International Journal of Engineering & Technology, 7 (3.20) (2018) 284-290

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Facial Action Units Analysis using Rule-Based Algorithm

[1]Hamimah Ujir, [2]Irwandi Hipiny and [3]D.N.F. Awang Iskandar

[123] Universiti Malaysia Sarawak, Malaysia
**Corresponding Author Email: uhamimah@unimas.my*

Abstract

Most works in quantifying facial deformation are based on action units (AUs) provided by the Facial Action Coding System (FACS) which describes facial expressions in terms of forty-six component movements. AU corresponds to the movements of individual facial muscles. This paper presents a rule based approach to classify the AU which depends on certain facial features. This work only covers deformation of facial features based on posed Happy and the Sad expression obtained from the BU-4DFE database. Different studies refer to different combination of AUs that form Happy and Sad expression. According to the FACS rules lined in this work, an AU has more than one facial property that need to be observed. The intensity comparison and analysis on the AUs involved in Sad and Happy expression are presented. Additionally, dynamic analysis for AUs is studied to determine the temporal segment of expressions, i.e. duration of onset, apex and offset time. Our findings show that AU15, for sad expression, and AU12, for happy expression, show facial features deformation consistency for all properties during the expression period. However for AU1 and AU4, their properties' intensity is different during the expression period.

Keywords: *facial action units; temporal analysis, facial expressions, dynamic analysis*

1. Introduction

Research areas such as user profiling, human psychology and augmented as well as virtual reality require efficient expression classification. Facial expression is about the deformation of facial features. It is a highly dynamical processes and looking at the sequences of faces instances rather than to still images can help to improve facial expression classification performance (Berretti et al., 2012). The Facial Action Coding System (FACS), introduced by Ekman and Friesen (1978), is the leading method to measure facial deformation in psychological research. FACS provides the descriptive power necessary to describe the details of facial expression (Tian et al, 2001). Action units (AUs) defined by FACS represent the facial muscle activity that produces facial appearance changes (Ekman and Friesen, 1978).

Face is by nature dynamic where it demonstrates universe set of facial expressions that involve 3D space and temporal dimension (3D plus time). Posed and spontaneous facial dynamics can be explicitly analyzed by detecting a sequence of temporal segments (i.e. natural, onset, apex, offset, natural) (Hess and Kleck, 1990).

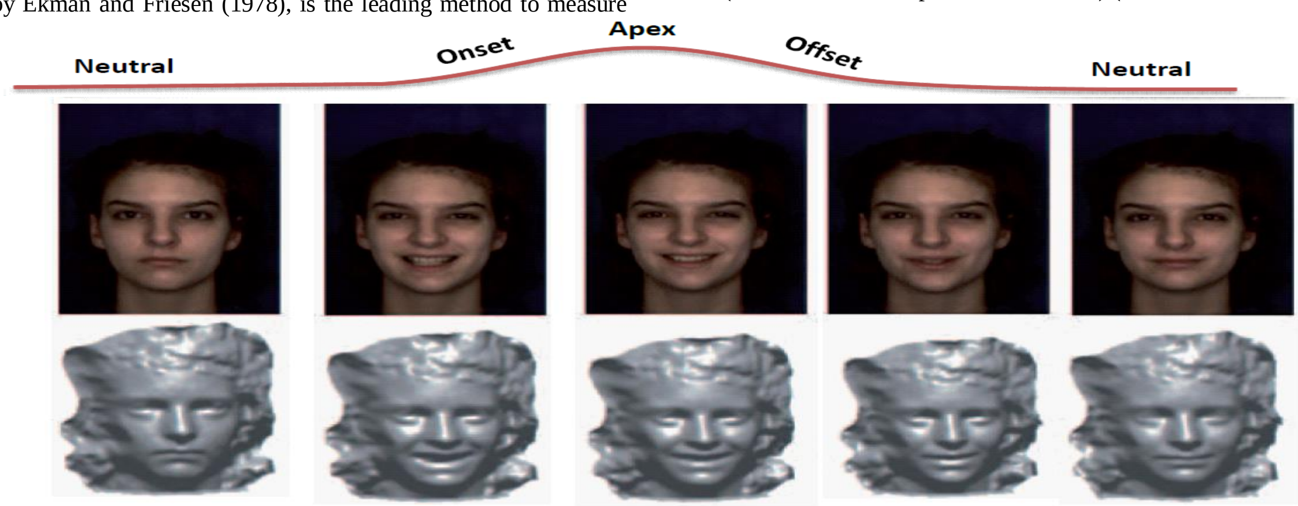

Figure 1: Sample expression image and model sequences from BU-4DFE database (Yin et al., 2008).

