EFFECT OF EYES DETECTION AND POSITION ESTIMATION METHODS ON THE ACCURACY OF COMPARATIVE TESTING OF FACE DETECTION ALGORITHMS[1]

N. Degtyarev, O. Seredin

Tula State University, 92 Lenin Ave., Tula 300600, Russian Federation; Phone: +7(4872)353637; E-mail: n.a.degtyarev@gmail.com

Many published comparisons of face detection algorithms used different evaluation procedures for each algorithm or even contain only a summary of the originally reported performance among several face detection algorithms on the pair of small datasets. Degtyarev et al. have proposed the FD algorithm evaluation procedure containing model of face representation conversion unifying the FD algorithms comparison procedures, which makes such evaluation more reliable. However, there is no evidence that such "conversion" does not diminish the localization accuracy. The aim of this work is to examined the effects of two different face representation conversion techniques - eyes estimation model proposed by Degtyarev et al. and highly scored eyes detection method proposed by Bolme et al. and based on ASE filters - via routine testing.

Introduction

Face detection (FD) algorithms are getting widely used in the modern world: security systems, interactive user interfaces, advertisement industry, entertainment services, video coding, etc. However, there is no universal and comprehensive algorithm that can be used for all possible tasks and under any circumstances. Therefore users should carefully choose FD algorithm based on published comparisons of FD algorithms such as made by Hjelmas et al. [4], by Yang et al. [9], by Beveridge et al. [1] and Degtyarev et al. [3]. It should be emphasized that the most of such researches used different evaluation procedures for each algorithm or even contain only a summary of the originally reported performance among several face detection algorithms on the pair of small datasets, nevertheless that such problem has been emphasized by Hjelmas et al. [4] in 2001.

Degtyarev et al. have aimed in [3] to propose parameters of FD algorithms quality evaluation and methodology of their objective comparison, to show the current state of the art in face detection (Short overview of this work and proposed model of face representation conversion is given below) unifying the FD algorithms comparison procedures.

The aim of this work is to examine the effects of two different face representation conversion techniques - *eyes estimation* model used in [3] and well known and publicly available *eyes detection* method, proposed by Bolme et al. [2] (ASE filters), from the highly scored by Scheirer et al. [8] correlation filters class. Such study can help end users of FD algorithms to understand and to interpret the results of its comparative testing better.

Comparative Testing of Face Detection Algorithms

The main problem of fair comparative testing of face detection algorithms is the abundance of different models of face representation in images: by the center of the face and its radius, by rectangle (OCV, FDLib, FoI), by coordinates of the centers of eyes (SIF, UniS, FSDK, VL), by ellipse, etc.

In recent FD algorithms' comparative testing made by Degtyarev et al. [3] authors supposed faces to be represented by coordinates of the centers of the eyes (i.e. centers of the pupils), because first, this representation looks to be more opportune in terms of the results compar-

[1] This work is supported by Russian Foundation for Basic Research (Grant 09-07-00394)

ison; second, usually face recognition algorithms require the centers of eyes matching for learning samples; third, experts mark eyes faster, easier and more precisely than they mark faces by rectangles. Thus, to unify the resulting comparison method authors suggested eyes position reconstruction model, which receives a face location in rectangle form and returns estimated coordinates of the centers of eyes.

In their work, Degtyarev et al. have tested following algorithms: Intel© OpenCV (OCV), Luxand© FaceSDK (FSDK), Face Detection Library (FDLib), SIFinder (SIF), University of Surrey (UniS), FaceOnIt(FoI), Neurotechnology© VeriLook (VL) - where OCV, FDLib, FoI represent results by rectangle and SIF, UniS, FSDK, VL by coordinates of the centers of eyes. And the key result of their work can be represented as a chart of ROC curves (see Fig. 1)

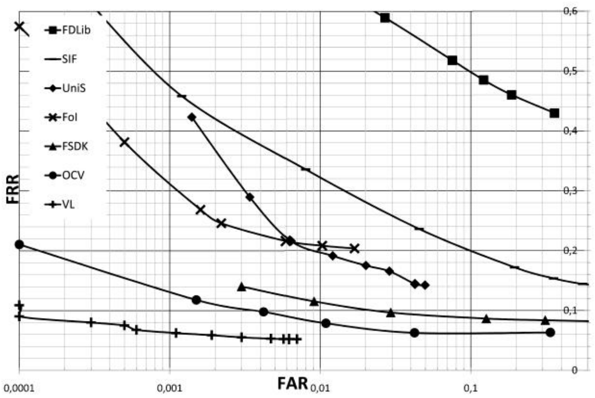

Fig. 1. The key result of comparative testing of Face Detection algorithms; ROC curves.

