A. Uhl and P. Wild. Combining Face with Face-Part Detectors under Gaussian Assumption. In A. Campilho and M. Kamel, editors, *Proceedings of the 9th International Conference on Image Analysis and Recognition (ICIAR'12)*, volume 7325 of *LNCS*, pages 80–89, Aveiro, Portugal, June 25–27, 2012. © Springer. doi: 10.1007/978-3-642-31298-4_10. The original publication is available at www.springerlink.com.

Combining Face with Face-part Detectors under Gaussian Assumption*

Andreas Uhl and Peter Wild

Multimedia Signal Processing and Security Lab
Department of Computer Sciences, University of Salzburg, Austria
{uhl,pwild}@cosy.sbg.ac.at

Abstract. This paper addresses a simple and effective approach of face and face-part classifier fusion under Gaussian assumption, which is able to process heterogeneous visible wavelength (VW) and near infrared (NIR) image data. Evaluations using existing and publicly available AdaBoost-based individual classifiers on the recently released CASIA-V4 iris distance database of close-up portrait images as well as on YaleB indicate, that (1) single classifiers are largely affected by the type of training data, especially for NIR and VW data, and therefore prone to errors, (2) by combining individual classifiers a more robust classifier is obtained, (3) processing time overhead is negligible, if individual classifiers exhibit a low false positive rate, and (4) the proposed fusion approach is not only able to reduce false positives, but also false negative detections.

Keywords: Face detection, eye localization, biometric fusion

1 Introduction

Biometric systems without active participation of users by means of which people can be identified in surveillance scenarios, are an active research topic. A new generation of portal-based iris-and-face recognition devices [12], [21] operating on full portrait images is about to replace traditional *stop and stare* high-cost iris recognition cameras with low throughput and narrow depth of field.

Fusion of face and iris biometric modalities for personal identification is initially proposed in [20] for Eigenfaces and dyadic wavelet-transform based iris recognition with reported perfect separation on good-quality randomly paired datasets (90 subjects). Fisher linear discriminant analysis is employed in [7] on good-quality face and iris matching scores gained from also randomly paired samples (40 subjects) and is shown to improve accuracy. The first single-sensor face and iris multibiometric system based on Eigenfaces and IrisCode-based iris recognition is proposed in [24] using a high-resolution dataset (76 subjects) acquired by a 10 megapixel 850nm NIR camera at 60-80cm distance. Score-combination of left-eye, right-eye and face yielded 0.25% False Rejection Rate (FRR) at 0.1% False Acceptance Rate (FAR), significantly better than face (2.65% FRR) and iris (4.58% FRR) as single features.

* Supported by the Austrian FIT-IT Trust in IT-Systems, project no. 819382.

All such systems depend on efficient and robust detection of eyes in face images before iris segmentation is able to extract the highly unique texture. Coarse segmentation errors in early steps make subsequent finer normalization impossible. While various different approaches for detecting faces, as well as face-parts, for example eye, nose, mouth or eyepairs, exist (see [23] for a survey), the high variability in recording conditions of surveillance-type imagery suggests to combine successful approaches to achieve even higher recognition accuracy, e.g. [6], [16], [4], [22], [14]. While face recognition typically operates on very low resolution VR input [18], it is iris recognition with its demands for high resolution focused NIR images of human eyes which challenges current state-of-the art systems, often solved by employing separate cameras for face and eye extraction [21]. However, recent challenges, like MBGC[1] and NICE[2] emphasize both face recognition in NIR as well as iris recognition under VW to merge both technologies. While we have presented robust iris segmentation techniques able to operate under heterogeneous conditions in [17], this work aims at combining object detection approaches to solve the heterogeneous face detection and eye localization problem before segmentation.

The remainder of this paper is organized as follows: Section 2 reviews related work regarding face and face-part detection. Subsequently, the proposed fusion framework is described in Sect. 3. Experiments are outlined in Sect. 4 using two different open face databases and comparing results with basic AdaBoost-based reference systems. Finally, Sect. 5 forms the conclusion of this work.

