A novel efficient Image Compression System based on Independent Component Analysis

Zafar SHAHID, Florent DUPONT, Atilla BASKURT

LIRIS UMR5205 CNRS, Université de Lyon, INSA-Lyon, Université Lyon 1,
F-69622 Villeurbanne Cedex, France

ABSTRACT

Next generation image compression system should be optimized the way human vision system (HVS) works. HVS has been evolved over millions of years for the images which exist in our environment. This idea is reinforced by the fact that sparse codes extracted from natural images resemble the primary visual cortex of HVS. We have introduced a novel technique in which basis functions trained by Independent Component Analysis (ICA) have been used to transform the image. ICA has been used to extract the independent features (basis functions) which are localized, bandlimited and oriented like HVS and resemble wavelet and Gabor bases. A greedy algorithm named matching pursuit (MP) has been used to transform the image in the ICA domain which is followed by quantization and multistage entropy coding. We have compared our codec with JPEG from the DCT family and JPEG2000 from the wavelets family. For fingerprint images, results are also compared with wavelet scalar quantization (WSQ) codec which has been especially tailored for this type of images. Our codec outperforms JPEG and WSQ and also performs comparable to JPEG2000 with lower complexity than the latter.

Keywords: Image Compression, Redundant Dictionaries, Matching Pursuit, Independent Component Analysis

1. INTRODUCTION

Independent component analysis (ICA)[1] presents a probabilistic image model in which an observed random vector x containing pixels from an image can be decomposed as:

$$x = \mathbf{A}s. \tag{1}$$

Here s is a vector containing independent sources, which are linearly combined into the observations x through the basis function $\mathbf{A}^i$ where the superscript i denotes the i-th column of $\mathbf{A}$. In non-orthogonal paradigm, collection of atoms (or basis functions) is termed as a dictionary which may be incomplete, complete or overcomplete. The terms *basis functions* and *atoms* are used interchangeably in this paper.

Basis functions trained by ICA have already been studied in literature. Ferreira *et al.*[2,3] have presented ICA based image compression system. They used a small image-database containing four images to train the basis functions and the results they got were inferior to JPEG2000[4] (2dB PSNR difference). We have used more general image database containing wildlife images. Moreover, our entropy coding consists of multiple stages to better exploit the redundancy of the quantized transformed coefficients. As a result, distortion and perceptual quality of our codec is better than JPEG and comparable to that of JPEG2000.

We organize our work as follows. In section 2, we introduce transform coding, the method which is used to transform the data with non-orthogonal dictionaries and ICA based image model. We discuss the image compression problem and explain why ICA dictionaries are suitable for image compression in section 3. Architecture of our proposed ICA based codec is presented in section 4. Section 5 contains its performance analysis including its energy compaction capability, feature capturing capability, quantization and comparison with other codecs. In the end, we present the concluding remarks about the proposed algorithm.

zafar.shahid@lirmm.fr, florent.dupont@liris.cnrs.fr, atilla.baskurt@liris.cnrs.fr

2. TRANSFORM CODING

Common audio and video sources have a lot of redundancy that is not removed by source coding. It is well-known that removing redundancy by using transform coding before quantization generates much better codes. In transform coding, data is transformed to some other space where it has sparse representation with high peak and heavy tails.

Historically, transforms have been linear and complete. Being orthogonal, they guarantee energy conservation and can be implemented by a simple matrix multiplication e.g. DCT and DWT. Now the paradigm is shifting towards non-orthogonal dictionaries which may be incomplete, complete or overcomplete. If $D = \{f_1, f_2, f_3, ..., f_m\}$ is a dictionary, the dictionary is said to be complete if the number of atoms is equal to number of pixels (or the rank of the observation vector) and the matrix $[f_1, f_2, f_3, ..., f_m]$ is full column rank. It is incomplete or overcomplete based on whether the number of atoms are less or greater than the number of pixels. Iterative algorithms like MP[5] and Basis Pursuit[6] are used to transform the image with these dictionaries. In this paper, we have used MP and its variant orthogonal matching pursuit (OMP) to transform the image in ICA domain.

