

Un nuevo método para la detección de caras basado en Integrales Difusas

Antonio Rama⁽¹⁾, Francesc Tarrés⁽¹⁾

arama@gps.tsc.upc.edu, tarres@gps.tsc.upc.edu

⁽¹⁾ Dept. Teoria del Senyal i Comunicacions - Universitat Politècnica de Catalunya, Barcelona, Spain

Abstract- A novel face detector using the non-linear Fuzzy Integral operator is presented in this paper. The proposed approach is composed of four different stages combined in a cascade scheme that gradually rejects the non-face objects. Each of these stages use only a few features following an strategy very similar to the well-known Adaboost technique. The main advantage of this proposal in front of Adaboost is that it has a significantly lower false detection rate when using the same set of optimal features as the state-of-the art Adaboost face detector. Simulations show a positive face detection rate higher than the 92% having a false detection rates lower than the 1% when using a four cascade scheme.

I. INTRODUCCIÓN

La detección de caras es una etapa fundamental en cualquier aplicación donde se realice algún tipo de análisis facial como por ejemplo reconocimiento de caras, codificación de vídeo en videoconferencias, interfaces inteligentes hombre-máquina, etc. El objetivo de esta etapa consiste en detectar y localizar la posición de un número indefinido de caras en una imagen. En general, la detección de caras es un problema muy complejo ya que los objetos a detectar pueden ser de diferentes colores, expresiones, poses, tamaños relativos o tener condiciones de iluminación muy dispares.

Se han propuesto numerosas técnicas para la detección de caras, pero básicamente se podrían agrupar en dos grandes categorías:

- **Métodos basados en reglas.** Se establecen relaciones (reglas) entre las diferentes características faciales como por ejemplo la distancia entre ojos o “simetría” de la cara [1, 2].
- **Métodos estadísticos.** Este tipo de algoritmos no asumen ningún tipo de información previa de la tipología de una cara; sino, que a partir de un conjunto de muestras (imágenes de caras y de no-caras) de entrenamiento extraen la información relevante que diferencia un objeto *cara* de un objeto *no cara* [3, 4]. Este grupo incluye uno de los métodos más referenciados y utilizados actualmente: el detector de caras *AdaBoost* [5].

Se puede encontrar una clasificación más detallada de los métodos existentes para la detección de caras en las referencias [2, 6], que son un resumen muy completo de las tecnologías actuales utilizadas para la resolución de este problema.

En este trabajo, se presenta un nuevo detector de caras basado en un operador no lineal denominado Integral Difusa (*Fuzzy Integral*). Los resultados obtenidos indican que este

método puede representar una mejora significativa respecto al detector de caras *AdaBoost*.

El resto del artículo está organizado de la siguiente manera: en los apartados 2 y 3, se explican brevemente las bases fundamentales sobre el método *Adaboost* y el operador no lineal basado en la Integral Difusa (*Fuzzy Integral*). En el apartado 4 se detalla el algoritmo de detección de caras propuesto en este trabajo, mientras que el siguiente apartado se describen las simulaciones realizadas y algunos resultados preliminares. Finalmente, se presentan las conclusiones y líneas de trabajo futuras.

II. DETECCIÓN DE CARAS MEDIANTE ADABOOST

El método conocido como Adaboost fue presentado originalmente por Viola y Jones [5] y es una extensión de un clasificador genérico al problema de la detección de objetos en imágenes. Éste método demuestra como a partir de características locales basadas en el cambio de intensidad se podía desarrollar un detector de caras muy robusto. La idea básica es la siguiente. Se determinan una serie de características basadas en las sumas y restas de los niveles de intensidad en la imagen. Para ello se utilizan filtros de *Haar* de un cierto tamaño y calculados para las posiciones concretas de la sub-imagen que se quiere clasificar. Dichas características son evaluadas por un “clasificador débil” para decidir si la sub-imagen corresponde a una cara (aceptada) o no (rechazada) tal y como muestra la Fig. 1. Si el valor de la característica está por encima de un cierto umbral θ , entonces la ventana se clasificará como cara. Este tipo de “clasificadores débiles” suelen conseguir unos resultados muy pobres. No obstante, combinando varios módulos como el de la Fig. 1, se pueden generar clasificadores más robustos (“clasificador fuerte”) cuya tasa de detección crece exponencialmente. El rectángulo punteado de la Fig. 2 corresponde a un “clasificador fuerte”. De todas maneras, a pesar que la tasa de detección de caras de un “clasificador fuerte” puede llegar hasta el 99%, presenta la desventaja de que la tasa de falsas detecciones (aceptar una sub-imagen como cara cuando no lo es) es todavía inaceptable al estar por encima del 30%. Por este motivo Viola y Jones propusieron un esquema basado en una cascada de