The temporal modeling starts with an initial natural segment, followed by onset which refers to an activation of the expression, refer to figure 1. Next is the maximum intensity of facial expression which is apex and followed by offset, the deactivation of the expression. The closing of the temporal segment will be back to natural segment once again. The dynamic characteristic of the database apparently crucial for expression recognition since it describe the temporal information between the peak and the natural state of the emotions. Existing works analyze temporal facial expression in video by tracking the facial features and measuring the movement in the frames (Suja et al., 2015).
Valstar and Pantic (2006) used data from Cohn-Kanade and MMI databases in their 15 AUs recognition experiment. They used boosting techniques to select important features and Support Vector Machine (SVM) to classify the AUs. In this work, the data (i.e. 3D facial points) is obtained from BU-4DFE database and a rule-based algorithm is used to classify AUs.
Russell and Bullock (1986) described human facial expressions in two-dimensional where pleasantness is plotted versus the level of arousal. The happy expression is on the pleasant side and the level of arousal is between average to high. On the other hand, sad is on the low arousal and unpleasant side of the graph. In this work, happy expression is chosen due to positive emotional expression and an expression of sad as the negative emotional expression.
Lien et al. (2000) presented an automated recognition of fine-grained changes in facial expression to capture the subtlety of human emotion and paralinguistic communication. Only 15 AUs that occur in the upper and lower face and that are common in emotion and paralinguistic communication are considered in their work. Their system extracted three type of feature information module such as dense-flow extraction, facial-feature tracking, and edge and line extraction. The feature information extracted is fed to discriminant classifiers or hidden Markov models that classify it into FACS action units.
The function of facial expression intensity is to convey the level of psychological arousal to other people. The estimation of human facial expression intensity is an important step in enhancing the capability of human-robot interfaces. In (Mahoor et al., 2009), an automated intensity measurement of naturally occurring facial actions is presented, which focused only on AU12 (lip corner puller) and AU6 (cheek raiser). Images of infants in a live face-to-face communication are used as the data. Using 4D data, the important facial expression episode which is the onset-apex-offset will be highlighted. Suja et al. (2015) proposed a dynamic method to detect apex frame of a video sequence. The Euclidean distance between feature points in apex and neutral frame is determined and their difference in corresponding neutral and apex frame is calculated. The calculated distances between a set of frames is summed up and this is repeated for all the sets. The frame that has maximum value of sum implies that, it has more movements of the features on the face and it is identified as the apex frame. The calculated distances is used to form the feature vector which later is given to classifier for recognizing facial expressions.
This paper presents an analysis of 3D facial action units' detection with temporal analysis and recognition of Happy and Sad expression based on the activated AUs. Our objectives are: (1) to analyze AUs properties; (2) to analyze AUs intensity and (3) to analyze the dynamic segment of facial expressions.

2. Action Unit Extraction and Detection

Human faces are characterized by high degrees of variation (Hjelmas et al., 2001). To automatically recognize facial expression, variation across human facial features is an issue and facial expression intensity variation is another concern. The most prolific way to measure the facial expression intensity is to track the 3D facial properties that deform in 4D space. These facial properties should be the significant properties that involves in the deformation of facial expressions. The establishment of the 4D (3D data plus time) database has made it possible to track 3D facial features. One of many publicly available face databases is BU-4DFE database. BU-4DFE database is the extended version of BU-3DFE which showing the collection of facial behavior from static 3D space to a dynamic 3D space.

The most frequently used 3D facial features in 3D face processing are the 3D point, 3D feature distance (Soyel and Demirel, 2007), curvature-based descriptors (Gökberk et al. 2006), local binary pattern (Sandbach et al., 2012), surface normals (Ujir et al., 2014) and facial profile curves and 3D shape analysis (Wang et al., 2006; Maalej et al., 2010). In this work, 3D point, also called as point cloud feature, is used.

