

An experimental digital VCR with 40 mm drum, single actuator and DCT-based bit-rate reduction

Citation for published version (APA):

Borgers, S. M. C., Heijnemans, W. A. L., Niet, de, E., & With, de, P. H. N. (1988). An experimental digital VCR with 40 mm drum, single actuator and DCT-based bit-rate reduction. *IEEE Transactions on Consumer Electronics*, 34(3), 597-605. <https://doi.org/10.1109/30.20159>

DOI:

[10.1109/30.20159](https://doi.org/10.1109/30.20159)

Document status and date:

Published: 01/01/1988

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

AN EXPERIMENTAL DIGITAL VCR WITH 40 MM DRUM, SINGLE ACTUATOR AND DCT-BASED BIT-RATE REDUCTION

S. M. C. Borgers

Philips Consumer Electronics,
Nederlandse Philips Bedrijven B.V.,
Building SFJ 7, 5600 MD Eindhoven, The Netherlands

W. A. L. Heijnemans, E. de Niet, P. H. N. de With
Philips Research Laboratories Eindhoven
Nederlandse Philips Bedrijven B.V.,
P.O. Box 80.000, 5600 JA Eindhoven, The Netherlands

ABSTRACT

An experimental digital video-recording system based on small recording mechanics is described. High picture quality at moderate recording bit rates is obtained with component coded video and application of an advanced bit-rate reduction technique, which is based on the Discrete Cosine Transform (DCT). A single actuator allows for high tracking accuracy and multispeed playback.

1. INTRODUCTION

Digital video recording is expected to become particularly important in those applications where multiple copying is often required, such as electronic imaging. We therefore investigated the feasibility of digital recording on the basis of small recording mechanics suitable for portable use.

For successful use of digital recording in these applications, it is of paramount importance that the bit rate to be recorded is as low as possible, i.e. preferably only a fraction of the say 200 Mbit/s which would be required for straight PCM recording. This gives more manageable requirements on speed of scanner revolution, recording bandwidth and/or number of parallel recording channels, thus making digital recording more suitable for consumer use. It is also the only way to obtain sufficient playing time with a small cassette.

The advantage gained by reducing the complexity of the actual recording system outweighs many times the additional processing circuitry which is needed for bit-rate reduction of the video signal. VLSI technology offers the opportunity for cost-effective implementation of advanced bit-rate reduction schemes with high reduction factors.

The work reported in this paper was carried out in the framework of the European DVS project on digital video recording, and was sponsored by the Dutch Ministry of Economic Affairs. However, the responsibility for the contents of the paper is carried by the authors.

Table 1: Characteristics of the experimental digital recording system

Video coding algorithm	8 x 8 DCT intraframe threshold coding
Video input	Y,U,V (625/50)
Input sampling Y	10.125 MHz
Input sampling U,V	3.375 MHz
Encoder input bit rate	104 Mbit/s
Encoder output bit rate	19 Mbit/s
Overhead (SYNC + ERCO)	7 Mbit/s
Channel modulation (with embedded tracking tones)	8-to-10
Tape	Mp
Tape width	8 mm
Drum diameter	40 mm
Speed of revolution	75 rps
Head	Amorphous ribbon
Head speed	9.4 m/sec
Trackwidth (double head)	2 x 10 μ m
Track pitch	11 μ m
One picture (25 Hz) in	3 x 2 tracks
Average recording bit rate	32 Mbit/s
Instantaneous recording bit rate	2 x 27.4 Mbit/s
Recording wavelength	.68 μ m
Wrap angle	221 $^{\circ}$
Effective wrap angle	210 $^{\circ}$
Playing time (8mm cassette)	\approx 1 3/4 hours

In recent years different experimental digital recording systems have been described, ranging from a system without data reduction [1] to sub-Nyquist sampling with DPCM [2] [3], and Hadamard transform coding with vector quantization [4].

