An Improved Method of Embedding Data into Pictures by Modulo Masking

KAZUHIKO HARA, TADASHI SHIMOMURA, TAKAAKI HASEGAWA, MEMBER, IEEE, AND MASAO NAKAGAWA, MEMBER, IEEE

Abstract—**An improved scheme with vertical block allocation of embedding data into industrial quality monochrome analog pictures by modulo masking is investigated. The video signal on each scan line is sampled, and a data bit is inserted into a block of three pels by an improved modulo masking scrambling technique of the luminance level of only one pel in the block. The performance of the system proposed here is compared to that of a conventional system. The number of data bits embedded in an image for the proposed system is about 1.3 times as large as that of the conventional system. In addition the SNR of the recovered image in the new system is increased by about 3-4 dB.**

I. INTRODUCTION

ANALOG picture or speech signals are highly correlated. For example the luminance levels of a picture cell (pel) and adjacent pels are highly correlated with each other. This characteristic can be exploited to enable the analog signal to be an unwitting data carrier. The process of inserting data into the picture or speech signal causes the transmitted signals to be different from the original signals. This signal modification, far from being undesirable, may be welcomed, making it fatiguing or difficult for an eavesdropper to comprehend the information being transmitted. The analog signals can be recovered at the receiver with an acceptably small perceptual degradation, and the data can be regenerated with a bit error rate (BER) below a specified level.

Some systems which transmit analog signals, such as speech, facsimile, and television, with multiplexed digital data, have been proposed [1]-[13]. Steele and Vitello embedded data into speech by using scrambling techniques [1], [2]. Wong, Steele, and Xydeas have inserted data onto the phase of speech signals [4]. Feher *et al.* have described methods for incorporating data with microwave analog signals. These are known as data-above-voice (DAV), data-under-voice (DUV) [6], and data-above-video (DAVID) [7]. Nakagawa and Saito proposed a simple hybrid multiplexing method for analog and digital signals using amplitude multiplexing with a signal separator at the receiver [11], [12]. Recently, starting with the concepts of Steele and Vitello [1], [2], Xydeas, Kostic, and Steele proposed methods of embedding data into monochrome pictures by modulo masking [3].

In the system of [3], contiguous blocks of N pels along the scan lines (one-dimensional processing) are used to support data. One bit may be inserted into a block by scrambling the luminance level of only one pel, rather than N pels. The method of scrambling is modulo masking where the luminance level of the pel to be scrambled is modulo added to a fixed number.

In the conventional system, the block consists of contiguous pels along the scan line (horizontal line). In this paper, we turn our attention to the point that not only horizontally, but also vertically, contiguous pels are highly correlated and the point that the transmitted pel's direction (horizontal direction) needs not agree with the block's direction. We propose a new system [13] using vertical block allocation with better performance, i.e., the amount of transmitted data for the new system becomes about 1.3 times as large as for the conventional system, while the SNR of the received picture is superior, about 3–4 dB better than that of the conventional system.

II. CONVENTIONAL SYSTEM [3]

In this section, we describe the data embedding and extraction algorithm of the conventional scheme [3] so as to facilitate the understanding of the new algorithm. The transmitter and receiver algorithms of the conventional system are shown in Fig. 1.

The video signal, band limited to f_c Hz, is sampled at a rate $f_s > 2f_c$, and uniformly quantized by an 8-bit quantizer to yield a sequence of 256-level (0-255) luminance samples. When a picture consists of n scan lines and the luminance level is sampled m times on each scan line, the picture is divided into $m \times n$ picture elements (pels). Data are then embedded into this sequence, subject to criteria to be described. The combined picture and data multilevel samples are then converted into a continuous analog signal by an interpolating filter which also band limits this signal to f_c.

Let x_{ij} be the luminance level of the jth pel on the ith scan line. Three horizontally contiguous pels are taken as a block.

In the first block, having luminance levels x_{11}, x_{12}, and x_{13}, the parameter A_1 is calculated using the following expression:

$$A_1 = \frac{(x_{11}-x_{12})^2+(x_{12}-x_{13})^2}{2}. \tag{1}$$

A_1 may be considered as a measure of local picture activity, and if

$$A_1 < T_{a,1} \tag{2}$$

where $T_{a,1}$ is a system threshold, it is considered that the block has a sufficiently low activity for it to support one bit of data. Prior to data insertion, the value of x_{12} is modified to

$$\dot{x}_{12} = \frac{x_{11}+x_{12}+x_{13}}{3} \tag{3}$$

and then data are inserted into the block of three pels by a scrambling technique.

Paper approved by the Editor for CATV of the IEEE Communications Society. Manuscript received September 20, 1985; revised July 3, 1987. This paper was presented at GLOBECOM '85, New Orleans, LA, December 1985.

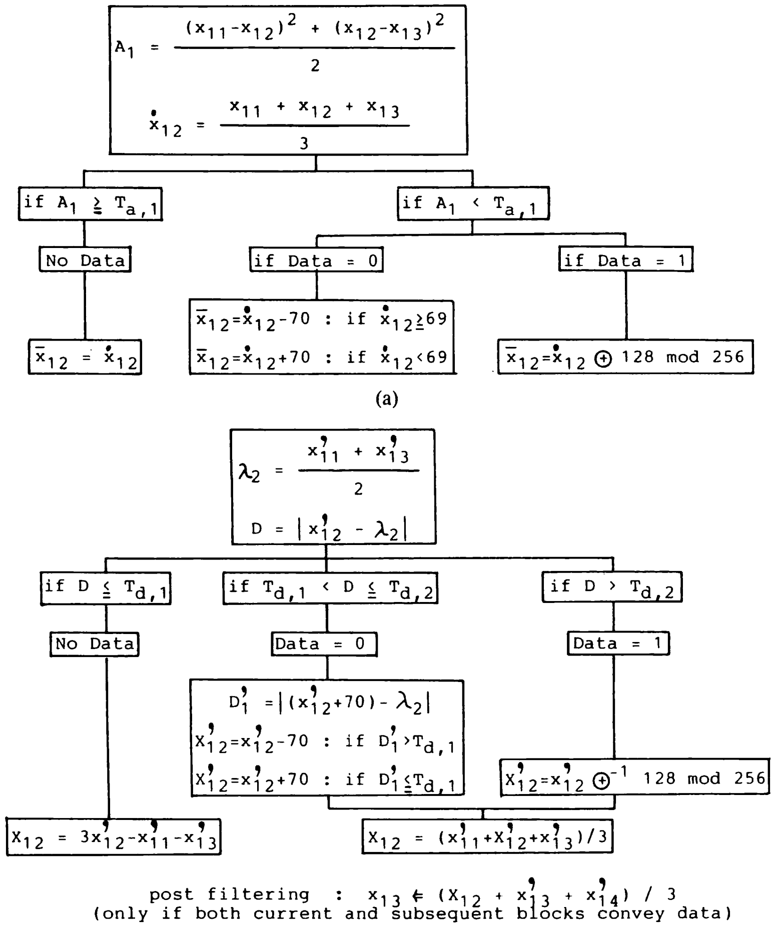
K. Hara was with the Department of Electrical Engineering, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama-shi, 223 Japan. He is now with Canon, Inc., 53, Imaikami-cho, Nakahara-ku, Kawasaki-shi, 211 Japan.

T. Shimomura and M. Nakagawa are with the Department of Electrical Engineering, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama-shi, 223 Japan.

T. Hasegawa was with the Department of Electrical Engineering, Keio University, 3-14-1, Kiyoshi, Kohoku-ku, Yokohama-shi, 223 Japan. He is now with Saitama University, 255, Shimoohkubo, Urawa-shi, Saitama-ken, 338 Japan.

IEEE Log Number 8719003.

0090-6778/88/0300-0315$01.00



(b)

Fig. 1. Conventional system. (a) The transmitter. (b) The receiver.

The representation of modulo addition is

$$u = v \oplus w \bmod z \tag{4}$$

which means v is added to w, modulo z, to give a number u, and the reverse operation is defined by the following equation:

$$v = u \oplus^{-1} w \bmod z. \tag{5}$$

The data are embedded into the block of pels by scrambling the luminance level $\dot{x}_{12}$ according to

$$\bar{x}_{12} = \begin{cases} \dot{x}_{12} \oplus 128 \bmod 256; & \text{data logical 1} \quad (6\text{-}1) \\ \dot{x}_{12} - 70, \text{ if } \dot{x}_{12} > 69; & \text{data logical 0} \quad (6\text{-}2) \\ \dot{x}_{12} + 70, \text{ if } \dot{x}_{12} \leqq 69; & \text{data logical 0.} \quad (6\text{-}3) \end{cases}$$

When the activity of luminance levels in the block is considered too high, i.e.,

$$A_1 \geqq T_{a,1}, \tag{7}$$

no data are transmitted. However, to assist the receiver in its task of deciding if the block of pels contains data, the activity A_1 of this block is lowered by means of altering the center pel to $\dot{x}_{12}$ using (3).

After these operations, the output sequence S_T is

$$S_T = \begin{cases} x_{11}, \bar{x}_{12}, x_{13} & \text{(data exists)} \quad (8) \\ x_{11}, \dot{x}_{12}, x_{13} & \text{(no data).} \quad (9) \end{cases}$$

Having processed the block of the first three pels on the first scan line, the same operation is done on the following blocks of three pels on the same line. The second scan line in the picture is next processed in an identical way to that of the first scan line, and the same operation is done in the entire frame. The multilevel samples are then converted into a continuous analog signal, band-limited to f_c, by an interpolating filter.

At the receiver, the continuous combined picture and data signal is sampled at f_c Hz to yield the received sequence. From this sequence,

$$S_R = x'_{11}, x'_{12}, x'_{13} \tag{10}$$

where a prime above the symbol implies its presence at the receiver

$$\lambda_2 = \frac{x'_{11} + x'_{13}}{2} \tag{11}$$

is formed. A distance measure

$$D = |x'_{12} - \lambda_2| \tag{12}$$

is made and compared to two distance thresholds $T_{d,1}$ and $T_{d,2}$.

