Claims:

1. A method for mixing two input audio signals into a single, mixed audio signal while maintaining a perceived sound level of the mixed audio signal, the method comprising:

receiving a main input audio signal;

receiving an associated input audio signal; wherein the associated input audio signal is coupled with the main input audio signal;

receiving mixing metadata, which contains scaling information for scaling the main input audio signal and which specifies how the main input audio signal and the associated input audio signal should be mixed, in order to generate a mixed audio signal at the perceived sound level; wherein the scaling information from the mixing metadata comprises a metadata scale factor for the main input audio signal, for scaling the main input audio signal relative to the associated input audio signal;

receiving a mixing balance input, which denotes an adjustable balance between the main input audio signal and the associated input audio signal, wherein the mixing balance input comprises scaling information which allows a deviation from a weighting of the main input audio signal and the associated input audio signal in the mixed audio signal as specified in the mixing metadata;

identifying a dominant signal as either the main input audio signal or the associated input audio signal from the scaling information provided by the mixing metadata and from the mixing balance input, wherein the respective other input audio signal is then identified as a non-dominant signal, and wherein the dominant signal is identified by comparing the mixing balance input with the metadata scale factor for the main input audio signal;

scaling the non-dominant signal in relation to the dominant signal; and

combining the scaled non-dominant signal with the dominant signal to yield the mixed audio signal.

2. The method of claim 1 further comprising:

determining a scale factor for the scaling of the non-dominant signal based on the scaling information from the mixing metadata and based on the mixing balance input;

wherein the determined scale factor is used for the scaling of the non-dominant signal.

3. The method of claims 1-2, wherein

if the value of the mixing balance input is greater than the metadata scale factor for the main input audio signal from the mixing metadata,

then the associated input audio signal is determined to be the dominant signal;

otherwise, the main input audio signal is determined to be the dominant signal.

4. The method of any of the claims 1-3, wherein if the associated input audio signal is identified to be the dominant signal, the method further comprises:

calculating the scale factor for the main input audio signal as the metadata scale factor for the main input audio signal from the mixing metadata minus the value of the mixing balance input.

5. The method of any of the claims 1-3, wherein if the main input audio signal is identified to be the dominant signal, the method further comprises:

calculating the scale factor for the associated input audio signal as the value of the mixing balance input minus the metadata scale factor from the mixing metadata for the main input audio signal.

6. The method of any of the claims 1-5 where the mixing metadata is contained in the associated input audio signal.

7. The method of any of the claims 1-6, wherein the main input audio signal comprises multiple audio channels.

8. The method of claim 7, wherein the mixing metadata further comprises:

a primary metadata scale factor for the main input audio signal relative to the associated input audio signal; and

one or more channel metadata scale factors for different channels of the main input audio signal relative to the associated input audio signal.

9. The method of any of the claims 1-8, wherein the metadata scale factor comprises a dB value.

10. The method of any of the claims 8-9, wherein the dominant signal is identified by comparing the value of the mixing balance input with the primary and channel metadata scale factors for the main input audio signal, and

if the value of the mixing balance input is greater than the sum of the primary metadata scale factor for the main input audio signal plus the maximum of all the channel metadata scale factors for all the channels of the main input audio signal,

then the associated input audio signal is determined to be the dominant signal;

otherwise, the main input audio signal is determined to be the dominant signal.

11. The method of claim 10, wherein if the associated input audio signal is identified to be the dominant signal, the method further comprises:

calculating a channel scale factor for each channel of the main input audio signal, wherein the channel scale factor for a channel is the primary metadata scale factor for the main input audio signal plus the channel metadata scale factor for the corresponding channel of the main input audio signal minus the value of the mixing balance input.

12. The method of claim 10, wherein if the main input audio signal is identified to be the dominant signal, the method further comprises:

determining a scale factor for the associated input audio signal, calculated as the value of the mixing balance input minus the sum of the primary metadata scale factor for the main input audio signal and the maximum of all the channel metadata scale factors for all the channels of the main input audio signal; and

determining a channel scale factor for each channel of the main input audio signal, calculated as the primary scale factor for the main input audio signal plus the channel scale factor for the corresponding channel of the main input audio signal minus the sum of the primary metadata scale factor and the maximum of all the channel scale factors for the channels of the main input audio signal.

13. The method of any of the claims 10-12, wherein the metadata scale factors for determining the dominant signal further include a scale factor for the associated input audio signal, which may be used as an adjustment, wherein the scale factor for the associated input audio signal is subtracted from the metadata scale factor for the main input audio signal for the determination of the dominant signal.

14. The method of any of the claims 1-13, wherein the perceived sound level is based on an average weighted level of the main input audio signal and the associated input audio signal.

15. The method of any of the claims 1-13, wherein the perceived sound level is based on a dialog level of the main input audio signal and the associated input audio signal.

16. The method of claim 15, wherein the dialog level is measured as dialnorm.

17. The method of any of the claims 1-16, wherein the main input audio signal and the associated input audio signal are encoded as Dolby Digital Plus "DD+" or Dolby Pulse signals.

