

APROXIMACIÓN AL RECONOCIMIENTO DE EMOCIONES FACIALES BASADO EN POSICIÓN DE PUNTOS DE INTERÉS

Víctor Manuel Álvarez Pato

Universidad Panamericana
valvarez@up.edu.mx

Ramiro Velázquez Guerrero

Universidad Panamericana
rvelazquez@up.edu.mx

Resumen

Con las técnicas actuales de reconocimiento facial, es posible descubrir automáticamente las emociones de una persona a través de una imagen de su rostro. Este estudio se vale de una aplicación en línea para detectar algunos puntos de interés en imágenes de rostros que expresan alguna emoción y compara sus posiciones con las de una expresión considerada neutral. Se busca establecer una relación entre el resultado obtenido y el propuesto por la herramienta FACS de Paul Ekman para determinar la viabilidad de un algoritmo de reconocimiento de emociones, así como posibles pautas para su desarrollo.

Palabras Claves: Face++, FACS, reconocimiento de emociones, reconocimiento facial.

Abstract

With the current facial recognition techniques, it is possible to automatically determine an individual's emotions through a digital image of his face. The present study employs an online API to detect certain landmarks in images of faces affected by some emotion and compares their positions with those of a neutral expression. We seek to establish a relationship between the obtained results and the one proposed by Paul Ekman's FACS tool to determine the viability of an

emotion recognition algorithm, as well as some possible guidelines for its development.

Keywords: *Emotion recognition, Face++, face recognition, FACS.*

1. Introducción

El rostro humano es capaz de comunicar grandes cantidades de información: nos permite distinguir a una persona de otras, conocer sus sentimientos e incluso prever sus reacciones inmediatas por medio de un rápido análisis facial.

La capacidad de procesar toda esta información ha sido exclusiva de algunos organismos vivos hasta hace relativamente poco tiempo. El trabajo de cientos de investigadores ha dado origen a algoritmos capaces de reconocer y aprovechar parte de esos datos presentes en el rostro [Zhao, 2003], [Liong, 2016], [Giannakakis, 2017], a pesar de que en muchas ocasiones sean tan sutiles que resultan difíciles de describir en el lenguaje común.

Actualmente, el desarrollo de la ciencia ha permitido que las computadoras puedan determinar si una imagen digital contiene o no un rostro humano, además de localizar su posición [Zhan, 2016], [Yan, 2014], [Wang, 2017]. Esto se conoce como detección facial y suele ser la primera fase en muchos algoritmos de análisis facial. También se ha conseguido identificar rostros de manera automática, lo cual resulta útil por ejemplo en sistemas de seguridad. La disciplina que engloba este tipo de algoritmos suele recibir el nombre de reconocimiento facial.

El reconocimiento de expresiones faciales, por otro lado, es una rama que estudia el modo de analizar y reconocer movimientos faciales y cambios en los rasgos de la cara a partir de información de tipo visual. Esto representa un apoyo fundamental para otro objetivo que requiere aún mayor conocimiento: el análisis de emociones [Sariyanidi, 2015], [Tarnowski, 2017].

El análisis automatizado de emociones permite tener una idea de los sentimientos que experimenta una persona, aun cuando ella misma no sea completamente consciente de ellos. Así, un video que capture las expresiones de un voluntario al probar un producto nuevo en un estudio mercadológico puede analizarse por medio de una computadora para conocer con mayor detalle sus reacciones,

mismas que podrían pasar inadvertidas si se le pidiera contestar una encuesta por escrito [Hamelin, 2017], [Yu, 2017]. Otras aplicaciones ya estudiadas incluyen diagnosticar desórdenes neuropsiquiátricos [Hamm, 2011], *E-learning* [Krithika, 2016] y monitorizar avances en terapias psiquiátricas, por citar algunos ejemplos. Este documento muestra una primera aproximación al reconocimiento automático de emociones en el rostro humano, con la intención de explorar su posible aplicación como apoyo en el diagnóstico de enfermedades.

Buena parte de los estudios realizados en esta materia, se basan los trabajos del psicólogo Paul Ekman [Ekman, 1971], quien sostiene que los principios de expresión facial son prácticamente innatos e independientes de la influencia cultural: esto permite crear un sistema de reconocimiento de emociones basado únicamente en la información visual del rostro.

