

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
11 December 2003 (11.12.2003)

PCT

(10) International Publication Number
WO 03/102947 A1

(51) International Patent Classification⁷: **G11B 20/00**

(21) International Application Number: PCT/IB03/02319

(22) International Filing Date: 21 May 2003 (21.05.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
02077170.5 3 June 2002 (03.06.2002) EP

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL];
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LEMMA, Aweke, N.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **APREA, Javier, F.** [IT/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

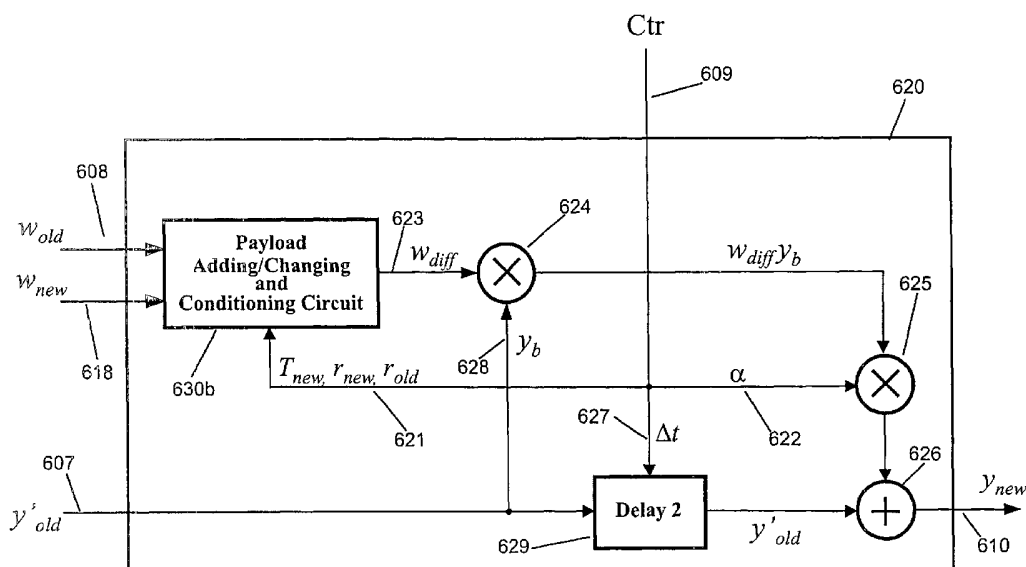
(74) Agent: **SCHMITZ, Herman, J., R.**; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: RE-EMBEDDING OF WATERMARKS IN MULTIMEDIA SIGNALS



(57) Abstract: Methods and apparatus for processing a multimedia signal comprising a watermark signal are described. The method includes the steps of: removing at least a portion of the watermark signal, and adding a new watermark signal to the multimedia signal so as to form a new watermarked multimedia signal.

Re-embedding of watermarks in multimedia signals

The present invention relates to apparatus and methods for re-embedding information in multimedia signals, such as audio, video or data signals.

5 Watermarking of multimedia signals is a technique for the transmission of additional data along with the multimedia signal. For instance, watermarking techniques can be used to embed copyright and copy control information into audio signals.

 The main requirement of a watermarking scheme is that it is not observable (i.e. in the case of an audio signal, it is inaudible) whilst being robust to attacks to remove the
10 watermark from the signal (e.g. removing the watermark will damage the signal). It will be appreciated that the robustness of a watermark will normally be a trade off against the quality of the signal in which the watermark is embedded. For instance, if a watermark is strongly embedded into an audio signal (and is thus difficult to remove) then it is likely that the quality of the audio signal will be reduced, if one tries to remove it without the knowledge of
15 the underlying technique and the secret key.

 The altering of a watermark by increasing the amount of embedded information in a watermark signal is known. In such instances, an extra watermark sequence is added to an existing watermarked signal. This is, for instance, implied in the 4C 12 bit watermark specification. A copy of this specification can be found at
20 <http://www.4centity.com/data/tech/4cspec.pdf>.

 It will be appreciated that when the watermark information needs to be changed repeatedly, such an approach not only degrades the quality of the original information signal (as additional watermark signals are added to change the payload) but also collisions between the individual embedded watermarks significantly degrade the watermark
25 robustness. For example, certain applications of copyright require that the copyright information embedded in a signal is changed repeatedly in order to assert proper copy control.

It is an object of the present invention to provide a watermarking scheme that substantially addresses at least one of the problems of the prior art, whether referred to herein or otherwise.

5

In a first aspect, the present invention provides a method of processing a multimedia signal comprising a watermark signal, the method comprising the steps of: removing at least a portion of an original watermark signal; and adding a new watermark signal to the multimedia signal so as to form a new watermarked multimedia signal.

10

Preferably, said original watermark signal is removed by applying a negative version of said original watermark signal to the multimedia signal.

Preferably, the method further comprises the step of determining the value of at least one of the parameters used to embed the original watermark in the multimedia signal.

15

Preferably, said parameter is utilized to remove at least a portion of said original watermark signal.

Preferably, said new signal is embedded in the multimedia signal using said embedding parameters having said determined value.

Suitably, said new signal is embedded in the multimedia signal using said embedding parameters having values other than the said determined values.

20

Preferably, said parameter comprises at least one of: embedding strength, synchronization information, time offset, time-scaling, an amount of a circular shift of a sequence, and a watermark symbol period.

Suitably, all of said original watermark is removed, and the new signal comprises a new watermark signal.

25

Preferably, said watermark comprises at least two sequences of values, at least one sequence of values being removed as said portion of the watermark signal, so as to leave at least one remaining sequence of values from the original watermark signal; and wherein said new signal comprises at least one further sequence of values which together with said remaining sequence forms a new watermark signal.

30

Preferably, all of said sequences of values are formed from a single sequence of values which has been circularly shifted by different amounts.

Preferably, said removed portion of the watermark signal was embedded with a predetermined strength into said multimedia signal, the method comprising the step of

embedding the new signal into the multimedia signal with preferably the same predetermined strength.

Preferably, the embedding strength is such that the degradation in the quality of the new watermarked multimedia signal is perceptible but not annoying.

5 Preferably, at least one of the original watermark signal and the new watermark signal comprises a smoothly varying signal formed by applying a window shaping function to a sequence of values, the integral over the window shaping function being zero.

Preferably, the window shaping function has a bi-phase behavior.

10 Preferably, the bi-phase window comprises at least two Hanning windows of opposite polarities.

In another aspect, the present invention provides a computer program arranged to perform any of the methods described above.

In a further aspect, the present invention provides a record carrier comprising a computer program described above.

15 In another aspect, the present invention provides a method of making available for downloading a computer program as described above.

In a further aspect, the present invention provides an apparatus for processing a multimedia signal comprising a watermark signal, the apparatus comprising: a deletion unit arrange to remove at least a portion of the watermark signal; and an embedder arranged to
20 add a new signal to the multimedia signal so as to form a new watermark signal.

Preferably, the apparatus further comprises a detector arranged to detect at least one value of a parameter of said watermark signal.

25 Preferably, the apparatus comprises a receiver of a multimedia signal comprising an apparatus.

For a better understanding of the invention, and to show how embodiments of the same may be carried into effect, reference will now be made, by way of example, to the accompanying diagrammatic drawings in which:

30 Figure 1 is a diagram illustrating a generalized watermark re-embedding apparatus according to a first embodiment of the present invention;

Figure 2 shows a schematic diagram of one type of watermark embedder;

Figure 3 is a schematic diagram showing the details of the payload adding and conditioning circuit (unit 630a of figure 2) used for embedding a primary watermark payload;

Figure 4 is a schematic diagram illustrating a watermark detector that can be used in one preferred embodiment of the invention;

Figure 5 is a schematic diagram showing the details of watermark generator unit 650 according to a preferred embodiment;

5 Figure 6 is a diagram illustrating a watermark embedder apparatus according to a preferred embodiment;

Figure 7 shows a signal portion extraction filter H ;

Figure 8 shows the payload adding/changing and watermark conditioning stage used for re-embedding a watermark;

10 Figure 9 is a diagram illustrating the details of the watermark conditioning apparatus H_c of figure 8, including charts of the associated signals at each stage;

Figure 10 shows a typical correlation function corresponding to the original payload of a watermark signal received by the apparatus of figure 1 and figure 4;

15 Figure 11 shows a correlation function after removal of one of the two sequences forming the original watermark; and

Figure 12 shows a correlation function of a new watermark having a new payload, formed by adding a second watermark sequence to the signal having the correlation function shown in figure 10.

20

Figure 1 illustrates a re-embedding apparatus 600 according to a first embodiment of the present invention. The apparatus includes an input 602 arranged to receive a watermarked information signal y'_{old} . Two copies of the input signal y'_{old} are formed, with one copy going to a delay unit 614, and the other being input to a detector 640.

25 The detector 640 is arranged to detect an estimate w'_{old} (output 612) of the watermark w_{old} that is embedded within the received signal y'_{old} , and to estimate the control parameters (output 609) needed for changing the watermark payload.

Information relating to the detected watermark w'_{old} is then passed to the watermark generator 650 and the watermark embedder 620. The information passed to the watermark generator can be a complete copy of the extracted watermark w'_{old} , or
30 alternatively sufficient information (such as, for watermarks comprising two circularly shifted sequences of a single series of values, the circular shift d_{old}) to allow the watermark generator to generate a copy w_{old} of w'_{old} . The output 608 of the watermark generator, also denoted by w_{old} , is preferably an error corrected version of the extracted watermark w'_{old} .

The watermark generator 650 additionally generates a new watermark signal w_{new} (output 618), for adding to the original information signal y'_{old} .

The delay unit 614 acts to delay the input signal y'_{old} whilst the intervening operations are carried out by the detector 640, and the watermark generator 650. The delayed signal y'_{old} is then passed (output 607) to the embedder 620.

In the watermark embedding unit 620, the old watermark signal w_{old} is removed from the signal y'_{old} , and the new watermark w_{new} added to y'_{old} so as to form an information signal y_{new} containing the new watermark w_{new} .

The above embodiment represents a particular generalized implementation of the present invention. Below is described a particular watermark re-embedding scheme in accordance with a further embodiment of the present invention for use in conjunction with a particular watermarking scheme.

Such a watermark can for example be embedded using the apparatus shown in figures 2 and 3.

Figure 2 shows an embedding apparatus 100 arranged to receive a host multimedia signal x at input 120, and output a watermarked version of the received signal (y_{old}) at output 132.

The apparatus 100 receives two sequences of values ($w_{old}[k]$ and $w_{ref}[k]$) at input 110 and 112. The combination of these two sequences of values is used to produce the watermark signal w_o . Each of the two sequences of values is a circularly shifted version of an original sequence of values. w_{old} is the original sequence w_s circularly shifted by d_{old} , and w_{ref} the original sequence w_s circularly shifted by d_{ref} i.e. w_{old} is hence a circularly shifted version of w_{ref} . The original sequence of values can, for instance, be generated using a Random Number Generator (RNG) using a predetermined seed value S .

