Detection of Facial Feature Points Using Anthropometric Face Model

Abu Sayeed Md. Sohail and Prabir Bhattacharya

Concordia University
1455 de Maisonneuve Blvd. West, Montréal, Québec H3G 1M8, Canada
E-mails: a_sohai@encs.concordia.ca, prabir@ciise.concordia.ca

Abstract. This paper presents an automatic technique for detecting the 18-most important facial feature points using a statistically developed anthropometric face model. Most of the important facial feature points are located just about the area of mouth, nose, eyes and eyebrows. After carefully observing the structural symmetry of human face and performing necessary anthropometric measurements, we have been able to build a model that can be used in isolating the above mentioned facial feature regions. In the proposed model, distance between the two eye centers serves as the principal parameter of measurement for locating the centers of other facial feature regions. Hence, our method works by detecting the two eye centers in every possible situation of eyes and isolating each of the facial feature regions using the proposed model. Combinations of differnt image processing techniques are then applied within the localized regions for detecting the 18-most important facial feature points. Experimental result shows that the developed system can detect the 18-feature points successfully in 90.44% cases when applied over the test databases.

Keywords: Facial Feature Point Detection, Anthropometric Face Model, Anthropometric Measurement.

1 Introduction

Identification of facial feature points plays an important role in many facial image applications like human computer interaction, video surveillance, face detection, face recognition, facial expression classification, face modeling and face animation. A large number of approaches have already been attempted towards addressing this problem, but complexities added by circumstances like inter-personal variation (i.e. gender, race), intra-personal changes (i.e. pose, expression) and inconsistency of acquisition conditions (i.e. lighting, image resolution) have made the task quite difficult and challenging. All the works that have addressed the problem of facial feature point detection so far can be grouped into several categories on the basis of their inherent techniques. Geometrical shape of facial features has been adopted in several works for facial feature point localization and detection [1][2]. Each feature is demonstrated as a geometrical shape; for example, the shape of the eyeball is circle and the shape of the eyelid is ellipse. This method can detect facial features very well in neutral faces, but fails to show good performance in handling the large variation in face images occurred due to pose and expression [3]. To overcome this limitation, a variation of shape-based approaches that looks for specific shape in the image adopting deformable and non deformable template matching [4][5], graph matching [6], snakes [7] or the Hough Transformation [8] has also been deployed. Due to the inherent difficulties of detecting facial feature points using only a single image, spatio-temporal information captured from subsequent frames of video sequence has been used in some other work for detection and tracking facial feature points [9][10]. Combination of color information of each of the facial features has been extracted and used to detect the feature points in some other works [11][12]. Color based algorithms are applicable for only color images and can't be used for gray scale images. Facial feature points detection approaches using machine learning techniques like Principle Component Analysis [13], Neural Network [3], Genetic Algorithm [14] and Haar classifiers [15] require a large number of face images and computational time for initial training. There are also some works that have used image intensity as the most important parameter for detection and localization of facial features [16][17].

Although anthropometric measurement of face provides useful information about the location of facial features, it has rarely been used in their detection and localization. In this paper, we have explored the approach of using a statistically developed, reusable anthropometric face model for localization of the facial feature regions as well as the detection of the18-most important facial feature points from these isolated regions using a hybrid image processing technique. The subsequent discussion has been organized into the following six sections: Section-II explains the proposed anthropometric face model, Section-III focuses on the isolation of facial feature regions using the anthropometric face model, Section-IV explains the techniques of detecting the 18-feature points from the identified face regions, experimental results are represented in Section-V and finally Section-VI concludes the paper.

2 Anthropometric Model for Facial Feature Region Localization

Anthropometry is a biological science that deals with the measurement of the human body and its different parts. Data obtained from anthropometric measurement informs a range of enterprises that depend on knowledge of the distribution of measurements across human populations. After carefully performing anthropometric measurement on 300 frontal face images taken from more than 150 subjects originated in different geographical territories, we have been able to build an anthropometric model of human face that can be used in localizing facial feature regions from face images [18]. Rather than using all the landmarks used by Farkas [19], we have used only a small subset of points and have added some new landmarks in our model. The landmarks points that have been used in our face anthropometric model for facial feature localization are represented in Fig. 1. It has been observed from the statistics of proportion evolved during our initial observation that, location of these points (P3, P4, P6, P7) can be obtained from the distance between two eye centers (P1 and P2) using midpoint of eyes (P5) as an intermediate point since distances between the pair of

