Prediction of Image Partitions Using Fourier Descriptors: Application to Segmentation-Based Coding Schemes

Ferran Marqués, *Member, IEEE,* Bernat Llorens, and Antoni Gasull, *Member, IEEE*

Abstract—This paper presents a prediction technique for partition sequences. It uses a region-by-region approach that consists of four steps: region parameterization, region prediction, region ordering, and partition creation. The time evolution of each region is divided into two types: regular motion and shape deformation. Both types of evolution are parameterized by means of the Fourier descriptors and they are separately predicted in the Fourier domain. The final predicted partition is built from the ordered combination of the predicted regions, using morphological tools. With this prediction technique, two different applications are addressed in the context of segmentation-based coding approaches. Noncausal partition prediction is applied to partition interpolation, and examples using complete partitions are presented. In turn, causal partition prediction is applied to partition extrapolation for coding purposes, and examples using complete partitions as well as sequences of binary images—shape information in video object planes (VOP's)—are presented.

Index Terms— Fourier descriptors, image prediction, mathematical morphology, partition interpolation, partition prediction, partition sequence coding, segmentation-based image sequence coding.

I. INTRODUCTION

IN THE framework of image sequence coding, there is a continuous need for new techniques to reach higher compression ratios. Toward this goal, second-generation coding techniques [12], [29] have been proposed. In this context, the study of segmentation-based coding techniques is, nowadays, a very active field of research [13], [16], [24]. The reason for such an activity is mainly twofold. First, segmentation-based coding techniques can increase the coding efficiency of current methods. Second, they set the basis for coding algorithms with content-based functionalities embedded [10]. Such techniques divide the image into a set of regions (partition) where each region fulfills a given homogeneity criterion. Therefore, the information to be transmitted is, on the one hand, the partition describing the regions and, on the other hand, the gray level or color information (called *texture* in the following) inside each region.

A usual approach to reduce the information to be sent is to exploit the temporal redundancy existing in image sequences. This way, the current image is predicted using the information of the previous images and the prediction error is computed. The necessary information to perform the prediction in the receiver side is transmitted as well as a simplification of the prediction error. Such a causal prediction is applied in Motion Picture Expert Group (MPEG) standards, where the so-called P-frames are predicted using the previous I-frames (*forward prediction*) [28].

Noncausal prediction is also used in the context of sequence coding. It can be utilized when only a subset of the total amount of frames from the image sequence is coded. In the receiver side, frames that have not been sent are interpolated from the transmitted information; that is, predicted using both past and future information. This technique is also applied in MPEG standards, where the so-called B-frames are predicted in a noncausal means, using past as well as future P-frames and I-frames (*bidirectional prediction* or *interpolation*) [28].

In segmentation-based coding schemes, images are described in terms of texture and partition information and both types of information have to be transmitted. As in classical coding schemes, coding efficiency can be improved using prediction techniques. In the case of arbitrarily shaped regions, texture prediction can be handled by extending the image prediction methods to the case of arbitrarily shaped areas [6]. However, specific tools are still necessary for causal and noncausal partition prediction.

The problem of partition prediction is stated in this work as follows. Given two known partitions P_i and P_j where $i < j$, each one composed of a set of regions $\{R_{ip}\}$ $\{R_{jp}\}$, respectively, new partitions have to be found representing the evolution beyond P_j (causal prediction) as well as from P_i to P_j (noncausal prediction). In this work, partitions are assumed to have labels that are coherent in time. That is, a region related to a given object should have assigned the same label in both partitions. This assumption can be made since there exist segmentation-based coding techniques that solve this problem by keeping track of the region labels through the time domain [16].

A. Causal Partition Prediction: Extrapolation

An approach to reduce the partition coding cost is by taking advantage of the temporal redundancy that exists between consecutive partitions. This way, the current partition can

be causally predicted relying on the previous partitions. The necessary information to perform the prediction in the receiver side is transmitted as well as a simplification of the prediction error.

Following this idea, the works presented in [5] and [9] propose to code the current partition in a region-by-region approach, by predicting it using only the previous partition. The motion of each region is computed as an estimate of its shape evolution. The prediction error is then simplified by means of morphological tools. In these approaches, the motion parameters obtained from the partition evolution are also used to compensate the texture information. This is done in order to avoid sending two different sets of motion parameters: one for the texture and another for the partition information. However, the use of a single set of motion parameters leads to an inaccurate compensation of the textures and, therefore, their coding cost increases.

A second approach is presented in [21]. In this work, the motion parameters computed for compensating the texture information are used to predict the evolution of the regions. Furthermore, each region is not coded individually but the partition is processed as a whole. As in the previous work, the prediction error is simplified and transmitted. This second approach reduces the global coding cost (partition plus texture information) with respect to the previous techniques. However, it requires an additional ordering information in order to recover the current partition in the receiver side. In addition, the evolution of the regions is predicted only based on texture information and regardless of the shape evolution. Therefore, the performance of this approach depends upon the occlusion effects that act on the shape of the region and, by this way, upon the technique used for estimating the motion from the texture information.

A technique allowing decoupling the problems of texture and partition prediction for coding is therefore needed. This can be obtained by computing the partition prediction in the receiver side. The evolution of every region can be separately predicted using only the information of its shape in the previous coded partitions. Once the evolution of each region has been predicted, the whole predicted partition is built. Therefore, as in the work reported in [21], the coding algorithm can deal with the complete partition as a whole. The main advantage of this approach is that the only information transmitted is the partition prediction error. Furthermore, textures can be coded using motion parameters that have been computed based only on texture information, which results into a more efficient coding algorithm.

B. Noncausal Partition Prediction: Interpolation

In order to improve coding efficiency, some images can be predicted in the receiver side using past and future information; that is, they can be interpolated. Even in the case of segmentation-based coding schemes, images can be interpolated without using partition information and only utilizing the decoded images. However, a more accurate interpolation can be achieved if the boundaries of the different textures in the intermediate images are known.

In addition, when dealing with coding algorithms addressing content-based functionalities [10], the image description should make possible the representation of each frame in terms of objects. Some functionalities demand this possibility to be extended to the case of interpolated frames. For example, when combining objects from various sequences that have been recorded at different frame rates, intermediate representations of the objects have to be produced. This leads to the necessity of producing intermediate representations of their shapes and, therefore, of the image partitions.

To interpolate partitions, the information of the interior of the regions yielded by the segmentation could be used. However, segmentations can be obtained using different criteria and, therefore, regions can be related to various concepts: homogeneous gray level [13], [24], homogeneous motion [3], [25], or combination of the previous ones [1], [22], etc. An interpolation technique independent of the segmentation criteria should use only partition information; that is, the shape and the position of the regions in the partition.

Following this idea, two different approaches have been proposed to tackle the problem of noncausal partition prediction. A first approach is to handle the partitions as a whole and jointly interpolate all the regions in the partition [18]. This approach estimates the evolution of the partition by computing a geodesic distance [18], [27] between the regions in the known partitions. Therefore, as pointed out in [4] and [18], this approach is only feasible when regions slightly vary their position between the two known partitions. If this is not the case, the evolution of each region has to be separately analyzed.

A second approach is to analyze the partition in a region-by-region basis [4]. The evolution of each region is predicted independently. Afterwards, predicted regions are combined in order to build the predicted partitions. This approach does not make any *a priori* assumption on the movement of the regions. However, in the work presented in [4], the procedure to predict the region evolution is based on the definition of a geodesic distance. This procedure does not allow causal prediction, since geodesic distances are only defined between the regions from the two known partitions. That is, partition extrapolation is not possible.

A technique able to predict, both in a causal and in a noncausal manner, the evolution of image partitions is, therefore, necessary. Such a technique should incorporate a region-by-region approach, since it permits to predict the evolution of image partitions without introducing any constrains in the motion of the regions.

C. Proposed Approach for Temporal Prediction of Partitions

In this paper, a technique for partition prediction (causal and noncausal) is proposed. This technique assumes that the partition sequence has coherent labels through the time domain. Each region in the known partitions is handled separately. The evolution of a region is divided into two different types: regular motion and shape deformation. Both types of evolution are parameterized and predicted using the Fourier descriptors that characterize the contour of the region.

