Video Sequence Segmentation Based On Rate-Distortion Theory

J.R. Morros, F. Marqués, M. Pardàs, P. Salembier

Dept. of Signal Theory and Communications
ETSETB - Universitat Politècnica de Catalunya
Campus Nord - Mòdul D5
C/ Gran Capità, 08034 Barcelona, Spain
Tel: (343) 401 74 04, Fax: (343) 401 64 47
E-mail: morros@gps.tsc.upc.es

ABSTRACT

This paper describes a coding-oriented segmentation technique for video schemes using an optimization strategy to address the problem of bit allocation. The optimization is based on the rate-distortion theory. Our purpose is to define a method to obtain an 'optimal' partition together with the best coding technique for each region of this partition so that the result is optimal in a rate-distortion sense.

KEYWORDS: Video coding, segmentation-based coding, rate-distortion theory, bit allocation

1 INTRODUCTION

Segmentation-based coding methods[8,16,10] rely on a segmentation of the image, that can be described as a set of textured regions. The texture and contours of these regions are coded and sent to the receiver. The goal is efficient sequence coding, i.e., maximal quality of the coded images as well as minimum coding cost. The segmentation has to produce a set of regions (partition) suitable for coding purposes. This can be done by ensuring that the regions are homogeneous in some sense (e.g. gray level, color or motion). Due to this homogeneity, the information of each region can be separately coded in an efficient manner.

Image segmentation is an ill-posed problem.[3] For a given image, different segmentations can be obtained depending on the homogeneity criterion. So, the segmentation technique has to be selected in function of the specific application where it is to be used. This means that to obtain a segmentation valid for coding purposes, the special characteristics of the coding techniques that are used should be taken into account, so that an 'optimal' partition results. This leads to minimum coding residue (difference between the coded and the original images) and also minimum coding cost. In addition, the segmentation algorithm should be flexible enough to allow the use of new coding techniques that may appear.

Video sequences are usually highly redundant in the temporal domain. Most of the information contained in a frame is already present in the previous frame. Motion information between two consecutive frames of the sequence can be used to take advantage of this redundancy. In the case of segmentation-based coding this can be done by making a prediction of the partition and the texture of the current frame being coded by motion compensating the previous ones. Then, a simplification of the resulting prediction error is coded. A problem to be addressed is how to maintain the temporal coherence between the partitions of successive coded frames. The partition of each frame must be related to the previous partition in order to minimize the amount of information to be sent. Moreover, for the coding algorithm to be completely general (no a priori assumptions on the contents of the images are made), it must be able to deal with new regions appearing on the image, while other regions may disappear.

Optimality can be defined in the framework of Shannon's rate-distortion theory,[6] where the problem of bit allocation can be stated as obtaining a representation of the sequence that leads to a minimum distortion of the coded signal for a given bit budget and a set of coding techniques. In the case of segmentation-based coding, this means to obtain an 'optimal' partition together with the best coding technique for each region of this partition so that the result is optimal in a rate-distortion sense. That is, an optimal repartition of the given bit budget among a set of regions.

Shoham and Gersho[17] solved the problem of optimal bit allocation for independent signal blocks and an arbitrary set of quantizers, using a fast algorithm based on the Lagrange-multiplier optimization method. This method inspired the work of Ramchandran and Vetterli,[12] where the goal was to obtain the optimum wavelet-packet decomposition in a rate-distortion sense jointly with optimum associated quantizers. This was done in a pure intra-frame approach. Also, they stated that their results could be applied to quadtree segmentation as well. In the work of Reusens[13] this was further developed, and a solution was provided for optimal quadtree partitioning in a rate-distortion sense together with optimal representation models. In this work, though a motion-compensation coding model was used, the definition of the quadtree relied in a purely intra-frame technique, so that the problem of temporal coherence of the segmentation was not addressed.

Our approach is based in the mentioned works. An optimal or very nearly optimal solution to the problem of bit allocation is provided, while ensuring the temporal coherence of the successive partitions. This is done by constructing a set of partition proposals that represent the image with various levels of detail. First, motion information is applied to the partition of the previous coded frame. The result is a new partition, the projected partition, that shows the time evolution of the regions from the last frame to the current frame. The set of partition proposals is derived from this projected partition, so that temporal coherence between successive frames is preserved. An optimization algorithm will choose the appropriate regions to form the final partition, as well as the best coding technique for each region of this partition. Fluctuations with respect to the projected partition can be properly introduced by selecting regions from the different partition proposals.

Although this bit allocation strategy has been tested in the framework of a very low bit-rate video coder, its generality allows it to deal with any kind of bit-rates. The algorithm can also work at constant-quality and variable bit rate as will be described in Section 5.

