## Box 6: Multiresolution analysis

The concept of multiresolution approximation of functions was introduced by Meyer and Mallat [MAL89a, MAL89c, MEY90] and provides a powerful framework to understand wavelet decompositions. The basic idea is that of successive approximation, together with that of "added detail" as one goes from one approximation to the next, finer one. We here give the intuition behind the construction.

Assume we have a ladder of spaces such that:

$$\ldots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \ldots$$

with the property that if $f(x) \in V_l$ then $f(x-2^{-l}k) \in V_l$, $k \in Z$, and $f(2x) \in V_{l-1}$. Call $W_l$ the orthogonal complement of $V_l$ in $V_{l-1}$. This is written

$$V_{l-1} = V_l \oplus W_l \qquad \text{(B6.1)}$$

Thus $W_l$ contains the "detail" necessary to go from $V_l$ to $V_{l-1}$. Iterating (B6.1) one has

$$V_{l-1} = W_l \oplus W_{l+1} \oplus W_{l+2} \oplus W_{l+3} \oplus \ldots \qquad \text{(B6.2)}$$

that is a given resolution can be obtained by a sum of added details.

Now, assume we have an orthonormal basis for $V_0$ made up of a function $\phi(x)$ and its integer translates. Because $V_0 \in V_{-1}$, $\phi(x)$ can be written in terms of the basis in $V_{-1}$, i.e. (15) is satisfied:

$$\phi(x) = \sum_n c_n g_0(2x-n)$$

then it can be verified that the function $h_0(x)$ (16) (with the relations (12)) and its integer translates form an orthonormal basis for $V_0$. And because of (B6.2), $h_0(x)$ and its scaled and translated versions, form a wavelet basis [MAL89a, MAL89c, MEY90].

The multiresolution idea is now quite intuitive. Assume we have an approximation of a signal at a resolution corresponding to $V_0$. Then a better approximation is obtained by adding the details corresponding to $W_0$, that is, the projection of the signal in $W_0$. This amounts to a weighted sum of wavelets at that scale. Thus by iterating this idea, a square integrable signal can be seen as the successive approximation or weighted sum of wavelets at finer and finer scale.

generate wavelet bases. The converse is also true. That is, orthonormal sets of scaling functions and wavelets can be used to generate perfect reconstruction filter banks [DAU88, MAL89a, MAL89c].

Extension of the wavelet concept to multiple dimensions, which is useful, e.g. for image coding, is shown in Box 7.

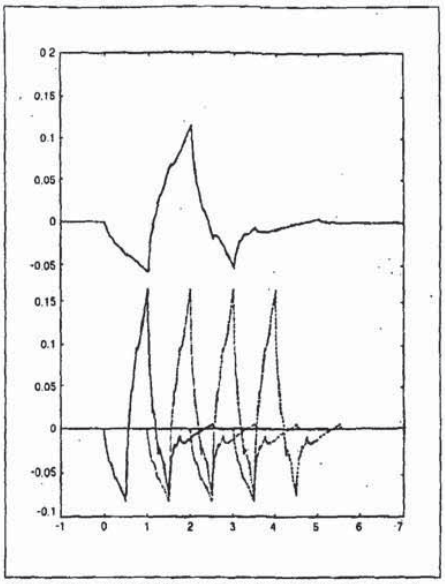


*Fig. 14. Two scales of the $D_4$ wavelet and shifts. This set of functions is orthogonal.*

## APPLICATIONS OF WAVELETS IN SIGNAL PROCESSING

From the derivation of the wavelet transform as an alternative to the STFT, it is clear that one of the main applications will be in non-stationary signal analysis. While conceptually, the CWT is a classical constant-Q analysis, its simple definition (based on a single function rather than multiple filters) allows powerful analytical derivations and has already lead both to new insights and new theoretical results [WAV89].

Applications of wavelet decompositions in numerical analysis, e.g. for solving partial differential equations, seem very promising because of the "zooming" property which allows a very good representation of discontinuities, unlike the Fourier transform [BEY89].

Perhaps the biggest potential of wavelets has been claimed for signal compression. Since discrete wavelet transforms are essentially subband coding systems, and since subband coders have been successful in speech and image compression, it is clear that wavelets will find immediate application in compression problems. The only difference with traditional subband coders is the fact that filters are designed to be regular (that is, they have many zeroes at $z = 0$ or $z = \pi$). Note that although classical subband filters are not regular (see Box 5 and Fig. 12), they have been designed to have good stopbands and thus are close to being "regular", at least for the first few octaves of subband decomposition.

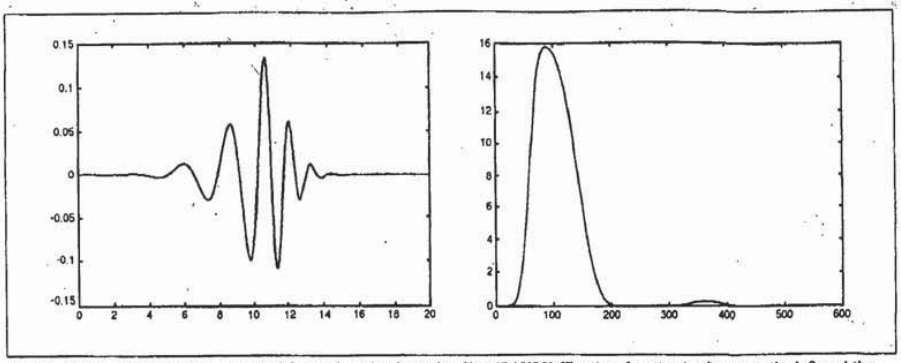

*Fig. 15. Orthonormal wavelet generated from a length-18 regular filter [DAU88]. The time function is shown on the left and the spectrum is on the right.*

It is therefore clear that drastic improvements of compression will not be achieved so easily simply because wavelets are used. However, wavelets bring new ideas and insights. In this respect, the use of wavelet decompositions in connection with *other* techniques (like vector quantization [ANI90] or multiscale edges [MAL89d]) are promising compression techniques which make use of the elegant theory of wavelets.

New developments, based on wavelet concepts, have already appeared. For example, statistical signal processing using wavelets is a promising field. Multiscale models of stochastic processes [BAS89], [CHO91], and analysis and synthesis of $1/f$ noise [GAC91], [WOR90] are examples where wavelet analysis has been successful. "Wavelet packets" [WIC89], which correspond to arbitrary adaptive tree-structured filter banks, are another promising example.

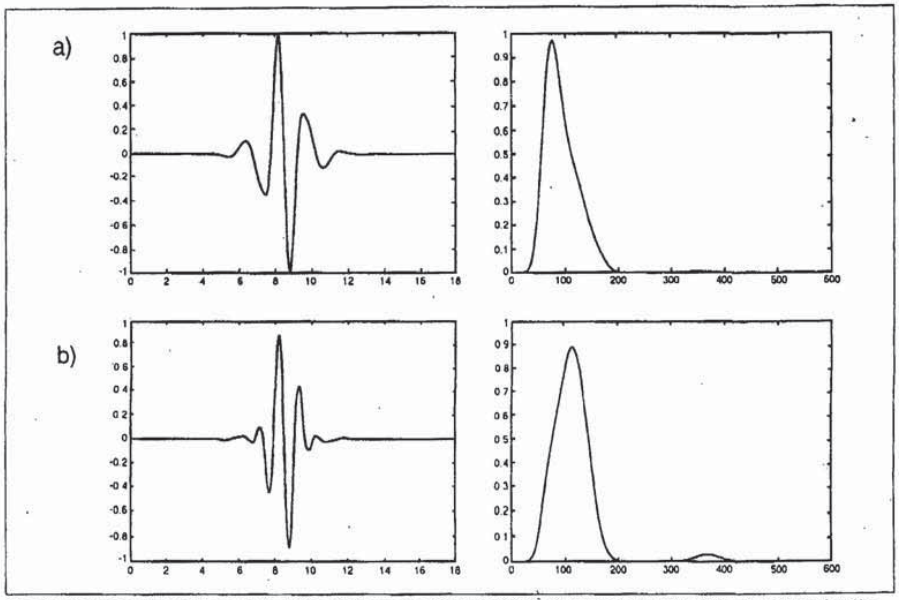


*Fig. 16. Biorthogonal wavelets generated from 18-tap regular filters [VET90b]. (a) Analysis wavelet. (b) Synthesis wavelet. The time function is shown on the left and the spectrum is on the right.*

## Box 7: Multidimensional filter banks and wavelets

In order to apply wavelet decompositions to multidimensional signals (e.g., images), multidimensional extensions of wavelets are required. An obvious way to do this is to use "separable wavelets" obtained from products of one-dimensional wavelets and scaling functions [MAL89a, MAL89c, MEY90]. Let us consider the two-dimensional case for its simplicity. Take a scaling function $g_d(x)$ (15) and a wavelet $h_d(x)$ (16). One can construct for two-dimensional functions :

$$g_d(x,y) = g_d(x)\cdot g_d(y)$$

$$h_s^{(1)}(x,y) = g_d(x)\cdot h_d(y)$$

$$h_s^{(2)}(x,y) = h_d(x)\cdot g_d(y)$$

$$h_s^{(3)}(x,y) = h_d(x)\cdot h_d(y)$$

which are orthogonal to each other with respect to integer shifts (this follows from the orthogonality of the one dimensional component). The function $g_d(x,y)$ is a separable two-dimensional scaling function (that is, a lowpass filter) while the functions $h_s^{(i)}(x,y)$ are wavelets. The set $\{h_s^{(i)}(2^j x-k,\ 2^j x-l),\ i = 1,2,3$ and $j,k,l \in \mathbf{Z}\}$ forms an orthonormal basis for square integrable functions over $\mathbf{R}^2$. This solution corresponds to a separable two-dimensional filter bank with subsampling by 2 in each dimension, that is, overall subsampling by 4 (see Fig. 17).

More interesting (that is, non-trivial) multidimensional wavelet schemes are obtained when non-separable subsampling is used [KOV92]. For example, a non-separable subsampling by 2 of a double indexed signal $x(n_1, n_2)$ is obtained by retaining only samples satisfying:

$$\begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},\ u_1, u_2 \in \mathbf{Z} \qquad \text{(B7.1)}$$

The resulting points are located on a so-called quincunx sublattice of $\mathbf{Z}^2$. Now, one can construct a perfect reconstruction filter bank involving such subsampling because it resembles its one-dimensional counterpart [KOV92]. The subsampling rate is 2 (equal to the determinant of the matrix in (B7.1)), and the filter bank has 2 channels. Iteration of the filter bank on the lowpass branch (see Fig. 18) leads to a discrete wavelet transform, and if the filter is regular (which now depends on the matrix representing the lattice [KOV92]), one can construct non-separable wavelet bases for square integrable functions over $\mathbf{R}^2$ with a resolution change by 2 (and not 4 as in the separable case). An example scaling function is pictured in Fig. 19.

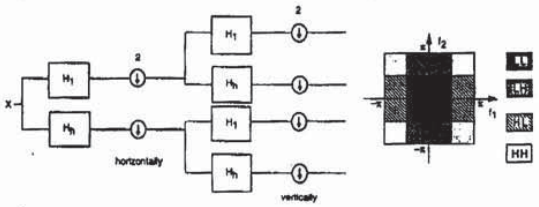


*Fig. 17. Separable two-dimensional filter bank corresponding to a separable wavelet basis with resolution change by 4 (2 in each dimension). The partition of the frequency plane is indicated on the right. $H_l$ and $H_h$ stand for low-pass and high-pass filter, respectively.*

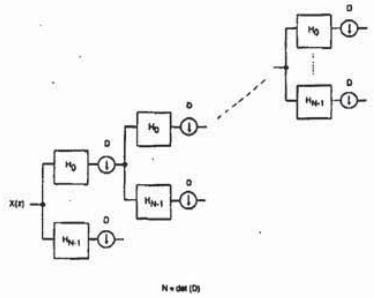


*Fig. 18. Iteration of a non-separable filter bank based on non-separable subsampling. This construction leads to non-separable wavelets.*

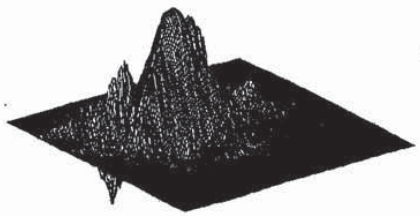

*Fig. 19. Two-dimensional non-separable orthonormal scaling function [KOV92] (orthogonality is with respect to integer shifts). The resolution change is by 2 ($\sqrt{2}$ in each dimension). The matrix used for the subsampling is the one given in (B7.1).*

## CONCLUSION

We have seen that the Short-Time Fourier Transform and the Wavelet Transform represent alternative ways to divide the time-frequency (or time-scale) plane. Two major advantages of the Wavelet Transform are that it can zoom in to time discontinuities and that orthonormal bases, localized in time and frequency, can be constructed. In the discrete case, the Wavelet Transform is equivalent to a logarithmic filter bank, with the added constraint of regularity on the lowpass filter.

The theory of wavelets can be seen as a common framework for techniques that had been developed independently in various fields. This conceptual unification furthers the understanding of the mechanisms involved, quantifies trade-offs, and points to new potential applications. A number of questions remain open, however, and will require further investigations (e.g., what is the "optimal" wavelet for a particular application?).

While some see wavelets as a very promising brand new theory [CIP90], others express some doubt that it represents a major breakthrough. One reason for skepticism is that the concepts have been around for some time, under different names. For example, wavelet transforms can be seen as constant-Q analysis [YOU78], wide-band cross-ambiguity functions [SPE67, AUS90], Frazier-Jawerth transforms [FRA86], perfect reconstruction octave-band filter banks [MIN85, SMI86], or a variation of Laplacian pyramid decomposition [BUR83], [BUR89]!

We think that the interest and merit of wavelet theory is to unify all this into a common framework, thereby allowing new ideas and developments.

## ACKNOWLEDGMENTS

The authors would like to thank C. Herley for many useful suggestions and for generating the continuous STFT and WT plots; and Profs. F. Boudreaux-Bartels, M.J.T. Smith and Dr. P. Duhamel for useful suggestions on the manuscript. We thank B. Shakib (IBM) for creating the three-dimensional color rendering of phase and magnitude of wavelet transforms (so-called "phasemagrams") used in the cover picture and elsewhere; C.A. Pickover (IBM) is thanked for his 3D display software and J.L. Mannion for his help on software tools. The second author would like to acknowledge support by NSF under grants ECD-88-11111 and MIP-90-14189.

**Olivier Rioul** was born in Strasbourg, France on July 4, 1964. He received diplomas in Electrical Engineering from the Ecole Polytechnique, Palaiseau, France, and from Telecom University, Paris, in 1987 and 1989, respectively.

Since 1989, he has been with the Centre National d'Etudes des Télécommunications (CNET), Issy-Les-Moulineaux, France, where he is completing work in the Ph.D. degree in Signal Processing at Télécom University, specializing in wavelet theory, image coding, and fast signal algorithms.

**Martin Vetterli** was born in Switzerland in 1957. He received the Dip. El.-Ing. degree from the Eidgenössische Technische Hochschule Zürich, Switzerland, in 1981; the Master of Science degree from Stanford University, Stanford CA, in 1982; and the Doctorat ès Science Degree from the Ecole Polytechnique Fédérale de Lausanne, Switzerland, in 1986.

In 1982, he was a Research Assistant at Stanford University, and from 1983 to 1986 he was a researcher at the Ecole Polytechnique. He has worked for Siemens and AT&T Bell Laboratories. In 1986, he joined Columbia University in New York where he is currently associate professor of Electrical Engineering, member of the Center for Telecommunications Research, and codirector of the Image and Advanced Television Laboratory.

He is a senior member of the IEEE, a member of SIAM and ACM, a member of the MDSP committee of the IEEE Signal Processing Society, and of the editorial boards of *Signal Processing* and *Image Communication*. He received the Best Paper Award of EURASIP in 1984 for his paper on multidimensional subband coding, and the Research Prize of the Brown Boveri Corporation (Switzerland) in 1986 for his thesis. His research interests include multirate signal processing, wavelets, computational complexity, signal processing for telecommunications and digital video processing.

## REFERENCES

[AKA90] A.N. Akansu, R.A. Haddad, and H. Caglar, "The Binomial QMF-Wavelet Transform for Multiresolution Signal Decomposition," submitted *IEEE Trans. Signal Proc.*, 1990.

[ALL77] J.B. Allen and L.R. Rabiner, "A Unified Approach to Short-Time Fourier Analysis and Synthesis," *Proc. IEEE*, Vol. 65, No. 11, pp. 1558-1564, 1977.

[ANI90] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image Coding Using Vector Quantization in the Wavelet Transform Domain," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, Apr.3-6, 1990, pp. 2297-2300.

[AUS90] L. Auslander and I. Gertner, "Wide-Band Ambiguity Function and a.x+b group," in Signal Processing, Part I: Signal Processing Theory, L. Auslander, T. Kailath, S. Mitter eds., Institute for Mathematics and its Applications, Vol. 22, Springer Verlag, New York, pp.1-12, 1990.

[BAS89] M. Basseville and A. Benveniste, "Multiscale Statistical Signal Processing," in Proc. 1989 IEEE Int. Conf. Acoust., Speech, Signal Proc., Glasgow, Scotland, pp. 2065-2068, May 23-26, 1989.

[BER88] J. Bertrand and P. Bertrand, "Time-Frequency Representations of Broad-Band Signals," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, NY, Apr.11-14, 1988, pp.2196-2199.

[BEY89] G. Beylkin, R. Coifman and V. Rokhlin, "Fast Wavelet Transforms and Numerical Algorithms, I," submitted, Dec. 1989.

[BOU85] G. F. Boudreaux-Bartels, "Time-Varying Signal Processing Using the Wigner Distribution Time-Frequency Signal Representation," in *Adv. in Geophysical Data Proc.*, Vol. 2, pp. 33-79, Jai Press Inc., 1985.

[BUR83] P.J. Burt and E.H. Adelson, "The Laplacian Pyramid as a Compact Image Code," IEEE Trans. on Com., Vol. 31, No.4, pp. 532-540, April 1983.

[BUR89] P.J. Burt, "Multiresolution Techniques for Image Representation, Analysis, and 'Smart' Transmission," Proc. SPIE Conf. on Visual Communication and Image Processing., pp. 2-15, Philadelphia, PA, Nov. 1989.

[CAL64] A. Calderón, "Intermediate Spaces and Interpolation, the Complex Method," Studia Math., Vol. 24, pp. 113-190, 1964.

[CHO91] K.C. Chou, S. Golden, and A.S. Willsky, "Modeling and Estimation of Multiscale Stochastic Processes," in Proc. *1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 1709-1712, May 14-17, 1991.

[CIP90] B.A. Cipra, "A New Wave in Applied Mathematics," *Science, Research News*, Vol. 249, August 24, 1990.

[CLA80] T.A.C.M. Classen and W.F.G. Mecklenbräuer, "The Wigner Distribution — A Tool for Time-Frequency Signal Analysis, Part I, II, III," *Philips J. Res.*, Vol.35, pp. 217-389, 1980.

[COH66] L. Cohen, "Generalized Phase-Space Distribution Functions," *J. Math.-Phys.*, Vol. 7, No. 5, pp. 781-786, 1966.

[COH89] L. Cohen, "Time-Frequency Distribution - A Review," *Proc. IEEE*, Vol. 77, No. 7, pp. 941-981, 1989.

[COH90a] A. Cohen, I. Daubechies and J.C. Feauveau, "Biorthogonal Bases of Compactly Supported Wavelets," to appear in *Comm. Pure and Applied Math.*

[COH90b] A. Cohen, "Construction de Bases d'Ondelettes Hölderiennes," *Revista Matematica Iberoamericana*, Vol.6, No. 3 y 4, 1990.

[CRI76] A. Croisier, D. Esteban, and C. Galand, "Perfect Channel Splitting by Use of Interpolation, Decimation, Tree Decomposition Techniques," Int. Conf. on Information Sciences/Systems, Patras, pp. 443-446, Aug. 1976.

[CRO76] R.E. Crochiere, S.A. Weber, and J.L. Flanagan, "Digital Coding of Speech in Subbands," *Bell Syst. Tech. J.*, Vol.55, pp.1069-1085, Oct.1976.

[CRO83] R.E. Crochiere, and L.R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[DAU88] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Comm. in Pure and Applied Math.*, Vol.41, No.7, pp.909-996, 1988.

[DAU90a] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. on Info. Theory*, Vol. 36, No.5, pp.961-1005, Sept. 1990.

[DAU90b] I. Daubechies, "Orthonormal Bases of Compactly Supported Wavelets II. Variations on a Theme," submitted to *SIAM J. Math. Anal.*

[DAU90c] I. Daubechies and J.C. Lagarias, "Two-Scale Difference Equations II. Local Regularity, Infinite Products of Matrices and Fractals," submitted to *SIAM J. Math. Anal*, 1990.

[DUF52] R.J. Duffin and A.C. Schaeffer, "A Class of Nonharmonic Fourier Series," *Trans. Am. Math. Soc.*, Vol. 72, pp. 341-366, 1952.

[EST77] D. Esteban and C. Galand, "Application of Quadrature Mirror Filters to Split-Band Voice Coding Schemes," Int. Conf. Acoust., Speech, Signal Proc., Hartford, Connecticut, pp. 191-195, May 1977.

[FLA89] P. Flandrin, "Some Aspects of Non-Stationary Signal Processing with Emphasis on Time-Frequency and Time-Scale Methods," in [WAV89], pp.68-98, 1989.

[FLA90] P. Flandrin and O. Rioul, "Wavelets and Affine Smoothing of the Wigner-Ville Distribution," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, April 3-6, 1990, pp. 2455-2458.

[FOU88] J. B. J. Fourier, "Théorie Analytique de la Chaleur," in *Oeuvres de Fourier*, tome premier, G.Darboux, Ed., Paris: Gauthiers-Villars, 1888.

[FRA28] P. Franklin, "A Set of Continuous Orthogonal Functions," *Math. Annal.*, Vol.100, pp.522-529, 1928.

[FRA86] M. Frazier and B. Jawerth, "The φ-Transform and Decomposition of Distributions," *Proc. Conf. Function Spaces and Appl.*, Lund 1986, Lect. Notes Math., Springer.

[GAB46] D. Gabor, "Theory of Communication," *J. of the IEE*, Vol.93, pp.429-457, 1946.

[GAC91] N. Gache, P. Flandrin, and D. Garreau, "Fractal Dimension Estimators for Fractional Brownian Motions," in Proc. *1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 3557-3560, May 14-17, 1991.

[GOU84] P. Goupillaud, A. Grossmann and J. Morlet, "Cycle-Octave and Related Transforms in Seismic Signal Analysis," *Geoexploration*, Vol.23, pp.85-102, Elsevier Science Pub., 1984/85.

[GRO84] A. Grossmann and J. Morlet, "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape," *SIAM J.Math.Anal.*, Vol.15, No.4, pp.723-736, July 1984.

[GRO89] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and Understanding Continuous Wavelet Transforms," in [WAV89], pp.2-20, 1989.

[HAA10] A. Haar, "Zur Theorie der Orthogonalen Funktionensysteme," [in German] *Math. Annal.*, Vol. 69, pp. 331-371, 1910.

[HEI90] C.E. Heil, "Wavelets and Frames," in Signal Processing, Part I: Signal Processing Theory, L. Auslander, et al. eds., IMA, Vol. 22, Springer , New York, pp.147-160, 1990.

[HER71] O. Herrmann, "On the Approximation Probem in NonRecursive Digital Filter Design," *IEEE Trans. Circuit Theory*, Vol. CT-18, No. 3, pp. 411-413, May 1971.

[IT92] *IEEE Transactions on Information Theory*, Special issue on Wavelet Transforms and Multiresolution Signal Analysis, to appear January, 1992.

[JOH80] J.D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," in *Proc. 1980 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, pp. 291-294, Apr. 1980.

[KAD91] S. Kadambe and G.F. Boudreaux-Bartels, "A Comparison of the Existence of 'Cross Terms' in the Wavelet Transform, Wigner Distribution and Short-Time Fourier Transform," Submitted *IEEE Trans. Signal Proc.*, Revised Jan. 1991.

[KOV92] J. Kovacevic and M. Vetterli, "Non-separable Multidimensional Perfect Reconstruction Filter Banks and Wavelet Bases for $R^n$," *IEEE Trans. on Info. Theory*, Special Issue on wavelet transforms and multiresolution signal analysis, to appear Jan. 1992.

[LIT37] J. Littlewood and R. Paley, "Theorems on Fourier Series and Power Series," *Proc. London Math. Soc.*, Vol.42, pp. 52-89, 1937.

[MAL89a] S. Mallat, "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation," *IEEE Trans. on Pattern Analysis and Machine Intell*, Vol.11, No. 7, pp.674-693, July 1989.

