

Caracterización de unidades de acción facial combinando métodos kernel y análisis de componentes independientes

Feature extraction of facial action units combining kernel methods and independent component analysis

Damián Alberto Álvarez, Juan Gabriel Fetecua, Álvaro Ángel Orozco, César Germán Castellanos*

Grupo de investigación en control e instrumentación, Laboratorio de investigación en instrumentación y medidas, E-106, Programa de Ingeniería Eléctrica, Universidad Tecnológica de Pereira, A.A. 97. Vereda la Julita, Pereira, Risaralda, Colombia

(Recibido el 10 de noviembre. Aceptado el 9 de marzo de 2010)

Resumen

El trabajo descrito en este documento presenta una metodología para la caracterización de unidades de acción facial (AUs), que representan cambios sutiles de las expresiones faciales. La metodología se basa en métodos Kernel, para realizar un mapeo no lineal de los datos y buscar las direcciones de las proyecciones de los datos en el espacio característico mediante análisis de componentes independientes (ICA). La validación se realiza sobre la base de datos Cohn-Kanade. Se hace un preprocesamiento de las imágenes a través de ecualización del histograma, un blanqueamiento de los datos con análisis de componentes principales basado en Kernel (KPCA), de esta forma el mapeo en el espacio característico busca una estructura lineal de los datos de entrada, finalmente se aplica ICA para hacer que la distribución de los datos proyectados sea lo menos Gaussiana posible. El desempeño alcanzado fue del $96.64\% \pm 0.54$ de exactitud para el reconocimiento promedio de tres combinaciones de AUs del rostro entero más rostros neutrales, se detectan principalmente cambios que ocurren entre transiciones rápidas de AUs que se manifiestan de forma instantánea. Adicionalmente la metodología planteada permite reducir el tamaño del espacio característico ya que se representan los datos en términos únicamente de sus componentes independientes (ICs) de tal

* Autor de correspondencia: teléfono: + 57 + 6 + 321 43 31, fax: + 57 + 6 + 313 71 22, correo electrónico: damianalvarez@ohm.utp.edu.co. (D. Álvarez)

manera que se utilizan tan solo las variables que aportan mayor información, lo que permite disminuir la complejidad del clasificador.

----- *Palabras clave:* Expresiones faciales, unidades de acción, análisis de componentes independientes, métodos kernel

Abstract

The work described in this paper presents a methodology for characterizing facial action units (AUs), which represents the subtle change of facial expressions, based on Kernels Methods perform a nonlinear mapping of data and looking for directions to the projections of the data in feature space through independent component analysis (ICA). The methodology validation was done on Cohn-Kande database. Image preprocessing was done through histogram equalization, a whitening on data with Kernel Principal Component Analysis (KPCA), for that the mapped in feature space search a lineal structure of the input data, finally we applied ICA for make the projected distribution of data is at least possible Gaussian. The results were 96.64% accurate for average recognition of three combinations of facial AUs of the whole face more neutral faces. Mainly changes that occur between rapid transitions of AUs instantly shown were detected. Additionally the proposed methodology can reduce the size of the feature space because data only in terms of independent components, are presented so as to use only those variables that provided greater information, which reduces the complexity of the classifier.

----- *Keywords:* Facial expression recognition, action unit, independent component analysis, kernel methods

Introducción

Los seres humanos expresan sus emociones, intenciones y opiniones de diversas maneras, siendo la más universal de estas las expresiones faciales. Aunque el reconocimiento de expresiones es natural para los seres humanos, dotar a una máquina de esta habilidad resulta bastante difícil [1].

Mediante la visión por computador es posible crear sistemas de medición automática de signos faciales a partir de video, que puedan ser usados en aplicaciones como: interacción hombre-máquina, salud mental, investigaciones comportamentales, incluyendo estudios de emoción, interacción social, comunicación, antropología y cualquier contexto en el cual sea importante monitorear las emociones, como también la comunicación paralingüística de las personas.

