PARTITION PREDICTION FOR SEGMENTATION-BASED CODING TECHNIQUES

Ferran Marqués, Bernat Llorens and Antoni Gasull
Universitat Politècnica de Catalunya
Campus Nord - Mòdul D5
C/ Gran Capità, 08034 Barcelona, Spain
E-mail: ferran@gps.tsc.upc.es

ABSTRACT

This paper presents a general partition prediction scheme. It consists of four steps: region parametrization, region prediction, region ordering and partition creation. The evolution of each region is separated into two types: regular motion and shape deformation. Fourier Descriptors are used to parametrized both types of evolution and they are separately predicted in the Fourier domain. The predicted partition is built from the ordered combination of the predicted regions, using morphological tools. This technique is applied in the framework of segmentation-based video coding techniques for coding sequences of complete partitions as well as sequences of binary images (shape information in Video Object Planes -VOP-).

1 INTRODUCTION

In segmentation-based image sequence coding techniques, the sequence of image partitions containing all the contour information usually requires a large amount of bits to be coded. A means to reduce the partition coding cost is by taking advantage of the temporal redundancy that exists between consecutive partitions. The current partition is predicted relying on the previous partitions and the prediction error is computed. The necessary information to perform the prediction in the receiver side is, therefore, transmitted as well as a simplification of the prediction error.

Following this idea, the works presented in [2, 3] propose to predict the current partition in a region-by-region approach, estimating the motion of the regions and simplifying the prediction error by means of morphological tools. Although the motion models used in [2, 3] are purely translational, more complex models can be used. In these works, the texture information is also compensated using the motion parameters that have been estimated for the region shapes. This leads to an unaccurate estimation and compensation of the textures and, therefore, their coding cost increases.

A second approach is presented in [7]. In this work, the motion parameters computed using texture information are used to predict the evolution of the regions of the previous partition. Nevertheless, each region is not processed individually but the partition is processed as a whole. Finally, the prediction error is simplified and transmitted. However, this approach requires an additional ordering to be sent to be able to recover the current partition in the receiver side.

The work presented in this paper proposes a partition prediction relying on the previous partitions which are available in the receiver side. The evolution of every region is separately predicted using only the information of its shape in the previous partitions. Once each region has been predicted, the whole predicted partition is built. Therefore, as in [7], the coding algorithm deals with the complete partition as a whole. The main advantage of this approach is that the only information which is transmitted is the partition prediction error. The same prediction that has been computed in the transmitter side can be conducted in the receiver side without additional information. Furthermore, textures can be coded using motion parameters which have been computed based only on texture information, which leads to a more efficient coding algorithm.

After this introduction, the paper is structured as follows. Section 2 deals with a global scheme for partition prediction. In Section 3, an implementation of the different parts of this scheme is presented. Section 4 is devoted to the application of this prediction method to the framework of partition coding. Finally, Section 5 shows some results of this technique.

2 A GLOBAL SCHEME FOR PARTITION PREDICTION

A general partition prediction scheme is proposed in this paper which extends the work on interpolation and extrapolation of partitions presented in [4, 1]. The complete scheme is illustrated in Figure 1.

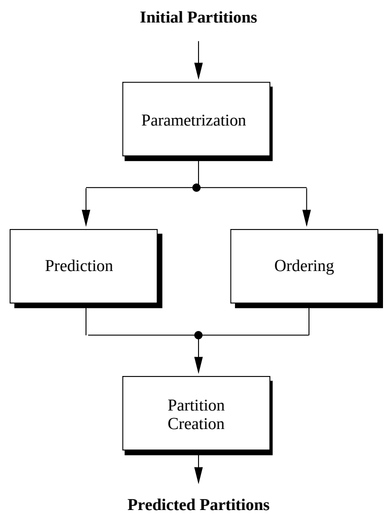


Figure 1: Block diagram of the partition prediction

This scheme relies on the shape and position information of each region. The region evolution through the time domain is divided into two types: regular motion and shape deformation. Regular motion is described by a given motion model (e.g.: translation, zoom and/or rotation). The region evolution that cannot be described by regular motion is said to be shape deformations. Both types of information are obtained directly from the shape and position of each region after their parametrization.

Once regions in both partitions have been parameterized, these parameters are used to predict their evolution. This prediction yields a separated representation of the evolution of each region. Predicted regions cannot be directly combined to obtained the final predicted partition. When combining them, two problems arise. First, predicted regions can overlap and, second, some parts of the space may not be covered by any region.