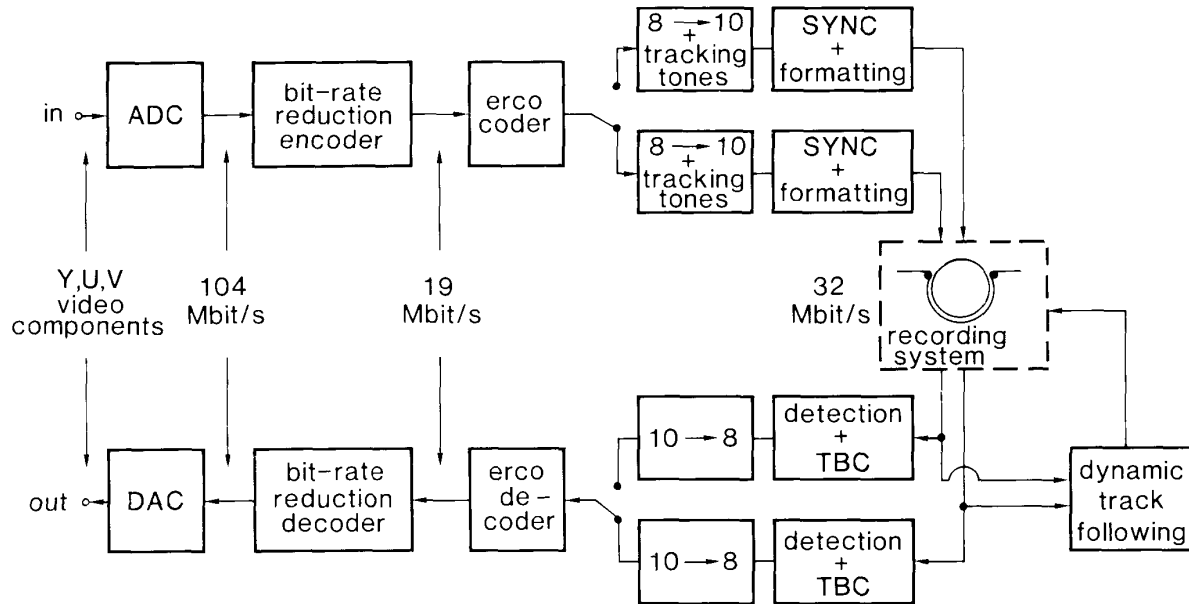


Figure 1. Block diagram of the digital VCR

Recently, a new system called Adaptive Dynamic Range Coding [5] has been proposed, which makes use of variable bit-rate encoding.

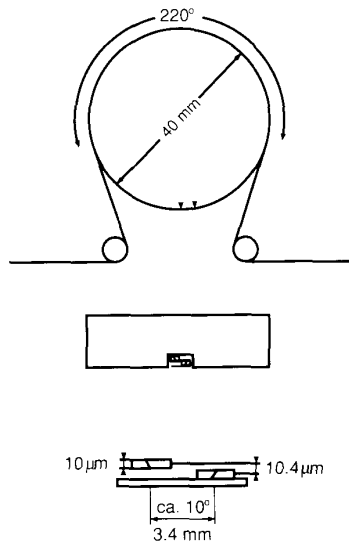


Figure 2. Diagram of scanner with single actuator containing two heads

For the experimental system described in this paper we have investigated a scheme which is based on the Discrete Cosine Transform (DCT) [6] and made it suitable for magnetic recording [7]. Component coded video is used, instead of PAL as was the case with our earlier work in this field [2].

At 19 Mbit/s at the output of the video encoder, the performance of this new scheme was found to be much higher than with the PAL/DPCM system [2]. The authors therefore believe that DCT forms a powerful basis for realizing high compression ratios while maintaining excellent picture quality.

The recording mechanism used is a modified 8mm scanning system. In order to match its recording capacity to the required channel bit rate of 32 Mbit/s (on average), we modified the scanner design to increase both the recording bandwidth and the recording density. On the basis of an 8mm cassette almost two hours of playing time has been obtained. In the new scanner design the concept of the single actuator system plays a substantial role. In addition to the fact that it is optimally suited for recording and tracking of very narrow tracks, it allows for flawless multispeed playback, since blocks of consecutive tracks can be read back at a different tape speed.

The recording system has been realized in real-time hardware. The experimental character is reflected by its great flexibility. Many parameters that control the bit-rate reduction are readily adjustable.