Con el fin de ayudar en el reconocimiento de expresiones en [2], se estableció una base de datos conocida como Cohn-Kanade (C-K), para clasificar AUs, acompañada de anotaciones del sistema de codificación de acciones faciales (FACS) en donde se codifican las imágenes del rostro en emociones específicas para cada expresión. Usando la base de datos C-K estudios como [3, 4] han trabajado en la clasificación de expresiones utilizando una representación en ICs y PCA como método de inicialización. La representación ICA tiene como ventaja que realiza una transformación a un espacio de menor dimensión, tarea a la que se le conoce como extracción de características [5], la cual dentro del reconocimiento de patrones debe ser abordada en condiciones de baja estadística. Utilizar un espacio de representación de alta dimensión puede presentar problemas en el

diseño del clasificador (como la maldición de la dimensionalidad), mientras que reduciendo la dimensionalidad del espacio de características mediante la transformación ICA es posible disminuir la complejidad del clasificador [6]. La desventaja de ICA radica en que la información de las expresiones representada en términos de AUs puede ser de naturaleza lineal o no lineal, por lo tanto un algoritmo ICA convencional que se basa en proyecciones lineales puede tener dificultad al tratar de representar combinaciones de AUs en donde la estructura de los datos es no lineal. Para sobrellevar este problema, en este trabajo se plantea una representación de las imágenes en un espacio característico KPCA donde se realiza un mapeo no lineal con métodos Kernel en busca de datos con estructura lineal [7, 8], de esta forma la tarea a realizar es aplicar ICA en el espacio transformado Kernel. Estudios como [9, 10] han combinado con éxito Kernel e ICA (KICA) para reconocer rostros, en este trabajo se pretende probar esta metodología para reconocer AUs.

Este documento se encuentra organizado así: inicialmente están los Antecedentes en donde se presenta un sistema de codificación de AUs y se hace una revisión de los estudios que las reconocen automáticamente, seguidamente esta el Marco Conceptual donde se exponen los métodos Kernels e ICA y una metodología que los combina, posteriormente hay un Marco Experimental donde se describe la base de datos y el experimento realizado y finalmente se discuten los resultados alcanzados, se presentan las conclusiones y trabajos futuros.

Antecedentes

En 1978 Ekman y Friesen diseñaron los FACS para caracterizar las expresiones faciales por medio de AUs. Este sistema está desarrollado basado en observaciones realizadas por expertos en el comportamiento humano, para explicar los cambios sutiles de las expresiones faciales. Cuenta con un total de 44 AUs. 30 de ellas se encuentran relacionados a la contracción de músculos faciales incluyendo 12 para la parte

inferior del rostro y 18 para la parte superior. Mientras las AUs restantes se deben a acciones combinadas. A partir de estos estudios y del avance de los computadores, investigadores en el campo de la visión por computador como [5, 6, 8, 11- 13], a finales de los 90 empezaron a crear métodos para su reconocimiento automático sobre secuencias de imágenes.

Entre los primeros trabajos, están [11, 12] donde desarrollaron sistemas de visión artificial sensibles a cambios sutiles en el rostro, para extraer automáticamente características y codificar una secuencia de imágenes en FACS. Realizan un análisis de las expresiones faciales de acuerdo a las regiones superior e inferior del rostro, la metodología con la que alcanzaron un mejor desempeño, realiza alineamiento de la imagen, extracción de características mediante seguimiento de flujo más PCA y clasificación con modelos ocultos de Markov (HMM), donde para cada AU se construye un HMM, alcanzando un acierto del 92% para el reconocimiento de ambas regiones.

Combinando tres tipos de análisis: espacial holístico, espacial localizado y movimientos faciales, en se [3] planteó un sistema para reconocer AUs de la parte superior del rostro, alcanzando un acierto del 91% para clasificar 6 AUs, Posteriormente en [13], se mejoró los porcentajes de acierto y extendió sus estudios a la clasificación de AUs de la parte inferior del rostro, alcanzando un acierto del 96% para el reconocimiento de mayor cantidad de AUs (12 tanto para la parte inferior del rostro como para la superior).

Un estudio bastante relacionado con [3] y [13] es el realizado en [4], donde se creó un sistema para el reconocimiento de AUs y se comparan diferentes técnicas para el análisis de expresiones faciales con la clasificación dada por expertos. Los mejores resultados se obtuvieron utilizando en forma separada dos representaciones una con wavelets y otra con ICA, alcanzando un 95.5% de acierto en la clasificación de 20 AUs tanto de la parte superior como inferior del rostro.

