

5

This Sheet Has Been Replaced



REPLACED SHEETS		SG
☒	Description : paged ...1... to paged ...14...	
☒	Chapter(s) : paged ...15... to paged ...17...	
☐	Figure(s) : paged ... to paged ...	
☒	Appendix : paged ...18... to paged ...	

This invention relates to digital signal encoders.

For efficient coding of audio signals or the like, there are coding techniques based on bit assignment, wherein an input audio signal is divided into signal components of a plurality of channels on the time axis or the frequency axis, and the number of bits is adaptively allocated to respective channels. For example, there are coding techniques such as band division coding (sub-band coding: SBC) to divide an audio signal on the time axis into signal components in a plurality of frequency bands to carry out coding thereof; adaptive transform coding (ATC) for transforming (orthogonal-transforming) a signal on the time axis to a signal on the frequency axis thereby to divide that signal into signal components in a plurality of frequency bands adaptively to carry out coding in respective bands; adaptive bit allocation (APC-AB) in which SBC and adaptive predictive coding (APC) are combined to divide a signal on the time axis into signal components in a plurality of bands to transform signal components in respective bands to those in a base band (low frequency band), thereafter to carry out a linear predictive analysis; and the like.

Efficient coding techniques in which the masking characteristic in the hearing of a human being is taken into consideration have been attempted. The masking effect is the phenomenon that a signal is masked by another signal, so it cannot be heard. The masking effect may occur on the time axis or on the frequency axis.

The masking phenomenon on the frequency axis is the effect that a signal component in a frequency band is masked by a signal component in another frequency band, so the former signal component cannot be heard. As the masking effect on the time axis, there are the temporal masking effect and the simultaneous masking effect. The simultaneous masking effect is the effect that a small sound occurring simultaneously with a loud sound is masked by the loud sound, so it cannot be heard. The temporal masking effect is the effect that small sounds preceding and succeeding a loud sound are masked by the loud sound, so they cannot be heard. The masking backward in time by the loud sound is called forward masking, and the masking forward in time by the loud sound is called backward masking. In temporal masking, from the viewpoint of human hearing, the effect of forward masking is effective for a long time (for example, about 100 msec), whereas the

S-2803022

effect of backward masking is effective for a short time (for example, about 5 msec). In addition, the level of the masking effect is about 20 dB for forward masking, and is about 30 dB for backward masking. As stated above, since sound subjected to masking cannot be heard, then
5 even if the number of bits allocated for quantization of a masked signal component is reduced in quantization of an audio signal, perceived deterioration can be lessened.

For more efficient coding the bit compression ratio must be increased. However, generally, in carrying out bit compression making
10 use of the masking effect, only the masking effect in signals on the frequency axis or the masking effect in signals on the time axis can be used.

In addition, in reducing the bit rate due to the masking effect on the time axis, it is desirable for effectively utilizing the masking
15 effect on the time axis to allow the processing time block length to be long. However, this makes real time processing difficult.

According to the present invention there is provided a digital signal encoder comprising:

first frequency analysis means of the non-block type for carrying out
20 frequency analysis of an input digital signal;

second frequency analysis means of the block type for further carrying out frequency analysis of frequency components analyzed by said first frequency analysis means;

first noise level setting means for setting a first allowable noise
25 level on the basis of the energy of each frequency component within a predetermined time period of outputs from said second frequency analysis means;

second noise level setting means for setting a second allowable noise level of each frequency component within said predetermined time period
30 on the basis of the energy of at least one of data preceding in time and data succeeding in time outputs from said second frequency analysis means; and

quantization means for quantizing an output from said second frequency analysis means;

35 a quantization characteristic of said quantization means being set on the basis of outputs from said first and second noise level setting means.





Embodiments of the invention can provide a digital signal encoder which can carry out real time processing with little deterioration in sound quality, by effectively utilizing both the masking effect with respect to signals on the frequency axis and the masking effect with respect to signals on the time axis, thus permitting the bit rate to be reduced.

The invention will now be described by way of example with reference to the accompanying drawings, throughout which like parts are referred to by like references, and in which:

10 Figure 1 is a block diagram of a digital signal encoder according to this invention;

Figure 2 is a block circuit of a filter bank;

Figure 3 is a diagram for explaining blocks and predetermined time periods in the embodiment;

15 Figure 4 is a diagram showing the bark spectrum;

Figure 5 is a circuit diagram showing the filter circuit;

Figure 6 is a diagram showing a masking spectrum; and

Figure 7 is a diagram in which the minimum audible curve and the masking spectrum are synthesized.

20 An embodiment of digital signal encoder comprises, as shown in Figure 1, a filter bank 10 serving as first frequency analysis means of the non-block type for carrying out frequency analysis of an input digital signal supplied to an input terminal 1, fast Fourier transform (FFT) circuits 20, 40 and 60 serving as second frequency analysis means of the block type for further carrying out frequency analysis of frequency components analyzed by the filter bank 10, and quantization circuits 29, 49 and 69 for quantizing respective outputs (FFT coefficient data) from the FFT circuits 20, 40 and 60. A masking spectrum calculation circuit 75 serves as first noise level setting means for setting a first allowable noise level on the basis of energies of respective frequency components within predetermined time periods (for example, time periods B1 to B3 ... of Figure 3 which will be described later) of outputs from the FFT circuits 20, 40 and 60; and time period delay circuits 23, 43 and 63, synthesis circuits 50 and 70, 30
35 a 5 ms delay (DL) circuit 51, 2.5 ms delay circuits 71, 72 and 73, a selection circuit 52, a synthesis/selection circuit 74, and weighting synthesis circuits 24, 44 and 64 serve as second noise level setting



means for setting second allowable noise levels of respective frequency components for the predetermined time periods on the basis of an energy of at least one of data preceding in time and data succeeding in time the predetermined time periods, thus to set quantization circuits 29, 5 49 and 69 on the basis of outputs from the first and second noise level setting means. The filter bank 10 comprises band division means for dividing the input digital signal into signal components in a plurality of frequency bands (three bands in this embodiment), including at least one filter (for example, a mirror filter such as a so-called QMF). The 10 FFT circuits 20, 40 and 60 carry out fast Fourier transformation of signals every block. The band division means (filter bank 10) carries out a division such that the bandwidth becomes larger as the frequency band shifts to a higher frequency. The FFT circuits 20, 40 and 60 carry out processing such that the block time length where the fast 15 Fourier transform processing is carried out becomes smaller as the frequency band shifts to a higher frequency. At least one datum at the second noise level setting means is caused to be data preceding in time. In addition, the second noise level setting means sets the second noise level on the basis of energies of data which are the same 20 in terms of the frequency components for periods different in time. An output from the encoder is derived from an output terminal 2.

An input digital audio signal of DC to 22 KHz obtained by sampling, for example at the sampling frequency $f_s = 44.1$ KHz is supplied to the input terminal 1 and thence to the filter bank 10. The 25 filter bank 10 includes, as shown in Figure 2, two stages of QMFs 11 and 12 cascade-connected. The input digital audio signal is supplied to the QMF 11 which divides it into signal components in two frequency bands meeting at 11 KHz. Accordingly, output signals in the frequency bands of DC to 11 KHz and 11 to 22 KHz are provided from the QMF 11. 30 The output signal 11 to 22 KHz is supplied to the FFT circuit 60 through a terminal 13. On the other hand, the output signal DC to 11 KHz is supplied to the QMF 12 which divides an input signal into signal components in two frequency bands meeting at 5.5 KHz. Accordingly, output signals in the frequency bands of DC to 5.5 KHz and 35 5.5 to 11 KHz are provided from the QMF 12. The signal 5.5 to 11 KHz is supplied to the FFT circuit 40 through a terminal 14, and the signal DC to 5.5 KHz is supplied to the FFT circuit 20 through a terminal 15.



Thus, in the filter bank 10 serving as the first frequency analysis means, an input digital sound signal is divided into signal components in three frequency bands on a non-block basis so that the bandwidth becomes larger as the frequency band shifts to a higher frequency. It is to be noted that while the filter comprises QMFs in the example of Figure 2, there may be employed a structure using band-pass filters.

In the FFT circuit 20 supplied with a signal in the frequency band of DC to 5.5 KHz from the filter bank 10, the signal is divided into signal components in the form of blocks every 10 ms to carry out FFT processing every block. In the FFT circuit 40 supplied with a signal in the frequency band of 5.5 to 11 KHz, FFT processing on a block basis every 5 ms is carried out. In the FFT circuit 60 supplied with a signal in the frequency band of 11 to 22 KHz, FFT processing on a block basis every 2.5 ms is carried out. Thus, in the FFT circuits 20, 40 and 60 serving as the second frequency analysis means, processing such that the block length subjected to FFT becomes smaller as the frequency band shifts to a higher frequency is carried out. As stated above, in this embodiment, in forming blocks at respective FFT circuits 20, 40 and 60, the time block length is smaller at a higher frequency band thereby to allow the time resolution at the higher frequency band to be high and allow the time resolution at the lower frequency band to be low to increase the number of samples within one block, thus allowing the frequency resolution to be high. Thus, since ordinary audio signals have short steady state periods in a high frequency band, it is appropriate that the time resolution in a high frequency band is caused to be high as described above. Moreover, since the frequency resolution of human hearing is generally high in a low frequency band, it is also appropriate that the frequency resolution in a low frequency band is caused to be high as described above.

The time block of the processing by the filter bank 10 and the respective FFT circuits 20, 40 and 60 is shown in Figure 3. Thus, in Figure 3, there are shown the band division and respective processing units (blocks) of FFT, etc, and respective blocks are designated by three parameters p, q and r, shown b (p, q, r) in the figure, where p represents elapsed time, q represents a band, and r represents a time block. In Figure 3, it is indicated that one time block in each band



has a time length (time resolution) of 10 ms in the lower frequency band DC to 5.5 KHz. It is also indicated that one time block length is 5 ms in the medium frequency band 5.5 to 11 KHz, and that one time block length is 2.5 ms in the high frequency band 11 to 22 KHz.

5 FFT coefficient data of every band obtained by FFT processing in the FFT circuits 20, 40 and 60 are supplied to the quantization circuits 29, 49 and 69, where they are quantized. At this time, for example, in every respective predetermined time period B1 to B3 ... of Figure 3, quantization processing is carried out. In this
10 quantization, there is carried out an adaptive quantization in which the quantization characteristic (bit allocation) of quantization is changed on the basis of outputs from the first and second noise level setting means obtained in consideration of the masking effect with respect to signals on the frequency axis, and the masking effect with
15 respect to signals on the time axis as described later. It is here noted that the respective predetermined time periods B1 to B3 ... are set to 10 ms which is the minimum unit of the processing at the FFT circuit 20.

20 The first and second allowable noise levels for carrying out such an adaptive quantization are determined in practice as follows.

 Output data of the FFT circuits 20, 40 and 60 are further divided into data in so-called critical bands. The critical band here mentioned is a band in which human hearing is taken into consideration. More particularly, when a pure sound is masked by a narrow band noise
25 having the same intensity including a pitch of the pure sound, the critical band refers to a band of a noise such that the bandwidth becomes broad as the frequency band shifts to a higher frequency. In this embodiment, such a critical band division is employed thereby to allow the bandwidth to be broader as the frequency band shifts to a
30 higher frequency, thus to divide the band concerned into, for example, 25 bands. To carry this out, output data (frequency band DC to 5.5 KHz) from the FFT circuit 20 are further divided into data in, for example, 20 bands on the side of the low frequency band of the critical band by the critical band division circuit 21. Moreover, output data
35 (5.5 to 11 KHz) from the FFT circuit 40 are further divided into data in, for example, 3 bands in a medium frequency band of the critical band by the critical band division circuit 41. In addition, output



data (11 to 22 KHz) from the FFT circuit 60 are further divided into data in, for example, two bands in a high frequency band of the critical band by the critical band division circuit 61.

5 Outputs from the critical band division circuits 21, 41 and 61 are supplied to energy detection circuits 22, 42 and 62, respectively. In the respective energy detection circuits 22, 42 and 62, energies (spectrum intensities in respective bands) of data every respective time block, and respective critical bands at the FFT circuits 20, 40 and 60 are determined, for example, by taking a sum total of amplitude
10 values in respective bands (the peak or average of the amplitude value, or energy sum total). Outputs from the energy detection circuits 22, 42 and 62, that is, the spectrum of the sum total every respective critical band is generally called a bark spectrum. The bark spectrum SB in each band is as shown in Figure 4, for example. It is to be
15 noted that the above-mentioned critical band is represented as 12 bands for brevity of explanation in Figure 4.

