

iFR: Interactively Pose Corrected Face Recognition

Simon Nash, Mark Rhodes and Joanna Isabelle Olszewska

School of Computing and Technology, University of Gloucestershire, The Park, Cheltenham, GL50 2RH, U.K.

Keywords: Face Detection, Face Recognition, AdaBoost, Template, Facial Features, Eigenfaces, Pose Correction, Human-Computer Interactive Systems.

Abstract: Although face recognition applications are growing, robust face recognition is still a challenging task due e.g. to variations in face poses, facial expressions, or lighting conditions. In this paper, we propose a new method which allows both automatic face detection and recognition and incorporates an interactive selection of facial features in conjunction with our new pose-correction algorithm. Our resulting system we called iFR successfully recognizes faces across pose, while being computationally efficient and outperforming standard approaches, as demonstrated in tests carried out on publicly available standard datasets.

1 INTRODUCTION

In recent years, computational face recognition has attracted an increasing attention because of its potential use in biometric systems (Jain et al., 2012), video-surveillance (Rambach et al., 2015), automated behavior analysis (Han et al., 2013), social networks (Choi et al., 2011), or human-machine interactions (Tasli et al., 2015).

For this purpose, different techniques for automatic face recognition have been developed. These methods could rely on facial features (Brunelli and Poggio, 1993), active appearance models (Cootes et al., 2001), active shape models (Lindner et al., 2015), or facial representations like Local Binary Patterns (LBP) (Ahonen et al., 2006), (Montagne et al., 2013). Other approaches use face decomposition in eigenfaces (Turk and Pentland, 1991) or empirical mode decomposition (EMD) (Gallego-Jutgla et al., 2013). Face recognition techniques could also involve neural networks (Nagi et al., 2008).

Despite these methods have been proven to be efficient, many challenges remain. In particular, pose variation has been identified as of great impact on the face recognition scores (Du and Ward, 2006). Indeed, face appearance and/or shape could change due to head's movements which can be described by the egocentric rotation angles, namely, pitch, roll, and yaw (Murphy-Chutorian and Trivedi, 2009), leading among others to intra-subject variations and making the face recognition across pose a task difficult (Zhang and Gao, 2009).

In the literature, few works have proposed to study face recognition under pose variations. In (Bordoloi and Sarma, 2009), they applied a pose correction by re-enforced multi-level decision support, using a generalized feed forward artificial neural network (GF-FANN) designed to detect the rotation angle of an input image and to correct it by applying the three pass shear rotation. (Nakamura and Takano, 2011) developed a rotation and size spreading associative neural network (RS-SAN), while (Huang et al., 2012) adopted a deep learning approach, using a deep belief network (DBN). These methods have performance in high 80% at extensive computational cost of dozens of seconds, whereas (Bordoloi and Sarma, 2009) achieve recognition success rates in 90% within several hundreds of seconds.

In this work, we propose to tackle with pose variations by introducing a new algorithm for automatic pose correction in order to efficiently process face recognition. Our algorithm makes decision based on feature points such as the nose tip and the nose top rather than on block facial features as in (Bordoloi and Sarma, 2009).

Our face recognition method deals with recognition of faces in arbitrary, in-depth rotations.

Our iFR system provides both flexibility in the choice of facial points by integrating a user-driven, interactive phase, and automation in the phase of pose correction through the new algorithm. On the other hand, our face recognition approach combines advantages of both facial features and template approaches as demonstrated in Section 3.

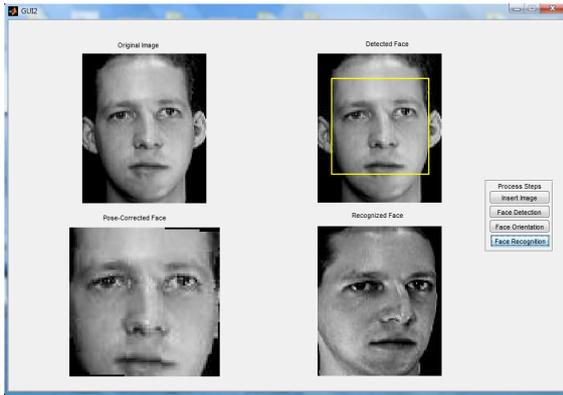


Figure 1: Example of our iFR system in action, featuring automatic face detection, interactive pose correction, and automatic face recognition.