2 Related work

The detection of people from facial images combining iris and face as biometric modalities requires the tasks of *face detection* and *eye localization* to be solved. A face (object) detector is a function $f : I \mapsto \{b_1, \ldots, b_{k(I)}\}$ assigning each image I the set of $k(I)$ regions $b_i = (x_i, y_i, w_i, h_i)$ with origin x_i, y_i and size $w_i \times h_i$ containing a face (object) within the image. The task of eye (object-part) localization is usually much simpler, i.e. a function $e : (b, I) \to (b_1, \ldots, b_l)$ assigning an image I and region of interest (face) b the l (fixed, e.g. $l = 2$ for left and right eye) locations of instances of eyes (object-parts) within b in I [10]. A lot of highly accurate specialized object detectors exist [23]: (1) knowledge-based, (2) feature invariant, (3) template matching and, most notably, (4) apprarance-based methods, learning face models from training images widely adapting Viola-Jones' [18] approach and subject of this work.

In applications not only the face, but multiple face-part objects with spatial relationships need to be detected in order to (1) increase recognition accuracy by biometric fusion, e.g. [24], or (2) benefit of more accurate face alignment, e.g. [10], [19]. Depending on the amount of information available, detector fusion approaches may be grouped into feature-level fusion (having access to individual features exploited by function f), score-level fusion (in architectures, where f

[1] Multiple Biometric Grand Challenge, `http://face.nist.gov/mbgc/`
[2] Noisy Iris Challenge Evaluation Part I, `http://nice1.di.ubi.pt/`

evaluates candidate boxes b by means of score-based metrics) and decision-level fusion (where the only information available is $f(I)$, i.e. the result).

Belaroussi *et al.* [2], [3] is an example of the first feature-level fusion type, their proposed face detection scheme is a weighted sum of different location maps based on an apprarance-based associative multi-layer perceptron (Diabolo) map, an ellipse general Hough transform map, and a skin map, with a product-combination of eye detectors. Jin *et al.* [9] present a hybrid eye detection approach integrating characteristics of single-eye and eyepair-detectors. In a first stage eye candidates are determined by a cascade projecting normalized eye images onto a weighted eigenspace and further filtered by eyepair templates. They also employ Gaussians to model the probability distribution of left-eye classes within the eigenspace.

Most fusion methods are score-based: Ma *et al.* [11] examine a novel approach for finding eyes in faces using a probabilistic framework, i.e. after face detection an appearance-based eye detector is applied to derive detection scores for varous locations, which are then fused by pairwise assessment of candidates. Xiao [22] employs weighted sum rule fusion on the detection scores to combine multiple detection algorithms on the same face image for the target application of face detection in dark environments (illumination by monitor light source). Weights are estimated depending on image quality (contrast, sharpness, color bit depth) by applying fuzzy adaptive fusion. Belhumeur *et al.* [4] train a global model of part-locations as hidden variables of a Bayesian objective function and combine the outcome of local detectors with a non-parametric set of prior models of face shape. Each local detector returns a score at each evaluation point used as likelihood that the desired part is located at this position. They keep track of the best global models and adopt a RANSAC-like generate-and-test approach to find a list of models that maximize the conditional probability. In their work they refer to [5], who also detect multiple candidate facial features and select the most face-like constellation using a statistical model, also able to deal with incomplete constellations. Aarabi *et al.* [1] present a fusion scheme of different frontal face detectors for live web-based face detection. They combine computationally inefficient, but reliable template-based techniques, popular neural networks, and subspace methods by moving a window of varying size over the image and optimizing for a fusion of face detection metrics. While these methods show good performance, they require access to detection scores, which are not always available. We therefore target the latter decision-level fusion scenario.

Decision-level fusion techniques in face detection are typically nested approaches reducing false negatives of face detectors by applying face-part detection on the found face objects. Wang *et al.* [19] study a hierarchical system detecting first a face and then locates eyes within the face boundary. Cristinacce *et al.* [6] propose a nested variant of locating facial features after face detection and combine individual locator results using a novel algorithm they call Pairwise Reinforcement of Feature Responses (PRFR) learning the pairwise distribution of all true feature locations relative to the best match of each individual feature detector. Face detection is used to predict approximate locations for each facial feature and local detectors are constrained to these locations. Nanni and Lumini [14] combine multiple face and eye detection systems using a serial scheme. In the first stage they employ two face detectors and eliminate false positive detections in the second stage with a proposed novel eye detector yielding 99.3% detection rate on BioID (faces detected with errors less than 25% offset with respect to the inter-eye distance) instead of 67%-97.8% for partial combinations. In case the second stage does not find eyes, the window is rejected as false positive. Micilotta *et al.* [13] present a coarse-to-fine approach for combining AdaBoost-based single body part detectors (face, torso, legs and hand-based classifier) using a Gaussian Mixture Model and RANSAC for assembly (selection of random body configurations), heuristics to eliminate outliers, a comparison of configurations with a trained a priori mixture model of upper-body configurations and selection of the configuration maximizing the likelihood.