Here, in order to allow the influence in the masking on the frequency axis of the bark spectrum SB to be taken into consideration, a predetermined weighting function is convoluted onto the bark spectrum
20 SB. To realize this, outputs from the energy detection circuits 22, 42 and 46, that is, respective values of the bark spectrum SB are supplied to the masking spectrum calculation circuit 75 serving as the first noise level setting means. The masking spectrum calculation circuit 75 includes a filter circuit 76, a function generation circuit 77, a
25 subtracter 78, and a divider 79 respectively corresponding to the energy detection circuits 22, 42 and 62. Accordingly, the outputs from each of the energy detection circuits 22, 42 and 62 is supplied to the filter circuit 76. The filter circuit 76 is, as shown in Figure 5, for example, composed of delay (Z^{-1}) elements $\dots 101_{n-3}$ to $101_{n+3} \dots$ for
30 delaying input data in sequence, multipliers $\dots 102_{n-3}$ to $102_{n+3} \dots$ for multiplying outputs from respective delay elements by a filter coefficient (weighting function), and a sum total adder 104. At this time, at the respective multipliers 102_{n-3} to 102_{n+3} , the convolution processing of the bark spectrum SB is carried out as follows. For
35 example, at the multiplier 102_{n-3} , outputs from respective delay elements are multiplied by the filter coefficient 0.0000086; at the



multiplier 102_{n-2} , those outputs are multiplied by the filter coefficient 0.0019; at the multiplier 102_{n-1} , those outputs are multiplied by the filter coefficient 0.15; at the multiplier 102_n , those outputs are multiplied by the filter coefficient 1; at the multiplier 102_{n+1} , those outputs are multiplied by the filter coefficient 0.4; at the multiplier 102_{n+2} , those outputs are multiplied by the filter coefficient 0.06; and at the multiplier 102_{n+3} , those outputs are multiplied by the filter coefficient 0.007. By the convolution processing mentioned above, the sum total of the portions indicated by dotted lines in Figure 4, added by the adder 104 is taken. The sum total thus provided is outputted from the output terminal 105.

Meanwhile, in the level a corresponding to the first allowable noise in the case of calculating the masking spectrum of the bark spectrum SB (allowable noise spectrum), if the level a is small, the masking spectrum (masking curve) with respect to signals on the frequency base will be lowered. As a result, the number of bits allocated in quantization at the quantization circuits 29, 49 and 69 must be increased. In contrast, if the above-mentioned level a is large, the masking spectrum will be raised. As a result, the number of bits allocated in quantization can be decreased. It is to be noted that the level a corresponding to the first allowable noise level refers to a level such that there results the first allowable noise level in every band of the critical band, by carrying out the deconvolution processing as described later. Moreover, generally in audio signals, etc, the spectrum intensity (energy) in the high frequency band portion is small. Accordingly, in this embodiment, a technique is adopted taking the above circumstances into consideration to allow the level a to be larger as the frequency band shifts to a higher frequency where the energy is small, thus to decrease the number of bits allocated at the high frequency band portion. From the above point of view, at the masking spectrum calculation circuit 75, the above-mentioned level a with respect to the same energy is set to a higher value as the frequency becomes higher.

Thus, in the encoder of this embodiment, the approach employed is to calculate the level a corresponding to the first allowable noise level, to effect a control such that the level a becomes high as the



frequency band shifts to a higher frequency. To realize this, an output from the filter circuit 76 is supplied to the subtracter 78 which serves to determine the level a in the convoluted region. Here, an allowable function (a function representing the masking level) for
5 determining the level a is supplied to the subtracter 78. By increasing and decreasing the allowable function, control of the level a is carried out. This allowable function is supplied from the function generation function 77.

When the number given in sequence from the low frequency band of
10 the critical band is assumed to be 1, the level a corresponding to the allowable noise level can be determined in accordance with the following equation (1):

$$a = S - (n - ai) \quad \dots\dots (1)$$

In equation (1), n and a are constants, respectively, where $a >$
15 0, S is the intensity of the convolution-processed bark spectrum, and $(n - ai)$ is an allowable function. Here, as described above, since a method of decreasing the number of bits from the high frequency band having less energy is advantageous to reduce the total number of bits, n is set to 38, and a is set to 1 in this embodiment. There is then no
20 deterioration in sound quality, and satisfactory coding is carried out.

In this way, the above-mentioned level a is determined. These data are transmitted to the divider 79 which carries out deconvolution of the level a in the above-mentioned convolution-processed region. Accordingly, by carrying out the deconvolution processing, a masking
25 spectrum can be provided from the level a . Namely, this masking spectrum can be provided from the level a . Namely, this masking spectrum becomes an allowable noise spectrum determined for every band. It is to be noted that while the above-mentioned deconvolution processing requires a complicated calculation, the simplified divider
30 79 is used in this embodiment to carry out the deconvolution.

In this encoder, the number of bits allocated for quantization in which the above-described masking on the frequency base is taken into consideration is determined, and a second allowable noise level of each frequency component for a predetermined time period is set on the basis
35 of an energy of at least one of data preceding in time and data succeeding in time for the predetermined period of outputs from the FFT circuits 20, 40 and 60. Thus, assuming that the above-mentioned



predetermined period is B2 of Figure 3, the data preceding in time becomes data for a predetermined time period B1, and the data succeeding in time becomes data for a predetermined time period B3. On the basis of at least one of data for these predetermined periods B1 and B3, an allowable noise level (masking level) with respect to each frequency component within the predetermined period B2 is set. Further, at least one data at the second noise level setting means is caused to be data preceding in time. Namely, by taking into consideration the forward masking in which the time of the masking effect is long in the temporal masking, on the basis of data for the predetermined time period B1 preceding in time, an allowable noise level for a predetermined time period B2 is determined. Further, the second noise level setting means sets the second allowable noise level on the basis of energies of data which are the same in terms of the frequency components for periods different in time. Namely, the second allowable noise level is set on the basis of energies of data preceding and succeeding in time of the same frequency band of the critical bands.

In other words, in the second noise level setting means, by taking into consideration the temporal masking by signals preceding and succeeding (for preceding and succeeding predetermined time periods B1 and B3) on the time base adjacent on the time base to a signal at a current time point (for example, a predetermined time B2) of an arbitrary band where the first allowable noise level is set by the masking spectrum calculation circuit 75 with respect to the signal at the current time point (the predetermined time B2) of the arbitrary band, an allowable noise level (second allowable noise level) for an arbitrary band at the current time point (predetermined time B2) is set. To realize this, each of outputs from the energy detection circuits 22, 42 and 62 is supplied to the time period delay circuits 23, 43, 63 and the 5 ms delay circuit 51, and the 2.5 ms delay circuit 71 of the second noise level setting means.

Here, these period delay circuits 23, 43 and 63 serve to delay data respectively supplied thereto every time period of, for example, 10 ms which is the processing unit for the predetermined time period. Further, outputs from the time period delay circuits 43 and 63 are supplied to synthesis circuits 50 and 70, respectively. The synthesis



circuits 50 and 70 synthesize data of the time blocks (5 ms, 2.5 ms blocks) at the FFT circuits 40 and 60 so that they become in correspondence with data of 10 ms, respectively. Further, the 5 ms delay circuit 51 carries out delay every 5 ms block. An output from the 5 ms delay circuit 51 is supplied to the selection circuit 52. The selection circuit 52 carries out a switching selection such that in the case where data of the 5 ms block supplied thereto are the forward block data within a predetermined time period being processed, those data are caused to be passed therethrough, and in the case where data of the 5 ms block supplied thereto are the backward block data within a preceding time period of the predetermined time period, those data are not caused to be passed therethrough. Thus, assuming that the predetermined period being processed is the period B2 of Figure 3, the selection circuit 52 carries out a selection such that when 5 ms block data supplied thereto are the block b (2, 2, 1) in Figure 3, that circuit is turned ON, and when the above data are the block b (1, 2, 2), that circuit is turned OFF. The 2.5 ms delay circuit 71 carries out delay every block of 2.5 ms. Outputs from the 2.5 ms delay circuit 71 are supplied in sequence to the 2.5 ms delay circuits 72 and 73. Each of the outputs from the 2.5 ms delay circuits 71, 72 and 73 is supplied to the synthesis/selection circuit 74. The synthesis/selection circuit 74 carries out a switching selection such that in the case where 2.5 ms block data supplied thereto are the backward block b (1, 3, 4) within a preceding predetermined time period of, for example, the predetermined period B2, that circuit is turned OFF, and in the case where those data are the blocks b (2, 3, 1), b (2, 3, 2) and b (2, 3, 3) within the predetermined time period B2, that circuit is turned ON. At the same time, when, for example, data of the block b (2, 3, 2) are supplied to the synthesis/selection circuit 74, the synthesis/selection circuit 74 synthesizes this block and the forward block (2, 3, 1); when data of the block b (2, 3, 3) are supplied to the synthesis/selection circuit 74, it synthesizes this block and the forward two blocks b (2, 3, 1) and b (2, 3, 2); and when data of the block b (2, 3, 4) are supplied to the synthesis/selection circuit 74, it synthesizes this block and the forward three blocks b (2, 3, 1), b (2, 3, 2) and b (2, 3, 3).

An output from the time period delay circuit 23 is supplied to



the weighting synthesis circuit 24, outputs from the synthesis circuit 50 and the selection circuit 52 are supplied to the weighting synthesis circuit 44, and outputs from the synthesis circuit 70 and the synthesis/selection circuit 74 are supplied to the weighting synthesis circuit 64. Further, data from the masking spectrum calculation circuit 75 are also supplied to respective weighting synthesis circuits 24, 44 and 64. Here, respective weighting synthesis circuits 24, 44 and 64 synthesize weighting coefficients in which the masking effects on the frequency base and the time base are taken into consideration with respect to data supplied thereto. Namely, these weighting coefficients are coefficients set by taking the masking effect into consideration. For example, in the case where a signal for a predetermined time period and time blocks preceding or succeeding a signal for a current predetermined period and respective time blocks is normalized, so that it takes a value of 1, a weighting coefficient corresponding to a level exerted on the signal for the current predetermined time period and the time blocks based on the masking (masking and temporal masking with respect to a signal on the frequency base, or the like) on the frequency axis and the time axis by the signal for the preceding or succeeding predetermined time period and time blocks is weighted with respect to the signal for the preceding or succeeding predetermined time period and time blocks. Thus, the allowable noise level (masking spectrum) utilizing the masking effects on the frequency axis and the time axis can be set.

It is to be noted that while the above-described masking spectrum in which the masking effect is taken into consideration is determined within the same critical band, that spectrum is permitted to be a spectrum in which masking between other critical bands is taken into consideration.

Outputs from the weighting synthesis circuits 24, 44 and 64 are further supplied through the synthesis circuits 25, 45 and 65 to subtractors 27, 47 and 67, respectively. To the subtractors 27, 47 and 67, outputs from the respective energy detection circuits 22, 42 and 62, that is, the previously described bark spectrum SB are also supplied through the delay circuits 31, 56 and 81, respectively. Accordingly, when subtractive calculations between the masking spectrum and the bark spectrum SB of the first and second allowable noise levels



are carried out at the subtracters 27, 47 and 67, the portions of the bark spectrum SB below the level indicated by each level of the masking spectrum MS are masked as shown in Figure 6.

5 Outputs from the subtracters 27, 47 and 67 are supplied to the
quantization circuits 29, 49 and 69 through ROMs 28, 48 and 68,
respectively. The ROMs 28, 48 and 68 store therein allocation bit
number information in quantization at the quantization circuits 29, 49
and 69 to output allocation bit number information corresponding to
10 outputs from the subtracters 27, 47 and 67, respectively. Thus, the
quantization circuits 29, 49 and 69 quantize FFT coefficient data
supplied through the delay circuits 30, 55 and 80 by allocating bit
numbers corresponding to outputs from the respective subtracters 27, 47
and 67. In other words, the quantization circuits 29, 49 and 69
15 quantize components of respective bands by the number of bits allocated
in dependence upon level differences between energies in every band of
the critical band allowable noise levels, in which the masking effects
on the frequency axis and the time axis are taken into consideration.
It is to be noted that this bit allocation is not carried out for the
predetermined time period, and the number of bits used for respective
20 predetermined time periods is determined in advance. Accordingly, bit
allocation is carried out within one predetermined time period. here
the delay circuits 30, 55 and 80 are provided by taking into account
delay quantities at respective circuits succeeding the critical band
division circuits 21, 41 and 61. Further, the delay circuits 31, 56
25 and 81 are provided by taking into consideration delay quantities at
respective circuits succeeding the time period delay circuits 23, 43
and 63, the 5 ms delay circuit 51, the 2.5 ms delay circuit 71, and the
masking spectrum calculation circuit 75.

