Jörn

Efficiency Analysis and Application of Uncovered Background Prediction in a Low Bit Rate Image Coder

Dietmar Hepper

published in:
IEEE Transactions on Communications,
Vol. 38, No. 9, September 1990, pp. 1578–1584

Efficiency Analysis and Application of Uncovered Background Prediction in a Low Bit Rate Image Coder

DIETMAR HEPPER

Abstract—**A fast algorithm for gaining scene background information from a TV sequence is presented. This picture information is used by an uncovered background predictor in a motion-compensating predictive coder for videoconference or picture telephone applications. In uncovered background regions, prediction error entropy is reduced close to the value of the frame difference entropy in static regions, which is regarded as a practical lower bound. Almost all these picture elements can therefore be reconstructed without transmission. A motion-compensating hybrid-DPCM/transform coder is supplemented by uncovered background prediction. The maximum bit rate reduction achieved in the experiments is 9% for a picture where 14% of the picture elements of temporally changed regions belong to uncovered background regions.**

I. INTRODUCTION

FOR the digital transmission of videoconference and video telephony signals, mainly low bit rates in the range of 384 kb/s to 64 kb/s are under consideration internationally. In order to reduce the data rate of the video camera signals, source coding schemes with motion-compensating prediction are applied [4], [9], [11]–[13], [15], [16]. This prediction method reduces prediction errors in static as well as in moving object regions; but it fails where the scene background is uncovered by a moving object since new picture information appears, for which no corresponding information can be found in the previous picture (see Fig. 1).

In order to reduce the data rate in these picture regions, uncovered background prediction is introduced [1], [2], [7], [10]. For that purpose, an algorithm is to be designed which gains picture information of the scene background and stores it in a special background memory, thus making it available to an uncovered background predictor. Wherever background that has been stored in background memory is uncovered in the scene, uncovered background prediction can be applied.

The algorithm for gaining background information presented in [10] pelwise compares the frame-difference signal to a certain threshold to decide whether a picture element belongs to a static or to a nonstatic region. For picture elements that are detected as static, the stored background signal is altered by $+1$, 0, or -1 out of 256 per picture toward the present reconstructed signal. Thus, it takes a large number of picture periods to adapt the background memory contents to new background information; on the other hand, the algorithm tends to keep stored background information within the background memory.

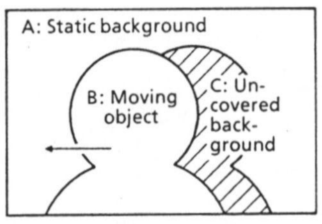

Fig. 1. Picture region types in case of a videoconference or picture telephone scene.

The algorithm in [1] and [2] is based on a more complex change detector. It is capable of instantaneously taking new background information into the background memory. But it does not prevent the replacement of stored background information by object information.

In Section II of this paper, a new algorithm for gaining background information is presented, which is developed from [7] and [10]. It is based on a reliable change detector and evaluates the change information of successive pictures. By this, it offers the capability of quickly taking new background information into the background memory, as an advantage over [10], as well as preserving stored background information from further deletion, as an advantage over [1] and [2].

The efficiency of the new algorithm in terms of prediction error reduction is uncovered background regions is analyzed in comparison to intraframe and motion-compensating prediction in Section III. Implementation of the uncovered background predictor in a hybrid-DPCM/transform coder environment is described in Section IV. Results regarding bit rate saving as well as subjective picture quality improvement are presented and discussed. Investigations have been carried out by means of computer simulations.

II. GAINING BACKGROUND INFORMATION FOR UNCOVERED BACKGROUND PREDICTION

Before picture information of the scene background can be used for uncovered background prediction, it first has to be detected and stored in a certain background memory.

The process of gaining background information consists of two phases: the initialization of the background memory at the beginning of a video sequence or after a scene change, and the further continuous updating with not yet stored background information during the sequence.

The new algorithm for gaining background information is based on:

- the detection of static scene background by means of a change detector which evaluates the reconstructed signal,
- the fast storage of background information into the background memory if not yet stored, and
- the update of stored background information by slowly adapting it to the real scene background.

Fig. 2 shows the general block diagram of an implementation. The first stage is formed by a reliable change detector, as described in [3]. The second stage is the actual background generator, i.e., the practical realization of the algorithm for gaining background infor-

---

Paper approved by the Editor for Image Communication Systems of the IEEE Communications Society. Manuscript received June 2, 1988; revised May 11, 1989. This paper was presented in part at the 1987 Picture Coding Symposium; and submitted in part to the 1990 Picture Coding Symposium.

The author was with the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover, Appelstr. 9 A, D-3000 Hannover 1, F. R. Germany. He is now with Thomson Consumer Electronics, R&D Laboratories, Hannover, Germany.

IEEE Log Number 9037894.

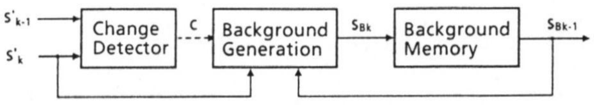

Fig. 2. Principal block diagram of the new algorithm for gaining background information. k = time index of present picture, s' = reconstructed signal, c = change detection signal (0 = static, 1 = changed), s_B = background signal.

mation. The third stage is the background memory, the output of which forms the uncovered background prediction signal.