Hierarchical methods are ideally suited for combination of methods with high positive rate, but very low false negatives and more accurate but possibly costly methods to eliminate false positives in the second stage. Our method is different in being able to account for false negatives as well, in contrast to e.g. [14]. While [6] make only use of position, we also take the detection size (w_i, h_i) into consideration.

3 Proposed system

We propose a fusion framework for object detectors $f_i : I \mapsto \{b_{i1}, \ldots b_{ik(I)}\}$ detecting faces and face-part objects, see Fig. 1. While our fusion method is generic in accepting arbitrary detectors, for experiments we employ Viola-Jones [18] based detectors for faces, eyes, eyepairs and noses, shipped with OpenCV[3] following Lienhart *et al.* [15]'s implementation. These detectors perform (1) *feature extraction* calculating properties of image regions, (2) *evaluation* selecting discriminative features with respect to the object to be detected and (3) *classification* judging whether a given window represents the object or not.

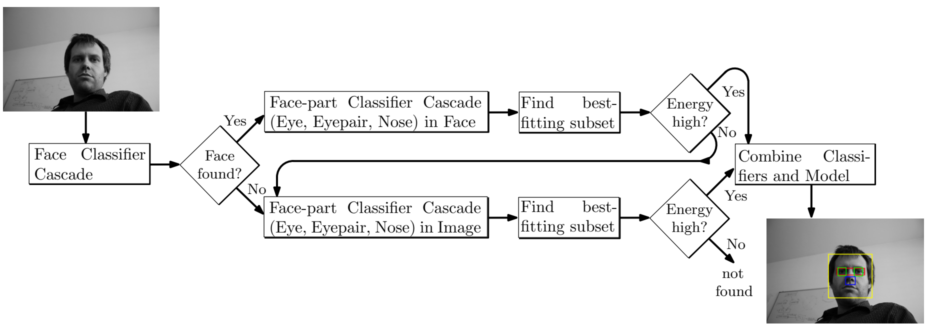


Fig. 1. Proposed model-driven face with face-part classifier fusion.

[3] OpenCV Library, `http://opencv.willowgarage.com`

3.1 Combining face and face-part classifiers

The combination of face and face-part classifiers requires the following steps: (1) *model estimation* builds a Gaussian model of priors and is computed beforehand, (2) *individual detection* employs single classifiers in parallel and/or serial combination, (3) *grouping* of results identifies the best-fitting subset of detection results with respect to an energy function taking spatial relationships into account, and (4) *refitting* of individual detection results with respect to the model.

Model estimation starts with a training phase. For each image I_j in a training set $\mathcal{I} = \{I_1, \ldots, I_n\}$, and each object detector $f_i \in \{f_1, \ldots, f_m\}$, we calculate the sets $\mathcal{C}^i = \cup_{j=1}^n f_i(I_j)$, from which all false detections are excluded (e.g. by rejecting all detection results with offsets greater than a fixed offset limit). In order to obtain scale-independent comparable results for regions $b \in f_i(I_j)$, position coordinates and size triples $b = (x, y, s)$ are given with respect to a coordinate system employing the midpoint between left and right eye R as origin, and the inter-eye distance S as unit value, see Fig. 2. We employ a single size parameter s instead of width and height, since typically classifiers employ a fixed w/h size ratio. Gaussian distributions $\mathcal{N}_x, \mathcal{N}_y, \mathcal{N}_s$ are fitted to the parameters of each detector f_i using its output on $\mathcal{I}$, i.e. we estimate the mean of parameters μ_x, μ_y, μ_s and standard deviations $\sigma_x, \sigma_y, \sigma_s$ for each f_i. Now, instead of predicting the position of the classifier, given the ground truth face position, we can use the output of f_i given an image I to exstimate the approximate localization R of the face according to the trained model. While for detectors in case of single localizations ($l = 1$, e.g. nose, mouth, eyepair) this procedure applies straightforward, detectors with multiple expected detections ($l > 1$, e.g. eyes) need multiple prediction models (for left and right eyes two reverse predictions are estimated, from which one is rejected during grouping). For training, μ_x, μ_y also consider the flipped classifier output for left and right eyes as well as the nose classifier. Also σ_y for eyes is corrected (by factor 3) to account for more variation with respect to head tilt. For face and eyepairs, location parameters μ_x and μ_y, respectively, are set to zero, to avoid overfitting with respect to the dataset. Finally, Z-normalization is applied to the Gaussians.