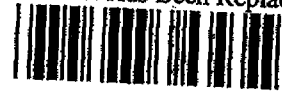
30 In synthesis at the synthesis circuits 25, 45 and 65, it is
possible to synthesize data indicating the so-called minimum audible
curve (equi-loudness curve) RC of human hearing, as shown in Figure 7,
supplied from minimum audible curve generation circuits 26, 46 and 66
and the above-mentioned masking spectrum. Accordingly, by both
synthesizing the minimum audible curve RC and the masking spectrum MS,
35 the allowable noise level can be down to the portion indicated by
slanting lines in the figure. Thus, the number of bits allocated to
the portion indicated by slanting lines in the figure can be reduced in



quantization. Figure 7 also shows a signal spectrum SS at the same time. Further, in this embodiment, there may be employed a configuration such that the above-described synthesis processing of the minimum audible curve is not carried out. In this case, the minimum
5 audible curve generation circuits 26, 46 and 66 and the synthesis circuits 25, 45 and 65 in the configuration of Figure 1 become unnecessary.

Respective outputs from the quantization circuits 29, 49 and 69 quantized by the adaptively allocated bit number in this way are
10 synthesized at a synthesis circuit 90, and an output thus synthesized is then outputted as a coded output from the output terminal 2.

As described above, the effect with the embodiment is that time resolution is improved in a higher frequency band, and the frequency resolution is improved in a lower frequency band. Thus, processing
15 suitable for human hearing can be carried out. In addition, at least one datum at the second noise level setting means is caused to be data preceding in time, and the second noise level setting means sets each second allowable noise level on the basis of energies of data which are the same in terms of the frequency components for periods different in
20 time. Thus, processing can be effected in a short time.

CLAIMS

1. A digital signal encoder comprising:
first frequency analysis means (10) of the non-block type for carrying
5 out frequency analysis of an input digital signal;
second frequency analysis means (20, 40, 60) of the block type for
further carrying out frequency analysis of frequency components
analyzed by said first frequency analysis means (10) ;
first noise level setting means (75) for setting a first allowable
10 noise level on the basis of the energy of each frequency component
within a predetermined time period of outputs from said second
frequency analysis means (20, 40, 60);
second noise level setting means (23, 43, 43, 50, 70, 51, 71, 72, 73,
52, 74, 22, 44, 64) for setting a second allowable noise level of each
15 frequency component within said predetermined time period on the basis
of the energy of at least one of data preceding in time and data
succeeding in time outputs from said second frequency analysis means
(20, 40, 60); and
quantization means (29, 49, 69) for quantizing an output from said
20 second frequency analysis means (20, 40, 60);
a quantization characteristic of said quantization means being set on
the basis of outputs from said first and second noise level setting
means (75, 23, etc.).
- 25 2. An encoder according to claim 1 wherein said first frequency
analysis means (10) comprises band division means for dividing said
input digital signal into signal components in a plurality of frequency
bands, including at least one filter (11, 12), and said second
frequency analysis means (20, 40, 60) comprises fast Fourier transform
30 means for carrying out fast Fourier transformation of signal components
every block.
- 35 3. An encoder according to claim 2 wherein said band division means
carries out a division such that the bandwidth becomes large as the
frequency band shifts to a higher frequency, said fast Fourier
transform means carrying out processing such that the block time length
where said fast Fourier transform processing is conducted becomes small



as the frequency band shifts to a higher frequency.

4. An encoder according to claim 1, claim 2 or claim 3 wherein said
at least one datum at said second noise level setting means (23, etc.)
5 is data preceding in time.

5. An encoder according to any one of the preceding claims wherein
said second noise level setting means (23, etc.) sets said second
allowable noise level on the basis of energies of data which are the
10 same in terms of frequency components for periods different in time.

6. An encoder according to any one of the preceding claims which
comprises a critical band division circuit (21, 41, 61) for further
dividing as a whole an output from said second frequency analysis means
15 (20, 40, 60) into signal components in critical bands.

7. An encoder according to claim 6 wherein said first noise level
setting means (23, etc.) sets said first allowable noise level on the
basis of an energy determined by a sum total of amplitude values in
20 respective bands of said critical bands.

8. An encoder according to claim 6 wherein said second noise level
setting means (23, etc.) sets said second allowable noise level every
said critical band.

25

9. An encoder according to claim 6 wherein said second allowable
noise level in an arbitrary band of said critical bands is set on the
basis of said second allowable noise level in another of said critical
bands.

30

10. An encoder according to claim 8 wherein the quantization
characteristic of said quantization means (29, 49, 69) is set in
dependence upon level differences between said energies in every band
of said critical bands, and said first and second allowable noise
35 levels.

This Sheet Has Been Replaced





11. An encoder according to any one of the preceding claims wherein the quantization characteristic of said quantization means (29, 49, 69) is such that as the frequency band shifts to a higher frequency, the number of bits allocated becomes smaller.

5

12. An encoder according to any one of the preceding claims wherein said first allowable noise level is a masking level on the frequency axis based on human hearing.

10 13. An encoder according to any one of the preceding claims wherein said second allowable noise level is a masking level on the time axis based on human hearing.

14. An encoder according to claims 12 and 13 wherein said first and
15 second allowable noise levels are obtained by synthesizing masking levels on said frequency and time bases and a minimum audible curve based on human hearing.

15. An encoder according to any one of the preceding wherein said at
20 least one filter of said first frequency analysis means (20, 40, 60) is a mirror filter.

16. An encoder according to any one of claims 1 to 14 wherein said at
25 least one filter of said first frequency analysis means (20, 40, 60) is a band-pass filter.

This Sheet Has Been Replaced



ABSTRACTDIGITAL SIGNAL ENCODERS

5 In a digital signal encoder for analyzing an input digital signal
by frequency analysis means (10) of the non-block type and frequency
analysis means (20, 40, 60) of the block type to quantize outputs
obtained, a quantization characteristic (29, 49, 69) is set in the
basis of an allowable noise level based on an energy within a
10 predetermined time period of an output from the block type frequency
analysis means (10) and an allowable noise level based on an energy of
at least one of data preceding and succeeding in time within the
predetermined time period, thereby to provide a real time processing.

15 Figure 1

This Sheet Has Been Replaced



✓
✓
✓
✓

1

1
15
1/4
18
14
12
1/6
—

REPLACEMENT SHEET

25 MAR 1997

25/3/97

DIGITAL SIGNAL ENCODERS

This invention relates to digital signal encoders.

For efficient coding of audio signals or the like, there are coding techniques based on bit assignment, wherein an input audio signal is divided into signal components of a plurality of channels on the time axis or the frequency axis, and the number of bits is adaptively allocated to respective channels. For example, there are coding techniques such as band division coding (sub-band coding: SBC) to divide an audio signal on the time axis into signal components in a plurality of frequency bands to carry out coding thereof; adaptive transform coding (ATC) for transforming (orthogonal-transforming) a signal on the time axis to a signal on the frequency axis thereby to divide that signal into signal components in a plurality of frequency bands adaptively to carry out coding in respective bands; adaptive bit allocation (APC-AB) in which SBC and adaptive predictive coding (APC) are combined to divide a signal on the time axis into signal components in a plurality of bands to transform signal components in respective bands to those in a base band (low frequency band), thereafter to carry out a linear predictive analysis; and the like.

Efficient coding techniques in which the masking characteristic in the hearing of a human being is taken into consideration have been attempted. The masking effect is the phenomenon that a signal is masked by another signal, so it cannot be heard. The masking effect may occur on the time axis or on the frequency axis.

The masking phenomenon on the frequency axis is the effect that a signal component in a frequency band is masked by a signal component in another frequency band, so the former signal component cannot be heard. As the masking effect on the time axis, there are the temporal masking effect and the simultaneous masking effect. The simultaneous masking effect is the effect that a small sound occurring simultaneously with a loud sound is masked by the loud sound, so it cannot be heard. The temporal masking effect is the effect that small sounds preceding and succeeding a loud sound are masked by the loud sound, so they cannot be heard. The masking backward in time by the loud sound is called forward masking, and the masking forward in time by the loud sound is called backward masking. In temporal masking, from the viewpoint of human hearing, the effect of forward masking is effective for a long time (for example, about 100 msec), whereas the

25 MAR 1997

effect of backward masking is effective for a short time (for example, about 5 msec). In addition, the level of the masking effect is about 20 dB for forward masking, and is about 30 dB for backward masking. As stated above, since sound subjected to masking cannot be heard, then
5 even if the number of bits allocated for quantization of a masked signal component is reduced in quantization of an audio signal, perceived deterioration can be lessened.

For more efficient coding the bit compression ratio must be increased. However, generally, in carrying out bit compression making
10 use of the masking effect, only the masking effect in signals on the frequency axis or the masking effect in signals on the time axis can be used.

In addition, in reducing the bit rate due to the masking effect on the time axis, it is desirable for effectively utilizing the masking
15 effect on the time axis to allow the processing time block length to be long. However, this makes real time processing difficult.

According to the present invention there is provided a digital signal encoder comprising:

first frequency analysis means of the non-block type for carrying out
20 frequency analysis of an input digital signal;

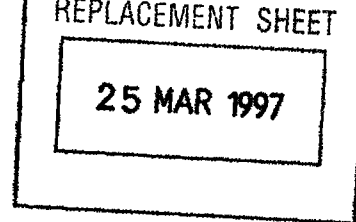
second frequency analysis means of the block type for further carrying out frequency analysis of frequency components analyzed by said first frequency analysis means;

first noise level setting means for setting a first allowable noise
25 level on the basis of the energy of each frequency component within a predetermined time period of outputs from said second frequency analysis means;

second noise level setting means for setting a second allowable noise level of each frequency component within said predetermined time period
30 on the basis of the energy of at least one of data preceding in time and data succeeding in time outputs from said second frequency analysis means; and

quantization means for quantizing an output from said second frequency analysis means;

35 a quantization characteristic of said quantization means being set on the basis of outputs from said first and second noise level setting means.



Embodiments of the invention can provide a digital signal encoder which can carry out real time processing with little deterioration in sound quality, by effectively utilizing both the masking effect with respect to signals on the frequency axis and the masking effect with respect to signals on the time axis, thus permitting the bit rate to be reduced.

The invention will now be described by way of example with reference to the accompanying drawings, throughout which like parts are referred to by like references, and in which:

10 Figure 1 is a block diagram of a digital signal encoder according to this invention;

 Figure 2 is a block circuit of a filter bank;

 Figure 3 is a diagram for explaining blocks and predetermined time periods in the embodiment;

15 Figure 4 is a diagram showing the bark spectrum;

 Figure 5 is a circuit diagram showing the filter circuit;

 Figure 6 is a diagram showing a masking spectrum; and

 Figure 7 is a diagram in which the minimum audible curve and the masking spectrum are synthesized.

