

A Thermal Facial Emotion Database and Its Analysis

Hung Nguyen¹, Kazunori Kotani¹, Fan Chen¹, and Bac Le²

¹ Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa, Japan
{[nvhung](mailto:nvhung@jaist.ac.jp), [ikko](mailto:ikko@jaist.ac.jp), [chen-fan](mailto:chen-fan@jaist.ac.jp)}@jaist.ac.jp
² University of Science, Ho Chi Minh city,
227 Nguyen Van Cu, Ho Chi Minh city, Vietnam
lhac@hcmuns.edu.vn

Abstract. In recent years, thermal image has extensively been used in many fields such as military (e.g., target acquisition, surveillance, night vision, homing and tracking) and civilian purposes (e.g., medical diagnosis, thermal efficiency analysis, environmental monitoring). It may be a promising alternative for investigation of facial expression and emotion. Currently there are very few database to support the research in facial expression and emotion, however most of them either only include posed thermal expression images or lack thermal information. For these reasons, we propose and establish a natural visible and thermal facial emotion database. The database contains seven spontaneous emotions of 26 subjects. We also analyze a visible database, a thermal database to recognize expression and thermal information to recognize emotion.

Keywords: Facial expression analysis, thermal image, visible image, spontaneous database, facial emotion, KTFE database.

1 Introduction

Although we have spent many years in doing research on facial expression analysis since the work of Darwin in 1872 [1], we have still not understood clearly how the human brain works to analyze the facial expression and how computer could reach the accuracy rate of automatic facial expression analysis as human. Facial expression analysis continues to be an active research topic for behavioral scientists. There is such a huge research on facial expression analysis, already described and surveyed in detail in [2],[3],[4]. Currently most research work uses visible images or videos and has achieved good result. However, under the lack of illumination, even darkness or exceeding of source of light, the result of visible expression analysis is not good. On the other hand, thermal images are not sensitive to light conditions. Consequently, using thermal images helps us to complete the gaps of visible images. Besides, the skin temperature changes are useful to classify the emotions [5] and facial expression is a good emotion-related behavior [6]. We can infer emotions from skin temperature and expressions from

several special emotions. Moreover, most of the current databases used for research are visible and posed. The expressions, from those databases, are usually obtained from unreal emotion and overlay features. Secondly, there are a few thermal facial image databases but they are posed thermal expression images. Even though a database is built in posed and spontaneous expressions, it still meets some mistakes such as when they designed data acquisition, they forgot about time lag phenomenon or expressions are elicited by asking participants to imitate sample expressions, exaggerated expressions. With these reasons, we propose and establish a thermal facial expression and emotion database to allow the research in facial expression analysis to be more realistic. In this paper, we describe in detail the materials and methods to design and collect the thermal facial emotion database - KTFE (Kotani Thermal Facial Emotion) database. To verify the effectiveness of our spontaneous database, we use PCA (Principal Component Analysis), EMC (Eigenspace Method based on Class features) and PCA-EMC to classify facial expressions of a visible and thermal facial image database. We have also used PCA and PCA-EMC to classify emotions of thermal facial emotion database leading to very attractive results.

2 Review of Existent Natural and Infrared Databases

Innumerable natural databases for facial expression analysis have been built since many years, such as Cohn-Kanade (CK) [7], UA-UIUC [8], MMI [9], UT-Dallas [10], Belfast [11], AAI [12] and so on. A comprehensive survey of these databases is given in [3].

Compared to the number of existing visible databases, only very few thermal face databases are available in the literature. Furthermore, these databases only include some posed thermal data and one spontaneous thermal data. In this

Table 1. Current thermal facial database

Ref	Size	Wave band	Education	Lightning	Exp Des
NIST Equinox [13]	600 subjects 1919 infrared images	8-12 μ m 3-5 μ m	Posed	Above, left and right	Smiling, frowning, surprise
IRIS [14]	30 subjects, 4228 pairs of thermal and visible images	7-14 μ m	Posed	Left, right, both lights, dark	Surprise, laughing, anger
USTC-UVIE [15]	215 subjects	8-14 μ m	Posed and spontaneous	Left, right and front	Happy, angry, neutral, disgusted, fearful, sad, and surprised

