

Applying Face Recognition in Video Surveillance Security Systems

Bauyrzhan Omarov¹, Batyrkhan Omarov^{2,3}, Shirinkyz Shekerbekova⁴, Farida Gusmanova¹, Nurzhamal Oshanova⁴, Alua Sarbasova¹, Zhanna Yessengaliyeva¹, Agyn Bedelbayev¹, Akmarzhan Maikhanova¹, Nurzhan Omarov⁵, Daniyar Sultan¹

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

²International Information Technology University, Almaty, Kazakhstan

³ Kazakhstan Innovations Lab Supported by UNICEF, Almaty, Kazakhstan

⁴Abay Kazakh National Pedagogical University

⁵Kazakh University of Railways and Communications

batyahan@gmail.com

Abstract. Face Detection and Recognition is an important surveillance problem to provide citizens' security. Nowadays, many citizen service areas as airports, railways, security services are starting to use face detection and recognition services because of their practicality and reliability. In our research, we explored face recognition algorithms and described facial recognition process applying Fisherface face recognition algorithm. This process is theoretically justified and tested with real-world outdoor video. The experimental results demonstrate practically applying of face detection from several foreshortenings and recognition results. The given system can be used in building a smart city as a smart city application, also in different organization to ensure security of people.

Keywords: Smart City, Video Surveillance, Face Recognition, Face Detection.

1 Introduction

If we talk about the concept of "smart city", first and foremost, it is improving the quality of life and creating comfortable living conditions for citizens. This is the combination of various technologies, management of communications, infrastructure, in the near future IOT.

The goal is the optimal use of modern technologies in each of the spheres of city life for more rational use of resources and improving the quality of life, doing business, etc. So, "Safe City" is the most important component of the "smart city" concept, besides video surveillance as part of a safe city, the state is becoming "the eyes" of a smart city.

Smart cities often intersect with a digital city, a wireless city, a safe city, an eco-city, a city with low carbon monoxide emissions, architectural perfection and other regional development concepts. This should be confused with the concepts of the industry of information technologies, electronic document management, electronic reporting, intellectual transport and an intelligent urban water / gas / power supply

network. Smart City is sharing data over the Internet, cloud services, geospatial infrastructure, dedicated telecommunication channels and other new generations of information technology. CCTV cameras, included in open or protected monitoring and control systems, ensure broadband cross-border interaction of all municipal structures, and facilitate the intellectual integration of applications into user innovations, open innovations, public innovations, and joint innovations. The process of transition to a smart city is characterized by a steady interest of both local enterprises and foreign investors. In this process, there are no templates for the use of video surveillance and network technologies. The main thing is an intelligent and cost-effective result. In addition, of course, there are increased requirements to the processing of video data streams, the quality of video surveillance equipment.

2 Literature Review

There are several approaches to create a face recognition algorithm.

The empirical approach was used at the very beginning of the development of computer vision. It is based on some of the rules that a person uses to detect a face. For example, the forehead is usually brighter than the central part of the face, which, in turn, is uniform in brightness and color. Another important feature is the presence of parts of the face in the image - the nose, mouth, eyes. To determine the faces, we did a significant reduction of the image area, where the presence of a face was assumed, or perpendicular histograms are constructed. These methods are easy to implement, but they are practically unsuitable in the presence of a large number of foreign objects in the background, several persons in the frame or when changing the angle.

The following approach uses invariant features characteristic of a face image. At its core, as in the previous method, lies the empiricist, that is, the attempt of the system to "think" as a person. The method reveals the characteristic parts of the face, its boundary, change in shape, contrast, etc., combines all these signs and verifies. This method can be used even when turning the head, but with the presence of other faces or a heterogeneous background, recognition becomes impossible.

The following algorithm is the detection of faces using patterns that are specified by the developer. A person appears to be a kind of template or standard, the purpose of the algorithm is to check each segment for the presence of this pattern, and the check can be made for different angles and scales. Such a system requires many time-consuming calculations.

All modern facial recognition technologies use systems that learn through test images. For training, bases with images containing faces and not containing faces are used. Each fragment of the investigated image is characterized as a feature vector, with which the classifiers (algorithms for determining an object in a frame) determine whether this part of the image is a face or not.

