

Subband Vector Quantization of Images Using Hexagonal Filter Banks¹

Osama S. Haddadin, V. John Mathews, and Thomas G. Stockham, Jr.

Department of Electrical Engineering
University of Utah
Salt Lake City, Utah 84112

Abstract

Results of psychophysical experiments on human vision conducted in the last three decades indicate that the eye performs a multichannel decomposition of the incident images. This paper presents a subband vector quantization algorithm that employs hexagonal filter banks. The hexagonal filter bank provides an image decomposition similar to what the eye is believed to do. Consequently, the image coder is able to make use of the properties of the human visual system and produce compressed images of high quality at low bit rates. We present a systematic approach for optimal allocation of available bits among the subbands and also for the selection of the size of the vectors in each of the subbands.

1. Introduction

Vector quantization (VQ) is a very powerful approach to data compression and has been successfully applied to a variety of signals [7] [17] [20]. In VQ a sequence of continuous or high rate discrete k -dimensional vectors is mapped into a sequence suitable for transmission over a digital channel or storage. The simplest form of vector quantizer operates as follows. First, a codebook of k -dimensional vectors is created using a training set representative of the waveforms. Once the codebook is designed and the representative vectors are stored, the process of encoding the incoming waveform can begin. k consecutive samples of the waveform are grouped together to form a k -dimensional vector at the input of the vector quantizer. The input vector is successively compared with each of the stored vectors and a metric or distance is computed in each case. The representative vector closest to the input vector is identified, and the index of the closest vector is available at the output for transmission or storage.

While direct VQ as described above will work reasonably well, much can be gained by combining vector quantizers with schemes such as predictive coding [10] [14] [21] [27] or subband coding [1] [2] [5] [6] [28] [30] [31]. This paper deals with

¹This work was supported in part by the National Science Foundation under Grant MIP-9016331

a particular type of subband vector quantizers that makes use of hexagonal filter banks.

The block diagram of a typical subband image coder is shown in Figure 1. It consists of an analysis filter bank that decomposes the image into several subband components, decimators, and quantizers for the decimated signals. At the receiver, the synthesis section reconstructs the coded image from quantized subband images by upsampling and filtering each of the subband images and then additively combining the results. A block diagram of a multichannel Analysis/Synthesis filter bank (ASFB) is shown in Figure 2.

The algorithm presented in the paper performs a subband decomposition of hexagonally sampled (or resampled [22] [23]) images using filter banks with hexagonal planes of support. There are several factors that motivate this approach. First, hexagonal sampling of images provides very efficient discrete representation of two-dimensional signals. Secondly, it is believed, on the bases of psychophysical experiments [3] [8] [9] [11] [12] [24] [25] [29], that the human eye decomposes the incident images into several channels based on the magnitude and orientation of the frequency components of the images. Subband decomposition using hexagonal filter banks provides a similar decomposition of the images and consequently, we can take advantage of the properties of the eye and design a quantizer in such a way that the distortions are in areas that the eye can tolerate. This will result in a subjectively more pleasing quantization for coded images. Finally, the filter bank we employ provides a multiresolution decomposition of the images which would be very useful in progressive transmission, browsing and other similar applications.

The rest of the paper is organized as follows. The next section describes in detail the subband vector quantizer. In our work, we have used vector quantization of each subband separately. Our experience has shown that forming vectors using elements belonging to each subband separately is better than forming vectors using elements belonging to different subbands. However, this will necessitate the need for developing a scheme for allocating the available bits in some optimal manner among all the subbands. This section also presents a systematic scheme for bit allocation as well as for optimal selection of the size of the vector dimension in each subband. The concluding remarks are made in Section 3.

