Feature-Adaptive Motion Energy Analysis for Facial Expression Recognition

Sungkyu Noh, Hanhoon Park, Yoonjong Jin, and Jong-Il Park

Department of Electrical and Computer Engineering, Hanyang University,
17 Haengdang-dong, Seongdong-gu, Seoul, Korea
{snoh,hanuni,lakobe8}@mr.hanyang.ac.kr, jipark@hanyang.ac.kr

Abstract. In this paper, we present a facial expression recognition method using feature-adaptive motion energy analysis. Our method is simplicity-oriented and avoids complicated face model representations or computationally expensive algorithms to estimate facial motions. Instead, the proposed method uses a simplified action-based face model to reduce the computational complexity of the entire facial expression analysis and recognition process. Feature-adaptive motion energy analysis estimates facial motions in a cost-effective manner by assigning more computational complexity on selected discriminative facial features. Facial motion intensity and orientation evaluation are then performed accordingly. Both facial motion intensity and orientation evaluation are based on simple calculations by exploiting a few motion energy values in the difference image, or optimizing the characteristics of feature-adaptive facial feature regions. For facial expression classification, a computationally inexpensive decision tree is used since the information gain heuristics of ID3 decision tree forces the classification to be done with minimal Boolean comparisons. The feasibility of the proposed method is shown through the experimental results as the proposed method recognized every facial expression in the JAFFE database by up to 75% with very low computational complexity.

Keywords: facial expression, feature-adaptive motion energy analysis, decision tree.

1 Introduction

Facial expression is an important element in human communication and studies have shown that facial expression reveals the underlying emotions [4]. However, it has always been a great challenge to build a machine that recognizes human facial expressions effectively and reliably. The main difficulty of automated facial expression recognition is that complex human facial features are represented with 'limited' verbal descriptions. We say 'limited' because the known verbal descriptions may not describe every little detail perceived by the human visual system. The best known facial expression analyzer i.e. human visual system is trained with tremendous amounts of data over a substantial time with the best known parallel learning system (i.e. human neurons and their network), and it may be safe to say that artificially

G. Bebis et al. (Eds.): ISVC 2007, Part I, LNCS 4841, pp. 452–463, 2007.

reproducing the complex facial feature representation used by human visual system is near impossible.

Despite the difficulties of automated facial expression recognition, most previous research has heavily relied on complex face models or complicated algorithms. Complicated face models for representing various facial motions have advantages for recognition accuracy but increase computational expense. FACS [3], the most widely used face model, has 46 action units that are generated by facial muscle movements. By combining 46 action units, FACS based face model [5, 2, 15, 19] can represent detailed 2D texture information of facial expressions. 3D modeling of the human face suggests that pose-invariant recognition can be achieved by using multiple instances of 2D texture information. There are mainly two ways to construct a 3D geometric face model from 2D face texture data: one is using a single image [17] and the other is using multiple images [12, 7, 18]. Constructing the 3D model from a single image faces restrictions such that the image has enough 2D texture data (i.e. frontal or limited pose-variants) to construct the 3D model. Constructing the 3D model from multiple images relieves such restrictions; however, multiple views are not always available. Even if multiple views are available, lack of matching features result in inaccurate synthesis of the 3D model. Moreover, considering many action units (FACS) or 3D vertices (3D model) inevitably increases the computational complexity of whole facial expression analysis and recognition process. In environments with limited computational capability, complex model and complicated methods are not suitable.

On the other hand, extremely simplicity-oriented face representation namely point-wise face model [13] has point-wise areas that loosely-fit rectangular regions around six facial features (each eye, each eyebrow, nose, and mouth). Of course, point-wise face model reduces the computational complexity dramatically by simply analyzing the motion intensity of point-wise areas; such a method is more suitable for relieving the burden of computational expenses. However, this method represents very limited, only intensity-based facial motions by neglecting the analysis of motion orientations of facial features. For the motion orientation of facial features, previous efforts used the tracking-based or template-based method such as optical flow [5], HLAC features and fisher weight map for the motion vector [14]. However, they tend to be computationally expensive by computing motion vector on each feature. The computation of the motion vector requires preprocessing, cost functions for motion displacements, etc. that tend to be costly. More reviews on the state of art in the field of facial expression recognition can be found in [11, 6].

