Claims

1. Apparatus for estimating an inter-channel time difference between a first channel signal and a second channel signal, comprising:

a signal analyzer for estimating a signal characteristic of the first channel signal or the second channel signal or both signals or a signal derived from the first channel signal or the second channel signal;

a calculator for calculating a cross-correlation spectrum for a time block from the first channel signal in the time block and the second channel signal in the time block;

a weighter for weighting a smoothed or non-smoothed cross-correlation spectrum to obtain a weighted cross correlation spectrum using a first weighting procedure or using a second weighting procedure depending on a signal characteristic estimated by the signal analyzer, wherein the first weighting procedure is different from the second weighting procedure; and

a processor for processing the weighted cross-correlation spectrum to obtain the inter-channel time difference.

2. Apparatus of claim 1, wherein the signal analyzer is configured as a noise estimator for estimating a noise level of the first channel signal or the second channel signal or both signals or a signal derived from the first channel signal or the second channel signal, and wherein a first signal characteristic is a first noise level and a second signal characteristic is a second noise level, or wherein the signal analyzer is configured to perform a speech/music analysis, an interfering-talker analysis, a background music analysis, a clean speech analysis or any other signal analysis in order to determine, whether a signal has a first characteristic or a second characteristic.

3. Apparatus of claim 1, wherein the first weighting procedure is selected for a first signal characteristic and the second weighting procedure is selected for a second signal characteristic, and wherein the first signal characteristic is different from the second signal characteristic.

4. Apparatus of one of the preceding claims, wherein the first weighting procedure comprises a weighting so that an amplitude is normalized and a phase is maintained, or wherein the second weighting procedure comprises a weighting factor derived from the smoothed or non-smoothed cross-correlation spectrum using a power operation having a power being lower than 1 or greater than 0 or a log function.

5. Apparatus of one of the preceding claims, wherein the second weighting procedure comprises a weighting so that an amplitude is normalized and a phase is maintained and additionally comprises a weighting factor derived from the smoothed or non-smoothed cross-correlation spectrum using a power operation having a power being lower than 1 or greater than 0 or between 0.79 and 0.82.

6. Apparatus of one of the preceding claims, wherein the first weighting procedure operates in accordance with the following equation:

$$\tilde{C}_{PHAT}(k,s) = \frac{\tilde{C}(k,s)}{|\tilde{C}(k,s)|}\text{ , or}$$

wherein the second weighting procedure operates in accordance with the following equation:

$$\tilde{C}_{MCSP}(k,s) = \frac{\tilde{C}(k,s)}{|\tilde{C}(k,s)|^{\rho}},$$

wherein $\tilde{C}_{PHAT}(k,s)$ is a weighted smoothed or non-smoothed cross correlation spectrum value for a frequency index k and a time index s obtained by applying the first weighting procedure,

wherein $\tilde{C}_{MCSP}(k,s)$ is a weighted smoothed or non-smoothed cross correlation spectrum value for a frequency index k and a time index s obtained by applying the second weighting procedure,

wherein $\tilde{C}(k,s)$ is a smoothed or non-smoothed cross correlation spectrum value for a frequency index k and a time index s, and

wherein ρ is a power value being different from 1.

7. Apparatus of one of the preceding claims, wherein the second weighting procedure comprises a normalization so that an output range of a second normalization procedure is in a range in which an output range of a first normalization procedure is positioned, or so that the output range of the second normalization procedure is the same as an output range of the first normalization procedure.

8. Apparatus of one of the preceding claims, wherein the second weighting procedure comprises a normalization based on the following equation:

$$\tilde{C}_{MCSP}(k,s) = \frac{\tilde{C}_{MCSP}(k,s)}{\frac{1}{N_{DFT}}\sum_{k=0}^{N_{DFT}-1}|\tilde{C}_{MCSP}(k,s)|},$$

wherein $\tilde{C}_{MCSP}(k,s)$ on the left side of the above equation represents a normalized cross correlation result for a certain k and a certain s, wherein $\tilde{C}_{MCSP}(k,s)$ on the right side of the above equation represents a cross correlation result before normalization for a certain k and a certain s, and wherein s is a frame index, k is a frequency index, and N_{DFT} is a DFT length.

9. Apparatus of one of the preceding claims, wherein the processor is configured to perform a first peak picking operation or a second peak picking operation depending on whether the first weighting procedure or the second weighting procedure has been used, wherein the first peak picking operation is different from the second peak picking operation.