This paper is organized as follows: Section 2 is a short overview of the problem of constrained bit allocation and its solution. Section 3 gives an overall description of the coding algorithm where the presented segmentation process is used. Section 4 shows how the set of partition proposals is constructed. Section 5 describes the optimization algorithm. In Section 6, experimental results are presented. Finally, in Section 7 some conclusions are provided.

2 CONSTRAINED BIT ALLOCATION

The constrained problem of minimizing the distortion D of independent signal blocks subject to a restriction on the coding cost R has been widely addressed in source coding literature.[17,12,13] An optimal solution to this problem can be obtained via dynamic programming,[2] but this approach leads to a very high computational complexity, so it is not appropriate for video sequence coding. The constrained problem can be formulated as: Given an image to code $\mathcal{I}$, a coding bit budget R_{budget}, a partition $\mathcal{P}$ of $\mathcal{I}$, and an available set of quantizers $\mathcal{Q}$, $\forall$ region $P_i \in \mathcal{P}$ find the mapping (P_i, Q_j) such that:

$$\min_{Q_j \in \mathcal{Q}} \sum_i D_i(Q_j) \quad \text{subject to} \quad \sum_i R_i(Q_j) \leq R_{budget} \tag{1}$$

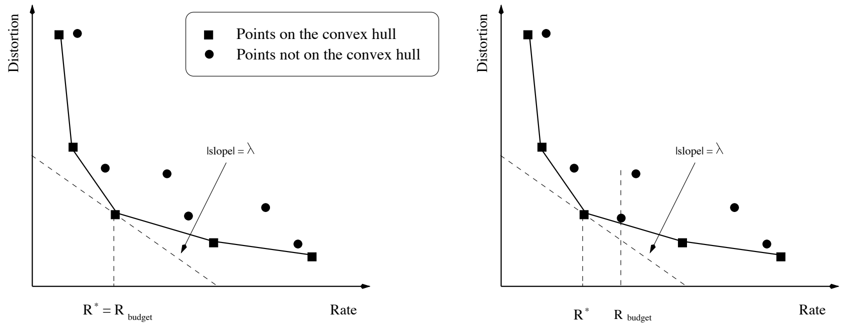
This constrained problem can be converted into an unconstrained problem by combining the distortion and the rate by means of a Lagrange multiplier λ, and then, minimizing the Lagrangian cost function $J(\lambda) = D + \lambda R$. This unconstrained problem is much easier to solve, and fast algorithms can be used. The unconstrained problem can be formulated as:

$$\min_{Q_j \in \mathcal{Q}} \sum_i (D_i(Q_j) + \lambda R_i(Q_j)) \qquad \lambda \in \Re, \quad \lambda \geq 0 \tag{2}$$

At optimality, λ is the same for all the signal blocks. As the signal blocks are independent, the problem can be solved separately for each block. Thus, equation 2 can be expressed as:

$$\sum_i \min_{Q_j \in \mathcal{Q}} (D_i(Q_j) + \lambda R_i(Q_j)) \tag{3}$$

If a $\lambda^* = \lambda^*(Q_1, \ldots Q_N)$ can be found such that $R_T(\lambda^*) = \sum R_i = R_{budget}$ the constrained and unconstrained problems are equivalent and a solution has been found. We have to note that this formulation does not guarantee to find always the solution to the constrained problem. Some times there is no λ^* such that $\sum_i R_i(\lambda^*) = R_{budget}$. In these cases, a near-to-optimal solution is found. It is possible to demonstrate that R(λ) is monotonically non-increasing with λ; that is, if $\lambda_2 \geq \lambda_1$, then $R(\lambda_2) \leq R(\lambda_1)$. This allows to determine the value of λ^* using an iterative fast convex search algorithm (bisection or Newton's method). This method traces out the points that are on the convex hull of all possible rate-distortion pairs. When the optimal solution to the constrained problem lies on the convex hull, solutions to the constrained and to the non-constrained problems are the same (Fig: 1 a). Otherwise, the result of the optimization is a point on the convex hull, and the solution is not strictly optimal. The results show, however, that the algorithm is highly efficient and can be considered essentially optimal.[17]



(a) Optimal solution

(b) Near-to-optimal solution

Figure 1: Rate-Distortion Curves

2.1 Application to segmentation-based coding

We want each frame of the sequence to be optimal in a rate-distortion sense. The optimization strategy, as stated in[13] will choose, from a set of candidate regions, and a set of coding techniques, the appropriate regions and the associated coding techniques so that the Lagrangian cost of the resulting coded image is minimal. This optimization problem can be formulated as follows:

> Given an image $\mathcal{I}$, a set of partitions of $\mathcal{I}$ $\mathcal{P}$, and a set of texture coders $\mathcal{C}$, find the optimal mapping of (P_{i_j}, C_k), such that $\{P_{i_j}\}$ is a valid partition of $\mathcal{I}$, optimal in a rate-distortion sense for a given bit budget R_{budget}.

