IMAGE COMPRESSION USING VARIABLE BLOCKSIZE VECTOR QUANTIZATION BASED ON RATE-DISTORTION DECOMPOSITION

Ching Yang Wang and Chyuan Huei Yang and Long Wen Chang

Department of Computer Science,
National Tsing Hua University,
Hsinchu 300, Taiwan, ROC
lchang@cs.nthu.edu.tw

ABSTRACT

In this paper, we propose an optimal quadtree segmentation of an image for variable blocksize vector quantization(VBVQ) such that the total distortion of the reconstructed image is minimal and the total required bits don't exceed the bit budget. The above constrain problem is converted into an equivalent unconstrained problem by using Lagrange multiplier. We prune the full quadtree by comparing the Lagrangian costs of the parent and four child nodes. If the adjacent subblocks merge into a larger block reduce the Lagrangian cost, these subblocks will be merged. Otherwise, these subblocks will be vector quantized. From our simulation results, we see that the reconstructed image of our proposed algorithm has 1-3 db higher than the fixed blocksize VQ and conventional VBVQ algorithms.

1. INTRODUCTION

The concept of variable blocksize vector quantization (VBVQ) was firstly introduced by Jacques Vaisey and Allen Gersho[2, 3]. The basic idea of VBVQ is that the smooth areas of an image are decomposed into larger blocks and the busy areas are decomposed into smaller areas. In [1, 2, 3, 4], the *block variance* is used to determine whether the block should be further decomposed into smaller subblocks. If the block variance is greater than a predefined threshold, the block is decomposed into smaller subblocks. The process goes on until the all block variances are less than the predefined threshold or the smallest block is reached. To use the block variance as a criterion for image segmentation has the following drawbacks. One drawback is that it only accords to the local activity of the block to determine whether decomposition is needed. Even though the decomposition of the block leads to larger distortion, the block with large block variance is still decomposed. The other is that the VBVQ can achieve variable bit rates by adjusting the threshold values. These threshold values are image dependent. It is difficult to determine these values to achieve the desired bitrate.

In this paper, we propose an optimal quadtree segmentation for VBVQ to compress an image based on rate and distortion consideration. We choose the quadtree segmentation to decompose the image into variable blocks because the overhead of the well-defined data structure is very trivial. For a given bit budget and a set of predefined codebooks, we find the best quadtree segmentation of an image for VBVQ such that the total distortion is minimal and the required number of bits does not exceed the bit budget.

2. THE OPTIMAL QUADTREE SEGMENTATION FOR VBVQ

A *quadtree* is a tree structure which each non-leaf node has four children. It is often used to segment an image into homogeneous blocks in some property such as intensity, color or texture[5]. The image is firstly checked whether it is homogeneous. If not, the image is decomposed into four subblocks. Each subblock is recursively split if it is not homogeneous until the smallest block is reached. In the quadtree the root node represents the entire image and the leaf nodes represents those homogeneous blocks which no more decomposition are needed. The root node is at the depth 0. For an image with size $N \times N$, the maximal possible depth of its corresponding quadtree is $d = \log_2 N$. If all leaf nodes on the depth d, the quadtree is a *full quadtree* and denote it by T. The node n_j^i of T represents the jth node of the depth i of T. There are total 4^i nodes in the depth i. For a node n_j^i of T, its parent node is $n_{\lfloor j/4 \rfloor}^{i-1}$, where $\lfloor . \rfloor$ is flooring function and its four children are $n_{4j}^{i+1}, n_{4j+1}^{i+1}, n_{4j+2}^{i+1}$, and n_{4j+3}^{i+1}, respectively. A pruned sub-quadtree S of T is a sub-quadtree of T with the same root node as the full quadtree T, a relation which we denote by $S \preceq T$. We will also consider the single root node is a pruned sub-quadtree of T. Let $\tilde{S}$ denote the set of leaf nodes of S. The quadtree has the following two properties:

1. For all $n_j^i \in \tilde{S}$, $\bigcup n_j^i$ constructs the entire image.
2. For n_j^i and $n_q^p \in \tilde{S}$, if $i \neq p$ or $j \neq q$ then $n_j^i \bigcap n_q^p = \phi$.

The first property implies that the all leaf nodes can cover the entire image. The second property means that the all leaf nodes don't overlap. Therefore, we only vector quantized the leaf nodes of the corresponding quadtree of the image.

Let us define the following terminologies which are used in this paper.

