Image Preprocessing for Illumination Invariant Face Verification

Mariusz Leszczyński

Institute of Radioelectronics, Warsaw University of Technology, Warsaw, Poland

Abstract—Performance of the face verification system depend on many conditions. One of the most problematic is varying illumination condition. In this paper 14 normalization algorithms based on histogram normalization, illumination properties and the human perception theory were compared using 3 verification methods. The results obtained from the experiments showed that the illumination preprocessing methods significantly improves the verification rate and it's a very important step in face verification system.

Keywords—DLDA, face verification, histogram normalization, homomorphic filtering, illumination normalization, LDA, PCA, preprocessing techniques, quotient image, retinex.

1. Introduction

Face is one of the most commonly used by people to recognize each other. Over the course of its evolution, the human brain has developed highly specialized areas dedicated to the analysis of the facial images [1]. In the past decades, face recognition has been an active research area and many types of algorithms and techniques has been proposed to equal this ability of human brain. It is however questioned whether the face itself is a sufficient basis for recognizing, a person from large population with great accuracy. Indeed, the human brain also relies on many contextual information and operate on limited population.

The most problematic perturbation affecting the performance of face recognition systems are strong variations in pose and illumination. Variation between images of different faces in general is smaller than taken from the same face in a variety of environments [2]. In face verification system authenticates a person's claimed identity and decide that claimed identity is correct or not. In this case we have limited user group and in the most cases we can forced or demand frontal pose orientations. Unfortunately we still have problems with illumination condition. Face recognition tests [3]–[6] revealed that the lighting variant is one of the bottlenecks in face recognition/verification. If lighting conditions are different from the gallery identity decision is wrong in many cases.

There are two approaches to this problem. Model-based [7], [8] and preprocessing-based. Model-based attempt to model the light variation. Unfortunately, this requires large amount of training data and sometimes fall when we have complicated lighting configuration.

The second approach using preprocessing methods to remove lighting influence effect without any additional knowledge. In this paper, we compare 14 normalization algorithms using 3 verification methods.

2. Histogram Normalization

Illumination preprocessing on 2D images can be divided into two groups: histogram transformation and photometric normalization.

2.1. Histogram Equalization (HQ)

Histogram normalization is one of the most commonly used methods. In image processing, the idea of equalizing a histogram is to stretch and/or redistribute the original histogram using the entire range of discrete levels of the image, in a way that an enhancement of image contrast is achieved. The most common used histogram normalization technique is histogram equalization where one attempts to change the image histogram into a histogram that is constant for all brightness values. This would correspond to a brightness distribution where all values are equally probable. For image $I(x,y)$ with discrete k gray values histogram is defined by:

$$p(i) = \frac{n_i}{N}, \tag{1}$$

where: $i \in 0, 1, \ldots, k-1$ grey level and N is total number of pixels in the image.

Transformation to a new intensity value is defined by:

$$i_{out} = \sum_{i=0}^{k-1} \frac{n_i}{N} = \sum_{i=0}^{k-1} p(i). \tag{2}$$

Output values are from domain of $[0,1]$. To obtain pixel values in to original domain, it must be rescaled by the $k-1$ value.

Figure 1 shows two face images with different light condition and preprocessed images with corresponding histograms.