20 An embodiment of digital signal encoder comprises, as shown in Figure 1, a filter bank 10 serving as first frequency analysis means of the non-block type for carrying out frequency analysis of an input digital signal supplied to an input terminal 1, fast Fourier transform (FFT) circuits 20, 40 and 60 serving as second frequency analysis means
25 of the block type for further carrying out frequency analysis of frequency components analyzed by the filter bank 10, and quantization circuits 29, 49 and 69 for quantizing respective outputs (FFT coefficient data) from the FFT circuits 20, 40 and 60. A masking spectrum calculation circuit 75 serves as first noise level setting
30 means for setting a first allowable noise level on the basis of energies of respective frequency components within predetermined time periods (for example, time periods B1 to B3 ... of Figure 3 which will be described later) of outputs from the FFT circuits 20, 40 and 60; and time period delay circuits 23, 43 and 63, synthesis circuits 50 and 70,
35 a 5 ms delay (DL) circuit 51, 2.5 ms delay circuits 71, 72 and 73, a selection circuit 52, a synthesis/selection circuit 74, and weighting synthesis circuits 24, 44 and 64 serve as second noise level setting

25 MAR 1997

means for setting second allowable noise levels of respective frequency components for the predetermined time periods on the basis of an energy of at least one of data preceding in time and data succeeding in time the predetermined time periods, thus to set quantization circuits 29, 49 and 69 on the basis of outputs from the first and second noise level setting means. The filter bank 10 comprises band division means for dividing the input digital signal into signal components in a plurality of frequency bands (three bands in this embodiment), including at least one filter (for example, a mirror filter such as a so-called QMF). The FFT circuits 20, 40 and 60 carry out fast Fourier transformation of signals every block. The band division means (filter bank 10) carries out a division such that the bandwidth becomes larger as the frequency band shifts to a higher frequency. The FFT circuits 20, 40 and 60 carry out processing such that the block time length where the fast Fourier transform processing is carried out becomes smaller as the frequency band shifts to a higher frequency. At least one datum at the second noise level setting means is caused to be data preceding in time. In addition, the second noise level setting means sets the second noise level on the basis of energies of data which are the same in terms of the frequency components for periods different in time. An output from the encoder is derived from an output terminal 2.

An input digital audio signal of DC to 22 KHz obtained by sampling, for example at the sampling frequency $f_s = 44.1$ KHz is supplied to the input terminal 1 and thence to the filter bank 10. The filter bank 10 includes, as shown in Figure 2, two stages of QMFs 11 and 12 cascade-connected. The input digital audio signal is supplied to the QMF 11 which divides it into signal components in two frequency bands meeting at 11 KHz. Accordingly, output signals in the frequency bands of DC to 11 KHz and 11 to 22 KHz are provided from the QMF 11. The output signal 11 to 22 KHz is supplied to the FFT circuit 60 through a terminal 13. On the other hand, the output signal DC to 11 KHz is supplied to the QMF 12 which divides an input signal into signal components in two frequency bands meeting at 5.5 KHz. Accordingly, output signals in the frequency bands of DC to 5.5 KHz and 5.5 to 11 KHz are provided from the QMF 12. The signal 5.5 to 11 KHz is supplied to the FFT circuit 40 through a terminal 14, and the signal DC to 5.5 KHz is supplied to the FFT circuit 20 through a terminal 15.

25 MAR 1997

Thus, in the filter bank 10 serving as the first frequency analysis means, an input digital sound signal is divided into signal components in three frequency bands on a non-block basis so that the bandwidth becomes larger as the frequency band shifts to a higher frequency. It is to be noted that while the filter comprises QMFs in the example of Figure 2, there may be employed a structure using band-pass filters.

In the FFT circuit 20 supplied with a signal in the frequency band of DC to 5.5 KHz from the filter bank 10, the signal is divided into signal components in the form of blocks every 10 ms to carry out FFT processing every block. In the FFT circuit 40 supplied with a signal in the frequency band of 5.5 to 11 KHz, FFT processing on a block basis every 5 ms is carried out. In the FFT circuit 60 supplied with a signal in the frequency band of 11 to 22 KHz, FFT processing on a block basis every 2.5 ms is carried out. Thus, in the FFT circuits 20, 40 and 60 serving as the second frequency analysis means, processing such that the block length subjected to FFT becomes smaller as the frequency band shifts to a higher frequency is carried out. As stated above, in this embodiment, in forming blocks at respective FFT circuits 20, 40 and 60, the time block length is smaller at a higher frequency band thereby to allow the time resolution at the higher frequency band to be high and allow the time resolution at the lower frequency band to be low to increase the number of samples within one block, thus allowing the frequency resolution to be high. Thus, since ordinary audio signals have short steady state periods in a high frequency band, it is appropriate that the time resolution in a high frequency band is caused to be high as described above. Moreover, since the frequency resolution of human hearing is generally high in a low frequency band, it is also appropriate that the frequency resolution in a low frequency band is caused to be high as described above.

The time block of the processing by the filter bank 10 and the respective FFT circuits 20, 40 and 60 is shown in Figure 3. Thus, in Figure 3, there are shown the band division and respective processing units (blocks) of FFT, etc., and respective blocks are designated by three parameters p, q and r, shown b (p, q, r) in the figure, where p represents elapsed time, q represents a band, and r represents a time block. In Figure 3, it is indicated that one time block in each band

25 MAR 1997

has a time length (time resolution) of 10 ms in the lower frequency band DC to 5.5 KHz. It is also indicated that one time block length is 5 ms in the medium frequency band 5.5 to 11 KHz, and that one time block length is 2.5 ms in the high frequency band 11 to 22 KHz.

5 FFT coefficient data of every band obtained by FFT processing in the FFT circuits 20, 40 and 60 are supplied to the quantization circuits 29, 49 and 69, where they are quantized. At this time, for example, in every respective predetermined time period B1 to B3 ... of Figure 3, quantization processing is carried out. In this
10 quantization, there is carried out an adaptive quantization in which the quantization characteristic (bit allocation) of quantization is changed on the basis of outputs from the first and second noise level setting means obtained in consideration of the masking effect with respect to signals on the frequency axis, and the masking effect with
15 respect to signals on the time axis as described later. It is here noted that the respective predetermined time periods B1 to B3 ... are set to 10 ms which is the minimum unit of the processing at the FFT circuit 20.

20 The first and second allowable noise levels for carrying out such an adaptive quantization are determined in practice as follows.

 Output data of the FFT circuits 20, 40 and 60 are further divided into data in so-called critical bands. The critical band here mentioned is a band in which human hearing is taken into consideration. More particularly, when a pure sound is masked by a narrow band noise
25 having the same intensity including a pitch of the pure sound, the critical band refers to a band of a noise such that the bandwidth becomes broad as the frequency band shifts to a higher frequency. In this embodiment, such a critical band division is employed thereby to allow the bandwidth to be broader as the frequency band shifts to a
30 higher frequency, thus to divide the band concerned into, for example, 25 bands. To carry this out, output data (frequency band DC to 5.5 KHz) from the FFT circuit 20 are further divided into data in, for example, 20 bands on the side of the low frequency band of the critical band by the critical band division circuit 21. Moreover, output data
35 (5.5 to 11 KHz) from the FFT circuit 40 are further divided into data in, for example, 3 bands in a medium frequency band of the critical band by the critical band division circuit 41. In addition, output

25 MAR 1997

data (11 to 22 KHz) from the FFT circuit 60 are further divided into data in, for example, two bands in a high frequency band of the critical band by the critical band division circuit 61.

5 Outputs from the critical band division circuits 21, 41 and 61 are supplied to energy detection circuits 22, 42 and 62, respectively. In the respective energy detection circuits 22, 42 and 62, energies (spectrum intensities in respective bands) of data every respective time block, and respective critical bands at the FFT circuits 20, 40 and 60 are determined, for example, by taking a sum total of amplitude
10 values in respective bands (the peak or average of the amplitude value, or energy sum total). Outputs from the energy detection circuits 22, 42 and 62, that is, the spectrum of the sum total every respective critical band is generally called a bark spectrum. The bark spectrum SB in each band is as shown in Figure 4, for example. It is to be
15 noted that the above-mentioned critical band is represented as 12 bands for brevity of explanation in Figure 4.

 Here, in order to allow the influence in the masking on the frequency axis of the bark spectrum SB to be taken into consideration, a predetermined weighting function is convoluted onto the bark spectrum
20 SB. To realize this, outputs from the energy detection circuits 22, 42 and 46, that is, respective values of the bark spectrum SB are supplied to the masking spectrum calculation circuit 75 serving as the first noise level setting means. The masking spectrum calculation circuit 75 includes a filter circuit 76, a function generation circuit 77, a
25 subtractor 78, and a divider 79 respectively corresponding to the energy detection circuits 22, 42 and 62. Accordingly, the outputs from each of the energy detection circuits 22, 42 and 62 is supplied to the filter circuit 76. The filter circuit 76 is, as shown in Figure 5, for example, composed of delay (Z^{-1}) elements $\dots 101_{n-3}$ to $101_{n,3}$ $\dots$ for
30 delaying input data in sequence, multipliers $\dots 102_{n-3}$ to $102_{n,3}$ $\dots$ for multiplying outputs from respective delay elements by a filter coefficient (weighting function), and a sum total adder 104. At this time, at the respective multipliers 102_{n-3} to $102_{n,3}$, the convolution processing of the bark spectrum SB is carried out as follows. For
35 example, at the multiplier $102_{n,3}$, outputs from respective delay elements are multiplied by the filter coefficient 0.0000086; at the

25 MAR 1997

multiplier 102_{n-2} , those outputs are multiplied by the filter coefficient 0.0019; at the multiplier 102_{n-1} , those outputs are multiplied by the filter coefficient 0.15; at the multiplier 102_n , those outputs are multiplied by the filter coefficient 1; at the multiplier
5 102_{n+1} , those outputs are multiplied by the filter coefficient 0.4; at the multiplier 102_{n+2} , those outputs are multiplied by the filter coefficient 0.06; and at the multiplier 102_{n+3} , those outputs are multiplied by the filter coefficient 0.007. By the convolution processing mentioned above, the sum total of the portions indicated by
10 dotted lines in Figure 4, added by the adder 104 is taken. The sum total thus provided is outputted from the output terminal 105.

Meanwhile, in the level a corresponding to the first allowable noise in the case of calculating the masking spectrum of the bark spectrum SB (allowable noise spectrum), if the level a is small, the
15 masking spectrum (masking curve) with respect to signals on the frequency base will be lowered. As a result, the number of bits allocated in quantization at the quantization circuits 29, 49 and 69 must be increased. In contrast, if the above-mentioned level a is large, the masking spectrum will be raised. As a result, the number of
20 bits allocated in quantization can be decreased. It is to be noted that the level a corresponding to the first allowable noise level refers to a level such that there results the first allowable noise level in every band of the critical band, by carrying out the deconvolution processing as described later. Moreover, generally in
25 audio signals, etc, the spectrum intensity (energy) in the high frequency band portion is small. Accordingly, in this embodiment, a technique is adopted taking the above circumstances into consideration to allow the level a to be larger as the frequency band shifts to a higher frequency where the energy is small, thus to decrease the number
30 of bits allocated at the high frequency band portion. From the above point of view, at the masking spectrum calculation circuit 75, the above-mentioned level a with respect to the same energy is set to a higher value as the frequency becomes higher.

Thus, in the encoder of this embodiment, the approach employed is
35 to calculate the level a corresponding to the first allowable noise level, to effect a control such that the level a becomes high as the

25 MAR 1997

frequency band shifts to a higher frequency. To realize this, an output from the filter circuit 76 is supplied to the subtracter 78 which serves to determine the level a in the convoluted region. Here, an allowable function (a function representing the masking level) for determining the level a is supplied to the subtracter 78. By increasing and decreasing the allowable function, control of the level a is carried out. This allowable function is supplied from the function generation function 77.

When the number given in sequence from the low frequency band of the critical band is assumed to be 1, the level a corresponding to the allowable noise level can be determined in accordance with the following equation (1):

$$a = S - (n - ai) \quad \dots (1)$$

In equation (1), n and a are constants, respectively, where $a > 0$, S is the intensity of the convolution-processed bark spectrum, and $(n - ai)$ is an allowable function. Here, as described above, since a method of decreasing the number of bits from the high frequency band having less energy is advantageous to reduce the total number of bits, n is set to 38, and a is set to 1 in this embodiment. There is then no deterioration in sound quality, and satisfactory coding is carried out.

In this way, the above-mentioned level a is determined. These data are transmitted to the divider 79 which carries out deconvolution of the level a in the above-mentioned convolution-processed region. Accordingly, by carrying out the deconvolution processing, a masking spectrum can be provided from the level a . Namely, this masking spectrum can be provided from the level a . Namely, this masking spectrum becomes an allowable noise spectrum determined for every band. It is to be noted that while the above-mentioned deconvolution processing requires a complicated calculation, the simplified divider 79 is used in this embodiment to carry out the deconvolution.