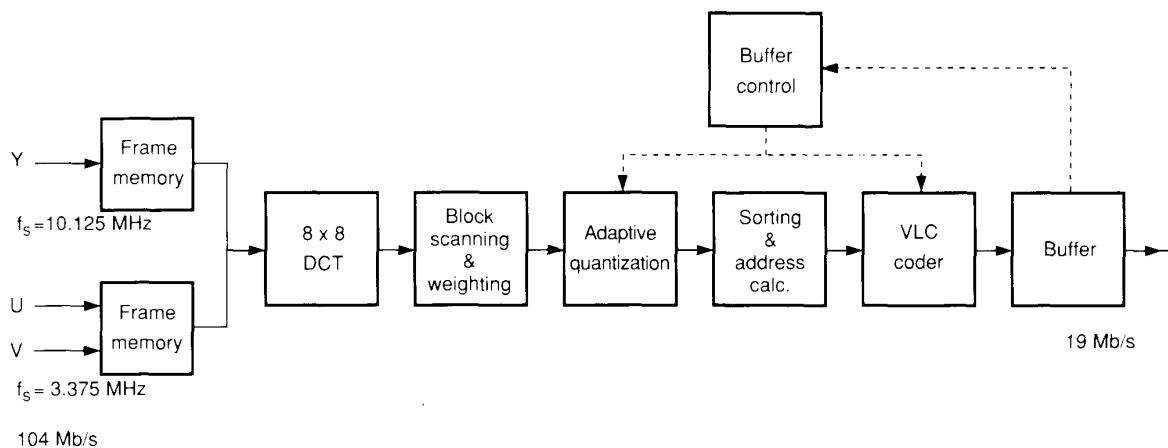


Figure 3. Block diagram of DCT encoder

and within the overall tape format we can experiment with different inner formats.

2. SYSTEM SUMMARY

A block diagram of the digital VCR is depicted in fig. 1, and its main parameters are listed in table 1. The luminance and chrominance signals of the video signal are sampled at 10.125 MHz and 3.375 MHz, respectively. Time-division multiplex of these signals is constructed so that the same data reduction encoder is used for all signal components. Disregarding line and field blanking, the input bit rate to the video encoder equals 104 Mbit/s.

Basically, bit-rate reduction is obtained in two steps. At the start, an orthogonal transform (DCT) is carried out on small two-dimensional blocks within a frame. The result after transformation is a block of spectral components (coefficients) with the signal energy concentrated in a limited number of coefficients. In the second step, these coefficients are quantized and coded statistically, thereby eliminating irrelevant and redundant information. It is emphasized here that each frame is independently coded allowing for individual frame access. This gives full flexibility in the realization of the multispeed modes.

The scanner has been derived from the 8mm system. For accurate track following, also during multispeed playback, a new actuator system has been designed containing a single piezo-ceramic element. The video data are recorded in two parallel channels formed by the two recording heads which are

mounted side by side on this actuator element (fig. 2.). The scanner rotates at 75 Hz, so that each frame is recorded in three segments of two parallel tracks. The channel bit rate is about 2×27 Mbit/s. The channel data are recorded in bursts with a duty cycle which is determined by an effective wrap angle of 210° . The actuator enables the reading of the three consecutive segments of separate frames in multispeed playback mode. Since variable length coding of the video data is used, special care must be taken to avoid overflow of information. This means that the output buffer of the bit-rate reduction encoder must be emptied at the end of each third segment. For very high-speed search-mode operation, segments of different frames must be combined.

3. BIT-RATE REDUCTION SYSTEM

In this section we focus on the applied bit-rate reduction technique. A detailed block diagram of the DCT encoder is depicted in Fig. 3. The key to the blocks indicated is as follows. A line-to-block conversion is carried out after complete video frames have been constructed. The luminance pixel blocks are multiplexed with the chrominance blocks in a 3:1 ratio. In the system, 8x8 pixel blocks are transformed with the DCT. After reordering (scanning) the coefficients are weighted individually to obtain frequency-dependent quantization. Moreover, the quantization is adaptive to the picture statistics. The variable-length coder sorts the quantized components and minimizes address information prior to bit assignment. Finally, the coded data is buffered giving

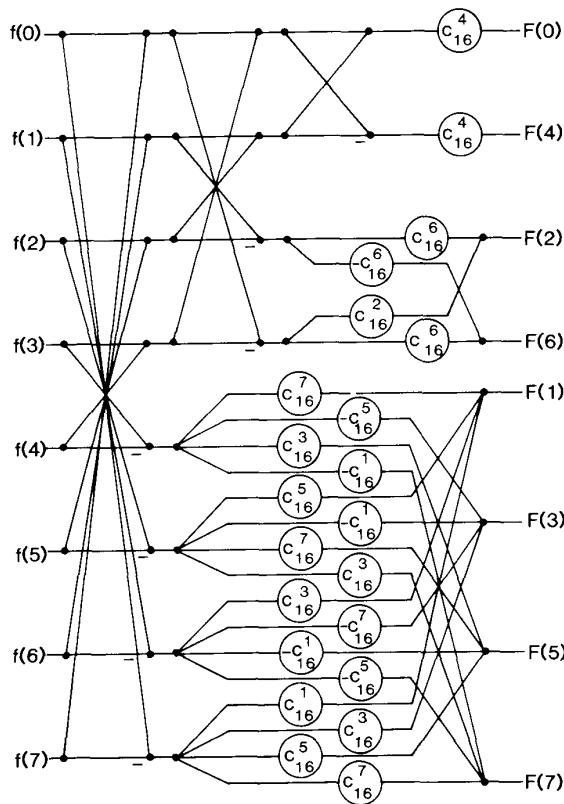


Figure 4. 8-point DCT butterfly diagram

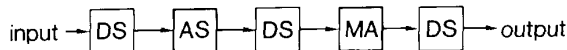


Figure 5. DCT transformer block diagram

a constant encoder output bit rate of 19 Mbit/s. Synchronization information has been added to improve the error resistibility of the compressed data. In addition, error correction is applied.

