IMAGE COMPRESSION USING ORTHOGONALIZED INDEPENDENT COMPONENTS BASES

Artur J. Ferreira
Inst. of Telecommunications, and
Inst. Sup. de Engenharia de Lisboa
R. Conselheiro Emidio Navarro, n. 1
1949-014 Lisboa, Portugal
arturj@cc.isel.ipl.pt

Mário A. T. Figueiredo
Inst. of Telecommunications, and
Inst. Superior Técnico
Av. Rovisco Pais
1049-001 Lisboa, Portugal
mtf@lx.it.pt

Abstract. In this paper we address the orthogonalization of *independent component analysis* (ICA) to obtain transform-based image coders. We consider several classes of training images, from which we extract the independent components, followed by orthogonalization, obtaining bases for image coding.

Experimental tests show the generalization ability of ICA of natural images, and the adaptation ability to specific classes.

The proposed fixed size block coders have lower transform complexity than JPEG. They outperform JPEG, on several classes of images, for a given range of compression ratios, according to both standard (SNR) and perceptual (picture quality scale – PQS) measures. For some image classes, the visual quality of the images obtained with our coders is similar to that obtained by JPEG2000, which is currently the state of the art still image coder. On fingerprint images, our fixed and variable size block coders perform competitively with the special-purpose wavelet-based coder developed by the FBI.

INTRODUCTION

Independent component analysis (ICA) considers a class of probabilistic generative models in which an observed random vector $\mathbf{X}$ is obtained according to $\mathbf{X} = \mathbf{AS}$, where $\mathbf{A}$ is an $N \times M$ unknown mixing matrix and $\mathbf{S}$ is a vector of independent sources [2, 7, 10]. The standard goal of ICA is to infer (learn) $\mathbf{A}$ from a set of samples of the random vector $\mathbf{X}$. To apply ICA to images, each sample of $\mathbf{X}$ usually contains the pixels in an image block.

It has been found that images of natural scenes are well modelled when the columns of $\mathbf{A}$ are wavelet-like filters, and the independent sources (elements of $\mathbf{S}$) have heavy-tailed distributions [1, 7, 10]. This means that, with high probability, only a small fraction of the components of $\mathbf{S}$ have significant values; this *sparse* nature of $\mathbf{S}$ underlies the potential usefulness of overcomplete ICA to compression [14] and denoising of natural images [7, 10].

Recently, we have shown that non-orthogonal ICA bases are suited for class-adapted image compression at low bit-rates using matching pursuit [3, 12] type algorithms [5]. In this paper, we exploit the data-dependent nature of the ICA decomposition, onto the basis defined by the orthogonalization of the columns of $\mathbf{A}$ (basis vectors). We use the FastICA algorithm [7, 8] to learn complete and overcomplete bases from training images. Since these bases are non-orthogonal, we apply orthogonalization methods to obtain fixed transforms for image coding, to be used like a data-independent transform.

The paper is organized as follows. First, we present basis vectors extracted from natural, fingerprint, faces, and synthetic images, using ICA. Next, we present the energy compaction ability of their orthogonalized versions, comparatively to other transforms. The coder architecture is then described, and experimental results are shown, along with the results of other coders. A final section presents some concluding remarks.

BASIS ESTIMATION

We apply the FastICA algorithm [6, 7, 8] to obtain both complete ($N = M$) and overcomplete ($M > N$) bases for randomly selected training sets of (8×8) image blocks (400 per image), after mean removal and sphering by principal component analysis [7, 9]. We consider a set of natural and synthetic images (from the University of Waterloo, Canada[1]), a set of fingerprint images[2], and a set of face images (from the Cambridge AT&T laboratories[3]). Fig. 1 displays the estimated complete ICA bases (the columns of $\mathbf{A}$) for these four classes. In the overcomplete case, the basis vectors are visually similar to these ones.