In this encoder, the number of bits allocated for quantization in which the above-described masking on the frequency base is taken into consideration is determined, and a second allowable noise level of each frequency component for a predetermined time period is set on the basis of an energy of at least one of data preceding in time and data succeeding in time for the predetermined period of outputs from the FFT circuits 20, 40 and 60. Thus, assuming that the above-mentioned

REPLACEMENT COPY
25 MAR 1997

predetermined period is B2 of Figure 3, the data preceding in time becomes data for a predetermined time period B1, and the data succeeding in time becomes data for a predetermined time period B3. On the basis of at least one of data for these predetermined periods B1 and B3, an allowable noise level (masking level) with respect to each frequency component within the predetermined period B2 is set. Further, at least one data at the second noise level setting means is caused to be data preceding in time. Namely, by taking into consideration the forward masking in which the time of the masking effect is long in the temporal masking, on the basis of data for the predetermined time period B1 preceding in time, an allowable noise level for a predetermined time period B2 is determined. Further, the second noise level setting means sets the second allowable noise level on the basis of energies of data which are the same in terms of the frequency components for periods different in time. Namely, the second allowable noise level is set on the basis of energies of data preceding and succeeding in time of the same frequency band of the critical bands.

In other words, in the second noise level setting means, by taking into consideration the temporal masking by signals preceding and succeeding (for preceding and succeeding predetermined time periods B1 and B3) on the time base adjacent on the time base to a signal at a current time point (for example, a predetermined time B2) of an arbitrary band where the first allowable noise level is set by the masking spectrum calculation circuit 75 with respect to the signal at the current time point (the predetermined time B2) of the arbitrary band, an allowable noise level (second allowable noise level) for an arbitrary band at the current time point (predetermined time B2) is set. To realize this, each of outputs from the energy detection circuits 22, 42 and 62 is supplied to the time period delay circuits 23, 43, 63 and the 5 ms delay circuit 51, and the 2.5 ms delay circuit 71 of the second noise level setting means.

Here, these period delay circuits 23, 43 and 63 serve to delay data respectively supplied thereto every time period of, for example, 10 ms which is the processing unit for the predetermined time period. Further, outputs from the time period delay circuits 43 and 63 are supplied to synthesis circuits 50 and 70, respectively. The synthesis

25 MAR 1997

circuits 50 and 70 synthesize data of the time blocks (5 ms, 2.5 ms blocks) at the FFT circuits 40 and 60 so that they become in correspondence with data of 10 ms, respectively. Further, the 5 ms delay circuit 51 carries out delay every 5 ms block. An output from the 5 ms delay circuit 51 is supplied to the selection circuit 52. The selection circuit 52 carries out a switching selection such that in the case where data of the 5 ms block supplied thereto are the forward block data within a predetermined time period being processed, those data are caused to be passed therethrough, and in the case where data of the 5 ms block supplied thereto are the backward block data within a preceding time period of the predetermined time period, those data are not caused to be passed therethrough. Thus, assuming that the predetermined period being processed is the period B2 of Figure 3, the selection circuit 52 carries out a selection such that when 5 ms block data supplied thereto are the block b (2, 2, 1) in Figure 3, that circuit is turned ON, and when the above data are the block b (1, 2, 2), that circuit is turned OFF. The 2.5 ms delay circuit 71 carries out delay every block of 2.5 ms. Outputs from the 2.5 ms delay circuit 71 are supplied in sequence to the 2.5 ms delay circuits 72 and 73. Each of the outputs from the 2.5 ms delay circuits 71, 72 and 73 is supplied to the synthesis/selection circuit 74. The synthesis/selection circuit 74 carries out a switching selection such that in the case where 2.5 ms block data supplied thereto are the backward block b (1, 3, 4) within a preceding predetermined time period of, for example, the predetermined period B2, that circuit is turned OFF, and in the case where those data are the blocks b (2, 3, 1), b (2, 3, 2) and b (2, 3, 3) within the predetermined time period B2, that circuit is turned ON. At the same time, when, for example, data of the block b (2, 3, 2) are supplied to the synthesis/selection circuit 74, the synthesis/selection circuit 74 synthesizes this block and the forward block (2, 3, 1); when data of the block b (2, 3, 3) are supplied to the synthesis/selection circuit 74, it synthesizes this block and the forward two blocks b (2, 3, 1) and b (2, 3, 2); and when data of the block b (2, 3, 4) are supplied to the synthesis/selection circuit 74, it synthesizes this block and the forward three blocks b (2, 3, 1), b (2, 3, 2) and b (2, 3, 3).

An output from the time period delay circuit 23 is supplied to

25 MAR 1997

the weighting synthesis circuit 24, outputs from the synthesis circuit 50 and the selection circuit 52 are supplied to the weighting synthesis circuit 44, and outputs from the synthesis circuit 70 and the synthesis/selection circuit 74 are supplied to the weighting synthesis circuit 64. Further, data from the masking spectrum calculation circuit 75 are also supplied to respective weighting synthesis circuits 24, 44 and 64. Here, respective weighting synthesis circuits 24, 44 and 64 synthesize weighting coefficients in which the masking effects on the frequency base and the time base are taken into consideration with respect to data supplied thereto. Namely, these weighting coefficients are coefficients set by taking the masking effect into consideration. For example, in the case where a signal for a predetermined time period and time blocks preceding or succeeding a signal for a current predetermined period and respective time blocks is normalized, so that it takes a value of 1, a weighting coefficient corresponding to a level exerted on the signal for the current predetermined time period and the time blocks based on the masking (masking and temporal masking with respect to a signal on the frequency base, or the like) on the frequency axis and the time axis by the signal for the preceding or succeeding predetermined time period and time blocks is weighted with respect to the signal for the preceding or succeeding predetermined time period and time blocks. Thus, the allowable noise level (masking spectrum) utilizing the masking effects on the frequency axis and the time axis can be set.

It is to be noted that while the above-described masking spectrum in which the masking effect is taken into consideration is determined within the same critical band, that spectrum is permitted to be a spectrum in which masking between other critical bands is taken into consideration.

Outputs from the weighting synthesis circuits 24, 44 and 64 are further supplied through the synthesis circuits 25, 45 and 65 to subtractors 27, 47 and 67, respectively. To the subtractors 27, 47 and 67, outputs from the respective energy detection circuits 22, 42 and 62, that is, the previously described bark spectrum SB are also supplied through the delay circuits 31, 56 and 81, respectively. Accordingly, when subtractive calculations between the masking spectrum and the bark spectrum SB of the first and second allowable noise levels

25 MAR 1997

are carried out at the subtracters 27, 47 and 67, the portions of the bark spectrum SB below the level indicated by each level of the masking spectrum MS are masked as shown in Figure 6.

5 Outputs from the subtracters 27, 47 and 67 are supplied to the
quantization circuits 29, 49 and 69 through ROMs 28, 48 and 68,
respectively. The ROMs 28, 48 and 68 store therein allocation bit
number information in quantization at the quantization circuits 29, 49
and 69 to output allocation bit number information corresponding to
10 outputs from the subtracters 27, 47 and 67, respectively. Thus, the
quantization circuits 29, 49 and 69 quantize FFT coefficient data
supplied through the delay circuits 30, 55 and 80 by allocating bit
numbers corresponding to outputs from the respective subtracters 27, 47
and 67. In other words, the quantization circuits 29, 49 and 69
15 quantize components of respective bands by the number of bits allocated
in dependence upon level differences between energies in every band of
the critical band allowable noise levels, in which the masking effects
on the frequency axis and the time axis are taken into consideration.
It is to be noted that this bit allocation is not carried out for the
predetermined time period, and the number of bits used for respective
20 predetermined time periods is determined in advance. Accordingly, bit
allocation is carried out within one predetermined time period. Here
the delay circuits 30, 55 and 80 are provided by taking into account
delay quantities at respective circuits succeeding the critical band
division circuits 21, 41 and 61. Further, the delay circuits 31, 56
25 and 81 are provided by taking into consideration delay quantities at
respective circuits succeeding the time period delay circuits 23, 43
and 63, the 5 ms delay circuit 51, the 2.5 ms delay circuit 71, and the
masking spectrum calculation circuit 75.

30 In synthesis at the synthesis circuits 25, 45 and 65, it is
possible to synthesize data indicating the so-called minimum audible
curve (equi-loudness curve) RC of human hearing, as shown in Figure 7,
supplied from minimum audible curve generation circuits 26, 46 and 66
and the above-mentioned masking spectrum. Accordingly, by both
synthesizing the minimum audible curve RC and the masking spectrum MS,
35 the allowable noise level can be down to the portion indicated by
slanting lines in the figure. Thus, the number of bits allocated to
the portion indicated by slanting lines in the figure can be reduced in

25 MAR 1997

14

quantization. Figure 7 also shows a signal spectrum SS at the same time. Further, in this embodiment, there may be employed a configuration such that the above-described synthesis processing of the minimum audible curve is not carried out. In this case, the minimum
5 audible curve generation circuits 26, 46 and 66 and the synthesis circuits 25, 45 and 65 in the configuration of Figure 1 become unnecessary.

Respective outputs from the quantization circuits 29, 49 and 69 quantized by the adaptively allocated bit number in this way are
10 synthesized at a synthesis circuit 90, and an output thus synthesized is then outputted as a coded output from the output terminal 2.

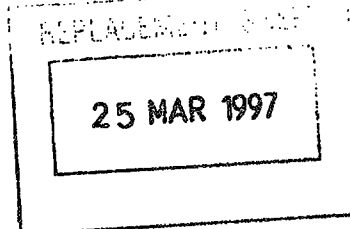
As described above, the effect with the embodiment is that time resolution is improved in a higher frequency band, and the frequency resolution is improved in a lower frequency band. Thus, processing
15 suitable for human hearing can be carried out. In addition, at least one datum at the second noise level setting means is caused to be data preceding in time, and the second noise level setting means sets each second allowable noise level on the basis of energies of data which are the same in terms of the frequency components for periods different in
20 time. Thus, processing can be effected in a short time.



25 MAR 1997

CLAIMS

1. A digital signal encoder comprising:
first frequency analysis means (10) of the non-block type for carrying
out frequency analysis of an input digital signal;
second frequency analysis means (20, 40, 60) of the block type for
further carrying out frequency analysis of frequency components
analyzed by said first frequency analysis means (10) ;
first noise level setting means (75) for setting a first allowable
noise level on the basis of the energy of each frequency component
within a predetermined time period of outputs from said second
frequency analysis means (20, 40, 60);
second noise level setting means (23, 43, 63, 50, 70, 51, 71, 72, 73,
52, 74, 22, 44, 64) for setting a second allowable noise level of each
frequency component within said predetermined time period on the basis
of the energy of at least one of data preceding in time and data
succeeding in time outputs from said second frequency analysis means
(20, 40, 60); and
quantization means (29, 49, 69) for quantizing an output from said
second frequency analysis means (20, 40, 60);
a quantization characteristic of said quantization means being set on
the basis of outputs from said first and second noise level setting
means (75, 23, etc.).
2. An encoder according to claim 1 wherein said first frequency
analysis means (10) comprises band division means for dividing said
input digital signal into signal components in a plurality of frequency
bands, including at least one filter (1, 12), and said second
frequency analysis means (20, 40, 60) comprises fast Fourier transform
means for carrying out fast Fourier transformation of signal components
every block.
3. An encoder according to claim 2 wherein said band division means
carries out a division such that the bandwidth becomes large as the
frequency band shifts to a higher frequency, said fast Fourier
transform means carrying out processing such that the block time length
where said fast Fourier transform processing is conducted becomes small



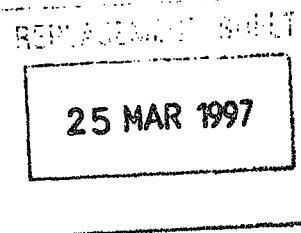
as the frequency band shifts to a higher frequency.

