Optimization and Efficiency of an Object-Oriented Analysis-Synthesis Coder

Michael Hötter

***Abstract*—In this paper, an object-oriented analysis-synthesis coder is presented the first time in its entity concentrating on an optimal relationship of its components image analysis, image synthesis and parameter coding and on a comparison of its coding efficiency to block-oriented hybrid coding. As block-oriented hybrid coder, the RM8 of the CCITT is used. The presented object-oriented analysis-synthesis coder is based on the source model of moving flexible 2D-objects and encodes arbitrarily shaped objects instead of rectangular blocks. The objects are described by three parameter sets defining their motion, shape and colour (colour parameters denoting luminance as well as chrominance values of the object surface). The parameter sets of each object are obtained by image analysis and coded by an object dependent parameter coding. Using the coded parameter sets, an image can be reconstructed by model-based image synthesis.**

Experimental results show that transmitting shape information and allowing small position errors (geometrical distortions) avoid the mosquito and blocking artefacts of a block-oriented coder. Furthermore, important image areas such as facial areas can be reconstructed with an image quality improvement up to 4 dB using the image analysis. As a whole, the reconstructed image of an object-oriented analysis-synthesis coder appears sharper compared to block-oriented hybrid coding.

***Index Terms*—Object-oriented image coding, object-based image coding, image analysis, image synthesis, parameter coding, motion estimation, predictive shape coding, object-oriented colour coding, videophone**

I. Introduction

In order to encode moving video signals at low bit rates, a coder concept has been proposed by the CCITT which is based on a block-oriented hybrid coding approach [3]. A block-oriented hybrid coder subdivides each image of a sequence into square blocks of $N \times N$ picture elements (pels) and encodes the luminance and chrominance signals of each block by motion compensated predictive and transform coding algorithms [3], [4], [7].

Besides block-oriented hybrid coding, object-oriented analysis-synthesis coding has been proposed in [18], [21], [1], [15], [20], [6]. These techniques encode arbitrarily shaped objects instead of square blocks. By the transmission of object shape fundamental to object-oriented coding, annoying coding errors of block-oriented hybrid coding known as mosquito and blocking artefacts can be avoided. Furthermore, with image analysis, important image areas such as facial details can be reconstructed with a higher image quality than with block-oriented hybrid coding.

Principally, object-oriented analysis-synthesis coding techniques aim at the efficient coding of scenes where

—only a few objects are moving,

—object motion is dominant and moderate,

—the moving objects cover up to 40–60% of the image area and

—no camera motion occurs (static background).

Based on these assumptions, investigations in object-oriented analysis-synthesis coding are—as in the paper presented here—at the moment restricted to typical videophone and videoconference applications.

From literature, two approaches for object-oriented analysis-synthesis coding are known. The first approach aims at efficient coding of one special object, e.g. a human face. Techniques which are based on this approach require an algorithm for recognizing the special object in a scene [1], [15], [20], [16], [17]. Therefore, these techniques are sometimes called knowledge-based analysis-synthesis coding techniques. These coding techniques have only partly been realized; a complete coder based on this approach does not exist until now.

The second approach for object-oriented analysis-synthesis coding is not restricted to only one type of object and therefore can be applied to a more general class of scenes [18], [13], [21]. Coding techniques which are based on this approach consist of four parts: the image analysis, the image synthesis, the memory of the object parameters and the parameter coding. The image analysis subdivides an image into moving objects and describes each object by three parameter sets defining its motion, shape and colour (colour denoting luminance as well as chrominance values of the object surface). The parameter sets are coded by an object dependent parameter coding. Using the encoded and transmitted parameters, an image can be reconstructed by image synthesis at the decoder as well as at the coder.

The object-oriented description can be used to control the coding of the parameter sets. For low bit rate transmission, irreversible coding techniques have to be applied in order to cut down the generated bit rate. In this situation, suppression of colour update information of an object introduces geometrical distortions instead of quantization error distortions. Assuming the same mean square distortion error, the geometrical distortions are less annoying than quantization error distortions as long as the modeling of an object is sufficiently exact.

Manuscript received June 22, 1992; revised January 17, 1994.

The author was with Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, Appelstraße 9A, 39167 Hannover, Germany; he is now with Robert Bosch GmbH, Research Institute for Communication (C/FOH), Robert-Bosch-Straße 200, Postfach 77 77 77, 31132 Hildesheim, Germany.

IEEE Log Number 9400885.

Coding techniques of the second approach can be based on different source models for describing the objects and their motion in a scene [18]. Depending on the source model, parameters with different information content and different bit-rates will be generated by the coder. In this contribution, the source model of flexible 2D-objects is used which results in object motion being described by a displacement vector field [13]. The displacements are measured in a fixed grid, e.g. 16×16 pel (picture element), and are then interpolated to guarantee a motion description for each pel of the object.

In order to realize an object-oriented analysis-synthesis coder based on the model of flexible 2D-objects, image analysis algorithms have been developed which automatically subdivide an image into moving objects and calculate their parameter sets [8], [9], [13]. Also, image synthesis algorithms which reconstruct an image by means of the encoded and transmitted parameters [12] and parameter coding algorithms which efficiently encode the motion, shape and colour parameters [10], [11], [14] have been evolved. In contrast to these papers which deal with important aspects and components of object-oriented analysis-synthesis coding, this paper presents a coder in its entity, concentrating not on a description of its components but on their relationship for optimal coding, i.e. for an optimal distribution of the available data rate to the motion, shape and colour parameters. This coder is compared with a block-oriented hybrid coder in terms of coding efficiency and image quality.

In Section II, the concept and structure of the coder is explained. The optimization of the coder, i.e. the optimal allocation of the data rate to the parameter sets, is described in Section III. Experimental results and a comparison of block-oriented hybrid coding and object-oriented analysis-synthesis coding are presented in Section IV.

II. Concept and Structure of the Object-Oriented Analysis-Synthesis Coder

In this section the concept and structure of an object-oriented analysis-synthesis coder are explained. First, a block diagram is used to describe the structure of the coder. Then, block-oriented hybrid coding is interpreted as a special realization of this structure and the development from the block-oriented to the object-oriented concept is motivated.

II.1. Structure of the Object-Oriented Analysis-Synthesis Coder

In the following, the block diagram of Fig. 1 is used to explain the structure of the object-oriented analysis-synthesis coder [18]. Input of the coder is a sequence of images. The image analysis subdivides each image of this sequence into moving objects and estimates for each object i three sets of parameters defining its motion, shape and colour parameters [9]. The colour parameters $\vec{S}_i$ denote the luminance and chrominance values of the object surface, the shape parameters $\vec{M}_i$ describe the position of the object i within the image plane and its boundary, and the motion parameters $\vec{A}_i$ determine the motion of object i within the image plane. These object parameters depend on the type of source model being applied-in this paper moving, flexible 2D-objects [14]. 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 i.e. caused by quantization of the parameter sets. 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 allows both the coder and decoder to identically reconstruct a transmitted image by image synthesis. The reconstructed image I'_k is displayed at the decoder and is also used for image analysis of the next input image I_{k+1} at the coder. Additionally, the stored parameters are used for the coding and decoding of the parameter sets of the next input image I_{k+1}.

In the following, the image analysis [8], [9], the image synthesis [12] and the parameter coding [10], [11], [14] of the object-oriented analysis-synthesis coder are briefly summarized.

II.1.A. Image Analysis

The goal of image analysis is to describe temporal signal changes between a synthesized image I'_k and the actual image I_{k+1} of the camera *by object motion*. For image analysis, an algorithm is applied which includes a hierarchical application of object motion and object shape estimation [9]. In the first step of hierarchy, temporally unchanged image regions, i.e. image areas of static background, are detected. In further steps of hierarchy, the changed image areas are analyzed and the temporal changes of each disjunct area are interpreted as the image area of one object described by three parameter sets $\vec{S}$, $\vec{M}$, $\vec{A}$. As the source model of moving, flexible 2D-objects is used here, object motion is described by a displacement vector field, i.e. it is assumed that temporal signal changes caused by moving 3D-objects in the scene can be described with sufficient accuracy by 2D-objects in the image plane moving locally varying (flexible) and translatorily. Hence, the motion parameters consist of displacement vectors. In our approach, these displacement vectors are calculated by hierarchical block matching [2], [13] where a reliable and accurate displacement estimation is achieved combining big and small measurement window sizes at different hierarchy levels. To improve displacement estimation at object boundaries where different displacement vectors have to be assigned to parts of the block, shape information of objects of previous images is used—if available—by evaluating only those parts of blocks for displacement estimation which belong to moving objects.

