Compression of Quincunx Subbands

JOSÉ ALEXANDRE NALON, JOÃO BAPTISTA T. YABU-UTI
Unicamp – Universidade Estadual de Campinas
DECOM – Departamento de Comunicações
Av. Albert Einstein, 400 – CEP 13081-970 – Campinas SP
{nalon, yabuuti}@decom.fee.unicamp.br

Abstract. Quincunx downsampling splits the spectrum of an image in two diamond-shaped subbands. This allows the information to which the human eye is most sensitive to be preserved, and also requiring less bits in a lossy compression. In this paper, we develop an algorithm that performs successive quincunx subband decompositions, thus allowing multi-level decompositions and efficient coding. Theory of lattices is briefly reviewed, and simulations using a simple subband coder were carried out. Results obtained are compared to the rectangular separable subband decomposition, showing that, without exploring redundancies due to quincunx simmetries, both strategies deliver simmillar results.

1 Introduction

In the last decade, subband coding, a coding technique started by Crochiére[1] and sometimes refered to as multiresolution analysis, has become greatly popular, given its simplicity, versatility and the broad spectrum of applications it provides, such as signal compression, feature detection and extraction, frequency detection and many more. Relation between wavelets and subband coding was revealed by Mallat, who also developed the successive decompositions algorithm, which increased efficiency of this technique.

Following it, Vetterli[2] expanded the concept to multidimensional signals, which made possible to use this technique with images. Concepts developed by Vetterli were extended by Woods and O'Neil[3]. These works only dealed with separable sampling of images. Viscito and Allebach[4] extended the analysis to arbitrary sampling lattices. Currently, non-separable sampling lattices have gained much attention [5, 6].

The use of subband coding in image compression was also investigated by a great number of authors, and also a great number of possibilities appeared. Some of the most efficient methods of image compression nowadays use subband coding, including the new standard JPEG 2000. We refer the reader to [7] for an extensive survey on the subject.

In this paper, we analyse quincunx subband coding of images. Quincunx subbands are non-separable diamond-shaped frequency bands. Since spatial frequencies to which human eye is most sensitive are located in the frequecy axes [8, 9], quincux subbands preserve information that is most relevant to the human eye. Also, mathematically, most of the analysis is a direct extension of the one-dimensional case.

One problem with quincunx downsampling is the nonexistence of an algorithm to perform successive decompositions. Here, we try to solve this problem by developing a technique that will allow any image to be decomposed at any desired level.

This paper is organized as follows: in section 2 we review the necessary concepts of multidimensional sampling with arbitrary lattices, and apply these concepts to the quincunx case; in section 3, we derive the conditions to design of perfect reconstruction filter banks for quincunx subbands; in section 4, we show how to perform successive decomposition using quincunx downsampling of images; in section 5 we apply the concepts to image compression; and in section 6 we present our conclusions and remarks.

2 Quincunx Sampling

Let $x_c(\mathbf{t})$, where $\mathbf{t} = [\ t_1\ t_2\ ]^T$, be a two-dimensional function in the space domain, and $X_c(\boldsymbol{\omega})$, where $\boldsymbol{\omega} = [\ \omega_1\ \omega_2\ ]^T$, be its representation in the frequency plane. The discretization operation of this function consists of obtaining samples of its values at fixed intervals of space. Given two linearly independent vectors $\mathbf{d}_1$ and $\mathbf{d}_2$ which define two directions along which samples are taken, we define $\mathbf{D} = [\ \mathbf{d}_1\ \mathbf{d}_2\ ]^T$ as the sampling matrix, which generates the sampling lattice of the discrete version of $x_c(\mathbf{t})$. We can express this operation with the equation:

$$x[\mathbf{n}] = x_c(\mathbf{Dn}) \qquad 1$$

where $\mathbf{n} = [\ n_1\ n_2\ ]^T$ is a vector of integer numbers. The most common sampling lattice is the separable one, which is described by the vectors $\mathbf{d}_1 = [\ 1\ 0\ ]$ and $\mathbf{d}_2 = [\ 0\ 1\ ]$. The sampling sublattice is determined by $\mathbf{D}$ as all the vectors $\mathbf{Dn}$. The coset of a lattice is the set of points obtained by shifting the entire lattice by an integer vector $\mathbf{k}$. For a given lattice, there will be always $N = |\det \mathbf{D}|$ cosets. In the case of subband coding, N will also determine the number of subbands in each decomposition.