In the following all parts of the block diagram are described in more detail.

DCT CALCULATION

First we consider the calculation of the DCT and its hardware implementation. From the definition of the two-dimensional DCT [6] it can be derived that

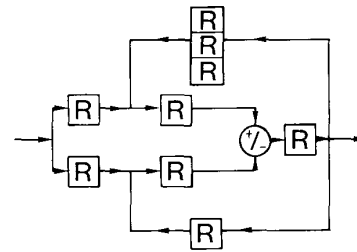


Figure 6. DCT adder/subtractor block

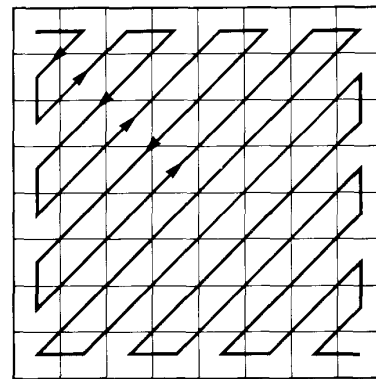


Figure 7. Zigzag scanning pattern

the calculation can be performed by applying two one-dimensional transforms. The number of multiplications and additions in the straightforward calculation of a one-dimensional transform on 8 picture elements is 64 and 56 respectively. Fortunately, fast algorithms [8] [9] [10] have been developed, which reduce the number of operations required substantially.

We have compared several fast algorithms [7], and found that, owing to the range of magnitudes of the cosine terms and the accumulation of multiplication errors, higher calculation accuracies are necessary, compared to the straightforward computation. We therefore designed a new fast algorithm with more multiplications but without the aforementioned disadvantage. The algorithm can be found by subsequently adding pairs of terms that will be multiplied by the same cosine terms [11]. After this, all multiplications are performed and some results have to be accumulated. Fig. 4 shows a

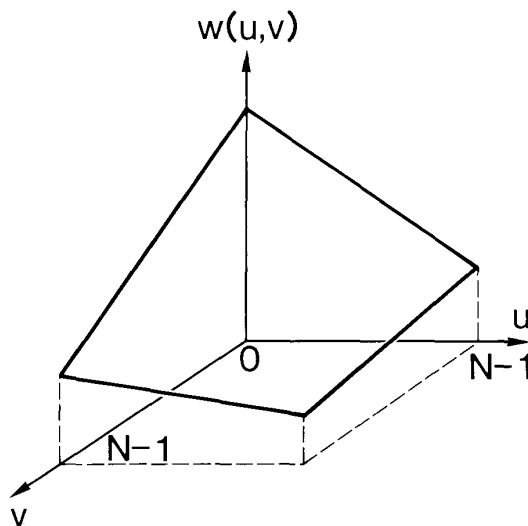


Figure 8. Weighting function

flowgraph of this fast algorithm. The input samples (pixels) are denoted by $f(i)$, $0 \leq i \leq 7$, and the corresponding coefficients at the output as $F(u)$, $0 \leq u \leq 7$.

The proposed algorithm has been implemented in hardware using standard components. The implementation as depicted in Fig. 5 comprises an adder/subtractor (AS), a multiplier/accumulator (MA), and three small register banks (DS) for data shuffling. The 14 additions at the start of each transformation (left side of Fig. 4) are performed serially by the adder/subtractor. Because of the number of additions, the internal speed of the adder/subtractor is doubled. The structure of the flowgraph of the algorithm requires feedback of intermediate calculation results. A detailed diagram of this block is shown in Fig. 6. The blocks (R) are registers for 12 or more bits. The remaining computations of the flowgraph are worked out with two multiplier/accumulators.

The Inverse Discrete Cosine Transform (IDCT) can be obtained by reversing the flowgraph of the algorithm. In the hardware implementation, the IDCT is obtained by interchanging the adder/subtractor and multiplier/accumulator blocks.

