

CAS(ME)²: A Database of Spontaneous Macro-expressions and Micro-expressions

Fangbing Qu^{1,2}, Su-Jing Wang³, Wen-Jing Yan⁴, and Xiaolan Fu¹(✉)

¹ State Key Laboratory of Brain and Cognitive Science, Institute of Psychology,
Chinese Academy of Sciences, Beijing, China

{qufb, fuxl}@psych.ac.cn

² University of Chinese Academy of Sciences, Beijing, China

³ Key Laboratory of Behavior Sciences, Institute of Psychology,
Chinese Academy of Sciences, Beijing, China

wangsujing@psych.ac.cn

⁴ Institute of Psychology and Behavioral Sciences,
Wenzhou University, Wenzhou, China

yanwj@wzu.edu.cn

Abstract. Micro-expressions are facial expressions that are characterized by short durations, involuntary generation and low intensity, and they are regarded as unique cues revealing one's hidden emotions. Although methods for the recognition of general facial expressions have been intensively investigated, little progress has been made in the automatic recognition of micro-expressions. To further facilitate development in this field, we present a new facial expression database, CAS(ME)², which includes 250 spontaneous macro-expression samples and 53 micro-expression samples. The CAS(ME)² database offers both macro-expression and micro-expression samples collected from the same participants. The emotion labels in the current database are based on a combination of Action Units (AUs), self-reports of every facial movement and the emotion types of the emotion-evoking videos to improve the validity of the labeling. Baseline evaluation was also provided. This database may provide more valid and ecological expression samples for the development of automatic facial recognition systems.

Keywords: Micro-expression recognition · Macro-expression recognition · Facial action coding system · Emotion labeling

1 Introduction

Facial expressions convey much information regarding individuals' affective states, statuses, attitudes, and their cooperative and competitive natures in social interactions [1–5]. Human facial expression recognition has been a widely studied topic in computer vision since the concept of affective computing was first proposed by Picard [6]. Numerous methods and algorithms for automatically recognizing emotions from human faces have been developed [7]. However, previous research has focused primarily on general facial expressions, usually called macro-expressions, which typically last for more than 1/2 of a second, up to 4 s [8] (although some researchers treat the

duration of macro-expressions as between 1/2 s and 2 s [9]). Recently, another type of facial expression, namely, micro-expressions, which are characterized by their involuntary occurrence, short duration and typically low intensity, has drawn the attention of affective computing researchers and psychologists. Micro-expressions are rapid and brief expressions that appear when individuals attempt to conceal their genuine emotions, especially in high-stakes situations [10, 11]. A micro-expression is characterized by its duration and spatial locality [12, 13]. Ekman has even claimed that micro-expressions may be the most promising cues for lie detection [11]. The potential applications in clinical diagnosis, national security and interviewing that derive from the possibility that micro-expressions may reveal genuine feelings and aid in the detection of lies have encouraged researchers from various fields to enter this area. However, little work on micro-expressions has yet been performed in the field of computer vision.

As mentioned above, much work has been conducted regarding the automatic recognition of macro-expressions. This progress would not have been possible without the construction of well-established facial expression databases, which have greatly facilitated the development of facial expression recognition (FER) systems. To date, numerous facial expression databases have been developed, such as the Japanese Female Facial Expression Database (JAFFE) [14]; the CMU Pose, Illumination and Expression (PIE) database [15] and its successor, the Multi-PIE database [16]; and the Genki-4 K database [17]. However, all of the databases mentioned above contain only still facial expression images that represent different emotional states.

It is clear that still facial expressions contain less information than do dynamic facial expression sequences. Therefore, researchers have shifted their attention to dynamic factors and have developed several databases that contain dynamic facial expressions, such as the RU-FACS database [18], the MMI facial expression database [19], and the Cohn-Kanade database (CK) [20]. These databases all contain facial expression image sequences, which are more efficient for training and testing than still images. However, all of the facial expression databases mentioned above collect only posed expressions (i.e., the participants were asked to present certain facial expressions, such as happy expressions, sad expressions and so on) rather than naturally expressed or spontaneous facial expressions. Previous literature reports have stated that posed expressions may differ in appearance and timing from spontaneously occurring expressions [21]. Therefore, a database containing spontaneous facial expressions would have more ecological validity.