18. The method of any of the claims 1-16, wherein the main input audio signal and the associated input audio signal are encoded as E-AC-3, MPEG-4 HE-AAC, aacPlus, AC-3, MPEG-1 Layer 2, MPEG-4 AAC or any derivation of MPEG-4 audio signals.

19. The method of claims 17 or 18, wherein the mixing metadata comprises mixing metadata according to ETSI TS 102 366 V1.2.1 Ch. E.1.2.2 for mixing DD+ signals.

20. The method of claim 19 referring back to claim 8, wherein the mixing metadata comprises scale factors for the multiple audio channels: extpgmscl, extpgmlscl, extpgmrscl, extpgmcscl, extpgmlsscl, extpgmrsscl, extpgmlfescl, corresponding to a primary scale factor and channel scale factors for a left channel, a right channel, a center channel, a left surround channel, a right surround channel, and a low-frequency effects, referred to as LFE, channel; and wherein the mixing metadata optionally further comprises panmean.

21. The method of claim 20, wherein the main input audio signal further comprises a Low Frequency Effect, referred to as LFE, channel, and a channel scale factor for the LFE channel is calculated as the minimum of the channel scale factor for the LFE channel and the maximum metadata scale factor for the other channels of the main input audio signal.

22. The method of claims 17 or 18, wherein the mixing metadata comprises mixing metadata according to ETSI TS 101 154 V1.9.1 Annex E2 for mixing input audio signals.

23. The method of claim 22 referring back to claim 8, wherein the mixing metadata comprises scale factors for the multiple audio channels: AD_fade_byte, AD_gain_byte_center, AD_gain_byte_front, AD_gain_byte_surround or AD_pan_byte, corresponding to a primary scale factor and channel scale factors for a center channel, one or more front channels, one or more surround channels and pan.

24. The method of claim 10, wherein the main input audio signal comprises a low frequency effects, referred to as LFE, channel, and wherein the LFE channel is excluded from the maximum of all the channel metadata scale factors for all the channels of the main input audio signal, such that:

if the value of the mixing balance input is greater than the sum of the primary metadata scale factor for the main input audio signal plus the maximum of all the channelmetadata scale factors for all the channels of the main input audio signal, except for LFE channel,

then the associated input audio signal is determined to be the dominant signal;

otherwise, the main input audio signal is determined to be the dominant signal.

25. The method of claim 12, wherein the main input audio signal comprises a low frequency effects, referred to as LFE, channel, and wherein the LFE channel is excluded from the maximum of all the metadata scale factors for all the channels of the main input audio signal, such that the main input audio signal is identified to be the dominant signal by:

determining the scale factor for the associated input audio signal, calculated as the value of the mixing balance input minus the sum of the primary metadata scale factor for the main input audio signal and the maximum of all the channel metadata scale factors for all the channels of the main input audio signal, except for the LFE channel; and

determining the scale factor for each channel of the main input audio signal, calculated as the primary scale factor for the main input audio signal plus the scale factor for the corresponding channel of the main input audio signal minus the sum of the primary metadata scale factor and the maximum of all the channel scale factors for the channels of the main input audio signal, except for the LFE channel.

26. The method of claims 8-25, where the mixing metadata for the associated input audio signal further comprises pan metadata information, the associated input audio signal comprises a mono signal and the main input audio signal comprises a stereo or multi-channel signal,

wherein the associated input audio signal is mixed with the channels of the main input audio signal according to the pan metadata information.

27. The method of any of the claims 1-26, wherein the mixing balance input comprises an external user input providing a value from negative values to positive values.

28. The method of claim 7, wherein the multiple channels of the main input audio signal and multiple channels of the associated input audio signal comprise surround signals with left, right, center, left-surround, right-surround and low frequency effects, referred to as LFE, channels.

29. The method of claim 28, wherein the multiple channels of the main input audio signal and the associated input audio signal comprise 5.1, 3.1 or 13.1 multi-channel signals.

30. The method of any of the claims 21-29, wherein the associated input audio signal comprises mono and a LFE channel, wherein the LFE channel of the associated input audio signal is mixed with the LFE channel of the main input audio signal, and the mono channel of the associated input audio signal is panned on the channels of main input audio signal according to pan metadata information.

31. A device for mixing signals that applies any of the methods of mixing signals of claims 1-30.

32. A decoder that applies any of the methods of mixing signals of claims 1-30 to decode two input audio signals into a single, mixed audio signal while maintaining a consistent perceived sound level.

33. A processor-readable storage medium that stores computer-readable instructions to execute the method of any of the claims 1-30.

34. A processor-readable program product comprising executable instructions for performing the method of any of the claims 1-30 when executed on a processor.

