Sanjukta Bhattacharya, Ananjan Maiti, Samhita Das
and Shristee Ganguly

2 Partial face recognition using image fusion

Abstract: Several biometric partial face recognition researches have been performed by many scientists. In this chapter, a novel technique has been recommended, which acknowledges students face to speed up the attendance procedures in a classroom. Students' partial pictures have been used to prepare the image set, and preprocessed different partial faces to gray-level images. Initially, the technique like discrete wavelet transform has been used to obtain local features. Afterward, the effectiveness of the approach has been improved by employing image fusion with the averaging method. The fusion technique along with the correlation technique was executed to the contrast between the fused images, and the test images were selected from the entire image set. Results revealed practically 90% of the instances that were matched. The acceptance rate on an overall analysis has been found to prevail between 86.67% and 87.5%.

Keywords: Image fusion , 2D correlation , binning , partial face , Internet of things

2.1 Introduction

The conventional way of taking the presence of students is strenuous and also lengthy. The lecture normally prolongs it, upkeeping the student's attendance. This technique is ineffective, particularly if it is a lecture with a large number of students. It also triggers a great deal of disorder and hindrance when an examination is hosted. Furthermore, the attendance sheet is subjected to damage and also loss while being handed down between different students and teachers. In some cases, the lecturers have the tendency to call a couple of students' full name randomly, which is an unfair student assessment procedure either. Eventually, these attendance records are used by the lecturers to keep track of the students' attendance rates. This process could be comfortable and useful with a small number of students; however, dealing with the records of a large number of students often leads to human error.

Therefore, to prevent this mistake, computerized attendance recording system is offered to every student thus saving time, initiative and reduces disruptions. These particular systems might consist of the utilization of biometrics like finger-mark, eye acknowledgment, retinal search and articulate acceptance. These kinds of systems are conclusively established in the current years; nevertheless, this is actually unpleasant and expense required when it comes to implementation on

huge scale gets increased significantly. Biometric iris acknowledgment uses template acceptance methods based on high-resolution and a distortion-free image of the irises from the human eyes. Iris is an organ, whose structure remains steady all throughout life. Therefore, this functions as an incredibly great biometric when it comes to developing identification of an individual. Speech identification is an one more individual recognition system through computer software application as well as an equipment tool along with the capability to decipher the individual's speech. In the event of this, we observe that an individual's speech could be rapidly captured and also utilized when it comes to an unexpected for any type of electronic device. This possesses low accuracy. A health problem like a cold could alter an individual's speech, making complete recognition challenging or difficult. To conquer these kinds of problems, the biometric function such as face identification could be utilized, which includes the stages like image acquisition, feature extraction, face classification [1] and also eventually marking the attendance. Likewise, some of the significant advantages that face recognition modern technology companies provide are the time participation monitoring, which enables leaving out the amount of time theft among the employees and also the paid hours start from this particular minute until the same check-out process. It is essential when it comes to entrepreneurs to rely on their employees, however, to watch on all of them. Face recognition describes a part of computer science which can determine individuals' facial expressions inside digital images. Face recognition technology may start through looking for human eyes. It could then utilize a genetic algorithm to spot face areas consisting of eyebrows, the mouth, nose, nostrils and also the iris. This particular technology is being used more often in digital photography since a way to assist cameras autofocus on individuals' faces. This system complements a person's face immediately against a data source of photos to develop the identity. Face detecting systems could utilize algorithms to forecast age, sex and also various other elements to serve up appropriate advertisements, and the acknowledgment explains a biometric innovation which heads way beyond identifying when a human face exists. The suggested work primarily handles the students face recognition within the class. Often, this similar procedure works using a computer application that catches a digital picture of an individual's face via video frame and also compares it to images in a database of stored records. Nevertheless, when it comes to the front or full face acknowledgment algorithm, there is a drawback that students in a class need to arrive near the electronic camera and provide their attention to obtain correct outcomes. It is because the students in a class do not regularly sit appropriately. Often it fails when the students are not in their frontal face and remain in a side face, that is, either left sided or right sided, the camera fails to determine the image of those students. The relative angle of the target faces affects the recognition score exceptionally. So we need to enroll faces in the recognition software; typically, several positions are used. The more direct the image and the higher its resolution, the

higher the score of any resulting matches. Finally, the subject must be holding a neutral expression and managing these specifications helps reduce variation between any two images.