To address the issue of facial expression spontaneity, Lucey et al. [22] developed the Extended Cohn-Kanade Dataset (CK+) by collecting a number of spontaneous facial expressions; however, this dataset included only happy expressions that spontaneously occurred between the participants' facial expression posing tasks (84 participants smiled at the experimenter one or more times between tasks, not in response to a task request). Recently, McDuff et al. [23] presented the Affectiva-MIT Facial Expression Dataset (AM-FED), which contains naturalistic and spontaneous facial expressions collected online from volunteers recruited on the internet who agreed to be videotaped while watching amusing Super Bowl commercials. This database achieved further improved validity through the collection of spontaneous expression samples recorded in natural settings. However, only facial expressions that were understood to

be related to a single emotional state (amusement) were recorded, and no self-reports on the facial movements were collected to exclude any unemotional movements that may contaminate the database, such as blowing of the nose, swallowing of saliva or rolling of the eyes. Zhang et al. [24] recently published a new facial expression database of 3D spontaneous elicited facial expressions, coded according to the Facial Action Coding System (FACS).

Compared with research on macro-expression recognition, however, studies related to micro-expression recognition are rare. Polikovsky et al. [25] used a 3D-gradient descriptor for micro-expression recognition. Wang et al. [26] treated a gray-scale micro-expression video clip as a 3rd-order tensor and applied Discriminant Tensor Subspace Analysis (DTSA) and Extreme Learning Machine (ELM) approaches to recognize micro-expressions. However, the subtle movements involved in micro-expressions may be lost in the process of DTSA. Pfister et al. [27] used a temporal interpolation model (TIM) and LBP-TOP [28] to extract the dynamic textures of micro-expressions. Wang et al. [29] used an independent color space to improve upon this work, and they also used Robust PCA [30] to extract subtle motion information regarding micro-expressions. These algorithms require micro-expression data for model training. To our knowledge, only six micro-expression datasets exist, each with different advantages and disadvantages (see Table 1): USF-HD [9]; Polikovsky’s database [25]; SMIC [27] and its successor, an extended version of SMIC [31]; and CASME [32] and its successor, CASME II [33]. The quality of these databases can be assessed based on two major factors.

These micro-databases have greatly facilitated the development of automatic micro-expression recognition. However, all these databases only include cropped micro-expression samples, which was not very suitable for automatic micro-expression spotting. Besides, the methods used for emotion labeling was not consistent, which usually labeled the emotion according to FACS or the emotion type of elicitation materials or both. These methods left the possibility that may include some meaningless

Table 1. Previous micro-expression databases

Database	Number	Estimated emotions	Elicitation	Tagging		
USD-HD	100	6 (basic emotions)	Posed	Micro/non-micro		
Polikovsky’s database	42	6 (basic emotions)	Posed	FACS		
SMIC	77	3 (positive/negative/surprise)	Spontaneous	Emotion category		
SMIC extension*	HS	164	3 (positive/negative/surprise)	Spontaneous	Emotion category	
	VIS					71
	NIR					71
CASME	195	7 (basic emotions)	Spontaneous	Emotion/FACS		
CASME II	247	5 (basic emotions)	Spontaneous	Emotion/FACS		

* The SMIC extension database includes three subsets containing expressions recorded using different cameras: HS represents expression samples recorded using a high-speed camera at 100 fps, VIS corresponds to a normal visible-light camera, and NIR corresponds to a near-infrared camera. The latter two cameras were both operated at 25 fps.

facial movements into the database (such as the blowing of the nose, swallowing of saliva or rolling of the eyes). In this database, together with the FACS and emotion type of elicitation material, we collected participants' self-reports on each of their facial movements, which to the best extent guaranteed the purity of the database.

