SEGMENTATION BASED CODING OF TEXTURES USING STOCHASTIC VECTOR QUANTIZATION

L. Torres   J. R. Casas   S. de Diego

Dept. of Signal Theory and Communications

Universitat Politècnica de Catalunya

P.O. Box 30.002 - 08071 Barcelona, SPAIN

ABSTRACT

In second generation image compression techniques the image to be compressed is first segmented. The pixels are divided into mutually exclusive spatial regions based on some criteria. After segmentation, the image consists of regions separated by contours. Then, the information is coded describing the shapes and interiors of the regions. The interiors of the regions are usually encoded using polynomials. The objective of this paper is to encode the interior of the regions by Stochastic Vector Quantization techniques. If the segmentation process has been well defined and the obtained regions are homogeneous, then it is possible to design a specific codebook suited to the statistics of each region. The approach is to design the codebook according to some previously defined model for the regions of the image found in the segmentation process. If the approach is combined with efficient contour coding techniques, good visual results for high compression rates are obtained.

1. INTRODUCTION

A major objective of image coding is to represent an image with as few bits as possible while preserving the level of quality and intelligibility required for the given application. Image coding has two major areas. One is the reduction of channel bandwidth required for image transmission systems. The other application is reduction of storage requirements. Although the standardization process in image compression is currently very well defined for a variety of applications, for instance, JPEG, H.261, and MPEG standards, the emergency of new visual communications systems and the constant need for bandwidth savings, indicate that there is a need for more efficient compression schemes that should provide very high compression rates while keeping

image quality. On the other hand, a major breakthrough in image coding should rely on the use of segmentation-based techniques (second generation image coding techniques) and not on the classical pixel-based approach. As an example, in the context of the future MPEG 4 standard, it is generally accepted that it will be based in techniques relying on the concept of objects.

In segmentation-based image compression techniques the image to be compressed is first segmented [1]. The pixels are divided into mutually exclusive spatial regions based on some criteria. After segmentation, the image consists of regions separated by contours. Then, the information is coded describing the shapes and interiors of the regions. This description forms the object-based representation of the image. In most segmentation-based compression schemes, the shape of the image regions are represented by encoding the segment boundaries. The interiors of the regions are represented by encoding, for example, the coefficients in the polynomial models describing each region, or for flat regions, the average gray level of the pixels in each region [2].

In our approach the interior of the regions is coded using vector quantization techniques. The vector quantization scheme has proven to be very effective in speech and image coding but has not been applied to segmented images because, among other reasons, not an easy method was available for the generation of the codebook. In order to overcome the intrinsic problem concerning the design of the codebook when using the LBG algorithm [3], a new simple fixed-rate VQ image coding scheme, called Stochastic Vector Quantization (SVQ), has been proposed and successfully applied to non-segmented images [4] [5] [6]. In this paper we have extended the previous reported results of the SVQ to segmentation-based image coding techniques. The paper is organized as follows: Section 2 presents the segmentation algorithm used in our approach. Section 3 reviews the basics of the Stochastic Vector Quantization

scheme. Section 4 shows the SVQ applied to segmented images and Section 5 presents some results. Finally conclusions are presented.

2. SEGMENTATION BASED CODING

A segmentation-based image coding scheme (also called object-based scheme) involves at least three main processing steps. The first one is the segmentation which splits the original data into various components having a certain degree of homogeneity and correspond as much as possible to semantic unities. Then, it is necessary to describe or represent each component or object. In order to be efficient, the description must be achieved in two complementary steps: description of the object contours (contour coding) and description of the inside (texture coding). These three steps are illustrated in Fig. 1.
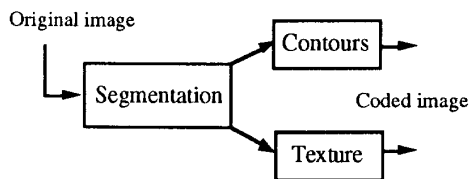

Figure 1. Segmentation based coding

The texture coding stage is usually done independently of the segmentation process [1]. However, if somehow the segmentation is made dependent on the texture coding, the quality of the overall scheme will improve. As the encoding of textures using Stochastic Vector Quantization techniques depends on the degree of homogeneity of the segmented regions, it is important to understand how the segmentation process is done. Besides, if progressive transmission or storage of data is necessary, a multiscale representation of the image is needed. Multiscale segmentation based image coding techniques have already been presented and proven to be useful to achieve a hierarchical decomposition of the image to be encoded [7] [8] [9]. The extension of the method to video sequences have been proposed in [10]. In the following a brief description of the segmentation process will be given for completeness purposes.

