A Spontaneous Micro-expression Database: Inducement, Collection and Baseline

Xiaobai Li, Tomas Pfister, Xiaohua Huang, Guoying Zhao, Matti Pietikäinen

Abstract— **Micro-expressions are short, involuntary facial expressions which reveal hidden emotions. Micro-expressions are important for understanding humans' deceitful behavior. Psychologists have been studying them since the 1960's. Currently the attention is elevated in both academic fields and in media. However, while general facial expression recognition (FER) has been intensively studied for years in computer vision, little research has been done in automatically analyzing micro-expressions. The biggest obstacle to date has been the lack of a suitable database. In this paper we present a novel Spontaneous Micro-expression Database SMIC, which includes 164 micro-expression video clips elicited from 16 participants. Micro-expression detection and recognition performance are provided as baselines. SMIC provides sufficient source material for comprehensive testing of automatic systems for analyzing micro-expressions, which has not been possible with any previously published database.**

I. INTRODUCTION

A. Micro-expression and Lie Detection

A micro-expression is a very brief facial expression which shows the true undergoing emotions that people try to conceal. The most significant characteristic of micro-expressions is that, compared to normal facial expressions, their durations are very short. Thus, usually it is hard to find them in real-time conversations [1].

Micro-expressions were first found by Haggard and Isaacs in 1966 [2]. Ekman and Friesen [3] also reported finding micro-expressions in their independent work three years later. They examined a film taken of a psychiatric patient, who later confessed that she lied to her doctor to conceal her plan to commit suicide. Although the patient seemed to be happy throughout the video, when the video was examined frame by frame, a hidden look of anguish was found which lasted for only two frames (1/12 s). In his next papers [4]–[6], Ekman indicated that micro-expressions are important clues for detecting lies, as they usually occur in high-stake situations where people would have serious consequences if they were caught lying or cheating –for example when the suspect is interrogated by a police officer. He also developed the Micro Expression Training Tool (METT) [7], [8] which can help to increase people' micro-expression detection performance. Later many other researchers joined in to analyze micro-expression and its relationship with lies. The public interest in micro-expressions has been growing

Xiaobai Li, Tomas Pfister, Xiaohua Huang, Guoying Zhao and Matti Pietikäinen ({lxiaobai, tomas.pfister, huang.xiaohua, gyzhao, mkp}@ee.oulu.fi) are with the Centre for Machine Vision Research, University of Oulu, Finland. Tomas Pfister (tp@robots.ox.ac.uk) is also with the Visual Geometry Group, University of Oxford, UK.

TABLE I: Popular databases for general facial expression recognition. AUs indicates action units of the Facial Action Coding System (FACS) [11].

Name	Participants	Size	Emotions	Elicitation
CK [12]	210 adults	480 videos	6 classes	Posed
MMI [13]	79 adults + 11 childrens	1250 videos + 600 images	6 classes + AUs	Posed + Spontaneous
JAFFE [14]	10 females	213 images	6 classes + neutral	Posed
RU-FACS [15]	100 adults	100 videos	AUs	Spontaneous
Multi-PIE [16]	100 adults	4200 images	6 classes	Posed
Oulu-CASIA [17]	80 adults	2880 videos	6 classes	Posed

both in scientific research and media. But in computer vision, micro-expression is still a quite new topic and not much work has been done.

B. The State of the Art in Computer Vision Research

After the idea of affective computing was brought up by Picard in 1998 [9], facial expression recognition has become a popular topic. Many good methods have been developed for automatically recognizing humans' emotional state from facial expressions. The development of these methods would not have been possible without well-established databases. For the general facial expression recognition (FER) problem, there are many widely used databases including both posed facial expressions and spontaneous (genuine) facial expressions. Some of the most popular databases are listed in Table 1. For more related work please refer to a review paper by Zeng et al. [10].

