Robust One-Shot Facial Expression Recognition with Sunglasses

H. Jiang, K. Huang, T. Mu, R. Zhang, T. O. Ting, and C. Wang

Abstract—**Automatic facial expression recognition (FER) is both interesting and important in computer vision and machine intelligence. While previous FER systems often focus on learning a classifier in a controlled environment, a more practical and robust scenario is now under consideration. More specifically, traditional FER requires collecting as many as possible facial photos so that they will accurately recognize expressions no matter a particular person wears sunglasses, hats, and other accessories or not. Such requirement is however inconvenient and could impose practical difficulties for users. To alleviate this problem, a robust one-shot FER system that only requires taking one single facial photo for each expression of each user is proposed in this paper. When taking the single photo, the user is free to choose whether to wear sunglasses or not. The sunglasses can even be of different shape and of various luminous transmittance. Such one-shot recognition improves the user-friendliness of the FER system. Importantly, a novel and practical sunglasses detection and recovery approach is developed, which obtains an obvious accuracy improvement of 6.09%, 5.86% and 4.33% with state-of-the-art classifiers including Support Vector Machine (SVM), Linear Discriminate Analysis (LDA) and K-Nearest Neighbors (KNN) respectively on the modified Japanese Female Facial Expression (JAFFE) benchmark database.**

Index Terms—**Facial expression recognition, sunglasses detection and recovery, one-shot recognition system.**

I. INTRODUCTION

Among various ways of human interpersonal communication, facial expression is mostly special since it directly represents human's emotion, idea and thought. It is both interesting and important to explore automatic facial expression recognition (FER). According to the literature, there have been many proposals in this field. For example, Kung *et al.* [1] proposed an algorithm where each facial image is decomposed into an identity part and an expression part represented by their corresponding nonnegative bases. By devising graph-embedding constraints on the expression subspace, the intra-class variation issue in the expression recognition procedures can be mostly tackled and resolved. Another study is researched by Siddiqi *et al.* [2], which

employed the stepwise linear discriminate analysis for feature extraction, where the hidden conditional random field model is applied for the recognition task. And also, Wang *et al.* [3] introduced a geometric alignment technique to preprocess the original expression images, in which Local Binary Pattern features are extracted before classifying with classifiers including SVM.

Despite the reported good performance, previous FER approaches have often focused on learning a classifier in a controlled environment. For example, if the classifier is trained with images without sunglasses, hats, or any other personal accessories, the test image is then strictly restricted. Obviously, these limitations significantly affect user-friendliness of the system and impose practical difficulties upon real application.

To alleviate this problem, a robust one-shot FER system that only requires taking only one single facial photo for each expression of each user is proposed in this paper, focusing on making the system robust enough to tackle various sunglasses styles. Namely, when taking a single photo, the user is free to wear or not wear sunglasses. The sunglasses can even be of different shape and of various luminous transmittance. In contrast with previous studies, the training facial expression of one specific person may include sunglasses, while the testing one may not, and vice versa. Such one-shot recognition system obviously promotes the user-friendliness of the FER system, and will be expected to be used in various real scenarios. To achieve this target, a novel, intelligent sunglasses recovery approach is now proposed. Firstly, a roll de-rotation operation based on relative positions between two eyes was applied, then a Canny Edge Detector was implemented to locate possible sunglasses area on the Region of Interest (ROI). By utilizing histogram matching that transforms darker grayscale value (representing detected sunglasses region) into brighter one (representing the facial area without sunglasses), the corrupted pixels due to sunglasses are excellently recovered. In addition, the proposed algorithm was tested on the modified dataset from JAFFE [4], a benchmark facial expression dataset, with the proposed robust scheme. Improvement with the recognition rate of 6.09%, 5.86% and 4.33% for classifiers of SVM, LDA and KNN respectively with the proposed one-shot recognition system is achieved.

The significance of this research is as follows: 1). A robust one-shot facial expression recognition approach, hardly seen in previous FER systems, is proposed; 2). Investigation on robust recognition with sunglasses (of different shapes and various luminous transmittance) is majorly focused on, and the proposed automatic sunglasses recovery is novel; 3). A one-shot facial expression dataset with sunglasses, which was

modified from the JAFFE and served as a basis for relevant researchers, will be discussed in the future paper.