document, we listed and compared several databases of infrared facial expression, along with the information related to the name, the number of subjects, wave band of thermal camera, lighting, illumination and expression description as table 1. Firstly, NIST Equinox [13] has been used in many researches of thermal image, which is not available anymore. Secondly, IRIS Thermal/Visible Face Database [14] is very useful only for face recognition because posed expressions are elicited by asking subjects to perform a series of emotional expressions in front of a camera. Thirdly, USTC-NVIE database [15] is a very good database and adaptable for a good posed and spontaneous thermal database. However, their procedure for data acquisition to induce emotions has a mistake. In their video clips to evoke emotion, the gaps between each emotion clip are 1-2 min long which is too short for participants to establish a neutral emotion status. They do not mention about the recording time before ending of each emotion clip. The changing of human temperate is later than the changing of emotion. Therefore, the time before ending of each emotion clips is very important. In a short, there is only one facial expression database using visible and thermal image, although many expression databases use visible or thermal only. Furthermore, there exist several unclear and non-suitable procedures in these databases. These reasons motivated us to propose and build up another natural visible and infrared facial emotion database.

Table 2. Information of participants in building the KTFE database

Number	Age	Sex	Education	Glasses	Nationality
2	32	2 M	Post Doc	2No	Vi
1	31	1M	Post Doc	1No	Vi
2	30	1M,1F	PhD	1Yes, 1No	Vi
1	29	1M	Master	1Yes	Vi
6	28	3M,3F	3Master, 3PhD	5Y,1No	Vi, Thai
1	27	1F	Master	1No	Vi
5	26	3M,2F	4Master, 1PhD	3Yes, 2No	Vi
5	25	4M,1F	2Master,3PhD	2Yes, 3No	Vi
2	24	2F	2Bachelor	2Y	Vi, Thai
1	12	1M	Pupil	1No	Jap

3 Materials and Method

3.1 Participants

The database contains 26 subjects from 11 year-old to 32 year-old as depicted in table 2. To ensure accuracy in results of the experiments, all of the participants were asked to take rest, maintain in good mood for 2 hours prior to the measurements and to avoid the presence of cosmetic substances on their face at the

time of experiment. Before taking the experiment, each participant consented to join the test and also signed the test agreement.

3.2 Measurement Devices and Environment

Room Setup. The room for conducting the experiment is L shaped with $8\text{m} \times 12\text{m} \times 3.5\text{m}$ and the omitted area is about 6m^2 . The experiment room was always kept quiet to ensure no effect to induce the participants emotions. During the data acquisition, the internal temperature of the room that is used for conducting experiments is maintained between 24°C and 26°C because of the sensitivity of the facial surface to the environmental temperature. To control the humidity and temperature of the room, we used the building air conditioning system, and the flow of air condition was not directed to the testing area. To keep the constant illumination between day and night, kept both the door and the curtains closed during the experiment. The experiment area had infrared camera, equipped with laptop, desk, chair, LCD screen, mass storage disk, head phone, and two special curtains. Two curtains separated the participant from the experimenter as a result of which the participants felt more comfortable and not shy, making it easier to induce their emotions. The view of the room and experiment area is depicted in Fig.1.

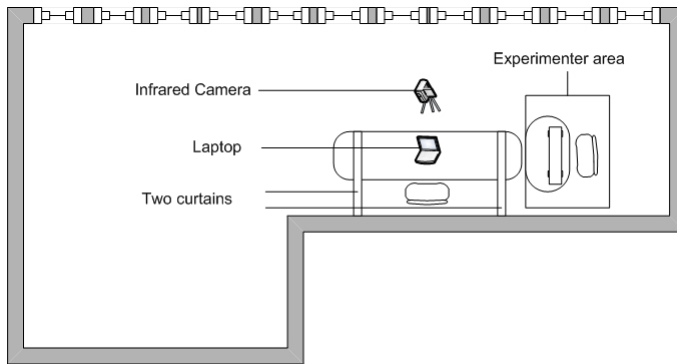


Fig. 1. Overview of experiment room

Camera Setup. We used an Infrared Camera NEC R300 to obtain the visible and thermal videos. The infrared camera has 3.1mega pixels visible camera capturing 5 frames per second and a long wavelength infrared (LWIR) camera opening from $8\mu\text{m}$ to $14\mu\text{m}$. The thermal sensitivity is 0.03°C at 30°C . Thermal infrared imaging data were captured at 5ft/s. The camera was placed at a height of 1.5m above the floor and 0.85m in front of the participants. To obtain the correct temperature of participants, the calibration was set up before each experiment and updated automatically per minute. We used NS9500 PRO, supporting real-time monitor, to capture both visible and thermal data and NS9500 STD to view, enhance, analyze, and extract the thermal data.