As shown in Figure 3, the payload adding and conditioning circuit 630a receives $w_{old}[k]$ and $w_{ref}[k]$ at respective inputs 110 and 112. Each sequence is multiplied by a respective sign bit r_{old} and r_{ref} (received at inputs 603 and 604) using multipliers 631 (where r_{old} and r_{ref} are respectively +1 or -1, and remain constant for any given payload). The resulting signals are then passed through signal conditioning units 632, which act to convolve each of the values within each sequence with a window shaping function of the period T_s . This produces smoothly varying output sequences $w_{old}[n]$ and $w_{ref}[n]$. Operation of a signal conditioning unit 632 is described later in more detail with reference to Figure 9.

Signal $w_{ref}[n]$ is then delayed by delay unit 634 by a predetermined delay T_r (where T_r is less than T_s , and preferably, $T_r = T_s/4$). Adder 635 then adds $w_{old}[n]$ and the delayed version of $w_{ref}[n]$, the result being the watermark signal $w_o[n]$ which is output from the circuit 630 at output 623.

- 5 Referring back to Figure 2, the received host signal x is split into two copies, a first copy going to adder 130, and a second copy going to multiplication unit 124. The copy sent to multiplication unit 124 may be filtered using a band pass filter (not shown) so as to form the signal x_b at input 122 to the multiplier 124.

$w_o[n]$ is multiplied with the possibly filtered version $x_b[n]$ of the host signal $x[n]$, scaled by gain factor α and added back to the host signal $x[n]$ to generate the watermark signal $y_{old}[n]$ given by

$$y_{old}[n] = x[n] + \alpha w_o[n] x_b[n]. \quad (1)$$

- The gain factor α is set by the variable gain device 128, and in the example shown is set in dependence upon the signal from the signal analyzer 126 which samples the host signal x using a psychoacoustic model. Such a model is, for instance, described in the paper by
- 15 E.Zwicker, "Audio Engineering and Psychoacoustics: Matching signals to the final receiver, the Human Auditory System", *Journal of the Audio Engineering Society*, Vol. 39, pp. Vol.115-126, March 1991. The gain factor α is normally chosen so as to minimize the impact of the watermark signal on the host signal quality.

- Such a watermarking scheme is characterized in that, during detection the
- 20 watermarked signal y_{old} generates two correlation peaks that are separated by pL_{old} (see figure 10), where the value pL_{old} is at least part of the watermark payload, and may be defined as

$$pL_{old} = |d_{ref} - d_{old}| \bmod \left(\left\lceil \frac{L_w}{2} \right\rceil \right) \quad (2)$$

where d_{ref} and $d_{old} \in [0, L_w - 1]$ are the relative positions of the correlation peaks as seen in figure 10, and L_w is the length (number of symbols or values) of each of the two watermark sequences.

- 25 In addition to pL_{old} , extra information is also encoded by changing the relative signs of the embedded watermarks. In the detector, this is seen as a relative sign r_{sign} between the correlation peaks. It will be seen that r_{sign} can take four possible values, and may be defined as:

$$r_{\text{sign}} = \frac{2 \cdot \mu_{\text{ref}} + \mu_{\text{old}} + 3}{2} \in \{0,1,2,3\}, \quad (3)$$

where μ_{old} and μ_{ref} are the signs of the correlation peaks, and correspond to the sign bits r_{old} and r_{ref} in figure 3, respectively. The overall watermark payload pL_w is then given as a combination of r_{sign} and pL_{old} :

$$pL_w = \langle r_{\text{sign}}, pL_{\text{old}} \rangle. \quad (4)$$

- 5 The maximum information (I_{max}), in number of bits, that can be carried by a watermark sequence of length L_w is thus given by:

$$I_{\text{max}} = \log_2 \left(4 \cdot \left\lceil \frac{L_w}{2} \right\rceil \right) \text{ bits} \quad (5)$$

Below is described a watermark re-embedding apparatus suitable for use in conjunction with the watermark embedder shown in figure 2 and the related watermarking
10 scheme. In this particular embodiment, the embedding apparatus is arranged to only partially remove the watermark, and replaces it with new information so as to change the watermark payload. Figure 1 illustrates the different functional blocks of the re-embedder.

Figure 4 shows a block diagram of a watermark detector 640 used to extract the watermark embedding parameters that are needed for the re-embedding process. The
15 detector consists of four major stages: (a) the watermark symbol extraction stage (200), (b) the buffering and interpolation stage (300), (c) the correlation and decision stage (400) and (d) the control signal generating stage (500).

In the symbol extraction stage (200), the received watermarked signal $y'_{\text{old}}[n]$ is processed to generate multiple (N_b) estimates of the watermarked sequence, which are
20 multiplexed into the signal $w_e[m]$. These estimates of the watermark sequence are required to resolve (compensate for) any time offset that may exist between the embedder and the detector, so that the watermark detector can synchronize to the watermark sequence inserted in the host signal.

In the buffering and interpolation stage (300), these estimates are de-
25 multiplexed into N_b separate buffers. An interpolation is subsequently applied to each buffer to resolve (compensate for) possible timescale modifications that may have occurred. For instance, a drift in sampling (clock) frequency may result in a stretch or shrink in the time domain signal (i.e. the watermark may have been stretched or shrunk).

In the correlation and decision stage (400), the content of each buffer is correlated with the reference watermark and the maximum correlation peaks are compared against a threshold to determine the likelihood of whether the watermark is indeed embedded within the received signal $y'_{old}[n]$.

5 In the control signal generating unit (500), the detection truth value, the corresponding watermark sequence, its buffer index, and the value of the new payload pL_{new} (input 616) are combined to generate parameters that are needed to changing the watermark payload. The outputs (609, 612) of the control signal generator are passed on to the watermark generator 650 and the watermark embedding unit 620.

10 Details of the watermark generating unit 650 are shown in figure 5. In this unit, w_{new} and w_{old} are generated using parameter information obtained from the detector 640 (inputs 609 and 612).

In one preferred embodiment, the watermark sequences w_{old} and w_{new} can be generated as follows. Firstly a finite length, preferably zero mean and uniformly distributed random sequence w_s is generated using a random number generator 651 with an initial seed S .
15 It will be appreciated that this initial seed S is preferably the same as that used in generating w_{old} during the first embedding. This results in the sequence of length L_w

$$w_s[k] \in [-1,1], \text{ for } k=0,1,2, \dots, L_w-1 \quad (6)$$

Then the sequence w_s is circularly shifted by the amounts d_{old} and d_{new} using the circularly shifting units 653a and 653b to obtain the random sequences w_{old} and w_{new} ,
20 respectively. It will be appreciated that these two sequences (w_{old} and w_{new}) are effectively a first sequence and a second sequence, with the second sequence being circularly shifted with respect to the first. These two sequences are then passed on to the watermark embedding unit 620 (outputs 608, 618).

25 Details of a watermark embedding unit 620 as part of the watermark re-embedding apparatus 600, and suitable for use with the considered particular watermarking scheme, is shown in figure 6.

The host signal y'_{old} (containing at least an initial watermark sequence w_{old}) is provided at input 607 of the apparatus. The host signal y'_{old} is passed in the direction of output 610 via the delay unit 629 and the adder 626. However, a replica y_b of the host signal
30 y'_{old} (input 628) is split off in the direction of the multiplier 624, for carrying the new watermark information.

The multiplier 624 is utilized to calculate the product of the watermark altering signal w_{diff} and the replica signal y_b . The watermark altering signal w_{diff} is obtained

from the payload changing and watermark conditioning apparatus 630b, and derived from the watermark random sequences w_{old} and w_{new} (inputs 608 and 618, respectively), which are input to the payload changing and watermark conditioning apparatus.

The resulting product, $w_{diff}y_b$ is then passed via a gain controller 625 to the adder 626. The gain factor α applied by the controller 625 controls the trade off between the audibility and the robustness of the watermark. It may be a constant, or variable in at least one of time, frequency and space. The apparatus in figure 6 shows that, when α is variable, it can be automatically adapted via a control signal 609 obtained from the watermark detector unit 640.

In another preferred embodiment, the gain factor α can be independently controlled via a signal analyzer based upon the properties of the host signal y_{old} . In the latter case, the gain α is automatically adapted, preferably so as to minimize the impact on the signal quality, according to a properly chosen perceptibility cost-function, such as a psychoacoustic model of the human auditory system (HAS). The same model may be used as that utilized to control the adaptive gain factor used to control the embedding strength of the original watermark signal (see Figure 2 and associated text).

In Figure 6, the resulting watermark audio signal y_{new} is then obtained at the output 610 of the embedding apparatus 620 by adding an appropriately scaled version of the product of w_{diff} and y_b to the host signal:

$$y_{new}[n] = y_{old}[n] + \alpha w_{diff}[n] y_b[n] \approx x[n] + \alpha w_c[n] x_b[n], \quad (7)$$

where w_c is derived from w_{ref} and w_{new} in the same way as w_o was derived from w_{ref} and w_{old} in figure 3.

Preferably, the parameters of the watermark $w_{diff}[n]$ are chosen such that when multiplied with y_b , it predominantly modifies the short time envelope of y_b .

Figure 7 shows one preferred embodiment in which the input 628 to the multiplier 624 in figure 6 is obtained by filtering the host signal y'_{old} using a filter H in the filtering unit 615. Preferably, the filter H is a linear phase band-pass filter characterized by its lower cut off frequency f_L and upper cut off frequency f_H . Preferably, the filter has the same properties as the filter utilized to extract x_b from x in the embedder 100.

In figure 8, the details of the payload changing and watermark-conditioning unit 630b are shown. In this particular unit, the watermark signals w_{old} and w_{new} are combined to generate the multi-bit watermark altering signal w_{diff} . The watermark altering signal w_{diff} ,

when combined with y'_{old} , generates a watermarked signal with a payload corresponding to w_{new} .

The sequences w_{old} and w_{new} are first multiplied with a respective sign bit r_{old} and r_{new} in the multiplying units 654a and 654b. The respective values of r_{old} and r_{new} are
 5 derived from the detection unit 640 and are passed on via the control input 609. The values of r_{old} and r_{new} remain constant (typically at either +1 or -1), and only change when the payload of the watermark is changed.

The difference $w_{diff}[k]$ between the signed sequences w_{old} and w_{new} is calculated using adder 635 to add the negative of the old watermark sequence w_{old} to the
 10 positive of the new sequence w_{new} . The result w_{diff} is then passed through the conditioning stage to generate the slowly varying multi-bit watermark $w_{diff}[n]$.

Figure 9 shows details of the watermark conditioning apparatus 632 used in the payload adding/changing and watermark conditioning apparatus 630. In the case of re-embedding, the watermark random sequence $w_{diff} = w_{new} - w_{old}$ is input to the conditioning
 15 apparatus 632.

In the conditioning circuit, the watermark signal sequence $w_{diff}[k]$ is first applied to the input of an up-sampler 180. Chart 181 illustrates one of the possible sequences w_{diff} as a sequence of values of random numbers between +1 and -1, with the sequence being of length L_w . The up-sampler adds $(T_s - 1)$ zeros between each sample so as to raise the
 20 sampling frequency by the factor T_s . T_s is referred to as the watermark symbol period and represents the span of the watermark symbol in the audio signal. In the case of the received signal y'_{old} having undergone a time scaling compared with the transmitted signal y_{old} , T_s is replaced with appropriately scaled sampling factor T_{new} that takes into account the scaling effect. Chart 183 shows the results of the signal illustrated in chart 181 once it has passed
 25 through the up-sampler 180.

