ИНФОРМАТИКА

COMPUTER SCIENCE

МРНТИ 28.23.15
УДК 004.93

https://doi.org/10.51889/2022-1.1728-7901.08

DEVELOPMENT OF A MODIFIED VIOLA-JONES ALGORITHM FOR FACE RECOGNITION

Amirgaliyeva Zh.[1], Sadykova A.[2], Kenshimov Ch.[1]*

[1]"Institute of Information and Computing Technologies" Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan, Almaty, Kazakhstan
[2]Al-Farabi Kazakh National University, Almaty, Kazakhstan
**e-mail: sadykova.akmanat@gmail.com*

Abstract

This article discusses a modified version of Viola-Jones' face detection algorithm. The research determined the influence of the orientation of a person's face in space on the operation of the Viola-Jones face recognition algorithm. The authors set the goal of this work to develop a method for finalizing the face recognition algorithm, it allows using a wide range of face rotation angles in the input image. All images in this work are in 3 dimension. In the process of work, the instability of the algorithm to a change in the position of the face is revealed and a method is proposed to increase the stability of the algorithm to this change. A technique for comparative analysis of the performance of face recognition algorithms is described. The article offers the results of a comparative analysis, as well as the test result of the developed algorithm is shown in the form of a table.

By the result, modified algorithm only slightly increased the percentage of image recognition at the boundaries of the interval of acceptable rotation angles for the Viola-Jones algorithm, the introduction of image segmentation based on skin color reduced the number of false alarms of the algorithm and reduced the execution time of the algorithm due to the fact that only part of the images.

Keywords: face recognition, Viola-Jones algorithm, CCTV, ICPR, recognition systems, segmentation.

Аңдатпа

Ж. Амиргалиева[1], А. Садыкова[2], Ч. Кеншимов [1]
[1] Қазақстан Республикасы Білім және ғылым министрлігінің Ғылым комитеті «Ақпараттық және есептеу технологиялары институты», Алматы қ., Қазақстан
әл-Фараби атындағы Қазақ Ұлттық Университеті, Алматы қ., Қазақстан

БЕТТІ ТАНУ ҮШІН МОДИФИКАЦИЯЛАНҒАН ВИОЛА-ДЖОНС АЛГОРИТМІН ДАМЫТУ

Бұл мақалада Виола-Джонстың бетті анықтау алгоритмінің өзгертілген нұсқасы қарастырылып отыр. Зерттеу барысында Виола-Джонстың бетті тану алгоритмінің жұмысына адамның бет-әлпетінің бағдарлануының әсері анықталды деп тұжырымдалады. Авторлар бұл жұмыстың мақсаты ретінде кескінінде тұлғаның айналу бұрыштарының кең ауқымын қолдануды енгізуге мүмкіндік беретін етіп бетті тану алгоритмін өңдеу әдісін әзірлеуді қойған. Бұл жұмыстағы барлық кескін 3D кеңістігінде болып табылады. Жұмыс барысында берілген Виола-Джонс алгоримінің өзгертілген нұсқасының тұлғаның позициясының өзгеруіне тұрақсыздығы анықталады және осы өзгеріске алгоритмнің тұрақтылығын арттыру әдісі сипатталып, ұсынылады. Бетті тану алгоритмдерінің өнімділігін салыстырмалы талдау әдісі сипатталған.

Мақалада салыстырмалы талдау нәтижелері ұсынылған, сонымен қатар әзірленген алгоритмнің тест нәтижесі кесте түрінде көрсетілген.

Түйін сөздер: бетті тану, Виола-Джонс алгоритмі, CCTV, ICPR, тану жүйелері, сегментация.

Аннотация

Ж. Амиргалиева[1], А.Садыкова[2], Ч. Кеншимов[1]
[1]Комитет науки Министерства образования и науки Республики Казахстан «Институт информационных и вычислительных технологий», г.Алматы, Казахстан
[2]Казахский национальный университет имени аль-Фараби, г.Алматы, Казахстан

РАЗРАБОТКА МОДИФИЦИРОВАННОГО АЛГОРИТМА ВИОЛЫ-ДЖОНСА ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ

В статье рассматривается модифицированная версия алгоритма обнаружения лица Виолы-Джонса. В исследовании определялась влияние ориентации лица человека в пространстве на работу алгоритма распознавания лиц Виолы-Джонса. Авторы поставили целью этой работы – разработать метод доработки алгоритма распознавания лиц, чтобы он позволял использовать большой диапазон углов поворота лица на входном изображении. Все изображения в данной работе являются в 3Д пространстве. В процессе работы выявляется неустойчивость алгоритма к изменению положения лица и предлагается методика повышения устойчивости алгоритма к данному изменению. Описана методика сравнительного анализа производительности алгоритмов распознавания лиц.