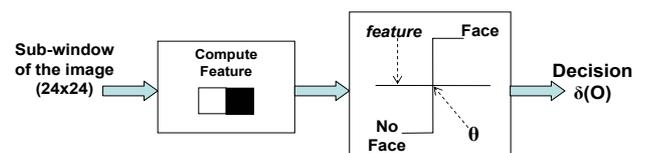


Fig. 1 Adaboost: Clasificador débil

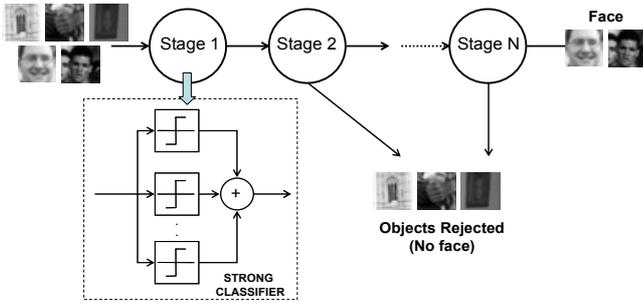


Fig. 2 Detector de caras Adaboost

“clasificadores fuertes” como el representado en la Fig. 2. Cada etapa corresponde a un “clasificador fuerte” y está entrenada con todos los ejemplos que la etapa anterior no ha podido clasificar correctamente más algunos nuevos. Por tanto, en la etapa de entrenamiento, cada etapa se entrena con un conjunto óptimo de características capaces de detectar cada vez ejemplos más complicados; es decir, las primeras etapas se encargan de descartar sub-imágenes que son muy diferentes de una cara, mientras que las últimas etapas pueden rechazar ejemplos mucho más complicados como pueden ser pelotas, globos, dibujos, etc...

III. FUNDAMENTOS SOBRE LA INTEGRAL DIFUSA

La teoría sobre *Medidas Difusas* está basada en el trabajo de Sugeno [7]. La introducción de conjuntos *difusos* [8], resultado de la generalización de la teoría clásica de conjuntos, dio resultado a la redefinición de las medidas probabilísticas clásicas. Sugeno propuso las medidas difusas, como elemento clave para la definición de la Integral Difusa. En este apartado, únicamente se presentan los conceptos elementales sobre la Integral Difusa y su posible aplicación en problemas de clasificación. Para más detalles sobre ésta, el lector puede consultar la siguiente literatura [9].

La idea básica que se persigue en éste artículo es la de utilizar un clasificador de caras basado en Integrales Difusas con un conjunto de características que midan los cambios de contraste existentes en la sub-imagen. Las Integrales Difusas son generalizaciones de los operadores integrales que incluyen operaciones no-lineales sobre los datos. En el contexto de clasificación de patrones, las Integrales Difusas más usadas son la Integral *Sugeno* y la Integral *Choquet*. En este artículo, se propone el uso de esta última para la fusión y clasificación de los datos.

