Algorithm for optimization of Viola–Jones object detection framework parameters

A D Egorov
National Research Nuclear University “MEPhI” (Moscow Engineering Physics Institute), Moscow
degorovalexeyd@gmail.com

Abstract. The Viola-Jones algorithm is one of the most popular algorithms for objects recognition in an image. This research paper deals with the possibilities of parametric optimization of the Viola-Jones algorithm to achieve maximum efficiency of the algorithm in specific environmental conditions. It is shown that with the use of additional modifications it is possible to increase the speed of the algorithm in a particular image by 2-5 times with the loss of accuracy and completeness of work by not more than 3-5%.

1. Introduction
Image recognition when analyzing a video stream or individual images is called computer vision. Computer vision is used in many modern technological solutions to ensure safety of objects and road traffic, personality analysis and recognition, building augmented reality. The mechanism for recognition a specific object in an image and video stream, for example, face, is an important element in the work of computer vision algorithms.

Recognition of a person's face in a complex task, including high optical interference and noise, is one of the main tasks to be solved when implementing remote technologies for monitoring their current functional and psychoemotional state [1, 2]. These technologies make it possible to carry out constant monitoring of the adequacy of behavior, first of all, of operational personnel managing hazardous objects to reduce the risk of technogenic accidents and catastrophes caused by the so-called human factor [3]. The developed algorithms and software for recognizing a person's face are subject to increased reliability and speed requirements, due to the need to process large video streams of data in real time [4]. The quality of the solution of this problem determines among other things the reliability of determining human bioparameters, related, for example, to the movement of the eyes, the level of microshivering of the facial areas, their temperature, and the reaction time [5].

2. About the Viola-Jones algorithm
One of the most common methods for object detection in an image is the Viola-Jones object detection framework. The Viola-Jones object detection framework [6, 7] was proposed by P. Viola and M. Jones in 2001. This method is based on the representation of the image in an integral form and the construction of a cascade of "weak" classifiers, which are called Haar-like features [8].

The Viola-Jones object detection framework consists in practice of the following steps:
1) Training a cascade classifier using an AdaBoost algorithm.
2) Preprocessing – constructing an integral image to optimize the work of the method.
3) Searching for an object in the image using Haar-like features.

The following parameters exist in the practical implementation of the algorithm:

1) Used cascade for objects search (trained before using the algorithm) – cas.
2) The change in the window size in the transition from one iteration of running the algorithm to another – sf. Min value of this parameter is 1.05. Maximum value is carried on maximum image size. The step of the parameter change is 0.05.
3) The number of objects recognized next to each other, on the basis of which it can be concluded that there is a sought object in this area – minn. Min value of this parameter is 1. Maximum is 10. The step of the parameter change is 1.
4) The initial size of the search box – mins. Min value of this parameter is carried on cascade min image size. In our case it is 20x20 pixels. Maximum value is carried on maximum image size. The step of the parameter change is 10.

Most of the papers in the study of the algorithm are devoted to improving the quality of cascade learning or the accuracy and speed of the comparison algorithm of the intended object. This research paper deals with the parameters of a particular implementation of the algorithm, as well as methods for optimizing them in order to obtain the highest accuracy with the smallest average operating time.

Empirical comparison of various parameters of the method is given in article [9].

Optimization in the algorithm is based on the method of enumeration of the above parameters.

3. Mechanism of algorithm parameters selection

An image database (test set) with the desired objects under the conditions in question is necessary to select the parameters of the Viola-Jones algorithm for specific conditions. The set of intended objects (if they are present in the image) with the coordinates in this image should be put in line to each image in this database.

The following parameters calculated on the application of the method with a specific set of parameters to the database are to be introduced to evaluate the effectiveness of the method:

1) Accuracy – P – the ratio of the number of correctly found objects to the total number of found objects.
2) Completeness – R – the ratio of the number of correctly found objects to the total number of objects in the test set.
3) Based on the accuracy and completeness, it is proposed to use the aggregate parameter – F-measure (1).

\[
F = \frac{2PR}{P + R}
\]  

(1)

4) It is proposed to introduce criteria (2,3,4) to work time tracking (T – the average time the method works on the test set)

\[
Q_1 = a \cdot (1 - P) + bT
\]  

(2)

\[
Q_2 = a \cdot (1 - R) + bT
\]  

(3)

\[
Q_3 = a \cdot (1 - F) + bT
\]  

(4)

5) It is proposed to use the parameter values for the considered criteria \( a = 1, b = 1/\max(T) \).

4. Analysis of the results

The analysis of the results is carried out on the basis of 808 images with 1202 objects marked. Each image size is 400 x 400 pixels. Accuracy

The criterion \( Q_1 \) uses accuracy of work for calculation. Table 1 shows top-8 results obtained.