Inside this particular succeeding study, in Section 2.3 we have talked about current research in biometric face recognition as well as its restrictions. The partial face recognition framework consisting of many collective actions are discussed in Section 2.4. Here, we have highlighted sequential phases of work in brief. The outcome of identification of the faces of students was compared in Section 2.5 and confirms the value of image fusion technique concerning student face recognition. Lastly, in Section 2.6, we have wrapped up study with limitation and potential scope.

2.2 Related work

Liao et al. [2] suggested a method that possesses absolutely no demand of face positioning and also eyeball coordinates. The analysis of the face images starts with identifying the descriptor via Multi-Keypoint Descriptors (MKD). By doing this, any probe face image, alternative or partial, could be sparsely represented through a comprehensive dictionary of gallery descriptors. Here, Gabor ternary pattern (GTP) produces a distinct keypoint descriptor for robust and biased face recognition. These experimental outcomes are stated on four public domain face data sources under each of the open-set recognition and verification scenarios. They dealt with the issue of identifying a face from its partial image and suggested an alignment-free approach called Multi-Key point Descriptors-Sparse Representation-based Classification (MKD-SRC). Their approach stands for every face image with a set of keypoint descriptors and constructs a large dictionary from all the gallery descriptors. By doing this descriptor of a partial probe image could be sparsely stood for by the dictionary, and the identity of the probe could be presumed appropriately.

Smeets et al. [3] offered the meshSIFT algorithm and also its utilization of 3D face acknowledgment. This approach consists of the neighborhood procedure, and every salient factor is explained in a feature vector, including concatenated histograms of shape indices and also slant angles. Afterwards, the feature vectors of two 3D facial surface areas are dependably fit through contrasting the angles in a feature space. As a result, the algorithm is robust to expression variations, missing information as well as outliers. It also showed that the amount of matching meshSIFT features is a reliable measure for expression-invariant face acknowledgment. It reveals the recognition rate of 93.7% and 89.6% for standard data sources, specifically.

Ding et al. [4] operated on recorded face images inside difficult scenarios that typically include essential position variant, which significantly deteriorates the efficiency

of algorithms developed to identify frontal faces. Here, the approach is of a novel face recognition framework that is efficient in dealing with the full variety of position variants inside 90 degrees. Finally, face matching was carried out in patch level rather than the holistic level. Extensive and also methodical testing on multi-PIE, CMU-PIE <http://www.cs.cmu.edu/afs/cs/project/PIE/MultiPie/Multi-Pie/Home.html> (last access date: 07.06.2019) and also The Facial Recognition Technology (FERET) database revealed that the proposed technique regularly surpasses single task-based standards in addition to cutting-edge techniques for the posed problem. Patch Based Partial Representation (PBPR) could be used to face images in the arbitrary position, which is a considerable benefit over presenting methods.

Lei et al. [5] suggested 3D face recognition with the accessibility of just partial information and also single training sampling is an extremely challenging task. This study offered an effective 3D face recognition technique to deal with this challenge. It examines face along with a collection of regional keypoint-based multiple triangle statistics (KMTS) that is robust to partial face information, large facial expressions as well as position variants. The suggested method stands for a 3D confront with a collection of regional geometric descriptors called KMTS. The suggested descriptor allows robust 3D PFR with inconsistent data, barricades and data corruptions. Finally, a durable prior classification database on the suggested local descriptor was the need for resolving the single sample issue.

Best-Rowden et al. [6] constructed an analytical system regarding a pair of mug-shot data sources, which are the most comprehensive facial aging data sources examined to this day concerning several topics. The longitudinal evaluation revealed that in spite of reducing initial scores, 99% of topics could still be identified in 0.01% fixed false accept rates. In future, they would attempt to monitor an accuracy along with a global threshold.

