tres rectángulos y nueve en cuatro rectángulos.

En la Figura 2.4 se muestra la suma de los píxeles dentro del rectángulo D, se pueden calcular cuatro referencias de matrices:

- 1) El valor de la imagen integral en la posición 1 es la suma de los píxeles en el rectángulo A.
- 2) El valor en la posición 2 es  $A + B$ .
- 3) El valor en la posición 3 es  $A + C$ .
- 4) El valor en la posición 4 es  $A + B + C + D$ .

Entonces la suma dentro de D se puede calcular como la diferencia y sumas de las siguientes posiciones  $4 + 1 - (2 + 3)$ .



**Figura 2.4:** Representación de píxeles [11]

Por un lado las características rectángulo pueden resultar un poco primitivas o toscas pero son capaces de dar una representación muy rica de la imagen si son eficientemente entrenadas y su alta eficiencia computacional compensa su falta de flexibilidad.

Además, gracias a la imagen integral cualquier característica rectángulo puede ser calculada en cualquier lugar de la imagen y a cualquier escala utilizando para ello unas pocas operaciones, esto hace que este detector sea extremadamente rápido.

La imagen integral es un espacio discreto<sup>14</sup> y por tanto puede ser escalado, se debe limitar a los números con lo que se perdería precisión. El principal problema del escalado es el umbral que en caso de escalar mucho los rectángulos pueden modificar mucho el valor de la característica. Para evitar este problema se debe modificar el peso de cada rectángulo teniendo en cuenta la diferencia de áreas entre el rectángulo original y el escalado utilizando la Ecuación 4.

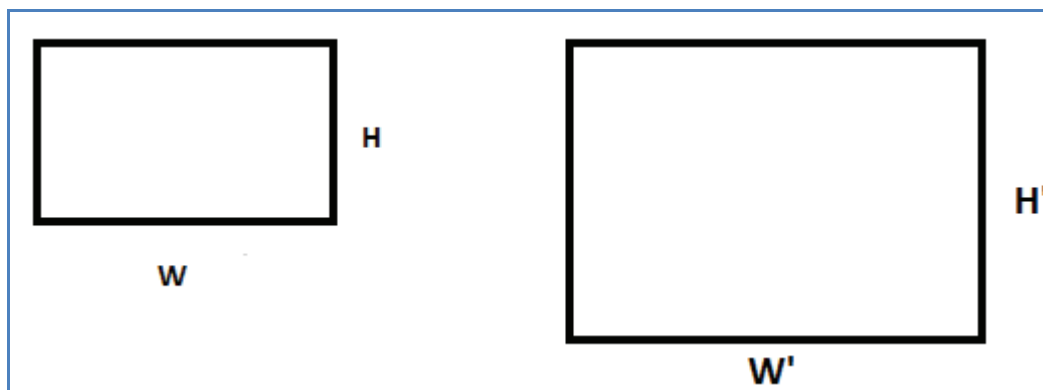
$$w' = \frac{\text{area original}}{\text{area escalada}}$$

**Ecuación 4:** Peso modificado

---

<sup>14</sup> Donde cada parte es analizada para facilitar el cálculo global del sistema.

En el entrenamiento se podrá trabajar con las características sin escalar, pero durante la detección es imprescindible aumentar de tamaño las características para poder abarcar toda la imagen como se muestra en la Figura 2.5.



**Figura 2.5:** Escalado de características [11]

A continuación, se describe el tercer paso del algoritmo de Viola&Jones.

#### 2.6.4.3 Etapa de Entrenamiento

Dado un conjunto de características y un conjunto de formación de imágenes positivas y negativas, cualquier número de métodos de aprendizaje de máquina podría ser utilizado para aprender una función de clasificación.

Viola&Jones entrena su detector de rostros con un algoritmo de Boosting que se encarga de seleccionar, de entre las miles de posibles características, las que mejor separan el espacio entre los rostros y no rostros.

El algoritmo de Boosting se basa en la unión de varios clasificadores sencillos (en este caso basados en las características tipo Haar) para crear un clasificador que tiene una menor tasa de error que los clasificadores individuales que lo forman. El algoritmo de Boosting utilizado por Viola&Jones es el AdaBoost.

A continuación, se describe el algoritmo de AdaBoost.

#### 2.6.4.3.1 AdaBoost

En su forma original el algoritmo de aprendizaje de Boosting se utiliza para aumentar el rendimiento de clasificación de un algoritmo de aprendizaje simple a veces llamado débil. El algoritmo de Boosting empleado por Viola&Jones en su trabajo es conocido como AdaBoost de aprendizaje débil, está diseñado para seleccionar la característica del rectángulo simple que separa los ejemplos positivos y negativos (esto es similar al acercamiento en el dominio de la recuperación de la base de datos de la imagen).

Freund y Schapire<sup>15</sup> demostraron que el error de entrenamiento del clasificador fuerte se aproxima a cero exponencialmente en el número de rondas.

El sistema AdaBoost se utiliza tanto para seleccionar un pequeño conjunto de características como para entrenar el clasificador [13]. Una serie de resultados ha demostrado que el rendimiento está relacionado con el margen de los ejemplos para el entrenamiento y que AdaBoost logra grandes márgenes rápidamente centrándose en el principal reto que es encontrar estas características.

Para cada característica el *clasificador débil*<sup>16</sup> determina la función de clasificación de umbral óptimo, de tal manera que exista un número mínimo de ejemplos clasificados erróneamente. Por lo tanto, un clasificador débil  $h_j(x)$  consiste en una característica  $f_j$ , un umbral  $\theta_j$  y una paridad  $p_j$  indicando la dirección del signo de la desigualdad como se muestra en la Ecuación 5.

---

<sup>15</sup> Freund y Schapire, inventaron el algoritmo AdaBoost en 1996. Recibieron el premio Gödel en 2003 por su trabajo.

<sup>16</sup> Son clasificadores menores, denominados “débiles” por su sencillez y escasa precisión

$$h_j(x) = \begin{cases} 1 & \text{if } p_j f_j(x) < p_j \theta_j \\ 0 & \text{por lo demás} \end{cases}$$

**Ecuación 5:** Clasificador débil

Donde  $x$  es una sub-ventana de 24X24 píxeles de una imagen. En la práctica, ninguna característica sola es capaz de realizar la tarea con bajo error, las primeras características resultantes del proceso de selección producen errores entre el 10% y 30%, mientras que las características seleccionadas más adelante realizan una tarea más complicada (debido a la naturaleza del AdaBoost), por lo tanto, producen errores entre el 40% y 50%.

A continuación, se muestra el proceso de selección y entrenamiento de AdaBoost.

- Dado un ejemplo de imágenes  $(x_1, y_1), \dots, (x_n, y_n)$  donde  $y_i = 0, 1$  para ejemplos negativo y positivos respectivamente.
- Inicializar pesos  $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$  para  $y_i = 0, 1$  respectivamente donde  $m$  y  $l$  son  $l_{ot}=1$ ,s números negativos y positivos respectivamente.

- Para  $t = 1, \dots, T$ :

1) Normalizar los pesos,  $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$ .

2) Seleccione el mejor clasificador débil con respecto al error ponderado como se muestra en la Ecuación 6.

$$\epsilon_t = \min_{f,p,\theta} \sum_i w_i |h(x_i, f, p, \theta) - y_i|.$$

**Ecuación 6:** Clasificador débil con respecto al error ponderado

- 3) Considerar  $h_t(x) = h(x, f_t, p_t, \theta_t)$  donde  $f_t$ ,  $p_t$ , y  $\theta_t$  son los mínimos de  $\epsilon_t$ .
- 4) Actualizar los pesos.

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$$

donde  $e_i = 0$  si por ejemplo  $x_i$  es clasificado correctamente,  $e_i = 1$  por otra parte, y  $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$

El clasificador fuerte final se muestra en la Ecuación 7; Donde  $\alpha_t = \log \frac{1}{\beta_t}$

$$C(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{por lo demás} \end{cases}$$

**Ecuación 7:** Clasificador Fuerte

El algoritmo AdaBoost continúa hasta que se han seleccionado  $T$  características. A cada clasificador débil se le asignará un valor  $\alpha$  que será más grande cuanto mejor sea la característica.

Una vez finalizado el algoritmo AdaBoost, el siguiente paso es clasificar una muestra, se considera un proceso sencillo donde se aplican las características del clasificador a la muestra y se suman los valores  $\alpha$  de aquellas que consideren que la muestra es un rostro. Si la suma es mayor que la mitad del total los valores  $\alpha$  se etiqueta la muestra como rostro.

A continuación, se describe el cuarto paso del algoritmo de Viola&Jones.

#### 2.6.4.4 Etapa Clasificadores en cascada

En esta sección se describe la forma de construir clasificadores en cascada que logra un mayor rendimiento de detección mientras reduce radicalmente el tiempo de cálculo. La clave es que se pueden construir clasificadores más eficientes

entrenados con AdaBoost que rechazan las sub-ventanas negativas, detectando la mayoría de las instancias positivas (es decir que el umbral de un clasificador potenciado puede ajustarse para que la tasa negativa falsa sea cerca de cero).

Los clasificadores más simples se utilizan para rechazar la mayoría de las sub-ventanas antes de que los clasificadores más complejos sean llamados a lograr bajas tasas de falsos positivos.

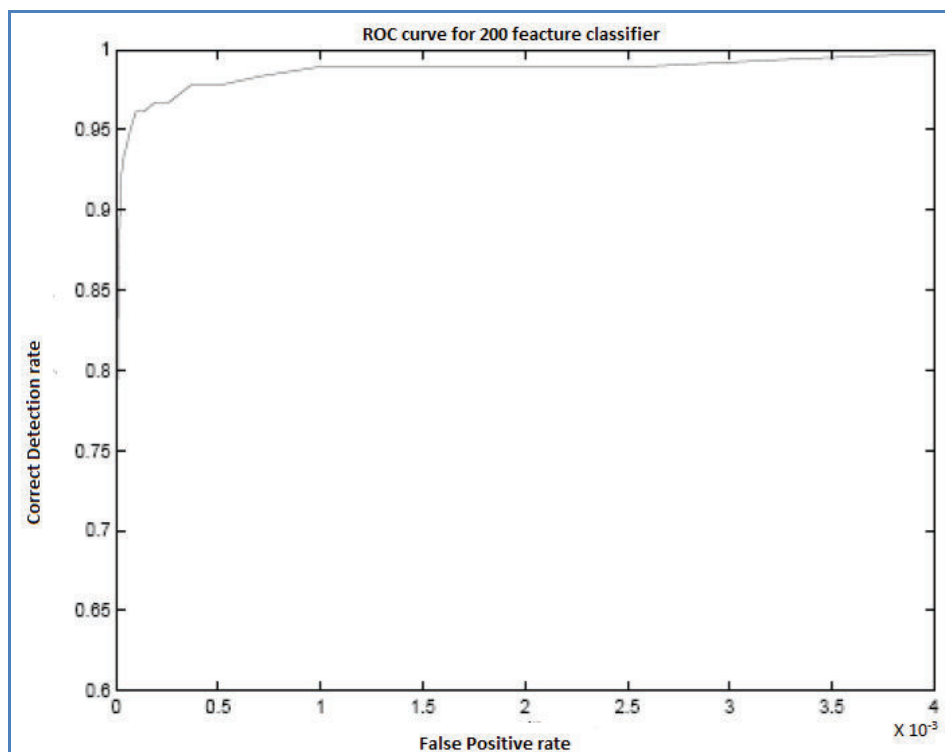
La forma general del proceso de detección es la de un árbol de decisiones degenerado a lo que se llama "cascada" como se puede observar en la Figura 2.7. Un resultado positivo del primer clasificador desencadena la evaluación de un segundo clasificador que también se ha ajustado para conseguir tasas de detección de rostros muy altas. Un resultado positivo del segundo clasificador desencadena un tercer clasificador, y así sucesivamente.

El primer clasificador contiene las características seleccionada por AdaBoost. El umbral inicial del clasificador es calculado para tener el mínimo error de entrenamiento, luego ese umbral es bajado de manera de obtener niveles de detección más altos, por supuesto esto implica niveles también más altos de falsos positivos. Se procede entonces a ajustar el umbral de este primer clasificador para obtener 100% de detecciones con un índice de falsos positivos del 40%.

Este proceso es capaz de reducir a menos de la mitad las ventanas a explorar en unas pocas operaciones, asegurando que todas los rostros son detectados.

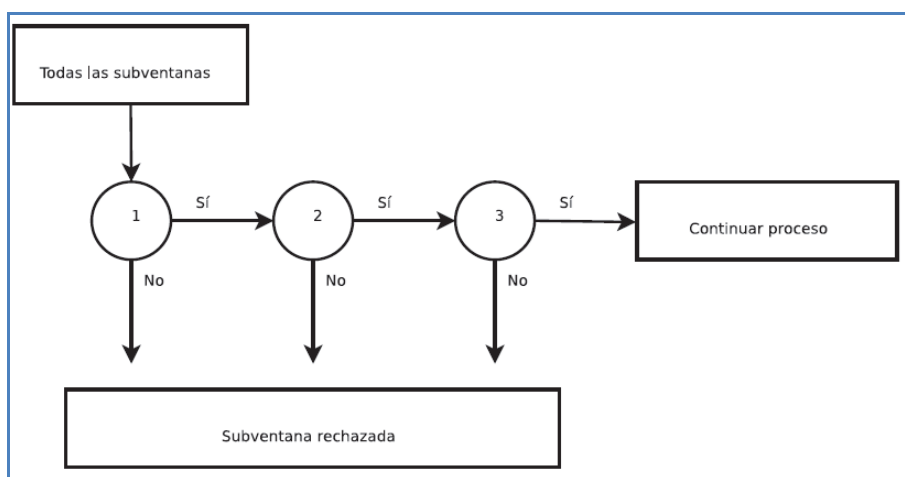
Los ejemplos más difíciles a los que se enfrentan los clasificadores más profundos empujan hacia abajo la curva característica de funcionamiento del receptor ROC (*Receiver Operating Characteristic*). A una tasa de detección dada, los clasificadores más profundos tienen tasas de falsos positivos, respectivamente, más altas como se puede observar en la Figura 2.6.





**Figura 2.6:** Curva ROC para un clasificador compuesto por 200 características rectángulo [11].

Cabe señalar que en la cascada un resultado negativo en cualquier punto conduce al rechazo inmediato de la sub-ventana.



**Figura 2.7:** Representación esquemática de una cascada de detección [11].

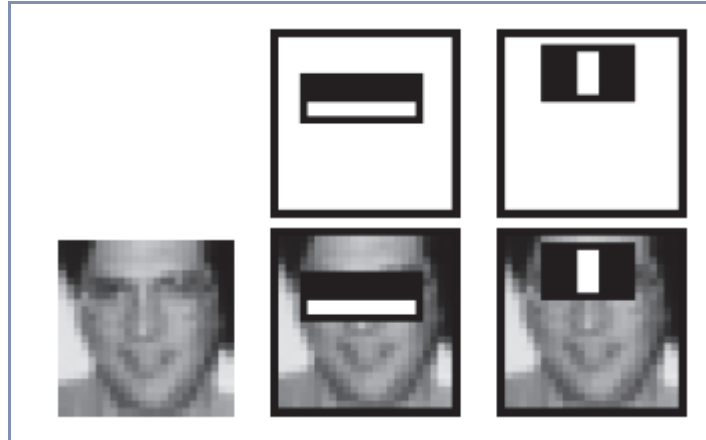
Una serie de clasificadores se aplican a cada sub-ventana. El clasificador inicial elimina un gran número de ejemplos negativos con muy poco procesamiento y las capas subsiguientes eliminan los negativos adicionales pero requieren un cálculo adicional.

Después de varias etapas de procesamiento el número de sub-ventanas se habrá reducido radicalmente. El procesamiento posterior puede adoptar cualquier forma como etapas adicionales a los clasificadores de cascada.

Las etapas en la cascada se construyen entrenando clasificadores usando AdaBoost y luego ajustando el umbral para minimizar los falsos negativos. Se debe tomar en cuenta que el umbral predeterminado de AdaBoost está diseñado para producir una tasa de error baja en los datos de formación. En general un umbral más bajo produce mayores tasas de detección y mayores tasas de falsos positivos.

Un clasificador de primera etapa excelente puede construirse a partir de un clasificador fuerte de dos características reduciendo el umbral para minimizar los falsos negativos. Medido contra un conjunto de entrenamiento de validación, el umbral puede ser ajustado para detectar 100% de los rostros con una tasa de falsos positivos del 40%.

La Figura 2.8 muestra la primera y segunda características seleccionadas por AdaBoost. Las dos características se muestran en la fila superior y luego se superponen en un rostro de entrenamiento típico en la fila inferior. La primera característica mide la diferencia de intensidad entre la región de los ojos y una región a través de las mejillas superiores. La característica indica en la observación que la región del ojo es a menudo más oscura que las mejillas. La segunda característica compara las intensidades en las regiones del ojo con la intensidad a través del puente de la nariz



**Figura 2.8:** Características utilizadas en Viola&Jones [11].

Para el diseño de la cascada es necesario fijar las metas de detección y performance. Lo usual sería lograr niveles de detección entre el 85 y 95%, con tasa de falso positivos muy bajas, del orden de  $10^{-5}$  o  $10^{-6}$ .

La idea es alcanzar esos registros pero con menor esfuerzo computacional.

Dado un clasificador en cascada entrenado, la tasa  $F$  de falsos positivos viene dada por la Ecuación 8.

$$F = \prod_{i=1}^K f_i$$

**Ecuación 8:** Cálculo de la tasa de falsos positivos

Donde  $f_i$  es la tasa de falsos positivos de la etapa  $i$  y  $K$  es el número de etapas del clasificador.

Por otro lado la tasa de detección  $D$  viene dada por la expresión mostrada en la Ecuación 9.

$$D = \prod_{i=1}^K d_i$$

**Ecuación 9:** Cálculo de la tasa de detección

Para una tasa de detección de 90% y un detector de 10 etapas, cada etapa debe alcanzar una tasa de detección de 0.99, ya que  $0.99^{10}$  es aproximadamente 0.9.

En principio esto es un poco complicado de lograr, la tarea se simplifica por la poca exigente tasa de falsos positivos  $F$ , ya que  $6 \times 10^{-6} \approx 0.3^{10}$ , por lo que se puede permitir una tasa de falsos positivos de 30%.