2.1. Matching pursuit

MP belongs to a class of iterative algorithms that approximate a signal by successive projections on the vectors of a non-orthogonal, overcomplete dictionary. After n number of iterations, residual is reduced to zero where n is equal to the size of dictionary, provided the dictionary is complete. Representations by overcomplete dictionaries are not unique. There exists many solutions and finding the best solution is NP-complete. For the decomposition of a 1D signal g, let us have atoms from a dictionary set given as:

$$D = \{f_1, f_2, f_3, ..., f_m\} \tag{2}$$

where all the atoms are normalized with $\|f_i\| = 1$. Decomposition begins by choosing to maximize the absolute value of the inner product:

$$\alpha = \arg \max_{\gamma \epsilon \{1,...,m\}} | < g_n, f_{n\gamma} > | \tag{3}$$

where α is an expansion coefficient for the signal onto the dictionary function, g_n is the residual signal in nth iteration and $f_{n\gamma}$ is the dictionary atom with index γ which maximizes α in this iteration. Residual signal is then computed as:

$$R_n = g_n - \alpha f_{n\gamma} \tag{4}$$

and in next iteration, this residual signal is then expanded again in the same way as the original signal until some stopping condition is met which is normally a compromise between maximum number of iterations and minimum energy residual. After n number of iterations, the signal can be approximated by:

$$\hat{g} = \sum_{i=0}^{n} \alpha_i f_{i\gamma} \tag{5}$$

where $\hat{g}$ is the approximated reconstructed version of the original signal. At each stage, dictionary element which minimizes mean square error (MSE) between original signal g and the coded signal $\hat{g}$ is chosen. So, MP bitstream is progressive and scalable. If the dictionary being used is derived by ICA, the components are coded in order of visual importance. Orthogonal matching pursuit (OMP) is a variant of MP in which each basis vector is orthogonalized with respect to all previously selected basis vectors at the time of selection.

2.2. ICA based image model

All overcomplete dictionaries like Gabor, anisotropically refined Gaussian[7] and B-spline,[8] are mathematical formulas approximating the image model. They target all the possible combinations of pixel values. But natural images, over which HVS has been optimized over millions of years, do not utilize all the possible pixel combinations and their characteristics have been exploited well by primary visual cortex of HVS.[9, 10]
Primary visual cortex has two very important characteristics. First, it contains many more neurons than image data which gives a notion of overcompleteness. Second, any particular neuron triggers rarely but with high

value[11] which gives a notion of sparseness. Here every basis function represents single neuron. So output of transform coding should be sparse (super-Gaussian). That is why ICA basis functions are extracted by maximizing the non-Gaussianity of independent components.

3. ICA FOR IMAGE COMPRESSION

Due to its resemblance with curvelets, it is natural to use ICA for image compression. ICA performs well for compression of natural images and it has also the capability to adapt to a certain image class. Concept of class specific image codecs already exist e.g. WSQ[12] was tailored for fingerprint images because JPEG2000 was not available at that time and DCT based codec had blocking artifacts. JPEG2000 is now the latest wavelet based codec and provides better compression than WSQ.[13]

3.1. Extraction of basis functions

For extraction of basis functions, FastICA[14] has been used for incomplete, complete and overcomplete dictionaries. *tanh* non-linearity has been used for maximization of super-Gaussianity. For overcomplete dictionaries it utilizes the supposition of quasi-orthogonality.[15] The algorithm is shown in Fig. 1 and works in the following steps:

- Images patches are extracted from random locations from a group of images having same statistical characteristics.

- As a preprocessing step, they are processed by approximative orthogonalization prewhitening.

- FastICA algorithm is then used to extract the basis functions which are as independent as possible.

- Basis functions, extracted as independent sources, are not ordered. For better performance of entropy encoder, we need to order them in a fashion that gives a long trail of zeros in the end and most of nonzero coefficients (NZs) in the beginning of every block. As ICA basis functions are trained for a specific class, energy of each basis is a clear indication how much the specific basis function exists in that class. So, energy level is an appropriate criteria for ordering of basis functions to get better compression. This ordering gives excellent results for basis functions of both natural images and database images. Thus, in post-processing step, ICA bases are ordered by their energy level and then are normalized.

