EXHIBIT P

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

Preprint# 2970
Session-Paper# F-II-4

Bitrate Reduction of High Quality Audio Signals by Modeling the Ears Masking Thresholds

Detlef Wiese and Gerhard Stoll
Institute für Rundfunktechnik, GmbH
Munich, West Germany

Presented at
the 89th Convention
1990 September 21–25
Los Angeles

AES

This preprint has been reproduced from the author's advance manuscript, without editing, corrections or consideration by the Review Board. The AES takes no responsibility for the contents.

Additional preprints may be obtained by sending request and remittance to the Audio Engineering Society, 60 East 42nd Street, New York, New York 10165 USA.

AN AUDIO ENGINEERING SOCIETY PREPRINT

BITRATE REDUCTION OF HIGH QUALITY AUDIO SIGNALS
BY MODELLING THE EARS MASKING THRESHOLDS

Detlef Wiese and Gerhard Stoll
Institut für Rundfunktechnik GmbH
D-8000 München 45
Floriansmühlstr. 60
Tel : +49 89 32 399-258
Fax : +49 89 32 399-351

Abstract

Since few years there are some different source-coding systems for high quality audio signals (CD-quality) which exploit the redundancy and irrelevancy of speech and music signals for an efficient data-reduction. Psychoacoustic research is the basis for the subband coding system *MUSICAM* (Masking-pattern Universal subband Integrated Coding And Multiplexing), which was developed in cooperation of french, dutch and german scientists. During the developement different coding and sound analyzing strategies were used and combined to optimize this system. A combination of two analysis methods, subband and transform, seems to be the best strategy to achieve the highest reduction of the bit-rate. Subband samples are coded using a bit-allocation, which is mainly derived from the spectral values of a Fourier transform. For the synthesis in the decoder only the subband method is used.

1. Introduction

In the future the developements in the field of audio will be dominated by digital technology with more and more signal processing power at its disposal. These capacities can be used for modelling known parts of the auditory system achieving a considerable bit-rate reduction for different applications.

An economic use of transmission channels and storage media for high quality digital audio signals requires a significant reduction of the bit-rate. A future high quality broadcasting system for radio or TV-sound has to attain the quality of the widely distributed compact discs. FM radio quality is no longer acceptable for the modern music listeners, who can perceive today nearly the same audio quality as used in digital production studios. The most important prerequisite for developing a digital audio broadcasting system is the spectrum efficiency, which is necessary to supply as much programs as possible. This is necessary particularly for a terrestrial system allowing not only stationary but also mobile reception. In addition, the sensitivity due to bit errors occuring most of all by mobile reception has to be minimized.

Broadcasting contribution links, which are another application for high quality audio bit-rate reduction systems are important for an international program exchange, as well as for reporting links. Normally, CD-quality requires digital transmission at over 700 kbit/s per monophonic channel, studio quality even more. The costs of the transmission links are proportional to the rate and therefore it might be desirable to have a stereo sound program exchange in studio quality all over the world on the basis of the future ISDN access level H0, which provides a capacity of 384 kbit/s. Another broadcasting requirement is the transmission of actual radio news and current affairs which depends on the telephone system with all its disadvantages, like the reduced bandwidth of only 3 kHz, the low signal to noise ratio and interferences. This transmission quality is capable of improvement. By combining data reduction schemes and the spread of the worldwide standard for the future ISDN, relative high quality audio channels can be established at reasonable costs using the 2*64 kbit/s ISDN basis connections.

The disseminated techniques for low bit-rate coding, subband coding and transform coding, are related to psychoacoustic phenomena of human sound perception. The subband coding system *MUSICAM* which is based on the proposals /1,2,3,4,5,6/ exploits most of all the irrelevancy, and partly the redundancy of speech and music signals for an efficient data-compression. The reduction is done mainly by considering the ears masking thresholds, such as spectral and temporal masking effects, as well as to a certain amount by considering the statistics of the audio signal. Spectrum analysis and signal-dependent bit-allocation are used to reduce the irrelevant parts of the audio signal, and to match the quantizing noise to the ear characteristics.

Due to the fact, that there are some important differences between the characteristics of measurement techniques and the ears properties, the sound quality assessment of source coding schemes, which exploit the irrelevancy, can't be based only on the conventional measurement methods nowadays in use. A high sophisticated psychoacoustic measuring instrument modelling the auditory system should be used to assess the quality. Up to now no suitable system is available and therefore carefully prepared listening tests have to be carried out to improve the coding algorithm and to guarantee a high audio quality system. First results of listening tests have shown that a bit-rate reduction from about 700 kbit/s (Compact Disc) down to 100 kbit/s per monophonic channel is possible without impairing the subjective quality for nearly any critical signals. Exploiting the dependencies between the two channels of a stereophonic program an additional reduction down to 128 kbit/s per stereophonic channel can be achieved nearly without or only minor perceptible impairments.

Many different applications in the area of communications are only conceivable by using low bit-rate coding, particularly those, which are in connection with future multi-channel systems (e.g. HDTV), electronic publishing, audio data bases, educational, business and entertainment applications. Due to this big range of applications, the requirements to the parameters quality, delay, complexity, error-sensitivity and bit-rate have to be considered carefully.

2. New aspects to basic principles of audio source coding

In recent years many different low bit-rate audio coding systems have received much attention. The reduction of bit-rate of these codecs is mainly determined by the reduction of the redundancy using statistical correlations. Additionally the high sophisticated codecs reduce the irrelevancy using psychoacoustical phenomena. The reduction can be constant or variable, wether a fixed or a dynamic bit-allocation is used. A considerable reduction isn't conceivable without using psychoacoustics.

But what should be the fundamental basis of the input-data for the calculation of the masking thresholds?