Model of Localization Accuracy

If detected faces are represented by the centers of the eyes (Fig.2.a), let's consider them to be correctly detected, if and only if detected eyes belong the area around the true eyes location with the diameter D_{Eyes}. Which depends on the distance between eyes' centers and, has been taken equal to 0.25 .(This criterion was originally used by Jesorsky et al. [5]), and calculates as $D_{Eyes} = 2\alpha \times l_{Eyes}$.

Otherwise, faces represented by rectangles should be "converted" to unified representation via eyes estimation or eyes detection. It should be noted that such "conversion" of face representation (from rectangle to eyes' centers) might deteriorate the localization accuracy for algorithms describing faces by rectangles. In this work we examined the effects of two different face representation conversion techniques - eyes estimation model proposed by Degtyarev et al. [3] and well known and highly scored eyes detection method proposed by Bolme et al. [2] and based on ASE filters.

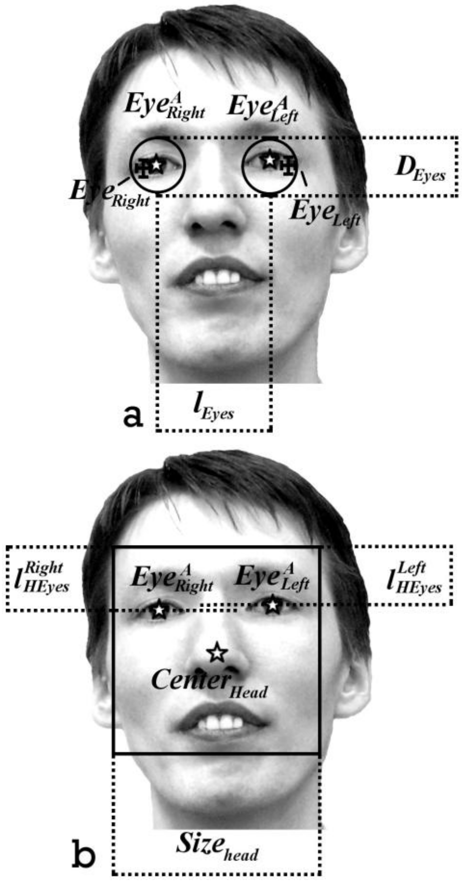

Fig. 2. Schematic face representation. Eye_{Left} and Eye_{Right} – absolute coordinates of detected left and right eye respectively; l_{Eyes} - distance between eyes' centers; l^{Left}_{HEyes}, l^{Right}_{HEyes}, l_{HEyes} - distance between top border of the face and center of the left or right eye; $Size_{Head}$ - size of the rectangle representing face; D_{Eyes} - diameter of the area of acceptable eyes' coordinates deviation from the true eyes location; $Center_{Head}$ - absolute coordinates of the found face.

Model of Eyes Position Estimation

Assume there is a full face portrait image with no incline (Fig. 2.b), and the algorithm has found its center and size - ($Center_{Head}$ and $Size_{Head}$ respectively). Obviously, the eyes on this image are located symmetrically about the vertical axis (i.e., at the half the distance between them: $l_{Eyes}/2$) and at the same distance

[1] This work is supported by Russian Foundation for Basic Research (Grant 09-07-00394)

(l_{HEyes}) from the top border of the face's rectangle. Thus the absolute coordinates of eyes can be estimated as:

$$Eye^{y} = Center^{y}_{Head} + l_{HEyes} - \tfrac{1}{2} Size_{Head},$$
$$Eye^{x}_{Right} = Center^{x}_{Head} - \tfrac{1}{2} l_{Eyes}, \qquad (1)$$
$$Eye^{x}_{Left} = Center^{x}_{Head} + \tfrac{1}{2} l_{Eyes}.$$