The regions have to be coded independently. The set of texture coders $\mathcal{C}$ is composed of different coding techniques. Each technique can be used with several levels of quantization. For the sake of simplicity, each quantization level will be considered as a a different coding technique.

3 OVERVIEW OF THE CODING ALGORITHM

The segmentation algorithm has been developed to work inside a complete video coding system, SESAME.[4] For the sake of completeness, a brief description of this algorithm is given in this section. Four basic blocks can be differenced, as shown in Figure 2.

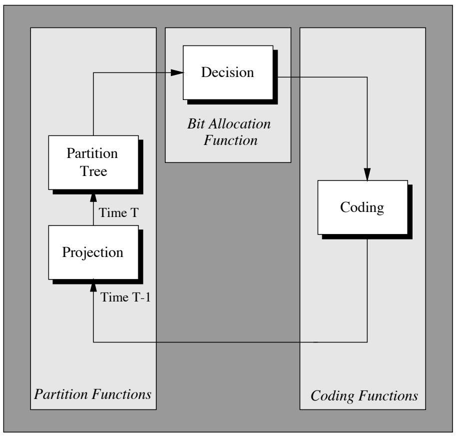


Figure 2: Scheme of the segmentation-based coder

The *Projection* block applies motion information to the partition of the previously coded frame #(n-1), to obtain a projected partition for the current frame to be coded #(n). The resulting partition is the basis for the current image segmentation. The projected partition provides the time evolution of the regions in the previous partition, i.e., it tracks the regions over each frame of the sequence. In the projection step some regions may disappear but new regions cannot appear.

For the intra-frame mode, the *Projection* step is not related to the previously coded partition. In this case, an initial partition with a predefined number of regions is constructed.

The *Partition Tree* block constructs a set of partition proposals from the projected partition, as shown in Figure 3. These partition proposals define a reduced set of regions which are candidate to belong to the final partition. The various levels of segmentation represent the image with various levels of detail, so fluctuations with respect to the projected partition can be properly introduced by selecting regions from the different partitions. As the partition proposals are derived from the projected partition, the temporal coherence between the partitions of successive frames is preserved.

On the one hand, a reduced set is necessary in order to limit the computational complexity of the decision that will be taken afterwards. On the other hand, the partition proposals should be carefully created to allow an efficient and pertinent decision. The final partition will in fact be composed of regions issued from the different levels of the Partition Tree.

The *Decision* block takes the proposals from the Partition Tree and makes a decision on which regions will belong to the final partition, and which coding technique will be used for every region. This step is based in the optimization algorithm described in the Section 2.

Finally, the *Coding* block takes the results of the decision block and codes the image. In the SESAME description, this block is divided into four sub-blocks since four different types of information have to be coded: the motion parameters, the texture parameters, the contours of the partition and the information related to the decision.

The segmentation procedure proposed in this paper is strongly related to the *Partition Tree* and *Decision* blocks. To discuss this segmentation procedure, we are going to further detail these blocks.

4 CONSTRUCTION OF THE PARTITION TREE

The Partition Tree is formed by the projected partition plus a set of hierarchical partitions that are generated from this projected partition (See Figure 3)

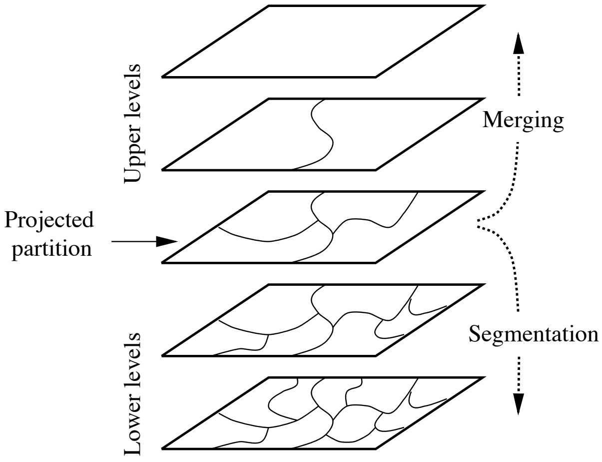


Figure 3: Partition Tree

The partitions are constructed by means of two different procedures: The upper levels, above the projected partition, consist on larger regions. These coarse partitions, are generated by merging regions from the projected partition. The criterion used to merge regions is motion similarity. Merged regions are more efficient in terms of coding cost, because they require only one set of motion parameters, one set of texture parameters, and there are less contours to code. However, this leads to an increase in the distortion.