For initializing the background memory, the first reconstructed picture of a scene is directly stored. Thus, certain picture regions already contain true background information.

From then on, the change detector evaluates the difference signal between the present and the previous reconstructed picture. For every picture element, absolute values of the difference signal are summed up over a square window; and by comparison of the sum to a certain threshold, a binary signal is generated which assigns the picture element either to a static or to a changed region. This signal is processed by median filtering and subsequent elimination of small singular regions.

Whenever a picture element is assigned to static background for exactly n successive pictures, the corresponding reconstructed signal value is directly stored into the background memory. Thus, compared to [10], background information is obtained quickly.

If a picture element is found to belong to a static region and background information has already been stored for it, the existing gray value is altered by -1, 0, or $+1$ out of 256 toward the new reconstructed signal value, as described in [10]; by this, possible wrong decisions are corrected and slight background luminance changes are tracked.

The background generation procedure can therefore be summarized by the following formula:

$$s_{Bk} = \begin{cases} s'_1 & \text{if } k = 1 \text{ (initialization)} \\ s'_k & \text{if } c = 0 \text{ for the } n\text{th time} \\ s_{Bk-1} - 1 & \text{if } c = 0 \text{ for at least the } n\text{th time} \\ & \quad \text{and } s_{Bk-1} > s'_k \\ s_{Bk-1} + 1 & \text{if } c = 0 \text{ for at least the } n\text{th time} \\ & \quad \text{and } s_{Bk-1} < s'_k \\ s_{Bk-1} & \text{if } c = 0 \text{ for at least the } n\text{th time} \\ & \quad \text{and } s_{Bk-1} = s'_k \\ & \text{or } c = 0 \text{ for less than } n \text{ times} \\ & \text{or } c = 1 \end{cases}$$

where

k = picture number as a time index
s' = reconstructed signal
s_B = background signal
c = change detector signal (0 = static, 1 = changed)
n = delay parameter in units of successive pictures.

The parameter n represents a delay in the process of gaining new background information. A value of $n = 1$ provides immediate action, whereas a value of $n \geq 2$ is more suitable where momentary wrong decisions in the change detection process can occur.

The luminance component of the CCITT video test sequence "Trevor White," sampled with 352 pels * 288 lines per picture and 10 pictures/s, and with a length of 3 s, has been processed. The sequence shows a head-and-shoulder view of a man sitting at a table, talking, and moving his arms, and is therefore suitable for investigations in the area of uncovered background prediction.

Fig. 3 shows the contents of the background memory at the end of the sequence for the new algorithm with $n = 1$ and for the algorithm described in [10]. Background information is obtained much more quickly by the new algorithm. The residual parts of the person in Fig. 3(a) indicate regions where no static scene background information could yet be detected and stored. Slight mistakes of object information taken as background information, such as close to the arm in the left part of Fig. 3(a), can be overcome by further alteration of the background memory contents by -1, 0, or $+1$ out of 256; they can be avoided by enlarging the value of n to 2—this increase in reliability is at the expense of slightly fewer picture elements for which background information is gained.


(a) (b)

Fig. 3. Background signal after 3 s (test sequence "Trevor White" at 10 pictures/s with 352 pels * 288 lines each), in case of (a) new algorithm, and (b) algorithm described in [10].

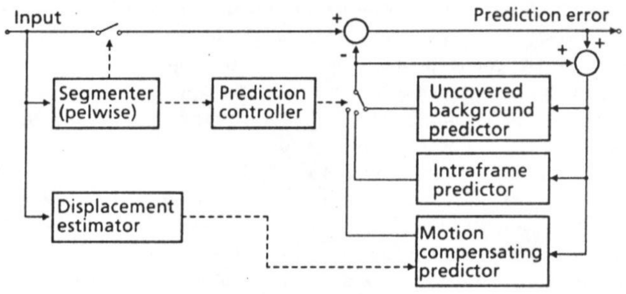

Fig. 4. Block diagram of the experimental coder for measuring the prediction error entropy without considering quantization effects.

III. Efficiency Analysis of Uncovered Background Prediction in Terms of Prediction Error Reduction

The efficiency of uncovered background prediction in terms of prediction error reduction has been investigated by means of an experimental predictive coder, and has been compared to that of motion-compensating and intraframe prediction in uncovered background regions.

A. Strategy for Prediction Error Measurement

A block diagram of the experimental coder is shown in Fig. 4. The first stage is formed by a segmenter, as described in [3], which assigns each picture element to one of the regions A, B, or C (see Fig. 1). The picture element is coded by DPCM without quantization.

Displacement vector estimation for motion-compensating prediction is done by 3-step search block matching [8], [11] with the sum of absolute differences as minimization criterion. For every 8-by-8 block, one vector is estimated with integer pel accuracy, with a search area of ±7 picture elements in horizontal and vertical directions.

For intraframe prediction of the present picture element, both the horizontally and the vertically adjacent elements in the causal neighborhood are weighted by 1/2 each.

For uncovered background prediction, background information is obtained in a first simulation run along the video test sequence by means of the new algorithm with $n = 1$. This information is then used for prediction in the subsequent experiments where the same sequence is coded.

Prediction error statistics in uncovered background regions are procured and evaluated for each of the three different types of predictors.