If

$$D \leqq T_{d,1}, \tag{13}$$

it is concluded that no data are present. From (3), the regenerated intensity of the center pel is given as

X_{12}[1] $= 3x'_{12} - x'_{11} - x'_{13}$. (14)

---

[1] X_{12} (capital letter) is not same as x_{12} (small letter).

If

$$D > T_{d,2} \tag{15}$$

where $T_{d,2} > T_{d,1}$, it is concluded that a logical 1 is embedded as defined in (6-1). From (5),

$$x'_{12} = x'_{12} \oplus^{-1} 128 \bmod 256 \tag{16}$$

and the value of x'_{12} is further smoothed according to

$$x_{12} = \frac{x'_{11} + X'_{12} + x'_{13}}{3}. \tag{17}$$

X_{12} is regarded as the regenerated value.

If

$$T_{d,1} < D \leqq T_{d,2}, \tag{18}$$

it is considered that a logical 0 is conveyed. x'_{12} is the received version of $\bar{x}_{12}$ determined by (6-2) or (6-3). To determine which equation has been used, the following procedure is performed:

$$X'_{12} = x'_{12} + 70 \tag{19}$$

and a new distance measure is formulated.

$$D'_1 = |X'_{12} - \lambda_2|. \tag{20}$$

If this X'_{12} is correct, D'_1 will be small; and (21) should be satisfied.

$$D'_1 \leqq T_{d,1}. \tag{21}$$

If (21) is not satisfied, it is known that the original addition of 70 to x'_{12} was wrong. So (22) is used.

$$X'_{12} = x'_{12} - 70. \tag{22}$$

The luminance level X'_{12} is then smoothed according to (17) to yield X_{12}.

Finally, the regenerated sequence becomes

$$L_1 = x'_{11}, \qquad X_{12}, x'_{13}. \tag{23}$$

The sequence of output pels for the first and second blocks is given by (23) and

$$L_2 = x'_{14}, \qquad X_{15}, x'_{16}, \tag{24}$$

respectively, where X_{15} is the modified value of the received pel x'_{15}. It was found experimentally that the effect of channel dispersion on a block of pels, whose center pel had been subjected to modulo masking scrambling as means of insertion, was to cause distortion in neighboring pels. The greatest distortion was inflicted on the last pel in the block, as it immediately followed the scrambled pel whose luminance level had been significantly altered by the modulo addition. Thus, for the first block of data, x'_{13} contains an error far greater than that in the next pel x'_{14}. The luminance errors in the picture are, therefore, reduced replacing x'_{13} with X_{13} as below:

$$X_{13} = \frac{X_{12} + x'_{13} + x'_{14}}{3}. \tag{25}$$

This postfiltering procedure is only applied if the subsequent block is also used to convey data.

III. NEW SYSTEM

In this section, we refer to problems of the conventional scheme, and propose a new algorithm.

A. Problems of the Conventional System

Problem 1): Distortion of X_{12} for ideal channel caused by an incomplete reverse procedure.

Let us consider the case that data of logical value 1 is embedded into the block (x_{11}, x_{12}, x_{13}). And if $x_{12} < 128$, the output sequence of the transmitter is

$$S_T = x_{11}, \qquad (x_{11} + x_{12} + x_{13})/3 + 128,\ x_{13}. \tag{26}$$

If the channel is ideal, the received sequence is the same as that transmitted, i.e., S_T. Now if it is correctly concluded from the value of the distance measure that data of logical value 1 have been embedded, the regenerated luminance level X_{12} is given from (16) and (17)

$$X_{12} = \frac{x_{11} + \dfrac{x_{11} + x_{12} + x_{13}}{3} + x_{13}}{3} = \frac{4x_{11} + x_{12} + 4x_{13}}{9}. \tag{27}$$

From (27), it is clear that the original luminance value of x_{12} is almost lost for large coefficients of x_{11} and x_{13}, although X_{12} (regenerated) is desired to be equal to x_{12} (original). Only when A_1, which indicates the local activity of the luminance levels x_{11}, x_{12}, and x_{13}, satisfies inequality (2) ($A_1 < T_{a,1}$), is data embedded. So, if the value of $T_{a,1}$ is very small, the following expression would be satisfied in many blocks:

$$x_{11} \fallingdotseq x_{12} \fallingdotseq x_{13} \tag{28}$$

and (27) becomes

$$X_{12} \fallingdotseq x_{12}. \tag{29}$$

Even when inequality (2) is satisfied, however, if the values of x_{11} and x_{13} are almost the same and only x_{12} is different, (28) is not satisfied. Then X_{12} has large distortion. Meanwhile, lowering the value of $T_{a,1}$ decreases the number of blocks which could convey data. If the value of $T_{a,1}$ is increased to increase the number of transmission data bits, (28) is not satisfied, and the distortion of X_{12} becomes large.

In short, distortion of X_{12} is caused by an incomplete reverse operation.

Problem 2): Distortion of x_{13} in dispersive channels.

Let us consider the case where data of logical value 1 is embedded into a block, as in the previous paragraph, and that the channel is dispersive. Fig. 2 shows examples of the transmitted and received signals on the assumption that x_{11}, x_{12}, and x_{13} are all equal to 100. E_1, E_2, and E_3 are distortions (errors) of x_{11}, $\bar{x}_{12}$ (($x_{11} + x_{12} + x_{13})/3 + 128$), and x_{13}, respectively. The regenerated luminance levels are calculated as

$$x'_{11} = x_{11} + E_1 \tag{30-1}$$

$$X_{12} = \frac{4x_{11} + x_{12} + 4x_{13}}{9} + \frac{E_1 + E_2 + E_3}{3} \tag{30-2}$$

$$x'_{13} = x_{13} + E_3. \tag{30-3}$$

It was found experimentally all the time that E_1 was far smaller than E_2 or E_3, and that E_2 and E_3 had opposite signs from each other. Therefore, it is known from (30-1) that the distortion of x'_{11} is small.

In the case that $x_{11} \fallingdotseq x_{12} \fallingdotseq x_{13}$, the distortion of X_{12} is found out to be

$$\frac{E_1 + E_2 + E_3}{3}$$

from (30-2). As E_2 and E_3 have opposite signs, they cancel each other out and the remainder is divided by three. Then the distortion of X_{12} indeed becomes small.

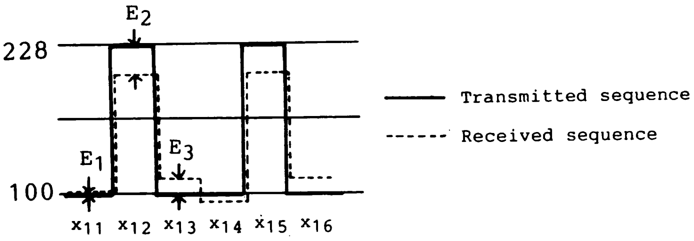


Fig. 2. Example of transmitted and received sequences (dispersive channel).

However, x'_{13} has distortion E_3, which may be relatively large. Although the distortion of x'_{13} is decreased by post-filtering which is applied only when the subsequent block is also used to convey data, x'_{13} is rather distorted because of being replaced with the mean value of three pels if the block has some degree of activity of luminance levels.

B. New System

To reduce the distortions which we have discussed in the previous section, and simultaneously to increase the transmission data rate, we have modified some points of the algorithm of the conventional system. The modified algorithm is shown in Fig. 5.

a) Arrangement of Blocks: In the conventional system, each block consists of three horizontal pels; in the new system, we arrange the blocks vertically as shown in Fig. 3. As a matter of course, even if any modification of blocks were done, the order of the transmitted pels is not modified. Because of this arrangement, the distortion is caused by a dispersive channel when data are embedded, exists only on the central line ($x_{21}, x_{22}, x_{23}, x_{24}, \cdots$). The upper line ($x_{11}, x_{12}, x_{13}, \cdots$) and the under line ($x_{31}, x_{32}, x_{33}, \cdots$) are only distorted a little. Furthermore, the use of vertical blocks makes it possible to overlap the blocks as shown in Fig. 3. For example, block-*A* (x_{11}, x_{21}, x_{31}) and block-*B* (x_{31}, x_{41}, x_{51}) have the common pel, x_{31}. Consequently, each block covers two pels in the new system, while each block covered three pels in the conventional system. The number of blocks that each picture requires in the new system is about 1.5 times as large as that in the conventional system.

The new system takes advantage of characteristic that distortion spreads horizontally, not vertically.

b) Regeneration by the Complete Reverse Procedure: Fig. 4 shows the transmitted and received signal waveforms of a data embedded line, which consists of only the central pels of blocks ($x_{21}, x_{22}, x_{23}, \cdots$), in the case of a dispersive channel (for $x_{21} = x_{22} = x_{23} = \cdots = 100$).

Equations (31) and (32) for the vertical block arrangement correspond to (3) and (17), respectively, for the conventional system.

$$\dot{x}_{21} = \frac{x_{11} + x_{21} + x_{31}}{3} \quad \text{(transmitter)} \tag{31}$$

$$X_{21} = \frac{x'_{11} + X'_{21} + x'_{31}}{3} \quad \text{(receiver).} \tag{32}$$

As a solution to Problem 1), a complete reverse regeneration procedure should be done; i.e., not equation (32) (incomplete reverse) but (33) below (complete reverse) should be used at the receiver:

$$X_{21} = 3X'_{21} - x'_{11} - x'_{31}. \tag{33}$$

However, from Fig. 4, the pels on this line may have large distortions if (33) is used. For example, X_{21} has rather large distortion on account of the term $3X'_{21}$. So, at the transmitter, when $A_1 \geqq T_{a,1}$, the signal is smoothed using (31). If $A_1 < T_{a,1}$, the signal is not smoothed and data are embedded. At the receiver, data are extracted by the same procedure as the conventional system and the value of x_{21} is regenerated by reverse operations.