Considerando que se tiene el conjunto de características $X = \{x_1, x_2, \dots, x_n\}$ donde x_i puede representar un píxel, una muestra de audio o, como en nuestro caso, una característica Haar en una posición determinada dentro de la sub-imagen que se desea clasificar. Dado este conjunto de características, se obtiene un número de M muestras para la fase de entrenamiento. Los atributos de cada característica para cada muestra vendrán representados por un vector

$$f = \{f(x_1), f(x_2), \dots, f(x_n)\} \quad (1)$$

La Integral *Choquet* está definida en términos de un número finito de medidas difusas que representan la importancia *a priori* de las características. Este conjunto de medidas difusas está definido para cada característica por separado y todas sus posibles combinaciones. Más concretamente, las medidas difusas son:

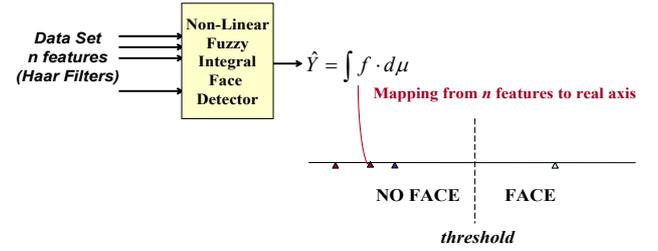


Fig. 3 Detector de caras basado en Integrales Difusas

$$\mu = \{\mu(x_1), \mu(x_2), \mu(x_3), \dots, \mu(x_1, x_2), \dots, \mu(x_1, x_2, x_3), \dots\} \quad (2)$$

La integral *Choquet* consiste en un proceso de dos etapas:

- 1) Reordenación de los valores de las características en orden no decreciente:

$$f(x_1') \leq f(x_2') \leq \dots \leq f(x_n') \quad (3)$$

donde $(x_1', x_2', \dots, x_n')$ es una cierta permutación de (x_1, x_2, \dots, x_n) .

- 2) La integral *Choquet* se obtiene calculando la siguiente expresión:

$$\int f \cdot d\mu = \sum_{i=1}^n [f(x_i') - f(x_{i-1}')] \cdot \mu(\{x_i', x_{i+1}', \dots, x_n'\}) \quad (4)$$

El entrenamiento del clasificador consiste en seleccionar el conjunto óptimo de medidas difusas que minimice el número de datos clasificados incorrectamente. Existen varias posibilidades para estimar las medidas difusas de la ecuación (2), pero casi todas están basadas en estrategias de *soft-computing*. En este trabajo, se ha seguido una estrategia basada en redes neuronales para estimar el conjunto de medidas difusas.

Una de las características de la Integral Difusa es que, una vez se ha realizado el entrenamiento y se dispone de una estimación de las medidas difusas el proceso de clasificación es computacionalmente muy eficiente. Tal y como muestra la Fig. 3, el clasificador basado en Integrales Difusas mapea el conjunto de datos (características) de entrada a un único escalar (eje real). Luego, dependiendo de un cierto umbral, este valor se clasificará como una cara o una no cara. En el siguiente apartado, se detalla el esquema propuesto, los experimentos realizados para evaluar dicho esquema y los resultados obtenidos comparados con el método de detección de caras *Adaboost*.

IV. ESQUEMA PROPUESTO PARA EL DETECTOR DE CARAS BASADO EN INTEGRALES DIFUSAS

A. Selección de las características

En este artículo se propone un detector de caras compuesto de una cascada de clasificadores basados en Integrales Difusas (Fig. 4). Uno de los principales inconvenientes al usar Integrales Difusas es que el número de parámetros (*Medidas Difusas*) crece exponencialmente con el número de características de entrada usadas para entrenar el sistema. Por este motivo, sería impracticable intentar entrenar el sistema con las mismas características

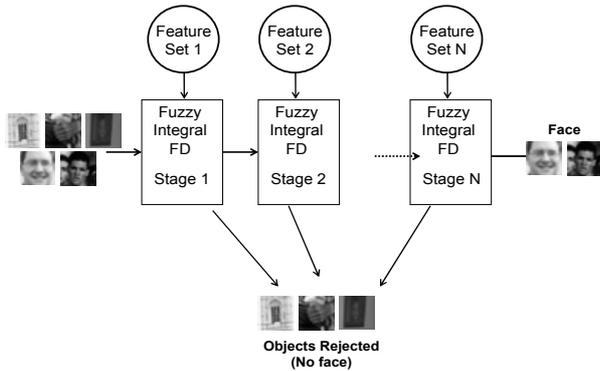


Fig. 4 Esquema para la detección de caras: Cascada de clasificadores basados en Integrales Difusas.