3.3 Procedures

Stimuli. In this experiment, we use selected emotional video clips to evoke the emotion of the participant. The video clips were gained by four persons from the internet and judged by the authors. There are four angry clips, four disgust clips, four fearful clips and one fearful game, six happy clips, seven sad clips, three surprised clips and two neutral clips. We further classified each emotion class into four sub-classes according to their intensive levels.

Data Acquisition. The experiment room had only one participant and experimenter during data collection. The participant was seated comfortably in an armchair in front of a laptop screen as shown in Fig.1. Fig.2 shows the data acquisition procedure, which fixes the former database mistake. Depending on the participants, we did not ask them to not wear on or take off their glasses. We also did not require them to keep their head fixed in one position because we wanted to obtain the spontaneous emotions. The participants were given an introduction of the purpose and procedure of the experiment prior to taking data and then they were asked to wear headphone. Before and after each session, the instrumental music turned on to help the participant return to the neutral feeling. In each session, we tested only one emotion and specially paid attention to the time lag phenomenon. When participants did not to take some fearful or angry clips, we stopped this session to protect the human right. After experiment of each person, we asked them to give the report contributed their feeling to each emotion video clips. These self-reported data were helped us to label the recorded videos.

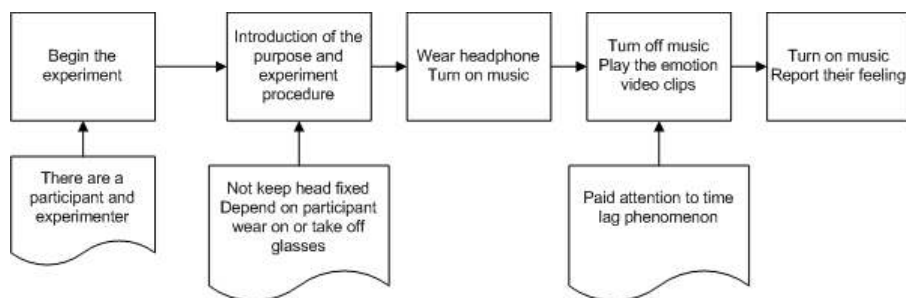


Fig. 2. Data acquisition procedure

3.4 KTFE Database Design

The first version of KTFE database, which contains seven spontaneous emotions of 26 subjects, includes 126 gigabyte of visible and thermal facial emotion videos, visible facial expression image database and thermal facial expression image database. To obtain the thermal expression image database, we manually

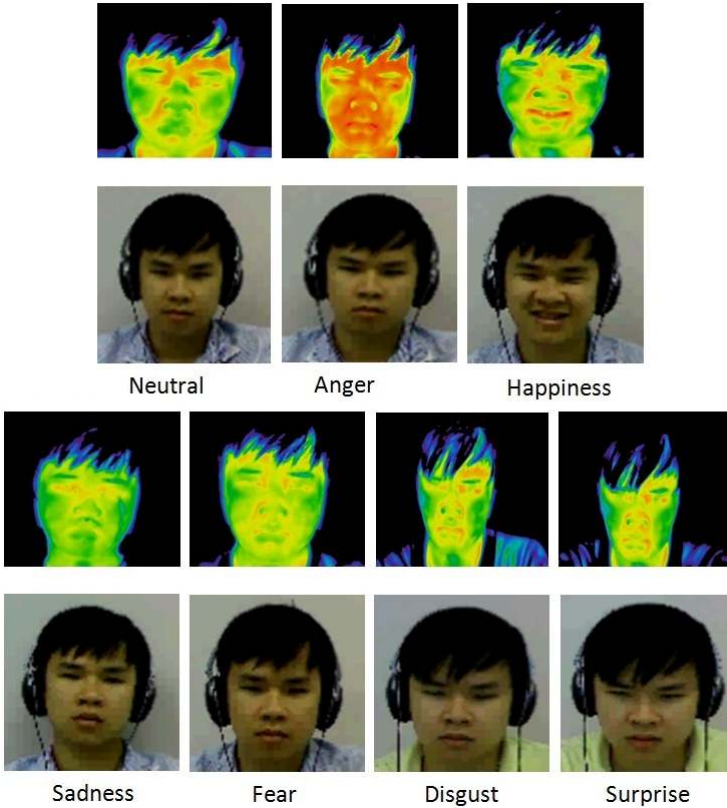


Fig. 3. Sample thermal and visible images of seven expressions

choose the expressions using NS9500STD software. There are three persons to select manually the suitable frames for every emotion of each person and extract into thermal images. The visible image database is also manually extracted and chosen by two persons.

