1991 Picture Coding Symposium
September 1991, Tokyo, Japan

CODING OF COLOR PARAMETERS IN AN ANALYSIS–SYNTHESIS CODER BASED ON MOVING 3D–OBJECTS

Jörn Ostermann, Hailian Li
Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Appelstraße 9A, 3000 Hannover 1, F.R.Germany
Phone: 49–511–7625308

Abstract

For object–oriented analysis–synthesis coding [1] an image analysis algorithm is required which automatically generates the parameter sets which allow to synthesize an image sequence. In this contribution it is assumed that 3–dimensional quasi–rigid objects are arbitrarily moving in 3–dimensional space. Thus, the 3–dimensional real world is modelled by a 3–dimensional model world [1,2] . It is the goal of the modelling to generate a model world which has a model image identical to the real image.

With this model, the image analysis [3,4,5] finds moving objects in the scene and describes each object by means of shape, motion and color parameters. Furthermore, model failures, i.e. areas where the model assumption of quasi–rigid 3D–object does not hold, have to be detected [6], since these areas cannot be synthesized.

Model failures do not obey the source model. They are described by shape and color parameters and have to be coded. Shape parameters for model failures consist of a 2d–mask, their coding is described in [6,7]. Efficient coding of color parameters for areas of model failure will be discussed here. Since areas of model failure are arbitrarily shaped, a generalized orthogonal discrete cosine transform is used for coding color parameters. Typically the image signal of model failures is not stationary. Therefore a segmentation algorithm is introduced in order to partition the areas of model failure into highly correlated regions. Experimental results for coding of model failures for an object–oriented analysis–synthesis coder based on moving 3D–objects aiming at transmission rates of 64 kbit/s will be presented.

Concept and structure of object–oriented analysis–synthesis coding

Fig. 1 explains the concept and structure of the object–oriented analysis–synthesis coder [1]. Input of the coder is a sequence of images. The image analysis subdivides each image of this sequence into moving objects and describes each object by three sets of parameters defining motion, shape and color parameters of the object. Color parameters include luminance and chrominance signals. The object parameters depend on the type of source model being applied [8]. Image areas which cannot be described by the source model are denoted as model failures. Since motion is not defined for these objects they are described by 2D–shape and color parameters only. The calculated parameter sets are coded by parameter coding. The parameter coding depends on a receiver model which includes assumptions about the visibility of coding errors. The coded parameter sets are transmitted to the receiver, decoded by parameter decoding and stored in a parameter memory. The parameter memory of the coder and decoder contains the same parameter information and enables both the coder and

Reprint from Jörn Ostermann, "Coding of color parameters in an analysis–synthesis coder based on moving 3d–objects", *Picture Coding Symposium 1991 (PCS 91)*, pp. 10.5–1 – 10.5–4, Tokyo, Japan, 2 – 2 September 1991

