数理解析研究所講究録
第 791 巻 1992 年 42-48

Wavelets and their Applications to Image Processing

(dedicated to my father for his 60th birthday)

Mei Kobayashi, IBM Tokyo Research Laboratory

Powerful computers are increasingly being used for the generation and processing of images. We may naively catagorize the images as: computer-generated, non-computer generated or non-computer-generated with computer-generated alterations. The first part of the talk consists of examples of computer-generated images to illustrate the capabilities of current graphics systems. Then we will move onto the primary subject of the talk, wavelets and image processing.

1. Four Graphics Projects

We begin by examining slides from four graphics projects at the IBM Tokyo Research Laboratory: Chemical Graphics, Airflow Simulation, Medical Systems, and Modeling and Rendering. A product developed in the first project by Koide et al [Koi] allows the user to see the 'shapes' of molecules using computer graphics; three types of visualization methods are available in its prototype model: wire mesh, density dots and solid ball/surface. Two-dimensional and three-dimensional translucent level sets indicating the probability of electron occupancy can also be generated. Results from a second, joint project with development and manufacturing show air and dust flow patterns in clean rooms, air heating systems and memory disks. In the first two applications FDM solvers were used by Koyamada and Miyazawa to compute flows, temperatures and pressures. An FEM model and solver were used by Kobayashi and Koyamada in the DASD problem because of a large number of irregularly shaped objects. Our two-dimensional numerical approximation of the pressure map inside a DASD compares well with experimentally measured results; outstanding features appear in the same areas in both models [KK]. Graphics tools developed at IBM illustrate streamlines and level sets in these very rough prototype models [KM]. The objective of a third project, CliPPS (Clinical Planning Support System), is to develop a visual and intelligent support system to aid in surgical and radiotherapy treatment planning [MIMY]. Special system features are scheduled to include three-dimensional translucent dynamic imaging and graphics. The goals of the last project are described by its name, 'Modeling and Rendering', i.e. to model and represent objects in a verbal or artistic form; depict. M. Miyata developed a fractal-based texture generation algorithm to draw realistic images of clouds [Myt1], carpets, lawns, grains in wood flooring and architectural interiors. A three minute-video film, *Edo* [Myt2], highlights stone wall [Myt3], pavement and pebble patterns generated using the algorithm and gives the viewer a glimpse into the life of old Tokyo and Edo Castle. These examples show the how current computers may generate near-realistic images. The next section will discuss how to compress, reconstruct and analyze image data like those discussed above as well those from photographic scanning.

2. Wavelets and their Applications to Image Processing

Wavelets have come to enjoy great attention and study in recent years for their ability to serve as an efficient basis set for approximating functions and operators arising in a variety of scientific and engineering applications. For detailed references on audio and image signal processing, solutions to differential equations and inverse problems, mathematical analysis and physics using wavelets, see [CGT],[FJW],[D1],[D2],[Le],[Me1],[RV]. We follow the notation of Daubechies [D1] and define wavelets as families of functions $h_{a,b}$

$$h_{a,b} = |a|^{-1/2} h\left(\frac{x-b}{a}\right) \quad ; \quad a, b \in R \quad , \quad a \neq 0$$

generated from a single function h by dilations and translations. One of the applications of the theory is to construct a basis set $\{h_{a,b}\}$ for efficient and accurate approximation of functions. In signal analysis [D2], the parameters a, b are restricted to a discrete sublattice, where we fix a dilation step $a_0 > 1$ and a translation step $b_0 \neq 0$. The corresponding wavelet family is $h_{m,n}(x) = |a_0|^{-m/2}\ h(a_0^{-m}x - nb_0)$. Here $a = a_0^m$ and $b = nb_0a_0^m$. We note that if the translation parameter b_0 is small, then the basis elements lie closer together (in some cases it may lead to overlap or redundancy) and the approximation would be of a fine resolution. In the extreme case $b_0 \to 0$, a continuous 'band' of these elements may be used to approximate a given function. The role of the dilation and translation parameters a and b as well as the parallels between the wavelet and Fourier methods will be discussed in more detail through an example given below.