It is possible to relate frequency representation of the discrete function $x[\mathbf{n}]$ with the representation in frequency

of $x_c(\mathbf{t})$ [8]. By combining both representations and Equation 1, we find

$$X(\boldsymbol{\omega}) = \frac{1}{N}\sum_{\mathbf{l}} X_c\left(\mathbf{D}^{-T}\boldsymbol{\omega} - 2\pi\mathbf{D}^{-T}\mathbf{l}\right) \qquad 2$$

where $\mathbf{D}^{-T}$ denotes the inverse transpose of $\mathbf{D}$.

This result is a direct extension of what is found in [10]. It is important to observe that Equation 2 represents the replication of the spectrum of x_c along the points represented by $2\pi\mathbf{D}^{-T}\mathbf{l}$, which shows that aliasing may occur if the signal is not band-limited. It is possible to show that aliasing will occur if the signal has components outside the region defined by

$$\left\{\mathbf{D}^{-T}\boldsymbol{\omega} : \boldsymbol{\omega} \in [-\pi,\pi]\times[-\pi,\pi]\right\} \qquad 3$$

The downsampling operation of a discrete signal on an arbitrary lattice is represented by

$$y[\mathbf{n}] = x[\mathbf{D}\mathbf{n}] \qquad 4$$

Following the same reasoning used in Equations 1-3, we can relate the spectrum of $y[\mathbf{n}]$ with that of $x[\mathbf{n}]$. We will find that

$$Y(\boldsymbol{\omega}) = \frac{1}{N}\sum_{i=0}^{N-1} X\left(\mathbf{D}^{-T}\boldsymbol{\omega} + 2\pi\mathbf{D}^{-T}\mathbf{k}_i\right) \qquad 5$$

where $\mathbf{k}_i$ is the integer vector that represents the i-th coset of the sampling lattice.

The upsampling operation is defined as the inverse operation

$$y[\mathbf{n}] = x\left[\mathbf{D}^{-T}\mathbf{n}\right] \qquad 6$$

with the constraint that this operation is only valid where $\mathbf{D}^{-T}\mathbf{n}$ maps to valid points of $x[\mathbf{n}]$, being defined as 0 when this doesn't occur. Its frequency representation is given by

$$Y(\boldsymbol{\omega}) = X\left(\mathbf{D}^{T}\boldsymbol{\omega}\right) \qquad 7$$

For a further discussion of effects of arbitrary sampling on images, we refer the reader to [4] and [8].

We are interested in quincunx sampling, which is described by vectors $\mathbf{q}_1 = [\ 1\ 1\ ]$ and $\mathbf{q}_2 = [\ 1\ -1\ ]$. We define the sampling matrix $\mathbf{Q} = [\ \mathbf{q}_1\ \mathbf{q}_2\ ]^T$, below, as the quincunx sampling matrix

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad 8$$

Quincunx sampling consists in sampling with diagonal regular periodicity, as shown in Figure 1. The vectors $\mathbf{q}_1$ and $\mathbf{q}_2$ defines the directions shown in the figure as t_1' and t_2', respectively. It is easy to see that diagonal periodicity of this sampling strategy is equivalent to rectangular sampling in a set of axis rotated 45°.

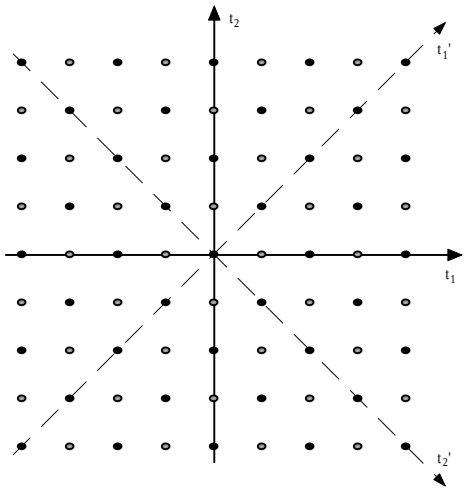


Figure 1: Quincunx sampling lattice. Shaded points represent the second coset of the lattice.

