

RECONOCIMIENTO FACIAL A TRAVÉS DE CÁMARAS DE VIDEO VIGILANCIA

Ing. Gladis Guzmán Guerrero¹, M.T.I. Rafael Hernández Reyna², M.T.I. Eloy Cadena Mendoza³.

Resumen— En los últimos años, el desarrollo de ingenios y programas informáticos para sistemas de seguridad ha experimentado un gran impulso. El reconocimiento facial es una herramienta importante para la seguridad, es por eso que el objetivo de este proyecto es diseñar una herramienta computacional que sea capaz de hacer reconocimiento facial de manera confiable a través de cámara de Video Vigilancia haciendo uso de un algoritmo de detección de objetos. El algoritmo de Viola Jones es un método de detección de objetos que se usa ampliamente en la detección de caras en imágenes y video. El algoritmo se basa en la comparación entre las intensidades luminosas de regiones rectangulares de las imágenes denominadas Haar-like features que calcula empleando una imagen integral.

Palabras clave— Detección facial, Reconocimiento Facial, Algoritmo Viola & Jones, Cámaras de Video Vigilancia, imagen Integral.

Introducción

El rostro humano es un objeto dinámico que tiene un alto grado de variabilidad en su apariencia lo cual hace que su detección sea un problema difícil de tratar en visión por computador. En los inicios de esta tecnología llamada “Reconocimiento facial” se usaba distintos algoritmos muy simples, el cual daba mayor oportunidad a que los errores se produjeran, ya que al ser así el mismo reconocimiento se podría dar para 2 personas diferentes. En la actualidad y con los avances logrados, además de los algoritmos que han sido exponencialmente mejorados, los errores son mínimos pues la manera en la que lleva a cabo el reconocimiento de rostros ha cambiado logrando obtener resultados más confiables. La detección de rostros es actualmente utilizada en diferentes aplicaciones informáticas y en diversas ramas de aplicación como son la seguridad de dispositivos, control de acceso a lugares restringidos, seguimiento de personas, entre otras. Los sistemas biométricos surgen como una solución real a los problemas de verificación. La biometría consiste de un conjunto de métodos automatizados para la verificación de individuos mediante el uso de características físicas o del comportamiento de la persona. Esta tecnología se basa en la premisa de que cada persona es única y posee rasgos distintivos que pueden ser utilizados para identificarla. El algoritmo que estamos utilizando tiene como finalidad, detectar los rostros humanos en una imagen para luego poder determinar la identidad de cada uno de ellos; proponiéndose una gran variedad de técnicas, desde algoritmos básicos de detección de bordes hasta algoritmos compuestos de alto nivel que utilizan métodos avanzados de reconocimiento de patrones. (Barroso, 2014)

Estas técnicas de detección se han abordado desde diferentes enfoques: Enfoques basados en rasgos faciales o características locales, en los que se buscan determinados elementos que componen el rostro, como los ojos, la nariz, la boca, Enfoques holísticos o basados en la imagen, en este caso los métodos trabajan con la imagen completa o zonas concretas de la misma de la cual se extraen características que puedan representar el objeto buscado, enfoques híbridos, estos métodos usan tanto la información local como la global para la detección, basándose en el hecho de que el sistema de percepción humano distingue tanto las características locales como globales del rostro, el objetivo principal del análisis de rostros es extraer información valiosa de las caras, como su posición en la imagen, las características, las expresiones, el género o la identidad de la persona. (Barroso, 2014)

El objetivo principal del análisis de rostros es extraer información valiosa de las caras, como su posición en la imagen, las características, las expresiones, el género o la identidad de la persona para poder lograr una alta efectividad y funcionamiento en la detección facial de nuestro algoritmo ya que es fundamental para obtener un reconocimiento de rostros preciso a través cámaras de video vigilancia.

El presente trabajo se apoya en el campo de la visión computacional, en especial en los métodos de verificación facial para la creación de un sistema de reconocimiento facial a través de cámaras de video vigilancia.