2.1 Multiscale Segmentation

The segmentation algorithm relies on tools based on Mathematical Morphology [11]. Mathematical Morphology is particularly suitable for an object-based approach because its methodology is

intimately based on the notion of object. Moreover, it leads to practical algorithms because it can be very efficiently implemented. The segmentation algorithm of interest is hierarchical following a Top-Down procedure. The algorithm starts with the most global point of view of the image and considers that the whole image represents a single region. The first level computes a simplified segmentation of the image. Then the successive levels will improve the current segmentation by introducing new regions which are judged to be pertinent at that particular stage.

All levels of the hierarchy have basically to perform a segmentation, therefore they involve the same elementary steps. Each level relies on three basic segmentation steps: preprocessing, feature extraction and decision. The objective of the preprocessing is to simplify the data to make them easier to segment. Feature extraction is used to assess the homogeneity of each local area. This processing step makes use of the simplified image and possibly of the information about the current segmentation. In morphology, the feature extraction is also known as marker extraction. A marker can be considered as a binary signal indicating the presence of an homogeneous area and thus of an object. It does not intend to solve the problem of contour localization. This is taken care of in the third step called decision. The decision step needs two different inputs: a signal to segment (for example the original data) and the set of markers defining the presence of objects. Its output is the segmentation result which can take the form of a contour or a label image. A contour image is a binary image with two classes of pixels: interior and contours. In the context of segmentation-based coding this representation is not the most appropriate because it leads to difficulties when a gray level or colour information has to be assigned to contour points on the receiver side. For this reason, a label representation is used. It consists of an image whose gray level values correspond to the partition classes of the image. An individual and arbitrary number is assigned to each region.

Then, the current segmentation is improved by the next level of the hierarchy which deals with the image components that have not been properly segmented in the previous levels. This remark leads to the use of a fourth elementary step aimed at the quality estimation of the current segmentation. Indeed, the quality estimation indicates to the following hierarchical level where the segmentation error is high and therefore where it has to be improved or refined. In our case, the ultimate goal of the segmentation is coding. This means that the quality estimation should indicate the areas which

have not been properly coded. It can be achieved by actually coding each region, that is by generating the image which will be reconstructed in the receiver side, and by computing the difference between the coded image and the original one. A high difference indicates the presence of a region which is poorly represented by the current segmentation and which should be better represented by the next levels. By contrast, a low difference characterizes a region which has been well represented by the current coding process. This difference image is called the coding residue. The choice of the model has a rather strong influence on the remaining levels. From our experience, a model able to represent complex gray level functions, leads to a final segmentation with fewer regions and to a more accurate representation of the image. It has become customary in object-based coding schemes to encode the texture using polynomials or transform methods. In our scheme, the Stochastic Vector Quantizer (SVQ) is used to encode the segmented regions what has proven to be an important improvement in visual quality at the expenses of obtaining smaller compression factors.

3. STOCHASTIC VECTOR QUANTIZATION

The Stochastic Vector Quantization (SVQ) scheme has been already presented to encode non-segmented images with good visual results [4] [5] [6]. The only difference of the SVQ with respect to the conventional vector quantizer is the design of the codebook. In the SVQ the codewords are generated by stochastic techniques instead of being generated by a training set representative of the expected input image. In order to generate the stochastic codewords, zero mean unity variance white-gaussian noise images of the same size as the subimages to be encoded, are passed through some shaping filter $H(z_1, z_2)$ whose output follows the selected model. Then, the filtered images are normalized to unity variance and used as codewords for the VQ scheme. In order to account for the variance information, a multiplicative scale factor has to be introduced aswell. This factor has to be encoded and sent to the receiver. Exactly as many stochastic images as the desired size of the codebook are filtered. With this approach, there is no need to use the LBG algorithm [3] for the design of the codebook, thus avoiding the need of a very computationally expensive way of generating the codebook. The model may be selected before the coding process or the coding scheme must first generate the model from the observed signal to be coded. In this context, the image data to be encoded, i.e., the subimages of size nxn, are considered as realizations of a certain stochastic process. If the codewords are designed according to realizations of this process, the subimages to be

encoded will find in the codebook, with a high degree of probability, a very similar codeword, thus resulting in an efficient coding. The quality of the coding scheme will rely on the quality of the model selected to design the codebook.