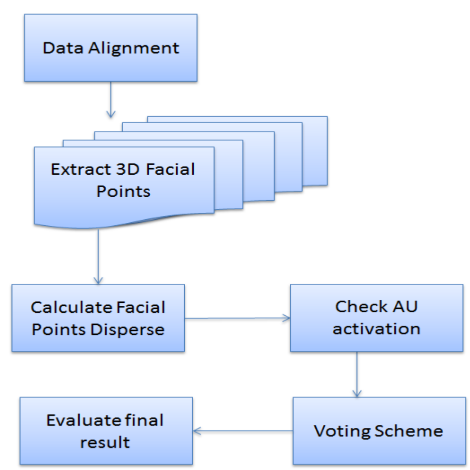

Figure 2 Overview of our work

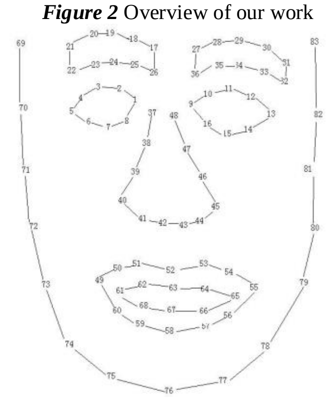

Figure 3 **83**: facial points provided by BU-4DFE database.

Figure 2 shows the workflow of our approach. The initial step is to extract 3D facial points from BU-4DFE database. 83 facial points are provided by BU-4DFE database, refer to figure 3 (Yin et al., 2008) and continue with aligning the facial points. The 3D alignment process followed the work in (Ujir. 2013). Then, the facial distance or angle for all AUs based on the FACS rule is computed. Next step is to check the AU activation status based on the fulfilment of FACS rule. Subsequently, the status of the AUs will be cross-checked with facial expression map where majority voting scheme is used. Finally, the results are corroborated with the ground truth.

An emotion label is quantified according to the assumption that each AU, forming a part of a certain basic expression, has an equal influence on that expression's intensity (Tian et al, 2001). Table 1 shows the combination of AUs that form Happy and Sad expression in different studies, which in our opinion is are due to a different level of facial expression intensity observed. Our work is based on (Ekman and Friesen, 1978). For example, happy expression is only shown if AU6 (cheek raiser) and AU12 (lip corner puller) are activated. While for sad expression, three AUs are important which is AU1 (inner brow raiser), AU4 (brow lowerer) and AU15 (lip corner depressor). However, in this paper, the results of facial mapping using the stated FACS rules are not discussed and we are only looking at the AUs that involve in the expression, the rest of the AUs are not observed.

Table 1 Comparison of facial action units for each expression in different studies

	Happy	**Sad**
Ekman and Friesen (1978)	6+12	1+4+15
Fasel et al. (2005)	6+12+25	1+2+4+15+17
Raouzaiou et al. (2002)	26+12+7+6+20	7+5+12
Deng and Noh (2008)	1+6+12+14	1+4+15+23
Zhang et al. (2008)	6+12	1+15+17
Primary		
Auxiliary	12+6+26+10+23	15+1+4+17+10
Savran et al. (2012)	1+6+12+14	1+ 4+ 15+23
Velusamy et al. (2011)	16+25+26	4+7+25+26

Table 2 Pseudo-codes of facial action unit rules**.**

AU	**AU Definition**	**FACS Rule**
1	Inner Brow Raiser	Prop A: Increased degree of P1, P5 and P26 **AND** Prop B: Increased degree of P9, P13 and P32
2	Outer Brow Raiser	Prop A: Increased degree of P1, P5 and P26 **OR** Prop B: Increased degree of P9, P13 and P32
4	Brow Lowerer	Prop A: P1 and P27 midpoint move downward **OR** Prop B: Distance P17 and P27 midpoint and P26 and P36 midpoint not increased
5	Upper Lid Raiser	Prop A: Increased of P3 **OR** Prop B: P11
6	Cheek Raiser	AU 12 activated.
7	Raised Lower Lid	Prop A: No AU9 **AND** Prop B: A12 activated, **OR** Prop C: Distance (P3 and P7) > 0 **OR** Prop D: Distance (P11 and P15) > 0 **OR** Prop E: decreased of P7 **OR** Prop F: decreased of P15.
9	Nose wrinkler	Distance of midpoint (P17 and P27) midpoint (P26 and P36) is increased
12	Lip Corner Puller	Prop A: Decreased distance (P49 and P1) **OR** Prop B: Decreased distance (P55 and P9) **OR** Prop C: Increased distance (P49 and P9) **OR** Prop D: Increased distance (P55 and P43)
15	Lip Corner Depressor	Prop A: Increased distance (P49 and P1) **OR** Prop B: increased distance (P55 and P9)
16	Lower Lip Depressor	Prop A: P67 downwards **OR** Prop B: P67 outwards **OR** Prop C: decreased distance (P67 and P76)
20	Lip Stretcher	Prop A: Increased distance (P49 and P55) **OR** Prop B: Increased distance ((midpoint P42 and P43) and P58)
23	Lip tightened but not pressed	Prop A: P63 and P67 are not absent **OR** Prop B: Decreased distance (P52 and P58) **OR** Prop C: Distance (P55 and P9) >0 **OR** Prop D: Distance (P49 and P55) not decreased **OR** Prop E: Distance (P49 and P1) not increased **OR** Prop F: Distance (P55 and P9) not increased
26	Jaw Drop	Threshold of distance ((midpoint P42 and P43) and P76)