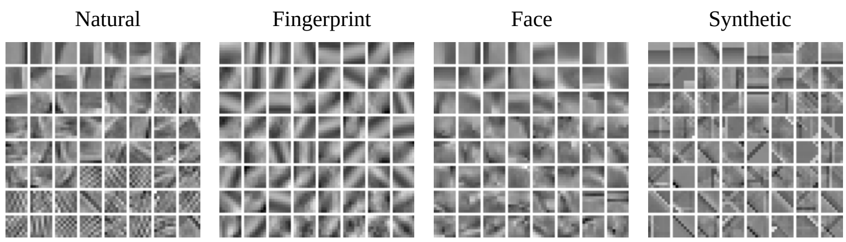

Figure 1: ICA bases from natural, fingerprint, face, and synthetic images.

For natural images, the resemblance between ICA, wavelet basis vectors,

[1] `http://links.uwaterloo.ca/bragzone.base.html`
[2] `http://bias.csr.unibo.it/fvc2000/databases.asp`
[3] `http://www.uk.research.att.com/facedatabase.html`

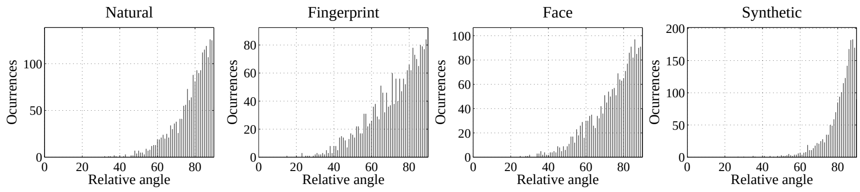

Figure 2: Relative angle histograms for the ICA bases in figure 1.

and Gabor functions has been noted [7, 10]. The basis vectors of each class are globally different, showing specific features. The histograms of relative angles between all the basis vectors, for the four bases considered, are displayed in Fig. 2. As we can see from these histograms, the ICA bases are non-orthogonal. To deal with this non-orthogonality, two options are available: apply matching pursuit type algorithms [3, 12] to perform the image decomposition; orthogonalize the ICA bases. The first approach was exploited in [5]. In this paper, we will follow the second option, depicted in Fig. 3.

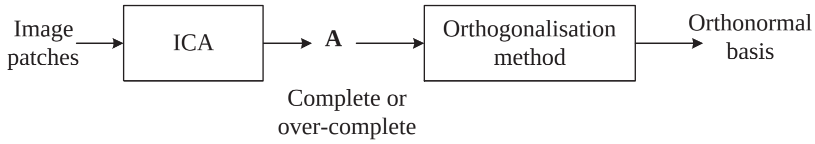

Figure 3: Orthonormal basis estimation.

The *Karhunen-Loève transform* (KLT) [9] is optimal in the sense of mean-square error, and decorrelation among the set of orthogonal transforms. The *discrete cosine transform* (DCT) [9], which is used in the JPEG standard, approximates the KLT under certain conditions [9], making it suitable for image coding. Thus, in Fig. 4 we present the KLT basis vectors for the same set of natural and fingerprint images, along with the orthogonalized ICA bases, obtained by KLT, and by the sequential *Gram-Schmidt* (GS) procedure [7]. The majority of DCT and KLT basis vectors are similar [9]. Notice the resemblance between KLT and ICA+KLT vectors, specially on the first row (low frequency basis vectors). Some ICA+GS vectors are similar to some KLT (and DCT) vectors.

ENERGY COMPACTION

In this section, we compare our orthogonal transforms with KLT and DCT on natural and fingerprint images. This comparison is performed by coding an image of each class (not used in the basis estimation step), using the first n coefficients, and measuring the corresponding SNR,