[MAL89b] S. Mallat, "Multifrequency Channel Decompositions of Images and Wavelet Models," *IEEE Trans. Acoust., Speech, Signal Proc.*, Vol. 37, No.12, pp.2091-2110, December 1989.

[MAL89c] S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(R)$,", *Trans. Amer. Math. Soc.*, Vol.315, No.1, pp.69-87, September 1989.

[MAL89d] S. Mallat and S. Zhong, "Complete Signal Representation with Multiscale Edges," submitted to *IEEE Trans. Pattern Analysis and Machine Intell*, 1989.

[MEY89] Y. Meyer, "Orthonormal Wavelets," in [WAV89], pp. 21-37, 1989.

[MEY90] Y. Meyer, *Ondelettes et Opérateurs*, Tome I. *Ondelettes*, Herrmann ed., Paris, 1990.

[MIN85] F. Mintzer, "Filters for Distortion-Free Two-Band Multirate Filter Banks," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.33, pp.626-630, June 1985.

[POR80] M.R. Portnoff, "Time-Frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analysis," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.28, pp.55-69, Feb. 1980.

[RAM88] T.A. Ramstad and T. Saramäki, "Efficient Multirate Realization for Narrow Transition-Band FIR Filters," in *Proc. 1988 IEEE Int. Symp. Circuits and Systems*, Helsinki, Finland, pp. 2019-2022, 1988.

[RIO90a] O. Rioul and P. Flandrin, "Time-Scale Energy Distributions: A General Class Extending Wavelet Transforms," to appear in *IEEE Trans. Signal Proc.*

[RIO90b] O. Rioul, "A Discrete-Time Multiresolution Theory Unifying Octave-Band Filter Banks, Pyramid and Wavelet Transforms," submitted to *IEEE Trans. Signal Proc.*

[RIO91a] O. Rioul and P. Duhamel, "Fast Algorithms for Discrete and Continuous Wavelet Transforms," submitted to *IEEE Trans. Information Theory*, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis.

[RIO91b] O. Rioul, "Dyadic Up-Scaling Schemes: Simple Criteria for Regularity," submitted to *SIAM J. Math. Anal.*

[ROS84] A. Rosenfeld ed., Multiresolution Techniques in Computer Vision, Springer-Verlag, New York 1984.

[SMI86] M.J.T. Smith and T.P. Barnwell, "Exact Reconstruction for Tree-Structured Subband Coders," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-34, pp. 434-441, June 1986.

[SHE90] M.J. Shensa, "The Discrete Wavelet Transform: Wedding the à Trous and Mallat Algorithms," submitted to *IEEE Trans. on Acoust., Speech, and Signal Proc.*, 1990.

[SPE67] J.M. Speiser, "Wide-Band Ambiguity Functions," *IEEE Trans. on Info. Theory*, pp. 122-123, 1967.

[VAI87] P.P. Vaidyanathan, "Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Techniques," *IEEE ASSP Magazine*, Vol. 4, No. 3, pp.4-20, July 1987.

[VAI88] P.P. Vaidyanathan and P.-Q. Hoang, "Lattice Structures for Optimal Design and Robust Implementation of Two-Band Perfect Reconstruction QMF Banks," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-36, No. 1, pp.81-94, Jan. 1988.

[VAI89] P.P. Vaidyanathan and Z. Doganata, "The Role of Lossless Systems in Modern Digital Signal Processing," *IEEE Trans. Education*, Special Issue on Circuits and Systems, Vol. 32, No.3, Aug. 1989, pp.181-197.

[VET86] M. Vetterli, "Filter Banks Allowing Perfect Reconstruction," *Signal Processing*, Vol.10, No.3, April 1986, pp.219-244.

[VET89] M. Vetterli and D. Le Gall, "Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations, *IEEE Trans. on Acoust., Speech Signal Proc.*, Vol.37, No.7, pp.1057-1071, July 1989.

[VET90a] M. Vetterli and C. Herley "Wavelets and Filter Banks: Relationships and New Results," in *Proc. 1990 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Albuquerque, NM, pp. 1723-1726, Apr. 3-6, 1990.

[VET90b] M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design," to appear in *IEEE Trans. on Signal Proc.*, 1992.

[WAV89] *Wavelets, Time-Frequency Methods and Phase Space*, Proc. Int. Conf. Marseille, France, Dec. 14-18, 1987, J.M. Combes et al. eds., Inverse Problems and Theoretical Imaging, Springer, 315 pp., 1989.

[WIC89] M.V. Wickerhauser, "Acoustic Signal Compression with Wave Packets," preprint Yale University, 1989.

[WOO53] P.M. Woodward, *Probability and Information Theory with Application to Radar*, Pergamon Press, London, 1953.

[WOR90] G.W. Wornell, "A Karhunen-Loève-like Expansion for 1/f Processes via Wavelets," *IEEE Trans. Info. Theory*, Vol. 36, No. 4, pp.859-861, July 1990.

[YOU78] J.E. Younberg, S.F. Boll, "Constant-Q Signal Analysis and Synthesis," in Proc. 1978 IEEE Int. Conf. on Acoust., Speech, and Signal Proc., Tulsa, OK, pp. 375-378, 1978.

## EXTENDED REFERENCES

**History of Wavelets:** see [HAA10, FRA28, LIT37, CAL64, YOU78, GOU84].

**Books on Wavelets:** (see also [WAV89, MEY90])

I. Daubechies, *Ten Lectures on Wavelets*, CBMS, SIAM publ., to appear.

*Wavelets and their Applications*, R.R. Coifman, I. Daubechies, S. Mallat, Y. Meyer scientific eds., L.A. Raphael, M.B. Ruskai managing eds., Jones and Bertlel pub., to appear, 1991.

**Tutorials on Wavelets:** (see also [FLA89, GRO89, MAL89b, MEY89])

R.R. Coifman, "Wavelet Analysis and Signal Processing," in Signal Processing, Part I: Signal Processing Theory, L. Auslander et al. eds., IMA, Vol. 22, Springer, New York, 1990.

C.E. Heil and D.F. Walnut, "Continuous and Discrete Wavelet Transforms," *SIAM Review*, Vol. 31, No. 4, pp 628-666, Dec. 1989.

Y. Meyer, S. Jaffard, O. Rioul, "L'Analyse par Ondelettes," [in French]*Pour La Science*, No.119, pp.28-37, Sept 1987.

G. Strang, "Wavelets and Dilation Equations: A Brief Introduction," *SIAM Review*, Vol. 31, No. 4, pp. 614-627, Dec. 1989.

**Mathematics, Mathematical Physics and Quantum Mechanics:** (see also [GRO84, DAU88, MEY90])

G. Battle, "A Block Spin Construction of Ondelettes, II. The Quantum Field Theory (QFT) Connection," Comm. Math. Phys., Vol. 114, pp. 93-102, 1988.

W.M. Lawton, "Necessary and Sufficient Conditions for Constructing Orthonormal Wavelet Bases," Aware Tech. Report # AD900402.

P.G. Lemarié and Y. Meyer, "Ondelettes et Bases Hilbertiennes," [in French] *Revista Matematica Iberoamericana*, Vol.2, No.1&2, pp.1-18, 1986.

T. Paul, "Affine Coherent States and the Radial Schrödinger Equation I. Radial Harmonic Oscillator and the Hydrogen Atom," to appear in *Ann. Inst. H.Poincaré*.

H.L. Resnikoff, "Foundations or Arithmeticum Analysis: Compactly Supported Wavelets and the Wavelet Group," Aware Tech. Report # AD890507.

**Regular Wavelets:** see [DAU88, DAU90b, DAU90c, COH90b, RIO91b]

Computer-Aided Geometric Design using Regular Interpolators:

S.Dubuc, "Interpolation Through an Iterative Scheme," *J. Math. Analysis Appl.*, Vol.114, pp.185-204, 1986.

N. Dyn and D. Levin, "Uniform Subdivision Schemes for the Generation of Curves and Surfaces," Constructive Approximation, to appear.

**Numerical Analysis:** (see also [BEY89])

R.R. Coifman, "Multiresolution Analysis in Nonhomogeneous Media," in [WAV89], pp. 259-262, 1989.

V. Perrier, "Toward a Method to Solve Partial Differential Equations Using Wavelet Bases," in [WAV89], pp. 269-283, 1989.

**Multiscale Statistical Signal Processing**: see [BAS89, CHO91].

Fractals, Turbulence: (see also [GAC91, WOR90])

A. Arnéodo, G. Grasseau, and M. Holschneider, "Wavelet Transform of Multifractals," *Phys. Review Letters*, Vol.61, No.20, pp.2281-2284, 1988.

F. Argoul, A. Arnéodo, G. Grasseau, Y. Gagne, E.J. Hopfinger, and U. Frisch, "Wavelet Analysis of Turbulence Reveals the Multifractal Nature of the Richardson Cascade," *Nature*, Vol.338, pp.51-53, March 1989.

**One-Dimensional Signal Analysis:** (see also [GRO89], [WIC-89])

C. D'Alessandro and J.S. Lienard, "Decomposition of the Speech Signal into Short-Time Waveforms Using Spectral Segmentation," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, Apr.11-14, 1988, pp.351-354.

S. Kadambe and G.F. Boudreaux-Bartels, "A Comparison of Wavelet Functions for Pitch Detection of Speech Signals," in *Proc. 1991 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, Toronto, Ontario, Canada, pp. 449-452, May. 14-17, 1991.

R. Kronland-Martinet, J. Morlet, and A. Grossmann, "Analysis of Sound Patterns Through Wavelet Transforms," *Int. J. Pattern Recognition and Artificial Intelligence*, Vol.1, No.2, pp. 273-302, pp.97-126, 1987.

J.L. Larsonneur and J. Morlet, "Wavelets and Seismic Interpretation," in [WAV89], pp.126-131, 1989.

F. B. Tuteur, "Wavelet Transformations in Signal Detection," in *Proc. 1988 IEEE Int. Conf. Acoust., Speech, Signal Proc.*, New York, NY, Apr. 11-14, 1988, pp.1435-1438. Also in [WAV89], pp. 132-138, 1989.

**Radar/Sonar, Ambiguity Functions:** see e.g. [AUS90, SPE67, WOO53]

**Time-Scale Representations:** see [BER88, FLA89, FLA90, RIO90a].

**Filter Bank Theory:** (see also [CRI76, CRO76, EST77, CRO83, MIN85, SMI86, VAI87, VAI88, VAI89, VET86, VET89])

J.D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," *Proc. ICASSP-80*, pp.291-294, April 1980.

T.Q. Nguyen and P.P. Vaidyanathan, "Two-Channel Perfect--Reconstruction FIR QMF Structures Which Yield Linear-Phase Analysis and Synthesis Filters," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol. ASSP-37, No. 5, pp.676-690, May 1989.

M.J.T. Smith and T.P. Barnwell, "A New Filter Bank Theory for Time-Frequency Representation," *IEEE Trans. on Acoust., Speech and Signal Proc.*, Vol. ASSP-35, No.3, March 1987, pp. 314-327.

P.P. Vaidyanathan, *Multirate Filter Banks*, Prentice Hall, to appear.

**Pyramid Transforms:** see [BUR83, BUR89, ROS84].

**Multidimensional Filter Banks:** (see also [KOV92])

G. Karlsson and M. Vetterli, "Theory of Two-Dimensional Multirate Filter Banks," *IEEE Trans. on Acoust., Speech, Signal Proc.*, Vol.38, No.6, pp.925-937, June 1990.

M. Vetterli, "Multi-Dimensional Subband Coding: Some Theory and Algorithms," *Signal Processing*, Vol. 6, No.2, pp. 97-112, Feb. 1984.

M. Vetterli, J. Kovacevic and D. Le Gall, "Perfect Reconstruction Filter Banks for HDTV Representation and Coding," *Image Communication*, Vol.2, No.3, Oct.1990, pp.349-364.

E. Viscito and J. Allebach, "The Analysis and Design of Multidimensional FIR Perfect Reconstruction Filter Banks for Arbitrary Sampling Lattices," *IEEE Trans. Circuits and Systems*, Vol.38, pp.29-42, Jan. 1991.

**Multidimensional Wavelets:** (see also [ANI90, KOV92, MAL89a, MAL89b, MAL89c, MAL89d]).

J.C. Feauveau, "Analyse Multirésolution pour les Images avec un Facteur de Résolution 2," [in French], Traitement du Signal, Vol. 7, No. 2, pp. 117-128, July 1990.

K. Gröchenig and W.R. Madych, "Multiresolution analysis, Haar bases and self-similar tilings of $R^n$," submitted to *IEEE Trans. on Info. Theory*, Special Issue on wavelet transforms and multiresolution signal analysis, Jan.1992.

# Biorthogonal Wavelets

Albert Cohen

**Abstract.** In this chapter, we study the construction of biorthogonal bases of wavelets which generalize the orthonormal bases and have interesting properties in signal processing. We describe the class of subband coding schemes associated with these wavelets and we give necessary and sufficient conditions for frame bounds which ensure the stability of the decomposition-reconstruction algorithm. We finally present the example of compactly supported spline wavelets which can be generated by this approach. The results presented in this chapter are mainly joint work with I. Daubechies and J. C. Feauveau.

## §1. Introduction

In recent years, orthonormal wavelet bases have revealed to be a powerful tool in applied mathematics and digital signal processing. The possibility of data compression offered by a multiscale decomposition leads to some very good results in speech [8] or image [1] coding or fast numerical analysis of operators [2].

One of the main reasons for this success is the existence of a Fast Wavelet Transform algorithm (FWT) which only requires a number of operations proportional to the size of the initial discrete data. This algorithm relates the orthonormal wavelet bases with more classic tools of digital signal processing such as subband coding schemes and discrete filters.

We can describe in four steps the connections between these different domains:

a) Wavelets bases are usually defined from the data of a multiresolution analysis; *i.e.*, a ladder of approximation subspaces of $L^2(\mathbb{R})$

$$\{0\} \to \cdots V_1 \subset V_0 \subset V_{-1} \cdots \to L^2(\mathbb{R}) \tag{1}$$

which satisfy the following properties

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j-1} \Leftrightarrow f(2^j x) \in V_0. \tag{2}$$

Wavelets–A Tutorial in Theory and Applications
C. K. Chui (ed.), pp. 123–152.

ISBN 0-12-174590-2

There exists a *scaling function* $\phi(x)$ in $V_0$ such that

$$\{\phi_k^j\}_{k\in\mathbb{Z}} = \{2^{-j/2}\phi(2^{-j}x-k)\}_{k\in\mathbb{Z}} \tag{3}$$

is an orthonormal basis for $V_j$. The function $\phi$ has to satisfy a two-scale difference equation which expresses the embedded structure of the $V_j$-spaces

$$\phi(x) = 2\sum_{n\in\mathbb{Z}} h_n\phi(2x-n). \tag{4}$$

The wavelet $\psi$ is then defined by

$$\psi(x) = 2\sum_{n\in\mathbb{Z}}(-1)^n h_{1-n}\phi(2x-n) = 2\sum_{n\in\mathbb{Z}} g_n\phi(2x-n) \tag{5}$$

and its integer translates $\{\psi(x-k)\}_{k\in\mathbb{Z}}$ form an orthonormal basis for the orthogonal complement $W_0$ of $V_0$ in $V_{-1}$. The functions

$$\{\psi_k^j\}_{k\in\mathbb{Z}} = \{2^{-j/2}\psi(2^{-j}x-k)\}_{k\in\mathbb{Z}},$$

thus, characterize the additional details between two levels of approximation ($V_j$ and $V_{j-1}$). By a *telescoping argument* using the ladder structure (1), the whole set $\{\psi_k^j\}_{j,k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{R})$. More details on multiresolution analysis can be found in [14] and [15].

b) In the Fourier domain, Equations (4) and (5) can be rewritten as

$$\hat{\phi}(2\omega) = m_0(\omega)\hat{\phi}(\omega) \text{ with } m_0(\omega) = \sum_{n\in\mathbb{Z}} h_n e^{-in\omega} \tag{6}$$

$$\hat{\psi}(2\omega) = m_1(\omega)\hat{\phi}(\omega) = e^{-i\omega}\overline{m_0(\omega+\pi)}\hat{\phi}(\omega) \tag{7}$$

where $m_0(\omega)$ is a $2\pi$ periodic function that satisfies the following two properties (due to the multiresolution analysis axioms)

$$|m_0(\omega)|^2 + |m_0(\omega+\pi)|^2 = 1; \tag{8}$$

$$m_0(0) = 1 \quad \text{and} \quad m_0(\pi) = 0. \tag{9}$$

Here, $m_0$ and $m_1$ are the transfer functions of a pair of low-pass and high-pass filters known in signal processing as Conjugate Quadrature Filters (CQF, see [16]). These discrete filters are the key of the FWT algorithm: To analyze a discrete signal $s_n$, one identifies it with the coordinates of a function in $V_0$; *i.e.*, $f_0 = \sum_{n\in\mathbb{Z}} s_n\phi(x-n)$. The coordinates of the signal in $V_1$ (resp. $W_1$) are then obtained by applying the discrete filter $\overline{m_0(\omega)}$ (resp. $\overline{m_1(\omega)}$) followed by a decimation of one sample out of two to keep the same total amount of information. It is then possible to iterate this decomposition process on the coarser approximation in the following way: $V_1 \to V_2 \overset{\perp}{\oplus} W_2$, $V_2 \to V_2 \overset{\perp}{\oplus} W_3, \ldots$.

The reconstruction stage consists of refining the decimated sequence of approximation (resp. detail) coefficients by using $m_0(\omega)$ (resp. $m_1(\omega)$) as interpolating filters and adding these two components multiplied by two to get the finer approximation. This sequence of operations — filtering, downsampling, interpolation, reconstruction — is known in signal processing as a (two-channel) subband coding scheme, as illustrated in Figure 1.

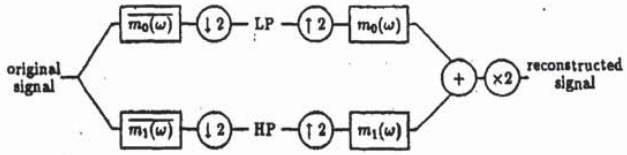


**Figure 1.** CQF subband coding scheme.

$\downarrow 2$ : removes one sample out of two

$\uparrow 2$ : insert a zero between each two samples.

c) Apparently the functions $\phi$ and $\psi$ do not play any role in this algorithm involving only the CQF pair. However, in many application, it is interesting to have these filters associated with a smooth wavelet and scaling function. Indeed, iterating (6) at each scale leads to

$$\hat{\phi}(\omega) = \prod_{k=1}^{+\infty} m_0(2^{-k}\omega). \tag{10}$$

This formula corresponds to an infinite number of iterations of the refinement process used in the reconstruction algorithm starting from a single low scale approximation coefficient. In the time domain, the limit of this process is the scaling function $\phi$ and if one starts the reconstruction on a detail coefficient, the limit is the wavelet $\psi$.

The smoothness of $\phi$ and $\psi$ will thus appear in the aspect of the low scale components which play an important role in data compression since many high scale details are thrown away.

d) In practice, the starting point to a multiscale analysis is a $2\pi$ periodic function $m_0(\omega)$ which satisfies Equations (8) and (9).

The scaling function is then defined by the infinite product (10). It has been shown in [7] and [12] that for a generic choice of $m_0(\omega)$, the scaling function is in $L^2(\mathbb{R})$ and satisfies

$$\langle \phi(x) | \phi(x-k) \rangle = \delta_{0,k}. \tag{11}$$

The wavelet is then derived from Equation (7). A particularly interesting class of CQF is the set of trigonometric polynomials $m_0(\omega)$ satisfying Equations (8)

and (9) since they correspond to finite impulse response (FIR) filters. They lead to compactly supported wavelets which have been constructed by I. Daubechies in [10] with the possibility of arbitrarily high regularity for $\phi$ and $\psi$ by choosing $m_0(\omega)$ in a smart way. ■

Unfortunately the CQF present some serious disadvantages for signal processing:

(1) They cannot be both FIR and linear phase (*i.e.*, with real and symmetrical coefficients) except the Haar filter which has no great interest since the associated wavelet $\psi = \psi_{[0,\frac{1}{2}]} - \chi_{[0,1]}$ is not continuous.
(2) Since they are the solution of the quadratic Equation (8), their coefficients are usually algebraic numbers with no simple expressions.
(3) Their design uses the Féjer-Riesz lemma (for FIR filters): $|m_0|^2$ is constructed at first and $m_0$ is derived using this lemma. However, this technique does not generalize to the multidimensional case.
(4) In the case of FIR filters, the subspaces $V_j$ have no simple and direct definition other than Span $\{\phi_k^j\}_{k\in\mathbb{Z}}$. For example, they cannot be composed of spline functions except for the Haar case.

For all these reasons, these filters are often rejected by the engineers for some specific applications. However, these disadvantages are not related to the structure of the subband coding scheme itself and they can be removed by using a more general class of filters. More precisely, we shall allow the decomposition and the reconstruction filters to be different. The result is a pair of *dual filters* $\{m_0, \tilde{m}_0\}$ which have to satisfy

$$\overline{m_0(\omega)}\tilde{m}_0(\omega) + \overline{m_0(\omega+\pi)}\tilde{m}_0(\omega+\pi) = 1. \tag{12}$$

These filters have been introduced in signal processing by M. Vetterli (see [17]).

Is it possible to mimic, in this more general setting, the construction of orthonormal wavelets from discrete filter that we describe previously? The answer is yes, but the orthonormality has been lost and the result is a pair of biorthogonal wavelet bases $\{\psi_k^j, \tilde{\psi}_k^j\}_{j,k\in\mathbb{Z}}$ which allow the following decomposition of any function $f$ in $L^2(\mathbf{R})$

$$\hat{f} = \sum_{j,k\in\mathbb{Z}} \langle f|\tilde{\psi}_k^j\rangle \psi_k^j = \sum_{j,k\in\mathbb{Z}} \langle f|\psi_k^j\rangle \tilde{\psi}_k^j. \tag{13}$$

In the next section of this chapter, we shall introduce the class of dual filters and their relation to biorthogonal wavelets.

In the third section, we shall discuss the additional conditions that must be filled by the filters to obtain biorthogonal wavelets bases. An important problem that does not occur in the orthonormal case is the frame bounds which relate the $L^2$ norm of a function and the $\ell^2$ norm of its coordinates in the [illegible] of Equation (13). These bounds are indeed crucial for the stability

of the decomposition-reconstruction algorithm. Two different strategies will be presented to check that these new wavelets form stable (or unconditional, or Riesz) bases.

Finally, we shall show in the last section that it is possible to build a (nonorthonormal) wavelet basis generated by a compactly supported spline function. Our approach is different from the technique developed by C. K. Chui in [3].

## §2. Dual filters and dual wavelets

### 2.1. Dual filters

Let us consider, in the most general sense, the subband coding scheme described in Figure 2. The decomposition is performed by the pair $\{\widetilde{m}_0, \widetilde{m}_1\}$; whereas $\{m_0, m_1\}$ are used for the reconstruction. A discrete signal $s_n$ can be represented by its discrete Fourier transform; *i.e.*,

$$s(\omega) = \sum_{n\in\mathbb{Z}} s_n e^{-in\omega}. \tag{14}$$

The decomposition stage transforms $s_n$ into an approximation sequence $a_n$ and a detail sequence $d_n$ defined by

$$a(2\omega) = \frac{1}{2}(\overline{\widetilde{m}_0(\omega)}s(\omega) + \overline{\widetilde{m}_0(\omega+\pi)}s(\omega+\pi)), \tag{15}$$

and

$$d(2\omega) = \frac{1}{2}(\overline{\widetilde{m}_1(\omega)}s(\omega) + \overline{\widetilde{m}_1(\omega+\pi)}s(\omega+\pi)). \tag{16}$$

And thus, the reconstructed signal $r_n$ can be written

$$r(\omega) = \alpha(\omega)s(\omega) + \beta(\omega)s(\omega+\pi) \tag{17}$$

with

$$\alpha(\omega) = m_0(\omega)\overline{\widetilde{m}_0(\omega)} + m_1(\omega)\overline{\widetilde{m}_1(\omega)}, \tag{18}$$

and

$$\beta(\omega) = m_0(\omega)\overline{\widetilde{m}_0(\omega+\pi)} + m_1(\omega)\overline{\widetilde{m}_1(\omega+\pi)}. \tag{19}$$

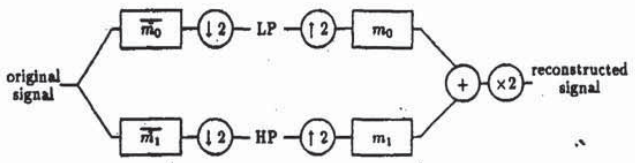


**Figure 2.** A general 2 channels subband coding scheme.

Perfect reconstruction is achieved for any signal if and only if $\alpha(\omega) = 1$ and $\beta(\omega) = 0$ for all $\omega$ in $[-\pi, \pi]$. This leads to the following system in which $m_0$ and $m_1$ can be considered as solutions and $\tilde{m}_0$ and $\tilde{m}_1$ as parameters namely:

$$\begin{cases} m_0(\omega)\overline{\tilde{m}_0(\omega)} + m_1(\omega)\overline{\tilde{m}_1(\omega)} = 1 \\ m_0(\omega)\overline{\tilde{m}_0(\omega+\pi)} + m_1(\omega)\overline{\tilde{m}_1(\omega+\pi)} = 0. \end{cases} \tag{20}$$

If we want to avoid the infinite impulse response solutions, we need to impose that the determinant is a monomial $\alpha e^{ik\omega}$, $\alpha \neq 0$, $k \in \mathbb{Z}$. Up to a shift and scalar multiplication on the filters, we choose, for sake of convenience, $\alpha = -1$ and $k = 1$. This leads to

$$\overline{m_0(\omega)}\tilde{m}_0(\omega) + \overline{m_0(\omega+\pi)}\tilde{m}_0(\omega+\pi) = 1, \tag{21}$$

and

$$m_1(\omega) = e^{-i\omega}\overline{\tilde{m}_0(\omega+\pi)}, \qquad \tilde{m}_1(\omega) = e^{-i\omega}\overline{m_0(\omega+\pi)}. \tag{22}$$

Equations (21) and (22) are thus the most general for finite impulse response subband coders with exact reconstruction (in the two channels case). We call $m_0$ and $\tilde{m}_0$ *dual filters*.

