Wavelet-Based Image Segment Representation

Liu Ying, Surendra Ranganath, Xiaofang Zhou

An efficient representation method for arbitrarily-shaped image segments is proposed. This method includes a smart way to select wavelet basis to approximate the given image segment, with improved image quality and reduced computational load.

Introduction: An important issue in object-based image coding is the efficient description of image segments having arbitrary shapes [1]. Kaup described an algorithm based on Matching Pursuits (MP) [2] to represent arbitrarily shaped image segments by selecting a small set of discrete cosine basis from a larger basis set defined on a rectangle circumscribing the region[3] (we denote it as MP-DCT). Due to the good energy compaction property of the discrete wavelet transform (DWT), we used this basis in the MP framework to formulate the MP-DWT. And we found an efficient basis selection method by restricting the basis functions in an area defined by Coefficient Selection Mask (CSMask). Simulation results showed that our algorithm gives good approximation of the given segment. To further reduce the computational load, smaller basis selection ranges are defined.

Successive Approximation Using Wavelet Basis: Given an image segment $f(m,n)$ containing M pixels. We start from a set of wavelet basis functions Φ defined on a rectangular region L circumscribing the given image segment A, $\Phi = \{\varphi_{kl}(m,n) \mid 0 \le k < M_0, 0 \le l < N_0, (m,n) \in L\}$, L is of size N=M_0*N_0. With a total of N=M_0*N_0 basis functions available, we aim to approximate $f(m,n)$ using only a small set of these basis functions by an iterative procedure. At each iteration, the basis function which best matches the residual signal is selected.

Suppose $g^{(v)}(m,n)$ is an approximation to $f(m,n)$ in the v^{th} step. According to linear approximation theory,

$$g^{(v)}(m,n) = \sum_{(k,l)\in K_v} c_{kl}^{(v)} \varphi_{kl}(m,n) \quad (1)$$

where K_v denotes the set of basis function indices used in the expansion for $g^{(v)}(m,n)$. The residual between the given image segment and the approximation, $r^{(v)}(m,n)$, is then approximated by selecting the basis function $\varphi_{pq}(m,n)$ which maximizes ΔE_A^v in (3)

$$r^{(v)}(m,n) = f(m,n) - g^{(v)}(m,n) \quad (2)$$

$$\Delta E_A^{(v)} = \frac{\left[\sum_{(m,n)\in A} r^{(v)}(m,n)\varphi_{pq}(m,n)\right]^2}{\sum_{(m,n)\in A} \varphi^2_{pq}(m,n)} \quad (3)$$

and the weight is given by

$$\Delta c_{pq}^{v} = \frac{\left[\sum_{(m,n)\in A} r^{(v)}(m,n)\varphi_{pq}(m,n)\right]}{\sum_{(m,n)\in A} \varphi^2_{pq}(m,n)} \quad (4)$$

The corresponding coefficient c_{pq} is then updated as

$$c_{pq}^{(v+1)} = c_{pq}^{v} + \Delta c^{v}_{pq} \quad (5)$$

The set of basis functions K_v is then expanded by the just selected basis function $\varphi_{pq}(m,n)$ if this function is not yet included in the list, and the new approximation function $g^{(v+1)}(m,n)$ is determined by evaluating (1).

The above steps are repeated for v =1,2,3…t, until the error energy drops below a pre-specified threshold or until it reaches a certain number of iterations. Then $g^{(t)}(m,n),(m,n)\in A$ approximates the image segment and provides an extrapolation for $(m,n)\in L\backslash A$. To use fast

DWT, M_0, N_0 are selected as the smallest powers of two to just enclose the segment. Initially, $f(m,n)$ is extended into $r^{(0)}(m,n)$ by zero padding [3].

Coefficient Selection Mask (CSMask): In the above procedure, if a wavelet basis function has a very small overlap region with *A,* the energy of the basis function inside *A* will be very small, resulting in very high value of (4). If such a basis is selected, the extrapolated region of the reconstructed image segment $g^{(v)}(m,n)$ will be highly unstable. To rectify this problem, we propose coefficient selection mask (CSMask). By selecting basis within CSMask only, we give priority to the basis that have larger overlap with *A*. Experimental results show that the extrapolation results are significantly improved (refer to Fig. 4(b)).

The specification of CSMask is illustrated by considering the segment in Fig.1(a). We first define a binary 'mask' image which has value '1' in *A*, and '0' in *L\A*, as shown in Fig.1(b). This is used to specify a set of masks at resolutions corresponding to various subbands in the wavelet decomposition. For example, a sub-mask of size *N*/4 is obtained for wavelet level 1 with one pixel for every 2*2 block in the 'mask'. For blocks containing all zero pixels, the corresponding pixel in sub-mask is set to '0'; otherwise, it is set to '1'. This yields the binary sub-mask as shown in Fig.1(c). Four of these sub-masks are used together to describe the segment shape in the 4 subbands from one level of wavelet decomposition, shown in Fig.1(d).

This process is repeated iteratively on the submask in the upper left hand corner of Fig.1(d) to generate a set of masks describing the segment shape in different subbands. An example is shown in Fig.1(e) to describe segment shape for 3-level wavelet decomposition. All the masks as in Fig.1(e) are collectively called CSMask. Thus, CSMask is related to the shape of the image segment as well as to the wavelet decomposition depth.