4 Data Analysis

In this section, we bring a fundamental evaluation of usability of the visible facial database, thermal facial database and conduct the thermal data to analyze emotions. Before evaluating of these databases, to avoid any undesired noise in the thermal images, visible images, we use median smoothing filter to reduce noises and blurring. The preprocessing for visible image is to normalize image to facial space containing only face. To analyze the facial expression using thermal infrared data of expressions, after reducing the effect of noise, we use three convention methods PCA, EMC, PCA-EMC. With PCA, the aim is to build a face space, including the basis vectors called principal components, which better

describe the face images [16]. To estimate emotions using PCA, we divide the training set into five classes and compute the eigen-space as following:

Step 1: Concatenating each row of a thermal data of each frame by row, the thermal data can be transformed to a column vector. Given M frames of thermal data as training data, we convert these datum to corresponding column vectors

Step 2: The mean of training data has to be calculated and then subtracted from each original thermal data in the datum.

Step 3: Calculate the eigenvectors and eigen-values of the covariant matrix.

For each test thermal data, we project it to the eigen-space of each class and derive the reconstruction thermal data from each eigen-space. Using mean square error, measuring the similarity, between input thermal data and reconstruction thermal data, we can choose a suitable class for input thermal data which is a minimum of mean square errors [16].

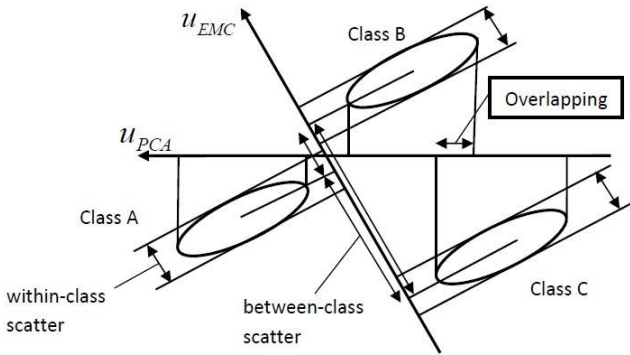


Fig. 4. Examples of a eigenvector of PCA and EMC [18]

In reference [17], the authors proposed eigen-space method based on class-features (EMC) to analyze the facial expressions. The difference between PCA and EMC is that PCA finds the eigenvector to maximize the total variance of the projection to line, while EMC is obtained eigenvector to maximize the difference between the within-class and between-class variance. The Fig.4 shows the advantage of EMC over PCA. The difference between the within-class and between-class covariance is calculated as following:

$$S = S_B - S_W. \tag{1}$$

$$S_B = \frac{1}{M} \sum_{f \in F} M_f (\bar{x}_f - \bar{x})(\bar{x}_f - \bar{x})^T. \tag{2}$$

$$S_W = \frac{1}{M} \sum_{f \in F} \sum_{f \in F}^{M_f} M_f (\bar{x}_{fm} - \bar{x}_f)(\bar{x}_{fm} - \bar{x}_f)^T. \tag{3}$$

$$\bar{x}_f = \frac{1}{M} \sum_{m=1}^{M_f} x_{fm}; \bar{x} = \frac{1}{M} \sum_{f \in F} \sum_{m=1}^{M_f} x_{fm}. \quad (4)$$

where F is a set of expression classes, M_f facial-patterns are given for each class $f \in F$ and x_{fm} is an N -dimension vector of the m -th facial patterns, $m = 1, M_f$

To estimate emotion using EMC, we divide the training set into five classes then calculate the eigenvectors and eigen-values of the matrix S and compute the eigen-space. For each test thermal data, we project it into the eigen-space of each class and an emotion is chosen if it gets maximum of cosine of angle between obtained vector after projection and eigenvector of each class. With PCA-EMC, we use PCA to reduce the dimension and then apply EMC to the obtain data.

