# Analysis of Spatio-chromatic Decorrelation for Colour Image Reconstruction

***Mark S. Drew and Steven Bergner***
***School of Computing Science,***
***Simon Fraser University***
***Vancouver, British Columbia***
***Canada V5A 1S6***
***{mark,sbergner}@cs.sfu.ca***

## Abstract

We investigate the implications of a unified spatio-chromatic basis for image compression and reconstruction. Different adaptive and general methods (PCA, ICA, and DCT) are applied to generate bases. While typically such bases with spatial extent are investigated in terms of their correspondence to human visual perception, we are interested in their applicability to multimedia encoding. The performance of the extracted spatio-chromatic spatial patch bases is evaluated in terms of quality of reconstruction with respect to their potential for data compression. The results discussed here are intended to provide another path towards perceptually-based encoding of visual data. This leads to a deeper understanding of the role played by chromatic features in data reduction.

## 1. Introduction

Decorrelation for redundancy reduction has a long history in image processing. In particular variants of the Principal Component Analysis (PCA) [1] for orthogonal decorrelation have been part of the arsenal of data reduction for many years. The main idea here is to account for most of the variance in the data using the first several principal axes, and then reduce the influence of further terms either by directly omitting these or by adopting a bit allocation scheme to deprecate their influence.

PCA can tell us how Nature processes vision, if we consider natural images. In particular, we expect to see colour-opponent channels arise in a natural fashion, simply by automatic inspection of the data. But as well, we hope to glean evidence of how spatial processing operates. And in fact, Ruderman et al. [2] found not only such colour-opponent structures but also spatial derivative-like filters, operating similarly and independently in each opponent-colour and luminance channel.

This work was based on a very simple scheme: first, using natural images derived from hyperspectral data and transformed to log space, Ruderman et al. formed small tiles, 3 pixels by 3 pixels, from an assembly of such images. Treating these 9 pixels as a vector, with each pixel containing 3-vector colour information, leads to 27-vector data for PCA analysis. Re-assembling the colour information, these colour patches could also be reconstituted as $3 \times 3$ colour squares, for viewing. In terms of spatial components, for each of the colour channels the spatial structure of the bases resembled the derivative-like and frequency-analysis-like structures arising in a Fourier analysis of greyscale images.

The latter result was not surprising (although the decorrelation from colour was): Olshausen and Field's seminal work on receptive field properties [3] implied that the receptive fields in mammalian primary visual cortex simple cells are spatially localized, oriented, and spatially band-pass in the sense of being selective to structure at different spatial scales, for non-colour luminance inputs. Visually, these fields resemble a 2-dimensional Discrete Cosine Transform basis in an $N \times N$ checkerboard structure (see, e.g., [4], and below), but with diagonal as well as rectangular basis images.

PCA has also been applied to non-interpolated, raw Colour Filter Array (CFA) data [5], with the result that the recovered basis finds only colour information, not luminance information, so is not appropriate for modelling spatial information.

## 2. Related Work

As opposed to an orthogonal PCA basis, some workers have also considered an Independent Component Analysis (ICA) of natural images [6]. ICA proceeds by producing a minimally redundant set of basis functions. To do so, a set of maximally statistically *independent* basis vectors is found. To understand what this means let us consider two stochastic variables $x_1$ and $x_2$, e.g. the coefficients of our data projected on two different axes. Their joint probability density is $p(x_1, x_2)$. The separate probability densities for $x_1$ or $x_2$, their so-called marginal probability, can be computed by $p_1(x_1) = \int p(x_1, x_2) dx_2$. Then $x_1$ and $x_2$ are independent if and only if

$$E\{h_1(x_1)h_2(x_2)\} = E\{h_1(x_1)\}\, E\{h_2(x_2)\}, \qquad (1)$$

where $E$ means expectation and $h_1, h_2$ can be essentially any two integrable scalar functions; the above is therefore a very strong condition. It says that any nonlinear transforms of the independent components are uncorrelated — their covariance is zero. In comparison, PCA decorrelates but does not guarantee independence. That is, projecting the data to the decorrelated axes the distribution of two resulting coefficients $b_1$ and $b_2$ fulfill

$$E\{b_1 b_2\} = E\{b_1\}\, E\{b_2\}. \qquad (2)$$

This corresponds to Eq.1 with $h_1, h_2$ being linear. However, PCA ("whitening") is still useful as a pre-processing step for ICA, and we use it that way here.