The contribution of this paper is twofold. It consists, on one hand, in a new face recognition system which enables a user to select interactively facial feature points, such as nose tip and nose top, in order to feed the automatic process correcting the face pose. On the other hand, the pose correction itself is done by a new algorithm which we present in this paper and which contributes to improve face recognition performance.

The paper is structured as follows. In Section 2, we describe our interactively pose corrected face recognition method (see Fig. 1). The resulting system, that automatically detects and recognizes a face which pose has been corrected by our algorithm, has been successfully tested on standard dataset, as reported and discussed in Section 3. Conclusions are drawn up in Section 4.

2 PROPOSED APPROACH

Our pose-corrected face recognition (iFR) system consists of three phases. At first, the automatic face detection, described in Section 2.1, performs the detection and location of a face in an image. Secondly, the located face is cropped automatically and serves as input to the second step. This latter one is the interactive phase of the system as explained in Section 2.2. Indeed, it is the user selecting two facial feature points based on which the pose correction is computed automatically. Thirdly, the recognition phase processes that pose-corrected face image, and provides the system output which could be either the recognized face or the ‘unrecognized pictogram’, as detailed in Section 2.3.

2.1 Automatic Face Detection

In our iFR system, the first phase aims to automatically detect and locate face(s) in a given image $I(x, y)$ with M and N , its width and height, respectively. For this purpose, we implemented the Viola & Jones approach (Viola and Jones, 2004), which is based firstly on the detection of Haar-like features and secondly on the processing of the AdaBoost classifier.

Local Haar-like features f encode the existence of oriented contrasts between regions in the processed image. They are computed by subtracting the sum of all the pixels of a subregion of the local feature from the sum of the remaining region of the local feature, using the integral image representation $II(i, j)$ which is defined as follows:

$$II(i, j) = II(i-1, j) + II(i, j-1) - II(i-1, j-1) + I(i, j), \quad (1)$$

where $I(i, j)$ is the pixel value of the original image at the position (i, j) (Wood and Olszewska, 2012).

The AdaBoost classifier requires at first a training phase during which it builds strong classifiers based on cascades of weak classifiers. In particular, to form a T -stage cascade, T weak classifiers are selected using the AdaBoost algorithm (Viola and Jones, 2004). In fact at a t stage of this cascade, a sub-window u of an image from the training set is computed by eq. (1) and it is passed to the corresponding t^{th} weak classifier. If the region is classified as a non-face, the sub-window is rejected. If not, it is passed to the $t+1$ stage, and so forth. Consequently, more stages the cascade owns, more selective it is, i.e. less false positive detections occur. However, that could lead to the increase in the number of false negative detection.

In order to select at each t level (with $1 < t < T$) the best weak classifier, an optimum threshold θ_t is computed by minimizing the classification error due to the selection of a particular Haar-like feature value $f_t(u)$. The resulting weak classifier k_t is thus obtained as follows:

$$k_t(u, f_t, p_t, \theta_t) = \begin{cases} 1 & \text{if } p_t f_t(u) < p_t \theta_t, \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where p_t is the polarity indicating the direction of the inequality.

Then, a strong classifier $K_T(u)$ is constructed by taking a weighted combination of the selected weak classifiers k_t according to

$$K_T(u) = \begin{cases} 1 & \text{if } \sum_{t=1}^T \alpha_t k_t(u) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where $\frac{1}{2} \sum_{t=1}^T \alpha_t$ is the AdaBoost threshold and α_t is a voting coefficient computed based on the classification error in each stage t of the T stages of the cascade (Viola and Jones, 2004).

Once the training phase is completed, the Adaboost classifier could be applied to a *test* image. The Viola & Jones' face detector (Viola and Jones, 2004) outputs then a bounding box surrounding the detected face (Wood and Olszewska, 2012), rather than its contour (Olszewska, 2013). Next, this bounding box' size and location are used to automatically crop out the face from the image, and our iFR system saves it as a new file.

2.2 Interactive Pose Correction

This phase of pose correction consists of two steps, one interactive and one automatic.

Hence, once the cropped face is uploaded into the iFR system (see bottom left of Fig. 1), the user at first determines the nose tip point (see Fig. 2(a)) and then the nose top point (see Fig. 2(b)) interactively. For that, the user clicks each time on the displayed face image to determine the nose tip point and the nose top point, respectively.