The results of masking threshold measurements for narrow band signals considering tone masking noise and vice versa, the dependency on the frequency and the level are the basis for the estimation of the masking threshold of a temporal variant complex audio signal. Very special masker/test-tone(noise) relations are described in the literature and the worst case results regarding the upper and lower slopes of the masking curves have to be considered for the assumption, that the same masking thresholds can be used for both, complex and simple situations.

The disseminated well known techniques are subband and transform coding, often linked with ADPCM, vector quantizing, entropy coding, pre- and post-processing /2,7,8,9/. The basic problem is the design of an optimal analysis-/synthesis strategy which provides the contrary requirement to serve both, a high frequency and a high time resolution combined with a low complexity implementation necessary for most applications. It's plausible, that there is no simple and low complexity way to model the auditory system. There are proposals to combine the two fundamental methods, subband and transform coding /5,6,10/ and a parallel and serial combination are shown in figure 1. There are different advantages and disadvantages regarding the bit-rate, complexity and delay.

The serial use of filters and transform as shown in figure 1 is an efficient strategy to obtain a frequency resolution which is similar to the characteristics of the human ear. The redundancy can be reduced using entropy coding for the spectral values of the transform within the individual lower subbands. This is an additional reduction in comparison to subband coding, because statistical investigations regarding the subband samples have shown, that in nearly all cases a gain in bit-rate reduction couldn't be achieved by entropy coding, e.g. like Huffman. The quantizing noise can be adapted optimally to a masking-pattern, because the noise is limited to a smaller frequency range coding the spectral components than coding the subband samples.

The serial combination has the disadvantage that frequency-aliasing is produced by the filterbank and additional time-aliasing is produced by the transform. The frequency-aliasing components are processed by the transform and due to the fact, that the necessary distinction between original and distortion signal is not possible, problems occur during psychoacoustic weighting because the input data for the threshold estimation are not exactly right. The possibility to match the quantizing noise very close to the masking-pattern can not be exploited optimally.

The advantages of the parallel combination is given by two independent ways of analyzing the audio signal. Only filtering is used for coding and this allows later improvements of the analysis in the encoder without any change of the decoder. Regarding the delay there is no addition of transform and filterbank which is of interest for many applications inpartlcular for the decoder. For the analysis, both, the wanted frequency and time resolution can be derived from the frequency transform and from the filterbank with an accuracy corresponding to the ears characteristics. Unfortunately, a parallel structure with short delay and low complexity decoder doesn't allow an optimum matching of the quantizing noise to the masking thresholds in the lower frequency region due to the lack of relatively small subbands.

As an example, the calculated masking thresholds using subband levels, that means scalefactors, are compared schematically to the thresholds calculated from the spectral components of a transform for inband, as well as for outband masking. The comparison is adopted to the characteristics of the MUSICAM-encoder. The calculated simultaneous masking threshold inside the third subband is shown in figure 2c and 2d. If only the subband levels are used for the calculation (2c), we have to assume the worst case situation, that means a masker with the subband level at the lower or upper cut-off frequency of the subband, according to the steepest slope, which is typically the lower. Using a transform on the other hand (2d), the position of the masker is known very exactly and the estimation of the curves can be approximated much more accurately. The decisive argument to prefer a transform with a high frequency resolution is the relevant difference regarding the minimum of the masking threshold, which is responsible for the bit-allocation. The necessary dynamic range can be calculated by the difference between the maximum level and minimum of the masking threshold inside the subbands (signal-to-mask ratio). Looking at figure 2c a bit-allocation of about 13 bit, looking at figure 2d a bit-allocation of only about 5 bit can be derived from the necessary signal-to-mask ratio. The figures 2a and 2b show the masking of adjacent subbands. For instance, the calculated bit-allocation in the adjacent fourth subband would be much less using a transform than using subband levels. The conclusion is to calculate the simultaneous masking threshold based on a transform with a sufficient frequency resolution. The masking index, which presents the difference between masker level and masking threshold at the maskers frequency, depends on the frequency and the type of the masker. The information about the structure of the maskers can be derived from the spectrum analysis /11/ and improves in addition the precision of the masking threshold estimation.

3. MUSICAM

The spectrum of the digital broadband audio signal is divided into a definite number of subbands using a suitable filterbank. Adapted to the ears masking effects in time, successive samples of each subband are combined to a block. Once in each frame the peak-value is represênted by a quantized scalefactor. The global masking threshold of the audio signal is estimated and the quantization of each subband is determined by the actual calculated signal-to-mask ratio in each subband for a certain time interval (Dynamic Bit Allocation).

3.1 Encoder

The blockdiagram of the MUSICAM-encoder is shown in figure 3. The mean part of a subband coding system is a suitable filterbank. Attaching importance to a perfect reconstruction of the signal and achieving a total cancellation of aliasing distortions, the filterbank is of a fundamental significance for the low bit-rate scheme. The requirements of many applications regarding the parameters complexity and delay are contrary to the demand of a high quality design of the filters. Typical designs of filterbanks as mentioned in /12,13/ for audio coding are tree-structured Quadrature-Mirror or Wave-Digital-Filters, which approximate bandwidths of the individual subbands according to the critical bands of the ear. The disadvantages of these structures are the high complexity and the long delay due to the cascading of many filter-stages. A new design of a windowed filterbank with a polyphase matrix /14/ in combination with a Fourier transform analysis /5, 6, 19/ seems to be the best compromise for a low bit-rate scheme.

The gain of the Polyphase filter network in comparison to a Quadrature Mirror Filter structure (QMF) is shown in figure 4. Achieving the same bandwidth as realized with QMF-filters the delay can be reduced from 53.33ms to only 10.66ms for analysis and synthesis filtering together. The computational power can be diminished from about 49MOPS to 5MOPS.