Otro tipo de sistema es el desarrollado en [1], donde se realiza un modelamiento multiestado de características faciales como: boca, ojos, cejas y mejillas; se extrae una descripción paramétrica de dichas características y se realiza un seguimiento de estos parámetros para introducirlos a una red neuronal, el reconocimiento promedio alcanzado fue del 96.4% y del 96.7% para AUs superiores e inferiores respectivamente.

En [14] se trabaja en la identificación de mayor cantidad de AUs (18), alcanzando un reconocimiento promedio del 94.5% sobre secuencias de imágenes de la base de datos C-K, extrayendo características con filtros Gabor y clasificando con máquinas de vectores de soporte (SVM). También utilizando Gabor, pero con Kernels para la caracterización en [15] se propone un sistema para el reconocimiento de AUs usando como clasificador redes neuronales. Obteniendo un reconocimiento promedio de 86.6% para 7 AUs de la parte superior del rostro y 81.6% para 7 AUs de la parte inferior, sobre la base de datos C-K.

Utilizando modelamiento facial al igual que en [1], en [16], se realiza el reconocimiento de AUs construyendo modelos temporales en 2D (imágenes que muestran cuando y donde ocurre el movimiento en una secuencia) basándose en el concepto de historia del movimiento en la imagen, el reconocimiento se completa combinando el clasificador del vecino más cercano con un sistema basado en reglas, el sistema desarrollado alcanza un 76.2% de acierto para reconocer 9 AUs de la parte inferior del rostro y 6 AUs de la parte superior, sobre la base de datos C-K. Para mejorar estos porcentajes en [17], se incorpora un esquema de factorización de vecindades al seguimiento con filtrado de partículas y un estudio de modelos rígidos y morfológicos y realizan la clasificación con SVM, alcanzando un 84.0% de acierto en el reconocimiento.

Posteriormente en [18], se incorpora al estudio [17], características espacio temporales extraídas de una representación wavelet y mejoran el

desempeño del clasificador con la técnica de AdaBoost, subiendo los porcentajes a un 90.2% de acierto para el reconocimiento de 15 AUs sobre la base de datos C-K.

Uno de los trabajos más recientes y que mejores resultados registran, es [19], donde se plantea que la combinación de componentes faciales puede producir mejores resultados que componentes individuales en el reconocimiento de expresiones faciales y adicionalmente realizan una comparación de sistemas entre los que se encuentran [1,4,15] y proponen un sistema que usa ecualización del histograma como pre-procesamiento, ICA para extracción de características y clasificación con SVM, alcanzando un 97% de acierto tanto para AUs superiores como inferiores y porcentajes de acierto más altos para el reconocimiento de AUs del rostro entero, sobre la base de datos C-K.

A partir de la revisión de los trabajos previos se observa que la información holística, juega un papel importante en el procesamiento de rostros y en el reconocimiento de expresiones faciales. Tal es así que trabajos que usan información global del rostro combinando AUs superiores e inferiores como [19] han mostrado ser más efectivos que los estudios que usan regiones localizadas como en el caso de [14,18] y los basados en características específicas como en [1].

Marco conceptual

Antes de plantear una metodología KICA se definirán dos conceptos de interés, los Kernels e ICA: Considere un grupo de datos $x \in R^N$, un Kernel k se presenta como una medida de similitud que puede ser vista como un producto punto del mapeo ϕ en un espacio característico H , como se presenta en la Ecuación (1):

$$\begin{aligned} \phi: R^N &\rightarrow H \\ x &\rightarrow \phi(x) \end{aligned} \quad (1)$$

Donde es posible definir el Kernel mediante la Ecuación (2) como la función equivalente a:

$$k(x, x') = \langle \phi(x), \phi(x') \rangle \quad (2)$$

Considere que se tiene $x_1, x_2, \dots, x_m \in X$, la matriz de Gram ó matriz Kernel de $m \times m$ elementos se define por la Ecuación (3), así:

$$K_{ij} = k(x_i, x_j) \quad (3)$$

La cual es definida positiva y simétrica. Así cualquier función que cumpla esta condición y pueda representarse por un producto punto, puede ser utilizada como Kernel [4], donde no es necesario conocer el mapeo para obtener la representación de los datos en el espacio característico. Entre las principales funciones Kernels se tienen, el Kernel Gaussiano y el Kernel polinomial no homogéneo, estas se explican en detalle en [20].