After individual face(part) detection, the best-fitting subset of detection results has to be identified with respect to the trained model. First, for m single classifiers, we select all combinations that can be formed for a subset containing i classifiers, $i \in \{1, \ldots, m\}$. By iterating from $c = 1$ to $c = 2^m$ and selecting the t-th classifier to be part of the subset, if the t-th bit in c is set, $c[t] = 1$, all possible combinations of detection results for these detectors are considered. There may be no combinations for a selected subset, if the classifier finds no objects of its type. Second, the location (reference position R) and size (inter-eye distance S) of the face with respect to a detection result (x, y, s) of the t-th detector is predicted using the trained classifier models $\mathcal{N}_x, \mathcal{N}_y, \mathcal{N}_s$ for detector type t. Origin $R = (X, Y)$ and inter-eye distance S are predicted as:

$$X := x - \mu_x \cdot s,\ Y := y - \mu_y \cdot s,\ S := \frac{s}{\mu_s} \tag{1}$$

Each of the three model parameters is assumed to exhibit a scattering of $\alpha := \sigma_x \cdot S, \beta := \sigma_y \cdot S, \gamma := \sigma_s \cdot S$. For a set $\mathcal{L} := \{L_1, \dots, L_m\}, L_i = (X_i, Y_i, S_i, \alpha_i, \beta_i, \gamma_i)$ of model locations we define the average model location $Avg(\mathcal{L})$ as follows:

$$Avg(\mathcal{L}) = (X_{Avg}, Y_{Avg}, S_{Avg}),\ X_{Avg} := (\sum_{i=1}^{m} \frac{1}{\alpha_i}) \cdot (\sum_{i=1}^{m} X_i \cdot \frac{1}{\alpha_i}),$$

$$Y_{Avg} := (\sum_{i=1}^{m} \frac{1}{\beta_i}) \cdot (\sum_{i=1}^{m} Y_i \cdot \frac{1}{\beta_i}),\ S_{Avg} := (\sum_{i=1}^{m} \frac{1}{\gamma_i}) \cdot (\sum_{i=1}^{m} S_i \cdot \frac{1}{\gamma_i}). \quad (2)$$

Location and size are weighted with the inverse deviation of each individual model prediction. Finally, it is desirable to estimate, how likely a subset represents a face. For $\mathcal{L}$ and a corresponding average model location $Avg(\mathcal{L}) = (X_{Avg}, Y_{Avg}, S_{Avg})$ we define the location energy:

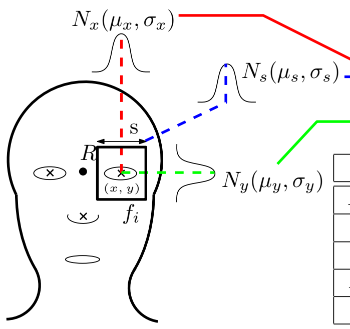
$$E(\mathcal{L}) := \frac{1}{m}\sqrt{1 + \sum_{i=1}^{m} \max(\frac{|X_i - X_{Avg}|}{\alpha_i}, \frac{|Y_i - Y_{Avg}|}{\beta_i}, \frac{|S_i - S_{Avg}|}{\gamma_i})} \quad (3)$$

This energy function turned out to be a good compromise between few well-fitting classifiers and many worse-fitting classifiers. Small values (less than 1) represent good model fits.

Finally, detection results of individual classifiers and the best-fitted model have to be combined for the face-part location task. Once, we have identified the best-fitting average model $Avg(\mathcal{L})$, we can now reset the position of individual classifiers. Each classifier (left eye, right eye, eyepair, nose, face), which is part of the best-fitting model is left at its original position. Each classifier not participating in the best average model is reconstructed based on the model, i.e. from μ_x, μ_y, μ_s of the corresponding classifier type.