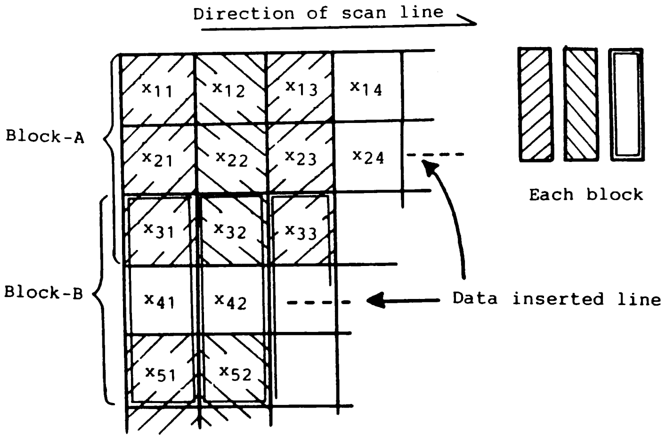


Fig. 3. Block allocation for the new system.

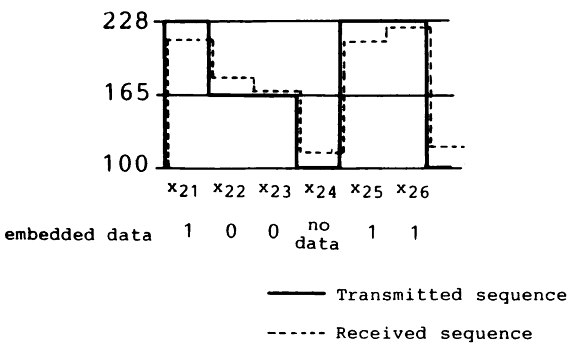


Fig. 4. Example of the transmitted and received sequences on data inserted line in the new system (dispersive channel).

By this complete reverse regeneration, the distortion of the regenerated center pel of the data containing block is not trebled. Although averaging of data embedding blocks is omitted, as such blocks have sufficiently low activity, the receiver can correctly regenerate the data.

c) Reducing Distortion: It is clear from Fig. 4 that in the case when the data of a block are different from that of the previous block, the distortions of the central pels of the blocks become large. When a previous block has embedded data and the central pel of a no-data block is regenerated by a reverse operation, i.e.,

$$X_{21} = 3x'_{21} - x'_{11} - x'_{31}, \tag{34}$$

x'_{21} has large distortion. Furthermore, X_{21} has a distortion about three times that. So, the following procedure is done at the transmitter. Although $A_1 < T_{a,1}$, if $A'_1 \geqq T_{a,1}$ in the following block, data are not embedded and the only operation performed is

$$\dot{x}_{21} = x_{21}. \tag{35}$$

In the receiver, if the block does not have embedded data and the previous block has data, X_{21} is regenerated using

$$X_{21} = x'_{21} \tag{36}$$

instead of (34).

In most cases, this procedure makes the complete reverse regeneration possible. By this procedure, an incomplete reverse regeneration occurs only if a block which has small activity ($A_1 < T_{a,1}$) and a block which has large activity ($A_1 \geqq T_{a,1}$) are transmitted alternately. In this case, the block which must be regenerated by (36) would be wrongly

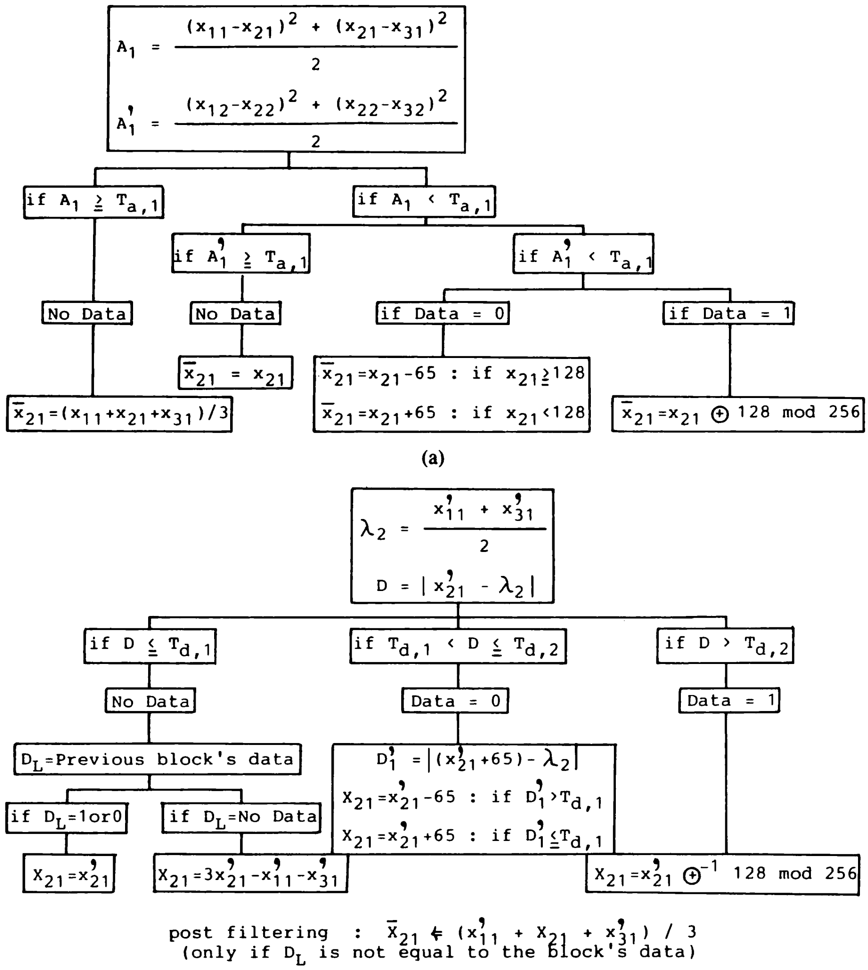


Fig. 5. New system. (a) The transmitter. (b) The receiver.

regenerated by (34). But the probability that such blocks are transmitted alternately is small; and, even if such a wrong regeneration is done, the distortion would not become large as such blocks' activities are small.

In the new system, the value added to or subtracted from x_{21}, when data of logical value 0 are embedded, is changed to 65 from 70 of the conventional system [(6-2) and (6-3)]. The determination of addition or subtraction depends on whether x_{21} is smaller than 128 or not.

d) Postfiltering: Distortions become large when the data of one block are different from that of the previous block. In such blocks, we calculate $\bar{X}_{21}$ as

$$\bar{X}_{21}=\frac{x'_{11}+X_{21}+x'_{31}}{3} \tag{37}$$

for the sake of reducing distortions. Unless a data error happens, the smoothing operation does not often make the distortion large because $A_1 < T_{a,1}$ is always satisfied in the block which conveys different data from the previous block at the transmitter.

IV. EXPERIMENTAL PROCEDURE

The experiments were performed by means of digital computer simulation on a monochrome picture, band-limited to 2.5 MHz, sampled at 5.5 MHz, and composed of 256 lines with 256 pels/line in the same way as Xydeas' procedure [3]. Then a picture consists of 65536 (= 256 × 256) pels. We used the well-known "girl" and "couple" images in our experiments.

Because picture signals tend to be conveyed over coaxial quality channels, we opted for a mildly dispersive channel in the form of a second-order low-pass Butterworth filter having a cutoff frequency of 2.5 MHz plus additive Gaussian white noise. We did experiments with a channel SNR (SNR_C), which was the ratio of the power of the transmitted signal to that of the Gaussian white noise of 40 dB as Xydeas did. In the simulation, the sampling instants at the receiver had an exact relative correspondence to those at the transmitter.

The estimation of the picture quality was performed by signal-to-noise ratio (SNR_L). SNR_L was defined as the ratio of the power of the original picture signal to that of the noise component of the regenerated picture signal (= regenerated signal–original signal).