El número esperado de características a evaluar en un proceso de medición está dado por la Ecuación 10.

$$N = n_0 + \sum_{i=1}^K (n_i \prod_{j=1}^i p_j)$$

**Ecuación 10:** Número de características

Donde  $K$  es el número de etapas o clasificadores,  $p_j$  es la tasa de detección del  $i$ -ésimo clasificador y  $n_i$  es la cantidad de características de  $i$ -ésimo clasificador.

El diseño y entrenamiento del sistema de detección de rostros implica que cuantas más características tenga un clasificador se obtendrá tasas más altas de detección así como mejores tasas de falsos positivos, pero cuanto más características se evalúen mayor tiempo de recursos computacionales se utilizará. Esto conlleva a que se debe minimizar  $N$ , para un  $D$  y  $F$  fijados, entre el número de etapas.

Una forma práctica de implementar es la siguiente:

- El usuario fija  $d_i$  y  $f_i$ , se entrena el clasificador utilizando AdaBoost, se van agregando características hasta alcanzar la performance requerida, se testea el sistema con un conjunto de validación. Si la tasa de falsos positivos no es alcanzada, se agrega otra etapa. Los ejemplos de imágenes negativas que no contienen un rostro para el entrenamiento se obtiene de recolectar todos los falsos positivos del clasificador anterior.

Una vez revisado el primer algoritmo de Viola&Jones para detección de rostros, a continuación se revisará el segundo algoritmo para detección de rostros utilizando la distancia de Hausdorff.

## 2.7 DISTANCIA DE HAUSDORFF

### 2.7.1 DEFINICIÓN

La distancia de Felix Hausdorff<sup>17</sup> es una medida que permite comparar conjuntos de puntos, usado en reconocimiento de patrones para determinar el grado de similitud entre dos objetos.

La distancia de Hausdorff proporciona una metodología matemática para determinar la semejanza de un conjunto de puntos con otro, examinando la fracción de puntos en un conjunto que se encuentran cerca de los puntos en el otro conjunto. Hay dos parámetros que se utilizan para decidir si dos conjuntos de puntos se parecen unos a otros usando esta distancia:

- 1) La distancia máxima que los puntos pueden separarse y todavía deben considerarse juntos.
- 2) La fracción de los puntos en un conjunto son como máximo esta distancia de los puntos del otro conjunto [16].

Las medidas de distancia basadas en Hausdorff difieren de las técnicas de correspondencia tales como los métodos de coincidencia de puntos<sup>18</sup> y la correlación binaria<sup>19</sup> porque no hay apareamiento de puntos en los dos conjuntos comparados.

---

<sup>17</sup> Felix Hausdorff fue un matemático alemán. Es considerado uno de los fundadores de la Topología moderna

<sup>18</sup> Método que trabaja en base a la coincidencia de puntos con un patrón

<sup>19</sup> Relación matemática definida entre los elementos de dos conjuntos

La distancia de Hausdorff es la distancia máxima de un conjunto A al punto más cercano del otro conjunto B, es una función máxima definida en la Ecuación 11.

$$h(A, B) = \max_{a \in A} \left\{ \min_{b \in B} (d(a, b)) \right\}$$

**Ecuación 11:** Distancia de Hausdorff

Donde a y b son puntos de los conjuntos A y B respectivamente y la distancia **d(ab)** es cualquier distancia entre estos puntos.

Por simplicidad se toma **d(ab)** como la distancia euclidiana entre a y b. Si por ejemplo A y B son dos conjuntos de puntos un algoritmo<sup>20</sup> sería.

Algoritmo1.

- $h = 0$
- para cada punto  $a_i$  de A
  - más corto = Inf ;
  - para cada punto  $b_j$  de B
 
$$d_{ij} = d(a_i, b_j)$$

Si  $d_{ij} < d$  más corta entonces

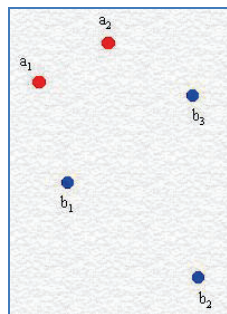
$$d \text{ más corta} = d_{ij}$$
  - Si  $d \text{ más corta} > h$  entonces
 
$$h = d \text{ más corta}$$

A continuación, se explica el algoritmo.

---

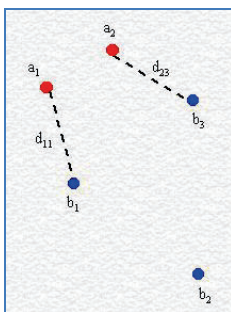
<sup>20</sup> Es la distancia entre dos puntos que se deduce a partir del teorema de Pitágoras.

En la Figura 2.9 se muestra un grupo de puntos correspondientes al conjunto A y otro grupo de puntos correspondiente al grupo B.



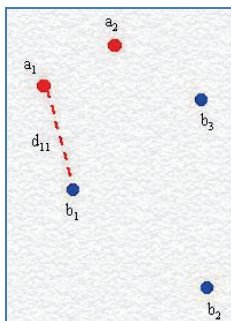
**Figura 2.9:** Conjunto de puntos de A y B [16]

De cada uno de los puntos del grupo A se busca la distancia más corta hacia uno de los puntos del grupo B obteniendo como resultado la Figura 2.10.



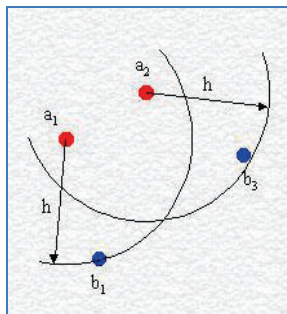
**Figura 2.10:** Distancias más cortas desde A hacia B [16]

De las dos distancias encontradas de los puntos A hacia los puntos B se busca la distancia mayor y esta se denomina  $h = (A, B)$  como se muestra en la Figura 2.11.



**Figura 2.11:**  $h = (A, B)$  Distancia de A hacia B [16]

Así se puede concluir que cualquier punto del grupo A se encuentra como máximo a la distancia  $h(A, B) = d(a_1, b_1)$  de algún punto de B como se muestra en la Figura 2.12.



**Figura 2.12:**  $h(A, B) = d(a_1, b_1)$  Distancia más lejana [16]

Con esto se puede señalar que la distancia de Hausdorff es con dirección así como también asimétrica<sup>21</sup>, lo que significa que la mayoría de las veces  $h(A, B)$  no es igual a  $h(B, A)$ . Esta condición general también es válida para el ejemplo de la Figura 2.11, como  $h(A, B) = d(a_1, b_1)$  no es igual a  $h(B, A) = d(b_2, a_1)$ . Esta asimetría es una propiedad de las funciones *maximin*<sup>22</sup>.

Una definición más general de la distancia de Hausdorff se define en la Ecuación 12.

$$H(A, B) = \max\{h(A, B), h(B, A)\}$$

**Ecuación 12:** Distancia de Hausdorff general

La Ecuación 12 define la distancia de Hausdorff *desde A hacia B* también llamada distancia Hausdorff dirigida. Las dos distancias  $h(A, B)$  y  $h(B, A)$  son algunas veces denominadas como distancias de Hausdorff hacia adelante y hacia atrás.

<sup>21</sup> Que no tiene correspondencia exacta en forma, tamaño y posición de las partes de un todo.

<sup>22</sup> Devuelve el valor mínimo o máximo de un conjunto de valores contenidos en un campo especificado



Si los conjuntos A y B están hechos de líneas o polígonos en lugar de puntos individuales entonces  $H(A, B)$  se aplica a todos los puntos que definen estas líneas o polígonos y no sólo a sus vértices, por tanto, el algoritmo 1 ya no podría utilizarse para calcular la distancia de Hausdorff entre tales conjuntos ya que implican un número infinito de puntos.

En los problemas de detección y reconocimiento de rostros, se permite a las dos imágenes experimentar algún tipo de transformación geométrica en el proceso de coincidencia. En este caso el objetivo es encontrar las transformaciones de una imagen que produzcan buenas coincidencias con la otra imagen.

### 2.7.2 DETECCIÓN DE ROSTROS BASADO EN LA DISTANCIA DE HAUSDORFF

Un enfoque basado en modelos geométricos que trabaja en imágenes fijas en escala de grises para la detección de rostros se basa en la distancia de Hausdorff. Este método realiza una detección de rostros sólida y precisa y su eficiencia lo hace adecuado para aplicaciones en tiempo real.

La distancia de Hausdorff (HD) utilizado para la detección de rostros de imágenes digitales plantea el siguiente problema.

Dada una imagen de entrada el algoritmo deberá decidir si hay un rostro en la imagen o no. De ser positiva la detección deberá devolver las coordenadas del rostro dentro de la imagen, que por lo general son los centros del ojo izquierdo y derecho, lo cual, es suficiente si el problema se restringe a los rostros de vista frontal.

Utiliza un método basado en bordes para detección de rostros primero se calcula una imagen de magnitud de borde con el operador de Sobel<sup>23</sup>. Los puntos de característica de borde relevantes son extraídos por un filtro local de umbral

---

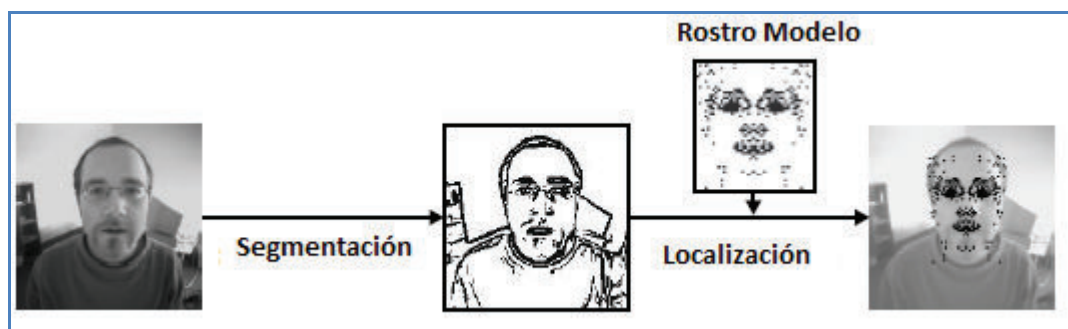
<sup>23</sup> El operador Sobel es utilizado en procesamiento de imágenes, especialmente en algoritmos de detección de bordes. Técnicamente es un operador diferencial discreto que calcula una aproximación al gradiente de la función de intensidad de una imagen.

adaptable para compensar la iluminación variable. Se asume que este procedimiento producirá una disposición característica de los puntos de segmentación en el área facial. Basado en el diseño de segmentación se utiliza un modelo de rostro que consiste en un conjunto de puntos y que puede representarse como una imagen binaria [79].

La característica de los puntos del modelo del rostro se elige de una manera que el patrón almacenado en el modelo sea de alguna manera similar a los patrones típicamente observados en el área del rostro de las imágenes.

Para detectar un rostro el modelo se superpone sobre la imagen en varias posiciones discretas. En cada posición se calcula la similitud entre el modelo traducido y la parte cubierta de la imagen donde se considera la posición que da la mayor similitud.

El procedimiento se ilustra en la Figura 2.13 y puede ser ajustado para permitir la detección de rostros de tamaños diferentes.



**Figura 2.13:** Procedimiento para buscar rostros [79]

Se utiliza una medida ligeramente adaptada denominada la distancia modificada de Hausdorff (MHD) [19] para calcular la similitud entre la imagen y el modelo. Dados los dos conjuntos de puntos  $A$  y  $B$ , la MHD se define como;  $A = \{a_1, \dots, a_m\}$  y  $B = \{b_1, \dots, b_n\}$  dos conjuntos de puntos finitos. Entonces, la distancia de Hausdorff se define en la Ecuación 13.

$$H(A, B) = \max(h(A, B), h(B, A)) \text{ donde}$$

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|, \text{ donde } a \in A \text{ y } b \in B$$

**Ecuación 13:** Distancia de Hausdorff dirigida

Por lo tanto,  $h(A, B)$  se llama la distancia Hausdorff dirigida desde el conjunto A a B con alguna norma subyacente  $\|\cdot\|$  en los puntos de A y B.

Para las aplicaciones de procesamiento de imágenes, existe una medida ligeramente diferente, la distancia de Hausdorff modificada (MHD), se define en la Ecuación 14.

$$h_{mod}(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \|a - b\|$$

**Ecuación 14:** Distancia de Hausdorff dirigida modificada

Tomando la media de las distancias de un solo punto, esta versión disminuye el impacto de los valores atípicos haciéndolo más adecuado para detección y reconocimiento de patrones.

Con la distancia de Hausdorff entendida el siguiente paso es revisar la detección de rostros que a continuación se detalla.

### 2.7.3 DETECCIÓN BASADO EN MODELOS GEOMÉTRICOS

Los conjuntos de puntos bidimensionales A y B son representaciones de la imagen y del objeto. Por lo tanto, cada punto del conjunto representa una característica de la imagen y por lo general puede ser un punto de borde del modelo geométrico como se muestra en la Figura 2.14 [21].

El objetivo se centra en encontrar los parámetros de transformación  $p \in P$  de manera que la distancia de Hausdorff entre el modelo transformado  $T_p(B)$  y A sea

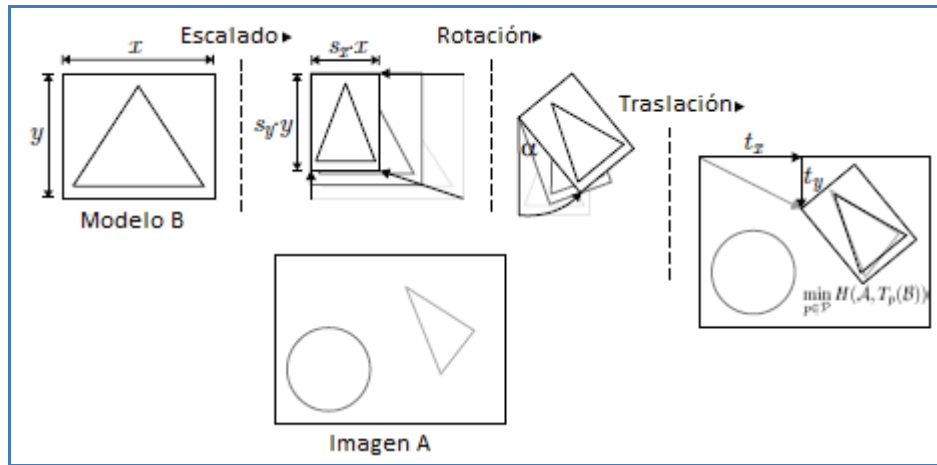
minimizada como se muestra en la Figura 2.14. La elección de las transformaciones permitidas como escala, traslado y su espacio de parámetros  $P$  depende de la aplicación. El cálculo de la distancia de Hausdorff eficiente permite una búsqueda exhaustiva en un espacio de transformación discretizado<sup>24</sup>.

El problema de detección puede formularse como se muestra en la Ecuación 15.

$$d_p = \min_{p \in P} h_{mod}(A, T_p(B))$$

**Ecuación 15:** Formula detección utilizando la distancia de Hausdorff

Entonces se llama  $h(T_p(B), A)$  a la distancia hacia adelante y  $h(A, T_p(B))$  la distancia inversa. Para considerar sólo la parte de la imagen que está cubierta por el modelo se reemplaza la distancia inversa por la distancia Box-inversa  $h_{box}$ .



**Figura 2.14:** Modelo de ajuste por escala, traslado y rotación [20].

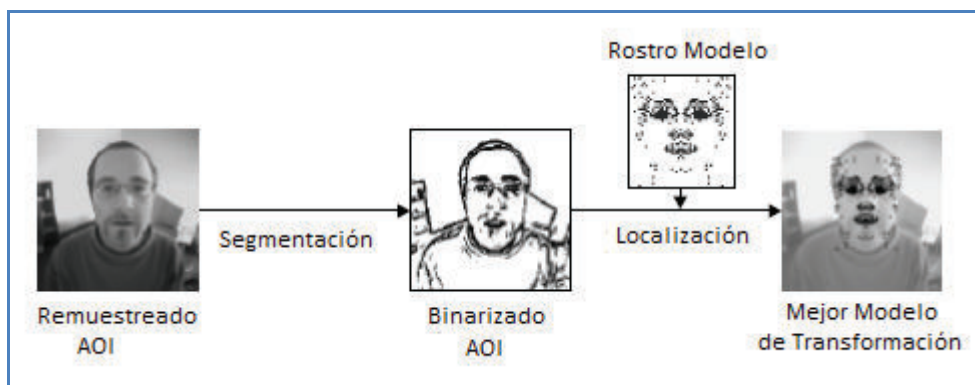
Para realizar la detección de rostros es importante conocer las etapas del algoritmo que se analizan a continuación.

<sup>24</sup> Dividir el todo en partes con menos complejidad, con el fin de facilitar los cálculos, el término se utiliza para expresar la individualización (discretización) de una unidad.

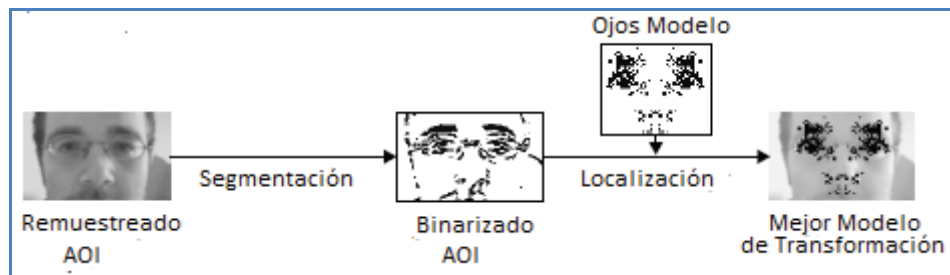
### 2.7.4 ETAPAS DEL ALGORITMO PARA LA DETECCIÓN [21]

El algoritmo de detección de rostros que utiliza la distancia de Hausdorff considera dos etapas; una detección gruesa o general que se muestra en la Figura 2.15 y una etapa de refinamiento que se muestra en la Figura 2.16.

Cada una de las etapas contiene fases de segmentación y localización internas que se revisarán más adelante, AOI<sup>25</sup> = área de interés.



**Figura 2.15:** Etapa de detección gruesa o general [20]



**Figura 2.16:** Etapa de detección de refinamiento [20]

#### 2.7.4.1 Detección gruesa o General

Antes de aplicar cualquier paso de segmentación se define un área de interés (AOI) con una relación de ancho/altura predeterminada para cada imagen entrante  $f$ . Esta AOI es re-muestreada a un tamaño fijo para ser independiente de las dimensiones de  $f$  [21].

