Robust Face Recognition Under Partial Occlusion Based on Local Generic Features

Amit Kumar Yadav, GLA University, Mathura, India
Neeraj Gupta, GLA University, Mathura, India
Aamir Khan, Independent Researcher, India
Anand Singh Jalal, GLA University, Mathura, India

https://orcid.org/0000-0002-7469-6608

ABSTRACT

Face recognition has drawn significant attention due to its potential use in biometric authentication, surveillance, security, robotics, and so on. It is a challenging task in the field of computer vision. Although the various state-of-the-art methods of face recognition in constrained environments have achieved satisfactory results, there are still many issues which are untouched in unconstrained environments, such as partial occlusions, large pose variations, etc. In this paper, the authors have proposed an approach which utilized the local generic feature (LGF) to recognize the face in the partial occlusion by fusing features scale invariant feature transform (SIFT) and multi-block local binary pattern (MB-LBP). It also utilizes robust kernel method for classification of the query image. They have validated the effectiveness of the proposed approach on the benchmark AR face database. The experimental outcomes illustrate that the proposed approach outperformed the state-of-art methods for robust face recognition.

KEYWORDS

Local Generic Feature, MB-LBP, Partial Occlusion, Robust Kernel Method, SIFT

INTRODUCTION

The face recognition task in the computer vision system can be described as follows: Given an input face image and a database of face images of known individuals, how can we verify or determine the identity of the person in the input image? A robust face recognition system is used in a biometric application for verifying and identifying an individual of interest. Researchers from the areas of image analysis, computer vision, machine learning, pattern recognition, and many others are working cooperatively, inspired not only by the fundamental challenges but also by various real-life applications in which person of interest identification is required. The interest of researchers is also increased by the fact that with the rising public concern for security, the need for identity verification such as face recognition is more apparent. Also, advance technology, such as in mobile devices and digital cameras made face recognition more important and easier to approach. The various state-of-art methods of face recognition in constrained environments have achieved satisfactory results. There are still many issues which are untouched in unconstrained environments, such as partial occlusions, large pose variations, etc.
In this paper, we have proposed an approach which utilized the Local Generic Feature (LGF) to recognize the face in the partial occlusion by fusing features Scale Invariant Feature Transform (SIFT) & Multi-Block Local Binary Pattern (MB-LBP). It also utilizes a Robust Kernel method for classification of the query image.

The following is the organization of this paper. In Section 2, we have briefly described the related work. In Section 3, the proposed model for robust face recognition is described. Section 4 shows the result and discussion of experimental results on standard datasets. Finally, section 5 elaborates the conclusion and future scope.

Related Work

There are various state-of-art approaches proposed in the literature for automatic and robust face identification. Yuan et al. (2016) proposed face identification under occlusion using an occlusion dictionary and support vector discrimination dictionary. The support vector machine scheme is used to train the non-occlusion dictionary, guaranteeing that the resulting dictionary can not only express testing samples from the same class but also can reduce the interference of samples of different classes. Huang et al. (2017) proposed kernel extended dictionary (KED) for face recognition, which shows efficient and effective direction for fusing KDA and Sparse Representation based Classification (SRC). Duan et al. (2017) proposed a partial face recognition approach which is approximated a transformation encoding based on non-rigid graphical structure and computing the correspondence between edges & node.

Liao et al. (2013) proposed a face recognition system under partial occlusion approach without proper alignment. They design the face representation method based on multi-key point descriptor without alignment, where the feature vector size of a face is determined by the original information of the facial image. Ying Wen (2017) proposed a robust face recognition system based on discriminative common vector dictionary. The discriminative common vector technique has utilized the feature of variations of fisher’s linear discriminate analysis for smallest query sample. Cheheb et al. (2017) proposed an approach to solve the partial occlusion present in face recognition system using a single sample per person (SSPP). First, image is divided into multiple patches and local binary patterns are applied as a texture descriptor on each block separately. Then, a kernel principle component analysis is applied to dimensionality reduction of the resulting descriptors. Wu et al. (2016) proposed an approach known as SRC based occlusion pattern (OPSRC). They find contiguous occlusion region from the testing sample to construct an occlusion pattern. An occlusion dictionary is constructed by the addition of all facial image’s occlusion pattern. The original dictionary and occlusion dictionary are solved together. If the occlusion is detected in the query sample, and then applied the similar occlusion to the entire facial feature in the dictionary.