$$SNR_n = 10 \log_{10} \left(\sigma^2 / MSE_n \right) \qquad [dB], \tag{1}$$

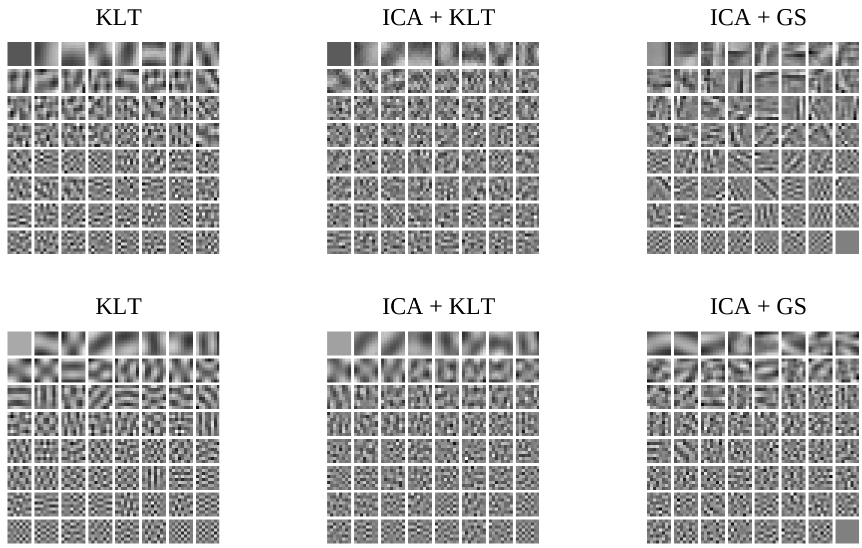

Figure 4: KLT, ICA+KLT, and ICA+GS bases obtained from natural and fingerprint images (on first and second row, respectively).

where σ^2 is the original image variance, and MSE_n is the mean squared error between the original image and its block by block n-term representation. Figure 5 shows the energy packing (SNR as a function of the number of coefficients), of each of these transforms, with orthogonalized complete and two times overcomplete ICA bases, extracted from natural and fingerprint images. We can see that the use of an orthogonalized two times overcomplete ICA basis does not improve much the SNR, when compared to the orthogonalization of a complete basis. The same happens when the degree of overcompleteness is higher (4 to 8 times). The DCT presents larger energy compaction than KLT, because the latter is extracted from a set of randomly selected blocks of several images.

One must be careful with the examination of the sequential GS procedure on an overcomplete ICA basis, because this procedure converges after using a number of vectors equal to dimension of the space. It does not really orthogonalizes all the vectors in the overcomplete ICA basis, and the result depends on which components are first considered. Because of this, and being aware of the ICA permutation ambiguity [2, 7, 10], the ICA basis vectors are sorted decreasingly according to their energy concentration on the spectra. After this procedure, the ICA low-pass vectors appear first (visually similar to the KLT first vectors). Symmetric orthogonalization [7] was also considered, but the energy compaction is much smaller than that obtained with the presented methods, because this method spreads energy through the coefficients (instead of concentrating it, as desirable).

Although KLT achieves better energy compaction than ICA+KLT and

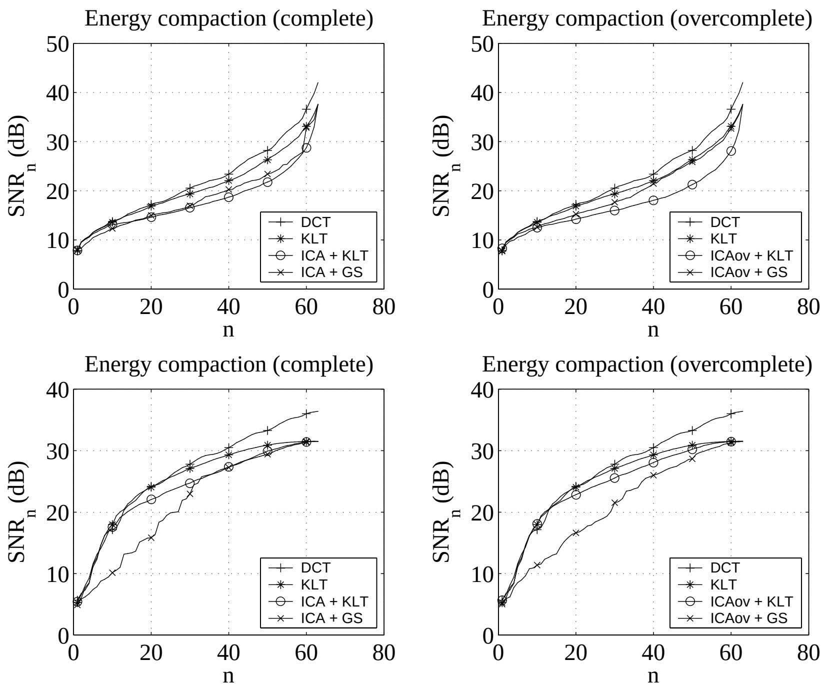

Figure 5: Energy compaction for the natural image *boat* (first row), and a fingerprint image (second row).