4. An encoder according to claim 1, claim 2 or claim 3 wherein said at least one datum at said second noise level setting means (23, etc.)
5 is data preceding in time.
5. An encoder according to any one of the preceding claims wherein said second noise level setting means (23, etc.) sets said second allowable noise level on the basis of energies of data which are the
10 same in terms of frequency components for periods different in time.
6. An encoder according to any one of the preceding claims which comprises a critical band division circuit (21, 41, 61) for further dividing as a whole an output from said second frequency analysis means
15 (20, 40, 60) into signal components in critical bands.
7. An encoder according to claim 6 wherein said first noise level setting means (23, etc.) sets said first allowable noise level on the basis of an energy determined by a sum total of amplitude values in
20 respective bands of said critical bands.
8. An encoder according to claim 6 wherein said second noise level setting means (23, etc.) sets said second allowable noise level every said critical band.
25
9. An encoder according to claim 6 wherein said second allowable noise level in an arbitrary band of said critical bands is set on the basis of said second allowable noise level in another of said critical bands.
30
10. An encoder according to claim 8 wherein the quantization characteristic of said quantization means (29, 49, 69) is set in dependence upon level differences between said energies in every band of said critical bands, and said first and second allowable noise
35 levels.



25 MAR 1997

11. An encoder according to any one of the preceding claims wherein the quantization characteristic of said quantization means (29, 49, 69) is such that as the frequency band shifts to a higher frequency, the number of bits allocated becomes smaller.
- 5
12. An encoder according to any one of the preceding claims wherein said first allowable noise level is a masking level on the frequency axis based on human hearing.
- 10
13. An encoder according to any one of the preceding claims wherein said second allowable noise level is a masking level on the time axis based on human hearing.
- 15
14. An encoder according to claims 12 and 13 wherein said first and second allowable noise levels are obtained by synthesizing masking levels on said frequency and time bases and a minimum audible curve based on human hearing.
- 20
15. An encoder according to any one of the preceding wherein said at least one filter of said first frequency analysis means (20, 40, 60) is a mirror filter.
- 25
16. An encoder according to any one of claims 1 to 14 wherein said at least one filter of said first frequency analysis means (20, 40, 60) is a band-pass filter.

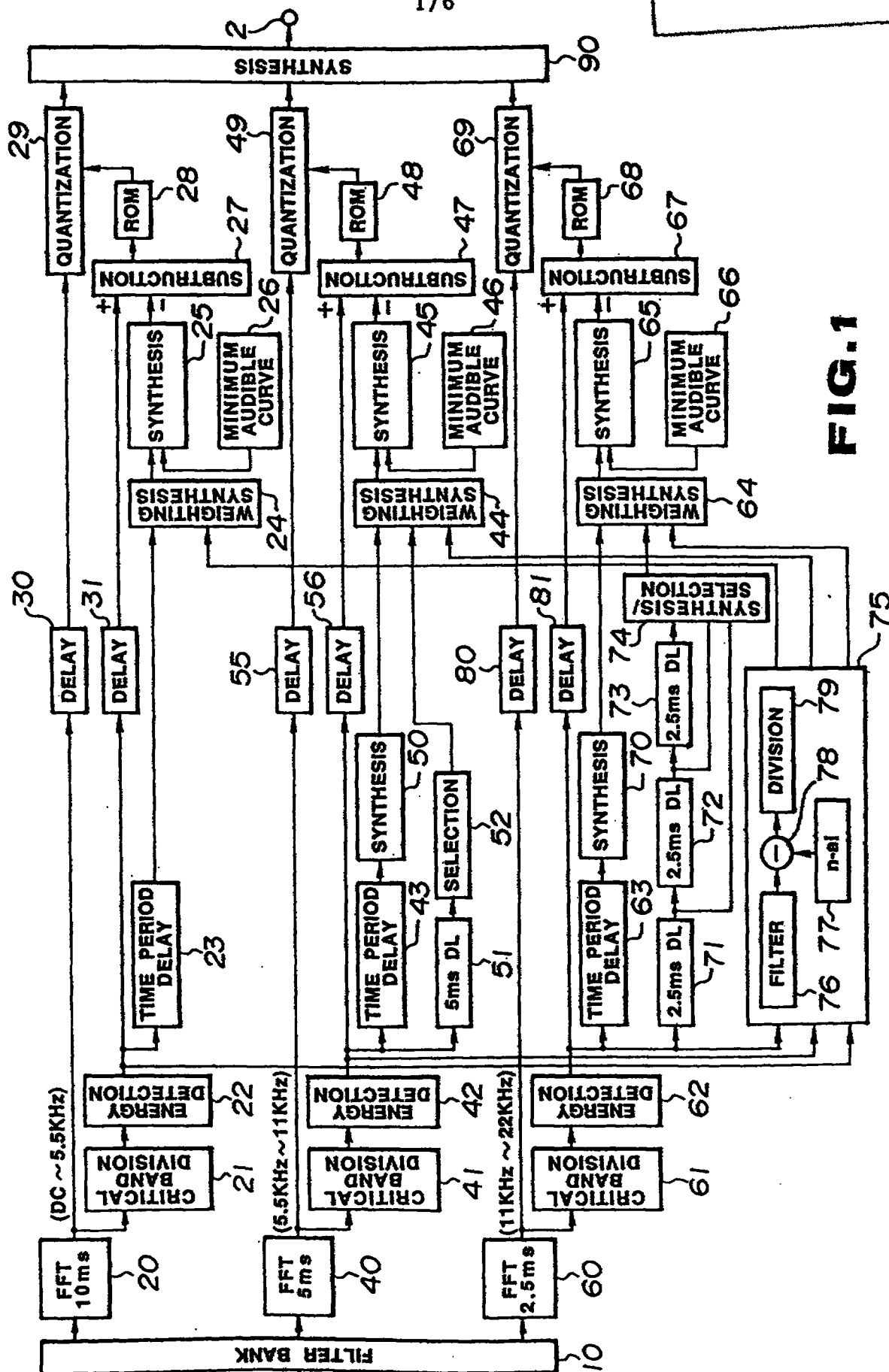
**ABSTRACT****DIGITAL SIGNAL ENCODERS**

5 In a digital signal encoder for analyzing an input digital signal
by frequency analysis means (10) of the non-block type and frequency
analysis means (20, 40, 60) of the block type to quantize outputs
obtained, a quantization characteristic (29, 49, 69) is set in the
basis of an allowable noise level based on an energy within a
10 predetermined time period of an output from the block type frequency
analysis means (10) and an allowable noise level based on an energy of
at least one of data preceding and succeeding in time within the
predetermined time period, thereby to provide a real time processing.

15 Figure 1

25 MAR 1997

1/6



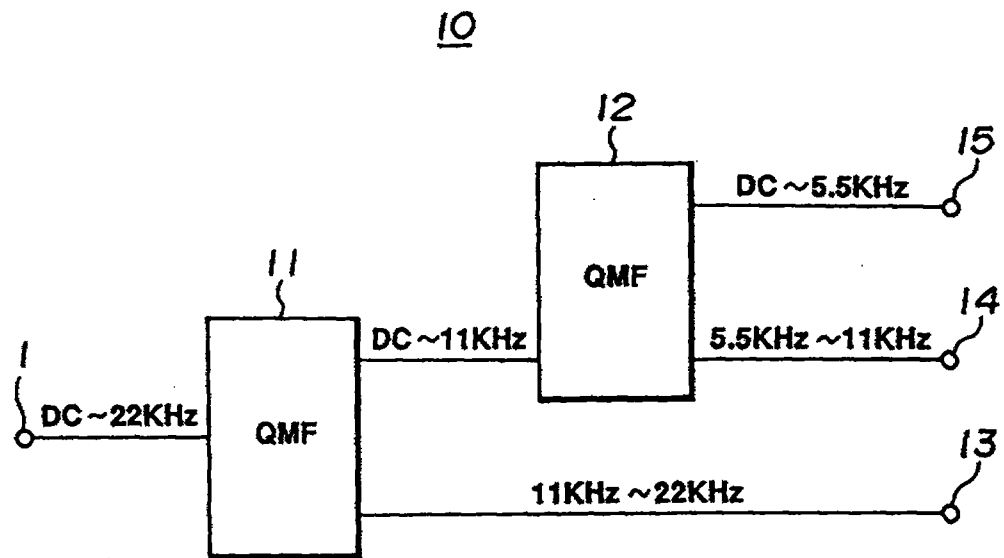
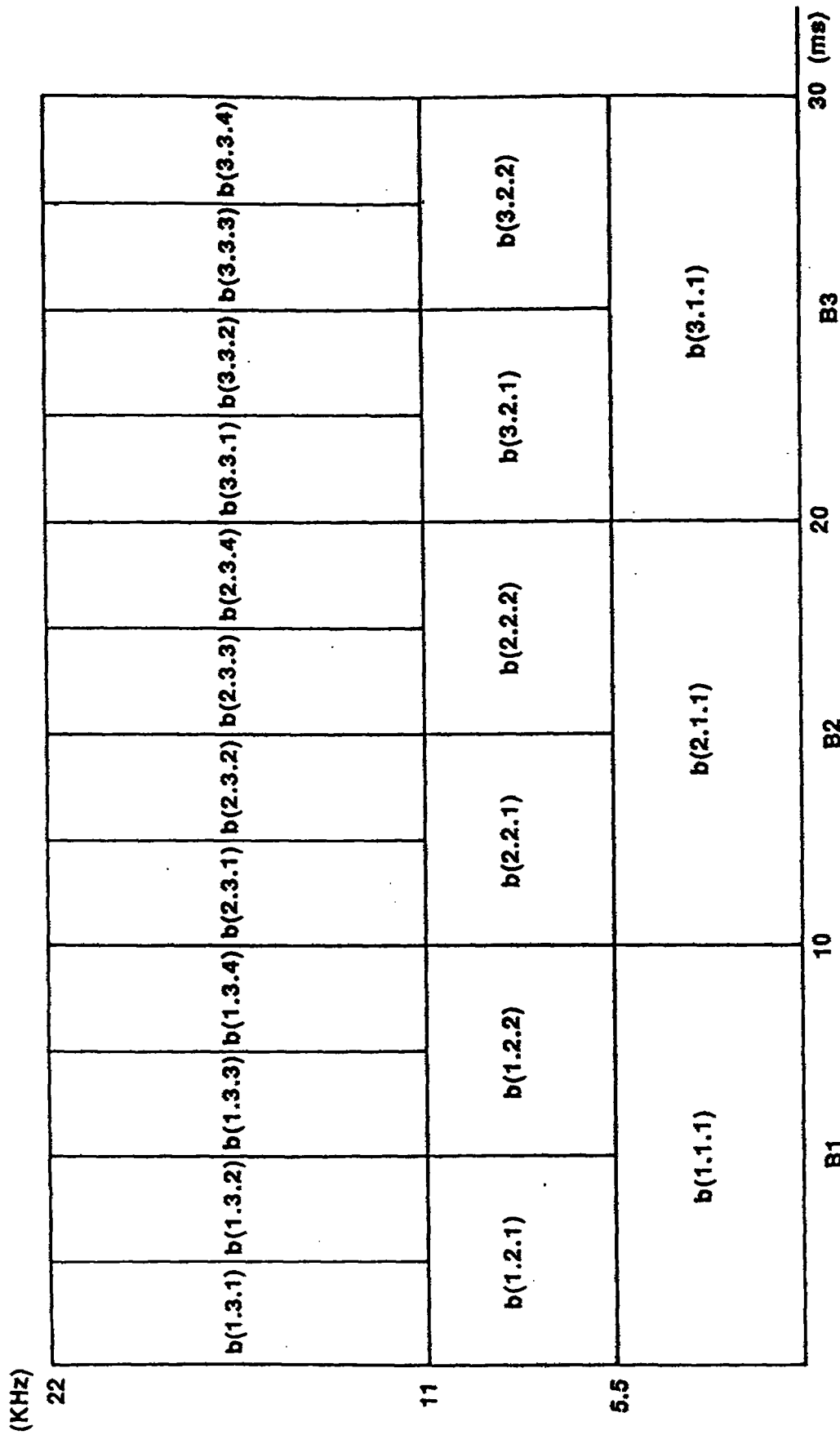


