Recognition of Blurred Faces Using Local Phase Quantization

Timo Ahonen, Esa Rahtu, Ville Ojansivu, Janne Heikkilä
Machine Vision Group, University of Oulu, PL 4500, FI-90014 Oulun yliopisto, FINLAND
{tahonen,erahtu,vpo,jth}@ee.oulu.fi

Abstract

In this paper, recognition of blurred faces using the recently introduced Local Phase Quantization (LPQ) operator is proposed. LPQ is based on quantizing the Fourier transform phase in local neighborhoods. The phase can be shown to be a blur invariant property under certain commonly fulfilled conditions. In face image analysis, histograms of LPQ labels computed within local regions are used as a face descriptor similarly to the widely used Local Binary Pattern (LBP) methodology for face image description. The experimental results on CMU PIE and FRGC 1.0.4 datasets show that the LPQ descriptor is highly tolerant to blur but still very descriptive outperforming LBP both with blurred and sharp images.

1. Introduction

In the last years, different problems in face image analysis, such as face detection, face recognition and facial expression recognition have received very much attention in computer vision research. These problems are interesting from the viewpoints of basic research aiming to efficient descriptors for facial images and of applications such as surveillance and human-computer interaction [2].

A key issue in face analysis is finding efficient descriptors for face appearance. Given the low inter-person variation in face images, ideal descriptors should be very discriminative. Still, they should be robust to different perturbations and changes such as illumination and pose changes, aging of the subjects, etc.

Despite the extensive research efforts towards face descriptors robust to the aforementioned disturbances, the problems caused by blur often present in the real-world face images have been mostly overlooked. Blur may be present in face images due to motion of the subject or the camera during the exposure, camera not being in focus, or low quality of the imaging device such as analog web camera. For example, in the Face Recognition Grand Challenge dataset [5], many of the "uncontrolled still" images appear blurred because the autofocus of the digital pocket camera did not focus into the face area.

Until now, there has been little work on blur invariant face descriptors. Facial image deblurring for recognition has been addressed in a few publications, see [8] and references therein, but to the best of authors' knowledge, explicitly constructing blur-invariant descriptors for face recognition has not been proposed before.

In this work we address the challenges caused by blur by applying the recently proposed blur tolerant Local Phase Quantization method to face description.

2. Face description using local phase quantization

The Local Phase Quantization operator was originally proposed by Ojansivu and Heikkilä for texture description [4]. The operator was shown to be robust to blur and outperform the Local Binary Pattern operator [3] in texture classification. In this work, we propose constructing a face descriptor using the LPQ operator.

2.1. Local phase quantization

The spatial blurring is represented by a convolution between the image intensity and a point spread function (PSF). In the frequency domain, this results in a multiplication $G = F \cdot H$, where G, F, and H are the Fourier transforms of the blurred image, the original image, and the PSF respectively. Further considering only the phase of the spectrum, the relation turns into a sum $\angle G = \angle F + \angle H$.

If the PSF is centrally symmetric, the transform H becomes real valued and the phase angle $\angle H$ must equal 0 or π. Furthermore, the shape of H for a regular PSF is close to a Gaussian or a sinc-function, which makes at least the low frequency values of H to be positive. At these frequencies, $\angle H = 0$ causing $\angle F$ to be a blur invariant property. This phenomenon is the basis of the local phase quantization (LPQ) method described in the following.

In LPQ the phase is examined in local M-by-M neighborhoods $\mathcal{N}_{\mathbf{x}}$ at each pixel position $\mathbf{x}$ of the image $f(\mathbf{x})$. These local spectra are computed using a short-term Fourier transform defined by

$$F(\mathbf{u}, \mathbf{x}) = \sum_{\mathbf{y} \in \mathcal{N}_{\mathbf{x}}} f(\mathbf{x} - \mathbf{y}) e^{-j2\pi \mathbf{u}^T \mathbf{y}}. \tag{1}$$

The transform (1) is efficiently evaluated for all image positions $\mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ using simply 1-D convolutions for the rows and columns successively.