A window shaping function $s[n]$, such as a bi-phase window, is then convolved with the up-sampled signal $w_i[n]$ so as to convert it into a slowly varying narrow-band signal $w_{diff}[n]$, whose behavior for the $w_{diff}[k]$ sequence of chart 181 is as shown in chart 185.

Chart 184 shows a typical bi-phase window shaping function. The window shaping function has support in the interval 0 to T_s only. The window function is applied to the watermark sequence in order to produce a smoothly varying signal, so as to minimize the decrease in the quality of the host signal.

Below is described in more detail the operation of the detection apparatus (200, 300, 400, 500) shown in figure 4.

In the watermark symbol extraction stage 200 shown in figure 4, the incoming watermark signal $y'_{old}[n]$ is input to the signal conditioning filter $H_b(210)$. This filter 210 is typically a band pass filter and has the same behavior as the corresponding filter H (615) shown in figure 7. The output of the filter H_b is $y'_b[n]$, and assuming linearity within the transmission channel, it follows from equation (1)

$$y'_b[n] \approx (1 + \alpha w_c[n])x_b[n] \quad (8)$$

Note that when no filter is used in the embedder (i.e., when $H=1$) then H_b in the detector can also be omitted, or it can still be included to improve the detection performance. If H_b is omitted, then y'_b in equation (8) is replaced with y'_{old} . The rest of the processing is the same.

For simplification, it is assumed that there is perfect synchronism between the embedder and the detector (i.e. no offset and no change in timescale), and that the audio signal is divided into frames of length T_s , and that $y'_{b,m}[n]$ is the n -th sample of the m -th frame of the filtered signal $y'_b[n]$. It should be noted that if there is not perfect synchronism between the embedder and the detector, then any deviation can be compensated for within the buffering and interpolation stage 300 utilizing techniques known to the skilled person e.g. iteratively searching through all possible scale and offset modifications until a best match is achieved.

The energy $E[m]$ corresponding to the $y'_{b,m}[n]$ frame is:

$$E[m] = \sum_{n=0}^{T_s-1} |y'_{b,m}[n]S[n]|^2 \quad (9)$$

where $S[n]$ is the same window shaping function used in the watermark conditioning circuit of figure 9. A person skilled in the art will appreciate that equation 9 represents a matched filter receiver, and is the optimum receiver when the symbol period is perfectly synchronized. Notwithstanding this fact, from now on, we set $S[n]=1$ in order to simplify subsequent explanations.

Combining this with equation 8, it follows that:

$$E[m] = \sum_{n=0}^{T_s-1} |y'_{b,m}[n]|^2 \approx \sum_{n=0}^{T_s-1} |(1 + \alpha w_e[m])x_{b,m}[n]|^2 \quad (10)$$

where $w_e[m]$ is the m -th extracted watermark symbol. and contains N_b time-multiplexed estimates of the embedded watermark sequences. Solving for $w_e[m]$ in equation 10 and ignoring higher order terms of α , gives the following approximation:

$$w_e[m] \approx \frac{1}{2\alpha} \left(\frac{\sum_{n=0}^{T_s-1} |y'_{b,m}[n]|^2}{\sum_{n=0}^{T_s-1} |x_{b,m}[n]|^2} - 1 \right) \quad (11)$$

In the watermark extraction stage 200 shown in figure 4, the output $y'_b[n]$ of the filter H_b is provided as an input to a frame divider 220, which divides the audio signal into frames of length T_s i.e. into $y'_{b,m}[n]$, with the energy calculating unit 230 then being used to calculate the energy corresponding to each of the framed signals as per equation (9). The output of this energy calculation unit 230 is then provided as an input to the whitening stage H_w (240) which performs the function shown in equation 11 so as to provide an output $w_e[m]$.

It will be realized that the denominator of equation 11 contains a term that requires knowledge of the host (original) signal x . As the signal x is not available to the detector, it means that in order to calculate $w_e[m]$ then the denominator of equation 11 must be estimated.

Below is described how such an estimation can be achieved for a bi-phase window shaping function, but it will be appreciated that the teaching could be extended to other window shaping functions.

It will be seen by examination of the bi-phase window function shown in figure 9 (chart 184), that when the envelope of an audio frame is modulated with such a window function, the first and the second halves of the frame are scaled in opposite directions. In the detector, this property is utilized to estimate the envelope energy of the host signal y'_{old} .

Consequently, within the detector, the audio frame is first sub-divided into two halves. The energy functions corresponding to the first and second half frames are hence given by

$$E_1[m] = \sum_{n=0}^{T_s/2-1} |y'_{b,m}[n]|^2 \quad (12)$$

and

$$E_2[m] = \sum_{n=T_s/2}^{T_s-1} |y'_{b,m}[n]|^2 \quad (13)$$

respectively. As the envelope of the original audio is modulated in opposite directions within the two sub-frames, the original audio envelope can be approximated as the mean of $E_1[m]$ and $E_2[m]$.

Further, the instantaneous modulation value can be taken as the difference between these two functions. Thus, for the bi-phase window function, the watermark $w_e[m]$ can be approximated by:

$$w_e[m] \approx \frac{1}{2\alpha} \left(\frac{E_1[m] - E_2[m]}{E_1[m] + E_2[m]} - 1 \right) \quad (14)$$

This output $w_e[m]$ is then passed to the buffering and interpolation stage 300, where the signal is de-multiplexed by a de-multiplexer 310, buffered in buffers 320 of length L_b so as to resolve any lack of synchronism between the embedder and the detector, and interpolated within the interpolation unit 330 so as to compensate for any time scale modification between the embedder and the detector. Such compensation can utilize known techniques, and hence is not described in any more detail within this specification.

During detection, in order to maximize the accuracy of the watermark detection, the watermark detection process is typically carried out over a length of received signal $y'_{old}[n]$ that is 3 to 4 times that of the watermark sequence length. Thus each watermark symbol to be detected can be constructed by taking the averages of several symbols. This averaging process is referred to as smoothing, and the number of times the averaging is done is referred to as the smoothing factor s_f . Thus, the detection window length L_D is the length of the audio segment (in number of samples) over which a watermark detection truth-value is reported. Consequently, $L_D = s_f L_w T_s$, where T_s is the symbol period and L_w the number of symbols within the watermark sequence. Typically, the length (L_b) of each buffer 320 within the buffering and interpolation stage is $L_b = s_f L_w$.

As shown in figure 4, outputs ($w_{D1}, w_{D2}, \dots, w_{DNb}$) from the buffering stage are passed to the interpolation stage and, after interpolation, the outputs ($w_{I1}, w_{I2}, \dots, w_{INb}$) of this stage, which correspond to the different estimates of the correctly re-scaled signal, are passed to the correlation and decision stage. If it is believed that no time scaling compensation is required, the values ($w_{D1}, w_{D2}, \dots, w_{DNb}$) can be passed directly to the correlation and decision stage 400 i.e. the interpolation stage 330 can be omitted from the apparatus.

The correlator 410 calculates the correlation of each estimate $w_{Ij}, j=1, \dots, N_b$ with respect to the reference watermark sequence $w_s[k]$. Each respective correlation output corresponding to each estimate is then applied to the maximum detection unit 420 which determines which two estimates provided the best fits for the circularly shifted versions w_{old} and w_{ref} of the reference watermark. The correlation values (the peak amplitudes and

positions) for these estimate sequences are passed to the threshold detector and payload extractor unit 430.

In another output of the correlation stage 410, the watermark sequences and the buffer indices corresponding to the two best fits for the circularly shifted versions w_{old} and w_{ref} of the reference watermark are passed on to the control signal generating unit 500.

If the interpolation stage is omitted, alternatively the correlator 410 calculates the correlation of each estimate w_{Dj} , $j=1, \dots, N_b$ with the reference watermark sequence $w_s[k]$ and the results are passed on for subsequent processing to the units 420 and 430 as outlined in the above paragraph.

The threshold detector and payload extractor unit 430 may be utilized to extract the payload (e.g. information content) from the detected watermark signal. Once the unit has estimated the two correlation peaks cL_1 and cL_2 that exceed the detection threshold, the distance pL between the peaks (as defined by equation (2)) is measured. Next, the signs μ_1 and μ_2 of the correlation peaks are determined, and hence r_{sign} calculated from equation (3). The overall watermark payload may then be calculated using equation (4).

For instance, it can be seen in Figure 10 that pL_{old} is the relative distance between the two peaks. Both peaks are positive i.e. $\mu_1 = +1$, and $\mu_2 = +1$. From equation (3), $r_{sign} = 3$. Consequently, the payload $pL_w = \langle 3, pL_{old} \rangle$.

The reference watermark sequence w_s used within the detector corresponds to (a possibly circularly shifted version of) the original watermark sequence applied to the host signal. For instance, if the watermark signal was calculated using a random number generator with seed S within the embedder, then equally the detector can calculate the same random number sequence using the same random number generation algorithm and the same initial seed so as to determine the watermark signal. Alternatively, the watermark signal originally applied in the embedder and utilized by the detector as a reference could simply be any predetermined sequence.

Figure 10 shows a typical shape of a correlation function as output from the correlator 410. The horizontal scale shows the correlation delay (in terms of the sequence bins). The vertical scale on the left-hand side (referred to as the confidence level cL) represents the value of the correlation peak normalized with respect to the standard deviation of the (typically normally distributed) correlation function.

As can be seen, the typical correlation is relatively flat with respect to cL , and centered about $cL = 0$. However, the function contains two peaks, which are separated by pL_{old} (see equation 2) and extend upwards to cL values that are above the detection threshold

when a watermark is present. When the correlation peaks are negative, the above statement applies to their absolute values.

A horizontal line represents the detection threshold. The detection threshold value controls the false alarm rate.

5 Two kinds of false alarms exist: the false positive rate, defined as the probability of detecting a watermark in non watermarked items, and the false negative rate, which is defined as the probability of not detecting a watermark in watermarked items. Generally, the requirement of the false positive alarm is more stringent than that of the false negative.

10 After each detection interval, the detector determines whether the original watermark is present or whether it is not present, and on this basis output a “yes” or a “no” decision to the outputting device and at the same time to the control signal generating unit 500.

If desired, to improve this decision making process, a number of detection
15 windows may be considered. In such an instance, the false positive probability is a combination of the individual probabilities for each detection window considered, dependent upon the desired criteria. For instance, it could be determined that if the correlation function has two peaks above a threshold of $cL = 7$ on any two out of three detection intervals, then the watermark is deemed to be present. Obviously, such detection criteria can be altered
20 depending upon the desired use of the watermark signal and to take into account factors such as the original quality of the host signal and how badly the signal is likely to be corrupted during normal transmission.

