



The Analysis of Facial Feature Deformation using Optical Flow Algorithm

Dayang Nur Zuhijah Awang Jesemi, Hamimah Ujir,
Irwandi Hipiny and Sarah Flora Samson Juan

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

November 7, 2020

The Analysis of Facial Feature Deformation using Optical Flow Algorithm

Dayang Nur Zulhijah Awang Jesemi, Hamimah Ujir, Irwandi Hipiny and
Sarah Flora Samson Juan
Universiti Malaysia Sarawak, Malaysia.

Abstract

Facial features deformed according to the intended facial expression. Specific facial features are associated with specific facial expression, i.e. happy means the deformation of mouth. This paper presents the study of facial feature deformation for each facial expression by using an optical flow algorithm and segmented into three different regions of interest. The deformation of facial features shows the relation between facial the and facial expression. Based on the experiments, the deformations of eye and mouth are significant in all expressions except happy. For happy expression, cheeks and mouths are the significant regions. This work also suggests that different facial features' intensity varies in the way that they contribute to the recognition of the different facial expression intensity. The maximum magnitude across all expressions is shown by the mouth for surprise expression which is 9×10^{-4} . While the minimum magnitude is shown by the mouth for angry expression which is 0.4×10^{-4} .

Keywords: Facial expression, Facial features, Facial deformation and Optical Flow.

1. Introduction

Facial expression is one of the responses to human emotional stimuli. Recognizing facial expression is one of the ongoing works in face processing area of study. Face processing brings benefits to several fields such as security, medicine, and marketing. In the security field, facial expression contributes to detection of the crowd's mood in a controlled environment. For medicine area, facial expression recognition could help to assess the pain and identify the depression state of a patient during treatment. While in marketing, by knowing the current mood of the crowd in public places, certain measurement can be taken to improve the crowd's mood and a different marketing strategy can be offered by the expert. It brings benefits in human centred multimodal human-computer interaction (HCI) where the affective computing in HCI domain employs human emotion to build more flexible and natural multimodal [1].

Facial features are the important features which are needed as the input in the development of a human-machine interaction, specifically for facial based technology. In face processing area, facial features such as eye, nose, mouth, cheeks and eyebrows are related to the facial expression interaction. Facial features will be deformed according to the intended facial expression. By tracking the facial features when a facial expression occurred, the dynamic of the emotion shown can be understood. Although the facial expression changes can be observed by the human eye, the movement of the facial feature is hard to observe. Thus, the intention of this study is to identify which facial features are deformed when these facial expressions occurred. Besides, the level of the facial expression intensity could be identified. Expression can be measured from low to a high level based on the sequence of expression through facial feature deformation. By detecting the movement of facial features, the intensity of the emotional state can be measured.

In this paper, the optical flow algorithm used to track facial feature deformation for six basic facial expressions (angry, disgust, fear, happy, sad, surprise). The intensity of facial expression is also measured. Section 2 describes the research method used in this study and followed by a discussion of results and analysis in Section 3. Finally, the conclusions are drawn.

2. Related Work

Facial expression study attracts researchers in artificial intelligent, computer vision and several other fields of studies, which includes the psychologist. Different researchers used different facial features in face processing studies. [18] proposed 7 most expressive facial regions excluding the eyes and mouth. Facial distance vectors are used in 3D facial expression analysis and classification [19][20][21][25].

In [16], participants were asked to identify expressions from female and male faces displaying six basic universal expressions (anger, disgust, fear, happiness, sadness, and surprise), each with three levels of intensity (low, moderate, and normal). Only two facial regions are considered in [1] which is the eyes and mouth region. One of their findings is the mouth region was the most significant for fear, angry, and disgust expressions and least for surprise. Expression changes models are developed by [2] and they form distinctive quantitative profiles of expressions, known as a Regional Volumetric Difference (RVD) function. Expression classifiers for the four-universal facial expression, anger, fear, happy and sad, are designed by training on RVD functions of expression changes.

In another domain, [3] proposed a framework based on Dynamic Bayesian Network (DBN) to systematically model relationships among spontaneous AUs for measuring their intensities. [4] analyze facial expressions with four intensities in 3D space by exploring the facial distances. The extracted distance measurements provide information for the facial expression intensity measurement as well as the classification of facial expressions. This work reveals that it is not necessary to rely on all facial feature points in estimating facial expression intensity.