propuestas para el método de detección de caras *Adaboost* [5, 11] ya que sería necesario un conjunto extremadamente grande de datos de entrenamiento. Si se tienen en cuenta todas las posibles posiciones de todas las características Haar de la Fig. 5 [5], en cada sub-imagen de 24x24 píxeles se obtendrían un total de 117,941 características de entrada en el sistema. Por esta razón, el detector de caras propuesto utilizará el subconjunto de características calculadas previamente para un detector de caras *Adaboost*. Después de varias pruebas exhaustivas [11], se ha seleccionado la siguiente configuración del método *Adaboost*:

- 11 etapas, es decir, 11 “clasificadores fuertes” combinados en cascada.
- 3325 ejemplos de caras (el mismo conjunto para todas las etapas) + 4500 imágenes que no contienen cara alguna por etapa.
- Todo el conjunto de características presentado en la Fig. 5.
- Mínima tasa de detección de caras del 99.5% en cada etapa.
- Máxima tasa de falsa detecciones del 30% en cada etapa.

Usando esta configuración inicial, el método *Adaboost* se entrena, obteniendo así un subconjunto de filtros Haar para cada etapa del esquema representado en la Fig. 2. Por ejemplo, el primer “clasificador fuerte” (primera etapa de la cascada) del detector de caras *Adaboost* utiliza únicamente un total de 6 características: Haar-Y2 en 3 posiciones diferentes, Haar-X4 en 2 posiciones, y Haar-Y4 en una posición. Estas 6 características se utilizarán también para entrenar la primera etapa del detector de caras basado en Integrales Difusas (ver Fig. 4).

B. Entrenamiento de los clasificadores basados en Integrales Difusas

En [9, 7] los autores propusieron algoritmos genéticos para entrenar el sistema. En este caso, se usará un algoritmo de aprendizaje basado en la siguiente ecuación de control

$$\mu^{i+1}(\overline{f(x')}) = \mu^i(\overline{f(x')}) + \sigma \cdot error \cdot \overline{\Delta f(x')} \quad (5)$$

donde $\overline{f(x')}$ son las características normalizadas previamente reordenadas en orden decreciente tal y como se menciona en la ecuación (3), σ es el incremento adaptativo y $error$ es un parámetro que puede tomar los valores -1, 0 ó 1 dependiendo de si la decisión del clasificador (0 significa que la muestra de entrenamiento ha sido clasificada de manera correcta). Y finalmente, $\overline{\Delta f(x')}$ es la diferencia entre

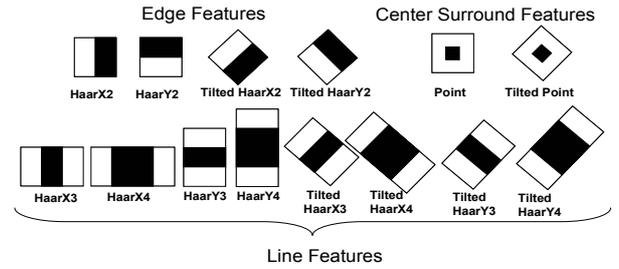


Fig. 5 Conjunto de características utilizadas para entrenar al sistema.

todas las características involucradas en la medida difusa que se está actualizando actualmente.

V. PRUEBAS EXPERIMENTALES Y RESULTADOS

A. Base de datos utilizada para realizar las pruebas

Para todas las simulaciones y pruebas experimentales que se han realizado, se ha utilizado una base de datos propia compuesta por un total de 3325 imágenes de caras que han sido previamente normalizadas a una resolución de 24x24 píxeles. Para los ejemplos negativos (no caras) se han utilizado más de 2000 imágenes de diferentes resoluciones que no incluyen ninguna cara (paisajes, objetos, animales, etc...). Si estas 2000 imágenes se dividen en sub-imágenes de 24x24 píxeles se obtienen más de 2 millones de ejemplos negativos.