2. System Description

The subband vector quantizer with hexagonal filter banks comprises of the following components:

- i) An ASFB that uses hexagonal sampling. The resulting subbands have hexagonal shapes as shown in Figure 3. Also, note that the bandwidths are different (the higher subbands have twice the bandwidth as lower subbands) for different subbands. It is necessary to use nonseparable filter banks to obtain such a decomposition. Use of hexagonal sampling has many advantages: They offer many

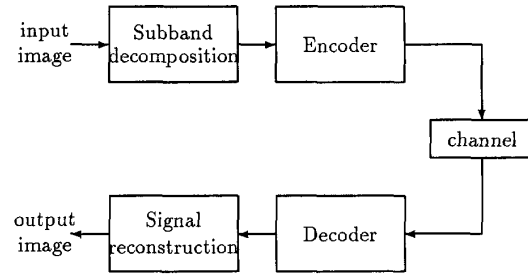


Figure 1: Subband coding of digital images.

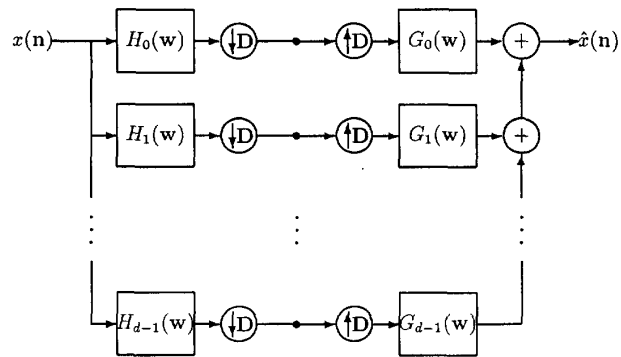


Figure 2: Multichannel analysis and synthesis filter bank

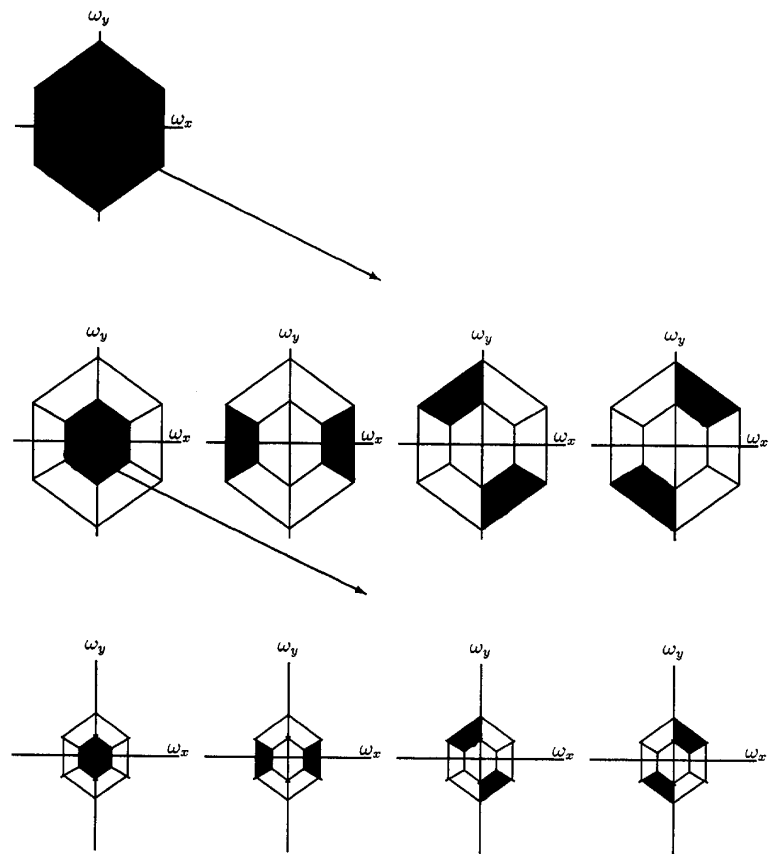


Figure 3: Idealistic frequency partitioning as indicated by the shaded regions in the middle level. The last level indicates further decomposition of lowest frequency subband into four more subbands. This process may be repeated a desired number of times.

advantages in computational simplicity. They provide the subbands with orientation properties that are not limited to vertical and horizontal directions.