¹ La Ing. Gladis Guzmán Guerrero, es Alumna de la Maestría en Sistemas Computacionales en el Tecnológico nacional de México, Instituto Tecnológico de Acapulco, Guerrero. gladis_guzman@hotmail.com (autor corresponsal)

² El M.T.I. Rafael Hernández Reyna, es Docente en la Maestría en Sistemas Computacionales en el Tecnológico nacional de México, Instituto Tecnológico de Acapulco, Guerrero. rhernan7@yahoo.com.mx

³ El M.T.I. Eloy Cadena Mendoza es Docente en la Maestría en Sistemas Computacionales en el Tecnológico nacional de México, Instituto Tecnológico de Acapulco, Guerrero. eloy_cadena@yahoo.com.mx

Planteamiento del Problema

La inseguridad en nuestro país se ha convertido en un problema. Hoy en día, es una de las principales características de todas las sociedades modernas, Guerrero es uno de los 5 estados más inseguros y delictivos del país según la revista Forbes [Solis, 2018]. Desafortunadamente esta situación ha pasado a afectar a instituciones educativas, es por eso que se deben tomar medidas de seguridad no solo de personal capacitado sino también haciendo uso de la tecnología inteligente. Hoy en día, es una de las principales características de todas las sociedades modernas, actualmente la extensión de la violencia se ha desbordado en un clima generalizado de criminalidad. En este trabajo se pretende hacer uso de la inteligencia artificial, utilizando algoritmos de reconocimiento facial para resultados de identidad confiable.

Objetivo General

El objetivo general de este proyecto, es el desarrollo de un sistema que sea capaz de obtener un reconocimiento facial confiable en imágenes, obtenidas a través de cámaras de video vigilancia, haciendo uso del algoritmo Viola-Jones.

Descripción del Método

El reconocimiento facial es una herramienta que nos permite identificar a una persona automáticamente por medio de una imagen digital. Para realizar un reconocimiento facial se deben analizar las características faciales de la persona, las cuales se pueden extraer ya sea de una fotografía o desde un fotograma en una fuente de video, esta se convierte en una plantilla y luego se compara con las imágenes en una base de datos para verificar la identidad de la persona. (Alfredo et al., 2017)

El reconocimiento facial se obtiene por medio de un proceso, este proceso se divide en 7 etapas como se muestra en la figura 1, como lo son: Pre-procesamiento, Segmentación, Extracción de Características, Detección facial y Clasificación para finalmente identificar la persona.

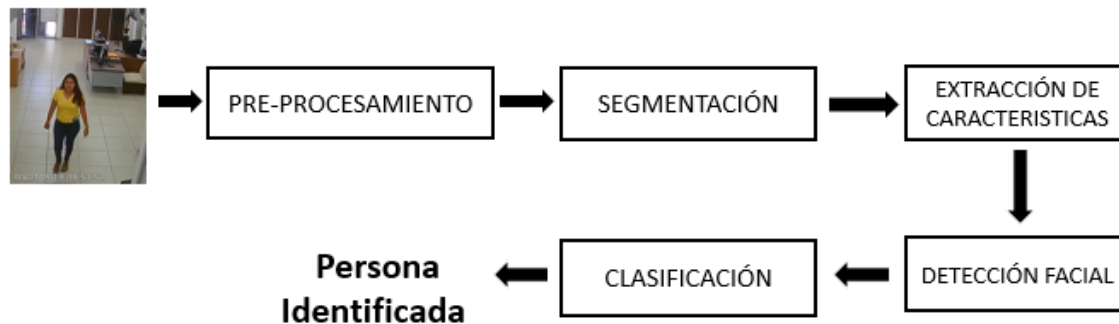


Figura 1. Metodología de Software.

Pre-procesamiento

El procesamiento de imágenes digitales es el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar la calidad o facilitar la búsqueda de información. Su objetivo es reducir el entorno que no es de interés para el problema. Fondo, ruido, entre otras. (Iii, 2004)

Segmentación

Consiste en partir una imagen de entrada en sus partes u objetos constituyentes. En general, la segmentación autónoma es una de las labores más difíciles del tratamiento digital de imágenes. Es el proceso en que se divide una escena percibida en sus partes para extraer objetos para su posterior reconocimiento y análisis. Por lo que una subdivisión del espacio de parámetros llamados células de acumulación donde (a_{\max}, a_{\min}) y (b_{\max}, b_{\min}) son los intervalos. (Iii, 2004)

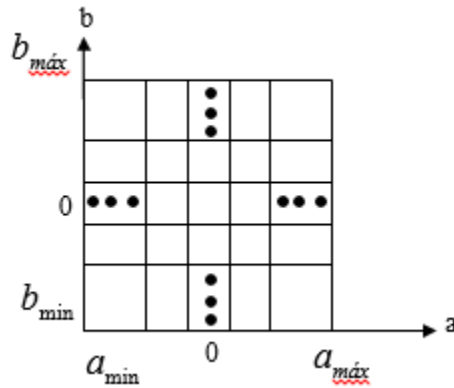


Figura 2. División del plano de parámetros en células. (Iii, 2004)