Since |det $\mathbf{Q}$| = 2, the cosets of the quincunx sampling lattice can be represented by two vectors

$$\mathbf{k}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \mathbf{k}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad 9$$

That also means that each decomposition of the signal will result in two subbands. We substitute the definitions in Equations 8 and 9 in Equation 5 to obtain the frequency representation of the quincunx downsampled signal from a rectangular one, and obtain

$$X_Q(\boldsymbol{\omega}) = \frac{1}{2}\left[X\left(\mathbf{Q}^{-T}\boldsymbol{\omega}\right) + X\left(\mathbf{Q}^{-T}\boldsymbol{\omega} + \boldsymbol{\pi}\right)\right] \qquad 10$$

where $\boldsymbol{\pi} = [\ \pi\ \pi\ ]^T = 2\pi\mathbf{Q}^{-T}\mathbf{k}_1$. The second term inside the brackets in Equation 10 is the aliased component in the quincunx downsampling, which needs to be made null in the reconstruction process. From that, and also from Equation 3, it's easy to see that aliasing will not occur if the spectrum of the signal is band-limited according to

$$|\omega_1| + |\omega_2| < \pi \qquad 11$$

Equation 11 represents a diamond shaped spectrum, as can be seen in Figure 2.

3 Filter Design

Tipically, a subband decomposition system consists of a filter bank followed by downsampling by the appropriate sampling periods, generating one subband for each analysis filter $H_i(\boldsymbol{\omega})$ in the filter bank. Each signal is then processed – *eg*, compressed, used for feature detection or extraction, transmitted, etc. – and after processing, if needed, the original signal is reconstructed from its subbands. The reconstruction process is a combination of

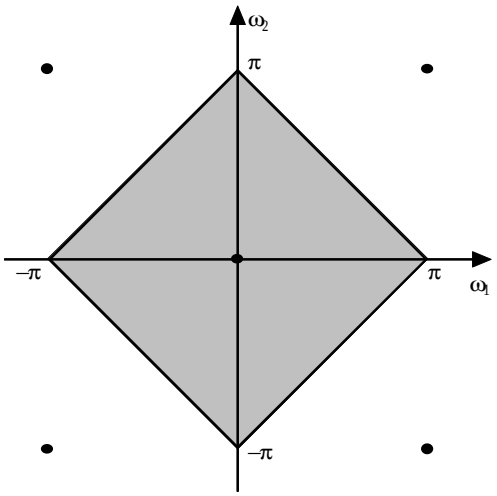


Figure 2: Frequency band of a quincunx downsampled image.

the upsampled and filtered version of all the subbands.

The filters $G_i(\omega)$ in the reconstruction portion are called synthesis filters. If the reconstructed signal at the end of the process is a perfect copy of the original, we say that perfect reconstruction is provided.

Figure 3 shows the subband decomposition of a signal for a quincunx sampling lattice. The symbol Q↓ represents quincunx downsampling, and Q↑ represents quincunx upsampling. There will be only two subbands, so only two filters are needed in the filter bank.

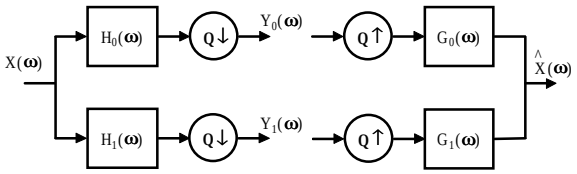


Figure 3: Quincunx subband decomposition and reconstruction of a signal

Each subband $Y_i(\omega)$ resulting from the analysis of the signal will be given by

$$Y_i(\omega)=\frac{1}{2}\left[H_i\left(\mathbf{Q}^{-T}\omega\right)X\left(\mathbf{Q}^{-T}\omega\right) + H_i\left(\mathbf{Q}^{-T}\omega+\pi\right)X\left(\mathbf{Q}^{-T}\omega+\pi\right)\right] \quad (12)$$

Following directly from Figure 3, the reconstructed signal will be

$$\hat{X}(\omega)=G_0(\omega)Y_0\left(\mathbf{Q}^{T}\omega\right)+G_1(\omega)Y_1\left(\mathbf{Q}^{T}\omega\right) \quad (13)$$