For bases estimation of natural images, we have used a dataset of natural images containing 13 wildlife images. For face and fingerprint databases, a set of images from respective databases have been used. Basis functions extracted from these datasets for 8x8 image block are shown in Fig. 2. These basis functions have been extracted after subtracting the arithmetic mean from the image pixels(so there are 63 basis functions for a block size of 8x8 pixels).
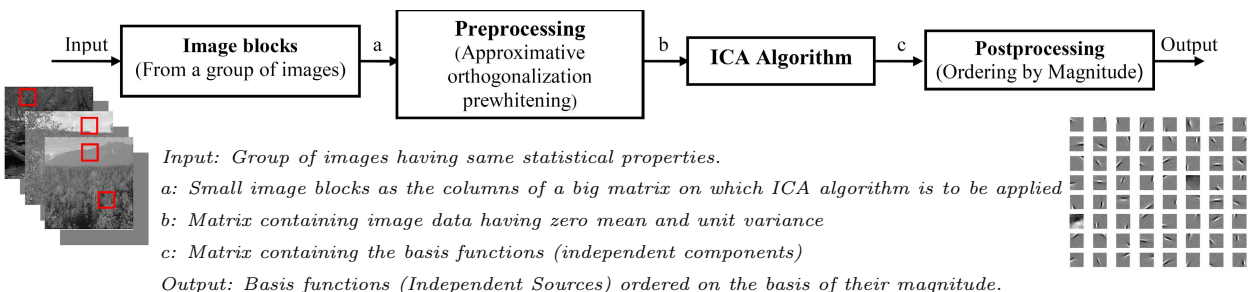

Input: Group of images having same statistical properties.
a: Small image blocks as the columns of a big matrix on which ICA algorithm is to be applied
b: Matrix containing image data having zero mean and unit variance
c: Matrix containing the basis functions (independent components)
Output: Basis functions (Independent Sources) ordered on the basis of their magnitude.

Figure 1. Block diagram for extraction of ICA basis functions.

3.2. Characteristics of ICA dictionaries

By ordering the ICA basis functions in a descending order of magnitude, they are automatically ordered in ascending order by frequency. It confirms that low frequencies are of greater importance for HVS. In the ICA domain, most of the NZs are in the beginning irrespective of their orientation and we get a large trail of zeros which can be well encoded by an entropy encoder. From Fig. 2, one can note that ICA bases are well localized in space and orientation like Gabor bases. Like wavelets, they represent more or less the same features in different scales. But they have many more degrees of freedom than wavelets and have different phases. They are non-separable and represent curves and edges.
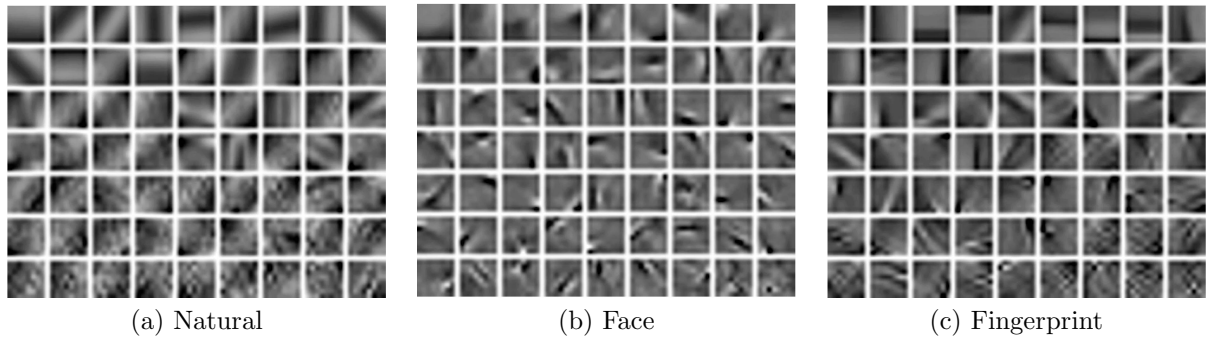

(a) Natural   (b) Face   (c) Fingerprint

Figure 2. Complete Basis Functions for 8x8 blocks from different image datasets.

4. CODEC ARCHITECTURE

Transform coding can be applied either on full image or small blocks. Other implementations of MP normally use a full frame as a single block. It gives better compression without any blocking artifacts but it is not good for error resilience over noisy channels. Additively, computation is very heavy as MP is an iterative algorithm. To enhance error resilience capability alongwith reduced computational load, we have processed the image in blocks of 16x16 pixels. These blocks are encoded using variable number of coefficients until either of the stopping criteria is met which are minimum error threshold and maximum number of encoded coefficients. We have used MP and OMP to transform the image in ICA domain. These transform coefficients are then uniformly quantized. Lastly, entropy coding is performed to exploit the statistical redundancy of quantized coefficients as shown in Fig. 3.