The rest of this paper is organized as follows. In the next section, the proposed novel robust sunglasses detection and recovery algorithm will be introduced in details. In Section III, the robust detection and recovery algorithm in a one-shot facial expression recognition system will be applied, and results of a series of experiments will be reported. In Section IV, final remarks conclusions will be given.

II. ROBUST SUNGLASSES DETECTION AND REGION RECOVERY

In this part, the proposed novel robust sunglasses detection and region recovery algorithm is introduced. When a new image is input, the algorithm will detect if sunglasses exist in the facial region. If presents, the recovery algorithm is utilized to remove the sunglasses and recover the facial region blocked by the sunglasses. Detail steps are described in the following.

A. Correction for Roll Rotation

As suggested in [3], a roll difference (of which the axis is perpendicular to the paper) will influence the recognition performance. Hence, a de-rotation operation, mainly based on the difference of centers of the two eyes, is a necessary procedure. Fig. 1 illustrates the comparison for this operation.

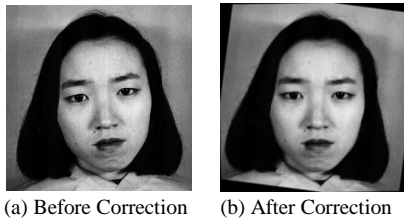(a) Before Correction   (b) After Correction
Fig. 1. Roll rotation correction.

To de-rotate the image, the two landmarks representing left and right eye centers generated from STASM are applied for rotation reference, as described in (1), where x_r and x_l represent column indices for the left and right eye center points respectively, with y_r and y_l the row ones. The image is then rotated by the angle of $-\theta$ with respect to the image center.

$$-y_l)/(x_r - x_l) \rfloor \qquad (1)$$

B. Sunglasses Region Detection

To perform a recovery for the sunglasses region, it is essential to exactly localize the sunglasses region firstly. In the proposed approach, the Canny Edge Detector is utilized [5].

TABLE I: SUNGLASSES REGION DETECTION

ORIGINAL IMAGE	ORIGINAL ROI	THRESHOLD ROI	COARSE EDGE	FINAL CONTOUR

As seen in Table I, the ROI is generated based on STASM [8], [9] (an Active Shape Model based facial components landmarks detector, which will be demonstrated in Section III) landmark positions of the corresponding eye and eyebrow. Since the Canny Edge Detector detects sharp edges easily, it leads to undesirable performance on the eye part (as well as the eyebrow part) due to their similar color compared with sunglasses. It is shown in the Original ROI column of TABLE. To tackle this defect, pixel values from both eye and eyebrow parts are transformed to the ones from their neighboring pixels. Before this, a simple average filter can be implemented to further weaken the surrounding edges. By applying a grayscale value thresholding operation to the target region, further enhancement will be achieved.

C. Grayscale Value Histogram Shifting

Wearing sunglasses can be regarded as an overall dropping operation on grayscale values for all pixels locating in sunglasses regions, since the grayscale value dropping rate (GVDR), is set to less than one (maximum grayscale value). Additionally, the distribution of the original values is reduced gradually. From Fig. 2, it is found that the histogram is shifted towards the lower grayscale section as GVDR increased while the histogram distribution becomes centralized.
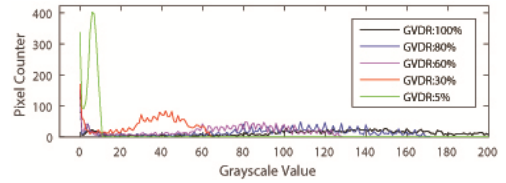

Fig. 1. Histogram comparison among sunglasses with different GVDR.

D. Histogram Matching for Sunglasses Recovery

Observed from Fig. 2, it can be obtained that mathematically, by transforming histograms with multiple GVDR values to the one at 100% GVDR, the image region corrupted by sunglasses will be recovered. The proposed transformation, majorly based on histogram matching [6], is designed to match the histogram distribution from the source image to the one from the target image. In this research, the matching target is the corresponding image with GVDR at 100% or the facial image without wearing sunglasses, since if GDVR is set to 100%, no grayscale change has ever been applied apart from the sunglasses frame. The target histogram model is calculated from the original face images, and then the matching can be divided into histogram equalization and matching.