According to [5], the decomposition of a face into several modules promotes the learning of a facial local structure and therefore the most discriminative variation of the facial features in each module is emphasized. There are several studies that employed face decomposition in their work [15][22][23][24]. The work in [4] considers three face regions while work in [16] only measured two regions. According to [5], each face region has its own impact level to the six basic expressions and the simplest way to find which modules are affecting facial expression is to find their priority weighting. [5] computed the weighting of each face region using Adaboost. Works in [15] subdivided a face into several modules where each of the modules contain different regions of facial features such as forehead, eyes, eyebrows, cheeks, nose and mouth. From the experiments, the region with the heaviest weight is the eyebrows, followed by mouth, eyes, cheeks, nose and forehead. Several works in face processing used optical flow, and that includes, face tracking [6], facial feature tracking [10] [14], facial expression classification [10] and face recognition [11]. [6] proposed a novel real-time face tracker which utilizes a modified version of the Viola-Jones algorithm for face detection. Forty-five uniform light condition videos from Boston Head Tracking Database are used in the experiment. Their proposed method easily outperformed the Viola-Jones face detector. A head tracker which is a combination of an optical flow and a template-based tracker is proposed in [10]. The tracking is robust as it works well even the subject's face gets partially occluded. [10] also proposed two efficient approaches to track the mouth and eyebrow movements and this has allowed them to track and capture facial expressions. Normally, optical flow is used onto data with motion information. When a real-time face recognition involves varieties of facial expression, optical flow also has been applied. In [11] work, an integrated face recognition system that is robust against facial expressions by combining information from the computed intra person optical flow and the synthesized face image in a probabilistic framework, is proposed.

In the existing work of optical flow in face processing, the study of facial features' intensity with the optical flow has not been carried out yet. With the advantage of optical flow, the tracking of facial features' intensity will help to contribute to the automated recognition of the different facial expression intensity.

3. Our Work

In this study, video data from BU3DFE [7] database is used, which comprises of six basic facial expressions. BU3DFE database provides video of subjects with audio video interleaved (avi.) file format. All 90 videos (for 90 different subjects) display the sequences of facial expression in frontal face view. The facial expression video consists of approximately about 100 frames per second with a length of 4 seconds into 640 x 480 resolutions. The image sequences of BU3DFE database begin with the neutral state to the highest intensity level and then back to the neutral state of expression, refer to figure 1. The intensity of facial feature deformation is measured using optical flow.



Figure 1: Facial Expression Sequences for Happy Expression [7]

Optical flow is the estimation of the apparent motion for an object between an observer and the scene [8]. The optical flow makes the motion measurement between two images frames possible. The method used is gradient based where it is based on local Taylor's series approximation to the image signal. The input for the optical flow is the images of a subject's face with the facial points. Each time the optical flow function is called, two video frames are extracted, which then will be used to measure the object's motion. To construct the optical flow between two images, equation 1, which is the optical flow constrained must be solved.

$$I_x u + I_y v + I_t = 0 \quad (1)$$

where:

I_x , I_y , and I_t are the spatiotemporal image brightness derivatives

u is the horizontal optical flow

v is the vertical optical flow

At the initial phase, the video of facial expression from BU-3DFE database is converted to grayscale format. The conversion of RGB video into a grayscale format is to avoid the redundant pixel while extracting the information. Besides, the conversion into grayscale can simplify a large amount of the calculation in the next process [9]. Optical flow for the video is computed due to its ability to measure the motion between two video frames. Two video frames are used at one time, and one of the video frames is the original state of subject's faces, which is the neutral face and the other one is the non-neutral facial expression with a different level of expression intensity. By concentrating on a region of interest (ROI), help to observe the motion of the facial features. For example, to measure human task performance, ROI of gaze is used [12] [13] [15]. The segmentation process, i.e. segment the face into several ROIs, needs to be done to measure the magnitude of optical flow in order. The video frame is segmented into the 6X4 grid. As a result, there are

three big regions: (1) eyes and eyebrow's region, (2) cheeks' region and (3) mouth region
Figure 2 shows segmentation of subjects in this study.



Figure 2: Segmented frame with optical flow. Image sourced from [7].