Using a sampling-frequency of 48kHz, the polyphase filter divides the audio signal into 32 subbands with a constant bandwidth of 750Hz. In each subband 12 successive samples are combined to a block. This is equivalent to a duration of 8ms and was choosen in view to the ears masking thresholds in time domain, because the possibly arising quantizing-noise has to be less than the ears pre-masking thresholds. In each block the absolut peak-value is determined and quantized as a 6-bit scalefactor with a dynamic range of 120dB. This range, which has to be represented by the scalefactors, depends on the dynamic range of the input signal, normally 16-bit (about 98dB), on the number of subbands and on a headroom, required for the filters. The scalefactors guarantee an optimal exploitation of the actual dynamic range of the quantizers. The bit-rate of the scalefactors is reduced by the elimination of irrelevancy and redundancy using a special coding technique as desribed in chapter 4.

One important advantage of filtering the audio signal into subbands is given by adopting the audio blocks optimally to the requirements of the temporal masking effects. A disadvantage of the filterbank can be seen in the lack of frequency resolution. A high resolution, that means small subbands in the lower frequency region and a lower resolution, that means wide subbands in the upper frequency region should be the basis for an adequate calculation of the masking thresholds in the frequency domain. The polyphase filter network is a socalled N-band filter with a constant bandwidth of the subbands over the whole frequency range and it provides no simple way to adopt another division. These subbands with a constant bandwidth of 750Hz are not sufficient for an estimation of the masking thresholds, particularly in the lower frequency region (see chapter 2).

To compensate the lack of accuracy of the spectrum estimation performed by the filterbank, a Fast Fourier Transformation is done parallel to the process of filtering the audio signal. The used FFT calculates 512 spectral components in a time-intervals of 24ms and with a spectral resolution of 46Hz. This analysis is necessary only in the encoder and it should be noticed, that the spectral values are not transmitted, the basic principle keeps to subband coding. The information about the temporal structure of the audio signal can be derived from the broadband and/or filtered audio signal, as well as from the scalefactors of the individual subbands. The spectral values are evaluated and combined with the scalefactors to estimate the ears masking thresholds in time and frequency with a very high accuracy for a certain time-slot /15,16/. Different combinations of transform length, scalefactor blocklength and bit-allocation time-slot have been investigated. Due to the fact, that the transform analysis is only responsible for the frequency resolution and that the spectrum of audio signals normally doesn't change within few milliseconds, it is completely sufficient to use a windowed 1024-FFT(21.3ms), which is calculated each 24ms. This timing was confirmed by statistical investigations, which have been carried out regarding the resulting bit-allocation. The result was a distinctive propability of only small differences of successive bit-alloactions. Unlike this, high differences were found in the seldom cases of coding the attack segments of audio signals like triangle or castagnettes. The information about the maximum level in the individual subbands can't be derived in each case from the inert transform, but from the scalefactors, providing a time resolution of 8ms.

The calculated masking threshold for one time-slot and the resulting multiplex-signal are shown in figure 5. As shown in chapter 2, the bit-allocation can be derived from the difference of the maximum level and the minimum of the masking threshold in each subband (signal-to-mask ratio). It determines the available bits per subband for quantizing. No transmitting capacity is neccessary for those subbands, which are completely masked by important components of adjacent subbands.

Because the total bit-rate of the encoder should be constant, e.g. 128 kbit/s, the dynamically-varying bit-rate margin can be used for different purposes /2/:

- Increased mask/noise-ratio
- Dynamic bit-error protection
- Transmission of non time critical arbitrary data

At the output of the encoder the multiplex-signal, which consists of the coded bit-allocation, scalefactors, subband samples and error-protection is transmitted.

3.2 Decoder

The blockdiagram of the MUSICAM-decoder is shown in figure 3. After separating the side-information, that means the bit-allocation and the scalefactors from the received multiplex-signal, the 12 successive samples of one block of each subband are decoded. Using the inverse filter network, which was used in the encoder, the complete broadband audio signal can be reconstructed. The decoding process needs significant less computation power than the encoding process. A real time hardware using either one Motorola DSP 56001 or one AT&T DSP32C has been realized. Due to the fact of the simple structure of the filterbank, it is easy to implement the decoder into a special VLSI.

4. Important Side-Information

The side-information consists of the bit-allocation, which gives the decoder the information about the quantization of the individual subbands and the scalefactors, which indicate the level of each subband and exploit optimally the actual dynamic range of the quantizers.

For use in a digital broadcasting system, as well as in contribution links more or less strong distortions of the encoded signal must be taken into account. A codec, which exploits mainly the irrelevancy is much more resistent against bit-errors, than a codec, which reduces most of all the redundancy of the audio signal. The sensitivity due to bit-errors depends on the type of the information and on the data format of the multiplex-signal. There is a significant difference regarding the sensitivity in the order of bit-allocation, scalefactors, MSB's to LSB's of the subband samples. The amount of error-protection is required in the same order /17/. Investigations have shown, that an additional protection of only a few MSB's of lower frequency subbands is sufficient to improve the quality of a disturbed signal.

A bit-rate reduction of the side-information, which needs, for instance, using the Golay-code with the code-rate 1/2 the same amount of bits for error-protection, is very efficient, because the gain of bit-rate is doubled. It increases with the necessary error-protection. Before estimating the gain, a special coding technique of the side-information would provide, it should be noticed, what strategy of error-protection would be necessary for the different parts of the side-information. A error in the bit-allocation, for instance, has to be corrected, because the encoded bitstream of the audio signal can not be decoded without this information for one whole time-slot. Considering disturbed scalefactors, only a recognition of bit-errors is necessary, because there exist different possibilities for error-concealment/18/. Therefore only a parity-bit for each scalefactor is sufficient.