Predicted regions are combined in order to obtain a predicted partition. This combination may give rise to two problems. First, some predicted regions may overlap in a given area. Second, some areas may not be covered by any predicted region. To solve the case of overlapping regions, an ordering among regions is defined. In turn, areas not covered by any region are handled by using region expansion techniques based on morphological tools.

This partition prediction approach only relies on partition information and, as a consequence, the problem of texture prediction is decoupled from that of partition prediction. This feature makes it very suitable for partition coding. In addition, the partition prediction is handled in a region-by-region basis, which makes possible the correct prediction of the evolution of regions with large motion. Finally, the evolution of the regions is characterized by their Fourier descriptors, which enable causal and noncausal prediction.

The above technique is further described in the sequel. After this introduction, Section II is devoted to the general partition prediction approach. It is composed of four steps: *region parameterization*, *region prediction*, *region ordering*, and *partition creation*. Section III details the specific parameterization that is used in this work. In Section IV, different methods for predicting the evolution of the previous parameters are discussed. Section V presents the region ordering that is utilized in the case of overlapping regions. Section VI describes different methods to solve the problem of uncovered regions. To evaluate the performance of the complete technique, this partition prediction is applied in two different contexts. This way, Section VII presents a partition interpolation method, whereas Section VIII deals with the use of this prediction technique in a partition coding approach. Conclusions are given in Section IX.

II. GLOBAL SCHEME FOR PARTITION PREDICTION

The prediction technique proposed in this work predicts the evolution of each region independently. Afterwards, the predicted regions are combined in order to build the predicted partitions. The complete prediction technique can be split into four steps [14]: region parameterization, region prediction, region ordering, and partition creation.

The general prediction scheme only uses the shape and position information of each region. This information is parameterized in order to make easier its prediction. Region parameterization should be defined taking into account that its final goal is to help the description of the region evolution. Toward this goal, the region evolution 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) and a set of parameters related to this model is defined. The region evolution that cannot be described as regular motion is said to be shape deformation and, as in the previous case, it is described by a set of parameters.

Once regions in both partitions have been parameterized, the evolution of their parameters is predicted. This parameter prediction yields a separate representation of the evolution of each region. The region prediction has to handle the problem

Known Partitions

Region Parametrization

Region Prediction

Region Ordering

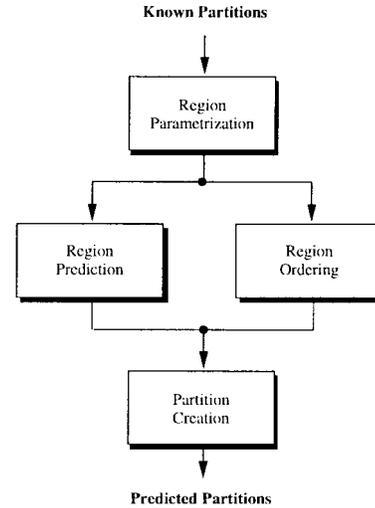Partition Creation

Predicted Partitions

Fig. 1. Global scheme.

of those regions that appear only in one of the two known partitions.

In addition, 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 partition space may not be covered by any region. To solve the problem of overlapping regions, regions are ordered. This ordering gives priority to one region with respect to its neighbors so that, in case of overlapping, the overlapping area 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 an actual partition is finally achieved. The complete scheme is illustrated in Fig. 1.

In the sequel, specific implementations of the above four steps are further described. For each block, the special case of regions only appearing in one of the known partitions (conflictive regions) is also discussed. Conflictive regions are produced by objects appearing or disappearing from the scene, as well as by errors in the segmentation procedure. These errors yield cases of lack of coherence in the labels of the known partitions.

III. REGION PARAMETERIZATION

Typical representations of regions and partitions used in coding applications (e.g., chain code approaches) are not suitable for prediction purposes. In order to deal with a region representation useful for prediction purposes, the concept of position function is used [30].

A. Position Function

A given region (e.g., region with label p) present in both known partitions (R_{ip} and R_{jp}) has to be separately parameterized. Contours are described as two complex functions $z_{ip}[n]$ and $z_{jp}[n]$, namely position functions, where both contours have been normalized to contain the same number of samples N. To obtain such functions, the exterior boundary of a region
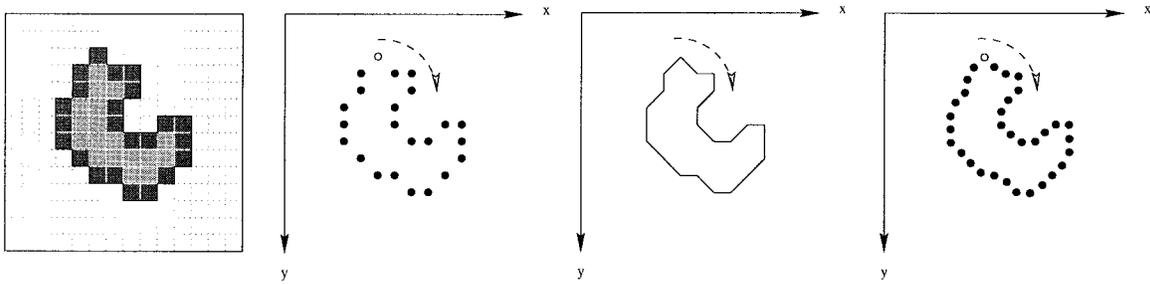
Fig. 2.   Example of the procedure to obtain the exterior boundary and the position function of a region.

is defined. The exterior boundary of a region (e.g., R_{ip}) is formed by those pixels that belong to R_{ip} and have at least a neighbor pixel that belongs to a different region that is not completely interior to R_{ip}.

A four-connected neighborhood is used since it yields simpler contours than an eight-connected one. The position of each contour point is denoted by its coordinates, as a complex number. In order to normalize the contour representation in both images, a continuous function is obtained by a polygonal interpolation of the contour points and N equidistant samples are taken from this continuous function. In practice, N is selected to be the smallest power of two larger than the number of original contour points. Fig. 2 illustrates the procedure for obtaining the exterior boundary and the position function from a given region.

Position functions could be used as parameterization of the regions to be predicted. However, the information conveyed by each point of a position function is local. As a consequence, the prediction that can be obtained with such a parameterization cannot easily characterize rotations and, very likely, yield entangled contours. Therefore, a set of parameters conveying a more global information is necessary.

B. Use of the Fourier Descriptors

The Fourier descriptors of a contour have already been used for extracting global information of an object description in pattern recognition [8] and shape approximation [7], as well as in graphic animation [2]. There are several ways to define the Fourier descriptors of a region [30]. For prediction purposes, the definition that directly relates the position function $z[n]$ to its Fourier transform $Z[k]$ is very useful:

$$Z[k] = \frac{1}{N} \sum_{n=0}^{N-1} z[n] e^{-j(2\pi/N)kn}. \tag{1}$$

All the information contained in the Fourier descriptors is not equally relevant for the current application. The set of parameters to be used in region prediction should describe correctly and naturally the contour evolution through the time domain. This evolution can be divided into two types: regular motion and shape deformation. The natural evolution of a region is closely related to a correct prediction of its regular motion. Considering this, a new set of parameters can be defined from the Fourier descriptors.

The information related to regular motion is extracted first from the Fourier descriptors and treated separately. Afterward,

the Fourier descriptors are normalized with respect to the parameters associated to the regular motion so that a set of parameters associated to the shape deformation is obtained.

In this work, the model of regular motion that is assumed represents translation, zooming, and rotation. Each of these types of motion is related to the evolution of a different parameter. For instance, translation is associated to the evolution of the gravity center of the region. Although in order to characterize the regular motion only the evolution of these parameters has to be known, their actual values are also needed. They are necessary to normalize the Fourier descriptors so that a set of parameters related to the shape deformation is obtained.

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 [30]: 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 parameters related to these concepts are as follows.

1) Gravity Center ζ: It is associated to the translation. This parameter is obtained directly from the first Fourier descriptor. The contour representation is normalized $\mathcal{D}_\zeta$ by subtraction of this sample from the Fourier descriptors

$$\zeta = Z[0], \ \mathcal{D}_\zeta Z[k] = Z[k] - \zeta \delta[k]. \tag{2}$$

The translation between regions R_{ip} and R_{jp} is computed as the evolution of the gravity center

$$\Delta \zeta = \zeta_{jp} - \zeta_{ip}. \tag{3}$$

2) Size β^{-1}: It is associated to the zooming. This parameter is not directly related to a single Fourier descriptor, but to a combination of all of them. The size parameter β^{-1}, as defined in [30], is computed after applying to the Fourier descriptors the previous normalization $\mathcal{D}_\zeta$, in order to obtain a size parameter invariant by translation.