В статье предлагается результаты сравнительного анализа, так же результат тестирования разработаного алгоритма показан в виде таблицы.

Ключевые слова: распознавание лиц, алгоритм Виолы-Джонса, CCTV, ICPR, системы распознавания, сегментация.

Introduction

Due to the fact that the incidence of terrorism has become more frequent all over the world, to ensure security in transport and in office premises, identification of a person, especially by face, becomes important. The solution of the problem of face recognition includes the stages of image acquisition, preprocessing, face detection and identification taking into account the identified features.

In this paper, we will consider the problem of face detection, which does not include matching a face with a known image from a database [1].

The solution to the problem of face detection is especially important when using video surveillance systems (such as CCTV) and in security complexes. Due to the growth in the computing power of personal computers and mobile devices, face detection is gaining popularity as a way of organizing human-machine interaction. Social networks such as Facebook detect faces in the photos uploaded by the user and offer to associate them with the user's online account. Also, there are applications using "augmented reality", such as video games, where the player can interact with objects in the virtual world through movements and gestures captured by the camera.

Today, the field of application of face detection algorithms is dynamically developing[2]. These algorithms are used in various embedded systems, and the conditions for using these systems cause significant differences in image quality. Real-time processing requirements make it impossible to post-process images or engage an operator for, so it is important to develop image defect-resistant algorithms that are computationally efficient. Thus, the problem of face detection is one of the priority directions in the development of machine learning and computer vision algorithms.

One of the factors that have a significant impact on correct face detection is the orientation of the head in space. We will call an algorithm resistant to rotations at a certain angle of rotation if it correctly detects a face, provided that the head is rotated around a vertical axis by such an angle.

The stability threshold for most of the algorithms most commonly used at the moment is within 10 degrees, which is not enough to detect faces without first positioning them in front of the camera. This makes such algorithms inapplicable for the task of searching for faces in CCTV systems or in a video stream from a mobile phone camera [3].

The aim of this work is to develop and implement an algorithm with improved cornering stability. A prerequisite for determining the direction of research is the analysis of the existing algorithm in order to obtain accurate numerical values of the ranges of the rotation angle, as well as performance (by performance we mean the time required to detect faces in test images. Since absolute performance depends on many factors, such as configuration hardware and software, we will be interested in relative performance - that is, the values obtained when running different algorithms on the same machine in an identical runtime environment).

In this paper, we propose a method for increasing the stability of the algorithm to rotations, which consists in preprocessing the input image in order to compensate for the rotation of the heads. Various approaches to assessing such inverse transformations have been implemented and analyzed. The Viola-Jones algorithm [4] is used as the basic detection algorithm. This algorithm was chosen due to the fact that it is the most frequently used algorithm today. In addition, it is relatively simple to implement, and can serve as a basis for further implementation of multi-stage face detection methods.

Stages of work

To implement the tasks described above, the following steps were performed.

- Search for publicly available databases of images of faces taken at different angles of rotation. The database of images of the INRIA Institute was selected [5]
- Writing an application to automatically prepare images for use at the stage of collecting statistics. Preparation includes the following steps:
- Retrieve image archives from relevant websites
- Bringing data from various databases to a unified catalog format.
- Convert images from various formats to the desired one. This task was solved by means of the Qt library, which is also used to display an image on the screen.
- Implementation of a program for collecting statistics on face detection algorithms. The program automatically executes algorithms using test images as input data and generates a report.

Viola-Jones algorithm implementation

The final stage is the implementation of head-rotation-resistant face detection algorithms. It consists of implementing the two approaches to increasing resilience described earlier. The first stage implements an algorithm that processes the image with the camera in order to compensate for the change in head position. Thus, most face detection algorithms can operate over a wider range of acceptable angles without modifying the original algorithm. The second stage is the development and implementation of an algorithm that uses a set of rotation-invariant features to search for faces.

Modified Viola-Jones algorithm

This chapter proposes the refinement of the Viola-Jones algorithm to increase the stability to face turns around the vertical axis. The modified algorithm consists of three stages.

