



Encuentro Internacional de
Educación en Ingeniería ACOFI

**GESTIÓN, CALIDAD Y DESARROLLO
EN LAS FACULTADES DE INGENIERÍA**

**CARTAGENA, COLOMBIA
18 al 21 de septiembre de 2018**



IDENTIFICACIÓN DE EMOCIONES A PARTIR DE GESTOS

Dayana Lucía Verdugo Alejo, Juan Manuel López López

**Universidad del Rosario
Bogotá, Colombia**

Resumen

Las emociones se pueden entender, de manera muy simple, como la forma en que las personas responden a estímulos externos o internos. Este tipo de respuestas pueden ser psicológicas y fisiológicas, y representan un mecanismo evolutivo de protección y comunicación. Para la mayoría de personas, la identificación de las emociones a partir del gesto facial, se realiza de manera casi inconsciente: la interacción se modula dependiendo del estado emocional percibido. Es diferente la interacción con una persona percibida como enfadada, que con una persona percibida como feliz.

El desarrollo de este proyecto se enfoca en la identificación automática de estados emocionales, en tiempo real, a partir de imágenes de rostros. Particularmente, se analizó una de las tres dimensiones en que se puede entender una emoción, llamada valencia. Se utilizaron herramientas de mejoramiento y análisis de imágenes en conjunto con técnicas de aprendizaje automático para obtener, a partir de una base de datos, un modelo de identificación de valencia emocional. Luego, se utilizó el modelo creado para realizar pruebas en tiempo real.

Los resultados son prometedores en cuanto que el esquema de procesamiento y generación del modelo indican la factibilidad de identificar emociones, a partir de gestos faciales en tiempo real. Este proyecto tiene una segunda etapa que está en desarrollo, la cual utiliza esta herramienta para la estimulación cognitiva en participantes con espectro autista, de la Clínica Howard Gardner, en Bogotá.

Palabras clave: identificación de emociones; gestos faciales; aprendizaje automático

Abstract

Emotions may be described, in a very simplistic manner, as the response to external or internal stimuli. These responses include the psychological and physiological aspects and represent an evolutionary mechanism for protection and communication. For most of the people, the identification of emotions from the facial gesture, is an unconscious process: interactions are modulated depending on the perceived emotional state. The interaction is different when one of the interlocutors is perceived as angry than when the perceived emotion is happiness.

This project aims to the automatic identification of emotional states, in real time, from face images. Particularly, one of the three emotional dimensions, called valence, was analyzed. Image enhancing and analysis tools were used along with machine learning techniques to generate, from a dataset of images, a model to identify the emotional valence. Then, the model was assessed in real time.

Results are promising. Processing and model generation schemes indicate the feasibility of the emotional identification, from facial gestures, in real time. This project has a second stage, currently in progress, in which the tool for emotional identification is used in cognitive stimulation protocols of autism-spectrum participants from the Howard Garnder Clinic, in Bogotá.

Keywords: *emotion identification; facial gestures; machine learning*

1. Introducción

Las emociones son una característica muy importante de la vida humana ya que permiten mediar y facilitar las interacciones entre seres humanos y el entorno. Las emociones se pueden estudiar a través de tres dimensiones:

- Valencia. Indica qué tan positiva o negativa es la emoción.
- Activación. Da cuenta de la pasividad o el nivel de actividad que esa emoción requiere.
- Dominancia. Hace referencia a qué tanto control se tiene sobre la emoción.

Una forma de estudiar las emociones es a través del análisis de los gestos faciales asociados. Hoy en día existe gran variedad de técnicas de procesamiento digital de imágenes gracias al auge de las cámaras digitales (como las incorporadas en los dispositivos móviles), al almacenamiento y transferencia en sistemas en la nube y redes sociales. Estas innovaciones han llevado a que la investigación sobre las emociones se centre en técnicas automáticas para identificar y clasificar las mismas a partir de imágenes.

En esta investigación se busca identificar en tiempo real las emociones empleando gestos faciales para una aplicación clínica. Se encuentra dividido en dos etapas. La primera etapa consistió en desarrollar un modelo de identificación de valencia emocional empleando una base de datos y haciendo pruebas en tiempo real con una cámara web. En este documento se explica detalladamente esta primera etapa. La segunda etapa, que aún está en desarrollo consiste en utilizar este sistema en un contexto de robótica social enfocado en protocolos de estimulación

cognitiva para pacientes con espectro autista (ASD por sus siglas en inglés) de la Clínica Howard Gardner, permitiendo un proceso de rehabilitación innovador en el área de la robótica social, siguiendo los lineamientos planteados en investigaciones previas como (Calvo et al., 2015; Palestra, Esposito, & De Carolis, 2017; Truschzinski & Müller, 2014). Este proyecto inició en el período 2017-2 y finalizará en diciembre del presente año.