El problema que intenta solucionar ICA formulado en notación matricial en la Ecuación (4) consiste en recuperar un vector de fuentes S , disponiendo, únicamente, de un vector de observaciones X :

$$S = WX \quad (4)$$

Por lo tanto el primer objetivo es calcular una matriz de pesos W que representa la inversa de una matriz de mezcla A (desconocida) con algunos cambios en escala y permutaciones.

Entre los primeros algoritmos para aproximar ICA se encuentra el *Infomax* [21]. Aunque es comúnmente usado, no fuerza una decorrelación de los datos. Con el fin de cumplir la restricción de ortogonalidad en forma exacta en [22], se introdujo el algoritmo *FastICA*, con el que según [23], es posible realizar una transformación ortonormal de los datos.

El algoritmo *FastICA* usa una nueva función objetivo para estimar las ICs, a diferencia de la aproximación realizada con *Infomax* donde se maximiza la entropía, este algoritmo usa la negentropía, para maximizar la independencia [23], es decir los pesos que se usan para obtener

las ICs se encuentran maximizando la Ecuación (5):

$$J_G(\mathbf{W}) = \left[E(E(\mathbf{W}^T \mathbf{X})) - E(G(v)) \right]^2 \quad (5)$$

donde G es una función objetivo unitaria no cuadrática que en la práctica es elegida de forma que sea rápida de calcular y que considere el orden en que las componentes son estimadas, v es una variable Gaussiana aleatoria que se asumen de media cero y varianza unitaria y W es un vector de pesos restringido a que $E((W^T X)^2) = 1$.

Para calcular todo el arreglo W se debe extender el uso de la función G . Para hacer esto la información mutua es minimizada cuando la suma de la negentropía J_G es maximizada, además se considera la restricción de no correlación para obtener el problema:

$$\begin{aligned} \max \sum_{i=1}^n J_G(w_i) \\ E\left(\left(w_k^T X\right)\left(w_j^T X\right)\right) = \delta_{kj} \end{aligned} \quad (6)$$

Maximizando la negentropía como se muestra en la Ecuación (6) se obtiene cada uno de los vectores w_i que conforma las filas de la matriz W , con la cual se calcula la transformación ICA. Esta maximización es llevada a cabo usando el algoritmo iterativo de punto fijo, el cual tiene propiedades de convergencia deseables y encuentra ICs de prácticamente cualquier distribución no Gaussiana [22].

Como se ha expuesto ICA encuentra una matriz de pesos mediante el uso de una función objetivo, lo que no se ha dicho es que se debe incluir un paso previo al aprendizaje conocido como blanqueamiento con el fin de reducir las estadísticas de primer y segundo orden de los datos. En este paso la media de las filas de X son sustraídas y seguidamente X se pasa a través de una matriz de blanqueamiento dada por la Ecuación (7):

$$W_z = 2(\text{Cov}(X))^{-\frac{1}{2}} = 2(XX^T)^{-\frac{1}{2}} \quad (7)$$

Cuando las entradas a ICA son los datos blanqueados, la matriz de pesos es $W_i = WW_z$.

KICA

La representación ICA dada por la Ecuación (4) es intrínsecamente lineal por lo que puede fallar para separar datos con estructura no lineal. Con el fin de sobrellevar estas limitaciones se plantea una metodología que considere la no linealidad de los datos, para esto se combina métodos Kernel e ICA, se asume un vector aleatorio x , el cual es posiblemente una mezcla no lineal, se realiza un mapeo no lineal a un espacio característico de mayor dimensión H , a partir de la Ecuación (1). Se realiza la transformación ICA en el espacio característico mediante el operador lineal matricial W^ϕ , por lo que la transformación definida en la Ecuación (4), queda de la forma mostrada en la Ecuación (8):

$$s = W^\phi \phi(x) \quad (8)$$

En la metodología KICA, también es necesario realizar una etapa previa de blanqueamiento de los datos, pero a diferencia de ICA esta no se calcula con la matriz de covarianza si no usando métodos Kernel, para esto se asume que se tiene un conjunto de datos $x_1, x_2, \dots, x_m \in R^N$ y se realiza su mapeo al espacio característico donde se debe garantizar, la condición establecida en la Ecuación (9):

$$\sum_{j=1}^M \phi(x_j) = 0 \quad (9)$$

Sea $Q = [\phi(x_1), \phi(x_2), \dots, \phi(x_M)]$ la matriz de covarianza en el espacio característico H se define por medio de la Ecuación (10), según [24]:

