Acquiface Interface as a Device for Acquisition of User's Facial Expressions in Game Expression Hunter

Sugiyanto[1,2]
1st Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
sugiyanto.1907@mhs.its.ac.id
2nd Faculty of Computer Science
Universitas Dian Nuswantoro
Semarang, Indonesia
sugiyanto@dsn.dinus.ac.id

I Ketut Eddy Purnama
Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
ketut@ee.its.ac.id

Eko Mulyanto Yuniarno
Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
ekomulyanto@ee.its.ac.id

Hanny Haryanto
Faculty of Computer Science
Universitas Dian Nuswantoro
Semarang, Indonesia
hanny.haryanto@dsn.dinus.ac.id

Nur Rokhman
Faculty of Computer Science
Universitas Dian Nuswantoro
Semarang, Indonesia
nur.rokhman@dsn.dinus.ac.id

Mauridhi Hery Purnomo
Faculty of Intelligent Electrical and Informatics Technology
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
hery@ee.its.ac.id

***Abstract*—The need test data on facial expression recognition for Indonesian people are not easily obtained and not available in existing databases. Race and ethnic differences can affect the results of facial expression recognition. Testing on facial expressions using untrained data will give more objective results. The purpose of this study is to utilize the Acquiface interface feature in the game Expressions Hunter to get data in the form of images of facial expressions from Indonesian players. Finite State Machine is used to provide behavior to the enemy according to a user action to collect facial expressions by hunting and killing enemies. Acquiface interface is a prototype of one of the features in the game Expression Hunter which functions to record user expressions in accordance with the collection of expressions obtained by users after killing regular enemy monsters or enemy Bosses. RADIATE facial stimulus sets are used as stimuli to generate facial expressions according to the type of expression. The result of recording is a facial expression video file which is then extracted into a sequence of facial expressions. The final image of facial expression is selected from the most expressive sequence of facial expressions (having peak expression).**

***Index Terms*—acquiface, radiate, fsm, expression hunter, expressions**

I. Introduction

Data requirements for facial expression recognition applications can be done in several ways following the required facial expression data criteria. Several public facial expression databases can be used, including FERET, JAFFE, Cohhen-Kanade (CK +), FER2013, OULU-CASIA, Bosphorus, etc. [1], [2]. If the facial expression data needed is more general in nature, it can utilize facial images taken from online video services [3], online television broadcasts [4], social media applications [5], or images from the internet [6]. However, if the facial expression data needed must have specific criteria, then it must take pictures directly by preparing participants who are willing to take facial expressions. These specific criteria, for example, are related to the age of the participant, race or certain ethnic groups.

In the process of collecting data directly on participants who have been prepared, needed a stimulus that can stimulate the participant's facial expressions according to the emotions needed. Several types of media can function as a stimulus to generate facial expressions according to emotions, including music [7], pictures [8], videos [9], or games [10]. The use of images as a stimulus for facial expressions has been done in several studies, including : POFA (The Pictures of Facial Affect) [11], FEEST stimuli (EmotionMegamixes, Caricatures, and Continua) [12], The Geneva Affective Picture Database (GAPED) [13], WSEFEP (Warsaw Set of Emotional Facial Expression Pictures) [8]. In this study, the authors used the image stimulus from the RADIATE (The Racially Diverse Affective Expression) Database [14].

Human emotions consist of complex interactions between subjective feelings, physiological responses, and behaviors

978-1-7281-4533-4/20/$31.00 ©2020 IEEE

that are triggered by external stimuli, which can be monitored from changes in emotional states through investigation of facial expression behavior [15]. Ekman and Izard get evidence that seven facial expressions are universal namely: disgust, anger, fear, sadness, happiness, surprise, and contempt (hate) [16]. However, Jack et al., argues that culture influences the expression of emotions displayed on a person's face [17]. Indonesia has a variety of cultures and ethnicities, this will affect the results of facial expression recognition when compared with the use of existing databases.