Las áreas de aplicación las cuales tiene este proyecto abarcan: Componente para procedimientos de interacción humano-robot en entornos clínicos, análisis en tiempo real de la evolución de las expresiones y entretenimiento.

Componente para procedimientos de interacción humano-robot en entornos clínicos: Las interacciones humano-robot buscan, en ocasiones, asemejarse a las interacciones humano-humano. Debido a esto, una identificación automática de las emociones permitiría a la máquina acercarse a su objetivo. Diversas investigaciones han mostrado el impacto positivo que tiene la incorporación de robots en protocolos de estimulación y rehabilitación (Calvo et al., 2015; Palestra et al., 2017; Truschzinski & Müller, 2014).

Para el caso en particular de pacientes con ASD, debido a que estos presentan un alto nivel de dificultad para prestar atención y concentrarse en una actividad en específico (Soto, 2002), se ha evidenciado significativamente que estas interacciones son más efectivas en comparación con terapias tradicionales, (Cruz Ardila & Salazar, 2014; Dautenhahn & Billard, 2002; Robins, Dautenhahn, te Boekhorst, & Billard, 2004; Sato et al., 2017).

Análisis en tiempo real de la evolución de las expresiones: El seguimiento y el reconocimiento de las actividades faciales de imágenes o videos han atraído gran atención en el campo de la computación, lo que ha dado paso al desarrollo de nuevas técnicas de procesamiento digital de imágenes y de video, tales como redes neuronales convolucionales y otras técnicas de aprendizaje profundo (deep learning) como por ejemplo las desarrolladas por (Harandi, Ahmadabadi, & Araabi, 2004; Li & Yu, 2015). Dichas técnicas se han enfocado en los procesos de extracción automática de características, además de la disminución del costo computacional para la identificación y análisis de rostros de personas. Esto lleva a analizar desde puntos de características faciales alrededor de cada componente facial (ceja, boca, nariz, ojos, entre otros) hasta diseñar y simular prototipos de expresiones faciales con el objetivo de representar el movimiento global del músculo facial (Tavakoli, Yanulevskaya, Rahtu, Heikkila, & Sebe, 2014; Yongqiang Li, Shangfei Wang, Yongping Zhao, & Qiang Ji, 2013).

Entretenimiento: En este campo, sistemas de reconocimiento automático de emociones, se aplican en entornos de realidad virtual reaccionan a las emociones del participante, brindando una sensación superior de inmersión y naturalidad del ambiente virtual. Esto permite ampliar el campo de desarrollo, llegando a asociar las emociones del participante con los eventos del entorno (Hazlett, 2006).

2. Métodos y Materiales

La metodología de trabajo para esta investigación parte de un estudio piloto (Primera etapa) para la validación del modelo de detección de emociones y posteriormente (Segunda etapa) su implementación y evaluación en un protocolo de estimulación cognitiva, en una población con ASD perteneciente a la Clínica Howard Gardner.

2.1. Primera Etapa

- a) **Base de datos:** Para la generación del modelo de identificación de emociones, se partió de la base de datos Averaged Karolinska Directes Emotional Database (AKDEF). Esta presenta un total de 4900 imágenes, las cuales están conformadas por 70 individuos, cada uno mostrando 7 expresiones emocionales diferentes. Cada expresión se encuentra fotografiada dos veces, desde 5 ángulos diferentes (Goeleven, De Raedt, Leyman, & Verschuere, 2008).
- b) **Pre-procesamiento de la base de datos:** En primera instancia, se procede a convertir el esquema de codificación de la imagen a escala de grises. Posteriormente se detecta la región en donde se encuentra el rostro en la imagen. La detección del rostro en las imágenes se basó en el algoritmo de Viola-Jones (Viola & Jones, 2001), el cual consiste en mover una ventana de detección alrededor de las imágenes con el fin de determinar si existe o no un rostro en cada ubicación.
- c) **Extracción de características:** Una vez pre-procesadas las imágenes, de las regiones donde se encuentran los rostros, se hace un proceso de extracción de características, enfocado en aquellas que dan cuenta de los gestos faciales asociados a las emociones (Majumder, Behera, & Subramanian, 2014; Niese, Al-Hamadi, Panning, & Michaelis, 2010; Rathee & Ganotra, 2017). En la tabla 1 se muestran dichas características.

