

## Development System for Emotion Detection Based on Brain Signals and Facial Images

Suprijanto<sup>(1)</sup>, Linda Sari<sup>(1)</sup>, Vebi Nadhira<sup>(1)</sup>, IGN. Merthayasa<sup>(2)</sup>, Farida I.M<sup>(1)</sup>

<sup>(1)</sup> Medical Laboratory-Departement of Engineering Physics

<sup>(2)</sup> Acoustic Laboratory- Departement of Engineering Physics

Institute of Technology Bandung-Indonesia

e-mail : supri@tf.itb.ac.id

**Abstract**— Detection of human emotions has many potential applications. One of application is to quantify attentiveness audience in order evaluate acoustic quality in concern hall. The subjective audio preference that based on from audience is used. To obtain fairness evaluation of acoustic quality, the research proposed system for multimodal emotion detection; one modality based on brain signals that measured using electroencephalogram (EEG) and the second modality is sequences of facial images.

In the experiment, an audio signal was customized which consist of normal and disorder sounds. Furthermore, an audio signal was played in order to stimulate positive/negative emotion feedback of volunteers. EEG signal from temporal lobes, i.e. T3 and T4 was used to measured brain response and sequence of facial image was used to monitoring facial expression during volunteer hearing audio signal. On EEG signal, feature was extracted from change information in brain wave, particularly in alpha and beta wave.

Feature of facial expression was extracted based on analysis of motion images. We implement an advance optical flow method to detect the most active facial muscle form normal to other emotion expression that represented in vector flow maps. The reduce problem on detection of emotion state, vector flow maps are transformed into compass mapping that represents major directions and velocities of facial movement.

The results showed that the power of beta wave is increasing when disorder sound stimulation was given, however for each volunteer was giving different emotion feedback. Based on features derived from facial face images, an optical flow compass mapping was promising to use as additional information to make decision about emotion feedback.

**Keywords**— Multimodal Emotion Detection, EEG, Facial Image, Optical Flow, compass mapping, Brain Wave

### I. INTRODUCTION

The state of human emotions could be used to quantify the human perception due to some sensation and stimulation. As well know, music is one of source human stimula-

tion that often could be generating different emotion expression on different person [ 1,2].

Human emotion expression has potential approach to find an optimum criterion related with theory of subjective preference of the sound field in architectural acoustic design or for music therapies [8 ].

Expression of human emotion could be identifying from external and internal physiological signal. One of source of external physiological signal is facial expression. Rapid facial signals were believed as facial expressions that used to communicate emotions. Its signals represent temporal changes in neuromuscular activity that may lead to visually detectable changes in facial appearance, including blushing and tears [1,6].

Electric potential of brain activities that measured thought electroencephalogram (EEG) has a potential source internal physiological signal for emotion detection. The common method to justify subject condition is based on power of EEG signal in the specific bandwidth that called brain wave. Based on brain wave signal analysis, positive or negative expression could be evaluated, particularly the change of power in alpha and beta wave [9].

The paper reported a preliminary research on development system for emotion detection based on multimodal emotion detection; one modality based on brain signals that measured using electroencephalogram (EEG) and the second modality is sequences of facial images. The feature as input for emotion detection was derived from vector flow of facial image and the change power of brain wave in range alpha and beta. On the first experiment, the method of feature facial expression was evaluated using "actor" facial image. On the second experiment, an audio signal was customized which consist of normal and disorder sounds. Furthermore, an audio signal was played in order to stimulate positive/negative emotion feedback of volunteers. EEG signal from temporal lobes, i.e. T3 and T4 was used to measured brain response and sequence of facial image was used to monitoring facial expression during volunteer hearing audio signal.