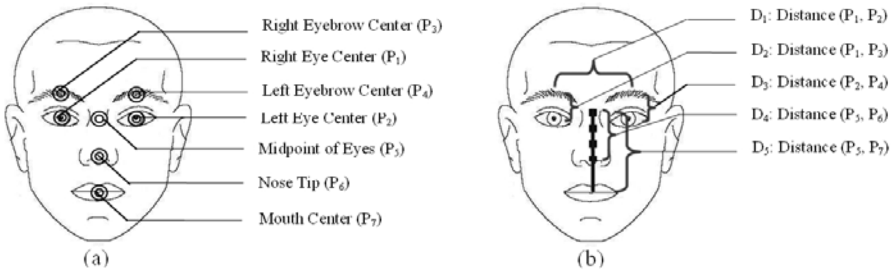

Fig. 1. Proposed Anthropometric Face Model for facial feature region localization (a) Landmarks used in our Anthropometric Face Model (b) Distances (Anthropometric Measurements)

points (P1 and P3), (P2 and P4), (P5 and P6), (P5 and P7) maintain nearly constant proportions with the distance between the center of left and right eyes (P1 and P2).Our proposed anthropometric face model of facial feature region localization has been developed from these proportional constants (Table 1) using distance between the center of left and right eyes (P1 and P2) as the principle parameter of measurement.

Table 1. Proportion of distances (D2, D3, D4, and D5) to D1 measured from subjects of different geographical territories

Proportion	Description	Constant
D_2/D_1	Proportion of the distance between right eye center and right eyebrow center to the distance between eye centers	≈ 0.33
D_3/D_1	Proportion of the distance between left eye center and left eyebrow center to the distance between eye centers	≈ 0.33
D_4/D_1	Proportion of the distance between midpoint of eye centers and nose tip to the distance between eye centers	≈ 0.60
D_5/D_1	Proportion of the distance between midpoint of eye centers and mouth center to the distance between eye centers	≈ 1.10

3 Identification of Facial Feature Region

Since distance between the two eye centers serves as the principal parameter for measuring the center locations of other facial feature regions, implementation of our Automatic Facial Feature Point Detection System begins with the detection of two eye centers using the generative framework for object detection and classification proposed by I. Fasel et al. [20]. This method has been able to point out the eye centers correctly almost in 99% cases for our dataset. Once right and left eye centers (P1 and P2) are detected, we step forward for measuring the rotation angle of the face in face image over horizontal axis (x-axis). For this purpose, we have imagined a right angled triangle formed with the right eye center (x1, y1), left eye center (x2, y2) and the third point that serves as the crossing point of the horizontal and vertical lines passing through the right and left eye centers respectively (Fig. 2).

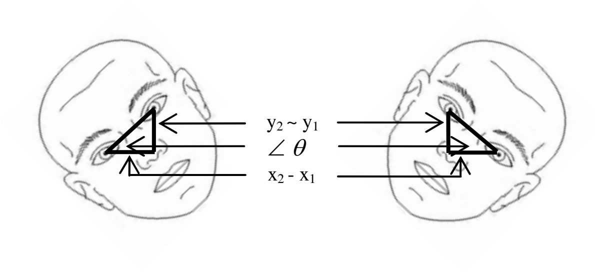

Fig. 2. Horizontal rotation (rotation over x-axis) of the face in a face image is corrected by determining the rotation angle. Co-ordinate values of the eye centers are used to calculate the amount of rotation to be performed.

The amount of horizontal rotation of the face, θ is then determined using the following equation:

$$\theta = \tan^{-1}\left(\frac{|y_1 - y_2|}{x_2 - x_1}\right)$$

For fitting the face with our model, the whole image is then rotated by the angle θ. The direction of the rotation (clock-wise or counter clock-wise) is determined by the polarity of difference of vertical distances (y_1 - y_2) between the right and left eye centers. If the difference is positive, the image is rotated in clock-wise direction, otherwise it is rotated to counter clock-wise direction. The new location of the right eye center (x_1, y_1) and left eye center (x_2, y_2) are then updated by detecting them once again over the rotated face image.