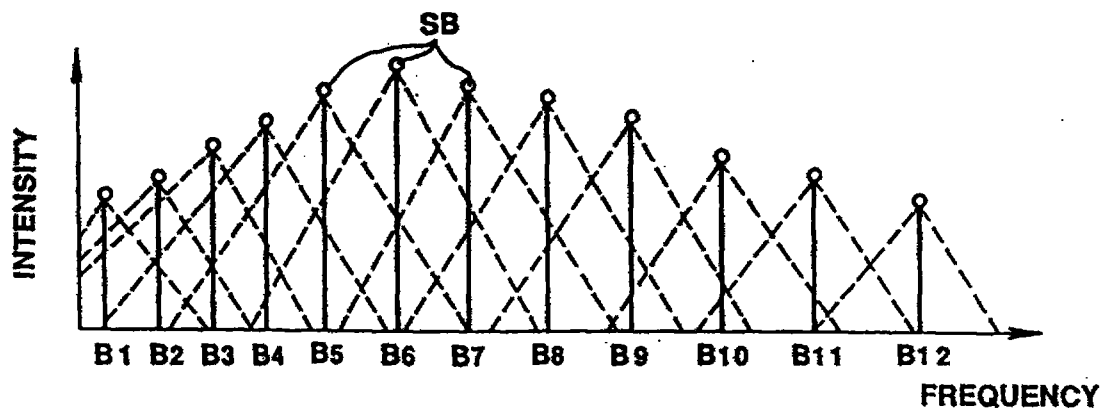
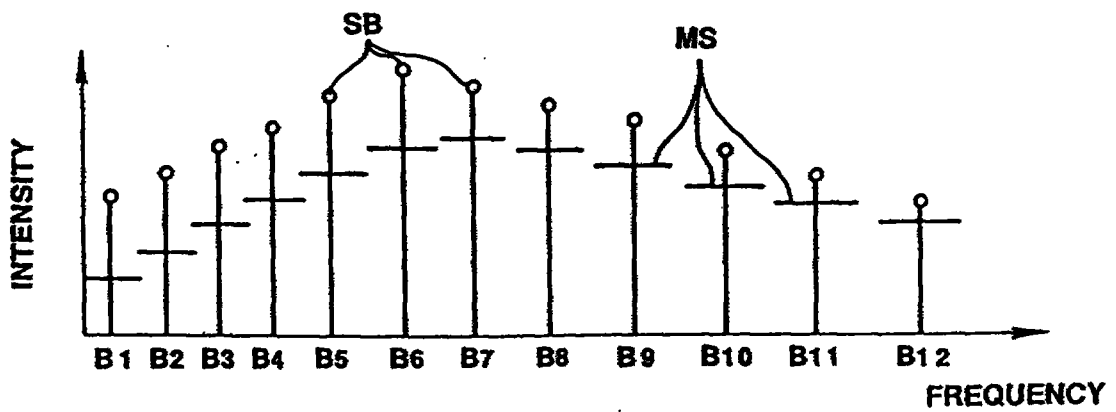
FIG. 2

25 MAR 1997



25 MAR 1997

4/6

**FIG. 4****FIG. 6**

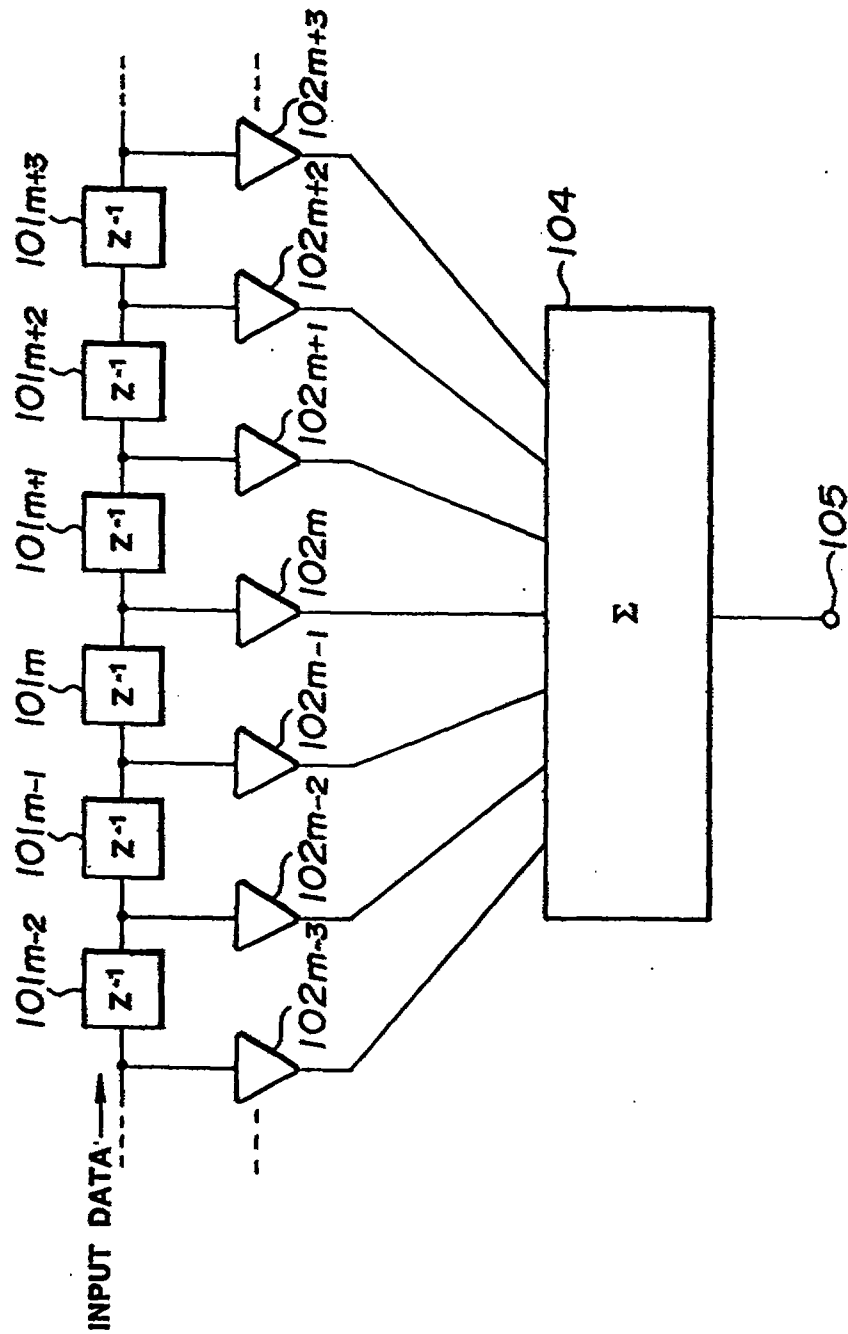
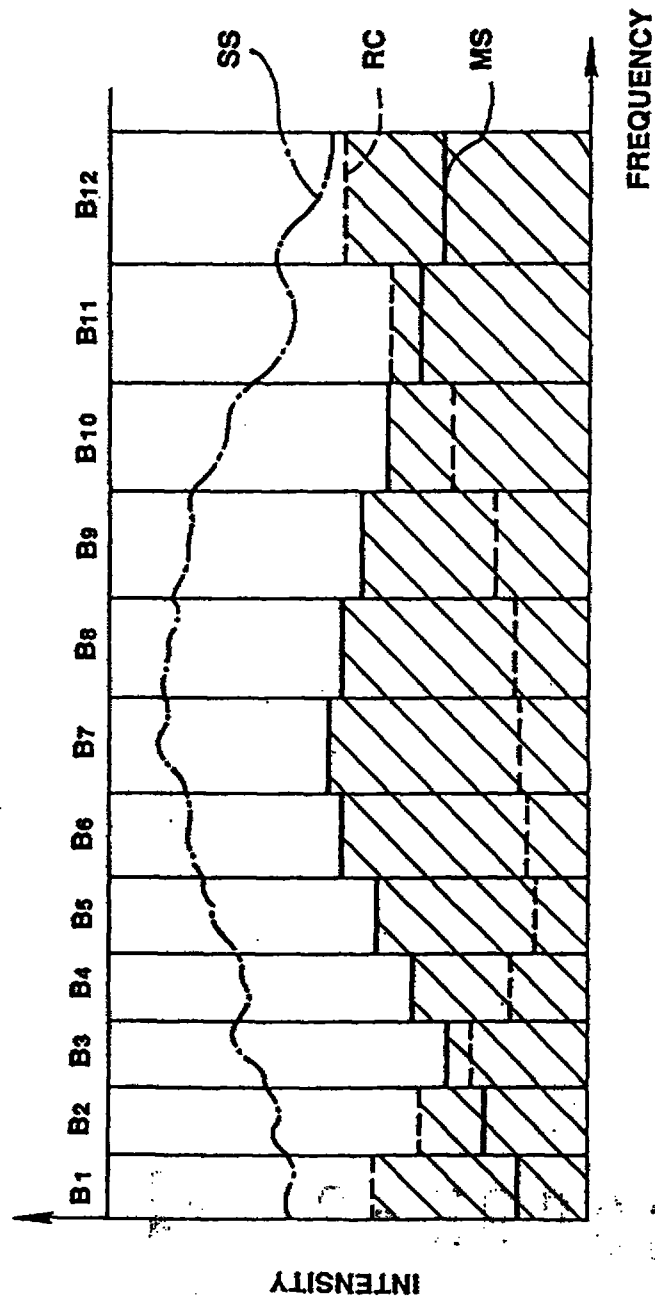
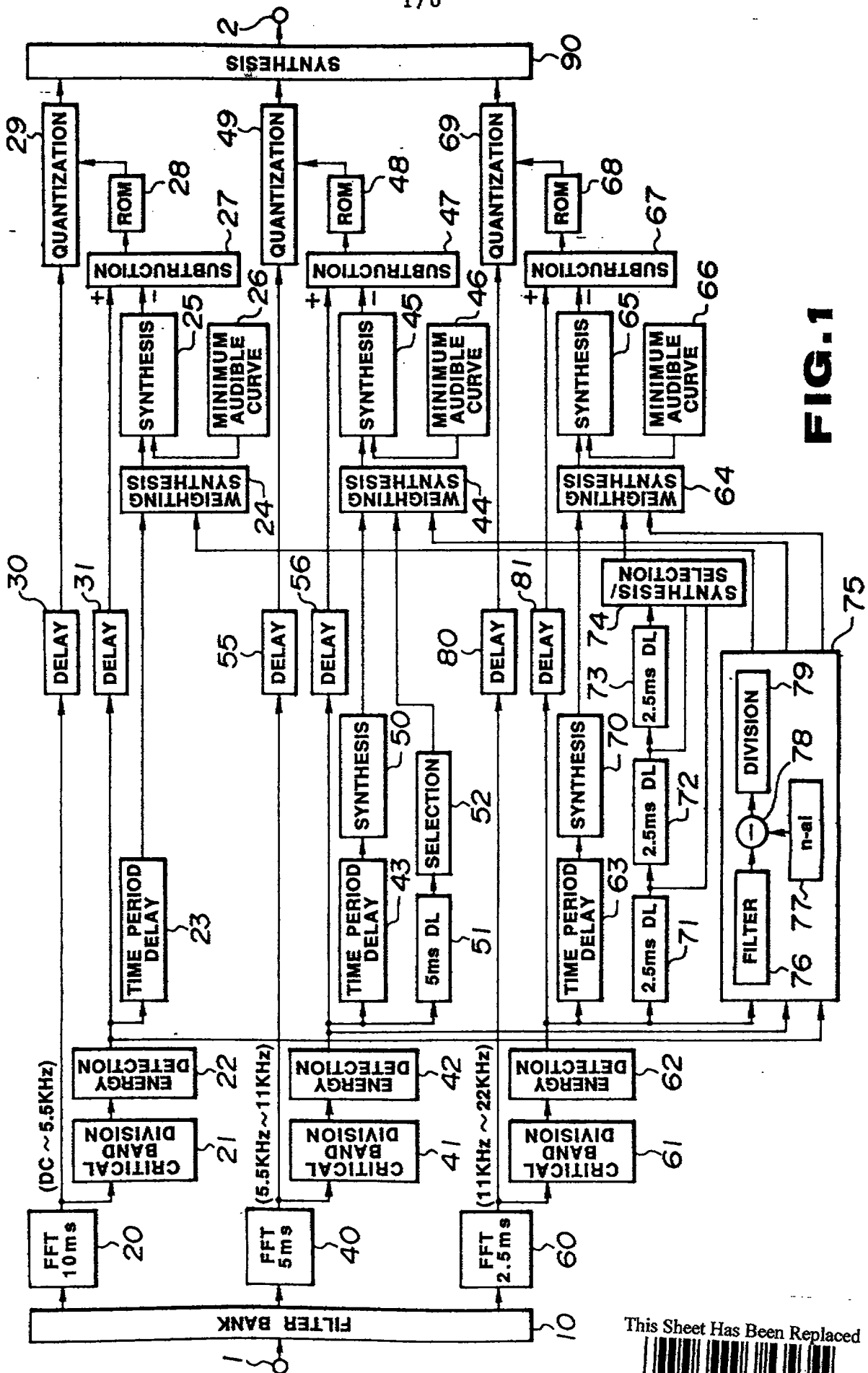