2.2. Histogram Truncation and Stretching (HT)

Histogram stretching consists in distributing the pixel appearance frequencies over the entire width of the histogram.
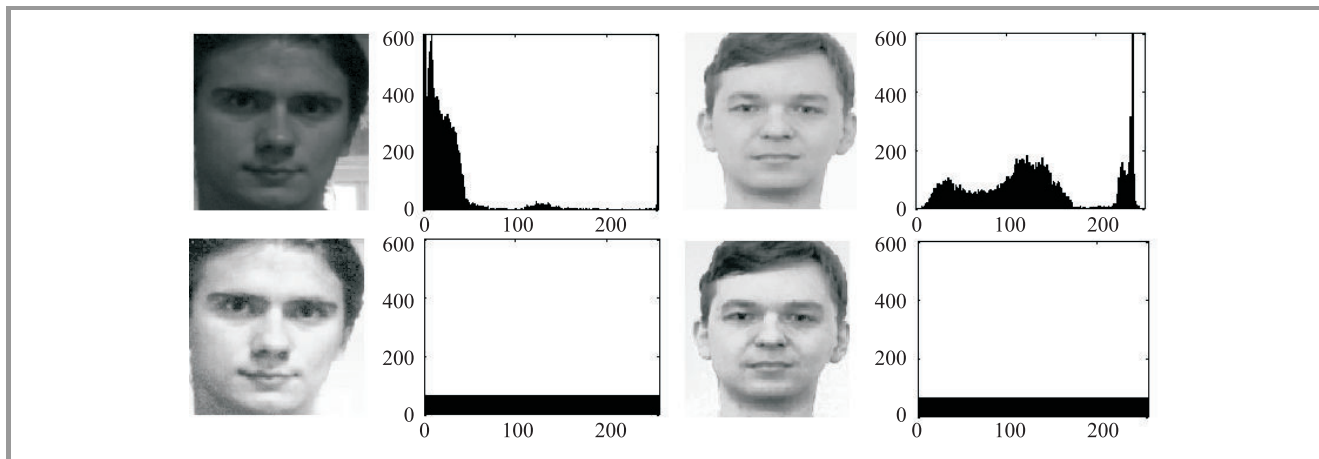
Fig. 1. Two sample images with histogram before (upper) and after (lower).
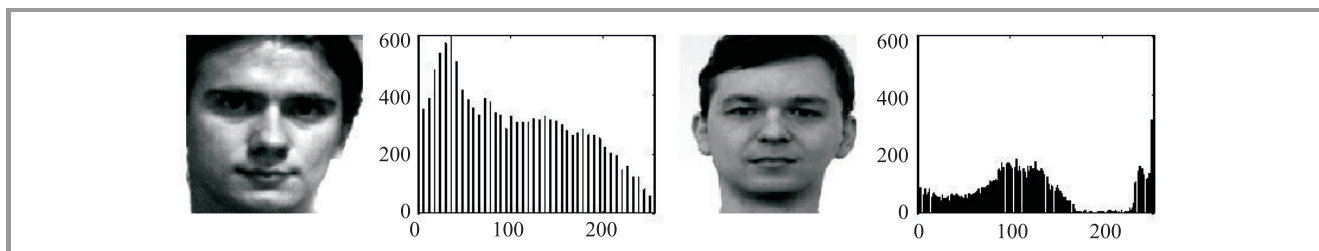

Fig. 2. Two sample images with histogram after histogram truncation and stretching.

Thus, it is an operation that consists in modifying the histogram in such a way as to distribute the intensities on the scale of values available as well as possible. This amounts to extending the histogram so that the value of the lowest intensity is zero and that of the highest is the maximum value. In this way, if the values of the histogram are very close to each other, the stretching will make it possible to provide a better distribution in order to make light pixels even lighter and dark pixels closer to black. Additional ten percentage of the lower and upper ends of an image histogram are truncated. This solves the problem when few very bright or dark pixels have the overall effect of darkening or brightening the rest of the image after rescaling (see Fig. 2).

2.3. Histogram Modeling

Histogram preprocessing is not only limited for HQ and stretching. We can model it with different density function where brightness distribution will be mapped to a specific probability distribution. According to [9], [10] general mapping function for the distribution function $f(x)$ may be calculated from:

$$\frac{N-R+0.5}{N} = \int_{x=-\infty}^{t} f(x)dx, \qquad (3)$$

where: R is rank of the pixels ordered from smallest intensity to the largest intensity value with assigned rank from 1 to N.

The right side of Eq. (3) represents target cumulative distribution function (CDF). The searching t parameter will be computed by from the inverse CDF of the left side Eq. (3).

2.4. Normal Distribution (ND)

The first consider distribution is normal distribution, which is the most commonly observed probability distribution. It was first described by De Moivre in 1733. Laplace used the normal curve in 1783 to describe the distribution of errors. Subsequently, Gauss used the normal curve to analyze astronomical data in 1809. The normal curve is often called the Gaussian distribution and its defined by the following equation:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \qquad (4)$$

where: μ is the mean and the second, σ is the standard deviation.
In our experiments (Fig. 3) we use the standard normal distribution, where $\mu = 0$ and $\sigma^2 = 1$.