QUANTIZATION AND CODING

The coding technique applied in our system is a threshold coding algorithm [12] [13]. This means that only coefficients that have an amplitude above a pre-defined threshold are transmitted to the decoder. Consequently, the decoder needs 'address' information to identify the spectral coordinates of the

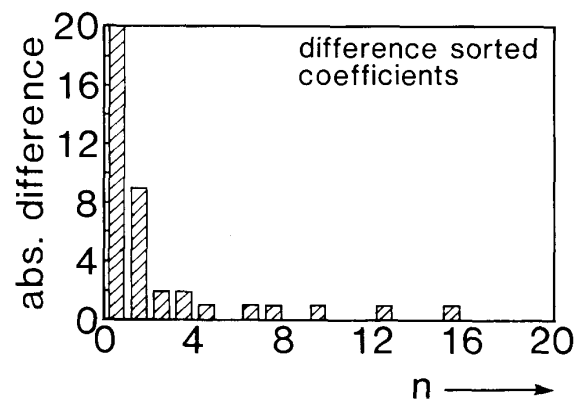
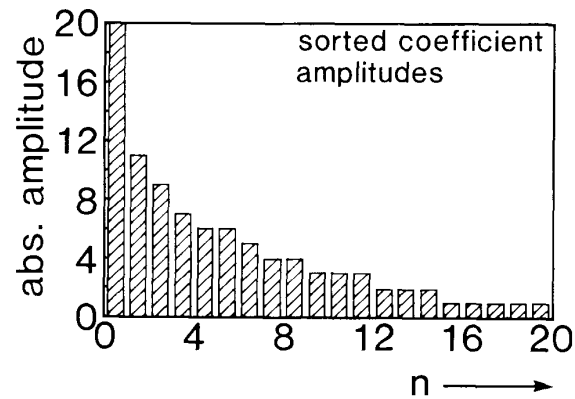


Figure 9. Example of sorted coefficients (upper part) and difference information (lower part)

coefficients received. On average, a major part of the signal energy is located in the low spatial frequencies. The average address information is therefore reduced by assigning low address values to corresponding coefficients. This is obtained by diagonal scanning of the coefficient block [13]. The scanning pattern is shown in Fig. 7.

Quantization of the coefficients is performed in two steps: first by weighting and then by adapting to the picture statistics. It is known from earlier experiments in picture coding [14] that the human eye is less sensitive for observing higher spatial frequencies. To exploit this phenomenon, a weighting function is applied. Each coefficient $F(u,v)$ is multiplied by its weighting factor $w(u,v)$ (see Fig. 8). It can be seen from Fig. 8 that high spatial frequencies have less weight (more amplitude

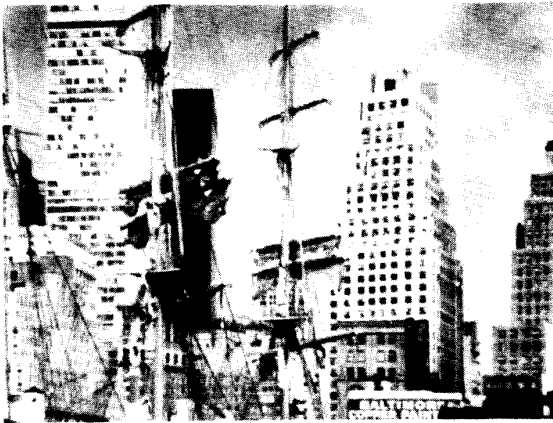


Figure 10. 'Baltimore' scene

suppression) than low spatial frequencies, resulting in a coarser quantization.

The second step in quantization depends on the pictorial data. We have employed a set of linear quantizers with a variable threshold and step size. Adaptation of these parameters is required due to the highly non-stationary character of the video signal. For example, blocks containing signal transients with a high contrast are more coarsely quantized than blocks with low amplitude details. Moreover, the buffer regulation affects the average quantization of the coefficients. The regulation monitors the amount of bits spent during encoding. Buffer overflow is avoided by quantizing the coefficients more coarsely.

The data after quantization are non-uniformly distributed and this motivates the application of variable-wordlength coding. Another reason is the goal for a low bit rate in combination with a high picture quality (high coding efficiency). Several algorithms have been compared [7] and a choice has been made for Differential Decreasing Amplitude Coding. The major algorithm properties can best be illustrated with an example. Let us consider the following block of coefficients (starting with the dc coefficient) after scanning and quantization: 287, 9, -1, 3, 0, 20, -6, -11, 0, 0, 6, -1, 1, 0, ... 0. The ac coefficients that are non-zero are selected for coding but first they are ranked according to their magnitude. Ranking of coefficients in bit planes was initially proposed [15] but we apply a full sort. The contents of the amplitude stack after sorting is depicted in the upper part of Fig. 9. The first coefficient for coding is the maximum amplitude,