In summary of the general operation of this particular re-embedding process, the re-embedding apparatus 600 is arranged to receive a signal y'_{old} containing a watermark at
25 input 602. The signal y'_{old} in this instance has been generated by the watermark embedding apparatus shown in figures 2 and 3, and includes a watermark comprising two circularly shifted versions w_{old} and w_{ref} of a single sequence w_s of values. A copy of the received signal y_{old} is passed to the detector 640.

As described above, the detector 640 is arranged to detect the presence of the
30 watermark within the signal y'_{old} , and to estimate the watermark embedding parameters (e.g. the amounts d by which the watermark sequences were circularly shifted, and the gain factor α used to control the trade off between the audibility and the robustness of the watermark).

Figure 10 illustrates a correlation function of the watermark embedded within y'_{old} , with the original sequence of values (w_s) used to form w_{old} . As can be seen, the payload

pL_{old} of the watermark in y'_{old} is, at least in part, defined by the two amounts by which the sequences comprising the watermark have been circularly shifted, d_{old} and d_{ref} .

In this preferred embodiment, only a portion of the original watermark signal is removed (the sequence of values which have been circularly shifted by the amount d_{old}).

- 5 The same sequence of values (w_s) as utilised in the original watermark is then circularly shifted by a new amount (d_{new}), and embedded within the information signal, using the detected embedding parameters.

Figure 11 illustrates a correlation function of the same watermark signal shown in figure 10, but in which the sequence of values corresponding to delay d_{old} has been removed (i.e. as if “ $-w_{old}$ ” had been added to y'_{old}). The result is a single correlation peak at d_{ref} .

Figure 11 shows the correlation function of the same signal shown in Figure 10, but after the same sequence of values w_s with a new circular shift delay (d_{new}) has been inserted. This results in two correlation peaks, separated by a payload pL_{new} (see equation 2 and 4) i.e. a new watermark signal. It will be appreciated that by only removing one half of the original watermark signal, and subsequently adding in a replacement half of the watermark signal, a new watermark has been generated but with minimum impact upon the quality of the information signal into which the watermark is embedded.

The detector 640 also provides synchronization information such as the time offset (Δt) to a delay unit 629. The copy of the input signal y'_{old} passed to the delay unit 629 is then appropriately synchronized to take into account any time offsets, rescaling and also the intrinsic delay caused by the various operations carried out by the units in the re-embedder (e.g. 630b, 624, 625, 615), and to ensure that y'_{old} is synchronized with y_b at the adder 626. One copy of the signal y'_{old} is passed to the filter H 615. This is similar to the filter shown in figure 7, and can be omitted. Such a filter, can for instance, be a band-pass filter, and gives an output y_b .

It will be appreciated that the above embodiments are provided by way of example only. Various modifications will be apparent to the skilled person.

For instance, whilst the preferred embodiment has described the partial removal of the original watermark, it will be appreciated that the whole of the original watermark could be removed and replaced. Equally, whilst only one sequence has been described as being removed and replaced by one sequence, it will be appreciated that a single sequence could be replaced by two or more sequences. Alternatively if the original payload

comprised three circularly shifted sequences or more, two such sequences could be replaced by a single sequence, or a plurality of sequences.

Whilst the new watermark has been described as utilizing the same values of embedding parameters as originally used, the new watermark could of course use alternative
5 values for any one of more of the embedding parameters.

For instance, whilst the above embodiment has described the watermark altering signal w_c as being scaled by a factor α , it will be appreciated that the two components forming w_c (ie. w_{old} and w_{new}) could be scaled by different amounts before being added together, or before being separately added directly to the received signal y'_{old} . In order
10 for the portion of the watermark to be removed, it is desirable that the embedding strength of the negative version of w_{old} is similar to that originally used to embed w_{old} in the host signal. However, w_{new} can obviously be embedded into y'_{old} with any desired strength.

It is desirable that watermarks are embedded into the host signal without unduly affecting the quality of the host signal. Preferably, all embedded watermark signals
15 are imperceptible to an observer i.e. in an audio signal, the effect of the watermark signal can not be heard, or in a video signal, the effect of the watermark signal can not be seen.

The ITU standard "Method for objective Measurements of Perceived audio quality", International Telecommunication Union, Geneva Switzerland (1999), defines a five grid scoring system (which is in conformation to ITU-R Rec BS.1116 (rev. 1) (1997) and
20 ITU-R Rec. BS.562-3 (1990) standards). The various scores are: 5 = Imperceptible, 4 = Perceptible but not annoying, 3 = Slightly annoying, 2 = Annoying, 1 = Very annoying.

Whilst it is preferable that all watermark signals are imperceptible, equally, a scoring of 4 on the ITU scale ("perceptible but not annoying") is acceptable in most systems.

Whilst the above embodiment describes the implementation of the present
25 invention with respect to one particular watermarking scheme, it will be appreciated that the invention can in fact be implemented using many other types of watermarking schemes.

For instance, one type of audio watermarking scheme is to use temporal correlation techniques to embed the desired data (e.g. copyright information) into the audio
signal.

30 This technique is effectively an echo-hiding algorithm, in which the strength of the echo is determined by solving a quadratic equation. The quadratic equation is generated by auto-correlation values at two positions: one at delay equal to τ , and one at delay equal to 0. In such a scheme, as echoes of the audio signal are added to the original audio signal, the resulting signal is in fact both an amplitude and a phase modulated version

of the original audio signal. At the detector, the watermark is extracted by determining the ratio of the auto correlation function at the two delay positions.

Also known are watermarking schemes based on the amplitude modulation of DFT (Discrete Fourier Transform) co-efficients, that require the calculation of DFT's at both
5 the encoder and the decoder.

Similarly, WO 98/53565, US 6,175,627 and WO 00/00969 describe alternative techniques, to which the present invention could be applied, for embedding or encoding auxiliary signals (such as copyright information) into a multimedia host or cover signal. As detailed in WO 00/00969, a replica of the cover signal, or a portion of the cover signal in a
10 particular domain (time, frequency or space), is generated according to a stego key, which specifies modification values to the parameters of the cover signal. The replica signal is then modified by an auxiliary signal corresponding to the information to be embedded, and inserted back into the cover signal so as to form the stego signal.

At the decoder, in order to extract the original auxiliary data, a replica of the
15 stego signal is generated in the same manner as the replica of the original cover signal, and requires the use of the same stego key. The resulting replica is then correlated with the received stego signal, so as to extract the auxiliary signal.

Using an alternative embodiment of the present invention, the extracted auxiliary signal can be replaced by a new one. This can be achieved by appropriately
20 subtracting the auxiliary information from the received signal using the stego key and the embedding parameters estimated using the detection unit. In relation to figure 1, this can be put into effect by utilizing a detector 640, a watermark generator 650, and a watermark re-embedder 620 responsive to the underlying embedding algorithm.

It will be appreciated by the skilled person that various implementations not
25 specifically described would be understood as falling within the scope of the present invention. For instance, whilst only the functionality of the embedding and detecting apparatus has been described, it will be appreciated that the apparatus could be realized as a digital circuit, an analog circuit, a computer program, or a combination thereof.

Within the specification it will be appreciated that the word "comprising" does
30 not exclude other elements or steps, that "a" or "an" does not exclude a plurality, and that a single processor or other unit may fulfil the functions of several means recited in the claims.

CLAIMS:

1. A method of processing a multimedia signal comprising a watermark signal, the method comprising the steps of:
removing at least a portion of an original watermark signal; and
adding a new watermark signal to the multimedia signal so as to form a new watermarked
5 multimedia signal.
2. A method as claimed in claim 1 wherein said original watermark signal is removed by applying a negative version of said original watermark signal to the multimedia signal.
- 10 3. A method as claimed in claim 1, further comprising the step of determining the value of at least one of the parameters used to embed the original watermark in the multimedia signal.
- 15 4. A method as claimed in claim 3, wherein said value of said parameter is utilized to remove at least a portion of said original watermark signal.
5. A method as claimed in claim 3, wherein said new signal is embedded in the multimedia signal using said embedding parameters having said determined value.
- 20 6. A method as claimed in claim 3, wherein said new signal is embedded in the multimedia signal using said embedding parameters having values other than the said determined values.
- 25 7. A method as claimed in claim 3, wherein said parameter comprises at least one of: embedding strength, synchronization information, time offset, time-scaling, an amount of a circular shift of a sequence, and a watermark symbol period.

8. A method as claimed in claim 1, wherein all of said original watermark is removed, and the new signal comprises a new watermark signal.

9. A method as claimed in claim 1, wherein said watermark comprises at least two sequences of values, at least one sequence of values being removed as said portion of the watermark signal, so as to leave at least one remaining sequence of values from the original watermark signal; and wherein said new signal comprises at least one further sequence of values which together with said remaining sequence forms a new watermark signal.

10

10. A method as claimed in claim 9, wherein all of said sequences of values are formed from a single sequence of values which has been circularly shifted by different amounts.

15

11. A method as claimed in claim 1, wherein said removed portion of the watermark signal was embedded with a predetermined strength into said multimedia signal, the method comprising the step of embedding the new signal into the multimedia signal with preferably the same predetermined strength.

20

12. A method as claimed in claim 11, wherein the embedding strength is such that the degradation in the quality of the new watermarked multimedia signal is perceptible but not annoying.

25

13. A method as claimed in claim 1, wherein at least one of the original watermark signal and the new watermark signal comprises a smoothly varying signal formed by applying a window shaping function to a sequence of values, the integral over the window shaping function being zero.

30

14. A method as claimed in claim 13, wherein the window shaping function has a bi-phase behavior.

15. A method as claimed in claim 14, wherein the bi-phase window comprises at least two Hanning windows of opposite polarities.

16. A computer program arranged to perform the method of claim 1.
17. A record carrier comprising a computer program as claimed in claim 16.
- 5 18. A method of making available for downloading a computer program as claimed in claim 16.
19. An apparatus for processing a multimedia signal comprising a watermark signal, the apparatus comprising:
- 10 a deletion unit arranged to remove at least a portion of the watermark signal; and
an embedder arranged to add a new signal to the multimedia signal so as to form a new watermark signal.
20. An apparatus as claimed in claim 19, further comprising a detector arranged to
15 detect at least one value of a parameter of said watermark signal.
21. A receiver of a multimedia signal comprising an apparatus as claimed in claim 19.