2.5. Lognormal Distribution (LN)

The lognormal distribution is an asymmetric distribution. Many physical, chemical, biological, toxicological, and statistical processes tend to create random variables that follow
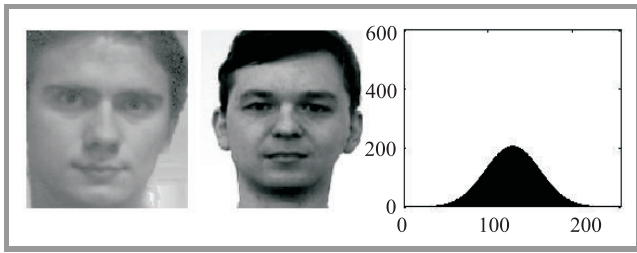
Fig. 3. Two sample images with histogram after mapping the histogram to a normal distribution.

lognormal distributions. For example, lognormal distributions can model certain instances, such as the change in price distribution of a stock or survival rates of cancer patients or failure rates in product tests.
Density function for this distribution is defined by:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(\frac{-(\ln x - \mu)^2}{2\sigma^2}\right), \qquad (5)$$

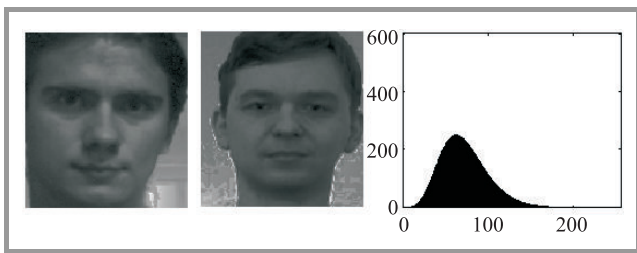for our experiments (Fig. 4) mean $\mu = 0$ and standard deviation $\sigma = 0.25$.

Fig. 4. Two sample images with histogram after mapping the histogram to a lognormal distribution.

2.6. Extreme Value Distribution (EV)

The third distribution is called extreme value distribution and appropriate for modeling many rare events, and has the following probability density function:

$$f(x) = \sigma^{-1} \exp\left(\frac{x - \mu}{\sigma}\right) \exp\left(-\exp\left(\frac{x - \mu}{\sigma}\right)\right), \qquad (6)$$

where: μ is the location parameter, and σ is the distribution scale (set to 0 and 1 in the experiments, see Fig. 5).
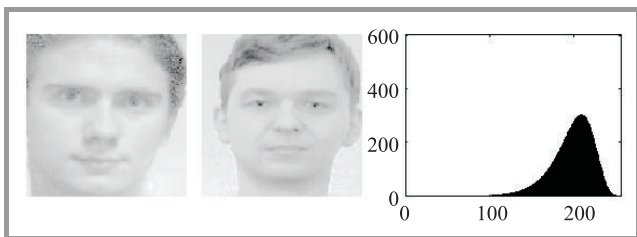

Fig. 5. Two sample images with histogram after mapping the histogram to a extreme value distribution.

2.7. Exponential Distribution (EN)

The exponential distribution is a commonly used distribution in reliability engineering. Density function for this distribution is defined by:

$$f(x) - \frac{1}{\beta} \exp\left(\frac{-x}{\beta}\right), \qquad (7)$$

where β is the scale parameter. In our experiments (Fig. 6) we use the standard exponential distribution, where $\beta = 1$.
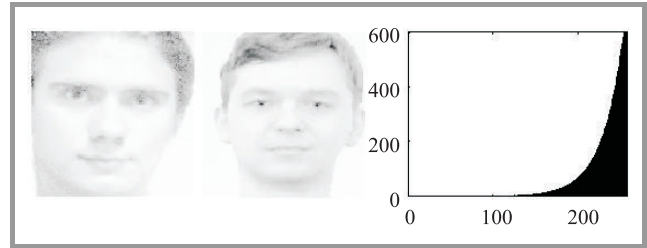

Fig. 6. Two sample images with histogram after mapping the histogram to a exponential distribution.

3. Photometric Normalization

The second approach for illumination normalization is based on human perception theory and illumination properties.