Currently, several dozens of computer methods for face recognition are actively used: methods based on neural networks [1]; the main components (own persons) [2], [3-5]; based on linear discriminant analysis [6-7]; elastic graph method [8]; a method

based on hidden Markov models [9-13]; method based on flexible contour models of the face; method of comparison of standards; optical flux method; methods based on lines of the same intensity; algebraic moments; Karunen-Loeve decomposition; fuzzy logic; Gabor filters, etc. A good overview of these methods can be found in [14].

One of the first developed methods of facial recognition is the method of main components (own faces). Its distinguishing feature is that the main components carry information about the signs of a certain generalized face. Face recognition using linear discriminant analysis is based on the assumption of linear separability of classes (persons) in image space. Neural network methods have a good generalizing ability.

3 Facial Recognition Problem

Recognition of objects is an easy task for people, the experiments conducted in [15] showed that even children aged one to three days are able to distinguish between remembered faces. Since a person sees the world not as a set of separate parts, our brain must somehow combine various sources of information into useful patterns. The task of automatic face recognition is to isolate these significant features from an image, transforming them into a useful presentation and producing some kind of classification.

The process of face recognition, which is based on geometrical features of the face, is probably the most intuitive approach to the problem of face recognition [15-16]. Experiments on a large data set have shown that, alone, geometric features cannot provide enough information for face recognition.

In this work, we explore face detection and recognition process, describe their mathematical representation and do experiments with facial recognition using Fisherface algorithm.

3.1 Development Overview

The solution as proposed in this research work consists of two parts as recovering low resolution image and the identity of object using the recovered high resolution image.

Image restoration part consists of three subtasks as

1. Converting the low resolution image to digital form
2. Image enhancement and recovery
3. Converting to graphical image from digits.

3.2 Face Detection

At the first stage, the face is detected and localized in the image. At the recognition stage, the image of the face is aligned (geometric and luminance), the calculation of the signs and the direct recognition - the comparison of the calculated signs with the standards embedded in the database. The main difference of all the algorithms presented will be the calculation of signs and the comparison of their aggregates among themselves. Such face detection system types shown in Figure 1.

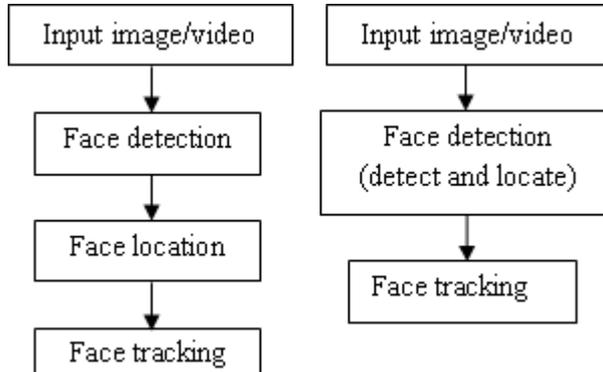


Fig.1. Face detection architecture.

3.3 Face Recognition

There are several different face recognition algorithms as correlation, eigenfaces, linear subspaces and fisherfaces. There were several experiments on identification the effectiveness of those algorithms where the FisherFaces algorithm was chosen as the best one with the lowest error rate in human face recognition. In accordance with experiment results made before we decided to choose the FisherFaces algorithm for face identification and recognition processes due to its fast and guaranteed recognition of the human. Figure 2 shows the plot that illustrates the error rate depending on the number of principal components.

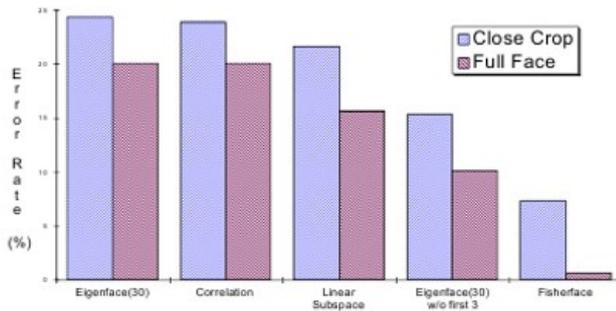


Fig.2. Face detection architecture.