Liu et al. (2016) introduced artificial occlusions into the training set. Then a class-specific dictionary is created by the incremental training set, and a dense hybrid and sparse representation framework that represent sparse corruption and noise and without a class-specific dictionary. Tan et al. (2017) introduced KLS-GSRC by fusion of kernel trick, group sparsity, and data locality. Then nonlinear & structure feature is combined in the training and test set that is a more discriminative representation and better exploited. McLaughlin et al. (2017) proposed a new approach for robust face recognition that is reliable to illumination variation and partial occlusion. The method was based on a single training sample per person. They introduce a new approach for comparison of the facial image similarity that is based on largest matching area (LMA) at each facial image location. LMA optimizes the accuracy of lighting conditions and at the same time, it improves the discriminative power between different individual.

Alrjebi et al. (2017) proposed a technique to overcome the issue of face recognition under partial occlusions. The given method is based on color image representation utilizing more than three color components in each patch rather than on whole face image. Xiaolin Chen et al (2016) gives a method
known as error detection and kernel principal component analysis (KPCA) approach for robust face identification under the partially occluded face.

Although face recognition in constrained environments has achieved satisfactory results, there are still many issues and problems in unconstrained environments, such as partial occlusions, large pose variations, etc. Duan et al. (2017) proposed an approach for partial-face-recognition by estimation of the non-rigid transformation vector that encodes the graphical structure and measures the make correspondence between edges and nodes. In this, they use SIFTSURF LBP to detect and extract Local feature then apply TPGM for classification. The TPGM perform better than non-spares representation approaches. However, proper alignment of the facial image with respect to the predefined facial landmark is the issue of deprecating the recognition accuracy of the TPGM approach. Inspired by this issue, in the proposed work we have detected and extracted local generic feature and construct feature descriptor which makes robustness to partial facial occlusion. In the proposed method, we utilize the discriminative power of SIFT and MB-LBP for feature extraction that deals with occlusion in face recognition problem. Local binary pattern (LBP) is effective for local feature extraction. However, we have used MB-LBP which is more robust than LBP and encodes not only microstructures but also macrostructures of image patterns, and hence provides a more complete image representation than the basic LBP operator by author Liao et. al. (2007). The proposed method successfully resolved the issue of partial occlusion of human faces.

Proposed Methodology
In the proposed approach, we have detected & extracted local generic feature and construct feature descriptor. These can be utilized for partial occlusion to robustly recognize the faces. The proposed approach has three main steps as shown in figure 1. The first step is preprocessing and proper alignment of the input facial image. The second step is the extraction of local feature and construction of the feature vector by fusion of Scale Invariant Feature Transform (SIFT) and Multi Block- Local Binary Pattern (MB-LBP). Finally, in third step robust kernel method is used for classification.

Preprocessing
In this phase of the robust face recognition system, it is needed to be pre-processing of a facial image. An improvement of the image data that suppress the irrelevant distortions is the purpose of the pre-processing and enhances some image features which are useful for further processing. Using image pre-processing that provide advantages and solve problems that ultimately lead to better global and local feature detection. Pre-processing is the stage where images are extracted and converted to grey color if obtained images are in RGB color scale and resized to 83X60. A grey image is matrix information whose value shows grey shades. The factors of a matrix of the gray image have integral value and range between 0 and 255. We have also used geometrically independent approaches such as logarithmic compression and gamma adjustment.

Alignment of facial images is very important for the further feature extraction and facial recognition phase. Since, SIFT features are invariant to image rotation, translation, scaling, light changes. Therefore, in this work, SIFT technique is used for image alignment for change in perspective.

Extraction of Local Generic Feature
The extraction of local feature is used for correspondence matching. In the feature space the extracted local feature is separated to produce effective and efficient discrimination between facial images. In this work, the SIFT and MB-LBP are used for local generic feature extraction. Finally, a robust kernel method is used for classification.