3.1. Single Scale Retinex (SSR)

In 1971 Land and McCann introduce the idea that image $I(x,y)$ is the product of two components, illumination $L(x,y)$ and reflectance $R(x,y)$ [11].

$$I(x,y) = L(x,y) R(x,y). \qquad (8)$$

Illumination contains geometric properties of the scene (i.e., the surface normals and the light source position) and reflectance contains information about the object. Based on the assumption that the illumination varies slowly across different locations of the image and the local reflectance may change rapidly across different location, the processed illumination should be drastically reduced due to the high-pass filtering, while the reflectance after this filtering should still be very close to the original reflectance. The reflectance can be also finding by dividing the image by the low pass version if the original image, which is representing illumination components.

Land proposed a technique called retinex, which is a combination of the words retina and cortex. Its try to explain model of the human visual system. The most interesting point for illumination normalization is the assumption, that perception depends on the relative or surrounding illumination. It means that reflectance $R(x,y)$ equals the quotient of image $I(x,y)$ and the illumination $L(x,y)$ calculated by the neighborhood of $I(x,y)$. It improves the visibility of

dark object while maintaining the visual different of the light area.

Single scale retinex algorithm proposed by Jobson and Woodell [12] defines a Gaussian kernel to estimate the neighborhood illumination. Additional the logarithmic transformation is employed to compress the dynamic range. Reflectance image is takes from the form:

$$R_{SSR}(x,y) = \log I(x,y) - \log\left[F(x,y) * I(x,y)\right], \quad (9)$$

where: $*$ denotes the convolution operation and $F(x,y)$ is the surround Gaussian function.

Figure 7 shows two sample face images received from single scale retinex.

Fig. 7. Two sample images received from SSR.

3.2. Mutli Scale Retinex (MSR)

Rahman [13] improved previous method by estimating illumination as a combination of several weighting (ω_n) Gaussian filters with different scales (N). Reflectance image is defined by:

$$R_{MSR}(x,y) = \sum_{n=1}^{N} \omega_n \left\{ \log I(x,y) - \log\left[F(x,y) * I(x,y)\right] \right\}. \quad (10)$$

Two sample face images received from multi scale retinex are shown in Fig. 8.

Fig. 8. Two sample images received from MSR.

3.3. Adaptive Single Scale Retinex (ASR)

ASR was presented by Park in [14]. The proposed method estimates illumination by iteratively convolving the input image with a 3×3 smoothing mask weighted by a coefficient via combining two measures of the illumination discontinuity at each pixel, see Fig. 9.

Fig. 9. Two sample images received from ASR.

3.4. Homomorphic Filtering (HOMO)

Homomorphic filtering [15] using the same properties as previous methods, that reflectance is connected with high frequency. In this case high-pass filtering is performed in frequency domain using Fourier transform. The processed image can be found by following equation:

$$I' = e^{\text{Re}\left(IFT\left(FT(\log I) * H\right)\right)}, \quad (11)$$

where: H is a high-pass Butterwoth's filter, FT – Fourier transform, IFT – inverse Fourier transform.

In Fig. 10 are shown two sample images received from homomorphic filtering.

Fig. 10. Two sample images received from HOMO.

3.5. Single Scale Self Quotient Image (SSQ)

The self quotient image was developed by Wang [16] in 2004 and is based on Land's human vision model. From Eq. (8) it can be derived that the reflectance is given by:

$$I(x,ky)\frac{1}{L(x,y)} = R(x,y). \quad (12)$$

Because illumination can be consider as the low frequency component then, it can be estimated as:

$$L(x,y) \approx F(x,y) * I(x,y), \quad (13)$$

with $F(x,y)$ is a low pass filter.

From Eqs. (12) and (13) the self quotient image $Q(x,y)$ is defined as:

$$Q(x,y) = \frac{I(x,y)}{F(x,y) * I(x,y)} \approx R(x,y). \quad (14)$$

Two sample face images received from single scale self quotient are shown in Fig. 11.

Fig. 11. Two sample images received from SSQ.

3.6. Multi Scale Self Quotient Image (MSQ)

Properties of the previous $Q(x,y)$ are dependent on the kernel size of filter $F(x,y)$. If it will be to small than $Q \approx 1$ and all reflectance information will be lost. On the other hand if kernel size will be too large then will appear halo effects near edges. To avoid this problems Wang propose multi scale approach where:

$$Q(x,y) = \sum_{k=1}^{n} m_k T\{Q_k(x,y)\}, \qquad (15)$$

where: m_k are weighting factors, T is nonlinear function and Q_k are quotient images corresponding to k scale.

$$Q_k(x,y) = \frac{I(x,y)}{\left(\frac{1}{N}W_k G_k\right)*I(x,y)}, \qquad k = 1,\ldots,n, \qquad (16)$$

where: N is normalization factor, $W_k G_k$ are weighted Gaussian kernels.