Several operators for size computation and normalization have been proposed in the literature. The reader is referred to [30] for an exhaustive review. Here, the definition of the size parameter as the norm of the Fourier descriptors has been chosen, mainly due to its robustness in front of noisy and peaky contours. The contour representation is normalized $\mathcal{S}_\beta$ by dividing the Fourier descriptors by their Euclidean norm L_2

$$\beta = \frac{1}{\|Z[k]\|_2}, \ \mathcal{S}_\beta Z[k] = \beta Z[k]. \tag{4}$$

The zooming between regions R_{ip} and R_{jp} is computed as the evolution of the size

$$\Delta\beta = \frac{\beta_{ip}}{\beta_{jp}}. \tag{5}$$

3) Angle of Orientation α: It is associated to the rotation. As the previous parameter, it is related to a combination of all Fourier descriptors. The orientation information is contained on the phase of the Fourier descriptors: two position functions representing the same object but having as unique difference a rotation $\Delta\alpha = \alpha_{jp} - \alpha_{ip}$ lead to the same Fourier descriptors up to a multiplicative, constant term of phase $e^{j\Delta\alpha}$. However, the computation of the parameter α is not straightforward given that the phase of the Fourier descriptors is, in real cases, rather noisy. Therefore, the value of the α parameter is usually estimated. In addition, the starting point of the position function, τ, is also directly related to the phase of the Fourier descriptors: two position functions representing the same object but with different starting points $\Delta\tau = \tau_{jp} - \tau_{ip}$ lead to the same Fourier descriptors up to a linear term of phase $e^{j\Delta\tau k}$. As a consequence, classical techniques jointly estimate the value of both parameters, α and τ and their related normalizations $\mathcal{R}_\alpha$ and $\mathcal{T}_\tau$, respectively. Although this approach is followed in this work, in the application of partition prediction, the value of $\Delta\tau$ is not relevant and, therefore, it will not be used for prediction.

Several techniques can be found in the literature for the joint estimation of the α and τ parameters, usually in the context of pattern recognition. All of them share the same basic structure: A term of phase $e^{j(\alpha+\tau k)}$ is added to the phase of the Fourier descriptors so that two previously selected descriptors, $Z[k_1]$ and $Z[k_2]$, become real after normalization. That is, for $k = k_1, k_2$

$$Z^*[k] = \mathcal{T}_\tau \mathcal{R}_\alpha \mathcal{S}_\beta \mathcal{D}_\zeta Z[k] \in R \tag{6}$$

where $Z^*[k]$ denotes the normalized Fourier descriptors that contain the shape deformation information. Such techniques differ in the means to select both indexes (k_1 and k_2) as well as in the procedure to ensure that a unique solution is achieved. In [20], indexes are fixed to be $k_1 = 1$ and $k_2 = -1$, since their associated Fourier descriptors have usually the largest magnitude and, therefore, are the most representative. However, this selection leads to eight different solutions. The same indexes are selected in [11], but under the constrain of yielding real and positive normalized descriptors, which reduces to two the number of possible solutions. Solutions in the case of objects with rotational symmetry are reported in [30]. However, this constrain is very seldom fulfilled in our application. In [31], it is proposed to select $k_1 = 1$ and k_2 as the index of the Fourier descriptor with largest magnitude

$$\|\mathcal{S}_\beta \mathcal{D}_\zeta Z[k_2]\|_2 = \max_{k-\{1\}} \|\mathcal{S}_\beta \mathcal{D}_\zeta Z[k]\|_2. \tag{7}$$

This approach leads to $|k_1 - k_2|$ possible solutions ($Z^*[k_1], Z^*[k_2] \in R^+$) and, in order to obtain a unique one, several maximization procedures are proposed [31]. Finally, in [30] a method is presented that directly selects the two indexes of the Fourier descriptors with largest magnitudes.
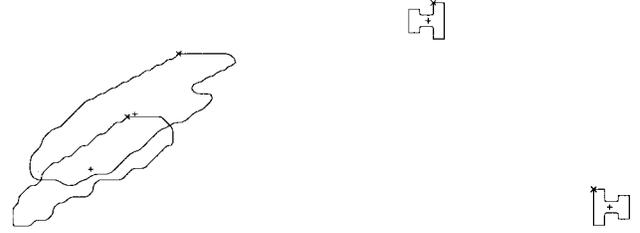

Fig. 3. Correct estimation of the angle of rotation for both a real and a synthetic region.

In our case, the join estimation of the α and τ parameters can be done so that it already accounts for the contour evolution. Parameters related to the rotation of a region can be computed using information from the contours in both partitions, $Z_{ip}[k]$ and $Z_{jp}[k]$. Therefore, the pairs (α_{ip}, τ_{ip}) and (α_{jp}, τ_{jp}) are jointly estimated. This is implemented by selecting for normalization the two indexes k_1 and k_2 leading to the two greatest magnitudes of the product:

$$MP[k] = \|Z_{ip}[k]Z_{jp}[k]\|_2 \tag{8}$$

where $MP[k_1] > MP[k_2]$ is assumed. Actually, following a conservative policy, more than two indexes are initially selected. In order to preserve possible correct solutions, all cases where $0.1MP[k_2] < MP[k]$ are analyzed. Using (8), each pair of possible coefficients results in a set of pairs of normalization parameters (α, τ). To obtain a unique solution, a measure of dissimilarity on the normalized descriptors is utilized as follows:

$$d[Z^*[k]_{ip}, Z^*[k]_{jp}] = \|Z^*[k]_{ip} - Z^*[k]_{jp}\|_2. \tag{9}$$

The pairs (α_{ip}, τ_{ip}) and (α_{jp}, τ_{jp}) leading to the normalized descriptors with minimum dissimilarity are chosen. This measure may yield solutions representing very large rotations. Such rotations are not usual in the addressed application and, therefore, they are penalized. The (α_{ip}, τ_{ip}) and (α_{jp}, τ_{jp}) leading to a solution close to the minimum one but involving a smaller rotation are finally selected:

$$d[Z^*[k]_{ip}, Z^*[k]_{jp}]_{\min \alpha} < 2d[Z^*[k]_{ip}, Z^*[k]_{jp}]_{minimum} \tag{10}$$

where the factor two has been found heuristically. The need for this final step can be assessed in the examples of Fig. 3. The first example shows a region obtained from the segmentation of two frames of the "table tennis" sequence (a finger of the player). The minimization of (9) yields $\Delta\alpha = \alpha_{ip} - \alpha_{jp} = 180°$, which clearly does not correspond with a natural motion in this case. When applying the minimization procedure of (10), $\Delta\alpha = 0.03°$ is obtained. Nevertheless, if the region evolution strongly corresponds to a large rotation, this minimization procedure still assigns to the region the correct α value. In the second example, a synthetic region that has been rotated is shown. In this case the estimated value is $\Delta\alpha = 180°$, which corresponds with the observed motion.

After this final normalization, 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
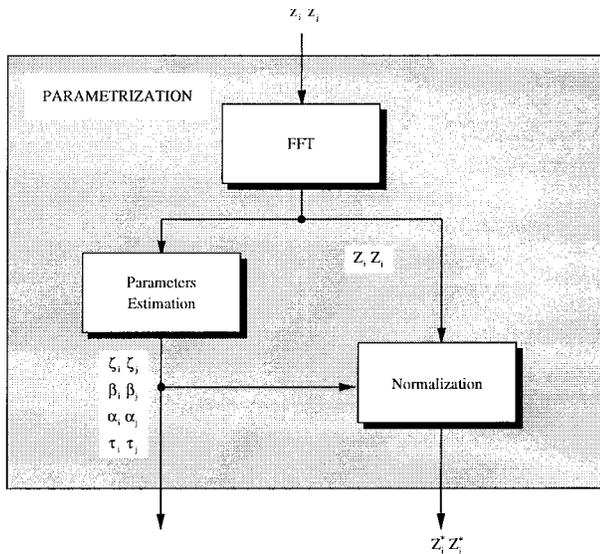
Fig. 4.   Block diagram of the parameterization procedure.