Considering the issues mentioned above, we developed a new database, CAS (ME)², for automatic micro-expression recognition training and evaluation. The main contributions of this database can be summarized as follows:

- This database is the first to contain both macro-expressions and micro-expressions collected from the same participants and under the same experimental conditions. This will allow researchers to develop more efficient algorithms to extract features that are better able to discriminate between macro-expressions and micro-expressions and to compare the differences in feature vectors between them.
- This database will allow the development of algorithms for detecting micro-expressions from video streams.
- The difference in AUs between macro-expressions and micro-expressions will also be able to be acquired through algorithms tested on this database.
- The method used to elicit both macro-expressions and micro-expressions for inclusion in this database has proven to be valid in previous work [33]. The participants were asked to neutralize their facial expressions while watching emotion-evoking videos. All expression samples are dynamic and ecologically valid.
- The database provides the AUs for each sample. In addition, after the expression-inducing phase, the participants were asked to watch the videos of their recorded facial expressions and provide a self-report on each expression. This procedure allowed us to exclude almost all emotion-irrelevant facial movements, resulting in relatively pure expression samples. The main emotions associated with the emotion-evoking videos were also considered in the emotion labeling process.

2 The CAS(ME)² Database Profile

The CAS(ME)² database contains 303 expressions — 250 macro-expressions and 53 micro-expressions — filmed at a camera frame rate of 30 fps. The expression samples were selected from more than 600 elicited facial movements and were coded with the onset, apex, and offset frames¹, with AUs marked and emotions labeled [35]. Moreover, to enhance the reliability of the emotion labeling, we obtained an additional emotion label by asking the participants to review their facial movements and report the emotion associated with each.

Macro-expressions with durations of more than 500 ms and less than 4 s were selected for inclusion in this database [8]. Micro-expressions with durations of no more than 500 ms were also selected.

¹ The onset frame is the first to change from the baseline (usually a neutral facial expression). The apex frame is the first to reach the highest observed intensity of the facial expression, and the offset frame is the first to reach the baseline.

The expression samples were recorded using a Logitech Pro C920 camera at 30 fps, and the resolution was set to 640×480 pixels. The participants were recorded in a room with two LED lights to maintain a fixed level of illumination. The steps of the data acquisition and coding processes are presented in following sections. Table 2 presents the descriptive statistics for the expression samples of different durations, which include 250 macro-expressions and 53 micro-expressions, defined in terms of total duration. Figure 1 shows examples of a micro-expression (a) and a macro-expression (b).

Table 2. Descriptive statistics for macro-expressions and micro-expressions

Expression type	N	Total duration	
		Mean (ms)	Standard deviation
Macro-expression	250	1305.87	657.6
Micro-expression	53	418.82	65.87

2.1 Participants and Elicitation Materials

Twenty-two participants (16 females), with a mean age of 22.59 years (standard deviation = 2.2), were recruited. All provided informed consent to the use of their video images for scientific research.

Seventeen video episodes used in previous studies [36] and 3 new video episodes downloaded from the internet were rated with regard to their ability to elicit facial expressions. Twenty participants rated the main emotions associated with these video clips and assigned them arousal intensity scores from 1 to 9 (1 represents the weakest intensity, and 9 represents the strongest intensity). Based on the ability to elicit micro-expressions observed in previous studies, only three types of emotion-evoking videos (those evoking disgust, anger, happiness) were used in this study, and they ranged in length from 1 min to approximately 2 and a half minutes. Five happiness-evoking video clips were chosen for their relatively low ability to elicit micro-expressions (see Table 3). Each episode predominantly elicited one type of emotion.

Table 3. Participant's ratings of the 9 video clips

Clip no.	Duration	Main emotion(s)	Rate of selection	Mean score
1	1'07"	Disgust	0.86	4.14
2	1'35"	Disgust	0.92	4.33
3	1'57"	Anger	0.75	4.5
4	2'24"	Anger	0.77	3.92
5	1'18"	Happiness	0.92	4
6	1'32"	Happiness	0.86	3.07
7	1'16"	Happiness	0.86	3.28
8	1'48"	Happiness	0.73	3.64
9	1'09"	Happiness	0.71	3.17

