

## Review Paper on Face Clustering

Rakhi<sup>1\*</sup>, Suman<sup>2</sup>

<sup>1,2</sup>Dept. of Computer Science, Deenbandhu Chhotu Ram University of Science and Technology, Haryana, India

\*Corresponding Author: [sharma.rakhi1641@gmail.com](mailto:sharma.rakhi1641@gmail.com), Tel.: +91-9467371646

DOI: <https://doi.org/10.26438/ijcse/v8i8.1316> | Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

Received: 04/Aug/2020, Accepted: 22/Aug/2020, Published: 31/Aug/2020

**Abstract-** In recent years video surveillance system has been broadly established and nowadays we can see the billions of cameras are embedded in smart phones and other devices. Face investigation from videos and images is most predominant task for various government agencies and industry alike. The process in which unlabeled facial images are assembled according to individual characteristics is called face clustering. Clustering is required in numerous applications for example, law enforcement, disruptive design, construction, social media, health care and surveillance applications etc. Here, in this paper we reviewed some of the face clustering works done by different researchers using different methods for different applications and also gives the brief knowledge about the work done by the previous researcher.

**Keywords:** Face Recognition, Face Clustering, Video segmentation, Video Organization etc.

### I. INTRODUCTION

There is drastic increase in the use of electronic gadgets in past few years. A lot of information is generated day after day by these devices. It is very challenging to extract meaningful material from this vast available data. The process of observing the activities or pattern of specific behavior is known as surveillance [24].

In recent times we can see that other than verification based on biometric and identity proof related applications human being have found courtesy in investigating video content [17].

According to their requirement different governments are using surveillance [24][4]. They may use it for investigating crimes. It can be used to prevent criminal activities. Using different surveillance devices we can keep record of evidence. Further it can also be used to meeting intelligence requirements and analyzing statistical data. As we already know everything has two aspects likewise, this technology also has advantages as well as disadvantages. Sometimes it may be misused for planning and committing crimes by criminal organizations. It may be helpful to businesses and other private investigators for investigating and gathering intelligence. Now a days CCTV cameras is most used device for surveillance [25]. face recognition is done by many methods. We also need computers to connect these cameras to the network. Cameras have very important role in safe keeping of home and in sensing and tracing different objects [24]. Specifically it is used for monitoring an area. Other types include Analyzing social network, attendance system based on biometric applications and data mining and profiling based on this. Satellite network is also used for the surveillance purpose. Among all these surveillance which is based on videos is one of the profound tasks.

Now a days security concern is more so the task of grouping or clustering of faces in videos is significant but simultaneously it is full of challenges [12] [6] [1][7]. Specifically, when it is considered that the faces are taken out from a portion of video which may be a movie or television series or episode. Our main motive here was to divide the faces taken into a specified number of clusters. This division is carried out considering that faces allocated to each cluster being the part of the same subject. Clustering of faces is prerequisite in countless associated applications, such as video-organization, video-segmentation and retrieval of useful data or content from the videos [18]. The problem of face clustering is storming our brain because most of the time the faces are captured in uncontrolled environment. Due to that we can find a lot of variations in the poses, brightness and face related expressions. There may be the faces which are covered with hands, glasses or any other object [6].

Face grouping plays a very imperative role in extracting movie's features related semantic. It may be helpful in determining the primordial actors, creating databases' references or detecting dialogues etc. [15]. Moreover, the process of face grouping [19] we can apply for untrained or unsupervised training of face recognition algorithms. In general, it can assist as a prior or pre-processing phase in any human based appearance and task which is related to processing videos. This is done so that we can generate a human wise categorization of the data. Face recognition is done by the extracting features from the images related to the face of particular individual. On the basis of certain image similarity conditions facial clustering assemble facial appearances that refer to the same person [13].

While applying face clustering our prime motive is assembling faces into groups corresponding to persons look like many in videos based on real life [20][23]. If this can be done successfully we can apply this process to

fields, which includes video retrieval depending on content, fast growing activities like browsing and association of video assemblages etc. However, this task does not seem less than a challenge. In real-world videos, there are a lot of factors involved which may be lighting conditions, expressions of faces and pose etc. These factors can significantly alter the look of faces. There may also be full or partial blocking caused by various items such as glasses, hair and other objects that may also cause problems. The most of methods are time consuming because they test the data before make any decisions[27]. Additionally, there may be blur effect which may be triggered by gesture and low spatial resolution [11].