Table 3. Confusion matrices of expression analysis of visible images with PCA

	Ag	Ha	Fe	Ne	Sa
Ag	78.23	6.06	7.51	8.20	-
Ha	2.30	46.62	8.35	40.08	2.65
Fe	5.46	5.39	82.80	6.35	-
Ne	5.96	34.79	5.92	36.31	17.02
Sa	-	12.68	3.65	14.09	69.58
Avg	62.71				

Table 4. Confusion matrices of expression analysis of visible images with PCA-EMC

	Ag	Ha	Fe	Ne	Sa
Ag	84.42	5.00	9.33	1.25	-
Ha	6.73	51.80	2.67	36.26	2.54
Fe	3.30	2.69	86.23	5.83	1.95
Ne	6.78	32.73	7.05	42.62	10.82
Sa	-	3.83	0.67	5.71	89.79
Avg	70.97				

4.1 Evaluation of the Visible Image Database

The preprocessing for visible image is to normalize image to facial space containing only face. The three well-know algorithms are used to classify images to emotions. There are PCA [16], EMC[8] and PCA-EMC methods. We extracted the visible image database from KTFE database. It includes 330 images of 22 subjects for 5 expressions. The rate of testing and training is 40 percent and 60 percent of total images, respectively. Table 3, 4, 5, including confusion matrices and average recognitions, shows the performance of these algorithms with respect to our visible image database. According to confusion matrices, anger

Table 5. Confusion matrices of expression analysis of visible images with EMC

	Ag	Ha	Fe	Ne	Sa
Ag	71.88	6.05	8.84	13.23	-
Ha	5.04	42.36	7.76	41.03	3.81
Fe	0.98	9.22	74.51	12.35	2.94
Ne	9.08	39.44	8.58	28.13	14.77
Sa	-	7.08	1.76	13.7	77.46
Avg	58.87				

Table 6. Confusion matrices of emotion classification of thermal data with PCA

	Ag	Ha	Fe	Ne	Sa
Ag	59.17	0.00	22.50	0.00	18.33
Ha	0.00	61.86	17.61	6.80	13.73
Fe	9.92	13.58	57.34	1.54	17.62
Ne	10.19	9.51	3.08	69.82	7.40
Sa	5.96	22.21	2.35	4.35	65.12
Avg	62.66				

Table 7. Confusion matrices of emotion classification of thermal data with PCA-EMC

	Ag	Ha	Fe	Ne	Sa
Ag	81.33	0.00	18.67	0.00	0.00
Ha	0.00	65.98	10.49	10.76	12.77
Fe	9.60	11.46	54.35	7.83	16.76
Ne	17.01	16.05	5.17	56.95	4.82
Sa	8.31	14.07	2.23	7.79	67.60
Avg	65.24				

expression do not recognize by sadness and vice versa. There are some confusion between happiness and neutral because some persons do not reveal their smiling.

4.2 Evaluation of the Thermal Data

To classify emissions using thermal data, we used PCA, PCA-EMC. The thermal data was extracted from KTFE. It includes 3.5GB thermal data for five emotions. The training and testing data are 40 percent and 60 percent of the total thermal data. From table 6, 7, 8, no instance of anger is incorrectly recognized as happiness. Neutral emotion and anger emotion are the most recognized by PCA and PCA-EMC, respectively. Based on the results, we confirm that thermal data are important addition information to support for expressions and emotions analysis.

Table 8. Confusion matrices of emotion classification of thermal data with EMC

	Ag	Ha	Fe	Ne	Sa
Ag	81.33	0.00	14.67	4.00	0.00
Ha	2.71	64.64	12.03	4.12	16.50
Fe	5.88	12.28	55.60	1.18	25.06
Ne	5.17	10.67	9.97	58.29	15.90
Sa	1.36	11.26	12.83	10.26	64.29
Avg	64.83				

Table 9. Confusion matrices of expression analysis of thermal images with PCA

	Ag	Ha	Fe	Ne	Sa
Ag	64.14	3	8.57	3	21.29
Ha	2.73	74.79	3.97	12.44	6.07
Fe	3.01	3.72	78.37	11.49	3.41
Ne	1.25	7.20	5.08	86.47	-
Sa	5.84	-	-	4.82	89.34
Avg	78.62				

Table 10. Confusion matrices of expression analysis of thermal images with PCA-EMC

	Ag	Ha	Fe	Ne	Sa
Ag	92.28	-	4.29	-	3.43
Ha	-	83.82	3.03	9.26	3.89
Fe	-	3.86	90.58	5.56	-
Ne	-	9.90	1.36	88.74	-
Sa	3.76	6.29	-	7.80	82.15
Avg	87.51				

Table 11. Confusion matrices of expression analysis of thermal images with EMC

	Ag	Ha	Fe	Ne	Sa
Ag	83.08	-	-	-	16.92
Ha	-	99.36	0.06	0.5	0.08
Fe	-	-	100	-	-
Ne	-	-	-	100	-
Sa	6.54	-	-	0.07	93.39
Avg	95.17				

4.3 Evaluation of the Thermal Image Database

Before classifying images to emotions, the preprocessing is similar to the preprocessing for the visible images. The PCA, EMC, PCA-EMC also used to classify the emotions. In this experiment, we use 330 images of 22 subjects for 5 expressions. We used 40 percent images and 60 percent images in total images to test

and train, respectively. From table 9, 10, 11, EMC method, suitable for thermal images, gave very high classify rate.