Recently, there have been increased demands on intelligent mobile or embedded systems with vision intelligence such as a cell phone recognizing characters, a robot recognizing facial expressions, etc. Accordingly, such intelligent systems have limited computational capacity compared to conventional PCs. This paper addresses the issue of reducing the computational complexity of automated facial expression analysis while recognizing various facial motions; we present feature-adaptive motion energy analysis for recognizing facial expressions.

Unlike previous efforts, our method is simplicity-oriented and yet cost-effective in terms of speed and accuracy. For the cost-effective model, we introduce a new action-based face model: (1) it has reduced 'action units,' and (2) it groups facial features that share the same 'action units' and similar appearances with symmetry into a *facial unit*.

Feature-adaptive motion energy analysis estimates facial motions in a cost-effective manner by first detecting the *adaptive facial feature regions* (preprocessing). The detected facial feature regions tightly surround the features adaptively in a rectangular shape; therefore, they have more accurate size and location, but smaller region than point-wise areas. Next, *discriminative facial features* are selected in order to minimize the computational expenses for the facial motion estimation. Motion intensity or orientation analysis is then performed adaptively on the detected facial feature regions. Both motion intensity and orientation analysis are based on simple calculations by exploiting few motion energy values within the detected feature region or optimizing the characteristics of the detected feature region. Therefore, a very small computational complexity increase is achieved for the facial motion estimation. After each feature's action is analyzed and known, the corresponding *facial unit*'s action is determined and a *facial motion descriptor* is generated for the classification. Computationally inexpensive ID3 decision tree classifies the facial expression with minimal Boolean comparisons using the *facial motion descriptor*.

The remainder of this paper will be presented in the following order. In Section 2, our action-based face model is introduced. Then, Section 3 presents our overall framework for facial expression analysis and recognition method. Experiments and discussions are covered in Section 4, and Section 5 draws the conclusion of this paper.

2 Action-Based Face Model

Table 1 shows our action-based face model, which is based on FACS [3], and the verbal descriptions of facial expressions from DataFace [4, 20].

Table 1. Action-Based Face Model

Facial Unit	Facial Feature	Action State	
1. Forehead	Forehead	Neutral	Expressed
2. Eyebrow	Left Eyebrow	Neutral	Expressed
			- Down - Up
	Right Eyebrow	Neutral	Expressed
			- Down - Up
3. Glabella	Glabella	Neutral	Expressed
4. Eye	Left Eye	Neutral	Expressed
			- Widened - Narrowed
	Right Eye	Neutral	Expressed
			- Widened - Narrowed
5. Outer Eye Corners	Left Eye Corner	Neutral	Expressed
	Right Eye Corner	Neutral	Expressed
6. Nose	Nose	Neutral	Expressed
7. Cheeks	Left Cheek	Neutral	Expressed
	Right Cheek	Neutral	Expressed

Table 1. (*continued*)

8. Mouth	Mouth	Neutral	Expressed
			- Thinned - Corners Up - Corners Down - Upper Lip Drawn Up - Open Wide
9. Chin	Chin	Neutral	Expressed

A facial unit is defined as a group of a single unique facial feature or two facial features that shares same action state and similar appearances with symmetry. Facial features in the same facial unit shares same action states, because they tend to move together toward same direction when expressed. By defining a facial feature as a subset of a facial unit, our classifier (ID3 tree) considers fewer attributes and thus the reduction in computational complexity of the classification is achieved (details in Section 3). Each facial feature has two default action states, neutral and expressed state; the expressed state is further categorized into sub-action states accordingly depending on the features.

3 Overall Framework

3.1 Preprocessing

Face Detection. Viola et al. [16] proposed the rapid object detection, including the human face, using a boosted cascade of simple features. Lienhart et al. [8] improved the performance by extending the rapid object detection framework in [16]. We used the method of face detection in an arbitrary scene as proposed in [8].