The magnitude of each facial feature is measured by computing the displacement of the initial position and previous position for each frame using Eq (2).

$$\bar{v} = \sqrt{(X_i - X)^2 + (Y_i + Y)^2} \quad (2)$$

Finally, the facial expression in the video is analysed. The framework can be shown in the Figure 3.

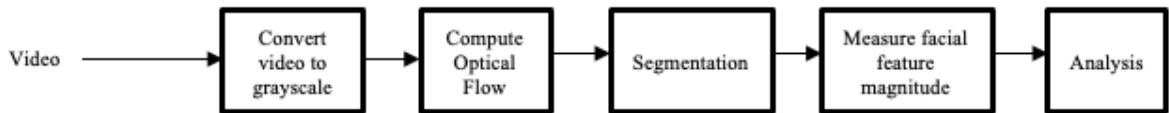


Figure 3: The facial expression analysis framework

4. Results and Analysis

In this section, the intensity of the facial features is analyzed where the magnitude of vector is calculated. The intensity of the facial feature versus sequence of frames is presented.

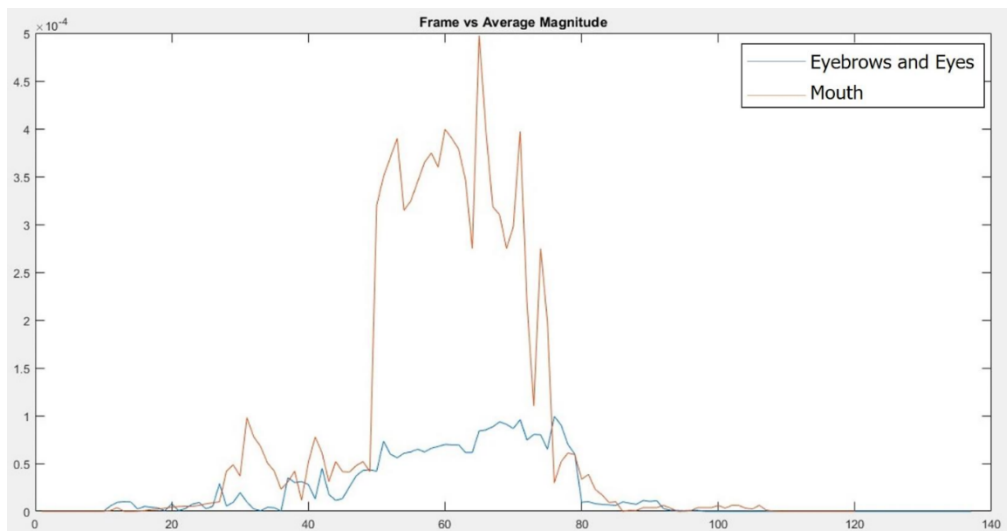


Figure 4: The facial features intensity of sad expression.

Figure 4 shows the sequence of frames versus average of magnitude of optical flow for sad expression. It shows that only eyes and mouth regions are deformed when sad expression occurred. There are no deformation happening in cheek region for sad

expression is shown. In addition, it also shows that eyes region is deformed first before mouth. The intensity of mouth and eyes increasing concurrently. Then, the graph also shows that at certain frames, the intensity of the facial feature will decrease over the time where it means that the facial expression of the subject goes back to neutral state. In line with works in [17], the deformation of facial muscles for Sad expression consists of both eyebrows and lips. However, studies conducted in [4] which is based on facial feature distance could not capture this deformation as the distance for facial features involved in sad expression has a random mean and standard deviation results.