To get Indonesian people's facial expressions, a device that can be used to capture facial expressions is needed. Game is a software with a very large number of users in Indonesia. Game users in Indonesia totaled 43.7 million with game sales of 879.7 million dollars [18]. This considers game selection as a tool for facial expression acquisition. Some researchers also use games related to facial expression data, including facial expressions as a game navigation tool [19], [20], facial expression recognition using games [21], and learning facial expressions through games [22]. In this research, Game Expression Hunter using the Aquiface interface that will capture the player's facial expression according to the stimulus provided.

In the game design, an intelligent agent is added to enhance game playing experience with various challenges, thus increasing the player's interest in playing the game [23]. The artificial intelligence model used for scenario behavior in this game is the Finite State Machine (FSM). In addition to control systems, FSM is a model commonly used to design intelligent agent behavior in games that has advantages in the simplicity of computation and ease of understanding and implementation [24]. FSM will recognize and record the actions of characters in the game (players and enemies). The amount of emotions owned by enemies determines the amount of energy. FSM will also determine the reward system that will be obtained by players after killing monsters according to how many collections of facial expressions.

The acquisition interface of the player's facial expressions in the game Expression Hunter will appear after the player kills the enemy. This interface will access the camera device on the computer to be able to record the player's facial expression. Players will get an image of the expression stimulus (from the RADIATE database) displayed on the same screen. This Expression Hunter game was created to get data in the form of images of facial emotion expressions from users, which could later be used as data testing in research activities on identifying Indonesian facial expressions.

The success of players killing enemies will take over the emotions the enemies have. Next, the player must make a recording of facial expressions according to the collection of expressions to get an award. From this action, we will get a recording of the player's facial expression in the form of a video. Then the videos will be extracted into images with 7 expressions: disgust, anger, fear, sad, happy, surprise, and neutral. The purpose of this research is to utilize the Acquiface interface feature in the Expressions Hunter game to get images of Indonesian facial expressions. This Indonesian facial expression image will be used as input for the facial expression recognition project.

II. Literature Review

A. Facial Expression

Facial or facial expression is the result of one or more muscle movements or positions on the face. facial expressions can convey the emotional state of a person to those who observe it. Ekman defines 6 universal facial expressions, including happiness, sadness, anger, fear, surprise, and disgust [25]. According to Matsumoto and Hwang there is strong evidence about the facial expressions of 7 emotions, namely: anger, contempt, disgust, fear, joy, sadness, and surprise [26]. In some facial expression databases using 7 basic emotion references, including JAFFE, CK +, and FER2013. The JAFFE database uses 7 expressions: Neutral, Happiness, Sadness, Surprise, Anger, Fear, and Disgust [27]. The Cohen-Kanade (CK +) facial expression database also uses 7 basic emotions, including Anger, Contempt, Disgust, Fear, Happy, Sadness and Surprise [28]. Whereas the FER2013 database uses emotional expressions: Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Neutral [29].

B. Finite State Machine

Finite State Machine is an abstract machine that can move from one state to another predetermined state [30]. Artificial intelligence using Finite State Machine (FSM) to support interaction with users by using 3 things: State, Event and Action. Finite State Machine will determine the conditions that must be met to change the status. FSM consists of a series of states that determine decisions.

According to Fig.1, each state is represented by a circle. There

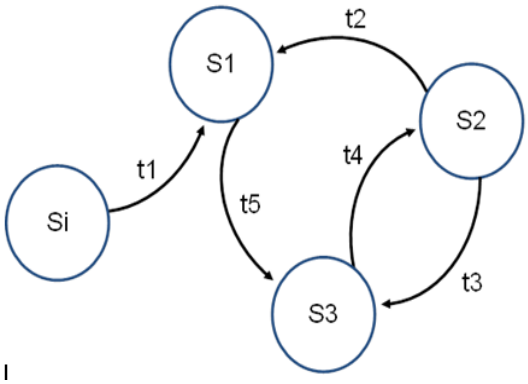


Fig. 1. Finite State Machine [30].

are four states (Si, S1, S2, and S3). Transitions between states are represented by arrows (t1, t2, t3, t4, and t5). FSM starts with Si as the initial state. If condition t1 is met, then Si state will move to state S1. Therefore, every condition (t1, t2, t3, t4, and t5) will result in a state transition. In Si condition, only condition t1 can result in a state transition, in conditions

S2 and S3, there are two conditions that can result in a state transition.