Adaptive Facial Feature Region Detection. We have adopted the facial feature detection framework proposed in [13] and modified it to detect more facial features and accurate facial features' locations and sizes. Our facial feature detection framework is shown in Fig.1 and proceeds as follows.

(1) *Person-independent*: Detected face region is rescaled into a 100x100 image to minimize individual differences.
(2) *Search region restriction*: A valley image [1] of the rescaled image and a generic feature template [9] roughly estimate the locations of main features, i.e. eyes, nose, and mouth. The valley image (See Fig.1c) clearly show main facial features and the generic feature template (See Fig.1d) segments the rescaled face image into 4 sub-regions for main features: R1 for left eye, R2 for right eye, R3 for nose, and R4 for mouth. Therefore, the generic feature template restricts the search regions of the scene where main facial features are most likely to occur.
(3) *Feature-adaptive search space regions*: To restrict even more search spaces for detecting main facial features, we assigned different sizes of rectangular search space regions (See Fig.1e) on top of the generic feature template. For example,

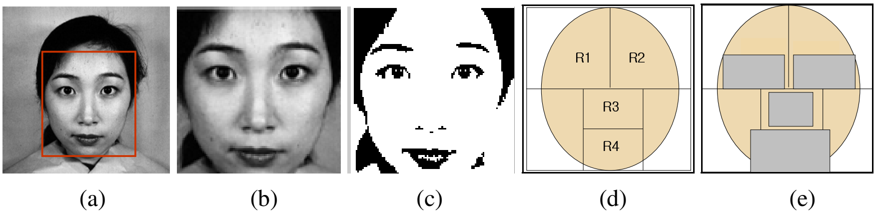


Fig. 1. Facial Feature Detection. (a) Face detection. (b) Rescaled face image. (c) Valley image. (d) Generic feature temple. (e) Rectangular search space regions for main features.

35x20 for each eye, 25x20 for nose, and 45x25 for mouth. By scanning through the specified search space regions and finding the start and end point locations, main facial feature regions are detected in fast and effective manner. Remaining facial features are detected in a same manner.

While scanning through the rectangular search space, size and location of the search space region adjusts to that of actual facial feature (See Fig.2). For the facial features whose actual sizes cannot be known from the valley image, the rectangular search space region becomes the detected feature region. Unlike the method in [13] where 20x20 rectangular regions are assigned on every detected feature, different sizes of rectangular regions that tightly surround facial features are assigned on the detected features. By doing so, noises from the neighboring regions can be reduced.

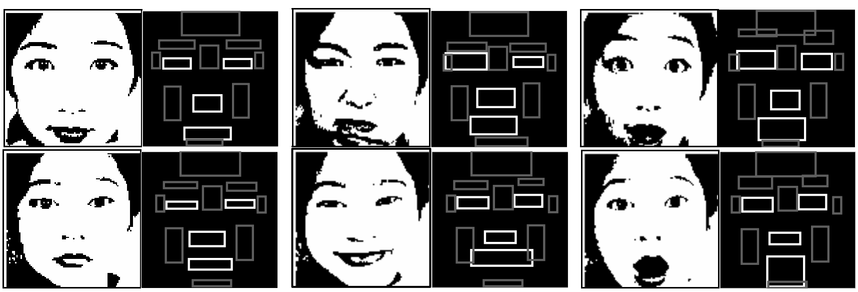


Fig. 2. Detected facial feature regions (main features: white, remaining features: gray)

3.2 Feature-Adaptive Motion Energy Analysis

Facial Feature Selection. Each facial feature's action states, as shown in Table 1, indicate the motion variability of the feature. Depending on the motion variability of the feature, facial features are divided into two categories: non-discriminative (i.e. 2 basic action states) and discriminative facial features (i.e. more than 2 basic action states). Non-discriminative facial features are selected for the feature-adaptive motion intensity evaluation while discriminative features are selected for the feature-adaptive motion orientation evaluation.