Lahasan et al. [7] suggested optimized symmetric partial facegraphs that were a memetic-based framework to determine faces that are prone to unfavorable circumstances. Those situations could be a facial feeling, occlusions and lighting changes. This study combined an enhanced harmony search algorithm and a sensible single particle optimizer to have the benefit of their global as well as regional search aptitudes. These features further functioned as the foundation so as to without effort design the partial facegraph.

Fu et al. [8] mentioned the effectiveness of the sparse representation classifier that might perform classification through assessing which class results in the minimal representation inaccuracy. For that reason, they had created virtual samples through making use of original training samples along with the aim of enhancing the variety of training samples. The respective representation scores utilized weighted score-level fusion of the virtual samples, and also the original training samples are merged with each other to acquire the finished classification outcomes with an accuracy of 76.61.

2.3 Methodology

Image fusion is a process that is used to gather maximum useful information from various images and converts the inclusion of all images into a single image (sometimes more than one image). This individual fused image is more accurate and contains all the useful information from different images [9]. The image fusion method is used for reducing the data amount as well as this method is also used for constructing much more appropriate images that are easily understandable by human beings as well as machines. Moreover, it can be said that image fusion process is capable enough to integrate multiple sources of image [10]. The resultant fused image contains fully spectral and spatial resolution features. Image fusion technique must fulfill two very basis requirements: (1) fused resultant image always have the capability to reverse each and every required information that is obtained from input source image; and (2) image fusion technique does not allow any incorrect or inaccurate diagnosis. There are three levels, where image fusion can occur. The levels are (1) pixel level, (2) feature level and (3) decision level. In case of pixel-level method the pixels of the source images are processed and preserve all the required image information. The pixel-level fusion is defined as low-level fusion. This is a very simple and common method that is done at the stage of image preprocessing [11]. This pixel-level image fusion can be divided into two groups: image point fusion and single-level fusion. In case of single-level fusion, a set of high category signals from different sensors that are consistent with the actual image are obtained to amalgamate, where in case of image point fusion each and every point that is obtained from different images are amalgamated directly. Feature [12]-level fusion is also known as medium-level image fusion that occurred at the feature extraction phase. In this fusion level, extracting image features and at the same time synthesizing the similar features from various images is the major task. The different types of features that are commonly extracted are angle, texture, similar lighting area, shape, profile, similar depth of focus area and others.

Decision level is defined as the highest level of fusion. All the useful information that are extracted from either feature-level image fusion or pixel-level image fusion are used to get best decision to successfully achieve the desired objective. Decision-level image fusion results are responsible to take any type of control and decision for the next stage.

Image fusion technique is categorized into three major divisions such as spatial domain fusion, transform domain fusion and statistical domain fusion. The spatial domain fusion technique handles with the pixel values of images directly. This technique is used to manipulate the image pixel value to successfully achieve the desired objective. There are different and most commonly used spatial domain fusion methods available such as Intensity-hue-saturation (IHS)-based methods, averaging, principle component analysis, Bovey transform and high-pass filtering method. In case of frequency domain technique, first conversion of an image into frequency domain

occurs with the help of Fourier transform. The necessary operations occur on Fourier transform of the particular image and after that to achieve the desired result inverse Fourier transform can be applied. There are also available commonly used transform domain fusion methods such as pyramid transform and wavelet transform [13]. Statistical domain fusion uses statistics to get the desired output. The statistical variables, for example, local correlation average, local correlation average with respect to variation and least squares are most frequently used. Drawbacks of image fusion such as dataset or operator dependency and color distortion can be solved through statistical domain fusion method. Some of the most common methods that fall under the category of statistical domain fusion are local mean and variance matching, local correlation modeling, local mean matching, regression variable substitution and others. Averaging fusion method is used to achieve the resultant final image where all the regions are focused. The average value is found out by applying the summation of all pixel values of every input images, which is divided by the total number of input images. This average value is used to give the corresponding resultant pixel images. The speed of this averaging fusion method is very fast. Averaging fusion method is defined as a simplest technique that is used to calculate the weighted average value from all the images that are taken as input and then the results are reflected on final fused image. The application of image fusion can be seen in various fields, where the image analysis is needed such as satellite image analysis, robotics, medical image analysis, remote sensing application, microscopic imaging and computer vision.