4. SVQ FOR SEGMENTED IMAGES

When applying the SVQ to segmented images, a model is found for each region of the segmented image. In our approach a very simple third order autoregressive model has been used to model each region. The coefficients of the filter are encoded using uniform 8 bit quantizers. We have noticed that better results are obtained if the coefficients of each filter are quantized before generating the stochastic codebook. In this way, we guarantee that exactly the same codewords are obtained in the emitter and in the receiver. The multiplicative factors needed to keep the variance information are DPCM encoded with an average of 3 bits. The mean of each region is approximated using a first order polynomial and encoding each polynomial coefficient with an average of 5 bits.

5. RESULTS

The SVQ scheme has been applied to the multilevel segmentation scheme explained above. The contours have been encoded using a modified lossless version of the well known chain code technique where 1.34 bits/contour pixel have been used [12]. In order to take advantage of the compression properties of the SVQ scheme, a codeword size of 8x8 has been selected. The square shape of the codewords have been matched to the irregular shape of the segmented regions. The mean square error has been used as the distortion distance between the original image and the codewords. Image 1 presents the original 256x256 8 bits image Cammeraman. Image 2 presents the encoded segmented image at a bit rate of 0.36 bits/pixel which means a compression rate of 22. Notice the excellent rendering of the contours due to the segmentation scheme and the good quality of the texture.

6. CONCLUSIONS

A new vector quantization encoding scheme has been presented to encode textures in a object-based second generation image coding technique. The design of the codebook has been based in a stochastic approach in order to match a previously defined model for the regions to be encoded. The segmentation scheme has been based in multiresolution scheme what improves the quality of the coded image. Compression factors of 22 have been achieved with good visual quality.

Current work is being performed to apply the proposed method to image sequences.

References

[1] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second Generation Image Coding Techniques", Proceedings of the IEEE, vol. 73, pp. 549-574, April 1985.

[2] M. Gilge, T. Engelhardt and R. Mehlan, "Coding of Arbitrarily Shaped Image Segments Based on a Generalized Orthogonal Transform", Image Communication, October 1989.

[3] Y.Linde, A.Buzo and R.M.Gray, "An Algorithm for Vector Quantizer Design". IEEE Transactions on Communications, Vol. 28, January 1980.

[4] L.Torres, E. Arias, "Stochastic Vector Quantization of Images", International Conference on Acoustics Speech and Signal Processing, San Francisco, March 24-27, 1992.

[5] L. Torres, J. Salillas, "Improvements on Stochastic Vector Quantization of Images", ICASSP 93, April 27-30, Minneapolis, USA.

[6] L. Torres, E. Arias, "Stochastic Generation of the Codebook in Vector Quantization of Images", Information Services & Use, IOS, Vol. 13, Number 1, 1993.

[7] P. Salembier and J. Serra, "Morphological Multiscale Segmentation of Images", Proc. of SPIE VCIP'92, Boston, MA, pp., 1992.

[8] L. Torres, P. Salembier, F. Marqués, P. Hierro, "Image Coding for Storage and Transmission Based on Morphological Segmentation", Europto, International Symposium on Video Communications and Fiber Optic Services, Berlin, Germany, April 4-9, 1993.

[9] P. Salembier, "Application of Mathematical Morphology to Picture Coding", Proceedings of General Assembly of URSI, Kyoto, Japan, 1993.

[10] MORPHECO, UPC - CMM - EPFL - LEP, "Morphological Segmentation Based Coding of Image Sequences", COST 211 ter Worhshop on New Techniques for Coding of Video Signals at Very Low Bit Rates, Hannover, Germany, December 1-2, 1993.

[11] J. Serra, "Image Analysis and Mathematical Morphology", Academic Press, 1982.

[12] F. Marqués, J. Sauleda, A. Gasull, "Shape and Location Coding for Contour Images" Proceedings of the Picture Coding Symposium 1993, Lausanne, March, 1993.

Image 1. Original image Cammeraman.

Image 2. Compressed image at 0.36 bits/pixel

V-600