ICA has been found useful for data reduction by 'sparse' coding, i.e., finding underlying sources such that any given image is naturally represented in terms of just a small number of these: ICA [6] reproduces results for optimizing sparseness [3].

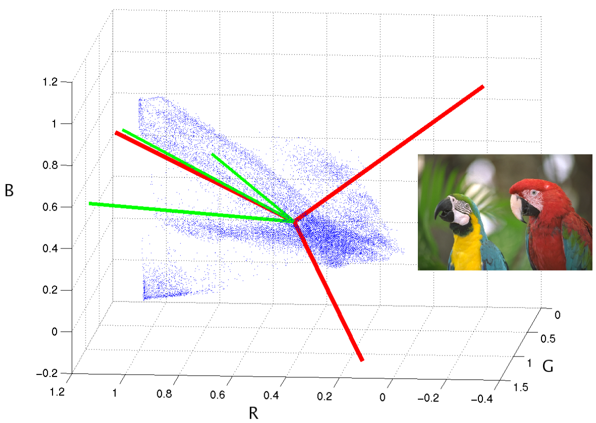



*Figure 1*: *New basis vectors for a given colour distribution from right image as found by PCA (red) and ICA (green).*

ICA is used for the extraction of hidden sources generating observed data. For example, consider the image in Fig. 1. The RGB values in this image form clusters, as in shown in the left of Fig. 1, with orthogonal PCA axes shown in red. In contrast, the ICA axes show that the image is actually comprised of just a few independent sources. It should be noted that ICA is data adaptive: we would like to develop a set of basis vectors that is different for each image. Of course, we could develop a universal set from a training set, but an adaptive model is bound to be more expressive for a particular image since, as we see from Fig. 1, the hidden characteristics of content are extracted.

The vocabulary used for ICA is somewhat different than that used for PCA. ICA is one way of solving the Blind Source Separation problem. Just as for PCA, if we have a feature-vector $\boldsymbol{x}$ , then if there are $k$ basis vectors we can approximate $\boldsymbol{x}$ as

$$\boldsymbol{x} \;\simeq\; \boldsymbol{B}\,\boldsymbol{b} \tag{3}$$

where the $k$ columns of $\boldsymbol{B}$ hold our ICA basis vectors, and $\boldsymbol{b}$ is the set of weights. In ICA, the matrix $\boldsymbol{b}$ is called the *mixing matrix* and $\boldsymbol{B}$ is a multidimensional stochastic variable of independent *sources*. E.g., in PCA, vector $\boldsymbol{x}$ could be a greyscale image, and $\boldsymbol{B}$ would consist of eigenimages. However, for ICA the columns of $\boldsymbol{B}$ are not orthogonal. Therefore to find the weights $\boldsymbol{b}$ we must use the Moore-Penrose pseudoinverse of $\boldsymbol{B}$ , applied to the target image. The pseudoinverse is referred to as the set of ICA *filters*. ICA has been used for multimedia data fusion as well [7, 8] — in this case, ICA recovers common latent subspaces for combined media. Typically, however, PCA is used [9, 10], applied to much smaller feature spaces than pixel-level.

The process of finding ICA vectors is based on the Central Limit Theorem, which states that a sum of non-Gaussian random variables is more like a Gaussian than are its individual components. But the independent sources sought can be written as sums of the observed data. Thus we can move toward independent sources by trying to find a sum of the observed data over vectors which have maximum non-Gaussianity. The latter property can be characterized by the kurtosis (the fourth-order cumulant), which is zero for a Gaussian. However, a more robust measure is formed by the negentropy, the difference between the entropy for a Gaussian and that for the current basis, using the observation that for a given standard deviation $\sigma$, a Gaussian $G_\sigma$ has maximum entropy compared to other probability distributions.

Thus one arrives at a gradient descent method for determining the IC basis, and since this can be phrased as a fixed-point problem, mechanisms similar to the Contraction Mapping Theorem can be brought to bear for existence, uniqueness, and convergence rate. The method we use here is the FastICA algorithm [11], which converges in cubic time.

