AUTOMATIC MICROEXPRESSION DETECTION. HISTORY AND DEVELOPMENTS

Denis SMOCVIN

Department of Software Engineering and Automation., Group FAF-212, Faculty of Computers Informatics and Microelectronics, Technical University of Moldova, Chișinău, Republic of Moldova

Corresponding author: Denis Smocvin, denis.smocvin@isa.utm.md

Abstract. *Emotions govern our life. Through facial expressions we can understand what other people feel. Microexpressions are a special kind of facial expressions that are hard to spot. Interest arises over the development of technologies for their automatic detection. Some problems arise in the development of models for detection microexpressions.*

Keywords: *facial expressions, emotion, machine learning, computer vision, FACS*

Introduction

Emotions govern our lives. They are more than fundamental impulses for survival. People may choose not to eat if disgusted by the food at hand. Sexual impulses are vulnerable to emotions too: there are people that are hesitating to make a sexual act because he or she may be disgusted or afraid. Sometimes emotions may drive someone into suicide, stepping over the willingness to live. It doesn't depend how much we want it; we cannot live without emotions. They are an important means of communication and a survival mechanism that can sometimes be inconvenient.

Studying emotions can be useful in many fields like psychology, animation, education, recruitment, law enforcement, childcare etc. One potential application of microexpression analysis is lie detection. When lying more contradictory behaviour could be found in verbal and non-verbal signals.

Emotions are expressed in multiple ways. It may be through facial expressions, voice, gestures, impulses that lead into action (for example it is usual to get closer to someone when we are angry), physiological changes and some invisible clues (like thoughts for example). We will focus on facial expressions. Precisely we will focus on microexpressions, a certain type of facial expressions, and their detection by humans and computers. But before explaining what it means, let us explore how facial expressions were studied.

History

It all started with Darwin's statement that emotions are universal, meaning they are not a cultural phenomenon, but a product of evolution. The question of universality was tackled by many scientists. There were many against it, like the American anthropologist Ray Birdwhistell, who specialized in the study of expressions and gestures. His arguments were based on cultural anthropology - every socially important thing, like facial expressions, are the product of education, therefore they differ for each culture [1]. His idea was indirectly supported by travellers, who thought that the expressions have different meanings across cultures based on their observations.

This problem can be solved by the idea of *display rules*, created by the most famous researcher in the domain of facial expressions Dr. Paul Ekman [2]. He stated that there are certain socially learned rules that vary across cultures. These rules state how you should express your emotions when you are near another person. Hence, in sport competitions we expect the losers to not express their sadness or disappointment. This was further studied in an experiment where people were instructed to watch a movie about chirurgical operations and car accidents. The Americans and Japanese presented the same expressions. But when a researcher was in the same room, the Japanese masked their negative expressions with a smile more than Americans did. This means that the expression of

emotions specifically may be universal, compared to some expressions like shaking the head from right to left as a means of saying "no".

The study had a vulnerability. It may have been the case that all the persons that were involved in the study had learned the meaning of facial expressions specific to the western culture from movies, or other media. This vulnerability spread to Ekman's other studies from Chile, Argentina, Brazil, the USA, Indonesia, USSR etc. To solve this problem, Ekman went to study the expressions of the tribes of New Guinea that had a Stone Age-style of life. They had no idea what a camera is. They were isolated from other cultures.

Initially Ekman studied the films of the neurologist Carleton Gajdusek, who studied a local pandemic (and later was awarded the Nobel Prize for discovering 'slow' viruses). There were clues that the facial expressions were universal. Ekman didn't spot any facial expressions he could not recognize from the video tapes. Even the expressions isolated from the social situations were correctly identified by his assistant, along with the context. His assistant's descriptions of what he had seen were the same as what Gajdusek described. Later, Ekman went to New Guinea to conduct a series of studies that indicated that facial expressions are universal. His results were accurately reproduced by the anthropologist Karl Heider. Other studies reflected the same results - facial expressions are universal [3-7].

The intercultural studies raised a series of questions: How many expressions do people have? Do the expressions contain frank or deceptive information ? Can every facial movement be considered an expression ? Can we lie with facial expressions like we do with words ? Now, these questions have answers thanks to the Facial Expression Coding System (FACS) developed by Paul Ekman [8]. With this tool Ekman found that a human can express more than ten thousand expressions and identified the ones that are relevant to emotions. FACS describes facial movements in anatomical terms using words, images and videos. It is currently used to study facial movements by researchers and by animation studios. Also, there is an interest in the domain of informatics for automating and accelerating the measurements.

