Claims

1. A time scaler for providing a time scaled version of an input audio signal,

wherein the time scaler is configured to compute or estimate a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and

wherein the time scaler is configured to perform the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;

wherein the time scaler is configured to time-shift a second block of samples with respect to a first block of samples, and to overlap-and-add the first block of samples and the time-shifted second block of samples, to thereby obtain the time-scaled version of the input audio signal, if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates a quality which is larger than or equal to a quality threshold value; and

wherein the time scaler is configured to determine a time shift of the second block of samples with respect to the first block of samples in dependence on a determination of a level of similarity, evaluated using a first similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples,

wherein the determined time shift is an information describing a position of highest similarity; and

wherein the time scaler is configured to compute or estimate a quality of the time scaled version of the input audio signal obtainable by a time scaling of the input audio signal on the basis of an information about the level of similarity~~, evaluated using a second similarity measure,~~ between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the

determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift,

wherein the information about the level of similarity between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift, is evaluated using a second similarity measure.

~~2. The time scaler according to claim 1, wherein the time scaler is configured to perform an overlap-and-add operation using a first block of samples of the input audio signal and a second block of samples of the input audio signal,~~

~~wherein the time scaler is configured to time-shift the second block of samples with respect to the first block of samples, and to overlap-and-add the first block of samples and the time-shifted second block of samples, to thereby obtain the time-scaled version of the input audio signal.~~

~~3~~2. The time-scaler according to claim ~~2~~1, wherein the time scaler is configured to compute or estimate a quality of the overlap-and-add operation between the first block of samples and the time-shifted second block of samples, in order to compute or estimate the quality of the time scaled version of the input audio signal obtainable by the time scaling.

~~4. The time scaler according to claim 2 or claim 3, wherein the time scaler is configured to determine the time shift of the second block of samples with respect to the first block of samples in dependence on a determination of a level of similarity between the first block of samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples.~~

~~5~~3. The time scaler according to claim ~~4~~1 or 2, wherein the time scaler is configured to determine an information about a level of similarity between the first block of

samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples, for a plurality of different time shifts between the first block of samples and the second block of samples, and to determine a time shift to be used for the overlap-and-add operation on the basis of the information about the level of similarity for the plurality of different time shifts.

~~6~~4. The time scaler according to ~~claim 4 or claim 5~~any one of claims 1 to 3, wherein the time scaler is configured to determine the time shift of the second block of samples with respect to the first block of samples, which time shift is to be used for the overlap-and-add operation, in dependence on a target time shift information.

~~7. The time scaler according to one of claims 4 to 6, wherein the time scaler is configured to compute or estimate a quality of the time scaled version of the input audio signal obtainable by a time scaling of the input audio signal on the basis of an information about the level of similarity between the first block of samples, or a portion of the first block of samples, and the second block of samples, time shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift.~~

~~8~~5. The time scaler according to ~~claim 7~~any one of claims 1 to 4, wherein the time scaler is configured to decide, on the basis of the information about the level of similarity between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift, whether a time scaling is actually performed.

~~9. The time scaler according to one of claims 1 to 8, wherein the second similarity measure is computationally more complex than the first similarity measure.~~

106. The time scaler according to any one of claims 1 to 95, wherein the first similarity measure is a cross correlation or a normalized cross correlation, or an average magnitude difference function or a sum of squared errors, and

wherein the second similarity measure is a combination of a cross correlations or of normalized cross correlations for a plurality of different time shifts.

117. The time scaler according to any one of claims 1 to 106, wherein the second similarity measure is a combination of cross correlations for at least four different time shifts.

128. The time scaler according to claim 117, wherein the second similarity measure is a combination of a first cross correlation value and of a second cross correlation value, which are obtained for time shifts which are spaced by an integer multiple of a period duration of a fundamental frequency of an audio content of the first block of samples or of the second block of samples, and of a third cross correlation value and a fourth cross correlation value, which are obtained for time shifts which are spaced by an integer multiple of the period duration of the fundamental frequency of the audio content,

wherein a time shift for which the first cross correlation value is obtained is spaced from a time shift for which the third cross correlation value is obtained, by an odd multiple of half the period duration of the fundamental frequency of the audio content.

139. The time scaler according to any one of claims 1 to 128, wherein the second similarity measure q is obtained according to

$$q=c(p)*c(2*p)+c(3/2*p)*c(1/2*p)$$

or according to

$$q=c(p)*c(-p)+c(-1/2*p)*c(1/2*p),$$

wherein c(p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by a period duration p of a

fundamental frequency of an audio content of the first block of samples or of the second block of samples;

wherein c(2*p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by 2*p;

wherein c(3/2*p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by 3/2*p;

wherein c(1/2*p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by ½*p;

wherein c(-p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by -p; and

wherein c(-1/2*p) is a cross correlation value between a first block of samples and a second block of samples, which are shifted in time by -½*p.