5 Conclusions

We proposed a KTFE database in this research for expression and emotion recognition, which has several advantages: Firstly, this is one of the first natural spontaneous visible and thermal videos. These databases will allow researchers on facial expressions and emotions to have more approaches more realistic; Secondly, this database already fixed some mistakes which the former database met when they did experiment settings such as the time lag phenomenon; Thirdly, we also had several researches in our data and obtained some results to support researchers using our database. The results on thermal data give us a promising future on facial research. In future, we will continue developing our data and working on thermal data to contribute better results.

References

1. Darwin, C.: The expression of the emotions in man and animals. Oxford University, New York (1872)
2. Tian, Y., Kanade, T., Cohn, J.: Facial Expression Analysis, Handbook of Face Recognition, pp. 247–275. Springer, New York (2005)
3. Zeng, Z., Pantic, M., Roisman, G.T., Huang, T.S.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31(1), 39–58 (2009)
4. Fasel, B., Luetttin, J.: Automatic facial expression analysis: a survey. *Pattern Recognition* 36, 259–275 (2003)
5. Khan, M.M., Ward, R.D., Ingleby, M.: Classifying pretended and evoked facial expression of positive and negative affective states using infrared measurement of skin temperature. *Trans. Appl. Percept.* 6(1), 1–22 (2009)
6. Nakanishi, R., Matsumura, K.I.: Facial skin temperature decreases in infants with joyful expression. *Infant Behavior and Development* 31, 137–144 (2008)
7. Kanade, T., Cohn, J., Tian, Y.: Comprehensive database for facial expression analysis. In: *Proc. IEEE Intl. Conf. Face and Gesture Recognition*, pp. 46–53 (2000)
8. Sebe, N., Lew, M.S., Cohen, I., Sun, Y., Gevers, T., Huang, T.S.: Authentic Facial Expression Analysis. In: *Proc. IEEE Intl. Conf. Automatic Face and Gesture Recognition, AFGR* (2004)
9. Pantic, M., Bartlett, M.S.: Machine Analysis of Facial Expressions, *Face Recognition*, pp. 377–416. I-Tech Education and Publishing, Vienna (2007)
10. O’Toole, A.J., Harms, J., Snow, S.L., Hurst, D.R., Pappas, M.R., Ayyad, J.H., Abdi, H.: A video database of moving faces and people. *IEEE Trans. Pattern Anal. Mach. Intell.* 27(5), 812–816 (2005)
11. Douglas-Cowie, E., Cowie, R., Schroeder, M.: The description of naturally occurring emotional speech. In: *Proc. 15th Int. Conf. Phonetic Sciences, Barcelona, Spain*, pp. 2877–2880 (2003)
12. Roisman, G.I., Tsai, J.L., Chiang, K.S.: The emotional integration of childhood experience: Physiological, facial expressive, and self-reported emotional response during the adult attachment interview. *Development Psychol.* 40(5), 776–789 (2004)

13. NIST Equinox database, <http://www.equinoxsensors.com/products/HID.html>
14. IRIS database, <http://www.cse.ohio-state.edu/otcbvs-bench/Data/02/download.html>
15. Wang, S., Liu, Z., Lv, S., Lv, Y., Wu, G., Peng, P., Chen, F., Wang, X.: A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference. *IEEE Transactions on Multimedia* 12(7), 682–691 (2010)
16. Lin, D.T.: Facial Expression Classification Using PCA and Hierarchical Radial Basis Function Network. *Journal of Information Science and Engineering* 22, 1033–1046 (2006)
17. Kurozumi, T., Shinza, Y., Kenmochi, Y., Kotani, K.: Facial Individuality and Expression Analysis by Eigenspace Method Based on Class Features or Multiple Discriminant Analysis. In: *ICIP* (1999)
18. Yabui, T., Kenmochi, Y., Kotani, K.: Facial expression analysis from 3D range images; comparison with the analysis from 2D images and their integration. In: *ICIP* (2003)