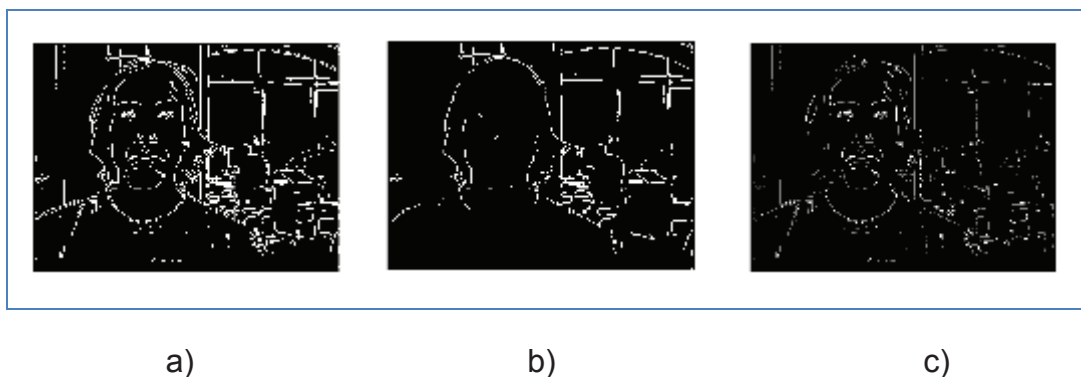
<sup>25</sup> AOI, Área de Interés de una imagen.

#### 2.7.4.1.1 Fase de segmentación

Primero se realiza un escalado de la imagen de entrada antes de que comience la fase de segmentación donde la imagen de entrada podía ser escalada hasta un 40% de su tamaño original sin afectar el resultado final de la detección.

La imagen de entrada debe ser transformada en una imagen con puntos en 2D que se consideran aristas del modelo geométrico. Esta imagen debería contener suficiente información de puntos para que el proceso de localización sea lo suficientemente robusto y más acertado. El operador extractor de aristas Sobel es el más usado en el algoritmo de detección de modelos geométricos basados en distancia de Hausdorff. Sin embargo, si se aplica un umbral bajo como se muestra en la Figura 2.17a se obtendrá como resultado una imagen con demasiadas aristas con muchos puntos, si se utiliza un umbral mayor como se muestra en la Figura 2.17b el número detectado de aristas disminuiría pero también desaparecerán puntos muy importantes [22].

Para conseguir una segmentación adecuada, en la que no haya tantos segmentos correspondientes al fondo y no se pierdan los puntos correspondientes a los elementos del rostro, se calcula la diferencia entre la imagen obtenida tras aplicar un extractor de aristas Sobel con un umbral bajo y la imagen obtenida tras aplicar el mismo extractor con un umbral alto como se muestra en la Figura 2.17c.



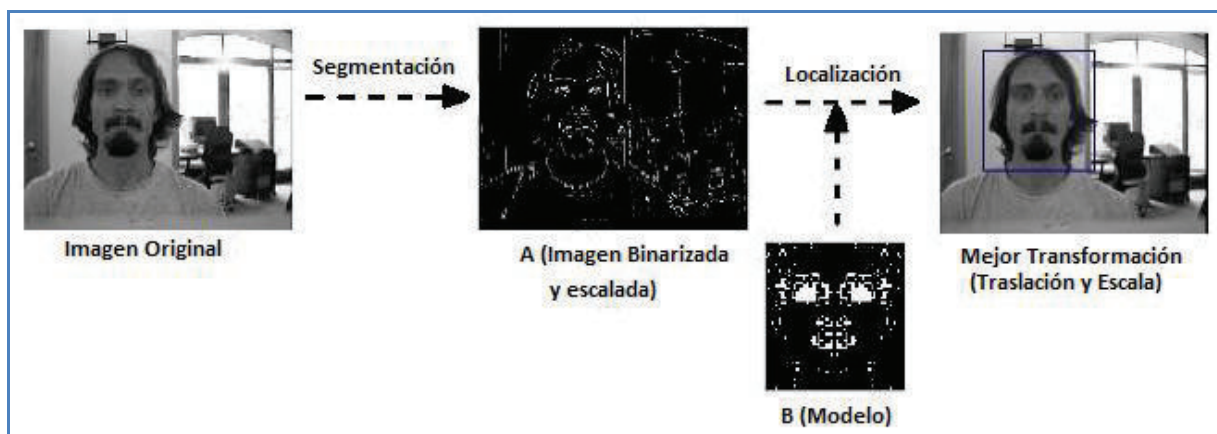
**Figura 2.17:** Detección de segmentos

- a) Detección de segmentos mediante el operador de Sobel con un umbral de 0.04
- b) Detección de segmentos mediante el operador de Sobel con un umbral de 0.09
- c) Detección de segmentos tras combinar los dos resultados anteriores [22].

Una vez terminada la fase de segmentación se analiza la fase de locación como se muestra a continuación.

#### 2.7.4.1.2 Fase de localización

Como resultado de la fase de segmentación el algoritmo dispone de una imagen en 2D escalada donde se pueden encontrar diferentes patrones de puntos que podrían representar a un rostro como se muestra en la Figura 2.18. Esta plantilla de puntos se creó como la media de un conjunto de rostros a partir de patrones genéticos, puntos que representan elementos faciales como ojos, boca y nariz correspondiéndose algunos otros con los límites del rostro [22].



**Figura 2.18:** Esquema del funcionamiento de la fase de localización facial [22].

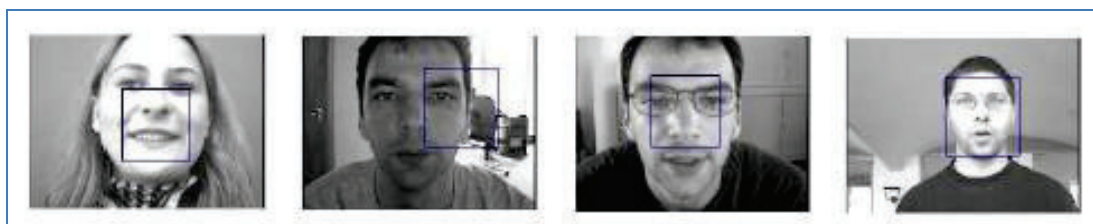
A partir de la plantilla B que se muestra en la Figura 2.18 que representa un rostro medio y la imagen A 2D que se muestra en la Figura 2.18 obtenida tras la segmentación, se busca el rostro que está presente en la imagen utilizando la Ecuación 15. Para ello, se aplican diferentes valores para los parámetros de escala y

traslación de la plantilla B. Los valores del conjunto de posibles transformaciones  $\hat{p}$  que minimizan  $H(A, T_p(B))$  proporcionan la posición y el tamaño del rostro.

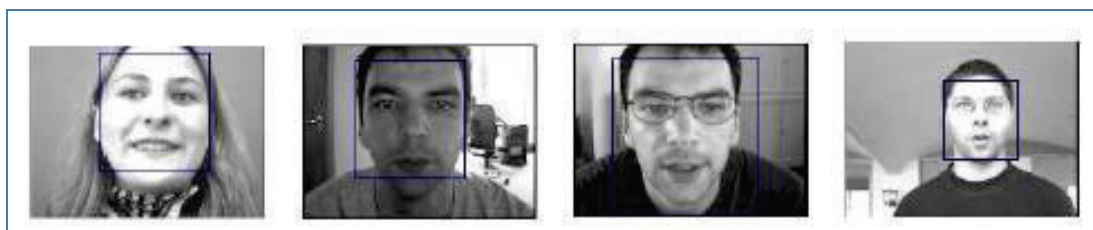
La escala de la plantilla B que se muestra en la Figura 2.18 es uno de los principales problemas prácticos a la hora de aplicar la distancia de Hausdorff, en el procesamiento de imágenes. Debido a que el tamaño de los rostros que se podrán encontrar en las imágenes de entrada puede ser muy variable, durante la búsqueda se deberían probar un gran rango de valores de escala  $(s_x; s_y)$  para la plantilla B.

Los resultados de buscar una escala adecuada se pueden observar en la Figura 2.19a. Por lo que es necesario aplicar un factor de normalización para adaptar la distancia de Hausdorff a diferentes tamaños de plantilla.

Este factor de normalización debe ser seleccionado con cuidado, pues modificar demasiado la distancia puede producir el efecto contrario, y sería muy complicado detectar los rostros más pequeños.



a) Detección de rostros sin factor de normalización



b) Detección de rostros usando factor de normalización

**Figura 2.19:** Problema de escala. a) Detección de rostros sin factor de normalización. b) Detección de rostros usando factor de normalización [22].



Después de un proceso de experimentación se determinó que al dividir  $h(T_p(B), A)$  y  $h(A, T_p(B))$  entre  $\sqrt{\frac{s_x + s_y}{2}}$  el algoritmo detecta la posición y el tamaño de los rostros de forma más efectiva.

#### 2.7.4.2 Refinamiento

Basado en el conjunto de parámetros  $\hat{p}$ , se define una segunda AOI = área de interés que cubre el área esperada del rostro. Esta área se vuelve a muestrear a partir de la imagen original  $f$ , dando como resultado una imagen en escala de grises  $h$  del área del rostro. La segmentación y localización son iguales a la etapa de detección gruesa excepto que se utiliza una plantilla B' más detallado de la región ocular.

Los valores de la distancia modificada de Box-inversa  $h_{box}(A', T_{\hat{p}}(B'))$  en la mejor posición, especialmente cuando se multiplica con la distancia hacia adelante modificada  $h(T_{\hat{p}}(B'), A')$ , Esto es útil si se desea una decisión rostro/sin rostro.

Las posiciones oculares se calculan a partir de los conjuntos de parámetros  $\hat{p}$  y  $\hat{p}'$ .

En comparación con las posiciones manuales fijadas de los ojos son usadas para clasificar la calidad del algoritmo [21].

Una vez analizados los dos algoritmos para detección de rostros actualmente más utilizados se realiza una Tabla comparativa de las principales características.

## 2.8 TABLA COMPARATIVA DE LOS ALGORITMOS

A continuación, se detalla en la Tabla 2.1 una comparación de los dos algoritmos presentados en este capítulo para la detección de rostros.

**Tabla 2.1:** Comparación entre los algoritmos Viola&Jones y la distancia de Hausdorff

CARACTERÍSTICA	ALGORITMO BASADO EN CASCADAS DE VIOLA-JONES	ALGORITMO BASADO EN MODELOS USANDO LA DISTANCIA DE HAUSDORFF
<b>Método</b>	El algoritmo divide la imagen integral en subregiones de tamaños diferentes y utiliza una serie de clasificadores en cascada, cada uno con un conjunto de características visuales.	El método está basado en bordes y funciona en imágenes fijas en escala de grises. La distancia de Hausdorff se utiliza como medida de similitud entre un modelo de cara general y las posibles ilustraciones del objeto dentro de la imagen.
<b>Etapas</b>	Usa cuatro etapas: la primera y segunda de ellas son conocidos como la Imagen Integral que permite que las características tipo Haar utilizadas por este detector se puedan computar muy rápido. La tercera es un algoritmo de aprendizaje automático, Adaboost, que selecciona sólo las características importantes de todo el conjunto. La cuarta es la creación de una estructura en cascada.	Se presenta un proceso de dos pasos que permite tanto la detección gruesa (segmentación y localización) y el refinamiento exacto de las caras (segmentación y localización).
<b>Uso de la imagen</b>	No utiliza directamente la imagen sino que utiliza una representación de la imagen llamada imagen integral. Procesa sólo la información presente en una imagen en escala de grises.	Usa imágenes fijas en escala de grises. La imagen de entrada debe ser transformada en una imagen binaria de aristas.
<b>Tiempo de procesamiento</b>	Tiene un tiempo de procesamiento muy bajo, en el orden de milisegundos.	El tiempo de procesamiento de la distancia de Hausdorff depende del número de puntos presente en los conjuntos A y B por lo que se realiza un escalado de la imagen de entrada de hasta el 40% de su tamaño original para aumentar la velocidad del sistema.
<b>Verdaderos positivos</b>	Tiene una probabilidad de verdaderos positivos del 99,9%	Verdaderos positivos aproximados del 97,7%
<b>Falsos positivos</b>	Tiene una probabilidad de falso positivos del 3,33%	Falsos positivos aproximados del 2,22%
<b>Factores adversos</b>	La imagen durante el proceso de detección puede suponer algunos problemas cuando la luminosidad no es la adecuada, aparecen elementos extraños, las caras están de perfil, tapadas por algún elemento o por alguna otra cara o en un ángulo complicado	La imagen presenta problemas cuando la luminosidad no es la adecuada y existen fondos complejos
<b>Ventajas</b>	Este detector se ha hecho muy popular debido a su velocidad a la hora de detectar las caras en imágenes y para su implementación. Cuenta con una selección eficiente de características, en lugar de escalar la imagen escala las características	Los buenos resultados de localización muestran que el sistema es robusto frente a diferentes condiciones de fondo y cambiando la iluminación. El comportamiento en tiempo de ejecución permite su uso en aplicaciones de video en tiempo real
<b>Desventajas</b>	El detector es más efectivo sólo en imágenes frontales de caras y presentar dificultades con una rotación de la cara a 90° alrededor del eje vertical y horizontal. Se puede obtener múltiples detecciones de la misma cara, debido a la superposición de sub-ventanas	Los resultados de este algoritmo de detección se ven muy afectados en el caso de que existan fondos complejos y rotaciones de caras
<b>Aplicaciones</b>	Entretenimiento (videojuegos, realidad virtual, humano-robótica, maquillaje virtual, aplicaciones de <i>smartphone</i> ) Tarjetas Inteligentes (licencia de conducir, pasaportes, dispositivo personal de acceso, seguridad de base de datos) video vigilancia (portal de control, seguimiento sospechoso e investigación)	La detección de rostros es sólida y precisa y su eficiencia lo hace adecuado para aplicaciones en tiempo real. Aplicaciones biométricas como sistemas de video-vigilancia o reconocimiento facial.

Finalmente, se puede concluir que los dos algoritmos cuentan con altas prestaciones en cuanto a la detección de un rostro y ciertas desventajas frente a factores externos como iluminación, sin embargo, la detección de rostros propuesto por Paul Viola y Michael Jones, es uno de los métodos más usados en la actualidad ya que permite detectar rostros en una imagen con tiempos de procesamiento bajos, es decir, a una mejor velocidad.

Con el análisis de los dos algoritmos, en el capítulo 3 se llevará a cabo la implementación del código de Viola&Jones a un ambiente de simulación en interiores controlado para detección de rostros en tiempo real en el programa MATLAB y se realizará un análisis del mismo para determinar su tasa de detección, falsos positivos y falso negativos que determinarán % de efectividad.

## **CAPÍTULO 3**

### **IMPLEMENTACIÓN DEL ALGORITMO Y ANÁLISIS DE RESULTADOS**

#### **3.1 INTRODUCCIÓN**

Existen muchos programas computacionales que sirven como herramientas de análisis e investigación donde se puede programar y simular ambientes de operación de sistemas electrónicos muy acercados a la realidad.

En el presente capítulo se presenta el uso de *Matlab* para realizar la implementación del algoritmo para detección de rostros que presentó las mejores características en el análisis del capítulo anterior.

Adicionalmente se realizará la comparación con una herramienta comercial disponible para detección y seguimiento de rostros.

Finalmente el capítulo abordará la recopilación de los resultados de la ejecución del código y el análisis de los mismos.

#### **3.2 SOFTWARE DE SIMULACIÓN**

Se selecciona MATLAB R2015b ya que proporciona un entorno de desarrollo integrado (IDE), porque es una aplicación informática que proporciona servicios integrales para facilitar el desarrollo al programador a través de un editor de código fuente, herramientas de construcción automáticas, un depurador e interfaz gráfica de usuario (GUI) [81].

MATLAB cuenta con un lenguaje de programación propio (lenguaje M) y está disponible para sistemas Unix, Windows, Mac OS X y GNU/Linux considerando algunos requerimientos mínimos de recursos del sistema para la versión utilizada R2015b [81].

- *Windows.- Windows 7 SP1 en adelante, Windows Server 2008 SP2 en adelante.*
- *Mac.- macOS 10.10 - 10.11.*
- *Linux.- Kernel 2.6 o superior, glibc 2.11 o superior.*
- *Procesador.- Intel o AMD x86-64 con soporte de instrucciones AVX2.*
- *Disco.- 2 GB solo para MATLAB, 4-6 GB para una instalación típica.*
- *RAM.- 1GB mínimo, 4 GB recomendado.*
- *Tarjeta gráfica.- Soporte para OpenGL 3.3 recomendado con 1 GB en GPU.*

MATLAB simplifica la operación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos [81].

MATLAB cuenta con dos herramientas adicionales como: Simulink (plataforma de simulación multi-dominio) y GUIDE (editor de interfaces de usuario - GUI). Además se pueden ampliar sus capacidades utilizando su conjunto de herramientas (*toolboxes*) y paquetes de bloques (*blocksets*) [81].

Herramientas y paquetes de bloques para Simulink:

- Modelado basado en eventos.
- Modelado Físico.
- Sistemas de control.
- Procesamiento de Señales y Comunicaciones.
- Código de Generación.
- Simulación y pruebas en tiempo real.
- Verificación, validación y prueba.
- Simulación Gráficos e Informes.

Es necesaria la revisión de estas herramientas y paquetes debido a su utilidad en las siguientes etapas del presente capítulo, como el ambiente de simulación que a continuación se detalla.

### 3.3 AMBIENTE DE SIMULACIÓN

La simulación tiene como principal objetivo anticiparse a la implementación física del sistema consiguiendo mostrar lo que posiblemente sucederá bajo determinadas condiciones internas y externas.

La simulación involucra dos facetas:

- 1) Construir y contar con todos los recursos intervinientes en el modelo del ambiente de simulación.
- 2) Ensayar diversas alternativas con el fin de elegir y adoptar la mejor a la que responde el sistema, procurando que sea la óptima o que por lo menos sea lo suficientemente aproximada a su máximo desempeño [82].

Los recursos que intervienen directamente en el ambiente de simulación se detallan a continuación:

- *Software* de simulación Matlab R2015b.
- Computadora con características mínimas requeridas con las que se garantiza que el *software* de simulación aprovechará todas sus funcionalidades.
- *Webcam* VGA integrada del dispositivo.
- Salón de pruebas de área requerida para pruebas a ciertas distancias de detección de rostros.
- Personas para pruebas en múltiples escenarios o alternativas de pruebas del algoritmo.
- Otros objetos con rostros para pruebas.