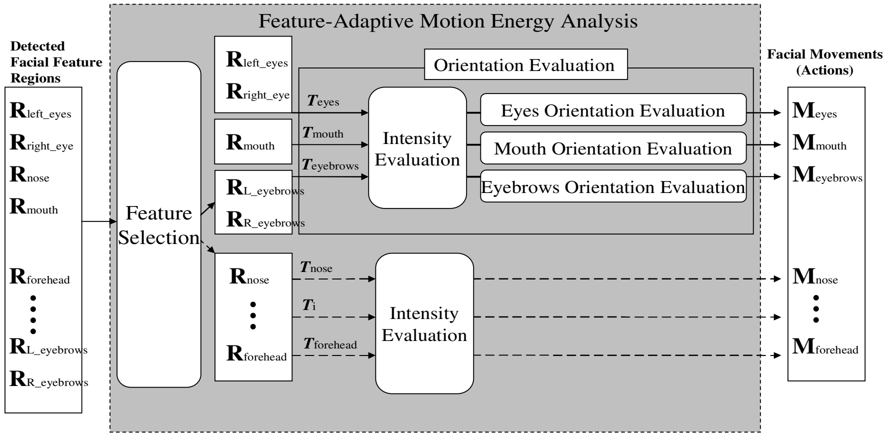
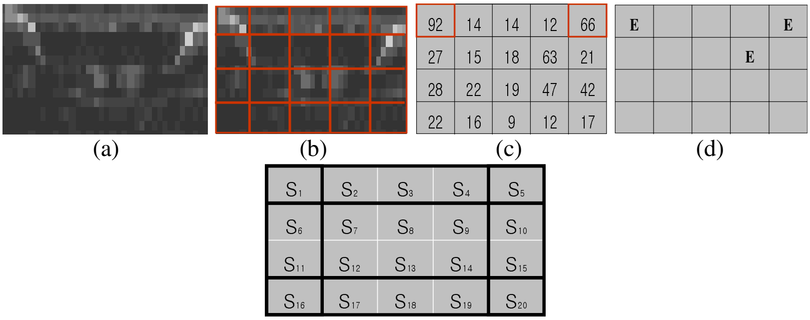


Fig. 3. Feature-adaptive motion energy analysis

Feature-Adaptive Motion Intensity Evaluation. For the non-discriminative facial features, we adopted the framework of point-wise motion energy analysis. Point-wise motion energy analysis uses the difference image and adaptive threshold values to measure the motion intensity of the specified region. The motion intensity is acquired in a fast manner by simply exploiting the few motion energy values within the specified region. More details on point-wise motion energy analysis are found in [13]. The difference is that our method uses the feature-adaptive regions that tightly surround facial features where the method in [13] uses 20x20 regions for every facial feature.

Feature-Adaptive Motion Orientation Evaluation. Feature-adaptive motion orientation evaluations for the discriminative facial features are discussed in detail here. Our approach extended the framework in [13] in three ways. First, we consider not only the motion intensity of features but also the motion orientation of features. Second, the motion orientation evaluation is performed only on selected discriminative features to reduce the computational complexity of the facial motion estimation. Third, the motion orientation of selected discriminative features optimizes the position and size of the detected feature-adaptive regions as well as the motion intensities of sub-regions in the detected feature-adaptive region. There are five discriminative features but three feature-adaptive motion orientation evaluation methods as shown in Fig. 3, because facial features within the same facial unit shares same action states (See Table 1).

Motion orientation evaluation optimizes the motion intensities of sub-regions of the detected mouth region and proceeds as follows. First of all, the location of the mouth is detected using the detected mouth region in the expressed image, R^{e}_{mouth} . Second, divide R^{e}_{mouth} into sub-regions $S_1 \ldots S_n$ where n = 20, and then find each region's average intensity ($\bar{I}_1 \ldots \bar{I}_n$). Third, find average mouth intensities of $\bar{I}_{mouth}$ by adding $\bar{I}_1$ through $\bar{I}_n$ and divide them by n where n = 20. Fourth, find two



(a) (b) (c) (d)

(e)

