DETECTION AND RECOVERY OF OCCLUDED FACE IMAGES BASED ON CORRELATION BETWEEN PIXELS

Ji-eun Lee and Nojun Kwak

Department of Electrical & Computer Engineering, Ajou University, Suwon, 443-749, South Korea

Keywords: Correlation coefficient, Occluded face, Occlusion detection, Face image recovery.

Abstract: In this paper, we propose a method to detect and recover the occluded parts of face images using the correlation between pairs of pixels. In the training stage, correlation coefficients between every pairs of pixels are calculated using the occlusion-free training face images. Once a new face image is shown, the occluded area is detected and recovered using correlation coefficients obtained in the training stage. We compare the performance of the proposed method with the conventional method based on PCA. The results show that the proposed method detects and recovers occluded area with much smaller noises than the conventional PCA based method.

1 INTRODUCTION

The problem of detecting and recovering the occluded parts of a face image is a very important task to solve. There has been works on detecting and recovering the occlusion of a face and the representative approach is to apply PCA (principal component analysis). These methods includes automatic eyeglasses removal from face images (Wu, 2004), reconstruction of the occluded parts by fast recursive PCA (Wang, 2007), application of probabilistic PCA (XiaoFeng, 2010), image completion (Efros, 1999; Sun, 2005; Komodakis, 2006) and so on (Lin, 2007).

The conventional PCA based methods use a weight matrix composed of the eigenvectors of the training data consisting of non-occluded face images to detect and recover the occluded parts.

In this paper, we propose to use correlation between pixels. The proposed algorithm can be outlined as follows. Firstly, the proposed method calculates the correlation coefficients between every pairs of pixels using the training images. Secondly, we estimate the pixel values of occluded parts of each test images using the correlation coefficients. Then we detect the occluded parts and recover them until the difference between the reconstructed and the occluded images are small.

The paper is organized as follows. In Section 2, the method based on correlation between pairs of pixels is proposed. Then we will compare the result of the proposed method with that of the conventional PCA based method through experiments in Section 3. Finally, conclusions follow in Section 4.

2 THE PROPOSED METHOD

2.1 Correlation Coefficient

The basic idea of the proposed method is that the pixel values of the occluded part can be predicted from the pixel values of the non-occluded part which are highly related to the ones in question.

Suppose each pixel value be a random variable. Then the relationship between a pair of pixels can be measured by the correlation coefficient.

The correlation coefficient ρ_{ij} between two pixels x_i and x_j is defined as

$$\rho_{ij} = \frac{C_{ij}}{\sigma_i \sigma_j} = \frac{E[(x_i-\mu_i)(x_j-\mu_j)]}{\sqrt{E[(x_i-\mu_i)^2]}\ \sqrt{E\left[\left((x_j-\mu_j)^2\right)\right]}}. \quad (1)$$

Here, C_{ij} is the covariance between two random variables and σ_i, σ_j are the standard deviations of each random variable. The μ_i and μ_j are the mean values of x_i and x_j respectively and $E[\cdot]$ is the expectation operation. If the correlation coefficient is close to 1 or -1, those two random variables are much related to each other and if it is close to 0, those two are hardly related at all.

Lee J. and Kwak N. (2012).
DETECTION AND RECOVERY OF OCCLUDED FACE IMAGES BASED ON CORRELATION BETWEEN PIXELS.
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, pages 568-571
DOI: 10.5220/0003772805680571

Normally, the expectation of random variables cannot be calculated without knowing the underlying joint distribution. However, it can be estimated by the samples and the estimated value of the correlation coefficient (1) using N samples can be calculated as follows.

$$\hat{\rho}_{ij} = \frac{\sum_{n=1}^{N}(x_{in} - \overline{x_i})(x_{jn} - \overline{x_j})}{\sqrt{\sum_{n=1}^{N}(x_{in} - \overline{x_i})^2}\sqrt{\sum_{n=1}^{N}(x_{jn} - \overline{x_j})^2}} \quad (2)$$

Here, x_{in} is the i^{th} pixel value of the n^{th} training image and $\overline{x_i}$ and $\overline{x_j}$ are i^{th} and j^{th} sample mean of pixel values respectively.

Figure 1 shows correlation coefficient maps of a face. The pixel marked X is the reference pixel (cheek and eye). Looking at the correlation coefficient map of an eye, we can see that the pixels in the right eye are highly related to the one in the left eye. This shows that in addition to the local pixels, far away pixels from the reference point can also provide important information in estimating the pixel value of the reference point.

2.2 Jointly Gaussian Distribution

In this part, we present a way to estimate an unknown pixel value from the other known ones.

Assume that the pixel value of x_i is not known, while that of x_j is known. We also assumed that pixel i and j follows jointly Gaussian distribution. Then, the pixel value of x_i can be obtained using a set of pixels $S_i = \{\, j \mid$ indexs of m number of x_j which are highly related to $x_i\}$ that are highly related to x_i.