- T: the full quadtree of an image.
- $S \preceq T$: a pruned sub-quadtree of T.
- $\tilde{S}$: the set of leaf nodes of S.
- n_j^i: the jth node of the depth i of T.
- C: the set of codebooks, $C = \{C_{2^k \times 2^k}; 0 \leq k \leq d\}$, where $C_{2^k \times 2^k}$ is the codebook contains codevectors whose size are $2^k \times 2^k$.
- $D(S)$, $R(S)$: the total distortion and required bits for the quadtree segmentation S for VBVQ, respectively. The distortion is measured by total square error.

Note that the $D(S)$ and $R(S)$ are *liner tree functionals* [6, 7]. That is,

$$D(S) = \sum_{n_j^i \in \tilde{S}} D(n_j^i), \tag{1}$$

and

$$R(S) = \sum_{n_j^i \in \tilde{S}} R(n_j^i). \tag{2}$$

Now, the problem is to find the best quadtree segmentation S^* of an image for VBVQ, such that the total distortion of the reconstructed image is minimal and the total required bits don't exceed the given bit budget. That is, we try to find an optimal quadtree segmentation S^* with minimal distortion

$$D(S^*) = \min_{S \preceq T}\{D(S)\}, \tag{3}$$

such that the total required bits

$$R(S^*) \leq R_{budget}. \tag{4}$$

The above constrained problem of seeking the optimal quadtree segmentation for VBVQ can be converted into an easy equivalent unconstrained problem by merging the distortion and rate through the Lagrange multiplier[8, 9]

$$J(\lambda) = D + \lambda R. \tag{5}$$

Thus the unconstrained problem becomes the minimization of the Lagrangian cost function define as

$$\begin{aligned} J(S^*, \lambda) &= \min_{S \preceq T}\{D(S) + \lambda R(S)\} && (6) \\ &= D(S^*) + \lambda R(S^*). && (7) \end{aligned}$$

The following theorem proves the constrained and unconstrained problem have the same solution space.

Theorem 1 *If the best quadtree segmentation S^* is the solution of unconstrained problem of (7) corresponding to some fixed positive λ and $R(S^*) = R_{budget}$, then S^* is also the solution of constrained problem of (3).*

For the proof of this theorem, interested readers can refer to [8, 10]

The search of the best quadtree segmentation of an image consists of two algorithm. Algorithm I finds the optimal pruned sub-quadtree for a given λ. Algorithm II uses the convex-searching bisection iteration algorithm for finding the optimal λ[8]. The Algorithm I can consider as a subroutine called by Algorithm II. The data structure associated with each node n_j^i of T has the form: $\{R_j^i, D_j^i, J_j^i, Split(n_j^i)\}$, where R_j^i, D_j^i and J_j^i mean the bitrate, distortion and Lagrangian cost of node n_j^i, respectively. $Split(n_j^i)$ is a logical variable which indicates whether the node n_j^i is split or not. $Split(n_j^i) = Yes$ means the node n_j^i is split; otherwise the node n_j^i is not split. That is, the node n_j^i is a leaf node of the quadtree.

Algorithm I: Find the optimal quadtree segmentation for a given λ.

Step 1: Given: an image with $N \times N$ pixels,
the full quadtree depth $d = \log_2 N$,
a Lagrange multiplier, λ.

Step 2: Generate the full quadtree segmentation of the image T.

Step 3: For each n_j^i of T, compute R_j^i and D_j^i
$R_j^i = \log_2 ||C_{N/2^i \times N/2^i}||$,
where $||C_{N/2^i \times N/2^i}||$ means the number of codevectors in the codebook $C_{N/2^i \times N/2^i}$ and
$D_j^i = |n_j^i - c_{i,k}|^2$,
where the $c_{i,k}$ is the nearest codevector in codebook $C_{N/2^i \times N/2^i}$.

Step 4: For the given λ, compute the Lagrangian cost J_j^i of node n_j^i

$J_j^i = D_j^i + \lambda R_j^i$.

Step 5: i=depth-1;

WHILE $(i \geq 0)$

FOR $j = 0, 1, \cdots, 4^i - 1$

IF $J_j^i(\lambda) \leq \sum_{k=0}^{3} J_{4j+k}^{i+1}$

$Spilt(n_j^i) \leftarrow$ **NO**;

ELSE

$Split(n_j^i) \leftarrow$ **YES**;

$R_j^i \leftarrow \sum_{k=0}^{3} R_{4j+k}^{i+1}$;

$D_j^i \leftarrow \sum_{k=0}^{3} D_{4j+k}^{i+1}$;

$J_j^i \leftarrow \sum_{k=0}^{3} J_{4J+k}^{i+1}$;

ENDFOR

i=i-1;

ENDWHILE

The Algorithm I compares the Lagrangian cost J_j^i and the total cost of its four children, $\sum_{k=0}^{3} J_{4j+k}^{i+1}$. If the Lagrangian cost J_j^i is smaller than the total cost of its four children, the block should not be split into four smaller subblocks. Therefore, we prune the child nodes of n_j^i and the $Split(n_j^i) = NO$. Otherwise, we split the node n_j^i, the $Split(n_j^i) = Yes$ and the R_j^i, D_j^i and J_j^i will be updated as in Step 5. This process will continue until the root rode is reached. Note that the total required bits, the total reconstructed distortion and the Lagrangian cost based on the optimal quadtree segmentation $S^*(\lambda)$ are stored in the elements R_0^0, D_0^0 and J_0^0 of the root node after running the Algorithm I.