The description of temporal signal changes between a synthesized image I'_k and the actual image I_{k+1} of the camera by object motion is not valid in image areas where the applied source model fails. In order to judge the validity of the source model, the motion description of each object is verified during image analysis. First, an image synthesis is performed by the motion parameters of each object, i.e. by means of the calculated displacement vectors. If there are e.g. covered and uncovered image regions or image areas of an opening mouth or closing eyes, the assumption that object motion can locally be described by displacement vectors is

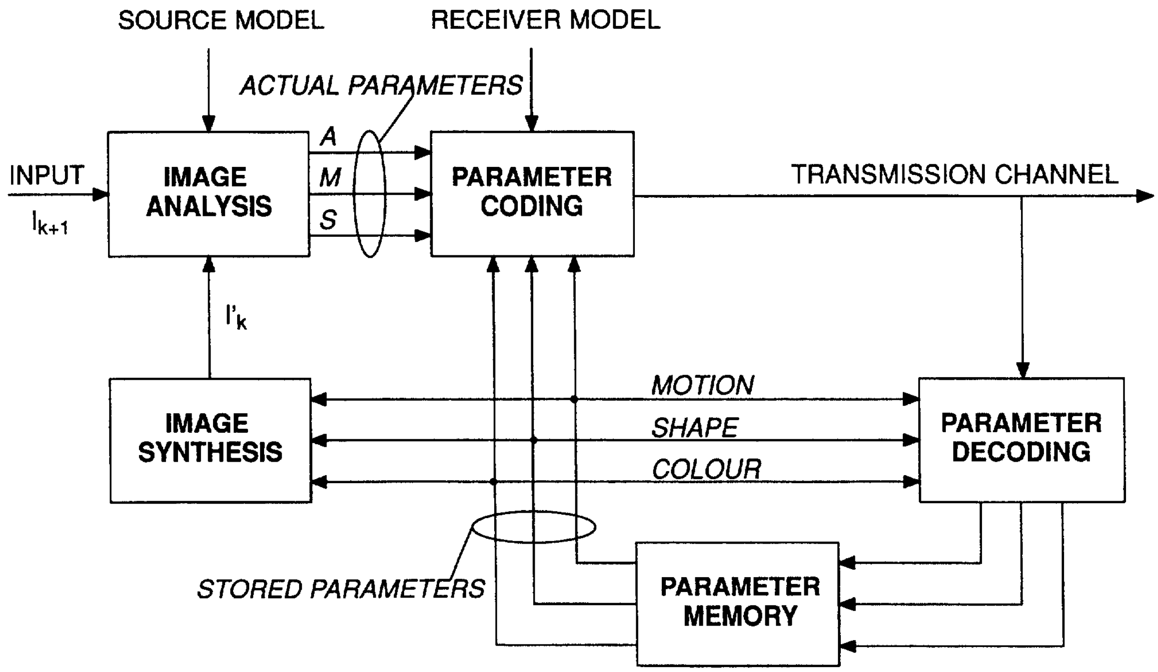


Fig. 1. Block diagram of an object-oriented analysis-synthesis coder.

not fulfilled and hence, the motion description is only valid in a part of the object. For that reason, the verification of the motion description in the image analysis part evaluating differences of the synthesized and the actual image [22] then yields a detection of those objects whose motion cannot be described with sufficient accuracy by displacements. These objects are denoted as Model Failure Objects *(MF-objects)* [22]. In contrast, all objects whose motion parameters describe the temporal changes with sufficient accuracy are classified as Model Compliance Objects *(MC-objects)* [14]. An example of the estimated parameter sets of the objects and of their classification is shown in Fig. 2.

II.1.B. Image Synthesis

Using the coded and transmitted parameter sets, an image can be reconstructed by image synthesis both, at the coder and decoder side. Principally, there exist two techniques for image synthesis: image synthesis by filter concatenation [7] and image synthesis by parameter concatenation [12]. These techniques and their properties will be briefly discussed in the following.

(i) Image synthesis by filter concatenation

Fig. 3 shows a one-dimensional example of image synthesis by filter concatenation. Assuming that all temporal signal changes are only due to motion, an image I_k to be synthesized is identical to the image I_0 at a displaced position. If displacement vectors with subpel amplitude resolution are applied, the displaced position does not necessarily coincide with the sampling grid of I_0 (see positions y_1, y_2, y_3 in Fig. 3). In this case, the signal s_0 of image I_0 has to be spatially filtered in order to interpolate the missing samples. The next image I_{k+1}

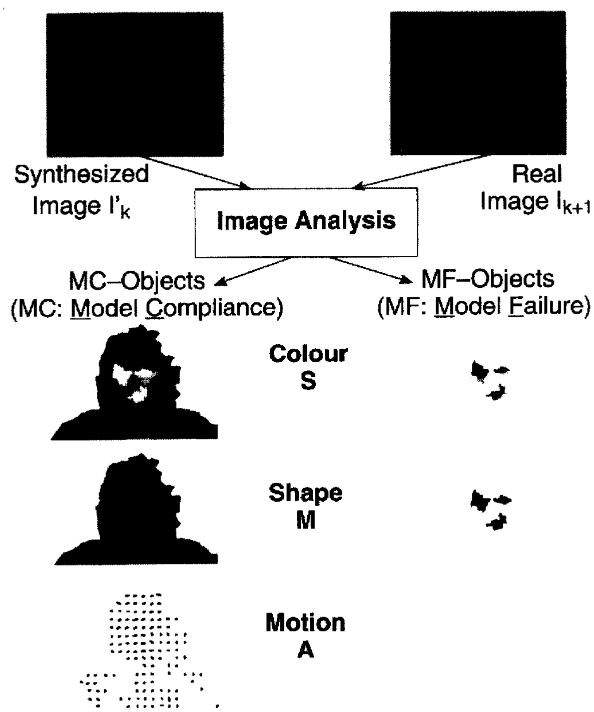
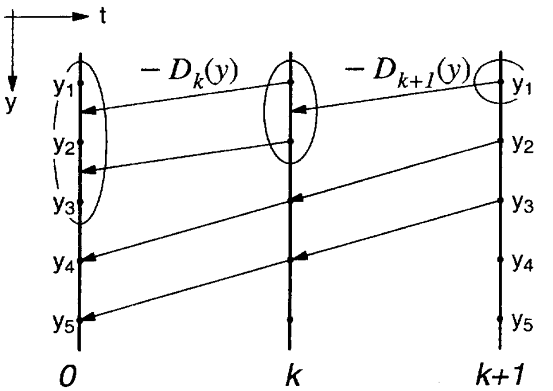


Fig. 2. One example of image analysis (Source model: moving, flexible 2D-objects; Parameter sets: colour S, shape M, motion A; Object types; MC- and MF-objects; Test sequence: *MISS AMERICA*).

to be synthesized is then calculated by means of the previously synthesized image I_k and the actual motion parameters (in Fig.



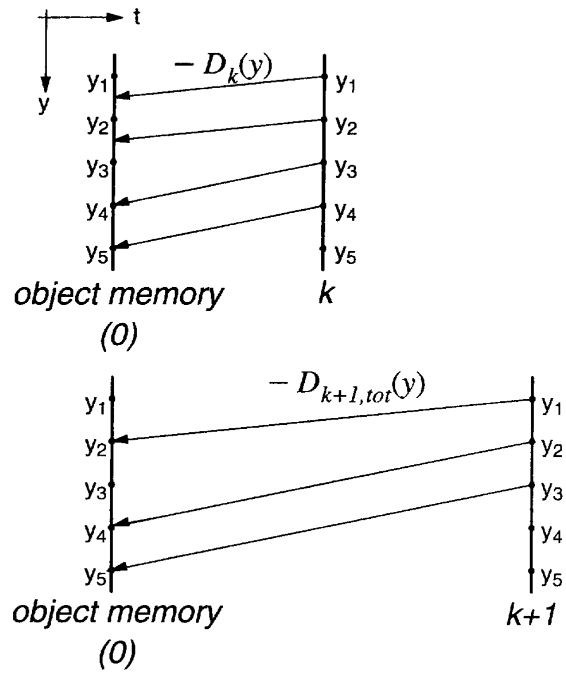
$$s'_{k+1}(y) = s'_k(y) * h(y - D_{k+1}(y))$$
$$= s_0(y) * h(y - D_{k+1}(y)) * h(y - D_k(y))$$

Fig. 3. Image synthesis by filter concatenation (one-dimensional case).

3 the displacement vector field $\vec{D}_{k+1}(y)$). This technique of image synthesis is usually applied in block-oriented hybrid coding techniques if displacement vectors are measured and transmitted with subpel amplitude resolution [7]. The main disadvantage of image synthesis by filter concatenation is that synthesis errors due to the interpolation filter accumulate (see position y_1 in Fig. 3).