Combining Equations 12 and 13, we get

$$\hat{X}(\omega)=\frac{1}{2}[H_0(\omega)G_0(\omega)+H_1(\omega)G_1(\omega)]X(\omega) + \frac{1}{2}[H_0(\omega+\pi)G_0(\omega)+H_1(\omega+\pi)G_1(\omega)]X(\omega+\pi) \quad (14)$$

From Equation 14 is possible to derive the conditions of perfect reconstruction for a filter bank. The first parcel of the right hand of the equation represents the desired part of the signal, while the second parcel represents the aliased terms. If we satisfy

$$H_0(\omega)G_0(\omega)+H_1(\omega)G_1(\omega)=e^{-j\omega_0} \quad (15)$$

$$H_0(\omega+\pi)G_0(\omega)+H_1(\omega+\pi)G_1(\omega)=0 \quad (16)$$

then the aliasing will be cancelled and the signal will be perfect reconstructed. The term in the right hand of Equation 15 means that the signal may suffer a delay. We will be interested in situations where the delay equals zero in both axes, to simplify the analysis.

The choice of filters that satisfy Equations 15 and 16 are numerous and difficult. There are numerous techniques to filter design, including, but not limited to, optimization of a cost function [4, 11, 12], mathematical definitions and simmetries [6, 13], and transformations on one-dimension filters to generate a two dimensional one [14, 15]. For any technique, the conditions stablished in Equations 15 and 16 are enough to obtain a perfect reconstruction filter bank. Also, note that solving those equations will also satisfy the anti-aliasing condition determined by Equation 11.

A classical solution to this set of equations is given by the quadrature mirror filters (QMF) [16, 17]. By making

$$\begin{aligned} H_1(\omega)&=H_0(\omega-\pi)\\ G_0(\omega)&=H_0(\omega)\\ G_1(\omega)&=-H_1(\omega) \end{aligned} \quad (17)$$

it can be proved that aliasing will be cancelled. Thus, it is needed to design only the lowpass analysis filter, and the others can be derived from it. This choice of filters leads to the following cost function:

$$E=\int_{passband}\left(1-|H_0(\omega)|\right)^2 d\omega+\int_{stopband}|H_0(\omega)|^2 d\omega \quad (18)$$

where passband, for a low-pass quincunx filter, is given by Equation 11, and stopband is its complement in the frequency plane (*ie*, $|\omega_1|+|\omega_2|>\pi$). Equation 18 is a direct extension of [1].

Cost function optimization can benefit from simmetry caused by the quincunx sampling. It is possible to divide the filter mask in rings [6] where the values of

the coefficients are the same. This simmetry can also be applied while implementing the quincunx subband decomposition, raising the computational efficiency of the algorithm. Figure 4 shows this simmetry for odd length filters, but it is also possible to apply similar simmetries to even-length filter masks.

15	14	13	12	11	12	13	14	15
14	10	9	8	7	8	9	10	14
13	9	6	5	4	5	6	9	13
12	8	5	3	2	3	5	8	12
11	7	4	2	1	2	4	7	11
12	8	5	3	2	3	5	8	12
13	9	6	5	4	5	6	9	13
14	10	9	8	7	8	9	10	14
15	14	13	12	11	12	13	14	15

Figure 4: Rings of simmetry of a quincunx filter coefficients.

The optimization of the cost function stated in Equation 18 is difficult if the first estimate of the coefficients is not appropriate. We used as initial values the filters obtained by frequency sampling, transforming directly from a one-dimensional QMF $H(\omega)$. By choosing samples at:

$$H_0(\boldsymbol{\omega}) = H\left(|\omega_1| + |\omega_2|\right) \qquad 19$$

and taking the inverse fourier transform, we found a good set of initial values for the filter coefficients, and proceeded with the otimization. In Table 1, we summarize coefficients obtained using a 5 coefficient QMF as a starting point. To satisfy the no-delay requirement in Equation 15, we forced zero-phase of the filter. Coefficients are ordered as in Figure 3. Also, Figure 5 shows the frequency response of filter $H_0(\omega)$ obtained.

