

Selection of Face Detection and Recognition Algorithms for the E-learning System

Oleh Shkodzinsky^a, Mykhailo Lutskiv^a and Marian Smoliy^a

^a Ternopil Ivan Puluj National Technical University, Ternopil, 46001 Ukraine

Abstract

The relevance of the development and implementation of photo fixation and face recognition tools in e-learning systems is substantiated, and an extended list of problems that should be solved is formulated. The main algorithms and approaches to the detection and recognition of faces were considered, as a result of which an effective combination of algorithms was chosen, suitable for use in the means of testing knowledge in learning management systems.

Keywords 1

Face recognition, photo fixation, knowledge testing, image recognition algorithms, person identification, identification accuracy

1. Introduction

The development of modern society is characterized by the powerful influence of computer technologies that permeate all spheres of human activity. Informatization of education is an important component of these processes. Computer technologies are becoming an integral part of the holistic educational process, which creates the prerequisites for its transformation and significant increase in efficiency.

Two leading trends have the greatest use in informatization of the educational process:

- use of information technology products as a means of teaching and cognition;
- use of information technologies to control the student's educational and cognitive activity, that together became the foundation for building a virtual educational environment of a modern educational institution [1].

Computer-based knowledge testing is increasingly used as a form of monitoring students' performance. The introduction of first testing, and later computer-based testing into the educational process became a purposeful step on the way to eliminating subjectivism in the assessment of the student's cognitive activity. Forms of student knowledge assessment in the form of tests have proven themselves as one of the most promising means of increasing the effectiveness of quality management of the educational process, despite the presence of both supporters and opponents of knowledge testing. The ratio between supporters and opponents is in a significant correlation to the quality indicators of the test material and implemented algorithms of knowledge testing.

The development of qualitative tests is a time-consuming task and requires compliance with such requirements as: significance; scientific credibility; representativeness (presence in the test of the main structural elements of the course content in required amount for assessment); increasing complexity of educational material; variability depending on the content of the studied material and the amount of hours; content systematicity; validity; comprehensiveness and balance of the test; relationship between content and form.

But there is another side of the issue, which opponents of testing use as an argument in their favor - it is the emergence of new opportunities for an unfair attitude towards passing knowledge test,

ITTAP'2022: 2nd International Workshop on Information Technologies: Theoretical and Applied Problems, November 22–24, 2022, Ternopil, Ukraine

EMAIL: shkod@tntu.edu.ua (A. 1); msilence2009@gmail.com (A. 2); email3@mail.com (A. 3)

ORCID: 0000-0002-9983-0471 (A. 1)



© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

especially in the conditions of distance learning, when the test-takers are dispersed in space and are out of examiner's visual control. This requires additional tools and measures for monitoring and verifying the progress of the testing process, which would provide confirmation of the integrity of the testing for each participant.

The most common abuse is the unauthorized use of the Internet resources or electronic resources in another operating system or browser and handy tools (books, notes, mobile devices, etc.), as well as the replacement of the test person. Effective solutions to these problems can be the development and use of means for activity monitoring of the browser tab where test takes place and the introduction of photo fixation and recognition tools implemented on the basis of modern information technologies.

Another feature of solving this problem is that the vast majority of educational institutions already use educational content management systems based on web technologies, and photo fixation and recognition tools should be able to be integrated into such systems and not require significant funds for the implementation of appropriate software and hardware, it can be implemented on standard user platforms, including mobile.

2. The formulation of the problem

Face detection, which is a variation of the general object detection problem, can be defined as determining whether a given image contains a face, and if it does, finding the location of each face [2]. Face detection is a key task because it is a prerequisite for other tasks such as face locating [3], face recognition [4], face analysis [5], face verification [6], face labeling and extraction [7], face tracking [8], recognition of emotions and facial expressions [9].

Facial recognition is the main task in the identification or verification of an individual using the features of his face. This is one of the most important problems of computer vision with great commercial interest.

Nowadays, there are several dozens of computer-based face detection and recognition methods [10]. However, these methods do not provide 100% reliability of identification and, at the same time, have limitations in recognition performance. Most of the early algorithms (before 2012) could not provide sufficient performance and accuracy due to high variability of images. Among the main challenges and problems of recognition are:

- illumination – some images may have very high or low illumination relative to the background (contrast) and may also be partially shaded;
- facial expressions – can vary widely, representing the full range of human emotions and states;
- skin types – different shades of facial skin;
- distance - if the distance to the camera is too large, the size of the face image may be too small;
- orientation - the orientation and angle of the face relative to the camera can vary significantly;
- complex background - a significant number of objects in the scene reduces the accuracy and speed of detection;
- several faces in one image - images with a large number of faces are very difficult for accurate and quick detection;
- overlapping faces - faces can be partially hidden by objects such as glasses, scarves, hands, hair, hats, medical masks and other objects, which affects the accuracy and speed of detection;
- low resolution - images with low resolution, as well as with high visual noise, are recognized with less accuracy and more slowly.