Histogram Equalization: The histogram equalization is firstly applied to both the source image (image to be matched) and the target image (matching target). In (2) and (3), s_k and v_k represent two equalized images of source image and target one respectively, n_j/n and m_i/m represent statistical probabilities for gray value r_k and z_k, $T(r_j)$ and $G(z_i)$ represent the transformations that will achieve the equalized images. It is illustrated in Fig. 3, where red and blue lines represent the source histogram and target one respectively.

$$s_k = T(r_k) = \sum_{j=0}^{k} p_r(r_j) = \sum_{j=0}^{k}(n_j/n) \qquad (2)$$

$$v_k = G(z_k) = \sum_{i=0}^{k} p_z(z_i) = \sum_{i=0}^{k} (m_i / m) \qquad (3)$$

Matching from Equalized Histograms: After the equalization, histograms from both regions are stretched with the value range of [0, 255]. As the result, they now exist in one identical space. It can be considered as $s_k = v_k$, which means:

$$\sum_{j=0}^{k} (n_j / n) = \sum_{j=0}^{k} (m_i / m) \qquad (4)$$

$$T(r_k) = G(z_k) \qquad (5)$$

The histogram matching scheme is given by the following:

$$z_k = G^{-1}(z_k) = G^{-1}[T(r_k)] \qquad (6)$$

(a) Original histogram
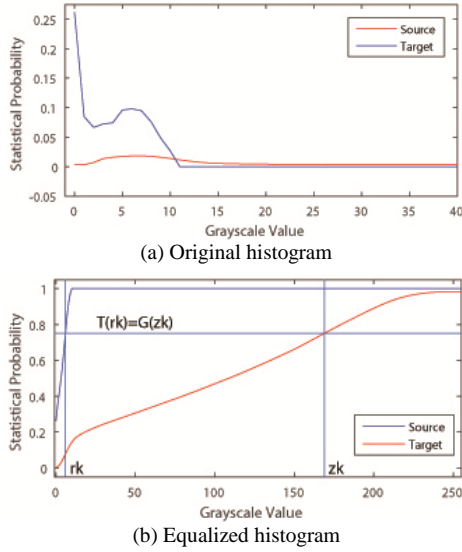

(b) Equalized histogram

Fig. 3. Histogram equalization.

In Table II, by locating the smallest difference between statistical probability values for both the original and target image in equalized histogram, and performing an inverse transformation ($G^{-1}(z_k)$) based on the target image, a mapping function can be simply found from the source grayscale value of r_k to the target one z_k [6], and Table II gives several examples of this matching performance.

TABLE II: HISTOGRAM MATCHING EXAMPLES

GVDR SHAPE & ID	5% CIRCULAR KA.AN1.39	30% RECTANGULAR KL.FE1.174	60% CIRCULAR NA.SA1.205	80% RECTANGULAR UY.HA2.138
SUNGLASSES IMAGES				
RECOVERED IMAGES				
ORIGINAL CORRESPONDING IMAGES				

III. EXPERIMENTS

In the following section, experiments to verify the proposed algorithm are described. Since there is no available facial expression dataset with sunglasses, firstly a modified new dataset from the original JAFFE data was generated by automatically inserting different shapes of sunglasses with various luminance transmittance. The whole generation process is introduced in the following subsection.

A. Data Preparation

The original JAFFE dataset contains 213 images of seven facial expressions from 10 females. For each one, there are three or four pictures involved. All portraits in the original database are wearing no sunglasses, so various styles of sunglasses around each eye section were firstly added manually.

By utilizing an Active Shape Model [7] based software package STASM [8], [9], in total 77 landmarks representing different facial component positions such as eyes, eyebrows, mouth, nose were detected. Eye regions were extracted with related landmarks automatically, generating positions to add sunglasses. Two shapes of sunglasses, including rectangular and circular ones, were added. By changing the grayscale values in the sunglasses region at a rate, namely, Grayscale Value Dropping Rate (GVDR) no greater than 100%, various luminance transmittance was created. Thus, for each image in the original dataset, there were in total 192 corresponding images with circular or rectangular sunglasses added manually with GVDR from 5% to 100%.

