# Optimal Blockwise Dependent Quantization for Stereo Image Coding

Woontack Woo and Antonio Ortega

***Abstract*— Research in coding of stereo images has focused mostly on the issue of disparity estimation to exploit the redundancy between the two images in a stereo pair, with less attention being devoted to the equally important problem of allocating bits between the two images. This bit-allocation problem is complicated by the dependencies arising from using a prediction based on the quantized reference images. In this paper, we address the problem of blockwise bit allocation for coding of stereo images and show how, given the special characteristics of the disparity field, one can achieve an optimal solution with reasonable complexity, whereas in similar problems in motion-compensated video only approximate solutions are feasible. We present algorithms based on dynamic programming that provide the optimal blockwise bit allocation. Our experiments based on a modified JPEG coder show that the proposed scheme achieves higher mean peak signal-to-noise ratio over the two frames (0.2–0.5 dB improvements) as compared with blockwise independent quantization. We also propose a fast algorithm that provides most of the gain at a fraction of the complexity.**

***Index Terms*— Blockwise quantization, dependent bit allocation, stereo image coding, Viterbi algorithm.**

## I. Introduction

MOST research efforts on stereo image/video coding have been devoted to investigating efficient disparity estimation/compensation (DE/DC) schemes to improve the encoding performance [2]–[7]. As in other coding scenarios, stereo images/video can be compressed by taking advantage of spatial/temporal redundancies in each monocular image/video. However, the coding efficiency for stereo images/video can be improved further by exploiting an additional redundancy associated with the similarity between the two images in a stereo pair, i.e., the "binocular" dependency. The central idea of stereo image coding based on DE/DC is to use one of the images in the stereo pair as a reference and to estimate the other image (the target).[1]

Manuscript received April 27, 1998; revised February 15, 1999. This work was supported in part by the National Science Foundation under Grant MIP-9502227 (CAREER), in part by the Integrated Media Systems Center (a National Science Foundation Engineering Research Center), in part by the Annenberg Center for Communication, University of Southern California, and in part by the California Trade and Commerce Agency. Parts of this paper were presented at the SPIE EI-VCIP, January 1998. This paper was recommended by Associate Editor Y. Wang.

W. Woo was with the Integrated Media Systems Center, Department of Electrical Engineering–Systems, University of Southern California, Los Angeles, CA 90089-2564 USA. He is now with Media Integration and Communications Labs, Advanced Telecommunication Research Institute, Kyoto 619-0288 Japan.

A. Ortega is with the Integrated Media Systems Center, Department of Electrical Engineering–Systems, University of Southern California, Los Angeles, CA 90089-2564 USA.

Publisher Item Identifier S 1051-8215(99)07030-5.

[1]Many of the intuitions and techniques used in motion estimation/compensation (ME/MC) are applicable to DE/DC due to the similarities between ME/MC and DE/DC.

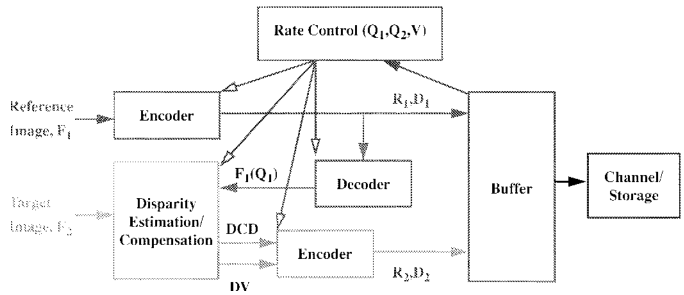



Fig. 1. Block diagram of a general encoder for stereo images, where the encoder consists of disparity estimation/compensation, transform/quantization, and entropy coding.

With few exceptions (e.g., [8]), quantization and bit-allocation issues specific to stereo coding have rarely been considered. Obviously, as shown in Fig. 1, the encoding performance also depends on making a "proper" choice of quantizers $(Q_1, Q_2)$ and not just on the choice of the disparity vectors $(V)$. However, available bit-allocation (or quantization) schemes are mainly developed based on the assumption of completely decoupled encoding steps, e.g., target and reference frames are quantized independently, and thus overall optimality cannot be guaranteed.

Here, we study the problem of optimal bit allocation for stereo image coding. Our proposed bit-allocation scheme is aimed at block-based, rather than segmentation-based, DE/DC techniques due to the comparative simplicity and robustness of block-based techniques. Note that we assume that the disparity vector (DV) field is estimated in "open loop," i.e., based on the original image rather than the quantized version, and then focus on the *quantizer allocation* to the reference and the residue images.[2]

The main novelty of our work is the introduction of an algorithm for *optimal blockwise dependent bit allocation* for stereo image coding. The binocular dependency has to be taken into account because the target image $(F_2)$ is compensated based on the quantized reference image $F_1(Q_1)$. Thus, each choice of quantizer for the reference frame results in different residual energy levels in the difference frame [3]. Given that the epipolar constraint is met,[3] the binocular dependency

[2]Note that techniques developed in rate-distortion (RD)-based ME in video coding [9]–[13] could also be used in conjunction with our algorithm to estimate an optimal (in an RD sense) disparity field.