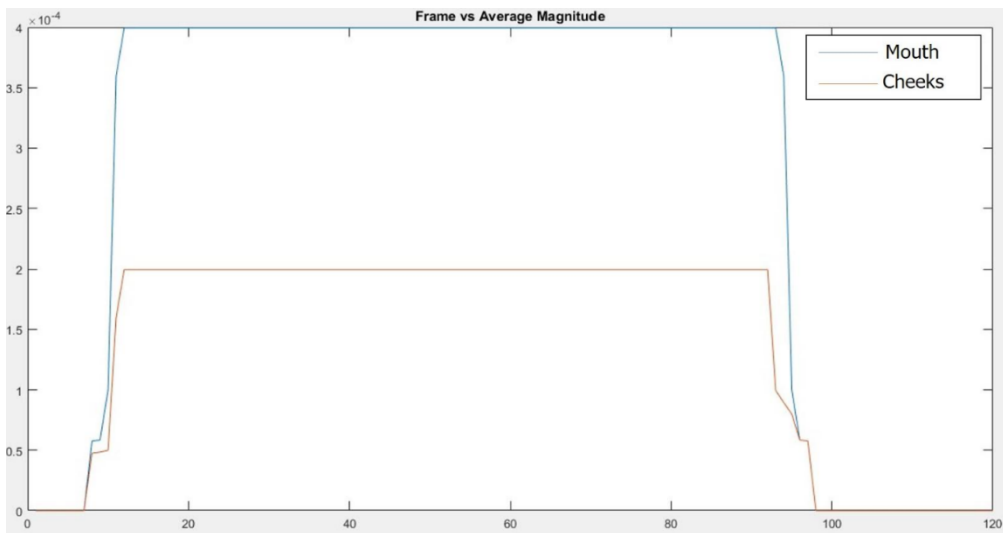


Figure 5: The facial features intensity of happy expression.

Figure 5 shows that only mouth and cheeks involve in happy expression. It also shows that mouth and cheeks segments are concurrently deformed together when happy expression are shown, and it started at frame 8. The mouth has a higher intensity rather than cheek. The intensity of facial features at both regions drop at frame 93.

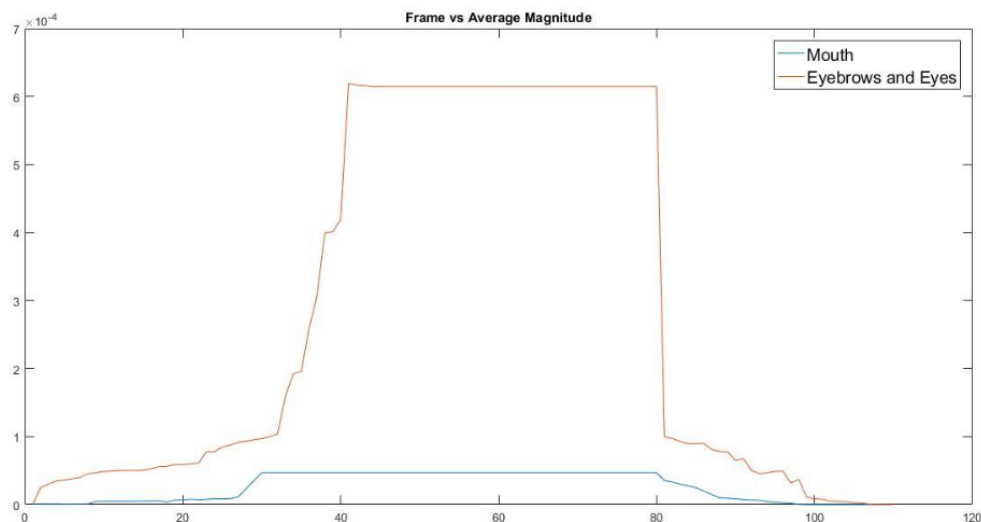


Figure 6: The facial features intensity of angry expression.

Figure 6 shows the facial features intensity of angry expression. From this graph, two regions have significant deformation when angry expression occurred, which are eyes and mouth region. It also shows that the eyes region has a higher intensity rather than mouth.

The intensity of facial features in both regions starts to increase and decrease at an approximately similar time.