For greyscale imagery [3, 6], PCA indicates that a mutually orthogonal spatial basis for imagery consists of band-pass filters similar to 2-dimensional Discrete Cosine Transform (DCT) basis images, but with some non-rectangular orientation present (cf. Fig. 2). How one creates such an image is by randomly selecting $N$-pixel by $N$-pixel square patches from a greyscale image, vectorizing these $N^2$ values, and identifying the basis as the eigenvectors of the mean-subtracted covariance matrix. In contrast, ICA of greyscale imagery produces basis functions that are again bandpass, but are more obviously oriented and are similar to Gabor functions — Gaussian-windowed sine waves [3, 6].

### 2.1. Basis functions generated by ICA on natural images

A survey of applications of ICA to the processing of different media (image/video, multimodal brain data, audio, text, and combined data) is provided by [12]. In their survey it becomes apparent that ICA has been widely used for classification but implications of ICA for multimedia compression have not been studied as of yet.

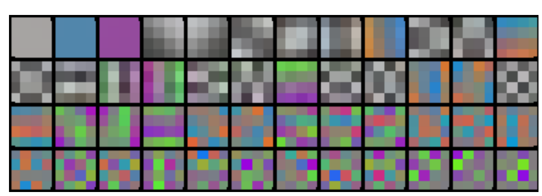


*Figure 2*: *Spatio-chromatic basis obtained from PCA on $4 \times 4 \times RGB$ image patches of the example in Fig. 1.*

When colour is included, our patch vectors become $N \times N \times 3$ structures, for RGB images. PCA proceeds as stated above, but for these longer vectors, and the resulting structure can be visualized as in Fig. 2 (for $N = 4$). Note that the PCA basis is adaptive to the image (but in fact does not change much, for natural imagery we have tried).

In [2], Ruderman et al. extend a 3D colour PCA, as in the red vectors in Fig. 1, to a spatial patch domain by using

$3 \times 3$ patches of 3-vector colour data. In comparison, in a sense Fig. 1 shows results for $1 \times 1$ patches. Ruderman et al. conclude that for foliage images, PCA of log L,M,S data tends to decorrelate spatial processes from chromatic ones, leading to 9 spatial features times 3 colour ones. The latter are, in order, luminance, blue-yellow, and red-green. This result was extended by Wachtler et al. [13, 14] by replacing PCA with ICA, again for L,M,S data but now using $7 \times 7$ patches.

Colour images and stereo vision have been investigated by [15] showing that the derived independent components also yield a separation of basis vectors into luminance and opponent colours.

Besides allowing for conclusions regarding human visual perception, these chromatic bases with spatial extent are very interesting from an image compression point of view. In the following we are going to take a closer look at the implications of encoding visual data with respect to these bases.

## 3. Data specific basis functions

The goal of our analysis is to compare the suitability of different data-adaptive basis functions for compressing visual data. Therefore, a set of colour images is chosen that spans a variety of outdoor scenes containing plants, animals, humans and artificial objects. The perspectives of the images range from detail shots to panoramas[1].

In the following we will consider three different sets of bases — two data-adaptive methods, namely PCA and ICA, and one general basis, the discrete cosine transform DCT, moved here into a colour domain such that it resembles the PCA.

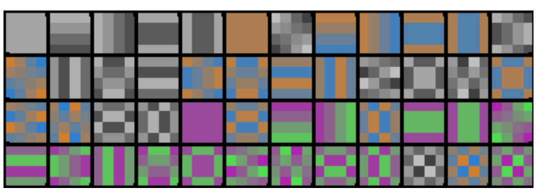


*Figure 3*: *Basis patches for DCT decomposition of spatio-chromatic $4 \times 4 \times RGB$ patches. Sorted in order of decreasing variance-accounted-for from left to right and top to bottom.*

The DCT basis is a descendant of the Fourier transform. To apply it to the spatio-chromatic setting we simply treat the two spatial patch extents and the index of RGB colour components as a three dimensional rectangular prism, e.g., a $3 \times 3 \times 3$ cube. Fig. 3 shows the resulting basis re-assembled as a colour picture. Note the similarity of this artificially generated basis to that obtained by PCA in Fig. 2. The most prominent difference is a slightly changed orientation of the directional frequencies that are axis-aligned in Fig. 3. A similar explanation applies to the fact that the main colour axes are chosen differently. Also note that the ordering by the variance-accounted-for is quite different because of the different alignment of the basis. In the DCT case the pure colour dimensions appear later in the sequence after several luminance frequencies. In the PCA case all three of them appear as the most significant vector.