Table 2 shows the pseudo-codes of facial action unit rules which is based on Pantic and Rothkrantz (2000). According to Valstar and Pantic (2006), out of 44 AUs available, only these AUs have a high relevance for inter-human communication Also, these AUs are sufficient to detect six basic emotions (Fasel et al., 2005). Using these rules, the particular facial features needed in the AUs measurement are calculated.

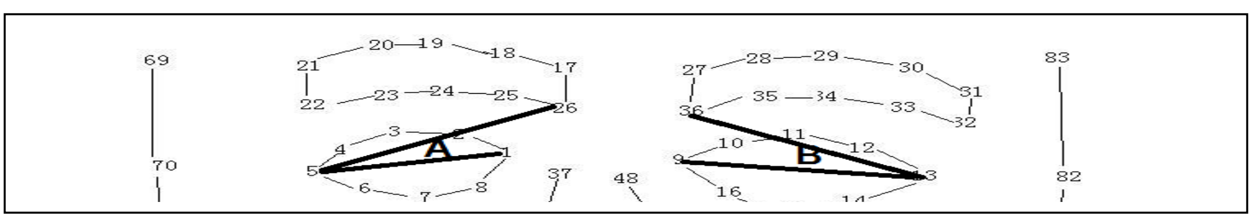

Figure 4 Property A and B for AU1 and AU2.

For certain AUs, the distance and direction between facial features are calculated while for other AUs, the degree between facial points is needed. 3D Euclidean distance is used in this work, as denoted in equation 1.0.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \qquad (1.0)$$

Each AU has its unique set of rules that have to be fulfilled before that particular AU is considered activated. Figure 4 shows two properties, A and B, for action units AU1 and AU2. These properties are referring to the angle between facial points. In another word, for AU1 and AU2, property A is on the right side and property B is on the left side. The difference between both AUs is AU1 is activated if both properties are detected. For AU2, it is considered activated if at least one property is detected. Figure 5 shows the distance between mid-points which is a concern for action unit AU9. First, a midpoint for P17 and P27 as well as a midpoint for P26 and P36 are calculated. Then, the distance between these midpoints is calculated and observed. If the distance between these midpoints is not increased, property B of AU4 is considered activated.

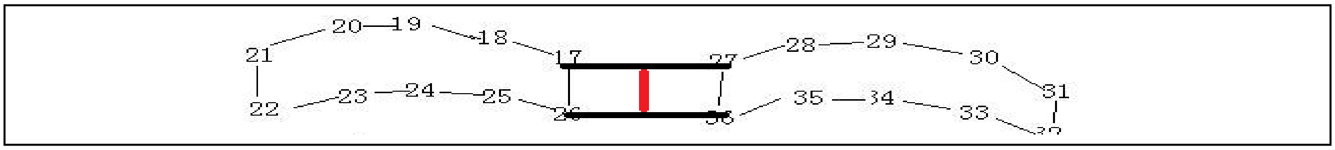

Figure 5 Property A and B for AU9

3. Experiments and Results

From the experiments carried out in this work, the results can be presented as follows:

1.0 Analysis on AUs properties involved in Sad and Happy expression

Figure 6 - 8 shows the comparison of facial properties that involved in deformation for AU1, 4 and AU15 for of sad expression. The *x*-axis is the time and *y*-axis is the intensity of the facial deformation. For figure 6, the red line, which denotes the property B for AU1, shows a rather intense deformation compared to the property A. Based on this, the left features of AU1 deforms a lot compared to the right features.

