1000-9825/2001/12(05)0645-05 ©2001 Journal of Software 软件学报 Vol.12, No.5

Improvement of λ-Iterate Bit Allocation Algorithm in Wavelet-Based Image Coding*

CHEN Yi-song, XIAO Rong, SUN Zheng-xing, ZHANG Fu-yan

(*Department of Computer Science and Technology, Nanjing University, Nanjing* 210093, *China*);
(*State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing* 210093, *China*)
E-mail: yschen@graphics.nju.edu.cn
http://www.nju.edu.cn

Received January 21, 2000; accepted July 3, 2000

Abstract: This paper gives an improved algorithm of λ-iterate bit allocation algorithm in wavelet-based image coding, which is based on the statistical characteristic of sub-band block after wavelet transform. The algorithm uses the concept of λ-sequence to simplify the calculation in search for optimal value and significantly speeds up the convergence process of the algorithm.

Key words: wavelet; image coding; bit allocation strategy; objective function; λ-iterate algorithm

Wavelet-based image coding has received popular attention in recent years[1~3]. In a wavelet-based image coding application where the compression ratio is restricted, the efficient bit allocation algorithm given by Yair Shoham and Allen Gersho[4] (called λ-iterate algorithm later) is often used to do quantization and bit allocation work for image blocks after wavelet transform[5]. In a concrete coding process based on the characteristic of blocks after discrete wavelet transform, we propose in the paper an improved algorithm to reduce the time needed to find an optimal bit allocation method.

1 λ-Iterate Algorithm

Wavelet-based image coding could be realized through 2-dimentional Mallat algorithm[6]. Figure 1 is an explanation of the 3-layer decomposition of an image. Each LL_j, LH_j, HL_j or HH_j block represents the layer decomposition in different scales, directions and frequency bands. A compression is realized by doing different quantization and bit allocation work for different blocks.

A common way of finding an efficient quantization and allocation scheme is the method of objective function[6,7], that is, the best scheme is got when an objective function reaches its extremum. A general mathematical description of the question is as follows.

Let S be the finite set of all admissible allocation vectors. Let $H(B)$ be some real-valued function called the

* Supported by the National Natural Science Foundation of China under Grant No. 69903006 (国家自然科学基金)

CHEN Yi-song was born in 1973. He received the B.S. degree from Xi'an Jiaotong University in 1996. Now he is a Ph.D. candidate in Nanjing University. His research interests are image and video processing. **XIAO Rong** was born in 1972. He received the M.S. degree from Nanjing University in 1998. Now he is a Ph.D. candidate in Nanjing University. His research interests are machine learning and data mining. **SUN Zheng-xing** was born in 1964. He is an associate professor in the Department of Computer Science and Technology, Nanjing University. His research areas are CAD and digital library. **ZHANG Fu-yan** was born in 1939. He is a professor in the Department of Computer Science and Technology, Nanjing University. His research areas are multimedia, CG and CAD, and digital library.

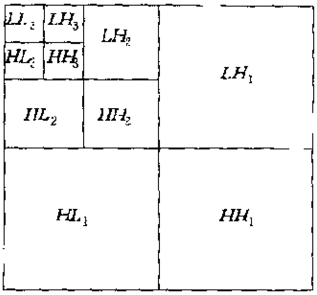

Fig. 1 An explanation of a 3-layer decomposition of an image

objective function of B, defined for all allocation vectors B in S. Consider the following problem (called the constrained problem).

Given a constraint R_c, find

$$R(B) \leqslant R_c \tag{1a}$$

subject to

$$H(B^*) = \min_{B\in S}\{H(B)\} \tag{1b}$$

λ-iterate algorithm could be used to find the optimal strategy. The algorithm is based on the following theorem:

Theorem. Find any $\lambda \geqslant 0$, the solution $B*(\lambda)$ to the unconstrained problem

$$\min_{B\in S}\{H(B) + \lambda R(B)\} \tag{2}$$

is also the solution to the constrained problem (1) with the constraint $R_c = R(B*(\lambda))$, that is, with $R(B) \leqslant R(B*(\lambda))$. To simplify notation, we define $R*(\lambda) = R(B*(\lambda))$.

What the theorem says is that to every nonnegative λ, there is a corresponding constrained problem whose solution is identical to that of the unconstrained problem. Then, if we find a proper λ, we also find the optimal bit allocation strategy.