[3]This constraint implies that the focal rays of the two cameras are parallel and perpendicular to the stereo baseline. Thus, if the cameras meet the epipolar constraint, then the disparity is *exactly* one-dimensional (1-D), i.e., a particular object will appear in the two images with a horizontal shift between its respective positions. Even if the parallel axis constraint is not strictly met (i.e., the disparity is not *exactly* 1-D), blocks in one row in the target image can be predicted fairly accurately from blocks located in the corresponding row in the reference frame because the vertical disparity is confined only to ± a few pixels.

becomes relatively simple, i.e., it occurs predominantly along the horizontal direction. This property not only simplifies the disparity estimation process but also allows us to find an optimal solution for our allocation problem. We first demonstrate how the optimal set of quantizers can be determined using the Viterbi algorithm (VA) and then introduce a novel method that approximates the optimal solution with limited loss in performance but much faster operation.

Note that our results may also provide some ideas for the related problem of blockwise dependent bit allocation in video coding, where choices of quantization for a reference frame affect the frames that are motion predicted from it [14]. In the case of video coding, it is difficult to take into account blockwise dependencies because motion vectors can have any direction in the two-dimensional (2-D) plane.[4] Accordingly, an optimal solution for the video case is not available, and thus much of the work has concentrated on analyses of framewise dependency, i.e., where a single quantizer is allocated per frame [14], [15]. Note that schemes such as [16] have addressed blockwise bit allocation but without taking into account the temporal dependency, i.e., the effect of a particular allocation on future frames.

Our experimental results demonstrate that the proposed scheme provides higher *mean* peak signal-to-noise ratio (PSNR), about 0.2–0.5 dB for the two images in a stereo pair, as compared with an optimal blockwise independent quantization. Note that finer quantization for the reference image tends to allow more efficient encoding for the disparity compensated difference frame [14]. We use this so-called *monotonicity* property as a starting point to propose a fast algorithm that further reduces the computational complexity without significant loss of quality. This blockwise dependent bit allocation can be a benchmark for faster allocation schemes, or can be used in off-line encoding applications or in applications where encoding is performed just once but decoding is performed many times.

## II. Dependent Bit Allocation

### A. Definitions and Notations

Let an image $F_l$ be segmented into $N$ square blocks, $F_l = \{B_m, 0 \le m \le N-1\}$, where $B_m$ represents the $m$th block in the image. In the case of a stereo pair, $F_1$ and $F_2$ denote the reference and the target images, respectively. Blocks in $F_2$ are denoted $B'_m$ to differentiate them from blocks in $F_1$. Let a quantizer (or quantization scale) be assigned to each block (from a finite set of available quantization choices). Then, a set of blockwise quantizers for $F_1$ can be represented as $Q_1 = \{q_m, 0 \le m \le N-1\}$, where $q_m$ denotes a quantization index. Similarly, a set of quantizers for the disparity compensated difference (DCD) frame is represented as $Q_2 = \{p_m, 0 \le m \le N-1\}$. The DV field is defined as $V = \{v_m, 0 \le m \le N-1\}$, where $v_m$ corresponds to the disparity for the $m$th block in $F_2$. In the same way, blockwise rate and distortion can be defined, and the overall rate and distortion are represented as the sum of the individual rates and distortions of the blocks.

In our experiments, we use simple objective measures such as mean square error (MSE) and PSNR. Note that subjective evaluation of three-dimensional quality is still an open problem and is not very reliable and repeatable yet. Therefore, we measure the distortions of $F_1$ and $F_2$ using MSE, i.e., $D_1 = (F_1 - F_1(Q_1))^2$ and $D_2 = (F_2 - \hat{F}_2(Q_1, Q_2, V))^2$, where $F(Q)$ denotes the decoded image when quantizer $Q$ is used. The decoded target image, $\hat{F}_2(Q_1, Q_2, V)$, can be reconstructed by adding the compensated target image with the DV field and the decoded DCD, i.e., $\hat{F}_2(Q_1, Q_2, V) = F_1(Q_1, V) + E(Q_2)$, where $E = F_2 - F_1(Q_1, V)$. We also measure the performance using mean PSNR for the stereo pair, defined as follows:

$$\mathrm{PSNR}_{\mathrm{mean}} = 10 \times \log_{10} \left\{ \frac{255^2}{(D_1 + D_2)/2} \right\} \tag{1}$$

where $D_1$ and $D_2$ denote the MSE's of the reconstructed images $\hat{F}_1$ and $\hat{F}_2$, respectively.

### B. Optimal Blockwise Dependent Quantization

For simplicity, we assume that the quantizer indexes are encoded with a constant number of overhead bits per block. Note that other 1-D dependencies such as those resulting from differential pulse code modulation (DPCM) encoding of quantization indexes could also be incorporated easily into our scheme.

Let $R_{\mathrm{budget}}$ be the remaining bit budget after allocating bits to the DV field. For a given DV field $V$, the optimal dependent bit-allocation problem can be formulated as follows[5]:

$$\begin{array}{ll}
\text{Given} & F_1, F_2, V, R_{\mathrm{budget}} \\
\text{find} & \hat{X} = (Q_1, Q_2) \\
\text{such that} & \hat{X} = \arg\min_{X} \{D_1(Q_1) + D_2(Q_2|Q_1)\} \\
\text{subject to} & R_1(Q_1) + R_2(Q_2|Q_1) \le R_{\mathrm{budget}}
\end{array}$$

where we would have an independent bit-allocation problem in the particular case where $D_2(Q_2|Q_1) = D_2(Q_2)$ and $R_2(Q_2|Q_1) = R_2(Q_2)$.