Entropy coding is lot more sophisticated as compared to that used in previous ICA based works[3] and is performed in a number of steps. As ICA coefficients are very sparse and have lot of zeros, arithmetic coding cannot compress them efficiently. Thus, runlength coding has been employed firstly to produce separate streams of DC coefficients, AC coefficients and their indexes. The correlation among DC coefficients is exploited by using differential pulse code modulation (DPCM). Similarly indexes of the AC coefficients are also decorrelated by DPCM. In the end, all the three are coded using adaptive arithmetic coding.[16] Three different histograms, trained on a set of coefficients of the same image class, are used for DC coefficients, AC coefficients and indexes.

5. PERFORMANCE ANALYSIS

To measure the image quality, we have used both objective and subjective quality measures. Signal to noise ratio (SNR) and peak signal to noise ratio (PSNR) have been used as objective quality measures while picture quality scale (PQS)[17] has been employed for subjective quality measure. In contrast to SNR and PSNR which examines differences for every single pixel, PQS takes into account both local and global image features that are important for HVS. Mean opinion score (MOS) ranges from 0 (worst quality) to 5 (best quality). PQS provides a numerical measure of image quality well correlated with MOS at the middle range(2-3). For low quality images, it has negative value and for high quality images, it is greater than 5.
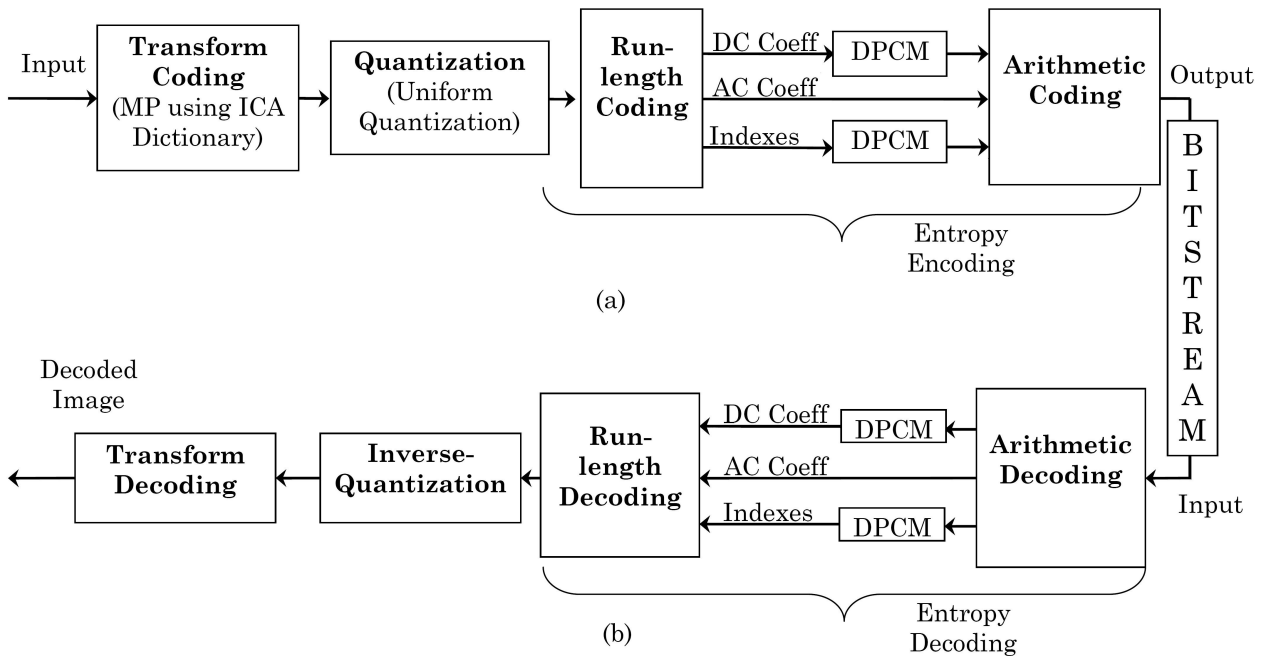
Figure 3. ICA based image compression system (a) Encoder, (b) Decoder.

5.1. Optimum ICA dictionary size

A comparison of incomplete, complete and overcomplete dictionaries for different image classes is presented as shown in Fig. 4. Incomplete dictionary contains 160 bases for 16x16 block. For Lena image, there is a difference of 6dB of PSNR when it is compressed by incomplete and 4-times overcomplete dictionaries. But for face and fingerprint images, this difference is only 2dB.