In contrast to the large number of FER research publications, only a few studies have been done on recognizing micro-expressions. Shreve et al. [18], [19] used strain patterns as a feature descriptor for spotting micro-expressions in videos. They claimed that their algorithms successfully spotted seven out of seven micro-expression samples with one false alarm. However, their dataset contains posed rather than natural micro-expressions, and their criteria set for micro-expressions (2/3 s) is longer than most accepted durations. Although there is no strict rule about the duration of micro-expressions, most agreed that they should be no longer than 1/2 s [20]. Polikovsky et al. [21] recorded micro-expressions from ten participants using a 200fps high speed camera, and used gradient orientation histograms as a descriptor for micro-expression recognition. But their data was also posed. They asked participants to perform seven basic emotions with

low facial muscle intensity and to go back to neutral as fast as possible. In another more recent work, Wu et al. [22] used GentleSVM, which is a combination of Gentleboost algorithm and SVM classifier, for spotting and recognising micro-expressions. But again, their methods were not tested on real micro-expressions. Instead, they used training videos from METT [7], [8], which are synthetic clips made by embedding one emotional frame in a sequence of neutral frames.

Although the authors in the above-mentioned micro-expression studies claimed to make progress in automatically analyzing micro-expressions, none of their works were actually tested on genuine (spontaneous) micro-expressions. As micro-expressions are involuntary behavior, theoretically they cannot be controlled or imitated [6]. One can obtain a 'posed micro-expression' by asking participants to pose, or by synthesizing or down-sampling the video somehow, but the result would be different from genuine micro-expressions, both spatially and temporally. In our previous work [23], spontaneous micro-expression clips were used, but only six participants' data were collected at that time. To better promote research on automatic micro-expression recognition a larger spontaneous micro-expression database is needed.

C. Motivation for Building a Spontaneous Micro-expression Database

Studying micro-expressions is important for understanding human's affective behavior, particularly when deceit is involved. For most people it is very difficult to recognize micro-expressions even after special training [24]. One reason is that micro-expressions are very short; another is that during a conversation, many other social factors are competing for attention at the same time – for example the speaker's words, gestures and also one's own plan for what to say next. For a computer none of these issues is a problem. Hence a method for automatically recognizing micro-expressions would be very valuable if it can be developed.

So far, the lack of an accessible testing resource has impeded the research. A well-established database with a sufficient amount of data is in need. To address this problem, in the following sections we will first describe the method used to collect the Spontaneous Micro-expression Database (SMIC) and outline its composition. We will then provide some basic testing results of the data as the baseline performance measure.

II. THE SPONTANEOUS MICRO-EXPRESSION DATABASE (SMIC)

Many methods have been employed to build facial expression databases, some of which were described in the previous section. But none of those methods is suitable for building a spontaneous micro-expression database. Two factors must be considered for this purpose: 1) Since micro-expressions are involuntary, genuine facial expressions must be elicited. This means that the participant needs to genuinely experience the emotion. 2) The facial expressions must also be very

brief so that they meet the criteria of a micro-expression. Genuine expressions can be induced by different material like images, movies, music etc. [25]. According to Ekman's research [4]–[6], micro-expressions appear when people are lying or trying to conceal their true feelings, especially when the consequences of being caught will be very serious (high-stake lies).

Thus, the experimental design for getting spontaneous micro-expressions must satisfy the following two conditions: 1) Stimuli used for eliciting participants' emotions must be strong enough, so that most participants will feel the emotion and show it on their faces. 2) There must be enough pressure to motivate participants to conceal their true feelings. Based on these considerations we designed an inhibited emotion inducing paradigm for building the SMIC database.

A. Acquisition Setup and Procedure

1) Setup: The database was recorded in an indoor bunker environment designed to resemble an interrogation room. Indoor illumination was controlled stable through the whole data recording period with four lights from the four upper corners of the room. 16 carefully selected movie clips (more details in the following section), which can induce strong emotions, were shown to participants on a computer monitor together with a speaker for audio output. Participants sat about 50cm from the computer monitor. While participants were watching the film clips, a camera fixed on top of the monitor recorded their facial reactions. The setup is illustrated in Figure 1.

20 participants participated in the recording experiment. For the recording of the first ten participants, a high speed (HS) camera (PixeLINK PL-B774U, 640×480) of 100fps was used to record the short duration of micro-expressions. For the recording of the latter ten participants (recorded 5 months later than the first ten participants), in addition to the high speed camera another integrated camera box was added (details about the camera provided in [26]), which consists of a normal visual camera (VIS) and a near-infrared (NIR) camera, both with 25 fps and resolution of 640×480. The VIS and NIR cameras were added for two reasons: first, to improve the diversity of the database; second, to investigate whether the current method can also be used on normal speed cameras of 25 fps. In contrast to a down-sampled version of the 100 fps data, the 25 fps data yields data similar to standard web cameras, including their limitations such as motions blurs. When multiple cameras were used, they were put parallel to each other and fixed on the middle top of the monitor to ensure frontal view recording. Due to technical issues there was a time delay about 3-5 seconds between the starting points of the three cameras. VIS and NIR clips were manually synchronized with the reference to the HS data.