But first, one may ask about the need for wavelet technology in signal processing when the Fourier Transform (FT) techniques have been developed and international standardization is under discussion. We quote from a recent article from the New York Times [Kol]: *'Researchers say a method, called wavelets, can provide clearer sound transmission and more bands for cellular telephones and better images for high definition television. In addition, it should make it possible to compress data that people thought could not be squeezed without loss of crucial details. The new method can also help mathematicians solve matrix equations, involving enourmous rows and columns of numbers, by allowing them to transform the equations into simpler forms.'* The article goes on to discuss problems which require massive data compression and reconstruction, such as weather and scientific satellite signal processing, 3D imaging of medical data, fingerprint storage for the F.B.I. and sound transmission, including music and voice. Kawahara of NTT has demonstrated how male and female voice data can be compressed and restored using less data than that required by conventional methods [Ka1],[Ka2]. In addition, we note that short-time FT techniques (STFT) experience a limitation known as the Heisenberg uncertainty principle [D2],[RV]; there is a trade-off between the time and frequency resolutions: $\Delta t \Delta f \geq 1/4\pi$. And the STFT are inflexible in time-frequency resolution; once a window is chosen, it remains fixed. There is no such limitation in the wavelet method. Some researchers believe that *' the wavelets method appears to resemble the way the human eye and human ear process data'* [Kol], thereby overcomes some of the problems associated with FT methods [RV]. It is only fair to comment that other researchers have expressed some reservations regarding practical applications; they believe that the primary advantage of using the wavelet technique appears to come from the overcompleteness property of wavelet bases, a characteristic which is not unique to wavelets [S]. There is yet a vast area of unexplored territory in the study of wavelets, and, hopefully, new and significant findings will enhance our scientific and technological knowledge.

For the remainder of this talk we will concentrate on the mathematics of wavelets and its applications to image processing. Wavelet and Fourier methods share many of the same properties and techniques. In the Fourier method, the basis set $\{\cos n\pi x, \sin n\pi x\}$ for approximating functions is generated from dilations of $\cos \pi x$ and $\sin \pi x$ on $[-1, 1]$, where n is the dilation parameter. Then translations of this basis set are used to approximate functions in intervals of length two. In the wavelet method, we generate a basis set by starting with any function $h(x)$. In our example, let $\chi_0(x)$ denote the characteristic function, which takes on the value one on the unit interval and zero elsewhere. To generate a basis set, we construct the mother functions $\Psi(x) = \chi(2x) - \chi(2x-1)$ which give us the Haar wavelets and $\Phi(x) = \chi(2x) + \chi(2x-1)$, the Haar scaling functions.

The Haar wavelet basis set for our reference resolution is $\{\phi_{0,n}(x)\}$, where $\phi_{0,n}(x) = 1$ on $[n-\frac{1}{2}, n+\frac{1}{2})$ and zero elsewhere. To generate a basis set on a finer resolution level, we set the mother function to $\phi_{-m,0}(x) = (\sqrt{2})^m$ on the interval $[-2^{-m-1}, 2^{-m-1})$ and zero elsewhere for postive integers m. For a coarser resolution level, set m to be a negative integer. The scaling functions allow us to move between different resolution levels. Sets of scaling functions are also generated by dilations and translations, where the dilationand translation parameters are determined by the resolution levels of the wavelets to be scaled. The basis set $\{\psi_{0,n}(x)\}$ for the Haar scaling functions which introduce a finer resolution level to the Haar wavelets $\{\phi_{0,n}(x)\}$ is described by $\psi_{0,n}(x) = -1$ on $[n-\frac{1}{2}, n)$, $\psi_{0,n}(x) = 1$ on $[n, n+\frac{1}{2})$ and zero elsewhere, for all $n \in Z$. As with the Haar wavelets, the support of the scaling functions is halved and height multiplied by a factor of $\sqrt{2}$ for finer resolution levels; the support of the scaling functions is doubled and height reduced by a factor of $\sqrt{2}$ for coarser resolution levels. In some ways, ϕ and ψ are analogous to the cosine and sine functions in Fourier expansions. The set of characteristic functions is a well-suited basis for image processing because this choice corresponds to sampling pixels at evenly spaced points on the two-dimensional plane, and assigning it the measured constant value on the sampling interval.