4.1. Completeness
The criterion $Q_1$ uses completeness of work for calculation. Table 2 shows top-8 results obtained.

| cas         | sf | minn | mins | true | false | time | P   | R   | $Q_1$ |
|-------------|----|------|------|------|-------|------|-----|-----|-------|
| lbp_improved| 1.4| 5    | 70   | 417  | 0     | 785  | 2   | 1.000| 0.347 | 0.012 |
| lbp_improved| 1.4| 3    | 70   | 616  | 1     | 586  | 2   | 0.998| 0.512 | 0.013 |
| lbp         | 1.3| 5    | 70   | 575  | 0     | 627  | 3   | 1.000| 0.478 | 0.018 |
| lbp         | 1.3| 3    | 70   | 709  | 1     | 493  | 3   | 0.999| 0.590 | 0.019 |
| lbp         | 1.3| 60   | 819  | 1    | 383   | 5    | 0.999| 0.681| 0.030 |
| lbp         | 1.3| 60   | 681  | 0    | 521   | 6    | 1.000| 0.567| 0.035 |
| lbp         | 1.2| 5    | 70   | 793  | 4     | 409  | 6   | 0.995| 0.660 | 0.040 |
| lbp_improved| 1.4| 3    | 40   | 714  | 3     | 488  | 7   | 0.996| 0.594 | 0.045 |

Table 2. Selection of parameters for the criterion $Q_2$

| cas           | sf | minn | mins | true | false | false | time | P   | R   | $Q_2$ |
|---------------|----|------|------|------|-------|-------|------|-----|-----|-------|
| lbp           | 1.1| 1    | 70   | 1061 | 196   | 141   | 9    | 0.844| 0.883| 0.175 |
| alt_improved  | 1.4| 1    | 60   | 1066 | 30    | 136   | 10   | 0.973| 0.887| 0.177 |
| lbp           | 1.1| 1    | 60   | 1096 | 296   | 106   | 15   | 0.787| 0.912| 0.184 |
| alt           | 1.2| 1    | 70   | 1116 | 52    | 86    | 18   | 0.955| 0.928| 0.187 |
| lbp           | 1.2| 1    | 70   | 1024 | 107   | 178   | 6    | 0.905| 0.852| 0.187 |
| alt2          | 1.2| 1    | 70   | 1125 | 82    | 77    | 21   | 0.932| 0.936| 0.199 |
| alt           | 1.2| 1    | 60   | 1142 | 71    | 60    | 26   | 0.941| 0.950| 0.217 |
| alt_improved  | 1.4| 1    | 70   | 1010 | 10    | 192   | 8    | 0.990| 0.840| 0.211 |

4.2. F-measure

The criterion $Q_3$ uses F-measure of work for calculation. Table 3 shows top-8 results obtained.

| cas  | sf | minn | mins | true | false | false | time | P   | R   | $Q_2$ | $Q_3$ | F     |
|------|----|------|------|------|-------|-------|------|-----|-----|-------|-------|-------|
| alt  | 1.1| 3    | 40   | 1165 | 57    | 37    | 97   | 0.953| 0.969| 0.640 | 0.961 | 0.102 |
| alt  | 1.1| 3    | 40   | 1165 | 58    | 37    | 91   | 0.953| 0.969| 0.606 | 0.961 | 0.102 |
| alt  | 1.1| 3    | 30   | 1166 | 73    | 36    | 156  | 0.941| 0.970| 0.971 | 0.955 | 0.106 |
| alt  | 1.1| 3    | 30   | 1166 | 74    | 36    | 151  | 0.940| 0.970| 0.943 | 0.955 | 0.106 |
| alt2 | 1.1| 3    | 40   | 1166 | 87    | 36    | 109  | 0.931| 0.970| 0.706 | 0.950 | 0.112 |
| alt2 | 1.1| 3    | 40   | 1166 | 88    | 36    | 104  | 0.930| 0.970| 0.677 | 0.950 | 0.112 |
| alt2 | 1.1| 3    | 30   | 1168 | 111   | 34    | 177  | 0.913| 0.972| 1.085 | 0.942 | 0.116 |
| alt2 | 1.1| 3    | 30   | 1168 | 112   | 34    | 171  | 0.913| 0.972| 1.051 | 0.941 | 0.117 |

5. Conclusions and recommendations

On the basis of the results, inference should be drawn that it is not expedient to use criterion of selection of parameters which is estimated only on accuracy of algorithm work, as the completeness does not exceed 31.8% for 8 best results.