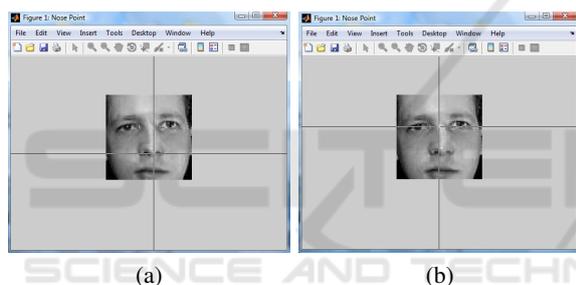


Figure 2: Interactive user's selection of the (a) nose tip; (b) nose top.

Secondly, these data are passed onto our system algorithm (Algorithm 1) in order to correct the face pose automatically.

Algorithm 1 performs first a flip of the face image, if need be (Fig. 3(a)). Indeed, in first instance, it compares the coordinates of the nose tip identified by the user with the value of half the height of the cropped image in order to determine whether the nose point is higher or lower than this point. If the nose tip is lower than the halfway line, then the image is already the right way up. However, if the nose tip is higher than the halfway line, the face is upside down, and then, is flipped 180° to turn the face the right way up.

Next, Algorithm 1 computes the angle to rotate the face in order to correct its pose (Fig. 3(b)). Hence, it automatically identifies the center point of the image, while the center point between both eyes, or nose top, has been interactively selected by the user. Then, it calculates, on one hand, the segment *line1* between the center point of the image and the nose top, and,

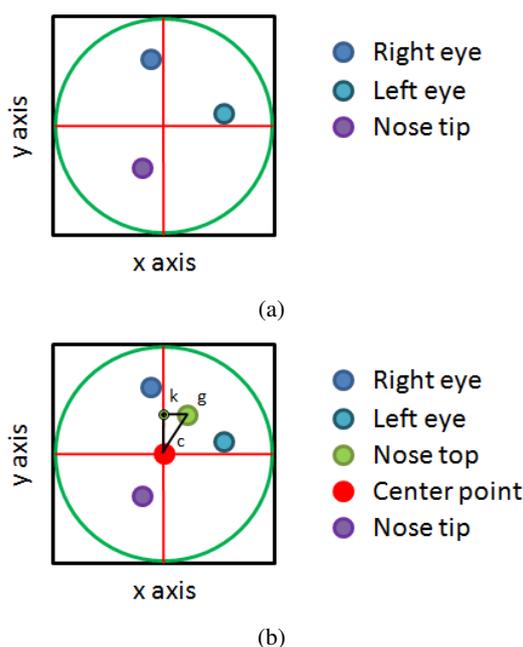


Figure 3: Illustration of the automatic pose correction method steps: (a) face flip; (b) face rotation. Best viewed in color.

on the other hand, the segment *line2* between the nose top and its projection on the vertical halfway line. Next, the lengths of the segments are computed as well as the in-between *angle* using trigonometric rules. Finally, the image is rotated according to that angle for pose correction and saved in a file to be used for automatic face recognition.

2.3 Automatic Face Recognition

The automated face recognition phase is a one-to-many process (Du and Ward, 2006). It compares a specific face image (*probe*) against all of the images in a face database (*gallery*). In particular, iFR system uses the eigenface approach (Turk and Pentland, 1991), which is based on a holistic region-based representation and pixel intensity features, and which introduces the principal component analysis (PCA) of the distribution of faces, or eigenvectors of the covariance matrix of the set of face images. These eigenvectors can be considered as a set of features which together characterize the variation between face images. Each image location contributes more or less to each eigenvector. So, an eigenvector can be seen as a sort of ghostly face called eigenface. Probe face classification, i.e. recognition of a test face, is then determined by applying the nearest neighbour approach to the probe face projection in face space.

Technically, eigenface-based face recognition is split in a training step and a testing step.