When $|S_i| = 1$ and the j^{th} pixel value is ν, the conditional probability of x_i given x_j is

$$P(x_i|x_j = \nu) = \frac{P(x_i, x_j = \nu)}{P(x_j = \nu)}. \quad (3)$$

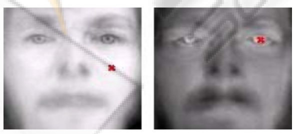

Figure 1: Correlation coefficient map (cheek and eye).

Because $P(x_j = \nu)$ is a constant, $P(x_i|x_j = \nu)$ is proportional to $P(x_i, x_j = \nu)$. With the assumption that the pixel i and j follows jointly Gaussian distribution, $P(x_i, x_j = \nu)$ becomes

$$P(x_i, x_j = \nu) \propto \exp\left(-\frac{1}{2}\begin{pmatrix}x_i\\ \nu\end{pmatrix}^T \Sigma^{-1}\begin{pmatrix}x_i\\ \nu\end{pmatrix}\right) = \exp\left(-\frac{1}{2}\frac{1}{(1-\rho_{ij}{}^2)\sigma_i{}^2}\left(x_i - \frac{\sigma_i}{\sigma_j}\rho_{ij}\nu\right)^2\right) \quad (4)$$

Here, Σ^{-1} is an inverse of the covariance matrix Σ of x_i and x_j.

$$\Sigma^{-1} = \frac{1}{(1-\rho_{ij}{}^2)\sigma_i{}^2\sigma_j{}^2}\begin{bmatrix}\sigma_i{}^2 & -\rho_{ij}\sigma_i\sigma_j\\ -\rho_{ij}\sigma_i\sigma_j & \sigma_j{}^2\end{bmatrix} \quad (5)$$

Using (4), the conditional mean and the variance of x_i are as follows:

$$E[x_i|x_j = \nu] = \frac{\sigma_i}{\sigma_j}\rho_{ij}\nu \quad (6)$$

$$\mathrm{Var}[x_i|x_j = \nu] = (1-\rho_{ij}{}^2)\sigma_i{}^2. \quad (7)$$

In case of $|S_i| = 2$ and assuming that the value of x_j is ν and x_k is ω, we can estimate the value of x_i using a weighted sum of $E[x_i|x_j = \nu]$ and $E[x_i|x_k = \omega]$ because $E[x_i|x_j = \nu, x_k = \omega]$ is hard to represent.

$$E[x_i|x_j = \nu, x_k = \omega] \simeq \alpha E[x_i|x_j = \nu] + \beta E[x_i|x_k = \omega] \quad (8)$$

In the above equation α and β should meet the condition $\alpha + \beta = 1$. The weight α can be regarded as the contribution of x_j in estimating the value of x_i. If the conditional variance of x_i given x_j decreases, the confidence of the estimation grows and the weight α should be increased. Therefore, the weight is set to be proportional to 1 divided by the standard deviation which is the square root of (7) in this paper. For example, when $|S_i| = 2$, $\alpha \propto 1/\sqrt{1-\rho_{ij}}$, $\beta \propto 1/\sqrt{1-\rho_{ik}}$ and it is made that $\alpha + \beta = 1$.

In case of $|S_i| > 2$, the weights are set in the similar way. If x_i and x_j are highly correlated to each other, we can assume high confidence resulting in a high weight. On the contrary, if the correlation between x_i and x_j are small, the corresponding weight is set to be small. In this paper, to alleviate computational complexity, we only make use of the pixels that have the highest correlation coefficient in estimating the unknown pixel value.

2.3 Detection and Recovery of Occluded Face Images

Obtaining Correlation Coefficient: The first step is to get the correlation coefficient for all the pixels using the non-occluded training face images X_{tr}. This is done by using (2). After getting the correlation coefficient of all the pairs of pixels, for each pixel, a list of highest correlation coefficient are stored.

Reconstruction: Assume that we do not know the first pixel value of the test face image which are centerized by the mean image of training data. The pixel value can be predicted using a set of pixels which are highly correlated with the one in question.

The method to calculate each pixel value x_i with the condition $|S_i| = 2$ is shown from (6) to (8). This is generalized to a case of $|S_i| = m$ as follows:

$$w_j = \frac{1}{\sqrt{1 - {\rho_{ij}}^2}} \tag{9}$$

$$\begin{aligned} z_i &= w_1 \frac{\sigma_i}{\sigma_1} \rho_{i1} x_1 + \cdots + w_m \frac{\sigma_i}{\sigma_m} \rho_{im} x_m \\ &= \sum_{j=1}^{m} w_j \frac{\sigma_i}{\sigma_j} \rho_{ij} x_j \end{aligned} \tag{10}$$

$$x_i' = \bar{x} + \frac{z_i}{\sum_{j=1}^{m} w_j} \tag{11}$$

Here, w_j means weight (or confidence) of x_j derived from the correlation coefficient between i^{th} and j^{th} pixels. The pixels x_j $(j = 1, \cdots, m)$ are the pixels that have the highest correlation coefficient with the i^{th} pixel.