In both the Fourier and wavelet methods, transforms are used to encode the approximation of a function. For $f(x) \in L^1(R)$, the Fourier expansion on $[-1, 1]$ is [Ru]

$$f(x) = \sum_n a_n \cos n\pi x + b_n \sin n\pi x \quad ,$$

where

$$a_n = \int_{-1}^{1} dx \cdot \cos n\pi x \cdot f(x) \quad , \quad b_n = \int_{-1}^{1} dx \cdot \sin n\pi x \cdot f(x) \quad .$$

For wavelets with mother function h, the continuous wavelet transform for $f \in L^2(R)$ is defined as [D1],[D2]

$$(Uf)(a,b) = < h_{a,b}, f > = |a|^{-1/2} \int dx \cdot h(\frac{x-b}{a}) \cdot f(x) \quad ; \qquad (a,b) \in (R\ 0) \times R \quad ,$$

and the discrete wavelet transform

$$(Tf)_m = < h_{m,n}, f > = |a_0|^{-m/2} \int dx \cdot h(a_0^{-m} x - nb_0) \cdot f(x) \quad ; \qquad a_0 > 1, b_0 \neq 0 \quad .$$

If f satisfies the *admissibility condition*

$$\int d\xi \cdot |\xi| \cdot |\hat{h}(\xi)|^2 < \infty \qquad ; \qquad \hat{h}(\xi) = \frac{1}{\sqrt{2\pi}} \int dx \cdot e^{-ix\xi} \cdot h(x)$$

has sufficient decay and T has a bounded inverse on its range, then for some $A > 0$ and $B < \infty$,

$$A||f|| < \sum_{m,n \in Z} | < h_{m,n}, f > |^2 < B||f||^2$$

for all $f \in L^2(R)$. And f may be approximated by its wavelet expansion as

$$f = \sum_{m,n} c_{m,n} h_{m,n} + R \qquad ; \qquad c_{m,n} = < h_{m,n}, f > \cdot \frac{2}{A+B}$$

where

$$||R|| \leq O\left(\frac{B}{A} - 1\right) \cdot ||f||$$

For more on the wavelet-Fourier analogy, see [D2],[Me1],[Ya].

Applications of the wavelet concept to image processing were first decribed in a work by Mallat [Ma1],[Ma2] whose algorithm is closely related to the Laplacian Pyramid Scheme for image processing by Burt and Adelson [BA1],[BA2]. The idea in both schemes is to compress sampled data through a series which yield less and less detail, corresponding to movement from the finer to coarser resolution spaces. Depending on the application, complete restoration of an image may not be necessary in the decomposition process so that only a fraction of the data set and computation steps may be needed. Furthermore, we note that the neighbourhood-based nature of of the algorithms allows for local refinements of an image. Both schemes can also be used for edge detection during the compression process.

On the most primitive level, coding of images takes place on a pixel-by-pixel basis. In very basic (or casual) predictive coding, the value of each pixel to be encoded is predicted using data from previously encoded pixels, and only the error in prediction is stored. Since neighbouring pixels in most images tends to be highly correlated, a prediction scheme based on a symmetric neighbourhood about each pixel is desirable. Unfortunately, neighbourhood-based prediction schemes often cannot be coded in a simple, sequential coding manner, and transform techniques are often used. Burt and Adelson's scheme takes combines the attractive features of predictive and transform methods; "*the predicted value for each pixel is computed as a local weighted average, using a unimodal Gaussian-like (or related trimodal) weighting function centered on the pixel itself. The predicted values for all pixels are first obtained by convolving this weighting function with the image. The result is a low-pass filtered image which is then subtracted from the original.* (In short,) *the technique is noncasual, yet computations are relatively simple and local* [BA1]. A one-dimensional example is given by Burt and Adelson to illustrate how the data about a pixel and its four, symmetrically located neighbours (two on each side) are used to generate a coarser image. Local weights used for "averaging" may be varied to reflect different degrees of correlation in the neighbouring pixels. During the coding process, the coarser representation and the difference between the original and coarser image will replace the original image data.