FIG. 5

25 MAR 1997

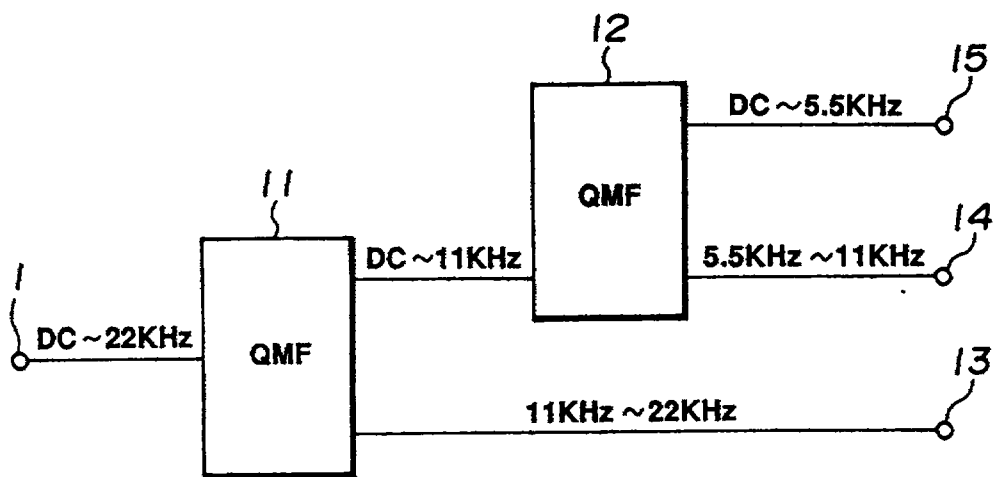
6/6

**FIG. 7**



This Sheet Has Been Replaced



10**FIG. 2**

This Sheet Has Been Replaced

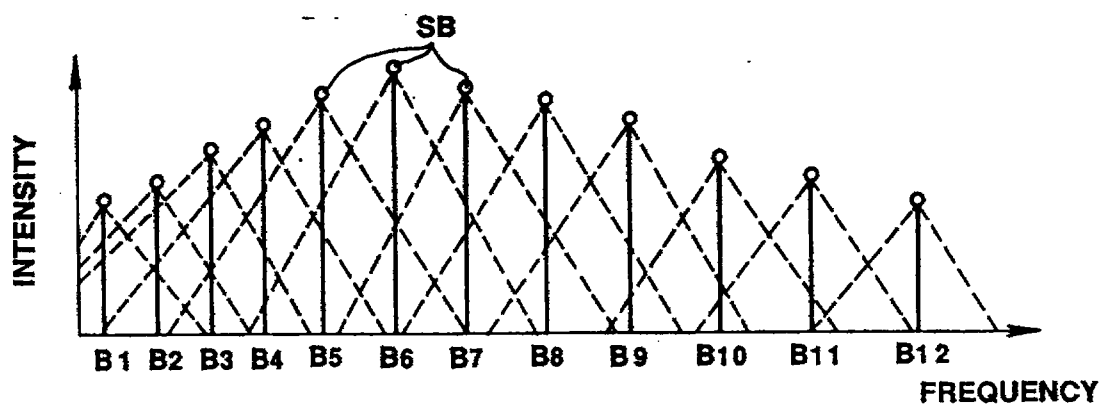
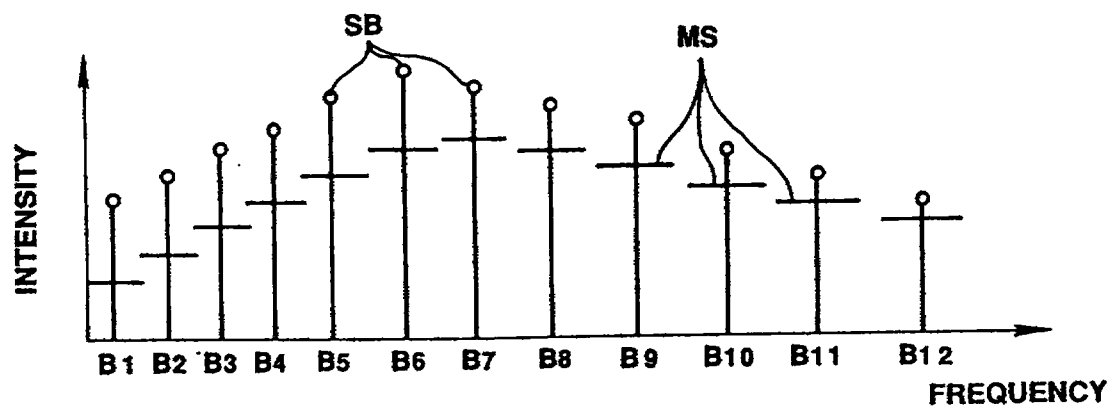


(KHz)													30 (ms)
22													B3
	b(1.3.1)	b(1.3.2)	b(1.3.3)	b(1.3.4)	b(2.3.1)	b(2.3.2)	b(2.3.3)	b(2.3.4)	b(3.3.1)	b(3.3.2)	b(3.3.3)	b(3.3.4)	
11													
	b(1.2.1)	b(1.2.2)			b(2.2.1)		b(2.2.2)		b(3.2.1)		b(3.2.2)		
5.5													
	b(1.1.1)				b(2.1.1)				b(3.1.1)				

FIG. 3

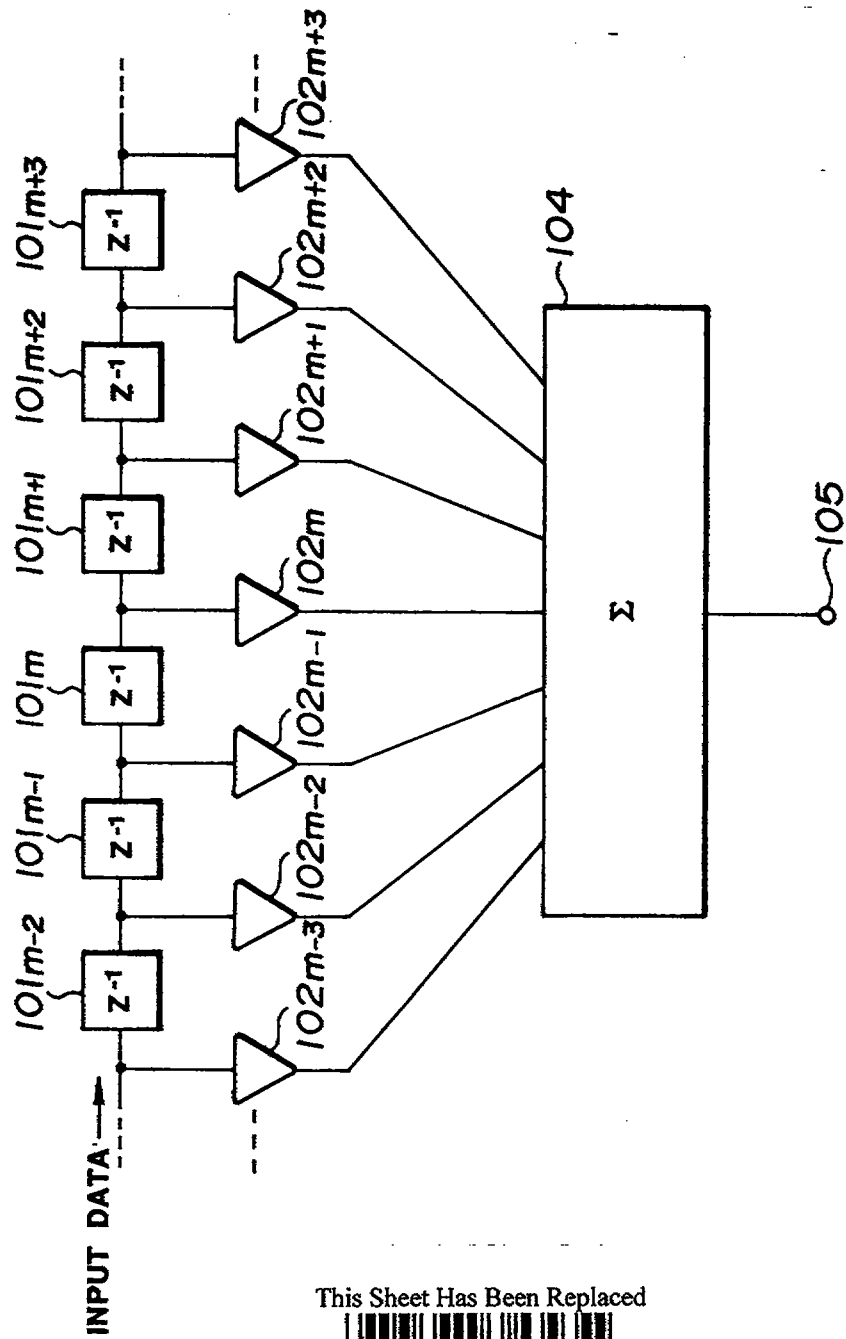
This Sheet Has Been Replaced



**FIG. 4****FIG. 6**

This Sheet Has Been Replaced



**FIG. 5**

This Sheet Has Been Replaced



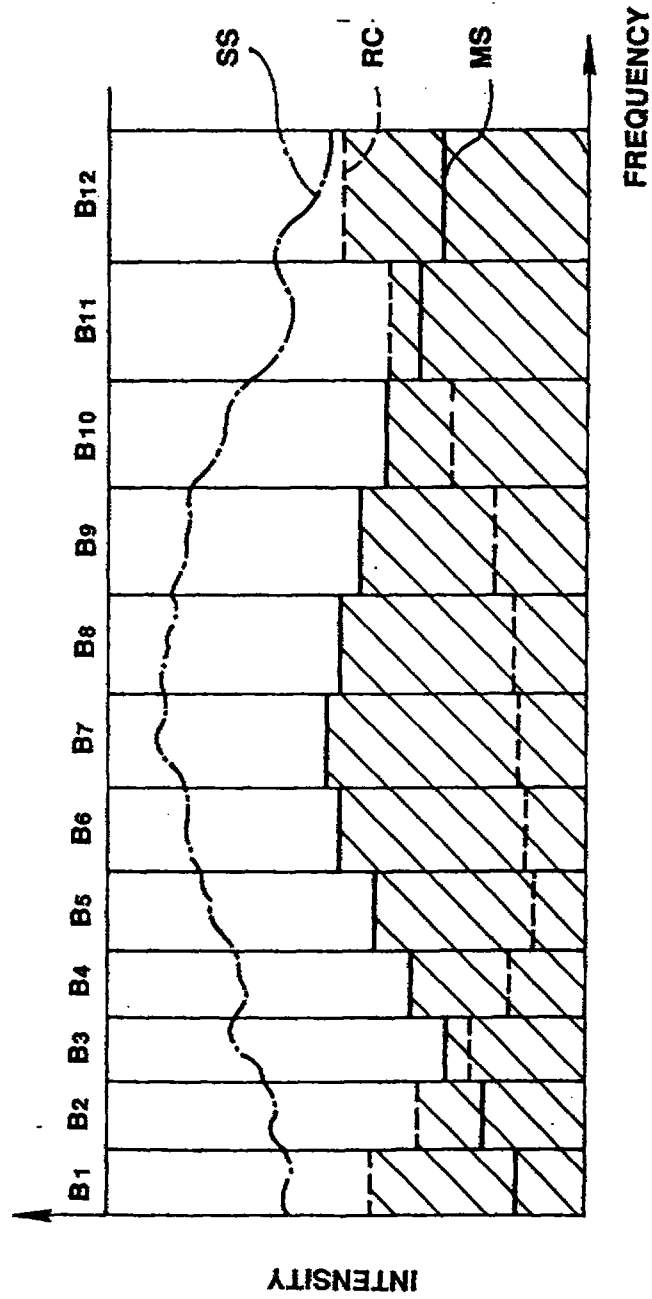


FIG. 7

This Sheet Has Been Replaced



2-222222