The test sequence used is again "Trevor White," with the

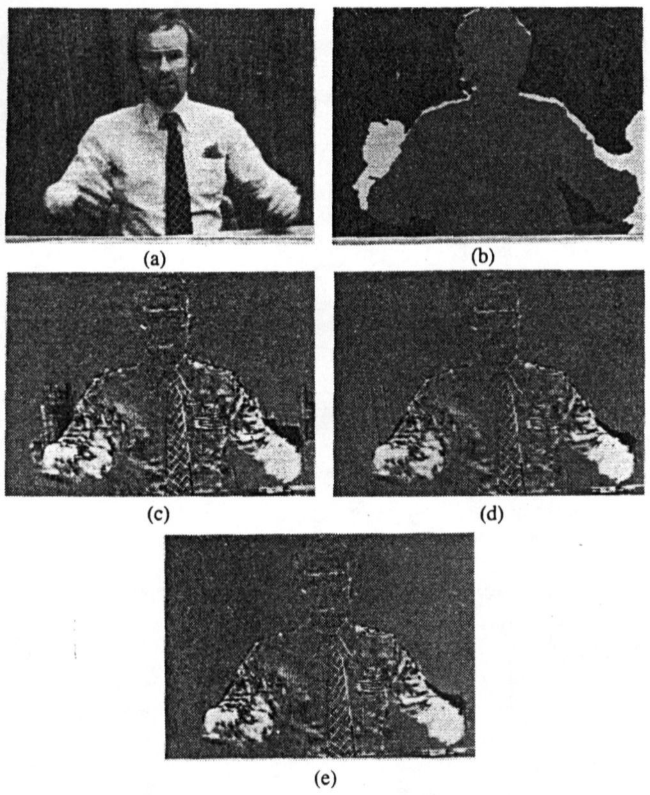

Fig. 5. Input signal, segmenter mask, and prediction error signal for the picture at $t = 2.7$ s of sequence "Trevor White:" (a) input signal, (b) segmenter mask (dark = static background regions [A], medium gray = moving object regions [B], light = uncovered background regions [C]); and (c)-(e) prediction error signal multiplied by 4 for: (c) motion-compensating prediction in regions B and C, (d) motion-compensating prediction in regions B, and intraframe prediction in regions C, (e) motion-compensating prediction in regions B, and uncovered background prediction in regions C.

sampling format as in Section II. Only the luminance signal is considered. To reduce the influence of camera noise on the results, the input signal is temporally filtered slightly by means of a simple nonlinear recursive noise reduction filter, as described in [6].

B. Comparison of Prediction Error Properties in Uncovered Background Regions

In order to demonstrate the prediction error properties in uncovered background regions for each of the three predictors, Fig. 5 shows the corresponding prediction error signals for a typical picture out of the test sequence where large areas of uncovered background occur. The prediction error amplitudes are magnified by a factor of 4, and an offset of 128 out of 256 is added. In addition, the input picture and the segmenter output are shown.

In uncovered background regions, prediction errors are large with motion-compensating prediction since the motion estimator cannot find picture information in the preceding picture corresponding to the present one. With intraframe prediction, prediction errors are reduced, but they are still significant and have to be transmitted. As a result of the 2-dimensional prediction method, the structure of the picture signal itself can be seen in the prediction error signal.

With uncovered background prediction, prediction errors are significantly reduced. Only some noise is present due to input signal noise amplitudes remaining after noise reduction filtering.

Regarding the prediction error entropy in uncovered background regions, Table I contains the values for the three different predictors. By uncovered background prediction, the entropy is reduced by 2.08 b/pel compared to motion-compensating prediction, and by 1.03 b/pel compared to intraframe prediction.

For comparison, the frame difference entropy in static background regions is regarded as an indication of the lower bound of

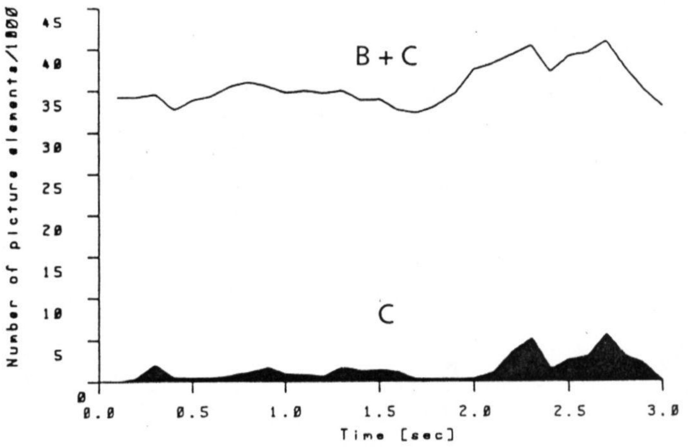

Fig. 6. Number of picture elements in regions C and B + C as a function of time for sequence "Trevor White" at 10 pictures/s (total number of picture elements per picture = 101 376).