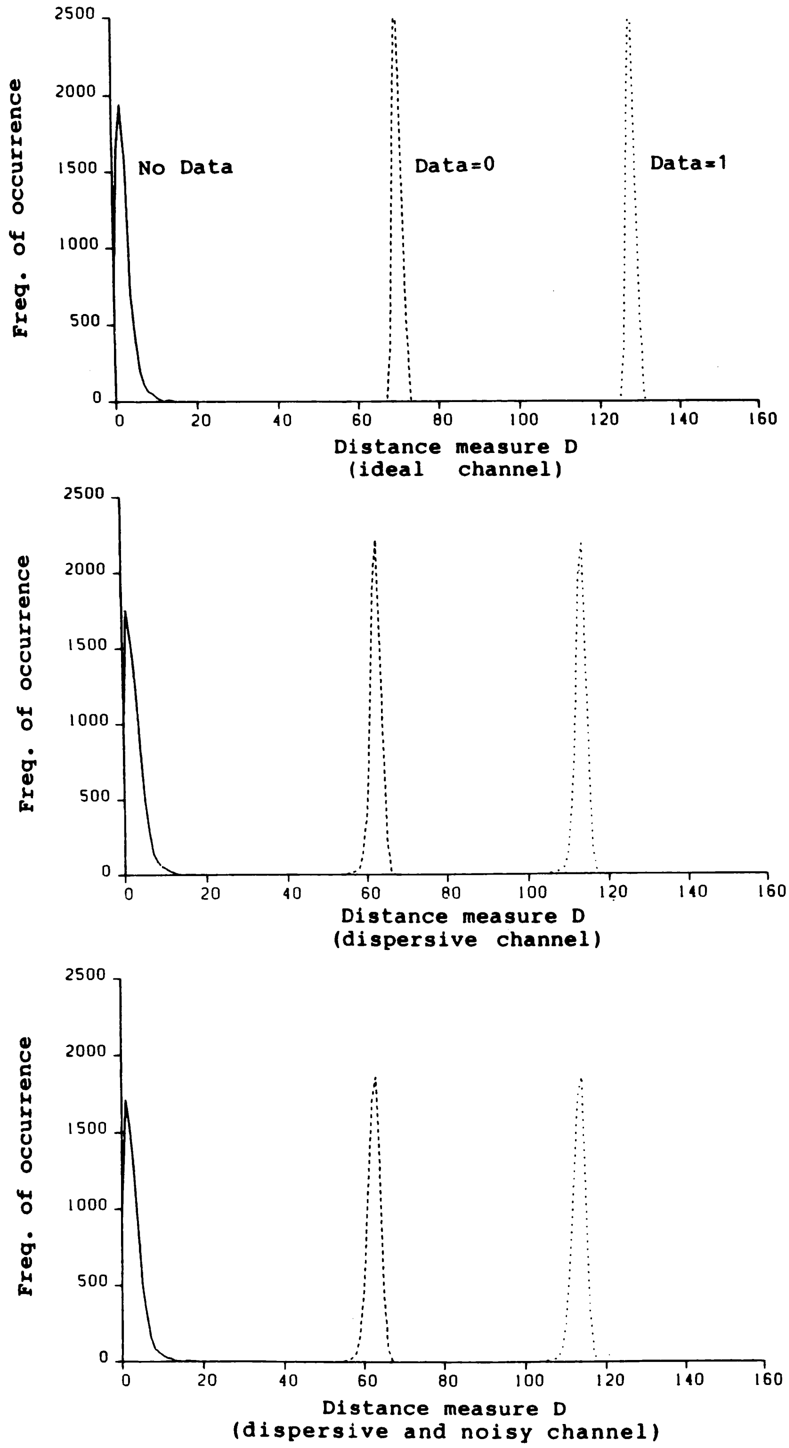
To ascertain the effect of modifying the embedding algorithm from that shown in Fig. 1 (conventional algorithm) to that shown in Fig. 5 (almost complete reverse algorithm), we simulated a system called system 1′. The arrangement of blocks was the same as that of the new system (Fig. 3) and the transmitter and receiver algorithms were the same as those of the conventional system (Fig. 1), except for the differences of subscripts where the postfiltering was same as that of the new system.

V. RESULTS

A. Selection of $T_{d,1}$, $T_{d,2}$

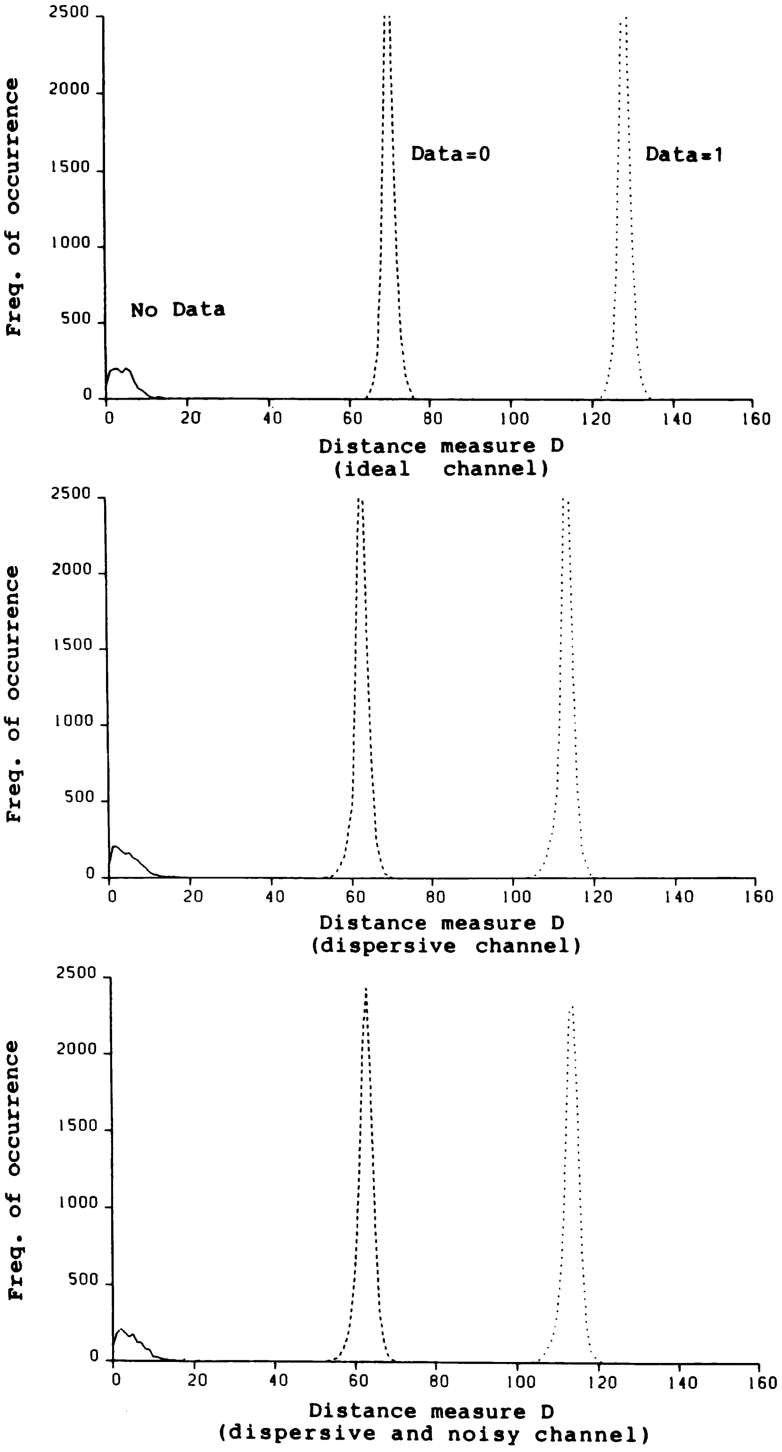
Each system has three system parameters. An activity threshold $T_{a,1}$ for the transmitter, and two distance thresholds $T_{d,1}$ and $T_{d,2}$ for the receiver. First, we determine $T_{d,1}$ and $T_{d,2}$.

Fig. 6 shows the histograms for the distance measure for the case of the ideal channel, the dispersive channel and the dispersive channel with its additive Gaussian white noise for the "girl" image. In Fig. 6, it can be shown that the histograms in the new system shifted to lower a value of D and



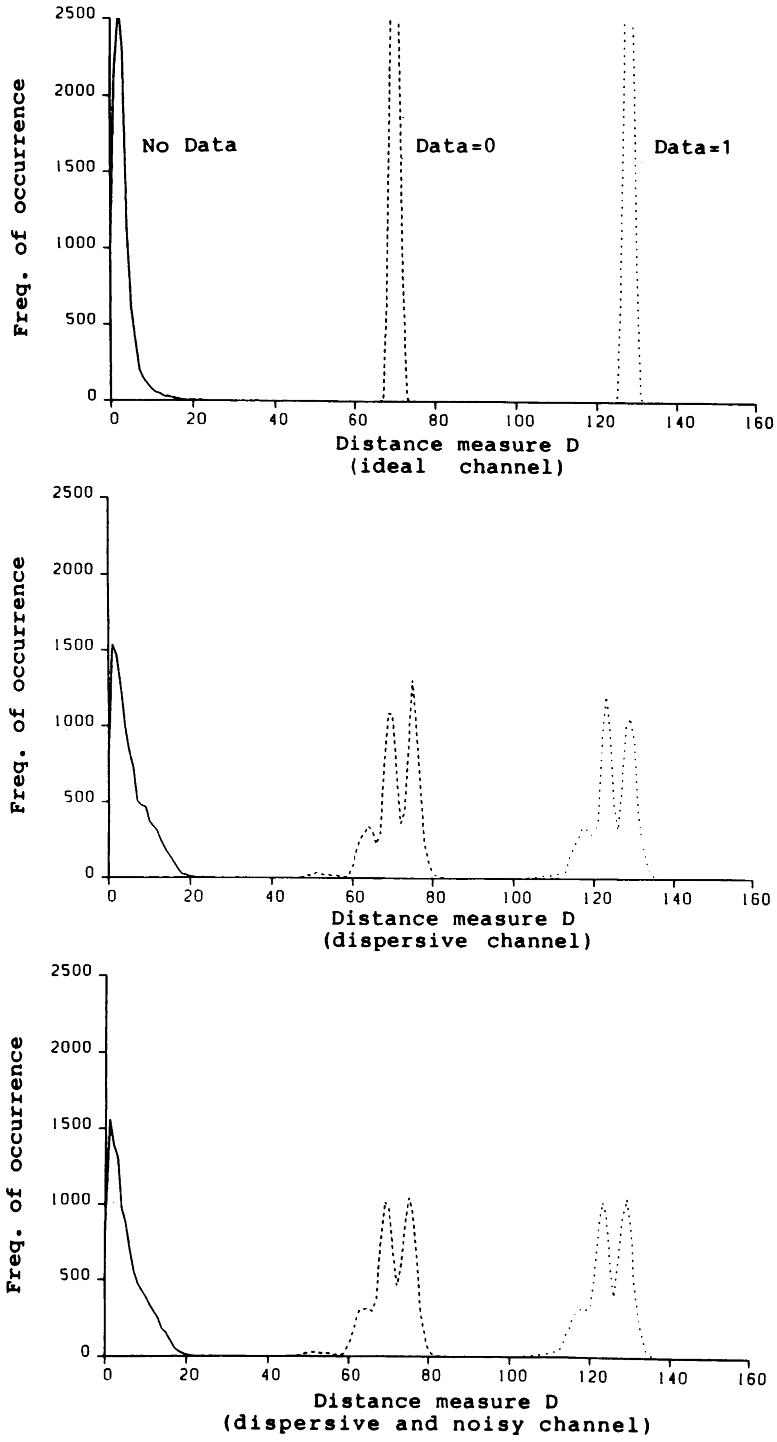
(a)

Fig. 6. (a) Histograms of distance measure for $T_{a,1} = 50$ (conventional system). (b) Histograms of distance measure for $T_{a,1} = 300$ (conventional system). (c) Histograms of distance measure for $T_{a,1} = 50$ (system 1′). (d) Histograms of distance measure for $T_{a,1} = 300$ (system 1′). (e) Histograms of distance measure for $T_{a,1} = 50$ (new system). (f) Histograms of distance measure for $T_{a,1} = 300$ (new system).