Scale-Invariant Feature Transform (SIFT)
In proposed approach local key point is detected using the SIFT by author Lowe et. al. (2004) (Scale-Invariant Feature Transform) algorithm, which is efficient for real-time object detection and
recognition. The SIFT algorithm extracts robust size, rotation, and occlusion invariant features. A feature vector from this is characterized by the rotations, occlusion and the stability of picture sizes.

SIFT consists of two main parts: the interest point detector and feature descriptor. To detect interest points in a given input image, \( I(x, y) \) the SIFT algorithm uses scale-space Difference-Of-Gaussian (DOG). As for an input image, \( I(x, y) \), the scale space is defined as a function, \( L(x, y, \sigma) \) produced from the convolution of a variable scale Gaussian \( G(x, y, \sigma) \) with the input image and the difference-of-Gaussian function \( D(x, y, \sigma) \) can be computed from the difference of two nearby scales separated by a multiplicative factor \( k \) as given in Eq. (1,2):

\[
D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \tag{1}
\]

\[
D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}
\]

Then local maxima and minima \( D(x, y, \sigma) \) are computed based on comparing each sample point to its eight neighbors in the current image and nine neighbors in the scale above and below. At this scale, the gradient magnitude, \( m(x, y) \), and orientation, \( \theta(x, y) \) is computed using pixel differences in Eq.(3,4). Thereafter, orientation is determined by building a histogram of gradient orientations weighted by the gradient magnitudes from the key points neighborhood and it is assigned to each interest point combined with the scale above and provides a scale and rotation invariant coordinate system for the descriptor.

\[
m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{3}
\]
\[
\theta(x, y) = \tan^{-1}\left(\frac{(L(x, y + 1) - L(x, y - 1))}{(L(x + 1, y) - L(x - 1, y))}\right)
\]

(4)

After detecting the interest points in each image, the SIFT descriptor computes the gradient magnitude and orientation at each image sample point in a region around the key-point location weighted by a Gaussian window. The coordinates of the descriptor and the gradient orientations are rotated relative to the key-point orientation to achieve orientation invariance.

**Multi-block Local Binary Patterns (MB-LBP) Operator**

MB-LBP is an extension of LBP that can be computed on multiple scales in constant time using the integral image. MB-LBP which is more robust than LBP and encodes not only microstructures but also macrostructures of image patterns, and hence provides a more complete image representation than the basic LBP operator. The calculation is done based on average values of block sub-region in MB-LBP instead of individual pixels by authors Liao et. al. (2007). Using MB-LBP we have to divide the image into multiple blocks. Each sub-blocks containing neighboring pixels is a square block. The whole image is composed of 9 blocks. The local generic feature was inherited by Multi-Block Local Binary Patterns (MB-LBP). Received by inheriting a histogram feature from every cell of the decomposed input image for the representation of every facial image for concatenated and geometrical enhancement. After preprocessing facial image is partitioned into sub-blocks then histograms feature is constructed and finally concatenated into a single feature vector. This feature vector is used to measure similarities by computing the distance between images and plays a big role in the efficient representation of the facial image.

In order to detect the micro-pattern and macro-pattern, MB-LBP is used. MB-LBP labeling is a pixel of the image by comparing the mean gray-values of neighborhood sub-blocks with the mean of gray value of center sub block. The scale of the MB-LBP technique is represented as \(S \times S\), where \(S\) represent the size of each sub blocks. Average is computed of scalar values over each sub blocks using the integral image.

\[
MB - LBP(b_c) = \sum_{n=0}^{7} S(b_n - b_c).2^n
\]

(5)

Where \(b_c\) is the average sum retrieved at middle block and \(b_n\) shows the average sum obtained at neighborhood blocks, and the function \(S(x)\) is denoted as:

\[
S(x) = \begin{cases} 
0, & x < 0 \\
1, & x \geq 0 
\end{cases}
\]

(6)

Figure 2 given examples of MB-LBP histogram feature of the filtered facial image. Where the facial image is partitioned into multiple sub-blocks and textural feature are extracted from each sub-blocks and finally concatenated into one unique histogram feature vector. The large scale operator gives complementary information to small scale details and reduces the noises and mean values over the sub-blocks makes the representation more robustness.