TABLE I
PREDICTION ERROR ENTROPY VALUES IN UNCOVERED BACKGROUND REGIONS IN CASE OF MOTION-COMPENSATING PREDICTION (MCP), INTRAFRAME PREDICTION (IaFP), AND UNCOVERED BACKGROUND PREDICTION (UBP)

Predictor	H [bits/pel]	H-difference [bits/pel]
MCP	4.74	2.08
IaFP	3.69	1.03
UBP	2.66 *	

* for comparison: frame difference entropy in static background regions: 2.49 bits/pel

prediction error entropy in case of uncovered background prediction in uncovered background regions; for the test sequence used, it is 2.49 b/pel. The prediction error entropy of 2.66 b/pel obtained by uncovered background prediction is close to this value. This demonstrates the efficiency of uncovered background prediction, in general, as well as the performance of the new algorithm for gaining background information.

Naturally, the gain is somewhat less in case of n being 2 instead of 1, because due to the longer delay, less background information is gained during the sequence; on the other hand, the reliability is increased as required in practical implementation.

For further comparison, a prediction error entropy of 4.33 b/pel was obtained by means of the algorithm of [10], and 4.01 b/pel were yielded using the algorithm after [1] and [2] with the change detector as above.

The prediction error entropy gain with uncovered background prediction is achieved in uncovered background regions only. Bit rate reduction for the whole picture strongly depends on the relative size of those areas. Fig. 6 shows the number of picture elements in uncovered background regions (C), compared to that in temporally changed regions (B + C), as a function of time. The difference between the two curves, i.e., the number of moving object picture elements varies only slightly, whereas the number of picture elements of region C varies significantly according to the amount of translatory motion and, therefore, to the size of uncovered background areas.

As mentioned before, the prediction errors in uncovered background regions show a noiselike structure in the case of uncovered background prediction [see Fig. 5(e)], which would appear in the reconstructed picture, if transmitted. Instead, if background information is available in the background memory, control information could be transmitted, which is far less bit consuming.

In order to examine if that is permissible, the number of picture elements in region C which are "predictable" by uncovered background prediction is estimated. For this purpose, the prediction

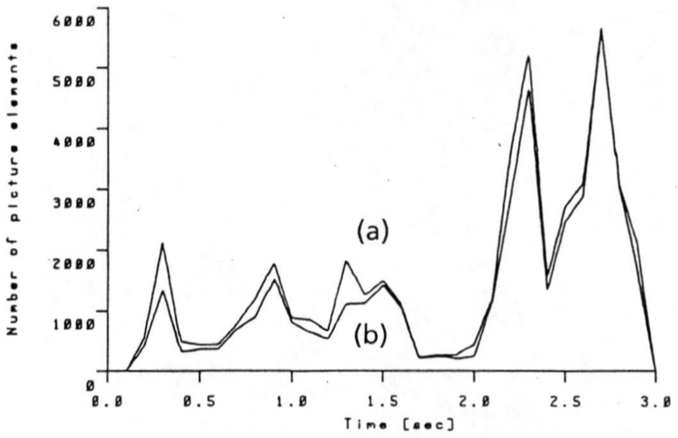

Fig. 7. Number of picture elements in uncovered background regions (sequence "Trevor White" at 10 pictures/s, total number of picture elements per picture = 101 376): (a) total, and (b) predictable by uncovered background prediction without prediction error transmission.

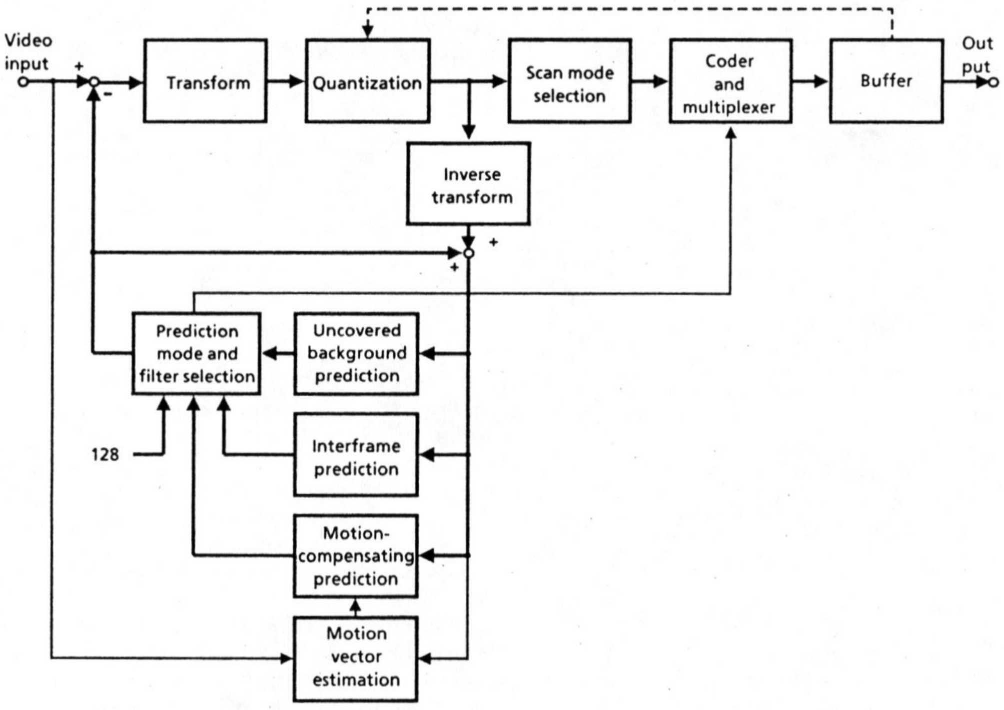

Fig. 8. Block diagram of the hybrid-DPCM/DCT source coder (128 = offset for intramode [1/256]).