For the intra-frame mode, the merging of the upper levels cannot be based on a motion similarity criterion. Instead, a gray level similarity criterion is used.

The lower levels, below the projected partition, are constructed by successively segmenting the projected partition. A constrained morphological segmentation is used to ensure that the contours of previous regions are preserved.[15] Regions of a given level are split to produce the regions of the next lower level. Each level has to refine the segmentation of the previous level by introducing new significant regions. The procedure is purely spatial: it does not take into account any motion information. As a consequence, the contours of new objects are based on spatial discontinuities.

5 DECISION: CHOICE OF THE PARTITION AND CODING STRATEGY

The function of the *Decision* step is to select a set of regions that form a valid partition of the current frame, and the appropriate texture coders so that the coded image is optimal in a rate distortion sense.

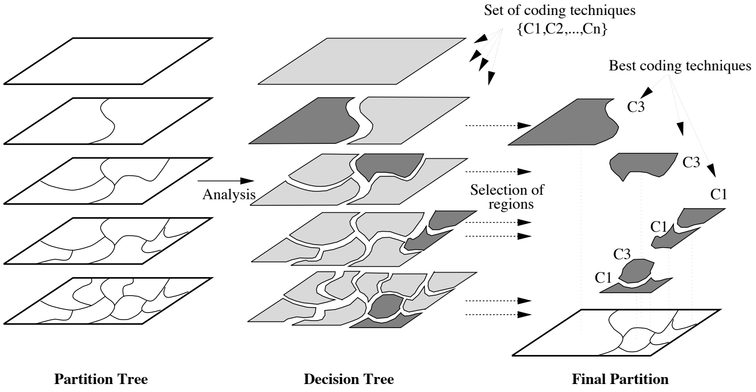


Figure 4: Decision Process

The fact that the contours of higher levels of the Partition Tree are preserved in the lower levels allows us to define the regions in the Partition Tree as a hierarchical structure, where each region in a given level is connected with the regions in the next lower level that originate from it in a father/children relationship.

5.1 Creation of the Decision Tree

The hierarchical structure of the Partition Tree can be represented by a tree-structure, that from now on will be called Decision Tree. Each region of the Partition Tree is represented by a node in the Decision Tree. In order to define the coding strategy in the Rate-Distortion sense, the Decision Tree should also convey the information about the coding cost (Rate) and quality (Distortion) of all possible texture coding techniques. Every node has associated a list of the rates and distortions that result from the coding of the corresponding region with each of the available coding choices (See Figure 5).

This structure allows to compare the Lagrangian cost of a region with the sum of the costs of the branches emanating from that region for a given λ. To preserve the additivity of the rate and distortion among the levels of the tree, the contours of the parent regions must not change in the oversegmentation steps.

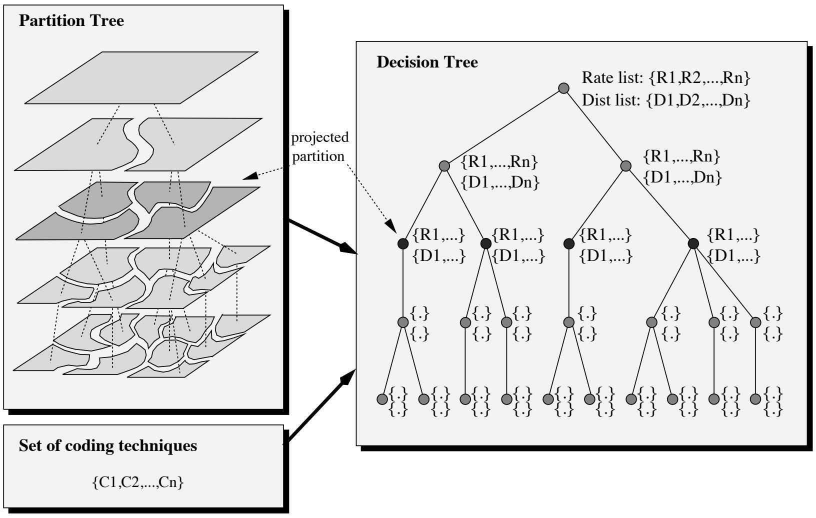


Figure 5: Construction of the Decision Tree

The texture of each region in each level is coded by all the available coding techniques, and the resulting rates and distortions are stored in the nodes of the Decision Tree. In this work, various coding techniques have been used: a shape-adaptive 2-D transform,[18,7] a generalized orthogonal transform,[5] a region based wavelet decomposition,[1,4] and the gray level average of the regions, which is a special case of these techniques. Various quantization levels are available for each technique.