$(Z^*[k])$ is associated to the shape deformation. The complete parameterization procedure is shown in a block diagram in Fig. 4.

C. Parameterization of Conflictive Regions

The goal in the case of conflictive regions is to find a way to parameterize them so that they can be handled as nonconflictive regions in the posterior steps of the prediction procedure. That is, a set of parameters has to be given to every conflictive region even in the partition where they do not appear.

A conflictive region can be produced by several mechanisms: a global movement of the camera (panning, zoom, tilt, ...), the specific motion of an object that appears in (disappears from) the scene and covers (uncovers) an area, or even errors in the segmentation procedure. From this set of possibilities, the only one that allows assuming an *a priori* behavior is the case of global movement. In this case, a conflictive region is assumed to follow a coherent motion with respect to the motion of its neighbor regions. Thus, if the region with label p is present in the partition P_i (R_{ip}) and absent in the partition P_j, the parameters of R_{jp} will be computed relying on

- the parameters of R_{ip};
- the evolution of the parameters of its neighbor regions from P_i to P_j.

This concept is implemented by computing first the regular motion of the neighbor regions ($\Delta\zeta$, $\Delta\beta$, $\Delta\alpha$). Note that $\Delta\tau$ does not characterize any real motion, and it has been computed only to correctly estimated $\Delta\alpha$. In order to decide which is the dominant type of motion of the neighbor regions, the variance of these parameters among the neighbor regions is calculated. The parameter that has the smallest variance is said to be associated to the dominant motion.

To assign a regular motion to the conflictive regions, a weighted median of the dominant motion values is used. A median filer is used so that the conflictive region adopts one of the already existing motion values. Weights linearly depend

on the amount of contour points that are shared between the conflictive region and its neighbors. Note that no assumption can be made on the shape deformation parameters in the case of conflictive regions. Therefore, these parameters are set to be constant.

IV. REGION PREDICTION

The evolution of the previous parameters is predicted in order to obtain the contours of the predicted regions. Several techniques have been analyzed for predicting the regular motion and the shape deformation.

A. Prediction of the Regular Motion Parameters

From the set of parameters (ζ, β, α, τ), τ does not need to be interpolated since the starting point of the position function does not affect to the final representation of the predicted contours. The other three parameters are related to translation, zooming and rotation, respectively. These types of motion are assumed to be constant between the two known partitions (P_i and P_j). Linear prediction is therefore used for regular motion parameters in order to fulfill the constant evolution requirement. The use of nonlinear prediction techniques has been also analyzed. However, such techniques translate into acceleration or deceleration of the objects.

Assuming that the two known partitions are P_0 and P_T, the values of the parameters of a region belonging to a predicted partition R_{tp} are computed as

$$\zeta_{tp} = \left(1 - \frac{t}{T}\right)\zeta_{0p} + \frac{t}{T}\zeta_{Tp} \tag{11}$$

$$\frac{1}{\beta_{tp}} = \left(1 - \frac{t}{T}\right)\frac{1}{\beta_{0p}} + \frac{t}{T}\frac{1}{\beta_{Tp}} \tag{12}$$

$$\alpha_{tp} = \left(1 - \frac{t}{T}\right)\alpha_{0p} + \frac{t}{T}\alpha_{Tp}. \tag{13}$$

The reason for using $1/\beta$ instead of β is that, following the definition of (4), the normalization is carried out with the inverse of the size parameter. Consequently, in order to have a constant evolution of the zooming, the inverse of the parameter has to be used.

The above expressions lead to a noncausal prediction (interpolation) in the case of $0 < t < T$. In turn, the case of $t > T$ defines a causal prediction (extrapolation). In the extrapolation case, the evolution of the size parameter has to be controlled since it may become negative, which does not correspond with any natural zooming. The fact that $1/\beta$ becomes negative indicates that the linear model for the size evolution of this region is unsuitable. A possible alternative would be to replace the linear model by a geometric one. However, given that this situation appears very seldom, and for the sake of simplicity, the chosen solution is to set the parameter to zero:

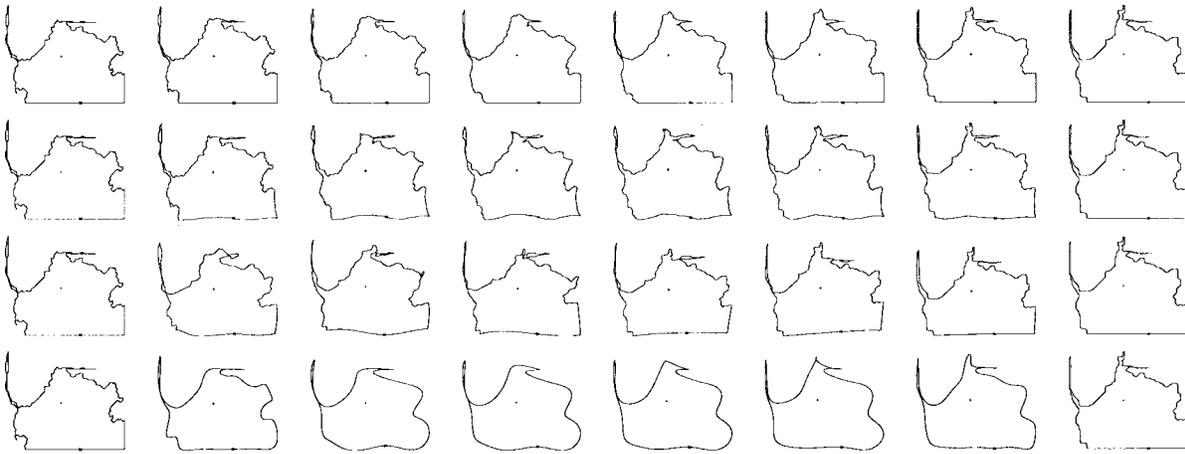$$\frac{1}{\beta_{tp}} < 0 \Longrightarrow \frac{1}{\beta_{tp}} = 0. \tag{14}$$

Fig. 5. Comparison between Cartesian (first row), polar (second row), high-frequency substitution (third row), and high-frequency elimination (fourth row) interpolation techniques for shape deformation.
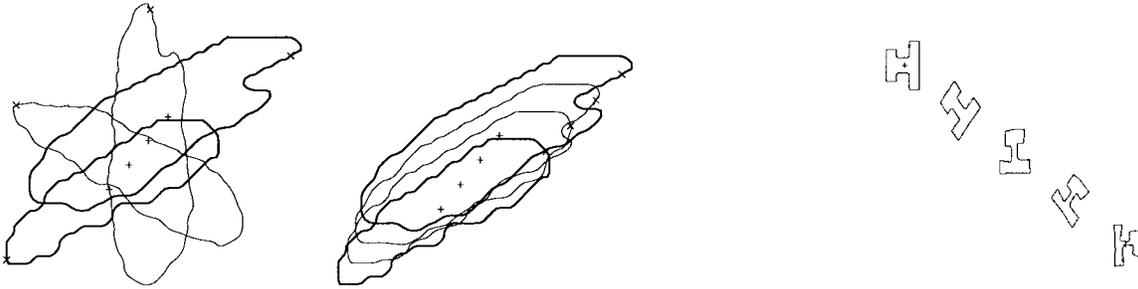

Fig. 6. Example of interpolation with and without constraints in rotation.

B. Prediction of the Shape Deformation Parameters

Four different techniques for predicting the normalized Fourier descriptors $Z^*[k]$ have been tested: Cartesian or linear prediction, polar prediction, high-frequency substitution [2], and high-frequency elimination [2]. Causal prediction cannot be done in the case of using either the high-frequency substitution or the high-frequency elimination techniques. In addition, best results, in terms of deformation conforming to natural object motion, have been obtained using the linear prediction

$$Z_{tp}^*[k] = \left(1 - \frac{t}{T}\right) Z_{0p}^*[k] + \frac{t}{T} Z_{Tp}^*[k]. \tag{15}$$

The performance of the four techniques are compared in the example of Fig. 5. Here, the different prediction results of a region from the "Miss America" sequence are analyzed. In all cases, regular motion has been linearly interpolated. Although the evolution of the regions with all techniques is quite natural, the only prediction technique that preserves line segments and corners is the Cartesian interpolation. Moreover, self-intersecting contours are more likely to appear in the polar and high-frequency substitution approaches.