Clearly the special case $m_0 = \tilde{m}_0$ corresponds to the CQF. However, the disadvantages of the CQF can be avoided:

(1) If $m_0$ is fixed, $\tilde{m}_0$ can be found as the solution of a Bezout problem, which is equivalent to a linear system on the coefficients. The Féjer-Riesz lemma is no more needed.
(2) The coefficients can be very simple numerically and, in particular, they can have finite binary expansions, which are very useful for the implementation. They can also be real and symmetrical.

We now want to mimic the construction of wavelets in this more general setting. For this, we shall assume that the dual trigonometric polynomials $m_0$ and $\tilde{m}_0$ satisfy

$$m_0(0) = \tilde{m}_0(0) = 1 \quad \text{and} \quad m_0(\pi) = \tilde{m}_0(\pi) = 0. \tag{23}$$

### 2.2. Dual wavelets and scaling functions

Let us define, first in the sense of tempered distribution, the dual scaling functions and wavelets by

$$\hat{\phi}(\omega) = \prod_{k=1}^{+\infty} m_0(2^{-k}\omega), \tag{24}$$

$$\hat{\tilde{\phi}}(\omega) = \prod_{k=1}^{+\infty} \tilde{m}_0(2^{-k}\omega), \tag{25}$$

$$\hat{\psi}(\omega) = m_1\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right), \tag{26}$$

$$\hat{\tilde{\psi}}(\omega) = \tilde{m}_1\left(\frac{\omega}{2}\right)\hat{\tilde{\phi}}\left(\frac{\omega}{2}\right). \tag{27}$$

Since we want to use these functions to analyze $L^2(\mathbb{R})$ it is necessary that $\phi$ and $\tilde{\phi}$ are both square integrable. Note that in the CQF-orthonormal case, this is always satisfied as a consequence of Equation (8) (see [7]). For dual filters this is false in general.

In this section we shall assume that $\phi$ and $\tilde{\phi}$ (and thus $\psi$ and $\tilde{\psi}$) are in $L^2(\mathbb{R})$. Some precise conditions on $m_0$ and $\tilde{m}_0$ for this square integrability to hold will be stated in the next section.

Starting from this first assumption, we are going to prove the following result:

**Theorem 1.** *For any function $f$ in $L^2(\mathbb{R})$, we have in the $L^2$ sense,*

$$f = \lim_{J \to +\infty} \sum_{j=-J}^{J} \sum_{k \in \mathbb{Z}} \langle f | \psi_k^j \rangle \tilde{\psi}_k^j = \lim_{J \to +\infty} \sum_{j=-J}^{J} \sum_{k \in \mathbb{Z}} \langle f | \tilde{\psi}_k^j \rangle \psi_k^j. \tag{28}$$

We remark that this does not mean that $\{\psi_k^j, \tilde{\psi}_k^j\}_{j,k \in \mathbb{Z}}$ are biorthogonal bases, or even frames, since the summations are made in a precise way. However, this is a first step toward the construction of biorthogonal Riesz bases.

The proof of Theorem 1 is based on several lemmas that we shall comment here and prove in the appendix. We first introduce formally two approximation operators,

$$P_j(f) = \sum_{k \in \mathbb{Z}} \langle f | \tilde{\phi}_k^j \rangle \phi_k^j, \tag{29}$$

and

$$\tilde{P}_j(f) = \sum_{k \in \mathbb{Z}} \langle f | \phi_k^j \rangle \tilde{\phi}_k^j, \tag{30}$$

and two detail operators,

$$\Delta_j(f) = \sum_{k \in \mathbb{Z}} \langle f | \tilde{\psi}_k^j \rangle \psi_k^j, \tag{31}$$

and

$$\tilde{\Delta}_j(f) = \sum_{k \in \mathbb{Z}} \langle f | \psi_k^j \rangle \tilde{\psi}_k^j. \tag{32}$$

**Lemma 2.** *The operators $P_j, \tilde{P}_j, \Delta_j$, and $\tilde{\Delta}_j$ are bounded on $L^2(\mathbb{R})$. Moreover, their norm is independent of $j$.*

This result gives a rigorous meaning to the definition of these operators on $L^2(\mathbb{R})$. We then havethe following.

**Lemma 3.** *For all $j$ in $\mathbb{Z}$,*

$$P_{j-1} = P_j + \Delta_j \quad \text{and} \quad \tilde{P}_{j-1} = \tilde{P}_j + \tilde{\Delta}_j. \tag{33}$$

These two identities are essentially equivalent to the perfect reconstruction property of the subband coding scheme. By a *telescoping argument* we immediately obtain, for $J > 0$,

$$P_{-J-1} = P_J + \sum_{j=-J}^{J} \Delta_j, \tag{34}$$

and

$$\tilde{P}_{-J-1} = \tilde{P}_J + \sum_{j=-J}^{J} \tilde{\Delta}_j. \tag{35}$$

The next stage consists of letting $J$ tend to $+\infty$ and using:

**Lemma 4.** *For all $f$ in $L^2(\mathbb{R})$,*

$$\lim_{j\to+\infty} \|P_j(f)\|_{L^2} = \lim_{j\to+\infty} \|\tilde{P}_j(f)\|_{L^2} = 0, \tag{36}$$

*and*

$$\lim_{j\to-\infty} \|P_j(f) - f\|_{L^2} = \lim_{j\to-\infty} \|\tilde{P}_j(f) - f\|_{L^2} = 0. \tag{37}$$

The first limit is just a consequence of $\phi, \tilde{\phi} \in L^2(\mathbb{R})$. The second limit is due to some specific identities satisfied by $\phi$ and $\tilde{\phi}$ because of the hypotheses on the dual filters $m_0$ and $\tilde{m}_0$.

Combining Equations (34), (35), (36), and (37), we clearly obtain the result of Theorem 1; *i.e.*, for any $f$ in $L^2(\mathbb{R})$, we have

$$\lim_{J\to+\infty} \left\| f - \sum_{j=-J}^{J} \sum_{k\in\mathbb{Z}} \langle f|\tilde{\psi}_k^j\rangle \psi_k^j \right\|_{L^2} = \lim_{J\to+\infty} \left\| f - \sum_{j=-J}^{J} \sum_{k\in\mathbb{Z}} \langle f|\psi_k^j\rangle \tilde{\psi}_k^j \right\|_{L^2} = 0. \tag{38}$$

We now examine *the gap* existing between this system of *dual wavelets* and a pair of biorthogonal Riesz bases which will be obtained in Section 3.

### 2.3. From dual wavelets to biorthogonal Riesz bases

Recall that a Riesz basis of a Hilbert space $H$ is a family of vectors $\{e_\lambda\}_{\lambda\in\Lambda}$ such that

(1) The finite linear combinations of $e_\lambda$ are dense in $H$.
(2) There exist two strictly positive constants $C_1$ and $C_2$ such that, for any finite family of coefficients $\{\alpha_\lambda\}_{\lambda\in\Lambda_f}$ $(\Lambda_f \subset \Lambda)$,

$$C_1 \sum_{\lambda\in\Lambda_f} |\alpha_\lambda|^2 \le \left\| \sum_{\lambda\in\Lambda_f} \alpha_\lambda e_\lambda \right\|_H^2 \le C_2 \sum_{\lambda\in\Lambda_f} |\alpha_\lambda|^2. \tag{39}$$

An equivalent definition for a Riesz basis (see [19]) is given by the following.:

(1) The vectors $\{e_\lambda\}_{\lambda\in\Lambda}$ are linearly independent, and
(2) $\{e_\lambda\}_{\lambda\in\Lambda}$ is a frame; *i.e.*, there exist two strictly positive constants $D_1$ and $D_2$ such that for any $f$ in $H$

$$D_1\|f\|_H^2 \le \sum_{\lambda\in\Lambda} |\langle f|e_\lambda\rangle|^2 \le D_2\|f\|_H^2. \tag{40}$$

We shall rather use this second definition for $\psi_k^j$ and $\tilde{\psi}_k^j$. For these dual wavelets, the following holds.

**Theorem 5.** *$\{\psi_k^j, \tilde{\psi}_k^j\}_{j,k\in\mathbb{Z}}$ are a pair of biorthogonal Riesz bases if and only if*

(1) *for all $j, j', k, k'$ in $\mathbb{Z}$,*

$$\langle\psi_k^j|\tilde{\psi}_{k'}^{j'}\rangle = \delta_{j,j'}\delta_{k,k'}. \tag{41}$$

(2) *There exist two strictly positive constants $C$ and $\tilde{C}$ such that, for all $f$ in $L^2(\mathbb{R})$*

$$\sum_{j,k\in\mathbb{Z}} |\langle f|\psi_k^j\rangle|^2 \le C\|f\|^2 \tag{42}$$

and

$$\sum_{j,k\in\mathbb{Z}} |\langle f|\tilde{\psi}_k^j\rangle|^2 \le \tilde{C}\|f\|^2. \tag{43}$$

**Proof:**

a) If Equation (41) is satisfied, then any $f$ in the closed linear span of the $\psi_k^j$ with $(j,k) \neq (j_0,k_0)$ satisfies $\langle f|\tilde{\psi}_{k_0}^{j_0}\rangle = 0$ and thus $\psi_{k_0}^{j_0}$ cannot be in this closed linear span. Thus, the $\psi_k^j$ are linearly independent and the proof is similar for $\tilde{\psi}_k^j$.

b) By Equation (28), we have

$$\psi_{k_0}^{j_0} = \lim_{J\to+\infty} \sum_{j=-J}^{J} \sum_{k\in\mathbb{Z}} \langle\psi_{k_0}^{j_0}|\tilde{\psi}_k^j\rangle\psi_k^j, \tag{44}$$

and, thus, $\psi_{k_0}^{j_0}(1 - \langle\psi_{k_0}^{j_0}|\tilde{\psi}_{k_0}^{j_0}\rangle)$ lies in the closure of the $\psi_k^j$ for $(j,k) \neq (j_0,k_0)$. By linear independence, this implies

$$\langle\psi_{k_0}^{j_0}|\tilde{\psi}_{k_0}^{j_0}\rangle = 1. \tag{45}$$

Isolating any $\langle\psi_{k_0}^{j_0}|\tilde{\psi}_k^j\rangle\psi_k^j$ in (44) we also obtain for any $(j,k) \neq (j_0,k_0)$, by linear independence

$$\langle\psi_{k_0}^{j_0}|\psi_k^j\rangle = 0, \tag{46}$$

and thus (41).

c) Combining Equations (28), (42), and (43), we obtain for any $f$ in $L^2(\mathbb{R})$,

$$\|f\|^2 = \lim_{J \to +\infty} \sum_{j=-J}^{J} \sum_{k \in \mathbb{Z}} \langle f|\tilde{\psi}_k^j\rangle\langle\psi_k^j|f\rangle$$

$$\leq \left( \sum_{j,k \in \mathbb{Z}} |\langle f|\psi_k^j\rangle|^2 \right)^{1/2} \left( \sum_{j,k \in \mathbb{Z}} |\langle f|\tilde{\psi}_k^j\rangle|^2 \right)^{1/2}$$

$$\leq C^{1/2}\|f\| \left( \sum_{j,k \in \mathbb{Z}} |\langle f|\tilde{\psi}_k^j\rangle|^2 \right)^{1/2}.$$

This leads to

$$C^{-1}\|f\|^2 \leq \sum_{j,k \in \mathbb{Z}} |\langle f|\tilde{\psi}_k^j\rangle|^2, \tag{47}$$

and similarly

$$\tilde{C}^{-1}\|f\|^2 \leq \sum_{j,k \in \mathbb{Z}} |\langle f|\psi_k^j\rangle|^2. \tag{48}$$

The lower frame bounds are thus directly deduced from the upper bounds. According to the second definition and to the identity (41), $\{\psi_k^j, \tilde{\psi}_k^j\}_{j,k \in \mathbb{Z}}$ are a pair of biorthogonal Riesz bases. Any $f$ in $L^2(\mathbb{R})$ can be written

$$f = \sum_{j,k \in \mathbb{Z}} \langle f|\tilde{\psi}_k^j\rangle\psi_k^j = \sum_{j,k \in \mathbb{Z}} \langle f|\psi_k^j\rangle\tilde{\psi}_k^j \tag{49}$$

where these expansions are unique and converge unconditionally. ■

Before examining the type of dual filters leading to such bases, we shall prove two technical results that we shall use in the next section to check the hypothesis of Lemma 5. The first one deals with Equation (41).

**Lemma 6.** *Let $\phi_n$ and $\tilde{\phi}_n$ be defined for $n > 0$ by*

$$\hat{\phi}_n(\omega) = \prod_{k=1}^{n} m_0(2^{-k}\omega)\chi_{[-2^n\pi, 2^n\pi]}(\omega), \tag{50}$$

*and*

$$\hat{\tilde{\phi}}_n(\omega) = \prod_{k=1}^{n} \tilde{m}_0(2^{-k}\omega)\chi_{[-2^n\pi, 2^n\pi]}(\omega). \tag{51}$$

*Then if $\phi_n$ and $\tilde{\phi}_n$ converge in $L^2(\mathbb{R})$ to $\phi$ and $\tilde{\phi}$,*

$$\langle\phi(x-p)|\tilde{\phi}(x-\ell)\rangle = \delta_{p,\ell}, \tag{52}$$

and

$$\langle \psi_k^j | \tilde{\psi}_{k'}^{j'} \rangle = \delta_{j,j'} \delta_{k,k'}. \tag{53}$$

**Proof:** By recursion we can establish

$$\langle \phi_n(x-p) | \tilde{\phi}_n(x-\ell) \rangle = \delta_{p,\ell}. \tag{54}$$

Indeed,

$$\begin{aligned}
\langle \phi_n(x-p) | \tilde{\phi}_n(x-\ell) \rangle &= \frac{1}{2\pi} \int_{-2^n\pi}^{2^n\pi} \left( \prod_{k=1}^{n} m_0(2^{-k}\omega) \overline{\tilde{m}_0(2^{-k}\omega)} \right) e^{i(\ell-p)\omega} d\omega \\
&= \frac{2^n}{2\pi} \int_{-\pi}^{\pi} \left( \prod_{k=0}^{n-1} m_0(2^k\omega) \overline{\tilde{m}_0(2^k\omega)} \right) e^{i2^n(\ell-p)\omega} d\omega \\
&= \frac{2^n}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left( \prod_{k=2}^{n-1} m_0(2^k\omega) \overline{\tilde{m}_0(2^k\omega)} \right) \\
&\quad [m_0(\omega)\overline{\tilde{m}_0(\omega)} + m_0(\omega+\pi)\overline{\tilde{m}_0(\omega+\pi)}] e^{i2^n(\ell-p)\omega} d\omega \\
&= \langle \phi_{n-1}(x-p) | \tilde{\phi}_{n-1}(x-\ell) \rangle \\
&= \cdots = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(\ell-p)\omega} d\omega = \delta_{p,\ell}.
\end{aligned}$$

Consequently the $L^2$ convergence of $\phi_n$ and $\tilde{\phi}_n$ implies

$$\langle \phi(x-p) | \tilde{\phi}(x-\ell) \rangle = \delta_{p,\ell} \tag{55}$$

which can also be written as

$$\sum_{k \in \mathbb{Z}} (\hat{\bar{\phi}} \hat{\tilde{\phi}})(\omega + 2k\pi) = 1. \tag{56}$$

We now have

$$\begin{aligned}
\sum_{k \in \mathbb{Z}} (\hat{\bar{\psi}} \hat{\tilde{\psi}})(\omega + 2k\pi) &= \sum_{k \in \mathbb{Z}} (\overline{m}_1 \tilde{m}_1 \hat{\bar{\phi}} \hat{\tilde{\phi}}) \left( \frac{\omega}{2} + k\pi \right) \\
&= \left[ \overline{m}_1 \tilde{m}_1 \left( \frac{\omega}{2} \right) + \overline{m_1} \tilde{m}_1 \left( \frac{\omega}{2} + \pi \right) \right] \sum_{k \in \mathbb{Z}} (\hat{\bar{\phi}} \hat{\tilde{\phi}}) \left( \frac{\omega}{2} + 2k\pi \right) \\
&= 1,
\end{aligned}$$

and

$$\begin{aligned}
\sum_{k \in \mathbb{Z}} (\hat{\bar{\psi}} \hat{\tilde{\phi}})(\omega + 2k\pi) &= \sum_{k \in \mathbb{Z}} (\overline{m}_1 \tilde{m}_0 \hat{\bar{\phi}} \hat{\tilde{\phi}}) \left( \frac{\omega}{2} + k\pi \right) \\
&= \left[ \overline{m}_1 \tilde{m}_0 \left( \frac{\omega}{2} \right) + \overline{m_1} \tilde{m}_0 \left( \frac{\omega}{2} + \pi \right) \right] \sum_{k \in \mathbb{Z}} (\hat{\bar{\phi}} \hat{\tilde{\phi}}) \left( \frac{\omega}{2} + 2k\pi \right) \\
&= 0.
\end{aligned}$$

Consequently,

$$\langle \psi(x-k) | \tilde{\psi}(x-\ell) \rangle = \delta_{k,\ell} \tag{57}$$

and

$$\langle \psi(x-k) | \tilde{\phi}(x-\ell) \rangle = 0. \tag{58}$$

If we now define $V_j = \text{Span}\{\phi_k^j\}_{k\in\mathbb{Z}}$, $W_j = \text{Span}\{\psi_k^j\}_{k\in\mathbb{Z}}$, and $\tilde{V}_j = \text{Span}\{\tilde{\phi}_k^j\}_{k\in\mathbb{Z}}$, $\tilde{W}_j = \text{Span}\{\tilde{\psi}_k^j\}_{k\in\mathbb{Z}}$, it follows from Equation (56) that for any $j$, $\tilde{V}_j$ is orthogonal to $W_j$. Since, for all $j' < j$, we have $\tilde{W}_{j'} \subset \tilde{V}_j$, it follows that $W_j$ and $\tilde{W}_{j'}$ are orthonormal when $j' < j$ and, by a symmetrical argument, when $j < j'$. Consequently, we obtain by Equation (55),

$$\langle \psi_k^j | \tilde{\psi}_{k'}^{j'} \rangle = \delta_{j,j'}\delta_{k,k'}, \tag{59}$$

and the lemma is proved. ■

The last lemma of this section deals with the upper frame bounds (42) and (43).

**Lemma 7.** *Suppose that the function $\phi$ satisfies*

$$\sup_{\omega\in\mathbb{R}} \sum_{k\in\mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^{2-\sigma} < +\infty, \tag{60}$$

*for some $\sigma > 0$, and*

$$\sup_{\omega\in\mathbb{R}} (1+|\omega|)^{\sigma} |\hat{\phi}(\omega)| < +\infty. \tag{61}$$

*Then there exists a constant $C$ such that, for all $f$ in $L^2(\mathbb{R})$,*

$$\sum_{j,k\in\mathbb{Z}} |\langle f | \psi_k^j \rangle|^2 \le C\|f\|^2. \tag{62}$$

*The same holds for $\tilde{\psi}$ and $\tilde{\phi}$.*

**Proof:** Since $\hat{\psi}(\omega) = m_1\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right)$ has at least a first order zero at the origin, using Equations (58) and (57), we may conclude

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}(2^j\omega)|^{\sigma} \le C_1, \tag{63}$$

and

$$\sum_{k\in\mathbb{Z}} |\hat{\psi}(\omega + 2k\pi)|^{2-\sigma} \le C_2 \tag{64}$$

uniformly in $\omega$.

Using the Plancherel and the Poisson formulas, we can derive for all $f$ in $L^2(\mathbb{R})$,

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} |\langle f|\psi_k^j\rangle|^2 &= \frac{1}{4\pi^2}\sum_{k\in\mathbb{Z}} 2^j \left|\int_{\mathbb{R}} \hat{f}(\omega)\overline{\hat{\psi}(2^j\omega)}e^{-i2^jk\omega}d\omega\right|^2 \\
&= \frac{1}{4\pi^2}\sum_{k\in\mathbb{Z}} 2^{-j} \left|\int_{\mathbb{R}} \hat{f}(2^{-j}\omega)\overline{\hat{\psi}(\omega)}e^{-ik\omega}d\omega\right|^2 \\
&= \frac{2^{-j}}{2\pi}\int_0^{2\pi} \left|\sum_{\ell\in\mathbb{Z}} \hat{f}(2^{-j}(\omega+2\ell\pi))\overline{\hat{\psi}(\omega+2\ell\pi)}\right|^2 d\omega \\
&\le \frac{2^{-j}}{2\pi}\int_0^{2\pi} \left(\sum_{\ell\in\mathbb{Z}} |\hat{f}(2^{-j}(\omega+2\ell\pi))|\,|\hat{\psi}(\omega+2\ell\pi)|^{\frac{\sigma}{2}}|\hat{\psi}(\omega+2\ell\pi)|^{1-\frac{\sigma}{2}}\right)^2 d\omega \\
&\le \frac{2^{-j}}{2\pi}\int_0^{2\pi} \left(\sum_{\ell\in\mathbb{Z}} |\hat{f}(2^{-j}(\omega+2\ell\pi))|^2|\hat{\psi}(\omega+2\ell\pi)|^{\sigma}\right)\left(\sum_{\ell\in\mathbb{Z}} |\hat{\psi}(\omega+2\ell\pi)|^{2-\sigma}\right) d\omega \\
&\le C_2\frac{2^{-j}}{2\pi}\int_{\mathbb{R}} |\hat{f}(2^{-j}\omega)|^2|\hat{\psi}(\omega)|^{\sigma}d\omega \\
&\le \frac{1}{2\pi}C_2\int_{\mathbb{R}} |\hat{f}(\omega)|^2|\hat{\psi}(2^j\omega)|^{\sigma}d\omega.
\end{aligned}$$

Summing over all the scales $j \in \mathbb{Z}$ and using Equation (59), we obtain

$$\begin{aligned}
\sum_{j,k\in\mathbb{Z}} |\langle f|\psi_k^j\rangle|^2 &\le \frac{C_1C_2}{2\pi}\int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega \\
&= C_1C_2\|f\|^2,
\end{aligned}$$

and this concludes the proof. ■

We are now ready to characterize the dual filters which lead to biorthogonal wavelet bases, such that the functions $\phi$ and $\tilde{\phi}$ satisfy the hypotheses of Lemmas 6 and 7.

## §3. Biorthogonal wavelet bases and stable subband coding schemes

In this section we shall present two strategies to design the dual filters so that the associated dual wavelets $\psi$ and $\tilde{\psi}$ generate a pair of biorthogonal Riesz bases. In other words, we shall establish some conditions for the stability of the FWT algorithm since this is equivalent to the frame bounds inequalities that we require on our multiscale bases.

The first strategy uses estimates on the decay at infinity of the functions $\hat{\phi}(\omega)$ and $\hat{\tilde{\phi}}(\omega)$ that can also be found in [6]. It furnishes a sufficient condition for biorthogonality and stability.

The second strategy is based on the study of two operators associated with the dual filters. We show here that it leads to sufficient conditions and it was proved in [4] that this criterion is also necessary. It is thus a sharper strategy but it is only tractable for filters of reasonable size.