Ekman también desarrolló el Sistema de Codificación de Acciones Faciales (FACS por sus siglas en inglés), que reduce cualquier emoción a una serie de movimientos aislados de los músculos faciales [Ekman, 1983]. Por ejemplo, una expresión de alegría en un rostro puede traducirse como una contracción de la porción orbitaria del músculo orbicular en combinación con el músculo cigomático mayor. Este tipo de movimientos cambian la posición relativa de algunos puntos de la cara –como la comisura de los labios, o los bordes de las cejas– que hoy en día resultan fáciles de detectar con la tecnología disponible.

2. Métodos

Para los experimentos realizados, se utilizó la interfaz de programación (API) de Face++ [Face++, 2017], una tecnología de servicios cognitivos en la nube gestionada por la empresa china Megvii, cuyo algoritmo principal está basado en redes neuronales convolucionales [Fan, 2014]. Dicha API permite enviar a los servidores de Face++ un conjunto de imágenes y recibir en formato JSON las coordenadas (x,y) de 83 puntos de interés para cada rostro detectado. Por medio de un programa escrito en Java se automatizó la transformación y envío de imágenes, y después se tradujeron los archivos JSON al formato .mat para poder analizar la información en Matlab.

El programa fue alimentado con una serie de imágenes obtenidas de la base de datos de Expresiones Faciales de Mujeres Japonesas (JAFFE) [Lyons, 1998], ya que ésta es de fácil acceso y contiene fotografías de varios rostros y diversas emociones. De este universo, se seleccionaron seis individuos con siete expresiones diferentes (una de ellas se considera neutral).

La teoría de Ekman postula la existencia de siete expresiones faciales básicas: alegría, tristeza, sorpresa, miedo, enojo, disgusto y desdén. En el estudio se buscó comprobar la relación entre los cambios en los puntos de interés provistos por Face++ y la clasificación realizada por humanos de las expresiones disponibles en la base de datos. De las siete expresiones básicas citadas, no se incluyó la última en las pruebas por no encontrarse explícitamente entre las imágenes disponibles. Una vez obtenidas las coordenadas de cada punto de interés, se normalizaron todos los conjuntos de acuerdo a dos puntos considerados menos sujetos a cambios a lo largo de todas las fotografías: los correspondientes al contorno superior de la nariz. Este proceso permite reducir las alteraciones de escala y rotación entre las imágenes utilizadas.

Para cada uno de los puntos de interés, se evaluó la diferencia entre su posición en el rostro de un individuo cuando su expresión se muestra neutral y su posición en cada una de las seis expresiones restantes, figura 1; finalmente se graficaron estas diferencias, promediadas entre los seis individuos. En la figura 2 se pueden apreciar los vectores de cambio para cada punto.

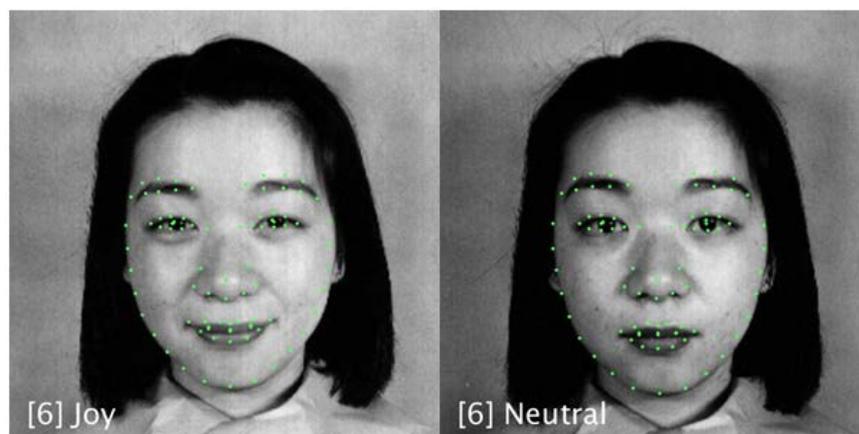


Figura 1 Localización de puntos de interés en dos expresiones del mismo individuo.

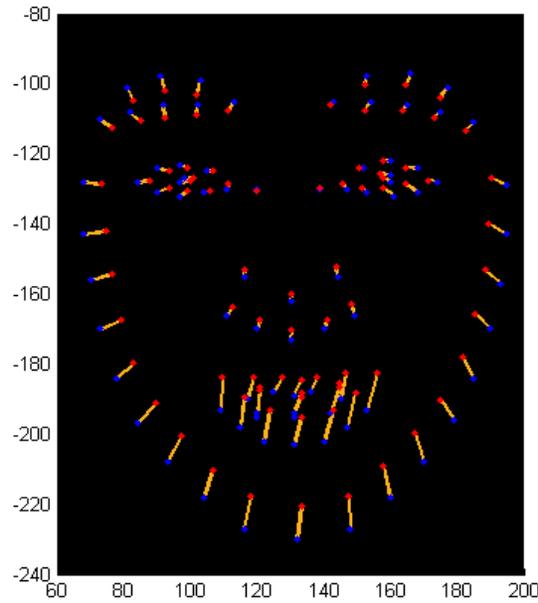


Figura 2 Diferencias en puntos de interés para las imágenes anteriores.