In fact, there exists an efficient λ-iterate algorithm to find the λ desired, as described in Ref. [1].

2 Improved Algorithm

In the application of wavelet-based image coding, assume the source image is divided into k frequency blocks, then (2) is converted to the following formula,

$$\min_{B\in S}\left\{\sum_{k=1}^{M} W_k(b_k) + \lambda \sum_{k=1}^{M} b_k\right\} = \min_{B\in S}\left\{\sum_{k=1}^{M} (W_k(b_k) + \lambda b_k)\right\} \tag{3}$$

where M is the number of the blocks and $W_k(b_k)$ is the objective function (normalized quantizer function) of the k-th block in some allocation scheme. b_k is the allocated bit number for the block. We can see from formula (3) that to make the whole objective function take minimal value, the unique objective function for each block should be minimized.

For different λ, finding an optimal value in each possible quantization scheme means finding the solution of the following formula.

$$\min_{i\in\{1..n\}}\{f_i \mid f_i = D_i + \lambda R_i\} \tag{4}$$

n is the number of the possible quantization schemes, D_i is the objective function of block i in some allocation scheme and R_i is the bit number for the block in the scheme.

Formula (4) is a regular straight line cluster in $f-\lambda$ coordinate plane, as shown in Table 1, the data of which are selected from LL block of the image Lena.

We can see from the table and can also mathematically prove[8], with the increment of scheme sequence number i, R_i increases proximately like an arithmetical progression and D_i decreases like a geometric progression. In $f-\lambda$ coordinate system, the cross point λ_i of two straight lines is got by the following formula:

$$\lambda_i = \frac{D_i - D_{i+1}}{R_{i+1} - R_i} \approx \frac{(1-r)D_i}{d} \tag{5}$$

d and r are respectively tolerance and ratio of the two progressions. We can see decrease property of λ from the formula. That is, cross point of lines with bigger slope has smaller λ value, as shown in Fig. 2. We can see from the figure that the envelope of the line clusters is constructed by a sequence of seams, which gives the final solution of formula (2). We can improve the algorithm based on above property.

Table 1 Relationship of i, N_i, R_i and D_i

Scheme i	Quantize phase N_i	Bit allocated R_i	Objective function D_i
0	NQuant[0]=0	Rate=0.000000	Dist=66556338.738506
1	NQuant[1]=1	Rate=60.145522	Dist=26416724.417641
2	NQuant[2]=3	Rate=119.808869	Dist=5790731.454754
3	NQuant[3]=7	Rate=186.742111	Dist=1711893.428406
4	NQuant[4]=15	Rate=254.826522	Dist=505270.760397
5	NQuant[5]=31	Rate=324.426691	Dist=112349.159275
6	NQuant[6]=63	Rate=390.334164	Dist=27223.825644
7	NQuant[7]=127	Rate=459.226389	Dist=6309.574064
8	NQuant[8]=255	Rate=522.643054	Dist=1630.095237
9	NQuant[9]=511	Rate=588.304833	Dist=400.401301

First, formula (4) is represented as a simple curve cluster and λ value of all the cross points can be calculated easily when doing preprocessing. It is clear that λ value is a decreasing sequence when scheme sequence number i increases, as shown in Fig. 2. We call this sequence as the λ sequence of the block. And for a single block, since we care only about scheme code but not a concrete λ value, what we need to do is just to compare current λ with pre-calculated λ sequence of the block and stop searching when finding first sequence value that is smaller than current λ. This means we get our optimal solution.

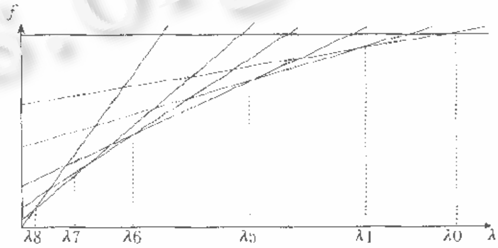

Fig. 2 Description of λ sequence

Second, because a bilateral approach algorithm is used in λ-iterate process, an optimal value for previous λ could help to decide the initial value of current λ. Due to the decrease property of λ sequence, for every single block, the optimal scheme code for current λ is surely between current λ high and λ low. So search can be done on scheme sequence number satisfying the following formule. This speeds up the process more.

$$\lambda_{low} \leqslant \lambda_i \leqslant \lambda_{high} \tag{6}$$

All the above discussion is from the angle of purely searching process of single λ. From the angle of reducing iteration times, we can still accelerate the iterative process. The fact that most energy of source image is in the block LL means that it needs more bits for quantization. So taking a λ sequence with relatively high precision of quantization is often a good choice.