(ii) Image synthesis by parameter concatenation

In order to avoid this accumulation of synthesis errors, an image synthesis by parameter concatenation is used in this contribution [12]. Fig. 4 gives a one dimensional example. In this approach, the signal s_0 of the first image I_0 of the image sequence is stored in an object memory. The image I_k is then synthesized by means of the stored signal s_0 of the object memory and the actual motion parameters (in Fig. 4: $\vec{D}_k(y)$). A comparison of Fig. 3 and Fig. 4 shows that image synthesis by filter concatenation and by parameter concatenation are identical for image I_k. In contrast to image synthesis by filter concatenation, the next image I_{k+1} is now synthesized by means of the stored signal s_0 of the object memory and an appropriate addition of all motion parameters which have been transmitted (in Fig. 4: $\vec{D}_{k+1,\,tot}(y)$). The main advantage of this technique over image synthesis by filter concatenation is that for each image to be synthesized the interpolation filter is only once applied.



$$s'_{k+1}(y) = s_0(y) * h(y - D_{k+1,tot}(y))$$
$$\text{with } D_{k+1,tot}(y) = D_{k+1}(y) + D_k(y - D_{k+1}(y))$$

Fig. 4. Image synthesis by parameter concatenation (one-dimensional case).

II.1.C. Parameter Coding

Parameter coding in an analysis-synthesis coder includes the coding of the motion, shape and colour parameters, the control of the coder modes and the priority control of the parameter transmission as shown in Fig. 5 [18]. At the output of image analysis the parameter sets of each object i are available in an uncompressed, numerical format. The parameters sets $\{\vec{S}_i\}$, $\{\vec{M}_i\}$, $\{\vec{A}_i\}$ have to be coded efficiently in order to increase the coding gain compared to block-oriented hybrid coding.

Therefore, individual parameter coding techniques are applied. The motion parameters are predictively coded using known DPCM techniques [13]. The shape parameters consist of object masks (see Fig. 2) which are approximated by a polygon and spline representation. First, the object shape is approximated by a polygon-representation. The number of calculated vertices depends on the chosen quality of the shape approximation. The quality is controlled by the absolute distance d_{max} between the approximated and the real object shape. The higher d_{max}, the coarser the approximation and the less vertices are calculated for the shape description. In order to combine this approach with the more natural looking spline functions, the vertices of the polygon-representation are used to calculate a spline-representation of the shape as well. Wherever the spline-representation fulfills the quality criterion d_{max}, it is used instead of the polygon, so that the final approximation of the shape is composed of spline and polygon parts. By this, a natural looking approximation of the object shape is achieved by a low number of vertices, where the quality measure d_{max} is guaranteed everywhere. Then, the vertices of this representation are predictively coded using the actual motion information and the stored shape information from previous images [11], [13]. Compared to known techniques [5], this shape coding technique reduces

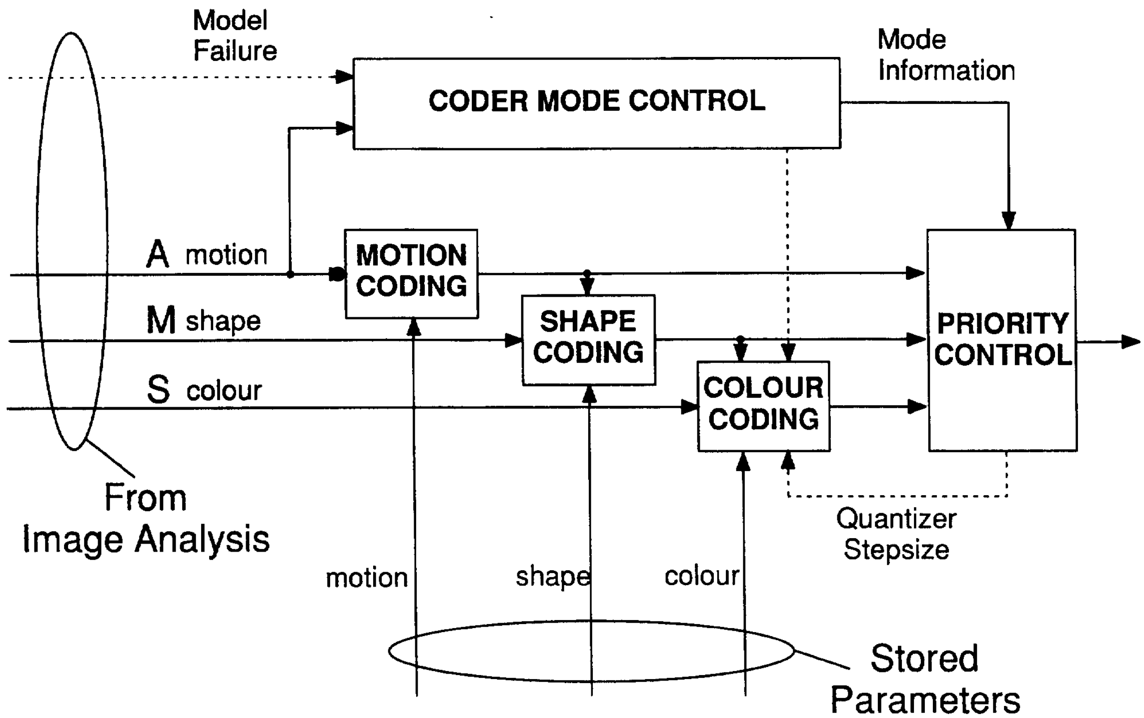


Fig. 5. Parameter coding.

the bit rate by a factor of about 4 to 6 ($d_{max} = 2.9$ pel), i.e., an average bit rate of only about 0.3 bit per shape pixel is needed [10]. The colour information is coded with a hybrid scheme, where, for each block, a DPCM technique is used instead of a DCT whenever it allows a more efficient coding [23]. The resulting information of each object i consists of three data amounts which represent the motion information R_{Ai}, the shape information R_{Mi} and the colour information R_{Si} of each object i.

The priority control decides which of the coded parameter sets of each object will be transmitted to the receiver. This decision is controlled by the mode information which is generated by the coder mode control. For this reason, the image analysis indicates to the coder mode control whether the object under consideration represents an MF-object or an MC-object.

Based on this information, the coder mode control selects the parameter sets to be transmitted (Table I) [18], [13], [14]. In the case of objects which cannot be described with sufficient accuracy by the source model, i.e. in the case of MF-objects, the motion information is omitted and only shape and colour information are transmitted (Mode 2). If the modeling of the object is sufficiently exact, i.e. in the case of MC-objects, only motion and shape information are transmitted to the receiver (Mode 1). This includes objects whose shape update information has been set to zero. The shape information of uncovered image areas is determined by the shape and motion parameters of the MC-objects. For these image areas only the colour parameters are coded and transmitted (Mode 3). Allowing small position and shape errors of MC-objects denoted as geometrical distortions, the image area of MF-objects can be decreased. These small position and shape errors are less annoying to a human observer than quantization error distortions which would occur otherwise. The coder mode control information has to be transmitted for each object. Principally, the motion parameters are transmitted first, then the shape update parameters of all objects and finally the colour parameters of the MF-objects and the uncovered image areas to guarantee a predictive coding of all parameter sets.

TABLE I
CODER MODES

Parameter sets to be transmitted \Modes	1 (MC-objects)	2 (MF-objects)	3 (Uncovered image areas)
Motion parameters	X		
Shape parameters	X	X	
Colour parameters		X	X
Distortions	geometrical	quantization error	

In the following, this coding strategy is compared to block-oriented hybrid coding and important differences between them are discussed.

Table II compares the selection of image areas to be updated by colour coding of an object-oriented analysis-synthesis coder and a block-oriented hybrid coder [13]. While for block-oriented hybrid coding this decision is based on the prediction errors and the chosen quantization characteristic of each block, for object-oriented analysis-synthesis coding this decision depends on the classification of the objects by image analysis. The resolution of the colour coding is chosen in

TABLE II
COMPARISON OF UPDATE PROCEDURES OF BLOCK-ORIENTED VERSUS OBJECT-ORIENTED CODER CONCEPTS

Technique	Selection Unit	Selection Criterion	Percentage of Updated Area
CCITT-Coder (RM8)	Block	Prediction Error + Quantization Characteristic	40%
Object-Oriented Analysis-Synthesis Coder*	Object	Model Failure + Data Rate	4%

*: source model of moving, flexible 2D-objects.

accordance with the available data rate. For an object-oriented analysis-synthesis coder based on the source model of moving, flexible 2D-objects the image area to be updated by colour coding can be decreased by a factor of 10 compared to a block-oriented hybrid coder in the case of typical videophone scenes [13].

A central problem for block-oriented hybrid coding with approximately constant quantizer stepsize is that a large area of the image is reconstructed by the transmission of colour parameters. Because of this large area, a coarse quantizer is applied in low bit rate coding so that the spatial resolution and hence the image quality of the decoded images is low and annoying coding errors known as mosquito and blocking artefacts become visible.

The mosquito artefacts are caused by the fact, that in block-oriented hybrid coding only one displacement vector is estimated for each block and used for the motion compensating prediction. This generates errors, if one block contains several objects moving differently. The errors have to be coded and transmitted to the receiver which is only possible in a lossy manner in low bit rate coding. Hence, visible coding errors remain at object boundaries which are called mosquito artefacts.