The first example of Fig. 6 shows the interpolation of the regions presented in the example of Fig. 3. This example illustrates the importance of estimating correctly α and τ. In the first case, the parameters have been obtained with the technique leading to the minimum of the measure of dissimilarity (9), whereas in the second case, the solution constrained in rotation (10) has been used. This second solution is also presented for the case of a synthetic region that has been rotated. Note that the second solution yields a more natural motion.

The second example of Fig. 6 presents the interpolation of the synthetic region presented in Fig. 3. In this case, one of the two initial regions have been deformed to show the correct estimation of the rotation angle in the presence of large rotations and deformations.

C. Complete Region Prediction

Once the parameters describing the regular motion and the shape deformation have been predicted separately, they have to be combined in order to obtain the final predicted regions. This is actually done as a denormalization procedure using the predicted parameters

$$Z_{tp}[k] = \mathcal{D}_{-\zeta_{tp}} \mathcal{S}_{1/\beta_{tp}} \mathcal{R}_{-\alpha_{tp}} Z_{tp}^*[k]. \tag{16}$$

The final position function $z_{tp}[n]$ is computed as the inverse Fourier transform of the Fourier descriptors $Z_{tp}[k]$. The complete prediction scheme is presented as a block diagram in Fig. 7. Note that this scheme can also be directly applied to conflictive regions since the parameterization procedure has given a complete set of parameters for every region. Nevertheless, conflictive regions are not extrapolated from the partition in which they do not appear.
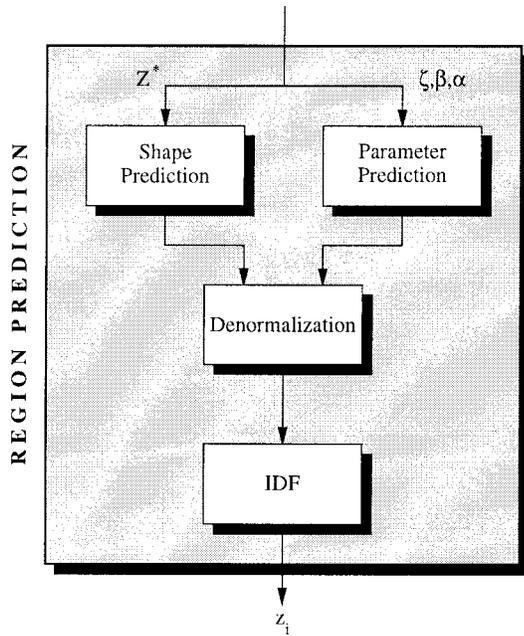
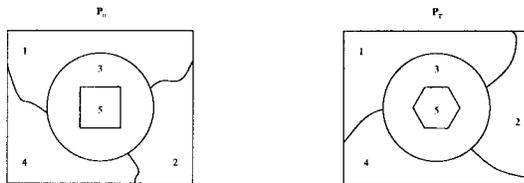Fig. 7. Block diagram of the region prediction procedure.

Fig. 9. Block diagram of the ordering procedure.

Fig. 8. Example of a partition evolution that would lead to an uncorrect ordering due to the presence of an interior region.

V. REGION ORDERING

In order to create the predicted partitions, regions that have been separately treated are combined. The direct combination of these regions rises two different problems. First, regions may partially overlap. Therefore, a priority has to be fixed to decide to which region have to be assigned those pixel laying on the overlapping area. Second, some areas of the predicted partition may not be covered by any region. As a consequence, a procedure has to be implemented to labeled every pixel in the partition. In this section, the problem of overlapping regions is discussed.

A. Ordering for Nonconflictive Regions

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.

The ordering is obtained by computing a distance on the normalized contours $z_{ip}^*[n]$ and $z_{jp}^*[n]$. Different distances have been analyzed and the best results, in terms of visual quality, have been achieved with the L_∞ norm. This distance
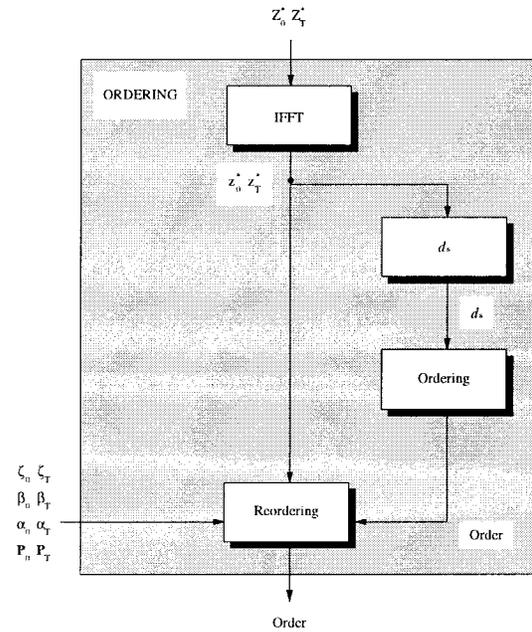
penalizes strong local differences between both shapes rather than small global variations of the regions:

$$d_s[z_{ip}^*, z_{jp}^*] = \|z_{ip}^*[n] - z_{jp}^*[n]\|_\infty. \tag{17}$$

The region whose evolution leads to the lowest value of expression (17) has the highest priority and receives an order level O_1.

B. Ordering for Conflictive Regions

A conflictive region should have always lower priority than any nonconflictive region. If the nonconflictive region with lowest priority has an order level $O_{\max}$, a conflictive region R_q will receive the order level $O[q] = O_{\max} + O[p]$. $O[p]$ represents the ordering of the nonconflictive region R_p from which the conflictive region R_q had obtained its regular motion parameters (see Section III-C). This ordering is, therefore, only used to give priority to a conflictive region with respect to other conflictive ones.

C. Ordering of Interior Regions

Interior regions should be handled as a special case in the ordering process. As stated in Section III, a region is characterized by its exterior boundary. With this characterization, the above ordering may result in the elimination of interior regions from the predicted partitions.

The problem is illustrated with an example. In Fig. 8, two partitions P_0 and P_T are shown. Region R_5 is interior to region R_3 and has suffered a strong deformation, whereas region R_3 has not. Therefore, the above ordering will penalize region R_5 in front of region R_3 and region R_5 will be covered by R_3 in all predicted partitions.

In order to solve the previous problem, a second step has been introduced in the ordering procedure. This step checks if a region has completely disappeared from the predicted partition

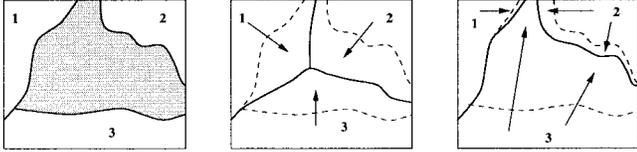Fig. 10. Example of areas that are not covered after combining the interpolated regions.

Fig. 11. Direct geodesic dilation of interpolated regions and its improvement by adapting the structuring element to the region priority.
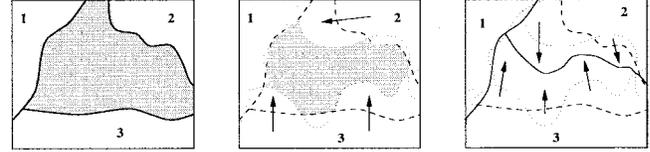

Fig. 12. Geodesic dilation of interpolated regions using masks and adapting the structuring element to the region priority.

due to its interior nature. If this is the case, a new priority level is assigned to the interior region higher than the priority level of the region that surrounds it. The complete ordering procedure is shown as a block diagram in Fig. 9.

VI. PARTITION CREATION

Predicted partitions are created by combination of the various predicted regions respecting the previous ordering. However, there may be holes after combining the regions. Such holes should be covered by neighbor regions to obtain a final partition. An example of this problem is presented in Fig. 10 where two partitions between frames 60 and 65 from the "car phone" sequence have been predicted as well as a new one corresponding to frame 67. In this example, holes appearing after combining the predicted regions are presented as black areas.

Different techniques have been studied to fill these holes. A simple solution is to dilate the predicted regions so that their labels cover the holes. For each interpolated region, its geodesic dilation of size one is performed [26]. This geodesic dilation uses as reference space the union of the region to be dilated and the area covered by its neighbor holes. It is applied first to the regions with lower level of priority in the previous ordering. That is, the more reliable is the region representation, the smaller its variation should be. This procedure is iterated until all the holes are covered.