followed by the maximum but one, etc. As a second step, the differences between the consecutive coefficient amplitudes are determined and then coded instead of the original amplitude values. The resulting amplitude difference information for coding is shown in the lower part of Fig. 9. It can be noticed that the difference information has smaller amplitudes than the original values and this applies to general data as well. As the coefficients are ranked according to their magnitude, the spatial coordinates are mixed up. Therefore, the corresponding address values of each coefficient are kept in a second stack during amplitude manipulations. The coefficient addresses are coded in a comparable way as the amplitudes. Consequently, the resulting bit stream is a multiplex of coded amplitude and address information.

The proposed algorithm intrinsically adapts to the coefficient structure that appears after transformation. Furthermore, as coding of every block starts with the most significant coefficients, it is possible to give additional protection to this part of the coded information. In addition, synchronization patterns are periodically inserted for decoder resynchronization and limitation of possible erroneous video signal reconstruction.

After the complete algorithm had been simulated for still pictures and moving sequences, the presented codec was realized in real-time hardware. The video signals applied were standard TV signals with 50 Hz field rate and 10.125 MHz and 3.375 MHz sampling frequency for luminance and chrominance, respectively. The subjective image quality obtained was greatly dependent on the pictures used. Therefore, we tested our algorithm with highly detailed pictures. An example of such a picture ('Baltimore') is shown in Fig. 10.

4. RECORDING CHANNEL AND TRACKING

For channel modulation NRZ-I with an 8-to-10 code is used. This modulation method has been combined with integrating detection. Instead of adding the required tracking tones to the binary data directly as was done in [2], the tones have been embedded in the channel modulation code itself. To this end a new DC-free code has been designed. The embedding technique is also suitable for thin Me media, where the direct addition of the tones to the binary data cannot be used because it yields a too low signal-to-noise ratio of the (low-frequency) tones. The embedding can be achieved by imposing a periodical variation on the Digital Sum Value (DSV), which is the integral sum of the disparity of subsequent 10-bit codewords. Disparity is the difference between the number of ones and zeros of a 10-bit word and

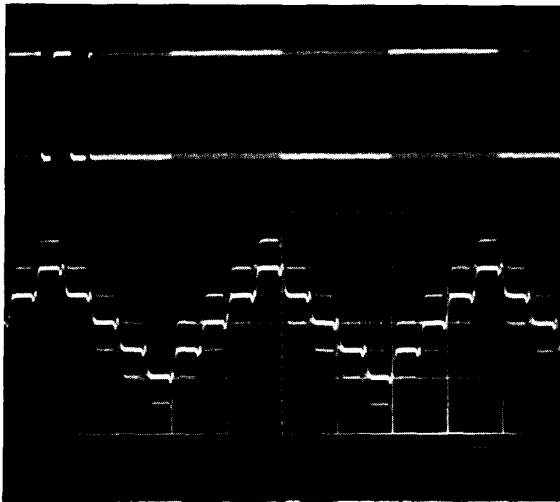


Figure 11. Embedding of tracking tones; upper part: encoded binary channel signal with square-wave tracking tone, lower part: triangular variation of the Digital Sum Value (DSV)

therefore equal to the DSV of that word. If we consider the signal-to-noise ratio of the tracking tone to be generated, the highest value is obtained if the disparity of the 10-bit code is modulated blockwise [16].

To explain the insertion of tracking signals in the encoded data signal, let us suppose that for each of the 256 input data words we can choose a pair of 10-bit output words which have equal disparity of opposite sign. A low-frequency tracking tone can be inserted by selecting words with a positive disparity

Table 2: Code tables used for the generation of embedded tracking tones

$T_{max} = 6$	
C_1	DSV
0 - 193	+2
194 - 255	+4
C_2	
0 - 179	-2
180 - 255	-4
C_3	
0 - 19	+2
20 - 255	0

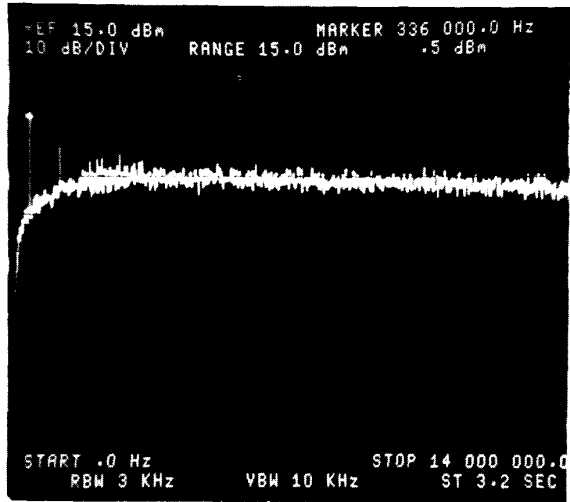


Figure 12. Frequency spectrum of the encoded binary channel signal; fundamental frequency and third harmonic of the embedded tracking tone are clearly visible

during a certain interval of the tone and a negative disparity during the next interval. If the desired wave