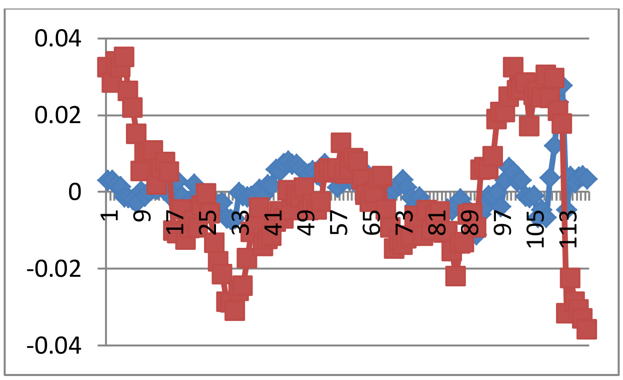

Figure 6 Comparison of AU1 for property A and B of Sad expression

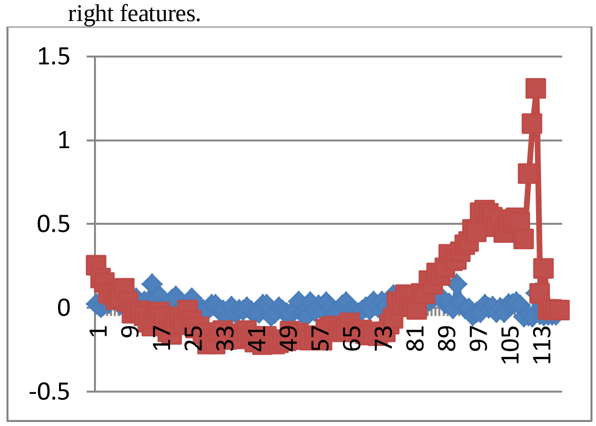

Figure 7 Comparison of AU4 for property A and B of Sad expression

Figure 7 shows the comparison of facial properties that involved in deformation of AU4 for sad expression. Property A, the blue line, has a low intense deformation across the time which in the opposite of property B. Property B for AU4 is referring to the distance between two midpoints as described in Figure 5. From the line chart in figure 7, the distance between the midpoints are getting intense towards the end of the expression.

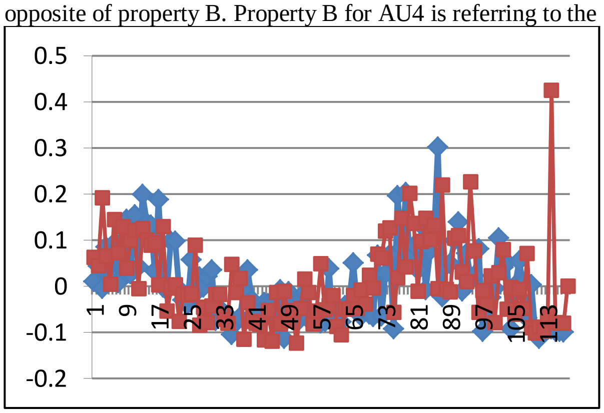

Figure 8 Comparison of AU15 for property A and B of Sad expression

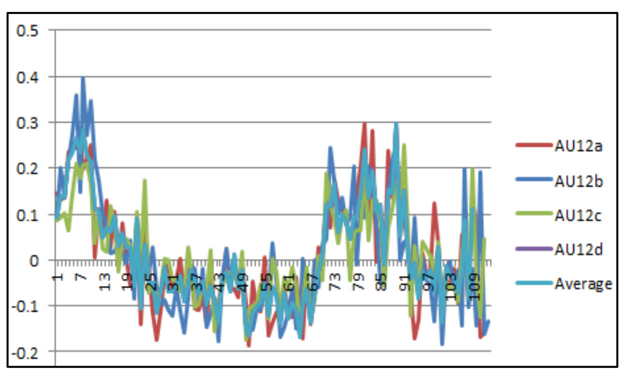

Figure 9 Property comparisons of AU12 for property A, B, C and D

A rather different deformation pattern can be seen for AU15 where both properties show the similar trajectory pattern, refer to figure 8. This means that both sides of lip corner are depressed at the same time, though with a different magnitude.

