Emotion Expression of Three Dimensional Face Model Using Naive Bayes and Fuzzy Logic

Muhammad Arief[1], Mitra Istitiar Wardhana[1,2],Surya Sumpeno[1], Mochamad Hariadi[1]
[1] Electrical Engineering Department, Institut Teknologi Sepuluh Nopember, Surabaya- Indonesia
[1,2] Game Technology Department, Universitas Negeri Malang, Indonesia

Summary
Various emotions such as happy, sad, angry, surprised, and disgust have been known for a long time and become an important aspect of human behavior. However, the application of emotion has not been used in human and computer interaction. Therefore a good system in human interaction and computer should be able to recognize, interpret and process human emotions. Research in the field of emotion is a complex process because it can change dynamically. Because of that, the study of text-based emotion mostly done due to text forms is relatively simple compared to other forms such as visual or sound. Various methods has been used to make the process of recognition of human emotions, such as SVM (Support Vector Machine), VSM (Vector Space Model), and Cauchy Naive Bayes. As for the process to produce emotional responses, method that have been used is Fuzzy Logic. In this research will be discussed visualization of emotional expression on the face of three-dimensional model using Naive Bayes and Fuzzy logic. A combination of both methods can produce emotional expressions of an Indonesian-language text. Results of Naive Bayes text classification accuracy for yield 64.83%. While the results of a Fuzzy Logic facial parameter values are dynamic, so they can show facial expressions that contain more than one type of emotion. From the distribution of questionnaires to obtain accurate values for facial expression 66.4%
Keywords :
Text classification, Naive Bayes, Fuzzy Logic, 3D Face Model

1. Introduction

Human interface and computer interface has been researched for a long time. Nowadays, many researchers give more attention to non-verbal information recognition [1]. Various type of features which usually used in emotional processing such as : voices intonation, face expressions, body posture, body movement, heartbeart, and blood pressure. Researches believe that emotional processing technology will be an important part in artificial intelegence, especially in communication among people. Although interaction between human and computer is different with interaction among people, some theories show that human interaction and computer has a similar basic character with interaction among people[2] .

Essentially, there are three types of media which usually use in emotional research such as : text, voice, and image. Text was choosed due to the emotions is a complex process because it changes dynamically. There are a lot of study about text-based emotion because the forms of text is relatively simple compared to other forms such as visual or sound. Although text is the simplest media, text has an important role in communication because text is not contain only description from information but also human behavior especially emotion. Text can be obtained from various resources i.e. : books, newspapers, websites, magazines, e-books, emails,etc. With all the advantages, text become the most popular media and contains lots of emotions.

There are a lot of method that can be used to process human emotion, i.e. : SVM (*Support Vector Machine*)[1] , VSM (*Vector Space Model*)[3], dan *Cauchy Naive Bayes[4]*, These methods can be used to recognize emotions in the text and the method to produce emotional response is Fuzzy Logic . In [5] Naive Bayes and Fuzzy logic is used to generate a response from the NPC's attributes change . In [6] naïve bayes is used to determine emotion in the text and show the emotional expression in two dimensional donkey face but the expression is crisp. From the result of previous research can be concluded that the existed technology are able to recognize emotions in the text. The emotion recognition process can not be separated from text classification process because classification process is used to analyze type of emotions on the text.

In this study, was discussed emotional processing from Indonesian text. Using text classification, the probability values can be obtained from the text for each type of emotions and the values are use to produce expression in 3D face model. This study used two methods namely Naive Bayes and Fuzzy Logic, Naive Bayes is method of text classification that used to identify the types of emotions while Fuzzy logic is used to process the results of text classification. The results of the Fuzzy Logic are the parameters of the model three-dimensional face. The purpose of the use of Fuzzy Logic is to obtain parameter values that contained more than one type of emotion.