Face clustering process is basically associated to face recognition [14] but it has several other aspects. It is considered that in face recognition one should have prior knowledge of the data set which includes the number of persons, training facial images etc. This data set may contain some categorized facial pictures; these pictures can be employed further for teaching a classifier. While we check or test images by using the classifier we will get the best result. Therefore face recognition process is supervised problem whereas face clustering is an unsupervised problem [11] which is able to detect the person from database[26].

When we apply clustering to the face images this involves a lot of challenges. Out of these a few are:

- Availability of largest possible data set size (millions of facial images).
- Large number of classes: a crowd may contain a huge number of individuals
- There may be high intra-class variability and very less inter-class separation
- Number of clusters is unknown
- Number of samples per cluster can vary [21].

## II. LITERATURE REVIEW

[1] **LING-FEI ZHANG et.al** used convolutional neural network for Extracting automatic features. They extracted these features by using Convolutional Neural Network, S-SOM, and a unified system of SOM. In the end, to verify the efficacy of the process they have used, they have done an experiment. The authors have also analyzed the influence of parameters which are related to S-SOM as a further research on the process of face grouping.

[2] **An Tien Vo et.al** proposed the image classification model that can be used for the online advertisement display detection. In proposal model the authors have used Convolutional Neural Networks. This model has two parameters (n, m). In this n is a count of layers and m is count of filters in Convolution layer. This model is abbreviated as nLmF-CNN.

By doing various experiments they have found out the suitable value of the parameters used. In this model they have online taken images as an input data. Using deep neural network's ConvNetJs library they found various

processing components. The outcome of this method was saying or displaying YES/NO. If the advertisement is displayed clearly then we will get outcome YES otherwise we will get NO. The meaning of NO may be either the advertisement is not displayed or it is not clear. Their tentative result was 86 %.

[3] **Siyu Xia et.al** presented a spectral theory based photo grouping method. Their basic motive was to increase the performance of photo grouping method, for that they wanted to use earlier or past information of photo albums of the family. They included that conditions, first any person can appear in one photo only once, which can work as the resemblance constraint when we will go to draw graph. Secondly, in each album we cannot show an individual more times than the count of or total number of photos. This shows that dimension or size of a cluster for an individual. It represents at the most the numeral of photos in any given album. They considered this parameter as a Minimum Cost Flow (MCF) linear network based optimization problem. For that reason after constructing graph they anticipated a constrained K-Means method for the purpose of clustering data. They have used two parameters to evaluate the performance of clustering. The first one is accuracy (AC) and second one is Normalized Mutual Information (NMI).

[4] **Yichun Shi et.al** designed a clustering procedure which is known as Conditional Pairwise Clustering (ConPaC). The estimation of the adjacency matrix is done based on the resemblances or similarity among face images. In their algorithm any number of clusters can be selected and the paired similarity between the faces can be retained. In this algorithm the authors have used Loopy Belief Propagation by taking in mind the clustering problem as a Conditional Random Field (CRF) model. The outcomes showed their method is giving better result than the other clustering methods. They have also proposed a new version of already existing ConPaC. This is appropriate for bulky data set.

[5] **Leonardo Chang et.al** proposed an effective method which is based on graph. This method can be used for grouping or clustering faces in the wild. Their procedure did not necessitate past or prior knowledge of the data. The authors have conducted experiments on four data sets. Their result showed that this method is suitable and near to market solutions.

[6] **Shijie Xiao et.al** anticipated a method that is known as weighted block-sparse low rank representation abbreviated as WBSLRR. This method deliberates the existing past info while getting a low-rank data representation. The authors have also established a modest approach which is also effective to achieve the clustering results. They found that this technique is apposite for the solution of large-scale problems.

[7] **Keisuke Inoue et.al** projected an automatic face clustering method. They have employed this process as a

pre-processing step of generating surface meshes. Authors have obtained final mesh which was the collection of sub – meshes. The result that they have got from this method proved its validity.

[8] **Panagiotis Antonopoulos et.al** suggested a process that can be used to cluster face images found in video series. For that the authors have presented a new mode of finding dissimilarity matrix. This they have accomplished using Scale Invariant Feature Transform (SIFT) image features. This dissimilarity matrix is provided as an input. For evaluating the eminence of the resultant clustering they have used three parameters, which are overall entropy, statistics and F measure.

[9] **Xiaochun Cao et.al** used tensors for face representation. They accomplished the clustering job centered on the tensor data. This recommended process is known as robust tensor clustering (RTC). This algorithm initially calculates the lower-rank approximation of the actual tensor data. For finding that the used a L1 norm optimization function. RTC is a robust method as in this L1 norms do not increase the consequence of noises compared with L2 norms. At that time, they computed the higher-order singular value decomposition of the estimated tensor to achieve the ultimate grouping or clustering outcomes. They utilized a non-greedy approach to acquire an improved solution.