Existing evidence in human vision suggest that the image information is processed by several frequency-selective channels (or bundles of cells) which combine to give the resulting human visual response. Each channel is a detector of a specific image feature. In particular experiments involving contrast threshold [8] [25] and electrical stimulation [9] imply the existence of elements which convey information about a region localized in space, frequency, and orientation. This discussion suggests that the basic structure of the hexagonal ASFB subband decomposition is similar to the processing that takes place in the human visual system [3] [32]. Therefore we believe that one can gain from designing the ASFB that attempts to mimic the behavior of the human visual system.

ii) Vector quantization of each subband separately. The bit allocation scheme is presented in this section.

iii) The synthesis section at the receiver that reconstructs the quantized images. Conditions for perfect reconstruction is discussed next.

Recent advances in the theory of ASFB provide conditions under which the designed nonseparable analysis and synthesis bank of filters satisfy perfect reconstruction of the original input signal. Aliasing and phase distortions are extremely disturbing to the human observer, and therefore it is required that the filters guarantee alias free and linear phase processing. Filter banks which provide a reconstruction that is alias free with negligible amplitude and phase distortions are referred to as Perfect Reconstruction Filter Banks (PRFB).

In this work, we restrict our study to maximally decimated ASFB systems. That is, the total number of samples in all the channels is equal to the number of samples in the original input signal. We also assume finite impulse response (FIR) filters of equal sizes for both the analysis and synthesis sections. FIR filters are generally preferred over infinite impulse response filters for image processing because they allow exact linear phase response, and because stability is guaranteed. Let the multidimensional input signal, $x(n)$, be defined on a hexagonal integer sampling grid, or lattice, denoted by Λ . The subband signals, $y_i(n)$ for $i = 0, \dots, d-1$, are then defined on a sublattice Λ_D , that is also hexagonal provided that the subsampling matrix D is chosen correctly.

Using the convolution property and the definition of downsampling, we describe the Fourier transform of the output signal of the i -th analysis channel as

$$Y_i(\omega) = \frac{1}{d} \sum_{j=0}^{d-1} H_i(D^{-T}\omega + 2\pi D^{-T}k_j) X_i(D^{-T}\omega + 2\pi D^{-T}k_j) \quad (1)$$

where d represents the number of subbands in the decomposition, and k_i is a shift vector associated with the sublattice of the i -th analysis channel. The output of the ASFB system (refer to Figure 2) is

$$\hat{X}(\omega) = \sum_{i=0}^{d-1} Y_i(D^T\omega) G_i(\omega) \quad (2)$$

The conditions for perfect reconstruction are given by [26]

$$\sum_{i=0}^{d-1} H_i(\omega) G_i(\omega) = d \quad (3)$$

$$\sum_{i=0}^{d-1} H_i(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_j) G_i(\omega) = 0 \quad j = 1, \dots, d-1. \quad (4)$$

A family of filters that do satisfy Equations 3 and 4 are the quadrature mirror filters [13] [15], defined for the hexagonal case [26] as

$$H_0(\omega) = G_0(-\omega) = F(\omega) = F(-\omega) \quad (5)$$

$$H_1(\omega) = G_1(-\omega) = e^{j\omega^T \mathbf{s}_1} F(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_1) \quad (6)$$

$$H_2(\omega) = G_2(-\omega) = e^{j\omega^T \mathbf{s}_2} F(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_2) \quad (7)$$

$$H_3(\omega) = G_3(-\omega) = e^{j\omega^T \mathbf{s}_3} F(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_3) \quad (8)$$

where the \mathbf{s}_i are known as the spatial shift vectors. \mathbf{s}_i for $i = 1, 2, 3$ are chosen such that the system is alias free, and F is a function that is invariant under negation of its argument. Using Equations 5 through 8 in Equations 3 and 4, for $d = 4$, reduces the design problem to that of finding a filter with $F(\omega)$ satisfying the constraints

$$\sum_{i=0}^3 |F(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_i)|^2 = 4. \quad (9)$$

and

$$F(\omega + 2\pi \mathbf{D}^{-T} \mathbf{k}_j) F(-\omega) = 0 \quad (10)$$

where \mathbf{k}_j can be any one of the non-zero shift vectors. Approaches to solving this optimization problem are discussed in [26] [31].