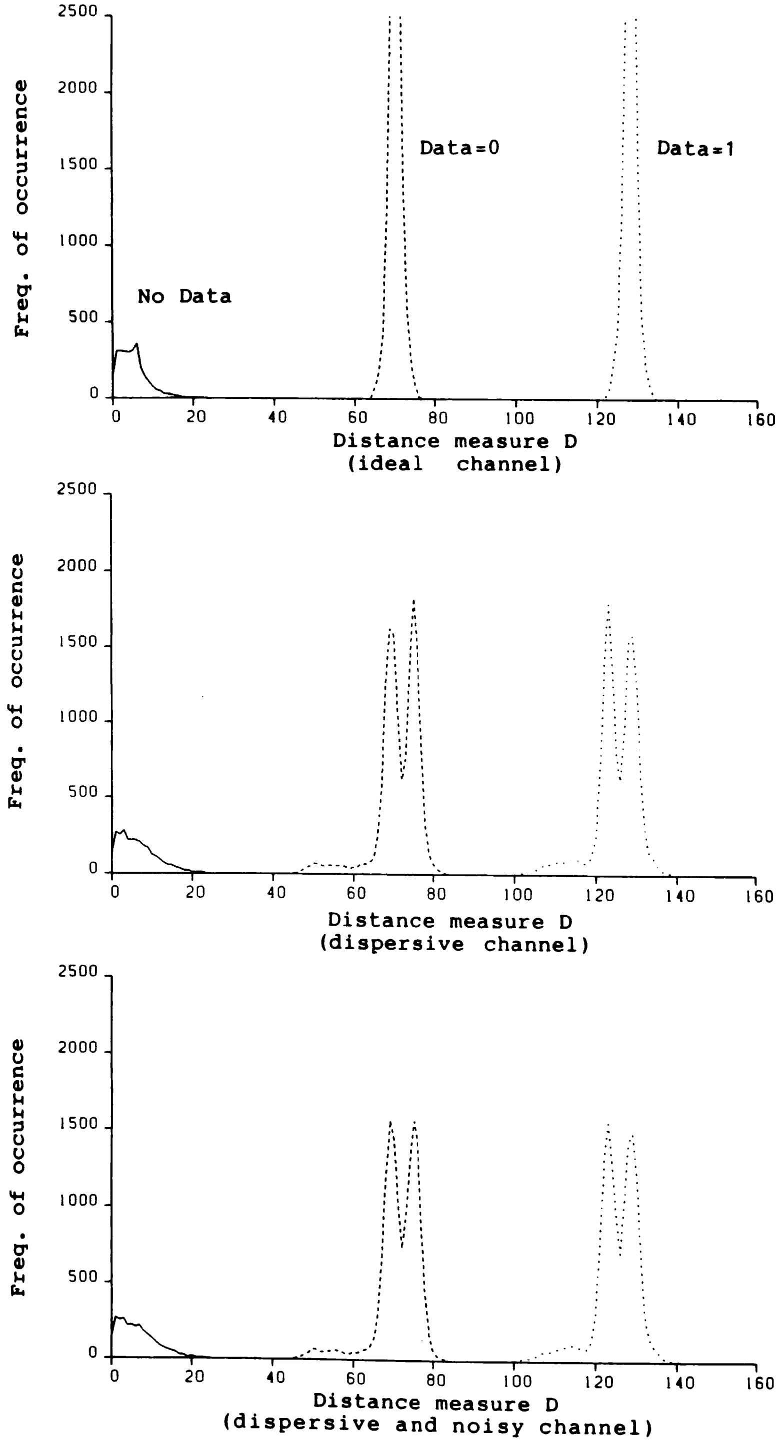
(b)

Fig. 6. (*Continued*).



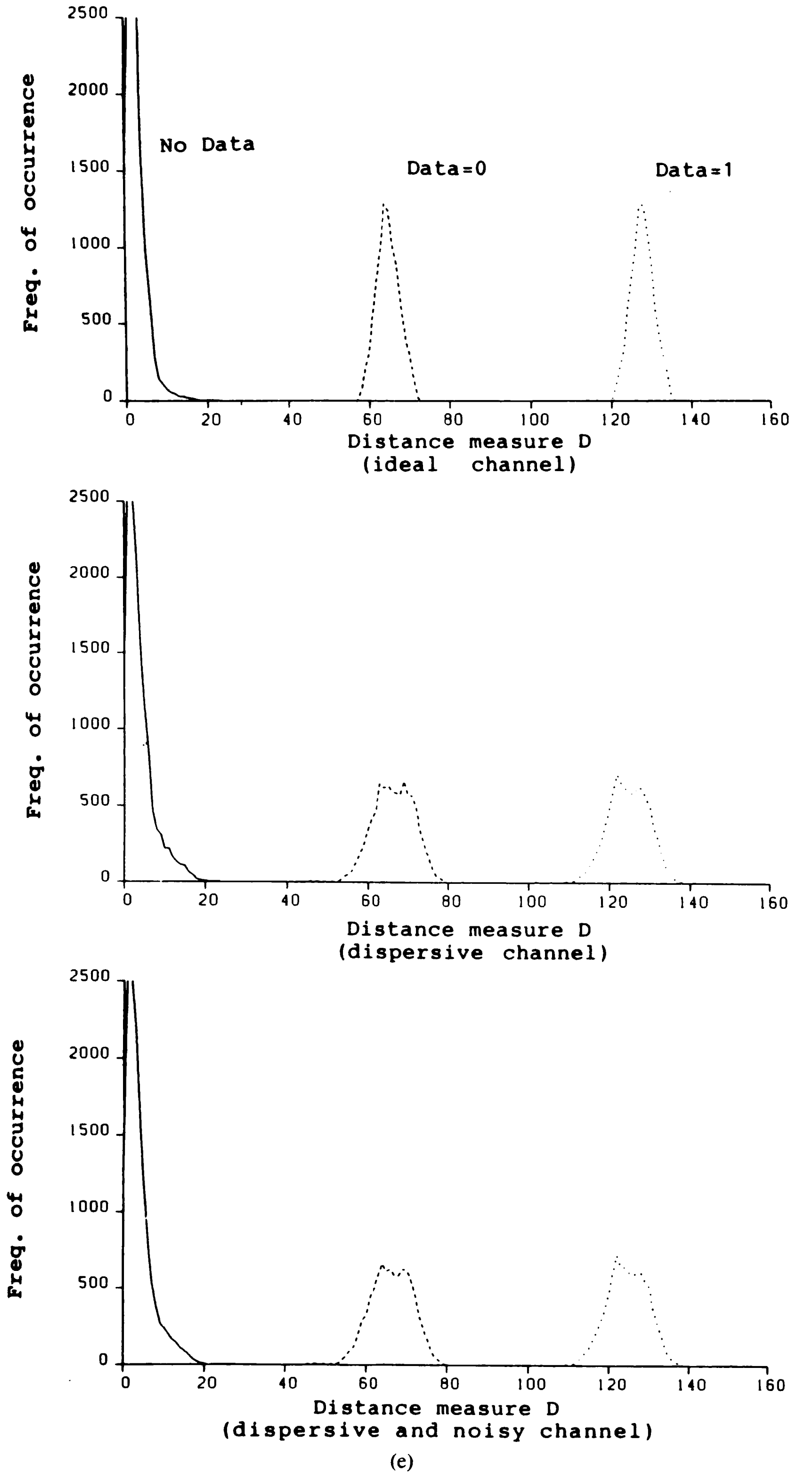
(c)

Fig. 6. (*Continued*).