Manuscript received May 5, 2010
Manuscript revised May 20, 2010

The advantages of using naïve bayes as a classifier are very easy to deal with missing attributes and fast to train. Despite the benefit of the naïve bayes, the weakness are the pitfalls usually made are due to a poor understanding of the central assumption behind naïve bayes, namely conditional independence.

2. Theoretical Consideration

2.1 Emotion

This study was used 5 types of emotion ,i.e : happy, sad, angry,afraid, and disgust which was obtained from ISSEAR research [3] and the data were obtained from expressions in English which represents a type of emotion then translated to Indonesian.

The resulting emotions which conducted in this study were facial expressions, thus should be clearly defined description of facial shape that contains a particular emotion. In [7] defined a textual description of several types of basic emotions of neutral, happy, sad, angry, scared and disgust.

2.2 Naive Bayes Classification [8]

Naive bayes theorem is one of guided approach method, derived from Bayes theorem. Bayes theorem itself is a fundamental statistical approach in pattern recognition. The probability model for a classifier is a conditional model

$$p(C|F_1,\ldots,F_n) \quad (1)$$

Over a dependent class variable C with a small number of outcomes or *classes*, conditional on several feature variables F_1 through F_n. The problem is that if the number of features n is large or when a feature can take on a large number of values, then basing such a model on probability tables is infeasible. We therefore reformulate the model to make it more tractable.

$$p(C|F_1,\ldots,F_n) = \frac{p(C)\,p(F_1,\ldots,F_n|C)}{p(F_1,\ldots,F_n)} \quad (2)$$

In practice we are only interested in the numerator of that fraction, since the denominator does not depend on C and the values of the features F_i are given, so that the denominator is effectively constant. The numerator is equivalent to the joint probability model

$$p(C,F_1,\ldots,F_n) \quad (3)$$

Which can be rewritten as follows, using repeated applications of the definition of conditional probability:

$$p(C)\,p(F_1,\ldots,F_n|C) \quad (4)$$

Now the "naïve" independence assumptions come into play: assume that each feature F_i is conditionally independent of every other feature F_j for $j \neq 1$. This means that

$$p(F_i|C,F_j) = p(F_i|C) \quad (5)$$

and so the joint model can be expressed as

$$\begin{aligned} p(C,F_1\ \ldots,F_n) &= p(C)\,p(F_1|C)\,p(F_2|C)\,p(F_3|C)\ \ldots \\ &= p(C)\prod_{i=i}^{n} p(F_i|C) \end{aligned} \quad (6)$$

2.3 Fuzzy Logic [9]

Fuzzy logic is a method for dealing with uncertainty. The feinition of uncertainty is a problem that has doubts, inaccuracies, lack of full information, and the value of truth is partial. The first process is fuzzification, to change the input that is certainly to be a form of fuzzy. In the fuzzification process, linguistic variable is defined from input values which the semantic determined based on membership functions. Membership functions, a way to show distribution membership membership degree alues from fuzzy logic input, in fuzzification process must be determined from input. Membership functions will be represented as a function graph. Then the following process is inference, in this process fuzzy rules are conducted. The process ended with defuzzification, in order to change fuzzy output output to be exact value.

3. Proposed Method

The flow of this research divided into three stages: Learning, Training, Fuzzy Logic process.Moreover the flowchart of this research can be seen in Figure 1.In order to understand the flow of the research, the variables in this study must be defined first and the variabels as defined as follows :

a. If document is a set of finite terms so that, document D = {w1,w2,….,wn)
b. Vocabulary V must be built as the list of all distinct words that appear in all the documents of the training set.
c. The words in the vocabulary becomes the attributes W_k, assuming that classification is independent of the positions of the word.
d. Let C be the number of targeted class, class is defined as type of emotion.
e. Let n_j be the total number of all distinct words in training document category C_j
f. Let N_k be the number of times Wk occurs in category C_j
g. Let $docs_j$ be the number of training documents in category C_j and $|examples|$ be the number of document in the training set of labeled documents.