This approach gives a great flexibility, because if a new coding technique is developed, it is possible to put it into competition with the previous ones, without modification of the decision algorithm. All kind of coding techniques appropriated for dealing with regions can be used. The simultaneous use of various coding techniques can deal efficiently with different kinds of textures.

The coding depends on whether the current frame is sent on intra or inter mode:

- Intra-frame mode: The texture of each region of the current frame is coded separately. Then, contours are coded by means of a chain-code algorithm.
- Inter-frame mode: There are two choices. Coding of the prediction error after motion compensation of the texture of each region, or coding of the original frame as in intra-frame mode. While the coding of the prediction error is generally more efficient, the possibility to code the gray level information of the current frame may result in a better coding of new regions appearing in the current frame. Then, contours are coded by means of a modified chain-code algorithm.[11] Motion-compensation followed by error coding of the contours may result in a further reduction of the bit rate.[14]

The distortion measure that has been used is squared error due to its simplicity. Anyway, any distortion measure that is additive in space can be used. Additivity in space means that if a region is split, the sum of the distortions of the resulting regions is equal to the distortion measured on the original region. In the optimization step, to prune the Decision Tree a comparison between the Lagrange cost of each node and the sum of Lagrange costs of all its child nodes is performed. This requires a distortion measure compatible with this comparison.

The rate is a measure of the global coding cost in bits of each region. For each region, the total cost is the sum of the costs of the texture coefficients, motion parameters (in inter-frame mode), and contour of the region.

To take into account the effect of the entropy coder, the cost of texture is estimated by measuring the entropy of the coefficients. The cost of the motion parameters is also given by entropy estimation of the model parameters. The cost of the contour is considered to be proportional to the perimeter of the region. Our tests show that these estimations give a very good approximation to the 'real' costs.

5.2 Optimization Algorithm

The optimization algorithm will determine the optimal subtree of the Decision Tree together with the optimal coding model for each node of the tree. The resulting subtree must define a valid partition for the current frame. As stated in Section 2, our approach is based on the Lagrange-multiplier method and on the work of Ramchandran and Vetterli and the work of Reusens.

For a given value of λ, a fast pruning criterion can be used that provides a valid partition and that minimizes the overall Lagrange cost J(λ). First, for each node, the Lagrange cost for each texture coding technique is computed. The technique giving the minimum J(λ) is considered as the optimal one for this node. Then, a set of local decisions are performed. The Lagrange cost of each node is compared with the sum of Lagrange costs of the nodes that emanate from it. If the cost of the 'father' node is less or equal than the cost of the 'child' nodes, the tree is pruned at this level. Otherwise, we look at the next level. This procedure is applied recursively until all the tree has been analyzed. The additivity in rate and in distortion is mandatory.

This pruning criterion is shown in figure 6. It can be formulated as follows:

$$\text{prune if} \quad (D_{child_1} + \cdots + D_{child_N}) + \lambda(R_{child_1} + \cdots + R_{child_N}) > (D_{parent} + \lambda R_{parent}) \qquad (4)$$

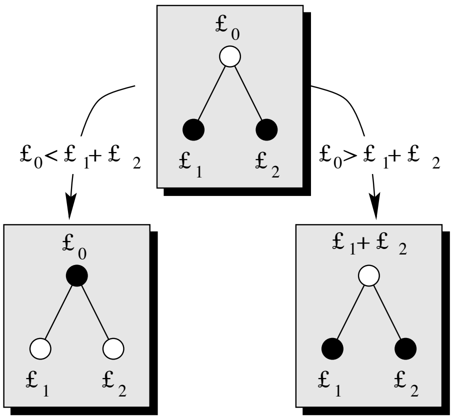


Figure 6: Local Decision

That is, if the Lagrange cost of the 'father' node is lower than the sum of Lagrange costs of the 'children', this node becomes active. Its associated region in the Partition Tree is selected to belong to the resulting partition. The 'children' nodes and all nodes that are below them are deactivated. If the Lagrange cost of the 'father' is larger than the sum of Lagrange cost of the 'children', these nodes are marked as active and the 'father' node remains deactivated. In a rate-distortion sense, it is better to split the region and to code the resulting regions. Note that this is only a step of the optimization algorithm. The resulting partition will be the optimal partition only if $J(\lambda)$ is minimum for all values of λ.