3. Resultados

Las gráficas obtenidas muestran por medio de una escala de colores el cambio en píxeles para la coordenada vertical de cada uno de los puntos de interés. Los puntos aparecen con sus descripciones respectivas y están agrupados por similitud en figuras 3 a la 7.

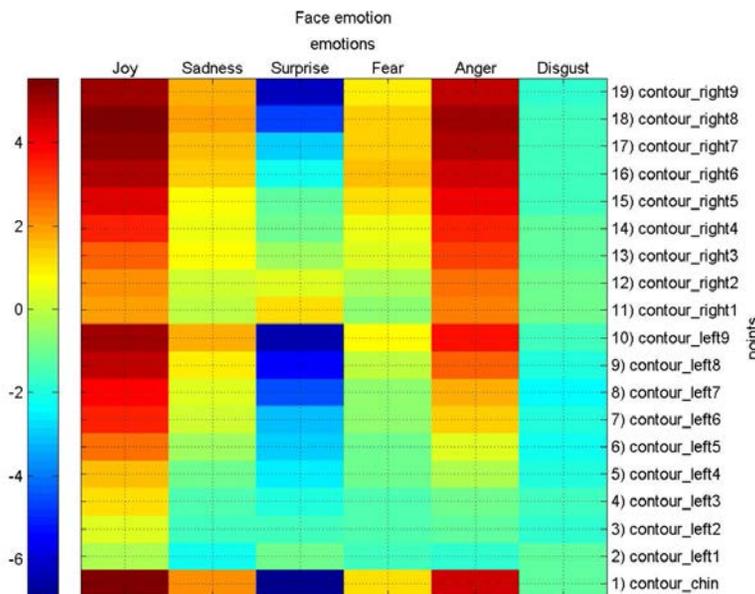


Figura 3 Desplazamientos verticales en los puntos asociados al contorno de la cara.

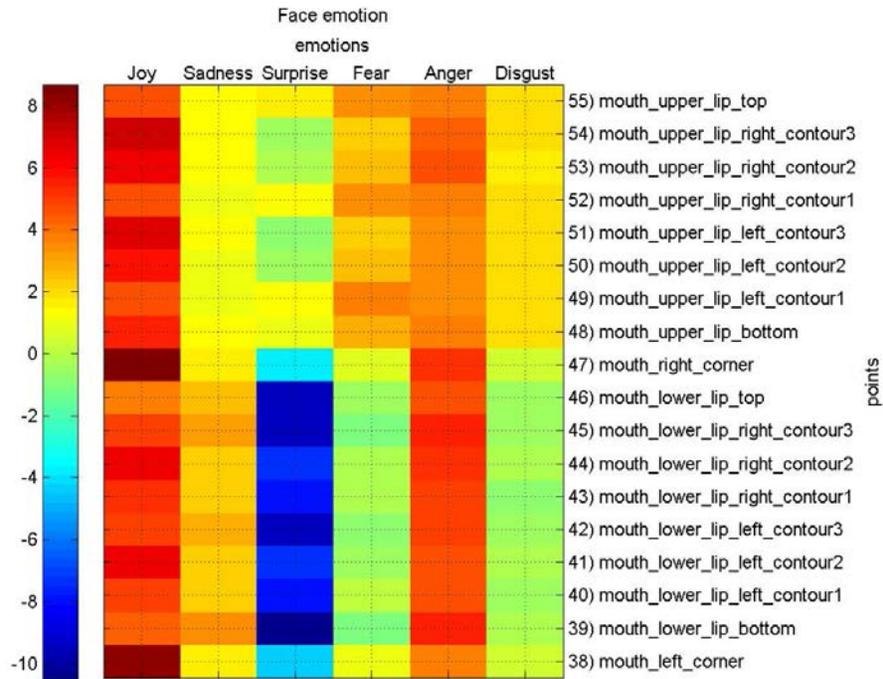


Figura 4 Desplazamientos verticales en los puntos correspondientes a la boca.