## II. EMOTION DETECTION BASED ON FACIAL IMAGES

### A. Facial Signal

The human face is the location for major communicative outputs. Rapid facial signals represent temporal changes in neuromuscular activity that may lead to visually detectable changes in facial appearance, including blushing and tears. Atomic facial signals underlie facial expressions [6]. Example of rapid facial signals that represented in change expression from normal to other expression, showed in Figure 1. Specifically, rapid facial signal could give special message on communication, such human emotion expression.



Fig. 1 Example of rapid facial signals

### B. Feature Extraction Methods for Emotion Expression

The basic of emotion expression can be categorized on positive (*happiness*) and negative expression (*sadness, anger, disgust*). On face images, emotion expression is mainly controlling the individual or combines motion of four facial feature: brows, eyes, nose and mouth. These are the most attractive features on the facial surface because they have high textures, and symbolize the underlying muscle activations. An observer may recognize easily and directly the messages transmitted from the movement of facial features.

Optical flow in an image sequence has been used to track highly textured regions reliably for extracting the motion information of facial features to be used in further recognition process. Optical flow provides an estimate of the movement of facial feature points that represented as a vector flow[3,4,5].

The basic assumption that used on optical flow is constant intensity on two consecutive facial images, although shift from one position to another occur in the images. However, this assumption is often failed in the real situation. Therefore, in this research was implemented an advance approach of optical flow that be able to compensate intensity variation [4].

The main problem that must be considered on optical flow for feature extraction for emotion expression is separation of non-rigid facial expression from rigid head motion, and facial geometric correspondence to keep face size constant across subjects. Both processes are necessary in order

to ensure that these variables do not interfere on resulted vector flows. For elimination of the above-mentioned rigid head motion from non-rigid facial expression, normalization process was done in order to normalize the face geometric position and maintain face magnification invariance. To do so, an affine transformation (which includes translation, scaling and rotation factors) was applied on two consecutive facial images.

### C. Optical Flow

Defined facial image with gray scale level denote  $f(x,y,t)$  as target of expression image and  $f(x',y',t-1)$  is normal of expression image. One of approach on the optical flow model can accommodate non-rigid deformation and also compensation of intensity variation:

$$m_7 f(x, y, t) + m_8 = f(m_1 x + m_2 y + m_5, m_3 x + m_4 y + m_6, t-1) \quad (1)$$

where  $m_1, m_2, m_3, m_4$  are the linear affine parameters,  $m_5$  and  $m_6$  are the translation parameters,  $m_7$  and  $m_8$  are two new (also spatially varying) parameters that embody a change in contrast and brightness. These parameters are estimated locally for each small spatial neighborhood, but for notational convenience their spatial parameters are dropped. In order to estimate these parameters, we define the following quadratic error function to be minimized:

$$E(m) \approx \sum_{x,y \in \Omega} [k - c^T m]^2 \quad (2)$$

where  $m = [m_1 \dots m_8]^T$ ,  $k = f_t - f + x f_x + y f_y$ ,  $c = (x f_x, y f_y, x f_y, y f_x, f_t - f)^T$  and  $\Omega$  is a small spatial neighborhood. This error function can now be minimized analytically by differentiating with respect to the unknown's parameters  $m$ :

$$\frac{dE(m)}{dm} = \sum_{x,y \in \Omega} -2c[k - c^T m] \quad (3)$$

setting the result equal to zero on Equation 3, and solving for  $m$  yielding:

$$m = \left[ \sum_{x,y \in \Omega} c c^T \right]^{-1} \left[ \sum_{x,y \in \Omega} c k \right] \quad (4)$$

The solution assumes that first term is invertible. This can be guaranteed by integrating over a large enough spatial neighborhood  $\Omega$  with enough image contents and estimation parameters  $m$  was done by iteration process. We define matrix  $L$  with dimension  $5 \times 5$ , where each element of  $L$  is  $\lambda_i$  as smoothness constants. Parameter  $m_i$  can be estimated on iteration process that written as

$$\mathbf{m}^{(j+1)} = (\mathbf{c}\mathbf{c}^T + \mathbf{L})^{-1}(\mathbf{c}\mathbf{k} + \mathbf{L}\bar{\mathbf{m}}^{(j)}) \quad (5)$$

The initial estimate  $\mathbf{m}^{(0)}$  is determined from the closed-form solution. On the  $j+1^{\text{st}}$  iteration  $\mathbf{m}^{(j+1)}$  is estimated from the previous estimate  $\mathbf{m}^{(j)}$ .