2) Procedure: Before recording started, the content of research was explained to participants and forms were signed if they agreed. Then, instructions were given as follows: 1) You will be shown several short emotional film clips, please try to keep your head position fixed and watch them carefully. 2) After each film clip you will have a short break. Please
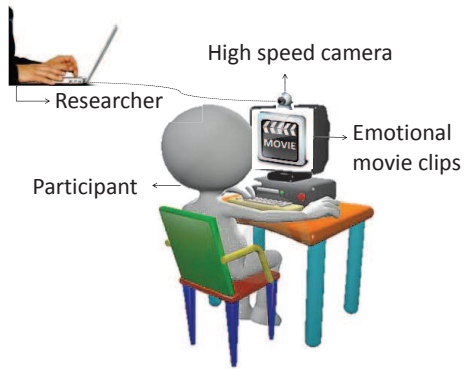
Fig. 1: SMIC recording system setup. Participants watch the movie alone in the recording room while the researcher monitors the participants' facial expressions remotely from another room.

TABLE II: Video clips selected as emotion inducers for the spontaneous micro-expression database SMIC.

Clip Name	Source	Emotion	Time
Pink flamingos	Film	Disgust	50s
Hellraiser	Film	Disgust	87s
The champ	Film	Sad	258s
Lion king	Film	Sad	120s
Shining	Film	Fear	343s
The thing	Film	Fear	204s
Capricorn	Film	Surprise	47s
Sea of love	Film	Surprise	9s
Funny moments 1	YouTube	Happy	112s
Funny cats	YouTube	Happy	141s
There's something about Mary	Film	Happy	141s
Church	YouTube	Happy	84s
When Harry met Sally	Film	Happy	173s
Italian went to Malta	YouTube	Happy	73s
Funny moments 2	YouTube	Happy	148s
Benny & Joon	Film	Happy	131s

fill in the questionnaire according to your true feelings about the film you just watched. 3) While you watch the films, I will stay in the other room observing your facial and body movements through the camera and try to guess which clip you are watching (clips are played in random order). Your task is to put on a poker face and NOT to reveal your true feelings. If you fail to hide your feelings so that I guess right, you will have to fill in a very long boring questionnaire of more than 500 questions.

After participants understood how to proceed, the researcher left and participants started formal experiments alone in the room. Questions in the self-report questionnaire include: 1) What kind(s) of emotion did you feel when watching the movie clip? Answers can be chosen from "happy, sad, disgust, fear, surprise, anger, or confused". 2) Do you feel pleasant or unpleasant when watching the movie clip? Answers were rated from 1 (most unpleasant) to 7 (most pleasant).

3) Stimuli Selection: Different sources and methods can be used for inducing spontaneous emotions. We chose to use short video clips for the SMIC database for three reasons: 1) Videos include both audio and visual information and are therefore more powerful emotion inducers. 2) Micro-expressions are more likely to occur if high-arousal emotions need to be hidden for a longer time [27], so videos are better than static images. 3) From the practical perspective of acquiring stable frontal face data, subjects watching movies are easier to control than subjects in interactive scenarios with multiple persons.

Before the formal experiment started, many emotional video clips were gathered from both commercial movies, which were demonstrated to elicit strong emotions in previous psychological studies [25], [28], [29], and YouTube clips. The strongest clips that induced six kinds of target emotions were selected. Anger movies were removed because the recordings did not elicit any micro-expressions of anger. Finally, 16 video clips were selected for the formal database recording. 11 clips were cut from commercial movies. Five funny videos were chosen from YouTube to even the number of positive (happy) and negative (sad, fear and disgust) video

clips in the database. Two clips were used for inducing surprise, but surprise can be induced by other clips as well which was demonstrated in our data labeling section. Details of selected video clips are given in Table 2. The recording time for each participant was about 50 minutes.