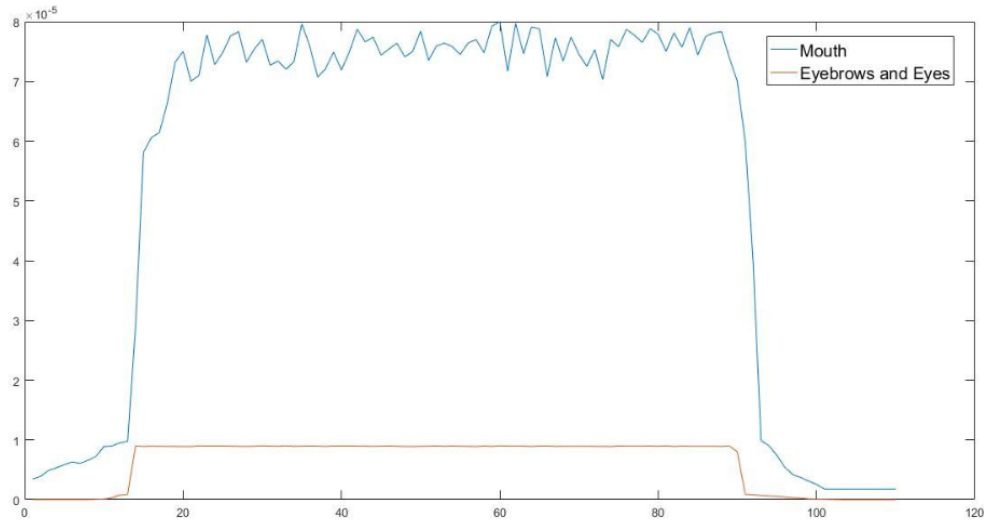


Figure 7: The facial features intensity of disgust expression.

Figure 7 shows the facial features intensity for disgust expression. The mouth region has the highest intensity rather than eyes' region. Based on the graph, it shows that the intensity of the facial features for both regions shows that they increase and decrease at an approximately similar time. However, at the highest intensity, the magnitude for eyes as well as the eyebrows is still compared to the mouth.

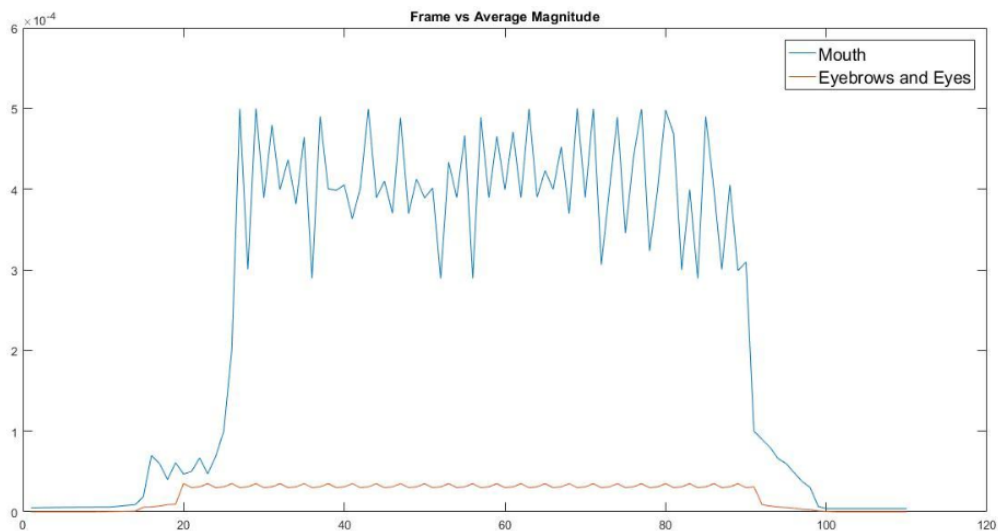


Figure8: The facial features intensity of fear expression

Figure 8 described the fear expression. The above graph shows that the mouth region has the highest intensity rather than the eyes' region. Mouth region deforms at the earliest frame compared with the eyes' region. The fluctuations of the mouth are significant across frames.