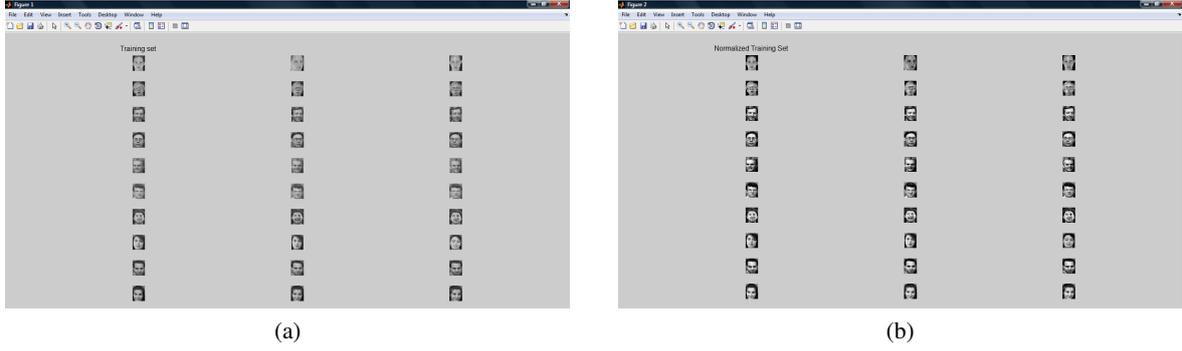


Figure 4: Example of the training dataset used for face recognition: (a) initial training dataset; (b) normalized training dataset.

Algorithm 1: Automatic Pose Correction.

Given $F = (F_x, F_y)$, the cropped face image,
 with its center point $c = (HF_x, HF_y)$;
 Given the selected nose tip point $p = (NP_x, NP_y)$;
 Given the selected nose top point $g = (CE_x, CE_y)$;
 k is the projection of g on the half-face axis,
 with $k = (HF_x, CE_y)$;

do

line1=pdist2(c,g)
line2=pdist2(g,k)

if $NP_y - HF_y < 0$ **then**
 imrotate(180°) ▷ Face Flip
end if

if $CE_x - HF_x > 0$ **then**
 if $(CE_y == HF_y)$ **then** *angle* = 90°
 else *angle*=rattodeg(sin(*line2*/*line1*))
 end if
else
 if $(CE_y == HF_y)$ **then** *angle* = 270°
 else *angle*=360°-rattodeg(sin(*line2*/*line1*))
 end if
end if
 imrotate(*angle*) ▷ Face Rotation

end do

At first, the training set is made of a set of face pictures, several poses per person as illustrated in Fig. 4(a). This training set is subject to facial pre-processing (Travieso-Gonzalez et al., 2011), which in this case consists in normalizing the training set by changing the mean and standard of all the testing images, in order to reduce errors due to lighting condition (see Fig. 4(b)).

Hence, let us consider the normalized training set

of face images be $\Gamma_1, \Gamma_2, \Gamma_3, \dots, \Gamma_F$. The average face of the set is defined by:

$$\Psi = \frac{1}{F} \sum_{i=1}^F \Gamma_i. \quad (4)$$

Each face differs from the average by the vector $\Phi_i = \Gamma_i - \Psi$. This set of very large vectors defines the face space, and is then subject to Principal Component Analysis (PCA) for dimension reduction. Thus, PCA seeks a set of P orthonormal vectors \mathbf{v}_i and their associated eigenvalues λ_i which best describes the distribution data. The vectors \mathbf{v}_i and scalars λ_i are the eigenvectors and eigenvalues, respectively, of the covariance matrix:

$$\begin{aligned} \Psi &= \frac{1}{F} \sum_{i=1}^F \Phi_i \Phi_i^T \\ &= AA^T, \end{aligned} \quad (5)$$

where the matrix $A = [\Phi_1 \Phi_2 \dots \Phi_F]$.

Secondly, in the testing phase, a new face (Γ) is transformed into its eigenface components, i.e. projected into the face space, by the operation $\omega_l = \mathbf{v}_l^T (\Gamma - \Psi)$ for $l = 1 \dots F'$.

The weights form a vector $\Omega^T = [\omega_1 \omega_2 \dots \omega_{F'}]$ that describes the contribution of each eigenface in representing the input face image, treating the eigenfaces as a basis set for face images.

To determine which face class provides the best description of an input face image, one can find the face class l that minimizes the Euclidean distance $\varepsilon_l = \|\Omega - \Omega_l\|$, where Ω_l is a vector describing the l^{th} face class (Turk and Pentland, 1991). Thus, a face is classified as belonging to class l according to the nearest neighbour approach, i.e. when the minimum ε_l is below a chosen threshold θ_ε , e.g. as illustrated for the True Positive (TP) case in Fig. 1. Otherwise, the face is classified as ‘unknown’ and an ‘unrecognized pictogram’ is displayed as for the True Negative (TN) case in Fig. 5.