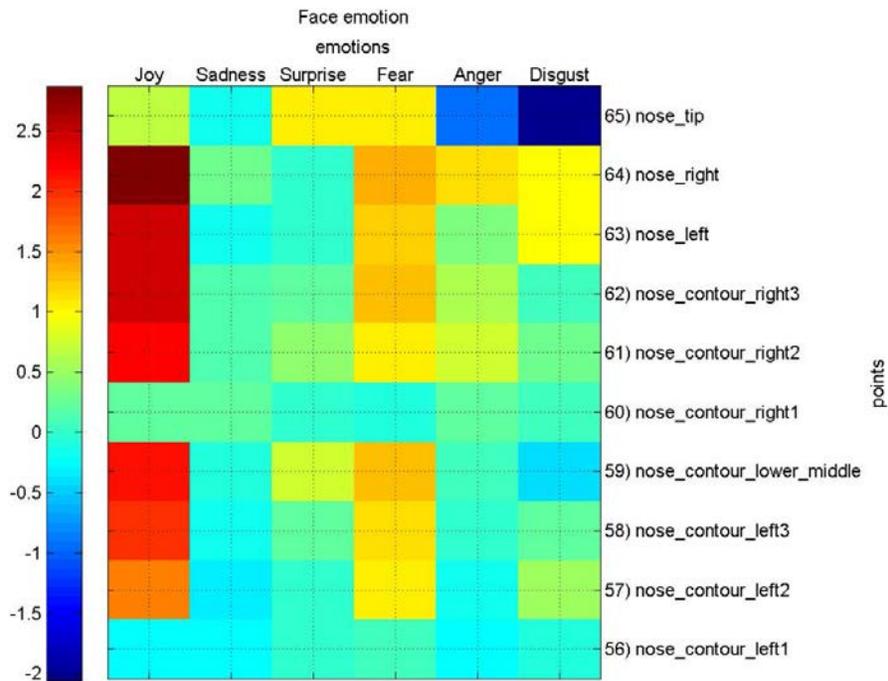


Figura 5 Desplazamientos verticales en los puntos correspondientes a la nariz.

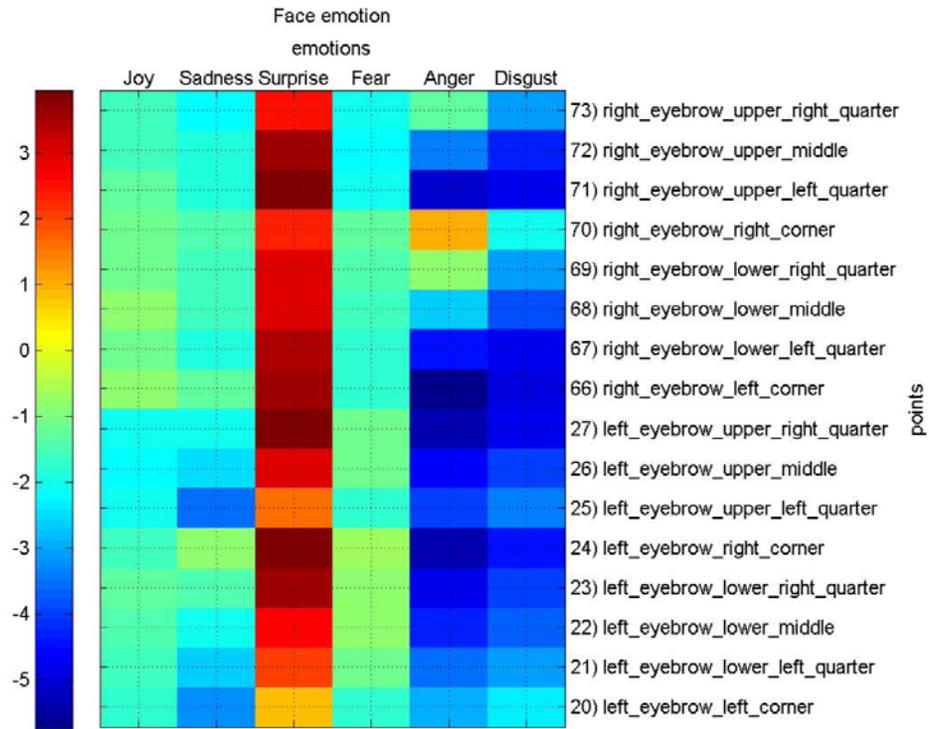


Figura 6 Desplazamientos verticales en los puntos correspondientes a las cejas.