2.4 Proposed method

There are two types of environment in which face detection can take place: controlled and uncontrolled environments. In a controlled environment, there is a camera mounted on the wall and the subjects are required to come in front of the camera to get their face recognized. The subjects need to stand in front of the camera in a specific angle, such as their full frontal face gets captured by the camera. There are other requirements that need to be confirmed to, for example, the proximity of the subject to the camera and proper lighting. In the uncontrolled environment, on the other hand, there are no face-to-face interaction between the camera and the subjects. The subjects are not required to stand in front of the camera in any specific angle as such. In the proposed method, the uncontrolled environment has been used. At first, the input video has been taken where the individual movement has been captured. Then the captured video is divided into frames. Binning technique is deployed on the whole set of frames. In our method, the bin size is set as 10. The middle frame is selected as a key frame from each and every bin. The selection is done in an adaptive way. Image fusion operation is taken place on all the key frames. At the next phase, the correlation is computed with the help of the

Figure 2.1: Proposed methodology.

fused key frame and the facial image gets stored in the database. The correlation values of each comparison are stored in a cell or array. Once all the comparisons are over, the image that has the highest correlation value is taken to be the matched image (Figure 2.1).

2.5 Experiments and results

In this proposed system, MATLAB R2013a software, 64-bit operating system, x64-based processor, Intel Core i3-4005U CPU @ 1.70 GHz has been used.

Figure 2.2: Database.

In the proposed method two database sets are used. The first database set that is shown in Figure 2.2 consists of 30 folders. Each folder consists of five images of an individual, so a total of 150 images are stored in the database. The images that are stored in the database are basically the frontal faces [14] as well as the partial faces of the individuals. The images are captured in an uniform lighting and same background environment. In the proposed method, the uncontrolled environment has been used. At first, the input video has been taken where the individual movement has been captured [15]. After that the captured video is divided into frames. Binning technique is deployed on the whole set of frames. In our method, the bin size is set as 10. The middle frame is selected as a key frame from each and every bin. The selection is done in an adaptive way. Image fusion operation is taken place on all the key frames. In this case, averaging method is used as an image fusion. At the next phase, the correlation is computed with the help of the fused key frame and the facial image is stored in the database. The correlation values of each comparison are stored in a cell or array. Once all the comparisons are done, the image that has the highest correlation value is taken to be the matched image. After performing the proposed method in the database 2, the acceptance rate is 86.67%.

Some of the accepted cases and some of the false cases are shown in Figures 2.3 and 2.4. Image fusion operation takes place on all key frames. In this case, averaging method is used as an image fusion.

At the next phase, the correlation is computed with the help of the fused key frame, and the facial image is stored in the database (Figure 2.5). The correlation

Figure 2.3: Accepted cases.

Figure 2.4: False cases.

Figure 2.5: Fused image.

values of each comparison are stored in a cell or array. After all the comparisons take place, the image that has the highest correlation value is taken to be the matched image. As the image area is quite large and the face position along with the facial expression varies for each image of an individual, the fused image gives a bit of a distorted image. This can be a reason for the low acceptance rate. That is, if the position of the face always remains constant, which is not the ideal case, the false acceptance rate can be reduced further. In the initial stage, the fusion of eye region of each image in the folder and then the correlation with the fused eye image has been done. The experimental result of eye region shows the lower acceptance rate than the current experiment with full face fusion. This is due to the fact that the fusion region, that is, the eye region, of all the images is very small and no noticeable changes are found in Figure 2.6; hence, the correlation value was high for all the fused images.

Figure 2.6: Eye sections.

We have created our own dataset, whose size is less in the pool to validate the robustness of the proposed system. In this database, the total number of images is 24. We have taken Official consent has been obtained from all participants in this study for fuse of their captured images. In this database set, both the partial faces along with the frontal faces are captured but the expressions of the individuals do not vary, and remain the same throughout. This database set makes use of uniform lighting and the images are captured against the same background (Figure 2.7).