B. Dataset Description

1) Annotation of Micro-expression Video Clips: The HS data were segmented and labeled by two annotators according to participants' self-reported emotions. The annotators followed the advice of Ekman to first view the video frame-by-frame and then with increasing speed [1]. The excerpt micro-candidate clips were compared with participants' self-reported feelings. Only labels that were consistent with the subjects' self-reports were included. The two annotators' choices were then cross-checked. Clips for which both annotators agreed on the label were included to the final SMIC database.

2) Participants: 20 participants participated in the SMIC database experiment. Among those six were female and 14 male. All participants were either undergraduate students or research staff. The mean age of all participants is 26.7 years ranging from 22 to 34. The participants are ethnically diverse: ten Asians, nine Caucasians and one African. Ten participants wore glasses.

3) Duration of Micro-expressions: There is an ongoing debate in psychology as to the acceptable duration of micro-expressions. For the SMIC database we use $1/2$ second as the cut line [30], [31]. For the HS dataset of 100fps, the longest micro-expression clips have 50 frames while for VIS and NIR datasets of 25fps, the longest micro-expression clips have 13 frames. Every micro-expression clip starts from a neutral (or relatively neutral) frame, with the second frame as the starting point of expressional facial muscle movement, and ends when facial expression turns back to neutral (or relatively neutral).

TABLE III: Data composition of SMIC database. The number of participants is described as N1/N2, N1 indicates the number of participants that generated micro-expression clips during recording, N2 indicates the total number of participants recorded. Po. = positive (happy); Ne. = negative (sad, fear, disgust); Sur. = surprise.

Data	Participants	Micro-expression clips			
		Po.	Ne.	Sur.	Total
HS	**16/20**	51	70	43	**164**
VIS	**8/10**	28	23	20	**71**
NIR	**8/10**	28	23	20	**71**

4) Not All Subjects Show Micro-expressions: Not every participant showed micro-expressions when recording SMIC. According to Ekman's research, when people are telling lies, about half of them might show micro-expressions, while the other half do not [1]. The reason why only some people show micro-expressions is for now unknown. In our study, four participants did not show any micro-expression at all throughout the 35 minutes of watching videos; for the rest of 16 participants, the number of micro-expression clips ranges from 2 to 39.

5) Dataset Composition: The final SMIC database includes 164 micro-expression clips from 16 participants (mean age is 28.1 years, six females and ten males, eight Caucasians and eight Asians). All clips are recorded in the HS dataset. 71 clips from eight participants were also recorded in the VIS and NIR datasets. In the following sections only the participants who showed micro-expression clips will be counted when referring to the number of participants. For each participant's micro-expression clips, the same number of clips without micro-expressions ("non-micro clips", including neutral faces, general facial expressions and other non-expressional facial movements) were randomly selected from the original video as the counterparts for the use of the micro-expression detection task (details described in the following section). The composition of SMIC database is shown in Table 3. Sample clips of the SMIC database can be found on this web page [1].

III. DATABASE EVALUATION

A. Experiments

Two experiments were carried out on the SMIC database for analyzing micro-expressions: (1) micro-expression detection and (2) micro-expressions recognition. Detection was done by distinguishing micro-expression clips from randomly selected non-micro clips in a two-category classification task. Recognition was a three-category classification task to discriminate three classes of micro-expressions (Positive vs. Negative vs. Surprise).

In detection, for each participant an equal number of non-micro clips were chosen as the number of micro-expression clips he/she has yielded. Two guidelines were used to ensure that the non-micro clips were representative of background noise: 1) The non-micro clips were randomly selected from

[1] http://www.cse.oulu.fi/SMICDatabase

the whole recording time with the exception of the labeled micro-expression clip frames. This means that the non-micro clips can include all kinds of facial movements, e.g. general facial expressions and other movements such as eye blinking, lip sipping or slight head tilting. 2) The length of non-micro clips were controlled to be of similar distribution as the length of the micro-expression clips (shown in Fig.2).
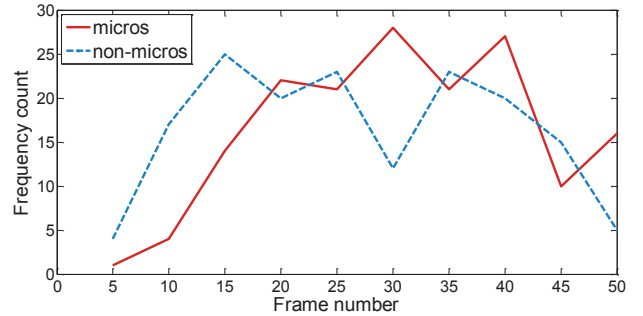
Fig. 2: Frame number distributions of micro-expression and non-micro clips in the HS dataset.

