

The invention relates to a method of transmitting an audio signal in accordance with the generic part of claim 1.

A method of transmitting an audio signal is known from application EP-A 0 193 143 (not a prior publication), wherein the analog signal is converted to a digital signal and is converted to a signal representing the short time spectrum before being transmitted, wherein the whole frequency band is divided into a plurality of frequency groups, and parts of signals in the frequency groups are given different weighting in their representational accuracy in the coding of the signal to be transmitted, based on psycho acoustic laws, and wherein the coded signal is transmitted and after transmission is decoded, transformed back and converted back to an analog signal.

Application EP-A 0 193 143 describes a method for transmitting a digital audio signal which manages with a lower mean bit rate per sample scanned, without appreciably affecting the quality in reproduction.

The invention in application EP-A 0 193 143 substantially comprises dividing the transformed amount values in frequency groups from the particular maximum amount into ranges of values, above a threshold, transmitting amount values within the same range as if they were the same value, and setting amount values below the threshold to zero.

In a method in accordance with the generic part of claim 1 the problem on which the invention is based is to improve the weighting of amplitude values in frequency values and the detection of sudden sound events, and in particular to make better allowance for psycho acoustic factors in the transmission process.

This problem is solved by the invention described in claim 1 or 2. Other advantageous embodiments of the invention are described in the sub-claims.

The threshold below which the values in a group are set to zero is not fixed at -30 dB below the group maximum; instead it can be shifted adaptively as far as -50 dB. This is necessary if the amounts clearly peak in the group, since spectral values have a clear linear dependency over this large dynamic range. If these values are set to zero an easily perceptible modulating disturbance will appear in the time signal after transformation and re-transformation owing to the dependency.

The distinctive feature for recognising such a situation is the ratio of the peak value to the mean value within a group.

To analyse whether there is a jump in the signal curve the energy of the signal, which has previously been taken through a high pass filter, is determined in sub-blocks (e.g. 64 values). In this way any jump is made steeper and can be detected more easily.

The rise before the jump, deduced therefrom, may take place over a wide band as in the past, but a better effect is obtained if the rise starts frequency-selectively. It can then be restricted to spectral ranges where the energy is less pronounced, thereby allowing higher lifting factors without appreciably changing the shape of the short time spectrum.

A hearing threshold is placed across all the groups and spectral values below it are set to zero. The threshold runs horizontally in the first 23 groups and rises by about 30 dB within the last three. It is shifted as a unit, so that a certain distance (typically 90 dB) from the maximum amount value is maintained in the lower range. In the case of shifts to lower values a stop is defined for the threshold, so that extremely small values are not considered. This threshold is to make allowance firstly for the absolute hearing threshold and secondly for masking effects covering more than one group.

For a better understanding of the invention an example will be explained in detail below with reference to the accompanying drawings, in which:

Figure 1 shows time windows with a sinusoidal signal oscillating in them

Figure 2 shows frequency spectra of a windowed sinusoidal signal

Figure 3 shows the progress of the process in time

Figure 4 illustrates the preprocessing of the signal for a special signal content

Figure 5 is a block circuit diagram of it and

Figure 6 shows a frequency spectrum with thresholds

Figure 1 shows a sinusoidal curve 4. Two time windows t_1 and t_2 are placed over the curve 4. Time window t_1 begins at zero point 1 of the sinus and ends at a second zero point 2 of the sinus. The zero points 1 and 2 are the intersections of the sinus and the abscissa, the time t being entered on the abscissa. The second time window t_2 begins at the zero point of the sinus but ends outside a zero point, at point 3.