value is compared to that of the original picture by means of the change detector described in [3]. Those picture elements in region C which, according to the change detector's decision, are unchanged, are regarded as being "predictable." Fig. 7 shows their number as a function of time along with the total number of uncovered background picture elements: the number of "predictable" picture elements is close to, and only slightly less than, the total number of uncovered background picture elements. It is obvious that, in case of uncovered background prediction with a reliable algorithm for gaining background information, prediction error transmission for uncovered background regions can be avoided as hitherto in case of conditional replenishment coders for static background regions only. Thus, the relative bit rate saving by uncovered background prediction is roughly indicated by the quotient:

$$\frac{\text{number of picture elements in region C}}{\text{number of picture elements in regions B + C}}$$

which, according to Fig. 6, is up to 14% in case of the sequence used.

IV. A Hybrid-DPCM/Transform Coder with Uncovered Background Prediction

The efficiency of uncovered background prediction in a low bit rate coder is investigated. A hybrid-DPCM/transform coder with a motion-compensating predictor, the structure of which currently is considered for low bit rate applications [12], is supplemented by an uncovered background predictor.

A. Coder Structure

Fig. 8 shows the coder block diagram. The video input signal in component form is fed into the source coder. The sampling of the input signal is: 352 pels * 288 lines per picture for the luminance; 176 pels * 144 lines per picture for the two chrominance components; and 7.5 pictures/s, noninterlaced.

Each picture is processed on a block-by-block basis with a blocksize of 16 pels by 16 lines.

The first stage of the source coder is formed by a DPCM loop. A motion-compensating predictor, a pure interframe predictor, and an

TABLE II
DERIVATION OF THE BLOCK MODE FOR THE LUMINANCE COMPONENT

Prediction/ intra mode	with filter	with non-zero coefficients
Intra mode *		
Motion-compensating prediction	yes	yes *
		no
	no	yes *
		no
Inter-frame prediction	yes	yes *
		no
	no	yes *
		no **
Uncovered background prediction	yes	yes *
		no
	no	yes *
		no

* "active" block ** "unmodified" block

uncovered background predictor are available for coding the luminance component. For the two chrominance components, pure interframe prediction is used, i.e., no motion compensation is carried out. Motion vector estimation for motion compensation is done by block matching with a three-step search procedure, with a maximum displacement of ±7 pels. Background information is obtained in a separate simulation run along the test sequence, as in Section III-A, by implementing the new algorithm with $n = 1$. A planar filter is applied to the prediction signal if it reduces the prediction errors [5].

The prediction error signal of a block is transformed by means of a DCT. The 12-bit transform coefficients are quantized and transmitted. For the luminance component, one out of four different scan orders (zig-zag, horizontal, vertical, or circular) is used, up to and including the last nonzero coefficient within the block; all these coefficients—the number of which is referred to as "scan length"—are entropy coded and transmitted as pairs of amplitudes and zero runlengths followed by a codeword marking the last transmitted coefficient as the "end of block." For the two chrominance components, only zig-zag scanning is used. The quantizer has a uniform characteristic; however, the input range for zero output is extended to twice the step size. The step size is selected according to the buffer fullness to control the amount of bits generated by the coder. If all coefficients in a block are zero after quantization, that block information will be coded separately. Blocks containing nonzero coefficients are referred to as "active blocks."

If motion-compensating prediction, interframe prediction, and uncovered background prediction generate large prediction errors, the picture signal itself with an offset of 128 out of 256, rather than the prediction error signal, is transformed and transmitted. This happens, e.g., if the motion estimation is inaccurate or the model assumptions for motion-compensating prediction, interframe prediction, and uncovered background prediction do not hold. This mode is referred to as the "intramode." It is also used to build up the first picture of a scene or after a scene change. In that case, twice the normal bit rate is provided, and the second picture is dropped at the transmitter side. At the receiver, the last picture before the scene cut is repeated once.

For every block, the information about prediction mode, possible planar filtering, and the presence of nonzero coefficients are jointly coded. In total, 13 different block modes are available according to Table II. For inactive blocks with interframe prediction and no filter, the reconstructed signal remains unchanged. For consecutive unchanged blocks, only addressing information is transmitted using a runlength code; the other 12 block modes are coded by a block-type entropy code. Thus, block mode coding comprises runlength coding and block-type entropy coding.

B. Experiments

The following experiments have been carried out with the source coder described above.

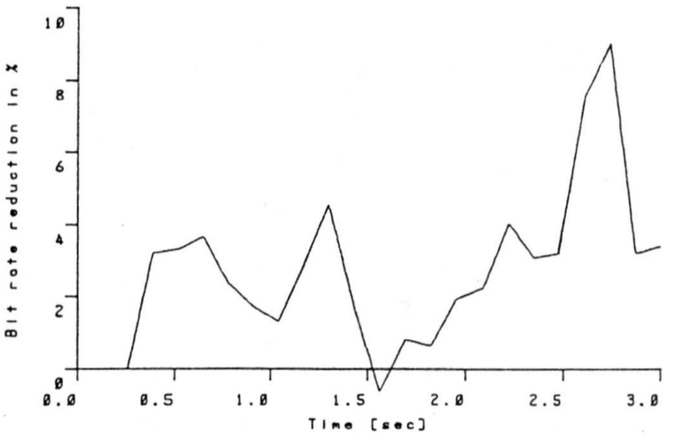

Fig. 9. Bit rate reduction in percent by uncovered background prediction for the coder in open loop (quantizer step size = 8 out of 2048).