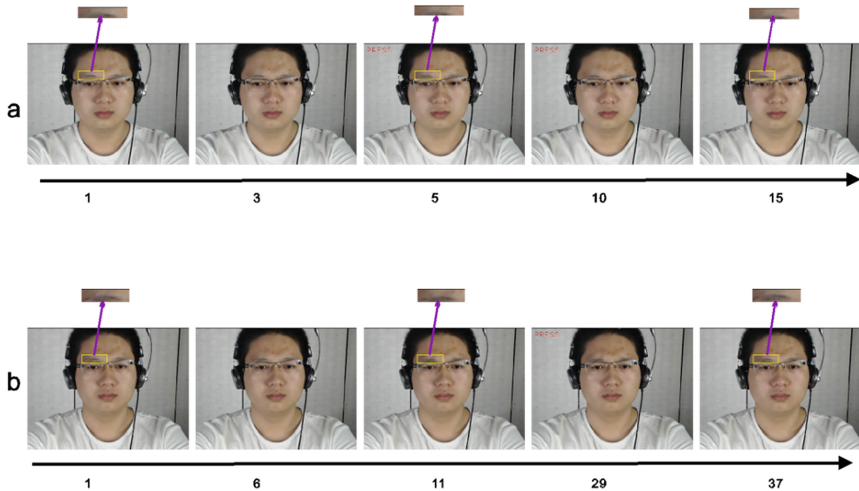


Fig. 1. Examples of a micro-expression (a) and a macro-expression (b). The apex frame occurs at approximately frame 5 for the micro-expression and frame 11 for the macro-expression, both of which represent the negative emotion anger. The AU related to these two expressions are all AU4 (inner brow).

2.2 Elicitation Procedure

Because the elicitation of micro-expressions requires rather strong motivation to conceal truly experienced emotions, motivation manipulation protocols were needed. Following a procedure used in previous studies [32], the participants were first instructed that the purpose of the experiment was to test their ability to control their emotions, which was stated to be strongly related to their social success. The participants were also told that their payment would be directly related to their performance. To better exclude noise arising from head movements, the participants were asked not to turn their eyes or heads away from the screen.

While watching the videos, the participants were asked to suppress their expressions to the best of their ability. They were informed that their monetary rewards would be reduced if they produced any noticeable expression. After watching all nine emotion-eliciting videos, the participants were asked to review the recorded videos of their faces and the originally presented videos to identify any facial movements and report the inner feelings that they had been experiencing when the expressions occurred. These self-reports of the inner feelings associated with each expression were collected and used as a separate tagging system, as described in the following section.

2.3 Coding Process

Two well-trained FACS coders coded each frame (duration = 1/30th of a second) of the videotaped clips for the presence and duration of emotional expressions in the upper and lower facial regions. This coding required classifying the emotion exhibited

in each facial region in each frame; recording the onset time, apex time and offset time of each expression; and arbitrating any disagreement that occurred between the coders. When the coders could not agree on the exact frame of the onset, apex or offset of an expression, the average of the values specified by both coders was used. The two coders achieved a coding reliability (frame agreement) of 0.82 (from the initial frame to the end). The coders also coded the Action Units (AUs) of each expression sample. The reliability between the two coders was 0.8, which was calculated as follows:

$$R = \frac{2 \times AU(C_1 C_2)}{AA}$$

where $AU(C_1 C_2)$ is the number of AUs on which Coder 1 and Coder 2 agreed and AA is the total number of AUs in the facial expressions scored by the two coders. The coders discussed and arbitrated any disagreements [33].

2.4 Emotion Labeling

Previous studies of micro-expressions have typically employed two types of expression tagging criteria, i.e., the emotion types associated with the emotion-evoking videos and the FACS. In databases with tagging based on the emotion types associated with the emotion-evoking videos, facial expression samples are typically tagged with the emotions happiness, sadness, surprise, disgust and anger [32] or with more general terms such as positive, negative and surprise [31]. The FACS is also used to tag expression databases. The FACS encodes facial expressions based on a combination of 38 elementary components, known as 32 AUs and 6 action descriptors (ADs) [34]. Different combinations of AUs describe different facial expressions: for example, AU6 + AU12 describes an expression of happiness. Micro-expressions are special combinations of AUs with specific local facial movements indicated by each AU. In the CAS(ME)² database, in addition to using the FACS and the emotion types associated with the emotion-evoking videos as tagging criteria, we also employed the self-reported emotions collected from the participants.