B. Methods

Since the muscle movements involved in spontaneous micro-expressions are very subtle, all faces were normalized and registered to a face model by using 68 feature points from the Active Shape Model [32] to address the variations of spatial appearance. After that, faces were cropped according to the eye positions from a Haar eye detector [33]. For feature extraction we used Local Binary Pattern histograms from Three Orthogonal Planes (LBP-TOP) [34] to describe the spatiotemporal local textures from cropped face sequences. Because the duration of micro-expressions is short, frame numbers can be very limited in some micro-expression clips especially for VIS and NIR datasets. That raises a problem when choosing parameters for LBP-TOP. To this end we used a temporal interpolation model (TIM) [35] which allows varying (up-sampling and down-sampling) the frame numbers to promote the feature extraction process. SVM was used as a classifier. The performance of the two experiments was computed using leave-one-subject-out cross validation. For more details on the methods refer to [23]. The past work on micro-expression recognition [23] applied this method on only six participants' data using the same methods for feature extraction as in this paper.

C. Results and Discussion

1) Baseline Performance: Both detection and recognition experiments were run on three datasets (HS, VIS and NIR) separately using the leave-one-subject-out cross validation. For the detection experiment the chance level is 50%; and for the recognition experiment the chance level is 33%. All

data (164 clips for HS, 71 clips for VIS and NIR) of the SMIC database were included in this part. Table 4 shows the performance of the three datasets. Parameter combinations that gave the best performance of each condition are listed.

TABLE IV: Leave-one-subject-out results on the SMIC database. Detection is classifying micro-expressions vs. non-micro; recognition is 3-category classification of Positive vs. Negative vs. Surprise. $x \times y \times t$ are the number of rows, columns and temporal blocks into which the LBP-TOP is divided. TIMn denotes that the original clips were temporally interpolated to n frames.

Dataset	Experiment	Method	Accuracy
HS	Detection	LBP-TOP($5 \times 5 \times 1$) + TIM10	65.55 %
	Recognition	LBP-TOP($8 \times 8 \times 1$) + TIM10	48.78 %
VIS	Detection	LBP-TOP($5 \times 5 \times 1$)	62.68 %
	Recognition	LBP-TOP($5 \times 5 \times 1$) + TIM10	52.11 %
NIR	Detection	LBP-TOP($8 \times 8 \times 1$) + TIM20	59.15 %
	Recognition	LBP-TOP($8 \times 8 \times 1$) + TIM10	38.03 %

2) Details of Experimental Results on Three Datasets: The second half of the data was also recorded using a near-infrared and standard frame-rate camera. This was to compare the performance of data that acquired 1) using high speed vs. normal speed recording and 2) within visual vs. near infrared spectrum. The first eight participants' clips in the HS dataset were excluded because there are no counterparts in VIS or NIR datasets for them, leaving 71 clips from the latter 8 participants. Same parameters were used for computing on all three datasets: all clips were interpolated to ten frames (TIM10), and the LBP-TOP features were extracted in $8 \times 8 \times 1$ blocks. Results are shown in Table 5. For both detection and recognition tasks, the performance on HS dataset is better than on the other two datasets that were recorded at normal 25fps speed. This indicates that the high speed camera did provide useful information for micro-expression analysis. When comparing the accuracy on the NIR and VIS datasets, the results on NIR dataset are much lower particularly for the recognition task. This may be because all videos of the SMIC database were recorded under normal indoor illumination conditions suitable for the VIS camera. A NIR camera shows its advantage over a VIS camera under darker illumination conditions [17].

TABLE V: Comparison of performance on the latter eight participants' data (71 clips) on three datasets. All clips were interpolated to ten frames (TIM10) and the LBP-TOP features were extracted in $8 \times 8 \times 1$ blocks. Tests are done using the leave-one-subject-out cross-validation.