#### D. Compass Mapping

The performance of any digital recognition system depends on input of a set feature data. On the first step of processing, a set feature data is derived by optical flow and resulting 2D vector maps. The dimension of set feature data is similar with image size. To reduce problem on recognition of emotion state, the dimension of feature data that represented on 2D vector maps must be reduced. Furthermore, the spatial information that contain on 2D vector map should be transformed into other simplify information.

The 2D vector map was transformed into compass mapping that that represents major directions and velocities of facial movement [11]. In other word, compass map will be transformed data sets that obtained from 2-D vector maps into the sum of all directional vectors from the centre origin.

This compass is not to be confused with a human face divided into segments, nor is this centre origin equivalent to the centre of the face. It simply represents the sum of all directional vectors from the centre origin.

### III. EMOTION DETECTION BASED ON BRAIN WAVE

EEG is a superposition of the volume-conductor fields produced by a variety of active neuronal current generators. EEG signal require quantitative techniques that can be validly applied to time series exhibiting ranges of non-stationary behavior [9]. The parameter EEG in the certain time may change due to internal and our external stimulation. The example of EEG signals in the time domain shown in Figure 3.

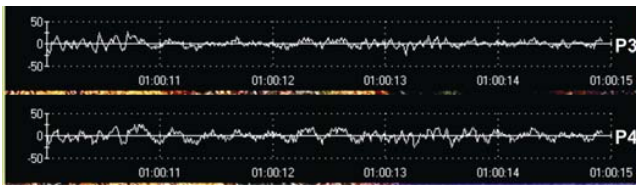


Fig. 3 EEG signals in the time domain from P3 and P4

The change of emotion condition can be occurring due to internal and external stimulation. In this work, we focused to investigate EEG respond that represented positive

or negative emotion feedback when volunteer stimulates a normal and strange music. However, EEG respond in time domains are relatively difficult to quantify emotion feedback because we predicted that the change EEG amplitude relatively indistinguishable. As alternative, analysis Power of EEG in the frequency domain was used. Since power spectral density (PSD) must be derived from certain duration of EEG signal, PSD of EEG was determined using pwelch method[8].

Given digital EEG signal  $X(N)$ , then data was divided with  $P$  segment along  $D$ , then shifting on each segment was done along  $S$  data ( $S \leq D$ ), resulting maximum  $P$  is adalah  $(N-D) / S + 1$ . Thus, data on each segment  $P$  is

$$x^{(p)}(k) = w(k)x(k + pS), \quad 0 \leq k \leq D-1 \quad (6)$$

where  $0 \leq p \leq P-1$ . From those data will be obtained:

$$\tilde{P}_{xx}^{(p)}(f) = \frac{1}{UDT} |X^{(p)}(f)|^2 \quad (7)$$

on the frequency range  $-1/2T \leq f \leq 1/2T$ , where

$$X^{(p)}(f) = T \sum_{k=0}^{D-1} x^{(p)}(k) \exp(-j2\pi f k T) \quad (8)$$

$$U = T \sum_{k=0}^{D-1} w^2(k) \quad (9)$$

The result of PSD estimation :

$$\hat{P}_B(f_i) = \frac{1}{P} \sum_{p=0}^{P-1} \tilde{P}_{xx}^{(p)}(f) \quad (10)$$

The example results of PSD estimation from EEG signal (in Figure 3) was shown in Figure 4.