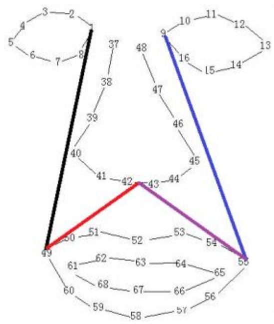

Figure 10 AU12 Properties

There are four properties of AU12 that needs to be observed for happy expression. For property A and B of AU12 to be labelled as activated, the distance of features involved must be decreased. However for property C and D, the distance must be increased. Figure 9 shows the properties pattern comparison for AU12. The four properties of AU12 is concerning the deformation of mouth, eyes and nose facial points, refer to figure 10.The trajectory for all properties across the time is consistent which imply all facial features involved deform at the same time when the subjects show happy expression.

2.0 Intensity Comparisons between AUs for Happy and Sad Facial Expression

Figure 11 shows the graph for AU involved in sad expression across the time. As described in Table 1, according to Ekman and Friesen (1978), only AU1, AU4 and AU15 are involved in the sad expression. From the graph, the intensity for AU15 (Lip Corner Depressor) as well as AU4 (Brow Lowerer) is significant compared to AU1 (Inner Brow Raiser). This tells that AU1 is much harder to detect in sad expression.

For happy expression, only two AUs are involved which is AU6 (Cheek Raiser) and AU12 (Lip Corner Puller). AU6 is activated when AU12 is activated. As described in figure 8, the similar pattern is found for AU12 in average with all AU12 properties. All properties intensified and weaken simultaneously implied that happy expression is easier to detect.

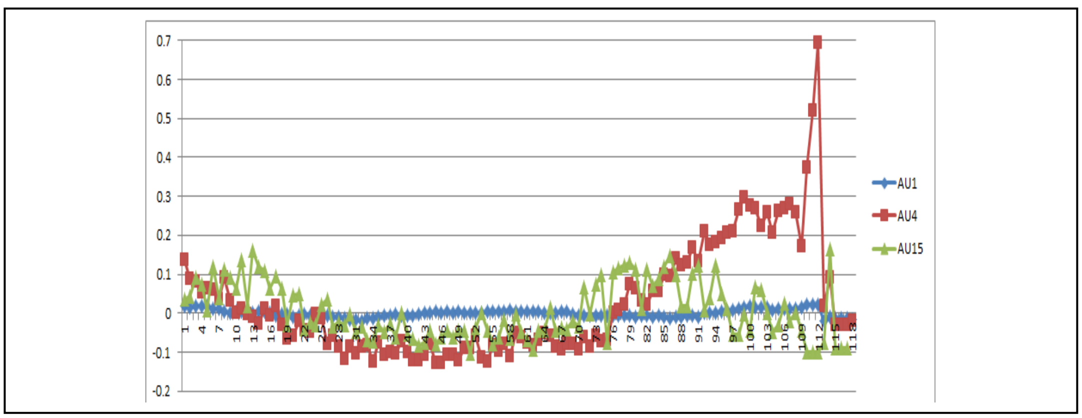

Figure 11 All AU properties across the time for sad expression

3.0 Dynamic Analysis for AUs

In human communication, the timing of a display is an important aspect of its meaning (Lien et al., 2000). A facial expression episode contains the onset time, offset time and time at an apex. Activated AUs are observed to detect the expression temporal segment, i.e. onset-apex-offset. This AUs activation depends on the facial muscular activity. The intensity level of facial expression is increasing from onset time to apex time. The offset time is the time from the first evidence of fading expression until it stops fading) (Hess and Kleck, 1990). For apex, the detection is quite easy as all of the rules for that particular AU activation are satisfied. Time at the apex is the amount of time the expression is held at the peak (Hess and Kleck, 1990). Not only we are able to detect the apex segment, we also successfully record the duration of the apex segment (from i^{th} frame till $i + m^{th}$, where $m < n$, n is the total number of frames). However, for onset time, the facial features that we observed should progress over t (i.e. decreased or increased according to the rules). As opposed to the onset, for offset, the facial features should be in the opposite progress.

Following Suja et al. (2015), we measured the deformation of facial features involved in AUs for every frame. Figure 12 shows the temporal segment for sad expression in average. All AUs involved in sad expression is observed to identify the temporal segment. Figure 13 shows the temporal segment for happy expression in average.