3 Evaluation

The following table gives a comparison between customary and improved algorithms. In our experiment we use 256-gray level image of 256 * 256 or 512 * 512 and do a wavelet decomposition of 5 layers. The data quantity is relatively small and the algorithm is rapid. The time value in the table is the sum of 5000 iterations.

Table 2 Experimental result

Image	Constrained ratio	Original time (seconds)	Improved time (seconds)
Lena (256 * 256)	8	13.77	2.20
Pep (256 * 256)	16	14.31	2.03
Lena (512 * 512)	32	12.85	2.09
Barbara (512 * 512)	64	12.52	2.21

We can see from the data in the table that the improved algorithm greatly speeds up iterative process with a gain of about 5 to 7 times, while coding effect (compression ratio & PSNR) is exactly the same as before, as can be seen in Fig. 3. Experiments on many other images give the similar result.

(a) Original image (b) Rebuilt image

Fig. 3 Original and rebuilt versions of the image Lena (512 * 512) compression ratio=32, PSNR=32.77

There is potential to improve the algorithm further. For example, using predictive technique in λ-iteration process may still reduce the iteration time.

4 Conclusion

The improved λ-iterate algorithm proposed in the paper significantly speeds up the process of bit allocation in wavelet-based image coding. If combined with rapid algorithm for other phases of coding, it may greatly reduce coding time. The algorithm can be used in not only image coding but also many other set allocation application fields, provided the set of quantizers and quantized items has the same data property as described in the paper, such as those described in Refs. [9,10].

References:

[1] Servetto, S.D., Ramchandran, K., Orchard, M.T. Image coding based on morphological representation of wavelet data. IEEE Transcations on Image Processing, 1999,8(9):1161~1174.

[2] Li, J., Lei, S. An embedded still image coder with rate-distortion optimization. IEEE Transcations on Image Processing, 1999,8(7):913~924.

[3] Tran, T.D., Nguyen, T.Q. A progressive transmission image coder using linear phase uniform filterbanks as block transforms. IEEE Transcations on Image Processing, 1999,8(11):1493~1507.

[4] Shoham, Y., Gersho, A. Efficient bit allocation for an arbitrary set of quantizers. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1988,36(9):1445~1453.

[5] Geoff, Davis. Baseline Wavelet Transform Coder Construction Kit. 1997, http://www.cs.dartmouth.edu/~gdavis.

[6] Mallat, S. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transcations on PAMI, 1989,11(7):674~693.

[7] Xiong, Z., Ramchandran, K., Orchard, M.T. Space-frequency quantization for wavelet image coding. IEEE Transcations on Image Processing, 1997,6(5):677~693.

[8] Lynch, T.J. Data Compression Techniques and Applications. Belmont, CA: Lifetime Learning Publications, 1985.

[9] Ramchandran, K., Vetterli, M. Rate-distortion optimal fast thresholding with complete JPEG/MPEG decoder compatibility. IEEE Transactions on Image Processing, Special Issue on Video Sequence Compression, 1994,3(5):700~704.

[10] Eleftheriadis, A., Anastassiou, D. Optimal Data Partitioning of MPEG-2 Coded Video. In: Proceedings of the 1st IEEE International Conference on Image Processing. Austin, Texas, 1994. I.273~I.277.

基于小波变换的图像编码中λ迭代比特分配算法的改进

陈毅松, 萧 嵘, 孙正兴, 张福炎

(南京大学 计算机科学与技术系,江苏 南京 210093)

(南京大学 计算机软件新技术国家重点实验室,江苏 南京 210093)

摘要:基于小波图像编码的金字塔式小波分解后各子带系数的统计特性,提出一种改进的λ迭代比特分配算法,该算法使用λ序列的概念简化了传统λ迭代法的最优值求解过程.实验证明该改进算法在限定压缩比的小波图像编码中能够大大加快比特分配算法的收敛速度.

关键词:小波;图像编码;比特分配策略;目标函数;λ迭代算法

中图法分类号:TP391 文献标识码:A