La imagen es dividida en una matriz de 3x3, inicialmente se desarrolla este tipo de segmentación, para disminuir problemas de confusión de piezas. Normalmente este uso de segmentación en células acumulativas se implementa en algoritmos básicos de segmentación de tipo similitud como es la transformada de Hough. El tipo de segmentación utilizado es Segmentación General. La mayoría de procesamiento de imágenes se aplica directamente sobre imágenes en escala de grises, debido al bajo consumo de cómputo de éste; la mayoría de métodos matemáticos deterministas que se utilizan en su procesamiento, están basados en la diferencia de niveles de grises, por lo que no existen muchos métodos para procesamiento de imágenes en color, aun cuando estos pueden ser utilizados en este tipo de formato de imagen. (Iii, 2004)

La detección del rostro se lleva a cabo aplicando el concepto de imagen integral implementada en el algoritmo desarrollado por Viola y Jones. Este algoritmo utiliza una imagen integral para extraer características de forma rápida y precisa, debido a que no trabaja directamente con los valores de intensidad de los píxeles, sino que lo hace a través de una imagen acumulativa que se va formando basada en las operaciones básicas que se realizan a medida que se va recorriendo la imagen. (Barroso, 2014)



Figura 3. Imagen integral. (Segmentación, n.d.)

La figura 3 ilustra la aplicación de este proceso con el fin de obtener la imagen integral (recuadro superior derecho) a partir de la imagen original $Im(x, y)$. La imagen integral se obtiene al realizar un desplazamiento de izquierda a derecha y de arriba hacia abajo realizando la suma de los píxeles a medida que se va desplazando en la localización de los puntos (x, y) , como se muestra en la figura 3. (Segmentación, n.d.)

Extracción de Características

La extracción de características en imágenes y secuencias de imágenes faciales, consiste en extraer información asociada con la activación de los diferentes músculos del rostro, esta tarea puede realizarse en forma global u holística en donde se analiza el rostro como un solo conjunto o localmente en donde se seleccionan regiones de interés del rostro como ojos cejas y boca. Adicionalmente los métodos de extracción de características faciales se pueden clasificar de acuerdo a su enfoque en: características faciales de movimiento o deformaciones del rostro. Inicialmente, el algoritmo necesita muchas imágenes positivas (imágenes de caras) e imágenes negativas (imágenes sin caras) para entrenar al clasificador. Entonces necesitamos extraer características de él. (Barroso, 2014)

En imágenes las características de cada objeto se extraen al aplicar ciertas funciones que permitan la representación y descripción de los objetos de interés de la imagen (patrones). La extracción de características es un paso en el reconocimiento de patrones en el cuál las medidas u observaciones son procesadas para encontrar atributos que puedan ser usados para asignar los objetos a determinada clase (Correa & Chichizola, 2001). La extracción de características es realizada aplicando a la imagen filtros con bases Haar. Estos filtros pueden ser calculados eficientemente sobre la imagen integral. Los filtros con bases Haar, realizan una codificación de diferencia de intensidades en la imagen, generando características de contornos, puntos y líneas, mediante la captura de contraste entre regiones.

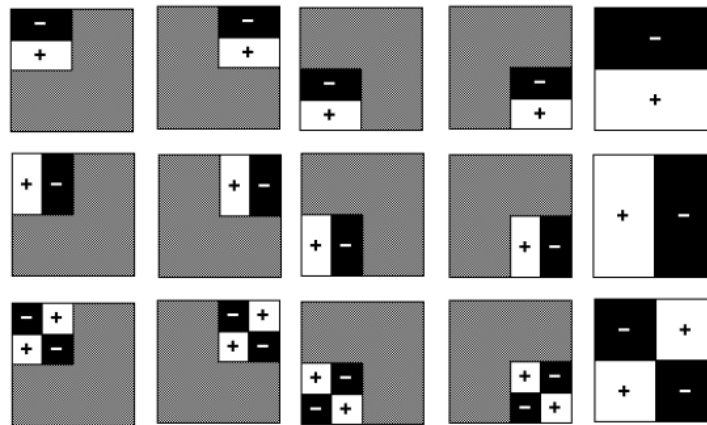


Figura 4. Filtros Haar rotados, trasladados y con cambios de escala

Inicialmente en esta etapa de extracción de características, el algoritmo necesita imágenes positivas (imágenes de caras) e imágenes negativas (imágenes sin caras) para entrenar al clasificador. Entonces necesitamos extraer características de él. Para esto, se utilizan las características de bases Haar, como se ilustra en la figura 4. En el trabajo de Viola-Jones, las características se definen sobre una ventana de búsqueda básica de 24×24 píxeles, lo que da lugar a más de 160,000 características posibles. Ahora todos los tamaños y ubicaciones posibles de cada ventana se usan para calcular muchas características. Para cada cálculo de características, necesitamos encontrar la suma de píxeles en rectángulos blanco y negro. Para resolver esto los autores del algoritmo Paul Viola y Michael Jones, introdujeron las imágenes integrales. (Alhamzawi, 2018)