The previous technique is illustrated in Fig. 11 where the dark area corresponds to nonassigned pixels and the region ordering follow $O[R_1] < O[R_2] < O[R_3]$. As it can be seen, this technique rises two main problems. First, in the case of a hole between two regions (R_1 and R_2 in the example), roughly half of the hole is covered by each region, regardless of their priority level. Second, dilated regions may cover areas very distant from the position of the predicted regions (R_3 in the example).

The first problem is solved by adapting the size of the geodesic dilation to be applied to each region with respect to its level of priority. Therefore, in the same iteration step, a region with low priority will be dilated by a structuring element

larger than the structuring element used for a region with high priority. The result of adapting the structuring element is also shown in Fig. 11. Note that this adaptation, although solving the first problem, may worsen the problem of dilated region covering areas very distant from its original position. That is, the shape of R_1 has almost remained the same whereas R_3 has extended up to distant areas.

To solve this second problem, the dilation is carried out in two steps. First, a mask is created marking the areas that more naturally a region can cover. To create these masks, regions from both known partitions are predicted keeping fixed their shape deformation parameters

$$Z_{(0 \to t)p}[k] = \mathcal{D}_{-\zeta_{tp}} \mathcal{S}_{1/\beta_{tp}} \mathcal{R}_{-\alpha_{tp}} Z_{0p}^*[k] \tag{18}$$

$$Z_{(T \to t)p}[k] = \mathcal{D}_{-\zeta_{tp}} \mathcal{S}_{1/\beta_{tp}} \mathcal{R}_{-\alpha_{tp}} Z_{Tp}^*[k] \tag{19}$$

$$Mask_{tp} = Area\{Z_{(0 \to t)p}[k]\} \cup Area\{Z_{(T \to t)p}[k]\}. \tag{20}$$

Only the holes inside the mask will be covered in the first step. The reference space for the geodesic dilation of region R_p is formed by the intersection of the holes and the mask $Mask_{tp}$. In the second step, the remaining holes are covered. In both steps, the structuring elements used in the geodesic dilations are adapted to the region priority. Fig. 12 shows the result of applying this procedure to the previous example. The expansion of region R_2 in the first step prevents the over-growing of region R_3 in the second step.

In the process of creating a partition, conflictive regions are not dilated. Note that the evolution of a conflictive region makes it disappear. Therefore, it does not seem natural that a conflictive region predicted in an intermediate partition could be larger than in the initial partition.

VII. APPLICATION TO PARTITION INTERPOLATION

The objective of this section is to show the capability of the proposed algorithm to interpolate partitions [14]. To assess the algorithm performance, interpolated partitions from various sequences are presented. The test sequences are "foreman," car phone, and "car-CCETT." In all cases, the initial segmentation has been performed by the algorithm described in [23]. Original sequences have a frame rate of 25 frames/s and the

Initial Partition P + 0	Interpolated Partition P + 1	Interpolated Partition P + 2	Interpolated Partition P + 3	Interpolated Partition P + 4	Initial Partition P + 5
Initial Partition P + 5	Interpolated Partition P + 6	Interpolated Partition P + 7	Interpolated Partition P + 8	Interpolated Partition P + 9	Initial Partition P + 10
Initial Partition P + 10	Interpolated Partition P + 11	Interpolated Partition P + 12	Interpolated Partition P + 13	Interpolated Partition P + 14	Initial Partition P + 15

Fig. 13.   Relation between new partitions and the technique used to create them.

Fig. 14.   Interpolation from four initial frames of the foreman sequence (210, 215, 220, and 225).

segmentation has been applied at five frames/s. That is, one out of five consecutive frames has been segmented.

In the examples, four partitions for each sequence have been obtained by means of segmentation and, between each two consecutive partitions, four new ones have been interpolated so that the final partition sequences has the original frame rate. The way to present these results is illustrated in Fig. 13.

In order to evaluate the interpolation results, each region in the partitions has been filled with the mean value of the gray-level value of its pixels. Note that in all cases the evolution of the images is soft and natural. For instance, the tilt of the head in the foreman sequence (Fig. 14) is perfectly represented in the interpolated images. In the car phone sequence (Fig. 15), the interpolation technique is able to cope with both the approximation of the head toward the camera as well as the motion of the mouth. Finally, Fig. 16 presents the interpolation results of the sequence car-CCETT. In this case, the turn of the car is correctly represented by the partition interpolation approach.

In the previous examples, the quality of the results depends on the performance of the segmentation algorithm that has been used. Another way to evaluate the algorithm is by using sequences segmented either by hand or by using chroma-keying techniques. With this kind of sequence, since the shape of the objects should conform to real shapes, the performance of the interpolation technique can be assessed from a new point of view. Fig. 17 shows the result of interpolating the masks of the "weather" sequence provided by MPEG. In the example, the person is moving backward while moving her hands. Note that, even with an interpolation factor of five, the evolution of the shape is naturally interpolated.

In addition, the use of the MPEG masks allows the illustration of the limitations of the algorithm. In Fig. 18, the interpolation of a frame from the weather sequence is presented. The shape of the person around this frame suffers a strong, local deformation. The woman is turning while moving both hands. In addition, some documents that the woman is holding appear in the mask, leading to a large deformation which is difficult to predict.

In Fig. 18, two different interpolations of the central frame, with respect to this motion, are presented. In the first case, the interpolation factor is five. It can be observed that this local deformation is not correctly interpolated and the complete shape is deformed. In the second case, an interpolation factor

Fig. 15. Interpolation from four initial frames of the sequence car phone (100, 105, 110, 115).
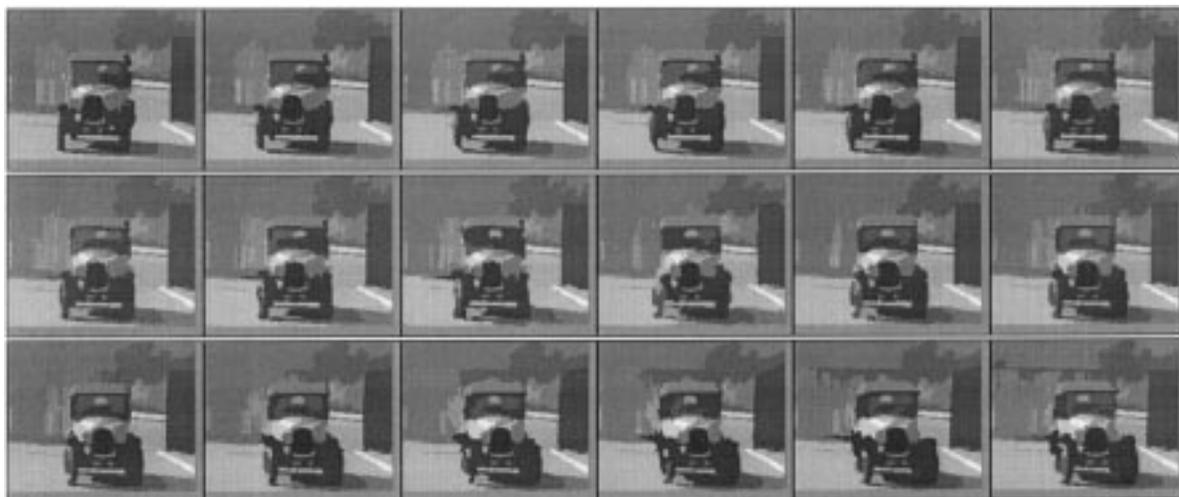

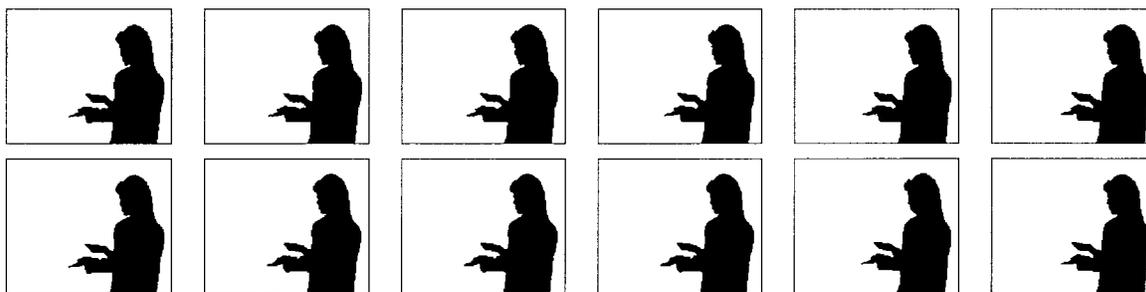Fig. 16. Interpolation from four initial frames of the sequence car-CCETT (50, 55, 60, 65).

