EE 123: Digital Signal Processing	**Spring 2007**

Lecture 31 — May 8

Lecturer: Prof. Anant Sahai *Scribe: Steffen Prince*

31.1 Outline

These notes cover the following topics:

- Conventional image encoding
- 1D wavelet and filter bank review
- 2D (separable) wavelets & filter banks
- Lossy encoding
- Problems with an orthogonal basis

31.2 Conventional Image Encoding

Before delving into 2D wavelets and JPEG2000, a survey of traditional image encoding is helpful. One of the oldest image encoding schemes used is the raster scan, for analog television broadcast. In a raster scan, an image is represented as rows of horizontal lines in a single vector. To reassemble the image, the data is "raked:" the rows are selected from the data and aligned vertically to form the 2D picture. For compression, the raster scan has a major disadvantage in that the image is represented as a 1D vector; therefore, compression cannot take advantage of similarities in adjacent pixels on different rows. This idea of relations between spatially adjacent pixels is called locality.

One of the most ubiquitous forms of image encoding on the Internet is the JPEG standard, which debuted in 1992. JPEG uses the discrete cosine transform (DCT) as its basis. It divides an image into 8x8 pixels blocks which it then compresses with a DCT. In the past, this approach was perfectly adequate since small images were the primary source. However, as the source image becomes larger, the 8x8 blocks do not capture enough detail for meaningful compression. As a consequence, using high compression rates on large images results in blocking artifacts.

The weaknesses of JPEG suggest the desire to capture locality without specifying the scale of locality a priori. The JPEG2000 standard, using a wavelet approach, does not rely on fixed blocks and therefore overcomes this limitation of JPEG. It also has other advantageous properties, but to see those, basic wavelet transformations must be discussed first.

31.3 1D Wavelet and Filter Bank Review

In practice, wavelets are implemented with filter banks. Filter banks iteratively decompose a signal into high- and low-frequency bands. This is illustrated in Fig. 1. The input signal is sent to two filters: a highpass filter and a lowpass filter. The outputs from these filters are then decimated. At this stage, it is easy to see no information is lost: the input signal has been decomposed into two signals of half the sampling rate containing the high- and low-frequency information of the input. In wavelet transformation, this first high-frequency band composes the finest scale information of the output. In this context, scale is the concept of picture information at different levels of detail. The decomposition proceeds by repeatedly extracting the high-frequency information, saving it, and sending the low-frequency band to the next step of filtering and decimation. This process is repeated for a finite number of levels.

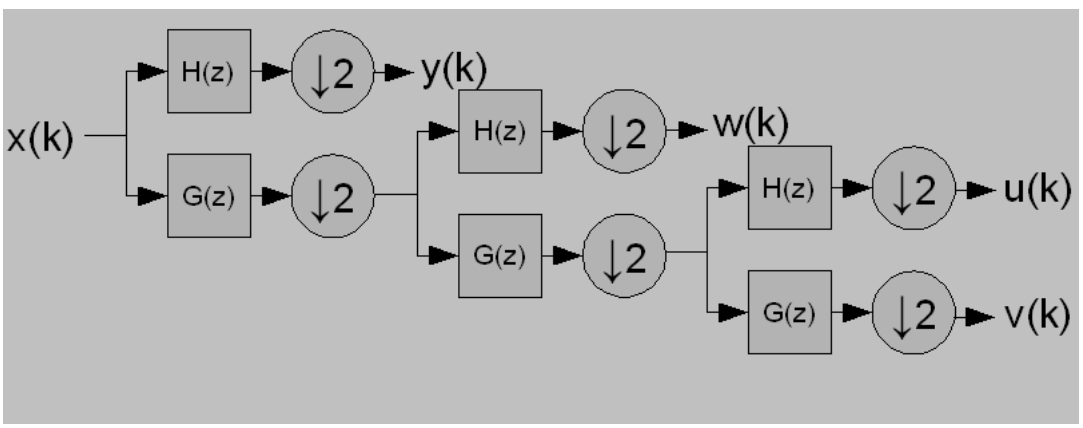

Fig. 1

The output of the filter bank in Fig. 2 is formed by placing the coarsest-scale piece first and appending the next finer-scale piece until the finest-scale piece is placed at the end. From Fig. 1, the output of the last lowpass branch is placed first, while the output of the first highpass branch is placed last. In this way, the original signal can be reconstructed to the desired level of detail without doing any reconstruction from higher detail pieces.