4.1 Bit-allocation

First of all, the existing bit-allocations for each subband were diminished to its essentials. It is plausible, that subbands in the lower frequency region need much more different quantizations than higher subbands, which are mainly coded with 0, 3 or 5 steps. This fact is established by considering the combination of the masking effect, which provides a higher masking power at higher than at lower frequencies and the 32 equidistant subbands. Statistical investigations of the bit-allocation were done to exploit the assumed redundancy for an additional data-reduction. Starting point was a bit-allocation which was calculated for a time-slot of 8 ms, the same length as used for one scalefactor frame. The probability density function was determined for three different considerations :

- the absolut value of the bit-allocation for each subband (absolut)
- the differences of adjacent bit-allocations at one time-slot (frequency)
- the differences of successive bit-allocations (time)

The redundancy of the bit-allocation, which can be derived from the distinctive gaussian shape of all three functions, was exploited for a bit-rate reduction by two different methods. An entropy coding technique with variable length of the codewords, proposed by Huffman and a variation of the time-slot for the bit-allocation have been compared regarding the bit-rate and concerning the parameters error-sensitivity, error-protection, complexity and delay. Using differential Huffman-coding for a time-slot of 8ms, the bit-rate of the bit-allocation can be reduced from 10.2kbit/s to a dynamic rate in the range of about 6kbit/s as a mean value. Unlike this advantage, large codeword-tables would be necessary in the encoder and decoder, because the propabilities for differences of successive bit-allocations are quite different for the individual subbands. In addition, the error-protection should be increased, because an bit-error of the bit-allocation would induce an error-spreading over more than one time-slot. With a variation of the time-slot a constant bit-rate of 5.1kbit/s is obtained for 16ms and 3.4kbit/s for 24ms. The consequence of a further extension would be a worse adapting of the quantizing noise. Therefore it seems to be the best compromise to implement a time-slot with a duration of 24ms. No large tables are necessary and a synchronization of 24ms is ensured.

4.2 Scalefactors

The same investigations as described in 4.1 were carried out in addition concerning the scalefactors and nearly the same results were obtained. Normally, a constant decay of the scalefactors towards higher subbands can be found, similar to the typical spectral slope of speech and music signals. Additional there is only a small deviation of the scalefactors between successive frames in each subband. The probability of a difference of successive scalefactors which is more than 2dB is less than 10%. Before discussing different strategies to exploit the probability density functions of the scalefactors, it has to be noticed that a linear coding with 6bits needs only a parity bit as error-protection, because a recognition of errors is more effective than a correction. This has to be considered carefully, because the gain of bit-rate depends on the used error strategy. The bit-rate of the scalefactors could be reduced from an average bit-rate of 14kbit/s to 7kbit/s using differential Huffman-coding. But at the same time the gain in bit-rate reduction could be lost, because the error strategy has to be changed from recognition (parity bit) to correction (Golay, code-rate 1/2). If one of the Huffman-codewords is disturbed, the whole frame can't be decoded.

Another coding strategy for the scalefactors, which exploits additional the irrelevancy of the inert behaviour of the human ear, has been studied. The three successive scalefactors of each subband of one time-slot are considered and classified into certain patterns. Depending on the pattern, one, two or three scalefactors are transmitted together with an additional select information, which indicates the pattern. If there are only small deviations from one scale-factor to the other, only one value has to be transmitted. Considering an attack, for instance triangle, two or all three scale-factors have to be transmitted. The obtained bit-rate is nearly the same as with Huffman-coding, but with the additional decisive advantage, that only a parity-bit is necessary for the error-protection and only small tables have to be used.

5. Listening tests

Subjective tests with MUSICAM-coded sound sequences were carried out to improve the coding system and to establish the quality at different bit-rates. The critical sound material was selected from digitally recorded and mastered CD's, particularly from the SQAM-CD (Sound Quality Assessment Material), a special CD which was produced by the EBU (European Broadcasting Union) for listening tests /19/.

Suitable sequences were selected concerning different requirements:

- time-critical sequences, like triangle, castagnettes or glockenspiel
- tonal sequences, like woman speaker and singer, basson, trumpet, violin and organ
- sequences with a high dynamic range, like the first part of 'Also sprach Zarathustra' (R. Strauss)

The tests were carried out using the bitrates 192, 128 and 96 kbit/s per monophonic channel.

5.1 Test methodology

Two different test procedures, the AB-comparison test and the ABX-test were applied. AB-comparison means the presentation of an ABAB-sequence. A is always the original, the reference and B can be either the original or the coded signal. The subjects have to assess B compared with A using the CCIR-5-grade impairment-scale /20/. The AB-comparison test was carried out for all twelve test sequences with the bit-rates 96 and 128kbit/s. Only a subset, the four most critical items of the AB-tests, was used for the ABX-tests at the bit-rates 192 and 128 kbit/s. The ABX-test is much more critical than the AB-test. After hearing three times the sequence ABX, the subjects have to decide, wether X is identical to A or to B. A can be the original and B the coded or vice versa. Sixteen subjects, most of all experienced, took part in the listening tests. The reproduction of the sequences was done via electrostatic headphones with diffuse-field frequency response according to /21/ and with a maximum sound pressure level of $88 dB_A$.