### 3.1. A Fourier criterion

Since we have assumed that $m_0$ and $\tilde{m}_0$ vanish at $\omega = \pi$, it is possible to express these filters in a factorized form

$$m_0(\omega) = \left(\frac{1+e^{i\omega}}{2}\right)^N p(\omega), \tag{65}$$

and

$$\tilde{m}_0(\omega) = \left(\frac{1+e^{i\omega}}{2}\right)^{\tilde{N}} \tilde{p}(\omega). \tag{66}$$

The following results give a sufficient criterion based on the properties of the trigonometric polynomials $p$ and $\tilde{p}$.

**Theorem 8.** *Suppose that the function $p(\omega)$ satisfies*

$$\inf_{j>0}\left[\max_{\omega\in\mathbb{R}}\left|\prod_{k=1}^{j} p(2^{-k}\omega)\right|^{\frac{1}{j}}\right] < 2^{N-\frac{1}{2}}. \tag{67}$$

*Then*

(1) $\hat{\phi}_n(\omega) = \prod_{k=1}^{n} m_0(2^{-k}\omega)\chi_{[-2^n\pi,2^n\pi]}(\omega)$ *converges to $\hat{\phi}(\omega)$ in $L^2(\mathbb{R})$, and*

(2) *The conditions in (60) and (61) of Lemma 7 are satisfied.*

*If $\tilde{p}(\omega)$ satisfies a similar hypothesis, then the dual filters $m_0$ and $\tilde{m}_0$ generate biorthogonal Riesz bases of wavelets.*

**Proof:** The hypothesis (67) implies that for some $j > 0$,

$$\max_{\omega\in\mathbb{Z}}\left|\prod_{k=1}^{j} p(2^{-j}\omega)\right| < 2^{j(N-\frac{1}{2}-\epsilon)}, \quad (\epsilon > 0). \tag{68}$$

We now write

$$\begin{aligned}|\hat{\phi}_n(\omega)| &= \left|\sum_{k=1}^{n}\left(\frac{1+e^{i2^{-k}\omega}}{2}\right)\right|^N \left|\prod_{k=1}^{n} p(2^{-k}\omega)\right| \chi_{[-2^n\pi,2^n\pi]}(\omega)\\ &= \left|\frac{\sin\left(\frac{\omega}{2}\right)}{2^n \sin(2^{-n-1}\omega)}\right|^N \left|\prod_{k=1}^{n} p(2^{-k}\omega)\right| \chi_{[-2^n\pi,2^n\pi]}(\omega).\end{aligned}$$

Since $|\sin(2^{-n-1}\omega)| > C2^{-n-1}|\omega|$ when $\omega$ is in $[-2^n\pi, 2^n\pi]$, we have

$$|\hat{\phi}_n(\omega)| \leq C(1+|\omega|)^{-N} \left| \prod_{k=1}^{n} p(2^{-k}\omega) \right| \chi_{[-2^n\pi, 2^n\pi]}(\omega). \tag{69}$$

We treat the second product in the following way: The contribution of the factors $p(2^{-k}\omega)$ for $k \geq \frac{\log(1+|\omega|)}{\log 2}$ can be globally majorated by a constant since the infinite product converges to a smooth function. We divide the other factors in packets of size $j$ that we can majorate using Equation (68). This leads to

$$\begin{aligned} |\hat{\phi}_n(\omega)| &\leq C(1+|\omega|)^{-N} \left(2^{j(N-\frac{1}{2}-\epsilon)}\right)^{\frac{\log(1+|\omega|)}{j\log 2}} \\ &\leq C(1+|\omega|)^{-\frac{1}{2}-\epsilon}. \end{aligned}$$

Since this bound is independent of $n$ and $\hat{\phi}_n(\omega)$ converges pointwise to $\hat{\phi}(\omega)$, we can apply Lebesgue's dominated convergence theorem to conclude that $\phi_n$ converges in $L^2(\mathbb{R})$ to $\phi$. At the limit, we also have

$$|\hat{\phi}(\omega)| \leq C(1+|\omega|)^{-1/2-\epsilon} \tag{70}$$

which immediately implies the conditions (60) and (61) of Lemma 7 by choosing $0 < \sigma < \min\left(\frac{1}{2}, \frac{4\epsilon}{2\epsilon+1}\right)$. If $\tilde{p}(\omega)$ satisfies a similar condition, then the results of Lemmas 6 and 7 can be applied to $\psi$ and $\tilde{\psi}$ and by Theorem 5, the families $\{\psi_k^j, \tilde{\psi}_k^j\}_{j,k\in\mathbb{Z}}$ are biorthogonal Riesz bases for $L^2(\mathbb{R})$. ■

This criterion was used in [6] to construct many biorthogonal wavelet bases but it is not sharp. In particular, Equation (70) is not strictly necessary to have $\phi$ and $\psi$ in $L^2(\mathbb{R})$. These functions can be very lacunary; *i.e.*, their Fourier transform can have a bad decay but only at some points which occur less and less frequently at infinity so that they are still square integrable. We now present a sharper criterion based on a different approach.

### 3.2. A matrix base criterion

Let us first introduce the basic tool which will be used in this approach.

**Definition 9.** *Let $m_0(\omega)$ be a trigonometric polynomial such that $m_0(0) = 1$ and $m_0(\pi) = 0$. The transition operator $T_0$ associated with this filter acts on $2\pi$ periodic functions in the following way*

$$T_0 f(\omega) = \left|m_0\left(\frac{\omega}{2}\right)\right|^2 f\left(\frac{\omega}{2}\right) + \left|m_0\left(\frac{\omega}{2}+\pi\right)\right|^2 f\left(\frac{\omega}{2}+\pi\right). \tag{71}$$

This operator appears in the works of W. Lawton [12] and J.P. Conze and A. Raugi [9] for the study of the orthonormal case. It can also be useful to estimate the Sobolev regularity of the scaling function associated to $m_0$

(see [11], [18], and [5]). Here, we will need two lemmas which give some basic properties of $T_0$.

**Lemma 10.** *Let* $m_0(\omega) = \sum_{k=0}^{N} c_k e^{ik\omega}$ *and* $T_0$ *the associated transition operator, and define the* $2N+1$ *dimensional space*

$$E_N = \left\{ \sum_{k=-N}^{N} c_k e^{ik\omega} \;\middle|\; (c_{-N}, \ldots, c_N) \in \mathbb{C}^{2N+1} \right\}, \tag{72}$$

*and its subspace*

$$F_N = \left\{ \sum_{k=-N}^{N} c_k e^{ik\omega} \;\middle|\; \sum_{k=-N}^{N} c_k = 0 \right\}. \tag{73}$$

*Then* $E_N$ *and* $F_N$ *are stable under the action of* $T_0$.

**Proof:** From the definition of $T_0$, it is clear that if $|m_0|^2$ and $f$ are two elements of $E_N$, then $T_0 f$ is also in $E_N$.

A trigonometric polynomial $f$ is in $F_N$ if and only if $f \in E_N$ and $f(0) = 0$. Consequently, if $f$ is in $F_N$, we have

$$T_0 f(0) = |m_0(0)|^2 f(0) + |m_0(\pi)|^2 f(\pi) = 0, \tag{74}$$

which proves that $F_N$ is stable under $T_0$. ■

**Remarks.** If we consider the Fourier expansion

$$|m_0(\omega)|^2 = \sum_{k=-N}^{N} H_k e^{ik\omega}, \tag{75}$$

then the matrix of $T_0$ restricted to $E_N$ is given by

$$M_0 = (2H_{i-2j})_{i,j=-N\ldots N} = 2 \begin{bmatrix} H_N & 0 & \cdots & \cdots & 0 \\ H_{N-2} & H_{N-1} & H_N & & \vdots \\ \vdots & \vdots & \vdots & & 0 \\ H_{-N} & H_{-N+1} & \vdots & & H_N \\ 0 & 0 & H_{-N} & & H_{N-2} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & \cdots & H_{-N} \end{bmatrix}. \tag{76}$$

Since $|m_0(0)|^2 = \sum_{k=-N}^{N} H_k = 1$ and $|m_0(\pi)|^2 = \sum_{k=-N}^{N} (-1)^k H_k = 0$, it follows that

$$\sum_k H_{2k} = \sum_k H_{2k+1} = \frac{1}{2} \tag{77}$$

and thus the row vector $\mu = (1, 1, \ldots, 1)$ satisfies

$$\mu T_0 = \mu. \tag{78}$$

This is another way to show that $F_N$ is stable since $F_N = (\mathbf{C}\mu)^\perp$ but it also shows that 1 is an eigenvalue of $T_0$ and if this eigenvalue is not degenerated then it is not in the spectrum of $T_0$ restricted to $F_N$.

The second lemma makes a connection between the iteration of the operator $T_0$ and the sequence $\phi_n$ which must converge to $\phi$ in $L^2(\mathbb{R})$.

**Lemma 11.** *For any $2\pi$ periodic function $f(\omega)$, it follows that*

$$\int_{-\pi}^{\pi} T_0^n f(\omega) d\omega = \int_{-2^n\pi}^{2^n\pi} f(2^{-n}\omega) \prod_{k=1}^{n} |m_0(2^{-k}\omega)|^2 d\omega = \int f(2^{-n}\omega) |\hat{\phi}_n(\omega)|^2 d\omega. \tag{79}$$

**Proof:** We prove it by induction. Equation (79) is trivial for $n = 0$ and if it is satisfied for some $n \geq 0$, then

$$\begin{aligned}
&\int_{-\pi}^{\pi} T_0^{n+1} f(\omega) d\omega = \int_{-\pi}^{\pi} T_0^n (T_0 f)(\omega) d\omega \\
&= \int_{-2^n\pi}^{2^n\pi} T_0 f(2^{-n}\omega) \prod_{k=1}^{n} |m_0(2^{-k}\omega)|^2 d\omega \\
&= \int_{-2^n\pi}^{2^n\pi} \left[ f(2^{-n-1}\omega) |m_0(2^{-n-1}\omega)|^2 + f(2^{-n-1}\omega + \pi) |m_0(2^{-n-1}\omega + \pi)|^2 \right] \\
&\quad \prod_{k=1}^{n} |m_0(2^{-k}\omega)|^2 d\omega \\
&= \int_{-2^{n+1}\pi}^{2^{n+1}\pi} f(2^{-n-1}\omega) \prod_{k=1}^{n+1} |m_0(2^{-k}\omega)|^2 d\omega. \quad \blacksquare
\end{aligned}$$

We are now ready to state a criterion based on transition operators.

**Theorem 12.** *Let $\lambda$ be the largest eigenvalue of $T_0$ restricted to $F_N$. If $|\lambda| < 1$, then*

(1) *$\phi_n$ converges to $\phi$ in $L^2(\mathbb{R})$, and*
(2) *the conditions (60) and (61) of Lemma 7 are satisfied.*

*If the same holds for the operator $\tilde{T}_0$ associated to $\tilde{m}_0$, then the dual filters $m_0$ and $\tilde{m}_0$ generate biorthogonal Riesz bases of wavelets.*

**Proof:** Let us define the trigonometric polynomial $c$ by

$$c(\omega) = 1 - \cos\omega. \tag{80}$$

It is clear that $c(\omega)$ is in $F_N$. Applying Lemma 11 and using the hypothesis $|\lambda| < 1$, we obtain

$$\begin{aligned}\int |\hat{\phi}_n(\omega)|^2 c(2^{-n}\omega)d\omega &= \int_{-\pi}^{\pi} T_0^n c(\omega)d\omega \\ &\leq \sqrt{2\pi}\|T_0^n c\|_{L^2} \\ &\leq C\left(\frac{1+|\lambda|}{2}\right)^n = C2^{-n\epsilon},\end{aligned}$$

with $\epsilon = \frac{1}{\log 2}[\log 2 - \log(1+|\lambda|)] > 0$ because $|\lambda| < 1$. This leads to a Littlewood-Paley type of estimate. Indeed, since $c(\omega)$ is positive and $c(\omega) \geq 1$ when $\frac{\pi}{2} \leq |\omega| \leq \pi$, we have

$$\int_{2^{n-1}\pi \leq |\omega| \leq 2^n\pi} |\hat{\phi}_n(\omega)|^2 d\omega \leq C2^{-n\epsilon}. \tag{81}$$

This estimate is also valid if we replace $\phi_n$ by $\phi$ because we have $|\hat{\phi}(\omega)| = |\hat{\phi}_n(\omega)\hat{\phi}(2^{-n}\omega)| \leq |\hat{\phi}_n(\omega)| \max_{|\omega|\leq\pi}(|\hat{\phi}(\omega)|)$ when $|\omega| \leq 2^n\pi$. Consequently

$$\int_{2^{n-1}\pi \leq |\omega| \leq 2^n\pi} |\hat{\phi}(\omega)|^2 d\omega \leq C2^{-n\epsilon}, \tag{82}$$

which means that $\phi$ is not only in $L^2(\mathbb{R})$ but also in the Besov space $B_2^{\epsilon,\infty}(\mathbb{R})$.

Let us now prove the $L^2$ convergence of $\phi_n$ to $\phi$. Since $m_0(0) = \hat{\phi}(0) = 1$, there exists an $\alpha$ in $(0;\pi]$ such that

$$|\omega| \leq \alpha \Rightarrow |\hat{\phi}(\omega)| \geq C > 0. \tag{83}$$

We now divide $\phi_n$ in two parts: $\phi_n = \phi_n^1 + \phi_n^2$ with

$$\hat{\phi}_n^1(\omega) = \hat{\phi}_n(\omega)\chi_{[-2^n\alpha,2^n\alpha]}(\omega), \tag{84}$$

and

$$\hat{\phi}_n^2(\omega) = \hat{\phi}_n(\omega)\left[\chi_{[-2^n\pi,2^n\pi]}(\omega) - \chi_{[-2^n\alpha,2^n\alpha]}(\omega)\right]. \tag{85}$$

Clearly $\hat{\phi}_n^1(\omega)$ converges pointwise to $\hat{\phi}(\omega)$ and by (83), we have

$$|\hat{\phi}_n^1(\omega)| \leq \frac{|\hat{\phi}(\omega)|}{C}. \tag{86}$$

By Lebesgue's theorem, $\phi_n^1$ converges to $\phi$ in $L^2$. We also have

$$\begin{aligned}\int |\hat{\phi}_n^2(\omega)|^2 d\omega &= \int_{2^n\alpha \leq |\omega| \leq 2^n\pi} |\hat{\phi}_n(\omega)|^2 d\omega \\ &\leq \frac{1}{c(\alpha)} \int |\hat{\phi}_n(\omega)|^2 c(2^{-n}\omega)d\omega \\ &\leq C2^{-n\epsilon},\end{aligned}$$

and thus $\phi_n^2$ tends to zero in $L^2$ and consequently $\phi$ is also the $L^2$ limit of $\phi_n$.

To prove Equations (60) and (61), we shall use the estimate (82). We first remark that since $m_0(\pi) = m_0(-\pi) = 0$, the scaling function satisfies

$$\hat{\phi}(2k\pi) = 0 \qquad \text{if} \quad k \in \mathbb{Z}/\{0\}. \tag{87}$$

Using a first order Taylor development, we can write

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^{2-\sigma} &\le \int_{\mathbf{R}} \left| \frac{d}{d\omega}[|\hat{\phi}|^{2-\sigma}] \right| d\omega \\
&= \int \left| \frac{d}{d\omega}[(|\hat{\phi}|^2)^{1-\frac{\sigma}{2}}] \right| d\omega \\
&= \int \left| \left(1 - \frac{\sigma}{2}\right) \frac{d}{d\omega}[|\hat{\phi}|^2] \right| |\hat{\phi}|^{-\frac{\sigma}{2}} d\omega \\
&\le \int |2-\sigma| \left| \frac{d\hat{\phi}}{d\omega} \right| |\hat{\phi}(\omega)|^{1-\frac{\sigma}{2}} d\omega \\
&\le C \left( \int \left| \frac{d\hat{\phi}}{d\omega} \right|^2 d\omega \right)^{1/2} \left( \int |\hat{\phi}(\omega)|^{2-\sigma} d\omega \right)^{1/2}.
\end{aligned}$$

The first factor proportional to the $L^2$ norm of $x\phi(x)$ which is finite since $\phi$ is square integrable and compactly supported.

To evaluate the second factor we compute the integral of $|\hat{\phi}(\omega)|^{2-\sigma}$ on a dyadic ring $2^{n-1}\pi \le |\omega| \le 2^n\pi$. By the Hölder inequality and (82) we obtain

$$\begin{aligned}
\int_{2^{n-1}\pi \le |\omega| \le 2^n\pi} |\hat{\phi}(\omega)|^{2-\sigma} d\omega &\le \left[ \int_{2^{n-1}\pi \le |\omega| \le 2^n\pi} |\hat{\phi}(\omega)|^2 d\omega \right]^{1-\frac{\sigma}{2}} (2^n\pi)^{\frac{\sigma}{2}} \\
&\le C 2^{-n\epsilon(1-\frac{\sigma}{2}) + n\frac{\sigma}{2}}.
\end{aligned}$$

The second factor will thus be finite and (57) will be satisfied if we choose $\sigma$ such that $\epsilon\left(1 - \frac{\sigma}{2}\right) - \frac{\sigma}{2} > 0$; *i.e.*, $\sigma < \frac{2\epsilon}{1+\epsilon}$. We apply the same method to show (58). If $|\omega|$ is in $[2^{n-1}\pi, 2^n\pi]$ for $n \ge 1$, we can write

$$\begin{aligned}
|\hat{\phi}(\omega)|^2 &\le \int_{2^{n-1}\pi \le |\omega| \le 2^n\pi} \left| \frac{d}{d\omega}[|\hat{\phi}|^2] \right| d\omega \\
&\le C \left( \int_{\mathbf{R}} \left| \frac{d\hat{\phi}}{d\omega} \right|^2 d\omega \right)^{1/2} \left( \int_{2^{n-1}\pi \le |\omega| \le 2^n\pi} |\hat{\phi}(\omega)|^2 d\omega \right)^{1/2} \\
&\le C 2^{-n\frac{1}{2}} \le C(1 + |\omega|)^{-\frac{1}{2}},
\end{aligned}$$

and thus (61) is satisfied if we choose $\sigma \le \frac{1}{4}$.

This concludes the proof of the theorem. ■

**Remarks.**

(1) In the proof we use estimates of $|\hat{\phi}|^2$ or its integral on dyadic rings $2^{n-1}\pi \leq |\omega| \leq 2^n\pi$, for $n \geq 1$. The case $|\omega| \leq \pi$ does not cause any problem since $\hat{\phi}(\omega)$ is a smooth function.

(2) The criterion that we have established is in fact sharp; *i.e.*, it is necessary and sufficient to have biorthogonal Riesz bases. The complete proof of this fact, which is very technical, can be found in [4].

Using our theoretical results, we shall now build an important family of biorthogonal wavelets where the functions $\phi$ and $\psi$ are piecewise polynomial.

## §4. Biorthogonal wavelets and splines

The connections between wavelet theory and spline theory come out naturally from the general framework of multiresolution analysis. Indeed, for a given $N \geq 1$, we can define a multiresolution analysis $\{V_j\}_{j\in\mathbb{Z}}$ by

$$V_j = \left\{ f \in L^2(\mathbb{R}) \cap C^{N-1}(\mathbb{R}) \Big| \left(\frac{d}{dx}\right)^{N+1} f = \sum_{k\in\mathbb{Z}} \alpha_k \delta_{2^j k} \right\}, \tag{88}$$

where $\delta_{2^j k}$ is the delta function localized at the point $2^j k$. In other words, $V_j$ is the space of square integrable functions which are piecewise polynomial of degree $N$ on the dyadic intervals $[2^j k, 2^j(k+1)]$ with $C^{N-1}$ continuity at the nod points $2^j k$.

A natural generator for these spaces is the $N$th degree $B$-spline function defined by

$$\phi^N(x) = (*)^{N+1}\chi_{[0,1]}. \tag{89}$$

It is then well known that the integer translates of $\phi$ generate $V_0$. For example, in the case of the linear splines ($N = 1$), $\phi$ is the *hat function*; *i.e.*, $\phi^1(x) = \max(0, 1 - |x - 1|)$.

However, it is clear that the translates of $\phi$ are not orthonormal. Still $\{\phi^N(x-k)\}_{k\in\mathbb{Z}}$ forms a Riesz bases for $V_0$; *i.e.*, there exists $C_2 \geq C_1 > 0$, such that for any sequence $\alpha_k$ in $\ell^2(\mathbb{Z})$,

$$C_1 \sum_{k\in\mathbb{Z}} |\alpha_k|^2 \leq \left\| \sum_{k\in\mathbb{Z}} \alpha_k \phi^N(x-k) \right\|_{L^2}^2 \leq C_2 \sum_{k\in\mathbb{Z}} |\alpha_k|^2 \tag{90}$$

which can also be expressed by

$$0 < C_1 \leq \sum_{k\in\mathbb{Z}} |\hat{\phi}^N(\omega + 2k\pi)|^2 \leq C_2. \tag{91}$$

This allows an orthonormalization process that preserves the structure of a family generated by the translates of a single function. The new scaling function is defined by

$$\hat{\phi}_o^N(\omega) = \hat{\phi}^N(\omega) \left( \sum_{k\in\mathbb{Z}} |\hat{\phi}^N(\omega + 2k\pi)|^2 \right)^{-1/2} \tag{92}$$

Its translates are orthonormal by construction but it is not compactly supported (it has exponential decay). The same holds for the associated wavelet. In this construction, due to Battle and Lemarié [13], the CQF filters do not have a finite impulse response, and this fact constitutes a major disadvantage for implementation.

Using the results of the two previous sections, we shall now construct biorthogonal bases of compactly supported wavelets where the functions $\phi^N$ and $\psi^N$ are spline functions of order $N$.

We simply choose for $\phi^N$ the $B$-spline function defined by (89) since we do not require anymore the orthonormality of its integer translates. The low pass filter associated with this scaling function is given by

$$m_0^N(\omega) = \left(\frac{1+e^{-i\omega}}{2}\right)^{N+1}. \tag{93}$$

To find a dual filter, it is useful to define the following polynomial

$$P_L(y) = \sum_{j=0}^{L-1} \binom{L-1+j}{j} y^j, \tag{94}$$

which is the solution of the Bezout equation

$$(1-y)^L P_L(y) + y^L P_L(1-y) = 1. \tag{95}$$

By a change of variable, we obtain

$$\left[\cos\left(\frac{\omega}{2}\right)\right]^{2L} P_L\left(\sin^2\left(\frac{\omega}{2}\right)\right) + \left[\sin\left(\frac{\omega}{2}\right)\right]^{2L} P_L\left(\cos^2\left(\frac{\omega}{2}\right)\right) = 1, \tag{96}$$

and by a shift

$$\left[\frac{1+e^{i\omega}}{2}\right]^{2L}\left[e^{-iL\omega} P_L\left(\sin^2\left(\frac{\omega}{2}\right)\right)\right] + \left[\frac{1-e^{i\omega}}{2}\right]^{2L}\left[(-1)^L e^{-iL\omega} P_L\left(\cos^2\left(\frac{\omega}{2}\right)\right)\right] = 1. \tag{97}$$

This formula gives a solution for all the values of $N$ smaller than $2L-1$. Indeed, if $N+1 \leq 2L$, we can take as a dual filter

$$\tilde{m}_0^{N,L}(\omega) = \left[\frac{1+e^{i\omega}}{2}\right]^{2L-N-1} P_L\left(\sin^2\left(\frac{\omega}{2}\right)\right) e^{-iL\omega}. \tag{98}$$

In other words, for a fixed $N$, $\tilde{m}_0^{N,L}$ is a dual filter for $m_0^N$ if $2L \geq N+1$.

How can we choose the parameter $L$ in an optimal way? At first, it seems natural to choose the smallest value of $L$ such that $2L \geq N+1$. Unfortunately,

this choice does not lead, except in the Haar case ($N = 0$), to a dual scaling function $\tilde{\phi}_{N,L}$ which is square integrable. More precisely the filter $\tilde{m}_0^{N,L}$ does not satisfy in this case the conditions of Theorems 8 and 12.