We will now very briefly discuss the vector quantizer. The codebook for each subband is created using a training set formed by the corresponding subband decomposition of a representative set of waveforms that the system will process. We have used the Linde, Buzo, Gray (LBG) algorithm [18] for codebook design. The size of the codebook and the vector dimension for each subband play an important role in the overall performance of the system. Techniques for selecting these parameters involve searching the bit-rate versus distortion plane for the optimal solution. However, most of such techniques ignore the dependency of the distortion on the vector dimension. In this paper, we will discuss, and employ, a new bit allocation technique which provides the complete solution (i.e. bit rate and vector dimension) to the problem. This technique is based on an adaptation of the scheme introduced by Bradley, Stockham and Mathews [6]. In the remainder of this section we provide a brief review of this scheme.

Our technique attempts to minimize a distortion function, $\mathbf{E}(\mathbf{r}, \mathbf{k})$, defined as

$$\mathbf{E}(\mathbf{r}, \mathbf{k}) = \sum_{i=0}^{d-1} e_i(r_i, k_i) \quad (11)$$

where e_i is the distortion introduced by quantizing the i -th channel, and is a function of that channel's rate r_i and vector size k_i . The vectors \mathbf{r} and \mathbf{k} are d -dimensional and are defined as

$$\mathbf{r} = [r_0 \dots r_{d-1}] \quad (12)$$

$$\mathbf{k} = [k_0 \dots k_{d-1}] \quad (13)$$

Bradley, et. al. [6] demonstrated that e_i can be approximated as

$$e_i(r_i, k_i) = \beta(k_i)e^{-\gamma(k_i)r_i} \quad (14)$$

when the Euclidean distortion measure is employed. The parameters $\beta(k_i)$ and $\gamma(k_i)$ are estimated empirically. Thus we may write Equation 11 as

$$\mathbf{E}(\mathbf{r}, \mathbf{k}) = \sum_{i=0}^{d-1} \beta(k_i)e^{-\gamma(k_i)r_i} \quad (15)$$

Now, the optimization problem may be stated as that of minimizing Equation 15 subject to a bit rate constraint given by

$$\sum_{i=0}^{d-1} r_i = \text{Total bit rate} \quad (16)$$

where $r_i \geq 0$ for $i = 0, \dots, d-1$, and a complexity constraint given by

$$\sum_{i=0}^{d-1} 2^{r_i k_i} \leq \text{Maximum codebook size.} \quad (17)$$

The constraint optimization problem is solved using a projected gradient algorithm [6] [16] [19].

3. Concluding Remarks

In this paper, we presented an algorithm for image coding that employs vector quantization of subband signals obtained using hexagonal filter banks and hexagonal sampling. The decomposition provided by the hexagonal filter bank shares several properties with the image decomposition that takes place in the visual cortex. Consequently, a compression scheme that makes use of such a decomposition can provide better subjective quality for the images. Experiments, the results of which are not reproduced here, have provided very good quality reproductions of coded monochrome still images at bit rates that are fractions of one bit per pixel.