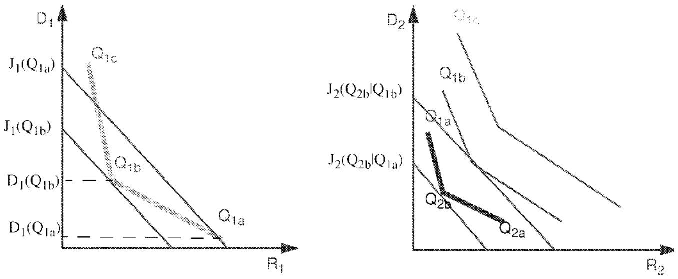
This constrained optimization problem can be transformed into an unconstrained problem using the Lagrange multiplier method [19]–[21] by introducing a Lagrangian cost

$$\begin{aligned} J(\lambda) &= J_1(Q_1) + J_2(Q_2|Q_1) \\ &= \{D_1(Q_1) + \lambda R_1(Q_1)\} \\ &\quad + \{D_2(Q_2|Q_1) + \lambda R_2(Q_2|Q_1)\} \end{aligned} \tag{2}$$

where the Lagrange multiplier $\lambda$ is a nonnegative constant.

[4] In the case of motion, each block in the predicted frame depends on up to four blocks in the reference, and conversely, blocks in the reference frame affect several blocks in the target image.

[5] The relative importance of $D_1$ and $D_2$ can be controlled using weighting constant $\alpha$, i.e., using $D_1 + \alpha D_2$ as our distortion measure. This allows us to support two different views of the depth-perception process: *fusion theory* and *suppression theory* [17], [18]. Note that this modification can easily be incorporated into our framework.




(a) (b)

Fig. 2. Operational RD plots in a typical dependent bit-allocation scenario: (a) reference image and (b) target image. Independent bit allocation: for a given $\lambda$, the quantizer $Q_{1b}$ is optimal because the Lagrangian cost $J_1(Q_{1b})$ is smaller than for the others. Dependent bit allocation: if stereo pairs are considered together, there is a chance for the quantizer $Q_{1a}$ to be optimal because the total Lagrangian cost $J_1(Q_{1a}) + J_2(Q_{2b}|Q_{1a})$ can be smaller than the cost $J_1(Q_{1b}) + J_2(Q_{2b}|Q_{1b})$.

Fig. 2 provides an example of why dependencies have to be taken into account [14].

For the blockwise quantizer assignments, the Lagrangian cost in (2) can be expressed as

$$J(\lambda) = \sum_{m=0}^{N-1} \{d(q_m) + \lambda r(q_m)\} + \sum_{n=0}^{N-1} \{d(p_n|q^{\eta_1}(v_n)) + \lambda r(p_n)\} \tag{3}$$

where $q^{\eta_1}$ is a vector that contains the quantizer indexes of those blocks in $F_1$ that are used to predict the current block in $F_2$. As shown in Fig. 3, $\eta_1$ denotes (at most) two consecutive blocks in $F_1$ along the DV. Given the disparity vector $v_1$, the selection of a quantizer for $B'_1$ in the DCD frame will be affected by the selection of quantizers for $B_2$ and $B_3$ in $F_1$. Thus, a block in the DCD frame depends only on the quantizers $p_n$ and $(q_m, q_{m+1})$, i.e., $d(p_n|q^{\eta_1}) = d(p_1|q_2, q_3)$ in Fig. 3. In general, the index $m$ can be denoted as $m = n + \lfloor (v_n/|B|) \rfloor$, where $\lfloor\,\rfloor$ and $|B|$ represent the floor function and the width of the block, respectively.

### C. Solution Using the Viterbi Algorithm

Due to the predominant 1-D dependency, a row of blocks (ROB) in the target image depends only on the ROB in the same position in the reference image. Therefore, we only need to consider the bit allocation for pairs of ROB's,[6] as other ROB's do not affect the result. Even if there is some small vertical disparity, this is a sufficiently good approximation.

Let $\text{ROB}_1$ and $\text{ROB}_2$ be the ROB's in the same position in the reference image and the DCD image, respectively. From now on, when we refer to the $k$th block, it should be clear that this is *within the particular* ROB. We first represent all possible allocations for each pair of ROB's by constructing a *trellis*. The costs of the nodes and branches of the trellis correspond, respectively, to the blocks in $\text{ROB}_1$ and $\text{ROB}_2$. Refer to Fig. 4 for the trellis corresponding to the example in Fig. 3.

[6] One from the reference and one from the target located at the same position.

We now define our method more formally. Let $k$ be the index of the stage.

*1) Stage:* The $k$th stage in the trellis corresponds to the $k$th block in $\text{ROB}_1$. Therefore, the number of stages $K$ is equal to the number of blocks in $\text{ROB}_1$.