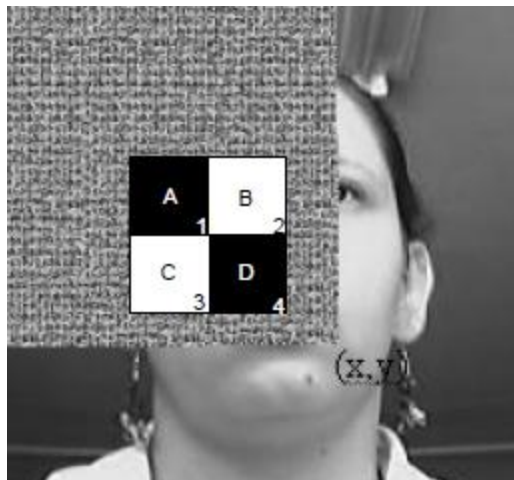


Figura 5. Convolución del filtro Haar con una imagen integral. (Alhamzawi, 2018)

De la convolución de un filtro Haar con la imagen integral, se puede extraer una característica en un tiempo constante sobre la imagen integral adicionando y sustrayendo los valores de los vértices para cada rectángulo como se muestra en la figura 5. Para mayor claridad, en la figura la suma de los píxeles que forman el rectángulo D se puede calcular como:

$$\text{sumD} = (4 + 1) - (2 + 3) \quad (2)$$

Donde 1, 2, 3, 4, son los valores dados en la imagen integral en dichas localizaciones. (Alhamzawi, 2018)

Detección facial

La detección facial es una tecnología de visión computarizada que determina el lugar y el tamaño de rostros humanos en imágenes o videos. Es uno de los subtipos de detección de objeto/clase, cuya tarea es encontrar la localización y el tamaño de los objetos en una imagen. Mientras que para el ser humano es trivial esta tarea, una computadora presenta dificultades debido a diversos factores como: variabilidad en la posición del rostro, presencia o ausencia de componentes estructurales (ejemplo: bigote), expresión facial, oclusión (uso de lentes, gorro, entre otras.) y condiciones del ambiente. La detección facial es una etapa importante en un sistema de reconocimiento facial, sin embargo influye de manera significativa en el resultado del proceso, ya sea dado un conjunto de imágenes o video en tiempo real. Por lo tanto, debe ser capaz de identificar los rostros independientemente de los factores que anteriormente se mencionaron. Los métodos de detección facial se dividen en cuatro categorías (no excluyentes):

- Métodos basados en conocimiento: Codifican el conocimiento humano mediante distancias y posiciones entre las características humanas (ojos, nariz, labios).
- Métodos basados en características invariantes: Las características invariantes son aquellas que no se modifican a eventuales cambios de luz, pose o ubicación de la cámara, tales como la ceja, nariz, textura de la piel y línea de pelo. Este método funciona detectando uno de estos componentes, construyendo un modelo estadístico y con los resultados, verificar la existencia de un rostro.
- Métodos basados en moldes (patrones): Es la relación entre una imagen de entrada y un patrón o molde previamente definido, cuyo objetivo es capturar características del rostro.
- Métodos basados en apariencia: Utilizan modelos obtenidos mediante entrenamiento de imágenes, tomando la imagen como un vector de características, es decir, es visto como una variable aleatoria. A diferencia de los métodos basados en moldes, donde el patrón es definido por un “experto”, los patrones en este modelo son determinados por el aprendizaje obtenido en el entrenamiento de imágenes.

Existen varios métodos para detectar rostros, sin embargo el algoritmo de Viola-Jones, es el más eficaz, obteniendo un mayor porcentaje de aciertos respecto a sus pares, además de más rapidez. Este algoritmo integra un nuevo concepto, la imagen integral, que junto con el algoritmo de boost como método de entrenamiento, forman un clasificador complejo y preciso (Alfredo et al., 2017).