In the transformation of time signals to a signal representing the short time spectrum different curves are obtained in the frequency spectrum. Figure 2 shows the frequency spectrum relating to the two windowed sinusoidal signals. For time window t_1 there is only a frequency line 5. For the sinusoidal signal which is windowed by means of the time window t_2 there is a frequency spectrum with a curved shape 6 which has a maximum at the frequency line 5 and a descending curve with several frequency lines 7. In a frequency group easily observed modulating disturbances may occur in the limitation and transmission of values which only come within a dynamic range of 30 dB below the maximum y_{max} , owing to the linear dependency of the spectral values. The dynamic range is therefore enlarged adaptively to up to 50 dB. The distinguishing feature for recognising such a situation is the ratio of the peak value to the quadratic mean value within a frequency group. The difference between the peak value and the mean value in a frequency group is determined and is multiplied by a factor, in this case 3. The product gives the new dynamic range. If the peak value is 60 dB and the associated mean value 45 dB, the difference is 15 dB. The difference of 15 dB is multiplied by the factor 3 giving a new dynamic range of 45 dB, i.e. an adaptation between -30 and -50 dB. If - in parallel with laying down the threshold which stipulates the group dynamics - it is found that the difference between the maximum and the

minimum amount value in a group is less than 30 dB, the dynamic range to be represented is made equal to that difference.

In figure 3 the analog signal $a(t)$, representing an audio signal such as speech or music, is converted to a corresponding digital audio signal in the analog/digital converter 11. At stage 12 so-called windowing of the signal takes place by time windows which succeed and overlap each other in time. The signal is divided into blocks which succeed each other in time each lasting about 20 milliseconds, particularly 23 milliseconds, in such a way that the signal of a block can in each case be further processed separately. At stage 13 preprocessing of the signal takes place; its significance will be explained later. At stage 14 the digital signal of a time window or of a block is converted to a frequency spectrum by transformation. So during the blocks which succeed each other in time there is in each case a signal at the output of stage 14, representing the spectral components of the signal over the whole frequency band for the duration of a time window or block. Thus stage 14 effects conversion of the signal from the time range to the signal representing the spectrum in the frequency range.

The signal from stage 14 reaches the coder 15. Here coding is effected in accordance with psycho-acoustic factors. This means that spectral components which are in any case not observed in reproduction owing to masking effects in particular are given lower weighting or omitted in coding. Such processing of the short time spectrum is possible, e.g. with a computer.

The signal thus coded passes through the transmitter 16 to the traffic channel 17. The reduction of the mean bit rate obtained enables the traffic channel to have corresponding narrow band dimensions. The channel 17 is followed by the receiver 18 which carries out substantially the reverse functions to the transmitter. The signal then passes first to a decoder 19 which carries out decoding corresponding to the coder 15. At stage 20 the resultant spectrum-representing signal in the frequency range is converted back to a digital signal in the time range. At stage 21 the signal is reconstituted into a uniform

continuous digital signal and allowance is made for the preprocessing at stage 13. The signal is then fed to the digital/analog converter 22. The converter 22 supplies the analog signal $\underline{b}(t)$ again. This is not identical with the signal $\underline{a}(t)$ because spectral components were given different weighting or suppressed during coding in the coder 15. But the difference between the analog signals $\underline{b}(t)$ and $\underline{a}(t)$ is such that it is not noticed by the hearer when the signals are reproduced. Thus only irrelevant matter is removed from the signal, information inaudible to the hearer, so that the necessary bit rate can be reduced during transmission along the traffic channel 17; the decision content in particular is reduced. On the signal path 23 the preprocessing stage 13 informs the transmitter 16 whether preprocessing has taken place. If it has, additional information is inserted or multiplexed into the coded signal, is recognised by the receiver 18 and passed onto the adaptation stage 21 via the signal path 24. A low pass filter to fulfil the scanning theorem is inserted upstream of the analog-digital converter 11. A second low pass filter for reconstruction is arranged downstream of the digital-analog converter 22.