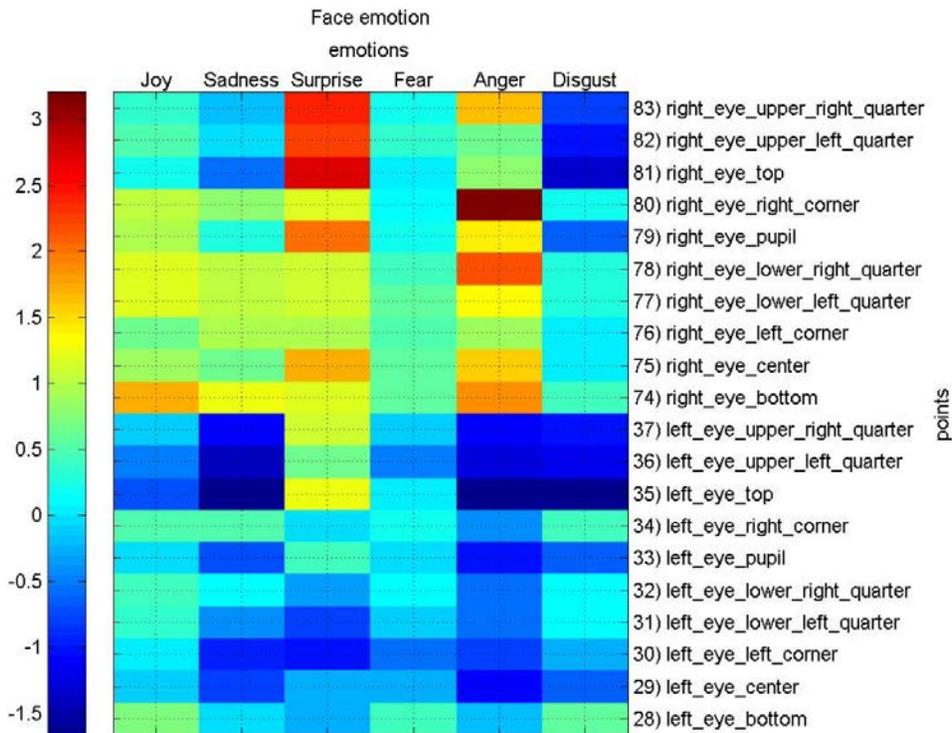


Figura 7 Desplazamientos verticales en los puntos correspondientes a los ojos.

4. Discusión

Puede observarse la congruencia entre los datos obtenidos y lo predicho por Ekman en las siguientes emociones:

- **Alegría:** se percibe un leve levantamiento de los párpados inferiores en los puntos 74 y 28 y levantamiento en las comisuras de los labios en los puntos 47 y 38. Los resultados coinciden con la predicción.
- **Tristeza:** el efecto del conjunto de músculos elevadores internos de las cejas puede apreciarse en los puntos 24 y 66, aunque deberían subir, parecen bajar ligeramente, es probable que por la acción del siguiente grupo muscular. Puede verse que los depresores de las cejas actúan en los puntos 26 y 25, los cuales claramente bajan, también en 73 y 72 se aprecia un descenso, aunque con menor claridad. Las comisuras de la boca 47 y 38 parecen subir ligeramente, cuando cabría esperar que bajaran por la actividad del grupo depresor de las comisuras.
- **Sorpresa:** los puntos 24 y 66 suben claramente, confirmando la previsión para los elevadores internos de las cejas. Lo mismo sucede con todos los puntos correspondientes a las cejas. Los músculos elevadores del párpado superior (35 y 81) presentan ascenso, con más claridad del lado derecho que del izquierdo. Esto concuerda con el supuesto de que los ojos se abren para expresar sorpresa, al igual que la boca, movimiento que puede distinguirse en los puntos del contorno facial: 1 a 19.
- **Miedo:** esta es la expresión más complicada de detectar, ya que actúan varios grupos de músculos simultáneamente, a veces cancelando entre sí los efectos observables. Por ejemplo, los músculos elevadores y depresores de las cejas afectan los puntos 24, 25, 26, 66 y 72 y 73 en sentidos opuestos, por lo que las lecturas obtenidas no son concluyentes. Lo mismo puede decirse de los puntos 28, 35, 74 y 81, que son jalados por los elevadores de párpados superiores y empujados por el músculo orbicular.
- **Enojo:** las cejas bajan claramente, como lo muestran los puntos 25, 26 y 72 los párpados superiores bajan en los puntos 35 y 81, de acuerdo con lo

previsto. Las coordenadas verticales no parecen ser suficientes para confirmar la tensión esperada en los labios ni en los orbiculares. Tal vez haga falta un análisis ligeramente distinto.