A very simple example to illustrate this problem is the case of the linear splines $N = 1$. If we choose $L = 1$, then the dual filter is given by

$$\tilde{m}_0^{1,1}(\omega) = e^{-i\omega}. \tag{99}$$

By formula (25), we see that

$$\hat{\tilde{\phi}}^{1,1}(\omega) = e^{-i\omega}, \tag{100}$$

which is not square integrable. Observe that $\phi^{1,1}$ is a delta function centered on 1 and that in the distribution sense we still have

$$\langle \phi^1(x-k) | \tilde{\phi}^{1,1}(x-\ell) \rangle = \delta_{k,\ell}, \tag{101}$$

but the subband coding scheme has no chances to be numerically stable (in the $\ell^2$ sense). For this reason, we need to choose a larger $L$, such that $\tilde{m}_0^{N,L}(\omega)$ satisfies the conditions of Theorem 8 or 12. In [6], we prove that $\tilde{\phi}^{N,L}$ can be made arbitrarily regular if we choose a large enough $L$ for a fixed $N$. This means that the Fourier transform of $\hat{\tilde{\phi}}^{N,L}$ has an arbitrarily high rate of decay at infinity and that $\tilde{m}_0^{N,L}$ will satisfy the conditions of Theorem 8. Note that the regularity of the functions $\tilde{\phi}$ and $\tilde{\psi}$ is not very important in applications since we only use them in the decomposition stage by inner product with the function to be analyzed. As explained in the introduction, smoothness is mostly important in the reconstruction process and we shall rather take $\phi^N$ and $\psi^{N,L}$ as synthesis functions since they are piecewise polynomials. For a given value of $N$, the best choice for $L$ is thus the smallest value such that $\tilde{m}_0^{N,L}$ satisfies the conditions of Theorem 12 (which are sharp).

For $N = 1$, this value is $L = 2$; *i.e.*,

$$\tilde{m}_0^{1,2}(\omega) = \left(\frac{1+e^{i\omega}}{2}\right)^2 e^{-2i\omega}\left(1 + 2\sin^2\left(\frac{\omega}{2}\right)\right). \tag{102}$$

For $N = 2$, this value is also $L = 2$; *i.e.*,

$$\tilde{m}_0^{2,2}(\omega) = \left(\frac{1+e^{i\omega}}{2}\right) e^{-2i\omega}\left(1 + 2\sin^2\left(\frac{\omega}{2}\right)\right). \tag{103}$$

For $N = 3$, this value is $L = 4$; *i.e.*,

$$\tilde{m}_0^{3,3}(\omega) = \left(\frac{1+e^{i\omega}}{2}\right)^4 e^{-4i\omega}\left(1 + 4\sin^2\left(\frac{\omega}{2}\right) + 10\sin^4\left(\frac{\omega}{2}\right) + 20\sin^6\left(\frac{\omega}{2}\right)\right). \tag{104}$$

We illustrate in Figures 3 and 4 the first two cases (linear and quadratic). Note that, for even $N$, the wavelets are shifted odd functions, whereas for odd $N$, they are shifted even functions. We show in these figures the results for the minimal value of $L$ (in the first column) and for the next value (in the second column).

Observe the chaotic aspect of the functions $\tilde{\phi}$ and $\tilde{\psi}$ in the quadratic case. One can check that these functions, although square integrable, do not satisfy the decay condition $(1 + |\omega|)^{-1/2-\epsilon}(|\hat{\tilde{\phi}}| + |\hat{\tilde{\psi}}|) < +\infty$. This means that only Theorem 12 can be applied here. This also explains the fractal aspect of these graphs: To be square integrable, $\tilde{\phi}$ and $\tilde{\psi}$ must have a lacunary structure, typical of these fractal figures.

Finally, let us mention a slightly different construction of compactly supported spline wavelet which is due to C. K. Chui and J. Z. Wang [9]. In this construction, the spaces $W_j$ are kept orthogonal and the wavelet are nonorthogonal only inside a given scale. The advantages of this framework is that the wavelet decomposition is still orthogonal and $\phi$ and $\psi$ are still symmetric or antisymmetric. Furthermore, the discussion on frame bounds is much simpler and that the dual wavelet is also a spline function. However the dual wavelet and filter are no longer compactly supported.

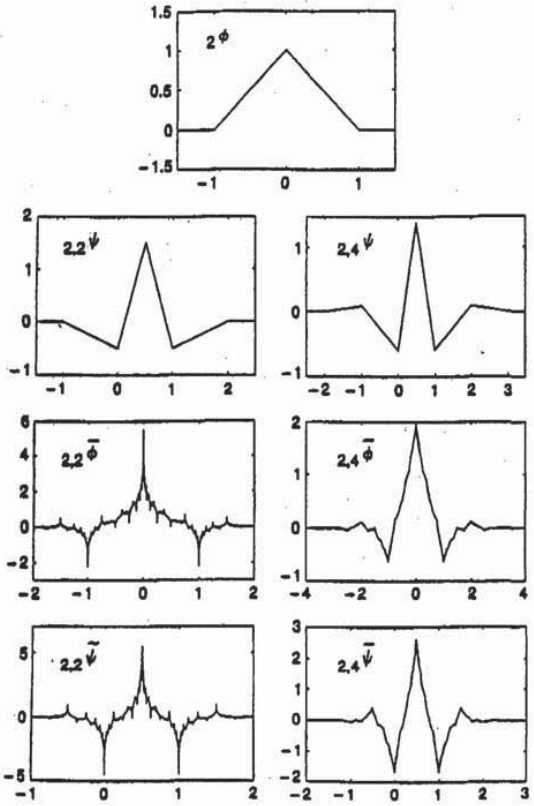


**Figure 3.** Graphs of the functions $\phi, \psi, \bar{\phi}$, and $\tilde{\psi}$ in the spline case.

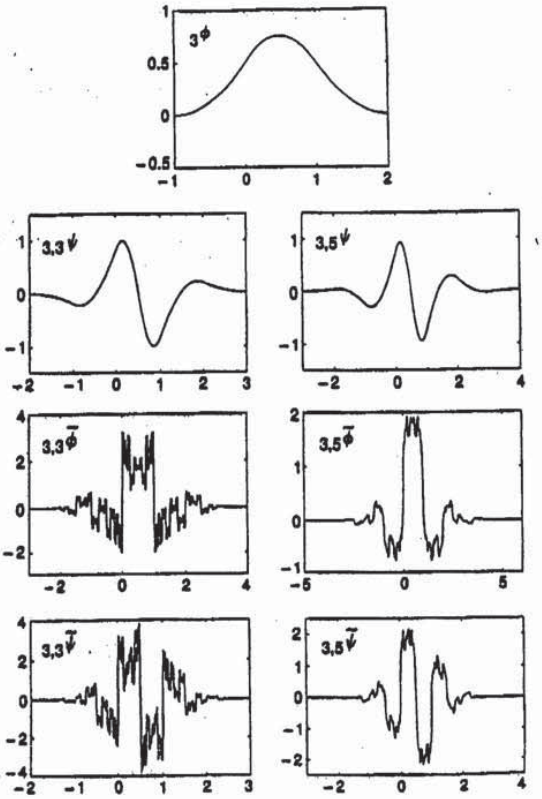


**Figure 4.** Graphs of the functions $\phi, \psi, \tilde{\phi}$, and $\tilde{\psi}$ in the quadratic case.

## Appendix: Proof of some lemmas

**Proof of Lemma 2**

We shall prove this result for $P_j$, the argument being the same for $\tilde{P}_j$, $\Delta_j$, and $\tilde{\Delta}_j$.

First, from the definition (29) of $P_j$, we have

$$P_j(f(y))(x) = P_0(f(2^j y))(2^{-j}x). \tag{105}$$

It is thus sufficient to prove that $P_0$ is a continuity operator on $L^2(\mathbb{R})$. We will then have $\|P_0\| = \|P_j\|$ for all $j$. Since both $\phi$ and $\tilde{\phi}$ are in $L^2$ and compactly supported, we can define the two trigonometric polynomials

$$t(\omega) = \sum_{k\in\mathbb{Z}} |\hat{\phi}(\omega + 2k\pi)|^2 = \sum_{n\in\mathbb{Z}} \langle \phi(x)|\phi(x-n)\rangle e^{in\omega}, \tag{106}$$

and

$$\tilde{t}(\omega) = \sum_{k\in\mathbb{Z}} |\hat{\tilde{\phi}}(\omega + 2k\pi)|^2 = \sum_{n\in\mathbb{Z}} \langle \tilde{\phi}(x)|\tilde{\phi}(x-n)\rangle e^{in\omega}. \tag{107}$$

For any sequence $\{e_n\}$ in $\ell^2(\mathbb{Z})$ we have

$$\begin{aligned}\int_{\mathbb{R}} \left|\sum_{n\in\mathbb{Z}} e_n \phi(x-n)\right|^2 dx &= \frac{1}{2\pi}\int_{\mathbb{R}} \left|\sum_{n\in\mathbb{Z}} e_n e^{-in\omega}\right|^2 |\hat{\phi}(\omega)|^2 d\omega \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\sum_{n\in\mathbb{Z}} e_n e^{-in\omega}\right|^2 \sum_{k\in\mathbb{Z}} |\hat{\phi}(\omega+2k\pi)|^2 d\omega \\ &\le \max(t(\omega)) \sum_{n\in\mathbb{Z}} |e_n|^2.\end{aligned}$$

For all $f$ in $L^2(\mathbb{R})$, we can estimate the $\ell^2$ norm of the sequence $\langle f|\tilde{\phi}(x-n)\rangle$ in the following way.

$$\begin{aligned}\sum_{n\in\mathbb{Z}} |\langle f|\tilde{\phi}(x-n)\rangle|^2 &= \frac{1}{4\pi^2}\sum_{n\in\mathbb{Z}} \left|\int \hat{f}(\omega)\overline{\hat{\tilde{\phi}}(\omega)} e^{in\omega} d\omega\right|^2 \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|\sum_{k\in\mathbb{Z}} \hat{f}(\omega+2k\pi)\overline{\hat{\tilde{\phi}}(\omega+2k\pi)}\right|^2 d\omega \\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\sum_{k\in\mathbb{Z}} |\hat{f}(\omega+2k\pi)|^2\right)\left(\sum_{k\in\mathbb{Z}} |\hat{\tilde{\phi}}(\omega+2k\pi)|^2\right) d\omega \\ &\le \max(\tilde{t}(\omega))\|f\|^2.\end{aligned}$$

It follows from these two estimates that for all $f$ in $L^2(\mathbb{R})$

$$\|P_0 f\|_{L^2}^2 \le \max(\tilde{t}(\omega)) \max(t(\omega)) \|f\|_{L^2}^2, \tag{108}$$

and thus $\|P_0\| \leq [\max(t)\max(\tilde{t})]^{1/2}$. ■

**Proof of Lemma 3**

We shall use the Fourier coefficients $h_n, g_n, \tilde{h}_n, \tilde{g}_n$ of the filters $m_0, m_1, \tilde{m}_0$, and $\tilde{m}_1$. For any $f$ in $L^2(\mathbb{R})$, we have

$$\begin{aligned} P_0(f) + \Delta_0(f) &= \sum_{k\in\mathbb{Z}} \langle f|\tilde{\phi}(x-k)\rangle \phi(x-k) + \sum_{k\in\mathbb{Z}} \langle f|\tilde{\psi}(x-k)\rangle \psi(x-k) \\ &= \sum_{n,m\in\mathbb{Z}} \left( \sum_{k\in\mathbb{Z}} \bar{\bar{h}}_{n-2k} h_{m-2k} + \sum_{k\in\mathbb{Z}} \bar{\bar{g}}_{n-2k} g_{m-2k} \right) \\ &\quad 4\langle f|\tilde{\phi}(2x-h)\rangle \phi(2x-m) \\ &= \sum_{n,m\in\mathbb{Z}} \left( \sum_{k\in\mathbb{Z}} [\bar{\bar{h}}_{n-2k} h_{m-2k} + (-1)^{n+m} \bar{\bar{h}}_{1-m+2k} h_{1-n+2k}] \right) \\ &\quad 4\langle f|\tilde{\phi}(2x-n)\rangle \phi(2x-m) \\ &= \sum_{m,n\in\mathbb{Z}} 4c_{m,n} \langle f|\tilde{\phi}(2x-n)\rangle \phi(2x-m). \end{aligned}$$

When $m - n = 2p$, we have

$$c_{m,n} = \sum_{k\in\mathbb{Z}} \bar{\bar{h}}_k h_{k+m-n} = \sum_{k\in\mathbb{Z}} \bar{\bar{h}}_k h_{k+2p} = \frac{1}{2}\delta_{0,p} \tag{109}$$

because of the duality relation (21). When $m - n = 2p + 1$, then

$$\begin{aligned} c_{n,m} &= \sum_{k\in\mathbb{Z}} \bar{\bar{h}}_{n-2k} h_{m-2k} - \sum_{k\in\mathbb{Z}} \bar{\bar{h}}_{1-m+2k} h_{1-n+2k} \\ &= 0 \end{aligned}$$

because the two sums contain exactly the same terms. We thus have

$$c_{m,n} = \frac{1}{2}\delta_{n,m} \tag{110}$$

and

$$P_0(f) + \Delta_0(f) = 2\sum_{k\in\mathbb{Z}} \langle f|\tilde{\phi}(2x-k)\rangle \phi(2x-k) = P_{-1}(f). \tag{111}$$

By Equation (105) we can rescale this identity to obtain for all $j$

$$P_j + \Delta_j = P_{j-1}; \tag{112}$$

the same holds for $\tilde{P}_j$, $\tilde{\Delta}_j$, and $\tilde{P}_{j-1}$. ■

**Proof of Lemma 4**

We will use the following well known approximation result: If $f$ is in $L^2(\mathbb{R})$, for all $\epsilon > 0$, there exists a function $g$ which is a finite linear combination of intervals characteristic functions $\left(g = \sum_{k=1}^{n} \alpha_k \chi_{I_k}\right)$ such that $\|f-g\| \leq \epsilon$. Using Lemma 2, this also gives

$$\|P_j(f)\| = \|P_j(f-g+g)\| \leq \|P_j(g)\| + C\epsilon, \tag{113}$$

and

$$\|P_j(f) - f\| \leq \|P_j(g) - g\| + (C+1)\epsilon. \tag{114}$$

It is thus sufficient to prove the limits (36) and (37) in the case where $f = \chi_{[a,b]}$.

We first prove Equation (36). Using the same argument as in the proof of Lemma 2, we obtain

$$\begin{aligned}
\|P_j(f)\|^2 &\leq C \sum_{k\in\mathbb{Z}} |\langle f|\tilde{\phi}_k^j\rangle|^2 \\
&= C \sum_{k\in\mathbb{Z}} \left| \int_a^b \overline{\tilde{\phi}_k^j(x)} dx \right|^2 \\
&\leq C \sum_{k\in\mathbb{Z}} \int_a^b 2^{-j} |\tilde{\phi}(2^{-j}x - k)|^2 dx \\
&= C \sum_{k\in\mathbb{Z}} \int_{2^{-j}a-k}^{2^{-j}b-k} |\tilde{\phi}(x)|^2 dx
\end{aligned}$$

when $j$ goes to $+\infty$. It is clear that this expression tends to zero if $\tilde{\phi}$ is in $L^2(\mathbb{R})$.

To prove (37), we shall directly evaluate $P_j(f) = P_j(\chi_{[a,b]})$. We first remark that, because of the hypothesis (23), we have

$$\hat{\phi}(2k\pi) = \hat{\tilde{\phi}}(2k\pi) = \delta_{0,k}. \tag{115}$$

This leads to the following summation formulas which are valid for almost all $x$:

$$\sum_{k\in\mathbb{Z}} \phi(x-k) = \sum_{k\in\mathbb{Z}} \tilde{\phi}(x-k) = 1, \tag{116}$$

and

$$\int \phi(x) = \int \tilde{\phi}(x) = 1. \tag{117}$$

We know that $\phi$ and $\tilde{\phi}$ are compactly supported; *i.e.*, both vanish out of an interval $[-s, s]$. For $j < 0$ and $|j|$ large enough, we can derive a pointwise estimate of

$$P_j(f)(x) = \sum_{k\in\mathbb{Z}} \langle f|\tilde{\phi}_k^j\rangle \phi_k^j(x). \tag{118}$$

(1) If $x$ is out of $[a-2^{j+1}s, b+2^{j+1}s]$, then either $\phi_k^j(x) = 0$ or $\langle f|\tilde{\phi}_k^j\rangle = 0$ and thus $P_j(f)(x) = 0$.
(2) If $x$ is inside $[a+2^{j+1}s, b-2^{j+1}s]$, then either $\phi_k^j(x) = 0$, or $\langle f|\tilde{\phi}_k^j\rangle = \int \overline{\tilde{\phi}_k^j} dx = 2^{j/2}$ and thus

$$P_j(f)(x) = \sum_{k\in\mathbb{Z}} 2^{j/2}\phi_k^j(x) = \sum_{k\in\mathbb{Z}} \phi(2^{-j}x - k) = 1. \tag{119}$$

In these two cases, $P_j(f)(x)$ is exactly $f(x)$. We still have to evaluate $\int_{R_j} |P_j(f) - f|^2$ where $R_j$ is the residual domain $R_j = [a-2^{j+1}s, a+2^{j+1}s] \cup [b-2^{j+1}s, b+2^{j+1}s]$. Now note that when $x$ is in $R_j$, we have

$$f(x) = \chi_{R_{j+1}}(x)f(x) = f_j(x) \quad (R_j \subset R_{j+1}), \tag{120}$$

and

$$P_j(f)(x) = P_j(f_j)(x). \tag{121}$$

Consequently

$$\begin{aligned}\int_{R_j} |P_j(f)(x) - f(x)|^2 dx &= \int_{R_j} |P_j(f_j)(x) - f_j(x)|^2 dx \\ &\le \int_{\mathbb{R}} |P_j(f_j) - f_j|^2 dx \\ &\le C\|f_j\|^2,\end{aligned}$$

and $\|f_j\|^2$ tends to zero when $j$ goes to $-\infty$. This concludes the proof of the lemma. ■

## References

1. Antonini, M., M. Barlaud, P. Mathieu, and I. Daubechies, Image coding using wavelet transform, *IEEE Trans. ASSP*, to appear.
2. Beylkin, G., R. Coifman, and V. Rokhlin, Fast wavelet transform, and numerical algorithms I, *Comm. Pure and Appl. Math.* **44** (1991), 141–183.
3. Chui, C. K. and J. Z. Wang, A general frame work of compactly supported splines and wavelets, CAT Report #219, Texas A&M University, College Station, TX, 1990.
4. Cohen, A. and I. Daubechies, A stability criterion for biorthogonal wavelet bases and their related subband coding schemes, AT&T Bell Laboratories, 1991, preprint.
5. Cohen, A. and I. Daubechies, Nonseparable bidimensional wavelet bases, AT&T Bell Laboratories, 1991, preprint.
6. Cohen, A., I. Daubechies, and J. C. Feauveau, Biorthogonal bases of compactly supported wavelets, *Comm. Pure and Appl. Math.*, 1991, to appear.

7. Cohen, A., Ondelettes, analyses multirésolutions et filtres miroir en quadrature, *Annales de l'IHP, analyse non linéaire* **7** (5) (1990).
8. Coifman, R., Y. Meyer, S. Quacke, and M. V. Wickerhauser, Signal processing and compression with wave packets, *Proceedings of the Conference on Wavelets*, Marseilles, Spring 1989.
9. Conze, J. P. and A. Raugi, Fonctions harmoniques pour un opérateur de transition et applications, Dept. de Math., Université de Rennes, France, 1990, preprint.
10. Daubechies, I., Orthonormal bases of compactly supported wavelets, *Comm. Pure and Appl. Math.* **41** (1988), 909–996.
11. Eirola, T., Sobolev characterization of solutions of dilation equations, Helsinki University of Technology, submitted to *SIAM J. Math. Anal.*, 1991.
12. Lawton, W., Necessary and sufficient conditions for constructing orthonormal wavelets bases, *J. Math. Phys.* **32** (1) (1991), 57–61.
13. Lemarié, P. G., Ondelettes à localisation exponentialle, *J. Math. Pure et Appl.* **67** (1988), 227–236.
14. Mallat, S., A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Pattern Anal. and Machine Intell.* **11** (7) (1989), 674–693.
15. Meyer, Y., *Ondelettes et Opérateurs*, in two volumes, Hermann, Paris, 1990.
16. Smith, M. J. T. and T. P. Barnwell, Exact reconstruction techniques for tree structured subband coders, *IEEE ASSP* **34** (1986), 434–441.
17. Vetterili, M., Filter banks allowing perfect reconstruction, *Signal Processing* **10** (1986), 219–244.
18. Villemoes, L., Energy moments in time and frequency for two scale difference equation solutions and wavelets, Math. Institute, Technical Univ. of Denmark, DK2800 Lyngby, Denmark, 1991, preprint.
19. Young, R. M., *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.

*Albert Cohen*
CEREMADE
Université Paris IX Dauphine
Place du Maréchal de Lattre de Tassigny
75016 Paris, France

within a neighborhood around the reference function $G_i$. The reduced order warping matrix for the reference function $G_i$ is given by

$$W_i = R_i^{-1}. \tag{12}$$

The warping vector for the expansion coefficient $a_i$ is given by

$$\bar{w}_i = \bar{r}_i W_i \tag{13}$$

where $\bar{r}_i$ is an operator that selects the row associated with $a_i$ from the matrix $W_i$. The update equation for the coefficient $a_i$ becomes

$$a_i^{l+1} = a_i^l - \beta \bar{w}_i \nabla J_i \tag{14}$$

where $\nabla J_i$ is the gradient for the elementary functions included in $R_i$.

The update term in (14) can be expressed as

$$\bar{w}_i \nabla J_i = -\iint e(x, y) \left[ \sum_j w_i(j) G_j(x, y) \right] dx\, dy \tag{15}$$

where $w_i(j)$ is the $j$th element of the warping vector $\bar{w}_i$. The new function created by the warping vector is given by

$$P_i(x, y) = \sum_j w_i(j) G_j(x, y). \tag{16}$$

This function $P_i$ is orthonormal to all elementary functions included in $R_i$, and is referred to as the "locally biorthonormal projection function." The remainder of this section outlines a method for calculating the optimal truncated projection function for a given support width.

Let $Q_p$ denote a pseudo-Hessian matrix whose elements are given by

$$q_p(i, j) = \langle G_i, P_j \rangle_{xy}. \tag{17}$$

If the projection and elementary functions are normalized such that $q_p(i, i) = 1$, then the normalized pseudo-Hessian matrix $\hat{Q}_p$ can be written as

$$\hat{Q}_p = I + V \tag{18}$$

where $I$ is the identity matrix and $V$ is a matrix containing all the off-diagonal elements of $\hat{Q}_p$. If all the elements of $V$ are small, then

$$\hat{Q}_p^{-1} \approx I - V = 2I - \hat{Q}_p. \tag{19}$$

Using (19), an estimate of the locally biorthonormal projection function is recursively defined by

$$P_i^{l+1}(x, y) = 2P_i^l(x, y) - \sum_{j=1}^{N} \hat{q}_p(i, j) P_j^l(x, y). \tag{20}$$

The recursive estimate of $P_i$ will converge if the Gabor expansion is unique (if $Q^{-1}$ exists). The resulting set of projection functions can be used in a descent implementation if $G_i$ in (7) is replaced by $P_i$. In such an implementation, the lower bound of the optimal convergence factor is estimated by

$$\beta_{LB}^*(i) \approx \frac{1}{\sum_j |q_p(i, j)|}. \tag{21}$$

## IV. Comments

The previous section has introduced two methods that improve convergence: the first method uses warping vectors; the second method uses locally biorthonormal projection functions. The overlap characteristics of the members of $R_i$ determine the warping vector $w_i$ and the window of the projection function $P_i$. In general, a different warping vector or projection function is required for each expansion coefficient $a_i$. These methods are well suited to a massively parallel implementation which assigns a separate processor memory group to each elementary function $G_i$.

For the single processor/memory case, these methods are most efficient for lattices that exhibit overlap invariance: the overlap $q(i, j)$ depends on the distance between elementary functions $G_i$ and $G_j$, not on their absolute lattice positions. Consider the two lattices presented in [2]: the Cartesian lattice and the log-polar lattice. The overlap characteristics of the Cartesian lattice are completely invariant to shifts in position or frequency. Bastiaans used this invariance to calculate a biorthonormal window [3], [4], which provides a direct solution (no iterations) to the Gabor expansion. If the Cartesian lattice is implemented as an iterative Gabor expansion, the increase in complexity over Daugman's neural network is small because all warping vectors $w_i$ will be identical. The improvement in convergence increases with the size of the neighborhood $R_i$, approaching Bastiaans' direct solution as the neighborhood size become arbitrarily large.

The log-polar lattice is not invariant to shifts in frequency. Consequently, there is no equivalent to Bastiaan's biorthonormal window. In an iterative implementation, it is possible to improve the convergence by exploiting the position invariant overlap characteristic that exists between elementary functions with a common frequency.[1] If the neighborhood $R_i$ contains only elementary functions with the same frequency, all warping vectors $w_i$ will be identical. As a result, a limited (but noticeable) increase in convergence is obtained with little additional complexity. The rate of convergence can be further increased if the number of unique warping vectors (and the complexity of the implementation) is allowed to increase.