4 Successive Quincunx Decomposition

One of the strongest features of subband coding is the possibility of implement successive decompositions, in order to reduce the energy of each subband. In signal compression, this is a desirable characteristic, since low energy signals require less bits to be transmitted. Successive decomposition is achieved by continually applying subband decomposition on those subbands which still have much energy on them – usually the low frequency band. For one-dimensional signals, and also in images being coded with a separable model, it is easy to see that aplying the same set of filters will result in the

Coefficient Position	Value
1	0.6470
2	0.1360
3	-0.0248
4	-0.0134
5	-0.0089
6	0.0043
7	0.0112
8	0.0000
9	-0.0007
10	0.0000
11	-0.0025
12	-0.0020
13	0.0008
14	-0.0001
15	0.0001

Table 1: Filter coefficients

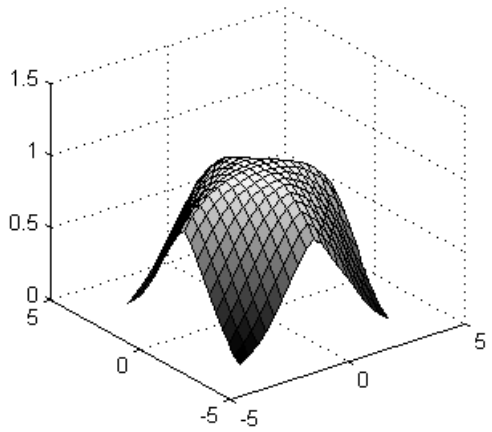


Figure 5: Lowpass analysis filter.

desired decomposition. We will show that this procedure can also be applied to quincunx decompositions.

First, we note that two successive quincunx decompositions result in a separable downsampling by a factor of two. This can easily be seen, as

$$\mathbf{Q}^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \mathbf{R} \qquad 19$$

Thus, after two quincunx decompositions, the resulting band will be rectangular, and the filter designed in the previous section can be applied again.

The band resulting from the first application of filter $H_0(\boldsymbol{\omega})$ and quincunx decimation will have the shape of a 45^o rotated square, which can easily be seen from Equation 3. If we apply a rotated version of filter $H_0(\boldsymbol{\omega})$,

$H_0(\mathbf{Q}^{-T}\boldsymbol{\omega})$ and proceed with quincunx decimation, the resulting band will have the shape

$$\left\{\mathbf{Q}^{-T}\mathbf{Q}^{-T}\boldsymbol{\omega} : \boldsymbol{\omega} \in [-\pi,\pi]\times[-\pi,\pi]\right\} \qquad 20$$

Since $\mathbf{Q}^{-T}\ \mathbf{Q}^{-T} = \mathbf{R}^{-T}$, this shows that as the low frequency analysis filter of the second quincunx decomposition, we can use a 45^o rotated version of original filter $H_0(\boldsymbol{\omega})$. It is interesting to note that $H_0(\mathbf{Q}^{-T}\boldsymbol{\omega})$ is the same filter as before, in respect with the rotated frequency set of axis $\boldsymbol{\omega}' = \mathbf{Q}^{-T}\boldsymbol{\omega}$. Thus, the successive application of $H_0(\boldsymbol{\omega})$ and quincunx downsampling will result in successive quincunx subband decomposition. Figure 6 shows the proccess.

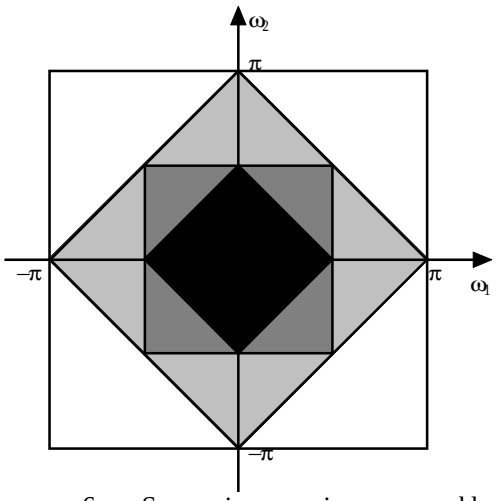


Figure 6: Successive quincunx subband decomposition.