Besides mosquito artefacts, the so called blocking artefacts which are visible all over the image decrease the subjective image quality for block-oriented hybrid coding. These coding errors are caused by the fact, that all prediction errors are coded by colour parameters without taking into account their visibility for a human observer. Also prediction errors which are caused by small position and shape errors of the objects are coded. These geometrical distortions generate line structured prediction errors which can only be coded with a high coding effort. As, on the other hand, geometrical distortions are not annoying, as a human observer pays attention to a natural looking of the moving object but not to its exact positioning, a big data amount is used for the updating of prediction errors in image areas a human observer is not interested in. In summary, the coder control of a block-oriented hybrid coder leads to a coding of the colour parameters by a combination of only a few DCT-basis functions such that in most image areas the block structure becomes visible and the subjective image quality is decreased.

In object-oriented analysis-synthesis coding, the coder control aims at the reduction of the large image area to be updated by colour coding without introducing subjectively annoying errors. Three important properties of object-oriented analysis-synthesis coding support this reduction:

First, the motion compensating prediction can be drastically improved at object boundaries as different displacement vectors can be assigned to parts of a block. Hence, the motion compensating prediction at object boundaries is error free within the object and static background such that colour updating is only necessary in image areas of uncovered background. Furthermore, the coding of the shape parameters needs only a small data amount, as the image analysis used here generates temporally smoothly varying object boundaries whose approximation vertices can be efficiently coded in a predictive manner [11]. Hence, the data rate which is available for the coding of the colour parameters in an object-oriented analysis-synthesis coder is nearly the same compared to block-oriented hybrid coding.

Secondly, the image analysis detects such image areas where prediction errors are caused by small position and shape errors of the objects [12]. As these errors do not degrade the subjective image quality, the coding of colour parameters and their transmission can be intentionally suppressed to save further bit rate.

Third, the image analysis detects those image areas where the motion description of the applied source model fails (image areas of MF-objects). In videophone applications these images areas often contain the areas of mouth and eyes, i.e. areas which are important for a human observer to judge the subjective image quality. Therefore, these image areas can be transmitted with a higher coding effort such that the overall subjective image quality is improved for a human observer.

II.2. Block-Oriented Hybrid Coding as a Special Realization of Object-Oriented Analysis-Synthesis Coding

The presented object-oriented coder concept represents a *generalization* of block-oriented hybrid coding; the *structure* of both concepts is *identical* [19]: assuming the source model of square, rigid objects with translatory motion and known size (i.e. a block of 16×16 pels) and replacing the image analysis by a motion estimator which assigns to each object i, i.e. to each block, one displacement vector, the block diagram of the object-oriented coder simplifies to that of a block-oriented hybrid coder (Fig. 6). For each block of the actual image I_{k+1} a motion parameter $\vec{A}_i$ is calculated with reference to the stored, reconstructed image I_k. This motion parameter $\vec{A}_i$ consists of one displacement vector whose components describe the horizontal and vertical displacement of the block in image I_{k+1}. The colour parameters $\vec{S}_i$ consist of the luminance and chrominance values of all discrete positions within the block lying on the sampling grid. The shape parameters are explicitly defined by the source model (i.e. block size of 16×16 pels), known to the receiver, and hence neither coded nor transmitted. The motion parameters of all objects of the actual image I_{k+1} are coded and used for hybrid coding of the colour parameters. The coded motion and colour parameters are transmitted to the receiver, where the decoded colour parameters $\{\vec{S}_i'\}$ are stored in the image memory and displayed as reconstructed image I_{k+1}'.

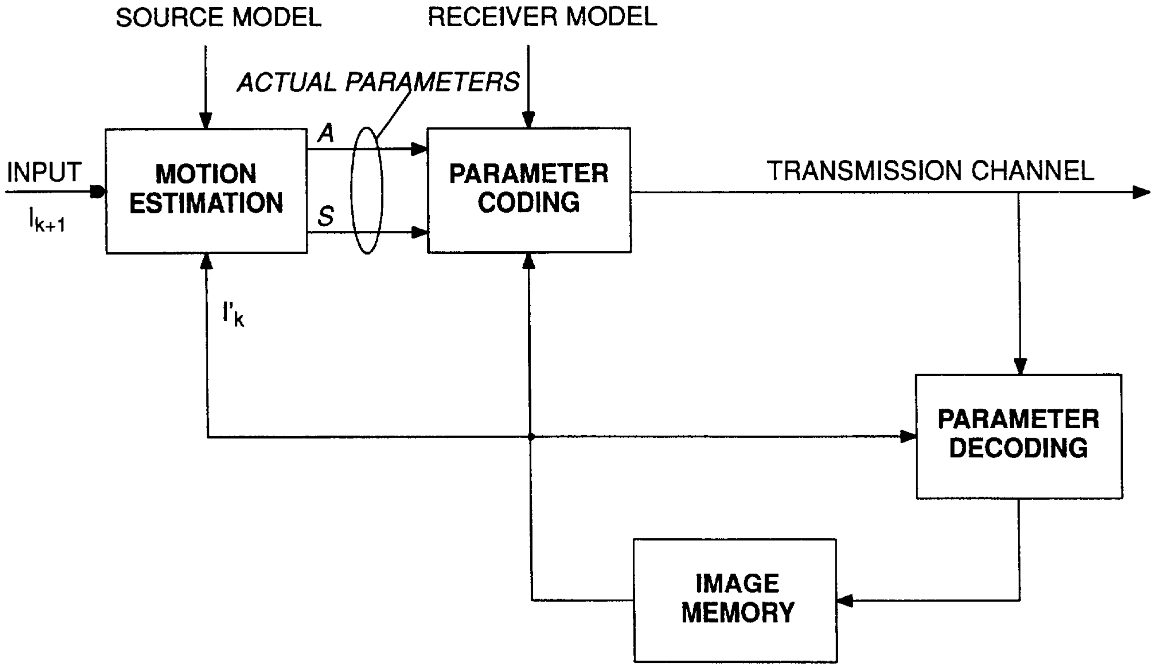


Fig. 6. Description of a block-oriented hybrid coder as a special realization of an object-oriented analysis-synthesis coder.

The generalization from the block-oriented to the object-oriented concept consists of the introduction of arbitrarily shaped objects instead of square blocks and the introduction of an image analysis which besides the motion estimation of objects includes the possibility to control the validity of the motion description and to control the parameter coding depending on the image contents.

While block-oriented hybrid coding techniques [3], [7] transmit only two parameter sets, namely the motion and colour information of each block, object-oriented analysis-synthesis coding additionally has to transmit the shape of each object. In order to obtain a higher coding gain for object-oriented coding compared to block-oriented hybrid coding, the additional bit rate R_M required for transmitting the shape information $\{\vec{M}_i\}$ has to be justified by a reduction of the bit rates R_A and R_S required for the motion parameters $\{\vec{A}_i\}$ and colour parameters $\{\vec{S}_i\}$.

Three properties of object-oriented analysis-synthesis coding contribute to this requirement. First, the shape coding algorithm used here is very efficient, i.e., the shape coding bit rate R_M is very low and hence, the decrease of data rate for coding the colour and motion parameters is small compared to block-oriented hybrid coding. Secondly, the synthesis of the colour information is improved at object boundaries using shape information as, in contrast to block-oriented hybrid coding, different displacement vectors may be assigned to parts of a block. Third, using image analysis for the detection of MF-objects and limiting the colour update transmission to these objects, the quality of image areas which are important for a human observer, e.g. facial details in videophone applications, can be improved.

III. Coder Optimization

In this section the minimization of the overall data rate by an optimal allocation of the data rate to the parameter sets is described. For this optimal allocation, the shape approximation accuracy and the spatial and amplitude resolution of the motion description are selected in such way, that for a given image quality the overall data rate is minimum [14]. In order to simplify the optimization, the multidimensional minimization problem is solved in two steps:

—optimization of the shape approximation accuracy and
—optimization of the resolution for the motion description.

The resolution of the shape approximation and hence the data amount for the shape parameters essentially influence the number of picture elements to be updated by colour coding, i.e. the transmission effort for the colour information directly depends on the resolution of the shape parameters. The shape parameters describe a concave hull of the object shape which includes independent of the shape approximation accuracy all picture elements of the object and which nestles with increasing shape approximation accuracy against the object shape. For this reason the number of picture elements to be updated by colour coding decreases with increasing accuracy of the shape approximation, i.e. the more accurate the shape approximation, the less pels have to be updated by colour coding. This relation of the two data amounts is—due to the shape approximation algorithm used here—approximately independent from the spatial and amplitude resolution of the motion description [22]. Hence, a separation of the multidimensional minimization problem as described above is allowed and useful. As investigations are concentrating on videophone sequences (see Section 1), the CCITT test sequence *CLAIRE* (see Section IV.1) has been used as a typical example for all experiments concerning the minimization of the data rate. Principally, the minimization should be performed for each image of a sequence separately dependent on the image contents. Here, this general problem of data distribution is not considered but it is assumed that the minimization achieved for a typical example of a videophone sequence is approximatively valid for the whole class of these scenes. As it is shown in Section IV, this assumption holds in our experiments.