When labeling the emotions associated with facial expressions, previous researchers have usually used the emotion types associated with the corresponding emotion-evoking videos as the ground truth [31]. However, an emotion-evoking video may consist of multiple emotion-evoking events. Therefore, the emotion types estimated according to the FACS and the emotion types of the evoking videos are not fully representative, and many facial movements such as blowing of the nose, eye blinks, and swallowing of saliva may also be included among the expression samples. In addition, micro-expressions differ from macro-expressions in that they may occur involuntarily, partially and in short durations; thus, the emotional labeling of micro-expressions based only on the FACS AUs and the emotion types of the evoking videos is incomplete. We must also consider the inner feelings that are reflected in the self-reported emotions of the participants when labeling micro-expressions. In this database, a combination of AUs, video-associated emotion types and self-reports was considered to enhance the validity of the emotion labels. Four emotion categories are used in this database: positive, negative, surprise and other (see Table 4).

Table 4. Criteria for labeling each type of emotion and their frequencies in the database

Emotion category	Criteria	Number	Macro-expressions	Micro-expressions
Positive	AUs needed for Happiness, at least AU6 or AU12	93	87	6
Negative	AUs needed for Anger (AU4 + AU5), Disgust (AU9, AU10 or AU4 + AU7), Sadness (AU1), or Fear (AU1 + AU2 + AU4 or AU20)	114	95	19
Surprise	AUs needed for Surprise (AU1 + AU2, AU25, or AU2)	22	13	9
Other	Other facial movements*	74	55	19

*The ‘other’ category includes facial expressions that cannot be classified as related to basic emotions, such as those associated with tension and control, hurt, sympathy, confusion and helplessness. Consistent with previous studies, the emotion labeling in this database is partially based on the AUs because micro-expressions are typically partial and of low intensity. In addition, the participants’ self-reports and the contents of the video episodes were also considered.

Table 4 predominantly presents the emotion labeling criteria based on the FACS coding results; however, the self-reported emotions and the emotion types associated with the emotion-evoking videos were also considered during the emotion labeling process and are included in the CAS(ME)² database.

3 Dataset Evaluation

To evaluate the database, we used Local Binary Pattern histograms from Three Orthogonal Planes (LBP-TOP) [28] to extract dynamic textures and used a Support Vector Machine (SVM) approach to classify these dynamic textures.

Among the 303 samples, the number of frames included in the shortest sample is 4, and the longest sample contains 118 frames. The frame numbers of all samples were normalized to 120 via linear interpolation. For the first frame of each clip, 68 feature landmarks were marked using the Discriminative Response Map Fitting (DRMF) method [37]. Based on these 68 feature landmarks, 36 Regions of Interest (ROIs) were drawn, as shown in Fig. 2. Here, we used Leave-One-Video-Out (LOVO) cross-validation, i.e., in each fold, one video clip was used as the test set and the others were used as the training set. After the analysis of 302 folds, each sample had been used as the test set once, and the final recognition accuracy was calculated based on all results.

We extracted LBP-TOP histograms to represent the dynamic texture features for each ROI. Then, these histograms were concatenated into a vector to serve as an input to the classifier. An SVM classifier was selected. For the SVM algorithm, we used LIBSVM with a polynomial kernel, $\mathcal{K}(x_i, x_j) = (\gamma x_i^T x_j + coef)^{degree}$, $\gamma = 0.1$, with

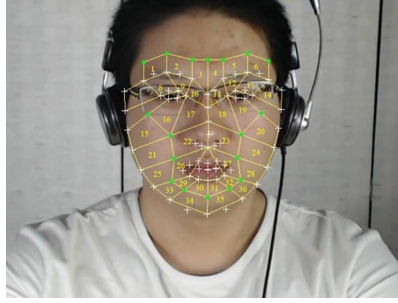


Fig. 2. Thirty-six Regions of Interest (ROIs).

$degree = 4$ and $coef = 1$. For the LBP-TOP analysis, the radii along the X and Y axes (denoted by R_x and R_y) were set to 1, and the radii along the T axis (denoted by R_t) were assigned various values from 2 to 4. The numbers of neighboring points (denoted by P) in the XY, XT and YT planes were all set to 4 or 8. The uniform pattern and the basic LBP were used in LBP coding. The results are listed in Table 5. As shown in the table, the best performance of 75.66 % was achieved using $R_x = 1$, $R_y = 1$, $R_t = 3$, $p = 4$, and the uniform pattern.