Despite the fact that today there is a great variety of methods, the general structure of the face detection and recognition process can be distinguished (see Fig. 1).

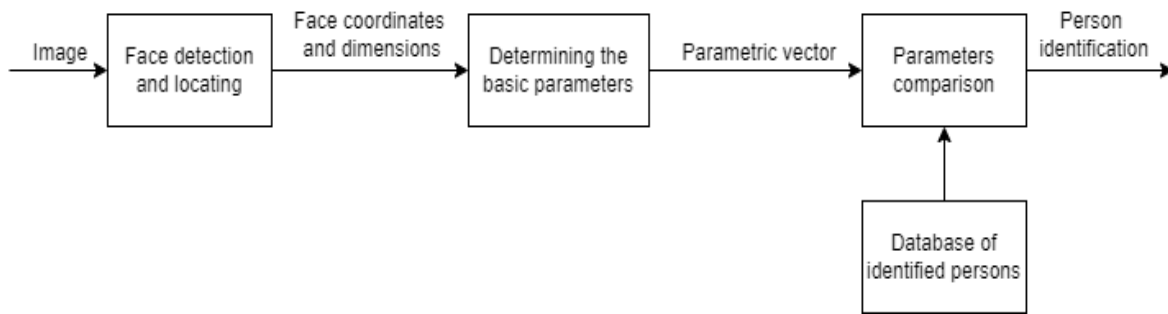


Figure 1: The general process of face image processing during recognition

At the first stage, detection and locating (finding coordinates) of the face on the image is performed. The best results are achieved when the person is looking directly into the camera when the image is captured, but modern algorithms also allow face detection in situations where the person is not looking directly into the camera (within certain limits, of course). The result of detection and locating are the found coordinates of the face(s) and its dimensions.

At the stage of determining the basic parameters, the face image is aligned and normalized (geometrically and by brightness), the face is encoded into a set of basic parameters formed by a parametric vector (array). After that, direct recognition takes place - a comparison of calculated parametric vectors with identified ones located in the database of identified persons. The main difference between all the considered algorithms will be the face detection mechanism itself and the mechanism for calculating the basic parameters - translating the face into a parametric vector.

The most common facial recognition algorithms available today include:

- principal component analysis, PCA
- Viola-Jones algorithm
- HOG algorithm and its combination with the SVM classifier
- deep convolutional neural networks (Deep CNNs): AlexNet, VGG, ResNet, etc.

3. Analysis of face detection and recognition methods

Let's consider the most promising of the listed methods of detecting and recognizing faces in images, analyze their advantages and disadvantages and choose the most optimal one for our project.

The requirements for algorithms for this project are:

- accuracy of identification - more than 95%;
- recognition speed - up to 1 second.

3.1. Principal component analysis

Principal component analysis (PCA) is a way to reduce the dimensionality of data while losing the minimum amount of information. The process of calculating the principal components is reduced to the calculation of the eigenvectors and eigenvalues of the covariance matrix of the input data, or to the singular decomposition (SVD) of the data matrix. The method of principal components is statistical and operates not with images, but with vectors in linear space [11]. In cases where there are significant changes in the level of illumination or the facial expression of a person in the image, the effectiveness of the method is significantly reduced [11].

Among the advantages of this algorithm, we can single out low requirements for computing power. However, it has significant disadvantages: high sensitivity (compared to other algorithms) to lighting, facial expressions, and facial angle; stricter requirements for image quality compared to other algorithms. These disadvantages causes low recognition accuracy (usually not higher than 80-90%).

3.2. Viola-Jones algorithm

The Viola-Jones algorithm is based on the integral representation of the image by Haar-like features, the construction of a classifier based on the adaptive boosting algorithm (AdaBoost), and the method of combining classifiers into a cascade structure. This method demonstrates high efficiency in solving the task of searching for objects in images and video streams in real time. Has a low probability of false face detection. Allows to detect faces at angles up to 30°. The accuracy of identification can reach values of more than 90% [12]. The method was developed in 2001, has a large number of implementations and is widely used, as it is simple and effective. The Viola-Jones algorithm is also implemented in the OpenCV free library.

Advantages over the PCA algorithm:

- low percentage of false positives;
- high speed of operation (up to several tens of milliseconds on modern CPUs);
- slightly higher accuracy;
- simplicity of software implementation (thanks to ready-made implementation in OpenCV).

3.3. HOG-algorithm with SVM-classifier

The Histogram of Oriented Gradients (HOG) algorithm combined with the Support Vector Machine (SVM) can be used to train highly accurate object classifiers, including human faces - first demonstrated by N. Dalal and B. Triggs in their paper Histogram of Oriented Gradients for Human Detection [13]. HOG counts the number of specific gradient orientations in local areas of the image. The idea is that the distribution of the local intensity and directionality of the gradient describes the local appearance and shape of the object [14].