Dataset	Experiment	Accuracy
HS	Detection	65.49 %
	Recognition	49.30 %
VIS	Detection	61.27 %
	Recognition	43.66 %
NIR	Detection	58.45 %
	Recognition	35.21 %

IV. FUTURE DIRECTIONS

We designed an emotion inducing protocol for building the spontaneous micro-expression database SMIC. Video clips that induce strong emotional reactions were shown to participants who were instructed not to reveal their true feelings on face – otherwise they would have to fill in a very long boring questionnaire as a punishment. This protocol simulated the high-stake lie condition and successfully induced micro-expressions in 16 out of 20 participants. Micro-expression clips were spotted in the recorded videos and labeled into three categories: positive, negative or surprise according to participants' validated self-reported emotions. SMIC provides sufficient sources for the study of micro-expression analysis, but there are still several improvements that could be done in the future.

A. Micro-expression Classes

Micro-expression clips in SMIC are now classified into three classes: positive, negative and surprise. While positive (happy) and surprise only includes one target emotion each, negative includes four emotion categories (sad, anger, fear and disgust). This kind of management was for the consideration of the data reliability. At the beginning of the database collection work we selected separate videos for eliciting each target emotion. We later found out that according to participants' self-report questionnaires, multiple kinds of emotion, especially negative ones, can appear together while watching one video. Since often only partial facial movements are involved in spontaneous micro-expressions, it is not as easy to judge which emotion it presents. For example, it is hard to say whether one clip represented fear or disgust if both emotions were reported by the participant. Under ambiguous conditions it is more convincing to classify a clip as a negative one. However, according to research in psychology [36] micro-expressions can express six universal emotions. Hence, future database construction could work towards more specific classes with experiments that give reliable evidence for ground truth.

B. Micro-expression Detection and Recognition Methods

Micro-expression detection was done by distinguishing micro-expressions from non-micro-expression clips. As a preliminary analysis of micro-expression data, this classification paradigm is necessary since it allows simple and direct descriptive presentations of the micro-expression data under laboratory controlled conditions. In the future, this needs to be extended to spotting micro-expressions from live-stream videos. That requires solutions for both detection and segmentation.

The performance reported in this paper serves as a baseline for future comparison. It uses basic methods and a single classifier for ease of repeatability. For the purpose of realizing an automatic lie detection application in scientific research or forensic interrogation, better methods are expected to be developed and tested on SMIC.

V. CONCLUSION

In this paper we have presented a novel spontaneous micro-expression database SMIC. It contains 164 micro-expression clips induced in 16 participants. Clips from all participants were recorded with a high speed 100fps camera. The clips of 8 participants (71 clips) were also recorded by normal-speed visual and near infrared cameras to enrich the variety of the data. SMIC contains sufficient amount of data to enable comprehensive testing of automatically analyzing micro-expressions. Baseline performance was also provided using LBP-TOP for feature description and SVM as classifier. Limitations were discussed and potential development directions were pointed out. SMIC is made available online[2] to promote research work in micro-expression analysis.

VI. ACKNOWLEDGMENTS

REFERENCES

[1] P. Ekman, "Lie catching and microexpressions," *The philosophy of deception*, pp. 118–133, 2009.

[2] E. Haggard and K. Isaacs, "Micromomentary facial expressions as indicators of ego mechanisms in psychotherapy," *Methods of research in psychotherapy. New York: Appleton-Century-Crofts*, pp. 154–165, 1966.

[3] P. Ekman and W. Friesen, "Nonverbal leakage and clues to deception," DTIC Document, Tech. Rep., 1969.

[4] P. Ekman and M. O'Sullivan, "Who can catch a liar?" *American psychologist*, vol. 46, no. 9, p. 913, 1991.

[5] M. Frank and P. Ekman, "The ability to detect deceit generalizes across different types of high-stake lies." *Journal of personality and social psychology*, vol. 72, no. 6, p. 1429, 1997.

[6] P. Ekman, "Darwin, deception, and facial expression," *Annals of the New York Academy of Sciences*, vol. 1000, no. 1, pp. 205–221, 2003.

[7] ——, "Microexpression training tool (METT)," *San Francisco: University of California*, 2002.

[8] ——, "METT micro expression training tool," *CD-ROM. Oakland*, 2003.

[9] R. Picard, *Affective Computing*. MIT Press, 1997.

[10] Z. Zeng, M. Pantic, G. Roisman, and T. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 1, pp. 39–58, 2009.

[11] P. Ekman and W. Friesen, "Facial action coding system: A technique for the measurement of facial movement," 1978.

[12] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*. IEEE, 2000, pp. 46–53.