Once the face is aligned to fit with our anthropometric model, the system works for detecting the centers of facial feature regions (P_3, P_4, P_6, P_7). This process begins with the detection of the midpoint of the two eye centers (P_5) that serves as the reference point for detecting P_6 and P_7, and the distance D_1 that is used as the principal parameter for measuring the location of other facial feature regions. Locations of the points P_3, P_4, P_6, P_7 are then identified by calculating the distances D_2, D_3, D_4, and D5 respectively (Fig. 1.b) using the proportionality constants proposed by our anthropometric face model (Table 1). Rectangular areas for confining the facial feature regions are then approximated using distance between two eyes as the measurement criteria. The areas are kept large enough to cover each of the feature regions completely.

4 Facial Feature Point Detection

Searching for the 18-facial feature points is done separately within each of the areas returned by the facial feature region identifier. Steps for the searching process are described below:

4.1 Feature Point Detection from Eye Region

The eye region is composed of dark upper eyelid with eyelash, lower eyelid, pupil, bright sclera and the skin region that surrounds the eye. The most continuous and non deformable part of the eye region is the upper eyelid, because both pupil and sclera change their shape with various possible situations of eyes; especially when the eye is closed or partially closed. So, inner and outer corner are determined by analyzing the shape of the upper eyelid.

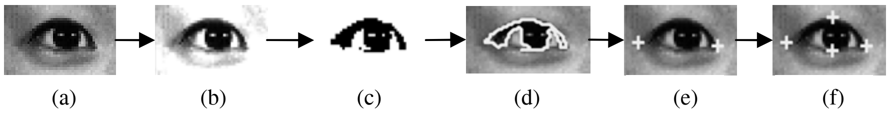

Fig. 3. Detection of feature points from the eye region. (a) Eye region (b) Contrast adjusted eye region (c) Binary version (d) Detected contour (e) Detected eye corners (f) Detected mid-points of upper and lower eyelids

To avoid the erroneous detection, discontinuity in upper eyelid region must be avoided. This can be done by changing the illumination of the upper eyelid in such a way that it differs significantly from its surrounding region. This was done by saturating the intensity values of all the pixels towards zero that constitutes the lower 50% of the image intensity cumulative distribution and forcing the rest of pixels to be saturated towards one (Fig. 3.b). The adjusted image is then converted to binary one (Fig. 3.c) using the threshold obtained from the following iterative procedure [21].

1. Pick an initial threshold value, *t*
2. Calculate the two mean image intensity values in the histogram (m_1 and m_2) below and above the threshold *t*.
3. Calculate new threshold. $t_{new} = (m_1 + m_2) / 2$.
4. If the threshold has stabilized ($t = t_{new}$), this is the appropriate threshold level. Otherwise, *t* become t_{new} and reiterate from step 2.

Contour that cover's the largest area (Fig. 3.d) is then isolated using the 4-connected contour following algorithm specified in [22]. For right eye, the inner eye corner is the right most point of the contour and outer eye corner is the leftmost point of the contour. For left eye, right most point over the contour becomes the inner corner and leftmost point becomes the outer corner. The whole eye contour region is then divided vertically into three equal parts and searching for the upper and lower mid eyelid is then done within the middle portion. For each value of x-coordinate $x_1, x_2, x_3, \ldots\ldots x_n$, that falls within this middle portion, there will be two values of y-coordinate; one from the upper portion of the eye contour $\{ y_{11}, y_{12}, y_{13}, \ldots\ldots y_{1n}\}$ and another from the lower portion of the eye contour$\{ y_{21}, y_{22}, y_{23}, \ldots\ldots y_{2n}\}$. Distance between each pair of points $\{(x_i, y_{1i}), (x_i, y_{2i})\}$ is then calculated. The maximum of the distances calculated from these two sets points and that lies closest to the midpoint of inner and outer eye corner is considered as the amount of eye opening. Mid upper eyelid and mid lower eyelid are simply the points that forms the maximum distance.

4.2 Eyebrow Corner Detection

Aside from the dark colored eyebrow, eyebrow image region also contains relatively bright skin portion. Sometimes, this region is also partially occulted with hair. Since dark pixels are considered as background in digital imaging technology, the original image is complemented to convert the eyebrow region as the foreground object and rest as the background (Fig. 4.b). Morphological image opening operation is then performed over the complemented image with a disk shaped structuring element having radius of 10 pixels for obtaining the background illumination (Fig. 4.c).