5.2 Results

The results of the AB-comparison tests are shown in figure 6 for the bit-rates 96kbit/s (white beams) and 128kbit/s (hatched beams). The beams indicate the mean value of all 16 subjects for each test sequence. Below, the standard deviations of the individual results are shown. In addition, the evaluation of the original signal is represented by the first (dotted) of the three grouped beams. Due to the uncertainty of the subjects a grade of 4.7 results for the original. The audio quality at a bit-rate of 128kbit/s with a grade of 4.6 as a mean value, also at a bit-rate of 96kbit/s with 4.3 grade is nearly the same as the quality of the original. The difference of 0.1 comparing the standard deviation of the original (0.6) with the 128kbit/s-coded signal (0.7) is negligible. A relative high standard deviation of 1.1 grade is obtained with 96kbit/s and it should be noticed, that there are some very critical items, such as no. 8 (soprano soloist), no. 9 (female pop vocalist) and no. 10 (violin), which show a significant degradation in quality.

The results of the ABX-tests are shown in figure 7 for the bit-rates 192kbit/s and 128kbit/s. The beams indicate the number of correct decisions. For both bit-rates the 0.05 significance level α is shown. The beams of the four signals are below this α-level for both bit-rates. That means, that there is no audible difference between the original and the coded signal.

6. Conclusion

New strategies developing a very flexible audio source coding system have been proposed. The combination of two analysis-methods, subband and transform, gives very good results concerning the important parameters quality, bit-rate, delay, complexity and error-sensitivity. Listening tests have shown that there are still a few very critical signals, which show a degradation of the quality at low bit-rates. Therefore two important parts have to be investigated in the future. Different parallel and serial combinations of transform and subbands should be investigated. In addition the correlation between the two channels of a stereophonic program, the stereophonic irrelevancy has to be exploited for a transparent coding at very low bit-rates without degradation of quality.

Bibliograhical references

/1/ Theile, G., Stoll, G. und Link, M.: Low bit-rate coding of high quality audio signals. An introduction to the MASCAM system. EBU Review-Technical No.230, August 1988, pp. 158-181.

/2/ Stoll, G., Link, M. und Theile, G.: Masking-pattern adapted subband coding: use of the dynamic bit-rate margin. 84th AES Convention, 1988, Paris, Preprint No.2585.

/3/ Veldhuis, R.N.J., Breeuwer, M. and van der Waal, R.G.: Subband coding of digital audio signals. Philips Journal of Research, 44(2-3), 1989, pp.329-343.

/4/ Dehery, Y. F.: Real-time software processing approach for digital sound broadcasting. Advanced digital techniques for UHF satellite sound broadcasting, EBU August 1988, pp. 95-99.

/5/ Wiese, D.: Optimierung eines Teilbandcodierverfahrens im Hinblick auf die Anwendung bei verschiedenen audio-visuellen Systemen (in german). 'Kleinheubacher Berichte 1989', pp. 613-622.

/6/ Stoll, G. and Wiese, D.: Optimizing of a subband coder for high quality audio signals by considering the ears masking thresholds. Presentation at the IEEE ASSP Workshop, New Paltz, New York, October 1989.

/7/ Johnston, J.: Transform Coding of Audio Signals Using Perceptual Noise Criteria. IEEE Journal on selected areàs in communication, Vol.6, No.2, February 1988, pp. 314-323.

/8/ Krahé, D.: Ein Verfahren zur Datenreduktion bei digitalen Audiosignalen unter Ausnutzung psychoakustischer Phänomene (in german). Rundfunktechnische Mitteilungen 30, 1986, pp. 117-123.

/9/ Brandenburg, K.: High quality sound coding at 2.5 bit/sample. 84th AES Convention, 1988, Paris, Preprint No. 2582.

/10/ Brandenburg, K. and Johnston, J.D.: Second Generation Perceptual Audio Coding: The Hybrid Coder. 88th AES Convention, 1990, Montreux, Preprint No. 2937.

/11/ Terhardt, E., Stoll, G. und Seewann, M.: Algorithm for extraction of pitch and pitch salience from complex tonal signals. Journal of the Acoustical Society, AM71, 1982, pp.679-688.

/12/ Fettweis, A.: Wave digital filters: Theory and practice. Proceedings of the IEEE, Vol. 74, 1986, pp. 270-327.

/13/ Spille, J. und Schröder, E. F.: Vergleich mehrerer für eine Teilbandcodierung geeigneter Filterfamilien unter besonderer Berücksichtigung der Wellendigitalfilter (in german). ITG-Fachbericht 106, Mainz 1988, pp. 229-234.

/14/ Rothweiler, J.H.: Polyphase Quadrature Filters - A new Subband Coding Technique. International Conference IEEE ASSP 1983, Boston, pp. 1280-1283.

/15/ Zwicker, E. und Feldtkeller, R.: Das Ohr als Nachrichtenempfänger (in german). S. Hirzel Verlag, Stuttgart, 1967.

/16/ Fastl, H.: Beschreibung dynamischer Hörempfindungen anhand von Mithörschwellen-Mustern (in german). Hochschulverlag Freiburg 1982.

/17/ Weck, Ch. und Theile, G.: Digital audio broadcasting: Optimizing of a combined concept of source and channel coding with respect of subjective criteria. International Broadcasting Conference, Brighton, September 1988.

/18/ Stoll, G. and Theile, G.: Terrestrischer digitaler Hörrundfunk - Optimierung der Quellencodierung im Hinblick auf eine effektive Übertragung durch Berücksichtigung der Gehöreigenschaften (in german). Bericht der 15.Tonmeistertagung, Mainz, 1988.

/19/ Stoll, G. and Dehery, Y.: High quality bitrate reduction system family for different applications. IEEE International Conference on Communications, April 1990, Nr.322.3, pp. 937-941.

/20/ Jakubowski, H. and Spikofski, G.: SQAM - The EBU Compact Disc for subjective assessments of audio systems. EBU Review, Technical No.227, February 1988.

/21/ Subjective assessment of sound quality. CCIR Recommendation 565-2; XVIth Plenary Assembly, Dubrovnik 1986, Vol.X, Part1, pp. 275-279.