**Classification Using Robust Kernel Method (RKM)**

In this section, we discuss Classification using Robust Kernel Method by the author Yang et. al. (2013) which is robustness to LGF-RKM based facial object representation and classification. In this approach, the image is decomposed into \(P\) patches or sub-regions. After extracting in each region, the LGF feature of each sub region, then combine the LGFs of total sub-regions into the Resultant
feature descriptor in the $i^{th}$ sub-block. After that combination of total feature descriptor extracted from each region, i.e., $t = [t_1, t_2, ..., t_p]$ would be represented as the image feature descriptor.

Similarly, extract sub feature vectors for each of the training samples, and denoted by $T_i$ the feature vector representation of all the sub-feature descriptors of the $i^{th}$ region from all training set. Now we have to kernel representation of query sample $t_i$ over the training sample $T_i$ and represented as

$$\min_{\alpha_i} \| \phi(t_i) - \phi(T_i)\alpha_i \|_p$$ \text{s.t.} $\|\alpha_i\|_p \leq \sigma$ \hspace{1cm} (7)

Where $\alpha_i$ is the coding coefficient vector mapped by the kernel function $\phi$ in the higher dimensional feature space. Assuming that there are different regions $t_i$ extracted from the same test sample has the same representation over their related vector $T_i$, then kernel representation of the query sample by concatenating each sub region features that are written as

$$\min_{\alpha} \left\| \phi(t_1)\phi(t_2)\ldots\phi(t_p) \right\|_p - \left[ \phi(T_1)\phi(T_2)\ldots\phi(T_p) \right]_p \text{s.t.} \|\alpha\|_p \leq \sigma$$ \hspace{1cm} (8)

Where $\alpha$ is the query sample coding coefficient vector? The above equation is seeking the regularized representation for a mapped feature within mapping with high dimensional space.

In the kernel based representation (3.4), Euclidean distance is used to compute the residual representation. This type of kernel representation is more effective when there are no outsiders in the query samples. To build the kernel representation reliable & robust to disguises or block occlusion, we have to use to robust fidelity term in the kernel representation. $e = [e_1, e_2, ..., e_p]$ Denoted as a residual vector, where $e_i$ is the $i^{th}$ blocks kernel representation residual i.e, $e_i = \sqrt{\|\phi(t_i) - \phi(T_i)\alpha_i \|^2}$

Then Robust Kernel Represented as

$$\min_{\alpha} \rho(e) \text{ s.t.} \|\alpha\|_p \leq \sigma$$ \hspace{1cm} (9)
Where \( \rho(e) = \sum_{i=1}^{P} \rho(e_i) \) and the cost function \( \rho(\cdot) \) is assumed to be non-sensitive to the outsiders in the testing sample.

Then query sample is classified by

\[
\text{identity}(t) = \min_{j} \left\{ \sum_{i=1}^{P} (w_i, \theta_i, j) \right\}
\]  

Where \( \theta_i, j = \| \phi(t_i) - \phi(T_i, j) \|_2^2 \) is the \( i \)-th region kernel based representation residual related with the \( j \)-th class. \( A_j \) is the sub matrix of \( A_i \) related with the \( j \)-th class and \( w_i \) is the some weight function i.e. range of normalized weight value is \([0,1]\).

**Experimental Results and Analysis**

In this section, we have discussed experimental results and their analysis on benchmark AR Martinez et. al. (1988) facial dataset. AR dataset is used to verify the accuracy of the proposed method. We compared the proposed approach with various state-of-the-art approaches on facial recognition under partial occlusion, which elaborate the robustness and effectiveness of the proposed method. The proposed approach is implemented in MATLAB (R2017a) at the Intel® core™ i3-6006U with 8 GB RAM and CPU @ 2.00 GHz.

**Data Set**

Frontal facial images of 126 individuals over the 4000 images contained in the AR dataset. In two separate sessions, images are taken under different variations, including facial occlusion/disguise, expression, and illumination. In each session, there are 4 samples with different expressions, 3 samples with different illumination conditions, and 3 samples wearing scarf and 3 sample wearing sunglasses respectively.