III.1. Optimization of the Shape Approximation Accuracy

The shape approximation is applied to the MC- and MF-object boundaries. As different parameter sets are coded and transmitted for both object types (see Table I), the optimization of the shape approximation and shape coding is separately achieved.

Caused by the predictive shape coding and the stationary behavior of the shape parameters of MC-objects, the shape approximation accuracy only insignificantly contributes to the data amount of their shape parameters. On the other hand, the quality of the image synthesis at object boundaries is mainly influenced by the resolution of the shape approximation. In order to avoid visible synthesis errors at object boundaries introduced by the shape approximation, the approximation error d_{max} has been refined in our experiments, until visible synthesis errors at object boundaries disappear. It has been found out in these experiments that the approximation error d_{max} should be smaller than 1.5 pel. The approximation error d_{max} describes the maximum absolute distance between the approximated and the object shape calculated by image analysis.

The resolution of the shape approximation of the MF-objects strongly influences the data amount for the coding of the colour parameters. The more accurate the shape approximation, the less pels have to be updated by colour coding. On the other hand, with increasing shape approximation accuracy the data amount for the shape parameters of the MF-objects increases, because the shapes cannot be predictively coded as the accuracy of the motion description in these image areas is insufficient. This problem is outlined in Fig. 7, where the data amounts for a shape approximation with high and low accuracy are compared. In the case of a shape approximation with high accuracy, the image area of MF-objects is small and hence the data amount for the colour information is low, whereas the required data rate for the shape information is quite high. In the case of a shape approximation with low accuracy, the data amount for the shape information decreases whereas the image area of MF-objects and hence the required data rate for coding their colour parameters increase. In order to minimize the data rate for coding the parameter sets of the MF-objects, i.e. the sum of their colour and shape information, the approximation error d_{max} is selected in such a way, that the sum of the colour and shape information is minimum for the MF-objects.

The experimental results of this optimization are shown in Fig. 8 where it can be seen that independent of the required

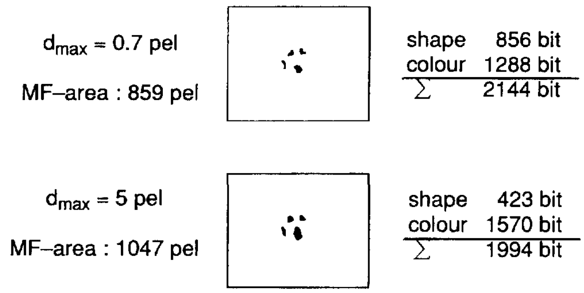


Fig. 7. The influence of shape approximation accuracy on shape and colour data rate (MF-objects, colour coding rate assumed 1.5 bit/pel).

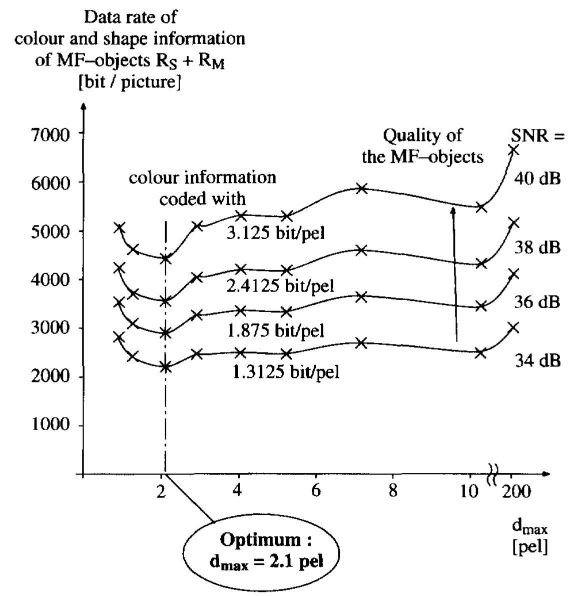


Fig. 8. Minimization of the data rate for coding the shape and colour information of MF-objects by optimal choice of shape approximation accuracy. (Simulations for d_{max} = 0.9, 1.1, 2.1, 2.9, 3.1, 4.1, 5.1, 7.1, 10.1, 200 [pel]; d_{max} = 200 [pel]: approximation by means of the four contour points with highest elongation.)

image quality in areas of MF-objects the data rate is minimum for an approximation error of $d_{max} = 2.1\ pel$. Furthermore, it can be seen, that the minimum is the more significant, the higher the required image quality for the MF-objects. This tendency is caused by the increasing influence of the data amount for the colour parameters: the higher the image quality in areas of MF-objects, the smaller the influence of the shape parameters to the required data rate for coding the parameter sets of the MF-objects.

III.2. Optimization of the Resolution for the Motion Description

With optimal values for the shape approximation accuracy fixed, the optimal spatial and amplitude resolution for the motion description has been found out experimentally. In our experiments, 1/1 pel and 1/2 pel amplitude resolution are used. The spatial resolution is defined by the height and width of the

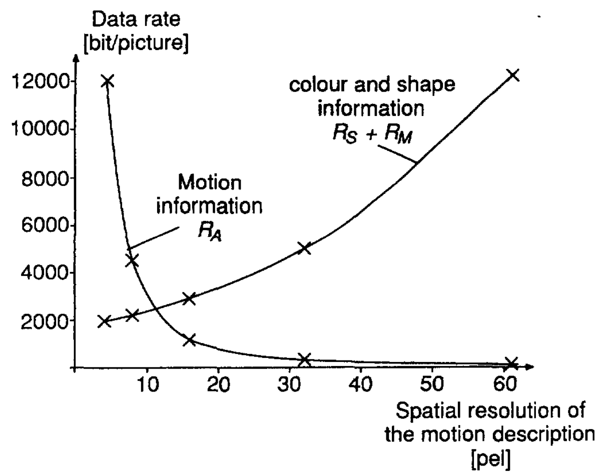


Fig. 9. Influence of the spatial resolution of the motion description on parameter coding. (Spatial resolution: 4 × 4, 8 × 8, 16 × 16, 32 × 32, 64 × 64 [pel].)

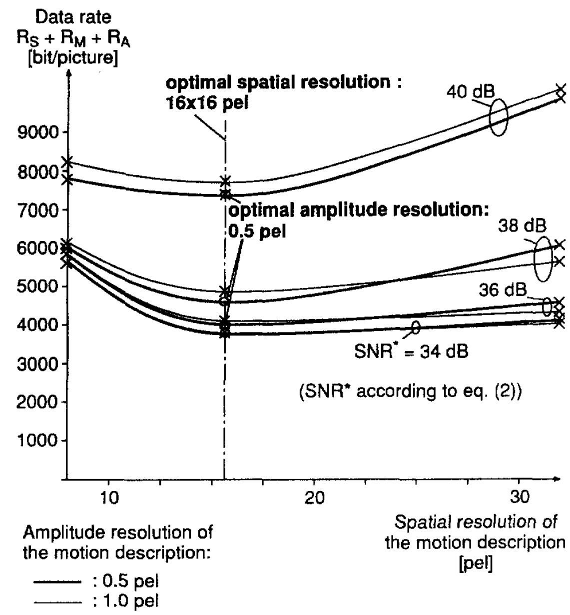


Fig. 10. Optimization of the spatial and amplitude resolution of the motion description. (Spatial resolution: 8 × 8, 16 × 16, 32 × 32 [pel].)

measurement grid of the displacement vectors. The spatial and amplitude resolution influence the number of picture elements to be updated by colour coding. Hence, the data amount for the colour information directly depends on the coding effort of the motion parameters. The finer the spatial and amplitude resolution, the less pels have to be coded by colour and shape parameters. On the other hand with increasing spatial and amplitude resolution of the motion description, the data amount for the motion information increases as well, as more displacement vectors have to be coded and transmitted to the receiver. This problem is outlined in Fig. 9 where the influence of resolution for the motion description on the parameter coding is shown for the example of spatial resolution. In order to minimize the overall data rate for coding the parameter sets of all objects, i.e. the colour, shape and motion information, the spatial and amplitude resolution has been varied using different bit rates for the colour coding. The experiments are restricted to a uniform spatial resolution of 8×8, 16×16 and 32×32 pels. Furthermore, the overall data rate is compared for an amplitude resolution of 1/1 pel and 1/2 pel. The experimental results of this optimization are shown in Fig. 10. For this optimization the subjectively weighted quality criteria SNR* is used for the assessment of the image quality as described in Section IV.2. It turned out that for a spatial resolution of 16 × 16 pels and an amplitude resolution of 1/2 pel, the overall data rate is minimum.

IV. Experimental Results

For a judgement of the object-oriented analysis-synthesis coder concept, the optimized coder based on the model of moving, flexible 2D-objects has been compared with a block-oriented hybrid coder in its image quality. For this comparison, typical videophone test sequences have been used which are described in Section IV.1. In Section IV.2 different objective quality criteria are introduced for the judgement of the image quality. Based on these quality criteria, the optimized coder is compared to a block-oriented hybrid coder in Section IV.3. The results of this comparison are finally discussed and judged.