- **Disgusto:** el músculo transversal de la nariz hace subir los puntos 63 y 64 como estaba pronosticado, pero los puntos 38 y 47 en las comisuras de los labios parecen subir cuando deberían bajar, por otro lado, el depresor del labio inferior asociado al punto 39 no parece moverse.

5. Conclusiones

Los resultados obtenidos confirman la teoría para tres de las seis emociones estudiadas.

Para las otras tres, dada su complejidad (el método FACS requiere un entrenamiento especializado) y la aparente acción contraria de algunos grupos de músculos cabría replantear el tipo de análisis realizado. En el futuro sería interesante alimentar los puntos de interés a algún tipo de inteligencia artificial (p.e. redes neuronales) para obviar el estudio pormenorizado de cada músculo individual.

No todos los puntos de interés provistos por la API de Face++ tienen la misma importancia. Podrían obtenerse resultados similares con un conjunto más reducido.

Las imágenes utilizadas corresponden a unas pocas personas, mujeres exclusivamente, y sus expresiones son actuadas. Sería útil contar con una base de datos más amplia y con expresiones naturales.

La normalización aplicada a los puntos de interés no es suficiente para corregir los errores que los movimientos de la cabeza introducen en los cálculos. Si uno de los sujetos levanta la barbilla o voltea hacia un lado, afecta a todos los puntos de interés.

6. Bibliografía y Referencias

- [1] Ekman, P., & Friesen, W. V., Constants Across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology*, 1971.

- [2] Ekman, P., Friesen, W. V., & Hager, J., The Facial Action Coding System (FACS): A technique for the measurement of facial action. Palo Alto, California: Consulting Psychologists Press, Inc. Ekman, P. Levenson. RW, & Friesen WV, Autonomic nervous system activity distinguishes among emotions. *Science*, 221, 1208-12, 1983.
- [3] Face++ - Cognitive Services - Leading Facial Recognition Technology: <https://www.faceplusplus.com/>, último acceso junio 2017.
- [4] Fan, H., Cao, Z., Jiang, Y., Yin, Q., & Doudou, C., Learning Deep Face Representation. ArXiv preprint arXiv: 1404, pp. 3840, 2014.
- [5] Giannakakis, G. et al. Stress and anxiety detection using facial cues from videos. *Biomedical Signal Processing and Control*, 31, pp. 89-101, 2017.
- [6] Hamelin, N., El Moujahid, O., Thaichon, P. Emotion and advertising effectiveness: A novel facial expression analysis approach. *Journal of Retailing and Consumer Services*. 2017. 36, pp. 103-111.
- [7] Hamm, J., Kohler, C.G., Gur, R.C., Verma, R. Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of Neuroscience Methods*, pp. 237-256, 2011.
- [8] Krithika L.B, Lakshmi P.G., Student Emotion Recognition System (SERS) for e-learning Improvement Based on Learner Concentration Metric. *Procedia Computer Science*, 85, pp. 767-776, 2016.
- [9] Liong, et al. Spontaneous subtle expression detection and recognition based on facial strain, *Signal Processing: Image Communication*, 47, pp. 170-182, 2016.
- [10] Lyons, M.J. et al. Coding Facial Expressions with Gabor Wavelets. *Proceedings, Third IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE Computer Society. Nara, Japan, April 1998.
- [11] Sariyanidi, E. et al. Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(6), pp. 1113-1133, 2015.

- [12] Tarnowski, P., Kołodziej, M., Majkowski, A., Rak, R.J., Emotion recognition using facial expressions, *Procedia Computer Science*, 108, pp. 1175-1184, 2017.
- [13] Wang, N., et al., Facial feature point detection: A comprehensive survey. *Neurocomputing*, 2017.
- [14] Yan, J., Zhang, X., Lei, Z., Li, S. Face detection by structural models, *Image and Vision Computing*, 32 (10), pp. 790-799, 2014.
- [15] Yu, C. & Ko, C., Applying FaceReader to Recognize Consumer Emotions in Graphic Styles, 60, pp.104-109, 2017.
- [16] Zhan, S., Tao, Q. & Li, X. Face detection using representation learning, *Neurocomputing*, 187, 19-26, 2016.
- [17] Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. Face recognition: A literature survey, *ACM computing surveys (CSUR)*, 35(4), 399-458, 2003.