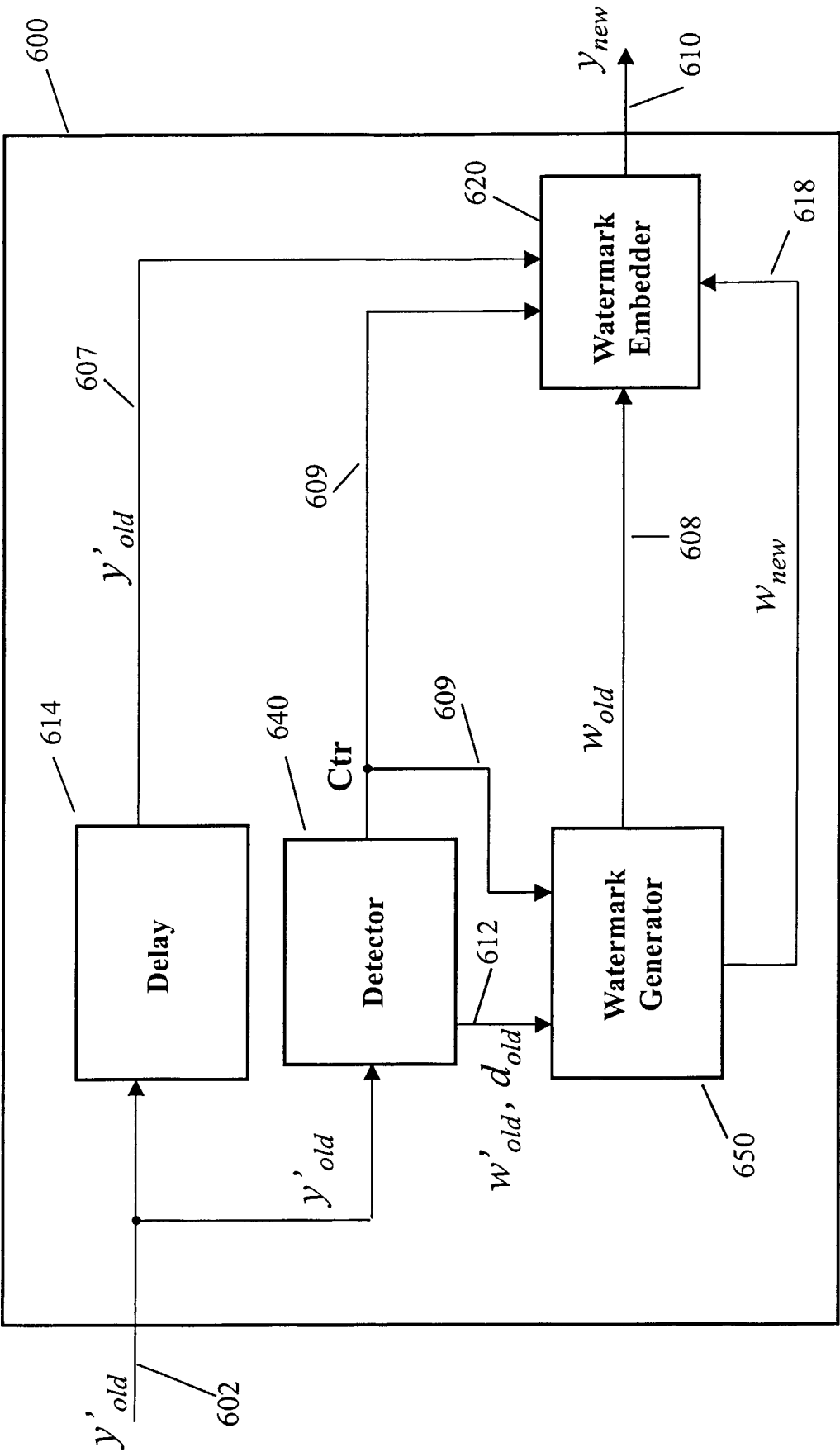


FIG.1

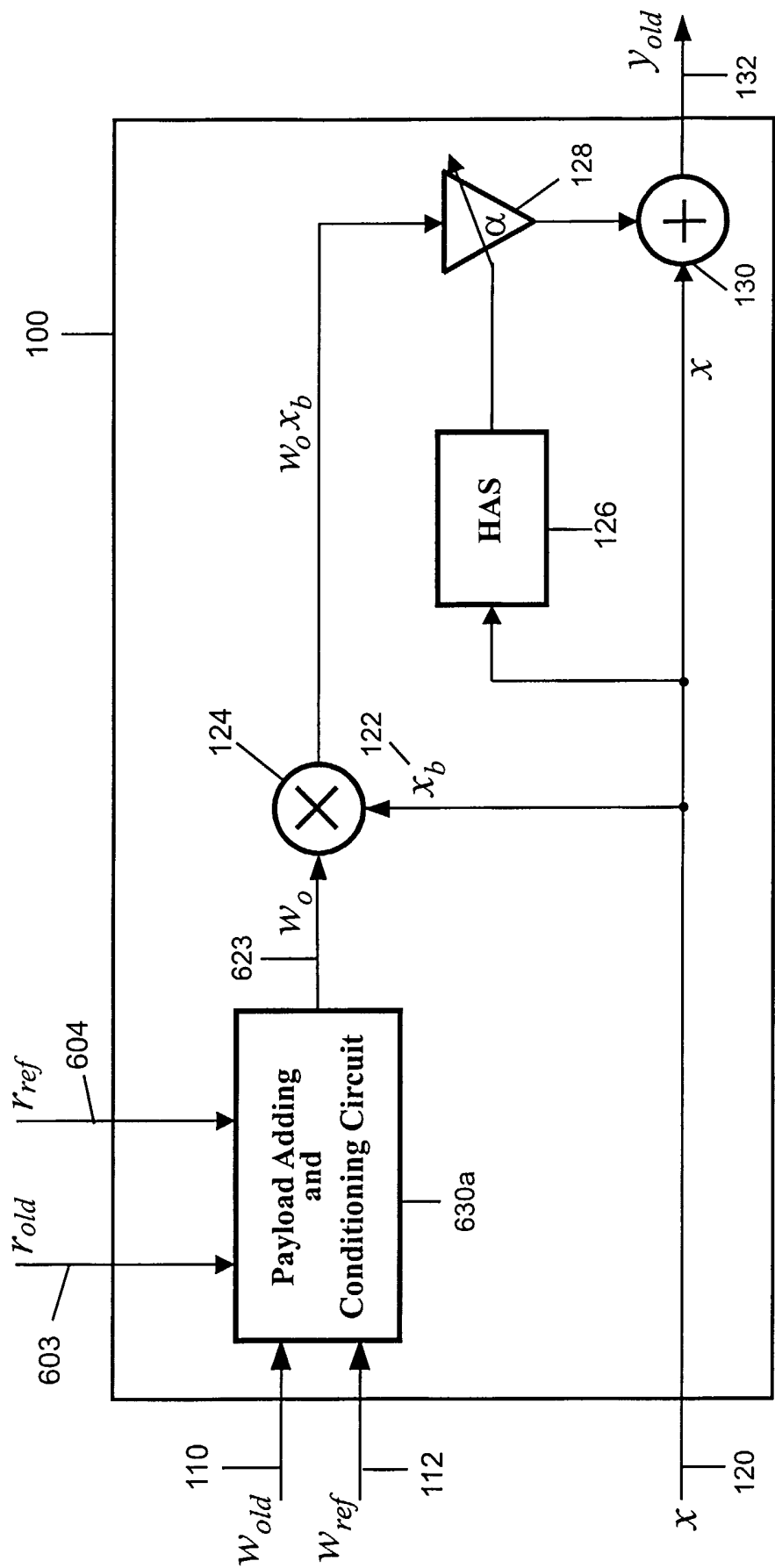


FIG.2

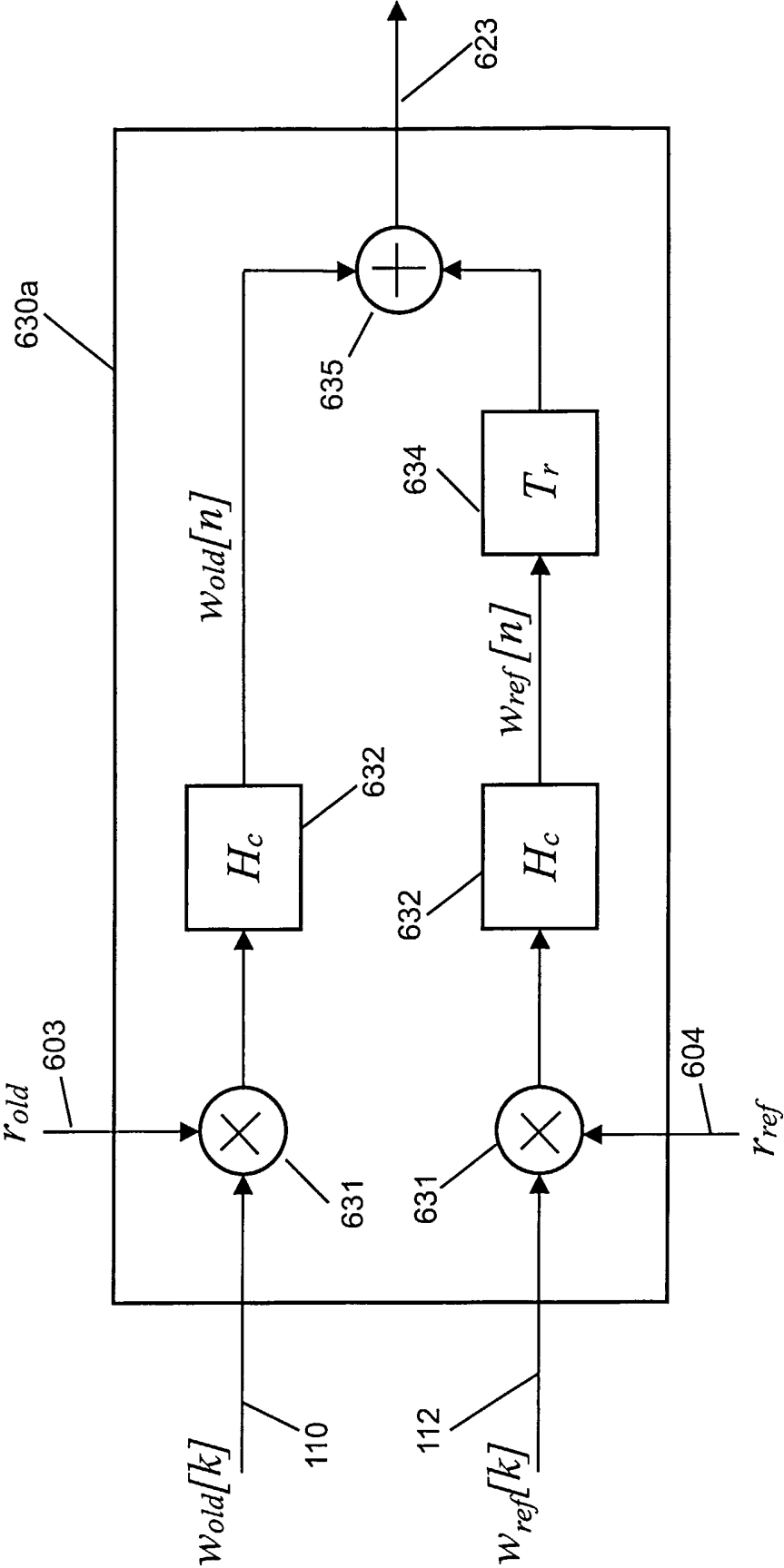


FIG.3

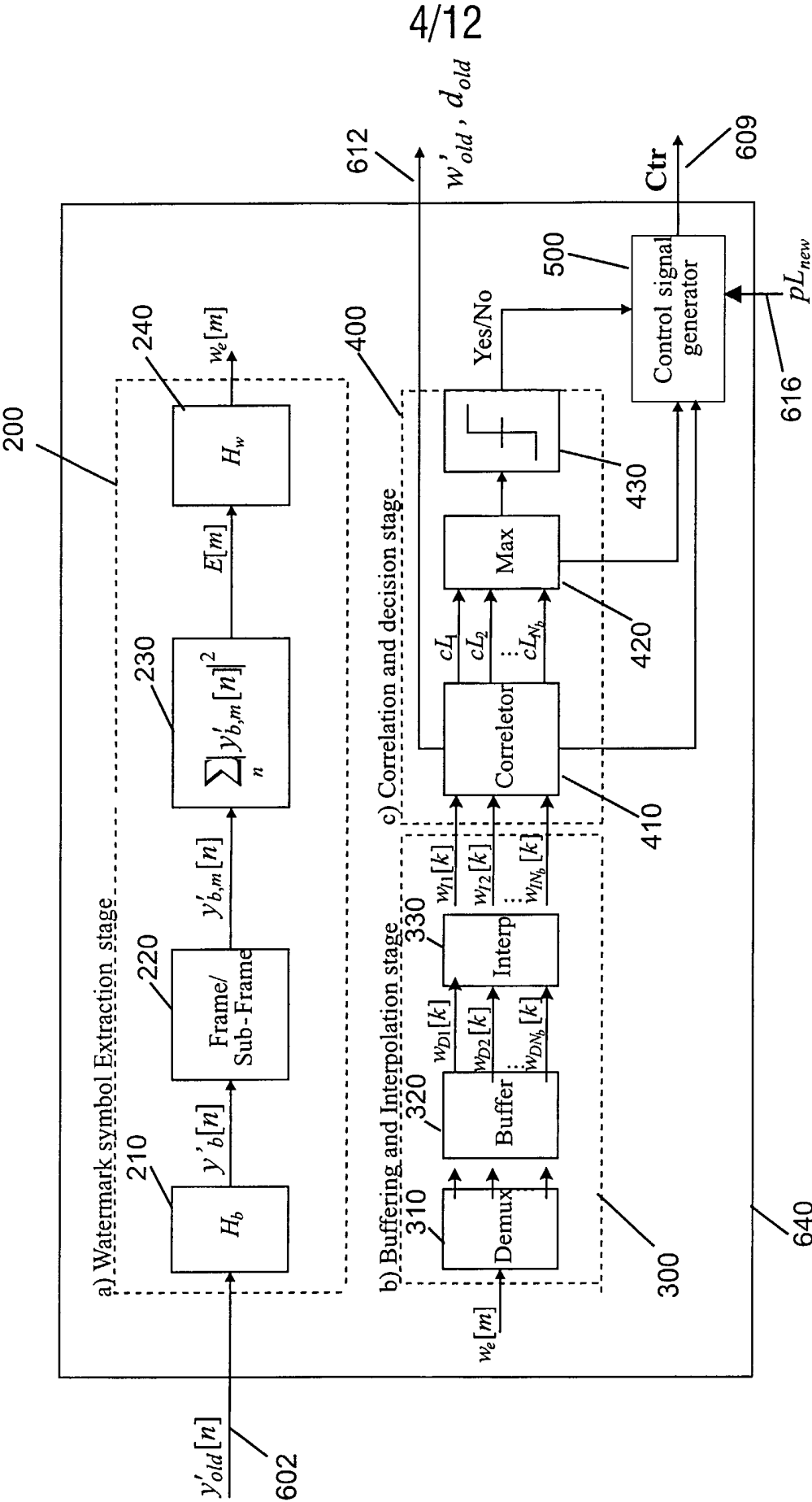


FIG.4

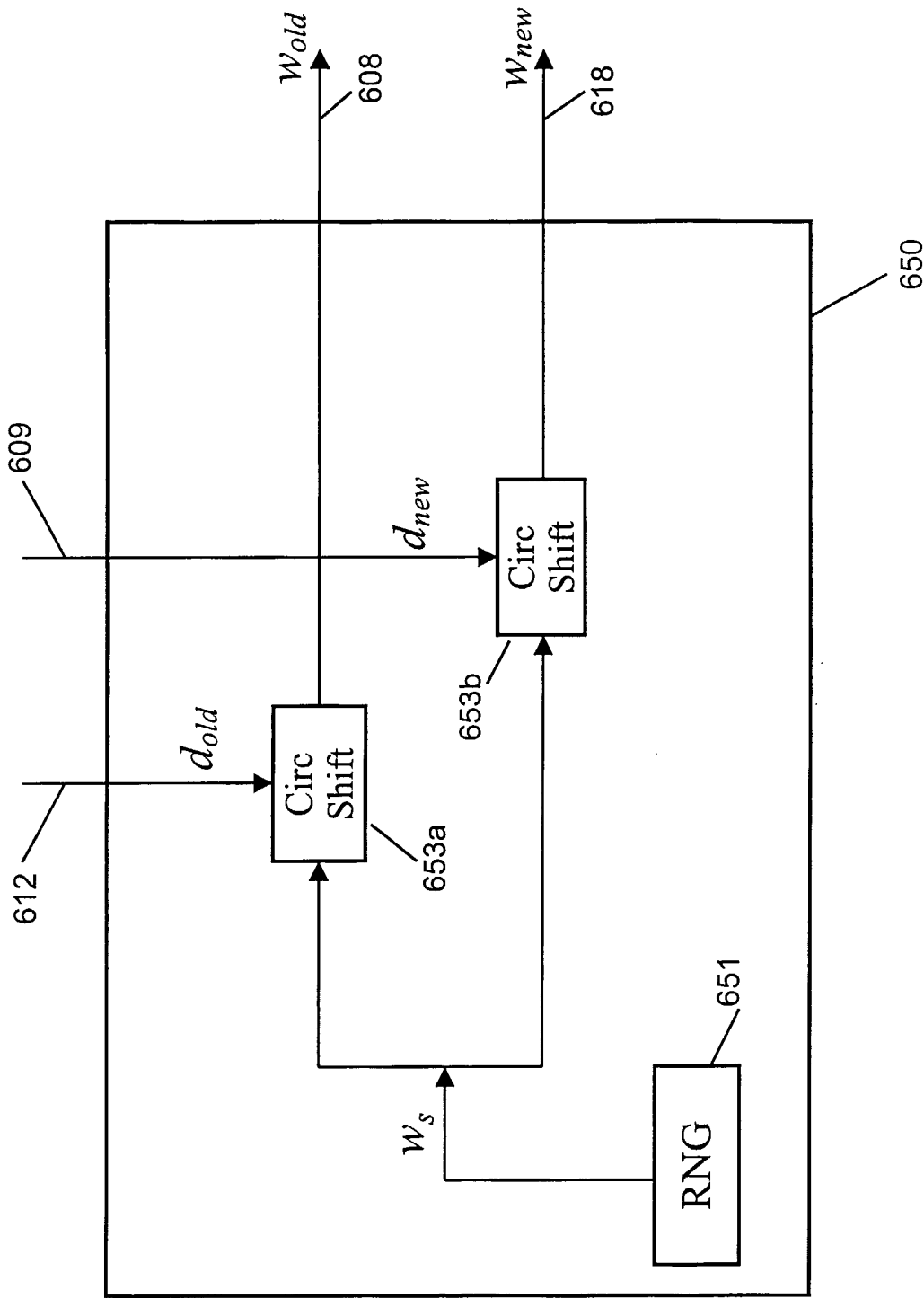


FIG.5

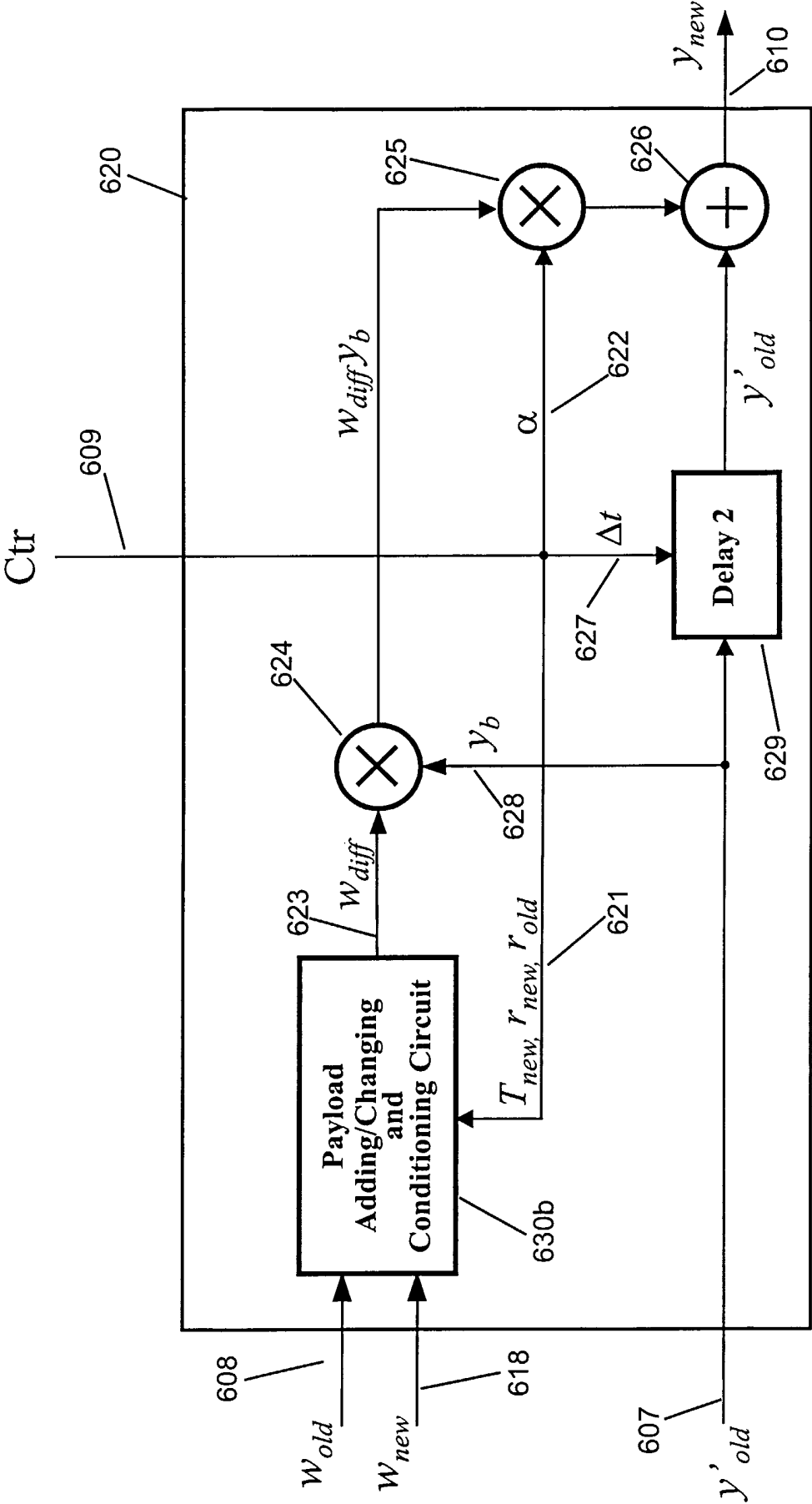


FIG.6

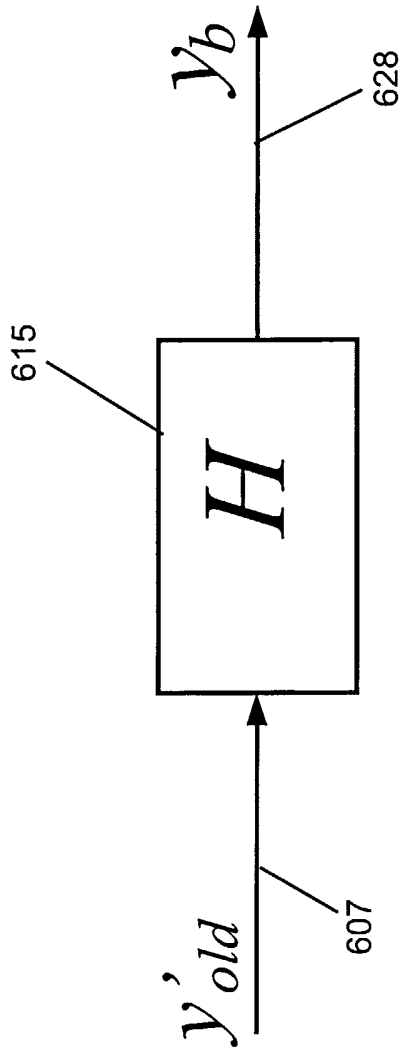


FIG.7

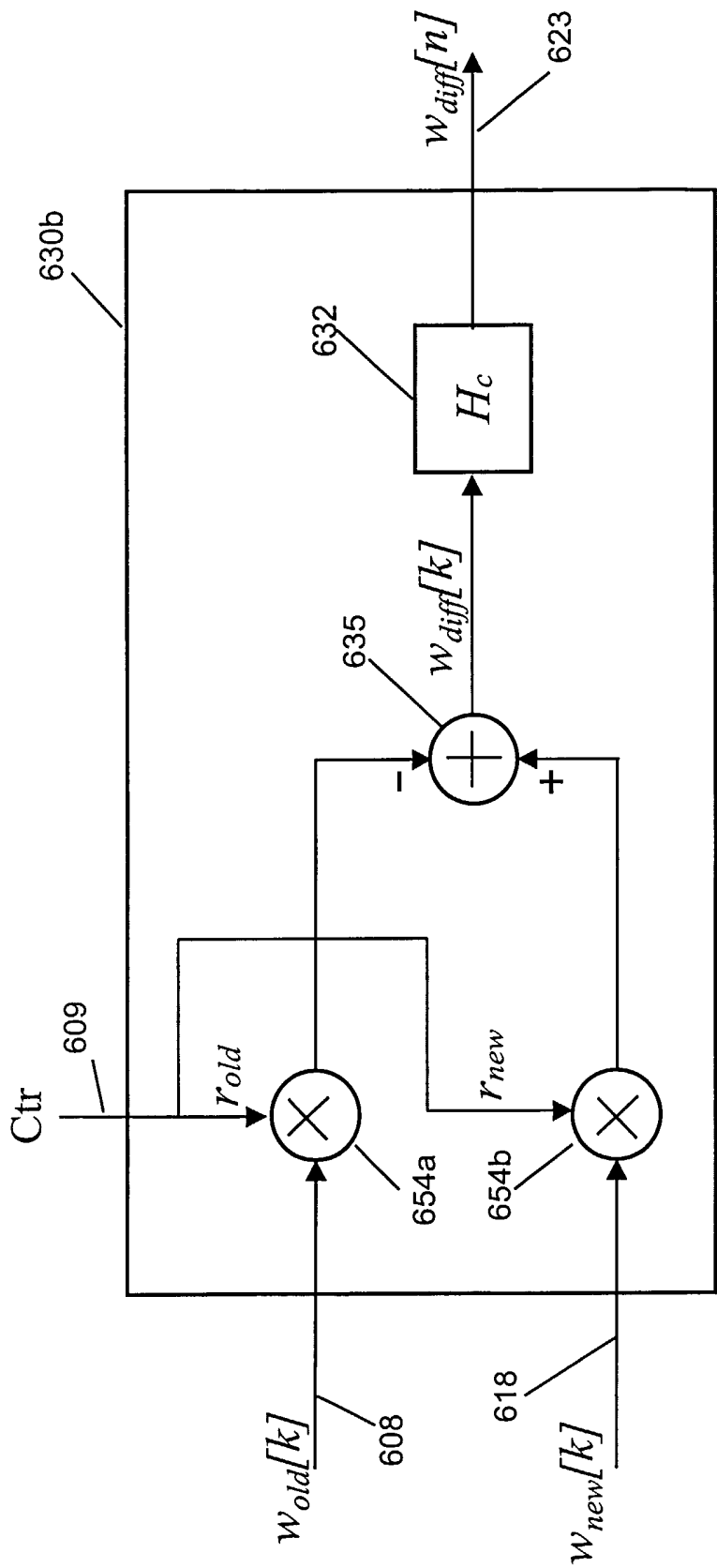


FIG. 8

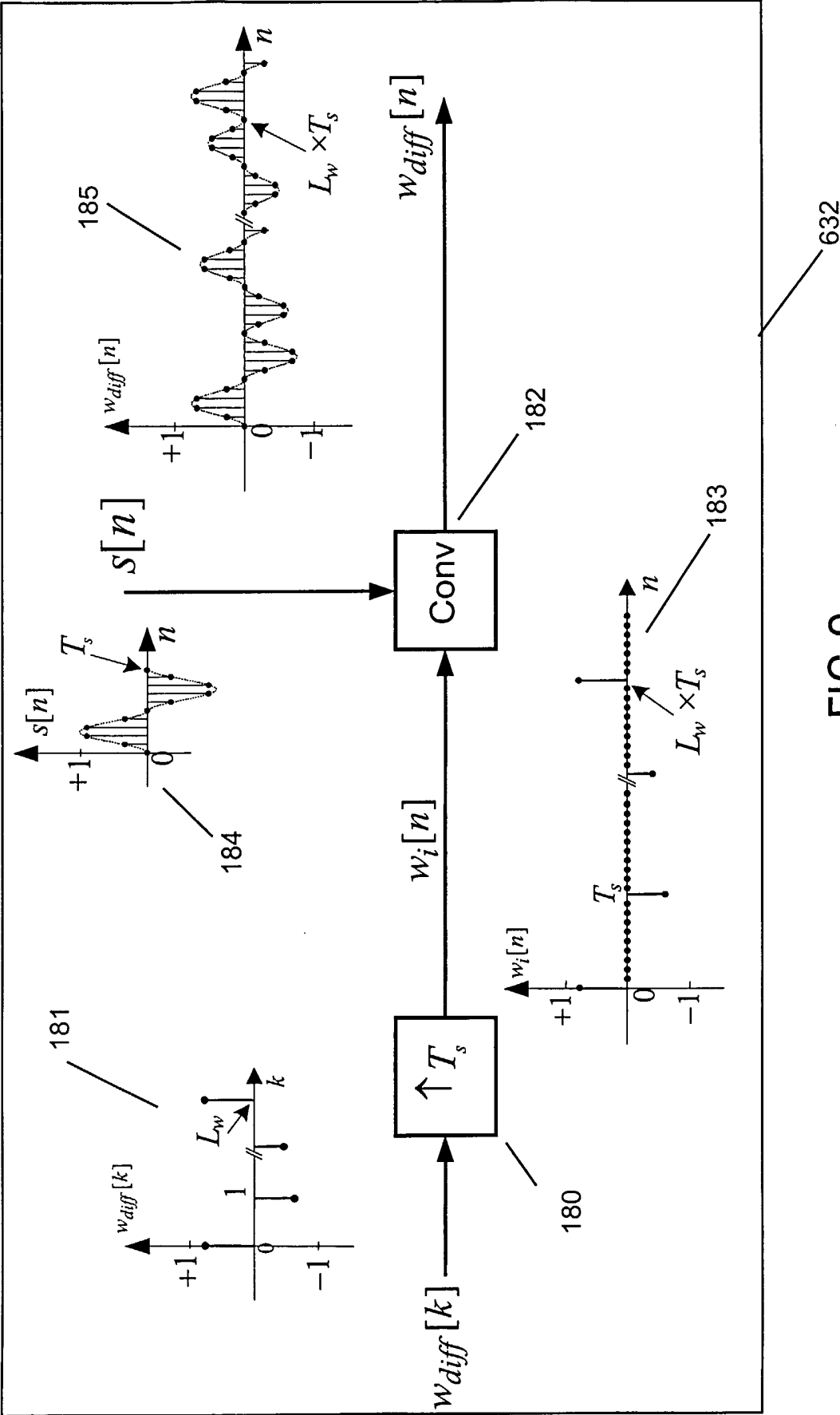


FIG.9

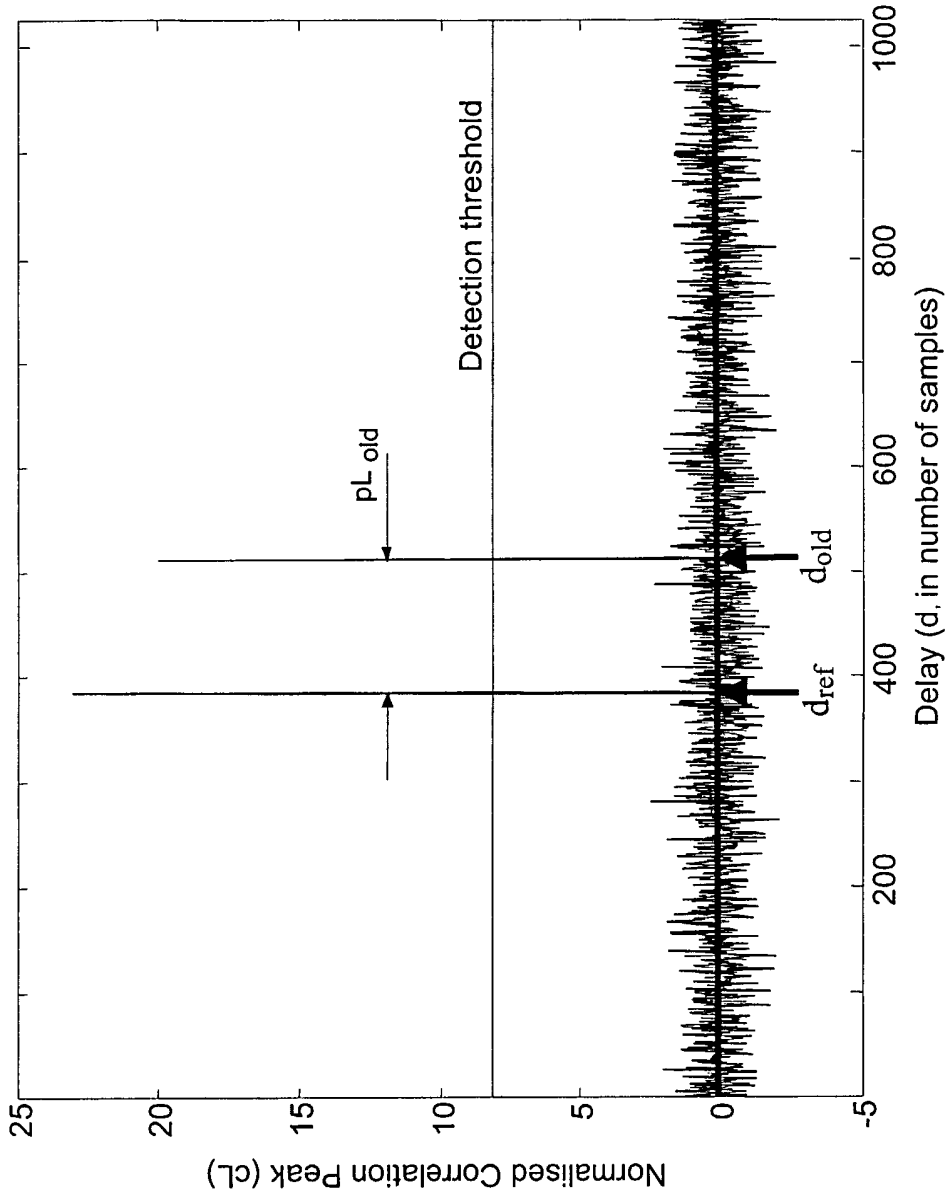


FIG.10

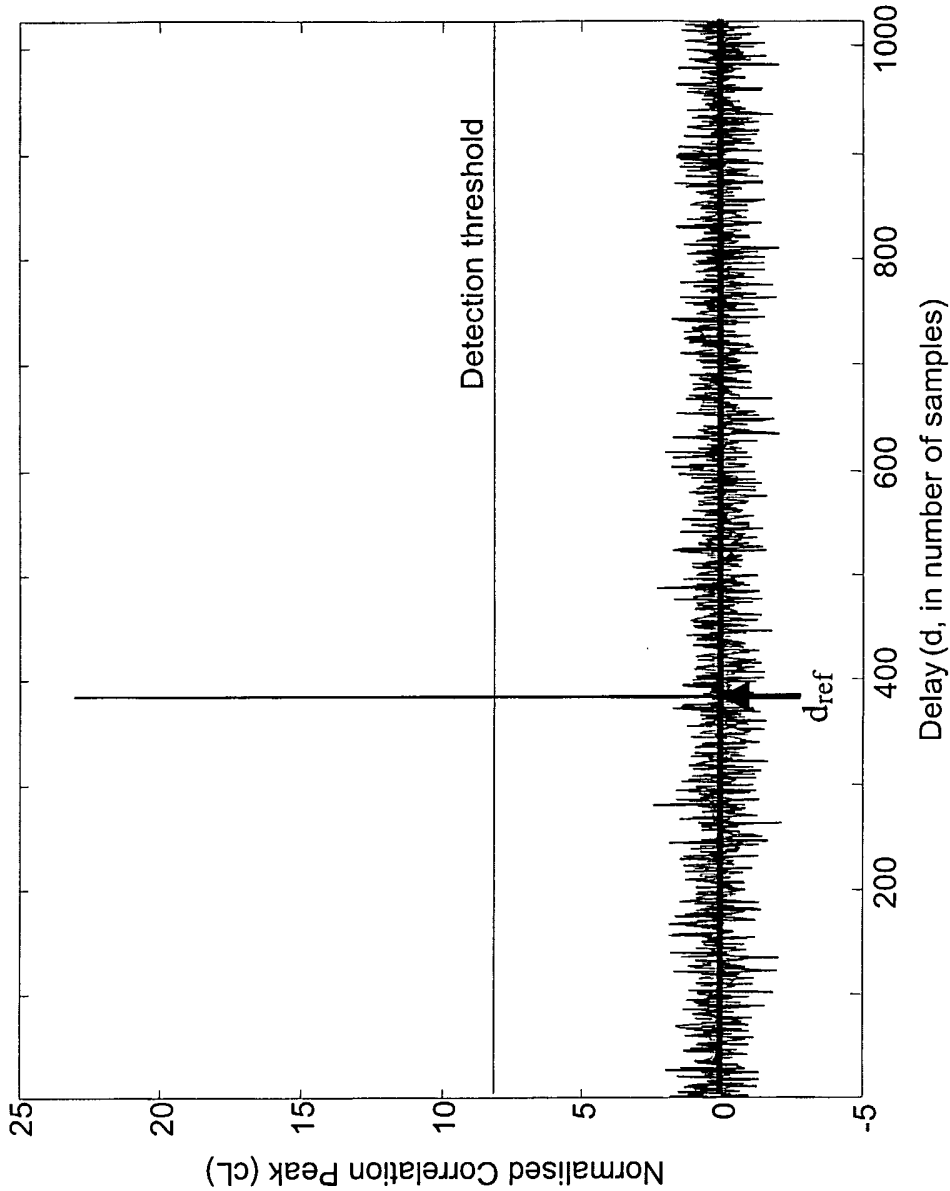


FIG.11

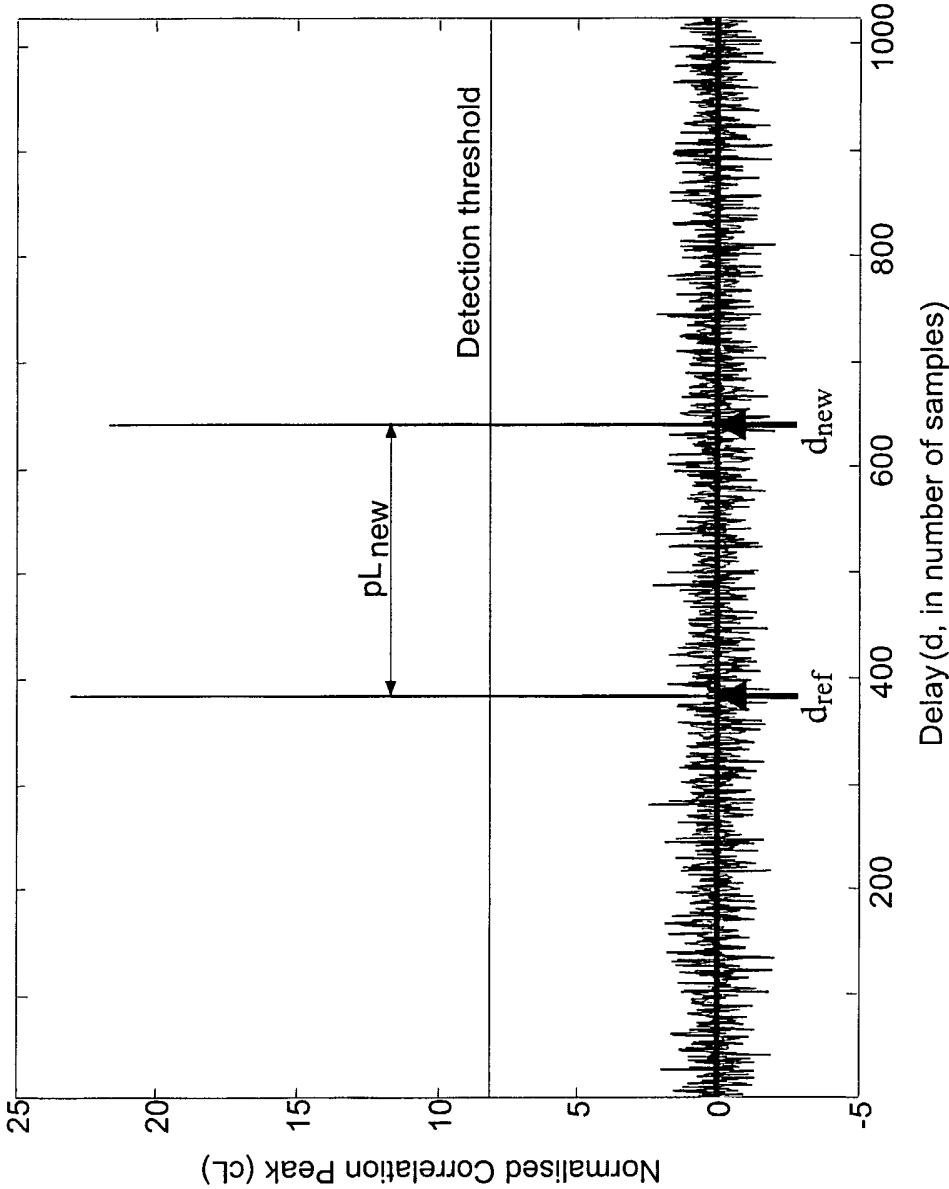


FIG.12

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 03/02319