Tabla 1
Características A Considerar

Medida	Característica
Longitud en pixeles	Ancho de los ojos
	Alto de los ojos
	Largo de la Boca
	Medida de la distancia de la ceja al ojo
	Ancho de la nariz
Normalizada (números enteros)	Forma de la cara (Circular, cuadrada, ovalada, diamante, rectangular y triangular)

- d) **Generación del modelo:** El modelo estadístico para la determinación de los gestos fue Máquina de soporte vectorial (SVM por sus siglas en inglés). Se seleccionó esta SVM debido a su robustez para realizar tareas similares, tales como las planteadas por (Li & Yu, 2015).

En la figura 1 se puede observar un ejemplo de una de las características: La región del ojo



Figura 1. Ejemplo de característica basada en longitudes medidas en píxeles.

Algunos ejemplos de características normalizadas a partir de la forma del rostro se pueden observar en la figura 2.

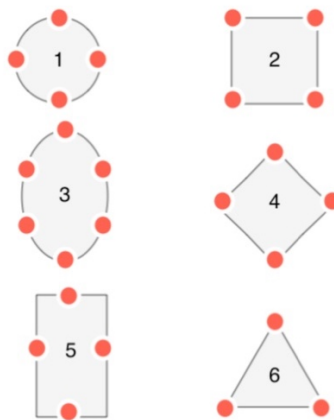


Figura 2. Ejemplo de característica basada en números enteros de la forma del rostro (Normalización).

Posteriormente se hizo uso de la *Statistics and Machine Learning Toolbox* de MATLAB® a través de la cual se generó un modelo entrenado utilizando las imágenes de la base de datos. Dicho modelo se entrenó para reconocer tres tipos de valencia: Positiva, negativa y neutra (Hazlett, 2006).

La SVM se generó a partir de dos pasos: Entrenar (Train) y Predecir (Predict). El primer paso se encarga de iterar varias veces un determinado proceso de aprendizaje que optimiza la separación de las muestras (80% de las muestras se utilizaron para este proceso de entrenamiento) entre los

diferentes tipos de valencia y el segundo paso se encarga de validar el modelo obtenido con muestras nuevas (20 % restante).

La evaluación de la efectividad del modelo se hizo a través de la sensibilidad y la especificidad. Estas métricas indican la capacidad del estimador para encontrar verdaderos positivos y verdaderos negativos respectivamente. Por otro lado, la precisión se encarga de medir la capacidad del modelo para predecir correctamente la categoría a la cual pertenece la muestra. Valores aceptables de estos parámetros son aquellos que están cercanos al 70 %.

El esquema de la generación del modelo se muestra en la fig. 3, mientras que en la fig. 4 se observa el proceso de predicción.

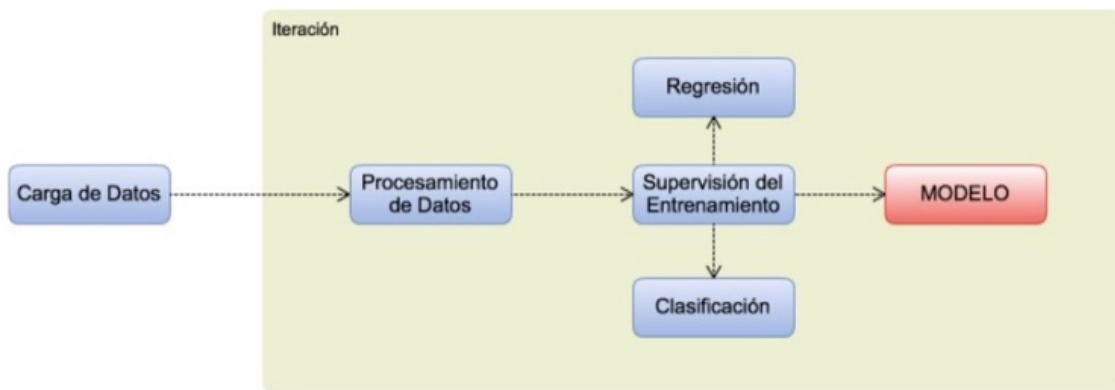


Figura 3. Proceso de entrenamiento.



Figura 4. Proceso de predicción.