3.2 The Viola-Jones approach

Single detectors in experiments are based on Viola-Jones' approach. Viola *et al.* [18] developed a very robust, accurate and real-time capable technique to the ob-



f_i	μ_x	σ_x	μ_y	σ_y	μ_s	σ_s
Eye_{right}	0.5	0.06	0	0.06	0.55	0.12
Eye_{left}	−0.5	0.06	0	0.06	0.55	0.12
Face	0	0.04	0.28	0.03	2.45	0.12
Eyepair	0	0.015	0	0.007	1.69	0.06
Nose	0	0.06	0.58	0.09	0.61	0.10

Fig. 2. Reference model under Gaussian assumption with trained parameters.

ject detection problem by employing rectangular Haar-features computed from the integral image for fast calculation (areas are calculated with 4 index operations). Each feature computed for fixed-sized windows x within I corresponds to a weak classifier h_j together with parity p_j (sign) and threshold θ_j (obtained by calculating means for the feature on both class sets and averaging) [18]:

$$h_j(x) = \begin{cases} 1 & \text{if } p_j \cdot f_j(x) < p_j \cdot \theta_j \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

In order to build strong classifiers from these simple weak ones, *AdaBoost* selects a small number of important features (at the drawback of long exhaustive classifier training and feature selection). This way, classifiers are trained by selecting one feature at a time and updating weights to produce a strong classifier using single features for T rounds [18]:

$$h(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Finally, a cascaded combination of classifiers is established, letting only windows pass on, which likely represent positive matches. This way, focus is given to promising regions, while retaining low false negative rate to speed up recognition (each node is trained with the false positives of the prior to quickly reject more likely non-face windows), i.e. iterative checks and rejection as soon as one classifier classifies the window as "non-face".

4 Experiments

Experiments target the application of eye detection in portrait images. We have compared results of the proposed model-driven face with face-part classifier fusion with the following alternatives: (1) *single* classifier, i.e. an eye-only classifier, and (2) a *nested* classifier performing a face cascade, i.e. eyes are detected in the most prominent face detection result. In order to judge the accuracy of detection results, we employ left-eye and right-eye offset (LO, RO) in percent of the inter-eye distance, the detection rate (DR) with respect to successful eye detections (detections with less than 20% LO and RO, respectively), and detection processing time (DT) in seconds. All results are given in Table 1.

Experiments are carried out using two different datasets: (1) *Casia-D* is a subset of 2 images per user (282 images) from the first publicly available long-range (3m) and high-quality NIR iris/face dataset with 2352×1728 pixel resolution, *CASIA-V4-Distance*[4] with manually selected eye and nose positions for ground truth; (2) *Yale-B* is a manually labeled subset of 252 images from challenging (varying illumination, different pose) VW 640×480 pixel resolution full-portrait face images (90 pixels eye distance) in *Yale Face Database B* [8].

From the detection rates in Tab. 1 we can see, that the proposed method delivers clearly the best results for the NIR dataset Casia-D with 5.95% LO and 5.74% RO and a total of 96.4% detection rate. Since the proposed method in

[4] CASIA Iris Image Database, `http://biometrics.idealtest.com`

Table 1. Offsets in percent of inter-eye distance (S) and Detection rates (Eye offset less than 20% of S) and Detection time (DT) per image of tested algorithms.

Algorithm	Left-eye Offset		Right-eye Offset		Detect Rate (%)		Detect Time (s)	
	Casia-D	YaleB	Casia-D	YaleB	Casia-D	YaleB	Casia-D	YaleB
Eye-only	19.48	14.06	32.73	17.13	65.8	87.3	0.60 s	0.61 s
Nested	81.6	3.78	139.9	5.26	14.6	97.6	0.65 s	0.28 s
Proposed	5.95	4.30	5.74	3.91	96.4	99.2	1.28 s	0.28 s

addition to performing a nested cascade for face detection tries each combination of detected face parts with the corresponding face to determine the combination that best matches a trained model, it is much more stable to single false detections. This is very likely the case, if the tested dataset differs in nature and/or recording conditions from the dataset used for feature training. Both eye-only with 19.48% LO, 32.73% RO and 65.8% DR and nested with 81.6% LO, 139.9% RO (these high rates occur, since frequently no eye can be detected) and 14.6% DR do not deliver satisfactory results. For YaleB, results are much closer, still the proposed variant delivers highest 99.2% DR (4.3% LO, 3.91% RO), before nested with 97.6% DR (3.78% LO, 5.26% RO) and eye-only with 87.3% DR (14.06% LO, 17.13% RO).