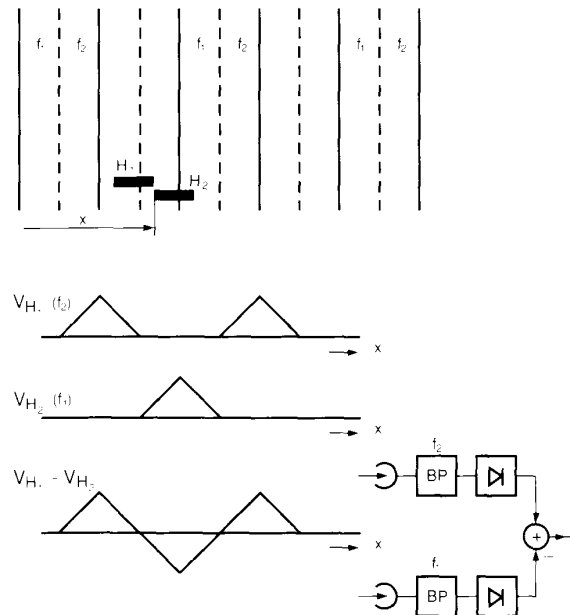


Figure 13. Allocation of tracking tones to subsequent pairs of recording tracks. Each pair containing f_1 and f_2 is followed by a pair without any tones, which is again followed by a pair with f_1 and f_2 , etc. The actuator control signal is derived from the two playback signals V_{H1} and V_{H2} after filtering at f_2 and f_1 , respectively

form of the tone is a square-wave, then the wave form of the DSV must be triangular.

With a maximal runlength $T_{\max} = 6$, the number of 10-bit word pairs with opposite disparity equal to 2 is insufficient for encoding all the 256 different input data words. Therefore, extra 10-bit words with disparities equal to +4, 0 and -4 have been included in the code tables (table 2). The code tables have been composed to give the best possible fit (yielding the least possible quantization noise) of the DSV wave form to the desired triangular shape. Figure 11 shows the generated code with the corresponding wave form of the DSV. A frequency spectrum of the former signal is shown in fig. 12.

For the actual track following two tones are used at $f_1 = f_{bit}/60$ and $f_2 = f_{bit}/80$, respectively. The allocation of these tones to the recording tracks is depicted in fig. 13. Each pair of tracks that contain the tones f_1 and f_2 is followed by a pair of tracks without tracking tones, which is again followed by a pair containing the tones, etc. By filtering the playback signals of heads H_1 and H_2 at f_2 and f_1 respectively, control signals for driving the piezo ceramic actuator can be derived.

In comparison to common double actuator systems, where each piezo carries a single head, an attractive advantage of the single actuator system is its insensitivity to mechanical hysteresis of the piezo. An active tracking system for the recording mode can therefore be omitted. A more obvious advantage is that the heads are always locked to a correct pair of tracks, whereas in usual systems there is an increased chance for false lock with increased track density. Fewer mechanical components and insensitivity of the track pattern to periodical errors of the scanner pivot bearing (no unwanted overwriting of neighbouring tracks) are other advantages of the single actuator system.

5. CONCLUSIONS

Digital video recording systems that are based on small cassettes can only obtain sufficient playing time if use is made of a bit-rate reduction technique. The investment to be made for the related video signal processing circuits is outweighed by the gain from simplification of scanner design and h.f. recording circuitry.

DCT forms a powerful basis for an efficient bit-rate reduction algorithm. At normal viewing distance (5-6 times the vertical display size) the following conclusions are drawn for highly detailed scenes. A reasonable picture quality has been obtained at 10 Mbit/sec. High picture quality, with

hardly noticeable coding noise, was obtained at 19 Mbit/sec. These pictures also showed a significant increase of the resolution compared to the pictures coded at 10 Mbit/sec. For scenes with critical motion, the use of motion-adaptive processing is required.