## References

[1] R. N. Braithwaite, "The use of the Gabor expansion in computer vision systems," Master thesis, Dep. Elec. Eng., Univ. of British Columbia, Vancouver, Canada, 1989.
[2] J. G. Daugman, "Complete discrete 2-D transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 7, pp. 1169–1179, 1988.
[3] M. J. Bastiaans, "Gabor's expansion of a signal into Gaussian elementary signals," *Proc. IEEE*, vol. 68, pp. 538–539, 1980.
[4] M. Porat and Y. Zeevi, "The generalized Gabor scheme of image representation in biological and machine vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 452–467, 1988.
[5] D. G. Luenberger, *Linear and Nonlinear Programming* Reading, MA: Addison-Wesley, 1984.

# Image Compression Using the 2-D Wavelet Transform

A. S. Lewis and G. Knowles

***Abstract*—The 2-D orthogonal wavelet transform decomposes images into both spatial and spectrally local coefficients. The transformed coef-**

Manuscript received November 15, 1990; revised August 5, 1991.

The authors were with the Department of Electrical Engineering, Imperial College, London, U.K. They are now with the Department of Information, U.I.B., 07071 Palma de Mallorca, Spain.

IEEE Log Number 9106077.

[1]Common frequency should be interpreted as elemeptary functions whose modulating waveform has the same radial frequency and the same orientation.

1057-7149/92$03.00 © 1992 IEEE

ficients were coded hierarchically and individually quantized in accordance with the local estimated noise sensitivity of the human visual system (HVS). The algorithm can be mapped easily onto VLSI. For the Miss America and Lena monochrome images, the technique gave high to acceptable quality reconstruction at compression ratios of 0.3–0.2 and 0.64–0.43 bits per pixel (bpp), respectively.

## I. Introduction

Traditional image compression techniques have been designed to exploit the statistical redundancy present within real world images. The discrete cosine transform (DCT), DPCM, and the entropy coding of subband images are all examples of this statistical approach. Removing redundancy can only give a limited amount of compresssion; to achieve high ratios, some of the nonredundant information must be removed. The statistical coders produce annoying visual degradation when operating in this mode because they introduce errors in visually important parts of the image structure (such as feature edges). By using methods of image decomposition that closely mimic the human visual system (HVS), compression can take into account the importance of each individual coefficient and code accordingly.

The HVS is an information processing system, receiving spatially sampled images from the cones and rods in the eye and deducing the nature of the objects it observes by processing this image data. At a low level, objects can be thought of as structures made up of surfaces of the same color or texture bounded by edges. The color and texture of an object is greatly affected by its orientation and illumination so the edges are usually the most important means of recognition, so a good image compression algorithm should try to minimize edge distortion.

Psychophysicists and visual psychologists have performed many experiments on the HVS to determine how it processes image data. They discovered that the eye filters the image into a number of bands, each approximately one octave wide in frequency. Further, in the spatial domain, the image should be considered to be composed of information at a number of different scales [6]. Marr formulated a *constraint of spatial localization* that *the physical phenomena that give rise to intensity changes in the image are spatially localized*.

In this paper we introduce a new approach to image compression based on decomposing the image using the orthogonal wavelet transform, and then apply a compression algorithm based on Marr's constraint of spatial locality. Firstly, we will introduce the wavelet transform, discuss its implementation by quadrature mirror filtering (QMF), and describe the ideas behind our compression method. In Section III we present an algorithm implementing our method which is suitable for real time use either on a DSP chip, or on dedicated hardware. In Section IV we give some results on the test images of Miss America and Lena for a range of compression ratios.

## II. Wavelet Transform and Image Compression

One-dimensional wavelet theory defines a function $\psi$, the wavelet, and its associated scaling function $\phi$, such that the family of functions $\{\psi^j(x)\}_{j\in Z}$, where $\psi^j(x) = \sqrt{2}^j\psi(2^j x)$, are orthonormal. The wavelet transform can be implemented by quadrature mirror filters [11], [12] $G = (g(n))$, and $H = (h(n))$, $n \in Z$, where $h(n) = 1/2\langle\phi(x/2), \phi(x-n)\rangle$, and $g(n) = (-1)^n h(1-n)$ ($\langle\ \rangle$ denotes $L^2$ inner product). $H$ corresponds to a low-pass filter, and $G$ is an octave wide high-pass filter. The reconstruction filters have impulse response $h^*(n) = h(1-n)$, and $g^*(n) = g(1-n)$. For a more detailed analysis of the relationship between wavelets and QMF see Mallat [7].

For images we use the hierarchical wavelet decomposition suggested by Mallat [7]. The $G$ and $H$ filters are applied to the image in both the horizontal and vertical directions, and the filter outputs subsampled by a factor of two, generating three orientation selective high-pass subbands, *GG*, *GH*, *HG*, and a low-pass subband *HH*. The process is then repeated on the *HH* band to generate the next level of the decomposition, etc. Four octaves of decomposition leads to thirteen subbands (Fig. 1). Fig. 2 shows the image of Lena decomposed in this way. This hierarchical approach to image decomposition fulfills the role of scales in Marr's decomposition. QMF subband coding algorithms for images have been explored by Vetterli [8], and for use in compression by Woods and O'Neil [9] using a flat decomposition and by Gharavi and Tabatabai [10] in a pyramidal decomposition akin to the wavelet transform.

In general the wavelet transform requires much less hardware to implement than Fourier methods, such as the discrete cosine transform DCT [5]. However, for the 4 coefficient Daubechies [1] wavelet, it is especially simple to calculate and invert the transform in hardware as no multipliers are needed to calculate the quantized coefficients ($h(0) = 11/32$, $h(1) = 19/32$, $h(2) = 5/32$, $h(3) = -3/32$) [4]. This allows us to incorporate the advantages of a multiresolution approach to decompose and reconstruct the image, without high hardware costs.

The wavelet decomposition is an alternative representation of image data but the number of bits used to store it has not changed. To compress the image data, we must decide which coefficients to send and how many bits to use to code them. Our compression algorithm consists of taking the low-pass subband in full, and then deciding which coefficients within the remaining subbands to keep. The problem is to decide which of the nonzero wavelet transform coefficients correspond to noise and which visually important details of the image, in particular, we want to preserve the edge-like information within the image. Simply thresholding each of the coefficients, would leave extra noise and visually unimportant information in the compressed image. We apply an algorithm based on Marr's constraint of spatial locality when deciding which coefficients to keep. So that if an important detail occurs at some place in the image we expect that the coefficients corresponding to that location will exceed a threshold in more than one octave; the orientation of the detail will be determined by which of the *GH*, *HG*, or *GG* subbands we use. In this way we utilize both the frequency and spatial locality of the wavelet transform to detect and code the image data efficiently. By examining the image at low resolutions, and hence, a low number of samples, an initial guess at the location of the edges can be made. These guesses can be confirmed or rejected by examining the higher resolution detail signal at the same spatial location. We shall use the same technique to not only save on processing time but also to compress the picture through rejecting the redundant areas; by controlling which areas are rejected we can aim for high compression and high quality in the final output. The threshold values are determined adaptively from a simple HVS model. Finally, the smaller the support of the wavelet, the less nonzero wavelet coefficients will correspond to an edge, so the more efficient will be our compression scheme. For this reason we chose the Daubechies-4 wavelet. The second stage of the compression process is to quantize the remaining nonzero coefficients; we use a linear mid-step quantizer with step-size derived from a HVS model, and then Huffman code the resulting values.

## III. The Compression Algorithm

In this section we describe the algorithm implementing the ideas discussed in the previous section, and combine this with a quantization model based on the HVS.

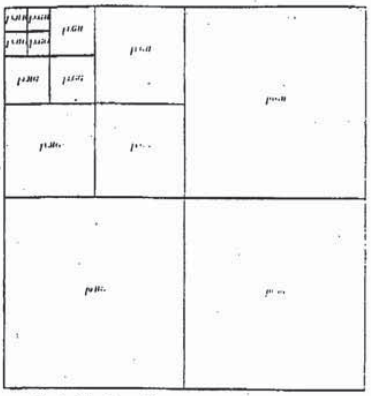

Fig. 1. Wavelet multiresolution image decomposition.

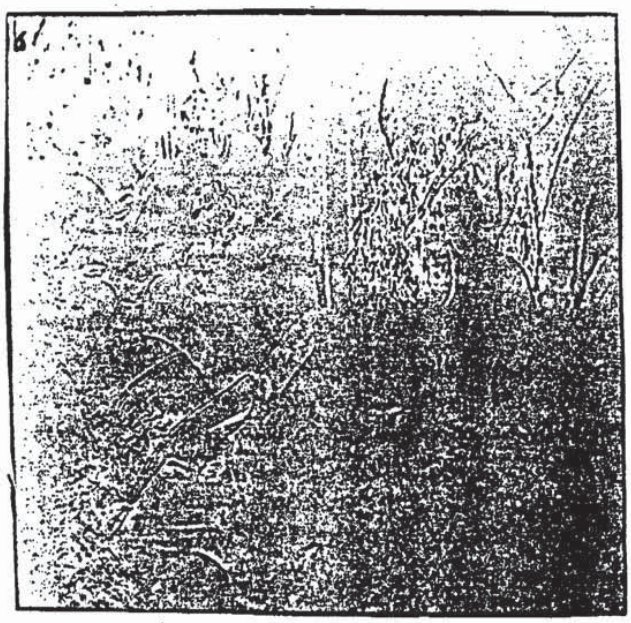

Fig. 2. Wavelet decomposition of Lena image (high-pass coefficients ×4).

Each coefficient in the high-pass bands of the wavelet transform for octaves ($r = 1 \cdots 3$) has four coefficients corresponding to its spatial position in the octave band above it in frequency. This suggests using a four branch tree to represent the structure of the decomposition. This data structure was used successfully in a previous video compression algorithm [2]. The tree is constructed by first considering each coefficient of all orientations ($x = GG, GH, HG$) of the lowest octave in the high-pass bands. The value is coded

and if its absolute value is greater than a fixed threshold value, the process is repeated for the four corresponding coefficients in the octave above. If not, the tree is rejected and zero coefficients are implied by the coding system for the tree above. This technique requires no control signals or addresses to be added by the coder and is efficiently implemented using stack-based hardware.

Our thresholding algorithm was far from optimal; though it followed edges and rejected flat areas very efficiently, mistakes happened too often and blurred parts of the image around edges.

A better measure of the likelihood of an edge at a point within a subband would be to calculate the energy in a small area encompassing the point. By using a small 2 by 2 block as the nodes of the tree structure, rather than single coefficient values, thresholding errors are greatly reduced. In addition, special block codes are introduced to increase the coding efficiency above that of the original method. Fig. 3 shows the new small block-based tree structure and its coding with the aid of the following recursive function, SendTree. Firstly, the low-pass image *HH* is coded in full(8 bits per pixel (bpp)), then the remaining subbands are coded recursively by applying the following subroutine SendTree to all the blocks within the lowest frequency subbands. (In the following, let $I^{r,s}(x, y)$ be the coefficient value of the decomposition $I$ at octave $r$, orientation $s$, and position $x$, $y$ within that subband.)

```
SendTree (I, r, s, x, y)
  If (THRESHOLD < ThresholdFunction(I^{r,s}(x + {0, 1}, y +
  {0, 1}))) {
    SendToken(BlockNotEmpty);
    SendCoefficients(I^{r,s}(x + {0, 1}, y + {0, 1}));
    SendTree(I, r − 1, s, 2x, 2y);
    SendTree(I, r − 1, s, 2(x + 1), 2y);
    SendTree(I, r − 1, s, 2x, 2(y + 1));
    SendTree(I, r − 1, s, 2(x + 1), 2(y + 1));
  } else SendToken(BlockEmpty);
```

The ThresholdFunction was chosen to be the sum of squares of the block's coefficients normalized to the HVS perceptual threshold discussed later.

Tree structuring and block coding have exploited interband and intraband correlation to facilitate data compression but so far we have not considered how to code the coefficient data itself (SendCoefficients). With 8 bpp we have 256 different levels to represent our transformed coefficients, but this may not be the optimum value. We can use our knowledge of the HVS to determine the smallest number of levels required at each point within the transformed image. The errors caused by using fewer levels can be considered as a noise source; visual psychophysics states that a number of factors effect the noise sensitivity of the eye: the background luminance, the proximity to an edge, the frequency band, and texture masking.

*Background luminance* is related to noise sensitivity by Weber's law. The eye is less sensitive to noise for brighter background luminances. The low-pass coefficients can be used to provide values of background luminance.

*Edge proximity* or spatial masking relates noise sensitivity to distance from and height of the edge. As the coefficients we send are supposed to be part of an edge, the spatial locality of the octave provides distance information, while the energy of the lower frequency coefficients indicates edge height. Sensitivity decreases for increasing edge height and at decreasing distances from the edge.

*Band sensitivity* is of the octave currently being coded. This is a fixed value for each octave, orientation, and luminance-chrominance channel. We have formulated an empirical model for this based on HVS experiments.

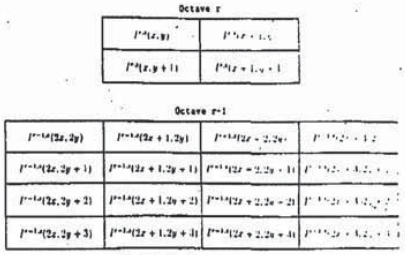


Fig. 3. Block tree structure.

*Texture masking* decreases the sensitivity to noise if there is high activity in the locality of the coefficient. The energy of lower frequency coefficients can indicate the texture activity level.

A mathematical model of the HVS can be constructed to allow the estimation of noise sensitivity for any part of the transformed image. Subband coding algorithms that exploit HVS properties usually quantize the whole subband on the basis of spectral response alone. Safranek and Johnston [13] combined band sensitivity, background luminance, and texture masking information to provide a perceptual threshold for each subband coefficient. They then used the minimum threshold value to quantize the entire band. In our algorithm we use a similar calculation to estimate the perceptual threshold of each 2 by 2 block, and then quantize each pixel in the block with this threshold using a linear mid-step quantizer. We also use this perceptual threshold figure to normalize the coefficient value before edge detection. It should be noted that only previously coded values are needed in the quantizer, hence, no additional side information is required.

The quantizer step-size for the coefficients in Fig. 3, *qstep*, is calculated as follows:

$$\text{qstep}(r, s, x, y) = q_0 * \text{frequency}(r, s) * \text{luminance}(r, x, y) * \text{texture}(r, x, y)^{0.034}$$

where $q_0$ is a normalization constant and

$$\text{frequency}(r, s) = \begin{cases} \sqrt{2}, & \text{if } s = GG \\ 1, & \text{otherwise} \end{cases} * \begin{cases} 1.00, & \text{if } r = 0 \\ 0.32, & \text{if } r = 1 \\ 0.16, & \text{if } r = 2 \\ 0.10, & \text{if } r = 3 \end{cases}$$

$$\text{luminance}(r, x, y) = 3 + \frac{1}{256} \sum_{i=0}^{1} \sum_{j=0}^{1} I^{3,HH}(i + 1 + x/2^{3-r}, j + 1 + y/2^{3-r})$$

$$\text{texture}(r, x, y) = \sum_{k=1}^{3-r} 16^{-k} \sum_{s}^{GG,GH,HG} \sum_{i=0}^{1} \sum_{j=0}^{1} (I^{k+r,s}(i + x/2^k, j + y/2^k))^2 + 16^{3-r} \operatorname{var}(I^{3,HH}(\{1, 2\} + x/2^{3-r}, \{1, 2\} + y/2^{3-r}))$$

where var is the variance of the four coefficient block, and the summation is zero when its lower limit exceeds its upper limit.

TABLE I
VARIABLE LENGTH CODES

| Quantized Level | Code Bits |
|---|---|
| . | . |
| . | . |
| . | . |
| −5 | 111001 |
| −4 | 11101 |
| −3 | 1111 |
| −2 | 101 |
| −1 | 011 |
| 0 | 00 |
| 1 | 010 |
| 2 | 100 |
| 3 | 1101 |
| 4 | 11001 |
| 5 | 110001 |
| . | . |
| . | . |
| . | . |

Fig. 4. Coding results for frame #1 of the Miss America test sequence. The original image is in the upper left. The upper right is coded at 0.30 bpp (SNR 37.09), the lower left at 0.25 bpp (SNR 36.56), and the lower right at 0.20 bpp (SNR 36.11).

Fig. 5. Coding results for Lena. The original image is in the upper left. The upper right is coded at 0.64 bpp (SNR 34.76), the lower left at 0.53 bpp (SNR 34.03), and the lower right at 0.43 bpp (SNR 33.18).

Our equations represent a crude approximation to the HVS noise sensitivity; more accurate models could be made to increase compression or picture quality.

After the number of quantization levels has been established and the coefficient has been quantized, further compression can be achieved by using the statistical properties of quantized coefficients. A well-known statistical property of subband coefficients is their Laplacian probability distribution function with a mode value of zero. Even after quantization this property is approximately true. A simple variable length or Huffman code can be constructed on the basis of this assumed property. Further compression can be achieved from the use of the variable length code as it exploits the entropy within the coefficient values. Table I shows some example code strings for a range of quantization levels.

The first two bits indicate the magnitude of the coefficient (00 – 0, 01 – 1, 10 – 2, and 11 – remaining). The third bit indicates the sign (0 – positive and 1 – negative). Trailing zeros terminated by a one indicate the magnitudes of the remaining values above two.

Our compression algorithm uses four "orthogonal" techniques to exploit redundancy within image data. Additional techniques could be used, such as DPCM, on the low-pass band or the entropy of a block's nonzero symbols, but we currently do not consider them to be cost effective.

## IV. Results

We have applied our method to two standard monochrome test images, Miss America and Lena; the coded reconstructions are shown in Figs. 4 and 5, respectively. The Miss America image is the first frame of a test sequence and is 352 by 288 pixels (CIF format) in size. The Lena image is 512 by 512 pixels in size. The SNR values are calculated over the whole of the image (including the edges).

Our results achieve better compression ratios for a given picture quality than the standard subband coding methods. Additionally, our algorithm is significantly simpler to implement in hardware to achieve real-time performance. This is important for the coding of video sequences [3]. We find that the errors introduced by our methods are less visually annoying than for DCT compressed images due to the lack of blocking effects. At very high compression

ration visual degradation is introduced, mainly as blotchiness in flat areas and slight fuzziness around sharp discontinuities. In particular, the background of the Miss America and Lena images, and Lena's face and shoulder become blotchy. Also their hair and Lena's feather lose their sharpness. Finally, our method is equally applicable to the compression of color images in YUV format.

In conclusion, our results indicate that the combination of HVS compatible filters with finite support, and a quantizer which introduces noise in the visually least important and noise insensitive parts of the image gives a significant improvement in compression/image quality over block-based transform methods. Finally, the 4-tap Daubechies filter and the coder we use are much simpler to implement in hardware than DCT, VQ or other subband coding methods [4]. We are currently incorporating this method in a video codec, and implementing it in VLSI([3]).

## References


[1] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *IEEE Trans. Inform. Theory*, vol. 36, pp. 961–1005, Sept. 1990.

[2] A. S. Lewis and G. Knowles, "Video compression using 3D wavelet transforms," *Electron. Lett.*, vol. 26, no. 6, pp. 396–397, 1990.

[3] ——, "A 64 Kb/s video codec using the 2-D wavelet transform," presented at IEEE Data Compression Conf., Snowbird, UT, 1991.

[4] ——, "A VLSI architecture for the 2D Daubechies wavelet transform without multipliers," *Electron. Lett.*, vol. 27, no. 2, pp. 171–173, 1991.

[5] G. Knowles, "VLSI architecture for the discrete wavelet transform," *Electron. Lett.*, vol. 26, no. 15, pp. 1184–1185, 1990.

[6] D. Marr, *Vision*. New York: Freeman, 1982.

[7] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, pp. 674–693, 1989.

[8] M. Vetterli, "Multi-dimensional sub-band coding: Some theory and algorithms," *Signal Proc.*, vol. 6, pp. 97–112, 1984.

[9] J. W. Woods and S. D. O'Neil, "Subband coding of images," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 34, pp. 1278–1288, Oct. 1986.

[10] H. Gharavi and A. Tabatabai, "Sub-band coding of monochrome and color images," *IEEE Trans. Circuit Syst.*, vol. 35, pp. 207–214, Feb. 1988.

[11] A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation, decimation, tree decomposition techniques," presented at Int. Conf. on Information Sciences/Systems, Patras, Greece 1976.

[12] M. J. T. Smith and T. P. Barnwell III, "Exact reconstruction techniques for tree-structured sub-band coders," *IEEE Trans. Acoust. Speech, Signal Processing*, pp. 434–441, June 1986.

[13] R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression," in *Proc. IEEE ASSP'89*, vol. 3, 1989, pp. 1945–1948.


# An Edge Preserving Differential Image Coding Scheme

Martin C. Rost and Khalid Sayood

***Abstract*—Differential encoding techniques are fast and easy to implement. However, a major problem with the use of differential encoding for images is the rapid edge degradation encountered when us such systems. This makes differential encoding techniques of limi utility especially when coding medical or scientific images, where e preservation is of utmost importance. We present a simple, easy implement differential image coding system with excellent edge pr ervation properties. The coding system can be used over variable r channels which makes it especially attractive for use in the packet n work environment.**

Manuscript received January 18, 1990; revised April 28, 1991. This work was supported by the NASA Goddard Space Flight Center under Grant NAG-5-916.

M. C. Rost is with Sandia National Laboratories, Albuquerque, NM 87185.

K. Sayood is with the Department of Electrical Engineering and the Center for Communication and Information Science, University of Nebraska, Lincoln, NE 68588-0511.

IEEE Log Number 9106076.

## I. Introduction

The transmission and storage of digital images requires an en mous expenditure of resources, necessitating the use of compr sion techniques. These techniques include relatively low compl ity predictive techniques such as adaptive differential pulse co modulation (ADPCM) and its variations, as well as relatively higl complexity techniques such as transform coding and vector qu tization [1], [2]. Most compression schemes were originally d veloped for speech and their application to images is at times pro lematic. This is especially true of the low complexity predicti techniques. A good example of this is the highly popular ADPC scheme. Originally designed for speech [3], it has been used wi other sources with varying degrees of success. A major proble with its use in image coding is the rapid degradation in quali whenever an edge is encountered. Edges are perceptually very i portant, and therefore, their degradation can be perceptually ve annoying. If the images under consideration contain medical or sc entific data, the problem becomes even more important, as edg provide position information which may be crucial to the viewe This poor edge reconstruction quality has been a major factor preventing ADPCM from becoming as popular for image codi as it is for speech coding. While good edge reconstruction cap bility is an important requirement for image coding schemes, a other requirement that is gaining in importance with the prolifer tion of packet switched networks is the ability to encode the imag at different rates. In a packet switched network, the available cha nel capacity is not a fixed quantity, but rather fluctuates as a fun tion of the load on the network. The compression scheme mus therefore, be capable of taking advantage of increased capaci when it becomes available while providing graceful degradatio when the rate decreases to match decreased available capacity.

In this paper we describe a DPCM-based coding scheme whic has the desired properties listed above. It is a low complexit scheme with excellent edge preservation in the reconstructed im age. It takes full advantage of the available channel capacity pr viding lossless compression when sufficient capacity is availabl and very graceful degradation when a reduction in rate is require

## II. Notation and Problem Formulation

The DPCM system consists of two main blocks, the quantiz and the predictor (see Fig. 1). The predictor uses the correlatio between samples of the waveform $s(k)$ to predict the next sampl value. This predicted value is removed from the waveform at th transmitter and reintroduced at the receiver. The prediction err is quantized to one of a finite number of values which is coded an transmitted to the receiver and is denoted by $e_q(k)$. The differenc between the prediction error and the quantized prediction error called the quantization error or the quantization noise. If the chan nel is error free, the reconstruction error at the receiver is simpl the quantization error. To see this, note (Fig. 1) that the predictio error $e(k)$ is given by

$$e(k) = s(k) - p(k) \qquad (1$$

1057-7149/92$03.00 © 1992 IEEE

# HIGH–RESOLUTION STILL PICTURE COMPRESSION

MLADEN VICTOR WICKERHAUSER

Department of Mathematics, Washington University
Campus Box 1146, St. Louis, Missouri 63130

April 19, 1992

## 1: INTRODUCTION

We shall consider the problem of storing, transmitting, and manipulating digital electronic images. Because of the file sizes involved, transmitting images will always consume large amounts of bandwidth, and storing images will always require hefty resources. Because of the large number $N$ of pixels in a high resolution image, manipulation of digital images is infeasible without low-complexity algorithms, i.e., $O(N)$ or $O(N\log(N))$. Our goal will be to describe some new methods which are firmly grounded in harmonic analysis and the mathematical theory of function spaces, which promise to combine effective image compression with low-complexity image processing. We shall take a broad perspective, but we shall also compare specific new algorithms to the state of the art.