C. RADIATE Database

In the process of recording facial expressions, it takes a stimulus to arouse emotions to produce facial expressions as desired. The RADIATE facial stimulus set has 16 different facial expressions consisting of more than 1,700 photographs, shown in Fig.2. This will provide a variety of emotions in a variety of stimulus circuits that are racially and ethnically diverse. Participants in the RADIATE database were 100 different racial and ethnic models (25% non-Hispanic White and 75% minority or ethnic group) [14].
RADIATE stimulus sets RADIATE overcomes potential limitations in existing stimulus packages, by providing 16 emotional expressions standardized by various racial and ethnic models [14].

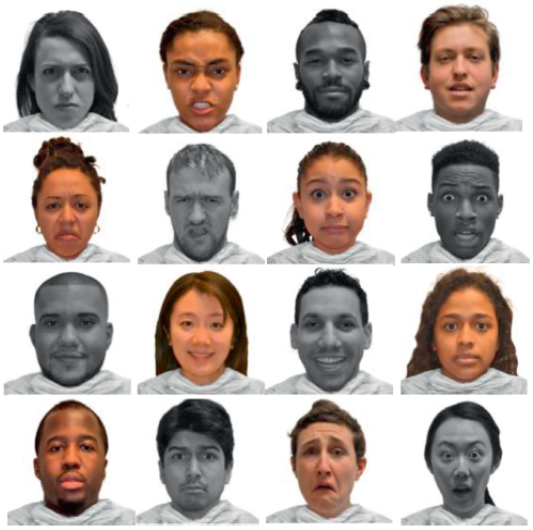


Fig. 2. 16 Expressions that are in the RADIATE Stimulus Set, from the top left : Angry (closed), Angry (open), Calm (closed), Calm (open), Disgust (open), Disgust (closed), Fear (closed), Fear (open), Happy (closed), Happy (open), Happy (exuberant), Neutral (closed), Neutral (open), Sad (closed), Sad (open), Surprise [14].

III. ACQUIFACE INTERFACE

A. The gameplay of Expressions Hunter Game

- The main enemy has a lot of monsters that can take the facial expressions of people found along their journey.
- Each monster can take more than one facial expression categories. Amount of facial expression categories also determines the amount of energy the monster.
- To be able to restore the victim's facial expression, the main character must kill each monster.
- Each time kill a monster, the player will get points and take all the facial expression categories that each monster has, as shown in Fig.3.
- Points earned can be used to upgrade weapons and add lives.
- After getting a facial expression category collection of monsters, players will be asked to record his facial expression by category collections owned.
- Recording the player's facial expressions using the Acquiface Interface, the result is a video file.

Fig. 3. Enemy monsters hold several expressions categories.

B. Finite State Machine

- FSM Regular Monster Enemies
  FSM for regular monster enemy behavior as shown in Fig.4. The behavior starts with a state of silence and changes to attack aggressively if the remaining emotions are still more than or equal to four. If emotions are left less than or equal to two, the enemy will run away. Finally, when emotions run out, the enemy dies.

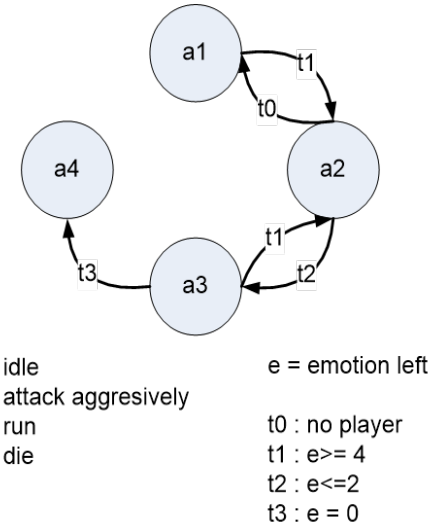
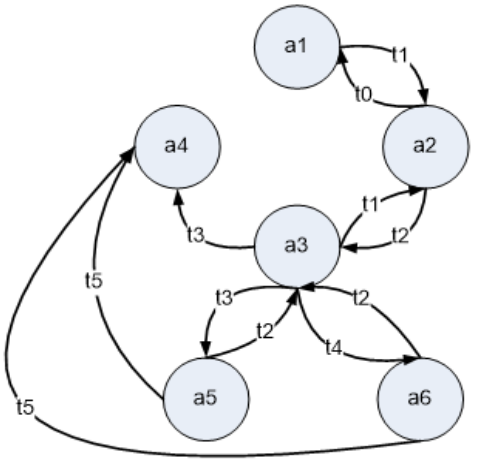


Fig. 4. FSM Regular Monster Enemies.