[10] **Xiaochun Cao et.al** recommended a Constrained Multi-view based Video Face Clustering procedure abbreviated as CMVFC. This is basically an integrated graph-based approach. Initially, they have used constraints pairwise in the entire video clustering practice. It is done in both representation based on sparse-subspace and the spectral clustering. In sparse subspace based representation they explored unidentified relationship. The parameters they used were used for guiding to learn more reasonable new representations in the constrained spectral clustering,. Second, their process ponders both the video face pairwise constraints along with the multi-view consistence concurrently. Precisely, the regularization of graph implements the constraints pairwise and the co-regularization castigates the deviation amid different graphs of numerous views.

### III. COMPARATIVE ANALYSIS

We have reviewed various papers on face clustering. In the following table we have listed some papers. A comparative analysis is done on the basis of the work done by various authors by using different algorithms or methods. We have also listed some of the advantages and disadvantages of these methods.

Table 1: Comparative Analysis

Year of publication	Author's name	Algorithm/method used	Advantages	Disadvantages
2018	LING-FEI ZHANG et.al	Integrated SOM and spectral clustering	Visually alike individuals are grouped together.	Low accuracy of clustering
2017.	An Tien Vo. et.al	nLmF-Convolutional Neural Network	Their process was suitable for advertisement image classification	Problem in identifying an appropriate pair parameter for particular dataset
2014	Siyu Xia et.al	Constrained K-Means	They have achieved a better outcome compared with other clustering procedures.	There is an effect of increasing album size on the clustering process
2018	Yichun Shi, et.al	Conditional Pairwise Clustering (ConPaC)	Performed good on the BLUFR face verification	It is difficult to find unary potentials for more robust face clustering
2019	Leonardo Chang et.al	Graph-based method for clustering	Their method gave outcome that can give improved performance than the other systems.	Not used prior knowledge data
2014	Shijie Xiao et.al.	WBSLRR	Appropriate for the solution of large-scale problems	Time consuming
2001	Keisuke Inoue et.al	Surface mesh generation based automatic procedure	customizable to numerous applications that have different necessities on the size and shape of a region	The faces with a enormous disparity of normal vector or with a long complex boundary can cause poorly shaped regions
2007	Panagiotis Antonopouloset.al	SIFT image features based calculation of dissimilarity matrix	clustering results are very satisfactory	It is not planned to calculate automatically the natural assemblage of the input data. As a result, the user has to give the number of clusters himself
2015	Xiaochun Cao et.al	RTC	Performed better than the other clustering processes and is more robust to noises.	It cannot cope-up with variations related to pose
2015	Xiaochun. Cao et.al	CMVFC	It has shown how to employ the intrinsic assistance of a video to help the face clustering	It is not sensible every time to implement the consistence through multiple views too much

#### IV. CONCLUSION

The main motive of this paper to reviewed different methods which is used for face clustering by different authors. Its also show the importance of face recognition and clustering of images. In the above table a comparative analysis of different methods used by different authors is shown. Nowadays researchers are paying more attention in face clustering methods because recognition of an image is very important for the security of the users. This technique is basically related to recognition process of faces but it has quite a lot of diverse features. A lot of effort has been done on face related recognition processes but very less attention have been given on face clustering. This paper show the different technique used in previous research and show the comparison among them. Face clustering finds its application in various areas.