Smaller Basis Selection Ranges: Due to the order in which basis functions are selected, we investigated restricting basis selection to low-frequency regions only at first and then to higher frequency regions. This was with a view towards reducing the computations further.

Let the decomposition level be $L_d \geq 3$. We classify the DWT coefficients into several frequency regions as defined below: Region R0, above decomposition level 2, as shown in Fig. 2(a); Region R2, the high frequency region in level 2 decomposition, Fig. 2(b); Region R1, the high frequency region from level 1 decomposition, Fig. 2(c).

Several basis selection areas were defined for experimentation. Range0: Unrestricted basis selection over the entire rectangular area; Range1: R0+R2, R2+R1. Here the first M/4 basis functions are chosen from the regions R0 and R2, others are selected from the region R2 and R1; Range2: R0+R2, R1. Here the first M/4 basis are chosen from R0 and R2, others are selected from R1; Range3: R0,R2,R1. Here the first M/16 basis are chosen from R0, the next 3M/16 are selected from R2, and the rest are selected from R1.

Using smaller basis selection ranges can reduce the computational load in basis selection. For the eye segment in Fig.1(a), with 5-level DWT using filter Db4, and 300 iterations, the computational load of basis selection are shown in Table1.

Combining CSMask and the smaller basis selection ranges, simulation results show that from Range0 through Range3, the PSNR gets lower, as restricting basis selection into smaller range may cause the loss of some useful information. We found that Range3 is unacceptable, while for Range0, Range1 and Range2, a trade-off can be sought between computational load and quality of reconstruction. Fig. 3 shows the results.

Comparison of MP-DWT and MP-DCT: We compared MP-DWT with MP-DCT, and found that: 1) MP-DWT provides better approximation result possibly because of its multiresolution nature. Fig.4(a) compares the PSNR values. 2) Without CSMask, extrapolation results of MP-DWT is less smoother than that of MP-DCT, but with CSMask, this disadvantage is overcome. Fig.4(b) shows some results. 3) Computational load of MP-DWT is lower. Suppose the given image segment contains M pixels and the extended rectangle is of size $N = N_0 * N_0$, f_L is the length of wavelet filter used. In one iteration, the computational load of MP-DWT is $o(N_0{}^2 f_L)$, that of MP-DCT is $o(N_0{}^2 \log_2 N_0)$.

References

1 Raj Talluri, et.al., “A Robust,Scalable, Object-Based Video Compression Technique for Very Low Bit-Rate Coding”. IEEE Trans. on Circuits and Systems for Video Technology, Vol 7, No.1, pp221-233, Feb. 1997.

2 S.G. Mallat, “Matching pursuits with time-frequency dictionaries”. IEEE Trans. Signal processing, vol 41, pp3397-3415, Dec. 1993.

3 A. Kaup, T.Aach, “Coding of Segmented Images Using Shape-Independent Basis Functions” Trans. on Image Processing, vol 7, No. 7, pp937-947, July 1998.

Author’s affiliations:

Liu Ying, Xiaofang Zhou (School of Information Technology &Electrical Engineering, University of Queensland, Queensland, 4072, Australia)

Ranganath Surrendra(Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 117576)

E-mail Address: liuy@itee.uq.edu.au

Figure captions:

Fig. 1 Procedure to get CSMask

Fig. 2 Smaller basis selection ranges
(The gray-color region is the region where basis selection will be done)

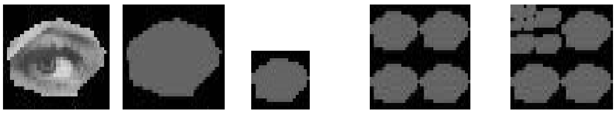
Fig. 3 Extrapolation results by combining CSMask with smaller basis selection ranges
(face segment, 6-level DWT with filter Db6, 300 iterations)

Fig. 4 Comparison of MP-DWT and MP-DCT

(a) PSNR values of MP-DCT and MP-DWT
(b) Extrapolation results of MP-DCT and MP-DWT

Table 1 Comparison of computational load in basis selection using different ranges

Fig. 1



(a) eye (b) Mask (c)SubMask (d) CSMask1 (e) CSMask

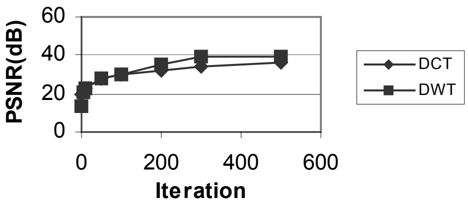
Fig. 2

(a) R0 (b) R2 (c) R1

Fig. 3

36.50dB Range0 | 35.44dB Range1 | 33.70dB Range2 | 22.65dB Range3

Fig. 4



(a)

MP-DWT without CSMask MP-DCT MP-DWT with CSMask

(b)

Table 1

Selection Ranges	Range0	Range1	Range2	Range3
Computational Load	$4.494*10^5$	$2.863*10^5$	$2.390*10^5$	$2.208*10^5$