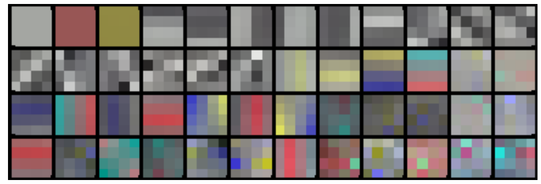


*Figure 4*: *Basis patches for ICA of spatio-chromatic $4 \times 4 \times RGB$ patches. Sorted in order of decreasing variance.*

Fig. 4 shows the result of ICA performed on the image set, for a particular patch size. While the results of DCT and PCA can be interpreted as a frequency decomposition of the data, the functions obtained by ICA exhibit a combined localization in space and frequency [15]. Apparently, this basis seems to again treat opponent colour separately from luminance — similar to the observation that is already illustrated in Fig. 2 for the PCA case.

Besides deciding on a method of basis generation, we have to make a choice about the size of the patches we will operate on. To create a basis, the analysis is performed on squared pixel neighbourhoods. We randomly sampled a total of 50000 patches over the images of the given set. The resulting basis functions then reflect the statistical properties of the presented data.

To use the basis for reconstruction, the images are regularly tiled into an arrangement of non-overlapping patches. As mentioned above in the discussion of Eq. 3, the coefficients for each patch of the image can be obtained by a linear transform using the filter patches. These are essentially the inverse of the basis patches. Respectively, going back from the coefficients to the actual image data is done by transforming the coefficients in a linear combination of the basis patches.

### 3.1. Variance thresholding and quantization of coefficients

After having projected the image data to the new basis the resulting coefficients have to be reduced in some way. If no reduction takes place no compression will apply. Therefore, we have implemented two different methods of reduction: pruning of components by variance thresholding and alternatively variance-based quantization.

#### *3.1.1. Variance thresholding*

The variance thresholding is the same method commonly applied in dimensionality reduction using PCA. The overall variance of the data is taken as the sum of the variances corresponding to each basis vector. This measure falls out from the eigenvalue decomposition of the covariance matrix. Another way of obtaining it is by looking at the variance of the coefficients after projecting the data to the basis. In order of decreasing variance, just as many basis vectors are preserved as needed to retain to a user-specified ratio of the total variance. The rest are considered to be insignificant and are set to zero, i.e. the coefficients

[1]We have used the 23 Kodak stills from CIPR at `http://www.cipr.rpi.edu/`

do not need to be stored. This is essentially the amount of compression that arises from this method.

ICA has a similar behaviour to PCA in terms of distributing variance among basis vectors. Its particular strength lies in minimising the mutual information among basis vectors or a sparse coding of coefficients. This advantage is particularly relevant when applying entropy encoding to the coefficients.

#### *3.1.2. Entropy coding*

The coefficients are given as continuous (floating point) numbers. To apply entropy coding, we have to quantize them first. An overview of different techniques for scalar and vector quantization is given in [16]. In our tests we have decided for the method of assigning each basis vector a number of bits proportional to its standard deviation. The proportionality factor is chosen to fulfil a given overall contingent of bits. Given the number of bits, each channel of coefficients is uniformly quantized from its minimum to its maximum occurring value, and ensuring the existence of one bin for zero.

Now, having expressed the coefficients as discrete numbers, we can apply an entropy-based encoding, e.g. Huffman coding, a form of variable length coding (VLC). The VLC compression indicated in the graphs below is a theoretical limit that can be computed from the sum of entropies for all channels by exponentiation to the base 2. The rate of compression then is the factor by which the estimated encoded data is smaller than the original data, which has been stored with 8 bits per channel.