Some of the accepted cases and some of the false cases are shown in Figures 2.8 and 2.9. After performing the proposed method in our own database, the acceptance rate is 87.5.

2.6 Face detection with Internet of things

At present, maximum technological advancements have taken place with the help of a concept named as Internet of things (IoT), where the various real-life physical devices are connected to the Internet [16, 17]. This IoT technology consists of very powerful, precise and riskless infrastructure along with features like minimum power consumption and cost effectiveness, where anything or anyone can be communicated or connected with each other anywhere or anytime to achieve the desired output. IoT can be implemented in different segments in our daily life, such as home automation, transportation, agriculture and healthcare segments. One of expandable areas using IoT infrastructure [18] is face recognition, which can be

Figure 2.7: Prepared dataset.

Figure 2.8: Accepted cases.

Figure 2.9: False cases.

applied in different fields such as biometric system [19], attendance system and security reasons. Nowadays, attendance system is a mandatory prerequisite for any organization but maintaining the attendance manually is a very time-consuming and difficult task. To resolve this issue, several automated methods have been implemented which is not only easy to handle, but also the accuracy level of these systems can be better [20]. Face recognition system is one of the popular human identification processes, where the ambiguities such as time consuming, fake attendance, high cost and human errors are resolved [21]. This system can be used in any organization like school, college, bank and corporate office to reduce the probability of any complication such as fake attendance by feature [22, 23] identification. In this system, the attendance of each and every person is automatically recorded whenever they enter into their workspace through the continuous detection of each and every person's face. After that the detected faces are simultaneously compared with the previously taken database. If the detected image is matched with the stored image, then the attendance is given to that particular person, otherwise not. The partial face recognition system is a part of face recognition which is also a very challenging and interesting process. In this case, the picture of a person's partial face is taken with the help of a camera and the database consists of the full face as well as partial face of that person. The intermediate and all the other steps are similar with any face recognition process. This partial face recognition can also be implemented with the help of IoT technology, where the camera is connected to a raspberry pi and then the comparison of images is done with the stored images in the database. The accuracy level of this type of attendance system is generally better when compared with the manual attendance system.

2.7 Conclusions

Partial face recognition algorithm is a modern and advanced technique. The process includes various advantages as well as limitations. It is also an efficient method of biometric identification technique. The proposed method is applied in a large data set, which gives the acceptance rate of 86.67%. We have also created our own dataset, whose size is less in the pool and the proposed method is also applied in that dataset which gives the acceptance rate of 87.5%. In this proposed method, averaging technique of image fusion and after that correlation method is used but other fusion techniques also exist through which the value of the correlation method may vary affecting the acceptance rate; hence, the accuracy may be varied.