To solve the first problem, regions are ordered. This ordering gives priority to one region with respect to its neighbors so that, in case of overlap, the area of overlapping is assigned to this region. Finally, after combining the predicted regions following the above ordering, uncovered areas should be assigned to some of the neighbor regions. This is done in order to ensure that a final partition is achieved.

3 IMPLEMENTATION OF THE PARTITION PREDICTION SCHEME

For each one of the previous blocks, a specific implementation is proposed. These implementations are extensions of the scheme presented in [4]. That is, they use Fourier Descriptors to represent the region shapes [9] and morphological tools [8] to construct the final partition. The evolution of each region is predicted using linear prediction on the previous parameters. Such a prediction implies a model of constant velocity of the region evolution.

For the sake of completeness, the scheme presented in [4] is briefly summarized in the sequel. The *Region prediction* step is further detailed since it contains the main differences with respect to the basic scheme.

3.1 Region parametrization

Fourier Descriptors are used for extracting global information of the regions. There are several ways to define the Fourier Descriptors of a region [9]. For prediction purposes, the definition that directly relates the position function $z_p[n]$ of region R_p to its Fourier transform $Z_p[k]$ is very useful:

$$Z_p[k] = \frac{1}{N}\sum_{n=0}^{N-1} z_p[n] \quad e^{-j\frac{2\pi}{N}kn} \tag{1}$$

From the Fourier Descriptors, a new set of parameters is defined which is more suitable for describing the evolution of the contours. This evolution can be divided into two types: regular motion and shape deformation. The information related to regular motion is extracted first from the Fourier Descriptors and treated separately. Afterwards, the Fourier Descriptors are normalized with respect to the parameters associated to the regular motion so that a set of parameters related to the shape deformation is obtained.

In this work, the model of regular motion that is assumed represents translation, zooming and rotation. Each type of motion is associated to the evolution of a different parameter. This way, four different normalizations are applied to each contour to obtain the normalized Fourier Descriptors which are associated to the shape deformation. These normalizations correspond to equiform transformations as defined in [9]: translation ($\mathcal{D}_\zeta, \zeta \in C$), zooming ($\mathcal{S}_\beta, \beta \in R^+$) and rotation ($\mathcal{R}_\alpha, \alpha \in [0, 2\pi]$) ($\mathcal{T}_\tau, \tau \in [0, N)$). The method to estimate each one of these parameters is completely described in [4].

After normalization with respect to these parameters, contours are represented by two sets of parameters. The first set is related to the regular motion of the region $(\zeta, \beta, \alpha, \tau)$ whereas the second set ($Z^*[k]$) is associated to the shape deformation.

$$Z_p^*[k] = \mathcal{T}_\tau \mathcal{R}_\alpha \mathcal{S}_\beta \mathcal{D}_\zeta Z_p[k] \tag{2}$$

3.2 Region prediction

The evolution of the previous parameters is predicted in order to obtain the contours of the predicted regions. From the set of parameters $(\zeta, \beta, \alpha, \tau)$, τ does not need to be predicted since it does not affect to the final representation of the predicted contours [4]. The other three parameters are linearly predicted in order to fulfill the constant evolution requirement.

$$\zeta_{tp} = (1 - \frac{t}{T})\zeta_{0p} + \frac{t}{T}\zeta_{Tp} \tag{3}$$

$$\frac{1}{\beta_{tp}} = (1 - \frac{t}{T})\frac{1}{\beta_{0p}} + \frac{t}{T}\frac{1}{\beta_{Tp}} \tag{4}$$

$$\alpha_{tp} = (1 - \frac{t}{T})\alpha_{0p} + \frac{t}{T}\alpha_{Tp} \tag{5}$$

In the previous expressions, the region R_p at time t is predicted using the information at the previous times 0 and T.

Linear prediction is also used for the normalized Fourier Descriptors $Z^*[k]$. However, the high frequency components of $Z^*[k]$ lead to predicted contours with unrealistic oscillations. To overcome this problem, the prediction of the shape deformations is carried out using only a set of low frequency normalized descriptors $Z^{*LF}[k]$.