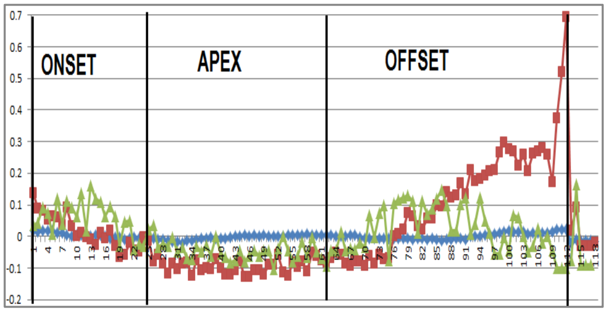

Figure 12 The temporal segment for sad expression

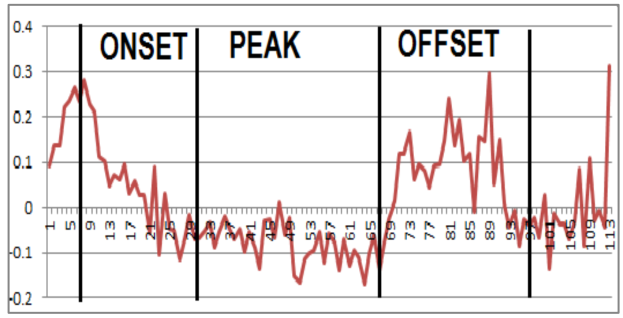

Figure 13 The temporal segment for happy expression

Table 3 shows the temporal analysis for four AUs for the selected expressions. Therefore, the temporal analysis for AU6 is not carried out as it only depends on AU12 activation. The onset time and apex time are quantitatively the same.

Table 3 The estimated average duration for onset, apex and offset for Happy and Sad expression (in milliseconds).

Expression/Temporal Segment	Onset	Apex	Offset
Happy	0.23	0.36	0.31
Sad	0.25	0.36	0.50

4. Conclusion

Any sign of the 3D facial features deformation with an ordered facial features' trajectory shall serve as a signal to the observer. This is the cue of emotional changes which will bring benefit to psychological studies as well as a start on a development of an emotional detection augmented reality tool.

The proposed approach encodes and quantifies 4 AUS which is sufficient to recognize happy and sad expression. The rule-based algorithm developed is based on the FACS rule introduced by Ekman and Friesen (1978). This work focused on the analysis of AUs which involved in happy and sad expression. The trajectory and relation between properties in AUs as well as the temporal segment of the expression episode is observed and discussed.

In our future work, we will be looking at the rest of basic expressions (i.e. Angry, Disgust, Fear and Surprise). We believe spatial variations due to intensity levels of a particular expression have an exclusive trajectory. Therefore, the intensity estimation problem will be looked at. With the findings, we believe this work could be extended to determine spontaneous and posed expressions.

Acknowledgments

This research is fully supported by Malaysian Ministry of Higher Education (MOHE) through Exploratory Research Grant Scheme (ERGS) 1020/2013(17). The authors fully acknowledged MOHE and Universiti Malaysia Sarawak for the approved fund which makes this important research viable and effective.

References

[1] Berretti, S., Bimbo, A.D. and Pala, P. 2012. Real-time Expression Recognition from Dynamic Sequences of 3D Facial Scans. In the Proceedings of the 5th Eurographics Conference on 3D Object Retrieval (EG 3DOR'12), pp. 85-92.

[2] Deng, Z. and Noh, J. 2008. Computer Facial Animation: a Survey in. Data-Driven 3D Facial Animation, London: Springer-Verlag, ch. 1, pp. 1-28.

[3] Ekman, P. and Friesen, W. 1978. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto,

[4] Fasel, L.R. Fortenberry, B. and Movellan, J.R. 2005. A Generative Framework for Real-Time Object Detection and Classification. International Journal of Computer Vision and Image Understanding, Vol.98, No.1, pp. 181—210.

[5] Gökberk, B., İrfanoğlu, M. O. and Akarun, L., 2006. 3D Shape-based Face Representation and Feature Extraction for Face Recognition. Journal of Image and Vision Computing, 24(8), pp. 857–869.

[6] Hess, U. and Kleck, R. 1990. Differentiating Emotion Elicited and Deliberate Emotional Facial Expressions. European Journal of Social Psychology, pp. 369-385.

[7] Hjelmäs, E. and Low, B. K. (2001). Face detection: A Survey. Computer Vision and Image Understanding. pp. 236–274.

[8] Jiang, B. Valstar, M., Martinez, B. and Pantic, M. (2011). A Dynamic Appearance Descriptor Approach to Facial Actions Temporal Modelling. IEEE Transactions on Biometrics Compendium Systems, Man, and Cybernetics, Vol. 44. pp. 161 - 174.