TABLE III: MANUALLY-ADDED SUNGLASSES EXAMPLES

ID	KA.AN.139		KL.AN1.167		YM.AN1.61	
ORIGINAL IMAGE						
GVDR	CIR	REC	CIR	REC	CIR	REC
30%						
60%						
100%						

In Table III, it is observed that higher GVDR leads to lower luminous transmittance for sunglasses (darker sunglasses), and hence, the more original information is erased. It is argued that if the GVDR is lower than 5%, there will be insufficient useful information remained. Hence such special cases will not be considered. This modified dataset is promised to be released.

B. Experimental Setting

In this section, the performance is evaluated based on the recognition improvement due to implementing the proposed recovery scheme. All images from both training and testing set were firstly recovered according to the scheme demonstrated in Section II, and then the recognition rate was recorded. However, experiments with non-recovered images for both training and testing sets were also be conducted for

comparison and evaluation. The recovery performance was quantified according to the recognition rate improvement.

For recognition procedures, Gabor Wavelet Features with five scales and eight orientations were firstly extracted [10], after which Kernel PCA [11] was performed to compress the feature data from 10240 dimensions into around 60-d with no less than 95% information preserved. Classifiers including SVM (LibSVM [12]), LDA and KNN were compared. It is also straightforward to evaluate other classifiers such as Bayesian Multinet [13] and different variants of SVM [14]. The final recognition rate was given as the average value over 50 times repeated independent experiments. The best parameters (the kernel width and balance parameter in SVM and neighbor number for KNN) were optimized with grid search.

Two major types of experiments were conducted. One was referred as Individual Training and Testing, where images from both training and testing sets were all from the identical GVDR group. The other one was called the Full Training and Testing, where images from the training set came from the full database with various styles of sunglasses, as did the same with those in the testing set. The Full Training and Testing type were the best simulation on real application cases.

When dividing the database into training and testing sets, two methods including the conventional five-fold scheme and the proposed one-shot one were utilized and compared.

Five-fold Scheme for Individual Training and Testing: When conducting the Individual Training and Testing with this scheme, 4/5 of images in the corresponding individual GVDR set were randomly chosen as the training set, while the remaining ones were put into the testing set.
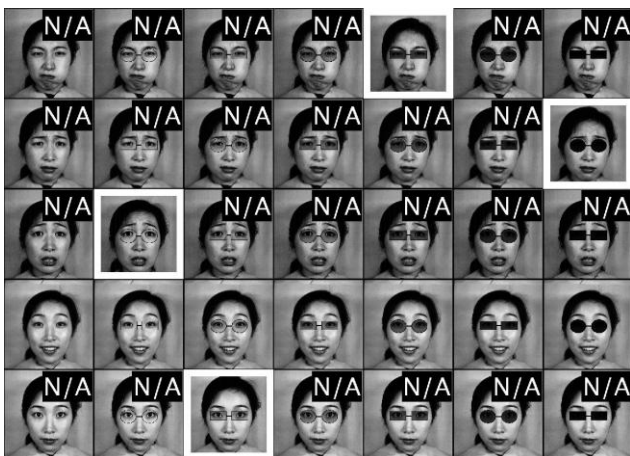

Fig. 4. Five-fold Separating Scheme Example for Full Training and Testing Portrait with the white frame is put into the training set; Portraits with "N/A" label are neither in the training nor testing set; Portraits with no special signs are all put into the testing sets.

Five-fold Scheme for Full Training and Testing: When performing the Full Training and Testing, the following conditions should be met: 1). A random binary switch is set to specify whether the specific image in the original JAFFE database or one of its corresponding images with various styles of sunglasses is to be put into the training set (4/5 of the original JAFFE database); 2). Another random binary switch is set to specify if the sunglasses are worn, namely, half of randomly selected portraits in the training set are wearing sunglasses; 3). A random integer value is set for ones wearing

sunglasses from 5 to 100 to specify the GVDR of source sunglasses for better simulation of real application scenario; 4). Also a random binary switch is set to specify whether the portrait is wearing the circular sunglasses or rectangular ones. For the last three conditions, they are only valid when the image is put into the training set. Otherwise, those pictures in the remaining 1/5 part of the original database are all put into the testing set with all corresponding images with or without various styles of sunglasses. So not only will the images in the training set never appear in the testing set, neither will their corresponding images. Fig. 4 gives several examples for better explanation, where images with a white frame were put into the training set, those with no signs were put into the testing set, and ones with "N/A" labels were neither in the training nor testing sets.