La mitad de los ejemplos positivos y solamente unas 50000 ejemplos negativos se han utilizado para entrenar un detector de caras basado en Integrales Difusas de 4 etapas. El resto de las muestras se han utilizado para la fase de test para extraer resultados.

B. Resultados para un detector de caras basado en Integrales Difusas de 4 etapas

Se ha implementado un esquema de 4 etapas basado en integrales difusas como el que se muestra en la Fig. 4. Cada una de las 4 etapas utilizará los mismos conjuntos de características *Haar* que se han obtenido al entrenar la configuración óptima del *Adaboost* presentada en la sección IV. De esta manera, en las etapas 1, 2, 3 y 4 se utilizarán 6, 9, 11 y 21 características Haar diferentes respectivamente. Para que una sub-imagen sea etiquetada como cara, ésta tendrá que pasar todas las etapas.

La tasa de detección de caras obtenida es superior al 92%, y se rechazan casi todas las sub-imágenes (99%) que no corresponden a caras. La primera etapa de la cascada por sí sola ya rechaza más del 90% de los objetos no-caras. La Fig. 6 representa un ejemplo extremo de este concepto. La imagen de test ha sido analizada con el detector de caras propuesto en este artículo, y con el detector de caras *Adaboost*. Los dos métodos han sido implementados sin ningún tipo de post procesamiento para eliminar ventanas solapadas.

Los resultados muestran que el detector mediante Integrales Difusas de 4 etapas detecta todas las caras y es capaz de rechazar casi la totalidad de sub-imágenes que no contienen una cara. Únicamente aparece una falsa detección que no está completamente solapada con una sub-imagen clasificada como cara.

Por el contrario el detector *Adaboost* de 4 etapas es capaz de detectar todas las caras (una no está totalmente

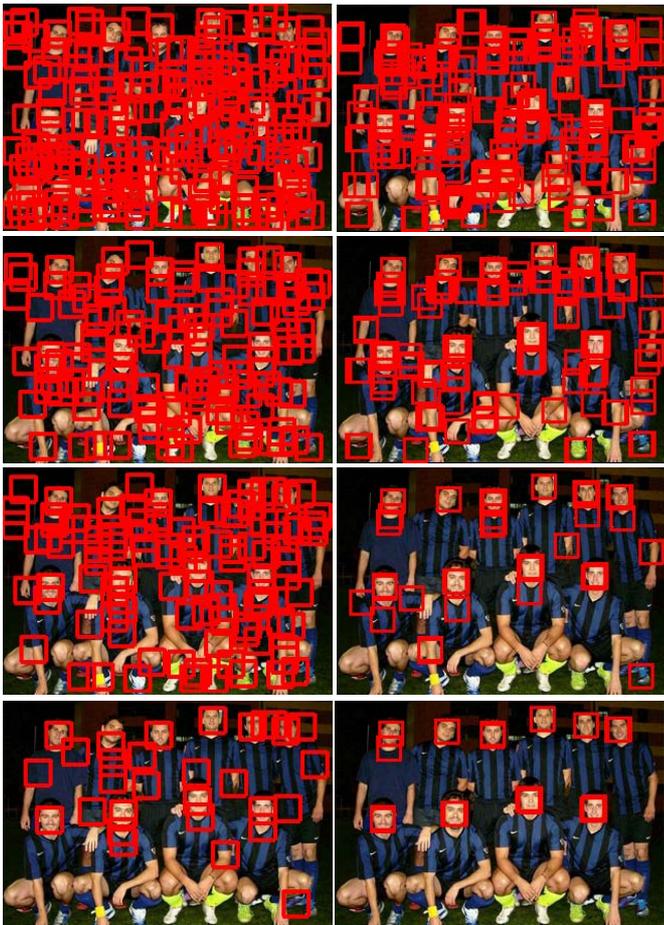


Fig. 6 Resultados comparativos. Detector de caras *Adaboost* (parte izquierda) y propuesto en este artículo (parte derecha) (De arriba a abajo las imágenes corresponden con las salidas de cada una de las 4 etapas de la cascada).

incluida en el rectángulo) pero todavía existen más de 25 sub-imágenes correspondientes a no-caras que han sido aceptadas (además únicamente la mitad de ellas están parcialmente solapadas para poderlas eliminar con un post procesado).