Fig. 4. Motion orientation evaluation of the expressed mouth: 'LipCornersUp' (a) Difference image of the mouth region. (b) The mouth region is divided into 20 sub-regions. (c) Average motion intensities of the sub-regions. S1 and S5 have highest motion intensities. (d) Evaluation of the sub-regions. (e) Sub-regions are organized into mouth regions. Upper mouth regions: S1 through S5. Middle mouth regions: S6 through S15. Lower mouth regions: S16 through S20. Left mouth corners: S1, S6, S11, S16. Right mouth corners: S5, S10, S15, S20.

sub-regions with highest intensities using $\overline{I}_1 \ldots \overline{I}_n$ and $\overline{I}_{mouth}$. Fifth, evaluate each sub-region and find expressed sub-regions E_i. The sub region S_i is expressed if and only if $\overline{I}_i > (\overline{I}_{mouth} + T_{mouth})$, where T_{mouth} is the feature-adaptive threshold for the mouth region (We let T_{mouth} = 10). Lastly, evaluate the mouth's motion orientation using expressed sub-regions E_i and highest two sub regions from the fourth step.

Note that sub- regions of the detected mouth region are further grouped into logical mouth regions (See Fig. 4e) for the mouth orientation evaluation, which is based on simple Boolean rules. The rule uses the properties that expressed sub-regions represent the shape of the expressed mouth and that the 2 high motion intensity regions indicate the motion orientation of the expressed mouth. Let STRONG(S_i) be strong motion intensity is detected in S_i, MID(S_i) be some motion intensity is detected in S_i, and LOW(S_i) be low motion intensity is detected in S_i. Then,

- *Lip Corners Up* if STRONG(S_1 and S_5).
- *Lip Corners Down* if STRONG(S_{16} and S_{20}).
- *Upper Lip Drawn Up* if and only if STRONG(S_2 through S_4) and MID(S_1 and S_5) and LOW(S_{16} through S_{20}).
- *Open Wide* if STRONG(S_3) and STRONG(S_{18}) and MID(S_2 and S_4) and MID(S_{17} and S_{19}) and LOW(S_7 through S_9) and LOW(S_{12} through S_{14}).
- *Thinned* if MID(S_7 through S_9) and MID(S_{12} through S_{14}) or LOW(S_1 through S_{20}).

The motion orientation evaluation of an eye is as follows. Let R^n_{eye} and R^e_{eye} be the eye region in neutral and expressed image respectively. We define the narrowness of an eye n^k_{eye} as a ratio of the detected eye region's width w^k_{eye} over h^k_{eye} where k is either n for the neutral image and e for the expressed image. Then, an eye is widened if n^e_{eye} is smaller than n^n_{eye}, narrowed if n^e_{eye} is greater than n^n_{eye}, otherwise it is neutral.

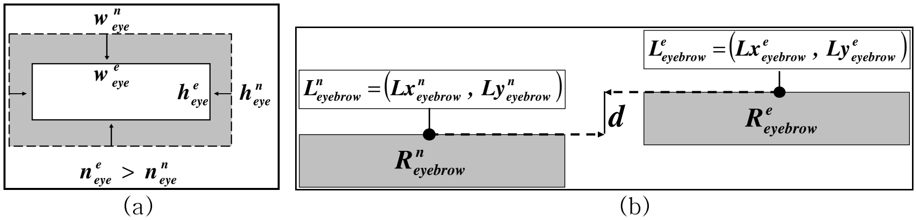


Fig. 5. (a)Eye motion orientation evaluation using the narrowness of a detected eye region. (b)Eyebrows motion orientation evaluation using the start point, a reference point to estimate the location of an eyebrow. The reference point is located at the center of upper end line of the detected eyebrow region and it is used to estimate the location of the eyebrow.

For the motion orientation estimation of eyebrows, let $L^n_{eyebrow} = (Lx^n_{eyebrow}, Ly^n_{eyebrow})$ and $L^e_{eyebrow} = (Lx^e_{eyebrow}, Ly^e_{eyebrow})$ be the reference location (See Fig.5b) of the detected eyebrow rectangular region in neutral and expressed image respectively. Then, an eyebrow is moved down if $Ly^n_{eyebrow}$ is greater than $Ly^e_{eyebrow}$, moved up if $Ly^n_{eyebrow}$ is smaller than $Ly^e_{eyebrow}$, otherwise it is neutral.