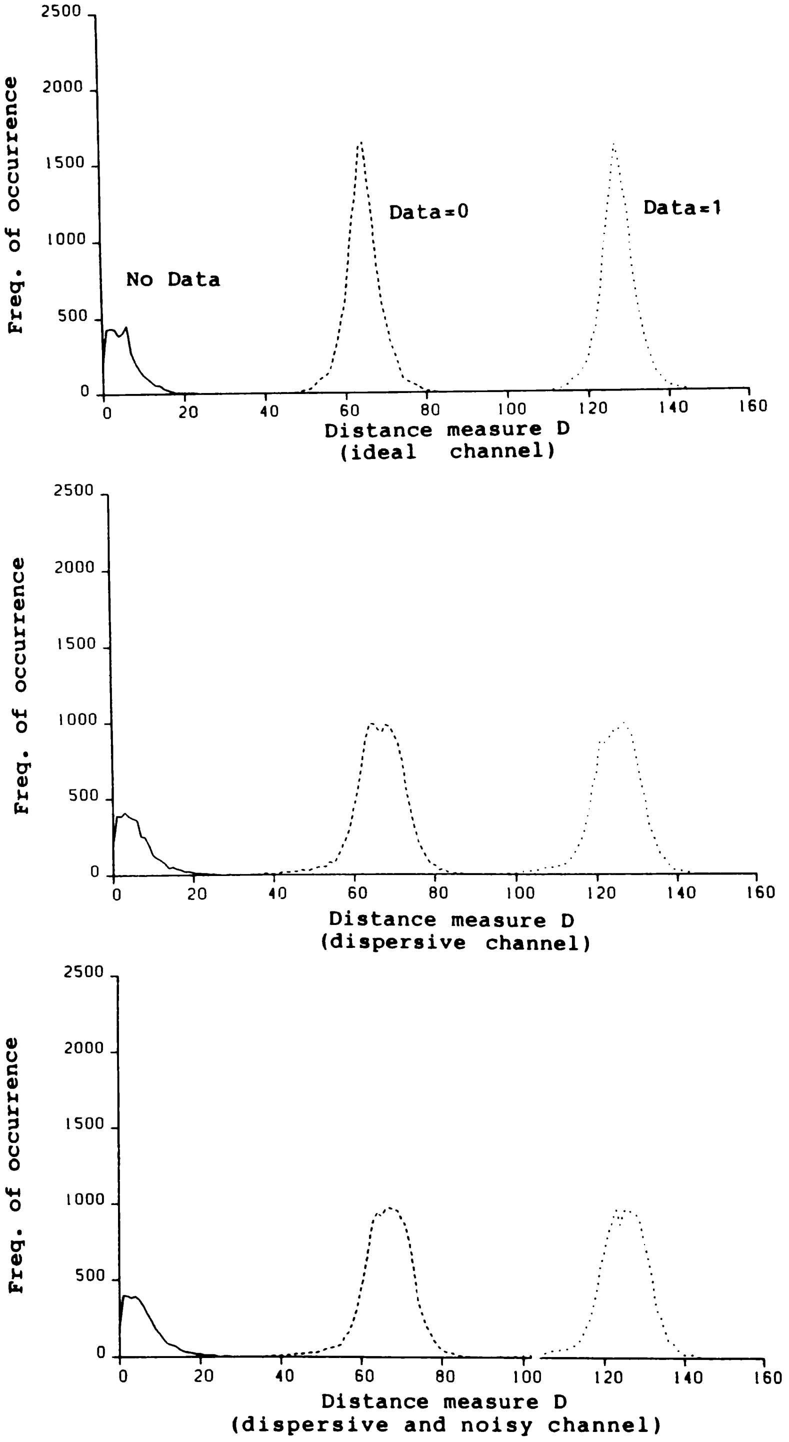
(d)

Fig. 6. (*Continued*).



(e)

Fig. 6. (*Continued*).



(f)

Fig. 6. (*Continued*).

TABLE I
SYSTEM PARAMETERS

	$T_{a,1}$	$T_{d,1}$	$T_{d,2}$
Conventional system	50	40	90
System 1'	50	40	90
New System	50	30	90

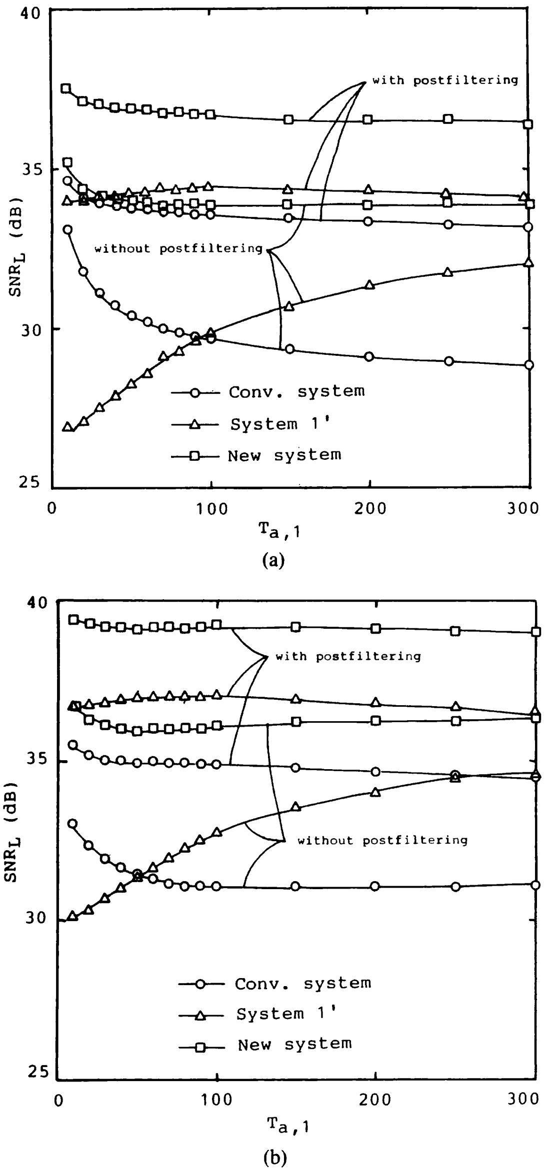


Fig. 7. (a) SNR_L as a function of $T_{a,1}$ (girl image). (b) SNR_L as a function of $T_{a,1}$ (couple image).

spread by the dispersive channel. When the channel had additive Gaussian white noise the histograms spread more widely.

When the histograms (no data, data logical 0, and data logical 1) do not overlap, data errors could not occur by selecting adequate values of $T_{d,1}$ and $T_{d,2}$. The values we determined for $T_{d,1}$, $T_{d,2}$ are given in Table I.

B. Selection of $T_{a,1}$

The larger the value of $T_{a,1}$, the more data can be conveyed. However, increasing the value of $T_{a,1}$ tends to make the histograms widely spread and overlap (see Fig. 6), and data errors occur.

The effect of $T_{a,1}$ on the SNR_L and the number of embedded data per picture are shown in Fig. 7 for the "girl" image and in Fig. 8 for the "couple" image.

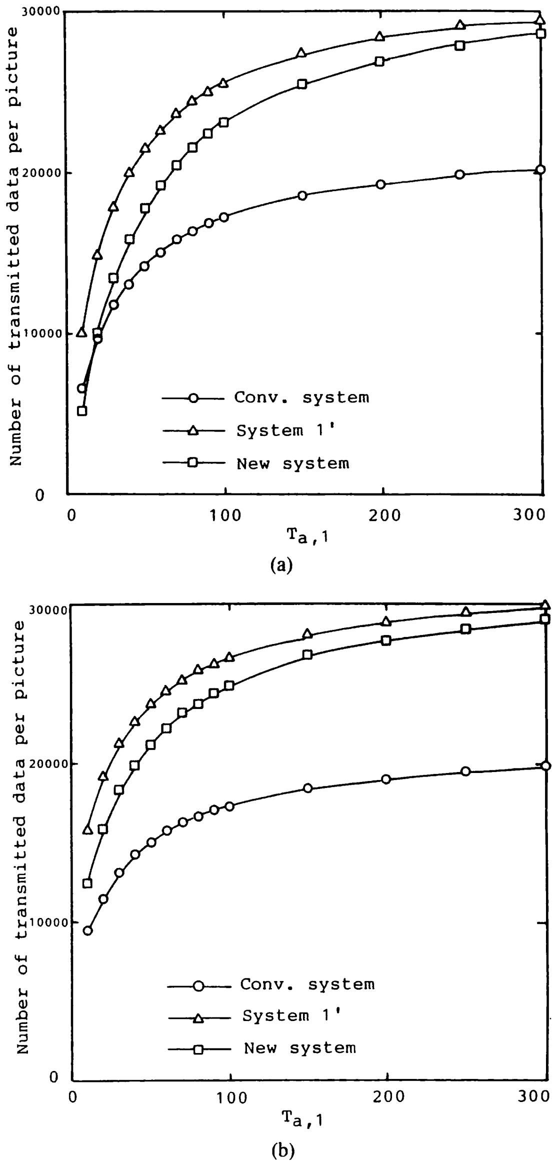
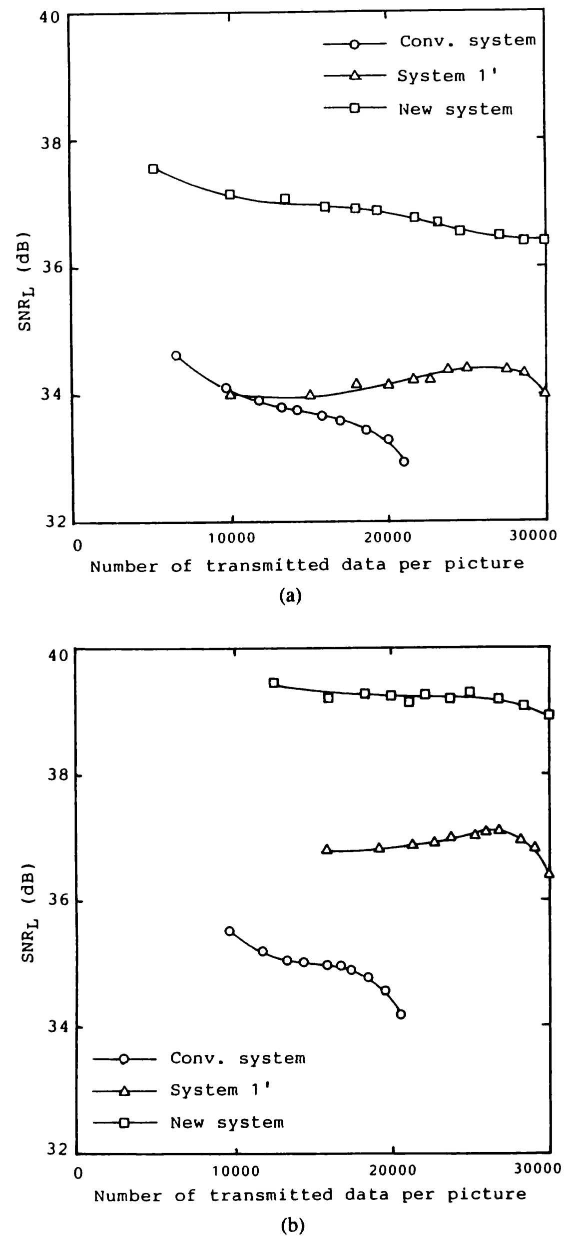


Fig. 8. (a) Number of transmitted data as a function of $T_{a,1}$ (girl image). (b) Number of transmitted data as a function of $T_{a,1}$ (couple image).