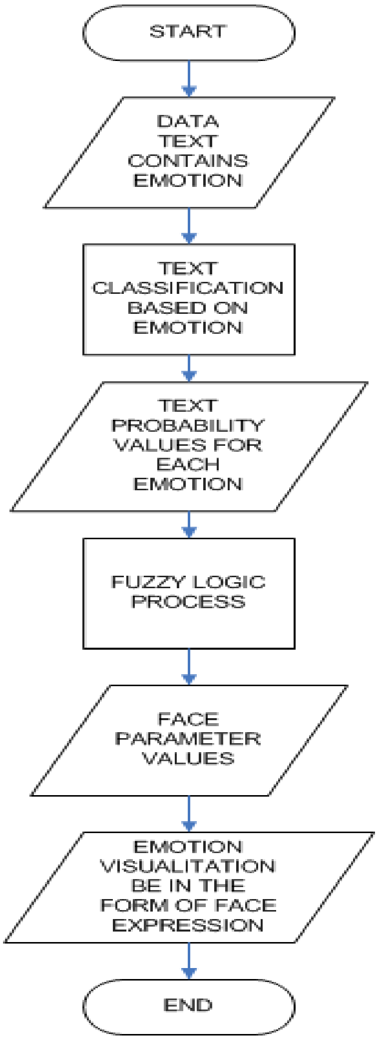


Figure 1 Flowchart

3.1 Learning

Learning was conducted to learn from a set of labeled documents wih predefined class in order to allow the learning machine to classify the emotion with obtain the probability values for each class in the newly encountered document D. The flow of learning process as defined as follows:

a. Collect vocabulary
b. For each class $P(C_j)$, the probability is

$$P(C_j) = \frac{|docs_j|}{|examples|} \quad (7)$$

c. For each attributes W_k,the probability is

$$P(W_k|C_j) = \frac{(Nk+1)}{(nj+|Text|)} \quad (8)$$

Where ,

$|Text|$ is a single documents generated by concenating all he training documents for the category C_j.

d. So that,

$$P(C_j/D) = \frac{P(Cj).P(D|Cj)}{P(D)} \; j = 1,2,\ldots,c \quad (9)$$

When comparing posterior probability for the same document *D*, *P(D)* is same for all categories thus will not affect the comparison.

e. The emotion in the text is determined based on the largest posterior probability and the equation is

$$AC(D) = argmax\ C_j \{P(C_j|D)\ j=1,2,c\} \quad (10)$$

Assumptions were needed in order to reduce the computation and the assumpation were :

f. The probability of encountering a specific word within document is the same regardless the word position.

$$P(D|C_j) = P(\{w_1, \ldots, w_n\}|C_j) \quad (11)$$

g. The probability of occurent of a word is independent of the occurrence of other words in the same document.

3.2 Training

Training was conducted to measure the classifier performance and cross – validation was used to check the performance. In cross validation, a fixed number of document is reserved for testing as if they were unlabelled documents and the remainder are used for training (as labeled documents,this process is repeates until the highest accurate value is obtained. The parameter which used in cross validation are class precision, and calss recall. The steps of cross validation is described below :

a. Let X be the entire data set of N documents
c =5 is the number of different categories
b. Fix the size s of the training set for to perform cross validation.
c. Set the number of trials T. If s=N-1, fix the number of trials T=N; else, T=40.
d. For trial r=1 to T

i. Select randomly s documents from X as labeled documents into training set X_T^l .
ii. Store the remaining documents (X-X_T^l) as unlabeled documents into X_T^u (as if they were unlabeled).
iii. Train NB using X_T^l . . (Compute Equation (7) and Equation (11).
iv. Use trained NB to compute the class of each element in X_T^u using Equation (11)

The performance of classification process depends on the data which used in training process. There are several methods to obtain the good data ,i.e:

a. Token
Is a process to break down long text to be a small part which called sentences
b. Stopword removal
Stopword means sentences which are not affect classification process. Th classification process that conducted in this study is emotion classification so the words such as : “yang”, “akan”,”di sana”, “apakah”,etc can be removed. In example a sentence “tidak senang” is different with “senang”.