Facial Motion Descriptor Generation. A facial motion descriptor is a vector of size 10 containing facial units' actions to represent the facial motion of the given facial expression. For facial units with a single facial feature, the feature's action state becomes the facial unit's action. For facial units with two facial features, a simple rule for generating the facial unit's action is required when corresponding facial features' actions mismatch. The rule has the following assumptions:

1. Movements of neighboring facial features affect each other. Therefore, neighboring facial features' action states are checked when the mismatch occurs. For example, forehead is expressed due to the raise of eyebrows; glabella and eyes are expressed and narrowed respectively due to the pulling down of eyebrows; etc.
2. When neighboring facial features do not provide enough clues, which happen rarely, stereotypical facial feature actions are used to generate the facial unit's action. For example, when action states of left and right eye do not match, it often is the case that the person is uncomfortable i.e. angry or disgust; therefore, facial unit action of eyes becomes narrowed, which is stereotypical eye action for angry or disgust.

3.3 Facial Expression Classification Using ID3 Decision Tree

Ordinary rule-based classifier needs 'trainer' or 'expert' to decide the order of comparisons (or rule); however, it is difficult to come up with an efficient rule when the rule is complicated. We used a modified rule-based classifier with ID3 decision tree as our 'expert' for deciding the order of comparisons. We trained the ID3 decision tree by carefully meeting the requirements for the sample data used by the tree.

There are four requirements that the sample training data used to generate ID3 decision tree must follow [21]. First, the same attributes must describe each example and have a fixed number of values. Second, an example's attributes must already be defined, i.e. they are not learned by ID3 decision tree. Third, discrete classes should be used and classes broken up into vague categories are not sufficient. Lastly, there must be an adequate number of training examples to distinguish valid patterns from chance occurrences.

Our training data set for the ID3 decision tree is generated by using a predefined facial expression recognition rule and 3456 possible action combinations of facial units. We have met all the requirements that the sample training data for the ID3 decision tree must follow because the tree is trained with total of 3456 possible samples (satisfy 4th condition) that are generated from the combination of nine predefined discrete facial units (satisfy 2nd and 3rd condition) and their fixed number of actions (satisfy 1st condition).

The training of the ID3 tree is a one time process for generating an efficient rule from the predefined rule. Since the decision tree is built by using the information gain heuristics of attributes in the training data sample, the attribute that best classifies the remaining samples are chosen when the tree splits. Thus, the classification is done with minimal Boolean comparisons by comparing the value of the root node to the corresponding facial unit action in the facial motion descriptor, and then going downward until it reaches the leaf node and then comparing the leaf node's value with the corresponding facial expression in the facial motion descriptor.

4 Experimental Results and Discussions

We validate the performance of the proposed method in terms of recognition accuracy and computational complexity.

For the recognition accuracy, the proposed method was tested on JAFFE [22] for quantitative analysis on six facial expressions. JAFFE has 213 images and they consist of 40 neutral and 183 expressed images from 10 individual models. For each individual model, one neutral image and all the expressed images were selected for the test; a total of 183 expressed images were tested.

The test was performed as follows. For an each expressed image in JAFFE, we first performed the preprocessing on both the expressed image and its reference neutral image to detect the face and facial features. Then, feature-adaptive motion energy analysis is performed to estimate facial motions of the given expressed image and a corresponding facial motion descriptor is generated. Finally, trained ID3 decision tree uses the generated facial motion descriptor to classify the given facial expression.