IV.1. Description of the Test Sequences

For the experiments, the test sequences *MISS AMERICA* and *CLAIRE* have been used which have been selected by international expert groups ("European COST 211bis Simulation Subgroup" and "CCITT Study Group XV") as typical videophone sequences showing head and shoulder in motion. The sequence *MISS AMERICA* has been provided by the British Telecom Research Laboratories ($BTRL$), Great Britain, the sequence *CLAIRE* by the Centre National d'Études des Télécommunications ($CNET$), France. Fig. 11 shows one picture of each test sequence. These test sequences exist in different formats for different application aspects. In these experiments, the Common Intermediate Format (CIF) has been used which has been accepted for videophone applications by the CCITT. Each picture consists of 288 lines and 352 pixels per line for the luminance component and of 144 lines and 176 pixels per line for the both chrominance components. Each sample is quantized with 8 bit, the sampling rate of the luminance component is 6.75 MHz. Both sequences have been used in a non-interlace format with a field frequency of 10 Hz.

It turned out in our experiments, that the data amount for shape, colour and motion information as well as the image quality measured by different criteria are very similar for both sequences. Thus, if not otherwise mentioned, all measurement results in the following represent an average of both sequences.

IV.2. Definition of Quality Criteria to Judge Image Quality

In order to compare the presented object-oriented coder with a block-oriented coder an optimal allocation of the data rate

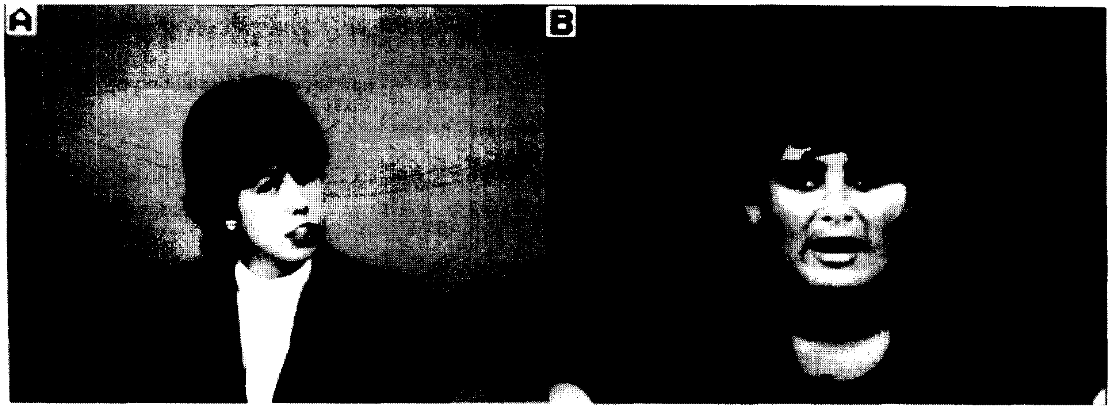


Fig. 11. Reference images of the test sequences. A: *CLAIRE* picture no. 33; B: *MISS AMERICA* picture no. 8.

to the parameter sets has to be found out. First, it is necessary to determine a quality criterion to judge the image quality. The quality criterion should judge the image quality by a comparison of numerical values. In the literature, the SNR (Signal to Noise Ratio) is often used:

$$SNR[dB] = 10 \cdot \log_{10}\left(\frac{255^2}{\sigma_e^2}\right), \qquad (1)$$

where σ_e^2 represents the variance of the coding error, i.e. the mean square error between the luminance samples of the original image and the corresponding samples of the coded image. σ_e^2 is calculated from the coding error of all samples of all images of a sequence. Hence, all coding errors are *equally weighted without taking into account their importance for a human observer.*

In object-oriented analysis-synthesis coding, geometrical distortions, i.e. small position errors of objects, are intentionally allowed, as they do not annoy the subjective image impression. As position errors may generate locally high synthesis errors (typically line structured), they decrease the SNR according to eq. (1) on one hand, but they do not diminish the subjective image quality on the other hand. In order to exclude these synthesis errors from the assessment of the image quality, a new subjectively weighted quality criterion according to

$$SNR^*[dB] = 10 \cdot \log_{10}\left(\frac{255^2}{\sigma_e^{*2}}\right), \qquad (2)$$

is introduced where in σ_e^{*2} the coding errors of only those samples are taken into account which do not belong to image areas of geometrical distortions. These areas are found during image analysis as described in [22]; the verification of the motion description by displacement vectors (see Section II) evaluates differences of the synthesized and actual image to detect those objects whose motion cannot be described with sufficient accuracy by displacement vectors and those image areas containing line structured synthesis errors assumed to be generated by geometrical distortions [14], [22]. These image areas of line structured synthesis errors are excluded of the assessment of image quality.

IV.3. Comparison of the Image Quality of the Object-Oriented Analysis-Synthesis Coder and a Block-Oriented Hybrid Coder

In order to assess the efficiency of the object-oriented analysis-synthesis coding, the image quality of the optimized coder "*OBJ.COD.*" according to Section III has been compared with the block-oriented hybrid coder "$RM8$" of the CCITT [4] based on objective and subjective criteria. This comparison has been achieved with the following restrictions:

- The test sequences *MISS AMERICA* and *CLAIRE* have been used for both coders applying various data rates from $r = 30$ kbit/s to $r = 75$ kbit/s. Hence, a data rate of 3000 to 7500 bits is available for the coding of one image (frame rate: 10 Hz).
- In order to control variations of the data rate needed for the coding of one image, a buffer is used whose size corresponds to the data amount which is available for the coding of one image.
- Both coders are initialized by an original image. Hence, the simulations assume that for the transmission of the first image an arbitrary data amount is available. Problems to reach the stable state after a scene cut are not considered.

IV.3.A. Comparison of the Image Quality by Objective Quality Criteria

As objective quality criteria, the signal-to-noise ratios SNR according to (1) and SNR* according to (2) have been used. The signal-to-noise ratio SNR is defined by the variance of the coding error σ_e^2 where σ_e^2 is measured by averaging the square coding error of all picture elements of all pictures of an image sequence. In contrast, for the measurement of SNR* the square coding error is averaged only over those picture elements which do not belong to image areas of geometrical distortions.

Fig. 12 shows the SNR of the optimized object-oriented analysis-synthesis coder *OBJ.COD.* and the block-oriented hybrid coder *RM8* for various data rates. Compared to the block-oriented coder the image quality of the object-oriented coder is about $0.3 - 1$ dB lower if measured by the SNR. This is caused by the fact that in block-oriented hybrid coding blocks with prediction errors exceeding a threshold

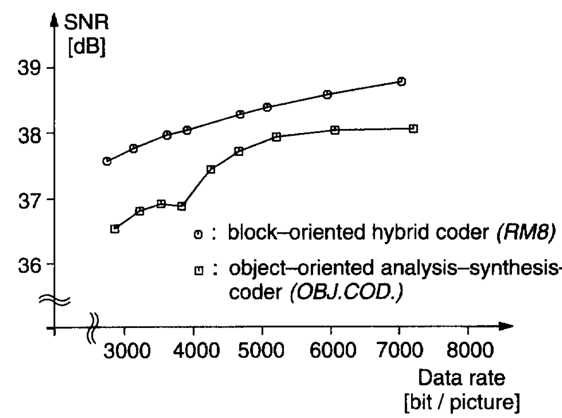


Fig. 12. Comparison of the presented object-oriented analysis-synthesis coder *OBJ.COD.* and the block-oriented hybrid coder *RM8* in terms of *SNR* (eq. (1)).

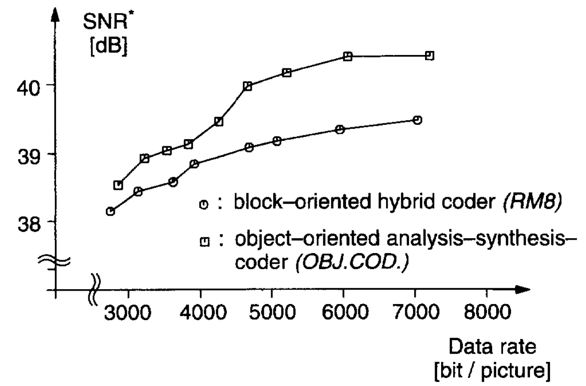


Fig. 13. Comparison of the presented object-oriented analysis-synthesis coder *OBJ.COD.* and the block-oriented hybrid coder *RM8* in terms of SNR^* (eq. (2)).

are updated by colour coding. Thus, this coding technique aims at optimizing the SNR. The prediction error denotes the error between an image I'_k after motion compensation and the original image I_{k+1}. Hence, the prediction error of block-oriented hybrid coding correspond to the synthesis error in object-oriented analysis-synthesis coding. In the case of object-oriented coding, synthesis errors caused by small position and shape errors of the objects are tolerated. These errors do not diminish the subjective image quality but they decrease the SNR.