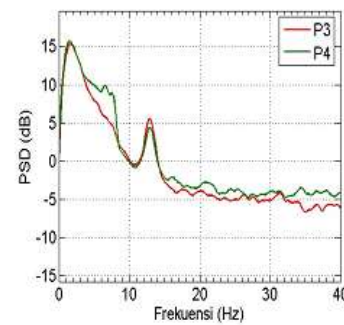






Fig. 4 Example results of PSD EEG signal using pwelch methods

Subject condition is often evaluated based on amplitude of PSD in a specific frequency range of EEG signal that called *brain wave*. The common type of brain wave is alpha, beta, theta and delta.

Alpha is a 10 Hz brain wave become the first electrical brain wave, which found by Hans Berger in 1929. Another next brain waves, was named as Greek alphabet (shown in Table 1).

Table 1 . Type of Brain wave

Wave		frequency	Voltage	Subjeck condition
genre	Pattern			
Betha		14 – 30 Hz	10-20 $\mu$ V	Activity, thinking
Alpha		8 - 13 Hz	Kids: 75 $\mu$ V Adult: 50 $\mu$ V	Relaks, closed eye
Tetha		4 – 7 Hz	Kids: 50 $\mu$ V Adult: 10 $\mu$ V	Light sleep/ emosional stress
Delta		0,5 – 3 Hz	10 mV	Profound sleep

We speculate that the negative or positive emotion feedback could be evaluated based on PSD change of alpha and beta waves.

#### IV. MATERIAL AND EXPERIMENTS

Three male university students, age in between 18-21 years, participated in the experiments. The first experiment, volunteers were instructed to imagine and to express the basic emotion (*happiness, surprise, anger*). Each facial expression was recorded using camera digital. This data was used to evaluated the preformed of feature extraction methods for emotion expression.

The second experiment, we prepared the audio stimulation that predicted can be induced positive/negative emotion feedback of volunteers. The audio signal about 10-15 minutes was customized which consist of normal and disorder sounds. The type of disorder sounds for each volunteer is not similar. EEG signal from temporal lobes, i.e. T3 and T4 was used to measured brain response and sequence of facial image was used to monitoring facial expression during volunteer hearing audio signal. The experiment setup was shown in Figure 5.

As bio-amplifier EEG, we used Flexcomp Infiniti (SA7550) that interfacing to computer using fiber optic to protect volunteer from electric shock.

Each volunteer was stimulated with two different audio signals. During acquired EEG and Facial image data, volunteer was instructed to minimize motion of head.

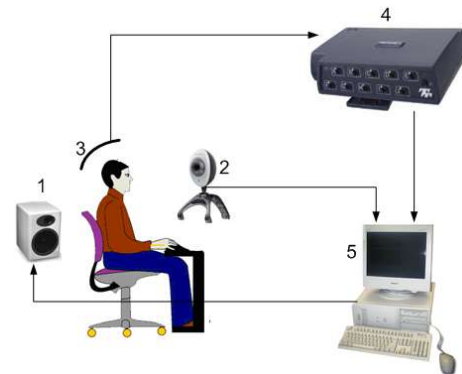


Fig. 5 Experiment setup. The audio was stimulated using speakers (1), where source controlled by computer (5). In the same time, a gold active electrodes attached on temporal lobes (T3 and T4) that connected to EEG amplifier (4) as well as the facial expression continues capture using web-came (2). All data was recorded on the computer (5).

#### V. RESULTS

##### A. The First Experiment

The objective of the first experiment is to evaluate our strategy to extract facial expression based on an advance optical flow. As we mentioned above, position moving of a facial feature was detected between normal expression and other expression. The feature sets that derived from optical flow were represented where a magnitude and direction of vector flow of change position of facial muscle due to changing of emotion expression.