As can be seen from the above graph, the FisherFace method learns the set of projections which perform well over a range of lighting variation, facial expression and even presence of glasses. Below, we explain the algorithmic description of the Fisherfaces method:

Let there be a random vector with samples drawn from classes:
 $X = \{X_1, X_2, \dots, X_n\}$

$$X_i = \{X_1, X_2, \dots, X_n\} \quad (1)$$

The scatter matrices S_B and S_W are calculated as:

$$S_B = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (2)$$

$\sum_{i=1}^c \sum_{x_i \in X_i} (x_i - \mu_i)(x_j - \mu_j)^T$, where μ is the total mean:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (3)$$

And μ_i is the mean of class $i \in \{1, \dots, c\}$:

$$\mu_i = \frac{1}{|x_i|} \sum_{x_j \in X_j} x_j \quad (4)$$

Fisher's classic algorithm now looks for a projection that maximizes the class separability criterion:

$$W_{opt} = \arg \max_W \frac{|W^T S_B W|}{|W^T S_W W|} \quad (5)$$

Following the method of Belhumer, Hespanha and Kriegman, a solution for this optimization problem is given by solving the General Eigenvalue Problem:

$$S_W^{-1} S_B v_i = \lambda_i v_i \quad (6)$$

There's one problem left to solve: The rank of S_W is at most (N-c), with N samples and classes. In pattern recognition problems the number of samples N is almost always smaller than the dimension of the input data (the number of pixels), so the scatter matrix S_W becomes singular. In [BHK97] this was solved by performing a Principal Component Analysis on the data and projecting the samples into the (N-c)-dimensional space. A Linear Discriminate Analysis was then performed on the reduced data, because S_W isn't singular anymore. The optimization problem can then be rewritten as:

$$W_{fld} = \arg \max_W \frac{|W^T W_{pca}^T S_B W_{pca} W|}{|W^T W_{pca}^T S_W W_{pca} W|} \quad (7)$$

The transformation matrix that projects a sample into the (c-1) dimensional space is then given by:

$$W = W_{fld}^T W_{pca}^T \quad (8)$$

Face detection, recognition and gender classification experiments carried out on the basis of facial images database [25]. Sample images are shown in Figure 3. In the formation of the database size of the images and the shooting conditions were the same. They used a 24-bit JPEG format. The base [25] contains pictures of people, male and female, of different nationalities and ages. It reflects changes in a person's appearance: different hairstyles, beards and glasses presence. In preparation for the experiment, two training samples have been created. The first of them contains five images of each person (only $5 \times 395 = 1975$ images). Second, 10 images of each person's individual learning ($10 \times 395 = 3950$ images). In addition, the dataset has several datasets as Face94, Face95, Face96, and Grimace that the characteristics are listed, below.



Fig. 3. Sample images of faces.

The approach that is used in this method finds out the facial features to discriminate between the persons. The performance of the system that uses the FisherFaces algorithm is highly depends on the input data. The FisherFaces provides a total reconstruction of the projected image by normalizing processing of the image [6, 26-28]. The total set of procedures is given in the Figure 4.

As can be seen from the Figure 4, the process of face verification starts with the detection stage, where the image is taken from the camera and is considered as an

input data. Then, there goes the normalization process in order to construct the proper image that can be used in FisherFace algorithm. Face normalization actually consists of geometry normalization, background removal and lighting normalization. The images of the face are normalized to a fixed size. If the face was in a wrong angle this angle is determined then is corrected in accordance with rules.

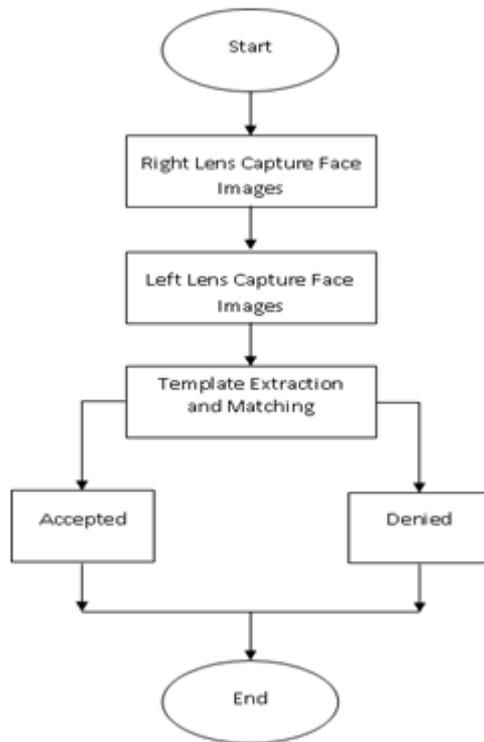


Fig. 4. Face detection architecture.