Therefore a more appropriate criterion for the assessment of the subjective image quality is the SNR* according to eq. (2) since synthesis errors which are caused by small position and shape errors of the objects are excluded from the measurements (in our experiments typically 2000–3000 picture elements corresponding to 2–3% of the image area). Fig. 13 shows the comparison of the two coders according to the quality criterion SNR*. Compared to the block-oriented coder, the image quality of the object-oriented coder is now about $0.3 - 1$ dB higher if measured by the SNR*.

Besides the global assessment of the image quality with and without taking into account geometrical distortions of objects, it is important for a human observer that image areas which the observer is interested in such as facial details are reconstructed in the highest possible image quality. In an object-oriented analysis-synthesis coder, these image areas are in most cases described by MF-objects. For a local assessment of the image quality in image areas which are relevant for a human observer, the SNR and SNR* have therefore been measured for both coders in image areas of MF-objects. As MF-objects do not contain geometrical distortions, SNR and SNR* are identical (see Section II). The image areas of MF-objects are detected by the image analysis of the object-oriented coder. As the relative amount of the image area of these regions is quite small and motion and shape parameters can be coded efficiently, the biggest amount of data is available for coding the colour parameters of the MF-objects. As can be seen in Fig. 14, the object-oriented coder yields an image quality in these areas

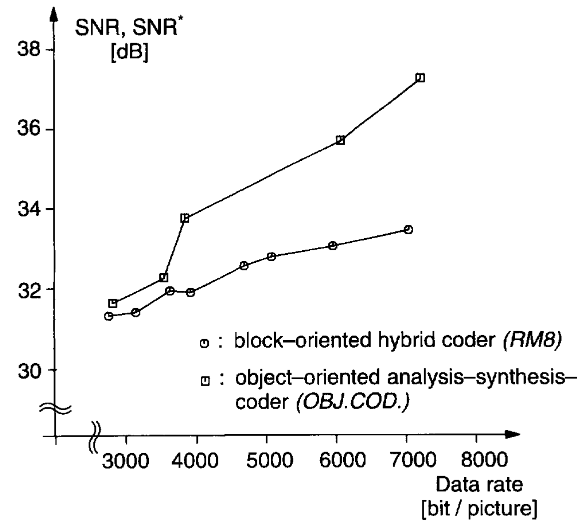


Fig. 14. Comparison of the presented object-oriented analysis-synthesis coder *OBJ.COD.* and the block-oriented hybrid coder *RM8* in terms of *SNR* (eq. (1)) and SNR^* (eq. (2)) in image areas of MF-objects.

which is up to 4 dB higher compared to block-oriented hybrid coding.

IV.3.B. Comparison of Image Quality by Subjective Quality Criteria

Besides the investigated objective quality criteria, subjective criteria are important for a human observer to judge image quality. A comparison of the decoded test sequences shows important properties of the object-oriented analysis-synthesis coder which are demonstrated and discussed in the following. Figs. 15 and 16 show the decoded versions ($r = 64$ kbit/s) of the original images in Fig. 11 applying the block-oriented hybrid code *RM8* and the object-oriented analysis-synthesis coder *OBJ.COD.*

Using the block-oriented hybrid coder, coding errors visible at object boundaries are caused by the fact that only one

Fig. 15. Coded images of Fig. 11A, data rate: 64 kbit/s. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

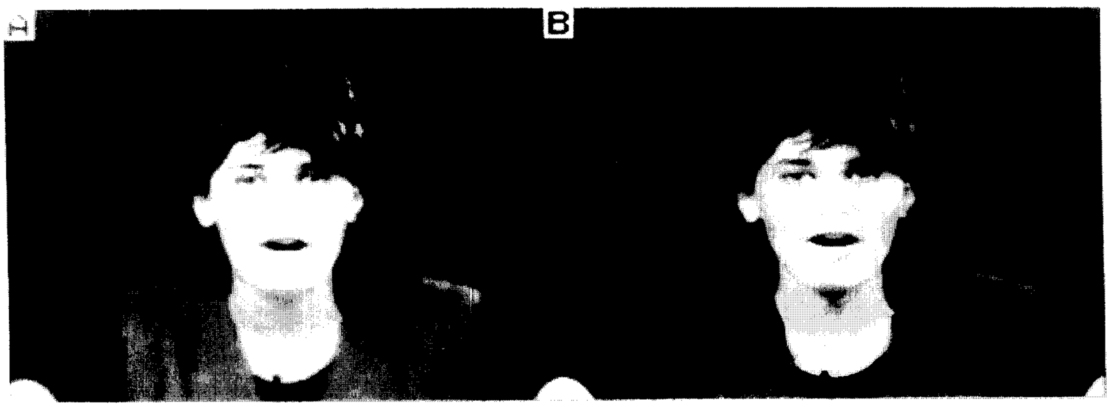


Fig. 16. Coded images of Fig. 11B, data rate: 64 kbit/s. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

displacement vector is assigned to blocks which contain differently moving objects. In these blocks, motion compensating prediction causes errors which are updated by the coding of the prediction error. Caused by the low data rate, the coding of the prediction error is irreversible and hence quantization errors known as mosquito artefacts arise (Figs. 17A and 18A). By transmitting the shape parameters of the objects, these errors can be avoided, because one displacement vector is now individually assigned to each pel. Hence, mosquito artefacts do not occur in object-oriented analysis-synthesis coding (Figs. 17B and 18B).

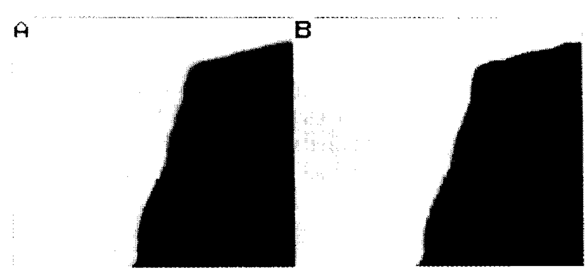


Fig. 17. Illustration of the mosquito artefacts for a detail of Fig. 15. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

Another problem of block-oriented hybrid coding is that all image areas which contain prediction errors are updated by colour coding independently of their importance for a human observer. In block-oriented hybrid coding at 64 kbit/s this area includes about 40% of the whole image area. Allowing small position and shape errors of the objects and transmitting the shape information, this amount can be decreased to less than 4% of the image area without introducing subjectively annoying distortions. Caused by the efficient coding of the motion and shape parameters in the object-oriented analysis-synthesis coder (average data amount for motion parameters: 1000 bits per image, average data amount for shape parameters: 900 bits), about 70% of the data rate (i.e. 4500 bits per image) is available for the coding of the colour parameters. This corresponds to an average data rate for the colour parameters of about 1 bit per pel in comparison to 0.1 bit per pel in case of the block-oriented hybrid coder. Therefore, the quality in image areas which are updated by colour coding is clearly increased for the object-oriented analysis-synthesis

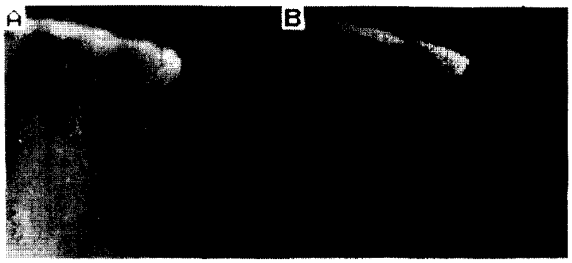


Fig. 18. Illustration of the mosquito artefacts for a detail of Fig. 16. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

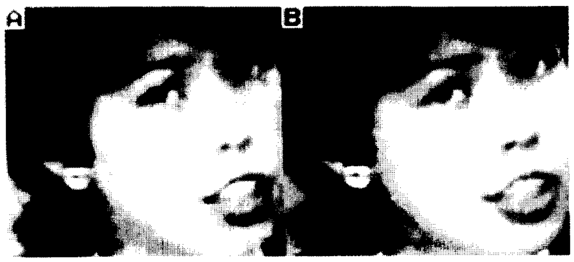


Fig. 19. Illustration of the blocking artefacts for a detail of Fig. 15. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

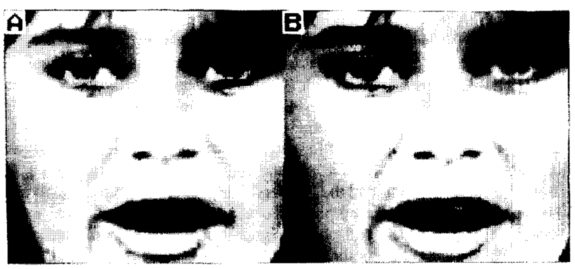


Fig. 20. Illustration of the blocking artefacts for a detail of Fig. 16. A: block-oriented hybrid coder *(RM8)*; B: object-oriented analysis-synthesis coder *(OBJ.COD.)*.

coder. Annoying coding errors which are known from block-oriented hybrid coding as blocking artefacts (Figs. 19A and 20A), do not occur for object-oriented coding. As a whole the decoded image of the object-oriented analysis-synthesis coder appears sharper compared to the block-oriented hybrid coder (Figs. 19B and 20B).