[13] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*. IEEE, 2005, pp. 5–pp.

[14] M. Kamachi, M. Lyons, and J. Gyoba, "The japanese female facial expression (jaffe) database," *URL http://www.kasrl.org/jaffe.html*, vol. 21, 1998.

[15] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan, "Automatic recognition of facial actions in spontaneous expressions," *Journal of Multimedia*, vol. 1, no. 6, pp. 22–35, 2006.

[16] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, "Multi-pie," *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010.

[17] G. Zhao, X. Huang, M. Taini, S. Li, and M. Pietikäinen, "Facial expression recognition from near-infrared videos," *Image and Vision Computing*, 2011.

[18] M. Shreve, S. Godavarthy, V. Manohar, D. Goldof, and S. Sarkar, "Towards macro-and micro-expression spotting in video using strain patterns," in *Applications of Computer Vision (WACV), 2009 Workshop on*. IEEE, 2009, pp. 1–6.

[19] M. Shreve, S. Godavarthy, D. Goldof, and S. Sarkar, "Macro-and micro-expression spotting in long videos using spatio-temporal strain," in *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*. IEEE, 2011, pp. 51–56.

[20] X. Shen, Q. Wu, and X. Fu, "Effects of the duration of expressions on the recognition of microexpressions," *Journal of Zhejiang University-Science B*, vol. 13, no. 3, pp. 221–230, 2012.

[21] S. Polikovsky, Y. Kameda, and Y. Ohta, "Facial micro-expressions recognition using high speed camera and 3d-gradient descriptor," in *Crime Detection and Prevention (ICDP 2009), 3rd International Conference on*. IET, 2009, pp. 1–6.

[22] Q. Wu, X. Shen, and X. Fu, "The machine knows what you are hiding: an automatic micro-expression recognition system," *Affective Computing and Intelligent Interaction*, pp. 152–162, 2011.

[23] T. Pfister, X. Li, G. Zhao, and M. Pietikainen, "Recognising spontaneous facial micro-expressions," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1449–1456.

[24] M. Frank, M. Herbasz, K. Sinuk, A. Keller, and C. Nolan, "I see how you feel: Training laypeople and professionals to recognize fleeting emotions," in *The Annual Meeting of the International Communication Association. Sheraton New York, New York City*, 2009.

[25] J. Coan and J. Allen, *Handbook of emotion elicitation and assessment*. Oxford University Press, USA, 2007.

[26] S. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 4, pp. 627–639, 2007.

[27] Q. Wu, X. Shen, and X. Fu, "Micro-expression and its applications," *Advances in Psychological Science*, vol. 18, no. 9, pp. 1359–1368, 2010.

[28] J. Gross and R. Levenson, "Emotion elicitation using films," *Cognition & Emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[29] P. Ekman, W. Friesen, and M. O'Sullivan, "Smiles when lying." *Journal of personality and social psychology*, vol. 54, no. 3, p. 414, 1988.

[30] D. Matsumoto, J. LeRoux, C. Wilson-Cohn, J. Raroque, K. Kooken, P. Ekman, N. Yrizarry, S. Loewinger, H. Uchida, A. Yee *et al.*, "A new test to measure emotion recognition ability: Matsumoto and ekman's japanese and caucasian brief affect recognition test (jacbart)," *Journal of Nonverbal Behavior*, vol. 24, no. 3, pp. 179–209, 2000.

[31] D. Matsumoto and H. Hwang, "Evidence for training the ability to read microexpressions of emotion," *Motivation and Emotion*, pp. 1–11, 2011.

[32] T. Cootes, C. Taylor, D. Cooper, J. Graham *et al.*, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[33] Z. Niu, S. Shan, S. Yan, X. Chen, and W. Gao, "2d cascaded adaboost for eye localization," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 2. IEEE, 2006, pp. 1216–1219.

[34] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 6, pp. 915–928, 2007.

[35] Z. Zhou, G. Zhao, Y. Guo, and M. Pietikainen, "An image-based visual speech animation system," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 10, pp. 1420 –1432, oct. 2012.

[36] P. Ekman, E. Rolls, D. Perrett, and H. Ellis, "Facial expressions of emotion: An old controversy and new findings [and discussion]," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 335, no. 1273, pp. 63–69, 1992.

[2]http://www.cse.oulu.fi/CMV/Downloads