Face Recognition using Tensors of Census Transform Histograms from Gaussian Features Maps

John A. Ruiz-Hernandez
john-alexander.ruiz-hernandez@inrialpes.fr

James L. Crowley
james.crowley@inrialpes.fr

Antoine Meler
antoine.meler@inrialpes.fr

Augustin Lux
augustin.lux@inrialpes.fr

INRIA Rhone-Alpes Research Center
LIG Laboratory
France

Face recognition is a challenging task due to the large variety of appearance that a face may exhibit under variations in illumination and viewing position as well as variations in facial expression. Many of the more successful approaches use Gabor wavelets as an image descriptor, resulting in relatively high computational cost.
In this work, we have explored the use of Gaussian derivative features calculated with a linear complexity half-octave Gaussian pyramid [2]. We propose a tensorial representation for Gaussian derivative histograms that associate Gaussian features by the nature of information that is encoded. This representation retains spatial structure without loss of information due to vectorization of features.

The over-all architecture of our face recognition system is showed in the Figure 1 (B). We can summarize our method as follows: first, local differences in the half-octave Gaussian pyramid are used to compute Gaussian feature maps. For each location in each Gaussian map, a Census Transform histogram is calculated and concatenated to form a local tensor. Only one tensor is computed for each class of Gaussian map. An MPCA (Multilinear Principal Component Analysis) [3] method is then applied to each tensor to reduce the dimensionality and correlation due of Census Transform. Finally we apply the KDCV (Kernel Discriminative Common Vectors) method [1] to generate a discriminative vector.

1 the Half-Octave pyramid

The half-octave Gaussian pyramid for an NxN pixels image is composed of up to $K = \log_2(N)$ images. Each image, $k \in [1,K]$ of the pyramid has been convolved with a Gaussian filter, $G(x,y,2^{\frac{k}{2}})$, and can be resampled with a sample distance of $2^{\frac{(k-1)}{2}}$, resulting in a constant ratio of scale over sample distance. The resulting "pyramid" represents the original M=NxN pixels image with a sequence of $K = 2\log_2 N$ images at a geometric progressions of scales each with half the number of samples of the previous, resulting in a total of 2xM samples. Gaussian derivatives are easily calculated in the row and column directions from these images by differences of adjacent pixels.

2 Census Transform of Gaussian Features Maps

In our approach Gaussian feature maps are obtained by computing Gaussian derivative features of first and second order at four different orientations $0, \pi/4, \pi/2$ and $3\pi/2$ at 6 levels from the half-octave pyramid, corresponding to $\sigma = \sqrt{2}, 2, 2\sqrt{2}, 4, 4\sqrt{2}, 8$, as well as the feature maps from the zeroth order Gaussian pyramid. This descriptor provide a textural description that complements the information given by Gaussian Derivatives Features.For each Gaussian map, we apply a Census Transform Operator (CT) [4]. This operator compares the intensity values of a pixel with their eight neighboring pixels to obtain an 8-binary string, which is converted in to a decimal number between 0 and 255.

3 Tensors of Census Transform Histograms

We divide each CT-Gaussian Feature map into non-overlapping rectangular sub-regions with a specific size and a set of histograms is computed for each sub-region. These histograms encode the most relevant textural and spatial information in a manner that is robust to illumination changes. We avoid the problem of loss of the spatial organization due to concatenation in only one single vector, we propose to organize the histograms in a Tensor, thereby conserving the structure of the Census Transform Gaussian maps. See Figure.1 (A).

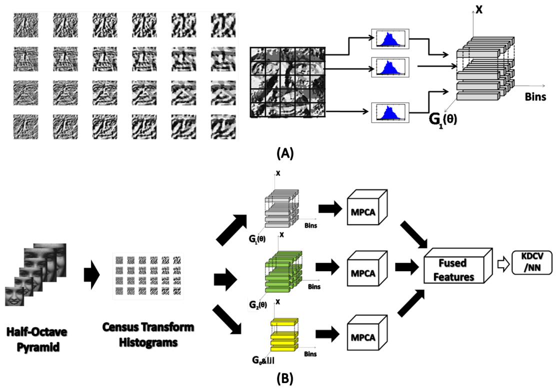

Figure 1: Example of construction for one tensor in our method (A), Overall architecture of our face recognition system (B)

4 Fusing CT Tensors Histograms with MPCA

our tensorial representation has a large redundancy, To reduce this correlation while conserving the spatial distribution of tensors, we propose the use of Multilinear Principal Component Analysis (MPCA) as suggested by Lu et al [3]. Their method determines a multilinear projection that captures most of the original tensorial input variation and supplants existing heterogeneous solutions such as the classical PCA and its 2-D variant 2-D PCA. In our approach we applied MPCA to each tensor and we retain the first m_k components from each vector. Generally the number of retained components in a PCA approach without lost of information is:

$$m_k \leq (ins, dim(v_k))$$

Where ins is the number of face images used to compute the tensors and $dim(v_k)$ is the dimension of each resultant vector. We then combine all of the resultant vectors to form a single normalized feature vector.

We have found that the discriminative power of each final vector v_{test} can be improved by projection onto an optimal discriminative space with a KDCV. The projected l_{test} is classified using the nearest neighbor rule and the cosine distance.

$$d_{cos}(l_{test}, l_{temp}) = -\frac{l_{test}^T l_{temp}}{||l_{test}|| ||l_{temp}||}$$

[1] H. Cevikalp, M. Neamtu, and M. Wilkes. Discriminative common vector method with kernels. *IEEE TNN*, 17(6):1550–1565, 2006.

[2] J.L. Crowley and O. Riff. Fast computation of scale normalised gaussian receptive fields. In *Proc. Scale Space Methods in Computer Vision*, pages 584–598, 2003.

[3] Haiping Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. Mpca: Multilinear principal component analysis of tensor objects. *IEEE TNN*, 19(1):18–39, 2008.

[4] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *ECCV '94*, pages 151–158, 1994.