The main problem that must be considered on optical flow for feature extraction for emotion expression is separation of non-rigid facial expression from rigid head motion, and facial geometric correspondence to keep face size constant across subjects. Both processes are necessary in order to ensure that these variables do not interfere on resulted vector flows. For elimination of the above-mentioned rigid head motion from non-rigid facial expression, **normalization process** was done in order to normalize the face geometric position and maintain face magnification invariance. To do so, an affine transformation (which includes translation, scaling and rotation factors) was applied on two con-

secutive facial images. The summary of processing steps were shown in Figure 6.

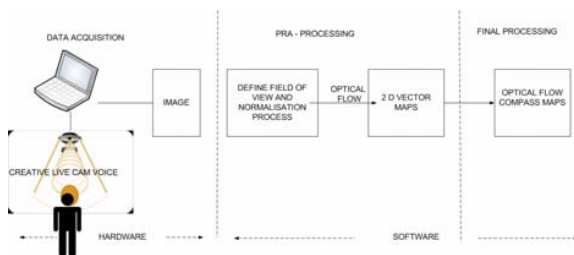


Fig 6. The summaries of processing steps for extracts facial expression

As shown in Figure 6, after normalization process, then compass mapping was derived from the output of optical flow. On Figure 7-a and Figure 7-b were shown facial image of subject that acting on normal and happiness expression. Change of facial feature was detected on the atomic facial. On Figure 7-c was shown motion vector maps where each vector detected motion in 5x5 pixel windows. The arrows length of vector is proportional to the motion of the facial feature. Feature of set data obtained from vector maps have dimension similar with facial image. A strategy that used to reduce dimension is based on the resultant of vector on similar direction of facial motion from the original neutral image. The obtained maps are called compass maps (see Figure 7-d). On the compass maps, vectors radiate from the origin (the centre of the compass – zero position). Its transformation was change Cartesian spatial domain (with dimension equal to image) into angle spatial domain (with dimension from angle  $0^{\circ}$ - $360^{\circ}$ ).

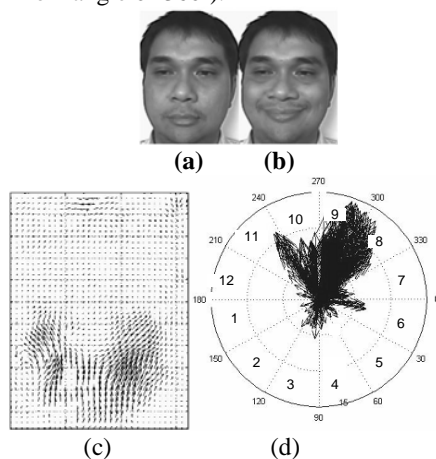


Fig. 7 . (a). Facial images on normal expression, and (b). Positive (happiness) expression. (c) mapping of motion vector that used to detect expression change compass maps represent represents the sum of all directional vectors from the centre origin.

The other results of compass maps from negative expression was shown in Figure 8.

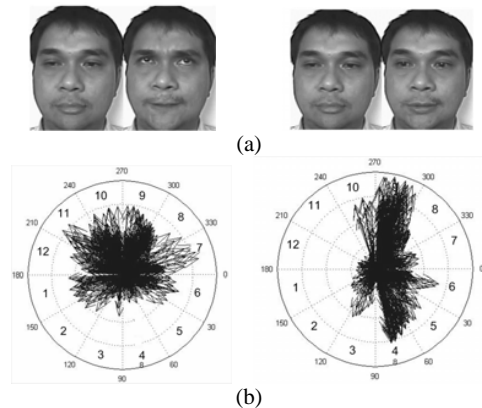


Fig. 8 . (a) Facial images on negative expression. (b) compass map of negative expression

The different spread of resultant vectors on similar direction of facial motion from the original neutral image to positive/negative expression can be shown on Figure 7-c and Figure 8-b.