For different image classes, an optimum dictionary size depends on the diversity of structures in that image class. For natural images which contain very diverse structures, overcomplete dictionary performs better than complete and incomplete dictionaries. For database images which contain limited number of structures, an incomplete ICA dictionary performs well. As the angle between ICA bases is normally above 70 degrees, OMP gives only a little quality improvement for overcomplete bases but for incomplete dictionaries, the improvement in quality is up to 0.2dB of PSNR.

5.2. Feature capturing capability of ICA

ICA basis functions capture the image in order of importance of features. Image structure is captured first and then fine details are encoded. Fig. 5 shows the result when images are coded using only 2, 4, 6 and 8 NZs per block, at maximum. After just 4 coefficients per block, SNR is more than 21dB and image quality is good. In Fig. 5.a, maximum number of NZs is set to 2 which imply that whole of the 512x512 image can take 2048 coefficients but this image has taken only 1874 NZs with 19.34dB of SNR.

5.3. Quantization of ICA coefficients

Quantization is the only lossy part in the whole compression system and results in decrease of PSNR value. Quantization also produces blocking artifacts which becomes visible when quantized by a higher value. In contrast to DCT coefficients, ICA coefficients have less quantization artifacts. Being very sparse, every coefficient is important and represents some feature. There is considerable loss of quality if any coefficient becomes zero
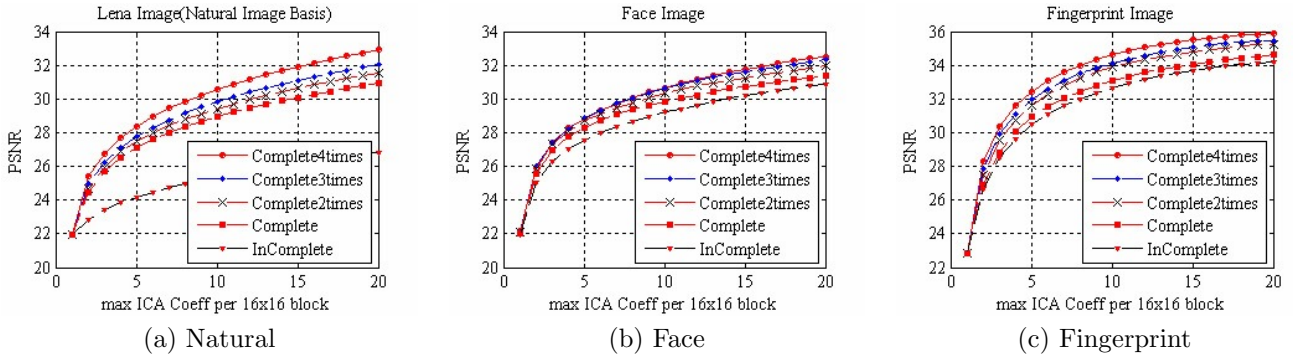
(a) Natural   (b) Face   (c) Fingerprint

Figure 4. Energy compaction of ICA bases with incomplete, complete and over complete dictionaries for both general and class specific bases functions
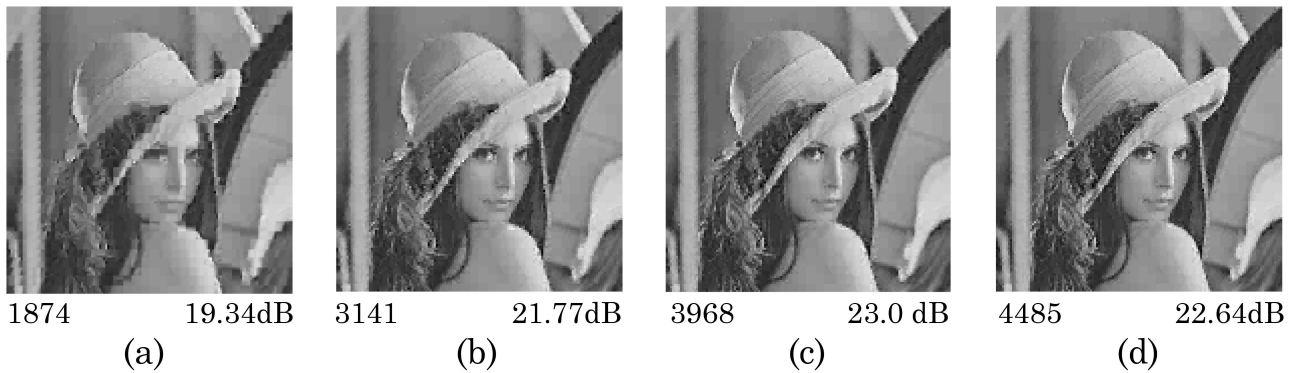

| 1874 | 19.34dB | 3141 | 21.77dB | 3968 | 23.0 dB | 4485 | 22.64dB |

(a)   (b)   (c)   (d)

