A Study on the Basics and Approaches of Identifying and Analyzing Human behaviour

Kshitij Patil, Dr. Shwetambari. A. Chiwhane

Computer Department, NBN Sinhgad School of Engineering

Abstract

With the development in fields of artificial intelligence and machine learning advancing rapidly, the field of human behaviour identification and its analysis has become a widely under research topic all over the globe. The cumulative aspect of processes involved in recognition of human affections is termed as affective computing. It is an integral part of sentiment recognition as well as computer-user interaction.It takes into consideration the emotional states of the user.The identification of human behaviour or feelings is a process which involves multiple domains working together in-order to acquire an inference.There are three sub-domains to be considered while computing human emotions:affect representation and recognition, affect interpretation and modeling, affect simulation while computer vision as well as pattern recognition which are complemented by artificial intelligence and machine learning play an important role in deriving a conclusion.

INTRODUCTION

Computers in today's life play a very important role. With concepts like artificial intelligence and machine learning spreading all over the globe, machines have become even more powerful in the way that they now have an ability to learn on their own when certain repetitive tasks are being performed. This has resulted in higher efficiency in terms of computational time and man hours. Due to this ability, new fields have been introduced. Analysis of human behaviour and sentiment is one of the examples. This analysis is termed as Affective Computing, wherein the machine can identify the user's emotions and affections and adjust its behaviour accordingly.

Facial expressions can be identified using two approaches namely, message judgement and sign judgement. Message judgement intents to draw conclusions by directly decoding the facial expression on the user's face. Message judgement results in computing the output as happy,sad,surprised,etc. Sign judgement on the other hand targets to study the physical signals such as raised eyebrows, depressed lips,etc.

The message judgement has a major drawback which makes sign judgement superior. The downside is that message judgement is incapable of interpreting the full range of emotions. It assumes that there is an obvious many-to-one mapping of facial expression to emotions which is not accurate as emotions corresponding to an expression vary from person to person. With that being said, this is where sign judgement demonstrates its mettle. Sign judgement uses descriptors. Commonly used descriptor is FACS which stands for Facial Action Coding System[5] which was developed by Carl-Herman Hjortsjö and then adopted by Paul Ekman.

Human expressions are coded using FACS. The revised version of the which defines 32 facial muscle actions termed as Action Units or AUs[2]. Since, each emotion is an outcome of stimulation of facial muscles, every emotion n can be identified as a cumulative outcome of AUs,which is objective and makes

ISSN: 2233-7857 IJFGCN

up the physical state or the person while subjective traits such as personality,emotional state,context,etc cannot be directly deduced as facial expressions are verified by semantic content of facial expression along with physical realizations. The 6 basic human emotions taken into consideration are-fear,happiness,surprise,disgust,anger and sadness as described in[4]. Hence, human emotions can be identified by using AUs. As we know human emotions directly reflect human behaviour in ideal cases, we can thereafter identify the subject's behaviour as well.

GENERIC PROCESSES IN COMPUTING THE BEHAVIOUR

Before the machine starts identifying the human behaviour it has to make some adjustments to the data for which computing is to be carried out. Suppose the machine has to compute the behaviour of a subject from one of its images. The machine will have to refine the input(subject's image) according to its requirements. The processes usually carried out are in the following sequence- Pre-processing, feature derivation,machine analysis of facial actions [2].

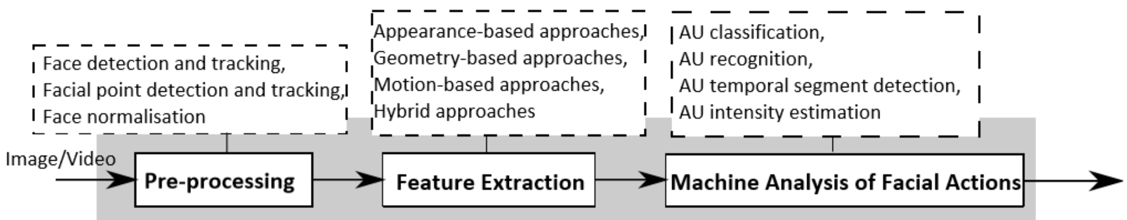

Fig. 1.Generic facial action recognition system configuration.

PRE-PROCESSING

The data provided to the machine must be pre-processed before initiating the elicitation of desired facial aspects to draw as accurate conclusions as possible. The intent of this step is to line up faces into a reference frame to extract aspects which will give reference to the same locations.It is a three step process-face detection, landmark localization and alignment.

Face Detection

This is the 1st step in analysis of faces. There are many algorithms used to detect faces. Most algorithms rely on frontal faces, making them the frame of reference resulting in a higher degree of reliability of the algorithms' computability[8 9]. The Viola&Jones face detector is most widely used. However, the recent algorithms perform better at detection but they come which increased cost.