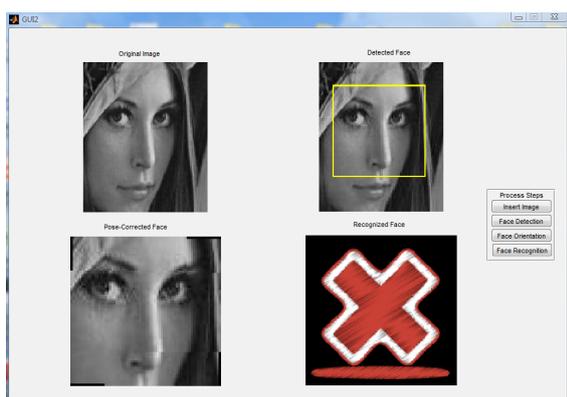


Figure 5: Example of our iFR system in action, featuring automatic face detection, interactive pose correction, and automatic face recognition.

Hence, the automatic recognition of a cropped, detected face is done quickly and precisely as demonstrated in Section 3.

3 EXPERIMENTS AND DISCUSSION

To validate our approach, we have applied our iFR system on the publicly available, well-established ORL dataset (Gross, 2005).

The ORL database is a set of faces taken between April 1992 and April 1994 at the Olivetti Research Laboratory in Cambridge, UK. This dataset contains 400 pictures of 40 distinct subjects. The images are organised in 40 directories (one for each subject) with 10 images per class. All the images are taken against a dark homogeneous background and the subjects are in up-right, frontal position (with tolerance for some side movements). The files are in portable gray map (pgm) format, and the size of each image is 92x112, 8-bit grey level.

This database owns challenges such as face variations over time since some of the pictures were taken at different times; gender variations since the dataset contains a mix of images with faces of different men and women; presence or absence of structural components (e.g. beard/no beard, etc.) and variation of facial details (e.g. glasses/no-glasses, etc.); variations in facial expressions (e.g. open/closed eyes, smiling/non-smiling, angry, neutral, etc.); variations of illuminations conditions; and variations in face poses such as face in-plane and/or out-plane rotations. In particular, ORL face poses are within $\pm 20^\circ$ in yaw and roll (Zhang and Gao, 2009). All these difficulties make the dataset challenging and interesting to test our approach on.

All the experiments have been run on a computer with Intel Core 2 Duo Pentium T9300, 2.5 GHz, 2Gb RAM, using MatLab.

We have assessed the performance of our automatic face recognition system by computing the face recognition rate defined as

$$recognition\ accuracy = \frac{TP + TN}{TP + TN + FN + FP}, \quad (6)$$

with TP, true positive, TN, true negative, FP, false positive, and FN, false negative. The obtained average face detection rate by our method has been reported and compared to state-of-art techniques in Table 1.

Experiments have been carried out to test our iFR system. It consisted in applying the face recognition method on the ORL database with a training set of 30 pictures (10 subject classes, 3 face poses per class) and a testing set of 100 pictures (10 subject classes, 10 face poses per class).

Examples of the results of the state-of-art recognition method (without any pose correction) as well as the results obtained using our face recognition method (with pose correction) are presented in Figs. 6. (b) and (c), respectively.

In Table 1, we have reported the face recognition scores for the state-of-art eigenfaces which do not include pose correction, and for our iFR which includes eigenfaces in conjunction with our Algorithm 1 for pose correction.

Table 1: Face recognition accuracy scores obtained by approaches such as \diamond (Viola and Jones, 2004), and our method.

	\diamond	our
recognition accuracy	92%	94%

From Table 1, we can observe that our iFR system shows excellent accuracy in face recognition, outperforming the state-of-the-art eigenface method.

The average computational speed of our approach is in the range of milliseconds, thus our developed system could be used in context of online face recognition.

Furthermore, for both speed and accuracy, iFR design outperforms other approaches using pose correction such as (Bordoloi and Sarma, 2009), (Nakamura and Takano, 2011), (Huang et al., 2012).