Since each quincunx downsampled subband will have half the number of samples of the previous one, a data structure is needed to represent these. Each line of a subband will also have half the number of samples of a line of the original image. We choose to colapse two lines into one, which allows us to build the image on Figure 7, which is a quincunx decomposition of level 2. This data structure, however, is useful to help visualize the coding process, and it is not necessary that it should be used. Since a quincunx subband has diagonal regularity, a sequence of matrices could be used.

5 Compression of Quincunx Subbands

We show now the use of quincunx subbands in the process of coding and compression of images. There are a lot of techniques to efficiently compress image subbands. We refer the reader to [7] for a complete survey on modern techniques of scalar and vector quantization applied to subband coding of images.

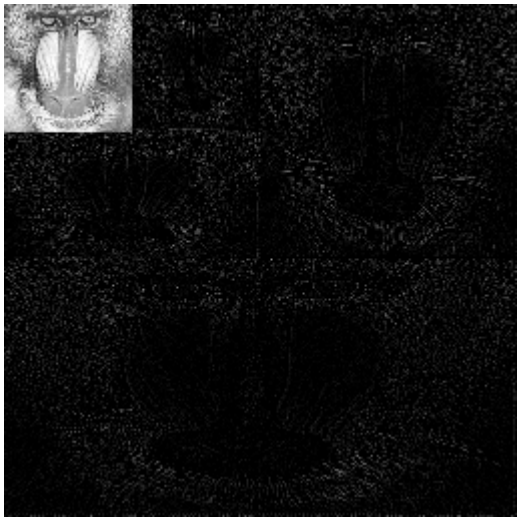


Figure 7: Four-level quincunx subband decomposition of image *baboon*.

Since our main interest is to compare performances of a quincunx subband coding with separable, we choose a simple technique that will allow us to obtain data about the process. Coefficients in each subband are quantized and then entropy coded. Depending on the level of the quantization, compression will achieve a greater efficiency, but with the effect of lowering signal-to-noise ratio.

Self-organizing maps (SOM) [18] were used in order to find a more appropriated set of quantization levels. Thus, quantization of the coefficients was not uniform. The process of converging a SOM applied to scalar quantization is simple. Let $\mathbf{w}$ be the set of all quantization levels. A specific subband coefficient $X(\mathbf{n})$ will then be closer to one of this values w_k, which is called the winner. The winner is then updated according to