Face Landmark Detection

It is the process in which the many face landmarks are used to the define the shape of the face. These landmarks are usually the peripherals of the eyes, tip of the nose, outermost point of the ears, hairline,etc. This results in better face recognition after face detection is carried out resulting in better performance of the machine.

ISSN: 2233-7857 IJFGCN

Face Alignment

After face landmark detection, the face is pretty much clear to the machine and read to process. Here, the face alignment with a common pre-defined reference co-ordinate system. By carrying out these measures things like transformation in facial expressions can be identified. The deviation in current frame with respect to the frame with no expressions is known as δ-image. These δ-image are used to track every bit of change in the facial features. Once the emotion or behaviour is identified these identified transformations can be used for future reference especially when machine learning is applied.

Feature Derivation

It is responsible for converting the data extracted from pixels to a higher level representation of motion and/or spatial arrangement of facial framework. Its aims towards minimizing the dimensionality of input space so as to lessen the discrepancy in the data caused by unnecessary constraints such as lighting, alignment errors or blur, and to minimize the sensitivity to circumstantial effects such as identity and head pose[6 7]. The feature extraction methodologies can be classified into three categories- Appearance features, geometric features, motion features[2].

Fig. 2.Upper and lower face AUs instances defined in the FACS

Appearance Features

They are the most commonly used features. The colour and texture of the faces are delineated by these features. These features can be susceptible to non-frontal head poses and lighting changes.

Geometric Features

The locations of facial landmarks are extracted and summarized by using these features. Change of expressions are a result of displacement caused in the facial muscles. Extracting facial point locations was an issue with these features but with recent research these have become more reliable.

Motion Features

These features capture flexible deformations in the skin due to the activation of facial muscles. In contrast to geometric features they are related to dense motion rather than to the motion of a discrete set of facial landmarks. Motion features are less person specific than appearance features. However, they require the full elimination of rigid motion. This means that they can be affected by misalignment and varying illumination conditions, and it is unclear how to apply them in the presence of non-frontal head poses.

The six basic emotions can be identified at this stage of computing. Basic emotions are mentioned in [4].

ISSN: 2233-7857 IJFGCN

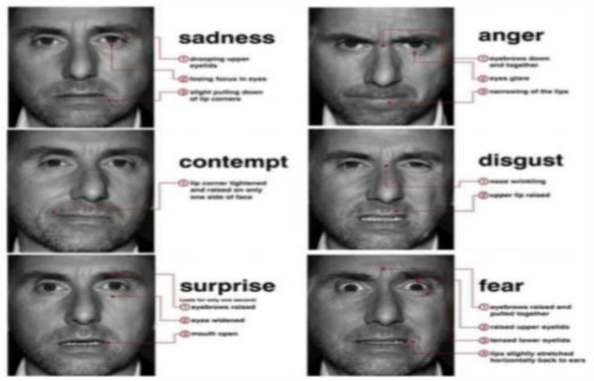

Fig. 3. Basic Emotions

Machine Analysis of Facial Features

In this step, various algorithms using machine learning are put to use in order to reach closer to the conclusion. The problems are classified into:Detection of AU, intensity estimation of AU, temporal segment detection of AU and classification of AU. Some machine learning algorithms commonly used are-

1. Naïve Bayes Classifier: It is based on the Bayes theorem. It works on conditional probabilities. It is most suitable in cases with high input dimensions. This algorithm is a popular method to categorize and retrieve text. Therefore, it can be used for gesture recognition mainly when text is involved.

2. K nearest neighbour (KNN): It is a simple algorithm for categorizing the data. It is very similar to the human nature of recalling events in the past and using them to make new decisions. It is mostly used in regression and classification problems. Majority voting by the neighbors is done in order to carry out classification of new data on the basis of similarity measures.

3. Support Vector Machine (SVM): Another algorithm used for regression and classification under supervised machine learning. It is mostly used to detect anxiety. It can be used to classify the facial expressions based on extracted features. It can solve non-linear problems with a scarce volume of data[10 11]. It can perform non-linear cassification problems by using kernel trick, i.e., implicitly mapping inputs to highly dimensional spaces.

There are many other algorithms such as Linear discriminant classifiers(LDC), Gaussian Mixture Model(GMM), Artificial Neural Networks(ANN), Decision Tree algorithms,etc. that can be used to make computing emotions easier, faster, efficient and more accurate.

BEHAVIOUR DETECTION CUES

Behaviour detection cues mostly include the cues in [2]. These cues with the help of some machine learning or artificial intelligence algorithms may prove to be of great assistance to identify he subject's resulting in a higher degree of accuracy of computation. Identifying and processing cues is a very vital step to derive a conclusion. The human behavior detection framework can be represented as shown in[3].