Mallat recognized that the Laplacian Pyramid Scheme is a special application of the wavelet concept [Ma1],[Ma 2]. The relationship between the finer and coarser resolution spaces is expressed mathematically in the definition of a multiresolution analysis [Ma1][Ma2][D1]: " *A Multiresolution analysis consists of (i) a family of embedded closed subspaces* $V_m \subset L^2(R)$, $m \in Z$; $\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots$ *such that*

$$(ii) \qquad \bigcap_{m \in Z} V_m = \{0\} \quad , \quad \overline{\bigcup_{m \in Z} V_m} = L^2(R),$$

and (iii) $f \in V_m \Leftrightarrow f(2\cdot) \in V_{m-1}$ *moreover, there is a* $\phi \in V_0$ *such that for all* $m \in z$, (iv) $\overline{V_m = \text{linear span}\{\phi_{m,n,}, n \in Z\}}$ *and there exist* $0 < A \leq B < \infty$ *such that for all* $(c_n)_{n \in Z} \in l^2(Z)$, $A \sum |c_n|^2 \leq \| \sum c_n \phi_{mn} \|^2 \leq B \sum |c_n|^2$. *Here* $\phi_{mn}(x) = 2^{-m/2}\phi(2^{-m}x - n)$. " The multiresolution analysis we use in image processsing is $V_m = \{\phi\}$, where $\phi \in L^2(R)$, nonzero constant on $[2^m n, 2^m(n+1)[$,$\forall n \in Z$ and zero otherwise (as described earlier). When one passes from a fine to a coarser resolution space, the data is, in a sense, 'averaged', and the information which is lost can only be retrieved by storing it as coefficients of basis elements for the space orthogonal to the coarse space. For the example above, we choose $\psi(x) = \phi(2x) - \phi(2x-1)$ to be this basis set. In image processing we may also use the $\psi(x)$ for edge detection.

Slides illustrating compression and edge detection for two- and three-dimensional images using a Haar-wavelet based algorithm is shown in figures 1 and 2. In our two-dimensional experiments, a $2^n \times 2^n$ pixel image is compressed to a a $2^{n-1} \times 2^{n-1}$ pixel image by averaging the four neighbouring pixels $(2^j - 1, 2^k - 1)$, $(2^j - 1, 2^k)$, $(2^j, 2^k - 1)$ and $(2^j, 2^k)$, for $j, k = 1, \ldots, n$ in a $2^n \times 2^n$ image data matrix. We store the compressed data in a matrix of size equivalent to that for our original, uncompressed data. The compressed image is given on the upper left and the horizonal, vertical and diagonal edge detection data on the upper right, lower left and lower right respectively. Results from a second compression are shown, from which one can see how data from succesive steps are stored in an analogous manner.

Preliminary results from some 3D image coding experiment are given in figure 2. The storage scheme is simply an extension of the two-dimensional case; one compressed image data set and seven edge detection sets are produced during each compression step. They are stored in octants, rather than quadrants. Generation of the images from the compressed data takes place in a matter of seconds.

The images in figure 1 were generated using image display software by Mr. Ioka [I1],[I2] of IBM and an IBM 6090 graphics terminal attached to a Canon Pixel Dio color copier using a prototype converter box. The images in figure 2 were produced using image display software by Mr. Miyazawa [Myz1],[Myz2] of IBM and an IBM RS6000 graphics terminal attached to a Canon Pixel Dio color copier using a prototype converter box. Because of the experimental nature of this set-up, some limitations on the photocopying capabilities of the screen images became apparent. Some of the screen details for the edge detection data were not captured by the Canon copier. A test to show inconsistencies in the screen and copier images was run. Every other pixel was colored black with the others set to white. Very fine, black lines could be seen on the screen, which looked like a light grey from far away. The copier produced oscillating patterns of dark and light bands, as shown in figure 3. The top figure is a "regular" A4-size copy of the screen.