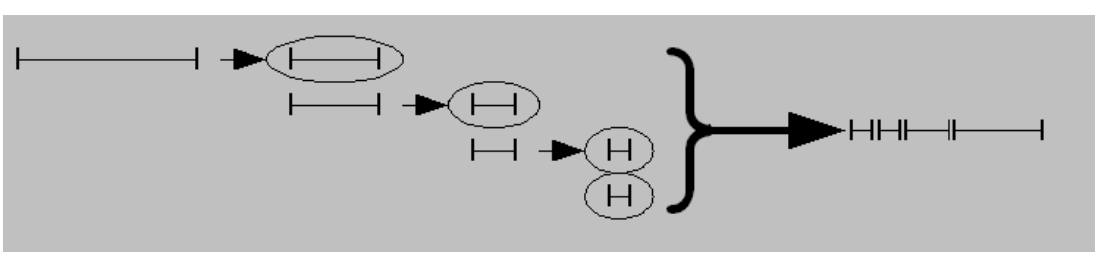

Fig. 2

If the filter bank uses orthogonal wavelets, the transform can be viewed as projecting the input signal onto a set of orthogonal basis functions. If, in addition, the filters are normalized, the resulting coefficients have the same total energy of the input signal.

31.4 2D (Separable) Wavelets and Filter Banks

The 1D wavelet transform can be applied twice to form a 2D wavelet transform. First, the rows of the input are transformed, then the columns. This approach only works if the filters are separable: that is, if the filter transfer functions are of the form $H(z_1, z_2) = H_1(z_1)H_2(z_2)$. Thus in 2D, the wavelet transform has 4 stages for every scale: filtering and decimation along the rows and then along the columns. To extend the transform to color, the image can be decomposed into its RGB components, essentially reducing the image to 3 separate grayscale images.

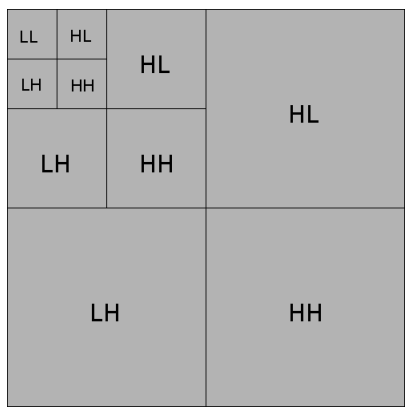

Fig. 3

The above figure shows how an image is successively decomposed into high- and low-freqency bands horizontally and vertically. The first letter refers to the horizontal component, the second to the vertical. Thus HL is a block containing a high-frequency horizontal band and a low-frequency vertical band.

31.5 Lossy Encoding

The goal of lossy encoding is to obtain good performance at high compression. Good performance here is defined as little perceived difference from the original. Compression simply refers to the number of bits per pixel needed for storage: higher compression means fewer bits per pixel.

Wavelets offer a way of culling detail from an image to reduce the amount of information to be stored. However, the wavelet coefficients themselves are compressable using lossless compression. From information theory, data is highly compressed if each bit looks independent of bits near it. Bits in long strings of 0's or 1's are obviously not independent of one another, and there are various lossless compression algorithms that can eliminate this. Another measure of compression is weight: the ratio of 1's to 0's. Ideal lossless compression has weight close to 1.

An important empirical observation of wavelet decomposition relevant to compression is that sharp edges and edgeless regions show up throughout the decomposition tree (Fig 3). Thus, an appropriate compression strategy would be to place all pieces of the wavelet decomposition into three categories:

1. Interesting
2. Boring
3. Temporarily boring

Boring means the piece *and its children* are uninteresting, and omitting them in the compression will not make much of a difference. Temporarily boring means that although one piece is uninteresting, its children may be interesting.

Now that the pieces meriting bits have been specified, a prioritization should be made. First, comparisons are made and the pieces deserving only one bit are selected. Then, pieces deserving two bits are selected, and so on.

31.6 Problems with an Orthogonal Basis

Until now, the filters used have not been specified. In practice, these filters are FIR. In order to avoid a large number of coefficients, the convolution result is truncated. This means, however, that the image must be extended in some way. A periodic extension (corresponding to circular convolution) results in edge artifacts, since the extensions may introduce large jumps. A symmetric extension is a better choice, since it guarantees continuity at the borders of the extensions. The cost is a doubling of the number of input samples.

There is still another problem. After filtering and decimation, the output is still not necessarily periodic, so no redundant coefficients can be eliminated. The solution is to impose the constraint of linear phase on the filter (symmetric or anti-symmetric). However, only the Haar basis is orthogonal and linear phase. Experimentally, the Haar basis is not very good for image compression. So some orthogonality must be sacrificed. Biorthogonal wavelets have some orthogonality relationships between their filters, but are not energy preserving. For JPEG2000, the Cohen-Daubechies-Feauveau wavelets are used for lossy compression.