It is necessary to increase the weight of completeness when calculating the criterion. For example, we take the coefficient $a$ for the criterion $Q_2$ equal to three and consider the results (table 4).

It is seen that the highest completeness is 0.963 with the accuracy of 0.894, taking into account the error of the criterion being calculated. This accuracy is less by 5.9% and the completeness is less than 0.6% than the optimal result, calculated using the criterion $Q_3$. However, the time the method works
when taking parameters from $Q_2$ differs in 2.2 times from the time the method works when the parameters are taken from $Q_3$. The greatest gain in time – 5.38 times – is achieved with the difference of the accuracy of 0.2% and the completeness of 4.1%.

Table 4. Selection of parameters for the criterion $Q_2$ with coefficient 3.

| cas | sf | minn | mins | true | false | false | time | P   | R   | $Q_2$ |
|-----|----|------|------|------|-------|-------|------|-----|-----|-------|
| alt | 1.2| 1    | 60   | 1142 | 71    | 60    | 26   | 0.941| 0.950| 0.297 |
| alt2| 1.2| 1    | 70   | 1125 | 82    | 77    | 21   | 0.932| 0.936| 0.311 |
| alt | 1.2| 1    | 70   | 1116 | 52    | 86    | 18   | 0.955| 0.928| 0.316 |
| alt2| 1.2| 1    | 60   | 1144 | 124   | 58    | 31   | 0.902| 0.952| 0.320 |
| alt | 1.1| 1    | 70   | 1131 | 95    | 71    | 26   | 0.923| 0.941| 0.324 |
| alt | 1.2| 1    | 50   | 1156 | 95    | 46    | 39   | 0.924| 0.962| 0.335 |
| alt | 1.1| 1    | 60   | 1158 | 138   | 44    | 40   | 0.894| 0.963| 0.336 |
| alt2| 1.2| 1    | 60   | 1144 | 123   | 58    | 36   | 0.903| 0.952| 0.348 |

It is recommended to introduce an additional check for the presence of objects on the found object to increase the accuracy of the work. For example, for a face it is the search for the eyes, mouth or nose on the supposed face found.

Thus, the algorithm for selecting parameters consists of the following steps:
1) Building a database of objects for the current environmental conditions.
2) Selection of parameters based on the criterion $Q_2$ or $Q_3$ depending on the business requirements of the system.
3) Increase the accuracy of work due to the modification of the additional search used.

Laboratory tests of developed software tools for optimizing human face recognition algorithms have confirmed their high efficiency.

The application of the developed algorithm for selecting parameters for face recognition in software and hardware complexes for remote monitoring of the current functional and psychoemotional state of a person will ensure continuous reliable registration of its bioparameters using optical technologies.

Acknowledgement

The research is carried out within the framework of the theme 00-2-026-4209 "Approval of the technology for reducing the risk of occurrence and mitigation of the consequences of man-induced disasters by minimizing the influence of the human factor on the reliability and trouble-free operation of NPPs and other hazardous objects" in accordance with the competitive part of the MEPhI government contract for 2017-2019.

References

[1] Aliushin M V and Kolobashkina L V 2014 Psychology questions Monitoring of bioparameters of the person on the basis of remote technologies №6 pp. 135–144.
[2] Aliushin M V, Aliushin A V, Andriushina L O, Kolobashkina L V and Pshenin V V 2013 Global nuclear safety Remote and not contact technologies of registration of bioparameters of operation personnel as control facility human factor and increases in the NPP safety., №3(8) pp. 69–77
[3] Alyushin M V, Alyushin V M, Kolobashkina L V 2016 Psychology issues Methodological aspects of the automated forecasting of emergency situations of technogenic origin №2, pp. 83–90
[4] Aliushin M V, Aliushin AV, Belopolskii V M, Kolobashkina L V and Ushakov V D 2013 Global nuclear safety Optical technologies for systems of monitoring of the current functional state of quick structure of management of nuclear power objects №2(7), pp. 69–
[5] Alyushin M V and Alyushin V M 2015 Psychology issues *Technique of measurement of time of operator reaction*, №5, pp. 157–165

[6] Viola P. and Jones M. 2001 proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001) *Rapid Object Detection using a Boosted Cascade of Simple Features*

[7] Viola P. and Jones M. 2004 International Journal of Computer Vision *Robust real-time face detection*, vol. 57, no. 2., pp.137–154

[8] Whitehill J. and Omlin C.W. 2006 Automatic Face and Gesture Recognition *Haar features for FACS AU recognition* pp. 101-106.

[9] Egorov A D, Shtanko A N and Minin P E 2015 Bulletin of the Lebedev Physics Institute *Selection of Viola-Jones Algorithm Parameters for Specific Conditions* 42 No 8 pp 244-248.