~~14~~10. The time scaler according to any one of claims 1 to ~~13~~9,

wherein the time scaler is configured to compare a quality value, which is based on a computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling, with a variable threshold value, to decide whether a time scaling should be performed or not.

~~15~~11. The time scaler according to claim ~~14~~10, wherein the time scaler is configured to reduce the variable threshold value, to thereby reduce a quality requirement, in response to a finding that a quality of a time scaling would have been insufficient for one or more previous blocks of samples.

~~16~~12. The time scaler according to ~~claim 14 or claim 15~~claim 10 or claim 11, wherein the time scaler is configured to increase the variable threshold value, to thereby

increase a quality requirement, in response to the fact that a time scaling has been applied to one or more previous blocks of samples.

~~17~~13. The time scaler according to any one of claims ~~14 to 16~~10 to 12,

wherein the time scaler comprises a range-limited first counter for counting a number of blocks of samples or a number of frames which have been time scaled because a respective quality requirement of the time scaled version of the input audio signal obtainable by the time scaling has been reached, and

wherein the time scaler comprises a range-limited second counter for counting a number of blocks of samples or a number of frames which have not been time-scaled because a respective quality requirement of the time scaled version of the input audio signal obtainable by the time scaling has not been reached; and

wherein the time scaler is configured to compute the variable threshold value in dependence on a value of the first counter and in dependence on a value of the second counter.

~~18~~14. The time scaler according to claim ~~17~~13, wherein the time scaler is configured to add a value which is proportional to the value of the first counter to an initial threshold value, and to subtract a value which is proportional to the value of the second counter therefrom, in order to obtain the variable threshold value.

~~19~~15. The time scaler according to any one of claims 1 to ~~18~~14, wherein the time scaler is configured to perform the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling, wherein the computation or estimation of the quality of the time scaled version of the input audio signal comprises an computation or estimation of artifacts in the time scaled version of the input audio signal which would be caused by a time scaling.

~~20~~16. The time scaler according to claim ~~19~~15, wherein the computation or estimation of the quality of the time scaled version of the input audio signal comprises an computation or estimation of artifacts in the time scaled version of the input audio

signal which would be caused by an overlap-and-add operation of subsequent blocks of samples of the input audio signal.

~~21~~17. The time scaler according to any one of claims 1 to ~~20~~16, wherein the time scaler is configured to compute or estimate the quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal in dependence on a level of similarity of subsequent blocks of samples of the input audio signal.

~~22~~18. The time scaler according to any one of claims 1 to ~~21~~17, wherein the time scaler is configured to compute or estimate whether there are audible artifacts in a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal.

~~23~~19. The time scaler according to any one of claims 1 to ~~22~~18, wherein the time scaler is configured to postpone a time scaling to a subsequent frame or to a subsequent block of samples if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates an insufficient quality.

~~24~~20. The time scaler according to any one of claims 1 to ~~23~~19, wherein the time scaler is configured to postpone a time scaling to a time when the time scaling is less audible if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates an insufficient quality.

~~25. The time scaler according to one of claims 1 to 24, wherein the second similarity measure provides more accuracy than the first similarity measure.~~

~~26~~21. The time scaler according to any one of claims 1 to ~~25~~20, wherein the first similarity measure is a cross correlation or a normalized cross correlation, or an average magnitude difference function or a sum of squared errors.

~~27. A time scaler for providing a time scaled version of an input audio signal,~~

~~wherein the time scaler is configured to compute or estimate a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and~~

~~wherein the time scaler is configured to perform the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;~~

~~wherein the time scaler is configured to compare a quality value, which is based on a computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling, with a variable threshold value, to decide whether a time scaling should be performed or not;~~

~~wherein the time scaler is configured to increase the variable threshold value, to thereby increase a quality requirement, in response to the fact that a time scaling has been applied to one or more previous blocks of samples, such that it is ensured that subsequent blocks of samples are only time scaled if a comparatively high quality level, higher than a normal quality level, can be reached.~~

~~28~~22. An audio decoder for providing a decoded audio content on the basis of an input audio content, the audio decoder comprising:

a jitter buffer configured to buffer a plurality of audio frames representing blocks of audio samples;

a decoder core configured to provide blocks of audio samples on the basis of audio frames received from the jitter buffer; and

a sample-based time scaler according to one of claims 1 to ~~27~~21, wherein the sample-based time scaler is configured to provide time-scaled blocks of audio samples on the basis of blocks of audio samples provided by the decoder core.