3.2.1 Training Data Text

In order to recognize emotion which contained in the text, training dataset is needed so the system can learn. This research used 5 type of basic emotion ,such as : happy, sad ,angry, ddisgust, and afraid, based on ISEAR (*International Survey on Emotion Antecedents and Reaction*) (3) dataset. This study used 200 training dataset from ISEAR dataset but the data must be translated to Indonesian.

3.3 Fuzzy Logic process

Fuzzy logic process was used to handle the input values from classification process which contained 5 types of emotion. The output from classification process are used to be input values for fuzzy logic in order to make face parameter values for 3D model face.With Using logic process, the 3D model can show expression that contains more than one eomotion. Variabels linguistic were divided into three types based on probability values :

a. Low : 0 – 0.4
b. Medium : 0.2 – 0.8
c. High : 0.6 - 1

In early process, fuzzification,the five input values from the classification process were mapped to be linguistic variabelsthen the linguistic variabels were used to decide membership degree in fuzzification set. Type of membership function that used in fuzzification process is trapezoid

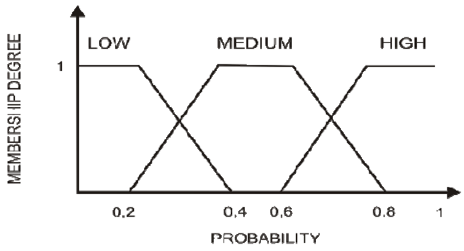


Figure 2 Membership Funtion In Fuzzification

The next process is inference. System inference value is value that given by system for each emotion and done based on fuzzy rules. IF section in fuzzy rule was explained emotion condition and THEN section was explained 3D face parameter. General form of fuzzy rule as defined as follows :

IF Emotion1 = A AND
Emotion2 = B AND
Emotion3 = C AND
Emotion4 = D AND
Emotion5 = E AND

THEN Parameter is F

Emotion1 - emotion5 are type of emotion that is used, A – E are variabels linguistic from each emotion, and F is 3D face parameter value. The number of fuzzy rule can be determined based on the number of input and linguistic variable, which is in this study are 5 for input and 3 for linguistic variable therefore the number of fuzzy rules is 3^5 = 243.

But all the fuzzy rules must be checked to qualify the requirement because the number of input values are never more than one. The input comes from classification process which is the total of probability values are never more than one then to fix this problem is changing every linguistic variables with the smallest value from the member which has membership value equal to one so that in variable linguistic “low” the value is changed to 0 (zero) and so on.

Before we made fuzzy rules for inference process, the seven membership functions had to made. The total of membership functions are correspond with the number of 3D face parameter that must be modified. . Each face parameter has upper limit and lower limit which will be the number of member in membership function.

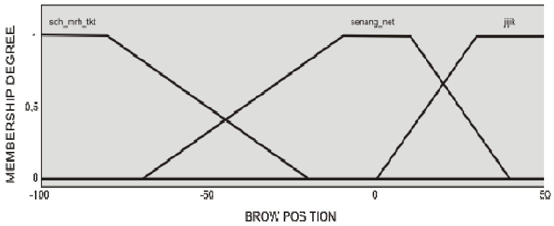


Figure 3 Membership Function For Brow Position

The last process in fuzzy logic is defuzzification,in this process the values in inference system is concenating to be a single value that to be used in 3D model face. The defuzzificaion use center of gravity method as defined as follows :

$$y* = \frac{\Sigma y\mu_R(y)}{\Sigma y\mu_R(y)}, \qquad (12)$$

Where,

y = Fuzzy output and $y\mu_R(y)$ is membership degree from y.