ICA+GS, this does not imply that all KLT-based coders achieve better SNR as a function of bit-rate. Generally, the coefficient distribution on orthogonalized ICA bases tends to be heavy-tailed, more suited for quantization and entropy coding than the coefficients of the KLT and DCT, which tend to a normal distribution. The orthogonalization by GS is better than KLT for natural images. The reverse happens for fingerprint images.

CODER ARCHITECTURE

The proposed image coder is transform-based, as shown in Fig. 6. The transform coefficients are obtained by orthogonal projection on the orthogonalized ICA basis. The basis is *not* transmitted with the image, thus being used like for a data-independent transform. Operation modes with fixed and variable size blocks are considered, using the first n transform coefficients. Thus, the coefficients vector dimension is less than the dimension of the space, resulting in a compact representation.

Coefficient quantization is performed using a Lloyd I [11] quantizer, learnt off-line from the transform coefficients. Each coefficient is quantized with 6 bits. Entropy coding of the quantizer output is carried out by an adaptive arithmetic coder [15], using source models (histograms obtained off-line from

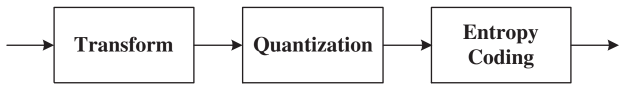

Figure 6: Transform-based coder architecture.

several test images of the specific class being considered). The mean value of each block is separately quantized (also with 6 bits), and entropy-coded.

Using 8×8 blocks, the proposed coders present coding and decoding complexity similar to JPEG. Calculating only the first n coefficients, the transform complexity is lower than for JPEG. On the other hand, our entropy coding scheme is more complex than that of JPEG. In the case of variable size blocks, image analysis is performed using blocks of sizes 16×16, 8×8, and 4×4, organized in a quad-tree structure [9]. Splitting each 16×16 or 8×8 block into its four sub-blocks is done when the maximum absolute difference between the original and the coded block exceeds a predefined threshold denoted Δ. The resulting tree decomposition is encoded using an adaptive arithmetic coder. In this case, the complexity is larger than for JPEG, but still remains below that of JPEG2000.

EXPERIMENTAL RESULTS

We now present results (SNR plots) of several coders, following the architecture in Fig. 6, considering the previously mentioned transforms: KLT, ICA+KLT, and ICA+GS. A Lloyd I [11] quantizer was designed for each of these transforms, and the corresponding coefficient source model was established for adaptive arithmetic entropy-coding. Fig. 7 plots SNR as a function of bit-rate, for several coders, on the natural image *boat*, a face, and a fingerprint image, using 8×8 blocks. The rate is varied by coding with an increasing number of coefficients, on consecutive tests (each coefficient is quantized with 6 bits). The plot on the left hand side refers to a complete ICA basis, while the one on the right corresponds to overcomplete bases. The bases were obtained from randomly extracted blocks of the well-known images *camera*, *barbara*, *bird*, and *goldhill*. For fingerprint and face images, the bases were extracted from a set of four images of the corresponding class (not including the test image). The JPEG[4] and JPEG2000[5] [16] coders results are also displayed for comparison. This comparison makes sense, because JPEG is the standard 8×8 DCT-based block coder, and JPEG2000 is currently the state of the art still image coder. For fingerprint images, we also display the results of WSQ (*wavelet scalar quantization* [4]) which is a wavelet-based special-purpose coder for fingerprint images.