- e) **Validación del modelo en tiempo real:** Finalmente se integra todo lo desarrollado previamente a una interfaz gráfica de usuario (GUI), empleando una cámara web con una resolución de 1280x720 pixeles para adquirir las imágenes. En la GUI se presenta un recuadro de las imágenes captadas en tiempo real en formato RGB y en la parte superior-derecha se observa dos pulsadores para regular el encendido y el apagado de la cámara. Además, un recuadro en la parte inferior muestra la realimentación del tipo de valencia presente a través de colores: Verde-positivo, azul-neutro y rojo-negativo.

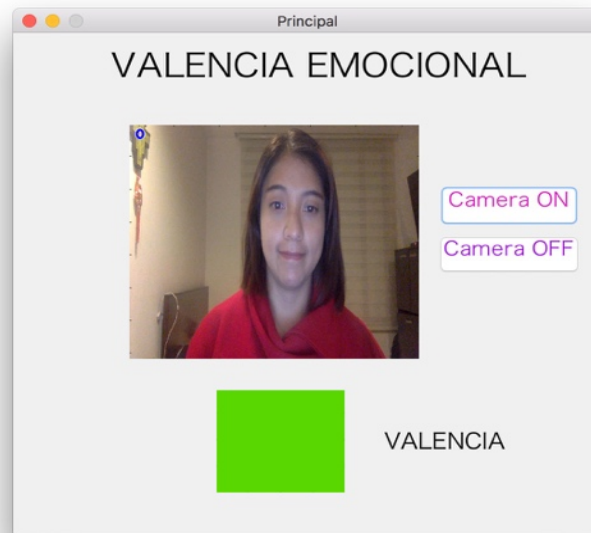


Figura 5. Interfaz gráfica de usuario (GUI).

2.2. Segunda etapa

Esta es una etapa en proceso de desarrollo. Actualmente se encuentra la definición del protocolo de estimulación, con base en la descripción del estado del arte, principalmente con material aplicado a nivel de Latinoamérica, teniendo mayor enfoque a nivel de Colombia, sumado a considerar las actividades y protocolos que ya se han llevado a cabo en la Clínica Howard Gardner para que de esta forma se pueda acoplar de manera efectiva y práctica el procedimiento a proponer.

3. Resultados Preliminares

El modelo generado (SVM) obtuvo una especificidad de 68,4% sensibilidad de 0.6 y precisión de 0.7 (Viola & Jones, 2001).

Por otro lado, se desea contribuir y dar elementos base para próximos proyectos e investigaciones que ayuden y complementen el tratamiento y el estudio en pacientes con ASD e incluso puedan suplementar proyectos cuya finalidad sea recreativa y didáctica. Sin olvidar que este proyecto contribuye a las investigaciones relacionadas con robótica social como se ha hecho con trabajos como (Matsumoto et al., 2016; Rosenblau, Kliemann, Dziobek, & Heekeren, 2016; Udayakumar, 2016), en donde se evidencian las ventajas de la inclusión de tecnología (principalmente robots) en poblaciones de niños con espectro autista, principalmente centrando los esfuerzos en tareas que involucran el contacto visual, la atención conjunta, la imitación y el reconocimiento y la producción y comprensión de emociones (Niese et al., 2010)

Finalmente, es de gran importancia que se continúe contribuyendo al estudio de temas como lo es el autismo por medio de la investigación para que de esta manera, sea posible innovar en diversas soluciones y ayudas para este tipo de población.