Se puede observar en todo momento como el detector de caras propuesto en este artículo, presenta en todas las etapas un número muy inferior de falsas alarmas que en el caso de la técnica *Adaboost*.

Además, continuando con la misma imagen de test, si se incluyen más etapas en la cascada del detector de caras *Adaboost*, el mejor resultado que es posible obtener, es el ilustrado en la Fig. 7, donde únicamente 8 de las 10 caras que aparecen en la imagen son detectadas. Estos resultados son claramente peores que los obtenidos con las 4 etapas del detector propuesto basado en Integrales Difusas.

VI. CONCLUSIONES Y TRABAJO FUTURO

En este artículo se ha presentado un nuevo detector de caras basados en Integrales Difusas. Los primeros resultados presentan un mejor compromiso entre detecciones correctas y falsas detecciones que el método de referencia *Adaboost*. De todas formas, la tasa de detección es similar en ambos algoritmos, por ello es necesario realizar un análisis más extenso para determinar en que condiciones o restricciones un método funciona mejor que otro.

Otro aspecto importante que se debería estudiar es el valor de las medidas difusas. Este punto puede ser muy



Fig. 7 Mejores resultados obtenidos por el clasificador *Adaboost* (7 etapas).

interesante ya que estas miden de alguna manera la importancia relativa de dichas características en la clasificación final. Esto ayudaría a elegir del conjunto total de características aquellas que son realmente importantes e influyen en la clasificación. De esta manera se reduciría el número de características a calcular en la fase de detección, disminuyendo así el coste computacional del algoritmo.

El nuevo detector de caras propuesto en este artículo, no solo presenta resultados muy esperanzadores comparables a la tecnología actual, sino que además abre una serie de nuevas posibilidades para mejorar dichos resultados.

AGRADECIMIENTOS

El trabajo de investigación presentado ha sido desarrollado dentro del proyecto VISNET-II, Red de Excelencia Europea financiada por el programa IST FP6 de la Comisión Europea y del proyecto de investigación CICYT TEC2005-07751-C02-02 financiado por el gobierno español.

The work presented was developed within VISNET II, a European Network of Excellence funded under the European Commission IST FP6 programme, and within the R+D Spanish project CICYT TEC2005-07751-C02-02.

REFERENCIAS

- [1] G. Yang and T. S. Huang, "Human Face Detection in Complex Background," *Pattern Recognition*, vol. 27, no. 1, pp. 53-63, 1994
- [2] M.-H. Yang, D. Kriegman and N. Ahuja, "Detecting Faces in images: a Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, n°1, pp.34-58, 2002
- [3] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, January 1998, 23-38
- [4] M. A. Turk, A. P. Pentland, "Face recognition using eigenfaces", *Proceedings of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp. 586-591, Maui, Hawaii 1991
- [5] Viola P., Jones M.: *Rapid Object Detection using a Boosted Cascade of Simple Features*, *Computer Vision and Pattern Recognition*, 2001
- [6] E. Hjelm and B. K. Low, "Face Detection: A Survey", *Computer Vision and Image Understanding*, vol. 83, no. 3, pp. 236-274, September 2001
- [7] M. Sugeno, "The Theory of Fuzzy Integrals and Its Applications", PhD thesis, Tokyo Institute of Technology, Japan, 1974.
- [8] L. A. Zadeh, *Fuzzy sets*, *Information Control*, (1965), pp. 338-353.
- [9] Aureli Soria-Frisch "Soft Data Fusion in Computer Vision", PdD thesis, Fraunhofer Institut fuer Produktionsanlagen und Konstruktionstechnik, Berlin, Germany, May 2004
- [10] R. Lienhart, and J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection". in *IEEE ICIP 2002*, Vol. 1, pp. 900-903, September 2002
- [11] J. M. Braup, F. Tarres, "Anàlisi de vídeo per la detecció automàtica de cares". Master Thesis in Catalan at the Technical University of Catalonia – Escola Politècnica Superior de Castelldefels. October 2006.