Then, the minimum Lagrange cost for the overall image is found with an iterative fast convex search algorithm, as stated in Section 2. The objective is to find a value of λ such that the cost of the resulting partition is less or equal than R_{budget}. The algorithm can be stated as:

- Find $\lambda_L > \lambda_U$ such that: $R_T(\lambda_L) \leq R_{budget} \leq R_T(\lambda_U)$.
- Apply the pruning criterion of Eq. 4 from full-depth tree to root. An optimal partition for the current λ is found.
- If $R(\lambda_L) = R_{budget}$ or $R(\lambda_U) = R_{budget}$ the solution is found and the algorithm stops. Otherwise, a new value of λ is computed:

$$\lambda_n \leftarrow \frac{|D_T(\lambda_L) - D_T(\lambda_U)|}{|R_T(\lambda_L) - R_T(\lambda_U)|} \tag{5}$$

The two last steps are iterated until $R(\lambda_n) = R_{budget}$ or until no further reduction in the quantity $R(\lambda) - R_{budget}$ is possible. The algorithm converges very fast. Usually, five to ten iterations are enough to obtain the optimal results. At the end of the algorithm, the result is a valid partition and the choice of coding techniques for each region that minimizes the Lagrange cost of the current frame (See Figure 4).

The algorithm can also work at constant-quality and variable bit rate. This is, a target quality value that must be reached with the minimum coding cost. If we define the Lagrange cost as J'(λ) = R + λ D, the same optimization algorithm can be used.

5.3 Optimization of the algorithm for very low bit rates

In the approach described up to here, every frame of the sequence is given the same bit budget. This results in frames with variable quality, so problems can arise when working at very low bit rates. Some frames may have very poor quality (scene changes, complex motion). To solve this problem, we propose a very simple method to give more bits to the frames with very bad quality.

The optimization works basically on a fixed nominal budget, but a minimum quality threshold is defined. If the coded image defined by the decision step does not reach this minimum quality, the budget is increased by steps of 5decision has found the optimum strategy:

- The distortion is minimal.
- The budget is at least equal to the nominal budget.
- The signal to noise ratio is everywhere above a given threshold.

This way, the frame have better quality, and it may be possible to use it for the compensation of the next frame. This improves the prediction of the future frames, so they can be coded with less bits. The coding budget for each frame is not constant, but the sequence overall bit rate is very homogeneous, as is shown in Section 6.

6 RESULTS

Experimental results are presented for *Foreman*, *News* and *Akiyo* sequences, used for the tests of MPEG-4 algorithm proposals. *Akiyo* has small motion and a lot of static background zones. The format of the sequence is QSIF (176 x 128 pixels). *Foreman* is a sequence with camera zoom, pan and a lot of complex motion, and important scene changes. The format is QCIF (176 x 144 pixels). *News* is a sequence with a foreground nearly static and a moving background. Only the first frame of these sequences has been coded in intra-frame mode. All sequences have been coded at a frame rate of 5 Hz. Figure 7 show original and coded frames from *Foreman* and *News* sequences, as well as the respective partitions and the choice of texture techniques. *Foreman* has been coded with a budget bit-rate of 48 kbits/s and a minimum PSNR per frame of 25 dB. *News* has been coded with a budget bit-rate of 24 kbits/s and a minimum PSNR per frame of 22 dB. *Akiyo* has been coded with a budget bit-rate of 12 kbits/s and a minimum PSNR per frame of 25 dB.