Existen un conjunto de alternativas para un ambiente de simulación, que constituyen la estrategia de simulación, cuyo objetivo es variar considerablemente los factores colocando posibles adversidades en las que el sistema podría fallar. Si el sistema supera favorablemente estas adversidades se concluirá que sobrepasa las

exigencias de un ambiente de instalación ideal donde aprovechará al 100% sus aplicaciones y funcionalidades.

Un ambiente de simulación cuenta con factores internos/externos como iluminación, temperatura, humedad relativa, los cuales son controlados o regulados para alcanzar condiciones ambientales normales.

A continuación, se detallan los parámetros de los factores internos/externos en condiciones normales.

- 1) Iluminación.- Dentro de un ambiente controlado se considera que el rango de luminosidad debe estar entre 20 Lux<sup>26</sup> a 10K Lux como se muestra en la Tabla 3.1.

Nivel de Iluminación	Área de Iluminación
1 Lux	Luna llena a gran altitud
50 Lux	Sala de vivienda familiar
80 Lux	Pasillo, Cuarto de Baño
320 Lux	Luz recomendada en oficina
400 Lux	Salida o puesta de sol en un día despejado
1000 Lux	Luz habitual en un estudio
10 – 25 KLux	Luz del día despejado
32 – 130 KLux	Luz directa hacia la cámara

**Tabla 3.1:** Ejemplos de Intensidad Luminosa [83]

- 2) Temperatura.- El rango de temperatura en el ambiente de pruebas depende directamente de la tolerancia de los dispositivos del sistema de videovigilancia

---

<sup>26</sup> Unidad de medida de la iluminancia

que generalmente varían entre los  $-30^{\circ}\text{C} \sim 50^{\circ}\text{C}$ . Sin embargo un rango aceptable de valores va desde  $6^{\circ}\text{C}$  a  $30^{\circ}\text{C}$  [88].

- 3) Humedad Relativa.- La humedad se refiere a la cantidad de vapor de agua que está presente en el aire. El rango puede variar entre el 10% y el 100% [88].

Hay que considerar que hoy en día los dispositivos de los sistemas de videovigilancia cuentan con sensores de lectura de los factores internos/externos, y que de alguna forma regulan estas diferencias automáticamente para mantenerse dentro de los valores de un ambiente controlado.

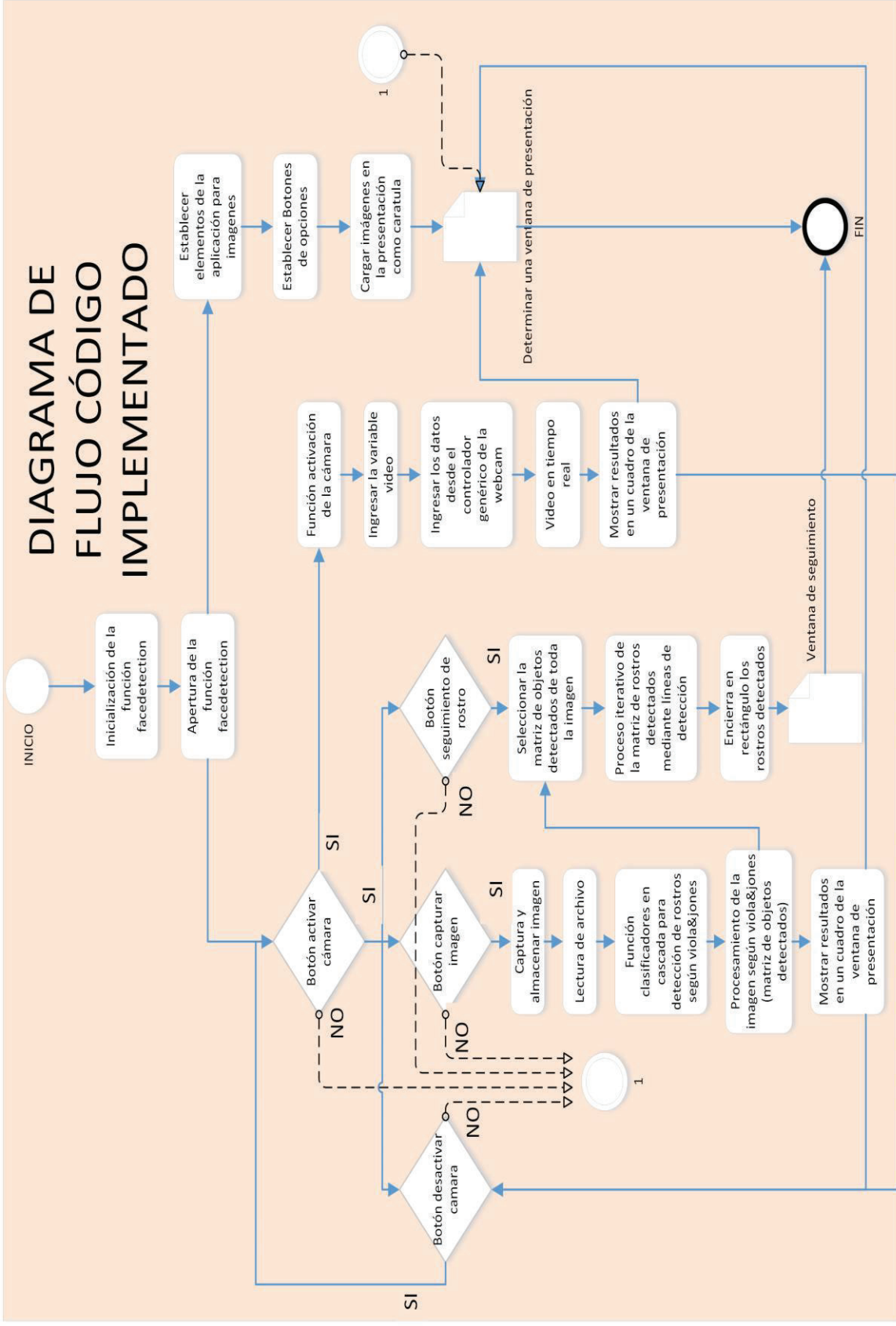
A continuación, se detallará el desarrollo del código para detección de rostros en MATLAB.

### **3.4 DESARROLLO DEL CÓDIGO**

Se ha desarrollado el código empleando clasificadores en cascada del algoritmo de Viola&Jones porque presenta las mejores características acorde al análisis realizado en el capítulo anterior frente al algoritmo de modelos geométricos que utiliza la distancia de Hausdorff.

La estructura del código está representada mediante el diagrama de flujo de la Figura 3.1, el cual permitirá la detección de múltiples rostros en tiempo real de un video que es generado por la cámara webcam integrada de una computadora en un ambiente controlado.





**Figura 3.1:** Diagrama de flujo del código implementado

El diagrama de flujo muestra la forma de ejecución del código desarrollado iniciando con la apertura de la función para detección de rostros y el establecimiento de una ventana de la interfaz gráfica de usuario con botones, recuadros y paneles.

Se crea la subfunción `activar_cámara` que consulta mediante un botón, si el usuario decide señalar el botón se puede visualizar en un panel el video en tiempo real y al mismo tiempo se despliegan botones de la subfunciones `desactivar_cámara`, `capturar_imagen` y `seguimiento de rostro`.

La subfunción `desactivar_cámara` permite que el usuario pueda regresar al estado anterior de activar la cámara deteniendo el video en tiempo real mostrado en el panel de la ventana de presentación.

La subfunción `capturar_imagen` se captura y almacena imágenes del video en tiempo real para que serán leídas y analizadas para detectar rostros mediante clasificadores en cascada del algoritmo de `viola&jones`. Posteriormente se realiza el procesamiento de los rostros detectados en una matriz para mostrar los resultados en el panel de la ventana de presentación.

La subfunción `seguimiento de rostros` se muestra en una nueva ventana de presentación donde los rostros detectados son seguidos y encerrados en un rectángulo.

Entonces el código para detección de rostros de acuerdo a la estructura del diagrama de flujo de la Figura 3.1 se muestra a continuación.

```
%-----ESCUELA POLITÉCNICA NACIONAL-----
%-----INGENIERIA EN ELECTRONICA Y TELECOMUNICACIONES---
%----- TRABAJO DE TITULACIÓN-----
%-- Tema: Detección de rostros -----
% Autor: Cristian Rene Sánchez Salazar -----
%*****
```

`function` varargout = faceDetection(varargin) % FACEDETECTION código MATLAB para faceDetection.fig, crea una nueva función FACEDETECTION o eleva el singleton existente \*.

`gui_Singleton` = 1; %Es una opción (0 o 1) que permite que solo se inicie una o varias copias de la interfaz gráfica de usuario al mismo tiempo. En este caso se escoge una sola copia.

%Se crea la siguiente estructura de la función principal, esta estructura contiene información de control, funciones manejadoras para algunas sub-funciones de la función principal.

```
gui_State = struct('gui_Name', mfilename, ...
    'gui_Singleton', gui_Singleton, ...
    'gui_OpeningFcn', @faceDetection_OpeningFcn, ...
    'gui_OutputFcn', @faceDetection_OutputFcn, ...
    'gui_LayoutFcn', [] , ...
    'gui_Callback', []);
if nargin && ischar(varargin{1})
    gui_State.gui_Callback = str2func(varargin{1});
end
```

%La función principal llama a la función Matlab `gui_mainfcn` y pasa la estructura y todos los argumentos de entrada a ella. Principalmente esta función crea la interfaz gráfica de usuario o llama a la sub-función adecuada cuando se le llama.

```
if nargin
    [varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
    gui_mainfcn(gui_State, varargin{:});
end
% Fin de la inicialización
```

% En la siguiente sección del código se realiza la apertura de la función `faceDetection` donde se ingresa los primeros argumentos que serán parte de la ventana de presentación del código.

`function` faceDetection\_OpeningFcn(hObject, eventdata, handles, varargin) % --- Se ejecuta la apertura de la función `faceDetection` justo antes de que se haga visible, Esta función no tiene argumentos de salida.

% Inicialización de los elementos de la aplicación -----

`axes(handles.axes1);` % Esta sentencia permite seleccionar el `axes1` sobre el que se va a trabajar en ese instante. Los `axes` son espacios de la interfaz gráfica de usuario para colocar las imágenes que se necesitan.

`axis off;` %desactiva los ejes de los `axes`

`axes(handles.axes2);` % Esta sentencia permite seleccionar el `axes2`

`axis off;` %desactiva los ejes de los `axes`

`axes(handles.axes3);` % Esta sentencia permite seleccionar el `axes3`

`axis off;` %desactiva los ejes de los `axes`

% Sentencias para colocar y desactivar de botones, axes y paneles para que no aparezcan en la ventana inicial

```
set(handles.desactivar_camara,'Visible','off'); % Botón
set(handles.capturar_imagen,'Visible','off'); % Botón
set(handles.face_tracking,'Visible','off'); % Botón
set(handles.axes1,'Visible','off'); % Axes
set(handles.axes2,'Visible','off'); % Axes
set(handles.axes3,'Visible','off'); % Axes
set(handles.panel_camara,'Visible','off'); % Panel
set(handles.panel_imagen,'Visible','off'); % Panel
set(handles.panel_face,'Visible','off'); % Panel
```

handles.output = hObject; % Elije la salida de línea de comandos predeterminada para faceDetection

guidata(hObject, handles); % Actualiza la estructura de manejadores

% UIWAIT o UIRESUME hace que faceDetection espere a la respuesta del usuario (vea UIRESUME) en Uwait (handles.figure1);

% Después de haber ingresado los Axes para la ventana de presentación donde se mostrará las regiones de las imágenes, se procede a ingresar la función de salida de argumentos para que estos puedan ser mostrados en la interfaz gráfica de usuario.

% --- Las salidas de esta función se devuelven a la línea de comandos.

```
function varargout = faceDetection_OutputFcn(hObject, eventdata, handles)
% Varargout matriz de celdas para devolver args de salida (ver VARARGOUT);
% HObject manejar a Figura
% eventdata reserved - que se definirá en una versión futura de MATLAB
% Maneja la estructura con manejadores y datos de usuario (vea GUIDATA)
```

```
axes(handles.axes9) % Esta sentencia permite seleccionar el axes9
imagenlogo = imread('foto1.jpg'); % Lee la imagen almacenada llamada foto1
axis off; %desactiva los ejes de los axes
imshow(imagenlogo); % Muestra la imagenlogo en el axes señalado
```

```
axes(handles.axes7) % Esta sentencia permite seleccionar el axes7
imagenlogo = imread('foto2.jpg'); % Lee la imagen almacenada llamada foto2
axis off; %desactiva los ejes de los axes
imshow(imagenlogo); % Muestra la imagenlogo en el axes señalado
```

```
axes(handles.axes8) % Esta sentencia permite seleccionar el axes8
imagenlogo = imread('foto3.jpg'); % Lee la imagen almacenada llamada foto3
axis off; %desactiva los ejes de los axes
imshow(imagenlogo); % Muestra la imagenlogo en el axes señalado
```

varargout{1} = handles.output; % Obtiene salida de la línea de comandos desde la estructura de controladores.

% Una vez establecida la presentación y argumentos de salida de la función, se llama a la sub-función activar\_cámara donde esta permitirá declarar una variable de vid que se utilizará en varias sub-funciones posteriores.

```
function activar_camara_Callback(hObject, eventdata, handles)
```

% Se activa botones y paneles en la aplicación -----

```
set(handles.desactivar_camara,'Visible','on'); % Botón
```

```
set(handles.capturar_imagen,'Visible','on'); % Botón
```

```
set(handles.axes1,'Visible','on'); % Axes
```

```
set(handles.panel_camara,'Visible','on'); % Panel
```

```
set(handles.panel_imagen,'Visible','on'); % Panel
```

```
set(handles.face_tracking,'Visible','on'); % Botón
```

global vid %defina la variable "vid" para que sea una variable global y pueda ser utilizada en diferentes funciones

```
axes(handles.axes1); activación de axes1
```

```
axis off; %desactiva los ejes de los axes
```

```
vid=videoinput('winvideo',1,'YUY2_640x480'); % "videoinput" crea un objeto de entrada de video (objeto de adquisición de imagen).
```

% Representa la conexión entre Matlab y un dispositivo de adquisición que es la cámara del computador. El adaptador de la cámara en este caso es de tipo genérico y se denomina "winvideo".

```
vid.ReturnedColorspace='rgb'; % Propiedad (método) que especifica el espacio de color en Matlab, en este caso MATLAB RGB color space.
```

```
vidRes=get(vid,'VideoResolution'); % Selecciona la resolución de video 640x480
```

```
nBands=get(vid,'NumberOfBands'); % Propiedad de adquisición del número de bandas de la imagen a ser adquirida, propiedad relacionada con la adquisición del video.
```

```
hImage=image(zeros(vidRes(2),vidRes(1),nBands)); % Muestra la imagen
```

```
preview(vid,hImage); % Muestra la vista de la información correspondiente al video en tiempo real.
```

% --- Botón desactivar\_cámara, sub-función desactivar\_cámara.

```
function desactivar_camara_Callback(hObject, eventdata, handles)
```

```
axes(handles.axes1); activación de axes1
```

```
axis off; % Desactiva los ejes de los axes
```

```
closepreview % cierra la adquisición del video en tiempo real
```

```
set(handles.axes1,'Visible','off'); % Oculta axes 1
```

```
set(handles.capturar_imagen,'Visible','off'); % Oculta botón
```

```
set(handles.face_tracking,'Visible','off'); % Oculta botón
```

% --- Botón en capturar\_imagen. Sub-función capturar\_imagen

```
function capturar_imagen_Callback(hObject, eventdata, handles)
```

% Se realiza la captura y almacenamiento de la imagen para luego leer el archivo y realizar la detección de rostros mediante clasificadores en cascada del algoritmo de viola&jones. Una vez realizada esta detección se procesa la información en una matriz de objetos detectados para mostrar en una región del cuadro de presentación.

global vid % Variable video de entrada de la webcam.

axes(handles.axes1) % activación de axes1

foto=getsnapshot(vid); %obtención de un snapshot del axes1, correspondiente al video en tiempo real.

axes(handles.axes2); % activación de axes2

imshow(foto) % Muestra la imagen en el axes2

imwrite(foto,'foto.jpg','jpg'); % Graba la imagen en un archivo

set(handles.panel\_face,'Visible','on'); % Activa el panel "panel\_face"

faceDetector = vision.CascadeObjectDetector(); % Detector de rostros en cascada que utiliza el algoritmo de Viola&Jones (Computer Vision System Toolbox)

data = imread('foto.jpg'); % Lectura del archivo grabado

bboxx = step(faceDetector, data); % Procesa la imagen utilizando el algoritmo Viola&Jones bboxx variable en la que se obtiene una matriz que contiene los rostros detectados.

axes(handles.axes3) % Activación de axes3

imshow(data) % Muestra la imagen considerando la detección de rostros

if(~isempty(bboxx))

hold on; % Mantiene la imagen anterior si ningún rostro es detectado

for i=1:size(bboxx,1) % Porción de código que hace referencia al recorrido de toda la imagen en una matriz

bbox=bboxx(i,:);

rc=bbox+[-bbox(3)/4,-bbox(4)/4,bbox(3)/2,bbox(4)/2];

ht=round(rc(4)/20);

wd=round(rc(3)/20);

xm=round(rc(1)+(rc(3)/2));

ym=round(rc(2)+(rc(4)/2));

rectangle('Position',rc,...

'Curvature',0,...

'LineWidth',2,...

'LineStyle','--',...

'EdgeColor','y')

% Crea el recuadro que encierra el rostro detectado.

% El comando line crea una línea

line([xm,xm],[rc(2),rc(2)+ht],... %

'LineWidth',2,...

'Color','y');

line([rc(1),rc(1)+wd],[ym,ym],...

'LineWidth',2,...