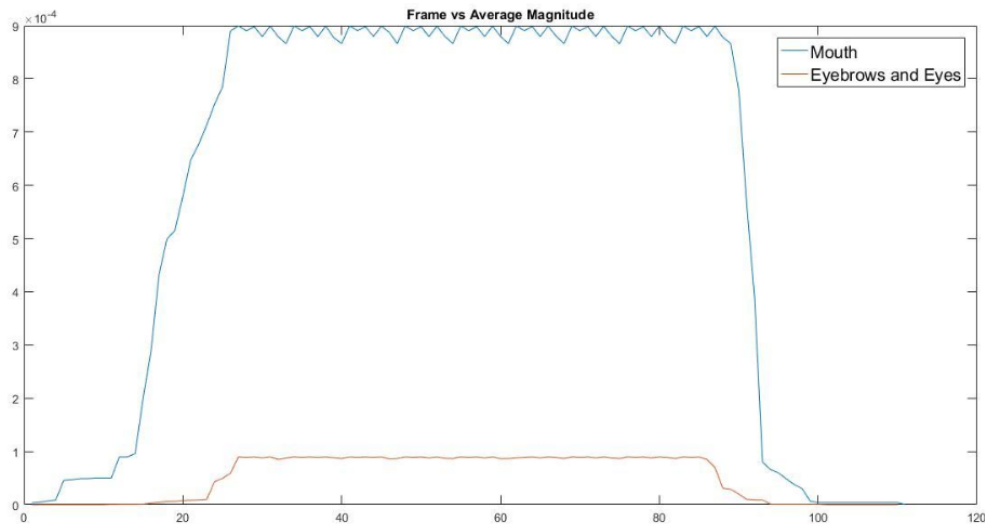


Figure 9: The facial features intensity of surprise expression

Figure 9 shows that the eyes and mouth regions are involved in surprise expression. Besides, it shows that the mouth region is more intense rather than the eyes' region during deformation. Mouth region also deforms first, and the facial features within the eyes' region deform a bit late and return to neutral state much earlier. According to [4], the significant facial muscle deformation for surprise expression is the mouth which is in line with our work.

For all expressions except happy, facial features in eyes and mouth region are significantly deformed when the emotion occurs. For happy expression, the facial features that involved are mouth and cheeks. The intensity within the mouth region is higher compared to the other facial features in all expressions except an angry expression. In addition, the intensity for mouth starts to deform first in angry, fear, surprise and disgust. For sad expression, the eyebrow and eyes deform at the earliest time frame. In the happy expression, both mouth and cheek are deformed together when emotion produced and based on the analysis, it shows that intensities of mouth are higher rather than cheeks.

5. Conclusion

In this study, optical flow is used to measure the motion of the facial expression by using Optical Flow Lucas Kanade algorithm. The magnitude of vector is measured to observe the facial features' intensities. In this study, the results supported the work in [4] that only certain facial features are deformed a particular facial expression is shown. Thus, it is not necessary to rely on all facial feature points in estimating facial expression intensity. However, results in [4] are different from our work.

Our next work is to combine other optical flow method which is Horn Schunk algorithm in our framework. The Horn Schunk algorithm is based on dense optical flow where we expect more interesting results. Also, in our future work, we will further our exploration in the pain assessment based on facial expression.

References

1. Jaimes, A. and Sebe, N. (2007), "Multimodal Human-Computer Interaction: A Survey", in *Computer Vision and Image Understanding*, vol. 108, Issue 1-2, pp. 116 – 134.