TABLE III
COMPARISON OF NUMERICAL RESULTS FOR THE CODER IN OPEN LOOP WITHOUT (α-CODER) AND WITH UNCOVERED BACKGROUND PREDICTION (β-CODER)

	Transmission bit rate [kbit/s]	Sequence average		Max-picture *	
		Value with β-coder	Comparison to α-coder	Value with β-coder	Comparison to α-coder
Number of active Y blocks per picture	64	137	-6.5%	144	-11.7%
	128	157	-12.3%	156	-24.3%
	304	223	-10.8%	230	-22.6%
Number of bits for Y coefficients per active block	64	41.06	+ 8.2%	44.03	+17.6%
	128	84.49	+14.8%	91.59	+32.9%
	304	146.02	+11.8%	145.10	+25.9%
Average quantizer step size [1/2048]	64	32.44	-1.2%	30.22	-3.5%
	128	18.38	-1.3%	16.22	-5.8%
	304	8.72	-1.4%	7.56	-6.8%
S/N, Y [dB]	64	34.88	+0.12	35.59	+0.27
	128	37.81	+0.12	38.65	+0.40
	304	41.72	+0.10	42.45	+0.47

* picture at t=2.7 sec with largest areas of uncovered background

1) Coder in Open Loop: A fixed quantizer step size is used, and the bit rates generated by the coder with and without uncovered background prediction are compared.

2) Coder in Closed Loop: The transmission bit rate is fixed; for three different values: a) 64 kb/s, b) 128 kb/s, and c) 304 kb/s (the considered video net bit rate in a 384-kb/s environment), picture quality and bit rate partitioning are compared.

The control range of quantizer step size is chosen from 8 to 64 out of 2048 at 64 kb/s, from 6 to 48 at 128 kb/s, and from 4 to 32 at 304 kb/s. For the coder in open loop, the quantizer step size is set to 8.

Test sequence "Trevor White" is processed.

C. Results

For convenience, the following terms are introduced:

α-Coder—Coder without uncovered background prediction; and

β-Coder—Coder with uncovered background prediction.

1) Coder in Open Loop: The simulation results obtained with the coder in open loop are gathered in Table III. With uncovered background prediction, the average bit rate is reduced by 3%—or, taking into account the picture frequency of 7.5 Hz, from 306 kb/s to 297 kb/s. In the case of relatively large uncovered background areas, occurring in the picture at $t = 2.7$ s—which is referred to as the "max-picture"—the reduction is 9.1%, or 10.5% for the luminance component only.

For a delay parameter in the process of gaining background

TABLE IV
NUMERICAL RESULTS FOR THE CODER WITH UNCOVERED BACKGROUND PREDICTION (β-CODER) IN CLOSED LOOP, AND COMPARISON TO THE CODER WITHOUT (α-CODER)

	Sequence average		Max-picture *	
Coder	α-coder	β-coder	α-coder	β-coder
Number of bits per picture:				
Total	40808	39577	40856	37151
Luminance	36429	35203	35724	31980
Coefficients	33003	32063	32223	29126
End of block + scan type	1558	1261	1596	1104
Motion vectors	950	901	912	744
Block mode	917	977	993	1006
Chrominance	4379	4374	5132	5171
Number of active Y blocks per picture	260	210	266	184

* picture at t=2.7 sec with largest areas of uncovered background

TABLE V
NUMERICAL RESULTS REGARDING LUMINANCE TRANSFORM COEFFICIENTS FOR THE CODER WITH UNCOVERED BACKGROUND PREDICTION (β-CODER) IN CLOSED LOOP, AND COMPARISON TO THE CODER WITHOUT (α-CODER)

Transmission bit rate [kbit/s]	Average number of non-zero coefficients per active block		Average scan length	
	Value with β-coder	Comparison to α-coder	Value with β-coder	Comparison to α-coder
64	7.13	+ 7.1 %	24.82	+ 6.0 %
128	12.28	+ 10.4 %	43.25	+ 9.8 %
304	19.51	+ 7.3 %	68.06	+ 7.3 %

information of $n = 2$, which provides greater reliability rather than immediate action, as with $n = 1$, bit rate saving would be about 0.3% less.

Fig. 9 shows the relative bit rate reduction, regarding all components, as a function of time. The peaks correspond to those pictures where a large amount of motion leads to large areas of uncovered background, and they demonstrate the gain yielded by uncovered background prediction.

Bit rate reduction is achieved by saving bits in coding active blocks, as well as by reducing the number of those blocks (see Table III). It is caused by smaller amplitudes in the prediction error signal, and fewer high-frequency components, or smaller coefficient amplitudes which are forced to zero by the quantizer. Thus, the bit rate for coding coefficient, end of block, and scan-type information is lessened (see Table III).