- FSM Enemy Boss
  The behavior of the enemy boss comes from ordinary enemy behavior with additional behavior, such as being able to use skills and summon other enemy monsters.

Boss enemies also have more emotions stored. The boss's enemy behavior is shown in Fig.5, starting with a state of silence and changing to attack aggressively if emotions are still more than or equal to nine. If emotions are left less than or equal to six then the enemy will run away. Boss enemies can use skills if the remaining emotions are five or more. When the remaining emotions are two, he can summon other regular enemies. Finally, when emotions run out, boss enemies die.



a1 : idle
a2 : attack aggresively
a3 : run
a4 : die
a5 : skill
a6 : call friend

e = emotion left

t0 : no player
t1 : e>= 9
t2 : e<=6
t3 : e<=5
t4 : e<=2
t5 : e=0

Fig. 5. FSM Enemy Boss.

- FSM Rewards
  Players can get rewards if they manage to capture emotions. The rewards depend on the time elapsed until the capture is successful. Fig.6 indicates if the time passes is less than two, the player will get an additional 50% prize if the time passes is less than 3, then the player will get + 20% prize. Players will get an additional 10% prize if the elapsed time is less than 4 and do not get a prize if the elapsed time is more than 5.

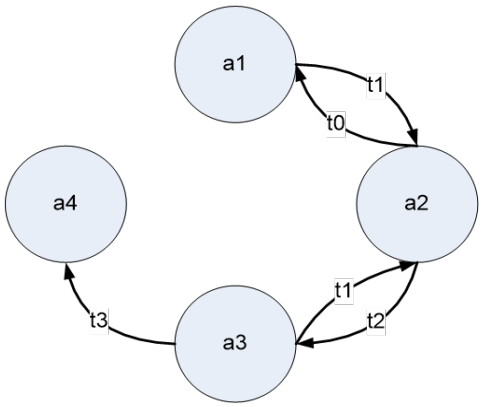
C. Stimulus Set

In the process of recording facial expressions, it takes a stimulus to arouse emotions to produce facial expressions as desired. The stimulus used at the Acquiface interface uses the RADIATE face set stimulus. Stimulus images of facial expressions will be displayed randomly according to the type of expression to be recorded. The position of the stimulus image on the Acquiface Interface is as shown in Fig.7.

D. Acquisition of Facial Expressions

The prototype testing of the Acquiface Interface was carried out with 40 participants aged 18-24 years (age mean = 19.90, SD = 12.50,and 10 of them women). Face expression acquisition will be done after the player can kill regular enemy monsters or Boss enemies. As shown in Fig.8, After the player kills enemy monsters, the player will get points and a collection of facial expressions owned by enemy monsters will switch to the player. Then the player will be asked to record his facial expression using the Acquiface Interface, which functions to record the player's facial expression according to each category of facial expression obtained. There are 7 types of expressions used, including Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise.

On the interface, there is a window that displays the picture taken by the camera on the device used by the user, the



a1 : no reward
a2 : +50% reward
a3 : +20% reward
a4 : +10% reward

t0 : time elapsed > 5
t1 : time elapsed < 2
t2 : time elapsed < 3
t3 : time elapsed < 4

Fig. 6. FSM Rewards.