One-shot Scheme for Individual Training and Testing: As previously demonstrated in Subsection A, three or four images were involved for each expression of each portrait. In this Individual Training and Testing, only one randomly selected image from three or four were put into the training set, while the remaining two or three were placed in the testing set.

One-shot Scheme for Full Training and Testing: For the Full Training and Testing with the One-shot Scheme, similar conditions were also met as described in the five-fold section, and the only difference is that the separating scheme here was one-shot scheme, rather than five-fold one. Fig. 5 gives several examples, in which two different expressions of two different portraits (Portrait: KA with Expression: Fear and Portrait: YM with Expression: Happy) were involved. For each expression of each portrait, only one image was selected for the training set (image with the white frame). For all the corresponding images of the selected training images, all other ones were neither in the training nor the testing sets. All images apart from those mentioned above were put into the testing set. Similarly, for images from the training set, their corresponding ones will never be part of the testing set. Additionally, the five-fold training set is larger than the one of one-shot.

Fig. 5. One-shot separating scheme examples for full training and testing.

C. Experimental Results

Individual Training and Testing with Five-fold Scheme: Table IV gives the performance by separating the proposed database with a five-fold cross-validation scheme. A significant improvement for recognition rate is obtained for lower GVDR groups when both training set and testing set are images with recovery scheme implemented, compared with those without. However, the improvement reduces gradually as GVDR increases, and when it reaches 90%, the improvement is weakened to negative values. Then there is in fact decrease rather than improvement brought due to the recovery scheme. It indicates that for higher GVDR level images, surely there is less information corrupted when sunglasses are added manually, and improvement on recognition rate due to the proposed recovery method is relatively decreased.

TABLE IV: INDIVIDUAL TRAINING & TESTING WITH FIVE-FOLD SCHEME

GROUP GVDR	SVM		LDA		KNN	
	NON[a]	REC[b]	NON	REC	NON	REC
10%[c]	75.833 ±6.991	82.071 ±5.525	61.000 ±8.744	70.048 ±7.735	65.524 ±5.308	73.000 ±5.834
20%	77.524 ±6.610	82.881 ±5.323	61.571 ±8.713	70.333 ±8.728	66.881 ±5.806	74.429 ±5.877
30%	79.262 ±5.513	84.286 ±5.319	61.833 ±7.880	70.095 ±7.747	69.310 ±5.790	74.833 ±6.340
40%	80.262 ±5.803	83.738 ±5.702	62.429 ±7.167	69.500 ±8.498	70.310 ±6.479	75.976 ±6.294
50%	80.952 ±6.607	84.143 ±5.586	63.524 ±8.441	68.143 ±8.540	72.833 ±6.361	76.690 ±5.967
60%	81.310 ±7.309	83.952 ±6.323	63.929 ±9.004	67.738 ±9.852	73.286 ±6.260	75.690 ±6.979
70%	81.857 ±7.009	83.619 ±5.912	65.238 ±8.918	66.881 ±9.415	74.738 ±7.012	75.238 ±7.079
80%	82.381 ±7.465	83.286 ±6.654	66.167 ±8.638	67.048 ±8.392	74.286 ±7.131	74.571 ±7.299
90%	83.167 ±7.143	82.976 ±6.646	66.476 ±8.259	66.214 ±7.996	74.905 ±7.054	74.357 ±6.999
100%	83.333 ±6.263	83.310 ±6.142	66.690 ±8.708	66.167 ±8.310	74.619 ±7.263	74.095 ±7.356
ORG[d]	86.000 ±6.911	85.714 ±6.477	67.762 ±8.414	67.667 ±9.099	84.286 ±7.063	84.143 ±7.920

a: No recovery applied for both training and testing sets;
b: Recovery applied for both training and testing sets;
c: Individual Training and Testing with images at 10% GVDR;
d: Individual Training and Testing with on original database;