Table 5. Performance in the recognition of micro-expressions with LBP-TOP feature extraction

		Basic LBP		Uniform pattern	
		P = 4	P = 8	P = 4	P = 8
$R_x = 1, R_y = 1$	Rt = 2	71.05	65.79	71.05	65.13
	Rt = 3	75.00	65.46	75.66	65.13
	Rt = 4	74.67	67.76	74.01	66.45
$R_x = 2, R_y = 2$	Rt = 2	71.38	68.75	71.38	68.09
	Rt = 3	72.37	67.43	72.37	67.76
	Rt = 4	72.37	68.42	72.37	69.74
$R_x = 3, R_y = 3$	Rt = 2	70.72	69.41	70.72	69.41
	Rt = 3	72.04	69.74	72.04	70.39
	Rt = 4	74.01	69.41	73.68	71.71
$R_x = 4, R_y = 4$	Rt = 2	71.05	70.72	71.05	70.39
	Rt = 3	73.03	70.39	73.03	71.71
	Rt = 4	73.36	69.74	73.36	73.36

4 Discussion and Conclusion

In this paper, we describe a new facial expression database that includes macro-expression and micro-expression samples collected from the same individuals under the same experimental conditions. This database contains 303 expression samples,

comprising 250 macro-expression samples and 53 micro-expression samples. This database may allow researchers to develop more efficient algorithms to extract features that are better able to discriminate between macro-expressions and micro-expressions.

Considering the unique features of micro-expressions, which typically occur involuntarily, rapidly, partially (on either the upper face or the lower face) and with low intensity, the emotional labeling of such facial expressions based only on the corresponding AUs and the emotion types associated with the videos that evoked them may not be sufficiently precise. To construct the presented database, we also collected self-reports of the subjects' emotions for each expression sample and performed the emotion labeling of each sample based on a combination of the AUs, the emotion type associated with the emotion-evoking video and the self-reported emotion. This labeling method should considerably enhance the precision of the emotion category assignment. In addition, all three labels are independent of one another and will be accessible when the database is published in the future, to allow researchers to access specific expressions in the database.

In the current version of the database, because of the difficulties encountered in micro-expression elicitation and the extremely time-consuming nature of manual coding, the size of the micro-expression sample pool may not be fully sufficient. We plan to enrich the sample pool by eliciting more micro-expression samples to provide researchers with sufficient testing and training data.

Acknowledgments. This project was partially supported by the National Natural Science Foundation of China (61375009, 61379095) and the Beijing Natural Science Foundation (4152055). We appreciate Yu-Hsin Chen's suggestions in language.

References

1. DePaulo, B.M.: Nonverbal behavior and self-presentation. *Psychol. Bull.* **111**(2), 203 (1992)
2. Ekman, P., Friesen, W.V.: Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* **17**(2), 124 (1971)
3. Friedman, H.S., Miller-Herringer, T.: Nonverbal display of emotion in public and in private: Self-monitoring, personality, and expressive cues. *J. Pers. Soc. Psychol.* **61**(5), 766 (1991)
4. North, M.S., Todorov, A., Osherson, D.N.: Inferring the preferences of others from spontaneous, low-emotional facial expressions. *J. Exp. Soc. Psychol.* **46**(6), 1109–1113 (2010)
5. North, M.S., Todorov, A., Osherson, D.N.: Accuracy of inferring self-and other-preferences from spontaneous facial expressions. *J. Nonverbal Behav.* **36**(4), 227–233 (2012)
6. Picard, R.W.: *A. Computing, and M. Editura.* MIT Press, Cambridge, MA (1997)
7. Tong, Y., Chen, J., Ji, Q.: A unified probabilistic framework for spontaneous facial action modeling and understanding. *Pattern Anal. Mach. Intell.* **32**(2), 258–273 (2010)
8. Ekman, P.: *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life.* Macmillan, New York (2007)
9. Shreve, M., et al.: Macro-and micro-expression spotting in long videos using spatio-temporal strain. In: 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops (FG 2011). IEEE (2011)