Clasificación

Esta etapa dentro del algoritmo de detección se encarga de asignar un conjunto de características dado a una clase con la que se encuentra una mayor similitud, de acuerdo a un modelo inducido durante el entrenamiento. Boosting es un método de clasificación que combina varios clasificadores básicos para formar un único clasificador más complejo y preciso. La idea se basa en la afirmación de que varios clasificadores sencillos, cada uno de ellos con una precisión ligeramente superior a una clasificación aleatoria, pueden combinarse para formar un clasificador de mayor precisión, siempre y cuando se disponga de un número suficiente de muestras de entrenamiento. La aplicación de clasificadores en cascada ha permitido obtener buenos resultados. En la figura 6, se muestra un esquema de un clasificador en cascada. (Alhamzawi, 2018)

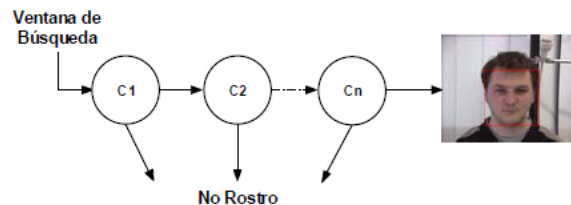


Figura 6. Clasificador en cascada

Para aplicar la técnica de boosting primero se debe establecer un algoritmo de aprendizaje sencillo (clasificador débil o base), que será llamado repetidas veces para crear diversos clasificadores base. Para el entrenamiento de los clasificadores base, se emplea, en cada iteración, un subconjunto diferente de muestras de entrenamiento y una distribución de pesos diferente sobre las muestras de entrenamiento (Alhamzawi, 2018). Finalmente, estos clasificadores base se combinan en un único clasificador que se espera sea mucho más preciso que cualquiera de los clasificadores base por separado.

Clasificador AdaBoost

AdaBoost es un algoritmo predictivo para clasificación y regresión. AdaBoost (refuerzo adaptativo) es un algoritmo de aprendizaje conjunto que puede utilizarse para la clasificación o regresión. Si bien AdaBoost es más resistente al sobreajuste que muchos algoritmos de aprendizaje automático, a menudo es sensible a datos ruidosos y valores atípicos. (The MathWorks, 2019)

AdaBoost se llama adaptativo porque utiliza múltiples iteraciones para generar un único aprendiz compuesto fuerte. AdaBoost crea el aprendiz fuerte (un clasificador que está bien correlacionado con el clasificador verdadero) agregando iterativamente los aprendices débiles (un clasificador que solo está ligeramente correlacionado con el clasificador verdadero). Durante cada ronda de entrenamiento, se agrega un nuevo alumno débil al conjunto y se ajusta un vector de ponderación para centrarse en los ejemplos que se clasificaron incorrectamente en las rondas anteriores. El resultado es un clasificador que tiene mayor precisión que los clasificadores de los alumnos débiles. (The MathWorks, 2019)

Referencias

- Alfredo, G., Lagla, F., Carlos, J., Chisag, C., Alejandro, O., & Pico, G. (2017). *Reconocimiento facial*.
- Alhamzawi, H. A. M. (2018). Faces and eyes Detection in Digital Images Using Cascade Classifiers. *Computer Engineering and Applications Journal*, 7(1), 57–66. <https://doi.org/10.18495/comengapp.v7i1.222>
- Barroso, C. (2014). *Implementación del algoritmo de detección de caras de viola y jones sobre una fpga*. 1–129.
- Correa, M. S., & Chichizola, F. (2001). *Diseño de Sistemas de Reconocimiento de Rostros*. Retrieved from http://sedici.unlp.edu.ar/bitstream/handle/10915/3874/Documento_completo__pdf-PDFA2.pdf?sequence=1
- Iii, C. (2004). *Procesamiento de imágenes capítulo iii* 33. 33–48. Retrieved from http://catarina.udlap.mx/u_dl_a/tales/documentos/msp/florencia_y_an/capitulo3.pdf
- Segmentación, P. De. (n.d.). *Prácticas de Robótica y Visión Artificial*. 0, 1–11.
- B. Fasel, J. Luetin. Automatic Facial Expression Analysis, *Pattern Recognition* 36 259 – 275, Science Direct, 2003.
- IEEE *Standar Glossary for Image Processing and Pattern recognition Terminology*. Published by the Institute of Electrical and Electronics Engineers, New York, USA, 1990. 4
- The MathWorks, I. (2019). *Matlab*. Obtenido de Matlab: <https://la.mathworks.com/discovery/adaboost.html>