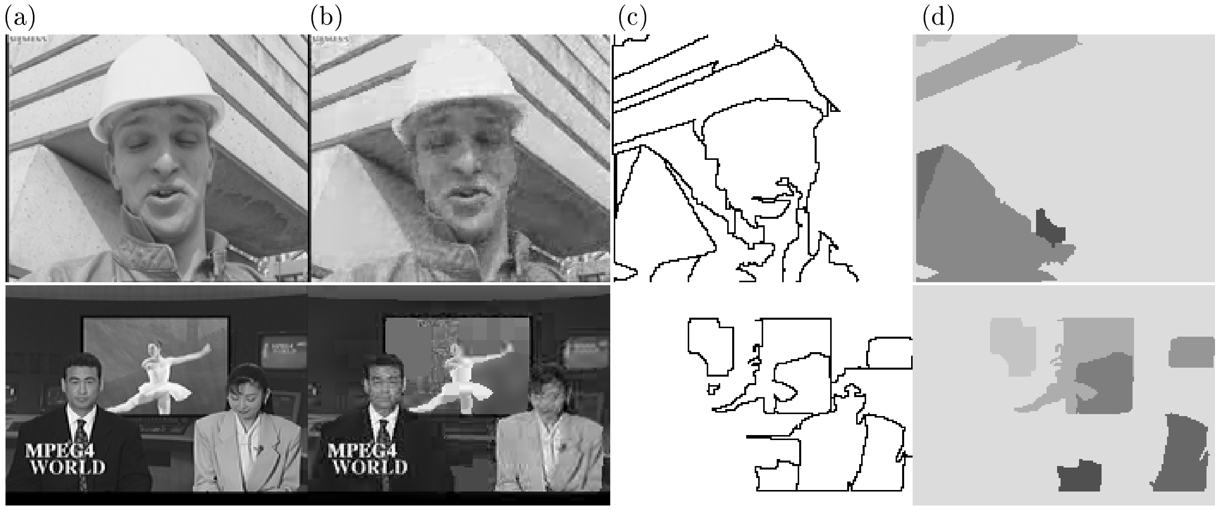
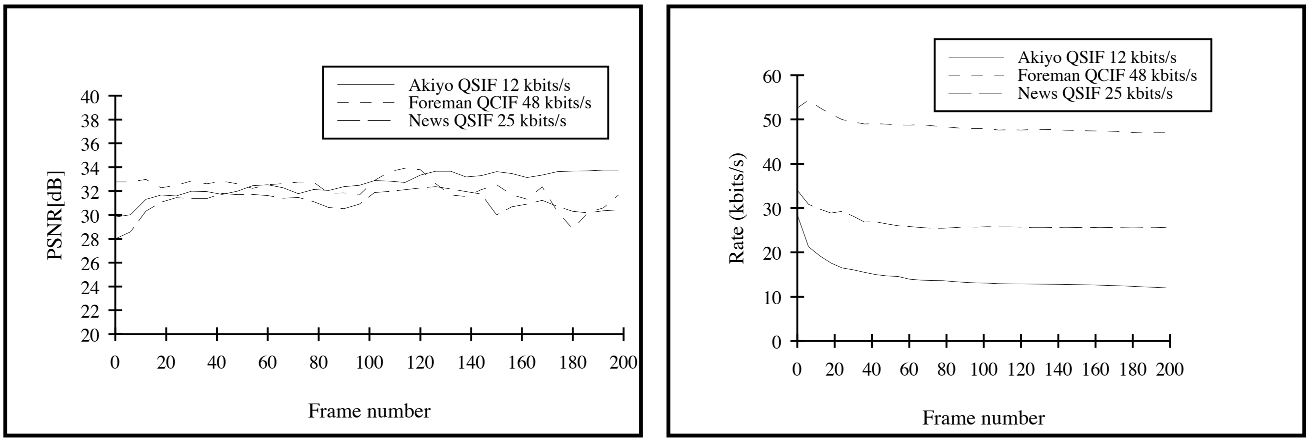


Figure 7: Coding results: First row: *Foreman* #(120). Second Row: *News* #(120). (a) Original images. (b) Coded images. (c) Partitions. (d) Decision mapping

Tough the bit-rate is not strictly constant, the results (see Figure 8) show that the variations are very small. As we can see, the quality of the resulting sequences is also near-to-constant. In the first frame the algorithm described in Section 5.3 increases the budget bit-rate until the quality reaches a given minimum. As it is coded in intra mode, more bits are needed to reach the minimum quality.



(a) (b)

Figure 8: PSNR and rate evolution

Figure 9 show the partitions that define the Partition Tree for the frame # 66 of Foreman, and the final partition.

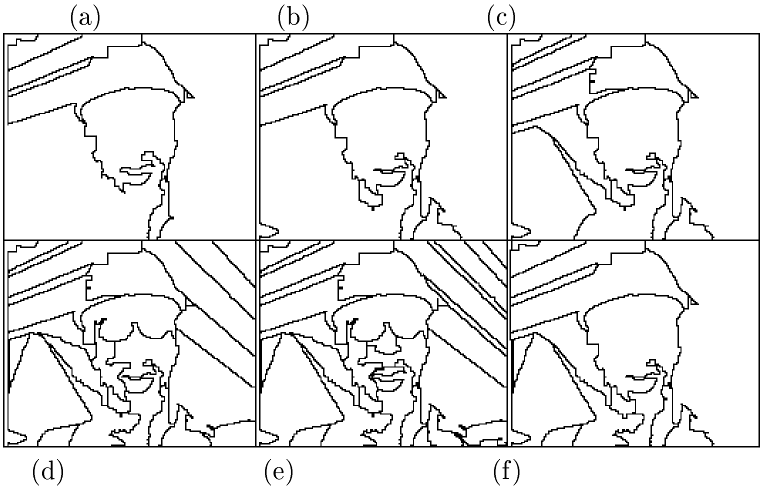


Figure 9: Segmentation Tree: Foreman #(66). (a)-(e):Levels of the tree. (f):Final segmentation