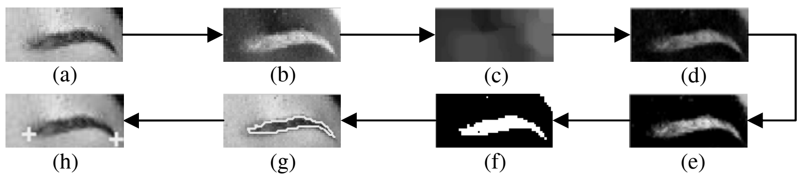

Fig. 4. Eyebrow corners detection (a) Eyebrow region (b) Complemented eyebrow image (c) Estimated background (d) Background Subtraction (e) Intensity adjustment (f) Binary eyebrow region (g) Eyebrow contour (h) detected eyebrow corner

The estimated background is then subtracted from the complemented image to get a brighter eyebrow over a uniform dark background (Fig. 4.d). Intensity of the resultant image is then adjusted on the basis of the pixels' cumulative distribution to increase the discrimination between the foreground and background (Fig. 4.e). We have then obtained the binary version of this adjusted image (Fig. 4.f) by thresholding it using Otsu's method [23] and all the available contours of the binary image are detected using the 4-connected contour-following algorithm specified in [22]. The eyebrow contour, which is usually the largest one, is then identified by calculating the area covered by each contour (Fig. 4.g). For left eyebrow, the point on the contour having the minimum values along x and y coordinates simultaneously, is considered as the inner eyebrow corner. Again, the point which has the maximum values along x and y co-ordinates together, is considered as the outer eyebrow corner (Fig. 4.h). Similar, but reverse procedure is applied for detecting the inner and outer corner of the right eyebrow.

4.3 Nostril Detection

Nostrils of a nose region are the two circular or parabolic objects having the darkest intensity (Fig. 5.a). For detecting the centre points of nostrils, filtering operation on the isolated nose region image is performed using a Laplacian of Gaussian as the filter. The filtered binary image complements the image intensity and thus changes nostrils as the brightest part of the image (Fig. 5.b). Searching for the peak of local maxima is then performed on the filtered image to obtain the centre points of the nostrils. To make the nostril detection technique independent of the size of image, the whole process is repeated varying the filter size starting form 10 pixel, until the number peaks of local maxima is reduced to two (Fig. 5.c).

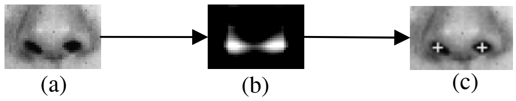

Fig. 5. Nostril detection from isolated nose region (a) Nose region (b) Nose region filtered by Laplacian of Gaussian filter (c) Detected nostrils

4.4 Feature Point Detection from Mouth Region

The simplest case of mouth feature points detection occurs when mouth is normally closed. However, complexities are added to this process when mouth is wide open or teeth are visible between upper and lower lips due to laughter or any other expression.

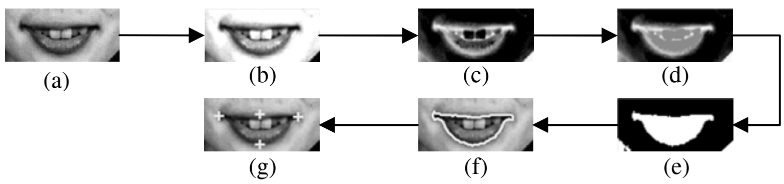

Fig. 6. Feature point detection from mouth region (a) Mouth region (b) Intensity adjustment (c) Complemented mouth region (d) Filled image (e) Binary mouth region (f) Detected mouth contour (g) Detected feature points of mouth region

These two situations provides additional dark and bright region respectively in the mouth contour and makes the feature point detention process quite complex. To handle these problems, contrast stretching on the basis of pixels' cumulative distribution is performed on the image for saturating the upper ½ fractions of the image pixels towards higher intensity value. As a result, lips and other darker region become darker and the skin region becomes brighter providing a clear separation boundary between the foreground and background (Fig. 6.b). A flood fill operation is then performed on the complemented image to fill-up the wholes of mouth region (Fig. 6.d). After this, the resultant image is converted to its binary version using the threshold value obtained using the procedure given in [21]. All the contours are then identified by applying the 4-connected contour following algorithm specified in [22] and mouth contour is isolated as the contour having the largest area (Fig. 6.f). The right mouth corner is then identified as a point over the mouth contour having the minimum x-coordinate value and the point which has the maximum x-coordinate values is considered as the left mouth corner. Middle point ($X_{mid,}$ $Y_{mid,}$) of the left and right mouth corner are then calculated and upper and lower mid points of the mouth are obtained by finding the two specific points over the mouth contour which has the same x-coordinate as that of ($X_{mid,}$ $Y_{mid,}$) but minimum and maximum y-coordinates respectively.