TABLE V: INDIVIDUAL TRAINING & TESTING WITH ONE-SHOT SCHEME

GROUP GVDR	SVM		LDA		KNN	
	NON	REC	NON	REC	NON	REC
10%	69.804 ±3.396	77.112 ±3.683	67.874 ±6.878	73.056 ±6.229	61.888 ±3.263	70.385 ±2.337
20%	71.112 ±3.246	78.643 ±3.468	68.147 ±6.474	72.154 ±7.551	63.580 ±3.438	72.399 ±2.260
30%	72.762 ±3.122	79.224 ±3.249	68.580 ±6.233	73.056 ±8.167	65.203 ±3.246	72.986 ±2.531
40%	74.238 ±3.561	79.986 ±3.053	69.748 ±6.398	73.280 ±7.296	67.126 ±2.949	73.706 ±2.637
50%	75.804 ±3.619	80.720 ±3.053	70.692 ±6.783	74.455 ±6.923	68.951 ±2.724	74.399 ±2.147
60%	77.056 ±3.340	80.538 ±3.549	71.524 ±7.365	74.294 ±6.770	70.524 ±2.276	74.371 ±2.630
70%	78.427 ±3.392	80.469 ±3.227	72.406 ±7.171	74.035 ±7.685	72.399 ±2.453	74.7558 ±2.425
80%	79.448 ±3.049	80.594 ±3.520	73.490 ±5.928	73.622 ±7.038	73.755 ±2.417	74.552 ±2.759
90%	80.007 ±3.125	80.357 ±3.281	74.371 ±60.88	73.315 ±6.428	74.434 ±2.490	74.371 ±2.255
100%	80.483 ±3.437	79.937 ±2.845	74.294 ±6.774	73.028 ±6.746	74.175 ±2.278	74.427 ±2.347
ORG	88.168 ±2.747	87.245 ±2.902	75.441 ±8.677	75.594 ±8.148	85.552 ±3.126	84.657 ±3.181

Individual Training and Testing with One-shot Scheme: Table V represents recognition rate with one-shot separating scheme for the Individual Training and Testing type. Still, consistent improvement for recognition rate can be acquired for lower GVDR groups when both training set and testing set are images with recovery scheme implemented. The similar decrease can also be found as GVDR increases, and when the GVDR is greater than 90%, the recognition performance is then weakened with the recovery scheme applied.

The performance for both separating schemes is given in Figs. 6, 7 and 8 for SVM, LDA and KNN respectively with the individual training and testing type.

Full Training and Testing with Both Schemes: Based on results given in Table IV and Fig. 9, a consistent promotion for SVM, LDA and KNN with both the five-fold scheme and the one-shot one is obtained.

When comparing with results from Individual Training and Testing experiments, although there exists decrease for higher GVDR group in individual types, when considering overall performance with Full Training and Testing, the benefits from the proposed recovery scheme shall never be underestimated. It can be concluded that introducing recovered images into both the training and testing set is effective in promoting the classification performance of sunglasses expression images with the One-shot Scheme. Improvements are 6.092%, 5.858% and 4.552% for SVM, LDA and KNN respectively, which is best simulates applications in real scenarios. Additionally, improvements are 6.512%, 6.979% and 4.821% respectively for SVM, LDA and KNN in the Five-fold Scheme.
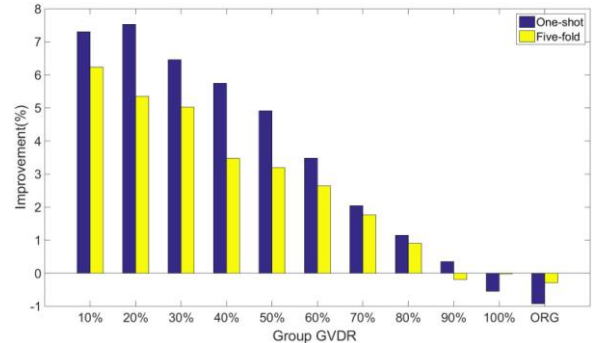

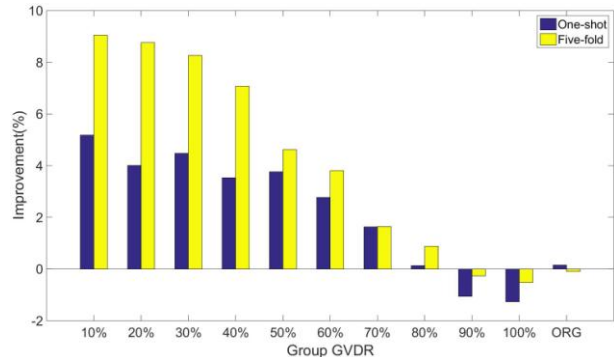Fig. 6. Improvement based on final recognition rate with SVM classifier.