Segmentation by color. The following conditions are used to highlight areas of the image that contain skin, where R, G, B are colors in the range [0; 255]: $\{to\ R >= 1.3 * G$

$$G <= 9 * B$$
$$G >= 0.8 * B$$

$$R <= 2.7 * G \quad (1)$$

These coefficients were chosen based on the analysis of the results presented in the work devoted to the detection of skin areas in conditions of different illumination [6]. After the selection of areas, adjacent pixels classified as "skin" are combined into clusters. Each cluster is bounded by a rectangular area. A face search is performed inside each rectangular area (the second and third stages of the algorithm) [7]. After that, areas that are less than 5.2 percent of the height of the image width are filtered (by comparing image area with area area). In addition, this formula is designed to detect Caucasians in images taken in daylight. Since the cascades of features used for testing were trained on a similar sample, this step should not lead to a false classification of the area containing the face.

Choosing a classifier. In each of the areas potentially containing a face, which were detected at the previous step of the algorithm, the search for parts of the face is performed using cascades of features trained to search for this part.

- The eyes are searched. Cascades are used to detect the left and right eyes. If one eye is found, a face search is performed using the face profile cascade (side view).
- When two eyes are found, the mouth is searched for in the image. In case a mouth is detected, the points are the centers of the rectangular regions bounding the eyes and mouth. The distances D_{right} and D_{left} from the

center of the right and left eyes to the center of the mouth are calculated. The real distance between the eyes D_{eyes} and the "expected" distance D_{exp} are calculated.

The "expected" distance is the distance that should be between the centers of the eyes in an image of a face looking directly into the camera. It is assumed that the image has been rotated only around the vertical axis, and the horizontal distances between parts of the face were not disturbed.

In this case, D_{exp} = D_l * D_r / 1.254 / 2. This coefficient was selected as an average value by analyzing 10 images from the database. The image is scaled horizontally with a factor $\frac{D_{exp}}{D_{eyes}}$.

If a face was not detected using the "profile" cascade, or a scaled image was obtained after the "eye" classifiers, a virtual image of a face is constructed containing an approximation of the original images before rotation (third stage of the algorithm).

Rotation compensation. At this stage, a "virtual image" is built, which is an approximation of the face image before rotation around the horizontal axis. For this, a face image is taken, which is a two-dimensional projection of the face onto the camera plane. We consider this image as a plane in three-dimensional space, assigning the coordinate $Z = 0$. We will rotate this plane around the vertical axis with a step of 6 degrees in the range from -40 to +40 degrees inclusive.

It is assumed that the combination of rotation and scaling in the previous step will make the image look more like the one we would get if the face is directed towards the camera, and the face detection algorithm should allow for a larger range of face rotation angles.

Analysis of the behavior of algorithms

This chapter describes the influence of affine transformations on face detection by the Viola-Jones algorithm and the modified algorithm. The ICPR base of the INRIA Institute is used as a test base [8].

Rotation around the vertical axis. Images from the ICPR database were used. Images of 13 person were selected from it. Turning the head around the vertical axis in the range [−100; +100] degrees in 13 degree increments.

Table 1. Viola Jones results

Angle, degrees	*Number of recognized*
-100	*1*
-85	*3*
-65	*9*
-55	*12*
-35	*13*
-15	*13*
0	*15*
20	*14*
25	*16*
35	*15*
50	*11*
70	*5*
95	*1*

From the table 1 it is clear that at 65 degrees more than half of the images are recognized, at 70 - less than 25%. Thus, the algorithm works stably at an angle less than 60%. When using the modified algorithm, one additional image was recognized at an angle of 60% and two at an angle of 55%, while at a lower angle of rotation there are no changes. This improvement is not significant, but suggests that the initial assumption about the possibility of an approximation is correct. Further direction - the development of a more accurate method for evaluating image rotation.

Rotate around the Z axis. The test consists of rotating the original image 360 degrees in 20-degree increments. Result: Face is detected when turning up to +/- 35 degrees. The behavior of the original and modified algorithms is similar, since the modified algorithm only compensates for rotations around the vertical axis.

Shear. The Shear transformation along the X and Y axes is specified by the following matrices:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{2}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3}$$

Shear test. The original image underwent Shear transformation with λ from 0.0 to 1.0 with a step of 0.1. The result is that more than 85% of images stop being recognized at λ> = 0.7.

Scaling test. The original image is sequentially enlarged by a factor of 1.3 until each of its dimensions is less than 778 pixels [9]. An algorithm runs on each scaled image.

The test results for the Viola-Jones algorithms and the modified algorithm do not differ from the original ones. Explanation - Viola-Jones algorithm is invariant to change sizes of the input image, since during the execution of the algorithm the classifier window is sequentially scaled. This test is part of the methodology for analyzing the behavior of the face detection algorithm under affine transformations, and, despite the fact that the Viola-Jones algorithm shows the same results [10], launching a test for this algorithm was necessary to check the correctness of the implementation of the test generation algorithm.