#### REFERENCES

- [1] Zhang, Ling-Fei, et al. "Research On Face Image Clustering Based On Integrating Som And Spectral Clustering Algorithm." *2018 International Conference on Machine Learning and Cybernetics (ICMLC)*. Vol. 2. IEEE, 2018.
- [2] Vo, An Tien, Hai Son Tran, and Thai Hoang Le. "Advertisement image classification using convolutional neural network." *2017 9th International Conference on Knowledge and Systems Engineering (KSE)*. IEEE, 2017.
- [3] Xia, Siyu, Hong Pan, and A. Kai Qin. "Face clustering in photo album." *2014 22nd International Conference on Pattern Recognition*. IEEE, 2014.
- [4] Shi, Yichun, Charles Otto, and Anil K. Jain. "Face clustering: representation and pairwise constraints." *IEEE Transactions on Information Forensics and Security* 13.7: 1626-1640, 2018
- [5] Chang, Leonardo, Aírel Pérez-Suárez, and Miguel González-Mendoza. "Effective and Generalizable Graph-Based Clustering for Faces in the Wild." *Computational Intelligence and Neuroscience* 2019 (2019).
- [6] Xiao, Shijie, Mingkui Tan, and Dong Xu. "Weighted block-sparse low rank representation for face clustering in videos." *European Conference on Computer Vision*. Springer, Cham, 2014.
- [7] Inoue, Keisuke, et al. "Face clustering of a large-scale cad model for surface mesh generation." *Computer-Aided Design* 33.3: 251-261, 2001.
- [8] Antonopoulos, Panagiotis, Nikos Nikolaidis, and Ioannis Pitas. "Hierarchical face clustering using sift image features." *2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing*. IEEE, 2007.
- [9] Cao, Xiaochun, et al. "Robust face clustering via tensor decomposition." *IEEE transactions on cybernetics* 45.11: 2546-2557, 2014
- [10] Cao, Xiaochun, et al. "Constrained multi-view video face clustering." *IEEE Transactions on Image Processing* 24.11: 4381-4393, 2015.
- [11] Wu, Baoyuan, et al. "Constrained clustering and its application to face clustering in videos." *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2013.
- [12] Sheffer, Alla. "Model simplification for meshing using face clustering." *Computer-Aided Design* 33.13: 925-934, 2001.
- [13] Vretos, Nicholas, Vassilios Solachidis, and Ioannis Pitas. "A mutual information based face clustering algorithm for movie content analysis." *Image and Vision Computing* 29.10: 693-705, 2011.
- [14] Li, Zhifeng, and Xiaou Tang. "Bayesian face recognition using support vector machine and face clustering." *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*. Vol. 2. IEEE, 2004.
- [15] Vretos, Nicholas, Vassilios Solachidis, and Ioannis Pitas. "A mutual information based face clustering algorithm for movies." *2006 IEEE International Conference on Multimedia and Expo*. IEEE, 2006.
- [16] Pei, Xiaobing, et al. "Manifold adaptive label propagation for face clustering." *IEEE transactions on cybernetics* 45.8: 1681-1691, 2014.
- [17] Tao, Ji, and Yap-Peng Tan. "Face clustering in videos using constraint propagation." *2008 IEEE International Symposium on Circuits and Systems*. IEEE, 2008.
- [18] Sharma, Vivek, M. Saquib Sarfraz, and Rainer Stiefelhagen. "A simple and effective technique for face clustering in tv series." 2017.
- [19] Tapaswi, Makarand, Marc T. Law, and Sanja Fidler. "Video face clustering with unknown number of clusters." *Proceedings of the IEEE International Conference on Computer Vision*. 2019.
- [20] Tang, Zhiqiang, et al. "Face clustering in videos with proportion prior." *Twenty-Fourth International Joint Conference on Artificial Intelligence*. 2015.
- [21] Otto, Charles, Brendan Klare, and Anil K. Jain. "An efficient approach for clustering face images." *2015 International Conference on Biometrics (ICB)*. IEEE, 2015.
- [22] Cai, Deng, Xiaofei He, and Jiawei Han. "Document clustering using locality preserving indexing." *IEEE Transactions on Knowledge and Data Engineering* 17.12: 1624-1637, 2005.
- [23] Kulshreshtha, Prakhar, and Tanaya Guha. "An Online Algorithm for Constrained Face Clustering in Videos." *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018.
- [24] Asad, Mahim-Ul, Rashed Mustafa, and Mohammad Shahadat Hossain. "An Efficient Strategy for Face Clustering use in Video Surveillance System." *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*. IEEE, 2019.
- [25] Chu, Wei-Ta, Ya-Lin Lee, and Jen-Yu Yu. "Visual language model for face clustering in consumer photos." *Proceedings of the 17th ACM international conference on Multimedia*. 2009.
- [26] Suma S L, Sarika Raga, "Real time face Recognition of human Faces by using LBPH and Viola Jones AlgorithmReal Time Face Recognition of human Faces by using LBPH and Viola Jones Algorithm", *International Journal of Scientific Research in computer Science and Engineering*, vol.6, Issue.5, pp.6-10, 2018.
- [27] A.I. Abughfa, A.B. Elmadani, "Offline Signature Verification Based on Image Processing and Hu Moment", *International Journal of Scientific Research in Network Security and Communication*. Vol.4, Issue.5, pp.1-7, 2016

#### AUTHORS PROFILE

Ms. Rakhi pursued bachelor of engineering in computer science of technology from Bhagat Singh Phool Mahila Vishwavidhyala Khanpur (Haryana) in 2018 and currently pursuing Master of technology in computer science and technology from Deenbandhu Chhotu Ram University of Science And Technology, Haryana, India