Let's try to estimate the parameters of the algorithm, namely l_{Eyes} and l_{HEyes}, as an average of the huge amount of images with experts' labeled eyes. Based on such analysis, the following coefficients have been founded: A - average proportion of distance between top border of the face and center of the eyes (l_{HEyes}) to the size of the face rectangle; and B - average proportion of the distance between eyes (l_{Eyes}) to the size of the face rectangle ($Size_{Head}$). They can be estimated using information about true eyes location on the images series:

$$A = \frac{1}{N}\sum_{i=1}^{N}\frac{l^{i}_{HEyes}}{Size^{i}_{Head}}, B = \frac{1}{N}\sum_{i=1}^{N}\frac{l^{i}_{Eyes}}{Size^{i}_{Head}},$$

where l^{i}_{HEyes}, l^{i}_{Eyes} and $Size^{i}_{Head}$ - respective parameters measured for the i-th image in the data set of N objects. Therefore the coordinates of the eyes for a given face size and the coefficient of proportions for the algorithms (1) are calculated according next equations:

$$Eye^{y} = Center^{y}_{Head} + Size_{Head}(A - \tfrac{1}{2}),$$
$$Eye^{x}_{Right} = Center^{x}_{Head} - Size_{Head}(\tfrac{1}{2}B), \qquad (2)$$
$$Eye^{x}_{Left} = Center^{x}_{Head} + Size_{Head}(\tfrac{1}{2}B).$$

If there is a full face portrait image with any incline, let's find l_{HEyes} as an average distance between center of each eye and top border of the face.

Experimental Testing: Eyes Position Estimation Model vs. Eyes Detection

The effects of the mentioned above face representation conversion techniques on the localization accuracy of FD algorithms can be examined via routine repeating of testing procedure used by Degtyarev et al. [3] for FD algorithms *simultaneously* describing founded faced by rectangles and by the centers of eyes (FSDK and VL according [3]) and eyes position reconstruction models based on them (Total test dataset size is 59888 images: 11677 faces and 48211 non-faces). Coefficients of eyes position estimation model were determined on *Georgia Tech Face Database*. In this work, the following criteria of face representation conversion method have been taken into account: variability of the model's coefficients under different conditions, detected eyes departures from true eyes position (according Jesorsky et al. [5] and Popovici et al. [6]) and ROC-curves distinctions.

Coefficients of the model for algorithms representing faces by rectangles (2) under different tuning parameter are given in the table below (FSDK and VL).

Table. Coefficients of the model of eyes position estimation.

FSDK	Coeff. A	Coeff. B
Average	0.4285	0.3650
RMSE	0.00035	0.00011

VL	Coeff. A	Coeff. B
Average	0.2506	0.4955
RMSE	0.00027	0.00051

As we can see the coefficients of the eyes position estimation model for each algorithm are different and have small spread around the average (RMSE) for different turning parameters of an algorithm that are expected because turning parameters may interfere with the size of detected faces. Therefore, proposed model is resistant to FD-algorithms' tuning parameters changes, but should be train for each face detectors individually.

The main criterion of localization accuracy according Jesorsky at el. [5] model is detected eyes departures from true eyes position. The comparisons of such departures for direct detected and estimated eyes demonstrate the almost equal localization accuracy at the given value of $\alpha = 0.25$ (Fig. 3). ASE Filter has demonstrated a bit lower results than other for α greater 0.25

To verify the closeness of eyes estimation and direct detection models in terms of localization accuracy, let's compare distinctions of ROC-curves for the algorithms with direct eyes centers detection and with eyes center estimation according proposed model.

[1] This work is supported by Russian Foundation for Basic Research (Grant 09-07-00394)

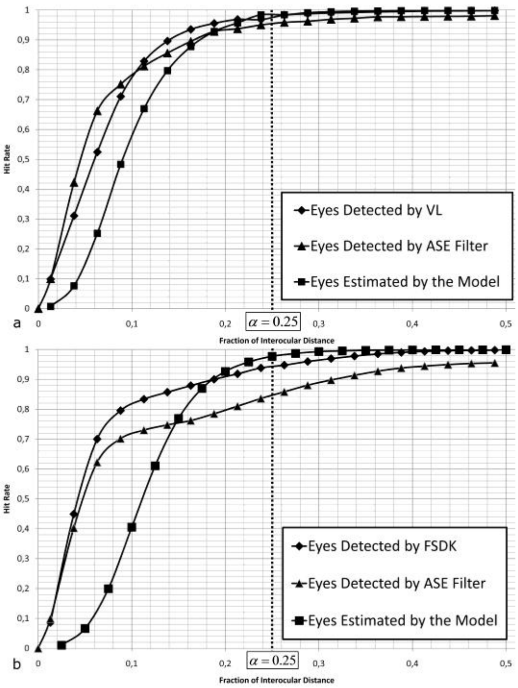

Fig. 3. Detected eyes departures from true eyes position: a- for VL rect. input; b – for FSDK rect. input.