$$s_i^\phi = \frac{1}{M} QQ^T \quad (10)$$

La matriz Kernel se determina por medio de la Ecuación (11):

$$K = Q^T Q \quad (11)$$

En virtud de los métodos Kernel es posible encontrar cada uno de los componentes de K sin necesidad de determinar el operador $\phi(\cdot)$, de la forma:

$$K_{ij} = \phi(x_i)^T \phi(x_j) = k(x_i, x_j) \quad (12)$$

A partir de los valores propios λ_i y vectores propios γ_i de la matriz K se define:

$$V = (\gamma_1, \gamma_2, \dots, \gamma_m) \quad (13)$$

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$$

Con los vectores de la Ecuación (13) es posible realizar el blanqueamiento de los datos, así:

$$y = P^T \phi(x) \quad (14)$$

Puesto que el mapeo se realiza implícitamente la Ecuación (14) se calcula como:

$$\begin{aligned} y &= \sqrt{M} \Lambda^{-1} V^T Q^T \phi(x) \\ &= \sqrt{M} \Lambda^{-1} V^T [k(x_1, x), k(x_2, x), \dots, k(x_m, x)]^T \\ &= \sqrt{M} \Lambda^{-1} V^T K_x \end{aligned} \quad (15)$$

Para que la estimación de ICA sea más simple y mejor condicionada es necesario que el vector Kernel K_x tenga sus datos centrados, si $C_M = (1/M)_{M \times M}$ y $C_1 = (1/M)_{M \times 1}$ se tiene:

$$\tilde{K}_x = K_x - C_M K_x - K C_1 + C_M K C_1 \quad (16)$$

Por lo tanto la Ecuación (15) queda de la forma mostrada en la Ecuación (17):

$$y = \sqrt{M} \Lambda^{-1} V^T \tilde{K}_x \quad (17)$$

En resumen la matriz de pesos W que se usa para obtener la transformación ICA en el espacio característico (KPCA blanqueado), se obtiene a partir de los datos que han pasado por dos etapas: un blanqueamiento, dado por la Ecuación (14) más un centrado dado por la Ecuación (16). Por lo que finalmente la Ecuación (8) queda de la forma [10]:

$$s = Wy \quad (18)$$

Experimentación

Base de datos

La base de datos utilizada en el experimento es la C-K [2], la cual es representativa y ofrece un banco de pruebas para realizar estudios comparativos de expresiones faciales. Consta de 500 secuencias de imágenes de 100 sujetos en un rango de edades entre 18 y 30 años. El 69% son mujeres y el 31% son hombres. De los cuales el 81% son Euro-Americanos, 13% son afro-americanos y el 6% pertenece a otros grupos como asiáticos o latinos. Las imágenes disponibles fueron tomadas de una cámara localizada directamente frente al sujeto. Los sujetos fueron instruidos experimentalmente para manifestar una serie de 23 distintas AUs que incluyen AUs simples y combinadas. Las imágenes fueron digitalizadas en escala de grises a 640*480 píxeles con 8 bits de precisión.

La tarea de reconocimiento de AUs puede ser llevada a cabo mediante técnicas de visión artificial prácticamente en cuatro etapas: preprocesamiento, segmentación o detección del rostro, extracción de características y clasificación. Estas se describen a continuación:

Preprocesamiento

Inicialmente la base de datos fue acondicionada tal como en [19], mediante ecualización adaptativa del histograma de contraste limitado, usando una distribución del histograma uniforme, con el fin de homogenizar las condiciones de iluminación. Esta técnica considera una ventana local para cada

píxel y calcula el valor de la nueva intensidad basado en el histograma local definido para cada ventana. Y una vez desarrollada la ecualización se combinan las vecindades de las ventanas usando interpolación bilineal para reducir las fronteras que se inducen.

Segmentación

Se realizó detección del rostro siguiendo la metodología expuesta en [25], esta se basa en la extracción de características de múltiples escalas con filtros de base Haar, los cuales se aplican a través de la convolución a una imagen transformada a la que se le conoce como imagen integral, que es construida a partir de operaciones básicas, por último se utilizan clasificadores en cascada mediante la técnica de *boosting* que combina varios clasificadores básicos para formar un único clasificador más complejo y preciso. Con esta etapa se consigue limitar el área de trabajo tan solo a la región de interés.