Advantages:

- much higher accuracy than Haar-like cascades (Viola-Jones method);
- more stable recognition than Haar-like cascades;
- good operating speed (less than a second on modern CPUs).

Disadvantages:

- sensitivity to the angle of the face, requires a frontal view;
- accuracy is inferior to deep convolutional networks.

This algorithm satisfies the set requirements, it has a fairly high accuracy with still sufficient speed, and can be used in this system.

3.4. Deep convolutional neural networks

Deep convolutional neural networks (Deep CNNs) are a type of artificial neural network (ANN) modeled after the mammalian visual cortex. The main components of a CNN are convolutional filters, an aggregation layer, a pooling layer, a fully connected (FC) layer, and a loss function layer. CNNs are used in a wide range of solutions for the tasks of object and action recognition, object detection, computational photography, and natural language processing. To date, deep CNNs have achieved outstanding success in most computer vision tasks and dominated many well-known competitions such as the ImageNet Large Scale Visual Recognition Challenge (ILSVRC)[13].

In 2015, deep convolutional neural networks surpassed the human level in image classification [15].

Advantages:

- very high recognition accuracy can be achieved (>99%);
- resistance to variability of data and input noise (angle of the face, lighting, shadows, etc.).

Disadvantages:

- high computational complexity, requires graphics processor (GPU) resources.

4. Conclusions

Considering all of the above, to improve the quality of the system under development, it may be relevant to create hybrid methods that use the advantages and disadvantages of several considered algorithms. Thus, combining the faster HOG+SVM method and the slower, but more accurate CNN method (for cases when the first one does not work satisfactorily) will allow to create an effective detector.

Considering that it is the type of CNNs - residual neural networks (ResNet) that have overcome the human level of image classification, and also dominated ImageNet competitions for several years in a row [15] and demonstrate high recognition accuracy at sufficient speed - it is reasonable to implement the recognition stage in this solution using ResNet.

5. References

- [1] Dyachuk S. F., Konovalenck I. V., Shkodzinsky O. K. Virtual educational environment of TNTU based on LMS ATutor // International scientific-practical seminar "Theory and practice of distance learning of foreign citizens: domestic and international experience" NURE 12: 11-15.
- [2] Yang, M. H., Kriegman, D. J., and Ahuja, N., 2002. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34-58. ISSN 01628828. doi:10.1109/34.982883.
- [3] Lam, K.-M. and Yan, H., 1994. Fast algorithm for locating head boundaries. *Journal of Electronic Imaging*, 3(4):351-359. ISSN 1017-9909. doi:10.1117/12.183806.
- [4] Chi, L., Zhang, H., and Chen, M., 2017. End-To-End Face Detection and Recognition. arXiv preprint, 1703.10818:1-9.
- [5] Moghaddam, B. and Pentland, A., 1997. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696-710. ISSN 01628828. doi:10.1109/34.598227.
- [6] Ou, W., You, X., Tao, D., Zhang, P., Tang, Y., and Zhu, Z., 2014. Robust face recognition via occlusion dictionary learning. *Pattern Recognition*, 47(4):1559-1572. ISSN 00313203. doi:10.1016/j.patcog.2013.10.017.
- [7] Gao, Y. and Qi, Y., 2005. Robust visual similarity retrieval in single model face databases. *Pattern Recognition*, 38(7):1009-1020. ISSN 00313203. doi:10.1016/j.patcog.2004.12.006.
- [8] Essa, I. and Pentland, A., 2002. Facial expression recognition using a dynamic model and motion energy. In *Proceedings of 5th IEEE International Conference on Computer Vision*, pp. 360-367. IEEE. doi: 10.1109/ iccv.1995.466916. 20-23 June 1995.
- [9] S. Agrawal and P. Khatri, "Facial Expression Detection Techniques: Based on Viola and Jones Algorithm and Principal Component Analysis," 2015 Fifth International Conference on Advanced Computing & Communication Technologies, 2015, pp. 108-112, doi: 10.1109/ACCT.2015.32.
- [10] K. Dang and S. Sharma, "Review and comparison of face detection algorithms," 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017, pp. 629-633, doi: 10.1109/CONFLUENCE.2017.7943228.
- [11] Kwang In Kim, Keechul Jung and Hang Joon Kim, "Face recognition using kernel principal component analysis," in *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 40-42, Feb. 2002, doi: 10.1109/97.991133.
- [12] Yi-Qing Wang, An Analysis of the Viola-Jones Face Detection Algorithm, *Image Processing On Line*, 4 (2014), pp. 128-148. <https://doi.org/10.5201/ipol.2014.104>
- [13] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 886-893 vol. 1, doi: 10.1109/CVPR.2005.177.
- [14] ImageNet Winning CNN Architectures (ILSVRC). URL: <https://www.kaggle.com/getting-started/149448>
- [15] ImageNet Large Scale Visual Recognition Challenge 2015. URL: <http://image-net.org/challenges/LSVRC/2015/>