[9] Lien, J.J.-J., Kanadea, T., Cohn, J, F., Li, C. C. 2000. Detection, Tracking, and Classification of Action Units in Facial Expression. Journal of Robotics and Autonomous Systems, pp. 131–146.

[10] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z. and Matthews, I.,2002. The Extended Cohn-Kanade Dataset (CK+): A Complete Dataset for Action Unit and Emotion-Specified Expression. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 94 – 101.

[11] Maalej, A., Ben Amor, B., Daoudi, M., Srivastava, A. and Berreti, S., 2010. Local 3D Shape Analysis for 3D Facial Expression Recognition. International Conference on Pattern Recognition, pp. 4129 – 4132.

[12] Mahoor, M.H., Cadavid, S., Messinger, D.S.and Cohn, J.F. 2009. A Framework for Automated Measurement of the Intensity of Non-Posed Facial Action Units. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 20-25 June 2009, pp. 74-80.

[13] Pantic, M. and Rothkrantz, L.J.M. 2000. An Expert System for Recognition of Facial Actions and Their Intensity. Proceedings of the Twelfth Innovative Applications of Artificial Intelligence International Conference on Artificial Intelligence (IAAI-00). Austin, USA, pp. 1026 - 1033, August 2000.

[14] Raouzaiou, A., Tsapatsoulis, N., Karpouzis, K. and Kollias, S., 2002. Parameterized Facial Expression Synthesis based on MPEG-4. EURASIP Journal Applied Signal Process. pp. 1021-1038.

[15] Russell, J. A. and Bullock.M., 1986. Fuzzy Concepts and the Perception of Emotion in Facial Expressions. Social Cognition, 4(3), pp. 309 – 341.

[16] Sandbach, G., Zafeiriou, S. and Pantic, M. 2012. Local Normal Binary Patterns for 3D Facial Action Unit Detection. Proceedings of the IEEE International Conference on Image Processing (ICIP 2012). Orlando, FL, USA, October 2012, pp. 1813 – 1816.

[17] Savran, A., Sankur, B., Bilge, M. T., 2012. Comparative Evaluation of 3D versus 2D Modality for Automatic Detection of Facial Action Units. Pattern Recognition, 45(2), pp. 767-782.

[18] Soyel, H. and Demirel, H., 2007. Facial Expression Recognition using 3D Facial Feature Distances. LNCS Book Series, Image and Analysis Recognition, Springer, pp.831 – 838.

[19] Suja, P., Kalyan Kumar V.P. and Tripathi, S. 2015. Dynamic Facial Emotion Recognition from 4D Video Sequences. The Eighth International Conference on Contemporary Computing (IC3), pp. 348-353.

[20] Tian, Y., Kanade, T. and Cohn, J.F. 2001. Recognizing Action Units for Facial Expression Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence.Volume 23(2). pp. 97-115.

[21] Ujir, H. 2013. 3D Facial Expression Classification Using a Statistical Model of Surface Normals and a Modular Approach, Ph.D. Dissertation, University of Birmingham, United Kingdom.

[22] Ujir, H., Spann,M. and Hipiny, I. 2014. 3D Facial Expression Classification Using 3D Facial Surface Normals. The 8th International Conference on Robotic, Vision, Signal Processing & Power Applications (ROVISP), pp. 245-253.

[23] Valstar, M. and Pantic, M. 2006. Fully Automatic Facial Action Unit Detection and Temporal Analysis. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'06), pp. 149-157.

[24] Velusamy, S. Kannan, H. Anand, B. Sharma, A. Navathe, B., 2011. A Method to Infer Emotions from Facial Action Units. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2028 – 2031.

[25] Wang, J., Yin, L., Wei, X., and Sun, Y., 2006. 3D Facial Expression Based on Primitive Surface Feature Distribution. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1399–1406.

[26] Yin, L.,Chen, X., Sun, Y., Worm, T, and Reale, M. 2008. A High-Resolution 3D Dynamic Facial Expression Database. The 8th International Conference on Automatic Face and Gesture Recognition, pp.

[27] Zhang, Y., Ji, Q., Zhu, Z. & Yi, B. 2008. Dynamic Facial Expression Analysis and Synthesis with MPEG-4 Facial Animation Parameters. IEEE Transaction Circuits System Video Technology, 18(10), pp. 1383-1396.