References

- [1] Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., and Dey, N. "A survey of image classification methods and techniques," in *Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, 2014, International Conference, on, 2014, pp. 554–557. IEEE.
- [2] Liao, S., Jain, A. K., and Li, S. Z. "Partial face recognition: alignment- free approach," *IEEE Transactions on pattern analysis and machine intelligence*, 35(5), 2013, pp. 1193–1205.
- [3] Smeets, D., Keustermans, J., Vandermeulen, D., and Suetens, P. "Mesh-SIFT: local surface features for 3d face recognition under expression variations and partial data," *Computer Vision and Image Understanding*, 117(2), 2013, pp. 158–169.
- [4] Ding, C., Xu, C., and Tao, D. "Multi-task pose-invariant face recognition," *IEEE Transactions on Image Processing*, 24(3), 2015, pp. 980–993.
- [5] Dey, N. (2019). Uneven illumination correction of digital images: A survey of the state-of-the-art. *Optik*, 183, 483–495.
- [6] Best-Rowden, L. and Jain, A. K. "Longitudinal study of automatic face recognition," *IEEE transactions on pattern analysis and machine intelligence*, 40(1), 2018, pp. 148–162.
- [7] Lahasan, B., Lutfi, S. L., Venkat, I., Al-Betar, M. A., and San-Segundo, R. "Optimized symmetric partial facegraphs for face recognition in adverse conditions," *Information Sciences*, 429, 2018, pp. 194–214.
- [8] Fu, L., Chen, D., Lin, K., & Li, A. (2018). An improved SRC method based on virtual samples for face recognition. *Journal of Modern Optics*, 65(13), 1565-1576.
- [9] Dey, N., Pal, M., and Das, A. "A session based blind watermarking technique within the NROI of retinal fundus images for authentication using DWT, spread spectrum and Harris corner detection," *arXiv preprint arXiv:1209.0053*, 2012.
- [10] Satapathy, S. C., Raja, N. S. M., Rajinikanth, V., Ashour, A. S., and Dey, N. "Multi-level image thresholding using Otsu and chaotic bat algorithm," *Neural Computing and Applications*, June 2018, Volume 29, Issue 12, pp. 1285–1307
- [11] Roy, P., Goswami, S., Chakraborty, S., Azar, A. T., and Dey, N. "Image segmentation using rough set theory: a review," *International Journal of Rough Sets and Data Analysis (IJRSDA)*, 1 (2), 2014, pp. 62–74.
- [12] Dey, N., Ashour, A., & Patra, P. K. (2017). *Feature Detectors and Motion Detection in Video Processing* (pp. 1-328). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-1025-3
- [13] Bhattacharya, T., Dey, N., and Chaudhuri, S. "A session based multiple image hiding technique using DWT and DCT," *arXiv preprint arXiv:1208.0950*, 2012.
- [14] Kamal, S., Dey, N., Ashour, A., Ripon, S., Balas, V., and Kaysar, M., "FbMapping: an automated system for monitoring facebook data," *Neural Network World*, 27(1), 2017, pp. 27.
- [15] Le, D. N., Nguyen, G. N., Van Chung, L., and Dey, N. MMAS algorithm for features selection using 1D-DWT for video-based face recognition in the online video contextual advertisement user-oriented system. *Journal of Global Information Management (JGIM)*, 25(4), 2017, pp. 103–124.
- [16] Sarowar, M. G., Kamal, M. S., and Dey, N. Internet of things and its impacts in computing intelligence: a comprehensive review–IoT application for big data. In *Big Data Analytics for Smart and Connected Cities* (pp. 103–136). IGI Global, 2019.
- [17] Kumar, P. M., Gandhi, U., Varatharajan, R., Manogaran, G., Jidhesh, R., and Vadivel, T. Intelligent face recognition and navigation system using neural learning for smart security in Internet of Things. *Cluster Computing*, 2017, pp. 1–12. <https://link.springer.com/article/10.1007/s10586-017-1323-4> (last access date: 07.06.2019)
- [18] Dharavath, K., Talukdar, F. A., Laskar, R. H., and Dey, N. Face recognition under dry and wet face conditions, In *Intelligent techniques in signal processing for multimedia security*, Dey,

- N.,V. Santhi (Eds.), *Studies in Computational Intelligence Series*, Springer, 2017, pp. 253–271.
- [19] Chaki, J., Dey, N., Shi, F., & Sherratt, R. S. (2019). Pattern mining approaches used in sensor-based biometric recognition: A review. *IEEE Sensors Journal*, 19(10), 3569–3580.
 - [20] Surekha, B., Nazare, K. J., Raju, S. V., and Dey, N. Attendance recording system using partial face recognition algorithm, In *Intelligent techniques in signal processing for multimedia security*, Dey, N.,V. Santhi (Eds.), *Studies in Computational Intelligence Series*, Springer, 2017, pp. 293–319.
 - [22] Dey, N., Bhatt, C., and Ashour, A. S. *Big data for remote sensing: Visualization, analysis and interpretation*, Cham: Springer, 2018.
 - [23] Chaki, J., and Dey, N. Pattern analysis of genetics and genomics: a survey of the state-of-art, *Multimedia Tools and Applications*, 2019, 1–32.