In Fig. 7, SNR_L was degraded as $T_{a,1}$ became large in both the conventional and the new system, but the degradation was greater in the conventional system. This is most noticeable in the systems without postfiltering. There are two reasons for this tendency.

First, in the conventional system, the luminance level of the central pel is replaced by the mean value of the block's pels. Therefore, if $T_{a,1}$ is large, the local picture activity also becomes large, and regenerated value becomes far from the original value. In the new system, the receiver regenerates luminance levels by means of the almost complete reverse procedure, and no distortion is caused by that kind of operation.

Second, in the conventional (horizontal block arrangement) system, the third pel of a data embedded block is most distorted due to the channel dispersion. Therefore, if the value of $T_{a,1}$ becomes large, the amount of the pel which has such distortion would increase in proportion to the amount of embedded data. In other words, the amount of embedded data directly determines the degree of the distortion. In the new (vertical block arrangement) system, the central pel of the



(a)

(b)

Fig. 9. (a) SNR$_L$ versus number of transmitted (girl image). (b) SNR$_L$ versus number of transmitted data (couple image).

TABLE II
SYSTEM PERFORMANCES

		SNR$_L$(dB)		Number of embedded data
			Post filtering	
Conv.	girl	30.42	33.76	14241
	couple	31.46	34.97	15128
System 1'	girl	28.29	34.23	21600
	couple	31.38	36.99	23800
New	girl	34.01	36.94	17855
	couple	35.98	39.16	21200

block that has embedded data which are different from the previous block's data are most distorted. As there is little distortion in the block that has embedded the same data as that of the previous block, the distortion would not always increase in proportion to the amount of embedded data. So, the degradation of SNR$_L$ would be greater in the conventional system than in the new system.

On the other hand, the SNR$_L$ of system 1' without postfiltering was increased as $T_{a,1}$ became large. The reason for this is as follows. In system 1', a data-free block is regenerated by (34). As mentioned in Section III-B c), if the previous block has embedded data, x'_{21} has large distortion, and, furthermore, X_{21} has larger distortion. As $T_{a,1}$ becomes large, few blocks will suffer from such distortion. So, SNR$_L$ would be increased.

The system parameter $T_{a,1}$ is chosen within the limits of no error by both the data transmitting efficiency and the SNR$_L$ of the regenerated picture. The value 50 was used for $T_{a,1}$ in [3]. As this value is appropriate for Figs. 7 and 8, our experiments were run with $T_{a,1} = 50$. The values of other parameters are shown in Table I.

C. Performance Comparisons

Fig. 9 shows the relation between SNR$_L$ and the amount of conveyed data. The higher SNR$_L$ and the larger the number of conveyed data, the better the system performance. Therefore, in Fig. 9, we can see that the new system has better performance than the older system on both points, i.e., SNR$_L$ and the amount of conveyed data.

The performance of each system with the system parameters shown in Table I are displayed in Table II. It is shown that the SNR$_L$ of the regenerated picture of the new system is 3–4 dB greater than that of the conventional system, and the number of embedded data per picture in the new system is about 1.3 times as large as that in the conventional system. By comparing the new system to the system 1', the effect of the algorithm modification is ascertained. In Table II, it is shown that the SNR$_L$ was improved about 2.5 dB by the almost complete reverse algorithm.

Fig. 10 shows the original and regenerated signals' waveforms for data embedded scan line, embedded data, and error signal between the original and regenerated signals for the "girl" image.

Finally, we refer to memories and the amount of calculation. In the conventional system, as the operation is done in sequence along the scan line, the minimum memory requirement for real-time operation is for three-pel operation. The new system needs memory for two scan-lines plus two-pel operation at least. However, that is not a large problem considering recent progress in semiconductor technology.

The amounts of calculation for each block of both systems are nearly equal.

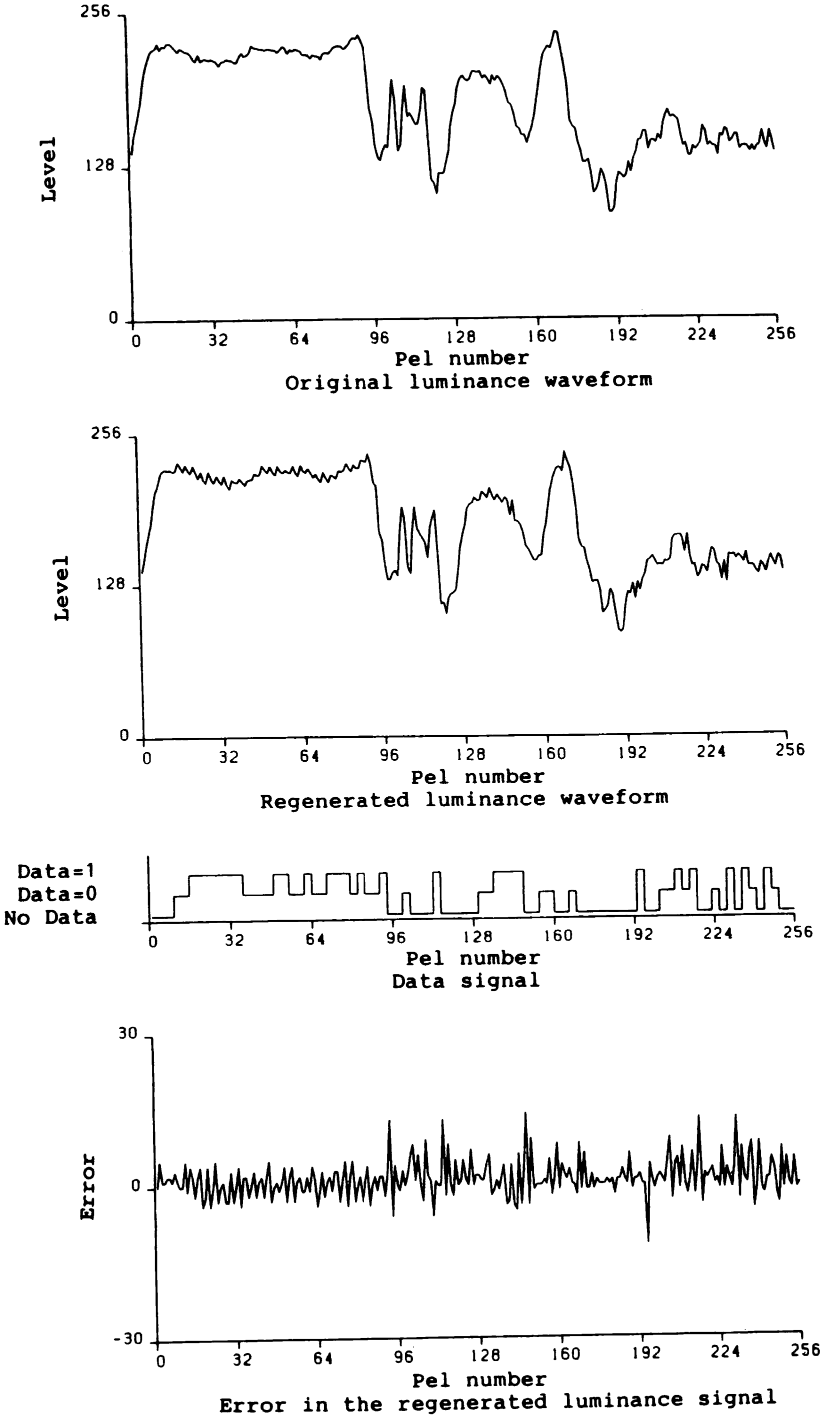
VI. CONCLUSION

By means of improving the system of embedding data into pictures by modulo masking, we proposed a new data embedding system with vertically allocated blocks. The vertical block allocation makes it possible not only to increase the amount of transmittable data, but also to operate an almost complete reverse procedure for regeneration of the picture. Also, we estimated the performance of the new system and compared it to the conventional system by computer simulation.

The amount of transmitted data in the new system was about 1.3 times as large as in the conventional, simultaneously the SNR$_L$ of the regenerated image was improved about 3–4 dB.

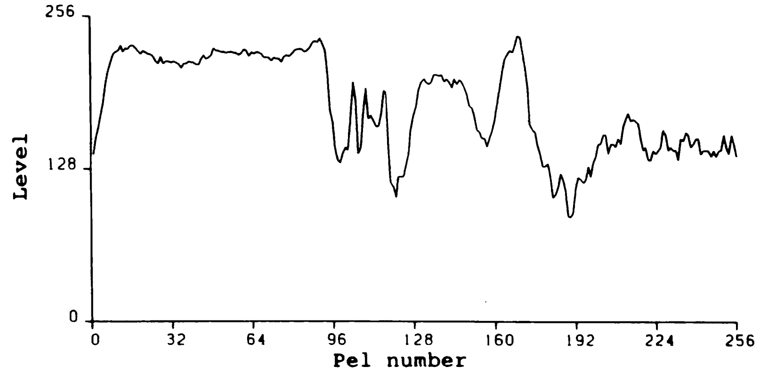
ACKNOWLEDGMENT

The authors would like to thank Prof. T. Tsunogae of Keio University for his support.