Fig. 7. Improvement based on final recognition rate with LDA classifier.

TABLE VI: FULL TRAINING & TESTING WITH BOTH SCHEMES

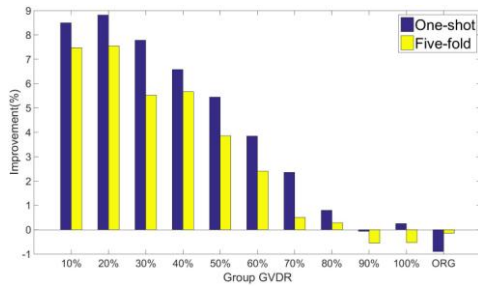| TYPE	SVM		LDA		KNN	
	NON	REC	NON	REC	NON	REC
FIVE-FOLD	69.090 ±2.986	75.182 ±3.145	64.302 ±6.270	70.160 ±7.258	63.855 ±2.238	68.407 ±2.638
ONE-SHOT	69.090 ±2.986	75.182 ±3.145	64.302 ±6.270	70.160 ±7.258	63.855 ±2.238	68.407 ±2.638

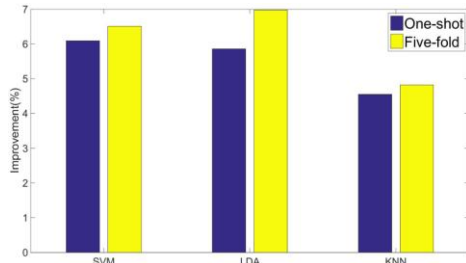Fig. 8. Improvement based on final recognition rate with KNN classifier.

Fig. 9. Overall performance improvement based on final recognition rate consistent improvement obtained for all of the 3 off-the-shelf classifiers.

IV. CONCLUSION

This paper has proposed a sunglasses recovery scheme based on Canny Edge Detection and histogram matching for automatic facial expression recognition. By randomly selecting one-shot images from a full database containing images with or without manually-added sunglasses with various luminous transmittance, the final recognition performance is greatly improved.

The advantage is that it is now possible to utilize only one-shot image for each expression of each portrait in the training phase. In comparison with experimental results without the recovery implemented, improvement of 6.092%, 5.858% and 4.33% on the final recognition rate has been achieved with classifiers of SVM, LDA and KNN respectively.

ACKNOWLEDGEMENT

REFERENCES

[1] H. Kung, Y. Tu, and C. Hsu, "Dual subspace nonnegative graph embedding for identity-independent expression recognition," *IEEE Trans. on Information Forensics and Security*, vol. 10, no. 3, pp. 626-639, March 2015.
[2] M. H. Siddiqi, R. Ali, A. M. Khan, Y. Park, and S. Lee, "Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields," *IEEE Trans. on Image Processing*, vol. 24, no. 4, pp. 1386-1398, April 2015.
[3] X. Wang, X. Liu, L. Lu, and Z. Shen, "A new facial expression recognition method based on geometric alignment and LBP features," in *Proc. IEEE 17th International Conference on Computational Science and Engineering (CSE)*, December 2014, pp. 1734-1737.
[4] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," in *Proc. IEEE 3rd Conference on Automatic Face and Gesture Recognition Proceeding*, April 1998, pp. 200-205.
[5] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679-698, November 1986.
[6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. India: Pearson Education, 2002, ch. 3.3, pp. 88-107.
[7] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models: their training and application," *IEEE Trans. on Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38-59, 1995.
[8] S. Milborrow and F. Nicolls, "Locating facial features with an extended active shape model," *Computer Vision-ECCV 2008*, pp. 504-513, Springer, 2008.
[9] S. Milborrow and F. Nicolls, "Active shape models with sift descriptors and mars," *IVISAPP*, vol. 1, no. 2, p. 5, 2014.
[10] W. Liu, C. Song, Y. Wang, and L. Jia, "Facial expression recognition based on gabor features and sparse representation," in *Proc. IEEE 12th Conference on Control Automation Robotics Vision (ICARCV)*, December 2012, pp. 1402-1406.
[11] B. Schlkopf, A. Smola, and K. Muller, "Kernel principal component analysis," in *Proc. IEEE Conference on Artificial Neural Networks*, Springer Berlin Heidelberg, 1997, pp. 583-588.
[12] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. on Intelligent Systems and Technology*, vol. 2, pp. 1-27, 2011.
[13] K. Huang, I. King, and M. R. Lyu, "Discriminative training of Bayesian Chow-Liu multinet classifiers," in *Proc. 2003 International Joint Conference on Neural Networks*, 2003, pp. 484-488.
[14] K. Huang, D. Zheng, J. Sun, Y. Hotta, K. Fujimoto, and S. Naoi, "Sparse learning for support vector classification," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 1944-1951.