/22/ Spikofski, G.: The diffuse-field probe transfer function of studio-quality headphones, EBU Review, Technical No. 229, pp. 111-126.

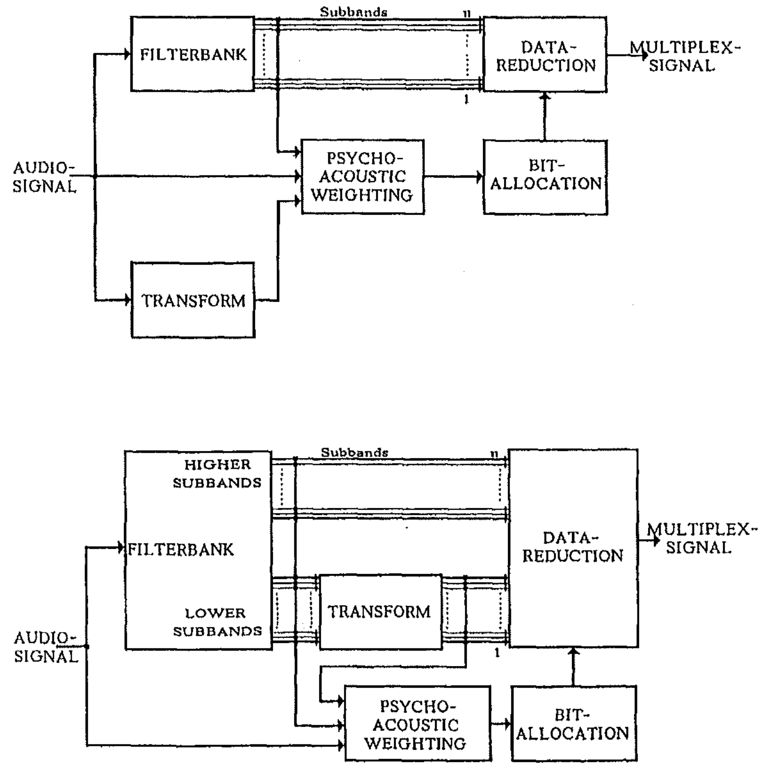

Fig. 1 :
Parallel and serial combination of filterbank and transform integrated in a typical encoder

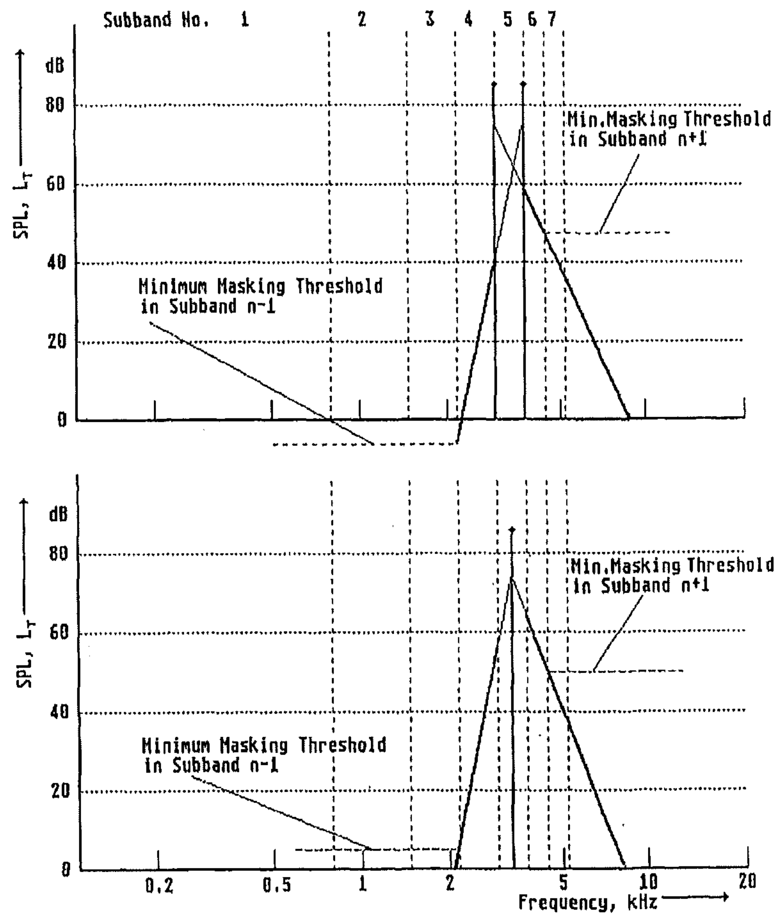

Fig. 2 a and b :
Comparison of the masking threshold estimation using scalefactors or spectral values for outband masking