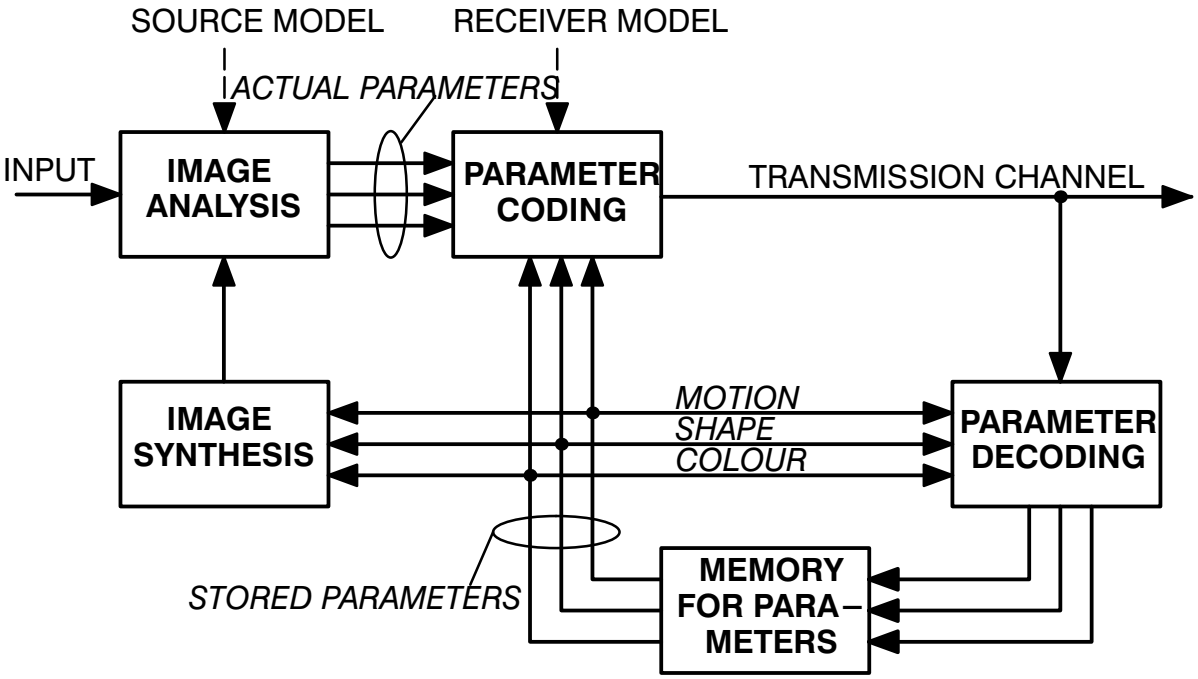

Fig. 1: Block diagramm of an object–oriented analysis–synthesis coder

decoder to identically reconstruct a transmitted image by image synthesis. The reconstructed image is displayed at the decoder and is also used for image analysis of the next input image at the coder. Additionally, the stored parameters are used for the coding and decoding of the parameter sets of the next input image. Here the source model of moving quasi–rigid 3D–objects is used.

Motion parameters are coded using 60 bits/object. Object shapes are currently described by silhouettes and require about 400 bits for each object silhouette. The number of bits needed for coding 3D–shape parameters is currently investigated. An object–oriented analysis–synthesis coder aiming at transmission rates of 64 kBit/s uses approximately 70% of the available data rate for parameter coding of model failures when applied to typical videophone test sequences.

Coding color information in areas of model failures

Model failure areas do not obey the source model and hence cannot be synthesized. They are described by 2D–shape parameters and color parameters. Usually model failures are image regions of arbitrary shape. An efficient transform coding technique for coding the image signal of these regions is the Discrete Cosine Transform for arbitrarily shaped image regions [9]. With typical videophone test sequences, model failures are mainly situated at opening and closing eyes and mouths of the speakers, i.e. model failures are areas with high contrast. Hence segmenting model failures into highly correlated regions can improve the efficiency of coding for model failures. If the area is segmented into regions, region boundaries have to be transmitted in addition to color para–meters.

Segmentation of color information into highly correlated regions

There are several algorithms for segmenting images into homogeneous regions [10,11]. Their use is determined by the application. Since the segmentation algorithm will be used to improve the coding efficiency of a DCT, its main goal is to subdivide areas of model failure into highly correlated regions. Graph–theoretic [12] and statistical [10] approaches have been investigated. It was found that the recursive Minimum Spanning

Reprint from Jörn Ostermann, "Coding of color parameters in an analysis–synthesis coder based on moving 3d–objects", *Picture Coding Symposium 1991 (PCS 91)*, pp. 10.5–1 – 10.5–4, Tokyo, Japan, 2 – 2 September 1991

Tree (MST) segmentation is well suited for the purpose of segmenting highly correlated regions for transform coding.

The segmentation algorithm contains an analysis and split part.
– In the analysis each pel of the image is first considered as one region. To each region the mean value of pels belonging to it is assigned as region value. Iterativly those two neighboring regions with the lowest difference in region values are merged to one new region with a new region value calculated as described above. Furthermore the position and size of the merged regions are noted in a sequential merging list, which can be interpreted as a minimum spanning tree (MST). This procedure stops as soon as one region describing the whole area to be segmented is left.
– In the split part, region merging is canceled in reverse order using the merging list. Canceling the last merge operation, the area is partitioned into two regions, canceling n merge steps results in a segmentation into n+1 regions.

A good quality measure of segmentation for this application is the mean square difference MSD_S between the original image and the segmented image where to each region of the segmented image the mean value of pels of the original image is assigned. A good segmentation will yield a lower MSD_S for a given number of regions than others. Given a segmentation into n regions MST segmentation ensures that refining this segmentation into n+1 regions is optimum with respect to the quality measure under the constraint that only one of the n regions may be subdivided into two.

Coding shape parameters of the segmentation

While the shape of model failures is coded using approximation techniques [6,7], segmentation boundaries have to be coded without errors. This is done using chain codes [13] and relative addressing of starting points for each chain [14].

Coding of DCT–coefficients

Segmented regions to be coded are transformed using a DCT for arbitrarily shaped image regions [9]. Coefficients are quantized using an adaptive quantizer requiring 2.5 bit/region side information. Quantized coefficients are coded using a two–dimensional variable length code [H261, CCITT].