### B. The Second Experiment

In these experiments, facial image and EEG signal were acquired simultaneously under sound stimulation. The setup of experiment was shown in Figure 5. The scheme of sound stimulation can be divided by three sections (see Fig. 9). For each volunteer, total duration of sound stimulation is about 15 minutes. The duration and starting of disorder sound (section 2) is random for each volunteer.

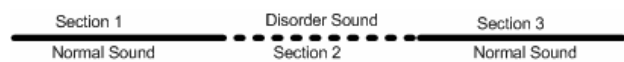


Fig. 9 . The scheme of sound stimulation

The EEG signal measured form point T3(left Hemisphere) and T4(right hemisphere) with reference on the mastoid. The PSD on each section was averaged for both points and for alpha and beta wave. Furthermore, from the sequence of facial image, we focused on the change of facial muscle where transition occurs from section 1 to section 2. The results of vector map that obtained from optical flow was transformed into compass mapping that that represent major directions and velocities of facial movement

The results from volunteer A, the sound stimulation in the section 1 was induced the power of alpha wave is relatively higher than the power of beta wave. Due to sound stimulation in the section 2, the power of beta wave is slightly increase than respond on the section 1. However, in the section 2, the power of alpha wave is increase for T3 and decrease for T4 (see in Figure 10-a).

With subjective criterion, positive emotion respond was reflected on facial image of volunteer(see in Figure 10-b). Based on the optical flow – compass maps, the significant of vector magnitude was indicated on specific segment of compass mapping. If the compass map of 360 degree is divided into 12 segments each of 30 degree, the most active of segment was detected on segment 9 and 10 (see in Figure 10-c)..

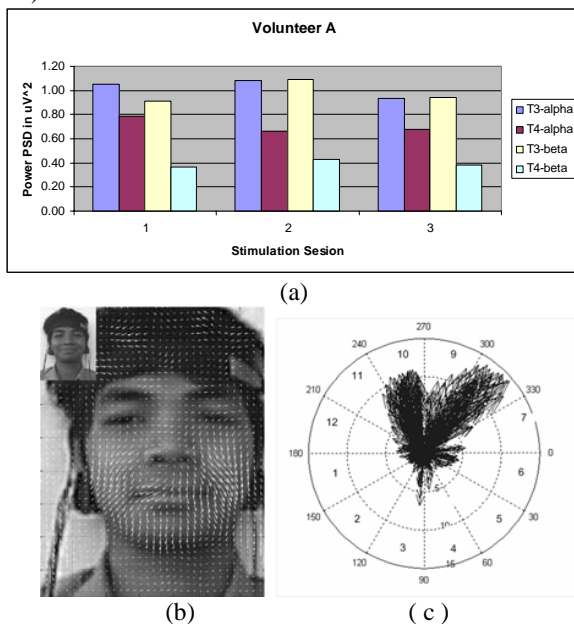
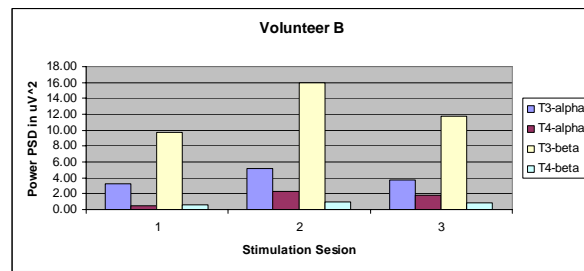


Fig.10 (a) The PSD of EEG signal due to audio stimulation and (b) the vector flow maps that detected the most active facial muscle (c). compass map of 360 degree is divided into 12 segments each of 30 degree. the most active of segment was detected on segment 9 and 10. The result is correlated with positive emotion respond based on subjective criterion.