4.1.2 Three Dimension Face Model Parameter

The Face model was taken from [10]. Before entering to fuzzy logic process, the values of face parameter must be defined for each types of emotion and the basic expression for each type of emotion can be seen below:

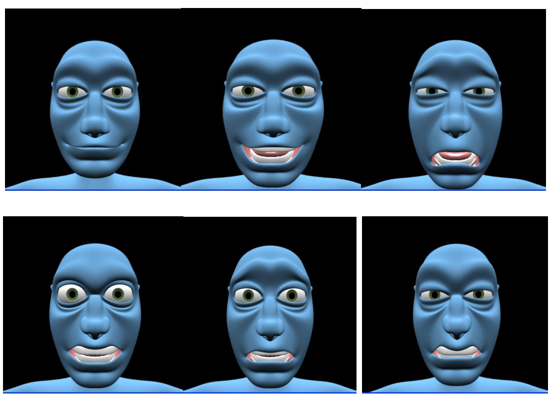


Figure 4 Basic Expression of Fave Model for every type of Emotion.Top position (left – right): neutral,happy, sad Bottom position (left – right) : angry, afraid, disgust

4. Experiment Result and Discussion

In this section will discussed experiment which is to produce face expression based on Indonesian text and the evaluation will be divided into three stages:evaluation of classification performance, evaluation of face expression experiment , and evaluation of questionnaire result.

4.1 Evaluation of Text Classification

Performance of classification prosess can be measured based on several parameters such as : class precision, class recall, and accuracy. The evaluation graph of each emotion class can be seen below :

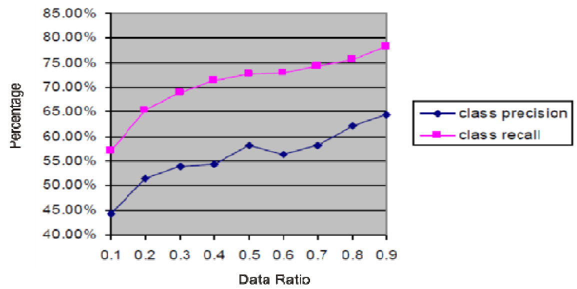


Figure 5 Result of Text Classification in Happy Class

The highest class precision and class recall value for happy class are obtained with 0.9 data ratio , the values are 64.38 % and 78.33 %.

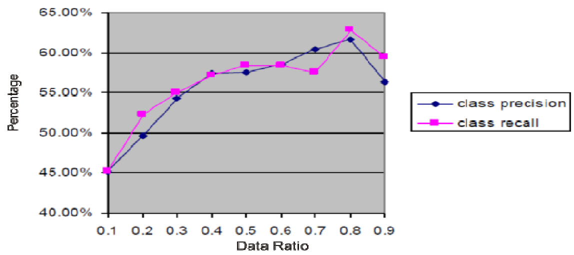


Figure 6 Result of Text Classification in Sad Class

The highest class precision and class recall value for sad class are obtained with 0.8 data ratio , the values are 61.58 % and 62.78 %.

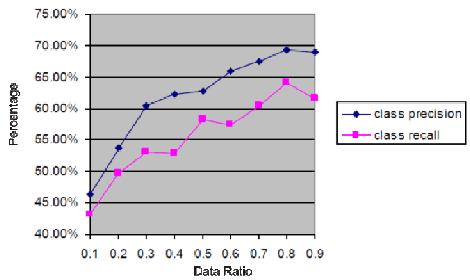


Figure 7 Result of Text Classification in Angry Class

The highest class precision and class recall value for angry class are obtained with 0.8 data ratio , the values are 69.37 % and 61.67 %.