The local Fourier coefficients are computed at four frequency points $\mathbf{u}_1 = [a, 0]^T$, $\mathbf{u}_2 = [0, a]^T$, $\mathbf{u}_3 = [a, a]^T$, and $\mathbf{u}_4 = [a, -a]^T$, where a is a sufficiently small scalar to satisfy $H(\mathbf{u}_i) > 0$. For each pixel position this results in a vector

$$\mathbf{F}_{\mathbf{x}} = [F(\mathbf{u}_1, \mathbf{x}), F(\mathbf{u}_2, \mathbf{x}), F(\mathbf{u}_3, \mathbf{x}), F(\mathbf{u}_4, \mathbf{x})].$$

The phase information in the Fourier coefficients is recorded by observing the signs of the real and imaginary parts of each component in $\mathbf{F}_{\mathbf{x}}$. This is done by using a simple scalar quantizer

$$q_j(\mathbf{x}) = \begin{cases} 1, & \text{if } g_j(\mathbf{x}) \geq 0 \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where $g_j(\mathbf{x})$ is the jth component of the vector $\mathbf{G}_{\mathbf{x}} = [\mathrm{Re}\{\mathbf{F}_{\mathbf{x}}\}, \mathrm{Im}\{\mathbf{F}_{\mathbf{x}}\}]$.

The resulting eight binary coefficients $q_j(\mathbf{x})$ are represented as integer values between 0-255 using binary coding

$$f_{LPQ}(\mathbf{x}) = \sum_{j=1}^{8} q_j(\mathbf{x}) 2^{j-1}.$$

As a result, we get the label image f_{LPQ} whose values are the blur invariant LPQ labels.

It can be shown that in quantization the information is maximally preserved if the samples to be quantized are statistically independent. Now in the case of real images the neighboring pixels are highly correlated, which will result in dependency between the Fourier coefficients g_j quantized in LPQ. We can however improve the situation by adding a simple decorrelation to LPQ.

Assume that the image function $f(\mathbf{x})$ is a result of a first-order Markov process, where the correlation coefficient between adjacent pixel values is ρ, and the variance of each sample is 1. As a result the covariance between positions $\mathbf{x}_i$ and $\mathbf{x}_j$ becomes $\sigma_{ij} = \rho^{||\mathbf{x}_i - \mathbf{x}_j||}$, where $||\cdot||$ denotes L_2 norm. With this information one can easily construct the covariance matrix $\mathbf{C}$ of all M^2 samples in the neighborhood $\mathcal{N}_{\mathbf{x}}$.

Rewrite the expression (1) as $F(\mathbf{u}, \mathbf{x}) = \mathbf{w}_{\mathbf{u}}^T \mathbf{f}_{\mathbf{x}}$, where $\mathbf{w}_{\mathbf{u}}$ is a vector containing the complex exponentials at frequency $\mathbf{u}$, and $\mathbf{f}_{\mathbf{x}}$ is another vector containing the values of f in $\mathcal{N}_{\mathbf{x}}$. Stacking the vectors $\mathbf{w}_{\mathbf{u_i}}$ into a matrix we have

$$\begin{aligned} \mathbf{W} = [&\mathrm{Re}\{\mathbf{w}_{\mathbf{u}_1}, \mathbf{w}_{\mathbf{u}_2}, \mathbf{w}_{\mathbf{u}_3}, \mathbf{w}_{\mathbf{u}_4}\}, \\ &\mathrm{Im}\{\mathbf{w}_{\mathbf{u}_1}, \mathbf{w}_{\mathbf{u}_2}, \mathbf{w}_{\mathbf{u}_3}, \mathbf{w}_{\mathbf{u}_4}\}]^T, \end{aligned}$$

so that $\mathbf{G}_{\mathbf{x}} = \mathbf{W}\mathbf{f}_{\mathbf{x}}$. Hence the covariance matrix $\mathbf{D}$ of $\mathbf{G}_{\mathbf{x}}$ can be obtained from $\mathbf{D} = \mathbf{W}\mathbf{C}\mathbf{W}^T$.

The covariance matrix $\mathbf{D}$ can be now used to derive a whitening transform $\mathbf{E}_{\mathbf{x}} = \mathbf{V}^T \mathbf{G}_{\mathbf{x}}$, where $\mathbf{V}$ is an orthogonal matrix derived from a singular value decomposition as $\mathbf{D} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$.