### 3.2. Quality of reconstruction

Besides looking at the size of the data after compression, we most importantly have to look at the quality of the reconstruction obtained from the reduced representations. A measure commonly used to assess this quality is the peak signal to noise ratio (PSNR), in decibels (dB)

$$PSNR = 10\log_{10}\left(\frac{G^2}{MSE}\right) \quad (4)$$

$$MSE = \frac{\sum_i^N \sum_j^M \sum_{k\in RGB}(p_k(i,j)-o_k(i,j))^2}{3MN} \quad (5)$$

where $G$ is the maximum representable value (e.g. 255) and $MSE$ is the mean squared error of the reconstructed picture $p_k(i,j)$ and the original $o_k(i,j)$. The entire term in the logarithm is inversely related to the reconstructed variance that we have referred to in the previous section when considering variance thresholding. A drawback of the PSNR is that it does not actually reflect the distortion as *perceived* by a human observer. Nevertheless, it is a convenient measure to illustrate the quality of reconstruction in a physical sense, and shall be sufficient for our subsequent evaluation.

## 4. Evaluation of different basis sets

The efficiency of a basis is understood as the relation between image quality retained for an achievable rate of compression (or vice versa). Thus, we have conducted a number of tests for different sets of bases (ICA, PCA, and DCT). Each basis is generated and applied separately over range of squared patch sizes from $1\times1$ to $16\times16$. The first case is similar to just interpreting the pixel colours. As patch size increases, the influence of neighbours is included more and more. Another variable in the comparison is the compression parameter. In the case of variance thresholding this is the ratio of reconstructed variance, and for the entropy coding this is the overall maximum number of bits for the stream. Note that the actual achievable compression then entirely depends on the entropy of the data.

### 4.1. Compression vs. quality using spatio-chromatic bases

The test results illustrated in Fig. 5(a,b) are for reducing the data by variance thresholding. The edge length of the square basis patches is given in pixels. The compression rate is the factor by which the encoded data is smaller than the original image. The comparison is made between colour and greyscale images, where in the first the idea of a spatio-chromatic basis is applied. The second test simply applies a greyscale PCA basis on a greyscale version of the images. The major observation that can be made is that the iso-lines of compression move towards higher PSNR for larger patch sizes. This trend is significantly stronger on the left, with colour included. Also, colour helps quality, compared to greyscale at equal compression. DCT is not shown since it is similar to PCA.

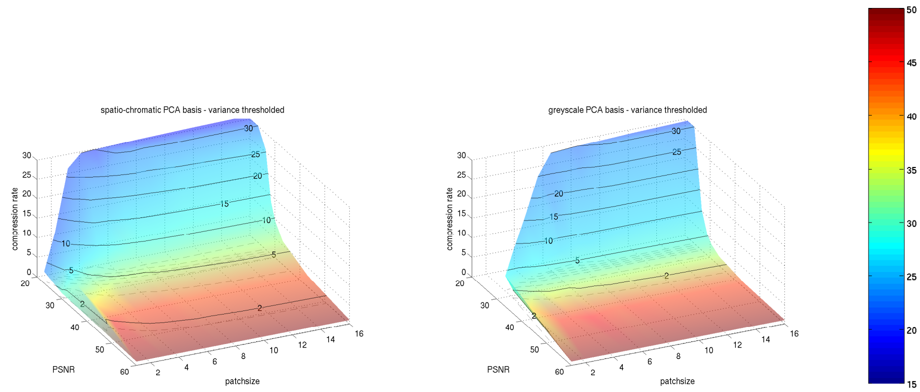
Fig. 6(a,b) provides a comparison of the entropy-based variable length coding of the spatio-chromatic coefficients of ICA vs. PCA. Here the property of ICA to result in sparsely coded coefficients becomes apparent. The lower entropy of the quantized data results in significantly higher compression rates. Nevertheless, as the surface for ICA is more bent to the back we note that the PSNR for the mid-range compression rates is lower than for PCA. Both plots show a significant improvement of the compression/error tradeoff as the patch size increases.

### 4.2. Performance of a specialized basis

The previous examples have shown the quality of reconstruction for general bases that were obtained from the entire set of images. The following analysis considers just one image (shown in Fig. 8). This allows us to tailor a *specialized* basis for this particular image.