We are to find how computers detect facial expressions, and specifically microexpressions. Before doing that, let's explore what kind of facial expressions exist.

Types of facial expressions

There are 4 types of facial expressions. A *complete* expression is visible on the whole face. A *partial* expression is visible on a part of the face. It may appear when people try to hide their feelings. A *diminished* expression happens when an emotion of low intensity is expressed. Here, the muscles don't contract much. When we have a partial or diminished expression it means that we are starting to express an emotion or we hide our feelings. The attempt to hide any emotion on the face may lead to a *microexpression* - a type of facial expression that lasts less than a fifth of a second. Microexpressions appear when someone consciously tries to inhibit the expression of any emotion. The inhibition may be unconsciously as well, when a person isn't aware of his or her sentiments.

Microexpressions may be complete and very short or very short and/or diminished. The combination of the three - micro (very short), partial (registered in a single zone of the face), and diminished (minimal muscular contraction) is the hardest to recognize [9].

Even people trained to detect microexpressions may find it hard to do so. Their more accurate detection happens not in-person, but by studying footage. One of Ekman's studies show that the majority of people don't seem to use the information from the more subtle facial expressions. In some conversations subtle expressions appear more often than the intense ones. Often, the subtle ones are more important, because they inform about the things not spoken yet that may not even be spoken [10-11]. Here arises the interest in the automatic detection of facial expressions and microexpressions.

Automatic microexpression detection systems

There are seven fundamental emotions: happiness, sadness, fear, anger, surprise, contempt and disgust. The figure below, represents the facial expressions of each of these emotions:

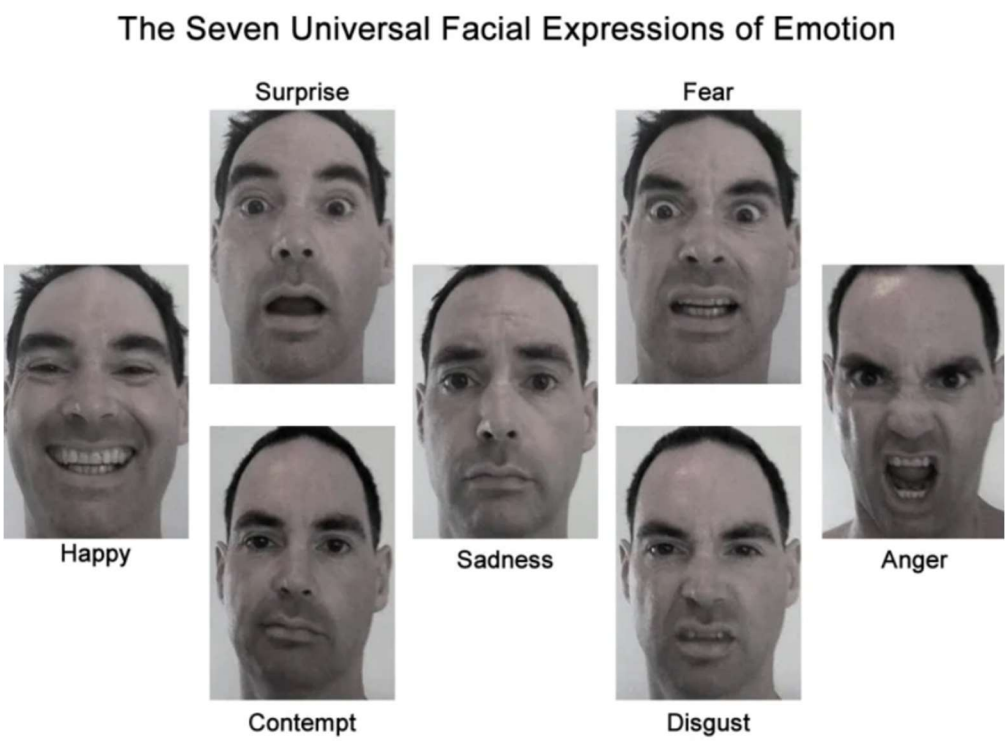

Figure 1: Facial expressions of the seven fundamental emotions [12]

Humans are trained to detect the subtle versions of the expressions above. The important aspect is that for each of the expressions you are allowed to look at them for a very short amount of time. Then you should write which emotion did the observed expression represent. There are a lot more configurations that may express the same emotion, but for training computers these basic ones are reproduced in order to train models for detection.

Microexpressions are too short and subtle for human eyes to perceive. According to Ekman, for microexpression recognition tasks, ordinary people without training only perform slightly better than chance on average. This is why automatic methods are being developed. Machine learning and computer vision algorithms play a big role in this process. Computer vision and deep learning techniques such as convolutional neural networks are used to create models for detecting emotion on human face.