4. Referencias

Artículos de revistas

- E. Goeleven, R. De Raedt, L. Leyman, and B. Verschuere, "The Karolinska Directed Emotional Faces: A validation study," *Cognition & Emotion*, vol. 22, no. 6, pp. 1094–1118, Sep. 2008.
- X. Li and Q. Yu, "Face Recognition based on Deep Neural Network," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no. 10, pp. 29–38, Oct. 2015.
- Majumder, L. Behera, and V. K. Subramanian, "Emotion recognition from geometric facial features using self-organizing map," *Pattern Recognition*, vol. 47, no. 3, pp. 1282–1293, Mar. 2014.
- K. Matsumoto et al., "Behavioral Study on Emotional Voice Perception in Children with Autism Spectrum Disorder," *Journal of Pediatric Neuropsychology*, vol. 2, no. 3–4, pp. 108–118, Nov. 2016.
- R. Niese, A.-H. Al-Hamadi, A. P. Panning, and M. Michaelis, "Emotion Recognition based on 2D-3D Facial Feature Extraction from Color Image Sequences," *JOURNAL OF MULTIMEDIA*, vol. 5, no. 5, pp. 488–500, Oct. 2010.
- N. Rathee and D. Ganotra, "Modelling Facial Features for Emotion Recognition and Synthesis," *IETE Journal of Research*, vol. 63, no. 6, pp. 845–852, Nov. 2017.
- G. Rosenblau, D. Kliemann, I. Dziobek, and H. R. Heekeren, "Emotional prosody processing in Autism Spectrum Disorder," *Social Cognitive and Affective Neuroscience*, p. nsw118, Aug. 2016.
- W. Sato et al., "Impaired detection of happy facial expressions in autism," *Scientific Reports*, vol. 7, no. 1, Dec. 2017.
- R. Soto, "EL SÍNDROME AUTISTA: UN ACERCAMIENTO A SUS CARACTERÍSTICAS Y GENERALIDADES," *Revista Educación*, vol. 26, pp. 47–61, 2002.
- N. Udayakumar, "Facial Expression Recognition System for Autistic Children in Virtual Reality Environment," *International Journal of Scientific and Research Publications*, vol. 6, no. 6, pp. 613–622, Jun. 2016.
- Yongqiang Li, Shangfei Wang, Yongping Zhao, and Qiang Ji, "Simultaneous Facial Feature Tracking and Facial Expression Recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2559–2573, Jul. 2013.

Libros

- R. Calvo et al., "Emotion Modeling for Social Robots," in *The Oxford Handbook of Affective Computing*, R. Calvo, S. D'Mello, J. Gratch, and A. Kappas, Eds. Oxford University Press, 2015.

- K. Dautenhahn and A. Billard, "Games Children with Autism Can Play with Robota, a Humanoid Robotic Doll," in *Universal Access and Assistive Technology*, S. Keates, P. Langdon, P. J. Clarkson, and P. Robinson, Eds. London: Springer London, 2002, pp. 179–190.
- B. Robins, K. Dautenhahn, R. te Boekhorst, and A. Billard, "Effects of Repeated Exposure to a Humanoid Robot on Children with Autism," in *Designing a More Inclusive World*, S. Keates, J. Clarkson, P. Langdon, and P. Robinson, Eds. London: Springer London, 2004, pp. 225–236.

Memorias de congresos

- J. C. Cruz Ardila and Y. Salazar, "APLICACIÓN ROBÓTICA PARA REALIZAR TERAPIAS EN NIÑOS CON AUTISMO," presented at the Twelfth LACCEI Latin American and Caribbean Conference for Engineering and Technology, Guayaquil, Ecuador., 2014.
- M. T. Harandi, M. N. Ahmadabadi, and B. N. Araabi, "Face recognition using reinforcement learning," 2004, vol. 4, pp. 2709–2712.
- R. L. Hazlett, "Measuring emotional valence during interactive experiences: boys at video game play," 2006, p. 1023.
- G. Palestra, F. Esposito, and B. De Carolis, "A Multimodal Interface for Robot-Children Interaction in Autism Treatment," presented at the 12th biannual Conference of the Italian SIGCHI Chapter, Cagliari, Italy, 2017.
- H. R. Tavakoli, V. Yanulevskaya, E. Rahtu, J. Heikkila, and N. Sebe, "Emotional Valence Recognition, Analysis of Saliency and Eye Movements," 2014, pp. 4666–4671.
- M. Truschzinski and N. Müller, "An emotional model for social robots: late-breaking report," 2014, pp. 304–305.
- P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," 2001, vol. 1, pp. 1-511-1-518.

Fuentes electrónicas

- Fundación Avante, "Autismo Colombia." Consultado el 27 de abril de 2018 en <https://www.fundavante.org/autismo-en-colombia>,

Sobre los Autores

- **Dayana Lucía Verdugo Alejo:** Ingeniero Biomédico (e), Universidad del Rosario, Escuela Colombiana de Ingeniería Julio Garavito. dayana.verdugo@mail.escuelaing.edu.co
- **Juan Manuel López López:** Ingeniero Electrónico, Universidad Nacional de Colombia. Magister en Ingeniería. Área: Electrónica y de Computadores, Doctor en Ingeniería, Universidad de los Andes. Profesor Asistente, juan.lopezl@escuelaing.edu.co

Los puntos de vista expresados en este artículo no reflejan necesariamente la opinión de la Asociación Colombiana de Facultades de Ingeniería.

Copyright © 2018 Asociación Colombiana de Facultades de Ingeniería (ACOFI)