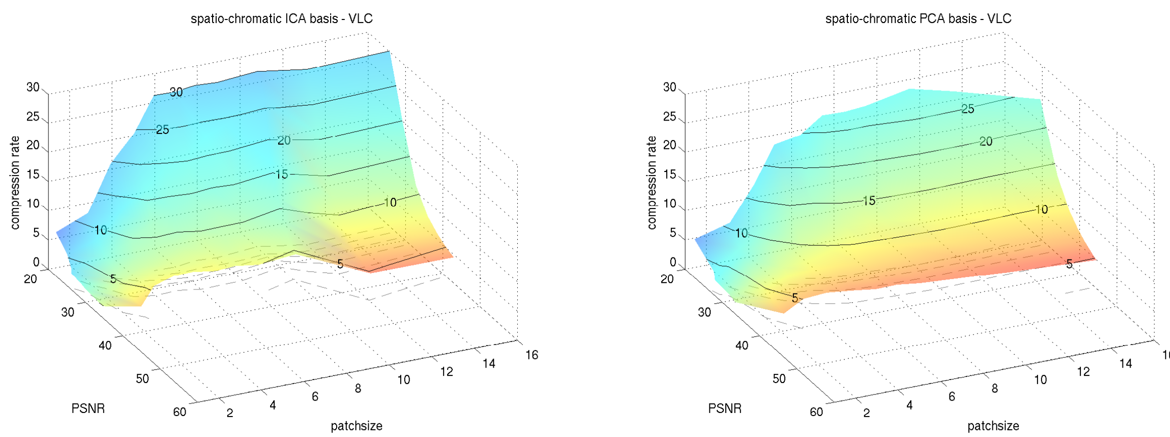
In the specific encoding test in Fig. 7, ICA performs significantly better than PCA for large patch sizes. It achieves higher compression for the same quality of reconstruction. Also, a trend can be noted that ICA tends toward increased quality with larger patch sizes while PCA roughly stays constant. The DCT statistics for this image are not included here because they are very similar to the outcome of the PCA for patch sizes larger than 4.

The image these statistics have been performed on is shown in Fig. 8a. The example shown is obtained after compression with a specialized ICA basis: this basis has been extracted from patch samples taken only from this particularly image. A visual comparison in Fig. 8b shows a detail of the left image. The lower half is the same image compressed to the same projected file size (1:12) using a



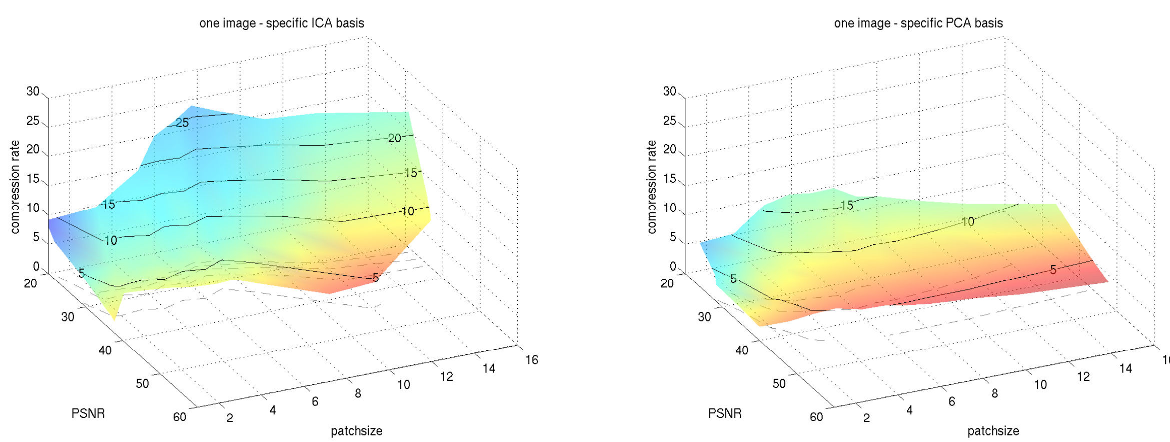

(a) (b) (c)

*Figure 5*: *Variance-thresholded compression using PCA generated (a) spatio-chromatic and (b) greyscale basis functions; colour indicates PSNR as per (c).*




(a) (b)

*Figure 6*: *Entropy-based compression of colour images using (a) ICA and (b) PCA generated spatio-chromatic basis functions.*




(a) (b)

*Figure 7*: *Entropy-based compression of Fig. 8 a) specific ICA, b) specific PCA.*


(a) (b)

*Figure 8*: *(a): Original image. (b): Compression ratio 1:12 using an image specific ICA basis (PSNR* 35.55*), and DCT compression (PSNR* 31.97*). Both are for* $16 \times 16$ *patches.*

DCT basis of the same block size ($16 \times 16$). DCT exhibits strong blocking artifacts while these are hardly noticeable in the ICA version.