10. Apparatus of claim 9, wherein the second peak picking operation is used when the second weighting procedure is used, and wherein the second peak picking operation is configured to apply a second threshold being lower than a first threshold used by the first peak picking operation.

11. Apparatus of claim 2 ~~and~~ or any one of claims ~~3~~ 4 to 10 when referring back to claim,

wherein the noise estimator is configured to estimate a level of a background noise or is configured to smooth an estimated noise level over time or is configured to use an IIR smoothing filter.

12. Apparatus of claim 2 ~~and~~ or any one of claims ~~3~~ 4 to 11 when referring back to claim 2,

wherein the noise estimator further comprises a signal activity detector for classifying the time block as active or inactive, wherein the noise estimator is configured to compute a signal level using one or more active time blocks, or wherein the noise estimator is configured to signal a high background noise level, when a signal to noise ratio is below a threshold, the threshold being in a range between 45 to 25 dB.

13. Apparatus of one of claims 1 to 12, further comprising:

a spectral characteristic estimator for estimating a characteristic of a spectrum of the first channel signal or the second channel signal for the time block;

a smoothing filter for smoothing the cross-correlation spectrum over time using the spectral characteristic to obtain a smoothed cross-correlation spectrum, and wherein the weighter is configured for weighting the smoothed cross-correlation spectrum.

14. Apparatus of one of claim 1 to 13,

wherein the processor is configured to normalize the smoothed cross-correlation spectrum using a magnitude of the smoothed cross-correlation spectrum.

15. Apparatus of one claims 1 to 14,

wherein the processor is configured

to calculate a time-domain representation of the smoothed cross-correlation spectrum or a normalized smoothed cross-correlation spectrum; and

to analyze the time-domain representation to determine the inter-channel time difference.

16. Apparatus of claim 15,

wherein the processor is configured to low-pass filter the time-domain representation and to further process a result of the low-pass filtering.

17. Apparatus of one of the preceding claims,

wherein the processor is configured to perform the inter-channel time difference determination by performing a peak searching or peak picking operation within a time-domain representation determined from the smoothed cross-correlation spectrum.

18. Apparatus of one of the claims 13 to 17,

wherein the spectral characteristic estimator is configured to determine, as the spectral characteristic, a noisiness or a tonality of the spectrum; and

wherein the smoothing filter is configured to apply a stronger smoothing over time with a first smoothing degree in case of a first less noisy characteristic or a first more tonal characteristic, or to apply a weaker smoothing over time with a second smoothing degree in case of a second more noisy characteristic or a second less tonal characteristic,

wherein the first smoothing degree is greater than the second smoothing degree, and wherein the first noisy characteristic is less noisy than the second noisy characteristic, or the first tonal characteristic is more tonal than the second tonal characteristic.

19. Apparatus of one of the claims 13 to 18,

wherein the spectral characteristics estimator is configured to calculate, as the spectral characteristic, a first spectral flatness measure of a spectrum of the first channel signal and a second spectral flatness measure of a second spectrum of the second channel signal, and to determine the spectral characteristic from the first spectral flatness measure and the second spectral flatness measure by selecting a maximum value from the first spectral flatness measure and the second spectral flatness measure, by determining a weighted average or an unweighted average between the first spectral flatness measure and the second spectral flatness measure, or by selecting a minimum value from the first spectral flatness measure and the second spectral flatness measure.

20. Apparatus of one of the claims 13 to 19,

wherein the smoothing filter is configured to calculate a smoothed cross-correlation spectrum value for a frequency by a weighted combination of the cross-correlation spectrum value for the frequency from the time block and a cross-correlation spectral value for the frequency from at least one past time block, wherein weighting factors for the weighted combination are determined by the characteristic of the spectrum.

21. Apparatus of one of the preceding claims,

wherein the processor is configured to determine a valid range and an invalid range within a time-domain representation derived from the weighted smoothed or non-smoothed cross-correlation spectrum,

wherein at least one maximum peak within the invalid range is detected and compared to a maximum peak within the valid range, wherein the inter-channel time difference is only determined, when the maximum peak within the valid range is greater than at least one maximum peak within the invalid range.

22. Apparatus of one of the preceding claims,

wherein the processor is configured

to perform a peak search operation within a time-domain representation derived from the smoothed cross-correlation spectrum,

to determine a variable threshold from the time-domain representation; and

to compare a peak to the variable threshold, wherein the inter-channel time difference is determined as a time lag associated with a peak being in a predetermined relation to the variable threshold.