Roughly speaking, most image compression algorithms split into three parts: invertible transformation, lossy quantization or rank-reduction, and entropy coding (or redundancy removal). There are a few algorithms which differ fundamentally from this scheme, e.g., the collage coding algorithm [Barnsley,Sloan], or pure vector quantization of the pixels. The former uses a deep observation that pictures of natural objects exhibit self-similarity at different scales; we prefer to avoid relying on this phenomenon, since our images may not be "natural." The latter uses a complex algorithm to build a super efficient empirical vocabulary to describe an ensemble of images; we prefer to avoid training our algorithm with any sample of images, to avoid the problem of producing a sufficiently large and suitable ensemble.

There has emerged an international standard for picture compression, promulgated by the Joint Photographic Experts Group (JPEG), which is remarkably effective in reducing the size of digitized image files. JPEG is 2-dimensional discrete cosine transform (DCT) coding of $8\times 8$ blocks of pixels, followed by a possibly proprietary quantization scheme on the DCT amplitudes, followed by either Huffman, Lempel–Ziv–Welch or arithmetic coding of the quantized coefficients. It has some drawbacks; for example, several incompatible implementations are allowed under the standard. Also, JPEG degrades ungracefully a high ultrahigh compression ratios, and it makes certain assumptions about the picture that are violated by zooming in or out, or other transformations. It works so well on typical photographs and many other images, however, that it has become the algorithm to beat in most applications. JPEG fails most noticeably on high resolution (i.e., oversampled) data, and on images which must be closely examined by humans or machines.

Alternatives to JPEG have recently appeared, and we shall discuss 3 of these: the fast discrete wavelet transform, the local trigonometric or lapped orthogonal transform, and the best-basis algorithm. These differ in the transform coding step, i.e., instead of DCT they first apply the wavelet transform, lapped orthogonal transform, or wavelet packet transform, possibly followed by a best-basis search. The resulting stream of amplitudes is then quantized and coded to remove redundancy.

Research supported in part by ONR Grant N00014-88-K0020 and by FBI contract A107183 .

Existing image processing algorithms work on the original pixels or else on the (2-dimensional) Fourier transform of the pixels. If the image has been compressed, it must be uncompressed prior to such processing. Alternatively, we can try to devise algorithms which transform the compressed parameters. If compression is accomplished by retaining only a low-rank approximation to the signal, then we can use more complex algorithms for subsequent processing. To put this idea into practice, we need to retain useful analytic properties such as the large derivatives used in edge detection. These will not be preserved by purely information-theoretic coding such as pure vector quantization, but we can choose transform coding methods whose mathematical properties combine efficient compression with good analytic behavior.

## 2: Transform Coding Image Compression

A digitally sampled image can only represent a band-limited function, since there is no way of resolving spatial frequencies higher than half the pixel pitch. Band limited functions are smooth; in fact they are entire analytic, which means that at each point they can be differentiated arbitrarily often and the resulting Taylor series converges arbitrarily far away. Since digitally sampled images faithfully reproduce the originals as far as our eyes can tell, we may confidently assume that our images are in fact smooth and well approximated by band-limited functions. Another way of saying this is that adjacent pixels are highly correlated, or that there is a much lower rank description of the image which captures virtually all of the independent features. In transform coding, we seek a basis of these features, in which the the coordinates are less highly correlated or even uncorrelated. These coordinates are then approximated to some precision, and that approximate representation is further passed through a lossless redundancy remover.

The figure below depicts a generic image compression transform coder. It embodies a three-step algorithm:

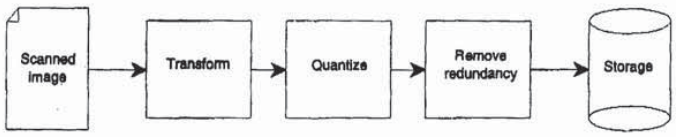


**Figure 2-1.**

**Idealized transform coder.**

The first block ("Transform") applies an invertible coordinate transformation to the image. We think of this transformation as implemented in real arithmetic, with enough precision to keep the truncation error below the quantization error introduced by the original sampling. The output of this block will be treated as a stream of real numbers, though in practice we are always limited to a fixed precision.

The second ("Quantize") block replaces the real number coordinates with lower-precision approximations which can be coded in a (small) finite number of digits. If the transform step is effective, then the new coordinates are mostly very small and can be set to zero, while only a few coordinates are large enough to survive. The output of this block is a stream of small integers, most of which are the same (namely 0). If our goal is to reduce the rank of the representation, we can now stop and take only the surviving amplitudes and tag them with some identifiers. If our goal is to reduce the number of bits we must transmit or store, then we should proceed to the next step.

The third block ("Remove redundancy") replaces the stream of small integers with some more efficient alphabet of variable-length characters. In this alphabet the frequently occurring letters (like "0") are represented more compactly than rare letters.

## 3: Decorrelation by transformation

We will consider six pixel transformations which have proven useful in decorrelating smooth pictures.

### 3.1: Karhunen–Loève.

Let us now fix an image size—say height $H$ and width $W$, with $N = H \times W$ pixels—and treat the individual pixels as random variables. Our probability space will consist of some collection of pictures $\mathcal{S} = \{S_1, S_2, \ldots, S_M\}$, where $M$ is a big number. The intensity of the $n$th pixel $S(n)$, $1 \le n \le N$, is a random variable that takes a nonnegative real value for each individual picture $S \in \mathcal{S}$. Nearby pixels in a smooth image are correlated, which means that the value of one pixel conveys information about the likelihood of its neighbors' values. This implies that having transmitted the one pixel value at full expense, we should be able to exploit this correlation to reduce the cost of transmitting the neighboring pixel values. This is done by transforming the picture into a new set of coordinates which are uncorrelated over the collection $\mathcal{S}$, and then transmitting the uncorrelated values.

More precisely, the collection of smooth pictures $\mathcal{S}$ has off-diagonal terms in the autocovariance matrix $A = \big(A(i,j)\big)_{i,j=1}^{N}$ of the pixels in $\mathcal{S}$:

$$A(i,j) = \frac{1}{M} \sum_{m=1}^{M} \hat{S}_m(i) \times \hat{S}_m(j), \tag{3.1-1}$$

where $\hat{S}_m = S_m - \frac{1}{M}\sum_m S_m$. $A$ can be diagonalized because it is symmetric (see [Apostol], theorem 5.4, page 120, for a proof of this very general fact). We can write $T$ for the orthogonal matrix that diagonalizes $A$; then $TAT^*$ is diagonal, and $T$ is called the *Karhunen–Loève transform*, or alternatively the *principal orthogonal decomposition*. The rows of $T$ are the vectors of the Karhunen–Loève basis for the collection $S$, or equivalently for the matrix $A$. The number of positive eigenvalues on the diagonal of $TAT^*$ is the actual number of uncorrelated parameters, or degrees of freedom, in the collection of pictures. Each eigenvalue is the variance of its degree of freedom. $TS_m$ is $S_m$ written in these uncorrelated parameters, which is what we should transmit.

Unfortunately, the above method is not practical because of the amount of computation required. For typical pictures, $N$ is in the range $10^4$–$10^6$. To diagonalize $A$ and find $T$ requires $O(N^3)$ operations in the general case. Furthermore, to apply $T$ to each picture requires $O(N^2)$ operations in general. Hence several simplifications are usually made.

### 3.2: DCT.

For smooth signals, the autocovariance matrix is assumed to be of the form

$$A(i,j) = r^{|i-j|} \tag{3.2-2}$$

where $r$ is the adjacent pixel correlation coefficient and is assumed to be just slightly less than 1. The expression $|i-j|$ should be interpreted as $|i_r - j_r| + |i_c - j_c|$, where $i_r$ and $i_c$ are respectively the row and column indices of pixel $i$, and similarly for $j$. Experience shows that this is quite close to the truth for small sections of large collections of finely sampled smooth pictures. It is possible to compute the Karhunen–Loève basis exactly in the limit $N \to \infty$: in that case $A$ is the matrix of a two-dimensional convolution with an even function, so it is diagonalized by the two-dimensional discrete cosine transform (DCT). In one dimension, this transform is an inner product with functions such as the one in the figure below:

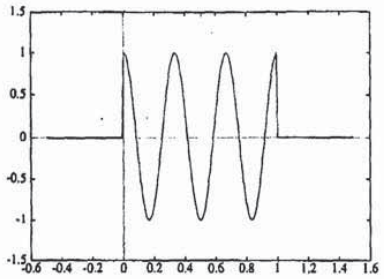


**Figure 3.2-2.**
**Example DCT basis function.**

This limit transform can be used instead of the exact Karhunen–Loève basis; it has the added advantage of being rapidly computable via the fast DCT derived from the fast Fourier transform. The Joint Photographic Experts Group (JPEG) algorithm uses this transform and one other simplification. $N$ is limited to 64 by taking $8 \times 8$ sub-blocks of the picture. JPEG applies two-dimensional DCT to the sub-blocks, then treats the 64 vectors of amplitudes individually in a manner we will discuss in the next section.

### 3.3: LCT or LOT.

Rather than use disjoint $8 \times 8$ blocks as in JPEG, it is possible to use "lapped" or "localized" (but still orthogonal) discrete cosine functions which are supported on overlapping patches of the picture. These *local cosine transforms* (LCT, as in [Coifman,Meyer]) or *lapped orthogonal transforms* (LOT, as in [Malvar]) are modifications of DCT which attempt to solve the blockiness problem by using smoothly overlapping blocks. This can be done in such a way that the overlapping blocks are still orthogonal, i.e., there is no added redundancy from using amplitudes in more than one block to represent a single pixel. For the smooth blocks to be orthogonal we must use DCT-IV, which is the discrete cosine transform using half-integer grid points and half-integer frequencies. The formulas for the smooth overlapping basis functions in two dimensions are derived from the following formulas in one dimension.

For definiteness we will use a particular symmetric bump function

$$b(x) = \begin{cases} \sin \frac{\pi}{4}(1 + \sin \pi x), & \text{if } -\frac{1}{2} < x < \frac{3}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.3-3}$$

This function is symmetric about the value $x = \frac{1}{2}$. It is smooth on $(-\frac{1}{2}, \frac{3}{2})$ with vanishing derivatives at the boundary points, so that it has a continuous derivative on $\mathbf{R}$. Notice that we can modify $b$ to obtain more continuous derivatives by iterating the innermost $\sin \pi x$. Let $b_1(x) = b(x)$ and define

$$b_{n+1}(x) = b_n\left(\tfrac{1}{2} \sin \pi x\right) \tag{3.3-4}$$

Then $b_n$ will have (use L'Hôpital's rule!) at least $2^{n-1}$ vanishing derivatives at $-\frac{1}{2}$ and $\frac{3}{2}$.

Now consider the interval of integers $I = \{0, 1, 2, \ldots, N-1\}$ where $N = 2^n$ is a positive integer power of 2. This can be regarded as the "current block" of $N$ samples in an array; there are previous samples $I' = \{\ldots, -2, -1\}$ and future samples $I'' = \{N, N+1, \ldots\}$ as well. The lapped orthogonal functions are

mainly supported on $I$, but they take values on $\{-N/2,\ldots,-1\} \subset I'$ and $\{N,\ldots,N/2-1\} \subset I''$ as well; those are the overlapping parts. For integers $k \in \{0,1,\ldots,N-1\}$, we can define the function

$$\psi_k(j) = \frac{1}{\sqrt{2N}}\, b\left(\frac{j+\frac{1}{2}}{N}\right) \cos\left(\pi\left(k+\frac{1}{2}\right)\left[\frac{j+\frac{1}{2}}{N}\right]\right) \tag{3.3-5}$$

Apart from $b$, these are evidently the basis functions for the so-called DCT-IV transform. The figure below shows one such function, with $N$ chosen large enough so that the smoothness is evident:

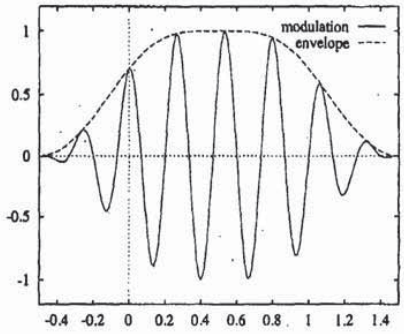


**Figure 3.3-3.**
**Example LCT basis function.**

The orthogonality of such functions may be checked by verifying the following equations:

$$\sum_{j=-\frac{3N}{2}}^{\frac{3N}{2}-1} \psi_k(j)\psi_{k'}(j) = \begin{cases} 1, & \text{if } k = k', \\ 0, & \text{if } k \neq k'. \end{cases} \tag{3.3-6}$$

The chosen window function or "bell" allows cosines on adjacent intervals to overlap while remaining orthogonal. For example, the function $\psi_k(j+N)$ is centered over the range $j \in \{-N,-N+1,\ldots,-1\}$ and overlaps the function $\psi_{k'}(j)$ at values $j \in \{-N/2,-N/2+1,\ldots,N/2-1\}$. Yet these two functions are orthogonal, which may be checked by verifying the equation below:

$$\sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} \psi_k(j+N)\psi_{k'}(j) = 0, \qquad \text{for all integers } k,\ k'. \tag{3.3-7}$$

Of course, rather than calculate inner products with the sequences $\psi_k$, we can preprocess data so that standard fast DCT-IV algorithms may be used. This may be visualized as "folding" the overlapping parts of the bells back into the interval. This folding can be transposed onto the data, and the result will be disjoint intervals of samples which can be "unfolded" to produce smooth overlapping segments. This is best illustrated by an example. Suppose we wish to fold a smooth function across 0, onto the intervals $\{-N/2,\ldots,-1\}$ and $\{0,1,\ldots,N/2-1\}$, using the bell $b$ defined above. Then folding replaces the function

$f = f(j)$ with the left and right parts $f_{0-}$ and $f_{0+}$:

$$(3.3\text{-}8)\qquad \begin{aligned} f_{0-}(j) &= \begin{cases} f(j), & \text{if } j < -N/2, \\ b(\frac{-j-\frac{1}{2}}{N})f(j) - b(\frac{j+\frac{1}{2}}{N})f(-j-1), & \text{if } j \in \{-N/2,\dots,-1\}, \end{cases} \\ f_{0+}(j) &= \begin{cases} b(\frac{j+\frac{1}{2}}{N})f(j) + b(\frac{-j-\frac{1}{2}}{N})f(-j-1), & \text{if } j \in \{0,1,\dots,N/2-1\}, \\ f(j), & \text{if } j \geq N/2. \end{cases} \end{aligned}$$

The symmetry of $b$ allows us to use $b(-x)$ instead of introducing the bell attached to the left interval. This action divides $f$ into two independent functions (the even and odd parts of $f$) which merge smoothly around the grid point 0. The process is an orthogonal transformation. We can fold the smooth function around the grid point $N$ in a similar manner:

$$(3.3\text{-}9)\qquad \begin{aligned} f_{1-}(j) &= \begin{cases} f(j), & \text{if } j < N/2; \\ b(\frac{-j-\frac{1}{2}}{N} - 1)f(j) - b(\frac{j+\frac{1}{2}}{N} - 1)f(2N-j-1), & \text{if } j \in \{N/2,\dots,N-1\}, \end{cases} \\ f_{1+}(j) &= \begin{cases} b(\frac{j+\frac{1}{2}}{N} - 1)f(j) + b(\frac{-j-\frac{1}{2}}{N} - 1)f(2N-j-1), & \text{if } j \in \{N,N+1,\dots,3N/2-1\}, \\ f(j), & \text{if } j \geq 3N/2. \end{cases} \end{aligned}$$

The new function $f_0$ defined below is a smooth, independent segment of the original smooth function $f$, restricted to the interval of values $\{0,1,\dots,N-1\}$:

$$(3.3\text{-}10)\qquad f_0(j) = \begin{cases} f_{0+}(j), & \text{if } j \in \{0,1,\dots,N/2-1\}, \\ f_{1-}(j), & \text{if } j \in \{N/2,N/2+1,\dots,N-1\}. \end{cases}$$

We can now apply the $N$-point DCT-IV transform directly to $f_0$.

We can likewise define $f_m(j)$ for the values $j \in \{mN, mN+1, \dots, (m+1)N-1\}$ by the same folding process, which segments a smooth function $f$ into smooth independent blocks. Folding to intervals of different lengths is easily defined as well. We can also generalize to two dimensions by separably folding in $x$ and then in $y$.

Unfolding reconstructs $f$ from $f_{0-}$ and $f_{0+}$ by the following formulas:

$$(3.3\text{-}11)\qquad f(j) = \begin{cases} b(\frac{-j-\frac{1}{2}}{N})f_{0-}(j) + b(\frac{j+\frac{1}{2}}{N})f_{0+}(-j-1), & \text{if } j \in \{-N/2,\dots,-1\}, \\ b(\frac{j+\frac{1}{2}}{N})f_{0+}(j) - b(\frac{-j-\frac{1}{2}}{N})f_{0-}(-j-1), & \text{if } j \in \{0,1,\dots,N/2-1\}. \end{cases}$$

Composing these relations yields $f(j) = \left[b(\frac{j+1/2}{N})^2 + b(\frac{-j-1/2}{N})^2\right] f(j)$. This equation is verified by the bell $b$ defined in Eq.(3.3-3), for which the sum of the squares is 1.

**3.4: Adapted block cosines.**

We can also build a library of block LCT bases (or block DCT bases) and search it for the minimum of some cost function. The chosen "best LCT basis" will be a patchwork of different-sized blocks, adapted to different-sized embedded textures in the picture. Again it will be necessary to encode the basis choice together with the amplitudes. A description of two versions of this algorithm and some experiments may be found in [Fang,Seré].

**3.5: Subband coding.**

A (one-dimensional) signal may be divided into frequency subbands by repeated application of convolution by a pair of digital filters, one high-pass and one low-pass, with mutual orthogonality properties.

Let $\{h_k\}_{k=0}^{M-1}, \{g_k\}_{k=0}^{M-1}$ be two finite sequences, and define two operators $H$ and $G$ as follows:

$$(3.5\text{-}12)\qquad (Hf)_k = \sum_{j=0}^{M-1} h_j f_{j+2k}, \qquad (Gf)_k = \sum_{j=0}^{M-1} g_j f_{j+2k}.$$

$H$ and $G$ are defined on square-summable signal sequences of any length. They are also be defined for periodic sequences of (even) period $P$, where we simply interpret the index of $f$ as $j + 2k \pmod P$. In that case, the filtered sequences will be periodic with period $P/2$.

The adjoints $H^*$ and $G^*$ of $H$ and $G$ are defined by

$$(H^*f)_k = \sum_{0 \le k-2j < M} h_{k-2j} f_j, \qquad (G^*f)_k = \sum_{0 \le k-2j < M} g_{k-2j} f_j. \tag{3.5-13}$$

$H$ and $G$ are called *(perfect reconstruction) quadrature mirror filters (or QMFs)* if they satisfy a pair of orthogonality conditions:

$$HG^* = GH^* = 0; \qquad H^*H + G^*G = I. \tag{3.5-14}$$

Here $I$ is the identity operator. These conditions translate to restrictions on the sequences $\{h_k\}, \{g_k\}$. Let $m_0, m_1$ be the bounded periodic functions defined by

$$m_0(\xi) = \sum_{k=0}^{M-1} h_k e^{ik\xi}, \qquad m_1(\xi) = \sum_{k=0}^{M-1} g_k e^{ik\xi}. \tag{3.5-15}$$

Then $H, G$ are quadrature mirror filters if and only if the matrix below is unitary for all $\xi$:

$$\begin{pmatrix} m_0(\xi) & m_0(\xi+\pi) \\ m_1(\xi) & m_1(\xi+\pi) \end{pmatrix} \tag{3.5-16}$$

This fact is proved in [Daubechies]. QMFs can be obtained by constructing a sequence $\{h_k\}$ with the desired low-pass filter response, and then putting $g_k = (-1)^k h_{M-1-k}$. That reference also contains an algorithm for constructing a family of such $\{h_k\}$, one for each even filter length $M$.

The frequency response of one particular pair of QMFs ("C30") is depicted below. We have plotted the absolute values of $m_0$ and $m_1$, respectively, over one period $[-\pi, \pi]$:

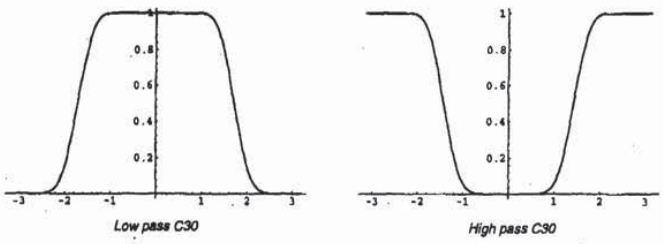


**Figure 3.5-4.**

**Absolute values of $m_0$ and $m_1$.**

Notice that $m_0$ attenuates frequencies away from 0, while $m_1$ attenuates those away from 0.

Below is the traditional block diagram describing the action of a pair of quadrature mirror filters. On the left is convolution and downsampling (by 2); on the right is upsampling (by 2) and adjoint convolution, followed by summing of the components. The broken lines in the middle represent either transmission or storage.

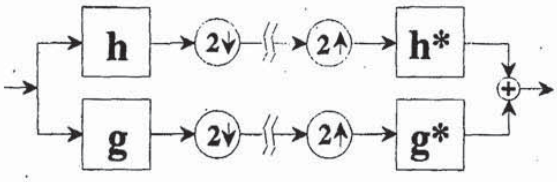


**Figure 3.5-5.**

**Block diagram of subband filtering.**

The underlying functions of subband filtering are produced by iterating $H^*$ and $G^*$ until we have enough points. For example, 10 iterations of $H^*$ applied to the sequence $e_0 = \{\dots, 0, 0, 1, 0, 0, \dots\}$ produces a 1024-point approximation to the smooth function whose translates span the lowest-frequency subband. Likewise, a single $G^*$ after 9 iterations of $H^*$ applied to $e_0$ produces a 1024-point approximation to the next lowest-frequency function. These are distinguished examples; the first is called the "scaling" or "father" function, while the second is called the "wavelet" or "mother" function. They are depicted below for a particular QMF ("C 30," with 30 taps):

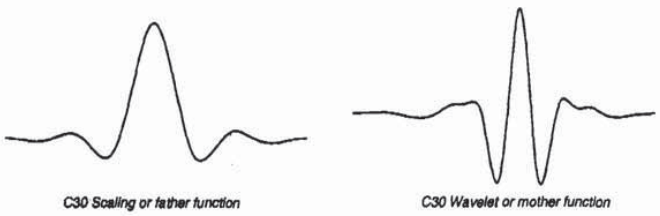


**Figure 3.5-6.**

**Lowest frequency subband basis functions.**

Higher-frequency subbands are spanned by functions with more oscillations, which are produced by using $G^*$ earlier in the iteration. The sequence of filters used to generate a function can be converted to an integer in binary notation as follows. Put $F_0^* = H^*$ and $F_1^* = G^*$ in the formula for the function (respectively, put $F_0 = H$ and $F_1 = G$). Then for any pair of integers $n$ and $L$ with $0 \le n < 2^L$ we can write $n = n_0 2^0 + n_1 2^1 + \cdots + n_{L-1} 2^{L-1}$, where $n_i \in \{0, 1\}$ for all $i = 0, 1, \dots, L-1$. To that combination $(n, L)$ we can associate a vector $F_{n_0}^* \circ F_{n_1}^* \circ \cdots \circ F_{n_{L-1}}^* e_0$.

For example, the following functions are 1024-point sequences $H^*G^*(H^*)^8 e_0$ and $G^*G^*(H^*)^8 e_0$, respectively given by $(n, L) = (2, 10)$ and $(n, L) = (3, 10)$:

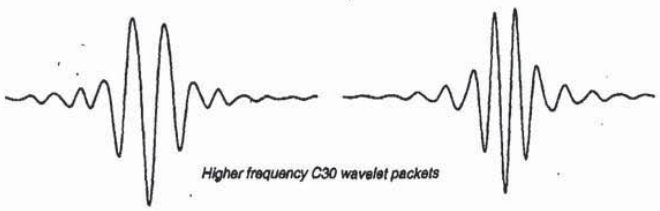


**Figure 3.5-7.**
**Higher-frequency subband basis functions.**

It is well known (and shown in [INRIA]) that the number of oscillations of the vector produced in this manner increases with $n'$, where $n$ is the Gray-code permutation of $n'$. The renumbering $n \rightarrow n'$ relates Paley order to sequency order for Walsh functions, and has an analogous effect in the smooth case. This fact can be used to analyze the spectrum of acoustic signals by measuring the amplitudes of wavelet packet coefficients. Acousfic signal compression by wavelet packet and best-basis methods was discussed in [W1].