7 CONCLUSIONS

We have shown an optimization strategy to address the problem of bit allocation within a segmentation-based coding scheme of video sequences. It provides optimal results in a rate distortion sense. The overall process takes into account the motion information. An inter-frame approach is used. The optimization is performed separately on each frame but the candidate regions are generated not only from the local statistics of the current frame, but from the information of the evolution of the regions from frame to frame. This results in a more efficient coding because the partition of the current frame can be predicted from the partition of the previous frame.

The flexibility of the algorithm makes also possible to extend its applications. For example, it can be adapted to code some selected regions in the partition with better quality than others by simply multiplying the distortion in the regions forming the area of interest.[9,4] Also, as stated in Section 5, the algorithm can also work at constant-quality and variable bit rate. In addition, new coding techniques can be used without modifying the structure of the algorithm.

Although this bit allocation strategy has been tested in the framework of a very low bit-rate video coder, its generality allows it to deal with any kind of bit-rates.

8 ACKNOWLEDGMENTS

This work was supported by the European Community through the MAVT project of the RACE 2072 program and by the project of the Spanish Government TIC95-1022-C05-05

9 REFERENCES

[1] H.J. Barnard. *Image and Video Coding Using a Wavelet decomposition.* PhD thesis, Technische Universiteit Delft, Faculteit der Elektrotechniek, Vakgroep Informatietheorie, 1994.

[2] R. Bellman. *Dynamic Programming.* Princeton University Press, Princeton, NJ, 1957.

[3] M. Bertero, T.A. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76:869–887, 1988.

[4] I. Corset et al. Technical description of SESAME. Technical Report MPEG 95/408, ISO/IEC JTC1/SC29/WG11, November 1995.

[5] M. Gilge, T. Engelhardt, and Mehlan R. Coding of arbitrarely shaped image segments based on a generalized orthogonal transform. *EURASIP, Image Communications*, 1(2):153–180, October 1989.

[6] R.M. Gray. *Source Coding Theory.* Kluwer Academic Publishers, Norwell, MA, 1990.

[7] K. Jensen, E.and Rijkse, I. Lagendijk, and P. van Beek. Coding of arbitrarily sahaped image segments. In *Proceedings of Workshop on Image Analysis and Synthesis in Image Coding*, Berlin, October 1994.

[8] M. Kunt, A. Ikonomopoulos, and M. Kocher. Second generation image coding techniques. *Proceedings of the IEEE*, 73(4):549–575, April 1985.

[9] F. Marqués, B. Marcotegui, and F. Meyer. Tracking areas of interest for content based functionalities in segmentation based video coding. In *Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing*, 1996. Accepted paper.

[10] F. Marqués, M. Pardàs, and P. Salembier. Coding-oriented segmentation of video sequences. In L. Torres and M. Kunt, editors, *Video Coding: The Second Generation Approach*, pages 79–124. Kluwer Academic Publishers, 1996. ISBN: 0 7923 9680 4.

[11] F. Marqués, J. Sauleda, and T. Gasull. Shape and location coding for contour images. In *Picture Coding Symposium*, pages 18.6.1–18.6.2, Lausanne, Switzerland, March 1993.

[12] K. Ramchandran and M. Vetterli. Best wavelet packet bases in a rate-distorsion sense. *IEEE Trans. on Image Processing*, 2(2):160–175, April 1993.

[13] E. Reusens. Joint optimization of representation model and frame segmentation for generic video compression. *Signal Processing*, 46:105–117, September 1995.

[14] P. Salembier. Motion compensated partition coding. In SPIE, editor, *Visual Communication and Image Processing, VCIP'96*, volume 2727, Orlando, USA, March 1996.

[15] P. Salembier and M. Pardàs. Hierarchical morphological segmentation for image sequence coding. *IEEE Transactions on Image Processing*, 3(5):639–651, September 1994.

[16] P. Salembier, L. Torres, F. Meyer, and C. Gu. Region-based video coding using mathematical morphology. *Proceedings of IEEE (Invited Paper)*, 83(6):843–857, June 1995.

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

[18] T. Sikora, S. Bauer, and B. Makai. Efficiency of shape-adaptive 2-d transforms for coding of arbitrarily shaped image segments. *IEEE Trans. on Circuits and Systems for Video Technology*, 1995.