*2) Node:* Each node in the $k$th stage corresponds to a possible quantizer choice for the $k$th block of $\text{ROB}_1$. The choices are ordered from top to bottom in order of finest to coarsest. Therefore, the number of state nodes per stage is $L = |q|$, i.e., the number of available quantizers for the reference image. Each node has a corresponding Lagrangian cost, $J_1(i;\, k)$ in (4), which depends only on the rate and the distortion of the $k$th block of $\text{ROB}_1$ when quantizer $i$ is used

$$J_1(i;\, k) = d\big(q_k^i\big) + \lambda r\big(q_k^i\big). \tag{4}$$

*3) Branch:* A branch, joining nodes $q_k^i$ and $q_{k+1}^j$, corresponds to the *optimal* vector of quantizers $p_n^{\text{ij}}$, for the (possibly more than one) blocks in $\text{ROB}_2$ that depend on blocks $k$ and $k+1$ in $\text{ROB}_1$. Each branch has a total Lagrangian cost

$$J_2(i,\, j; k) = \sum_{n \in \eta_2(k,\, k+1)} \Big\{ d\Big(p_n^{\text{ij}} | q_k^i,\, q_{k+1}^j\Big) + \lambda r\Big(p_n^{\text{ij}} | q_k^i,\, q_{k+1}^j\Big)\Big\} \tag{5}$$

which adds up the Lagrangian costs corresponding to each of the blocks $n$. Note that more than one block in $\text{ROB}_2$ can be assigned to a given branch (this will depend on how large the disparity search region is). For example, in Fig. 3, two blocks in $\text{ROB}_2$ are assigned to a branch, i.e., $B'_1$ and $B'_2$ both depend on $B_2$ and $B_3$, and thus the two Lagrangian costs corresponding to $B'_1$ and $B'_2$ would be added to each branch linking stages 2 and 3 in the trellis.

*4) Path:* A path is a concatenation of branches from the first stage to the final stage in the trellis. Each path corresponds to a set of quantization choices for both $\text{ROB}_1$ (nodes) and $\text{ROB}_2$ (branches). The cost of a path is the accumulated cost of branches and nodes along the path.

*5) Trellis:* The trellis is made of all possible paths linking the nodes in the first stage and the nodes in the last stage, i.e., all possible concatenated choices of quantizers for a given pair of ROB's in the stereo pair.

Once the trellis has been constructed, the optimal blockwise dependent quantization problem is equivalent to finding the smallest cost path from a node in the first stage to a terminal node in the last stage of the trellis. For a fixed $\lambda$, by applying the VA [22], we can obtain the best possible quantizer selection that minimizes the Lagrangian cost defined in (3). To find the optimal bit allocation for a given bit budget, we may need to iteratively change $\lambda$ until we find $\lambda^*$ such that $R(\lambda^*) - R_{\text{budget}} \le 0$. The desired $\lambda^*$ can be selected using a fast bisection search algorithm, as in [21]. For a fixed $\lambda$, the procedure is as follows.

Step 1) (Initialization) Add an initial node $B_0$ and a final node $B_T$ where $T = K + 1$, where $K$ denotes the number of stages. Set $k = 0$ and $J_{\text{acc}}(0;\, 0) = 0$.

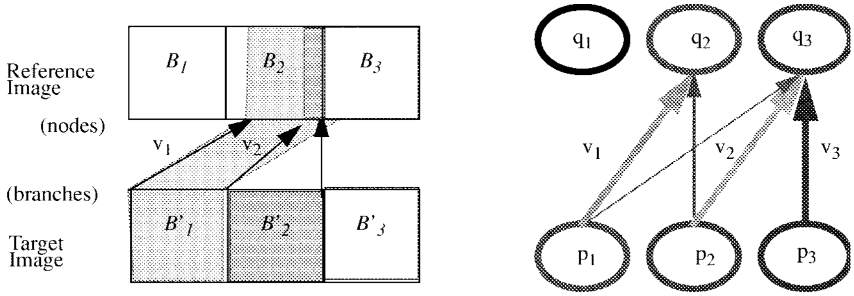



Fig. 3. Binocular dependency between corresponding blocks along the disparity vector. At most two consecutive blocks in the reference image are related to a block in the target image. For example, a block $B_1'$ in the target image is compensated from two consecutive blocks $B_2$ and $B_3$ in the reference image along the disparity vector $v_1$. Therefore, the distortion of the block in the DCD frame is a function of $p_1$, $q_2$, and $q_3$.

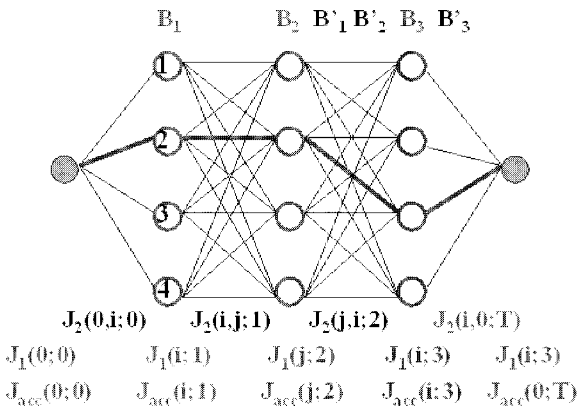



Fig. 4. Trellis structure for blockwise dependent bit allocation. Each node in the trellis corresponds to a quantizer choice for a block in $F_1$ and has a corresponding Lagrangian cost. The quantizer indexes are monotonically increasing from finest to coarsest. A branch linking two stages corresponds to a quantization assignment to all the dependent blocks in the DCD frame. The corresponding Lagrangian cost is attached to the branch. The darker path denotes selected quantizers using the Viterbi algorithm.