A large number of blocks is now inactive and coded by uncovered background prediction and without filtering; with the α-coder, these blocks mainly have been active and coded with motion-compensating prediction and with filtering or in the intramode, since these modes are appropriate for uncovered background regions in that case.

Furthermore, in case of the α-coder, there are a lot of active blocks in static background regions coded with interframe prediction and without filtering, due to camera noise. This causes visible temporal business in the reconstructed signal. With uncovered background prediction, the number of those blocks, and therefore temporal business, is widely reduced due to the noise reduction property of the background generation algorithm.

2) Coder in Closed Loop: Tables IV and V show numerical results for the β-coder at the three given fixed transmission bit rates, as well as a comparison to the α-coder. The absolute values of the quoted differences between the β-coder and the α-coder represent the gain yielded by uncovered background prediction.

As with the coder in open loop, the number of active blocks is reduced with uncovered background prediction. At a transmission bit rate of 64 kb/s, corresponding to 8533 b/picture, an average 6.5% less active luminance blocks occur; for the max-picture, the reduction is even about 12%. Fig. 10 shows the percentages of active blocks per picture as a function of time for both the α-coder and the β-coder at 64 kb/s. The percentage is reduced with uncovered background prediction, especially for pictures with large areas of uncovered background.

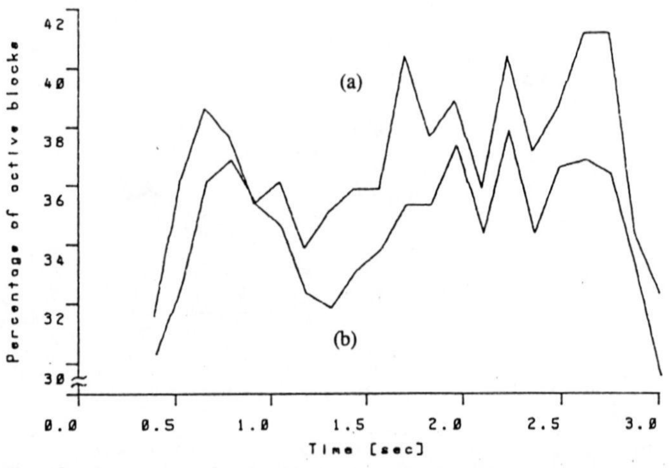

Fig. 10. Percentage of active blocks, i.e., blocks with nonzero transform coefficients, for the coder in closed loop at 64 kb/s: (a) coder without, and (b) coder with uncovered background prediction.

With this reduction, more bits are available per active block—8.2% on average, and 17.6% for the max-picture for coding luminance coefficients. Therefore, finer quantization is applied so that more nonzero coefficients per active block occur: the average number of nonzero coefficients per active block increases by 7.1%, the average scan length by 6% (see Table V).

At transmission bit rates of 128 and 304 kb/s, similar results are achieved.

Since motion vector coding and block mode coding consume nearly the same data rate at the given bit rates, the higher number of bits being available at 128 and 304 kb/s is mainly used in coding transform coefficients. Increasing the bit rate from 64 to 128 kb/s, i.e., by a factor of 2, causes an increase by a factor of about 2 in the number of bits for luminance transform coefficients per active block for both the α-coder and the β-coder. These bits are used for finer quantization and, thus, for transmitting more nonzero coefficients per active block. The number of active blocks also increases, since noise in the input signal along with finer quantization gives rise to active blocks in the background. Similar statements can be formulated for the transition from 128 to 304 kb/s.

With increasing transmission bit rate, the uncovered background predictor provides an increasing improvement of the signal-to-noise ratio, since the uncovered background prediction signal contains less quantization noise.

As regards subjective picture quality, at 128 and 304 kb/s, the reconstructed pictures appear sharper in case of the β-coder, since more nonzero coefficients per active block, i.e., more high-frequency components, are transmitted.

At 64 kb/s, however, only a little improvement can be observed in overall picture quality, since the test sequence "Trevor White" contains large, temporally changed areas as well as a lot of spatial detail; therefore, at a transmission rate of 64 kb/s, the coder can only provide a relatively poor picture quality; and the signal-to-noise ratio of the reconstructed signal compared to the original and related to 255^2 is about 35 dB.

At 128 kb/s, the signal-to-noise ratio is about 38 dB, and improvement in subjective picture quality by uncovered background prediction is evident, especially for pictures with large uncovered background areas, e.g., the max-picture. Pictures show less coding noise, i.e., there is less temporal business noticeable around the moving object (known as "mosquito effects"), and fewer block structures are visible (known as "blocking effects"). These statements also apply to a comparison of subjective picture quality at 304 kb/s.

V. CONCLUSION

An algorithm for gaining picture information of the scene background of a TV sequence is presented. It is based on detection of static regions by evaluating the reconstructed signal; picture information is stored into a background memory if it has been static for an adjustable number of successive picture periods, and if it has not yet been stored; possible wrong decisions are corrected, and slight changes of background luminance are tracked.