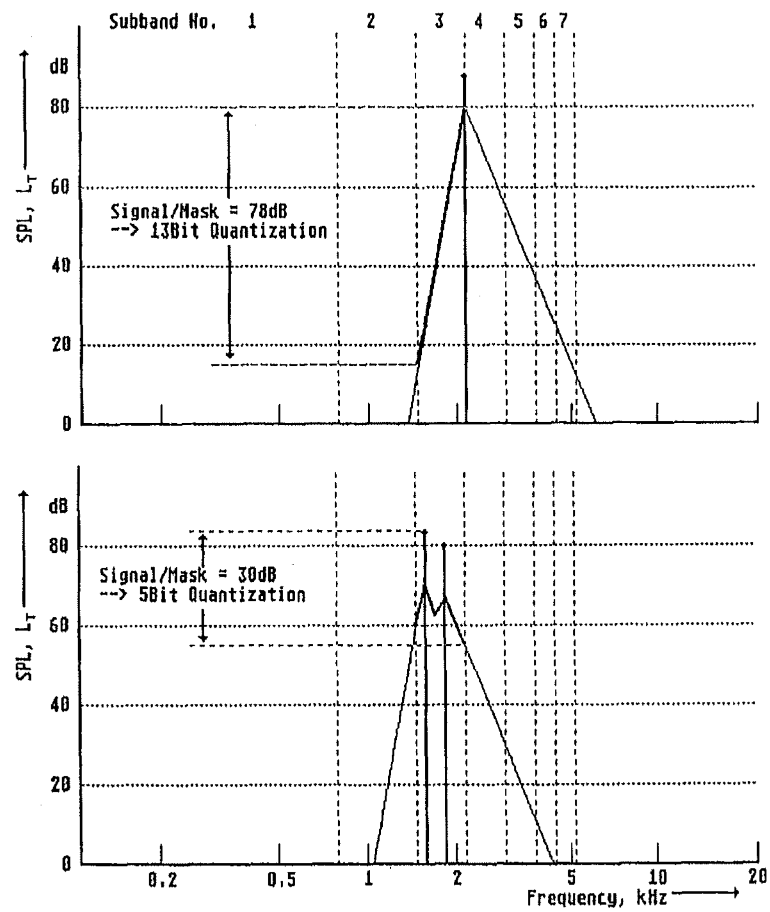

Fig. 2c and d :
Comparison of the masking threshold estimation using scalefactors or spectral values for inband masking

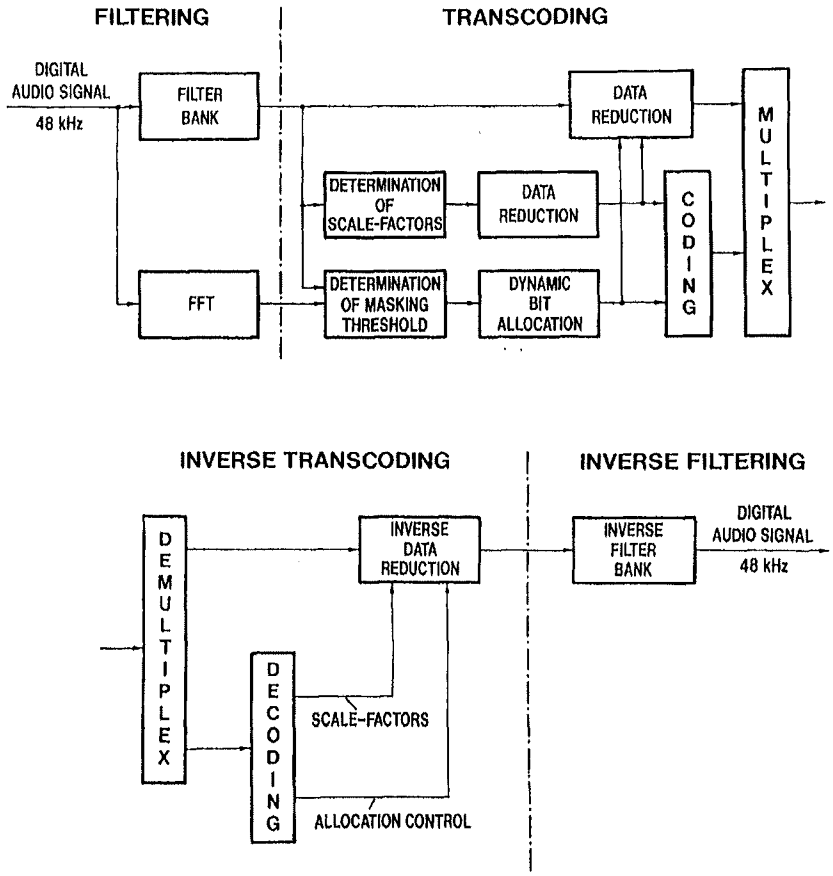

Fig. 3 :
Blockdiagram of the MUSICAM-encoder (above) and -decoder (below)