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G11B20/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G11B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2002/149976 A1 (SAKO YOICHIRO) 17 October 2002 (2002-10-17) abstract; figures 2,3A-B,5-8 paragraphs '0001!, '0006!, '0009!, '0011!-'0014! & WO 02 11140 A 7 February 2002 (2002-02-07) ---	1,3,6, 16-21 9-12
X	PIVA A ET AL: "Managing copyright in open networks" IEEE INTERNET COMPUTING, MAY-JUNE 2002, IEEE, USA, vol. 6, no. 3, pages 18-26, XP002255464 ISSN: 1089-7801 page 18, left-hand column, line 1 -right-hand column, line 11 * p.20 section "reversibility" ---	1,2,8, 16-19,21
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

24 September 2003

Date of mailing of the international search report

09/10/2003

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

INTERNATIONAL SEARCH REPORT

Int. Patent Application No.
PCT/IB 03/02319

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	VEEN VAN DER M ET AL: "ROBUST, MULTI-FUNCTIONAL AND HIGH-QUALITY AUDIO WATERMARKING TECHNOLOGY" PREPRINTS OF PAPERS PRESENTED AT THE AES CONVENTION, vol. 110, no. 5345, 12 May 2001 (2001-05-12), pages 1-9, XP001086463 the whole document ---	9-11
Y	WO 02 25662 A (KERKHOF LEON M VAN DE ;LEMMMA AWEKE N (NL); APREA JAVIER F (NL); KO) 28 March 2002 (2002-03-28) abstract ---	12
T	LEMMMA A N ET AL: "A temporal domain audio watermarking technique" IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE, INC. NEW YORK, US, vol. 51, no. 4, April 2003 (2003-04), pages 1088-1097, XP002242882 ISSN: 1053-587X the whole document -----	9-11, 13-15

INTERNATIONAL SEARCH REPORT

Int [REDACTED] International Application No
PCT/IB 03/02319

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002149976 A1	17-10-2002	WO 0211140 A1	07-02-2002
WO 0225662 A	28-03-2002	AU 2354402 A	02-04-2002
		CN 1395733 T	05-02-2003
		CZ 20021733 A3	14-08-2002
		WO 0225662 A1	28-03-2002
		EP 1323161 A2	02-07-2003
		US 2002172116 A1	21-11-2002