References

- [1] E. Adelson, E. Simoncelli and R. Hingorani, "Orthogonal pyramid transforms for image coding," *SPIE Visual Communication and Image Processing II*, vol. 845, pp. 50-58, 1987.
- [2] R. Baseri, Vector Quantization of Monochrome Images using a Visual Masking Function, *M. S. Thesis*, University of Utah 1991.
- [3] B. S. Baxter, Image Processing in the Human Visual System, *Ph. D. Dissertation*, University of Utah, 1973.
- [4] C. Blakemore and F. W. Campbell, "On the Existence of Neurones in the Human Visual System Selectively Sensitive to the Orientation and Size of Retinal Images," *Journal of Physiology*, London, vol. 203, pp. 237-260, 1969.
- [5] J. N. Bradley, Subband Image Coding with Vector Quantization, *Ph. D. Dissertation*, University of Utah 1989.
- [6] J. N. Bradley, T. G. Stockham, Jr., and V. J. Mathews, "Subband Coding with Vector Quantization," submitted to *IEEE Transactions on Communications*
- [7] A. Buzo, A. H. Gray, Jr., R. M. Gray, and J. D. Markel, "Speech coding based upon Vector Quantization," *IEEE Transactions on Information Theory*, vol. IT-28, no. 5, pp. 562-574, October 1980.
- [8] F. W. Campbell and J. G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *Journal of Physiology*, London, vol. 197, pp. 551, 1968.
- [9] F. W. Campbell, F. F. Cooper and C. Enroth-Cudgell, "The Spatial Cells of the Cat," *Journal of Physiology*, London, vol. 203, pp. 223, 1969.
- [10] P. Chang and R. M. Gray, "Gradient Algorithms for Designing Predictive Vector Quantizers," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, no. 4, August 1986.
- [11] P. Colas-Baudelaire, Digital Picture Processing and Psychophysics: A Study of Brightness Perception, *Ph. D. Dissertation*, University of Utah, 1973.
- [12] T. N. Cornsweet, *Visual Perception*, Prentice-Hall, 1970.
- [13] A. Croiser, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques," *International Conference on Information Sciences and Systems*, Patras, pp. 443-446, August 1976.
- [14] V. Cuperman and A. Gersho, "Vector Predictive Coding of Speech at 16 kbits/s," *IEEE Transactions on Communications*, vol. COMM-33, no. 7, pp. 685-696, July 1985.

- [15] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding scheme," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 191-195, 1977.
 - [16] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, 1981.
 - [17] R. M. Gray, "Vector Quantization," *IEEE Acoustics, Speech, and Signal Processing magazine*, vol. 1, pp. 4-29, April 1984.
 - [18] Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantization Design," *IEEE Transactions on Communications*, vol. COM-28, pp. 84-95, January 1980.
 - [19] D. C. Luenberger, *Linear and Nonlinear Programming*, Reading, Addison Wesley Publishing Company, 1984.
 - [20] J. Makhoul, S. Roucos and H. Gish, "Vector Quantization in Speech Coding," *Proceedings of the IEEE*, vol. 73, no. 11, pp. 1551-1558, November 1985.
 - [21] V. J. Mathews, R. W. Waite, and T. D. Tran, "Image Compression Using Vector Quantization of Linear (one-step) Prediction Errors," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 733-736, Dallas, Texas, April 1987.
 - [22] R. M. Mersereau, "The Processing of Hexagonally Sampled Two-Dimensional Signals," *Proceedings of the IEEE*, vol. 67, no. 6, pp. 930-949, June 1979.
 - [23] R. M. Mersereau and T. C. Speake, "The processing of periodically sampled multidimensional signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-31, no. 1, pp. 188-194, February 1983.
 - [24] D. A. Pollen and S. F. Ronner, "Phase relationship between adjacent simple cells in the visual cortex," *Science*, vol. 212, pp. 1409-1411, 1981.
 - [25] F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, Holden-Day, 1965.
 - [26] E. P. Simoncelli and E. Adelson, "Non-separable Extensions of Quadrature Mirror Filters to Multiple Dimensions," *Proceedings of the IEEE*, vol. 78, no. 4, pp. 652-664, April 1990.
 - [27] B. Vasudev, "Predictive VQ Schemes for Gray Scale Image Compression," *Proceedings of GLOBECOM'87*, pp. 436-441, Tokyo, Japan, November 1987.
 - [28] M. Vetterli, J. Kovacevic and D. Legall, "Perfect Reconstruction Filter Bank for HDTV Representation and Coding," *Signal Processing: Image Communication*, vol. 2, pp. 349-363, 1990.
-

- [29] A. B. Watson and A. J. Ahumada, Jr., "A Hexagonal Orthogonal-Oriented Pyramid as a Model of Image Representation in Visual Cortex," *IEEE Transactions on Biomedical Engineering*, vol. 36, no. 1, January 1989.
 - [30] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, October 1986.
 - [31] J. W. Woods, ed., *Subband Image Coding*, Klumer Academic Publishers, 1991.
 - [32] Z. Xie, Previsualized Image Vector Quantization, *Ph. D. Dissertation*, University of Utah, 1989.
-