Figure 5. Lena image encoded with (a) 2, (b) 4, (3) 6, (d) 8 NZs at maximum per 16x16 blocks. (Value at bottom left is NZs and SNR is in bottom right)

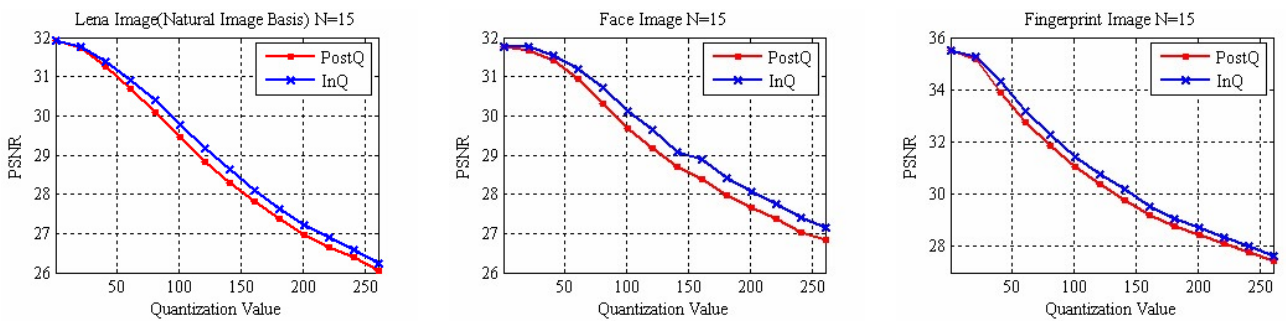because of quantization. Fig. 6 shows the decrease in PSNR value with increase in quantization.

Figure 6. Effect of quantization on the quality of ICA compressed image. The maximum number of allowed coefficients per block was set to be 15.

For iterative methods like MP, quantization can be done either inloop or a posteriori.[18] Fig. 6 shows the difference in quality when inloop and a posteriori quantization are used. Owing to non-orthogonal nature of overcomplete dictionaries, the quantization residual of one coefficient can be absorbed by other coefficients in inloop quantization. Hence inloop quantization performs better. When compared in terms of PSNR, the

difference between PSNR of images quantized by inloop and a posteriori is around 0.3dB. If it is analyzed in terms of PQS which is much correlated with HVS, the difference is about 0.2 over a scale of 5 which shows that the loss, because of a posteriori quantization, is considerable and should be avoided if possible.

5.4. Comparison with other codecs

Both for general and class specific compression, the images which are being compressed do not belong to group of images which are used to train the basis functions using ICA. They are not part of the image groups which are used to train histograms for adaptive arithmetic coding either.

5.4.1. General Image Compression

To show ICA capability to compress natural images, Lena image has been compressed with basis functions trained by ICA on a set of natural images* which contains 13 wildlife images. Natural images have a lot of diversity, so overcomplete dictionary gives better results. Lena image of 512x512 pixels is compressed with 2 times overcomplete ICA dictionary using MP as shown in Fig. 7. ICA based image compression system performs better than JPEG and comparable to JPEG2000.

	-1.27	-0.80	-1.09
Original	2308 27.98dB	2401 28.82dB	5600 27.44dB
(a)	(b)	(c)	(d)

Figure 7. Image compression comparison of Lena image (a) Original Image, Image compressed by (b) ICA, (c) JPEG2000, (d) JPEG (Image size is at bottom left(in bytes), PSNR is at bottom right and PQS is at top right).