Algorithm II: Iterating towards the optimal λ.

Step 1: Pick $\lambda_l \leq \lambda_u$

Call Algorithm I to get the optimal segmentation $S^*(\lambda_l)$ and $S^*(\lambda_u)$ which satisfy $R_0^0(\lambda_u) \leq R_{budget} \leq R_0^0(\lambda_l)$.

If it fails to find λ_l and λ_u which satisfy the above inequalities means that the given problem is unsolvable.

IF $(R_0^0(\lambda_u) = R_{budget} = R_0^0(\lambda_l))$

find the optimal $\lambda^* = \lambda_u$; **STOP**;

Step 2: $\lambda_{next} = |\frac{[D_0^0(\lambda_l) - D_0^0(\lambda_u)]}{[R_0^0(\lambda_l) - R_0^0(\lambda_u)]}| + \epsilon$,

where ϵ is small positive number to ensure to get the lower bitrate if λ_{next} is a singular slope value[8].

Step 3: Run the Algorithm I for λ_{next}

IF $(R_0^0(\lambda_{next}) = R_0^0(\lambda_u))$

$\lambda^* = \lambda_u$; **Stop**;

ELSE

IF$(R_0^0(\lambda_{next}) > R_{budget})$

$\lambda_l = \lambda_{next}$;

GOTO Step 2;

ELSE

$\lambda_u = \lambda_{next}$;

GOTO Step 2 ;

ENDIF

ENDIF

After finding the optimal λ^* by Algorithm II, we obtain the best quadtree segmentation of an image for VBVQ.

3. SIMULATION RESULTS

For each tested image contains 512×512 pixels, the maximal depth of the quadtree of the tested image is $d = \log_2 512 = 9$. Therefore, we need 9 codebooks. The codebook $C_{1\times1}$ has 64 codevectors,

$$C_{1\times1} = \{c_{1,k}(m,n) = 4k | 0 \leq m, n < 1, 0 \leq k < 64\}.$$

The codebook $C_{2\times2}$ and codebook $C_{4\times4}$ are generated by LBG's algorithm[11]. The images, LENNA, JET and PEPPER, are used as a training set. Each codebook contains 256 codevectors. The codebook $C_{8\times8}$ to $C_{512\times512}$ don't exist in fact. For these blocksize greater than 4×4 are coded by their block means. Therefore, their corresponding codebooks $C_{8\times8}, \ldots, C_{512\times512}$ are defined as follows.

$$C_{2^i\times2^i} = \{c_{2^i,k}(m,n) = k | 0 \leq m, n < 2^i, 0 \leq k < 256\},$$

where $3 \leq i \leq 9$.

The reconstructed image of JET, which is under the constrain of bit budget $R_{budget} = 0.5bpp$, is shown in Fig. 1(a) and its corresponding optimal quadtree segmentation is shown in Fig. 1(b). We can see that our proposed algorithm doesn't produce obvious block effect and the detail areas are preserved well, such as the number on the fin of the JET can be clearly recognized. Furthermore, the proposed algorithm is also robust when the tested image is outside the training set. The performance of our proposed algorithm is plotted in Fig. 2. However, it is still suffered from the block effect when the bitrate is below 0.3 bpp. Therefore, the post processing to alleviate the block effect will be necessary for low bitrate application[3].

4. CONCLUSION

In this paper, we propose an optimal quadtree segmentation which is based on the rate-distortion consideration for VBVQ to compress still images. Our proposed algorithm can find the optimal quadtree segmentation for VBVQ so that the distortion is minimal and the bitrate is under the bit budget. The simulation results show that the proposed algorithm not only reduces the block effect but also preserve the edge areas very well. It also gets higher PSNR than the fixed blocksize VQ and the conventional VBVQ which used the block mean to segment the image.

