Local SVD based Near-Infrared Face Retrieval
Shiv Ram Dubey1, Satish Kumar Singh2 and Rajat Kumar Singh2

1Indian Institute of Information Technology, Srity 2Indian Institute of Information Technology, Allahabad

Introduction

- Local descriptor such as Local Binary Pattern (LBP) [1] is accepted as a very prominent feature descriptor.
- The performance of such descriptors depends upon the local information of the image. The local information of the image can be enhanced using some preprocessing.
- The preprocessed images in the form of 4 sub-bands (i.e. S, U, V, and D sub-bands) are obtained by applying the Singular Value Decomposition (SVD) over the original image.
- The descriptors are computed over these sub-bands (mainly S sub-band) and termed as the SVD based descriptors.
- The experimental results confirm the superiority of using S band of SVD over NIR face databases.

Proposed SVD Sub-band Based NIR Face Retrieval

- The proposed framework of Near-Infrared (NIR) face retrieval is illustrated in Fig. 1. The main components are Singular Value Decomposition (SVD) sub-band formation, local descriptor extraction, and NIR face retrieval.
- A concept of SVD sub-band decomposition and multi-resolution representation [4] is used.
- Let \( f_{ij} \) is the intensity value of pixel at \( i^{th} \) row and \( j^{th} \) column of NIR face image \( M \) having dimension \( m \times n \).
- Let \( p_i \) is the input image of dimension \( n_i^2 \times n_j^2 \) for the \( l^{th} \) level of SVD factorization.
- The input image \( p_i \) is divided into \( 2 \times 2 \) non-overlapping blocks and SVD is applied over each block. Thus, a total \( n^2 = n_i^2 \times n_j^2 \) SVD is required, where \( n_i^2 = \lfloor n_i/2 \rfloor \) and \( n_j^2 = \lfloor n_j/2 \rfloor \).
- Let \( p_{l,t,x} \) represents the \( (t_x,t_y)\text{th} \) block of image \( p_i \), where \( t_x \in [1, n_i^2] \) and \( t_y \in [1, n_j^2] \) and given as follows,

\[
p_{l,t,x} = \begin{pmatrix} A_{l,t,x} \;& B_{l,t,x} \\ C_{l,t,x} \;& D_{l,t,x} \end{pmatrix} \quad (1)
\]

where \( A_{l,t,x} \) and \( C_{l,t,x} \) are \( 2 \times 2 \) matrices containing the orthogonal column vectors, and \( B_{l,t,x} \) is a \( 2 \times 2 \) diagonal matrix having singular values at the main diagonals. \( A_{l,t,x}, B_{l,t,x}, C_{l,t,x}, \) and \( D_{l,t,x} \) are given in the following format,

\[
Z_{l,t,x} = \begin{pmatrix} 2^{2t_x} & 2^{2t_x-1} \\ 2^{2t_y} & 2^{2t_y-1} \end{pmatrix} \quad (2)
\]

where \( Z \) represents \( P, A, B, \) and \( C \), and \( B_{2^{2t_x}+2^{2t_y}-1} = B_{2^{2t_x}-2^{2t_y}+1} = B_{2^{2t_y}-2^{2t_x}+1} = B_{2^{2t_y}-1} = 0 \).
- The four sub-bands at \( l^{th} \) level, namely \( S_1, U_1, V_1, \) and \( D_1 \) are formed from Eq. (3),

\[
\begin{align*}
S_1^{l,t,x} &= A_{2^{2t_x-1},2^{2t_y-1}} + B_{2^{2t_x},2^{2t_y}+1} + D_{2^{2t_y},2^{2t_x}+1} + C_{2^{2t_y},2^{2t_x}-1} \\
U_1^{l,t,x} &= A_{2^{2t_x},2^{2t_y}+1} + B_{2^{2t_x},2^{2t_y}-1} + C_{2^{2t_y},2^{2t_x}+1} + D_{2^{2t_y},2^{2t_x}-1} \\
V_1^{l,t,x} &= A_{2^{2t_x},2^{2t_y}+1} + B_{2^{2t_x},2^{2t_y}+1} + C_{2^{2t_y},2^{2t_x}-1} + D_{2^{2t_y},2^{2t_x}+1} \\
D_1^{l,t,x} &= A_{2^{2t_x}+1,2^{2t_y}+1} + B_{2^{2t_x},2^{2t_y}+1} + C_{2^{2t_y},2^{2t_x}+1} + D_{2^{2t_y},2^{2t_x}+1}
\end{align*}
\]

- The input image \( p_i \) for SVD at \( l^{th} \) level is defined recursively in terms of the original image \( I \) and S sub-band \( (S_{l-1}^{l-1}) \) at \( (L-1)^{th} \) level as follows,

\[
P_l = \begin{cases} I & \text{if } l = 1 \\ S_{l-1}^{l-1} & \text{Else} \end{cases} \quad (3)
\]

- For multi-resolution sub-bands the S sub-band obtained in the previous level is used as the input image.

Experiments and results

- The PolyU-NIR [2] and CASIA-NIR [3] databases are used for the face retrieval experiments. The PolyU-NIR face database consists of the total 7277 images from 55 subjects. The CASIA-NIR face database is comprised of the total 3940 images from the 197 subjects having 20 faces each.
- Four descriptors namely Local Binary Pattern (LBP) [1], Directional Binary Code (DBC) [5] and Local Gabor Binary Pattern (LGBP) [6] are computed over SVD sub-bands.
- The values of N and R are set to 8 and 1 for all descriptors.
- The SVD sub-bands at level 1 are used.
- The Chi-square (\( \chi^2 \)) distance is used as similarity measure.

- Fig. 3 shows the average retrieval precision (ARP) plots over PolyU-NIR face database and Fig. 4 depicts the average retrieval rate (ARR) over CASIA-NIR face database.

- The top 10 retrieved images using different descriptors over CASIA-NIR database is displayed in Fig. 5.

References