V. Conclusion

In this paper, the bit allocation for an object-oriented analysis-synthesis coder based on the source model of flexible 2D-objects with 2D-motion has been optimized first. The approximation error of the shape representation as well as the spatial and amplitude resolution of the motion description have been varied in order to minimize the overall data rate, i.e. the sum of the bit rates for the three parameter sets, at a given image quality. It has been found that for

— a measurement and transmission of the displacements in a grid of 16×16 pels and with an amplitude resolution of 0.5 pel for object motion description and
— a shape approximation which differs no more than 1.5 pels (MC-objects) and 2.1 pels (MF-objects) from the original shape,

the overall data rate is minimum.

Then, the image quality of the optimized object-oriented analysis-synthesis coder and the block-oriented hybrid coder have been compared based on subjective and objective criteria. The main advantages of the presented coder compared to a block-oriented hybrid coder like the "*RM8*" are:

— By the transmission of shape information, the quality of the image synthesis is drastically increased in image background areas in the vicinity of moving objects. Annoying coding errors known from the block-oriented coder as mosquito effects are eliminated by the new coder concept.
— In case of low bit rate coding where the available bit rate does not allow a perfect reconstruction, a block-oriented hybrid coder generates visible quantization errors while an object-oriented coder can suppress the updating of the colour information of an object. By transmitting only the motion and shape information, geometrical distortions due to small position errors of an object are allowed. These distortions are less annoying than quantization errors.
— The object-oriented priority control of the colour parameter transmission allows to emphasize the coding of important object features. It is possible e.g. to encode small objects like eyes more accurately than the other areas.
— In the block-oriented coder about 40 percent of the image area is updated by colour coding (at a data rate of 64 kbit/s). Allowing small position errors of objects and transmitting the shape information, this amount can be reduced to 4 percent of the image area without introducing annoying distortions.
— Owing to the efficient coding of the shape and motion information in the object-oriented analysis-synthesis coder, about 70 percent of the data rate ($r = 64$ kbit/s) is available for the coding of the colour updating. This corresponds to a data rate of about 1 bit per pixel compared to 0.1 bit per pixel for the block-oriented coder, i.e. the SNR of the image area updated by colour coding is higher than for the block-oriented coder.
— If the assessment of the image quality is based on a subjectively weighted criterion the image quality of the object-oriented analysis-synthesis coder is about 0.3–1 dB, in image areas of eyes and mouth up to 4 dB better compared to block-oriented hybrid coder.
— As a whole, the decoded image of the object-oriented analysis-synthesis coder appears sharper compared to the block-oriented hybrid coder.

Acknowledgment

The author wishes to thank Professor Musmann for encouraging this work and for many helpful discussions. Further, the author wishes to thank Mr. Schiller for his software support and his suggestions concerning the subject of colour coding and block-oriented hybrid coding. This work has been supported by the Forschungsinstitut der Deutschen Bundespost TELEKOM beim Fernmeldetechnischen Zentralamt, Darmstadt.

REFERENCES

[1] K. Aizawa, H. Harashima, T. Saito, "Model-based analysis-synthesis image coding (MBASIC) system for a person's face," *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 139–152, October 1989.

[2] M. Bierling, "Displacement estimation by hierarchical blockmatching," *Visual Communications and Image Processing '88*, T. Russell Hsing, Editor, Proc. 1001, pp. 942–951 (1988), Cambridge, U.S.A., November 1988.

[3] CCITT, "Draft revision of recommendation H.261: video codec for audio visual services at $p \times 64$ kbit/s," *Study Group XV, WP/1/Q4, Specialist Group on Coding for Visual Telephony*, Doc. No. 584, November 1989.

[4] CCITT SG XV WP XV/1, "Description of Reference Model 8 (RM8)," Doc. 525, July 1988.

[5] M. Eden, M. Kocher, "On the performance of a contour coding algorithm in the context of image coding. Part I: Contour segment coding," *Signal Processing*, vol. 8, no. 4, pp. 381–386, July 1985.

[6] W. Geuen, F. Kappei, "Principle strategy of model based source coding," *Picture Coding Symposium (PCS '87)*, Stockholm, Sweden, no. 12.2, June 1987.

[7] K. Grotz, J. Mayer, G. Süußmeier, "A 64 kbit/s videophone codec with forward analysis and control," *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 103–116, October 1989.

[8] M. Hötter, R. Thoma, "A selfadaptive, model based segmentation algorithm," *Picture Coding Symposium (PCS '88)*, Turino, Italy, no. 2.2, September 1988.

[9] M. Hötter, R. Thoma, "Image segmentation based on object oriented mapping parameter estimation," *Signal Processing*, vol. 15, no. 3, pp. 315–334, October 1988.

[10] M. Hötter, K.-W. Hahn, "Object-oriented analysis-synthesis coding based on moving 2D objects," *2nd International Workshop on 64 kbit/s Coding of Moving Video*, Hannover, Germany, no. 4.1, September 1989.

[11] M. Hötter, "Predictive contour coding for an object-oriented analysis-synthesis coder," *IEEE International Symposium on Information Theory*, San Diego, California, U.S.A., p. 75, January 1990.

[12] M. Hötter, "Object-oriented analysis-synthesis coding based on the model of flexible 2D-objects," *Picture Coding Symposium (PCS '90)*, Cambridge, U.S.A., no. 7.1, March 1990.

[13] M. Hötter, "Object-oriented analysis-synthesis coding based on moving two-dimensional objects," *Signal Processing: Image Communication*, vol. 2, no. 4, pp. 409–428, December 1990.

[14] M. Hötter, "Optimization of an object-oriented analysis-synthesis coder based on the model of flexible 2D-objects," *Picture Coding Symposium (PCS '91)*, Tokyo, Japan, no. 10.4, September 1991.

[15] R. Koch, "Adaptation of a 3D facial mask to human face in videophone sequences using model based image analysis," *Picture Coding Symposium (PCS '91)*, Tokyo, Japan, pp. 285–288, September 1991.

[16] A. Koike, M. Kaneko, Y. Hatori, "Model-based coding with 3-D motion estimation and shape change detection," *Picture Coding Symposium (PCS '90)*, Cambridge, U.S.A., no. 9.5, March 1990.

[17] M. Koyama, K. Ohyama, H. Fujiwara, "Model based real time synthesis system of facial expressions," *Picture Coding Symposium (PCS '90)*, Cambridge, U.S.A., no. 7.4, March 1990.

[18] H. G. Musmann, M. Hötter, J. Ostermann, "Object-oriented analysis-synthesis coding of moving images," *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 117–138, October 1989.

[19] H. G. Musmann, "Objektorientierte Analyse-Synthese-Codierung," *Workshop on "Objektorientierte Codierung und Objekterkennung in Bildfolgen"*, Hannover, Germany, pp. 1–2, February 1991.

[20] Y. Nakaya, H. Harashima, Y. C. Chiang, "Model-based/MC-DCT hybrid coding system," *Picture Coding Symposium (PCS '91)*, Tokyo, Japan, no. 10.3, September 1991.

[21] J. Ostermann, "Modelling of 3D-moving objects for an analysis-synthesis coder," *SPIE/SPSE Symposium on Sensing and Reconstruction of 3D Objects and Scenes '90*, B. Girod, Editor, Proc. SPIE 1260, pp. 240–250, Santa Clara, California, February 1990.

[22] J. Ostermann, H. Li, "Detection and Coding of Model Failures in an Analysis-Synthesis Coder Based on Moving 3D-Objects," *3rd International Workshop on 64 kbit/s Coding of Moving Video*, Rotterdam, The Netherlands, no. 4.4, September 1990.

[23] H. Schiller, M. Hötter, "Colour Coding in an Object-Oriented Analysis-Synthesis Coder," *submitted for publication in Signal Processing: Image Communication.*

Michael Hötter was born in Troisdorf, Germany, on February 6, 1961. He studied Electronic Engineering (Communications Technology branch) at the Technical University of Aachen (Germany). In 1986 he obtained the Dipl.-Ing. degree and joined the "Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung" at the University of Hannover (Germany), where he worked in the areas of motion estimation, image analysis and image coding. During that time, he was in charge of several internal projects and a research project together with the Deutsche Bundespost, Darmstadt (Germany) dealing with object-oriented analysis-synthesis image coding for low bit rates. In this research project for advanced image compression schemes, computer vision techniques (analysis of moving images) are employed for a decomposition of images into objects, which are described by colour, shape and motion parameters.

In 1992, Mr. Hötter obtained the Ph.D. (Dr.-Ing.) degree from the University of Hannover. He has published/presented about 15 papers on global motion estimation, image analysis and object-oriented image coding in international journals, workshops, and conferences.

Since 1992, Mr. Hötter is with the Research Institute for Communications of Robert Bosch GmbH, Hildesheim (Germany). His present interests and activities cover motion estimation, image analysis, object detection and description, statistical evaluation of measurement data and the application of these techniques in realtime vision systems.