On the face image, the ICA-based coders perform better than JPEG, at high compression ratios (until 0.8 bpp), being close to JPEG2000. On the fingerprint image, the ICA+KLT coder results are better than those of JPEG

[4] ftp://ftp.simtel.net/%2f/pub/simtelnet/msdos/graphics/jpegsr6.zip
[5] http://www.ee.unsw.edu.au/~taubman/kakadusoftware/

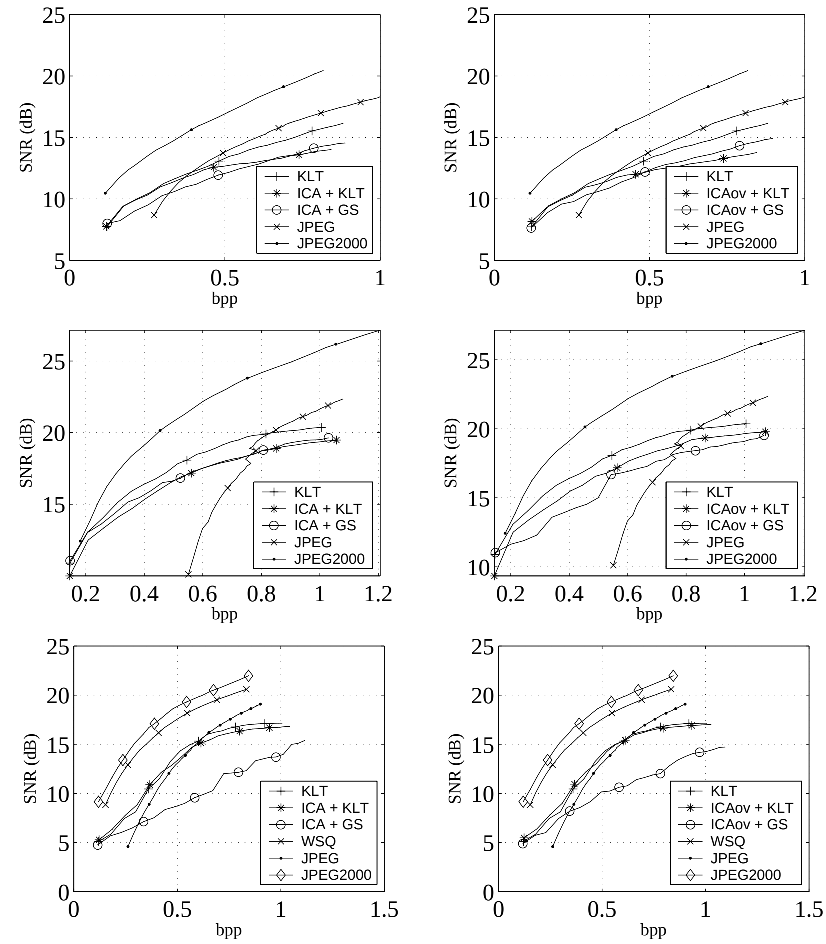

Figure 7: SNR plots for the natural image *boat*, a face image, and a fingerprint image, on the first, second, and third row respectively. The first column refers to a complete basis, and the second to an overcomplete basis.

(until 0.6 bpp), but worse than WSQ. The ICA+KLT based coder attains better results than ICA+GS.

We next present results concerning the generalization and adaptation abilities of the ICA bases learnt from the image classes considered.

Natural Images

Extracting the ICA bases from the *boat* image, we code the *bird* test image, at 0.3 bpp, using the first 6 coefficients. The results are displayed in Fig. 8, in which we present the corresponding SNR and PQS (*picture quality scale* [13]) values. The latter is an objective distortion measure of perceptual image

17.7 dB -0.73 12.8 dB -3.98 16.1 dB -1.76 23.2 dB 2.1

ICA + KLT ICA + GS JPEG JPEG2000

Figure 8: Natural image *bird*, coded at 0.3 bpp.

quality, based on a model of the human visual system [13]. As can be seen, extracting a basis from a single image gives a transform with generalization ability comparable to data-independent transforms, such as the DCT. It can also be noted that the orthogonalization by GS produces worse results than the KLT orthogonalization (specially on the image contours). The JPEG image is more blocky than the one obtained with ICA+KLT.

As a test for the generalization ability of the bases learnt from natural images, we have coded three images from different classes (see results in Fig. 9): synthetic, faces, and fingerprints. For the face image, the SNR, the PQS, and the visual quality achieved by the ICA-based coders are better than those obtained with JPEG, and close to JPEG2000. These results testify for the good generalization ability of the ICA bases of natural images.