The efficiency of uncovered background prediction in a quantization error-free DPCM coding environment is analyzed. The prediction-error entropy in uncovered background regions is reduced close to the frame difference entropy in static regions resulting from noise. Most of the picture elements in uncovered background regions can be reconstructed without transmission by the information from the background memory. Thus, the relative bit rate saving by uncovered background prediction is roughly indicated by the number of picture elements in uncovered background regions divided by the number of picture elements in temporally changed regions. This saving depends on the velocity of motion, and amounts to up to 14% for the test sequence used.

Implementing uncovered background prediction in a motion-compensating hybrid-DPCM/transform coder in open loop with fixed quantization, a maximum bit rate reduction of 9% is achieved. With the coder in closed loop, at a fixed transmission rate of 128 or 304 kb/s, the reconstructed pictures appear sharper and show less coding noise when uncovered background prediction is applied: "mosquito effects" and "blocking effects" are reduced, as well as temporal business in static background regions. Along with the increase in subjective picture quality, the signal-to-noise ratio is improved as well. This improvement increases with higher bit rates. At 64 kb/s, the improvement of subjective picture quality is marginal.

The capability of uncovered background prediction is not fully exploited in the described hybrid coder because of the block-oriented processing and coding. The coding gain which can be provided by uncovered background prediction may be obtained by an object-oriented coding technique [17] that distinguishes between moving object, static background, and uncovered background regions, and adapts the prediction mode correspondingly.

ACKNOWLEDGMENT

The author would like to thank Prof. Dr.-Ing. H. G. Musmann for encouraging and supporting this work, and for various stimulating discussions. He is also indebted to H. Schiller for software support, and to H. Li for his assistance in computer simulations.

REFERENCES

[1] S. Brofferio, "Predictive videophone coding using background knowledge," in *Conf. Rec. PCS,* 1986, pp. 87-88.

[2] S. Brofferio and V. Corradi, "Videophone coding using background prediction," in *Conf. Rec. EUSIPCO,* 1986, vol. 2, pp. 813-816.

[3] M. Bierling and R. Thoma, "Motion compensating field interpolation using a hierarchically structured displacement estimator," *Signal Processing,* vol. 11, no. 4, Dec. 1986.

[4] H. Brusewitz and P. Weiss, "A videoconference system at 384 kb/s," in *Conf. Rec. PCS,* 1986.

[5] H. Brusewitz, "Filtering in the hybrid coding loop," in *Conf. Rec. PCS,* 1987, pp. 196-197.

[6] D. J. Crawford, "Spatio-temporal filtering in television picture coding," Ph.D. dissertation, Univ. Essex, July 1983.

[7] D. Hepper and H. Li, "Analysis of uncovered background prediction for image sequence coding," in *Conf. Rec. PCS,* 1987, pp. 192-193.

[8] J. R. Jain and A. K. Jain, "Displacement measurement and its application in interframe image coding," *IEEE Trans. Commun.,* vol. COM-29, pp. 1799-1808, Dec. 1981.

[9] G. Kummerfeldt, F. May, and W. Wolf, "Coding television signals at 320 and 64 kb/s," in *Conf. Rec. 2nd Int. Tech. Symp. Opt. Electro-Opt. Appl. Sci. Eng., SPIE Conf. B 594 Image Coding,* Cannes, France, Dec. 1985.

[10] N. Mukawa and H. Kuroda, "Uncovered background prediction in interframe coding," *IEEE Trans. Commun.,* vol. COM-33, pp. 1227-1231, Nov. 1985.

[11] H. G. Musmann, P. Pirsch, and H.-J. Grallert, "Advances in picture coding," *Proc. IEEE,* vol. 73, pp. 523-548, Apr. 1985.

[12] S. Okubo, "Work progress in 'Specialists group on coding for visual telephony,'" in *Conf. Rec. PCS,* 1987, pp. 153-154.

[13] S. Sabri, K. Cuffling, and B. Prasada, "Coding of video signals at 50 kb/s using motion compensation techniques," in *Conf. Rec. IEEE Mil. Commun. Conf.,* Nov. 1983.

[14] H. Schiller and P. Gerken, "Progressive updating of unchanged picture areas in a hybrid coding system," in *Conf. Rec. PCS,* 1987, pp. 194-195.

[15] R. H. J. M. Plompen, J. G. P. Groenveld, J. Biemond, and F. Booman, "The application of a translation invariant transform for low bitrate video coding," in *Conf. Rec. EUSIPCO,* 1986, pp. 845-848.

[16] F. Sugiyama, K. Dachiku, and T. Watanabe, "64 kbps to 1.5 Mbps color video codec with transform domain motion detection," in *Conf. Rec. PCS,* 1987, pp. 155-156.

[17] H. G. Musmann, M. Hötter, and J. Ostermann, "Object oriented analysis-synthesis coding of moving images," *Signal Processing (Image Communications),* Special Issue on 64 kb/s Coding of Moving Video, vol. 1, no. 2, pp. 117-138, Oct. 1989.

★

Dietmar Hepper was born in Hamburg, Germany, in 1957. He studied electrical engineering at the University of Hannover, Germany, with emphasis on information processing, and received the Dipl.-Ing. degree in 1983.

He was with the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover, working on low bit rate image coding for videoconference and visual telephony applications. Since 1988 he has been with Thomson Consumer Electronics, R&D Laboratories, Hannover, Germany, where he is working on video data compression for digital storage media as well as TV and HDTV applications.