'Color','y');

```

line([xm,xm],[rc(2)+rc(4),rc(2)+rc(4)-ht],...
'LineWidth',2,...
'Color','y');
line([rc(1)+rc(3),rc(1)+rc(3)-wd],[ym,ym],...
'LineWidth',2,...
'Color','y');
line([rc(1),rc(1)+wd],[rc(2),rc(2)],...
'LineWidth',4,...
'Color','y');
line([rc(1),rc(1)],[rc(2),rc(2)+ht],...
'LineWidth',4,...
'Color','y');
line([rc(1)+rc(3),rc(1)+rc(3)],[rc(2),rc(2)+ht],...
'LineWidth',4,...
'Color','y');
line([rc(1)+rc(3),rc(1)+rc(3)-wd],[rc(2),rc(2)],...
'LineWidth',4,...
'Color','y');
line([rc(1),rc(1)],[rc(2)+rc(4)-ht,rc(2)+rc(4)],...
'LineWidth',4,...
'Color','y');
line([rc(1),rc(1)+wd],[rc(2)+rc(4),rc(2)+rc(4)],...
'LineWidth',4,...
'Color','y');
line([rc(1)+rc(3),rc(1)+rc(3)-wd],[rc(2)+rc(4),rc(2)+rc(4)],...
'LineWidth',4,...
'Color','y');
line([rc(1)+rc(3),rc(1)+rc(3)],[rc(2)+rc(4),rc(2)+rc(4)-ht],...
'LineWidth',4,...
'Color','y');
end
hold off; % Bloquea la imagen anterior si un rostro es detectado
end % Fin de la detección de rostros

% --- botón seguimiento de rostros en tiempo real de toda la imagen. Sub-función
face_tracking.
function face_tracking_Callback(hObject, eventdata, handles)
face_tracking
% En la ventana de presentación se puede escoger la opción de seguimiento de
rostros en tiempo real al utilizar la matriz de objetos detectados de toda la imagen
con la creación de líneas como referencia que se mostraran encerrando al rostro
en recuadro independiente a la ventana de presentación.

```

**Código 1:** Código para detección de rostros en tiempo real implementado

Una vez detallado el desarrollo del código se puede pasar a la siguiente etapa de ejecución detallada a continuación.



### 3.5 EJECUCIÓN DEL CÓDIGO EN MATLAB

#### 3.5.1 INICIALIZACIÓN Y PRESENTACIÓN DEL CÓDIGO

Para ejecutar el código es necesario revisar los factores del ambiente controlado y los recursos intervinientes con los cuales se realizarán los ensayos en diversos escenarios para obtener los mejores resultados.

Factores internos revisados:

- Revisión de Iluminación de la sala en dos aspectos.- Con luces apagadas y persianas cerradas muestra baja iluminación pudiendo confundir al sistema con sombras. Con luces encendidas y persianas abiertas muestra una iluminación adecuada normal.
- Revisión de la temperatura en la sala adecuada a las 3 pm cuando se realizaron las pruebas como se muestra en la Figura 3.2.



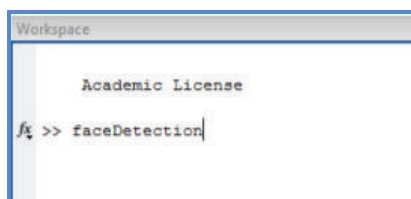
**Figura 3.2:** Registro de temperatura en la sala de pruebas

Los recursos intervinientes revisados:

- Laptop HP core I3/ Windows 7, con el software MATLAB 2015b instalado.
- Webcam VGA operativa incorporada de la laptop.
- Disponibilidad de la sala en el horario de la tarde.
- Tres personas con disponibilidad para hacer detecciones de sus rostros.
- Hojas impresas con rostros completos o incompletos con algunas características.

Con los factores y recursos debidamente seleccionados se ejecuta el archivo FaceDetection.m de Matlab. Se escribe el nombre del código para detección de rostros en la ventana de comandos como se muestra en la Figura 3.3.





**Figura 3.3:** Ventana de comandos corriendo código implementado

Posteriormente aparecerá la ventana de presentación como se muestra en la Figura 3.4.



**Figura 3.4:** Ventana de presentación del trabajo realizado

En la ventana de presentación se mostrará un botón para activar la cámara web integrada a la computadora para visualizar el video en tiempo real.

Adicional a esto se mostrarán los argumentos que se insertaron en la función activar cámara con la que se llamará posteriormente a las subfunciones desactivar\_cámara, capturar\_imagen y seguimiento\_de\_rostros como se muestra en la Figura 3.5.



**Figura 3. 5:** Presentación completa del trabajo realizado

A continuación, se describirán cada una de las subfunciones cuando son elegidas por el usuario.

- 1) La opción de desactivar\_cámara puede suspender el código regresándolo al estado anterior de activar\_cámara y en caso de que se desee cerrar la aplicación sin problemas.
- 2) La opción de capturar\_imagen toma una foto del video en tiempo real de la cámara *web*, la muestra en la región de imagen capturada de la presentación y la procesa a través del algoritmo de detección de rostros de Viola & Jones mostrando su resultado en la región de rostro detectado de la presentación.
- 3) La opción de seguimiento\_de\_rostro despliega una nueva ventana donde se detecta un rostro en tiempo real y realiza el seguimiento.

### **3.5.2 EJECUCIÓN DEL CÓDIGO PARA DETECCIÓN DE ROSTROS**

Como se indicó en el capítulo 2 los parámetros o indicadores para determinar un buen desempeño de un algoritmo son: la tasa de detección, tasa de falsos positivos y la tasa de falsos negativos encontrados en una imagen al momento de ejecutarlo.

En la ejecución del código se evaluarán varios escenarios de pruebas para determinar su comportamiento en ambientes con muy poca iluminación e iluminación normal.

A continuación se presentan las pruebas del ambiente con poca iluminación.

#### **3.5.2.1 Pruebas en Ambientes con Iluminación Limitada**

Se utiliza una sala con iluminación reducida y dentro de este escenario se realizan las siguientes otras consideraciones como: variar ángulos de inclinación, ángulo de enfoque de la cámara, distancia de separación entre la cámara y objeto, que generalmente se encuentran en los sistemas de videovigilancia.

Con lo mencionado se puede iniciar las pruebas en los diferentes escenarios descritos a continuación con iluminación limitada.

- Cámara web frontal a  $0^\circ$  de inclinación vertical u horizontal

Con la cámara apuntando directamente a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos, mostrados en la Figura 3.6 y la Tabla 3.2.



**Figura 3.6:** Detección frontal con poca Iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.6	3	100%	0	0

**Tabla 3.2:** Detección frontal con poca Iluminación

- Cámara web a  $45^\circ$  de inclinación horizontal

Con la cámara apuntando a los objetos con un ángulo de  $45^\circ$  a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.7 y la Tabla 3.3.



**Figura 3.7:** Cámara web con enfoque a  $45^\circ$  con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.7	3	100%	0	0

**Tabla 3.3:** Resultados con enfoque a 45° con poca iluminación

- Detección de rostros girados a 45°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 45° a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.8 y la Tabla 3.4.



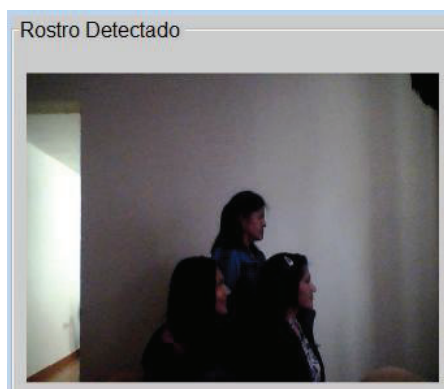
**Figura 3.8:** Detección de rostros girados a 45° con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.8	3	100%	0	0

**Tabla 3.4:** Detección de rostros girados a 45° con poca iluminación

- Detección de rostros girados a 90°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 90° a una distancia de 2 m no se detecta ninguno de los tres rostros dentro de la sala, sin falsos positivos y tres falsos negativos mostrados en la Figura 3.9 y la Tabla 3.5.



**Figura 3.9:** Detección de rostros girados a 90° con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.9	3	0%	0	3

**Tabla 3.5:** Detección de rostros girados a 90° con poca iluminación

A continuación, se presentan las pruebas del ambiente con iluminación normal controlada.

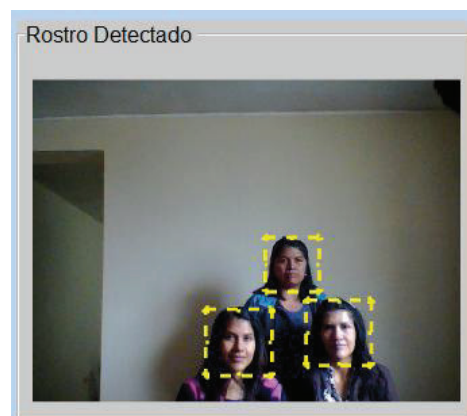
### 3.5.2.2 Pruebas en Ambientes con Iluminación Normal

Se utiliza una sala con iluminación normal controlada y dentro de este escenario se realizan las siguientes otras consideraciones como: variar ángulos de inclinación, ángulo de enfoque de la cámara, distancia de separación entre la cámara y objeto, que generalmente se encuentran en los sistemas de videovigilancia.

Con lo mencionado se puede iniciar las pruebas en los diferentes escenarios descritos a continuación con iluminación normal.

- Cámara web frontal a 0° de inclinación vertical u horizontal

Con la cámara apuntando directamente a una distancia de 2 m detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.10 y la Tabla 3.6.



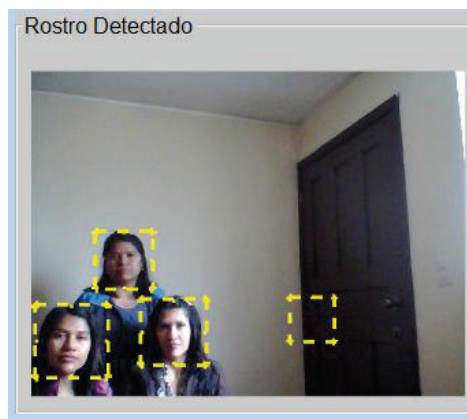
**Figura 3.10:** Detección frontal con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.10	3	100%	0	0

**Tabla 3.6:** Detección frontal con iluminación normal

- Cámara web a 45° de inclinación horizontal

Con la cámara apuntando a los objetos con un ángulo de 45° a una distancia de 2 m se detectan los tres rostros dentro de la sala, con un falso positivo en la puerta de acceso a la sala y cero falsos negativos mostrados en la Figura 3.11 y la Tabla 3.7.



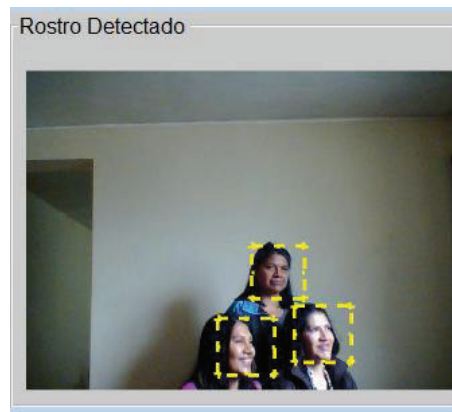
**Figura 3.11:** Cámara web con enfoque a 45° con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.11	3	100%	1	0

**Tabla 3.7:** Resultados con enfoque a 45° con iluminación normal

- Detección de rostros girados a  $45^\circ$

Con la cámara apuntando directamente a los objetos y estos giran su rostro a  $45^\circ$  a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.12 y la Tabla 3.8.



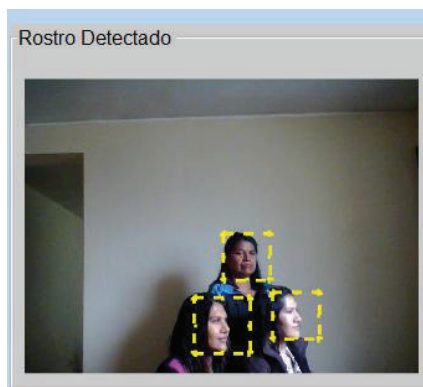
**Figura 3.12:** Detección de rostros girados a  $45^\circ$  con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.12	3	100%	0	0

**Tabla 3.8:** Detección de rostros girados a  $45^\circ$  con iluminación normal

- Detección de rostros girados a  $90^\circ$

Con la cámara apuntando directamente a los objetos y estos giran su rostro a  $90^\circ$  a una distancia de 2 m detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.13 y la Tabla 3.9.



**Figura 3.13:** Detección de rostros girados a  $90^\circ$  con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.13	3	100%	0	0

**Tabla 3.9:** Detección de rostros girados a 90° con iluminación normal

Además de realizar las pruebas con personas dentro de la sala, es importante realizar pruebas con objetos que se entienden como rostros que deben ser detectados. Estas pruebas se realizarán a continuación con rostros impresos en hojas que presentarán características completas, incompletas y confusas de un rostro.

### 3.5.2.3 Pruebas con otros objetos para detección de rostros

Otro tipo de pruebas se pueden realizar cuando las facciones de un rostro se muestran completas, incompletas, en escala grises y confusas. Los resultados obtenidos se muestran a continuación considerando que:

- 1) Cuando un rostro es encerrado en el recuadro ajustado a su tamaño de rostro es una detección satisfactoria.
- 2) Cuando un rostro es encerrado en el recuadro pero no se ajusta adecuadamente el tamaño de rostro es una detección buena.
- 3) Cuando un rostro no es encerrado en un recuadro o su recuadro identifica menos del 50% del rostro es una detección fallida.

- Detección de rostros con facciones completas

La detección es satisfactoria ya que encuentra todas las facciones del rostro como se observa en la Figura 3.14. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.10.





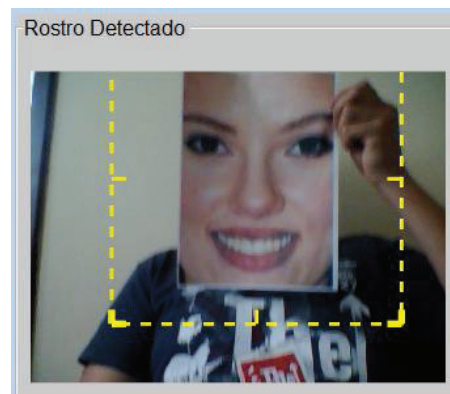
**Figura 3. 14:** Detección de rostro con facciones completas

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.14	1	100%	0	0

**Tabla 3.10:** Detección de rostro con facciones completas

- Detección de rostros con facciones incompletas

La detección es buena sin tener todas las facciones del rostro completas como se observa en la Figura 3.15. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.11.



**Figura 3.15:** Detección de facciones de un rostro sin facciones completas de rostro

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.15	1	100%	0	0

**Tabla 3.11:** Detección de facciones de un rostro sin facciones completas de rostro

- Detección de rostros blanco y negro

La detección es buena con un rostro en escala grises completa como se observa en la Figura 3.16. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.12.



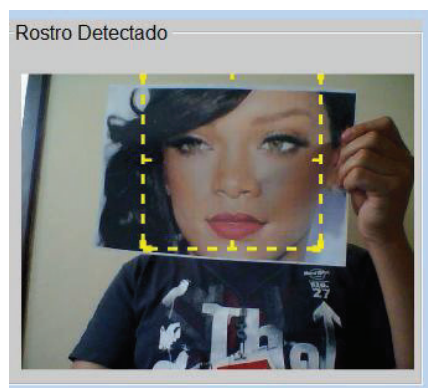
**Figura 3.16:** Detección de rostro blanco y negro

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.16	1	100%	0	0

**Tabla 3.12:** Detección de rostro blanco y negro

- Detección de rostros casi completo

La detección es buena sin tener todas las facciones del rostro completas como se observa en la Figura 3.17. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.13.



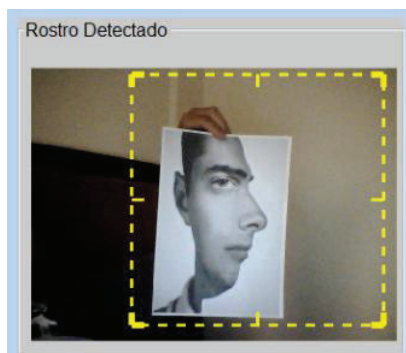
**Figura 3.17:** Detección de rostro casi completo

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.17	1	100%	0	0

**Tabla 3.13:** Detección de rostro casi completo

- Detección de rostros confuso

La detección es buena dado que es un rostro confuso como se observa en la Figura 3.18. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.14.



**Figura 3.18:** Detección de rostro confuso

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.18	1	100%	0	0

**Tabla 3.14:** Detección de rostro confuso

Los resultados de las pruebas de ejecución del código en diferentes escenarios son aceptables y se revisará su consolidado más adelante, lo usual sería lograr tasas de detección entre el 85 y 95%, con tasas de falso positivos muy bajas en el orden de  $10^{-5}$  o  $10^{-6}$  [13].

En la siguiente etapa se analizará un código disponible comercial de Mathwoks para detección y seguimiento de rostros.

### 3.6 CÓDIGO COMERCIAL DE MATHWORKS

MATLAB es uno de los lenguajes de programación más importantes que se utilizan en universidades, empresas y hogares de todo el mundo para la enseñanza, la investigación y el aprendizaje basado en proyectos [89].

La licencia home ayuda si como usuario final desea operar, instalar y administrar sus propios proyectos, registrada solo de manera personal disponibles para Windows, Linux, and Macintosh la cual tiene un costo asociado. El usuario puede registrarse y obtener una cuenta de Mathworks, con la que puede contar con

ayuda y compartir información con los demás miembros de la comunidad de Mathworks [96].

Dentro de esto se cuenta con ayuda de funciones y códigos que se pueden descargar gratuitamente, como es el caso del código para detección y seguimiento de rostros en la página de *Mathworks* de Matlab realizado por Dima Lisin que está disponible con el siguiente nombre *Detect and track multiple faces from live video*.

El código comercial detecta y hace seguimiento de múltiples rostros en el video de una webcam, este código utiliza la función *facetedetector* de clasificadores en cascada del algoritmo de viola&jones y la función *tracker* utiliza el algoritmo de seguidor de características Kanade-Lucas-Tomasi KLT, con el cual se realizarán comparaciones en las mismas condiciones que fue evaluado el código desarrollado.

