Local SVD based Near-Infrared Face Retrieval

Journal of Visual Communication and Image Representation, Elsevier, 2017

Shiv Ram Dubey[1], Satish Kumar Singh[2] and Rajat Kumar Singh[2]

[1]Indian Institute of Information Technology, Sricity [2]Indian 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.
- Four local descriptors over SVD sub-bands are tested for retrieval over PolyU-NIR [2] and CASIA-NIR [3] face databases
- The experimental results confirm the superiority of using S sub-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 $I^{i,j}$ is the intensity value of pixel at i^{th} row and j^{th} column of NIR face image M having dimension $m_x \times m_y$.
- Let P_L is the input image of dimension $m_x^L \times m_y^L$ for the L^{th} level of SVD factorization.
- The input image P_L is divided into 2×2 non-overlapping blocks and SVD is applied over each block. Thus, a total $n^L = n_x^L \times n_y^L$ SVD is required, where $n_x^L = \lfloor m_x^L/2 \rfloor$ and $n_y^L = \lfloor m_y^L/2 \rfloor$.
- Let P_{L,t_x,t_y} represents the $(t_x, t_y)^{th}$ block of image P_L, where $t_x \in [1, n_x^L]$ and $t_y \in [1, n_y^L]$ and given as follows,

$$P_{L,t_x,t_y} = \left(A_{L,t_x,t_y}\right)\left(B_{L,t_x,t_y}\right)\left(C_{L,t_x,t_y}\right)^T \quad (1)$$

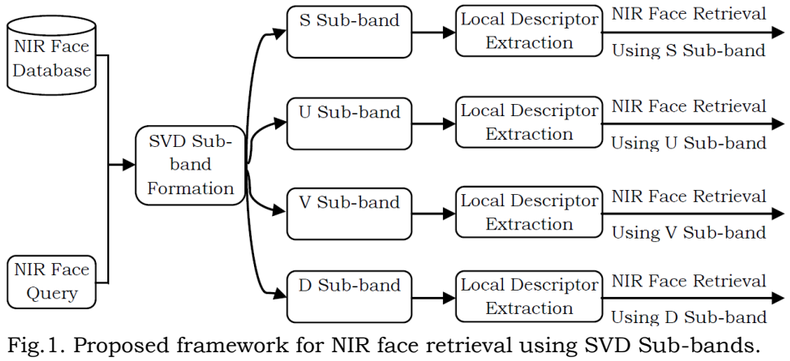


Fig.1. Proposed framework for NIR face retrieval using SVD Sub-bands.

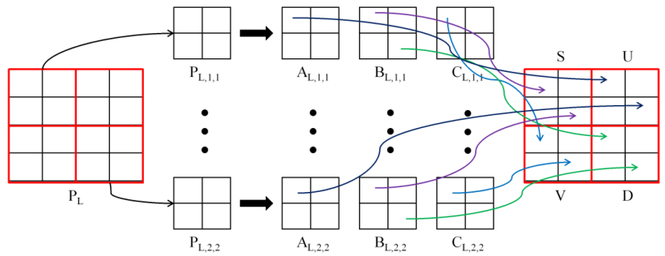


Fig.2. Illustration of S, U, V and D sub-band formation by the SVD factorization of any input P_L using an example of size 4×4.

where A_{L,t_x,t_y} and C_{L,t_x,t_y} are 2×2 matrices containing the orthogonal column vectors, and B_{L,t_x,t_y} is a 2×2 diagonal matrix having singular values at the main diagonals. P_{L,t_x,t_y}, A_{L,t_x,t_y}, B_{L,t_x,t_y}, and C_{L,t_x,t_y} are given in the following format,

$$Z_{L,t_x,t_y} = \begin{bmatrix} Z_L^{2t_x-1,2t_y-1} & Z_L^{2t_x-1,2t_y} \\ Z_L^{2t_x,2t_y-1} & Z_L^{2t_x,2t_y} \end{bmatrix} \quad (2)$$

where Z represents P, A, B, and C, and $B_L^{2t_x-1,2t_y} = B_L^{2t_x,2t_y-1} = 0$.