Fig. 17. Interpolation between the original masks of the frames 220 and 225 from the weather sequence.

of three has been used. Here, the quality of the interpolated shape has improved and the global deformation is acceptable.

VIII. APPLICATION TO PARTITION CODING

The objective of this section is to show the capability of the proposed algorithm to extrapolate partitions. The previous partition prediction can be used in the framework of partition coding [15]. As stated in the Introduction, the proposed

partition prediction used in the partition coding approach is presented in [21]. The complete coding scheme is here briefly summarized as follows.

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 [5], [9], and [21] because it does not imply the use
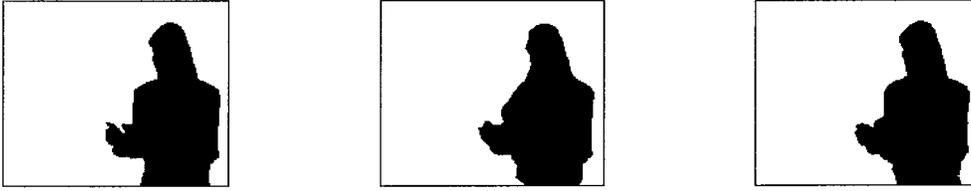
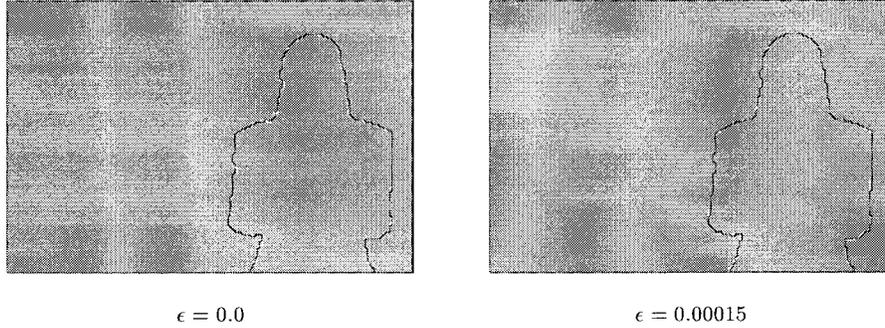Fig. 18.   Interpolation of frame 187 of the weather sequence using interpolation factor of five and three.

$\epsilon = 0.0$ $\epsilon = 0.00015$

Fig. 19.   Comparison between predictions using all ($\epsilon = 0.0$) or only the low-frequency ($\epsilon = 0.000\,15$) normalized descriptors.

of a single set of motion parameters to predict both the textures and the partition.

2) *Simplification of the prediction error:* An *overpartition* is defined by merging the predicted partition with the current partition. Each region of the overpartition 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 [17].

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. 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.

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 has to correctly represent the region. Therefore, only a small reduction on the total energy of the normalized descriptors is allowed, as follows:

$$\frac{\sum_{k=0}^{N^{LF}-1} \|Z_p^{*LF}[k]\|^2}{\sum_{k=0}^{N-1} \|Z_p^{*}[k]\|^2} \geq 1 - \epsilon. \tag{21}$$

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). Such a small 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 zero and T by means of the following expression:

$$Z_{tp}^{*LF}[k] = \left(1 - \frac{t}{T}\right) Z_{0p}^{*LF}[k] + \frac{t}{T} Z_{Tp}^{*LF}[k], \tag{22}$$

Fig. 19 shows the prediction obtained for the binary shapes of frames 61 of the weather sequence. 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 ($\epsilon = 0.000\,15$). In Fig. 19, the actual and predicted shapes are presented in black and white, respectively.

This partition prediction technique has been tested on binary shape sequences as well as on complete partition sequences.
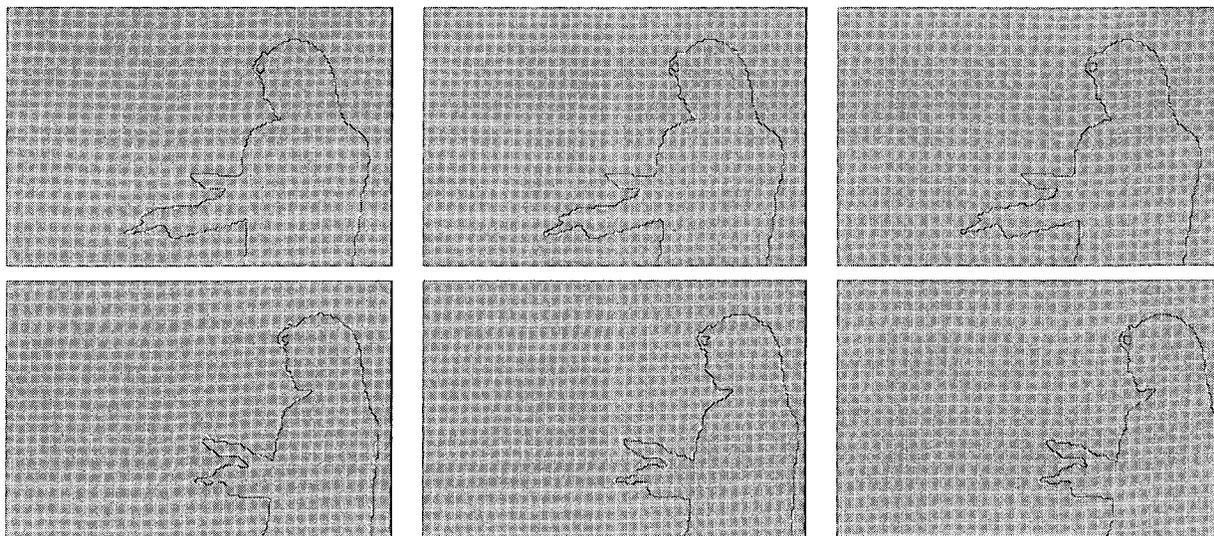
Fig. 20. Example of binary shape prediction (first row: frames 206, 208, 210; second row: frames 232, 234, 236).
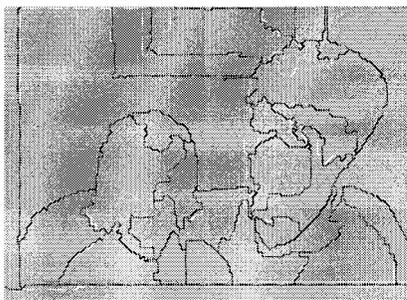

Fig. 21. Coding of the partition of frame 8 of the mother and daughter sequence.

Here, some results using the proposed technique on both kinds of sequences are presented.

A. Binary Shape Prediction

In Fig. 20, the prediction of six frames of the weather sequence is 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 Fig. 20 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.

B. Complete Partition Prediction

Fig. 21 shows the prediction of a complete partition of frame 8 from the "mother and daughter" sequence. The prediction relies on the partitions of frames 0 and 4, which have been segmented using the technique presented in [19]. In this case, the simplification has been applied so that all the prediction errors leading to regions with less than five 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 one frame skipped between used frames, while here three frames are skipped.

IX. CONCLUSIONS

The examples of the previous section illustrate the fact that the partition prediction technique that has been proposed performs correctly in different scenarios. This prediction technique has been tested on a large set of sequences and using different segmentation approaches. In all cases, the achieved results show similar quality as those above presented.

The prediction technique here discussed only uses partition information and, therefore, it is independent of the segmentation approach that is utilized to obtain the initial partitions. Nevertheless, the performance of the proposed technique improves in both visual quality and computational load when the temporal coherence of the initial partitions increases. Furthermore, the technique is totally unsupervised and no additional information is necessary apart from the initial partitions.

The visual quality improvement is due to the fact that if the initial partitions track correctly the evolution of objects in the scene, the prediction procedure yields smooth transitions that are more pleasant to the human visual system.

The reduction of the computational load comes from the fact that if the initial partitions present a strong time coherence, the number of holes after combining the interpolated regions decreases, and so does their size. In this case, a simplification of the algorithm presented in Section VII can be used, which further reduces the required computation.