Low frequency Fourier Descriptors are selected with respect to its magnitude. The set of selected normalized descriptors have to correctly represent the region. Therefore, a very small reduction on the total energy of the normalized descriptors is allowed:

$$\frac{\sum_{k=0}^{N^{LF}-1} \|Z_p^{*LF}[k]\|^2}{\sum_{k=0}^{N-1} \|Z_p^*[k]\|^2} \geq 1 - \epsilon \tag{6}$$

Nevertheless, even a very small reduction on the total energy (e.g.: $\epsilon = 0.0001$) results in a very large reduction of the number of coefficients (values around $N^{LF}/N = 8\%$ are usual). This reduction is due to the fact that the largest part of the energy is concentrated in the lowest frequency descriptors.

Shape deformation of region R_p at time t is finally predicted using the information at the previous times 0 and T by means of the following expression:

$$Z_{tp}^{*LF}[k] = (1 - \frac{t}{T})Z_{0p}^{*LF}[k] + \frac{t}{T}Z_{Tp}^{*LF}[k] \tag{7}$$

3.3 Region ordering

Once all the regions have been separately interpolated, they are ordered so that possible overlappings are solved. The ordering gives priority to the regions having a smaller amount of shape deformation with respect to their neighbors. This procedure assumes that if the evolution of a region R_p can be represented only relying on the regular motion in a more accurate way than its neighbors, this region should preserve the shape given by its prediction [4].

3.4 Partition creation

Predicted partitions are created by combination of the various predicted regions respecting the previous ordering. However, holes may appear after combining the regions. Such holes should be covered by neighbor regions to obtain a final partition. The proposed technique performs a geodesical dilation of the predicted regions using structuring elements whose size is adapted to each region with respect to its priority level [4].

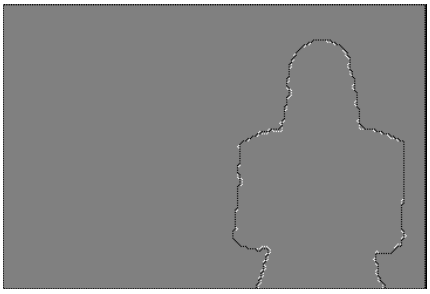
4 PARTITION PREDICTION FOR PARTITION CODING

The previous partition prediction can be used in the framework of partition coding .As stated in the introduction, the proposed partition coding algorithm is an extension of the method proposed in [7]:

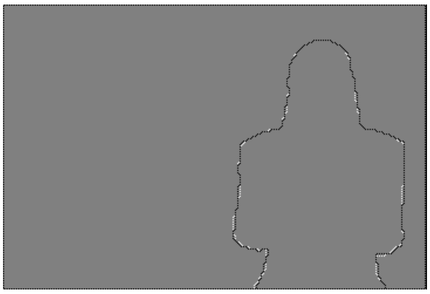
1. **Partition prediction:** The method above described is used so that a prediction of the partition in the current frame is obtained from the two previously coded partitions. The procedure is different from those used in [2, 3, 7] because it does not imply the use of a single set of motion parameters to predict both the textures and the partition.
2. **Simplification of the prediction error:** An 'over-partition' is defined by merging the predicted partition with the current partition. Each region of the over-partition is analyzed: if the predicted label (number identifying the region) does not correspond to the current label, the region is said to be part of the compensation error. Error regions are eliminated if they are very small or if they do not imply a meaningful grey level difference after coding of its texture.
3. **Coding of the error:** The error is coded by sending the information necessary to restore the over-partition in the receiver side. The receiver knows the contours of the predicted partition and some extra contours have to be sent. These contours are coded by an extended chain code technique [5].
4. **Coding of the region label:** Once the contours of the partition have been defined, one should assign the correct label to each region. For a large number of regions the correct label is defined by the compensation but for some regions the label should be actually sent.

The fact of using a simplification step to reduce the amount of prediction error makes necessary the elimination of high frequency normalized descriptors in the *Region prediction* step of the partition prediction ($Z^*[k] \rightarrow Z^{*LF}[k]$). The region prediction achieved using all normalized descriptors presents small oscillations with respect to the original region contour. Such oscillations would be removed by the simplification step; that is, they would not be handled as part of the prediction error but they would be assumed to be part of the correct contour. This way, such oscillations would propagate through the coded partition sequence.