Fig. 6 shows the proposed method's performance in terms of recognition accuracy on JAFFE database. Overall recognition accuracy for JAFFE database was 75.4% with the highest recognition rates on 'Surprise' expression (83.8 %) and the lowest recognition rates on 'Fear' expression (63.3 %). We also compared the performance of our method with other simplicity-oriented method that was tested on JAFFE database, namely point-wise motion energy analysis. As shown in Table 2, our method outperforms the method in [13]. This is due to the following factors: (1) More accurate (position and size) adaptive facial feature regions are detected and it avoids

Expression	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	**20**	1	0	3	2	0
Disgust	1	**14**	0	1	1	2
Fear	2	3	**19**	0	4	2
Happy	1	2	2	**36**	3	0
Sad	1	2	3	3	**18**	0
Surprise	1	0	2	1	2	**31**

Fig. 6. Confusion Matrix

misanalysis of the feature's motion orientation. (2) A greater variety of facial motions are recognized from the feature's motion orientation evaluation.

4.1 Computational Complexity

The computational complexity of feature-adaptive motion energy analysis is discussed here. Our approach does not require calculations of motion vectors. Instead, the facial motion is estimated by simply exploiting the motion energy values within the detected feature regions or the characteristics (position and size) of the detected feature region. Therefore, the computational complexity is measured mainly using the size of detected feature regions.

For a face image of size N (i.e. 10000 for a 100x100 face image), facial features, their surrounding rectangular regions, and their areas (i.e. size or number of pixels in the region) to be $F_1...F_n$, $R_1...R_n$, and $A_1...A_n$ respectively. Then, the time complexity of motion intensity evaluations for *non-discriminative facial feature regions* (=9) is

$$\sum_{f=1}^{9} O(A_f + c) \tag{1}$$

where c is a constant time to calculate an average intensity of the facial feature region and apply threshold to see if the region is expressed. After removing the constant time complexity, equation (1) can be rewritten as

$$\sum_{f=1}^{9} O(A_f) \tag{2}$$

For the mouth orientation evaluation, the original mouth detected region is divided into 20 sub-regions with each sub-region a_s where s = 1…20 and it has a size of A_{mouth} divided by 20. Then, the time complexity to evaluate the mouth orientation is

$$\sum_{s=1}^{20} O(a_s) \tag{3}$$

and it can be rewritten in its general form of $O(A_f)$ where f = mouth. Note that the time complexity of the mouth orientation evaluation does not increase from that of the motion intensity evaluation on the mouth region A_{mouth} when we disregard the constant term. Furthermore, it takes a constant time for the motion orientation evaluation on an eye or an eyebrow as well as the generation of a facial motion descriptor. Then, the time complexity of feature-adaptive motion energy analysis, after disregarding constant terms, is

$$\sum_{f=1}^{10} O(A_f) \tag{4}$$

In our experimental environment (Intel Pentium 4 3.0 Ghz, implemented using JAVA without code optimization), it takes an average of 96 milliseconds for each facial expression to perform the preprocessing, feature-adaptive motion analysis (when the average sum of A_f was about 2500 pixels for a 100x100 facial image), and facial expression classification (using the pre-trained ID3 decision tree). Note that more than 90% of the time is spent on the preprocessing. Since detected facial feature regions tightly surround facial features, a small increase in computational complexity is achieved for the motion orientation evaluation as shown Table 2.

Table 2. Comparison of simplicity-oriented methods. ***t*** = time it takes to read a pixel

Methods	**Scope**	**Tested Database**	**Overall Accuracy**	**Time Complexity (facial expression analysis on a 100x100 face image)**
[13]	-Point-wise face model -Motion intensity analysis -Rule-based classifier	JAFFE	70% (183 images)	$\sum_{f=1}^{6} O(A_f) = 2400t$
Proposed	-Action-based face model -Motion intensity and orientation analysis -Modified rule-based classifier using ID3 decision tree	JAFFE	75% (183 images)	$\sum_{f=1}^{10} O(A_f) \approx 2500t$

5 Conclusions

We propose a cost-effective method that is capable of recognizing various facial expressions with very low computational expenses. The proposed method introduces a simplified action-based model to reduce the computational expenses of facial expression analysis and recognition procedure. Feature-adaptive facial feature regions have information about facial feature's size, position, and motion intensities within the region; therefore, feature-adaptive motion energy analysis optimizes the feature region's characteristics to effectively estimate the facial motion of the given expression.