REFERENCES

[1] J. Benoia, L. Wu, and D. Barba, "Joint contour-based and motion-based image sequences segmentation for TV image coding at low bit rate," in *Proc. Visual Communication and Image Processing,* Chicago, IL, Sept. 1994, pp. 1074–1085.

[2] O. Bertran, R. Queval, and H. Maitre, "Shape interpolation using Fourier descriptors with application to animation graphics," *Signal Process.,* vol. 4, pp. 53–58, 1982.

[3] P. Bouthemy and E. François, "Motion segmentation and qualitative dynamic scene analysis from an image sequence," *Int. J. Comput. Vis.,* vol. 10, pp. 157–182, 1993.

[4] R. Bremond and F. Marqués, "Segmentation-based morphological interpolation of partition sequences," in *Mathematical Morphology and its Applications to Image and Signal Processing,* P. Maragos, R. Schafer, and M. Butt, Eds. Boston, MA: Kluwer, 1996, pp. 369–376.

[5] P. Brigger and M. Kunt, "Contour image sequence coding using the geodesic morphological skeleton," in *Proc. Int. Workshop on Coding Techniques for Very Low Bit-Rate Video,* Colchester, U.K., Apr. 1994, pp. 3.1–3.2.

[6] F. Dufaux and F. Moscheni, "Segmentation-based estimation for second generation video coding techniques," in *Video Coding: The Second Generation Approach,* L. Torres and M. Kunt, Eds. Boston, MA: Kluwer, 1996, pp. 219–264.

[7] F. Etesami and J. J. Uicker, "Automatic dimensional inspection of machine part cross-sections using Fourier analysis," *Comput. Vis., Graph., Image Process.,* vol. 29, pp. 216–247, 1985.

[8] G. H. Granlund, "Fourier preprocessing for hand print character recognition," *IEEE Trans. Comput.,* vol. C-21, pp. 195–201, 1972.

[9] C. Gu and M. Kunt, "Contour simplification and motion compensation for very low bit-rate video coding," in *Proc. 1st IEEE Int. Conf. Image Processing,* Austin, TX, Nov. 1994, pp. 423–427.

[10] ISO/IEC JTC1/SC29/WG11, MPEG-4 Proposal Package Description (PPD), July 1995.

[11] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Comput. Vis., Graph., Image Process.,* vol. 18, pp. 236–258, 1982.

[12] M. Kunt, A. Ikonomopouloa, and M. Kocher, "Second generation image coding techniques," *Proc. IEEE,* vol. 73, pp. 549–575, Apr. 1985.

[13] B. Marcotegui and F. Meyer, "Morphological segmentation of image sequences," in *Mathematical Morphology and its Applications to Image Processing,* J. Serra and P. Soille, Eds. Boston, MA: Kluwer, 1994, pp. 101–108.

[14] F. Marqués, B. Llorens, and A. Gasull, "Interpolation and extrapolation of image partitions using Fourier descriptors: Application to segmentation based coding schemes," in *Proc. IEEE Int. Conf. Image Processing,* Washington, DC, Oct. 1995, vol. III, pp. 684–687.

[15] ———, "Partition prediction for partition coding," in *Proc. EUSIPCO 96, VIII Europ. Signal Processing Conf.,* Trieste, Italy, Sept. 1996.

[16] F. Marqués, M. Pardàs, and P. Salembier, "Coding-oriented segmentation of video sequences," in *Video Coding: The Second Generation Approach,* L. Torres and M. Kunt, Eds. Boston, MA: Kluwer, 1996, pp. 79–124.

[17] F. Marqués, J. Sauleda, and A. Gasull, "Shape and location coding for contour images," in *Proc. Picture Coding Symp.,* Lausanne, Switzerland, Mar. 1993, pp. 18.6.1–18.6.2.

[18] F. Meyer, "A morphological interpolation method for mosaic images," in *Mathematical Morphology and its Applications to Image and Signal Processing,* P. Maragos, R. Schafer, and M. Butt, Eds. Boston, MA: Kluwer, 1996, pp. 337–344.

[19] M. Pardàs and P. Salembier, "Joint region and motion estimation with morphological tools," in *Proc. 2nd Workshop on Mathematical Morphology and its Applications to Signal Processing,* J. Serra and P. Soille, Eds., Fontainebleau, France, Sept. 1994, pp. 93–100.

[20] E. Persoon and K. S. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-7, pp. 170–179, 1977.

[21] P. Salembier, "Motion compensated partition coding," in *Proc. SPIE Visual Communication and Image Processing,* Orlando, FL, Mar. 1996, vol. 2727, pp. 403–415.

[22] P. Salembier *et al.,* "Segmentation-based video coding system allowing the manipulation of objects," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 7, pp. 60–74, Feb. 1997.

[23] P. Salembier and M. Párdas, "Hierarchical morphological segmentation for image sequence coding," *IEEE Trans. Image Processing,* vol. 3, pp. 639–651, Sept. 1994.

[24] P. Salembier *et al.,* "Region-based video coding using mathematical morphology," *Proc. IEEE,* vol. 83, pp. 843–857, June 1995, invited paper.

[25] H. Sanson, "Joint estimation and segmentation of motion video coding at very low bitrates," in *Proc. COST 211th Europ. Workshop on New Techniques for Coding of Video Signals at Very Low Bitrates,* Dec. 1993, pp. 2.2.1–2.2.8.

[26] J. Serra, *Image Analysis and Mathematical Morphology.* New York: Academic, 1982.

[27] P. Soille, "Generalized geodesic distances applied to interpolation and shape description," in *Mathematical Morphology and its Applications to Image Processing,* J. Serra and P. Soille, Eds. Boston, MA: Kluwer, 1994, pp. 193–200.

[28] ISO-IEC CD 13818 Information Technology, MPEG-2, "Generic coding of moving pictures and associated audio," Nov. 1993.

[29] L. Torres and M. Kunt, *Video Coding: The Second Generation Approach.* Boston, MA: Kluwer, 1996.

[30] P. Van Otterloo, *A Contour-Oriented Approach for Shape Analysis.* Englewood Cliffs, NJ: Prentice-Hall, 1991.

[31] T. P. Wallace and P. A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Comput. Graph. Image Process.,* vol. 13, pp. 99–126, 1980.

Ferran Marqués (S'91–M'93) graduated from the Polytechnic University of Catalonia (UPC), Barcelona, Spain, in December 1988. He received the Ph.D. degree from UPC in December 1992.

From 1989 to June 1990, he was with the Swiss Federal Institute of Technology, Lausanne working in digital image sequence processing and coding. In June 1990, he joined the Department of Signal Theory and Communications, UPC. From June 1991 to September 1991, he was with the Signal and Image Processing Institute, University of Southern California, Los Angeles. Since 1995, he has been an Associate Professor at UPC. He is also Vice-Dean of International Affairs in the Electrical Engineering School, UPC. He has served as Officer responsible for the task of membership development with the European Association for Signal Processing (EURASIP) since 1994. His main areas of interest are still image and sequence segmentation, still image and sequence object-based coding, motion estimation and compensation, mathematical morphology, and biomedical applications.

Dr. Marques received the Spanish Best Ph.D. Thesis on Electrical Engineering Award in 1992. Since 1996, he has served as Associate Editor in the area of image communications for the *Journal of Electronic Imaging.*

Bernat Llorens was born in Barcelona, Spain, on June 16, 1970. He received the Telecommunication Engineering diploma from the Polytechnic University of Catalonia , Barcelona, in 1995.

In 1996, he joined Indra Espacio S.A., where he is on a research team developing a CDMA-based communications modem. His interests include image processing and signal processing for communication applications.

Antoni Gasull (M'84) received the Telecommunication Engineering diploma and the Ph.D. degree in telecommunications, both from the Polytechnic University of Catalonia (UPC), in 1978 and 1983, respectively.

He was a Visiting Professor during the first semester of 1985–1986 at the University of Rhode Island, Kingston. In 1986, he co-founded the image processing group, Signal Theory and Communications Department, UPC, which has been involved in several national and European research projects, and several technical contributions in object-based image coding. He is currently Vice Chairman and Doctorate Program Director of the Signal Theory and Communications Department. His initial research interests were in spectral estimation and array processing. His current interests are image analysis, document image analysis, and biomedical image processing applications.