4 Facial Recognition Problem

Face recognition system generally involves two main stages as “Face Detection” and “Face Identification”. First one is face detection, where the system is searching for any faces then takes the image of this face. Following this, image processing cleans up the facial image into black-white colors. In our research, face can be detected from several foreshortenings. Implemented results are given in Figure 5.



Fig. 5. Face detection from several foreshortenings.

After detecting face, next step will be executed. In this step, feature extraction and verification process will be done. After recognizing the detected and processed facial image is compared to a database of faces in order to decide who that person is, Figure 6.

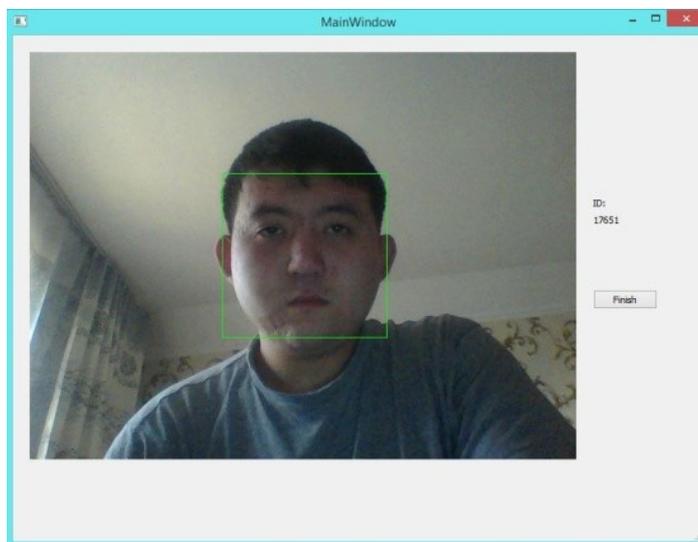


Fig.6. Face recognition. Identification of personal id

5 Conclusion

In this work, we applied Fisherface face recognition algorithm for facial recognition problem as a video surveillance system of Smart City application. We used Fisherface algorithm because of its practicality and high recognition rate. The mathematical representation of facial recognition problem and Fisherface algorithm was investigated. Experiment results demonstrate face detection and recognition results. Further, we are going to use the proposed system as an application of a Smart City Platform and for schools to identify pupils by faces.