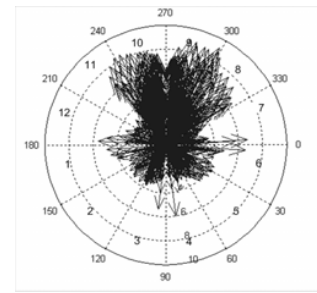
For volunteer B, we change type of sound stimulation and the duration and starting of disorder sound (section 2). The power of beta wave, particularly on T3 is relatively higher than the power of alpha wave for all section (see in Figure 11-a). Based on subjective criterion of facial images, we detected negative emotion respond (see in Figure 11-b). Based on compass mapping, the active segments are spreading on whole segment.



(a)



(b)



(c)

Fig. 11 (a) The PSD of EEG signal due to audio stimulation and (b) the vector flow maps that detected the most active facial muscle (c). compass map of 360 degree is divided into 12 segments each of 30 degree. the most active of segments are spreading on whole segment. The result is correlated with negative emotion respond based on subjective criterion

The other type of sound stimulation and the duration and starting of disorder sound (section 2) was used as stimulant for volunteer C. Due to sound stimulation in the section 2, the power of beta wave is slightly increase than respond on the section 1. Again, the significant increasing power was detected for T3. On the section 3, the power of alpha tends to higher than the power of beta wave on both T3 and T4(see in Figure 12-a).

Based on subjective criterion on facial images, the negative respond are expressed. The magnitude of vector flows is relatively lower than respond from volunteer B. From the results of compass mapping, we found that the active segments are also spreading on whole segment although the amplitude of resultant vector is relatively smaller than the results from volunteer B (see in Figure 12-c).

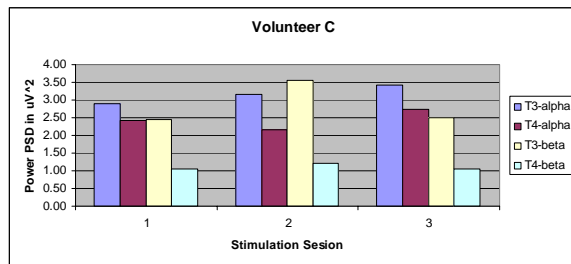
## VI. CONCLUSIONS AND DISCUSSION

The preliminary work as part our research to develop system for emotion detection based on multi-modality feature sets was presented in the paper. In the real life, human

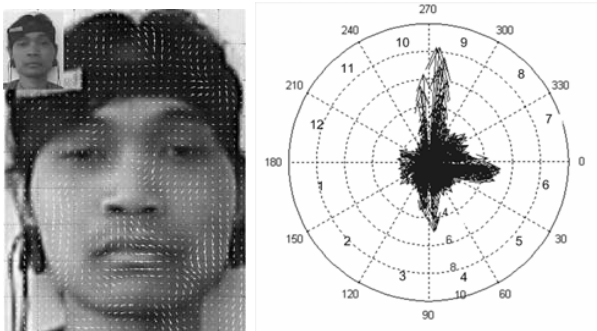
are capable of producing thousands of emotion expression that vary in complexity, intensity and meaning. However, for specific application, such as for automatic sleep analysis, lie detector system or subjective preference for architectural acoustic design, complexity system for emotion detection could be reduced. The source of feature sets that correspond to specific emotion criteria is feasible to define.

Related with automatic system to evaluate subjective preference for architectural acoustic design, results of the

sphere (T3) is relatively higher that the respond from right hemisphere (T4). As mentioned in [7], the respond of left hemisphere is more sensitive to temporal characteristic of sound stimulation. The stimulation sound that used in section 1 and 2 may be significant different in temporal characteristic than spatial characteristic. From the result our experiment, the change emotion respond of volunteer could be detected based on the increasing power of beta wave. However, the type of emotion respond (negative/positive) could not conclude from EEG data.



(a)



(b)

(c)

Fig. 12 (a) The PSD of EEG signal due to audio stimulation and (b) the vector flow maps that detected the most active facial muscle (c). compass map of 360 degree is divided into 12 segments each of 30 degree. the most active of segments are spreading on whole segment. The result is correlated with negative emotion respond based on subjective criterion

typical feature sets that derived from EEG signal and facial images was promising to developed as input for automatic system for subjective preference that related with room acoustic design. On its application, the target of emotion criterion is focused on justify the positive/negative emotion feedback.