5.4.2. Class specific image compression

To demonstrate ICA adaptability to class specific images, face† and fingerprint‡ images were compressed with ICA basis functions trained on a set of images from the respective databases. Incomplete ICA dictionary has been used with OMP for compression of a face image of 272x336 pixels and fingerprint image of 288x288 pixels. The results shown in Fig. 8 shows that ICA based image compression system performs better than JPEG and comparable to JPEG2000 for face images. For fingerprint image, ICA based image compression system outperforms both WSQ and JPEG2000 as shown in Fig. 9. Results recommend ICA based codec for compression of class specific images.

6. CONCLUSION

A novel set of basis functions and codec architecture, based on MP, has been evaluated and compared against standard codecs. The generalization ability of the ICA bases has been presented for natural images and its adaptation to specific class is explained for face and fingerprint images. For natural images which contain lot of diversity, overcomplete basis functions perform better than incomplete basis functions. In case of database

---

*www.cis.hut.fi/projects/ica/imageica

†www.bsp.brain.riken.jp/ICALAB/

‡http://www.cognaxon.com

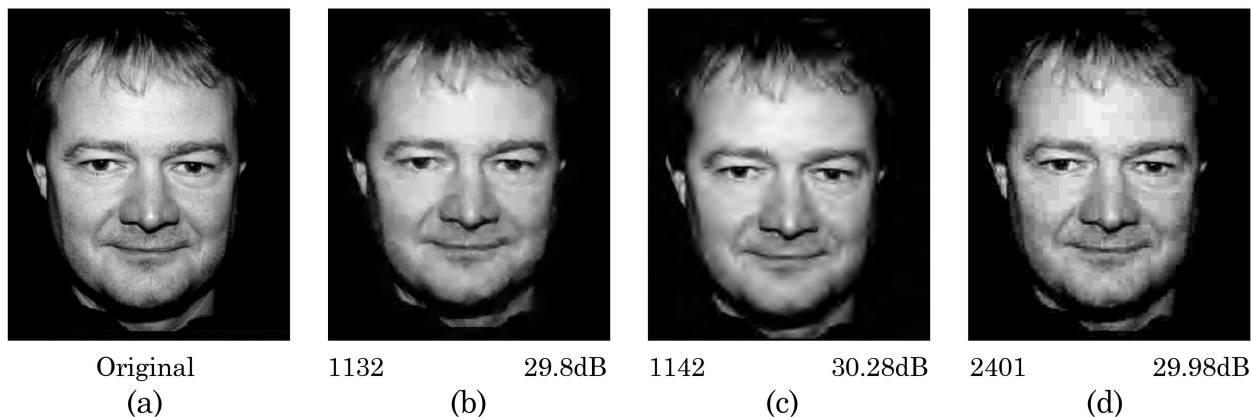| Original | 1132 | 29.8dB | 1142 | 30.28dB | 2401 | 29.98dB |
| (a) | (b) | | (c) | | (d) | |

Figure 8. Image compression comparison of face image (a) Original Image, Image compressed by (b) ICA, (c) JPEG2000, (d) JPEG (Image size is at bottom left(in bytes) and PSNR is at bottom right).
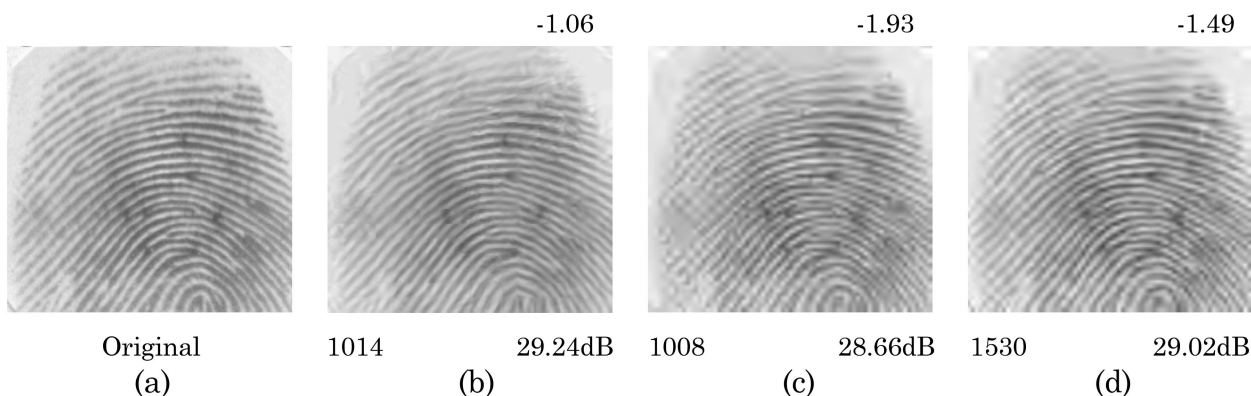