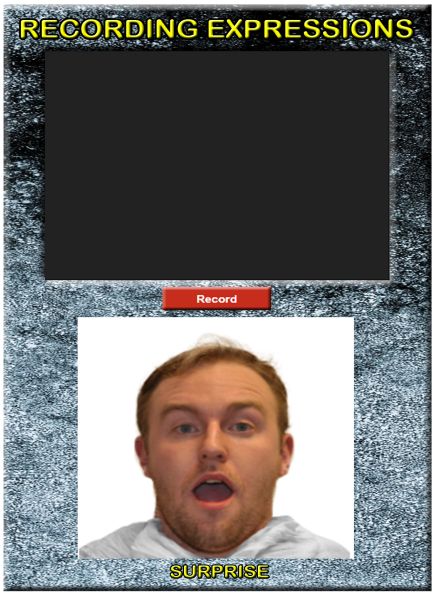


Fig. 7. Acquiface Interface

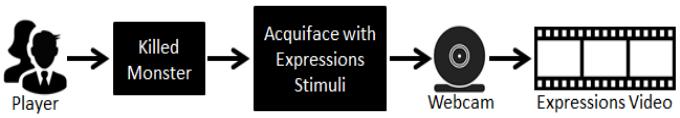


Fig. 8. Facial Expressions Acquisition Proses

Record button to record, and the image of facial expression stimulus taken from the RADIATE database. The stimulus image will be displayed randomly according to the type of player's expression to be recorded. The process of recording each type of expression will be done with a time duration of 5 seconds. The results of recording this expression will be saved as a video file of type. Webm. When playing this Game Expression Hunter, a player can produce more than one video for each emotion on facial expressions. This possibility occurs because the player can kill more than one monster, and each monster can store the same collection of expressions as other monsters. For this reason, the best facial expression video was chosen by selecting several facial expression videos with the same emotions on each player.

E. Facial Expression Video Extraction

After getting the results in the form of videos of player's facial expressions, the next step is extracting facial expression videos as shown in Fig.9. The process of extracting facial expression videos is done separately after getting the videos from Game Expression Hunter. In this extraction process, the facial expression video files are converted into sequences of picture frames. Then, from each sequence of the picture frames, the top frame will be selected manually by selecting the most expressive picture frame according to each emotion. The same thing will be done if there are facial expressions with the same emotions in a player. So that an image will be generated for each facial expression emotion on each player.

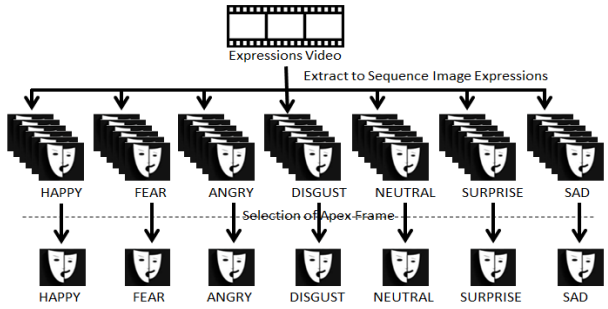


Fig. 9. Facial Expression Video Extraction to Facial Expression Image

F. Validation of Facial Expressions Images

The image files will then be used for the validation process. To validate facial expression images, performed using samples consisting of 30 raters (18 men and 12 women), ages between 17-45 years (M = 23.67, SD = 7.18). Raters labeled 280 facial expressions with 7 types of randomly displayed expressions. Labeling each facial expression image is done by marking the Labeling Form by rating them according to 7 types of expressions: Angry, Fear, Happy, Disgust, Neutral, Sad, and Surprise. At the time of the assessment, each facial expression image was displayed for 3 seconds using a projector connected to the computer. Meanwhile, to provide an assessment, raters were given 10 seconds.

IV. RESULT

The Acquiface interface prototype test was conducted with 40 participants, using 7 expressions, 280 video images of facial expressions were obtained. Extraction of each video is done to get a sequence of images of facial expressions. From the sequence of images of each expression, the most expressive image (peak expression) will be chosen according to the desired emotional expression. The results of video extraction of facial expressions to facial expressions images as shown in Fig.10, stored in a face expression image using the RGB color model with a .jpg file type.

Fig. 10. Extracted images according to the type of expression, starting from the top left : Happy, Fear, Neutral, Surprise Sad, Angry, dan Disgust.