El código fuente inicia llamando a dos funciones para detección y seguimiento: La función *facetedetector* y la función *tracker*.

```
% detectAndTrackFaces
% Automatically detects and tracks multiple faces in a webcam-acquired video
stream.
% Copyright 2013-2014 The Mathworks, Inc
clear classes;
% Instantiate video device, face detector, and KLT object tracker
vidObj = webcam;
faceDetector = vision.CascadeObjectDetector(); % Finds faces by default
tracker = MultiObjectTrackerKLT;
```

Se obtiene un tamaño de trama que se maneja y se crea una instancia de reproductor de video.

```
%% Get a frame for frame-size information
frame = snapshot(vidObj);
frameSize = size(frame);

%% Create a video player instance
videoPlayer = vision.VideoPlayer('Position',[200 100 flipr(frameSize(1:2)+30)]);
Con el video de entrada se crea un proceso iterativo hasta que satisfactoriamente se detecte los rostros, la aplicación seguirá activa hasta que el usuario presione ctrl+C para detenerlo.
```

```

%% Iterate until we have successfully detected a face
bboxes = [];
while isempty(bboxes)
    framergb = snapshot(vidObj);
    frame = rgb2gray(framergb);
    bboxes = faceDetector.step(frame);
end
tracker.addDetections(frame, bboxes);

%% And loop until the player is closed
frameNumber = 0;
keepRunning = true;
disp('Press Ctrl-C to exit...');
while keepRunning

    framergb = snapshot(vidObj);
    frame = rgb2gray(framergb);

    if mod(frameNumber, 10) == 0
        % (Re)detect faces.
        %
        % NOTE: face detection is more expensive than imresize; we can
        % speed up the implementation by reacquiring faces using a
        % downsampled frame:
        % bboxes = faceDetector.step(frame);
        bboxes = 2 * faceDetector.step(imresize(frame, 0.5));
        if ~isempty(bboxes)
            tracker.addDetections(frame, bboxes);
        end
    else
        % Track faces
        tracker.track(frame);
    end

    frameNumber = frameNumber + 1;
end

```

Una vez detectados los rostros del video se colocan rectángulos delimitadores de los puntos rastreados.

```

% Display bounding boxes and tracked points.
displayFrame = insertObjectAnnotation(framergb, 'rectangle',...
    tracker.Bboxes, tracker.BoxIds);
displayFrame = insertMarker(displayFrame, tracker.Points);
videoPlayer.step(displayFrame);
frameNumber = frameNumber + 1;
end
%% Clean up
release(videoPlayer);

```

**Código 2:** Código para detección de rostros de la herramienta comercial [5]

Para visualizar los resultados, el código muestra una ventana de la aplicación en la que se observarán los resultados de la ejecución, sometidos a los mismos ambientes y escenarios que el código desarrollado anteriormente.

### 3.7 EJECUCIÓN DEL CÓDIGO DE MATHWORKS

#### 3.7.1 EJECUCIÓN DE LA DETECCIÓN DE ROSTROS

En la ejecución del código se evaluarán varios escenarios de pruebas para determinar su comportamiento en ambientes con muy poca iluminación e iluminación normal.

A continuación, se presentan las pruebas del ambiente con poca iluminación.

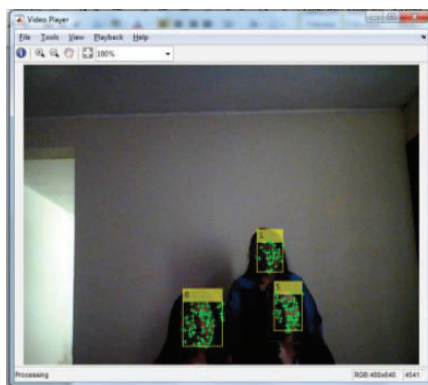
##### 3.7.1.1 Pruebas en Ambiente con Iluminación Limitada

Se utiliza una sala con iluminación reducida y dentro de este escenario se realizan las siguientes otras consideraciones como: variar ángulos de inclinación, ángulo de enfoque de la cámara, distancia de separación entre la cámara y objeto, que generalmente se encuentran en los sistemas de videovigilancia.

Con lo mencionado se puede iniciar las pruebas en los diferentes escenarios descritos a continuación con iluminación limitada.

- Cámara frontal a 0° de inclinación vertical u horizontal

Con la cámara apuntando directamente a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.19 y la Tabla 3.15.



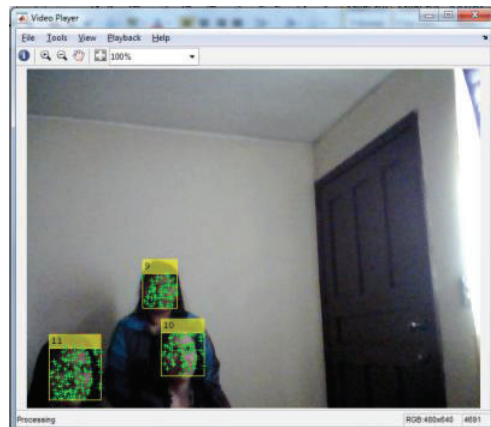
**Figura 3. 19:** Detección frontal con poca Iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.19	3	100%	0	0

**Tabla 3.15:** Detección frontal con poca Iluminación

- Cámara a 45° de inclinación horizontal

Con la cámara apuntando a los objetos con un ángulo de 45° a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.20 y la Tabla 3.16.



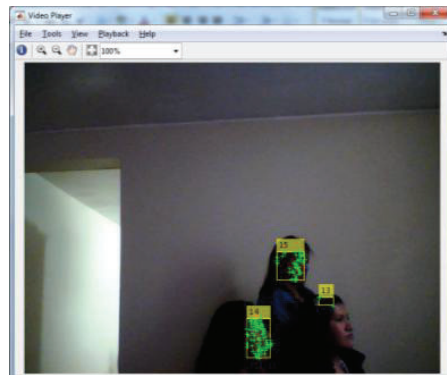
**Figura 3.20:** Cámara con enfoque a 45° con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.20	3	100%	0	0

**Tabla 3.16:** Resultados con enfoque a 45° con poca iluminación

- Detección de rostros girados a 45°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 45° a una distancia de 2 m se detectan dos rostros de los tres rostros dentro de la sala, con un falso positivo y un falso negativo mostrados en la Figura 3.21 y la Tabla 3.17.



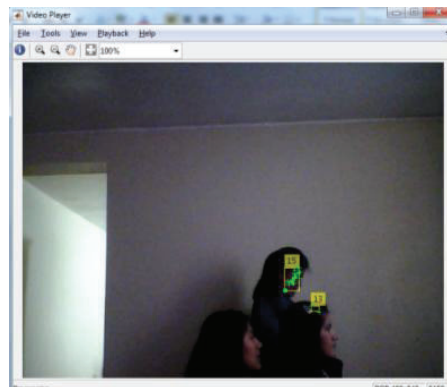
**Figura 3.21:** Detección de rostros girados a 45° con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.21	3	66.666%	1	1

**Tabla 3.17:** Detección de rostros girados a 45° con poca iluminación

- Detección de rostros girados a 90°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 90° a una distancia de 2 m no se detectan rostros de los tres rostros dentro de la sala, dos falsos positivos y tres falsos negativos mostrados en la Figura 3.22 y la Tabla 3.18.



**Figura 3.22:** Detección de rostros girados a 90° con poca iluminación

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.22	3	0%	2	3

**Tabla 3.18:** Detección de rostros girados a 90° con poca iluminación



A continuación, se presentan las pruebas del ambiente con iluminación normal controlada.

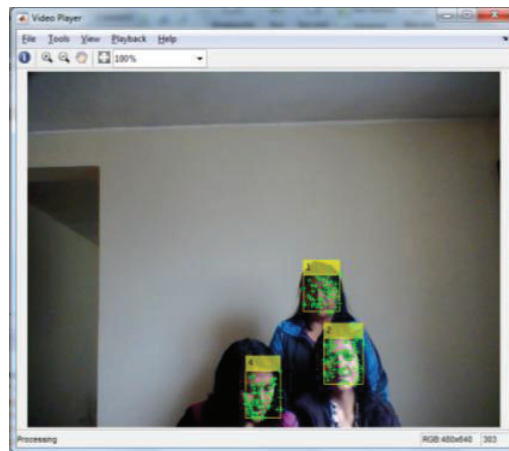
### 3.7.1.2 Pruebas en Ambientes con Iluminación Adecuada

Se utiliza una sala con iluminación normal controlada y dentro de este escenario se realizan las siguientes otras consideraciones como: variar ángulos de inclinación, ángulo de enfoque de la cámara, distancia de separación entre la cámara y objeto, que generalmente se encuentran en los sistemas de videovigilancia.

Con lo mencionado se puede iniciar las pruebas en los diferentes escenarios descritos a continuación con iluminación normal.

- Cámara frontal a 0° de inclinación vertical u horizontal

Con la cámara apuntando directamente a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.23 y la Tabla 3.19.



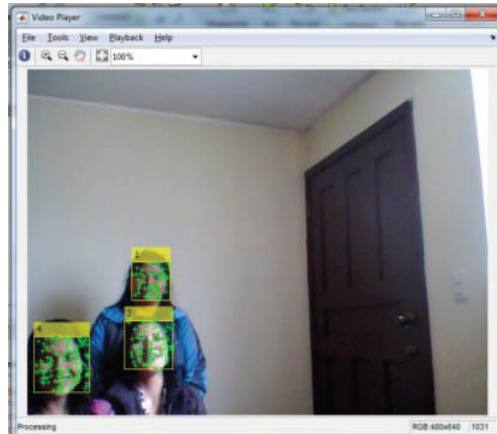
**Figura 3.23:** Detección frontal con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.23	3	100%	0	0

**Tabla 3.19:** Detección frontal con iluminación normal

- Cámara a 45° de inclinación horizontal

Con la cámara apuntando a los objetos con un ángulo de 45° a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.24 y la Tabla 3.20.



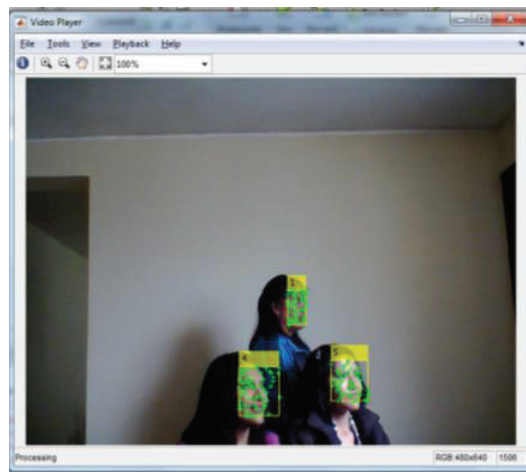
**Figura 3. 24:** Cámara con enfoque a 45° con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.24	3	100%	0	0

**Tabla 3.20:** Resultados con enfoque a 45° con iluminación normal

- Detección de rostros girados a 45°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 45° a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.25 y la Tabla 3.21.



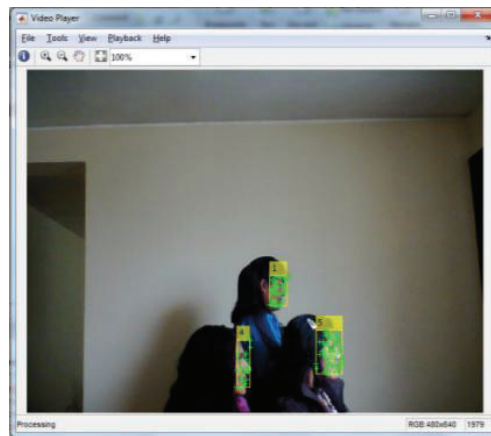
**Figura 3.25:** Detección de rostros girados a 45° con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.25	3	100%	0	0

**Tabla 3.21:** Detección de rostros girados a 45° con iluminación normal

- Detección de rostros girados a 90°

Con la cámara apuntando directamente a los objetos y estos giran su rostro a 90° a una distancia de 2 m se detectan los tres rostros dentro de la sala, sin falsos positivos ni falsos negativos mostrados en la Figura 3.26 y la Tabla 3.22.



**Figura 3.26:** Detección de rostros girados a 90° con iluminación normal

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.26	3	100%	0	0

**Tabla 3.22:** Detección de rostros girados a 90° con iluminación normal

Además de realizar las pruebas con personas dentro de la sala, es importante realizar pruebas con objetos que se entienden como rostros que deben ser detectados. Estas pruebas se realizarán a continuación con rostros impresos en hojas que presentarán características completas, incompletas y confusas de un rostro.

### 3.7.1.3 Pruebas con otros objetos para detección de rostros

Otro tipo de pruebas se pueden realizar cuando las facciones de un rostro se muestran completas, incompletas, en escala grises y confusas. Los resultados obtenidos se muestran a continuación considerando que:

- 1) Cuando un rostro es encerrado en el recuadro ajustado a su tamaño de rostro es una detección satisfactoria.
- 2) Cuando un rostro es encerrado en el recuadro pero no se ajusta adecuadamente el tamaño de rostro es una detección buena.
- 3) Cuando un rostro no es encerrado en un recuadro o su recuadro identifica menos del 50% del rostro es una detección fallida.

- Detección de rostros con facciones completas

La detección es satisfactoria ya que encuentra todas las facciones del rostro como se observa en la Figura 3.27. Un rostro detectado, sin falsos positivos ni falsos negativos de acuerdo a la Tabla 3.23.



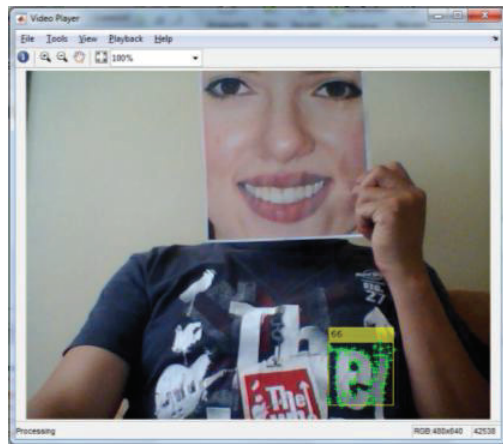
**Figura 3.27:** Detección de rostro con facciones completas

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.27	1	100%	0	0

**Tabla 3.23:** Detección de rostro con facciones completas

- Detección de rostros con facciones incompletas

La detección es fallida sin tener todas las facciones del rostro completas como se observa en la Figura 3.28. Cero rostros detectados, con un falso positivo y un falso negativo de acuerdo a la Tabla 3.24.



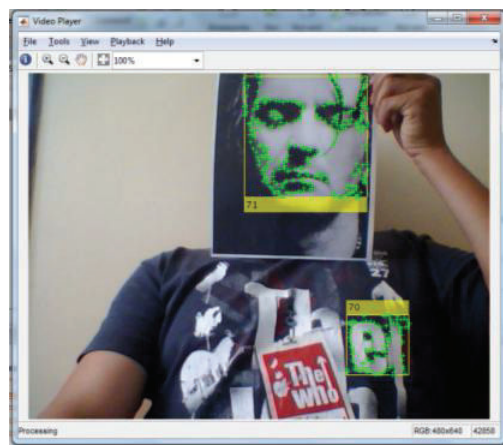
**Figura 3.28:** Detección de facciones de un rostro sin facciones completas de rostro

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.28	1	0%	1	1

**Tabla 3.24:** Detección de facciones de un rostro sin facciones completas de rostro

- Detección de rostros blanco y negro

La detección es buena con un rostro en escala grises completa como se observa en la Figura 3.29. Un rostro detectado, un falso positivo y cero falsos negativos de acuerdo a la Tabla 3.25.



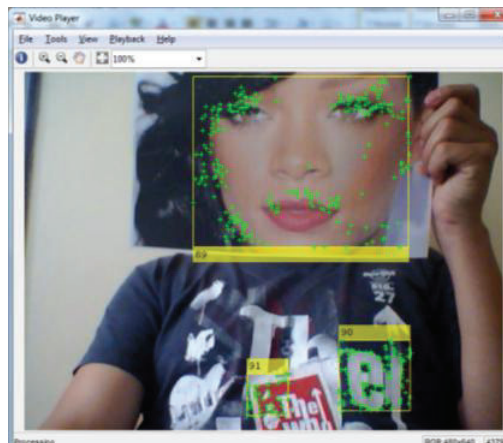
**Figura 3.29:** Detección de rostro blanco y negro

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.29	1	100%	1	0

**Tabla 3.25:** Detección de rostro blanco y negro

- Detección de rostros casi completo

La detección es buena sin tener todas las facciones del rostro completas como se observa en la Figura 3.30. Un rostro detectado, dos falsos positivos y cero falsos negativos de acuerdo a la Tabla 3.26.



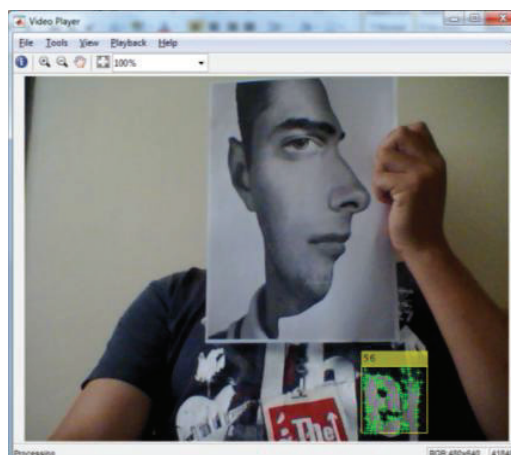
**Figura 3.30:** Detección de rostro casi completo

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.30	1	100%	2	0

**Tabla 3.26:** Detección de rostro casi completo

- Detección de rostros confuso

La detección es fallida dado que es un rostro muy incompleto y confuso como se observa en la Figura 3.31. Cero rostros detectados, con un falsos positivos y un falso negativo de acuerdo a la Tabla 3.27.



**Figura 3.31:** Detección de rostro confuso

Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.31	1	0%	1	1

**Tabla 3.27:** Detección de rostro confuso

### 3.8 ANÁLISIS DE LAS PRUEBAS REALIZADAS

Como se describió en el capítulo 2, para determinar la efectividad de un algoritmo para detección de rostros existen tres parámetros fundamentales de analizar:

- 1) La tasa de detección que son todos los rostros que han sido detectados y corresponden a un rostro.
- 2) Los falsos positivos que son los rostros que han sido detectados y que no corresponden a un rostro.
- 3) Falsos Negativos que son los rostros que no han sido detectados cuando correspondían a un rostro.

Con estos parámetros se podrá analizar y obtener un porcentaje ponderable del desempeño y efectividad de los algoritmos.