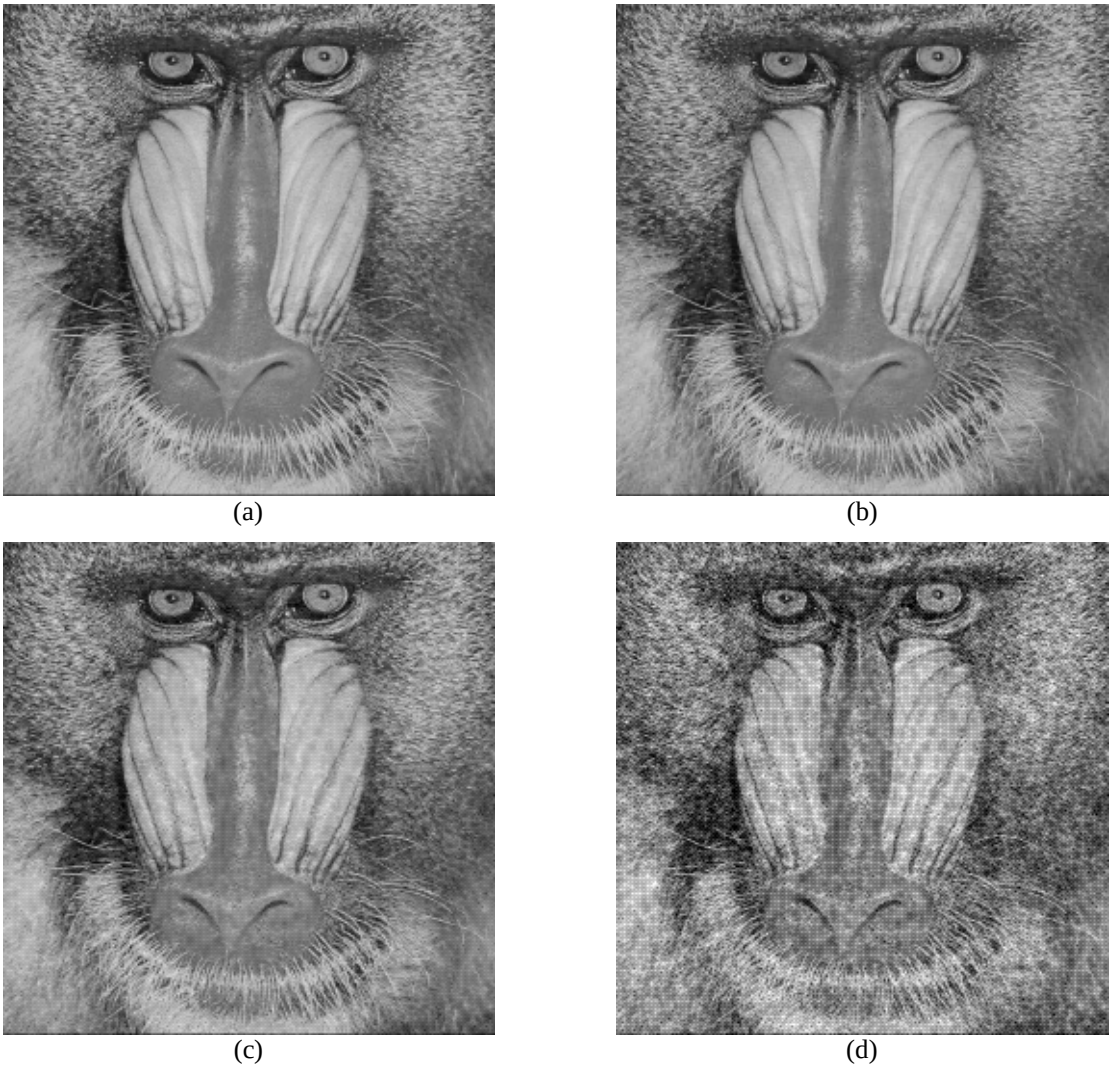
$$w_k = w_k + \delta\left(X(\mathbf{n}) - w_k\right) \qquad 21$$

where δ is the learning rate. By repeating this process to every coefficient in the subband, the result will be a more appropriate set $\mathbf{w}$ of quantization levels. Each subband can be used many times to converge the SOM to better the approximation. We repeated the process three times with each subband.

Entropy coding of the quantized levels was done using arithmetic coding.

Figure 8 shows the reconstructed images of the baboon image, coded respectively with 1, 0.5 and 0.25 bit per pixel. Signal-to-noise ratio to these images is given in the figure.

Tests were carried with a great variety of images. Figure 9 shows a plot of variation of average signal-to-noise ratio according to bits per pixel obtained in image coding, for rectangular and quincunx downsampling, and also for the standard cosine transform technique



(a) (b)

(c) (d)

Figure 8: a) Original baboon; b) Baboon coded at 1 bpp. SNR = 31.10 dB; c) Baboon coded at 0.5 bpp. SNR = 22.61 dB; d) Baboon coded at 0.25 bpp. SNR = 17.43 dB

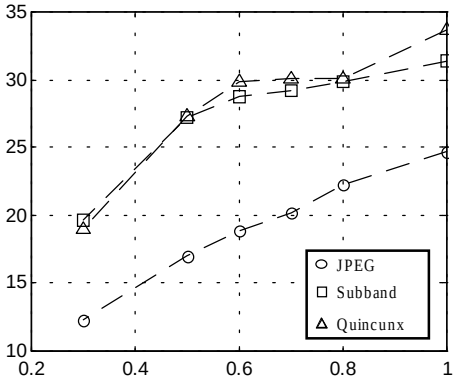


Figure 9: Variation of bit rate and signal-to-noise ratio for JPEG, sub-band and quincunx coding.

implemented by JPEG. Figure 9 shows that quincunx performance is pretty close to rectangular one, and both usually perform better than JPEG.