4 CONCLUSIONS

In this work, we have proposed a computationally efficient and robust face recognition system called iFR, consisting of three phases, namely face detection,



Figure 6: Examples of face recognition results for (a) given test images; (b) using state-of-art method; (c) our method.

pose correction, and face recognition. Hence, the face detection phase uses the Viola & Jones method, while the pose correction phase relies on our new, pose-correction algorithm fed by interactively selected facial features, such nose tip and nose top, in the detected face. The face recognition phase applies the eigenface approach on the pose-corrected face. Our iFR system combines the advantages of both template and feature approaches for face recognition. It outperforms the state-of-the-art methods, and is well suited system for online applications involving face recognition.

REFERENCES

- Ahonen, T., Hadid, A., and Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041.
- Bordoloi, H. and Sarma, K. K. (2009). Face recognition with image rotation detection, correction and reinforced decision using ANN. *International Journal of Electrical and Electronics Engineering*, 3(7):1205–1212.
- Brunelli, R. and Poggio, T. (1993). Face recognition: features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052.
- Choi, J. Y., Neve, W. D., Plataniotis, K. N., and Ro, Y. M. (2011). Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks. *IEEE Transactions on Multimedia*, 13(1):14–18.
- Cootes, T. F., Edwards, G. J., and Taylor, C. J. (2001). Active appearance model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685.
- Du, S. and Ward, R. (2006). Face recognition under pose variations. *Journal of the Franklin Institute*, 343(6):596–613.
- Gallego-Jutgla, E., de Ipina, K. L., Marti-Puig, P., and Sole-Casals, J. (2013). Empirical mode decomposition-based face recognition system. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 445–450.
- Gross, R. (2005). *Face Databases*, chapter Handbook of Face Recognition, pages 301–327. Springer-Verlag.
- Han, M.-J., Lin, C.-H., and Song, K.-T. (2013). Robotic emotional expression generation based on mood transition and personality model. *IEEE Transactions on Cybernetics*, 43(4):1290–1303.
- Huang, G. B., Mattat, M. A., Lee, H., and Learned-Miller, E. (2012). Learning to align from scratch. In *Proceedings of the Conference on Advances in Neural Information Processing Systems*, pages 764–772.
- Jain, A. K., Klare, B., and Park, U. (2012). Face matching and retrieval in Forensics applications. *IEEE Multi-Media*, 19(1):20.
- Lindner, C., Bromiley, P. A., Ionita, M. C., and Cootes, T. F. (2015). Robust and accurate shape model matching using random forest regression voting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9):1862–1874.
- Montagne, C., Kodewitz, A., Vigneron, V., Giraud, V., and Lelandaïs, S. (2013). 3D local binary pattern for PET image classification by SVM. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 145–4150.
- Murphy-Chutorian, E. and Trivedi, M. M. (2009). Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):607–626.
- Nagi, J., Ahmed, S. K., and Nagi, F. (2008). A Matlab-based face recognition system using image processing and neural networks. In *Proceedings of the International Colloquium on Signal Processing and its Applications*, pages 83–88.
- Nakamura, K. and Takano, H. (2011). *New Approaches to Characterization and Recognition of Faces*, chapter Face discrimination using the orientation and size recognition characteristics of the spreading associative neural network, pages 197–212. InTech.
- Olszewska, J. I. (2013). Multi-scale, multi-feature vector flow active contours for automatic multiple-face detection. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 429–435.
- Rambach, J., Huber, M. F., Balthasar, M. R., and Zoubir, A. M. (2015). Collaborative multi-camera face recog-

- nition and tracking. In *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance*, page 1.
- Tasli, H. E., den Uyl, T. M., Boujut, H., and Zaharia, T. (2015). Real-time facial character animation. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, page 1.
- Travieso-Gonzalez, C. M., del Pozo-Banos, M., and Alonso, J. B. (2011). *Biometric Systems, Design and Applications*, chapter Facial identification based on transform domains for images and videos, pages 293–316. InTech.
- Turk, J. and Pentland, A. (1991). Face recognition using eigenfaces. In *Proceedings of the IEEE International Conference on computer Vision and Pattern Recognition*, pages 586–591.
- Viola, P. and Jones, M. J. (2004). Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.
- Wood, R. and Olszewska, J. I. (2012). Lighting-variable AdaBoost based-on system for robust face detection. In *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 494–497.
- Zhang, X. and Gao, Y. (2009). Face recognition across pose: A review. *Pattern Recognition*, 42(11):2876–2896.