2923. The audio decoder according to claim 2822, wherein the audio decoder further comprises a jitter buffer control,

wherein the jitter buffer control is configured to provide a control information to the sample-based time scaler, wherein the control information indicates whether a sample-based time scaling should be performed or not, and/or wherein the control information indicates a desired amount of time scaling.

3024. A method for providing a time scaled version of an input audio signal,

wherein the method comprises computing or estimating a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and

wherein the method comprises performing the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;

wherein the method comprises time-shifting a second block of samples with respect to a first block of samples, and to overlap-and-add the first block of samples and the time-shifted second block of samples, to thereby obtain the time-scaled version of the input audio signal, if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates a quality which is larger than or equal to a quality threshold value; and

wherein the method comprises determining a time shift of the second block of samples with respect to the first block of samples in dependence on a determination of a level of similarity, evaluated using a first similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples,

wherein the determined time shift is an information describing a position of highest similarity; and

wherein the method comprises computing or estimating a quality of the time scaled version of the input audio signal obtainable by a time scaling of the input audio signal on the basis of an information about the level of similarity, evaluated using a second similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift.

~~31. A method for providing a time scaled version of an input audio signal,~~

~~wherein the method comprises computing or estimating a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and~~

~~wherein the method comprises performing the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;~~

~~wherein the method comprises comparing a quality value, which is based on a computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling, with a variable threshold value, to decide whether a time scaling should be performed or not;~~

~~wherein the method comprises increasing the variable threshold value, to thereby increase a quality requirement, in response to the fact that a time scaling has been applied to one or more previous blocks of samples such that it is ensured that subsequent blocks of samples are only time scaled if a comparatively high quality level, higher than a normal quality level, can be reached.~~

~~32~~25. A computer program for performing the method according to claim ~~30 or 31~~ 24 when the computer program is running on a computer.

~~33~~26. A time scaler for providing a time scaled version of an input audio signal,

wherein the time scaler is configured to compute or estimate a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and

wherein the time scaler is configured to perform the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;

wherein the time scaler is configured to time-shift a second block of samples with respect to a first block of samples, and to overlap-and-add the first block of samples and the time-shifted second block of samples, to thereby obtain the time-scaled version of the input audio signal, if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates a quality which is larger than or equal to a quality threshold value; and

wherein the time scaler is configured to determine a time shift of the second block of samples with respect to the first block of samples in dependence on a determination of a level of similarity, evaluated using a first similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples; and

wherein the time scaler is configured to compute or estimate a quality of the time scaled version of the input audio signal obtainable by a time scaling of the input audio signal on the basis of an information about the level of similarity, evaluated using a second similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift;

wherein the first similarity measure is a cross correlation or a normalized cross correlation, or an average magnitude difference function or a sum of squared errors, and

wherein the second similarity measure is a combination of a cross correlations or of normalized cross correlations for a plurality of different time shifts; or

wherein the second similarity measure is a combination of cross correlations for at least four different time shifts.

3427. A method for providing a time scaled version of an input audio signal,

wherein the method comprises computing or estimating a quality of a time scaled version of the input audio signal obtainable by a time scaling of the input audio signal, and

wherein the method comprises performing the time scaling of the input audio signal in dependence on the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling;

wherein the method comprises time-shifting a second block of samples with respect to a first block of samples, and to overlap-and-add the first block of samples and the time-shifted second block of samples, to thereby obtain the time-scaled version of the input audio signal, if the computation or estimation of the quality of the time scaled version of the input audio signal obtainable by the time scaling indicates a quality which is larger than or equal to a quality threshold value; and

wherein the method comprises determining a time shift of the second block of samples with respect to the first block of samples in dependence on a determination of a level of similarity, evaluated using a first similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, or a portion of the second block of samples; and

wherein the method comprises computing or estimating a quality of the time scaled version of the input audio signal obtainable by a time scaling of the input audio signal on the basis of an information about the level of similarity, evaluated using a

second similarity measure, between the first block of samples, or a portion of the first block of samples, and the second block of samples, time-shifted by the determined time shift, or a portion of the second block of samples, time-shifted by the determined time shift;

wherein the first similarity measure is a cross correlation or a normalized cross correlation, or an average magnitude difference function or a sum of squared errors, and

wherein the second similarity measure is a combination of a cross correlations or of normalized cross correlations for a plurality of different time shifts; or

wherein the second similarity measure is a combination of cross correlations for at least four different time shifts.

~~35~~28. A computer program for performing the method according to claim ~~34~~ 27 when the computer program is running on a computer.