Raters assessment results as shown in TABLE I, the data shows the percentage of raters who successfully labeled the facial expression images that have been displayed. The average suitability of the labeling with the overall facial expression displayed is 92.10% (SD = 0.027). While the results for each type of expression are: Angry (M = 92.33%, SD = 0.037), Fear (M = 92.25%, SD = 0.039), Happy (M = 94.92%, SD = 0.026), Disgust (M = 85.58% , SD = 0.052), Neutral (M = 96.08%, SD = 0.023), Sad (M = 91.83%, SD = 0.038), and Surprise (M = 91.67%, SD = 0.032).
Testing using One-Way ANOVA conducted to compare with the results of the assessment of participants. The result was significant F (7, 210) = 25.22, p ¡ 0.001. The post-hoc test showed that the level of recognition of all types of expressions did not differ, and the F-value was greater than the F-critical value for the alpha value = (0.05).

V. CONCLUSIONS AND FUTURE WORK

The prototype interface works well when testing by producing videos of the player's facial expressions according to the stimuli displayed. Each player can produce video facial expressions of more than one emotion because the monsters killed by the player allow having the same collection

TABLE I
ACCURACY OF IDENTIFICATION OF FACIAL EXPRESSIONS

Facial Expressions						
Angry	***Fear***	***Happy***	***Disgust***	***Normal***	***Sad***	***Surprise***
0.923 (M)	*0.923 (M)*	*0.949 (M)*	*0.856 (M)*	*0.961 (M)*	*0.918 (M)*	*0.917 (M)*
0.037 (SD)	*0.039 (SD)*	*0.026 (SD)*	*0.052 (SD)*	*0.023 (SD)*	*0.038 (SD)*	*0.032 (SD)*

of emotional expressions. From the results of the facial expression video extraction, it produces 280 images of facial expressions according to 7 basic emotions: Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise. The results of the validation process by labeling all images (280 images) of facial expressions produced by the Acquiface prototype get an accuracy rate of 92.10% (SD = 0.027). Labeling process errors occur for several reasons, including: poor lighting, change the orientation of the player's face, and the distance a player's face with the camera. Testing using One-Way ANOVA gave significant results F (7, 210) = 25.22, p ¡ 0.001, and the F-value are greater than the F-critical value for alpha = (0.05).

It can be concluded that the facial expression image produced by the Acquiface prototype can be used as an input image in the facial expression recognition application project. The next work is preparing Game Expression Hunter into an expression simulation game that can be used to build a database of Indonesian facial expressions.

ACKNOWLEDGMENT

We would like to offer our special thanks to BPPDN Scholarship Program of the Ministry of Research, Technology and Higher Education of Indonesia 2019.

REFERENCES

[1] R. Weber, C. Soladié, and R. Seguier, "A survey on databases for facial expression analysis," 01 2018.

[2] R. Weber, J. Li, C. Soladié, and R. Seguier, *A Survey on Databases of Facial Macro-expression and Micro-expression*, pp. 298–325. 07 2019.

[3] P. Hus'ak, J. Cech, and J. Matas, "Spotting facial micro-expressions " in the wild "," 2017.

[4] K. Zhang, E. Barati, E. Rashedi, and X.-W. Chen, "Long-term face tracking in the wild using deep learning," 08 2016.

[5] P. Indrawan, S. Budiyatno, N. Ridho, and R. Sari, "Face recognition for social media with mobile cloud computing," *International Journal on Cloud Computing: Services and Architecture*, vol. 3, pp. 23–35, 02 2013.

[6] R. R. Herrera and F. J. G. Funes, "Face detection for identification of people in images of internet," *International Journal of Computational Engineering Research*, vol. 04, pp. 9–17, 11 2014.

[7] B. Geethanjal, K. Adalarasu, and J. Mohan, "Music induced emotion and music processing in the brain– a review," *Journal of Clinical and Diagnostic Research*, vol. 12, 01 2018.

[8] M. Olszanowski, G. Pochwatko, K. Kuklinski, M. Scibor-Rylski, P. Lewinski, and R. Ohme, "Warsaw set of emotional facial expression pictures: A validation study of facial display photographs," *Frontiers in Psychology*, vol. 5, 12 2014.