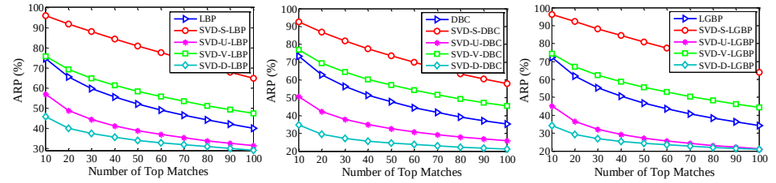
- The four sub-bands at Lth level, namely S_L, U_L, V_L and D_L are formed from Eq. (3) as, $S_L^{t_x,t_y} = B_L^{2t_x-1,2t_y-1}$, $U_L^{t_x,t_y} = A_L^{2t_x-1,2t_y-1}$, $V_L^{t_x,t_y} = C_L^{2t_x-1,2t_y-1}$, and $D_L^{t_x,t_y} = B_L^{2t_x,2t_y}$.
- The input image P_L for SVD at L^{th} level is defined recursively in terms of the original image (I) and S sub-band ($S_{L-1}^{t_x,t_y}$) at $(L-1)^{th}$ level as follows:

$$P_L = \begin{cases} I & if\ L = 1 \\ S_{L-1}^{t_x,t_y} & Else \end{cases} \quad (3)$$

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

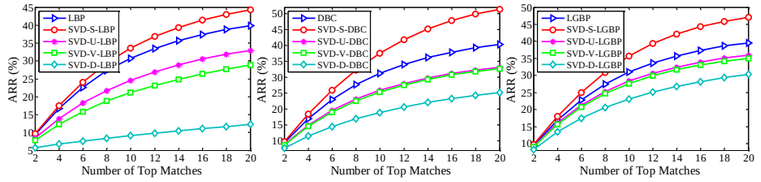
Experiments and results

- 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 (χ^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.



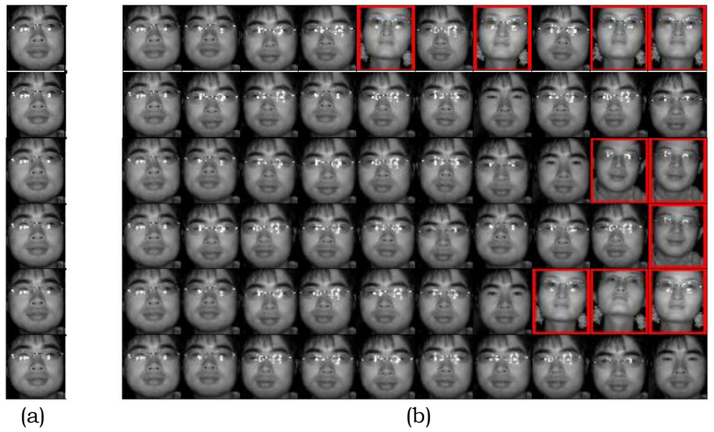
(a) (b) (c)

Fig.3. The results in terms of ARP (%) using (a) LBP, (b) DBC, and (c) LGBP using different SVD sub-bands over PolyU-NIR face database.



(a) (b) (c)

Fig.4. The results in terms of ARR (%) using (a) LBP, (b) DBC, and (c) LGBP using different SVD sub-bands over CASIS-NIR face database.



(a) (b)

Fig.5. Retrieval results from CASIA-NIR face database using LBP (1st row), SVD-S-LBP (2nd row), DBC (3th row), SVD-S-DBC (4th row), LGBP (5th row), and SVD-S-LGBP (6th row), (a) query face, and (b) top 10 retrieved faces. Faces in red rectangles are the false positives.

References

[1]. Ojala et al., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE TPAMI*, 24(7): 971-987, 2002.
[2]. PolyU-NIRFD, www.comp.polyu.edu.hk/~biometrics/NIRFace/polyudb_face.htm.
[3]. CASIA-NIRFD, www.cbsr.ia.ac.cn/english/NIR_face%20Databases.asp.
[4]. Singh and Kumar, "Singular value decomposition based sub-band decomposition and multi-resolution (SVD-SBD-MRR) representation of digital colour images", *Pertanika Journal of Science and Technology*, 19(2): 229-235, 2011.
[5]. Zhang et al., "Directional binary code with application to PolyU near-infrared face database", *Pattern Recognition Letters*, 31(14): 2337-2344, 2010.
[6]. Zhang et al., "LGBPHS: A novel non-statistical model for face representation and recognition," *IEEE ICCV*, pp. 786-791, 2005.