Transfer test. The original image is successively enlarged by a factor of 1.3 until its length and height are less than 778 pixels. An algorithm is run on each scaled image.

The test results for the Viola-Jones algorithms and the modified algorithm do not differ from the original ones. The Viola-Jones algorithm is shear invariant, since the classifier window collapses in the horizontal and vertical directions during the execution of the algorithm.

In the course of this work, all expected results were achieved:

- Described a methodology for testing face detection algorithms for resistance to affine transformations.
- Developed software implementation of this technique for automated testing.
- Analysis of the behavior of the ViolaJones face detection algorithm.
- The analysis of the behavior of the modified face detection algorithm has been carried out.

Despite the fact that the modified algorithm only slightly increased the percentage of image recognition at the boundaries of the interval of acceptable rotation angles for the Viola-Jones algorithm, the introduction of image segmentation based on skin color reduced the number of false alarms of the algorithm and reduced the execution time of the algorithm due to the fact that only part of the Images.

Recognition error - false positive, which is not observed in the right picture, where areas with faces were marked by color segmentation before detection. Based on these results, we can conclude that the Viola-Jones algorithm can be effectively combined with preprocessing, and the use of additional classifiers can improve the quality of face detection.

Conclusion

In the course of this work, all expected results were achieved:

- Described a methodology for testing face detection algorithms for resistance to affine transformations
- Developed software implementation of this technique for automated testing
- The analysis of the behavior of the ViolaJones face detection algorithm has been carried out.
- The analysis of the behavior of the modified face detection algorithm has been carried out.

Despite the fact that the modified algorithm only slightly increased the percentage of image recognition at the boundaries of the interval of acceptable rotation angles for the Viola-Jones algorithm, the introduction of image segmentation based on skin color reduced the number of false alarms of the algorithm and reduced the execution time of the algorithm due to the fact that only part of the images.

Acknowledgment

This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. GF AP08857573).

References:

1 K. Mallat and J. Dugelay, "A benchmark database of visible and thermal paired face images across multiple variations," 2018 International Conference of the Biometrics Special Interest Group (BIOSIG), 2018, pp. 1-5, doi: 10.23919/BIOSIG.2018.8553431.

2 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.

3 S. Sandeepa, A. Reyaz and M. Silpa, "An Efficient Approach to Recover CCTV Video from Proprietary DVR File System," 2018 International CET Conference on Control, Communication, and Computing (IC4), 2018, pp. 250-254, doi: 10.1109/CETIC4.2018.8531073.

4 W. LU and M. YANG, "Face Detection Based on Viola-Jones Algorithm Applying Composite Features," 2019 International Conference on Robots & Intelligent System (ICRIS), 2019, pp. 82-85, doi: 10.1109/ICRIS.2019.00029.

5 S. Silva, B. A. Gutman, E. Romero, P. M. Thompson, A. Altmann and M. Lorenzi, "Federated Learning in Distributed Medical Databases: Meta-Analysis of Large-Scale Subcortical Brain Data," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 270-274, doi: 10.1109/ISBI.2019.8759317.

6 S. R. Gudadhe, "Selection & Detection of Skin and Skin Color Background Under Complex Background," 2018 International Conference on Research in Intelligent and Computing in Engineering (RICE), 2018, pp. 1-3, doi: 10.1109/RICE.2018.8509033.

7 Y. He, Y. Ye, P. Hanhart and X. Xiu, "Geometry Padding for Motion Compensated Prediction in 360 Video Coding," 2017 Data Compression Conference (DCC), 2017, pp. 443-443, doi: 10.1109/DCC.2017.18.

8 X. Xu, L. Zhang and F. Li, "MSSVT: Multi-scale feature extraction for single face recognition," 2018 24th International Conference on Pattern Recognition (ICPR), 2018, pp. 1996-2001, doi: 10.1109/ICPR.2018.8545343.

9 C. Jiang, M. Wang, X. Tang and R. Mao, "Face recognition method based on sparse representation and feature fusion," 2019 Chinese Automation Congress (CAC), 2019, pp. 396-400, doi: 10.1109/CAC48633.2019.8997456.

10 K. Vikram and S. Padmavathi, "Facial parts detection using Viola Jones algorithm," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), 2017, pp. 1-4, doi: 10.1109/ICACCS.2017.8014636.