The values in $\mathbf{E}_{\mathbf{x}}$ are uncorrelated, which assuming Gaussian distributions is equal to independence. Using $\mathbf{E}_{\mathbf{x}}$ in the quantization (2) instead of $\mathbf{G}_{\mathbf{x}}$ results in more information to be stored in most cases, even though the correlation model or the Gaussian assumption would not hold exactly. It can be also shown that since decorrelation is purely a rotation of the vector $\mathbf{G}_{\mathbf{x}}$, it does not affect to the blur invariance. Note also that $\mathbf{V}$ can be solved in advance for a fixed ρ.

2.2. LPQ face descriptor

To use local phase quantization for face description, we apply the procedure of [1]. The face image is first labeled with the LPQ operator. Then, the label image is divided into non-overlapping rectangular regions of equal size and a histogram of labels is computed independently within each region. Finally, the histograms from different regions are concatenated to build a global description of the face.

3. Experiments

The efficiency of the proposed face descriptor in the face recognition scenario was tested using two datasets, the CMU PIE (Pose, Illumination, and Expression) dataset [6] and the Face Recognition Grand Challenge (FRGC) experiment 1.0.4 [5].

With both datasets, the images were first converted to gray-scale and then registered using given ground truth eye positions. In the CMU PIE experiment, the images were cropped to size 128×128 pixels and in the FRGC experiment the image size was 130×150 pixels.

In both the experiments, the local phase quantization operator was compared to LBP, which is one of the state-of-the-art methods in facial image analysis. More precisely, the $\mathrm{LBP}_{8,2}^{u2}$ operator [3, 1] was used as a control method. The parameters for the LPQ operator were $M = 7, a = 1/7$ and $\rho = 0.9$.

Figure 1. Example images from the CMU PIE database with increasing artificial blur. From left: no blur, $\sigma = 1$, $\sigma = 2$

.

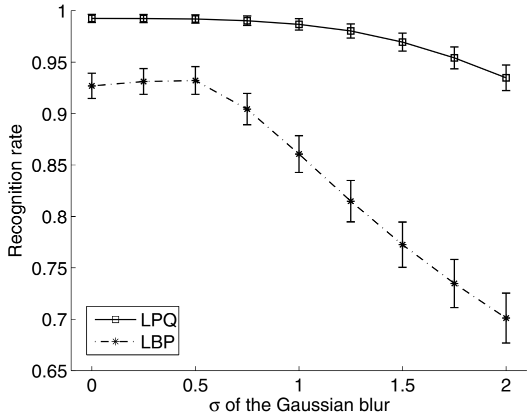

Figure 2. Recognition rates on the CMU PIE images with increasing Gaussian blur.

The LBP and LPQ histograms were extracted independently from the non-overlapping rectangular regions of size of 8×8 pixels and then concatenated to build a global description of the face. Recognition was performed with a nearest neighbor classifier with Chi square distance using the obtained histograms as feature vectors.

In the first experiment, the performance of the LPQ face descriptor was tested on images with artificially caused blur. The widely used CMU PIE database [6] was selected to serve as test material in this experiment. The images in the PIE database contain systematic lighting variation so it is possible to observe the joint effect of lighting changes and blur. For our experiments, we selected a set of 23 images of each of the 68 subjects in the database. Two of these are taken with the room lights on and the remaining 21 each with a flash at varying positions.

One sharp image per person was used as a gallery image, and the remaining 22 images were artificially blurred by convolving them with a Gaussian blur mask with $\sigma = \{0, 0.25, \ldots, 2\}$. These 68×22 blurred images were then used as probes for testing the methods. Examples of original and blurred images from the PIE dataset are shown in Fig. 1. The experiment was repeated with 10000 random partitions into gallery and probe data and the mean and standard deviations over the tests were used to assess the performance.

The mean recognition rates with standard deviations for LPQ and $\text{LBP}_{8,2}^{u2}$ are plotted in Fig. 2. As it can be seen in the results, LPQ produces better recognition result than LBP even with no blur (99.2 % vs. 92.7 %). The LBP descriptor tolerates slight blur very well but as blur increases from $\sigma = 0.5$, the recognition rate drops rapidly. At $\sigma = 2.0$, the achieved recognition rate is 70.1 %. Local phase quantization, on the other hand, tolerates increasing blur significantly better. At $\sigma = 1.0$, the recognition rate is 98.6 %, after which the rate starts to decrease slightly faster, but even at $\sigma = 2.0$ the recognition rate still is 93.5 %, higher than using LBP on images with no blur.