Specific classes

Finally, we study the adaptation ability of ICA to specific image classes, by extracting the ICA bases from face and fingerprint images. We then code an image of that class, *not* used in the basis estimation process. These tests (see results in Fig. 10 and Fig. 11) reveal the adaptation ability to specific image classes. On face images, the results are clearly better than that of JPEG, without block effects. On fingerprint images, the variable size block coder attains SNR and PQS values which are almost exactly the same as WSQ, and close to JPEG2000.

CONCLUSIONS

In this paper, we have considered the use of orthogonalized ICA bases for image compression, like for a data-independent transform. The proposed coder presents better results than JPEG, at the same (and lower) coding complexity, for a wide range of compression ratios, on several image classes. These coders also show less block effect than JPEG, and images with better visual quality, using 8×8 blocks. The use of complete and overcomplete ICA bases followed by orthogonalization produces roughly the same results.

We have shown the generalization ability of the ICA bases extracted from natural images and the adaptation to specific classes, exemplified on face and

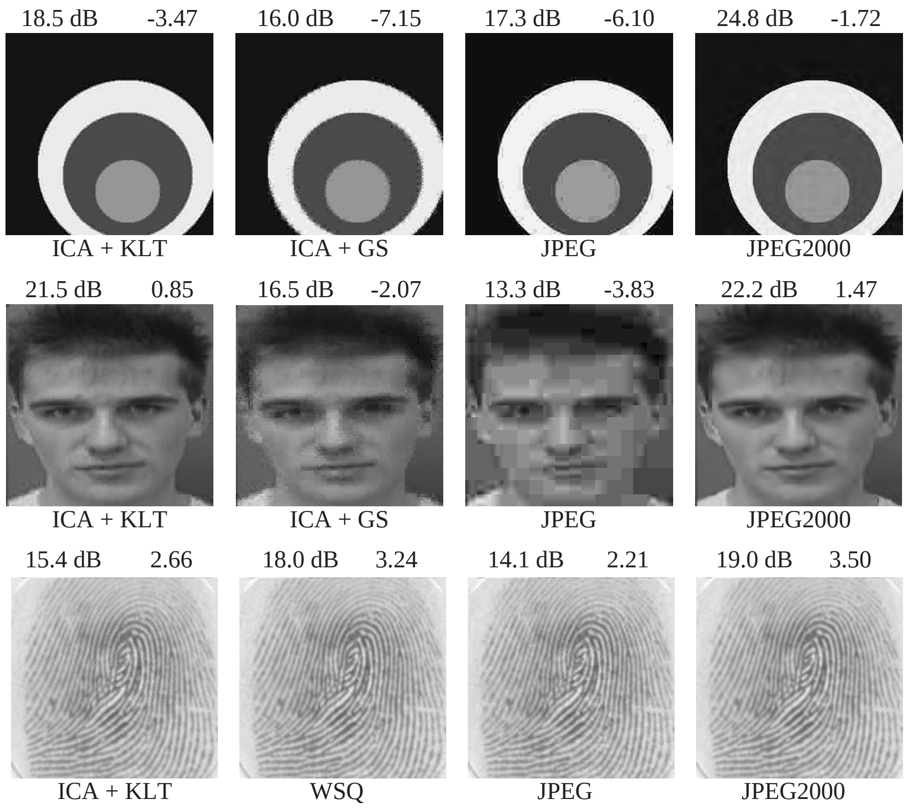

Figure 9: Synthetic image circles (at 0.3 bpp), a face image, and a fingerprint image (both at 0.6 bpp).

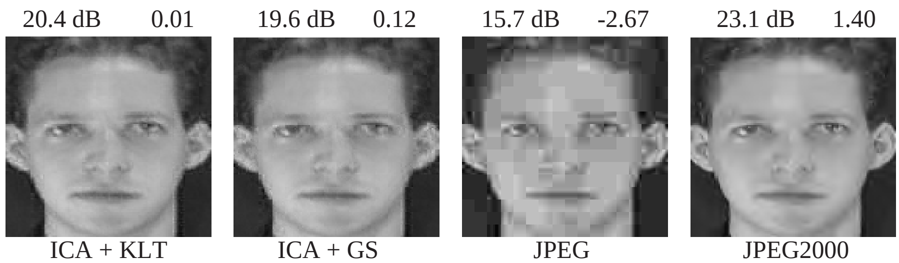

Figure 10: Face image, coded at 0.6 bpp.

fingerprint images. For the latter, a comparison with WSQ was carried out, and we concluded that using variable size blocks, one can attain distortion values similar to this special-purpose wavelet-based coder developed by the FBI. Thus, we can conclude that image coding with ICA is an appropriate and competitive choice for specific image classes.