Now we can define 4 2-dimensional convolution-decimation operators in terms of $H$ and $G$, namely the tensor products of the pair of quadrature mirror filters:

$$F_0 \stackrel{\text{def}}{=} H \otimes H, \qquad F_0 v(x,y) = \sum_{i,j} h_i h_j v(i+2x, j+2y) \tag{3.5-17}$$

$$F_1 \stackrel{\text{def}}{=} H \otimes G, \qquad F_1 v(x,y) = \sum_{i,j} h_i g_j v(i+2x, j+2y) \tag{3.5-18}$$

$$F_2 \stackrel{\text{def}}{=} G \otimes H, \qquad F_2 v(x,y) = \sum_{i,j} g_i h_j v(i+2x, j+2y) \tag{3.5-19}$$

$$F_3 \stackrel{\text{def}}{=} G \otimes G, \qquad F_3 v(x,y) = \sum_{i,j} g_i g_j v(i+2x, j+2y) \tag{3.5-20}$$

These convolution-decimations have the following adjoints:

$$F_0^* v(x,y) = \sum_{i,j} h_{x-2i} h_{y-2j} v(i,j) \tag{3.5-21}$$

$$F_1^* v(x,y) = \sum_{i,j} h_{x-2i} g_{y-2j} v(i,j) \tag{3.5-22}$$

$$F_2^* v(x,y) = \sum_{i,j} g_{x-2i} h_{y-2j} v(i,j) \tag{3.5-23}$$

$$F_3^* v(x,y) = \sum_{i,j} g_{x-2i} g_{y-2j} v(i,j) \tag{3.5-24}$$

The orthogonality relations for this collection are as follows:

$$F_n F_m^* = \delta_{nm} I; \tag{3.5-25}$$

$$I = F_0^* F_0 \oplus F_1^* F_1 \oplus F_2^* F_2 \oplus F_3^* F_3. \tag{3.5-26}$$

By a "picture" we will mean a finite sequence indexed by two coordinates $S = S(x, y)$. It is convenient to regard pictures as periodic in botyh $x$ and $y$, though this is not absolutely necessary. For simplicity of implementation, we shall also assume that the $x$-period (the "width" $N_x = 2^{n_x}$) and the $y$-period (the "height" $N_y = 2^{n_y}$) are both positive integer powers of 2, so that we can always decimate by two and get an integer period. The space of such pictures may be decomposed into a partially ordered set $\mathbf{W}$ of subspaces $W(n, m)$ called *subbands* (see below), where $m \geq 0$, and $0 \leq n < 4^m$. These are the images of orthogonal projections composed of products of convolution-decimations. Denote the space of $N_x \times N_y$ pictures by $W(0,0)$ (it is $N_x \times N_y$ dimensional), and define recursively

$$W(4n + i, m + 1) = F_i^* F_i W(n, m) \subset W(0,0), \qquad \text{for } i = 0, 1, 2, 3. \tag{3.5-27}$$

The orthogonality condition on the QMFs implies that the projections from $W(0,0)$ onto $W(n, m)$ are orthogonal, i.e., they conserve energy. The subspace $W(n, m)$ is $(N_x 2^{-m}) \times (N_y 2^{-m})$-dimensional. These subspaces may be partially ordered by a relation which we define recursively as follows. We say $W$ is a *precursor* of $W'$ (write $W \prec W'$) if they are equal or if $W' = F^* F W$ for a convolution-decimation $F$ in the set $\{F_0, F_1, F_2, F_3\}$. We also say that $W \prec W'$ if there is a finite sequence $V_1, \ldots, V_s$ of subspaces in $\mathbf{W}$ such that $W \prec V_1 \prec \cdots \prec V_s \prec W'$. This is well defined, since each application of $F^* F$ increases the index $m$.

Subspaces of a single precursor $W \in \mathbf{W}$ will be called its *descendents*, while the first generation of descendents will naturally be called *children*. By the orthogonality condition,

$$W = F_0^* F_0 W \oplus F_1^* F_1 W \oplus F_2^* F_2 W \oplus F_3^* F_3 W. \tag{3.5-28}$$

The right hand side contains all the children of $W$.

The subspaces $W(n, m)$ are called *subbands*, and the transform coding method that first transforms a signal into subband coordinates is called *subband coding*. If $S \in W(0,0)$ is a picture, then its orthogonal projection onto $W(n, m)$ can be computed in the standard coordinates of $W(n, m)$ by the formula $F_{(1)} \ldots F_{(m)} W(0,0)$, where the particular filters $F_{(1)} \ldots F_{(m)}$ are determined uniquely by $n$. Therefore we can express in standard coordinates the orthogonal projections of $W(0,0)$ onto the complete tree of subspaces $\mathbf{W}$ by recursively convolving and decimating with the filters.

The quadrature mirror filters $H$ and $G$ form a partition of unity in the Fourier transform (or wavenumber) space. The same is true for the separable filters $F_i$. They can be described as nominally dividing the support set of the Fourier transform $\hat{S}$ of the picture into dyadic squares. If the filters were perfectly sharp, then this would be literally true, and the children of $W$ would correspond to the 4 dyadic subsquares one scale smaller. We illustrate this in the figure below.

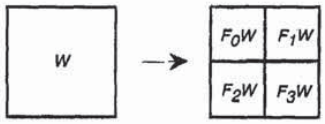


**Figure 3.5-8.**

**Four subband descendents of a picture.**

The next figure shows 2 generations of descendents, the complete decomposition of $\mathbf{R}^4 \times \mathbf{R}^4$. The subbands are labelled by the "$n$" index in $W(n, m)$. Within the dyadic squares are the $n$-indices of the corresponding

subspaces at that level. If we had started with a picture of $N \times N$ pixels, then we could repeat this decomposition process $\log_2(N)$ times.

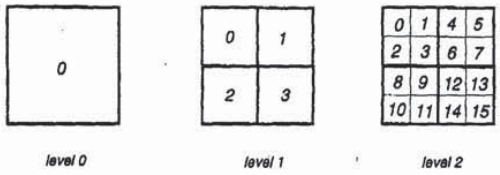


**Figure 3.5-9.**

**Two levels of subband decomposition.**

All subbands together form a quadtree, in which each subspace forms a node and directed edges go from precursors to descendents. The orthogonality relation among the subspaces implies that every connected subtree which contains the root $W(0,0)$ corresponds to an orthonormal subband decomposition of the original picture: the subbands correpond to the leaves of the subtree. Having stated this general nonsense result, let us consider specific examples.

The subbands $W(n,m)$ as we define them are in one-to-one correspondence with rectangular regions (in fact squares) in wavenumber space, and the quadtree stacks these regions one on top of the other. We can idealize various orthogonal subband bases as disjoint covers of wavenumber space. A few of these are schematically represented below:

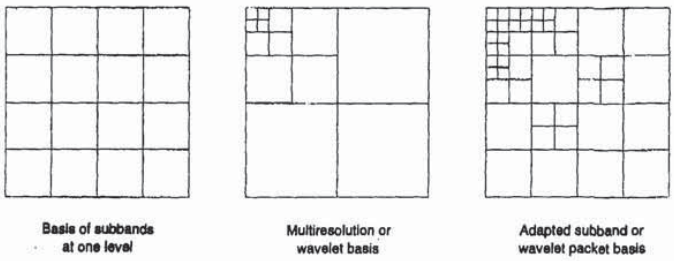


**Figure 3.5-10.**

**Various decompositions into subbands.**

The leftmost decomposition is a subband decomposition in which we have all the subbands at a fixed level—in this case, level 2. The subbands are labeled (0,2), (1,2), (2,2), (3,2), ... , (15,2) as in figure 3.5-9 above. The middle decomposition produces 2-dimensional "isotropic" wavelets, i.e., which have the same scale in both the $x$ and $y$ directions. The subbands in this decomposition are labelled by (0,4), (1,4), (2,4), (3,4), (1,3), (2,3), (3,3), (1,2), (2,2), (3,2), (1,1), (2,1), and (3,1).

The rightmost decomposition is an adapted subband basis such as might be discovered by minimizing

# File History Content Report

The following content is missing from the original file history record obtained from the United States Patent and Trademark Office. No additional information is available.

Document Date - 1996-07-08

Document Title - Non-Patent Literature

Page(s) - 12

obvious that a function $\psi$ with the requisite properties exists. The surprising and fortunate recent discovery of many such functions ([Daubechies], [Mallat], [Meyer] and others) also provided a fast $O(N)$ algorithm to compute the associated wavelet transforms.

The wavelet basis down to level $L$ consists of the elements spanning the subbands $W(1,1)$, $W(2,1)$, $W(3,1)$, $W(1,2)$, $W(2,2)$, $W(3,2)$, ... ,$W(1,L-1)$, $W(2,L-1)$, $W(3,L-1)$, $W(1,L)$, $W(2,L)$, $W(3,L)$ and the largest-scale average $W(0,L)$. The pixel values may be transformed into this basis via the 2-dimensional version of the "pyramid scheme" described in [Mallat]. Graphically, this is the following:

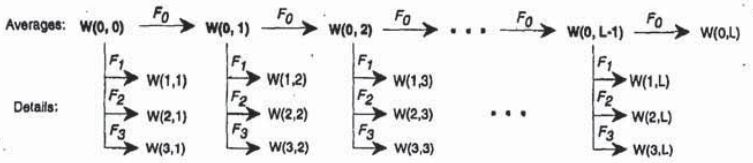


**Figure 3.5.2-12.**

**Two dimensional pyramid scheme.**

The pyramid scheme is also called a multiresolution analysis, and has been extensively studied. It provides an algorithm of complexity $O(N)$ for the transformation of an $N$-pixel picture.

*3.5.3: Custom subbands.* There may be features which are most efficiently described by expansion into certain subbands. For example, if the image contains textures of a particular size and frequency, then we will find large values in the corresponding subband. This is true for example in fingerprint image compression, where both the large-scale texture of the ridges and the fine-scale texture of the pores contribute large values. The first are not so important for identification since ridges can be easily deformed, while the latter must be preserved very accurately since pore shapes and distributions often provide strong clues for identification. In that case, we can choose to include both the ridge and pore subbands in our transform, de-emphasize the values of the first and amplify the values in the second.

**3.6: Adapted subband coding.**

Wavelets decorrelate pictures which are close to self-similar. Other subband bases decorrelate pictures composed of overlapping textures on different scales. Some combination is needed for pictures which are close to self-similar but contain embedded patches of textures of various sizes, but it is not clear that any fixed choice of subbands will contain suitable templates. But it is possible to use a library of bases of *wavelet packets* (this name is from [CMQW]), which are efficiently encoded superpositions of wavelets. These adapted subband bases come with a natural quad-tree organization and some remarkable orthogonality properties. It is possible to introduce a cost function and pick a "best" wavelet packet basis for one or many pictures. This basis and the resulting decorrelated pixel valuess can then be compactly coded, but also the analysis performed during the choice of representation provides some information about the picture and could be useful for feature recognition.

*3.6.1: Adapted subbands or wavelet packets.* As mentioned in an earlier paper [CMQW], we can build a large library of adapted subband bases by retaining all amplitudes in the quadtree. The amplitudes produced at each stage are correlations of the signal with compactly-supported oscillatory functions called *wavelet packets*. From the tree $\mathbf{W}$ of subspaces we may choose a *basis subset*, defined as a collection of mutually orthogonal subspaces $W \in \mathbf{W}$, or lists of pairs $(n,m)$, which together span the root. Basis subsets

are in one-to-one correspondence with dyadic decompositions of the unit square. Classical subband coding takes amplitudes from a fixed set of subbands, usually from a single level of the quadtree. Wavelet transform coding also extracts amplitudes from a fixed collection of blocks, the octave subbands.

Even for a small tree, the library of wavelet packet bases is very large:

**Proposition.** *The number of wavelet packet basis subsets for $N$-pixel pictures is greater than $2^N$. The number of operations needed to compute all the transformed pixel values in all these basis subsets, however, is no more than $N \log_4(N)$.*

*Proof.* A decomposition to level $n$ is only possible for a picture of size at least $N = 4^n$ pixels, and in such a tree there can therefore be at most $N \log_4(N)$) transformed pixel values. Let $A_n$ be the number of bases in the library corresponding to a tree of $1+n$ levels, namely levels $0, \ldots, n$. Then $A_0 = 1$, and we can calculate $A_{n+1} = 1 + A_n^4$, namely the root and combinations of the 4 children, which are independent subtrees with $A_n$ bases each. Simplifying this by discarding the 1 gives the estimate $A_{n+1} > 2^{4^n} = 2^N$ for $n > 1$. □

*3.6.2: The best-basis algorithm.* To each subspace $W \in \mathbf{W}$ we may assign an information cost $H_W$. The quantity $H_W(S)$ measures the expense of including $W$ in the decomposition used to represent the picture $S$. Define the *best basis* for representing $S$ (with respect to $H_W$) to be the basis subset $B_0$ which minimizes

( 3.6.2-29)
$$\sum_{W \in B} H_W(S)$$

over all basis subsets $B \subset \mathbf{W}$.

Some examples of information cost functions are listed in [CW]. The simplest is the number of elements above a predetermined threshold $\epsilon$, namely $H_W(S) = \#\{x \in S_W : |x| \geq \epsilon\}$, where $S_W$ is the sequence of the pixel values of $S$ as transformed into the standard basis of $W$. This sequence is $F_{i_m} \ldots F_{i_1} S$, where $W = F_{i_m}^* F_{i_m}^* \ldots F_{i_1}^* F_{i_1} W(0,0)$. The following algorithm finds the basis subset with the fewest amplitudes above the threshold.

Set a predetermined deepest level $L$. Label as "kept" each subspace at level $L$, i.e., the subspaces indexed by $(n, L)$ for $0 \leq n < 4^L$. Next, set the level index $m$ to $L-1$. Compare the information cost of the subspace $W(n,m)$ with the sum of the information costs of its children $W(4n, m+1)$, $W(4n+1, m+1)$, $W(4n+2, m+1)$, and $W(4n+3, m+1)$. If the parent is less than or equal to the sum of the children, then mark the parent as "kept." This means that by choosing the parent rather than the children, we will have fewer amplitudes above the threshold in the representation of $S$. On the other hand, if the sum of the children is less than the parent, leave the parent unmarked but attribute to her the sum of the children's information costs. By passing this along, prior generations will always have their information costs compared to the least costly collection of descendents.

After all the subspaces at level $m = L-1$ have been compared to their children, decrement the level index and continue the comparison. At each level, we are comparing the information cost of a node to the sum of the lowest information costs obtainable by any decompositions of its 4 children. We can proceed in this way until we have compared the root $W(0,0)$ to its 4 children. We claim that the topmost "kept" nodes in depth-first order constitute a best basis. I.e., the collection of "kept" nodes $W$ with no "kept" precursors is a basis subset which minimizes information cost. But this is easily proved by induction on the level index (see [CW] for the details).

If we think of the amplitudes below $\epsilon$ as negligible, we now have a basis in which the fewest amplitudes are non-negligible. This cost accounting requres deciding in advance what negligible means, which in some applications may not be feasible. The decision may be postponed by using a different measure of the concentration of energy into the amplitudes. For example, there is an additive analog of Shannon entropy, namely,

( 3.6.2-30)
$$H_W(S) = -\sum_{x \in S_W} x^2 \log x^2,$$

with $S_W$ as above. This is related to the classical measure of the concentration of a probability distribution function which we discussed in an earlier section.

### 3.7: Functions underlying the transforms.

Each of the transform methods correlates a picture with some underlying functions or "templates" and then stores the correlations. For JPEG, the templates are products of sampled cosine functions restricted to square blocks:

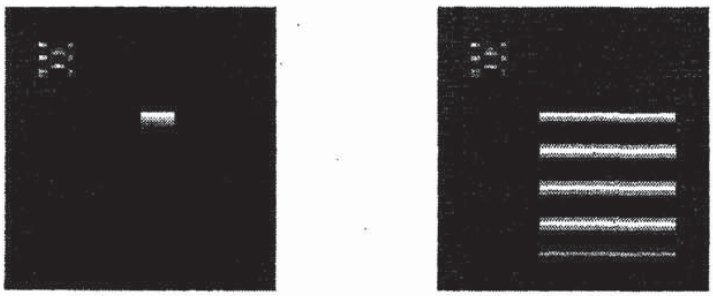

**Figure 3.7-13.**
**Density plots of fixed-block and adapted-block DCT functions.**

For LCT, the underlying functions are "smeared out," without the sharp edges of the DCT functions:

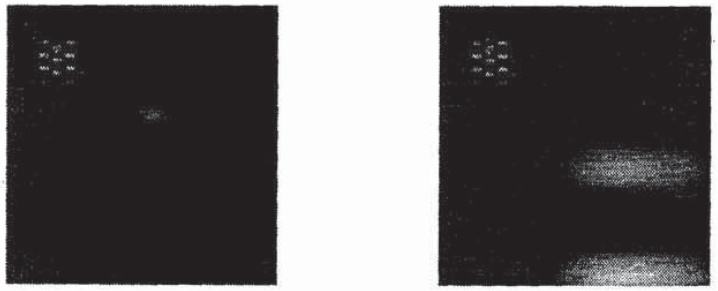

**Figure 3.7-14.**
**Density plots of fixed-block and adapted-block LCT functions.**

Two-dimensional wavelets and wavelet packets superpose in a different manner; wavelet packets of different scales can overlap in any manner, so long as their frequencies are distinct:

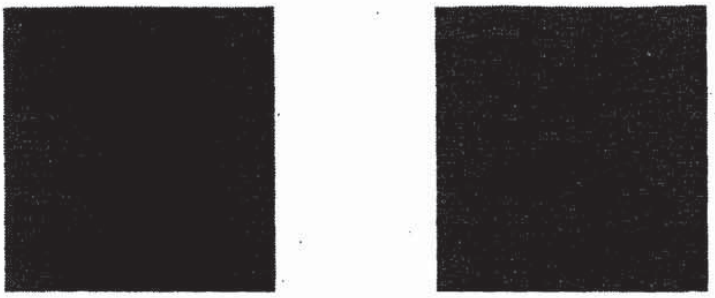

**Figure 3.7-15.**

**Density plots of one wavelet and a superposition of three wavelets.**

Wavelet packets, on the other hand, are superpositions of wavelets which are arranged so that they are efficient to describe. In particular, they correlate better with textures:

**Figure 3.7-16.**

**Density plots of one wavelet packet and a superposition of three.**

### 4: Quantization of transformed amplitudes

The transformed pixel values are real numbers which must be approximated in a (small) finite alphabet, or *quantized*, before they can be transmitted. All of the distortion from a lossy transform coding scheme is introduced at this step. The range of transformed values is divided up into numbered subintervals or *bins*. Any pixel value falling into a bin is approximated by the bin's index, as in the figure below:

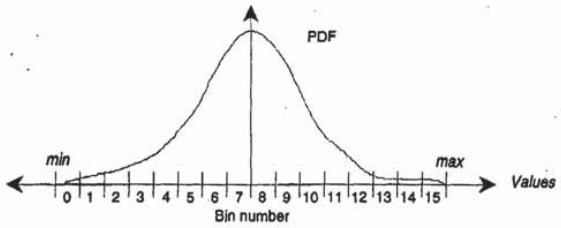


**Figure 4-17.**

**Quantization into 16 equal-size bins.**

In practice, the bin around 0 is often taken to be twice as wide as the other bins, and quantization is performed by scaling and truncation to the integer nearest to 0. The resulting integers are then biased into the range $[0, 1, \ldots, b-1]$, where $b$ is the number of bins.

Quantization is undone by replacing the bin index with the value at the center of the bin. When the transformed pixel values will always have the same sign, or when 0 plays no special role, then quantization is done by dividing the full range between the maximum and minimum values into $b$ bins. When the transformed pixel values can be both positive and negative and 0 is an important value, then it is vital that the bin around 0 be centered exactly at 0. Otherwise, the one-sided bias introduced by many values close to but not exactly equal to 0 will appear as artifacts in the reconstructed image.

The *quantization error* is the difference between the actual sequence of values and the sequence of bin center values, measured in some norm like mean-square-error or maximum absolute error. It is possible to vary the width of the bins, so that the more popular values (where the PDF is larger) are quantized more finely, into narrower bins. The distortion-minimizing choice of bin widths given a fixed number of bins can be found using the Lloyd-Max algorithm (see [Jayant,Noll]). This has the effect of reparametrizing the rate-distortion curve for the signal, but not improving the compression rate for a given degree of distortion. We can easily allow variable numbers of equal-sized bins, to adjust the compression rate.

**4.1: Uniform quantization.** In this method we use a single set of bins for all the transformed pixel values. This method is used when we have no *a priori* knowledge of the importance or relative visibility of a basis element.

**4.2: Visibility quantization.** In this method we use a model of the relative importance of a transformed pixel value to choose a weighting coefficient. The value is multiplied by this weight prior to quantization. It is known that the human eye is less sensitive to errors at certain spatial frequencies. When it is possible to determine the spatial frequency of the part of an image which will be reconstructed from an amplitude, this fact may be used to reduce the perceived distortion at a given level of compression.

**4.3: Bit allocation.** When transformed pixel values can be grouped, for example by subband, then we can allocate bits to the groups in a nonuniform manner to minimize the quantization error. The optimal allocation for a fixed number of subbands assigns bits in proportion to the variance within a subband (see [Jayant,Noll]). Another way to put this is, if $q_i$ is the number of quantization bins to be assigned to subband $W_i$, then we should have $q_i/\sigma(W_i)$ =constant for all $i$, where the constant depends upon the total variance of all the transformed pixel values as well as on the total number of bits we can afford to transmit or store.

Great competitive advantage may be gained from intelligent bit-allocation schemes, perhaps in combination with visibility weighted quantization based on acurate models of the kind of images to be transmitted. Such schemes are valuable property, and jealously guarded secrets in the industry.

## 5: Removing redundancy or entropy coding

Digitally coded data can often be reversibly transformed into a more efficient form, requiring fewer bits to store than its original representation. We shall refer to such an invertible transform as "lossless coding" because all the information in the original bits can be recovered. This is the only acceptable way to compress certain kinds of data sets such as compiled computer programs and data archives. Such methods have great practical significance and have been extensively studied.

It is a classical fact (see, for example, [Shannon,Weaver]) that there is a limiting rate of (lossless) compression achievable by such an alphabet substitution. Suppose we have an arbitrarily long message $M = (M(1), M(2), M(3), \ldots)$ composed in an alphabet with finitely many letters: $M(j) \in a = \{a_1, a_2, \ldots, a_n\}$ for all $j = 1, 2, \ldots$. Coding these letters in the obvious manner requires $\log_2(n)$ bits per letter, and so the first $L$ letters of the message will require $L \log_2(n)$ bits to transmit. Let $B(a, L)$ denote the number of bits required to transmit the first $L$ letters $M(1) \ldots M(L)$ using alphabet $a$; then we see that

$$\lim_{L \to \infty} \frac{B(a, L)}{L} = \log_2(n) \tag{5-31}$$

Shannon's theorem asserts that there is a nonnegative number $H$, the entropy of the probability distribution of the original alphabet, such that any new alphabet $\{b_i\}_{i=1}^{n}$ of variable-length characters satisfies:

$$H \leq \lim_{L \to \infty} \frac{B(b, L)}{L}, \tag{5-32}$$

and that for every $\epsilon > 0$ there in fact exists a particular alphabet $b^\epsilon$ which satisfies

$$\lim_{L \to \infty} \frac{B(b^\epsilon, L)}{L} \leq H + \epsilon. \tag{5-33}$$

Suppose that the original signal is written using the letters $a_1, a_2, \ldots, a_n$, each occurring with probability

$$P(a_i) = p_i = \lim_{L \to \infty} \frac{\#\{j : 1 \leq j \leq L, M(j) = a_i\}}{L} \tag{5-34}$$

Then the entropy of the message is $H = -\sum_{i=1}^{n} p_i \log_2(p_i)$, and it is not too hard to show (as in [Ash], Theorem 1.4.2, p.17) that

$$0 \leq H \leq \log_2(n). \tag{5-35}$$

The left equality (best compression) holds if and only if the original message consists of a single letter of the alphabet repeated forever. The right equality (no compression) holds if and only if the original letters $a_i$ are equally probable.

Several algorithms exist to construct the good alphabets, of which the earliest is probably static Huffman coding [Huffman]. In practice the probabilities $p_i$ are determined empirically as the message is being sent, so there are refinements such as dynamic Huffman coding [Storer, p.40], arithmetic or Q-coding [Storer, p.47], data dictionary methods, and so on.

We can refine the application of Shannon's theorem by taking different lengths $n$ for the initial alphabet. For example, it is natural to consider 8-bit characters ($n = 256$) for binary data emitted by a typical computer. Nevertheless, it may be that for a certain class of signals (such as Kanji text, which uses more