References

- [1] R. Collins et al. "A system for video surveillance and monitoring," VSAM Final Report, Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-00-12, May 2000.
- [2] Haritaoglu, "Harwood and Davis, W4: Real time surveillance of people and their activities," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, no. 8, pp. 809–830, Aug. 2000.
- [3] Remagnino, Jones, Paragios, and Regazzoni, Video Based Surveillance Systems Computer Vision and Distributed Processing. Norwell, MA: Kluwer , 2002.
- [4] G. Stauffer, "Learning patterns of activity using real-time tracking," IEEE Trans. Pattern Anal. Machine Intell., vol. 22, no. 8, pp. 747–757, Aug. 2000.
- [5] Remagnino, Jones, Paragios, and Regazzoni, Video Based Surveillance Systems Computer Vision and Distributed Processing. Norwell, MA: Kluwer , 2002.
- [6] VACE: Video Analysis and Content Exploitation [Online]. Available: <http://www.icarda.org/InfoExploit/vace/>
- [7] A. K. Jain, R. Bolle, and S. Pankanti, "Biometrics: Personal Identification in Networked Security," A. K. Jain, R. Bolle, and S. Pankanti, Eds.: Kluwer Academic Publishers, 1999.
- [8] Qianwen Wan, Shishir Paramathma Rao, Aleksandra Kaszowska, V. Voronin, Karen Panetta, Holly A. Taylor, Sos Agaian, "Face description using anisotropic gradient: thermal infrared to visible face recognition", Proc. SPIE 10668, Mobile Multimedia/Image Processing, Security, and Applications 2018, 106680V (14 May 2018); doi: 10.1117/12.2304898; <https://doi.org/10.1117/12.2304898>
- [9] Wolf M. (2018) Image and Video Analysis. In: Smart Camera Design. Springer, Cham
- [10] SantoshKumaraAmitPandey, K.Sai Ram Satwik. Deep learning framework for recognition of cattle using muzzle point image pattern. Measurement Volume 116, February 2018, Pages 1-17
- [11] Kumar S, Tiwari S., Singh S.K, Face recognition for cattle, in: Proceedings of 3rd IEEE International Conference on Image Information Processing (ICIIP), 2015, pp. 65–72.
- [12] M.A. Turk, A.P. Pentland, Face recognition using eigenfaces, in: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '91), 1991, pp. 586–591.
- [13] P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman Eigenfaces vs. fisherfaces: Recognition using class specific linear projection IEEE Trans. Pattern Anal. Mach. Intell., 19 (7) (1997), pp. 711-720
- [14] G. Baudat, F. Anouar Generalized discriminant analysis using a kernel approach Neural Comput., 12 (10) (2000), pp. 2385-2424
- [15] K.R. Muller, S. Mika, G. Ratsch, K. Tsuda, B. Scholkopf An introduction to kernel-based learning algorithms IEEE Trans Neural Netw., 12 (2) (2001), pp. 181-201
- [16] Kang M. G. Park S. C., Park M. K. Superresolution image reconstruction: a technical overview. IEEE Signal Processing Magazine, 20:21–36, 2013.
- [17] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips. Face recognition: A literature survey. ACM Computing Surveys, pages 399–458, 2003.
- [18] Ion Marqu'es. Face Recognition Algorithms. Universidad delPais Vasco. June 2010.

- [19] M. Turk. A random walk through eigenspace. *IEICE Transactions on Information and Systems*, E84-D(12):1586–1595, December 2001.
- [20] M.-H. Yang, D. Kriegman, and N. Ahuja. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34–58, January 2002.
- [21] S. Zafeiriou, C. Zhang, Z. Zhang, A Survey on Face Detection in the wild: past, present and future, *Computer Vision and Image Understanding* (2015).
- [22] Greenburg E (2004) Codifying new urbanism: how to reform municipal land development regulations. American Planning Association PAS Report Number 526
- [23] LEED, Leadership in Energy and Environmental Design <http://www.usgbc.org>. Accessed 22 April 2011EU Strategic Energy Technology (SET) plan http://ec.europa.eu/energy/technology/set_plan/set_plan_en.htm.
- [24] Dr Libor Spacek. 2016. Collection of Facial Images. Designed and maintained by Dr Libor Spacek. Electronic resource. <http://cswww.essex.ac.uk/mv/allfaces/index.html>, June 2016.
- [25] Suliman, A., & Omarov, B. S. (2018). Applying Bayesian Regularization for Acceleration of Levenberg-Marquardt based Neural Network Training. *International Journal of Interactive Multimedia & Artificial Intelligence*, 5(1).
- [26] Omarov, B., Altayeva, A., & Im Cho, Y. (2017, June). Smart building climate control considering indoor and outdoor parameters. In *IFIP International Conference on Computer Information Systems and Industrial Management* (pp. 412-422). Springer, Cham.
- [27] Altayeva, A., Omarov, B., & Im Cho, Y. (2018, January). Towards Smart City Platform Intelligence: PI Decoupling Math Model for Temperature and Humidity Control. In *2018 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 693-696). IEEE.
- [28] Altayeva, A., Omarov, B., & Im Cho, Y. (2017, December). Multi-objective optimization for smart building energy and comfort management as a case study of smart city platform. In *2017 IEEE 19th International Conference on High Performance Computing and Communications; IEEE 15th International Conference on Smart City; IEEE 3rd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)* (pp. 627-628). IEEE.