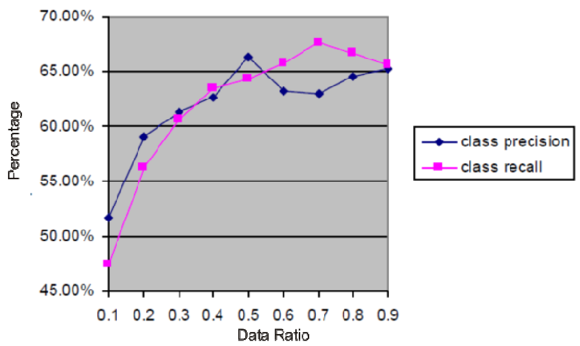


Figure 8 Result of Text Classification in Affraid Class

The highest class precision value for afraid class is obtained with 0.5 data ratio thus the value is 66.25 % and the highest class recall value is obtained with 0.7 data ratio thus the value is 67.59 %.

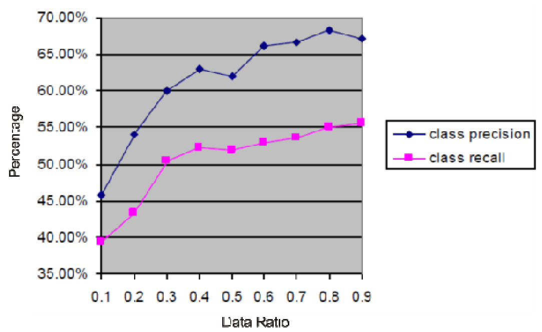


Figure 9 Result of Text Classification in Disgust Class

The highest class precision value for disgust class is obtained with 0.8 data ratio thus the value is 68.28% and the highest class recall value is obtained with 0.9 data ratio thus the value is 55.56 %.

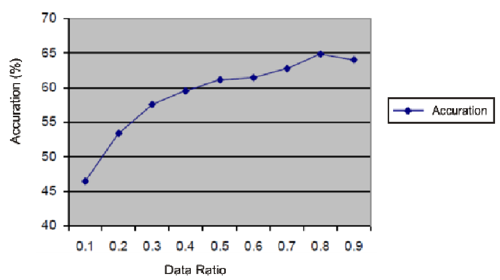


Figure 10 Result of Text Classification in Overall Class

The accuration of overall class rose from 0.1 data ratio until 0.8 data ratio but in the 0.9 data ratio , the accuration level decreased. The highest value reached at 0.8 data ratio with 64.83.

4.2 Evaluation of Face Expression

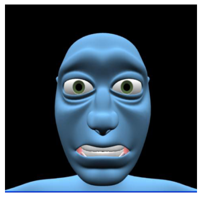
In this experiment, we used 5 data from ISEAR but the datas were not used as training dataset. Therefore, the datas were not recognizable by the system and the result of experiment can be seen below :

1st Experiment

Text1 : ketika saya berhasil mempertahankan hubungan saya dengan seorang gadis

Probability Values

Happy : 0.5208962169
Angry : 0.3022667838
Sad : 0.09009898541
Disgust : 0.06507883748
Afraid : 0.02165917644

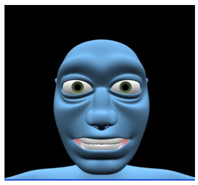


2nd Experiment

Text 2 : ketika saya bertengkar dengan teman dekat

Probability Values

Happy : 0.1707853052
Angry : 0.2744905029
Sad : 0.3215331076
Disgust : 0.213093341
Afraid : 0.02009774331

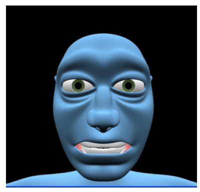


3rd Experiment

Text 3 : ketika teman dekat berbohong pada saya

Probability Values

Happy : 0.267174978
Angry : 0.4354774695
Sad : 0.1192149594
Disgust : 0.1461661016
Afraid : 0.03196649149

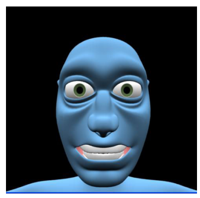


4th Experiment

Text 4 : ketika saya jatuh dan kaki saya patah

Probability Values

Happy : 0.1307485149
Angry : 0.2670615728
Sad : 0.04036394041
Disgust : 0.0685675256
Afraid : 0.4932584463