Unlike previous approaches, our method does not neglect facial feature's motion orientation for simplicity, nor does it require a complex face model and complicated algorithms for more reliable recognition. Experimental results show that the computational complexity of the proposed method is as low as the extremely simple facial expression recognition framework that neglects the motion orientation of facial features, and yet the proposed method recognizes a greater rage of facial expressions (up to 75%). For future research, we plan to further investigate more efficient detection and manipulations of adaptive facial feature region, and extend our work in real-scenarios where facial poses and illuminations cannot be controlled.

References

1. Chow, G., Li, X.: Toward a System for Automatic Facial Feature Detection. Proc. of Pattern Recognition 26, 1739–1755 (1993)
2. Donato, G., Bartlett, M., Hager, J., Ekman, P., Sejnowski, T.: Classifying facial actions. IEEE Transactions on PAMI 21, 974–989 (1999)
3. Ekman, P., Friesen, W.V.: Facial Action Coding System. Consulting Psychologists Press Inc. (1978)
4. Ekman, P.: Facial Expression and Emotion. American Psychologists Col. 48(4), 384–392 (1993)
5. Essa, I., Pentland, A.: Coding, Analysis, Interpretation, Recognition of Facial Expressions. IEEE Transactions on PAMI 19, 757–763 (1999)
6. Fasel, B., Luttin, J.: Automatic facial expression analysis: Survey. Pattern Recognition 36, 269–275 (2003)
7. Gokturk, S., Bouguet, J., Tomasi, C., Girod, B.: Model-based face tracking for view-independent facial expression recognition. In: Proc. of IEEE Int. Conf. on FGR, IEEE Computer Society Press, Los Alamitos (2002)
8. Lienhart, R., Maydt, J.: An Extended Set of Harr-Like Features for Rapid Object Detection. Proc. of IEEE ICIP 1, 900–903 (2002)
9. Lin, C., Wu, L.: Automatic Facial Feature Extraction by Genetic Algorithms. Image Processing 8, 834–845 (1999)
10. Mitchell, T.: Decision Tree Learning. In: Machine Learning, pp. 52–81, ch. 3, McGraw-Hill, New York (1997)
11. Pantic, M., Rothkrantz, L.J.M.: Automatic Analysis of Facial Expressions: the State of the Art. IEEE Transactions on PAMI 22, 1424–1445 (2000)
12. Pantic, M., Rothkrantz, L.: Facial action recognition for facial expression analysis from static face images. IEEE Transactions on SMC-Part B: Cybernetics 34, 1449–1461 (2004)
13. Park, H., Park, J.: Analysis and Recognition of Facial Expression Based on Point-Wise Motion Energy. In: Proc. of ICIAR, pp. 700–708 (2004)
14. Shinohara, Y., Otsu, N.: Facial Expression Recognition Using Fisher Weight Maps. In: Proc. of IEEE Int. Conf. on FGR, IEEE Computer Society Press, Los Alamitos (2004)
15. Tian, Y., Kanade, T., Cohn, J.: Recognizing action units for facial expression analysis. IEEE Transactions on PAMI 23, 1–9 (2001)
16. Viola, P., Jones, M.: Rapid Object Detection Using a Boosted Cascade of Simple Features. Proc. of IEEE CVPR 1, 511–518 (2001)
17. Yepeng, G.: Automatic 3D Face Reconstruction based on Single 2D Image. In: Proc. of Multimedia and Ubiquitous Engineering (2007)
18. Zalewski, L., Gong, S.: Synthesis and recognition of facial expressions virtual 3d views. In: Proc. of IEEE Int. Conf. on FGR, IEEE Computer Society Press, Los Alamitos (2004)
19. Zhang, Y., Ji, Q.: Active dynamic information fusion for facial expression understanding from image sequences. IEEE Transactions on PAMI 27, 699–714 (2005)
20. DataFace, Available at: http://face-and-emotion.com/dataface/emotion/expression.jsp
21. ID3 algorith, Available at: http://www.cise.ufl.edu/~ddd/cap6635/Fall-97/Short-papers/2.htm
22. JAFFE database, Available at: http://www.kasrl.org/jaffe.html