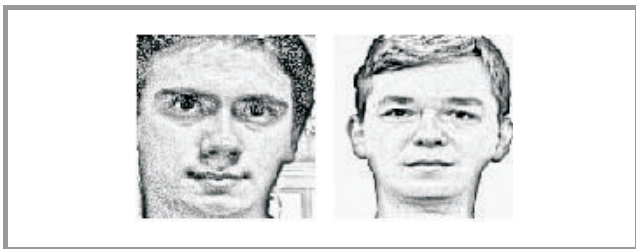Figure 12 shows two sample face images received from multi scale self quotient.

Fig. 12. Two sample images received from MSQ.

3.7. DCT-Based Normalization (DCT)

This technique [17] is based on fact that illumination can be consider as the low frequency component. First image

Fig. 13. Two sample images received from DCT-based normalization.

is transform into frequency domain using discrete cosine transform (DCT) and then some number of DCT coefficients are sets to zero. This removes some of the low-frequency information contained in the images and reduce illumination influence (see Fig. 13). Target image is obtained after applying inverse discrete cosine transform (IDCT).

3.8. Wavelet-Based Normalization (WAV)

Next method combine two approaches based on histogram normalization and illumination low frequency properties. In the first step discrete wavelet transform is used to decompose the facial image into approximation, horizontal, vertical and diagonal components. The approximation components represents low level image components. Next equalizes the histogram of the approximation coefficients matrix. As a final step it performs an inverse wavelet transform to recover the normalized image.

In Fig. 14 are shown two sample face images received after wavelet-based normalization.

Fig. 14. Two sample images received after wavelet-based normalization.

4. Feature Discrimination

Biometric pattern verification is conceptually different from traditional class membership verification. This is involving with following terms:

1. We deal always with a subset of the whole collection of classes.

2. The number of classes used in training time of recognition system is small and usually different from classes which are recognized in exploiting time.

Since natural human centered pattern classes cannot be used in person verification biometric systems, another categorization has to be sought. It appears that differences of human features for the biometric measurements of the same person (within-class differences) and for different persons (between-class features) create a consistent categorization including two specific classes. The specificity of this two classes follows from the fact that means of these two classes are both equal to zero. Moreover, for the within-class feature variation (var_w) could be sometimes greater than between-class feature variation (var_b), i.e., usually the squared within-class errors are of the same magnitude as

squared between-class errors. Therefore, it is natural to look for such a linear transformation $W : R^N \rightarrow R^n$ of original measurements $x \in R^N$ (e.g., vectorized pixel matrix of face image or its 2D frequency representation) into a target feature vector $z = W^t x$ for which intra-class differences are decreased while inter-class differences are increased. This is the problem of the classical linear discriminant analysis (LDA) [18]. However from the previous works described in [19] it is already known that in case of face verification the dual linear discrimination analysis (DLDA) leads to better results than the optimization of Fisher ratio (LDA). Difference between DLDA and Fisher LDA if a way we founding optimal W:

$$
\begin{aligned}
\text{Fisher } LDA \quad & W = \arg\max_w \frac{\text{var}_b(Z)}{\text{var}_w(Z)}, \\
\text{dual } LDA \quad & W = \arg\min_w \frac{\text{var}_w(Z)}{\text{var}_b(Z)},
\end{aligned}
\tag{17}
$$

where: $Z = [z_1, \ldots, z_L]$ and L number of images. In our next experiments we are using LDA, DLDA and oldest method principal components analysis [20]. PCA-based face recognition method was proposed in [21] and became very popular. Using PCA method we find a subset of principal directions (principal components) in a set of the training faces. Then like in LDA we project faces into the space of these principal components and get the feature vectors.

5. Experimental Results

The experiments are carried out on normalized images taken from the following databases (Fig. 15):

- Altkom (80 persons 1680 images),
- Banca (52 persons 474 images),
- Valid (106 persons 1575 images),
- WUT database (143 persons 769 images).