Comparatively to JPEG2000, which is now the state of the art still image coder, the distortion and perceptual quality of the ICA-based coders are close, within the specific image classes considered, specially at low bit rates.

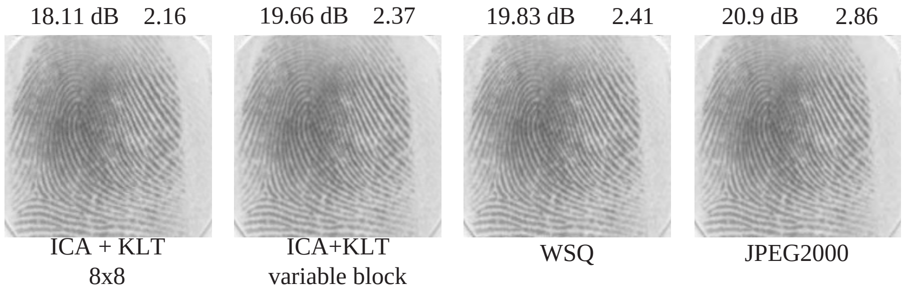

Figure 11: Fingerprint image, coded at 0.6 bpp.

REFERENCES

[1] A. Bell and T. Sejnowski, "The 'independent components' of natural scenes are edge filters," **Vision Research**, vol. 37, pp. 3327–3338, 1997.

[2] P. Comon, "Independent component analysis: a new concept?" **Signal Proc. Elsevier**, vol. 36, pp. 287–314, 1994.

[3] G. Davis, **Adaptive Nonlinear Approximations**, Ph.D. thesis, Courant Institute of Mathematical Sciences, New York University, 1994.

[4] Federal Bureau of Investigation, **WSQ gray-scale fingerprint image compression specification**, IAFIS-IC-0110v2 (rev. 2.0), 1993.

[5] A. Ferreira and M. Figueiredo, "Class adapted image compression using independent component analysis," in **Proc. International Conference on Image Processing (ICIP'2003)**, Barcelon, Spain, 2003.

[6] A. Hyvärinen, R. Cristescu and E. Oja, "A fast algorithm for estimating overcomplete ICA bases for image windows," in **Proc. Int. Joint Conf. on Neural Networks**, Washington, D.C., 1999, pp. 894–899.

[7] A. Hyvärinen, J. Karhunen and E. Oja, **Independent Component Analysis**, Wiley Interscience, 2001.

[8] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," **Neural Computation**, vol. 9, pp. 1483–1492, 1997.

[9] A. Jain, **Fundamentals of Digital Image Processing**, Prentice Hall, 3rd edn., 1989.

[10] T.-W. Lee, **Independent Component Analysis - Theory and Applications**, Kluwer Academic Publishers, 1998.

[11] S. Lloyd, "Least squares quantization in PCM," **IEEE Trans. on Information Theory**, vol. IT-28, pp. 127–135, 1982.

[12] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," **IEEE Trans. on Signal Processing**, vol. 41, no. 12, pp. 3397–3415, 1993.

[13] M. Miyahara, K. Kotani and V. Algazi, "Objective picture quality scale (PQS) for image coding," **IEEE Trans. on Comm.**, vol. 46, pp. 1215–1226, 1998.

[14] A. Puga and A. Alves, "An experiment on comparing PCA and ICA in classical transform image coding," in **ICA99**, 1999, pp. 105–108.

[15] J. Rissanen and G. Langdon, "Arithmetic coding," **IBM Journal of Research and Development**, vol. 23, pp. 149–162, 1979.

[16] D. Taubman and M. Marcellin, **JPEG2000: Image Compression Fundamentals, Standards, and Practice**, Kluwer Academic Publishers, 2001.