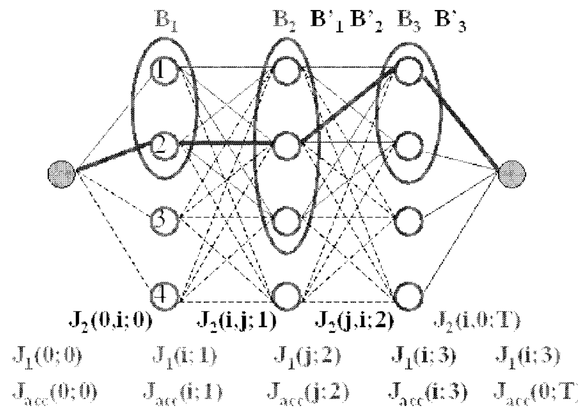



Fig. 5. A heuristic fast search. The trellis in Fig. 4 can be restricted using the proposed fast search algorithm. The search space is reduced to the (circled) nodes, selected by a blockwise optimization using two $\lambda$'s for the reference image only. If we choose the two $\lambda$'s as $(0, \lambda_2)$, then we keep the finest quantizer for $\mathrm{ROB}_1$. Then, we only need to calculate RD values of the blocks in $\mathrm{ROB}_2$ for the remaining (solid line) branches.

Step 2) At stage $k$, branches are added to the end of each node $i$ (of all surviving paths), and Lagrangian costs $J_1$ and $J_2$ are assigned to the node and the branch, respectively.

Step 3) At a stage $(k+1)$, for each node $j$, an accumulated transition cost from node $i$, $J_{\mathrm{tr}}(i, j;\ k)$, is calculated by summing the accumulated cost $J_{\mathrm{acc}}(i;\ k)$ and the transition cost $J_2(i, j;\ k)$. Of all arriving branches (at most $L$), the one with the lowest accumulated transition cost is chosen. The resulting cost is assigned to the accumulated cost $J_{\mathrm{acc}}(j;\ k+1)$, and the remaining branches are pruned

$$\begin{aligned} J_{\mathrm{tr}}(i, j;\ k) &= J_{\mathrm{acc}}(i;\ k) + J_2(i, j;\ k) \\ J_{\mathrm{acc}}(j;\ k+1) &= \min\{J_{\mathrm{tr}}(i, j;\ k)\}_{i=0}^{L-1} \\ &\quad + J_1(j;\ k+1). \end{aligned} \tag{6}$$

Step 4) If $k < K$, then $k = k+1$; go to Step 1) and repeat.

Step 5) The path with minimum total cost across all paths can be found by backtracking the surviving path.

In the proposed framework, the quantization choices for the $k$th block in the reference image and corresponding blocks in the DCD frame do not affect the choices for the future blocks. Thus, based on Bellman's optimality principle, the VA provides a globally optimal solution because suboptimal paths at a given node cannot be optimal overall and can thus be pruned. Similarly, overall optimality within the stereo pair can be achieved by assigning the same $\lambda$ to every pair of ROB's since each pair of ROB's is independent [14], [21].

### *D. Fast Algorithm Using Monotonicity*

We now propose a fast algorithm based on the monotonicity property, i.e., the observation that finer quantization of $F_1$ tends to allow more efficient coding, in the RD sense, for the DCD frame [14]. For example, $J(p|q^*) \leq J(p|q)$, for $q^* \leq q$, where $q^*$ is finer than $q$. If $\lambda = 0$, $d(p|q^*) \leq d(p|q)$, for $q^* \leq q$. To take advantage of this, we first consider an "$\mathrm{ROB}_1$-only optimization" and then only calculate RD values for the selected nodes and branches. Fig. 5 shows an example of the reduced trellis obtained from that of Fig. 4.

The proposed fast search algorithm is as follows.

Step 1) First, we select a pair of Lagrange multipliers, $\lambda$'s, e.g., $(\lambda_1, \lambda_2)$. For example, we can choose $(0, \lambda_2)$

so that we do not eliminate the finest quantizer for $\mathrm{ROB}_1$ (which tends to be good, given monotonicity).

Step 2) Then, for each $\lambda$, we set to zero the branch costs and then select, at each stage, the node that minimizes $D + \lambda R$. Each $\lambda$ will provide an optimal path (a set of nodes) in the trellis. We then restrict ourselves to only consider those paths that lie *in between* the paths selected using $\lambda_1$ and $\lambda_2$.

Step 3) Last, we use the algorithm outlined above except that we apply the VA on the pruned trellis so that only a subset of the branches representing blocks in $\mathrm{ROB}_2$ needs to be grown.

The proposed fast scheme reduces the computational complexity significantly. The most significant contribution to the complexity of the VA comes from having to compute RD values to assign the node and branch costs. For example, if a block in $\mathrm{ROB}_2$ depends on two blocks in $\mathrm{ROB}_1$, each combination of quantization choices for these blocks gives a different residue. Thus, we would need to compute the residues $L \times L$ times and to quantize them $L$ times, where $L$ is the number of quantizers. In other words, the required total number of RD values per trellis is in $O(L^3 K)$, because the total number of nodes and branches per trellis are $L \times K$ and $L^2 \times K$, respectively. Let the number of remaining nodes per stage in the pruned trellis be $\tilde{L}$. In our proposed fast scheme, we need only those RD values corresponding to the remaining nodes and branches in the trellis. Thus, the required total number of RD values for the pruned trellis is $O(K\tilde{L}^2 L)$. Based on the above, if $\tilde{L}$ is small, our pruning will result in much reduced complexity and, with good choices of $(\lambda_1, \lambda_2)$, will not affect much the final quality.