Figure 4 illustrates the preprocessing of a sudden sound event 29 occurring within a time window t_1 to t_7 at the point of time t_9 . Such an event may, for example, be the striking of a triangle. The preprocessing described takes place at stage 13 in figure 3. The event 29 is preceded by a preoscillation between t_8 and t_9 , although premasking makes this inaudible. During conversion to the frequency spectrum at stage 14 in figure 3 a signal in the frequency range is formed in each case, giving the spectral distribution in the window t_1 to t_7 . Since there is no longer any assignment of spectral lines to individual points of time within a time window in this signal, the event 29 would be averaged out over the whole time window t_1 to t_7 , i.e. quasi blurred. This can cause audible distortion.

To avoid this still possible fault a time window t_1 to t_7 or block is divided into 32 sub-blocks. The amplitudes of the individual sub-blocks are determined. As soon as the jump in amplitude between two sub-blocks exceeds a predetermined limit, determined in figure 4 by the event 29, an additional measure is triggered. The predetermined limit is

of the order of 20 dB. The measure is that before the jump in amplitude the signal is raised in amplitude on the transmission side by a compander process and correspondingly lowered on the receiving side. Thus the said fault caused by the blurring of the transient event over the whole time window is reduced. Figures 4D and 4E show the operation of a compander system. Before a signal jumps it is amplified with an amplification factor V_1 , in this case 5, during preprocessing in an expander. On the receiving side this is reversed in a compressor during adaptation 21 with a second amplification factor V_2 , in this case $1/5$. The total amplification for this signal is the whole time "1".

$$\text{Total amplification} = V_1 \times V_2 = 1$$

After a signal jump the amplification factors up to the next jump have the value 1 both in the expander and the compressor. In the transitions between amplification factors 1 and 5 there is a straight or curved progression for a period between half a sub-block or a whole sub-block but preferably for the duration of a sub-block. The aim is to exploit the premasking effect of the ear (under 1 millisecond).

The preprocessing 13 is shown in figure 5. The block 30 checks whether a signal rises by a predetermined limit between two of the 32 sub-blocks. So long as the rise between two blocks is below the predetermined limit (20 dB), the signal from the input 35 is switched directly via a time delay 36 and the two-way switch 33 to the output 34. If there is a signal jump of 20 dB between two sub-blocks, the signal from the input 35 is switched via an expander 32 and the two-way switch 33 to the output 34. The high pass filter 31 makes any jump in the signal steeper, so that the jump is detected better by the block 30. Then the signal is raised in amplitude by the expander 32 on the transmitting side, before the jump in amplitude, and correspondingly lowered again on the receiving side. This reduces faults caused by blurring of the transient event over the whole time window. The checking block 30 works continuously, even while the signal from the input 35 is switched directly via the two-way switch 33 to the output 34. The time delay 36 compensates for differences in time caused by the expander 32 and the checking block 30. The checking block 30 reports the timing of a jump or the exceeding of a limit to

the expander 32. This report is information on the number of the sub-block. The expander 32 automatically calculates a suitable amplification factor.

Figure 6 shows a coordinate system with the curve 41 of a frequency spectrum plotted on its abscissa over a frequency f . The curve 41 has a maximum $x_{\max}$. An absolute maximum $x_{\max 1}$ corresponds to a level A of zero dB. In a computer this zero dB corresponds to an absolute value of two exponent (exp) 15, which is additionally amplified with an amplification factor of 1024. The computer then calculates altogether with an absolute value of $2 \exp 25$. The values correspond to a voltage. A hearing threshold 42 is placed over the whole frequency, and below this threshold the spectral values are set to zero, i.e. are not considered. In the first 23 frequency groups the threshold 42 runs parallel with the abscissa; it rises by about 30 dB within the last three groups, the rise taking place from 10 kHz. Before a first rising point, which will be referred to as a salient point, the threshold 42 is always 90 dB from the (variable) maximum $x_{\max}$. That is to say, this threshold is always positioned dependent on the maximum $x_{\max}$. The threshold 42 cannot drop below a minimum value 43 though, which will be referred to as the stop. The stop 43 is -128 dB from the absolute maximum. The first salient point is advantageously at 10 kHz and a second salient point at 12 kHz. Between these two points the threshold 42 first rises 10 dB. From the second salient point, i.e. within the range from 12 to 22 kHz, it has a rise of 90 dB. Frequency groups and division of a frequency spectrum are explained, e.g. in "Psychoakustik" (Psycho Acoustics) by E. Zwicker, published by Springer Verlag, Berlin Heidelberg New York 1982.