6 Conclusion

In this paper, we developed an algorithm to perform successive quincunx downsampling of images, and analyzed the performance of a subband image coder based on that procedure. In our preliminary tests, we have found that, in the context of image coding, quincunx downsampling performs at least as well as standard rectangular downsampling. However, the coding procedure does not take into account redundancies that may appear in the subbands due to quincunx simmetries. Research in this topic is being conducted, as well as other

related topics, such as computational complexity and subjective performance.

This is a work still in development, and thus not all the data are available. The available results, however, show that quincunx subband coding performs very closely to the common rectangular subband coding, sometimes a little better, sometimes a little worse. These results, however, refer to objective quality of the reconstructed images, *ie*, noise energy and signal to noise ratio. That suggests us that size of file and noise obtained by this technique will be very similar to separable sampling. We expect that further research will show that will actually deliver better results.

7 References:

[1] R. E. Crochiere, S. A. Webber & J. L. Flanagan, *Digital Coding of Speech in Subbands*, Bell Syst. Tech. Journal, vol. 55, pp 1069-1085, Oct 1976.

[2] M. Vetterli, *Multidimensional subband coding: some theory and algorithms*, Signal Processing, vol. 6, pp 97-112, Apr 1984.

[3] J. W. Woods & S. D. O'Neil, *Subband Coding of Images*, IEEE Transactions on Acoustics, Speech and Signal Processing, vol. ASSP-34, no. 5, pp 1278-1288, Oct 1986.

[4] E. Viscito & J.P. Allebach, *The Analysis and Design of Multidimensional FIR Perfect Reconstruction Filter Banks for Arbitrary Sampling Lattices*, IEEE Transactions on Circuits and Systems vol. 38, pp 29-41, Jan 1991.

[5] J. Kovacevic & W. Sweldens, *Wavelet Families of Increasing Order of Arbitrary Dimensions*, IEEE Transactions on Image Processing, vol. 9, no. 32, pp 480-496, 2000.

[6] J. Kovacevic & M. Vetterli, *Nonseparable Multidimensional Perfect Reconstruction Filter Banks and Wavelet Bases for R^n*, IEEE Transactions on Information Theory, vol. 38, no. 2, pp 533-555, Mar 1992.

[7] P. C. Cosman, R.M. Gray & M. Vetterli, *Vector Quantization of Image Subbands: a Survey*, IEEE Transactions on Image Processing, vol. 5, no. 2, pp 202-225, Feb 1996.

[8] R.M. Meresereau & D. E. Dudgeon, *Multidimensional Digital Signal Processing*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1984.

[9] J. S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1990.

[10] A. V. Oppenheim & R. W. Schafer, *Discrete-Time Signal Processing*, Prentice-Hall, Inc, Englewood-Cliffs, New Jersey, 1989.

[11] Masaaki Ikehara, Takayuki Nagai & Truong Q. Nguyen, *Time-Domain Design and Lattice Structure for FIR Paraunitary Filter Banks with Linear Phase*, Signal Processing, vol. 80, pp 333-342, 2000.

[12] Hung-Ching Lu, Shian-Tang Tzeng, *Design of Two-Dimensional FIR Digital Filters for Sampling Structure Conversion by Genetic Algorithm Approach*, Signal Processing, vol. 80, pp 1445-1458, 2000.

[13] D. Stanhill & Y. Y. Zeevi, *Two-Dimensional Linear Phase Orthogonal Filter Banks and Wavelets*, IEEE Transactions on Signal Processing, vol. 46, no. 1, pp 183-190, Jan 1998.

[14] A. A. C. Kalker & I. A. Shah, *Ladder Structures for Multidimensional Linear Phase Perfect Reconstruction Filter Banks and Wavelets*, Proceedings of SPIE Conference onf Visual Communications and Image Processing, vol. 1818, pp 12-20, Boston, Nov 1992.

[15] I. A. Shah & A. A. C. Kalker, *Theory and Design of Multidimensional QMF Subband Filters from 1D Filters and Polynomials Using Transforms*, available at http://citeseer.nj.nec.com/shah92theory.html

[16] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1993.

[17] R. E. Crochiere & L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1983.

[18] S. Haykin, *Neural Networks: a Comprehensive Foundations*, 2nd ed, Prentice-Hall Inc. Englewood Cliffs, New Jersey, 1999.