## III. Experimental Results

In our experiments we use two stereo pairs, one synthesized and one natural.[7] The target image is segmented into blocks of size 8 $\times$ 8 pixels, and then disparity estimation is performed using fixed-size block matching. The search window sizes are (0, 15) and ($\pm$ 2, 15) for the synthesized and the natural pairs, respectively. For this particular selection of the block and search window sizes, two consecutive blocks in $F_1$ will affect *at most* two consecutive blocks in $F_2$, as shown in Fig. 4. Note that our algorithm can accommodate arbitrary search regions.

The resulting DV field is losslessly encoded using DPCM with a causal median predictor to exploit the spatial redundancy among neighboring DV's. The reference image and the DCD frame are encoded using a JPEG-like coder, with the only modification with respect to baseline JPEG [23] being that we allow each block to have a different quantization scale (QS). Consequently, the change of QS per block allows the encoder to assign different levels of quantization coarseness to each block. For each block one of among eight different QS's can be chosen from the set QS $= \{90, 80, \ldots, 20\}$, where increasing values indicate finer quantization. In our calculation of rate, we assume that a constant overhead is used for each block.

In our experiments, we compare the RD performance of the blockwise dependent quantization scheme with that of: 1) framewise constant quantization and 2) blockwise independent optimal allocation. Note that a constant quantization scale is used for all blocks in each image in 1), while the optimal quantization scale for each block in the two images is estimated with a fixed $\lambda$ in 2). The RD points we plot are obtained for $\lambda = \{0, 0.1, 0.5, 1, 2, 100\}$.

Fig. 6 compares RD performance of the synthesized image. Fig. 6(a) and (b) shows RD performances for the reference image and the target image in terms of the rate and the PSNR. Fig. 6(c) compares the mean RD performance of a pair of stereo images in terms of the overall bit rate and the mean PSNR.

As shown in Fig. 6, quantization selections for $F_1$ affect RD performance for $F_2$. Table I compares the resulting RD performances for the dependent bit-allocation scheme to those for the independent blockwise bit-allocation scheme. In methods, "IND" and "DEP" denote the blockwise independent and dependent bit-allocation schemes, respectively. Note that at the same rate, $R_{\mathrm{mean}} = 0.727$, the dependent bit-allocation scheme tends to assign relatively more bits to $F_1$ and achieve slightly higher PSNR gain for $F_2$ even using a lower rate, as compared to the independent scheme.

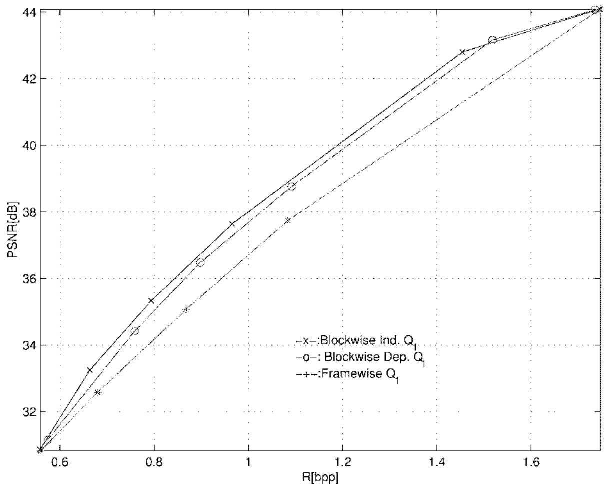
Fig. 7 shows the mean RD performance obtained for another stereo pair, for which small vertical disparity vectors are allowed.

According to the experimental results, at the same rate, the proposed blockwise dependent bit-allocation method resulted in 0.2–0.5 dB improvement in mean PSNR for the two images in a stereo pair, as compared to the optimal blockwise independent quantization. Note that the mean PSNR gains mainly arising from the fact that the finer quantization for the reference image ($F_1$), increasing the rate at the expense of decreasing the rate for the target image ($F_2$), improves the encoding efficiency for the target image ($F_2$).

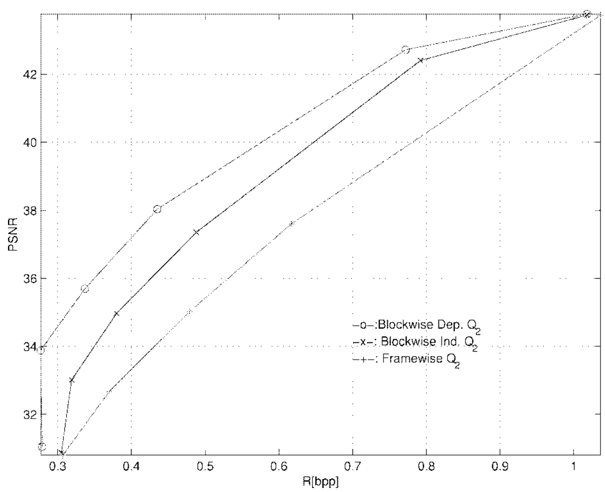
Fig. 8 shows the RD performance of the proposed fast algorithm. To keep the finest quantizers for the reference, we set $\lambda_1$ to be zero. The selected $\lambda$'s are $\{\lambda_1, \lambda_2\} = \{0, 0.5\}$ in our experiments. As explained in the previous section, we restrict the search space to the nodes in between two paths selected by the two $\lambda$'s. Last, the set of dependent quantization assignments is determined using the pruned trellis. In our example, only 61% of original nodes and 37.5% of branches remain in the pruned trellis, which results in significant savings in the computation of RD values. The resulting number of computations in the pruned trellis is about 37.5% of the original. The overall RD performance remains practically unchanged in this case, as compared to the original algorithm. Note, however, that we need to make a good choice for the $\lambda$ range, based on the expected quality level for the overall image. Thus, in the example, we show good performance at high rates, whereas the low rate points cannot be achieved since the corresponding nodes have been pruned out.