The bottom is a 90-degree rotated A4(R)-size copy.

In the near future I would like to experiment with a variety of 3D medical and 2D video data using different types of wavelets.

Acknowledgements I would like to thank Mr. Ioka and Mr. Miyazawa of the IBM Tokyo Research Laboratory for allowing me to use their programs and their patient help.

References

[**BA1**] P.J. Burt, E.H. Adelson, IEEE Trans. Comm. 31(4) (Apr. 1983) 532-540.

[**BA2**] P.J. Burt, E.H. Adelson, ACM Trans. Graph. 2(4) (Oct. 1983) 217-236.

[**CGT**] J. M. Combes et al., eds., Wavelets, Proc. Marseille 1987, Springer (1989).

[**D1**] I. Daubechies, Comm. Pure Appl. Math. 41(1988), 909-996.

[**D2**] I. Daubechies, IEEE Trans. Inf. Theory 36(1990), 961-1005.

[**FJW**] M. Frazier, B. Jawerth, G. Weiss, CBMS Reg. Conf. Ser. 79, AMS (1991).

[**I1**] M. Ioka, DIS90C FORTRAN A, IBM Tokyo Scientific Center.

[**I2**] M. Ioka, DIS90G FORTRAN A, IBM Tokyo Scientific Center.

[**Ka1**] H. Kawahara, demo, Nov. 14, 1991 SIAM Japan Wavelet Seminar

[**Ka2**] H. Kawahara, SIAM Japan Wavelet Seminar textbook, Ch.4 (1991) (Japanese).

[**KK**] M. Kobayashi, K. Koyamada, talk, Workshop, Tokushima U., (Aug. 1991).

[**Koi**] A. Koide, MolWorld (on IBM PS55), CU/6000 (on IBM RS/6000).

[**Kol**] G. Kolata, NY Times (Nov. 12, 1991) Science Section B1 and B9.

[**KM**] K. Koyamada, T. Miyazawa, IPSJ 32(5) (May 1991) 560-569.

[**Le**] P. G. Lemarie, ed., Les Ondelettes en 1989, Springer Math. 1438 (1990).

[**Ma1**] S.G. Mallat, GRASP Lab Report 103 MS-CIS-87-22 (May 1987).

[**Ma2**] S.G. Mallat, GRASP Lab Report 80 MS-CIS-87-87 (Sept. 1987).

[**Me1**] Y. Meyer, Ondelettes et Operateurs I, Hermann (1990) (French).

[**MIMY**] T. Miyazawa et al, Proc. IPSJ Graph. CAD Symp. (1989) 59-69.

[**Myt1**] M. Miyata, Trans. IEICE E72(6) (June 1989) 763-770.

[**Myt2**] M. Miyata *Edo* (video), Siggraph Ann. Conf. (1990).

[**Myt3**] M. Miyata, Computer Graphics 24(46) (Aug. 1990) 387-394.

[**Myz1**] T. Miyazawa, DISPIMG, IBM Tokyo Scientific Center (1991).

[**Myz2**] T. Miyazawa, VOXELVIEWER, IBM Tokyo Scientific Center (1991).

[**Ru**] W. Rudin, Real and Complex Analysis, McGraw-Hill (1966).

[**S**] M. Satoh (ATR), remarks, AVIRG Seminar (Jan. 1992).

[RV] O. Rioul, M. Vetterli, IEEE Signal Processing Magazine (Oct. 1991) 14-38.

[Ya] M. Yamaguti, SIAM Japan Wavelet Seminar textbook, Ch.1 (1991) (Japanese).