The feature sets that derived from EEG signal showed that disorder sound tend to stimulate increase power of beta wave. From three volunteer, the respond from left hemi-

The other feature sets that related with negative/positive emotion respond was attempted to detect from facial images. We implemented and evaluated the optical flow methods to extract feature sets that represent vector flows. On this method, the significant vector flows represent the most active facial muscle form normal to another expression. The reduce problem on recognition of emotion state, dimension reduction of feature set of 2D vector maps is required. An approach called compass mapping was used to represent major directions and velocities of facial movement. The compass diagram (360 degrees) is divided into 12 segments each of 30 degrees, and represents the total directional facial movement from the original neutral image. Vectors radiate from the origin (the centre of the compass – zero position). We found that the active segments location for negative emotion feedback is relatively spreading on whole segment than the active segments location for positive emotion feedback.

Although the power of beta wave is increasing when disorder sound stimulation was given, for each volunteer was giving different emotion feedback. From the results of vector flows, we can learn what typical negative or positive feedback from typical Indonesia peoples, which is important as criterion in automatic emotion recognition system.

Future work, we will increase the number of volunteer to develop better database of typical feature of positive/negative emotion feedback. In the sub system of recognition system, the various pattern recognition strategies, such as principal component analysis and artificial neural network will be implemented.

#### ACKNOWLEDGMENT

The research is partly funded by Faculty Research Program-Institut Teknologi Bandung-2008

# REFERENCES

1. C. Busso, Z. Deng \*, S. Yildirim, M. Bulut, C. Min Lee, A. Kazemzadeh, S. Lee, U. Neumann\*, S. Narayanan, "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information", the Proceedings of of ACM 6th International Conference on Multimodal Interfaces (ICMI 2004), State College, PA, Oct 2004
2. M. Pantic and L.J.M. Rothkrantz, "Toward an Affect-Sensitive Multimodal Human-Computer Interaction", Invited Speaker on the Proceedings of IEEE Vol. 91, No. 9, September 2003
3. D. DeCarlo, D. Metaxas, "Optical Flow Constraints on Deformable Models with Applications to Face Tracking", International Journal of Computer Vision, 38(2), pp. 99-127, July 2000
4. S. Periaswamy and H. Farid. Elastic registration in the presence of intensity variations. IEEE Transactions on Medical Imaging, 2003
5. Cohn, J.F., Zlochow, A.J., Lien, J.J., and Kanade, T., "Feature-Point Tracking by Optical Flow Discriminates Subtle Differences in Facial Expression," The Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition, Japan, April 1998
6. J. Lien, T. Kanade, J. Cohn, and C. Li, "Automatic Analysis of Facial Expressions: The State of the Art", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1424-1445, Vol.22, Issue 12, 2000
7. Sato, S., Nishio, K., and Ando, Y. 'Propagation of alpha waves corresponding to subjective preference from the right hemisphere to the left with changes in the IACC of a sound field'. J. Temporal Des. Arch. Environ. 3, 60-69. 2003
8. John L. Semmlow, Biosignal and Biomedical Image Processing: MATLAB-Based Applications, (Signal Processing and Communications, 22), Marcel Dekker, Inc. 2004
9. Webster, John G (1998), "Medical Instrumentation: Application and Design": John Wiley and Sons, Third Edition, USA.
10. James A. Cond, John J.B. Allen and E. Harmon, Voluntary facial expression and hemispheric asymmetry over the frontal cortex, Psychophysiology, 38 ~2001!, 912-925. Cambridge
11. Carmen J. Duthoit, et.al, "Optical Flow Image Analysis of Facial Expression of Human Emotion – Forensic Applications, Journal e-Forensic, January 2008,