As it shown in Fig. 4, the distinctions of ROC-curves are insufficient and can be ignored for FD algorithms comparison testing.

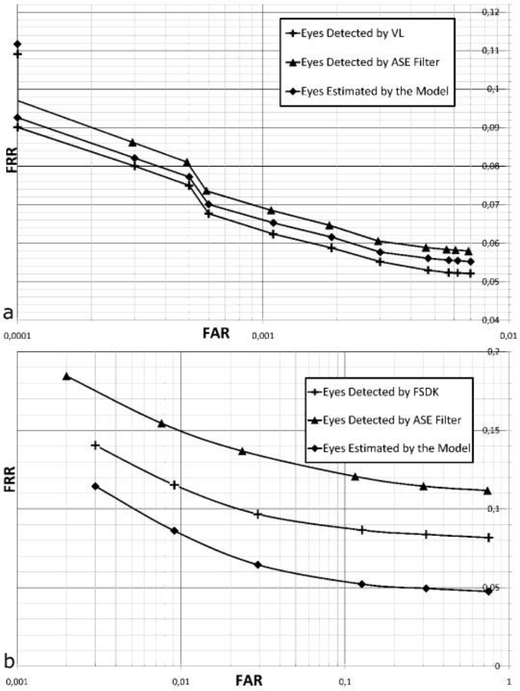

Fig. 4. The ROC plots for different eyes position reconstruction models. FAR (in log scale) against FRR. a - for VL; b - for FSDK.

Discussion and Conclusion

The main problem of fair comparative testing of face detection algorithms is many different models of face representation in images. The only one way to overcome it is unification of face representation via its "conversion". However any "conversion" of face representation might deteriorate the localization accuracy. In this work we examined the influence of two different face representation conversion techniques from rectangles to eyes' centers: eyes estimation model proposed by Degtyarev and well known and highly scored eyes detection method proposed by Bolme et al. [2] and based on ASE filters.

Degtyarev's model has been demonstrated to be fast, easy to implement and sufficiently precise for comparative testing of FD algorithms, despite its simplicity. In our experiments, Bolme's detection model (ASEF) has demonstrated worse performance than direct eyes detection and Degtyarev's model in terms of comparative testing of FD algorithms for $\alpha > 0.2$, but it has been more precise than other for $\alpha < 0.1$ (Fig.3).

However, authors thought that more sophisticated and comprehensive regression eyes' estimation model of higher orders will minimize introduced localization accuracy errors.

References

1. J.R. Beveridge, A. Alvarez, J. Saraf, W. Fisher, P.J. Flynn, J. Gentile. Face Detection Algorithm and Feature Performance on FRGC 2.0 Imagery. // First IEEE International Conference on Biometrics: Theory, Applications and Systems. - 2007.
2. D.S. Bolme, B.A. Draper, J.R. Beveridge. Average of Synthetic Exact Filters. // IEEE CVPR 2009 - 2009. – P. 2105-2112.
3. N. Degtyarev, O. Seredin. Comparative Testing of Face Detection Algorithms: // In A. Elmoataz et al. (Eds.): ICISP 2010, LNCS 6134, Springer, Heidelberg. -2010- P. 200-209.
4. E. Hjelmas, B.K. Low. Face Detection: A Survey. // Computer Vision and Image Understanding, Vol. 83, Issue 3. - 2001.- P. 236-274.
5. O. Jesorsky, K.J. Kirchberg, R.W. Frischholz. Robust face detection using the hausdorff distance. // LNCS, June – 2001 – P. 90–95.
6. V. Popovici, J.P. Thiran, Y. Rodriguez, S.Marcel. On performance evaluation of face detection and localization algorithms. // Proc. of the 17th Intern. Conf. on Pattern Recognition.-2004.- P. 313-317.
7. T. Sauquet, Y. Rodriguez, S. Marcel. Multiview face detection. // IDIAP-Research Report.-2005.
8. W.J. Scheirer, A. Rocha, B. Heflin, T.E. Boult. Difficult detection: A comparison of two different approaches to eye detection for unconstrained environments // Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems. – 2009 -P.176-183.
9. M.H. Yang, D.J. Kriegman, N. Ahuja. Detecting faces in images: A survey. // IEEE Transactions on Pattern analysis and Machine intel., Vol. 24, Issue 1. - 2002.- P. 34–58.

[1] This work is supported by Russian Foundation for Basic Research (Grant 09-07-00394)