**Figure 3. Images from AR dataset with various variations**
Result Analysis

We have tested the proposed approach by performing several experiments. To measure the accuracy of the proposed approach, the results are compared with existing works. We have used overall accuracy to performance measure. Image size is normalized to 83x60.

Quantitative Analysis

Accuracy is used as performance measure for comparison with the other approaches. The overall accuracy is calculated to measure the correct classification rate. The parameters are calculated as follows.

\[
\text{Accuracy(\%)} = \frac{TP + TN}{TP + TN + FP + FN}
\] (11)

\[
F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}
\] (12)

**Performance of Face Recognition under Partial Occlusion:** A comparison with existing methods shows that the results obtained using the proposed approach efficiently recognized facial image. Table 1 contains a comparative analysis of the existing method with that of the proposed method and shows the recognition accuracy of using the Local Generic Feature and Robust Kernel method on AR standard dataset. To describe the results, an overall accuracy shown in below table. For a comparative study with the related approach, all these query sample and training samples were properly aligned and normalized to 83x60. By looking at the results it is clear that there is better accuracy in the proposed approach over the state-of-the-art.

**Graphical Representation of Performance of Facial Occlusion:** The graphical representation of the results is shown figure 4. It shows the statistical representation of recognition accuracy of proposed method as well as various state-of-the-arts. It shows that proposed approach performs better than state-of-the-arts. MKD-SRC-GTP Liao et. al. (2013), RPSM Wang et. al. (2016), TPGM Duan et. al. (2017) have misalignment and intra-class variation issue. Proposed LGF-RKM method achieves the best performances for all sessions it illustrates that LGF-RKM not only does well in occlusion sample but also tackle non occluded images.

| Method       | Session1-SunGlass | Session1-Scarf | Session2-SunGlass | Session2-Scarf |
|--------------|-------------------|----------------|-------------------|---------------|
| CPD          | 71.00             | 75.67          | 49.33             | 61.00         |
| MLERPM       | 75.00             | 78.33          | 53.33             | 66.67         |
| LAIRPM       | 87.33             | 88.33          | 56.33             | 81.33         |
| MKD-SRC-GTP  | 82.33             | 83.33          | 57.67             | 76.33         |
| RPSM         | 88.67             | 90.33          | 63.67             | 85.67         |
| TPGM         | 89.33             | 91.00          | 65.00             | 86.67         |
| Proposed     | 93.33             | 94.01          | 69.66             | 89.33         |
CONCLUSION

In this paper, we have tended to the issue of robust and reliable face recognition under the partially occluded face. Proposed approach have used local generic feature descriptors SIFT and MB-LBP to detect and extract feature descriptor. A Robust Kernel Method is used for classification. The proposed approach is validated on the standard dataset and the results are compared with state-of-the-art approaches. The proposed approach achieves better performance for partial occlusion on the AR dataset. The quantitative results illustrate the effectiveness of the proposed technique and outperform over the state-of-the-art.
REFERENCES

Alrjebi, M. M., Pathirage, N., Liu, W., & Li, L. (2017). Face recognition against occlusions via colour fusion using 2D-MCF model and SRC. Pattern Recognition Letters, 95, 14–21. doi:10.1016/j.patrec.2017.05.013

Cheheb, I., Al-Maadeed, N., Al-Madeed, S., Bouridane, A., & Jiang, R. (2017). Random sampling for patch-based face recognition. Biometrics and Forensics (IWBF), 5th International Workshop on, 1-5. doi:10.1109/IWBF.2017.7953104

Chen, X., Wang, S., & Ruan, X. (2016). Recognition of partially occluded face by error detection with logarithmic operator and KPCA. Image and Signal Processing, BioMedical engineering and Informatics (CISP-BMEI), International Congress on, 460-464. doi:10.1109/CISP-BMEI.2016.7852755

Duan, Y., Lu, J., Feng, J., & Zhou, J. (2017). Topology preserving graph matching for partial face recognition. Multimedia and Expo (ICME), IEEE International Conference on, 1494-1499. doi:10.1109/ICME.2017.8019548