## 5. Summary

The computation of individual bases for restricted sets of images is interesting from a vision and a multimedia point of view. While the first point has been subject of previous work targeting analogies to human perception, we have tried to illuminate the latter. The results indicate a significant difference comparing the compressibility of coefficients from ICA and PCA. The sparse coding property of ICA bases has been shown to have a noticeable impact on the efficiency of subsequent entropy compression.

A problem inherent in the approach of specialized bases is that they first have to be generated in a computationally expensive preprocessing. Furthermore, a basis specific to one data set would have to be stored along with the coefficients to allow for decoding. This would certainly add overhead to the compressed data. Nevertheless, in a constrained domain it is possible to prepare basis functions that can be reused.

In future we plan the extension of the current framework to video data. The inclusion of motion will result in a temporal spatio-chromatic basis. As well as revealing implications for video compression, this may indeed also bear relation to human perception.

Because of the proximity of the outcome of independent component analysis to receptive fields of simple cells in the V1 visual cortex, it could be possible to derive a more perceptually-based error metric for evaluation of the quality of visual representations. Advances of research in the human perceptual system may lead the way to an error metric that more closely corresponds to the assessment by a human observer.

Another interesting property of the ICA basis is that it resembles expressive features of the image. This property also hints at the relationship between ICA filters and wavelet analysis. Taking this into account, it seems worthwhile to consider the compressed coefficients as a higher-level feature description of the image.

## References


[1] H.S. Fairman and M.H. Brill. The principal components of reflectances. *Color Research and Application*, 29:104–110, 2004.

[2] D.L. Ruderman, T.W. Cronin, and C.-C. Chiao. Statistics of cone responses to natural images: Implications for visual coding. *J. Opt. Soc. Am.*, 15(8):2036–2045, August 1998.

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

[4] Z.N. Li and M.S. Drew. *Fundamentals of Multimedia*. Prentice-Hall, 2004.

[5] D. Alleysson and S. Süsstrunk. Spatio-chromatic PCA of a mosaiced color image. In *2nd European Conf. on Color in Graphics, Imaging and Vision*. Society for Imaging Science & Technology (IS&T), 2004.

[6] J.H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. Roy. Soc. B*, 265:359–366, 1998.

[7] J.-Y. Pan and C. Faloutsos. Videocubes: A novel tool for video mining and classification. In *Int. Conf. on Asian Digital Libraries (ICADL)*, 2002.

[8] J. Larsen, L.K. Hansen, T. Kolenda, and F.A. Nielsen. Independent component analysis in multimedia modeling. In *4th Int. Symp. on Independent Component Analysis and Blind Source Separation (ICA)*, 2003.

[9] E. Sahouria and A. Zakhor. Content analysis of video using principal components. *IEEE Trans. on Circ. and Sys. for Video Tech.*, 9:1290–1298, 1998.

[10] M.S. Drew and J. Au. Clustering of compressed illumination-invariant chromaticity signatures for efficient video summarization. *Image and Vision Comput. J.*, 21:705–716, 2003.

[11] A. Hyvärinen and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Computation*, 9(7):1483–1492, 1997.

[12] J. Larsen, L. K. Hansen, T. Kolenda, and F. AA. Nielsen. Independent component analysis in multimedia modeling. In *Fourth Intl. Symposion on ICA and BSS*, pages 687–696, Nara, Japan, 2003.

[13] T. Wachtler, T.-W. Lee, and T.J. Sejnowski. Chromatic structure of natural scenes. *J. Opt. Soc. Am. A*, 18:65–77, 2001.

[14] T.-W. Lee, T. Wachtler, and T.J. Sejnowski. Color opponency is an efficient representation of spectral properties in natural scenes. *Vis. Res.*, 42:2095–2103, 2002.

[15] P.O. Hoyer and A. Hyvrinen. Independent component analysis applied to feature extraction from colour and stereo images. *Network: Computation in Neural Systems*, 11(3):191–210, 2000.

[16] P.C. Cosman, R.M. Gray, and M. Vetterli. Vector quantization of image subbands: A survey. *IEEE Trans. on Image Proc.*, 5(2):202–225, 1996.