23. Apparatus of claim 22,

wherein the processor is configured to determine the variable threshold as a value being equal to an integer multiple of a value among a largest portion such as 10 % of values of the time-domain representation.

24. Apparatus of one of claims 1 to 23,

wherein the processor is configured to determine a maximum peak amplitude in each subblock of a plurality of subblocks of a time-domain representation derived from the smoothed cross-correlation spectrum,

wherein the processor is configured to calculate a variable threshold based on a mean peak magnitude derived from the maximum peak magnitudes of the plurality of subblocks, and

wherein the processor is configured to determine the inter-channel time difference as a time lag value corresponding to a maximum peak of the plurality of subblocks being greater than the variable threshold.

25. Apparatus of claim 24,

wherein the processor is configured to calculate the variable threshold by a multiplication of the mean peak magnitude by a value, the mean peak magnitude being determined as an average of the maximum peak magnitudes of the plurality of subblocks,

wherein the value is determined by an SNR characteristic of the first channel signal and the second channel signal, wherein a first value is associated with a first SNR value and a second value is associated with a second SNR value, wherein the first value is greater than the second value, and wherein the first SNR value is greater than the second SNR value.

26. Apparatus of claim 25,

wherein the processor ~~-~~is configured to use a third value being lower than the second value in case of a third SNR value being lower than the second SNR value and when a difference between the threshold and a maximum peak is lower than a predetermined value.

27. Apparatus of claim 2 ~~and~~ or any one of claims ~~3~~ 4 to 26 when referring back to claim 2, wherein the noise estimator comprises a background noise estimator and a time-smother for providing a background noise estimate, or

wherein the noise estimator ~~comprise~~ comprises a signal activity detector, a frame selector for selecting an active frame only under control of the signal activity detector, and a signal level calculator for calculating a signal level in the active frame, and a time smoother for smoothing a result of the signal level calculator over time to provide a signal level estimate, or

wherein the noise estimator is configured to calculate a signal to noise ratio from a smoothed or non-smoothed signal level and a smoothed or non-smoothed background noise level for a frame, and a comparator for comparing the signal to noise ratio value to a threshold for the frame in order to provide the noise level for the frame.

28. Apparatus of one of the preceding claims, wherein the apparatus is configured

for performing a storage or a transmission of the estimated inter-channel time difference, or

for performing a stereo or multi-channel processing or encoding of the first and the second channel signals using the estimated inter-channel time difference, or

for performing a time alignment of the two channel signals using the inter-channel time difference, or

for performing a time difference of arrival estimation using the estimated inter-channel time difference, or

for performing a time difference of arrival estimation using the inter-channel time difference for the determination of a speaker position in a room with two microphones and a known microphone setup, or

for performing a beamforming using the estimated inter-channel time difference, or

for performing a spatial filtering using the estimated inter-channel time difference, or

for performing a foreground or background decomposition using the estimated inter-channel time difference, or

for performing a location operation of a sound source using the estimated inter-channel time difference, or

for performing a location of a sound source using the estimated inter-channel time difference by performing an acoustic triangulation based on time differences between the first channel signal and the second channel signal or the first channel signal, the second channel signal and at least one additional signal.

29. Method of estimating an inter-channel time difference between a first channel signal and a second channel signal, the method comprising:

estimating a signal characteristic of the first channel signal or the second channel signal or both signals or a signal derived from the first channel signal or the second channel signal;

calculating a cross-correlation spectrum for a time block from the first channel signal in the time block and the second channel signal in the time block;

weighting a smoothed or non-smoothed cross-correlation spectrum to obtain a weighted cross correlation spectrum using a first weighting procedure or using a second weighting procedure depending on a signal characteristic estimated, wherein the first weighting procedure is different from the second weighting procedure; and

processing the weighted cross-correlation spectrum to obtain the inter-channel time difference.

30. Method of claim 29, further comprising:

estimating a characteristic of a spectrum of the first channel signal or the second channel signal for the time block;

smoothing the cross-correlation spectrum over time using the spectral characteristic to obtain a smoothed cross-correlation spectrum, and wherein the weighting weights the smoothed cross-correlation spectrum.

31. Computer program for performing, when running on a computer or a processor, the method of claim 29 or 30.