Regarding speed, the proposed method with 1.28 seconds DT per image needs twice as much time for Casia-D than the other approaches, because of many false-detections of faces and the conducted re-run on the entire image, not performed by the nested variant with 0.65 seconds DT (eye-only is only slightly faster with 0.6 seconds DT). However, the additional time needed by grouping and refitting is negligible, as can be seen from the low difference in YaleB detection times between Nested and Proposed (both 0.28 seconds per image). This re-run is not needed for YaleB, as it contains full portraits, no close-up images, see Figs. 3, 4. Eye-only is much slower than Nested for YaleB, because nested detection operates on a very small image (due to low resolution input and pyramidal operation mode, i.e. detection always operates on the smallest downscaled version greater than 256×256 pixels).

Recapitulatory, while the idea of grouping and refitting needs a check of all subsets of classifier results, and thus exponential complexity in the number

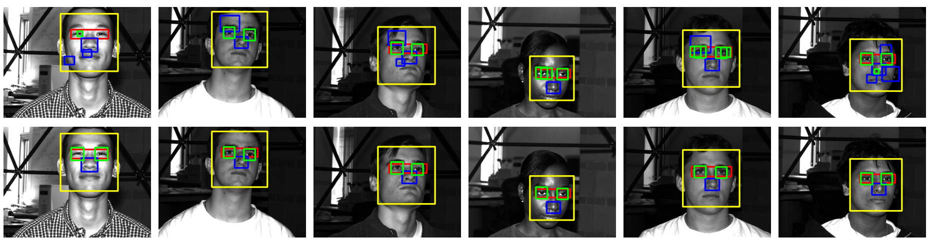


Fig. 3. YaleB Results: Nested (top) vs. Proposed (bottom)

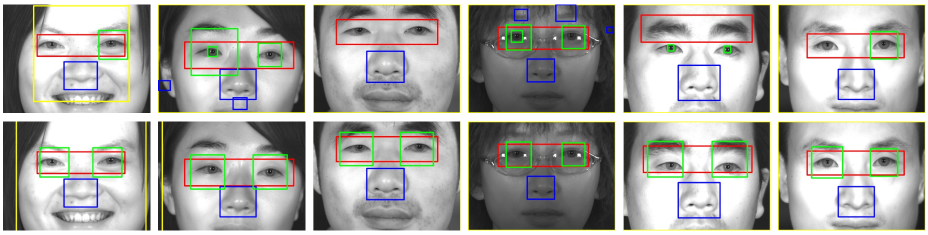


Fig. 4. Casia-D Results: Nested (top) vs. Proposed (bottom)

of returned detection results, practical evaluations show, that given a sufficient accuracy of individual responses the average case can be executed quite fast.

5 Summary

This work examined an approach to overcome limited capabilities of detection techniques with respect to database-independent face/eye detection. Experimental results indicated, that face and face part classifier fusion works relatively robust and does not consume significantly more time than nested approaches. By exploiting spatial relationships between single detectors, heterogeneous detection of faces and face-parts (eyes, nose, epairs, etc.) in NIR and WV images can be achieved, benefiting of the best currently available detectors for each type of imagery without suffering from a high individual false-detection rate because of high variability between training and testing datasets. Future topics of research include testing on different datasets, an optimization of models for the refitting task, and a combination with iris segmentation for combined iris and face recognition. While several approaches exist reducing false positive detections by incorporating multiple (face-part) detectors, the proposed technique is one of very few fusion techniques also able to reduce false negatives restaurating missing information (caused by, e.g., occlusions), alleviates simple integration into existing detectors, and does not require access to detection scores.

References

1. Aarabi, P., Lam, J., Keshavarz, A.: Face detection using information fusion. In: Proc. Int'l Conf. on Information Fusion. pp. 1–8 (2007)
2. Belaroussi, R., Milgram, M., Prevost, L.: Fusion of multiple detectors for face and eyes localization. In: Proc. Int'l Symp. Image and Signal Processing and Analysis (ISPA). pp. 24–29 (2005)
3. Belaroussi, R., Prevost, L., Milgram, M.: Multi-stage fusion for face localization. In: Proc. Int'l Conf. on Information Fusion. pp. 1–8 (2005)
4. Belhumeur, P., Jacobs, D., Kriegman, D., Kumar, N.: Localizing parts of faces using a consensus of exemplars. In: Proc. IEEE Conf. on Comp. Vis. and Pattern Rec. (CVPR). pp. 545 – 552 (2011)
5. Burl, M., Leung, T.K., Perona, P.: Face localization via shape statistics. In: Workshop on Automatic Face and Gesture Recognition. pp. 1–6 (1995)