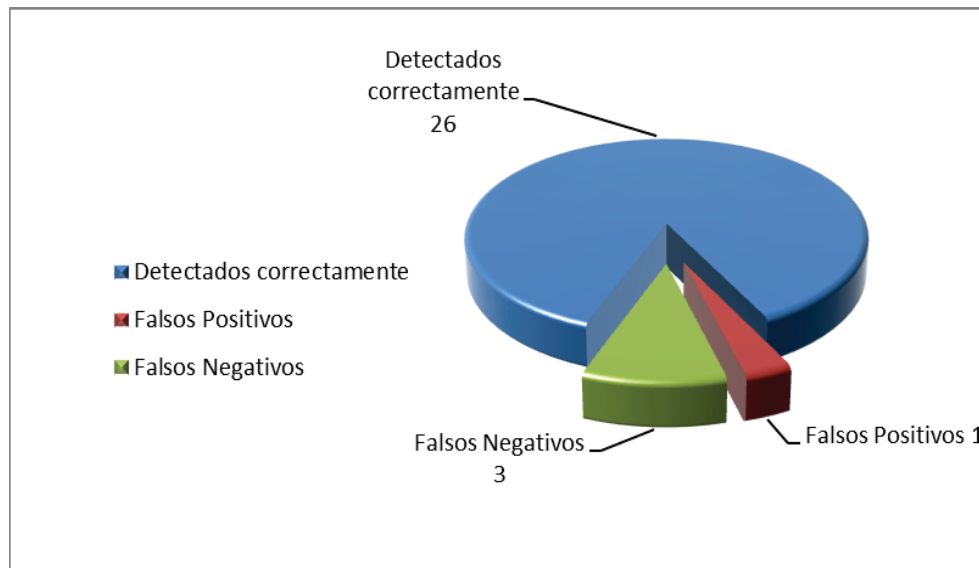
A continuación, se muestra en la Tabla 3.27 los resultados consolidados de las pruebas realizadas de cada código:

- El código desarrollado fue evaluado en 13 escenarios con 29 rostros, 26 rostros fueron detectados correctamente, 1 falso positivo fue encontrado, 3 falsos negativos como se muestra en la Figura 3.33. Se obtiene un porcentaje favorable de detección del 92% como se muestra en la Tabla 3.27 y se presentan problemas de detección en ambientes con poca iluminación.

Resultados del código desarrollado				
Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.6	3	100%	0	0
Figura 3.7	3	100%	0	0
Figura 3.8	3	100%	0	0

Figura 3.9	3	0%	0	3
Figura 3.10	3	100%	0	0
Figura 3.11	3	100%	1	0
Figura 3.12	3	100%	0	0
Figura 3.13	3	100%	0	0
Figura 3.14	1	100%	0	0
Figura 3.15	1	100%	0	0
Figura 3.16	1	100%	0	0
Figura 3.17	1	100%	0	0
Figura 3.18	1	100%	0	0
<b>TOTAL</b>	<b>29</b>	<b>92%</b>	<b>1</b>	<b>3</b>

**Tabla 3.28:** Tabla de resultados 1 código desarrollado



**Figura 3.32:** Análisis de resultados del código desarrollado

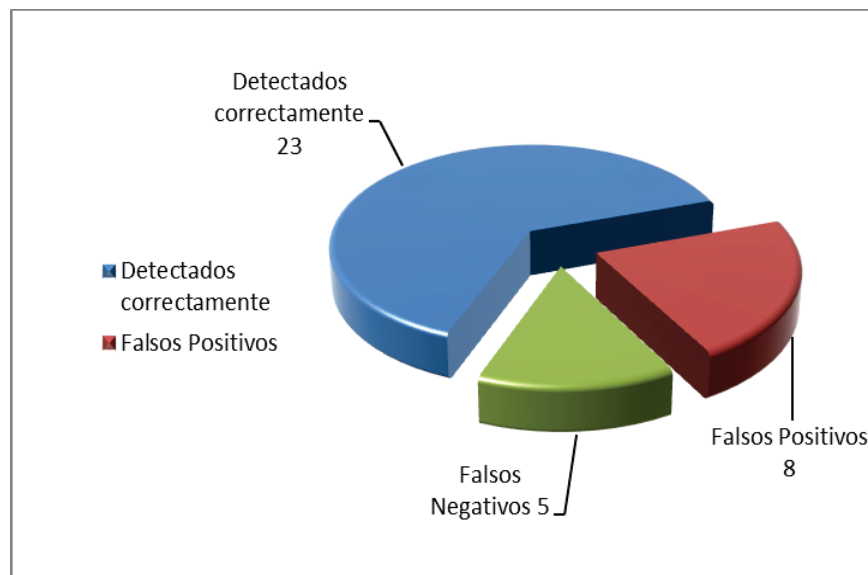
El código de Mathworks fue evaluado en 13 escenarios con 29 rostros, 23 rostros fueron detectados correctamente, 8 falsos positivos fueron encontrados, 5 falsos negativos como se muestra en la Figura 3.34. Se obtuvo un porcentaje de detección del 74% como se muestra en la Tabla 3.28, presentando mayor dificultad en la detección de rostros girados a más de 45°, con poca iluminación.

Además el código de Mathworks muestra que tiene una tasa de falsos positivos detectados en rostros u objetos del fondo de la imagen.



Resultados del código de Mathworks				
Imagen	# de Rostros	Tasa de detección	Falsos Positivos	Falsos Negativos
Figura 3.19	3	100%	0	0
Figura 3.20	3	100%	0	0
Figura 3.21	3	67%	1	1
Figura 3.22	3	0%	2	2
Figura 3.23	3	100%	0	0
Figura 3.24	3	100%	0	0
Figura 3.25	3	100%	0	0
Figura 3.26	3	100%	0	0
Figura 3.27	1	100%	0	0
Figura 3.28	1	0%	1	1
Figura 3.29	1	100%	1	0
Figura 3.30	1	100%	2	0
Figura 3.31	1	0%	1	1
<b>TOTAL</b>	<b>29</b>	<b>74%</b>	<b>8</b>	<b>5</b>

**Tabla 3.29:** Tabla de resultados 2 código de Mathworks



**Figura 3.33:** Análisis de resultados del código de Mathworks

En conclusión la poca iluminación puede ser un factor de riesgo a considerar para la detección de rostros, pero en sistemas de videovigilancia se corrige con el uso de dispositivos infrarrojos, con esto la imagen presentará mejores posibilidades para una favorable detección.

Los ángulos de inclinación y rotación de rostros en la implementación de los sistemas de videovigilancia son corregidos por la ubicación precisa de la cámara donde se considera el enfoque, distancia y alineación correcta.

Todos estos resultados del análisis realizado se tomarán en cuenta en las conclusiones y recomendaciones que se detallan en el capítulo siguiente que se presenta a continuación.

## **CAPÍTULO 4**

### **CONCLUSIONES Y RECOMENDACIONES**

#### **4.1 CONCLUSIONES**

- La analítica de video ha tenido un desarrollo importante en el campo de la electrónica brindando sus prestaciones en varios campos de aplicación. En la última década ha aportado significativamente en el campo de los sistemas de videovigilancia al introducir los algoritmos para detección, seguimiento y reconocimiento, con lo que permite al usuario obtener información útil de un determinado lugar.
- La detección de rostros es el primer paso necesario en los sistemas de seguimiento y reconocimiento con el propósito de localizar y extraer la región del rostro desde un fondo de la imagen. Esto tiene varias aplicaciones en áreas tales como recuperación de imágenes basada en contenido, codificación de video, videoconferencia, vigilancia de multitudes e interfaces inteligentes entre humanos y computadoras.
- Existe una gran cantidad de algoritmos de detección de rostros que han ido evolucionando para localizar un rostro humano en una imagen o escena. Permitiendo así realizar combinaciones de algoritmos entre algunos de ellos: detección por color, detección por el movimiento, detección basada en modelos geométricos, etc, y cada una cuenta con ciertas ventajas y desventajas frente a las demás donde continuamente se van desarrollando mejores técnicas que permiten minimizar los problemas ocasionados por factores externos.
- Los problemas más comunes en la detección de rostros que se dan en el proceso de adquisición de las imágenes en un sistema de videovigilancia son las condiciones de luminosidad, el bajo contraste, objetos extraños en

el fondo, rostros parcialmente capturados, posición de la persona, expresión facial y distorsión de la imagen.

- El algoritmo para detección de rostros de Viola&Jones es uno de los más robustos utilizados en la actualidad y aplica el uso de una cascada de clasificadores muy sencillos que se ejecutan uno detrás de otro.
- Las características básicas de Haar son utilizadas por el algoritmo y el determinar qué características son las adecuadas para utilizar en la detección de rostros es uno de los factores importantes del entrenamiento de una imagen. Estas características son elegidas de un conjunto de características sencillas y completas a la vez, debido a que se pueden establecer muchas características pero usar las más básicas ayudan a determinar y descartar más fácilmente regiones de una imagen.
- Cada clasificador de la cascada se entrena con el algoritmo AdaBoost. Los primeros clasificadores son muy sencillos y permiten rechazar una gran cantidad de no rostros mientras que aceptan un porcentaje muy alto de rostros. Los clasificadores posteriores se encargan de afinar la detección. La cascada permite desechar gran parte de las regiones de la imagen y sólo concentrarse en las zonas en las que es más probable que haya un rostro por lo que la detección es bastante rápida.
- Una dificultad de AdaBoost en la detección de rostro por el método de Viola&Jones se considera que no minimiza directamente los falsos negativos, si no que minimiza el error total (falsos negativos + falsos positivos). Esto hace que una tasa de falsos negativos sea mucho menor que la de falsos positivos al entrenar un clasificador con AdaBoost. Para alcanzar altas tasas de detección a costa de tolerantes tasas de falsos positivos se debe ajustar el umbral del perceptrón generado por AdaBoost. Umbrales más bajos darán como resultado tasas de detección más altas y

tasas de falsos positivos más altas, umbrales más altos darán tasas de falsos positivos más bajas pero peores tasas de detección.

- La distancia de Hausdorff también es uno de los métodos robustos para detección de rostros donde la técnica consiste en un sistema de detección basado en ajuste de modelos o bordes geométricos de imágenes en escala de grises y la distancia modificada de Hausdorff. Su forma de aplicación consiste primeramente en definir un patrón correspondiente a un rostro estándar normalmente frontal y se busca la zona de la imagen más similar a este patrón mediante la correlación. Su estructura es bastante sólida y se basa en realizar una detección gruesa de la región facial con etapas de segmentación y posicionamiento para luego pasar a un re-afinamiento en una segunda fase donde se aplicarán las mismas etapas antes indicadas.
- Al analizar y comparar los algoritmos para detección de rostros por el método de cascada de Viola&Jones y por modelos geométricos usando la distancia de Hausdorff se puede concluir que los dos cuentan con altas prestaciones ventajosas en cuanto a la detección de un rostro y ciertas desventajas frente a factores externos como iluminación, sin embargo, la detección de rostros propuesto por Paul Viola y Michael Jones es uno de los métodos más usados en la actualidad ya que ha permitido segmentar múltiples rostros en una imagen con tiempos de procesamiento más bajos lo cual se traduce en una mejor velocidad de detección.
- En la fase de implementación los parámetros para determinar si un algoritmo es adecuado: son la tasa de detección, tasa de falsos positivos y tasa de falsos negativos. Llevar el código de Viola&Jones a la implementación confirma que su tasa de detección y la velocidad de procesamiento de una imagen son efectivas para sistemas de videovigilancia pudiendo utilizarlo en un seguimiento de rostros en tiempo real. Las pruebas del sistema para determinar la tasa de detección de rostros del algoritmo implementado se realizaron en múltiples escenarios

de un ambiente controlado con uno o varios rostros para la detección frontal, a 45°, a 90°. Los buenos resultados de identificación muestran que el sistema es robusto frente a diferentes condiciones de fondo y cambiando la iluminación.

- De las pruebas realizadas y comparadas con el código comercial se puede concluir que ambos presentan dificultad en la detección de rostros en ambientes con poca iluminación y rostros que han sido rotados abruptamente. Sin embargo, estos aspectos son corregidos en videovigilancia con infrarrojos para la falta de iluminación de la cámara y una buena ubicación de la cámara para detectar rostros frontales, con giro o rotación del rostro máximo de 80°.
- La tasa de falsos positivos y falsos negativos de los códigos de seguimiento implementado y comercial presentan valores diferentes debido a que el algoritmo de seguimiento implementado realiza un posicionamiento de puntos tradicional y el algoritmo comercial utiliza el algoritmo de KLT. Esta técnica está basada en el seguimiento de los puntos característicos mediante las ecuaciones desarrolladas por Lucas-Kanade para el cálculo del flujo óptico, que a su vez implementan el método iterativo de Newton-Raphson para una búsqueda de ascenso por gradiente lo cual genera más falsos positivos. Por tanto este algoritmo es muy recomendado para la detección de objetos en general.

## 4.2 RECOMENDACIONES

- La seguridad es un campo en el que no se puede estar a la espera de que las cosas sucedan y lamentar lo sucedido. Este campo ha desarrollado varios sistemas de videovigilancia en los que se puede apoyar para determinar una solución preventiva y hasta correctiva ante eventos y situaciones que pueden ser parte de un control o de un proceso institucional, empresarial y hasta comercial.

- Es necesario implementar sistemas de videovigilancia con aplicaciones automáticas de analítica de video, pues esta presta muchas ventajas al administrador de la red reduciendo su espacio físico, recursos de almacenamiento, recursos de monitoreo con lo que mejora altamente su rendimiento y disminuye su costo operacional. Cada instante se mejoran estos sistemas que están en el mercado actual con varios modelos y costos accesibles,
- Para prevenir los actos vandálicos en establecimientos como centros comerciales, bancos, instituciones educativas, etc., es recomendable soluciones más autónomas que las tradicionales que adicional a la generación de un contenido de video, preste de manera automática la detección de múltiples objetos en tiempo real, de acontecimientos en una escena y mediante el análisis de la imagen con algoritmos especializados para cada actividad permitan generar alarmas para evitar los actos de vandalismo.
- En este trabajo desarrollado se ha visto que los algoritmos que se han analizado son los más robustos y que presentan tasas de detección altas, sin embargo, están sujetos a dificultades en las condiciones del ambiente como luminosidad y una buena ubicación del sistema, por lo que es recomendable contar con un especialista que evalúe estas condiciones primero para que todos los beneficios de un sistema sean aprovechados al máximo.
- Al momento de realizar detección de rostros utilizando el Algoritmo de clasificadores en cascada de Vioja&Jones es necesario fijar las metas primeramente de detección y desempeño del sistema, lo cual ayudará a partir desde ciertos porcentajes esperados de detección y lo recomendable es entre el 85 y 95%, con tasa de falso positivos muy bajas, del orden de  $10^{-5}$  o  $10^{-6}$ . Con esto también se estará considerando el menor esfuerzo de los recursos de procesamiento y por tanto menor tiempo de detección.

- Al utilizar el algoritmo de modelos geométricos utilizando la distancia de Hausdorff es importante considerar que este algoritmo presenta tasas altas de detección en rostros frontales planos por lo que es recomendable utilizarlo en dispositivos con aplicaciones de este tipo como lecturas faciales biométricas, cámaras fotográficas, cámaras de laptops, etc., es decir donde sean aprovechadas sus características de detección.
- Las dificultades encontradas se proyectan a la investigación futura para superar las restricciones de los algoritmos analizados, mejorándolos en tasas de transmisión, ambientes de operación, objetos de detección e incluso evolucionando a la detección, seguimiento y reconocimiento de patrones en tres dimensiones que está en proceso de pruebas y desarrollo.
- En la etapa de pruebas y análisis de resultados de los algoritmos simulados es recomendable ejecutar las pruebas en los ambientes que mayor dificultad presenten con lo que se obtiene los resultados en los que se analizarán las tasas de detección, falsos positivos y falsos negativos con los que se valorarán.



## REFERENCIAS BIBLIOGRÁFICAS

- [1] T. Sakai, M. Nagao, and T. Kanade, Computer analysis and classification of photographs of human faces, in Proc. First USA—Japan Computer Conference, 1972, p. 2.7.
- [2] J. Candamo, M. Shreve, Member, IEEE, Dmitry B. Goldgof, Fellow, IEEE, Deborah B. Sapper, and Rangachar Kasturi, Fellow, IEEE Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms, IEEE transactions on intelligent transportation systems, vol. 11, no. 1, march 2010
- [3] G.J.D. Smith. Behind the screens: Examining constructions of deviance and informal practices among CCTV control room operators in the UK. Surveillance & Society, 2((2/3)): 377-396, 2004.
- [4] W. Hu, T. Tan, Fellow, IEEE, Liang Wang, and Steve Maybank, A Survey on Visual Surveillance of Object Motion and Behaviors, IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 34, no. 3, august 2004
- [5] E. Hjelmas, B. K. Low, Face Detection: A Survey, E-mail: erikh@hig.no; boon@met.ed.ac.uk Received October 23, 2000; accepted April 17, 2001
- [6] J. Candamo, M. Shreve, Member, IEEE, Dmitry B. Goldgof, Fellow, IEEE, Deborah B. Sapper, and Rangachar Kasturi, Fellow, IEEE, Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms, IEEE transactions on intelligent transportation systems, vol. 11, no. 1, march 2010
- [7] A. González Marcos, F.J. Martínez de Pisón Ascacíba, A. V. Pernía Espinoza, F. Alba Elías, M. Castejón Limas, J. Ordieres Meré, and E. Vergara Gonzáles, Técnicas y Algoritmos Básicos de Visión Artificial, 1ª ed. Universidad de la Rioja: Servicio de Publicaciones, 2006.
- [8] R. C. Gonzalez, R. E. Woods, Digital Image Processing, 3rd edition
- [9] S.A.B.I.A. (2010, Dic.) Sistemas Adaptativos y Bioinspirados en Inteligencia Artificial. [Online]. Disponible en:  
<http://sabia.tic.udc.es/gc/Contenidos%20adicionales/trabajos/3D/VisionArtificial/>