5 Performance Analysis

For measuring the performance of our system, we have tested it on three publicly available face image databases namely, Caltech Face Database [24], BioID Face Database [25], and Japanese Female Facial Expression Database [26].

Table 2. Detection Accuracy (In Percent) of the Automatic Facial Feature Point Detector

Feature Point	Caltech Face Database	BioID Face Database	JAFFE Database	Average Accuracy
Right Eyebrow Inner Corner	95.41	94.16	98.57	96.05
Right Eyebrow Outer Corner	87.36	90.42	92.17	89.98
Left Eyebrow Inner Corner	96.20	93.52	96.38	95.37
Left Eyebrow Outer Corner	88.40	86.26	90.35	88.34
Right Eye Inner Corner	93.12	90.83	94.70	92.88
Right Eye Outer Corner	85.34	87.92	89.62	87.63
Mid Point of Right Upper Eyelid	84.49	86.71	88.4	86.53
Mid Point of Right Lower Eyelid	83.60	85.38	86.73	85.24
Left Eye Inner Corner	95.11	92.64	92.83	93.53
Left Eye Outer Corner	86.69	90.76	91.46	89.64
Mid Point of Left Upper Eyelid	85.77	88.26	89.61	87.88
Mid Point of Left Lower Eyelid	84.22	87.69	88.98	86.96
Right Nostril	97.23	93.19	98.34	96.25
Left Nostril	96.95	91.88	97.21	95.35
Right Mouth Corner	92.79	87.40	95.32	91.84
Left Mouth Corner	94.10	92.45	97.89	94.81
Mid Point of Upper Lip	85.73	83.91	91.20	86.95
Mid Point of Lower Lip	79.31	82.33	86.28	82.64
Average Detection Accuracy of the 18-feature Points over 3-Different Databases				90.44

The Frontal Face Image Database of Caltech consists of 450 face images taken from 27 unique people under different lighting, expressions and backgrounds. The BioID Face Database consists of 1521 gray level images with a resolution of 384x286 pixels. Each image in the database shows the frontal face view of one out of 23 different test persons. The Japanese Female Facial Expression (JAFFE) Database contains 213 images each representing 7 deferent facial expressions (6 basic facial expressions + 1 neutral) posed by 10 Japanese female models. Accuracy of our Automatic Facial

Feature Detector in detecting 18-facial feature points over these databases is summarized in Table 2. As the result indicates, our feature detector performs well both for the neutral face images and well as for the face images having various expressions. On an average, we have been able to detect each of the 18-feature points with a success rate of 90.44% using the developed system.

6 Conclusion

We have presented a completely automatic technique for detecting the 18-facial feature points where localization of the facial feature regions were performed using an Anthropometric Face Model [18]. One of the important features of our system is that rather than just locating each of the feature regions, it identifies the specific 18-featre points over these regions. Our system has been tested over three different face image databases and on an average we have been able to detect each of the 18-facial feature points with a success rate of 90.44%. The proposed technique is independent of the scale of the face image and performs well even at the presence of six basic emotional expressions. Since the system can correct the horizontal rotation of face over x-axis, it works quite accurately in case of horizontal rotation of the face. However the system has limitation in handling vertical rotation of face over y-axis and can perform satisfactorily only when vertical rotation is less then 25 degree. Use of Anthropometric Face Model in the localization of facial feature regions has also reduced the computational time by avoiding the image processing part which is usually required for detecting facial feature regions from a face image. As the distance between the centers of two eyes serves as the principal measuring parameter for facial feature regions localization, improvement in eye center detection technique can also further improve the performance of the whole Automatic Facial Feature Point Detection System.

Acknowledgments. The authors wish to thank the corresponding authorities of the databases for releasing necessary permission regarding their use in this research. This work was supported in part by grants from the NSERC and the Canada Research Chair Foundation.