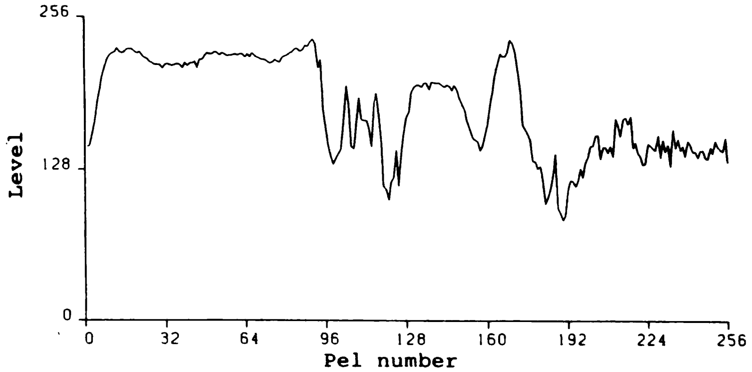


(a)

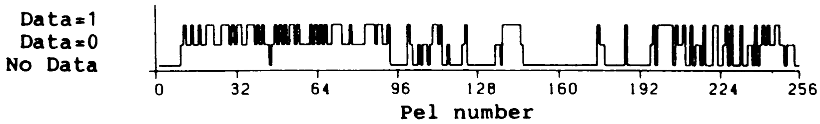
Fig. 10. (a) Waveforms of scan line duration (conventional system). (b) Waveforms of scan line duration (system 1′). (c) Waveforms of scan line duration (new system).



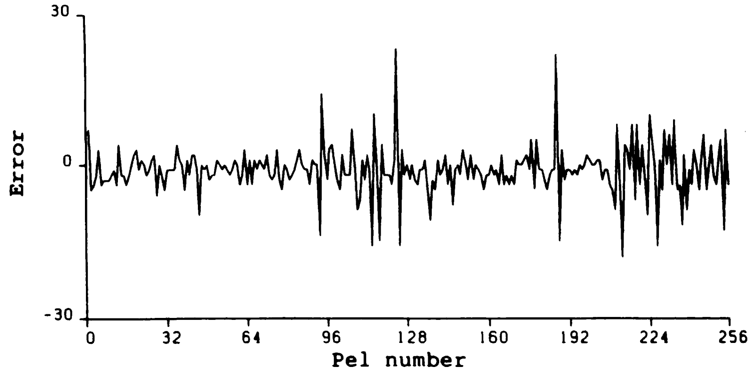
Original luminance waveform



Regenerated luminance waveform



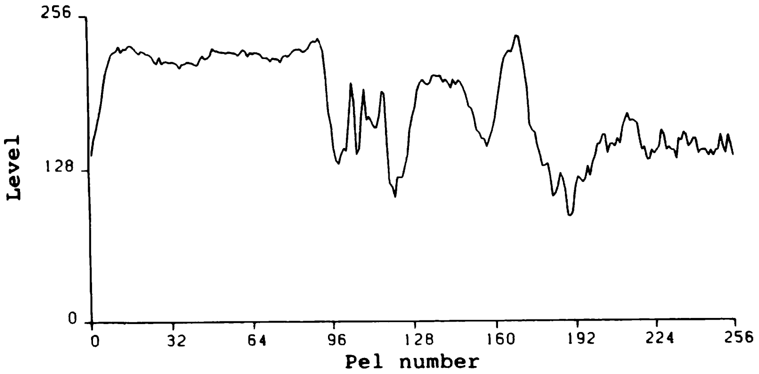
Data signal



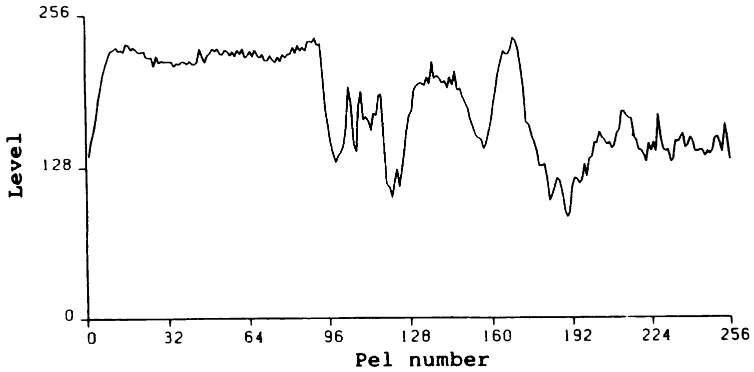
Error in the regenerated luminance signal

(b)

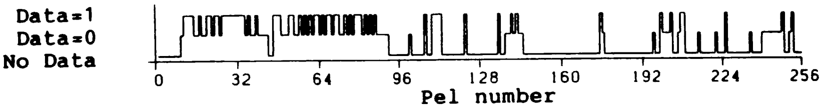
Fig. 10. (*Continued*).



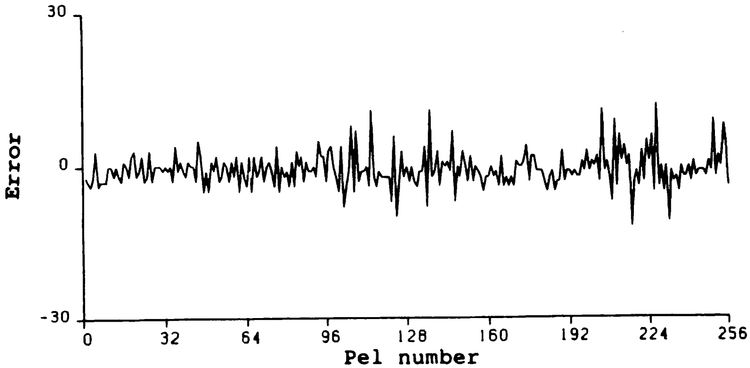
Original luminance waveform



Regenerated luminance waveform



Data signal



Error in the regenerated luminance signal

(c)

Fig. 10. (*Continued*).

REFERENCES

[1] R. Steele and D. Vitello, "Simultaneous transmission of speech and data using code-breaking techniques," *Bell Syst. Tech. J.*, vol. 60, pp. 2081-2105, Nov. 1981.

[2] ——, "Embedding data in speech using scrambling techniques," in *Proc. IEEE ICASSP-82,* Paris, France, May 3-5, 1982, Session DSP14, pp. 1801-1804.

[3] C. S. Xydeas, B. Kostic, and R. Steele, "Embedding data into pictures by modulo masking," *IEEE Trans. Commun.*, vol. COM-32, pp. 56-69, Jan. 1984.

[4] W. C. Wong, R. Steele, and C. S. Xydeas, "Transmitting data on the phase of speech signals," *Bell. Syst. Tech. J.*, vol. 61, pp. 2947-2970, Dec. 1982.

[5] E. F. Brown and A. N. Netravali, "Technique for transmitting digital data together with a video signal," U.S. Patent 4 237 484, Dec. 2, 1980.

[6] K. Feher, R. Goulet, and M. Morris, "1.544 Mbit/s data above FDM voice and data under FDM voice microwave transmission," *IEEE Trans. Commun.*, vol. COM-23, pp. 1321-1327, Nov. 1975.

[7] K. Feher and M. Morris, "Simultaneous transmission of digital phase-shift keying and of analog television signal," *IEEE Trans. Commun.*, vol. COM-23, pp. 1509-1514, Dec. 1975.

[8] G. B. Lockhart and Y. O. Al-jalili, "Method for superimposing data on amplitude-modulated signals," *Electron. Lett.*, vol. 18, pp. 379-380, Apr. 1982.

[9] T. L. Lim and M. S. Mueller, "Adaptive equalization and phase tracking for simultaneous analog/digital data transmission," *Bell Syst. Tech. J.*, vol. 60, pp. 2039-2063, Nov. 1981.

[10] F. Akashi, Y. Sato, and M. Eguchi, "High-speed digital and analog parallel transmission technique over single telephone channel," *IEEE Trans. Commun.*, vol. COM-30, pp. 1213-1218, May 1982.

[11] Y. Saito and M. Nakagawa, "A simple multiplex system for analog and digital signals," *Trans. IECE Japan,* vol. J65-B, no. 1, pp. 86-93, 1982.

[12] M. Nakagawa and Y. Saito, "A simple hybrid multiplexing method for analog and digital signals," in *Proc. IEEE ICC'82,* June 13-17, 1982, p. 7f.5.

[13] K. Hara, T. Hasegawa, and M. Nakagawa, "An improved method of embedding data into pictures by modulo masking," in *Proc. IEEE GLOBECOM '85,* New Orleans, LA, Dec. 2-5, 1985, p. 20.6.

★

Kazuhiko Hara was born in Niigata, Japan, on Dec. 6, 1962. He received the B.S. degree in electrical engineering from Keio University, Japan, in 1984.

He joined Canon Inc. in 1984, where he has been engaged in development and design of measuring equipments to test IC packages using ultrasonic sound.

Tadashi Shimomura was born in Nara, Japan, on October 21, 1963. He received the B.S. degree in electrical engineering from Keio University, Yokohama, Japan, in 1986. He is now in the Master course of Department of Electrical Engineering, Keio University.

His current research interest is in digital image processing.

★

Takaaki Hasegawa (S'82–M'86) was born in Kamakura, Japan, on July 4, 1957. He received the B.S., M.S., and Dr. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1981, 1983, and 1986, respectively.

Since 1986 he has been a Research Associate at Saitama University, Saitama, Japan. His research interests are spread-spectrum systems, multiplex operations, scrambling techniques, optical communications, signal processing, and information processing.

Dr. Hasegawa is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.

★

Masao Nakagawa (M'81) was born in Tokyo, Japan, on November 19, 1946. He received the B.S., M.S., and Dr. degrees in electrical engineering from Keio University, Yokohama Japan, in 1969, 1971, and 1974, respectively.

Since 1973, he has been with the Department of Electrical Engineering, Keio University, where he is now an Associate Professor. His research interests are in signal processing for communication systems, nonlinear phenomena, optical communications, spread-spectrum communications, speech processing, and nonlinear acoustics. He has published more than 40 journal papers, more than 20 international conference papers, and 6 books as an author and coauthor.

Dr. Nakagawa is a member of the Institute of Electronics, Information, and Communication Engineers of Japan.

T-COM/36/3//19003

An Improved Method of Embedding Data into Pictures by Modulo Masking

Kazuhiko Hara
Tadashi Shimomura
Takaaki Hasegawa
Masao Nakagawa

Reprinted from
IEEE TRANSACTIONS ON COMMUNICATIONS
Vol. 36, No. 3, March 1988