Extracción de características

El conjunto de imágenes preprocesadas y segmentadas obtenidas a partir de los procedimientos anteriormente descritos se utilizan como entrada a la etapa de caracterización en donde se combinan métodos Kernel e ICA, presentada en el marco conceptual como KICA. En esta etapa los valores de intensidad de las imágenes son mapeados mediante una función Kernel de la forma de la Ecuación (12). Seguidamente los datos son blanqueados y centrados usando las Ecuaciones (14) y (16) respectivamente. Finalmente se obtienen los vectores de características resultantes a través de la Ecuación (18), los cuales son utilizados en el clasificador final.

Clasificación

Como algoritmo de decisión se emplea un clasificador sencillo con el fin de dar mayor importancia a las características, en particular se usa un clasificador estadístico basado en el

vecino más cercano el cual visto de un modo práctico encuentra los k patrones del conjunto de entrenamiento más próximos al patrón observación con una métrica dada (para el caso de este estudio la distancia Euclidiana), anota las clases a las que pertenecen dichos patrones y decide por votación mayoritaria entre las clases de los k patrones.

Experimento realizado

De acuerdo a [19] los mejores porcentajes de reconocimiento de AUs se obtienen usando AUs combinadas de las regiones superiores e inferiores del rostro, en el experimento realizado usaron las siguientes 4 combinaciones de AUs de todo el rostro: Neutral, AU4 + AU17, AU6 + AU12 + AU25, AU1 + AU2 + AU5 + AU25 + AU27. En este trabajo se clasificarán estas mismas 4 combinaciones de AUs usando la misma base de datos con el fin de observar si con la metodología que se plantea en este estudio, la cual adiciona métodos Kernel es posible alcanzar resultados similares o superiores.

En el experimento realizado para reconocer AUs se seleccionó en forma aleatoria un conjunto de 26 sujetos de la base de datos C-K, compuesta por 18 mujeres y 8 hombres. Se tomaron las 3 primeras imágenes de cada secuencia como neutral y las últimas 4 como la combinación de AUs correspondientes. Se repite el experimento 10 veces seleccionando el 90% y 10% de los conjuntos para validación y entrenamiento respectivamente.

La aplicación de ICA a imágenes se realizó de la forma convencional, como se expone en [5], donde se plantea que el algoritmo puede usarse a través de dos arquitecturas: tratando las imágenes como variables aleatorias y los píxeles como observaciones (Arq I) o tratando los píxeles como variables aleatorias y las imágenes como observaciones (Arq II). Otra forma de realizar la aplicación del algoritmo ICA a imágenes es primero aplicándolo por filas y luego por


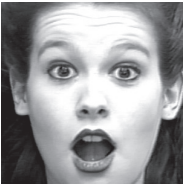


columnas tal como se planteo recientemente en [26].

En la metodología KICA expuesta en este trabajo se uso un Kernel polinomial no homogéneo de la forma: $k(\mathbf{x}, \mathbf{x}') = (a\langle \mathbf{x}, \mathbf{x}' \rangle + c)^d$ donde $d \in \mathbb{N}$, $a, c \geq 0$. Con $d = 3$ y se seleccionó sistemáticamente un vector de constantes $c = [1e - 3 \ 1 \ 3]$. Se utilizó una función objetivo $G = X^3$ y en el algoritmo *FastICA* se calcularon las ICs una por una.

Resultados y Discusión

Como se observa en la tabla 1, el algoritmo KICA utilizando la arquitectura II reconoce mejor las dos primeras clases, mientras que con el algoritmo ICA en la arquitectura I se obtiene mejores resultados para las dos clases restantes. Esto se debe a que con KICA se reconocieron principalmente cambios de forma o transiciones fuertes entre AUs que ocurren de forma instantánea como lo son las combinaciones AU4 +AU17 (que se asocian con la emoción de tristeza) y las AU1+AU2+AU5+AU25+AU27 (que se asocian con la emoción de sorpresa). Por el contrario al utilizar ICA se reconocen satisfactoriamente AUs que se presentan como resultado de transiciones suaves o estados emocionales en donde no se presentan cambios de forma considerables. Esto se relaciona directamente con el espacio mapeado sobre el cual se aplica el algoritmo *FastICA* para hacer que los datos proyectados sean lo menos Gaussiano posible, es decir si *FastICA* se aplica sobre datos decorrelacionados mediante un mapeo lineal utilizando por ejemplo PCA, como es el caso de la transformación ICA tradicional, se reconoce mejor combinaciones de AUs con cambios suaves o si *FastICA* se aplica sobre datos en un espacio de mucho mayor dimensión mapeados en forma no lineal con KPCA, como es el caso de la transformación KICA, se reconoce mejor combinaciones de AUs con cambios fuertes. Visto en forma general con la metodología KICA se alcanzó un reconocimiento promedio mayor para todas las clases.