	-1.06		-1.93		-1.49	
Original	1014	29.24dB	1008	28.66dB	1530	29.02dB
(a)	(b)		(c)		(d)	

Figure 9. Image compression comparison of fingerprint image (a) Original Image, Image compressed by (b) ICA, (c) JPEG2000, (d) WSQ (Image size is at bottom left(in bytes), PSNR is at bottom right and PQS is at top right).

images, incomplete, complete, and overcomplete bases produce roughly the same results and hence, incomplete bases are recommended for database images.

ICA based codec has given promising results both for class specific and general images at low and medium bitrates. ICA based image compression system encodes image features in order of visual importance. Image structure is captured first and then the finer details are encoded. Images compressed with this codec contain less blocking artifacts as ICA bases are localized both in space and frequency and present lower complexity than JPEG2000. Based on the results, ICA based codec is recommended both for general and class-specific images.

ACKNOWLEDGMENTS

REFERENCES

1. P. Comon, "Independent component analysis, a new concept?," *Signal Processing* **36**(3), pp. 287–314, 1994.
2. A. J. Ferreira and M. A. T. Figueiredo, "Class-adapted image compression using independent component analysis," in *International Conference on Image Processing (ICIP) Barcelona, Spain (September 2003).*, **1**, pp. 625–8, Sept. 2003.

3. A. J. Ferreira and M. A. T. Figueiredo, "On the use of independent component analysis for image compression," *Signal Processing: Image Communication* **21(5)**, pp. 378–389, 2006.

4. D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*, Kluwer Academic Publishers, Dordrecht, 2001.

5. S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing* **41**(12), pp. 3397–3415, 1993.

6. S. S. Chen, *Basis Pursuit.* PhD thesis, Department of Statistics, Stanford University, Stanford , CA http://www-stat.stanford.edu/ schen/, 1995.

7. P. Vandergheynst and P. Frossard, "Efficient image representation by anisotropic refinement in matching pursuit," in *ICASSP '01*, pp. 1757–1760, (Washington, DC, USA), 2001.

8. A. Z. Averbuch and V. A. Zheludev, "A new family of spline-based biorthogonal wavelet transforms and their application to image compression," *IEEE Transactions on Image Processing* **13**, pp. 993–1007, July 2004.

9. A. J. Bell and T. J. Sejnowski, "The 'independent components' of natural scenes are edge filters," *Vision Research* **37**(23), pp. 3327–3338, 1997.

10. B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature* **381**, pp. 607–609, June 1996.

11. A. E. C. Pece, "The problem of sparse image coding," *J. Math. Imaging Vis.* **17**(2), pp. 89–108, 2002.

12. J. N. Bradley and C. M. Brislawn, "The wavelet scalar quantization compression standard for digital fingerprint images," in *ISCAS*, pp. 205–208, 1994.

13. M. Figueroa-Villanueva, N. Ratha, and R. Bolle, "A comparative performance analysis of jpeg 2000 vs. wsq for fingerprint image compression," in *AVBPA03, Audio- and Video-Based Biometric Person Authentication* , pp. 385–392, 2003.

14. E. Bingham and A. Hyvärinen, "A fast fixed point algorithm for independent component analysis of complex valued signals," *International Journal of Neural Systems* **10**, pp. 1–8, February 2000.

15. A. Hyvrinen, R. Cristescu, and E. Oja, "A fast algorithm for estimating overcomplete ica bases for image windows," in *Int. Joint Conf. on Neural Networks*, pp. 894–899, (Washington, D.C.), 1999.

16. A. Moffat, R. M. Neal, and I. H. Witten, "Arithmetic coding revisited," *ACM Trans. Inf. Syst.* **16**(3), pp. 256–294, 1998.

17. M. Miyahara, K. Kotani, and V. Algazi, "Objective picture quality scale (pqs) for image coding," *IEEE Transaction on Communication* **46(9)**, pp. 1215–1226, 1998.

18. P. Frossard, P. Vandergheynst, R. Figueras i Ventura, and M. Kunt, "A Posteriori Quantization of Progressive Matching Pursuit Streams," *IEEE Transactions on Signal Processing* **52**(2), pp. 525–535, 2004.