[9] C. Holland, N. Ebner, T. Lin, and G. Samanez-Larkin, "Emotion identification across adulthood using the dynamic faces database of emotional expressions in younger, middle aged, and older adults," *Cognition and Emotion*, vol. 33, pp. 1–13, 03 2018.

[10] T.-W. Tsai and M.-Y. Lin, "An application of interactive game for facial expression of the autisms," pp. 204–211, 09 2011.

[11] P. Ekman and W. V. Friesen, *Pictures of facial affect*. Consulting Psychologists Press, 1976.

[12] Y. AW, D. Perrett, A. Calder, R. Sprengelmeyer, and P. Ekman, "Facial expressions of emotion: Stimuli and tests (feest)," *Thames Valley Test Company (TVTC)*, 01 2002.

[13] E. Dan-Glauser and K. Scherer, "The geneva affective picture database (gaped): A new 730-picture database focusing on valence and normative significance," *Behavior research methods*, vol. 43, pp. 468–77, 03 2011.

[14] M. I. Conley, D. V. Dellarco, E. Rubien-Thomas, A. O. Cohen, A. Cervera, N. Tottenham, and B. J. Casey, "The racially diverse affective expression (radiate) face stimulus set," *Psychiatry research*, vol. 270, pp. 1059–1067, 2018.

[15] R. E. Jack and P. G. Schyns, "The human face as a dynamic tool for social communication," *Current Biology*, vol. 25, pp. R621–R634, 2015.

[16] L. J. David Matsumoto, *Culture and Psychology*. Cengage Learning, 2007.

[17] R. Jack, O. Garrod, H. Yu, R. Caldara, and P. Schyns, "Facial expressions of emotion are not culturally universal," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, pp. 7241–4, 04 2012.

[18] newzoo.com, "The indonesian gamer 2017," 2017.

[19] M. N. Ilmi and A. Chowanda, "Navigation key using face expression in endless game," 03 2018.

[20] N. Müller, B. Eska, R. Schäffer, S. Völkel, M. Braun, G. Wiegand, and F. Alt, "Arch'n'smile: A jump'n'run game using facial expression recognition control for entertaining children during car journeys," pp. 335–339, 11 2018.

[21] M. T. Akbar, M. N. Ilmi, I. V. Rumayar, J. Moniaga, T.-K. Chen, and A. Chowanda, "Enhancing game experience with facial expression recognition as dynamic balancing," *Procedia Computer Science*, vol. 157, pp. 388 – 395, 2019. The 4th International Conference on Computer Science and Computational Intelligence (ICCSCI 2019) : Enabling Collaboration to Escalate Impact of Research Results for Society.

[22] K. Gülkesen, F. Isleyen, B. Cinemre, M. Samur, S. Kaya, and N. Zayim, "A web-based game for teaching facial expressions to schizophrenic patients," *Applied Clinical Informatics*, vol. 8, pp. 719–730, 07 2017.

[23] H. Haryanto, "Agen cerdas kompetitif berbasis finite state machine dalam game pembelajaran untuk anak," *Techno.com*, vol. 9, 2010.

[24] Y. P. Anung Rachman, Vincent Suhartono, "Agen cerdas animasi wajah untuk game tebak kata," *Jurnal Teknologi Informasi)*, vol. 6, 4 2010.

[25] P. Ekman, "Universal facial expressions of emotion," *California Mental Health Reasearch Digest)*, vol. 4, Autum 1970.

[26] D. Matsumoto and H. S. Hwang, "Reading facial expressions of emotion," *Psychological Science Agenda*, vol. 25, no. 5, 2011.

[27] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," vol. 1998, pp. 200 – 205, 05 1998.

[28] P. Lucey, J. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression," pp. 94 – 101, 07 2010.

[29] I. Goodfellow, D. Erhan, P. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, Y. Zhou, C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor, M. Milakov, J. Park, and Y. Bengio, "Challenges in representation learning: A report on three machine learning contests," *Neural Networks*, vol. 64, 07 2013.

[30] G. S. David Bourg, *AI For Game Developers*. O'Reilly Media, Inc., 2002.