In the second experiment we compared the descriptors inn the FRGC experiment 1.0.4. That dataset is divided into training, probe and gallery sets which have no overlap. The probe and gallery images represent 152 subjects, and there are 1 gallery image and 2–7 probes per subject, totaling 152 images in the gallery set and 608 images in the probe set. The gallery images were taken with good quality camera under controlled conditions whereas the probe images were taken with a pocket digital camera under uncontrolled conditions. Examples of the gallery and probe images are shown in Fig. 3. As can be seen in the sample images, the uncontrolled probes contain variation of lighting, facial expression, and blur. That the probes contain significant out-of-focus blur makes this dataset very interesting for comparing LPQ and LBP.

To reduce the effect of lighting variations in facial images, Tan and Triggs proposed an illumination normalization procedure consisting of gamma correction, difference of Gaussian filtering, and contrast equalization [7]. In their experiments, they showed that this method significantly increases the recognition results of LBP based face description on the FRGC dataset. They also proposed a variant of LBP named Local Ternary Patterns and a Distance Transform to replace the histogram based approach.

Figure 3. Example images from the FRGC 1.0.4 dataset. From left: gallery image and two probe images.

Table 1. Recognition rates on the FRGC 1.0.4 dataset.

Method	Without preproc.	With preproc.
LTP+DT [7]		68.4
$LBP^{u2}_{8,2}$	32.6	64.3
LPQ	45.9	**74.5**

In our experiments on the FRGC 1.0.4 dataset we applied the illumination normalization preprocessing proposed by Tan and Triggs. Moreover, we report the results obtained with Local Ternary Patterns and Distance Transform as another control method. The preprocessing and control results were computed using implementations made available by Tan and Triggs. Note that these figures differ from those reported in [7]. This is due to the different experimental setup that was utilized in [7]. In that work several gallery images per subject were used.

The recognition results in the FRGC experiment 1.0.4 are reported in Table 1. The results without preprocessing show that the LPQ operator handles the blur and other variation in the probe set better than LBP. The lighting changes still pose a challenge also to LPQ. However, as illumination normalizing preprocessing is introduced, the recognition rate of both the methods increase notably. With preprocessing, local phase quantization reaches clearly higher recognition rate than local binary pattern or local ternary pattern.

4 Discussion and conclusions

In this paper we proposed local phase quantization operator for face recognition. Prior to this work, the effectiveness of this blur insensitive operator was experimentally shown on texture images without blur or with artificial blur. Here we analyzed the applicability of the operator for the very challenging task of face recognition and showed that it reaches higher recognition rates than the widely used local binary pattern operator.

The tests on the CMU PIE database with artificially caused Gaussian blur showed that the LPQ operator is very tolerant to blur compared to LBP. Most interestingly, its recognition rate even with no blur was better than that of LBP. In the FRGC Experiment 1.0.4, the LPQ operator clearly outperformed both the control methods, LBP and LTP. This shows that it is very robust not only to blur but also to other challenges such as lighting and facial expression variations present in real-world images. The operator is also simple to implement and fast to compute requiring only image convolutions with small separable kernels and vector rotations. Despite its simpleness and robustness, LPQ was shown to be highly discriminative producing very good recognition results.

In this work, the LPQ descriptor was applied to the face recognition problem. However, we believe that this same methodology is applicable to other face description tasks as well as to local image description in other problem domains.

Acknowledgement. This work has been performed with support from the MOBIO project, which is a 7th Framework Research Programme of the European Union.

References

[1] T. Ahonen, A. Hadid, and M. Pietikäinen. Face description with local binary patterns: Application to face recognition. *IEEE T-PAMI*, 28(12):2037–2041, 2006.

[2] S. Z. Li and A. K. Jain, editors. *Handbook of Face Recognition*. Springer, 2005.

[3] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE T-PAMI*, 24(7):971–987, Jul 2002.

[4] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. In *ICISP 2008*.

[5] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek. Overview of the face recognition grand challenge. In *CVPR 05*, pages I: 947–954, 2005.

[6] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression database. *IEEE T-PAMI*, 25(12):1615–1618, 2003.

[7] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *AMFG 2007*.

[8] Y. Yao, B. R. Abidi, N. D. Kalka, N. A. Schmid, and M. A. Abidi. Improving long range and high magnification face recognition: Database acquisition, evaluation, and enhancement. *Computer Vision and Image Understanding*, 111(2):111–125, 2008.