2. Alvino C., Kohler C., Barrett F., Gur R.E., Gur R.C., and Verma R. (2007), "Computerized measurement of facial expression of emotions in schizophrenia," *Journal of Neuroscience Methods*, vol. 163, no. 2, pp. 350–361.
3. Li Y., Mavadati S.M., Mahoor M.H. and Ji Q. (2013), "A unified probabilistic framework for measuring the intensity of spontaneous facial action units", *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Shanghai, 2013, pp. 1-7, doi: 10.1109/FG.2013.6553757.
4. Ying, A.C.C., Ujir, H., Hipiny, I. (2017), "3D facial expression intensity measurement analysis". In: Zulikha, J.N.H., Zakaria (eds.) *Proceedings of the 6th International Conference of Computing and Informatics*, pp. 43–48. School of Computing, Sintok
5. Ujir H. (2013), "3D Facial Expression Classification using a Statistical Model of Surface Normals and a Modular Approach", University of Birmingham, 2013.
6. Ranftl A., Alonso-Fernandez F. and Karlsson S. (2015), "Face Tracking Using Optical Flow", *2015 International Conference of the Biometrics Special Interest Group (BIOSIG)*, Darmstadt, 2015, pp. 1-5, doi: 10.1109/BIOSIG.2015.7314604.
7. Yin L., Wei X., Sun Y., Wang J., and Rosato M.J. (2006), "A 3D facial expression database for facial behavior research", *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Southampton, 2006, pp. 211-216, doi: 10.1109/FGR.2006.6.
8. Psaltis, A. (2013), "Optical flow for dynamic facial expression recognition", Master's thesis, 2013.
9. Feroz K.F.C.M. (2014), "Vehicle Tracking and Speed Estimation For Traffic Surveillance", Universiti Tun Hussein Onn, Malaysia.
10. Gast E.R. and Lew M.S. (2011), "A Framework for Real-Time Face and Facial Feature Tracking using Optical Flow Pre-estimation and Template Tracking", *Computing Research Repository (CoRR)*.
11. Hsieh C., Lai S. and Chen Y. (2010), "An Optical Flow-Based Approach to Robust Face Recognition Under Expression Variations", in *IEEE Transactions on Image Processing*, vol. 19, no. 1, pp. 233-240, Jan. 2010, doi: 10.1109/TIP.2009.2031233.
12. Hipiny I. and Ujir H. (2015), "Measuring task performance using gaze regions", *2015 9th International Conference on IT in Asia (CITA)*, Kota Samarahan, 2015, pp. 1-8, doi: 10.1109/CITA.2015.7349836.
13. Hipiny I., Ujir H., Minoi J.L., Juan S.F.S., Khairuddin M.A. and Sunar M.S. (2017), "Unsupervised segmentation of action segments in egocentric videos using gaze", *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Kuching, 2017, pp. 351-356, doi: 10.1109/ICSIPA.2017.8120635.
14. Cui y. and Jin Z., "Facial Feature Points Tracking Based on AAM with Optical Flow Constrained Initialization", *Journal of Pattern Recognition Research* 7 (2012) 72-79.
15. Ujir H., Sing L.C. and Hipiny I. (2014), "A modular approach and voting scheme on 3D face recognition", *2014 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Kuching, 2014, pp. 196-199, doi: 10.1109/ISPACS.2014.7024451.
16. Wells L.J., Gillespie S.M., Rotshtein P. (2016). "Identification of Emotional Facial Expressions: Effects of Expression, Intensity, and Sex on Eye Gaze". *PLoS One*. 2016 Dec 12;11(12):e0168307. doi: 10.1371/journal.pone.0168307. PMID: 27942030; PMCID: PMC5152920.
17. Ekman P. and Friesen W.V. (1978). "Facial Action Coding System: A Technique for the Measurement of Facial Movement", *Consulting Psychologists Press*, Palo Alto.
18. Wang J., Yin L., Wei X. and Sun Y. (2006), "3D Facial Expression Recognition Based on Primitive Surface Feature Distribution", *2006 IEEE Computer Society Conference on*

- Computer Vision and Pattern Recognition (CVPR'06)*, New York, NY, USA, 2006, pp. 1399-1406, doi: 10.1109/CVPR.2006.14.
19. Soyel H., Demirel H. (2007), "Facial Expression Recognition Using 3D Facial Feature Distances", In: Kamel M., Campilho A. (eds) *Image Analysis and Recognition*. ICIAR 2007. Lecture Notes in Computer Science, vol 4633. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-74260-9_74.
 20. Tang H. and Huang T.S. (2008), "3D facial expression recognition based on properties of line segments connecting facial feature points", 2008 8th *IEEE International Conference on Automatic Face & Gesture Recognition*, Amsterdam, 2008, pp. 1-6, doi: 10.1109/AFGR.2008.4813304.
 21. Ujir, H., Hipiny, I., and Awang Iskandar, D.N.F. (2018), "Facial Action Units Analysis using Rule-Based Algorithm", *International Journal of Engineering & Technology*, 7(3.20), 284-290. doi:<http://dx.doi.org/10.14419/ijet.v7i3.20.19167>
 22. Gottumukkal R. and Asari, V.K. (2003), "An Improved Face Recognition Technique based on Modular PCA Approach", *Pattern Recognition Letter*, 25(4), pp. 429 – 436, 2003.
 23. Chiang C.C., Chen X. and Yang C. (2009), "A Component-Based Face Synthesizing Method", *APSIPA Annual Summit and Conference*, pp. 24-30, 2009.
 24. Tena J.R., De la Torre F. and Matthews I. (2011), "Interactive Region-based Linear 3D Face Models", *ACM Transactions Graph*. 30(4): 76, 2011.
 25. Albu F., Hagiiescu D., Vladutu L., Puica M. (2015), "Neural Network Approaches for Children's Emotion Recognition in Intelligent Learning Applications", *Proc. of EDULEARN 2015*, pp. 3229-3239, 2015.