ISSN: 2233-7857 IJFGCN

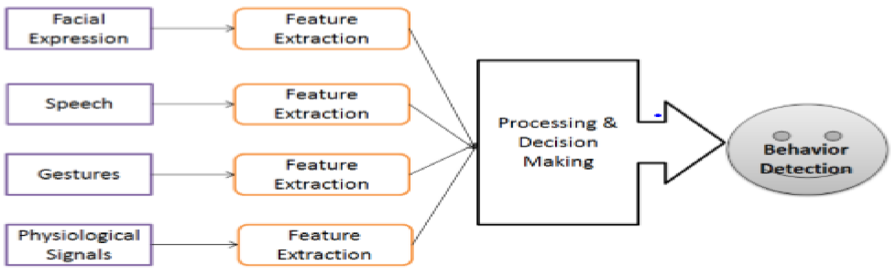

Fig. 4. Human Behavior Detection Framework with multiple modalities.

CHALLENGES

Affective computing is such a field where everything is easier said than done. The major challenge faced in analysis sentiment is that the AUs are not suitable for an environment with no restrictions. Some of these restrictions include handling occlusions, non-frontal head poses,varying illumination conditions,etc.

Non-availability of appropriate data sets covering full range of emotions for unconstrained environment is another issue. Emotions vary person to person therefore it is impossible to find a generic data set moreover it should cover all ranges of emotions in an unconstrained manner which makes it almost impossible as of now.

Frequent encountering of low intensity AUs are a smaller issue but these may influence the result of computation. These are not well computed as they do not provide enough information to deduce a particular conclusion. A low intensity AU is mostly encountered when the subject is intentionally controlling their facial behaviour in order to hide their emotions.

CONCLUSION

As Affective Computing is a newly emerging field it is under research. It is research to such an extent that there are no data sets covering the huge spectrum of inputs due to which there are no tried,tested and proven algorithms to computed the human behaviour. Moreover, scenarios like the subject intentionally controlling their facial actions adds even more problems to the already existing huge set of problems. As if the number of challenges faced are not enough to add to which low intensity AUs generated contribute to a great extent and make it even worse. They are nearly non-computable but are capable of influencing the deduced conclusion. To summarize, this field needs to undergo huge research and development inorder to reach a point where proposed techniques are approved and serve as generic techniques.

REFERENCES

[1] B. Dudzik et al., "Context in Human Emotion Perception for Automatic Affect Detection: A Survey of Audiovisual Databases," 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), Cambridge, United Kingdom, 2019, pp. 206-212.

[2] B. Martinez, M. F. Valstar, B. Jiang and M. Pantic, "Automatic Analysis of Facial Actions: A Survey," in IEEE Transactions on Affective Computing, vol. 10, no. 3, pp. 325-347, 1 July-Sept. 2019.

ISSN: 2233-7857 IJFGCN

[3] N. Jadhav and R. Sugandhi, "Survey on Human Behavior Recognition Using Affective Computing," 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN), Lonavala, India, 2018, pp. 98-103.

[4] T. S. Saini, M. Bedekar and S. Zahoor, "Analysing human feelings by Affective Computing - survey," 2016 International Conference on Computing Communication Control and automation (ICCUBEA), Pune, 2016, pp. 1-6.

[5]Hamm, C. Kohler, R. Gur and R. Verma, "Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders" , Journal of Neuroscience Methods, vol. 200, no. 2, pp. 237-256, 2011.

[6] Shwetambari Kharabe, C. Nalini," Robust ROI Localization Based Finger Vein Authentication Using Adaptive Thresholding Extraction with Deep Learning Technique", Journal of Advanced Research in Dynamical & Control Systems, Vol. 10, 07-Special Issue, 2018.

[7] Shwetambari Kharabe, C. Nalini," Using Adaptive Thresholding Extraction - Robust ROI Localization Based Finger Vein Authentication", Journal of Advanced Research in Dynamical & Control Systems, Vol. 10, 13-Special Issue, 2018.

[8] Shwetambari Kharabe, C. Nalini," Evaluation of Finger vein Identification Process", International Journal of Engineering and Advanced Technology (IJEAT), International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6S, August 2019.

[9] Udayan Birajdar, Sanket Gadhave, Shreyas Chikodikar, Shubham Dadhich, Shwetambari Chiwhane, "Detection and Classification of Diabetic Retinopathy Using AlexNet Architecture of Convolutional Neural Networks", Proceeding of International Conference on Computational Science and Application, online 05 January 2020, pp 245-253.

[10] Dr. C. Nalini, Shwetambari Kharabe, Sangeetha S," Efficient Notes Generation through Information Extraction", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6S2, August 2019.

[11] Shwetambari Kharabe, C. Nalini , R. Velvizhi," Application for 3D Interface using Augmented Reality", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-6S2,August 2019.

ISSN: 2233-7857 IJFGCN