Tabla 1 Resultados en el reconocimiento de AUS

<i>Patrones</i>	<i>Imágenes</i>	<i>Número de muestras</i>	<i>FastICA Arq I</i>	<i>KICA Arq II</i>
AU4 + AU17		16	90,00%	99,99%
AU1+AU2+AU5+AU25+AU27		72	98,52%	99,99%
Neutral		78	99,99%	98,57%
AU6+AU12+AU25		56	94,00%	88,00%
Total		222	95,64±1,14%	96,64±0,54%

Conclusiones

Se desarrolló una metodología efectiva para el reconocimiento de AUS utilizando ICA y KICA, encontrando que para esta tarea el algoritmo KICA aplicado sobre la arquitectura II con un Kernel polinomial entregó los mejores resultados, alcanzando un reconocimiento promedio de $96,64\% \pm 0,54$ para tres combinaciones de AUS más rostros neutrales.

Los resultados encontrados sugieren que el desempeño de la metodología implementada esta relacionado al tipo de arquitectura sobre el cual se aplican los métodos a las imágenes para la tarea de reconocimiento de AUS. Para el caso de la metodología basada en KICA, que fue con la

que se alcanzaron mejores resultados en general, se obtuvo mejor desempeño con la arquitectura II que entrega características globales, mientras que la metodología basada en ICA funcionó mejor sobre la arquitectura I que produce características localizadas. Esto refuerza el hecho que el reconocimiento de expresiones puede realizarse de manera más exacta si se considera información global del rostro y no información localizada de determinadas regiones de la imagen.

En la metodología de caracterización expuesta en este documento se representa un arreglo de imágenes el cual es una mezcla no lineal en términos de únicamente sus ICs, realizando primero un mapeo no lineal con un Kernel y posteriormente aplicando la transformación

ICA, obteniendo así un espacio característico de tamaño reducido donde solo permanecen las variables más discriminantes haciendo que este espacio sea lo más linealmente separable posible, razón por la cual tan solo se requiere de clasificadores básicos por ejemplo los basados en distancias, para reconocer las AUs.

En este estudio las características que se utilizaron para el reconocimiento de AUs provienen directamente de las representaciones ICA y KICA, como trabajo futuro se plantea aplicar métodos de selección de características que permitan conservar un menor conjunto de características y descartar información irrelevante, con el fin de tener mayor eficiencia computacional y mejorar la exactitud en la clasificación.