Additional research is required to achieve a more complete allocation algorithm including the disparity estimation. Further

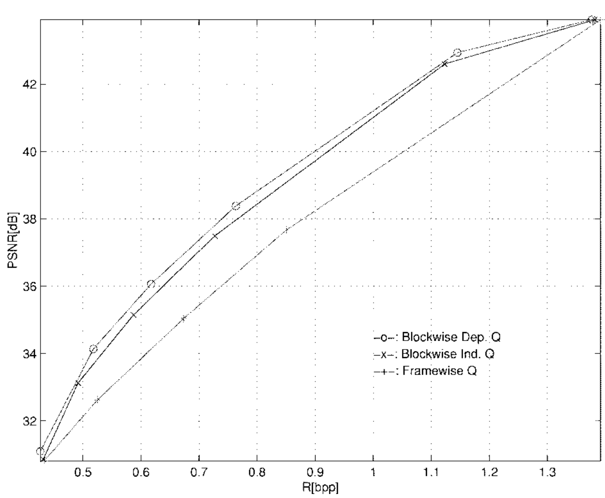
[7] The test images and resulting DV fields are available from http://escalus.usc.edu/~wwoo/Stereo. The original images where obtained from Room: http://www-dbv.cs.uni-bonn.de/~ft/stereo.html and Fruit: http://www.ius.cs.cmu.edu/idb/html/stereo/index.html.




(a)




(b)




(c)

Fig. 6. RD performance comparison (image: Room.256, block size $= 8 \times 8$, $SW = 16$, $|Q| = 8$, QS $= \{90, 80, \ldots, 20\}$, and $\lambda = \{0, 0.1, 0.5, 1, 2, 100\}$). The "+" mark denotes the DC with framewise quantization. The "$\times$" mark and "o" mark correspond to the DC with blockwise independent and dependent quantizations, respectively. Each point is generated with one different $\lambda$. (a) RD performance for the reference image is similar for both types of blockwise allocation. (b) Better RD performance for the target image can be achieved using the dependent bit-allocation approach. (c) Overall performance also improves when taking dependencies into account.

TABLE I
COMPARISON OF RD PERFORMANCE IN TERMS OF RATE (bpp) AND PSNR (dB) (ROOM.256)

| Method | $\lambda$ | $R_1$/PSNR$_1$ | $R_2$/PSNR$_2$ | $R_{\rm mean}$/PSNR$_{\rm mean}$ |
|---|---|---|---|---|
| IND | 0.500 | 0.966/37.64 | 0.488/37.35 | 0.727/37.49 |
| DEP | 0.588 | 1.044/38.23 | 0.410/37.47 | 0.727/37.83 |

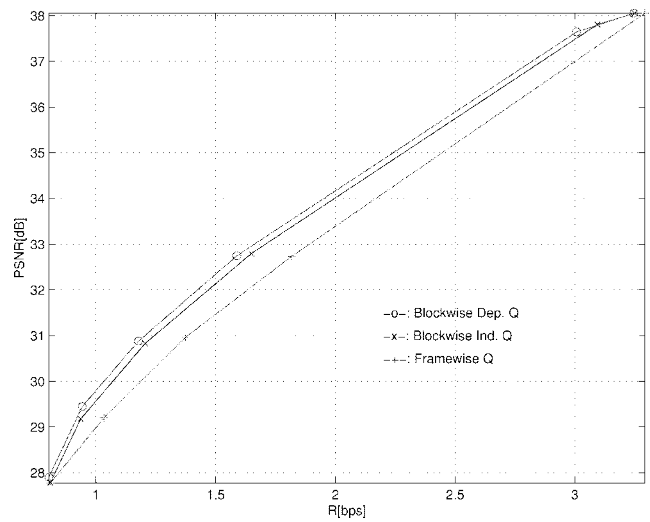



Fig. 7. RD performance comparison. (Fruit: block size $= 8 \times 8$, $SW = 10$, $|Q| = 8$, QS $= \{90, 80, \ldots, 20\}$, and $\lambda = \{0, 0.1, 0.5, 1, 2, 100\}$). The "+" mark denotes the DC with framewise quantization. The "$\times$" mark and "o" mark correspond to the DC with blockwise independent and dependent quantizations, respectively.