Which gives 391 persons with 4525 images. Picture from this databases were taken in different light conditions and except the Altkom database in some time interval. According to [22] images are normalized to the size 46×56 based on fixed eye center position.

To quantify verification performance we are using receiver operating characteristic (ROC). This characteristic shows the tradeoff between two types of verification's errors false rejection error against false acceptance error. To more clarity presentation in Table 1 we show only the equal error rate (EER), which is the value were false rejection and false acceptance errors are equal. Based on the results we can conclude that except homomorphic filtering all compared methods gives verification improvements, especially using DLDA as the discriminative algorithm. The best results was conducted using multi scale quotient images where we get 38% less errors.
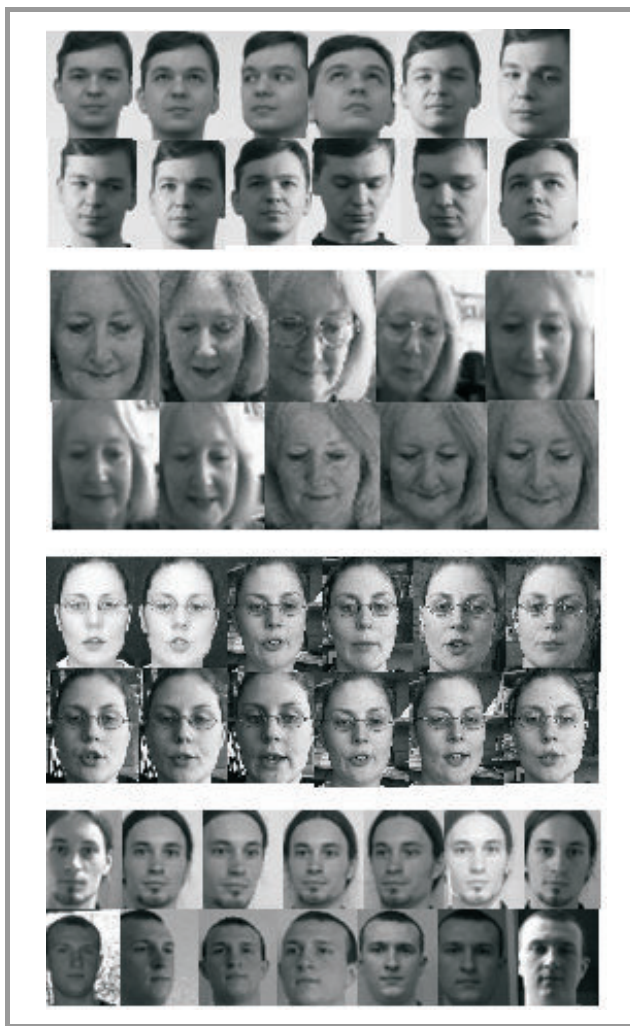

Fig. 15. Face databases – from upper Altkom, Banca, Valid, WUT.

Table 1
Performance comparison of different normalization methods

	EER		
	OCA	LDA	DLDA
ORG	0.2138	0.2571	0.2173
HQ	0.1737	0.1868	0.166
HT	0.1935	0.1976	0.1656
ND	0.1681	0.1746	0.164
LN	0.1772	0.2058	0.1857
EV	0.1585	0.1857	0.1521
EN	0.1535	0.1845	0.165
SSR	0.1941	0.2178	0.1896
MSR	0.1987	0.2073	0.1883
ASR	<u>0.1423</u>	0.1898	0.1458
HOMO	0.3464	0.3376	0.26
SSQ	0.1566	0.1957	0.15
MSQ	0.1494	0.1865	**0.1346**
DCT	0.226	0.2287	0.1854
WAV	0.1813	<u>0.17</u>	0.1435

6. Conclusion

In this paper we analyze 14 illumination invariant algorithms. The performance of the presented methods were compared on database contains 4525 images of 391 persons taken in different light conditions. The results obtained from the experiments showed that the illumination preprocessing methods significantly improves the verification rate. The best results were achieved using human perception related MSQ algorithm, with 38% less verification errors compared to the same DLDA discriminant method with using not preprocessed images. Very promising seems to be combination of both analyzing approaches (histogram and photometric normalization) as a field to future work.