Figure 2 shows the prediction obtained for the binary shapes of frames #61 of the sequence *Weather*. Two different predictions have been performed: in the first case, all the normalized descriptors have been used ($\epsilon = 0.0$), whereas, in the second case, only the normalized descriptors with lowest frequency have been utilized ($\epsilon = 0.00015$).



$\epsilon = 0.0$



$\epsilon = 0.00015$

Figure 2: Comparison between predictions using all or only the low frequency normalized descriptors

In these examples, the actual contours are shown in black, while the prediction error (before simplification) is shown in white. The small oscillations of the prediction error that can be observed in the case of $\epsilon = 0.0$ are removed when reducing the number of normalized descriptors ($\epsilon = 0.00015$).

5 RESULTS

This partition prediction technique has been tested on binary shape sequences as well as on complete partition sequences. In the sequel, some results using the proposed technique on both kind of sequences are presented.

5.1 Binary shape prediction

In Figure 3, the predictions of 6 frames of the sequence *Weather* are shown. The results are presented without simplification so that the quality of the prediction can be assessed. In order to allow larger motion, images have not been predicted using the two previous frames, but for the binary shape of frame $\#t$, the already coded shapes of frames $\#t-2$ and $\#t-4$ have been used. The first row in Figure 3 contains the predicted binary shapes of frames #206, #208 and #210; on turn, the second row contains those of frames #232, #234 and #236. Note that the global motion as well as the shape deformations are correctly predicted in both examples.

Figure 3: Example of binary shape prediction

5.2 Complete partition prediction

Figure 4 shows the prediction of a complete partition of frame #8 from the sequence *Mother and Daughter*. The prediction relies on the partitions of frames #0 and #4 which have been segmented using the technique presented in [6]. In this case, the simplification has been applied so that all the prediction errors leading to regions with less that 5 pixels are removed. In the example, the additional contours that have to be sent are shown in white while those contours that have been correctly predicted are shown in black.

Note that the prediction error is larger in this example than in the previous one. Errors mainly appear in areas where the segmentation result presents a lack of stability. Furthermore, the partition in this example is predicted using a longer term prediction than in the previous one: in the previous example there is 1 frame skipped between used frames, while here 3 frames are skipped.

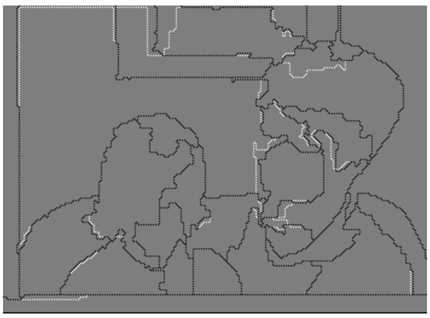


Figure 4: Coding of the partition of frame #8 of the sequence *Mother and Daughter*

6 ACKNOWLEDGMENTS

This work has been supported by the European Community through the *MAVT* project of the *RACE* program and by the Spanish Government: *TIC 95-1022-C05-05*.

References

[1] R. Bremond and F. Marqués. Segmentation-based interpolation of image sequences using mathematical morphology. In *Proceedings of International Symposium on Mathematical Morphology*, May 1996.

[2] P. Brigger and M. Kunt. Contour image sequence coding using the geodesic morphological skeleton. In *International Workshop on Coding Techniques for Very Low Bit-rate Video*, pages 3.1–3.2, Essex, Colchester, April 1994.

[3] C. Gu and M. Kunt. Contour simplification and motion compensation for very low bit-rate video coding. In *First IEEE International Conference on Image Processing*, pages 423–427, Texas, U.S.A., November 1994.

[4] F. Marqués, B. Llorens, and A. Gasull. Interpolation and extrapolation of image partitions using Fourier descriptors: Application to segmentation based coding schemes. In *IEEE International Conference on Image Processing*, volume III, pages 584–587, Washington, DC, October 1995.

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

[6] M. Pardàs and P. Salembier. Joint region and motion estimation with morphological tools. In J. Serra and P. Soille, editors, *Second Workshop on Mathematical Morphology and its Applications to Signal Processing*, pages 93–100, Fontainebleau, France, September 1994. Kluwer Academic Press.

[7] P. Salembier. Motion compensated partition coding. In *SPIE Visual Communication and Image Processing, VCIP'96*, volume 2727, pages 403–415, Orlando, USA, March 1996.

[8] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

[9] P. Van Otterloo. *A contour-oriented approach for shape analysis.* Prentice Hall International (UK), 1991.