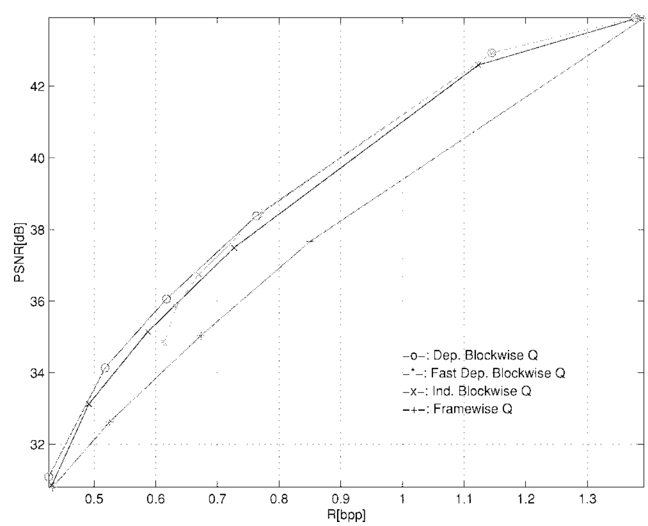



Fig. 8. RD performance comparison of fast algorithm. (Room: block size $= 8 \times 8$, $SW = 10$, $|Q| = 8$, $QS = \{90, 80, \ldots, 20\}$, $\lambda_1 = [0, 0.5]$, and $\lambda = \{0, 0.1, 0.5, 1, 2, 100\}$). The "*" mark denotes the proposed fast algorithm, which only uses 61% of the original nodes (the resulting computation corresponds to about 37.5% of the original). The "+" mark denotes the DC with framewise quantization. The "$\times$" mark and "o" mark correspond to the DC with blockwise independent and dependent quantizations, respectively.

study of our algorithm may lead to a better understanding of the similar issues in blockwise dependent allocation for video coding, where an optimal solution cannot be achieved due to the 2-D nature of the dependencies. The extension to stereo video coding [12], [24], [25], in which both temporal and binocular dependencies have to be taken into account, is another area of future work.

## References

[1] W. Woo and A. Ortega, "Dependent quantization for stereo image coding," in *Proc. SPIE EI-VCIP*, Jan. 1998, vol. 3309, pp. 902–913.
[2] M. E. Lukacs, "Predictive coding of multi-viewpoint image sets," in *Proc. ICASSP*, Oct. 1986, pp. 521–524.
[3] M. G. Perkins, "Data compression of stereopairs," *IEEE Trans. Commun.*, vol. 40, pp. 684–696, Apr. 1992.
[4] W. Woo and A. Ortega, "Stereo image compression based on the disparity compensation using the MRF model," in *Proc. SPIE VCIP*, Mar. 1996, vol. 2727, pp. 28–41.
[5] ——, "Stereo image compression based on the disparity field segmentation," in *Proc. SPIE EI-VCIP*, Feb. 1997, vol. 3024, pp. 391–402.
[6] M. B. Slima, J. Konrad, and A. Barwicz, "Improvement of stereo disparity estimation through balanced filtering: The sliding-block approach," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 913–920, Dec. 1997.
[7] H. Aydinoglu and M. H. Hayes, "Stereo image coding: A projection approach," *IEEE Trans. Image Processing*, vol. 7, pp. 506–516, Apr. 1998.
[8] M. S. Moellenhoff and M. W. Maier, "Transform coding of stereo image residuals," *IEEE Trans. Image Processing*, vol. 7, pp. 804–812, June 1998.
[9] W. C. Chung, F. Kossentini, and M. J. T. Smith, "Rate-distortion constrained statistical motion estimation for video coding," in *Proc. IEEE ICIP*, Oct. 1995, pp. 184–187.
[10] J. Ribas-Corbera and D. Neuhoff, "Optimal bit allocations for lossless video coders: Motion vectors vs. difference frames," in *Proc. IEEE ICIP*, Oct. 1995, pp. 180–183.
[11] G. M. Schuster and A. K. Katsaggelos, "A theory for the optimal bit allocation between displacement vector field and displaced frame difference," *IEEE J. Select. Areas Commun.*, vol. 15, pp. 13–26, Dec. 1997.
[12] D. Tzovaras and M. G. Strintzis, "Motion and disparity field estimation using rate-distortion optimization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 171–180, Apr. 1998.
[13] M. C. Chen and A. N. Willson, Jr., "Rate-distortion optimal motion estimation algorithm for motion-compensated transform video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 147–158, Apr. 1998.
[14] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. Image Processing*, vol. 3, pp. 533–545, Sept. 1994.
[15] L. J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, pp. 486–499, Aug. 1998.
[16] T. Wiegand, M. Lightstone, D. Mukherjee, T. G. Campbell, and S. K. Mitra, "Rate-distortion optimized mode selection for very low bit-rate video coding and the emerging H.263 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 182–190, Apr. 1996.
[17] B. Julesz, *Foundations of Cyclopeon Perception.* Chicago, IL: Univ. of Chicago Press, 1971.
[18] I. Dinstein, M. G. Kim, A. Henik, and J. Tzelgov, "Compression of stereo images using subsampling transform coding," *Opt. Eng.*, vol. 30, no. 9, pp. 1359–1364, Sept. 1991.
[19] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Oper. Res.*, vol. 11, pp. 399–417, 1963.
[20] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization.* New York: Wiley, 1988.
[21] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1445–1453, Sept. 1988.
[22] G. D. Forney, "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268–278, Mar. 1973.
[23] W. Pennebaker and J. Mitchell, *JPEG Still Image Compression Standard.* New York: Van Nostrand Reinhold, 1994.
[24] B. L. Tseng and D. Anastassiou, "Multi-viewpoint video coding with mpeg-2 compatibility," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 414–419, Aug. 1996.
[25] A. Puri, R. V. Kollarits, and B. G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *J. Signal Process.—Image Commun.*, vol. 10, pp. 201–234, 1997.