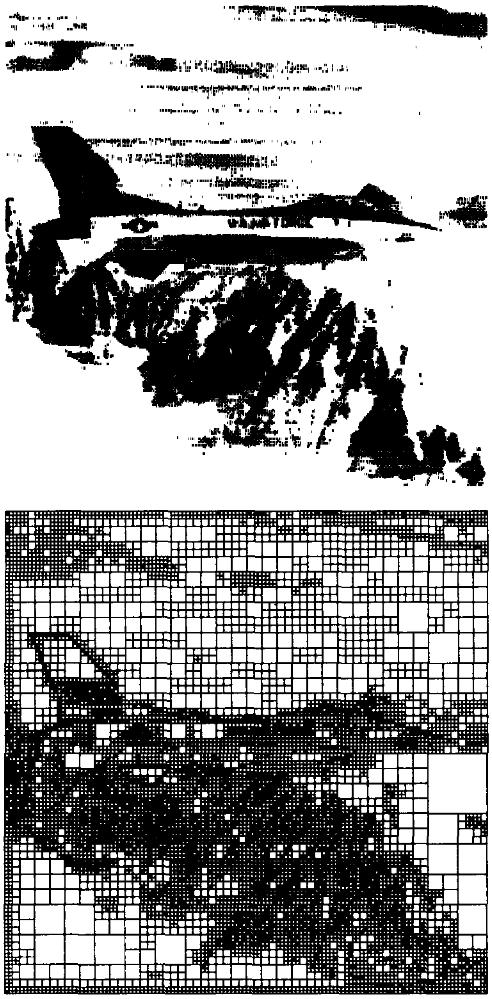

Figure 1: (a)The reconstructed images of JET using our proposed algorithm. Its bitrate is 0.5 bpp and its PSNR is 34.16 db. (b) The optimal quadtree segmentation of JET under the constrain of bit budget $R_{budget} = 0.5bpp$.

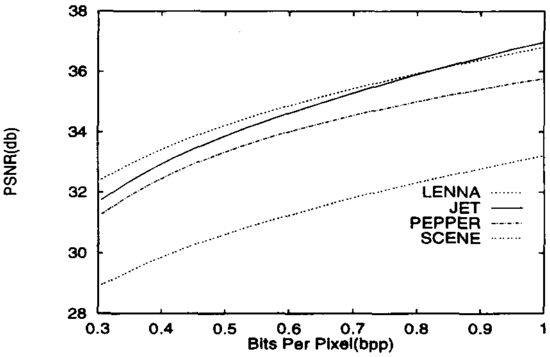

Figure 2: The performance of our proposed optimal quadtree segmentation for VBVQ. The tested image SCENE is outside the training set.

5. REFERENCES

[1] Ching Yang Wang and Long Wen Chang. "Color image coding using variable blocksize vector quantization in (R,G,B) domains". In *Visual Communications And Image Processing*, pages 512–524, Boston, MA, November 1991.

[2] Jacques Vaisey and Allen Gersho. "Variable blocksize image coding". In *ICASSP*, pages 1051–1054, Dallas, TX, April 1987.

[3] Jacques Vaisey and Allen Gersho. "Image compression with variable block size segmentation". *IEEE Trans. Signal Processing*, VOL. 40, NO. 8:pp. 2040–2060, August 1992.

[4] H. Sun and M. Goldberg. "Image sequence coding by adaptive vector quantization with variable block size". *Journal of Visual Communication And Image Representation*, VOL. 2, NO. 1:pp. 91–99, March 1991.

[5] Hanan Samet. "The quadtree and related hierarchical data structure". *ACM Computing Surveys*, VOL. 16, NO. 2:pp. 187–260, June 1984.

[6] Allen Gersho and Robert M. Gray. *"Vector Quantization and Signal Processing"*. Kluwer Academic Publishers, 1991.

[7] Philip A. Chou, Tom Lookabaugh, and Robert M. Gray. "Optimal pruning with applications to tree-structure source coding and modeling". *IEEE Trans. Image Processing*, VOL. 35, NO. 2:pp. 299–315, March 1989.

[8] Kannan Ramchandran and Martin Vetterli. "Best wavelet packet bases in a rate-distortion sence". *IEEE Trans. on ASSP*, VOL. 2, NO. 2:pp. 160–175, April 1993.

[9] Yair Shoham and Allen Gersho. "Efficient bit allocation for an arbitrary set of quantizers". *IEEE Trans. on ASSP*, VOL. 36, NO. 9:pp. 1445–1453, September 1988.

[10] Siu-Wai Wu and Allen Gersho. "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV". *IEEE Trans. on Circuits and Systems for Video Tech*, VOL. 1,NO 1: pp.100–112 Mar 1991

[11] Y. Linde, A. Buzo, and R. M. Gray. "An algorithm for vector quantizer design". *IEEE Trans. Communication*, VOL. 28:pp. 84–95, Jan 1980.