References

[1] N. Kanwisher, J. McDermott, and M. Chun, "The fusiform face areas: a module in human extrastriate cortex specialized for face perception", *J. Neurosci.*, vol. 17, no. 11, pp. 4302–4311, 1997.

[2] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for illumination changes", *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 19, pp. 721–732, 1997.

[3] P. J. Philips, H. Moon, S. A. Rizvi and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms", *Trans. PAMI*, vol. 22, pp. 1090–1104, 2000.

[4] P. J. Philips, P. Grother, R. Micheals, D. M. Blackburn, E. Tabassi and M. Bone, "Face recognition vendor test 2002: evaluation report", Tech. Rep. NISTIR 6965, National Institute of Standards and Technology, 2003.

[5] P. J. Philips, P. J. Flynn, T. Scruggs, K. W. Bowyer and W. Worek, "Preliminary Face Recognition Grand Challenge Results", in *Proc. 7th Int. Conf. Autom. Face Gesture Recogn.*, Southampton, UK, 2006, pp. 15–24.

[6] P. J. Philips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott and M. Sharpe, "FRVT 2006 and ICE 2006 large-scale results", NISTIR 7408, National Institute of Standards and Technology, 2007.

[7] A. S. Georghiades, P. N. Belhumeur and D. J. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose", *IEEE Trans. Patt. Anal. Mach. Intell.*, pp. 643–660, 2001.

[8] M. Vasilescu, D. Terzopoulos, "Multilinear analysis of image ensembles: tensor-faces", in *Proc. Eur. Conf. Comp. Vision ECCV 2002*, Copenhagen, Denmark, 2002, pp. 447–460.

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover, 1964.

[10] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*. 2nd ed., New York: Wiley, 1993.

[11] E. Land and J. McCann, "Lightiness and retinex theory", *J. Opt. Soc. America*, vol. 61, pp. 1–11, 1971.

[12] D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex", *IEEE Trans. Image Process*, vol. 6, no. 3, pp. 451–462, 1997.

[13] Z. Rahman, G. Woodell, D. Jobson, "A Comparison of the Multiscale Retinex with other Image Enhancement Techniques", in *Proc. 50th IS&T Anniv. Conf.*, Boston, USA, 1997.

[14] Y. K. Park, S. L. Park, and J. K. Kim, "Retinex method based on adaptive smoothing for illumination invariant face recognition," *Sig. Proces.*, vol. 88, no. 8, pp. 1929–1945, 2008.

[15] R. Gonzalez and R. Woods, *Digital Image Processing*. 2nd ed. Boston: Addison-Wesley Longman, 1992.

[16] H. Wang, S. Z. Li and Y. Wang, "Face recognition under varying lighting condition using self quotient image", in *Proc. 6th IEEE Int. Conf. Autom. Face Gesture Recog.*, Seoul, Korea, 2004, pp. 819–824.

[17] Štruc V., Pavešić, N. "Performance evaluation of photometric normalization techniques for illumination invariant face recognition", in *Advances in Face Image Analysis: Techniques and Technologies*, Y. J. Zhang, Ed. Hershey: IGI Global, 2010.

[18] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. San Diego: Academic Press, 1992.

[19] W. Skarbek, K. Kucharski, and M. Bober, "Dual LDA for face recognition", *Fundamenta Informaticae*, vol. 61, 303–334, 2004.

[20] K. Pearson, "On lines and planes of closest fit to systems of points in space", *Philosophical Mag.*, vol. 2, no. 6, pp. 559–572, 1901.

[21] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces", in *Proc. IEEE Comp. Soc. Conf. Comp. Vision Patt. Recogn. CVPR'91*, Maui, USA, 1991, pp. 586–591.

[22] ISO/IEC 15938-3:2002 (2002).

Mariusz Leszczyński was born in Kozienice, Poland, on April 1979. He received B.Sc. in radiocommunication and multimedia engineering and M.Sc. in telecomunications degrees from the Warsaw University of Technology (WUT) in 2002 and 2004, respectively. Currently he is a Ph.D. student. His research interest include image segmentation, image processing and face recognition and verification.
e-mail: mariuszleszczynski@gmail.com
Institute of Radioelectronics
Warsaw University of Technology
Nowowiejska st 15/19, 00-665 Warsaw, Poland