For recording with high track densities the single piezo actuator system with a double head has some clear advantages over the more conventional solution with double actuators. Problems due to hysteresis of the polarization have been eliminated as well as the sensitivity to periodic inaccuracies of the scanner pivot bearing. On the other hand, the flexibility of the actuator system for accurate following of subsequent tracks in multispeed playback mode has been maintained. Tracking tones with excellent signal-to-noise ratio can be embedded in the 8-to-10 channel modulation code.

REFERENCES

- [1] S. Mita et al., 'Digital video recording techniques using 1/2 inch metal particle tape', IEEE Trans. on Cons. Elec., Vol. CE-31, No. 3, pp. 386-397, August 1985.
- [2] L.M.H.E. Driessen et al., 'An experimental digital video recording system', IEEE Trans. on Cons Electr., Vol. CE-32, No.3, pp. 362-371, August 1986.
- [3] A. Hirota et al., 'Picture coding for home VTR', Proc. Pict. Coding Symp. Tokyo, pp. 136-137, April 1986.
- [4] C. Yamamitsu et al., 'An experimental digital VTR capable of 12 hour recording', IEEE Trans. on Cons. Elec., Vol. CE-33, No. 3, pp. 240-248, August 1987.
- [5] T. Kondo et al., 'Adaptive dynamic range coding scheme for future consumer digital VTR', IERE Proc. 7th Int. Conference on Video, Audio and Data Recording York (UK), pp. 219-226, March 1988.
- [6] N. Ahmed, T. Natarajan and K.R. Rao, 'Discrete cosine transform', IEEE Trans. Comp., Vol. C-23, pp. 90-93, January 1974.
- [7] P.H.N. de With and S.M.C. Borgers, 'On adaptive DCT coding techniques for digital video recording', IERE Proc. 7th Int. Conference on Video, Audio and Data Recording York (UK), pp. 199-204, March 1988.
- [8] W-H. Chen, C.H. Smith and S.C. Fralick, 'A fast computational algorithm for the discrete cosine transform', IEEE Trans. on Comm., Vol. COM-25, pp. 1004-1009, September 1977.
- [9] B.G. Lee, 'A new algorithm to compute the discrete cosine transform', IEEE Trans. Acoust.

- Speech Signal Proc., Vol. ASSP-32, pp. 1243-1245, December 1984.
- [10] M. Vetterli and H.J. Nussbaumer, 'Simple FFT and DCT algorithms with a reduced number of operations', Signal Processing, Vol. 6, No. 4, July 1984.
- [11] C.P. Visser, 'Discrete cosine transform: algorithms and hardware realization' (in Dutch), M. Thesis, Delft Univ. of Techn., January 1986.
- [12] A.G. Tescher, 'A dual coding algorithm', Proc. Nat. Telecomm. Conf., pp. 53.4.1-53.4.4, 1979.
- [13] W-H. Chen and W.K. Pratt, 'Scene adaptive coder', IEEE Trans. on Comm., Vol. COM-32, No. 3, pp. 225-232, March 1984.
- [14] A.K. Jain, 'Image data compression: a review', Proc. IEEE, Vol. 69, No. 3, pp. 349-389, March 1981.
- [15] W. Keesen, 'A new bit-assignment for transform coding', Proc. Pict. Cod. Symp. Tokyo, pp. 148-149, April 1986.
- [16] K.A.S. Immink, 'Signal-to-noise ratio of pilot tracking tones embedded in binary coded signals', IEEE Trans. Magn., vol. MAG-24, March 1988.

BIOGRAPHIES

Stephan M. C. Borgers was born in Rotterdam, The Netherlands in 1950. He graduated in electrical engineering from the Delft University of Technology in 1976. He joined Philips Consumer Electronics in 1979 where he was involved in digital signal processing for TV and video recording. At the moment he is involved in data reduction for digital magnetic consumer recording.



Werner A. L. Heijnemans was born in Breda, The Netherlands in 1947. He graduated in electrical engineering from the Eindhoven University of Technology in 1970. He joined Philips Research Laboratories Eindhoven in 1972 and was involved in magnetic deflection systems for colour CRTs. Later he was involved in other research topics related to TV display systems. At present he is head of the Magnetic Recording Department at the same laboratories.



Edmond de Niet was born in The Hague, The Netherlands in 1927. He joined Philips Research Laboratories in 1950 and was active in the audio field for a number of years. In later years he became a staff member of the same Laboratories and was involved in system aspects of magnetic bubble devices and magnetic recording for both audio and video applications.



Peter H.N. de With was born in Lexmond, The Netherlands in 1958. He graduated in electrical engineering from the Eindhoven University of Technology. He joined Philips Research Laboratories Eindhoven in 1984 where he became a member of the Magnetic Recording Department. He is working on digital data reduction techniques.