References

1. Xhang, L., Lenders, P.: Knowledge-based Eye Detection for Human Face Recognition. In: Fourth IEEE International Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies, Vol. 1 (2000) 117–120
2. Rizon, M., Kawaguchi, T.: Automatic Eye Detection Using Intensity and Edge Information. In: Proceedings TENCON, Vol. 2 (2000) 415–420
3. Phimoltares, S., Lursinsap, C., Chamnongthai, K.: Locating Essential Facial Features Using Neural Visual Model. In: First International Conference on Machine Learning and Cybernetics (2002) 1914–1919
4. Yuille, A.L., Hallinan, P.W., Cohen, D.S.: Feature Extraction from Faces Using Deformable Templates. International Journal of Computer Vision, Vol. 8, No. 2, (1992) 99–111

5. Brunelli, R., Poggio, T.: Face Recognition: Features Versus Templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 10 (1993) 1042–1062
6. Herpers, R., Sommer, G.: An Attentive Processing Strategy for the Analysis of Facial Features. In: Wechsler H. et al. (eds.): Face Recognition: From Theory to Applications, Springer-Verlag, Berlin Heidelberg New York (1998) 457–4687.
7. Pardas, M., Losada, M.: Facial Parameter Extraction System Based on Active Contours. In: International Conference on Image Processing, Thessaloniki, (2001) 1058–1061
8. Kawaguchi, T., Hidaka, D., Rizon, M.: Detection of Eyes from Human Faces by Hough Transform and Separability Filter. In: International Conference on Image Processing, Vancouver, Canada (2000) 49–52
9. Spors, S., Rebenstein, R.: A Real-time Face Tracker for Color Video. In: IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 3 (2001) 1493–1496
10. Perez, C. A., Palma, A., Holzmann C. A., Pena, C.: Face and Eye Tracking Algorithm Based on Digital Image Processing. In: IEEE International Conference on Systems, Man and Cybernetics, Vol. 2 (2001) 1178–1183
11. Hsu, R. L., Abdel-Mottaleb, M., Jain, A. K.: Face Detection in Color Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5 (2002) 696–706
12. Xin, Z., Yanjun, X., Limin, D.: Locating Facial Features with Color Information. In: IEEE International Conference on Signal Processing, Vol.2 (1998) 889-892
13. Kim, H.C., Kim, D., Bang, S. Y.: A PCA Mixture Model with an Efficient Model Selection Method. In: IEEE International Joint Conference on Neural Networks, Vol. 1 (2001) 430–435.
14. Lee, H. W., Kil, S. K, Han, Y., Hong, S. H.: Automatic Face and Facial Feature Detection. IEEE International Symposium on Industrial Electronics (2001) 254–259
15. Wilson, P. I., Fernandez, J.: Facial Feature Detection Using Haar Classifiers. Journal of Computing Sciences in Colleges, Vol. 21, No. 4 (2006) 127–133
16. Marini, R.: Subpixellic Eyes Detection. In: IEEE International Conference on Image Analysis and Processing (1999) 496–501
17. Chandrasekaran, V., Liu, Z. Q.: Facial Feature Detection Using Compact Vector-field Canonical Templates. In: IEEE International Conference on Systems, Man and Cybernetics, Vol. 3 (1997) 2022–2027
18. Sohail, A. S. M., Bhattacharya, P.: Localization of Facial Feature Regions Using Anthropometric Face Model. In: I International Conference on Multidisciplinary Information Sciences and Technologies, (2006)
19. Farkas, L.: Anthropometry of the Head and Face. Raven Press, New York (1994)
20. Fasel, I., Fortenberry, B., Movellan, J. R.: A Generative Framework for Real-time Object Detection and Classification. Computer Vision and Image Understanding, Vol.98 (2005) 182–210
21. Efford, N.: Digital Image Processing: A Practical Introduction Using Java. Addison-Wesley, Essex (2000)
22. Ritter, G. X., Wilson, J. N.: Handbook of Computer Vision Algorithms in Image Algebra. CRC Press, Boca Raton, USA (1996)
23. Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 9, No. 1 (1979) 62–66
24. The Caltech Frontal Face Dataset, collected by Markus Weber, California Institute of Technology, USA; available online at: http://www.vision.caltech.edu/html-files/archive.html
25. The BioID Face Database, developed by HumanScan AG, Grundstrasse 1, CH-6060 Sarnen, Switzerland; available online at: http://www.humanscan.de/support/downloads/facedb.php
26. Lyons, J., Akamatsu, S., Kamachi, M., Gyoba, J.: Coding Facial Expressions with Gabor Wavelets. In: Third IEEE International Conference on Automatic Face and Gesture Recognition (1998) 200–205.