Pfister et al (2011) [13] pioneered the research on spontaneous microexpression recognition with the first publicly available spontaneous microexpression dataset SMIC and achieved results that favourably compare with the accuracy of humans. Li X. et al (2018) [14] developed a system integrating microexpression recognition and detection toward reading hidden emotions in videos that achieved 80.28% for three classes: positive, negative and surprise. For five classes - happiness, disgust, surprise, repression, and other - 57.49% accuracy has been achieved. This is already a better accuracy than that of humans.

Instead of directly recognizing a certain number of prototypical expressions as in most of the previous research, Action Units (AU) can provide an intermediate meaningful abstraction of facial expressions, and carry lots of information which can help better detect and understand people's feelings. Action units are the fundamental actions of muscles (individual or group of muscles). AUs were used for macroexpression analysis including pain detection and pain intensity estimation. It may be beneficial to use them in microexpression analysis as well.

Another aspect is the datasets. When training a machine learning model for a specific task, a specific dataset should be composed. There have been some microexpression datasets collected for research: SMIC, CASME, CASME II, SAMM, MEVIEW, CAS(ME)2. There are hundreds of microexpression videos from 30 to 40 subjects in these datasets and there is still a need for higher

quality, including naturally collected and well-annotated microexpressions large scale datasets. Creating such a dataset from scratch is time-consuming and very challenging. A possibility is to use the publicly available online videos from different platforms and mining with video tagging techniques followed by manual annotation. Another option could be collaborative and parallel data collection and labelling through cloud sourcing. For lie detection purposes, audio may be necessary along video footage.

The available datasets do not represent realistic situations, because the expressions are represented from the frontal view with stable and bright light conditions, with no lighting variations, faces fully visible. In real-world applications it is unlikely to have such conditions. Algorithms must be developed in such a way as to adapt to different conditions.

Conclusions

Automatic Microexpression Detection Systems are being developed for improving the speed and quality of microexpression detection. The problems that we are facing today is the quality of algorithms, dataset quality and data volume, dataset creation, applicability of samples from available datasets. Solving these problems will give us the opportunity to create powerful models for detection microexpressions that can be used in lie detection, interviews, education and other fields.

References

1. BIRDWHISTELL, R. L. . *Kinesics and Context.* Philadelphia: University of Pennsylvania Press, 1970.
2. EKMAN, P. *Emotions Revealed, Second Edition: Recognizing Faces and Feelings to Improve Communication and Emotional Life*, 2003
3. EKMAN, P. “Universals and cultural differences in facial expressions of emotions"
4. THOMPSON, J., “Development of facial expressions of emotion in blind and seeing children”, *Archives of Psychology,* 1941, 37.
5. FULCHER, J. S.. “‘Voluntary’ facial expression in blind and seeing children” *Archives of Psychology,* 1941, 38.
6. EIBL-EIBSEFELDT, I.. *Ethology, The biology of behavior,* 1970.
7. HOLT, REINHART and WINSTON. GALATI, D., SCHERER, K.R. & RICCI-BITTI, P/E. “Voluntary facial expression of emotion: Comparing congenitally blind with normally sighted encoders” *Journal of Personality and Social Psychology*, 1997, 73:1636-79
8. Paul Ekman Group, Facial Action Coding System. [online]. [accessed 03.03.2022]. Available: https://www.paulekman.com/facial-action-coding-system/
9. EKMAN, P., “Emotions Revealed. Understanding Faces and Feelings”, 1997, 321-2
10. EKMAN, P., O’SULLIVAN, M. & FRANK, M. “A few can catch a liar” *Psychological Science* , 1997, 10:263-66
11. EKMAN, P. & O’SULLIVAN, M. “Who can catch a liar?” *American Psychologist,* 1997, 46:913-20
12. Try not to smile. The seven universal facial microexpressions of emotion. https://www.pinterest.com/pin/90494273737111288/ . [online]. [accesed 03.03.2022]
13. PFISTER, T., LI, X., ZHAO, G., and PIETIKAINEN, M. “Recognising spontaneous facial micro-expressions,” in *Proceedings of the IEEE International Conference on Computer Vision* 2011 (Barcelona)
14. LI, X., HONG, X., MOILANEN, A., HUANG, X., PFISTER, T., ZHAO, G., et al. Towards reading hidden emotions: a comparative study of spontaneous micro-expression spotting and recognition methods. *IEEE Trans. Affect. Comput.* 2018, 9, 563–577. doi: 10.1109/TAFFC.2017.2667642