Results

Experimental results obtained with coding color parameters of model failures in an object–oriented analysis–synthesis coder will be shown. Areas of model failures are segmented until the quality measure MSD_S of the segmentation is higher than 18 dB. Using the segmentation algorithm improves subjective image quality significantly. This is due to missing mosquito effects at contours in the coded image signal. The required bit rate does not increase when compared to coding model failures without segmentation. For a peak SNR of 34 dB in model failure areas an average of 1.3 bit/pel for shape and color are required. Assuming CIF, 10 Hz and 64 kbits/s, 30% of available data rate are required for coding motion and shape parameters of moving objects. 4500 bits/frame are available for coding color parameters of model failures. Hence, an object–oriented analysis–synthesis coder must keep the average area of model failures below 3.5% of image area.

Conclusion

An efficient algorithm for coding color parameters of arbitrarily shaped model failure areas has been presented. Prior to coding, model failures are segmented into highly correlated regions using a recursive MST segmentation. Coding color parameters require 1.3 bit/pel for a peak SNR of 34 dB. The use of the segmentation improves subjec-

Reprint from Jörn Ostermann, "Coding of color parameters in an analysis–synthesis coder based on moving 3d–objects", *Picture Coding Symposium 1991 (PCS 91)*, pp. 10.5–1 – 10.5–4, Tokyo, Japan, 2 – 2 September 1991

tive picture quality without increasing bit rate. Since the average area of model failure in an analysis–synthesis coder is below four percent of image area, a high quality coding algorithm for videophone applications seems to be within reach.

References:

[1] H.G. Musmann, M. Hötter, J. Ostermann, "Object–oriented analysis–synthesis coding of moving images",Signal Processing, Image Communication, No. 2, Nov. 1989.

[2] F. Kappei, C.–E. Liedtke, "Modelling of a natural 3–D scene consisting of moving objects from a sequence of monocular TV images", SPIE, vol. 860, Cannes 1987.

[3] J. Ostermann, "Modelling of 3D Moving Objects for an Analysis–synthesis Coder", SPIE/SPSE Symposium on Sensing and Reconstruction of 3D Objects and Scenes, Bernd Girod, editor, Proc. SPIE 1260, Santa Clara, California, 11–16 February 1990.

[4] C.–E. Liedtke, H. Busch, R. Koch, "Automatic modelling of 3D moving objects from a TV image sequence", SPIE/SPSE Symposium on Sensing and Reconstruction of 3D Objects and Scenes, Bernd Girod, editor, Proc. SPIE 1260, Santa Clara, California, 11–16 February 1990.

[5] R. Koch, "Adaptation of a 3D facial mask to human faces in videophone sequences using model based image analysis", to be published at PCS 91, Tokio, September 1991.

[6] J. Ostermann, H. Li, "Detection and coding of model failures in an analysis–synthesis coder based on moving 3D–objects", Third International Workshop on 64 kbit/s Coding of Moving Video, Rotterdam, The Netherlands, 4–6 September 1990.

[7] M. Hötter, K.–W. Hahn, "Predictive contour coding for an object–oriented analysis–synthesis coder", IEEE International Symposium on Information Theory, San Diego, California, USA, January 1990.

[8] M. Hötter, "Optimization of an object–oriented analysis–synthesis coder based on the model of flexible 2D–objects", to be published at PCS 91, Tokio, March 1991.

[9] M. Gilge, "Regionenorientierte Trnasformationscodierung in der Bildkommunikation", Fortschritt–Berichte VDI Reihe 10 no. 128, Düsseldorf, VDI–Verlag 1990.

[10] R. Mester, U. Franke, T.Aach, "Top–Down image segmentation using object detection and contour relaxation", proceedings ICASSP'89, Glasgow, Great Britain, May 1989.

[11] R.M. Haralick, L.G.Shirpio, "Image segmentation techniques", Computer Vision, Graphics, and Image Processing, 29, 1985.

[12] O.J. Morris, M. de J. Lee, A.G. Constantinides, "Graph theory for image analysis: an approach based on the shortest spanning tree", IEE Proceedings, No.2, April 1986.

[13] M. Eden, M. Kocher, "On the performance of a contour coding algorithm in the context of image coding part 1: Contour segment coding", Signal Processing, no. 8, 1985, pp.381–386.

[14] A. Friedrich, "Optimierung der Luminanzcodierung beliebig geformter Bildsegmente mittels Discreter–Cosinus–Transformation durch Segmentierung der Bildsegmente", diploma thesis, Universität Hannover, August 1990.

Reprint from Jörn Ostermann, "Coding of color parameters in an analysis–synthesis coder based on moving 3d–objects", *Picture Coding Symposium 1991 (PCS 91)*, pp. 10.5–1 – 10.5–4, Tokyo, Japan, 2 – 2 September 1991