Huang, K.-K., Dai, D.-Q., Ren, C.-X., & Lai, Z.-R. (2017). Learning kernel extended dictionary for face recognition. IEEE Transactions on Neural Networks and Learning Systems, 28(5), 1082–1094. doi:10.1109/TNNLS.2016.2522431 PMID:26890929

Li, Z. H., Hou, Y., & Liu, H. B. (2014). Alignment of face images based on SIFT feature. In Machine Learning and Cybernetics (ICMLC) International Conference on (Vol. 2, pp. 597-600). IEEE. doi:10.1109/ICMLC.2014.7009675

Liao, S., Jain, A. K., & Li, S. Z. (2013). Partial face recognition: Alignment-free approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(5), 1193–1205. doi:10.1109/TPAMI.2012.191 PMID:23520259

Liao, S., Zhu, X., Lei, Z., Zhang, L., & Li, S. Z. (2007). Learning Multi-scale Block Local Binary Patterns for Face Recognition. In S. W. Lee & S. Z. Li (Eds.), Lecture Notes in Computer Science: Vol. 4642. Advances in Biometrics. ICB. doi:10.1007/978-3-540-74549-5_87

Liao, S., Zhu, X., Lei, Z., Zhang, L., & Li, S. Z. (2007). Learning multi-scale block local binary patterns for face recognition. International Conference on Biometrics, 828-837. doi:10.1007/978-3-540-74549-5_87

Liu, H., Duan, H., Cui, H., & Yin, Y. (2016). Face recognition using training data with artificial occlusions. In Visual Communications and Image Processing (pp. 1–4). VCIP.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2), 91–110. doi:10.1023/B:VISI.0000029664.99615.94

Martinez, A. M., & Benavente, R. (1998). The AR face database (CVC Technical Report No. 24). Barcelona, Spain: Universitat Autonoma de Barcelona. Computer Vision Center.

McLaughlin, N., Ming, J., & Crookes, D. (2017). Largest matching areas for illumination and occlusion robust face recognition. IEEE Transactions on Cybernetics, 47(3), 796–808. doi:10.1109/TCYB.2016.2529300 PMID:26955057

Myronenko, A., & Song, X. (2010). Point set registration: Coherent point drift. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(12), 2262–2275. doi:10.1109/TPAMI.2010.46 PMID:20975122

Tan, S., Sun, X., Chan, W., Qu, L., & Shao, L. (2017). Robust Face Recognition with Kernelized Locality-Sensitive Group Sparsity Representation. IEEE Transactions on Image Processing, 26(10), 4661–4668. doi:10.1109/TIP.2017.2716180 PMID:28641251

Wen, Y. (2017). A novel dictionary based SRC for face recognition. Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference on, 2582-2586. doi:10.1109/ICASSP.2017.7952623

Weng, R., Lu, J., Hu, J., Yang, G., & Tan, Y.-P. (2013). Robust feature set matching for partial face recognition. Computer Vision (ICCV), IEEE International Conference on, 601-608. doi:10.1109/ICCV.2013.80

Weng, R., Lu, J., & Tan, Y.-P. (2016). Robust point set matching for partial face recognition. IEEE Transactions on Image Processing, 25(3), 1163–1176. doi:10.1109/TIP.2016.2515987 PMID:26761775

Wu, C.-Y., & Ding, J.-J. (2016). Occlusion pattern-based dictionary for robust face recognition. Multimedia and Expo (ICME), 2016 IEEE International Conference on, 1-6. doi:10.1109/ICME.2016.7552982
Yang, M., Zhang, L., Shiu, S. C.-K., & Zhang, D. (2013). Robust kernel representation with statistical local features for face recognition. *IEEE Transactions on Neural Networks and Learning Systems, 24*(6), 900–912. doi:10.1109/TNNLS.2013.2245340 PMID:24808472

Yuan, L., & Li, F. (2016). Face recognition with occlusion via support vector discrimination dictionary and occlusion dictionary based sparse representation classification. *Chinese Association of Automation (YAC), Youth Academic Annual Conference of, 110-115. doi:10.1109/YAC.2016.7804874*