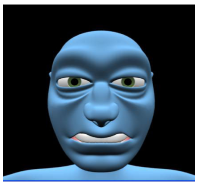


5th Experiment

Text 5 : anak kecil yang kencing sembarangan di depan umum

Probability Values

Happy : 0.008850187021
Angry : 0.08833650855
Sad : 0.009824129096
Disgust : 0.6849140535
Afraid : 0.2080751219



5. Conclusion and Future Work

5.1 Conclusion

Based on the experimental result can be inferred :

a. Using Naive Bayes as text classification method, the accuracy of the system to recognize emotions in the text is 64,83 %
b. Using Naïve Bayes and Fuzzy Logic can produces forming parameter of fave expression which affected with more than one type of emotion.

5.2Future Work

a. Add types of emotions into six types : happy, sad, angry,afraid, disgust, and shock
b. Add more text dataset to improve accuracy of text classification
c. Improve quality of expression visualitation with changing the face and add animation in every change of emotion
d. Add feature extraction such as : TDF – IF to give weight for a word and use word pairs in vocabulary to preserve the utterances.

References

[1] Chuang, Ze-Jing and Wu, Chung - Hsien. Multi Modal Emotion Recognition from Speech and Text. Computational Linguistic and Chinese Language Processing. Vol. 9, 2.
[2] Reeves, B. and C, Nass. The Media Equation : How People Treat Computers, Televison and New Media Like Real People and Places. s.l. : Cambridge University Press, 1996.
[3] Feeler : Emotion Classification of Text Using Vector Space Model. Taner, Danisman and Adil, Alpkocak. 2008.
[4] Emotion Recogniton Using a Cauchy Naive Bayes. Nicu, Sebe and S., Lew Michael. 2002.
[5] Wardhana, Istiar Mitra, Sumpeno, Surya and Hariadi, Mochamad. Artificial Intelegence in Game To Response the emotion From Indonesian Text Using Text Classification and Fuzzy Logic.Natioanl Electric, Informatic, and its Education Conference. 2009.
[6] http://www.convo.co.uk/x02/.
[7] Tekalp, M. Face and 2-D Mesh Animation in MPEG - 4. Tutorial Issue On The MPEG - 4 Standart. 1999.
[8] Wu, Xindong and Kumar, Vipin, [ed.]. The Top Ten Algorithms in Data Mining. s.l. : CRC Press.
[9] Jang, J.Roger, Chuen-Tsai and Mizutani, Eiji. Neuro - Fuzzy and Soft Computing. s.l. : Prentice - Hall Inc, 1997.
[10] www.blendermags.org.

Muhammad Arief received his bachelor degree in Electrical Engineering from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia in 2009. During studies, he was highly involved in student union acivities. He was a vice president of National Electronic – Game Maker 2008.

Mitra Istiar Wardhana received his bachelor degree in computer science from school of informatics and computer science Malang, Indonesia. He received his mster degree in game technology from of Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, in 2010.Now he is a lecture in State Univesity of Malang.

Surya Sumpeno received his bachelor degree in Electrical Engineering from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia in 1995. He received M.Sc. degrees in Graduate School of Information Sciences Tohoku University Japan, in 2007.Nowadays he is pursuing the Ph.D. degree at the Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia and he is the staff of Electrical Engineering Deparment of Institut Teknologi Sepuluh Nopember (ITS) Indonesia.

Mochamad Hariadi received his B.E.degree in Electrical Engineering Department of Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, in 1995. He received both M.Sc. and Ph. D. degrees in Graduate School of Information Sciences Tohoku University Japan, in 2003 and 2006 respectively. Currently, he is the staff of Electrical Engineering Deparment of Institut Teknologi Sepuluh Nopember (ITS) Indonesia. He is the project leader in joint research with PREDICT JICA project Japan and WINDS project Japan. His research interest is in Video and Image Processing, Data Mining and Intelligent System. He is a member of IEEE, and member of IEICE.