CLAIMS

1. A method of transmitting an audio signal, wherein the analog signal is converted to a digital signal, transmitted digitally and converted back to an analog signal, wherein the audio signal is converted into a signal representing the short time spectrum before being transmitted, wherein the whole frequency band is divided into a plurality of frequency groups, parts of signals in the frequency groups are given different weighting in their representational accuracy in the coding of the signal to be transmitted, based on psycho acoustic laws, wherein the coded signal is transmitted and after transmission is decoded, transformed back and converted back to an analog signal, and wherein a threshold is allocated to a frequency group and amount values below the threshold are set to zero, characterised in that the thresholds are shifted to allow for an absolute hearing threshold and masking effects covering more than one group and/or there is frequency-selective raising of the amount values.
2. A method of transmitting an audio signal, wherein the analog signal is converted to a digital signal, transmitted digitally and converted back to an analog signal, wherein the audio signal is converted before transmission into a signal representing the short time spectrum before being transmitted, wherein the whole frequency band is divided into a plurality of frequency groups, parts of signals in the frequency groups are given different weighting in their representational accuracy in the coding of the signal to be transmitted, based on psycho acoustic laws, wherein the coded signal is transmitted and after transmission is decoded, transformed back and converted back to an analog signal, and wherein a threshold is allocated to a frequency group and amount values below the threshold are set to zero, characterised in that the thresholds are shifted dependent on the ratio of the peak value to the mean value within a frequency group and/or there is frequency-selective raising of the amount values.
3. The method of claim 1 or 2 characterised in that the threshold below which the values are set to zero in a frequency group is shifted adaptively (figure 2).

4. The method of any of the preceding claims, characterised in that the threshold is shifted between -30 and -50 dB.
5. The method of claim 1 or 2, characterised in that a signal in preprocessing (13) is taken through a high pass filter (31) before the checking block (30).
6. The method of claim 1 or 2, characterised in that a hearing threshold (42) is placed across all the frequency groups.
7. The method of claim 6, characterised in that the hearing threshold (42) assumes a predetermined distance from a maximum (x_{max}).
8. The method of claim 7, characterised in that the hearing threshold (42) can be shifted only as far as a stop (43).
9. The method of claim 8, characterised in that the stop (43) is at about -128 dB from an absolute maximum (x_{max1}).
10. A receiver to receive an audio signal which is converted to a digital signal using an analog/digital converter (11), which signal is transmitted digitally and after being received is converted back to an analog signal using a digital/analog converter (22), wherein the whole frequency band of the audio signal is divided into a plurality of frequency groups, and parts of signals in the frequency groups are given different weighting in their representational accuracy in the coding of the signal to be transmitted, using a coder (15) based on psycho acoustic laws, wherein a threshold is allocated to a frequency group and amount values below the threshold are set to zero, and wherein a decoder (19) contained in the receiver (18) effects decoding corresponding to a coder (15) at the transmission side, and the decoded signal is supplied to the D/A converter (22), characterised in that the threshold in the coder (15) can be shifted to allow for an absolute hearing threshold and masking effects



covering more than one group and/or there is frequency-selective raising of the amount values, and irrelevant parts of the audio signal $(h(t))$ are completely or partially removed or greatly reduced.

11. The receiver of claim 10, characterised in that the receiver (18) recognises whether an additional piece of information, reporting whether preprocessing has taken place, has been inserted or multiplexed into the coded signal, that the audio signal transmitted is constituted into a uniform digital signal in an adaptation stage (21), and that the adaptation stage (21) allows for the preprocessing of the audio signal by the coder.