6. Cristinacce, D., Cootes, T., Scott, I.: A multi-stage approach to facial feature. In: Proc. Brit. Mach. Vis. Conf. (BMVC). pp. 231–240 (2004)
7. Gan, J.Y., Liang, Y.: A method for face and iris feature fusion in identity authentication. Int. J. Comp. Sci. Network Sec. 6(2), 135–138 (2006)
8. Georghiades, A.S., Belhumeur, P.N.: From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Mach. Intell. 23(6), 643–660 (2001)
9. Jin, L., Yuan, X., Satoh, S., Li, J., Xia, L.: A hybrid classifier for precise and robust eye detection. In: Proc. Int'l Conf. on Pattern Recognition (ICPR). pp. 731–735 (2006)
10. Kroon, B., Hanjalic, A., Maas, S.: Eye localization for face matching: is it always useful and under what conditions? In: Proc. Int'l Conf. on Content-based Image and Video Retrieval (CIVR). pp. 379–388 (2008)
11. Ma, Y., Ding, X., Wang, Z., Wang, N.: Robust precise eye location under probabilistic framework. In: Proc. Int'l. Conf. on Autom. Face and Gesture Rec. pp. 339–344 (2004)
12. Matey, J., Naroditsky, O., Hanna, K., Kolczynski, R., LoIacono, D., Mangru, S., Tinker, M., Zappia, T., Zhao, W.Y.: Iris on the move: Acquisition of images for iris recognition in less constrained environments. Proc. IEEE 94, 1936–1947 (2006)
13. Micilotta, A., Jon, E., Bowden, O.: Detection and tracking of humans by probabilistic body part assembly. In: Proc. of British Machine Vision Conference (BMVC). pp. 429–438 (2005)
14. Nanni, L., Lumini, A.: A combination of face/eye detectors for a high performance face detection system. IEEE Multimedia PP(99), 1–15 (2011)
15. R. Lienhart, A. Kuranov, V.P.: Empirical analysis of detection cascades of boosted classifiers for rapid object. Tech. rep., Microproc. Res. Lab, Intel Labs (2002)
16. Saragih, J., Lucey, S., Cohn, J.: Face alignment through subspace constrained mean-shifts. In: Proc. Int'l Conf. on Computer Vision (ICCV). pp. 1034 – 1041 (2009)
17. Uhl, A., Wild, P.: Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. In: Proc. Int'l Conf. on Biometrics (ICB). pp. 1–8 (2012), to appear
18. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. IEEE Conf. on Comp. Vis. and Pattern Rec. (CVPR). pp. 511–518 (2001)
19. Wang, P., Green, M., Ji, Q., Wayman, J.: Automatic eye detection and its validation. In: Proc. IEEE Conf. on Comp. Vis. and Pattern Rec. (CVPR). pp. 164–171 (2005)
20. Wang, Y., Tan, T., Jain, A.: Combining face and iris biometrics for identity verification. In: Proc. Int'l Conf. on Audio- and Video-based Biometric Person Auth. (AVBPA). pp. 805–813 (2003)
21. Wheeler, F., Perera, A., Abramovich, G., Yu, B., Tu, P.: Stand-off iris recognition system. In: Proc. IEEE Int'l Conf. on Biometrics: Theory, Applications and Systems (BTAS). pp. 1–7 (2008)
22. Xiao, Q.: Face detection using information fusion. In: IEEE Workshop on Comp. Intell. in Biometrics and Identity Mgmnt. (CIBIM). pp. 157–162 (2011)
23. Zhang, C., Zhang, Z.: A survey of recent advances in face detection. Tech. rep., Microsoft Research (2010), mSR-TR-2010-66
24. Zhang, Z., Wang, R., Pan, K., Li, S., Zhang, P.: Fusion of near infrared face and iris biometrics. In: Lee, S.W., Li, S. (eds.) Proc. Int'l Conf. on Biometrics (ICB). LNCS, vol. 4642, pp. 172–180. Springer (2007)