H. Jiang received the B.Eng degree in mechanical engineering from the School of Mechatronic Engineering of the China University of Mining and Technology, Xuzhou, P.R. China, in 2012. In 2013, he received the M.Sc. degree from the Department of Mechanical and Automation, the Chinese University of Hong Kong. Currently, he is working towards the Ph.D. degree in the Department of Electric and Electronic Engineering of Xi'an Jiaotong-Liverpool University, Suzhou, P.R. China. His major research interest includes pattern recognition, and machine learning.

K. Huang received the B.Eng in automation from the Department of Automation, Xi'an Jiaotong University, Xi'an, China in 1997, M.S. in pattern recognition and intelligent systems, from Institute of Automation, Chinese Academy of Sciences in 2000, and the Ph.D. degree from the Department of Computer Science and Engineering, the Chinese University of Hong Kong in 2004. He is now an affiliated full professor in University of Electrical Science and Technology of China, and an associate professor of the Department of Electric and Electronic Engineering of Xi'an Jiaotong-Liverpool University, SIP, Suzhou, 215123, P.R. China. He has published 6 books in Springer and more than 100 international papers including 35 premier international journal articles. His current research interests include machine learning, pattern recognition, information retrieval and image processing. He is now a member of IEEE and ACM.

T. Mu received the B.Eng in electrical engineering from the Special Class for the Gifted Young, University of Science and Technology of China, Hefei, P.R. China in 2004, and the Ph.D. degree from the Department of Electrical Engineering and Electronics, the University of Liverpool, Liverpool, UK in 2008. She joined the University of Liverpool as a lecturer in 2012. She has published 28 articles in premier international journals. Her current research interests include machine learning, data visualization and mathematical modeling, with applications to information retrieval, text mining and bioinformatics. She is now a member of IEEE.

R. Zhang received the First-class (Hons) degree in telecommunication engineering from Jilin University of China in 2001 and the Ph.D. degree in computer science and mathematics from University of Ulster, UK in 2007. She joined Xi'an Jiaotong-Liverpool University as a Lecturer in 2012. Her current research interests include machine learning, data mining and statistical analysis.

T. O. Ting obtained First-class (Hons) degree in electronic and telecommunication engineering from UNIMAS, Sarawak, Malaysia, MEng from Multimedia University, Malacca, Malaysia and Ph.D in electrical engineering from The Hong Kong Polytechnic University. He is currently a lecturer with the Dept. of Electrical and Electronic Engineering, Xi'an Jiaotong-Liverpool University. His main research is on the application of optimization techniques for a wide range of engineering problems.
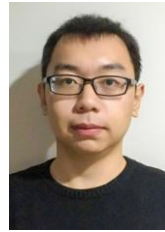
C. Wang received the B.Eng from the Department of the Electrical & Electronic Engineering of Xi'an Jiaotong-Liverpool University, Suzhou, P.R. China in 2015. He is now a master student in the Department of the Electrical & Electronic Engineering of the Xi'an Jiaotong-Liverpool University, Suzhou, P.R. China. His current research interests are pattern recognition and image processing.