- [10] E. Hjelm, B. K. Low, Face Detection: A Survey, E-mail: erikh@hig.no; boon@met.ed.ac.uk Received October 23, 2000; accepted April 17, 2001
- [11] Rapid Object Detection using a Boosted Cascade of Simple Features
- [12] C. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In International Conference on Computer Vision, 1998.
- [13] Y. Freund, R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. In Computational Learning Theory: Eurocolt '95, pages 23–37. Springer-Verlag, 1995.
- [14] R. E. Schapire, Y. Freund, P. Bartlett, and W. Sun Lee. Boosting the margin: A new explanation for the effectiveness of voting methods. In Proceedings of the Fourteenth International Conference on Machine Learning, 1997.
- [15] Rockafellar, R. Tyrrell; Variational Analysis. Springer-Verlag. p. 117. ISBN 3-540-62772-3. Wets, Roger J-B (2005).
- [16] G. Toussaint (1998) CS 507 Computational Geometry. [Online]. Disponible en: <http://cgm.cs.mcgill.ca/~godfried/teaching/cg-projects/98/normand/main.html>
- [17] P. Suau1, F. Pujol2, R. Rizo1, M. Pujol1, Detección facial basada en una distancia de Hausdorff normalizada.
- [18] W. Rucklidge. Efficient Visual Recognition Using the Hausdorff Distance, volume 1173 of Lecture notes in computer science. Springer, 1996.
- [19] M.P. Dubuisson, A.K. Jain, A modified Hausdorff distance for object matching. In ICPR94, pages A: 566-568, Jerusalem, Israel, 1994.
- [20] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, Robust Face Detection
- [21] W. Rucklidge. Efficient Visual Recognition Using the Hausdorff Distance, volume 1173 of Lecture notes in computer science. Springer, 1996.
- [22] P. Suau1, F. Pujol2, R. Rizo1, M. Pujol1, Detección facial basada en una distancia de Hausdorff normalizada
- [23] D. de Sevilla (2017, Ene) Las profesiones digitales ganarán fuerza en 2017 [Online]. Disponible en: <http://queaprendemoshoy.com/category/tecnologia-y-sm/>
- [24] F. J. García Mata, VIDEOVIGILANCIA: CCTV USANDO VIDEOS IP, Publicaciones Vértice, S.L, 1ª ed., 1ª imp. 21/06/2010
- [25] Itys (2016, Nov) Ventajas y Beneficios de Video IP [Online]. Disponible en: [http://www.itys.pe/video\\_ip.html](http://www.itys.pe/video_ip.html)

- [26] P. Gil, J. Pomares, F. Candelas, "Redes y Transmisión de D, universidad de Alicante.
- [27] I. Ruby, Tabla de Velocidades de Trasmisión de Cables utp, coaxial, fibra óptica, ESCORCIA 07-09-11
- [28] Test de velocidad (2015, Mar) Cables de Conexión [Online]. Disponible en: <http://www.testdevelocidad.es/redes/elegir-mejor-cable-red-aprovechar-al-maximo-nuestra-conexion/>
- [29] Radioenlace (2014, Dic) Tipos de fibra óptica Monomodo y Multimodo [Online]. Disponible en: <http://www.radio-enlace.com/tipos-de-fibra-optica-monomodo-y-multimodo/>
- [30] J. Vaca Reyes (2014, Sep) Tipos de comunicación de redes alámbricas e inalámbricas [Online]. Disponible en: <http://jennifervacaicas.blogspot.com/2014/09/tipos-de-comunicacion-de-redes.html>
- [31] Wikipedia.org (2017, Mar) RJ-45 [Online]. Disponible en: <https://es.wikipedia.org/wiki/RJ-45>
- [32] Wikipedia.org (2017, Mar) Cable\_coaxial [Online]. Disponible en: [https://es.wikipedia.org/wiki/Cable\\_coaxial](https://es.wikipedia.org/wiki/Cable_coaxial)
- [33] EA7AHG. (2017, Mar) Cables\_coaxiales [Online]. Disponible en: <https://sites.google.com/site/ea7ahg/antena/cables-coaxiales>
- [34] Wikipedia.org (2017, Ene) Fibra\_óptica\_multimodo [Online]. Disponible en: [https://es.wikipedia.org/wiki/Fibra\\_%C3%B3ptica\\_multimodo](https://es.wikipedia.org/wiki/Fibra_%C3%B3ptica_multimodo)
- [35] comnet, Mediaconverters [Online]. Disponible en: [http://www.comnet.net/products/MediaConverters/CNFE2MC2C\[M\]\\_esp.pdf](http://www.comnet.net/products/MediaConverters/CNFE2MC2C[M]_esp.pdf)
- [36] E. Vidal (2011, Mar) "¿Cuántos y cuáles tipos de redes inalámbricas podemos encontrar" [Online]. Disponible en: <http://redes603didisa.blogspot.com/2011/03/cuantos-y-cuales-tipos-de-redes.html>
- [37] I. Pellejero, F. Andreu, A, Lesta, "Fundamentos y aplicaciones de seguridad en redes WLAN"
- [38] Wikipedia.org (2017, Mar) Cámara de vídeo [Online]. Disponible en: [https://es.wikipedia.org/wiki/C%C3%A1mara\\_de\\_v%C3%ADdeo](https://es.wikipedia.org/wiki/C%C3%A1mara_de_v%C3%ADdeo)

- [39] informaticamoderna (2008-2016) “Cámara IP”, informaticamoderna [Online]. Disponible en: [http://www.informaticamoderna.com/Camara\\_IP.htm](http://www.informaticamoderna.com/Camara_IP.htm)
- [40] Probo (2010, Feb) “cctv-analogo-vs-ip”, CCTV, Seguridad Electrónica [Online]. Disponible en: <http://probo69.blogspot.com/2010/02/cctv-analogo-vs-ip.html>
- [41] L. Bryant (2010, Nov) , “CÁMARAS IP VS ANÁLOGAS”, Información de la actualidad en seguridad Electrónica (cámaras, cercos eléctricos, alarmas), Disponible en: <http://segurimaxperu.blogspot.com/2010/11/camaras-ip-vs-analogas.html>
- [42] INTPLUS (2017, Ene) , “Tipos de cámaras”, Seguridad Electrónica, Disponible en: <http://www.videovigilancia.com/tiposcamaras.htm>
- [43] CAMSEG (2013, Nov) , “Tipos de cámaras de seguridad más populares”, Cámaras de Seguridad, Disponible en: <http://camara-seguridad.com.mx/blog/consejos/140-tipos-de-camaras-de-seguridad-mas-populares>
- [44] System.Com (2017, Nov) , “NEWSURWAY”, Seguridad Industrial y Sistemas, Disponible en: <http://system-com.com.mx/index.php?cPath=44>
- [45] Coreuno (2014, May) , “Cámara DOMO con detector de movimiento”, Seguridad Sistemas, Disponible en: <http://www.coreuno.com/productos/producto/camara-domo-con-detector-de-movimiento/6937790219593/>
- [46] Todoelectrónica (2017, Ene) , “Cámaras de vigilancia domo, Cámaras de seguridad”, Disponible en: [https://www.todoelectronica.com/camaras-de-vigilancia-domo-camaras-de-seguridad-c-28\\_296.html](https://www.todoelectronica.com/camaras-de-vigilancia-domo-camaras-de-seguridad-c-28_296.html)
- [47] Telefoníatotal (2014, Dic) , “Cámaras inalámbricas”, Disponible en: <http://telefoniatotal.com/producto/kit-4-camaras-inalambricas/>
- [48] D. Reina (2015, May) , “HISTORIA DE LA VIDEOVIGILANCIA”, Disponible en: <http://racalarm.com/blog/cctv/historia-de-la-videovigilancia/>
- [49] N. Sosio (2014, Feb) , “Cámaras de Seguridad Analógicas”, SOS Seguridad, Disponible en: <http://www.seguridadsos.com.ar/camaras-de-seguridad-analogicas/>

- [50] Axis Communications (2009, Feb) , “El acceso más fácil al vídeo en red”, wp\_encoders, Disponible en: [http://www.axis.com/files/whitepaper/wp\\_encoders\\_34384\\_es\\_0902\\_lo.pdf](http://www.axis.com/files/whitepaper/wp_encoders_34384_es_0902_lo.pdf)
- [51] Dointech (2015) , “CCTV Video Vigilancia” , Automatización Seguridad y Control, Disponible en: <http://www.dointech.com.co/videovigilancia-ip.html>
- [52] S, Marty;. “Diseño de un Sistema de Tele-vigilancia sobre IP para el edificio CRAI de la escuela politécnica superior de Gandia”. Universidad Politécnica de Valencia (2013)
- [53] V. Ramirez (2009, Nov) , “Componentes de sistemas de videovigilancia” , Disponible en: [https://esemanal.mx/2009/11/componentes\\_de\\_sistemas\\_de\\_videovigilancia/](https://esemanal.mx/2009/11/componentes_de_sistemas_de_videovigilancia/)
- [54] Tredess (2010, Ene) , “SISTEMAS DE VÍDEO VIGILANCIA Y SEGURIDAD”, TRedess\_vv\_ Disponible en: [http://www.tredess.com/sites/default/files/TRedess\\_vv\\_2014.pdf](http://www.tredess.com/sites/default/files/TRedess_vv_2014.pdf)
- [55] TOTEM BLOG (2016, May) , “Importancia y beneficios de las cámaras CCTV”, Disponible en: <http://totem.com.ec/blog/cctv-importancia-y-beneficios/>
- [56] ESYS (2015, Ago) , “Importancia y beneficios de las cámaras CCTV”, seguridad y sociedad, Disponible en: <http://www.fundacionesys.com/es/system/files/VIDEOVIGILANCIA%202016.pdf>
- [57] 911Alarmas (2014) , “PROTEGEMOS LO MÁS IMPORTANTE: TU FAMILIA”, seguridad, Disponible en: <http://911alarmas.com/index.php?modulo=contenido&id=75>
- [58] Infogremio (2016, Abr) , “videovigilancia”, seguridad electrónica, Disponible en: <http://gremioseseguridad.com/electronica/2016/04/21/consejos-recomendar-sistema-videovigilancia/>
- [59] Scati (2015) , “Soluciones videovigilancia”, Disponible en: [http://www.scati.com/es/soluciones/retail\\_15.html](http://www.scati.com/es/soluciones/retail_15.html)
- [60] Accesor (2014), “Centros Comerciales, videovigilancia e identificación personal autorizado”, Soluciones en Control de Accesos y seguridad, Disponible en: [http://www.accesor.com/esp/art2\\_query.php?sec=21](http://www.accesor.com/esp/art2_query.php?sec=21)
- [61] TOTEM BLOG (2016, May) , “Cámaras de seguridad en la Universidad”, Disponible en: <http://totem.com.ec/blog/camaras-de-seguridad-en-la-universidad/>

- [62] Dallmeier (2015) , “Soluciones CCTV / IP para aeropuertos”, Disponible en: [https://www.dallmeier.com/fileadmin/user\\_upload/upload\\_dallmeier.com/PDFs/Downloads/Broschures/Branchenspezifisch/Spanisch/Soluciones\\_sectoriales-Aeropuertos\\_es.pdf](https://www.dallmeier.com/fileadmin/user_upload/upload_dallmeier.com/PDFs/Downloads/Broschures/Branchenspezifisch/Spanisch/Soluciones_sectoriales-Aeropuertos_es.pdf)
- [63] Alava ing (2016) , “Soluciones globales de seguridad para centros penitenciarios”, Area Seguridad, Disponible en: [soluciones-globales-para-seguridad-en-centros-penitenciarios](#)
- [64] S.A.B.I.A. (2010, Dic.) Sistemas Adaptativos y Bioinspirados en Inteligencia Artificial. [Online]. Disponible en: <http://sabia.tic.udc.es/gc/Contenidos%20adicionales/trabajos/3D/VisionArtificial/>
- [65] R. Castillo (2013, Jul) “El Sistema de Enfoque”. [Online]. Disponible en: <http://todo-fotografia.com/2013/el-sistema-de-enfoque/>
- [66] S.A.B.I.A. (2010, Dic.) Sistemas Adaptativos y Bioinspirados en Inteligencia Artificial. [Online]. Disponible en: <http://sabia.tic.udc.es/gc/Contenidos%20adicionales/trabajos/3D/VisionArtificial/>
- [67] Anixter Inc. (2016), “Video Analytics – Embedded vs Server vs Hybrid “. [Online]. Disponible en: [https://www.anixter.com/es\\_la/about-us/news-and-events/news/video-analytics-embedded-vs-server-vs-hybrid.html](https://www.anixter.com/es_la/about-us/news-and-events/news/video-analytics-embedded-vs-server-vs-hybrid.html)
- [68] Wikipedia.org (2017, Ene), “ADdeovigilancia\_IP” [Online]. Disponible en: [https://es.wikipedia.org/wiki/V%C3%ADdeo\\_de\\_vigilancia\\_IP](https://es.wikipedia.org/wiki/V%C3%ADdeo_de_vigilancia_IP)
- [69] Foscam. (2016), “Grabador mini NVR FN3004H “. [Online]. Disponible en: <http://www.foscam.es/grabador-mini-nvr/>
- [70] Wikipedia.org (2017, Mar), “Servidor\_de\_Video” [Online]. Disponible en: [https://es.wikipedia.org/wiki/Servidor\\_de\\_Video](https://es.wikipedia.org/wiki/Servidor_de_Video)
- [71] Axis (2002), “¿Qué es un Servidor de Vídeo? “. [Online]. Disponible en: <http://platea.pntic.mec.es/~lmarti2/axis/El%20Servidor%20de%20Video.pdf>
- [72] Wikipedia.org (2017, Mar), “Conmutador\_(dispositivo\_de\_red)” [Online]. Disponible en: [https://es.wikipedia.org/wiki/Conmutador\\_\(dispositivo\\_de\\_red\)](https://es.wikipedia.org/wiki/Conmutador_(dispositivo_de_red))
- [73] Wikipedia.org (2017, Mar), “Cámara\_IP” [Online]. Disponible en: [https://es.wikipedia.org/wiki/C%C3%A1mara\\_IP](https://es.wikipedia.org/wiki/C%C3%A1mara_IP)

- [74] Axis (2010), "Video compression ", Formatos de compresión. [Online]. Disponible en: <https://www.axis.com/pe/es/learning/web-articles/technical-guide-to-network-video/compression-formats>
- [75] Wikipedia.org (2015, Mar), "Medio\_de\_transmisión" [Online]. Disponible en: [https://es.wikipedia.org/wiki/Medio\\_de\\_transmisi%C3%B3n](https://es.wikipedia.org/wiki/Medio_de_transmisi%C3%B3n)
- [76] Slideshare.org (2016, Jul), "Fundamentos y conceptos básicos de Sistemas de C.C.T.V. con algoritmos de Video analítico." [Online]. Disponible en: [https://es.slideshare.net/german\\_cruz/7analtica-devideo](https://es.slideshare.net/german_cruz/7analtica-devideo)
- [77] A. Pazos (2015, May), "¿Qué es el video analítico? (I)." [Online]. Disponible en: <http://www.ventasdeseguridad.com/201104255842/articulos/enfoques-miscelaneos/ique-es-el-analisis-de-video-i.html>
- [78] Rivatech (2017, Ene), Analítica de vídeo." [Online]. Disponible en: <http://www.rivatech.de/es/analitica-de-video>
- [79] K. J. Kirchberg, O.Jesorsky, and R. W. Frischholz, Genetic Model Optimization for Hausdorff Distance-Based Face Localization
- [80] Innux (2016), "CCTV IP" [Online]. Disponible en: <https://www.innux.pt/hardware/cctv-ip/>
- [81] Wikipedia.org (2017, Feb), "MATLAB" [Online]. Disponible en: <https://es.wikipedia.org/wiki/MATLAB>
- [82] A. Ortiz, "El proceso de toma de decisiones en los sistemas Integrantes", INSTITUTO TECNOLOGICO DE REYNOSA, publicación 30/nov/2015.
- [83] J. Flores, and V. Arispe, "Iluminación Led, Área Tecnológica", Presentación .ppt, publicación 14/nov/2014.
- [84] D. Barragan, "Manual de Interfaz Gráfica de Usuario en Matlab", publicación 25/Mayo/2017
- [85] D. Lisin (2016, Sep), "Detect and Track Multiple Faces" [Online]. Disponible en: <https://www.Mathworks.com/matlabcentral/fileexchange/47105-detect-and-track-multiple-faces/content/detectAndTrackFaces.m>
- [86] Seguritec (2015), "Cámaras de vigilancia para restaurantes y hoteles" [Online]. Disponible en: <http://www.seguritec.cat/camaras-de-vigilancia-para-restaurantes-y-hoteles/>



- [87] A. G. FERNÁNDEZ (2015), “La videovigilancia en el ámbito laboral y el derecho a la intimidad” [Online]. Disponible en: [http://www.juntadeandalucia.es/institutodeadministracionpublica/aplicaciones/boletin/publico/boletin65/Articulos\\_65/Gude-Fernandez\(RGDC\\_20\).pdf](http://www.juntadeandalucia.es/institutodeadministracionpublica/aplicaciones/boletin/publico/boletin65/Articulos_65/Gude-Fernandez(RGDC_20).pdf)
- [88] Axis (2015), “Datasheet cámaras de red AXIS P33 “, Formatos de compresión. [Online]. Disponible en: [https://www.axis.com/files/datasheet/ds\\_p33v\\_59192\\_es\\_1502\\_lo.pdf](https://www.axis.com/files/datasheet/ds_p33v_59192_es_1502_lo.pdf)
- [89] Mathworks Inc. (2017), “¿Qué es MATLAB? “, [Online]. Disponible en: <https://es.mathworks.com/academia/matlab-campus.html>
- [90] G. Tocomá (2010), Detección de caras y análisis de expresiones faciales “, [Online]. Disponible en: <http://alojamientos.us.es/gtocomá/pid/pid10/deteccioncaras.htm>
- [91] CES 2008: “Face Detection New Trend for Camcorders”, [Online]. Disponible en: <http://www.camcorderinfo.com/content/CES-2008-Face-Detection-a-New-Trend-for-Camcorders-34250.htm> (en inglés), Enero 2008
- [92] besonders-sein (2017): “Face Detection & Recognition Homepage “, [Online]. Disponible en: <https://facedetection.com/algorithms/>
- [93] N, Paz, “AdaBoost con aplicación a detección de caras mediante el algoritmo de Viola&Jones”, Documento 1, Febrero de 2009
- [94] Wikipedia.org (2015, Feb), “Viola–Jones object detection framework” [Online]. Disponible en: [https://en.wikipedia.org/wiki/Viola%E2%80%93Jones\\_object\\_detection\\_framework](https://en.wikipedia.org/wiki/Viola%E2%80%93Jones_object_detection_framework)
- [95] J. P. Lerma, “Implementación del algoritmo de detección facial de Viola&Jones”, Documento 1, 23 de Marzo del 2009
- [96] Mathworks Inc. (2017), “¿Pricing and Licensing? “, [Online]. Disponible en: <https://es.mathworks.com/pricing-licensing.html?prodcode=ML&intendeduse=home>