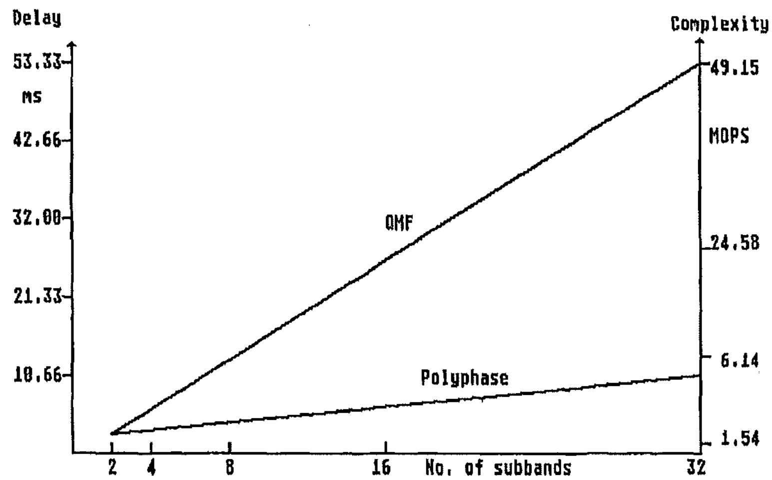

Fig. 4 :
Comparison of Quadrature-Mirror-Filters and Polyphase-Filters

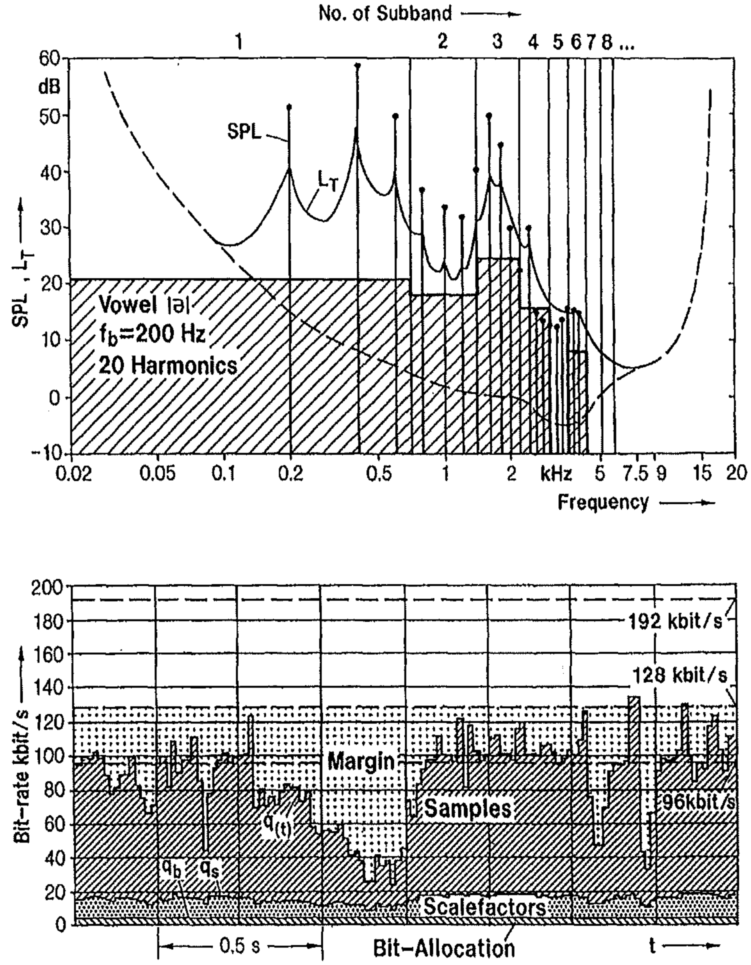

Fig. 5 :
Amplitude-spectrum, masking-pattern, imperceptible quantizing noise (above) and resulting bit-rate of the multiplex-signal (below)

Exhibit P Page 360

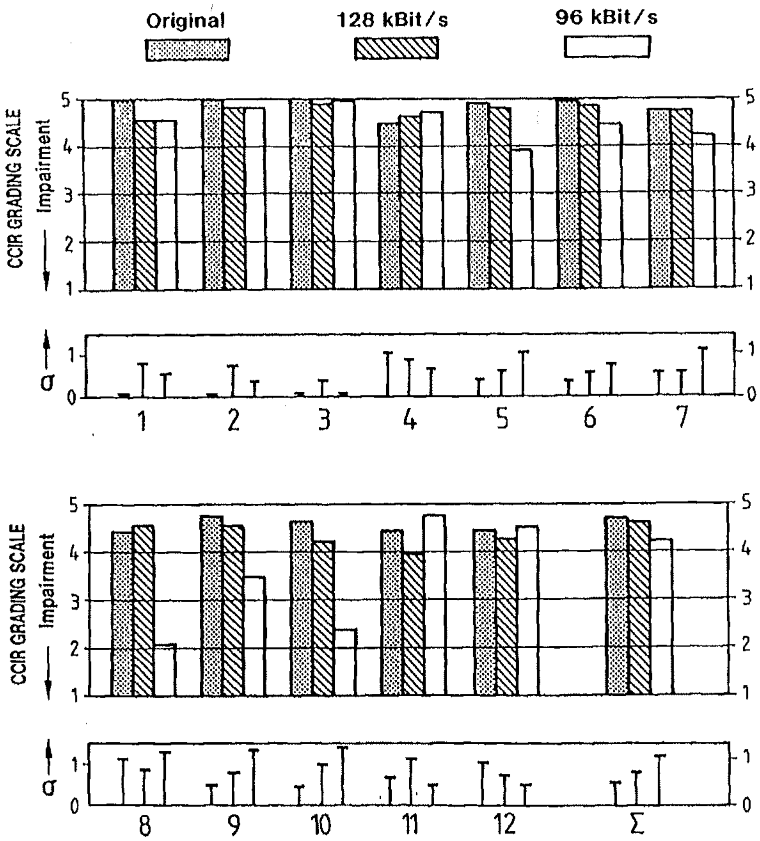

Fig. 6 :
Results of AB-comparison-listening-tests for twelve critical audio-signals

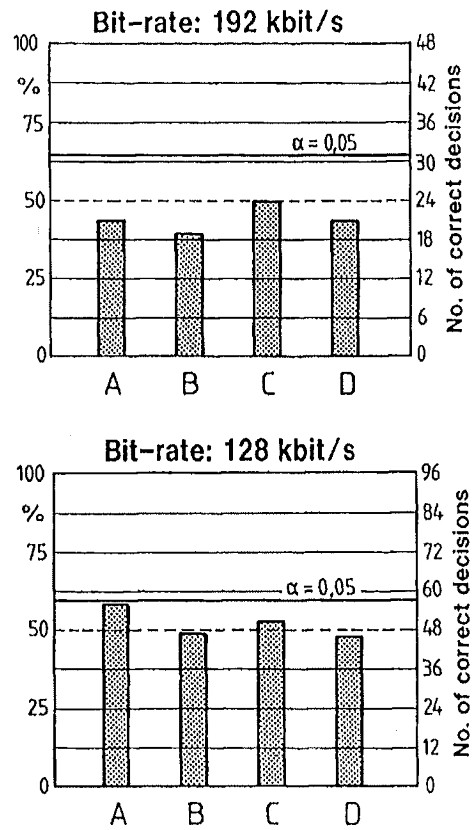

Fig. 7 :
Results of ABX-tests for four very critical audio-signals
(A: triangle, B: female singer, C: violin, D: female speech)