Occlusion Detection: The reconstructed face image x' and the original test face image x are compared to decide whether a pixel is occluded or not. This is achieved by checking the difference between pixel values of x and x' at the same pixel is more than a threshold ϵ or not. This procedure is described in the following equation:

$$|x_i' - x_i| \begin{array}{l} < \epsilon \quad \Rightarrow \quad i \in O^c \\ \geq \epsilon \quad \Rightarrow \quad i \in O. \end{array} \tag{12}$$

Here, i denotes the specific location of the pixel and x_i and x_i' are i-th pixel values of x and x' respectively. The set O is the set of occluded pixels and O^c is the complement of O which corresponds to the set of non-occluded pixels.

Recovery: After detecting occluded parts from the face image, occluded parts are filled with the reconstructed image x' while non-occluded parts remains the same as the original image x. The resulting recovered image x'' is as follows:

$$x_i'' = \begin{cases} x_i & \text{if} \quad i \in O^c \\ x_i' & \text{if} \quad i \in O \end{cases} \tag{13}$$

After the recovery stage, x is replaced with x'' and the process is repeated until the difference between x and x'' becomes small enough.

3 EXPERIMENTS

3.1 BioID Data

BioID data contains 1521 gray-scale face images. It consists of frontal face images of 23 people and the size of each image is 100 by 100. Among 1521 images 1000 images were used for training and the other 521 images were used for testing.

We covered each test image with a squared box which has a random height and width on a random position to make an occlusion like Figure 3.

3.1.1 Conventional PCA Based Method

The first row of figure 3 shows the difference of pixel values between the original occluded image and the recovered image. We represent the occluded parts as value 1 and the non-occluded parts as value 0. Final recovered images can be seen in second row of figure 3. As a result of the detection and recovery, we can see a lot of noise in the figures.

3.1.2 The Proposed Method

The first row of figure 4 shows the detected occluded area. In the figure, we can see some noise around the occluded parts, but the noise far from and within the occluded parts is not high compared to Figure 3. The second row of figure 4 shows the recovered image. As a result of recovery, we can see the continuity between the occluded parts and the non-occluded parts is enhanced resulting in smoother images than those in Figure 3.

Figure 2: Occluded face images for BioID data.

Figure 3: Detected and recovered face images for BioID data using the conventional PCA based method.

Figure 4: Detected and recovered face images for BioID data using the proposed method.

Table 1: MAE and MSE between recovered BioID data and non-occluded BioID data.

	MAE	MSE
PCA based	17052.45	788.06
Correlation based	10500.65	532.22

Table 2: Number of iterations and processing time for BioID data.

	# of iteration		Processing time	
	Avg.	Var.	Avg.	Var.
PCA based	3.117	1.226	0.092	0.051
Proposed	5.633	3.820	1.173	0.421

3.1.3 Numerical Comparison

The MAE and MSE are shown on the Table 1. As we expected, the proposed method showed less error than the conventional PCA-based method.

Both the number of iterations and the processing time of the proposed method are more than those of the PCA-based method in Table 2. The reason can be attributed to the fact that the conventional PCA based method reconstructs the image at once by multiplying the weight matrix W, while in the proposed method reconstruction is done pixel by pixel.

4 CONCLUSIONS

In this paper, we proposed a new method to recover the occluded face images using the correlation coefficient between pairs of pixels. The simple idea that a pixel value can be determined by the weighted sum of other pixel values which are highly correlated with the one in question.

The proposed method is compared with the conventional PCA based method and it showed better recovery performance in both qualitatively and quantitatively. The blurring of the recovered images is much less and the border lines between occluded and non-occluded parts are connected well. Moreover, the mean absolute error and the mean squared error value of the proposed method are smaller than the PCA-based method. However, the proposed method was comparatively slower than the PCA-based method and this should be enhanced in the future work.

ACKNOWLEDGEMENTS

This work was supported by Korea Research Foundation Grant funded by Korean Government (KRF-2011-0005324).

REFERENCES

Wu, C., Liu, C., Shum, H., Y., Xy, Y., Q., Zhang, Z., 2004. “Automatic eyegalsses removal from face images”. *Asian Conference on Computer Vision, IEEE Trans. Pattern Analysis and Machine Intelligence.,* Vol. 26, no. 3, pp. 322-336.

Wang, Z., M., Tao, J., H., 2007. “Reconstruction of partially occluded face by fast recursive PCA”. *International Conference on Computational Intelligence and Security Workshops.*, pp . 304-307.

XiaoFeng, S., Min, T., Deren, C., RuoFeng. T., 2010. “A two step method to recover occluded part of face”. *International Conference on Information Sciences and Interaction Sciences.*, pp . 169-173.

Lin, D., Tang, X., 2007. “Quality-driven face occlusion detection and recovery”. *IEEE Conference of Computer Vision and Pattern Recognition.,* pp. 1-7.

Efros, A., A., Leung, T., K., 1999. “Texture synthesis by non-parametric sampling”. *IEEE International Conference on Computer Vision.*

Sun, J., Yuan, L., Jia, J., Shum, H.,-Y., 2005. “Image Completion with Structure Propagation”. *SIGGRAPH.*

Komodakis, N., Tziritas, G., 2006. “Image Completion Using Global Optimization”. *IEEE CVPR.*