10. Ekman, P., Friesen, W.V.: Nonverbal leakage and clues to deception. *Psychiatry* **32**(1), 88–106 (1969)
11. Ekman, P.: *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. WW Norton & Company, New York (2009). (Revised Edition)
12. Porter, S., Ten Brinke, L.: Reading between the lies identifying concealed and falsified emotions in universal facial expressions. *Psychol. Sci.* **19**(5), 508–514 (2008)
13. Rothwell, J., et al.: Silent talker: a new computer-based system for the analysis of facial cues to deception. *Appl. Cogn. Psychol.* **20**(6), 757–777 (2006)
14. Lyons, M., et al.: Coding facial expressions with Gabor wavelets. In: 1998 Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition. IEEE (1998)
15. Sim, T., Baker, S., Bsat, M.: The CMU pose, illumination, and expression (PIE) database. In: 2002 Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition. IEEE (2002)
16. Gross, R., et al.: Multi-pie. *Image Vis. Comput.* **28**(5), 807–813 (2010)
17. Whitehill, J., et al.: Toward practical smile detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(11), 2106–2111 (2009)
18. Bartlett, M.S., et al.: Automatic recognition of facial actions in spontaneous expressions. *J. Multimedia* **1**(6), 22–35 (2006)
19. Pantic, M., et al.: Web-based database for facial expression analysis. In: 2005 IEEE International Conference on Multimedia and Expo ICME 2005. IEEE (2005)
20. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition. IEEE (2000)
21. Schmidt, K.L., Cohn, J.F.: Human facial expressions as adaptations: evolutionary questions in facial expression research. *Am. J. Phys. Anthropol.* **116**(S33), 3–24 (2001)
22. Lucey, P., et al.: The extended cohn-kanade dataset (CK +): a complete dataset for action unit and emotion-specified expression. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE (2010)
23. McDuff, D., et al.: Affectiva-MIT facial expression dataset (AM-FED): naturalistic and spontaneous facial expressions collected “In-the-Wild”. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE (2013)
24. Zhang, X., et al.: BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database. *Image Vis. Comput.* **32**(10), 692–706 (2014)
25. Polikovsky, S., Kameda, Y., Ohta, Y.: Facial micro-expressions recognition using high speed camera and 3D-gradient descriptor (2009)
26. Wang, S.-J., et al.: Face recognition and micro-expression recognition based on discriminant tensor subspace analysis plus extreme learning machine. *Neural Process. Lett.* **39**(1), 25–43 (2014)
27. Pfister, T., et al.: Recognising spontaneous facial micro-expressions. In: 2011 IEEE International Conference on Computer Vision (ICCV). IEEE (2011)
28. Zhao, G., Pietikainen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(6), 915–928 (2007)
29. Wang, S.-J., et al.: Micro-expression recognition using dynamic textures on tensor independent color space. In: Pattern Recognition (ICPR). IEEE (2014)
30. Wang, S.-J., Yan, W.-J., Zhao, G., Fu, X., Zhou, C.-G.: Micro-expression recognition using robust principal component analysis and local spatiotemporal directional features. In: Agapito, L., Bronstein, M.M., Rother, C. (eds.) ECCV 2014 Workshops. LNCS, vol. 8925, pp. 325–338. Springer, Heidelberg (2015)

31. Li, X., et al.: A spontaneous micro-expression database: Inducement, collection and baseline. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). IEEE (2013)
32. Yan, W.-J., et al. CASME database: a dataset of spontaneous micro-expressions collected from neutralized faces. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). IEEE (2013)
33. Yan, W.-J., et al.: CASME II: an improved spontaneous micro-expression database and the baseline evaluation. PLoS ONE **9**(1), e86041 (2014)
34. Ekman, P., Friesen, W., Hager, J.: Facial Action Coding System: The Manual on CD-ROM Instructor's Guide. Network Information Research Co, Salt Lake City (2002)
35. Yan, W.-J., et al.: For micro-expression recognition: database and suggestions. Neurocomputing **136**, 82–87 (2014)
36. Yan, W.-J., et al.: How fast are the leaked facial expressions: the duration of micro-expressions. J. Nonverbal Behav. **37**(4), 217–230 (2013)
37. Asthana, A., et al.: Robust discriminative response map fitting with constrained local models. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE (2013)