Referencias

1. Y. Tian, T. Kanade, J. Cohn. "Recognizing Action Units for Facial Expression Analysis". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 23. 2001. pp. 97-115.
2. T. Kanade, J. Cohn, Y. Tian. "Comprehensive database for facial expression analysis". *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition*. Grenoble (France). 2000. pp. 46-53.
3. M. S. Bartlett. *Face image analysis by unsupervised learning and redundancy reduction*. Ph.D. dissertation. Univ. California. San Diego. 1998. pp. 27-92.
4. G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, T. J. Sejnowski. "Classifying facial actions". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 21. 1999. pp. 974-989.
5. A. R. Webb. *Statistical Pattern Recognition*. 2ª ed. Ed. John Wiley & Sons. Indianapolis. USA. 2002. pp. 305-318.
6. R. O. Duda, M. E. Hart, D. G. Stork. *Pattern Classification*. 2ª ed. Ed. Wiley Interscience. Hoboken. USA. 2000. pp. 20-83.
7. B. Schölkopf, A. J. Smola. *Learning with Kernels*. MA. Ed. MIT Press. Cambridge. 2002. pp. 25-55.
8. T. Martirriggiano, M. Leo, P. Spagnolo, Dapos, T. Orazio. "Facial feature extraction by kernel independent component analysis". *AVBPS Advanced Video and Signal Based Surveillance*. IEEE Computer Society. Los Alamitos (CA). 2005. pp. 270-275.
9. T. Martirriggiano, M. Leo, T. D'Orazio, A. Distant. "Face Recognition by Kernel Independent Component Analysis". *IEA/AIE'2005: Proceedings of the 18th international conference on Innovations in Applied Artificial Intelligence*. Ed. Springer-Verlag. London. 2005. pp. 55-58.
10. X. Yang, X. Gao, D. Zhang, J. Yang. "Kernel ica: An alternative formulation and its application to face recognition". *Pattern Recognition*. Vol. 38. 2005. pp. 1784-1787.
11. J. J. Lien, T. Kanade, A. Zlochow, J. F. Cohn, C.-C. Li. "Automatically recognizing facial expressions in the spatio-temporal domain". *Workshop on Perceptual User Interfaces*. Vol. 36. 1997. pp. 94-7.
12. J. J. Lien, T. Kanade, J. F. Cohn, C.C. Li. "Automated facial expression recognition based on face action units". *FG '98: Proceedings of the 3rd. International Conference on Face & Gesture Recognition*. Nara (Japan). Vol. 14. 1998. pp. 390-395.
13. M. S. Bartlett, J. C. Hager, P. Ekman, T. J. Sejnowski. "Measuring facial expressions by computer image analysis". Research supported by NSF Grant No BS-9120868, Lawrence Livermore National Laboratories Intra-University Agreement B291436 and Howard Hughes Medical Institute. San Diego (CA). 1999. pp. 1-25.
14. M. S. Bartlett, G. Littlewort, C. Lainscsek, I. Fasel, J. Movellan. Machine learning methods for fully automatic recognition of facial expressions and facial actions. *Proc. IEEE Int'l Conf. Systems, Man and Cybernetics*. Hague (Netherlands). 2004. pp. 592-597.
15. J. Bazzo, M. Lamar. "Recognizing facial actions using gabor wavelets with neutral face average difference". *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*. Seoul. Korea. 2004. pp. 505-510.
16. M. Valstar, M. Pantic, I. Patras. "Motion history for facial action detection in video". *IEEE International Conference on Systems, Man and Cybernetics*. Delft (Netherlands). 2004. pp. 635-640.
17. M. Valstar, I. Patras, M. Pantic. "Facial Action Unit Recognition using Temporal Templates". *IEEE Int'l Workshop on Human-Robot Interaction*. Kurashiki Okayama (Japan). 2004. pp. 253-258.
18. M. Valstar, M. Pantic. "Fully automatic facial action unit detection and temporal analysis". *IEEE Int'l Conf.*

- on *Computer Vision and Pattern Recognition*. New York. Vol. 3. 2006. pp.1-8.
19. C. F. Chuang, F. Y. Shih. "Rapid and brief communication: Recognizing facial action units using independent component analysis and support vector machine". *Pattern Recognition*. Vol. 39. 2006. pp. 1795-1798.
 20. C. Campbell. "Kernel methods: a survey of current techniques". *Neurocomputing*. Vol. 48. 2000. pp. 63-84.
 21. A. J. Bell, T. J. Sejnowski. "An information maximization approach to blind separation and blind deconvolution". *Neural Computation*. Vol. 7. 1995. pp. 1129-1159.
 22. A. Hyvärinen. "Fast and robust fixed-point algorithms for independent component analysis". *IEEE Transactions on Neural Networks*. Vol. 10. 1999. pp. 626-634.
 23. M. A. Vicente, P. O. Hoyer, A. Hyvarinen. "Equivalence of some common linear feature extraction techniques for appearance-based object recognition tasks". *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 29. 2007. pp. 896-900.
 24. Schölkopf, A. Smola, K.R. Müller. "Nonlinear component analysis as a Kernel eigenvalue problem". *Neural Computation*. Vol. 10. 1998. pp. 1299-1319.
 25. M. L. Guevara, J. D. Echeverry, W. A. Uruña. "Detección de rostros en imágenes digitales usando clasificadores en cascada". *Scientia et Technica*. Vol. 38. 2008. pp. 1-6.
 26. Q. Gao, L. Zhang, D. Zhang. "Sequential row-column independent component analysis for face recognition". *Neurocomputing*. Vol. 72. 2009. pp. 1152-1159.