35. A device for mixing input audio signals into a single, mixed audio signal while maintaining a perceived sound level of the mixed audio signal, comprising:

a receiver for receiving a main input audio signal, an associated input audio signal and mixing metadata; wherein the associated input audio signal is coupled with the main input audio signal; wherein the mixing metadata contains scaling information for scaling the main input audio signal; wherein the scaling information specifies how the main input audio signal and the associated input audio signal should be mixed, in order to generate a mixed audio signal at the perceived sound level; wherein the scaling information from the mixing metadata comprises a metadata scale factor for the main input audio signal, for scaling the main input audio signal relative to the associated input audio signal;

a mixing balance input, which denotes the adjustable balance between the main and the associated input audio signals; wherein the mixing balance input comprises scaling information which allows a deviation from a weighting of the main input audio signal and the associated input audio signal in the mixed audio signal as specified in the mixing metadata; and

a mixing unit configured to identify a dominant signal as either the main input audio signal or the associated input audio signal from the scaling information provided by the mixing metadata and from the mixing balance input, wherein the respective other input audio signal is then identified as a non-dominant signal; and wherein the dominant signal is identified by comparing the mixing balance input with the metadata scale factor for the main input audio signal;

wherein the mixing unit is further configured to scale the non-dominant signal in relation to the dominant signal; and to combine the scaled non-dominant signal with the dominant signal to yield the mixed audio signal.

36. The device of claim 35, wherein the scale factor for the non-dominant signal is determined directly from the scaling information from the mixing metadata and the mixing balance input;

the non-dominant signal is scaled using the determined scale factor; and

the scaled non-dominant signal is combined with the dominant signal into the mixed signal.

37. The device of claim 35 or 36, wherein the mixing metadata is integrated with the associated input audio signal.

38. The device of any of the claims 35-37, wherein the mixing balance input further comprises an external user input providing a value from negative values to positive values.

39. The device of any of the claims 35-38, wherein the receiver is further configured to receive multi-channel main input audio and associated input audio signals and the mixing unit is configured to mix the multi-channel main input audio and associated input audio signals.

CLAIMS

1. A method for mixing two input audio signals into a single, mixed audio signal while maintaining a perceived sound level of the mixed audio signal, the method comprising:

receiving a main input audio signal;

receiving an associated input audio signal; wherein the associated input audio signal is coupled with the main input audio signal;

receiving mixing metadata, which contains scaling information for scaling the main input audio signal and which specifies how the main input audio signal and the associated input audio signal should be mixed, in order to generate a mixed audio signal at the perceived sound level; wherein the scaling information from the mixing metadata comprises a metadata scale factor for the main input audio signal, for scaling the main input audio signal relative to the associated input audio signal;

receiving a mixing balance input, which denotes an adjustable balance between the main input audio signal and the associated input audio signal, wherein the mixing balance input comprises scaling information which allows a deviation from a weighting of the main input audio signal and the associated input audio signal in the mixed audio signal as specified in the mixing metadata;

identifying a dominant signal as either the main input audio signal or the associated input audio signal from the scaling information provided by the mixing metadata and from the mixing balance input, wherein the respective other input audio signal is then identified as a non-dominant signal; and wherein the dominant signal is identified by comparing the mixing balance input with the metadata scale factor for the main input audio signal;

scaling the non-dominant signal in relation to the dominant signal; and

combining the scaled non-dominant signal with the dominant signal to yield the mixed audio signal.

2. The method of claim 1 further comprising:

determining a scale factor for the scaling of the non-dominant signal based on the scaling information from the mixing metadata and based on the mixing balance input;

wherein the determined scale factor is used for the scaling of the non-dominant signal.

3. The method of claims 1-2, wherein

if the value of the mixing balance input is greater than the metadata scale factor for the main input audio signal from the mixing metadata,

then the associated input audio signal is determined to be the dominant signal;

otherwise, the main input audio signal is determined to be the dominant signal.

4. The method of any of the claims 1-3, wherein if the associated input audio signal is identified to be the dominant signal, the method further comprises:

calculating the scale factor for the main input audio signal as the metadata scale factor for the main input audio signal from the mixing metadata minus the value of the mixing balance input.

5. The method of any of the claims 1-3, wherein if the main input audio signal is identified to be the dominant signal, the method further comprises:

calculating the scale factor for the associated input audio signal as the value of the mixing balance input minus the metadata scale factor from the mixing metadata for the main input audio signal.

6. The method of any of the claims 1-5 where the mixing metadata is contained in the associated input audio signal.

7. The method of any of the claims 1-6, wherein the perceived sound level is based on an average weighted level of the main input audio signal and the associated input audio signal.

8. The method of any of the claims 1-7, wherein the mixing balance input comprises an external user input providing a value from negative values to positive values.

9. A device for mixing signals that applies any of the methods of mixing signals of claims 1-8.

10. A processor-readable storage medium that stores computer-readable instructions to execute the method of any of the claims 1-8.

11. A device for mixing input audio signals into a single, mixed audio signal while maintaining a perceived sound level of the mixed audio signal, comprising:

a receiver for receiving a main input audio signal, an associated input audio signal and mixing metadata; wherein the associated input audio signal is coupled with the main input audio signal; wherein the mixing metadata contains scaling information for scaling the main input audio signal; wherein the scaling information specifies how the main input audio signal and the associated input audio signal should be mixed, in order to generate a mixed audio signal at the perceived sound level; wherein the scaling information from the mixing metadata comprises a metadata scale factor for the main input audio signal, for scaling the main input audio signal relative to the associated input audio signal;

a mixing balance input, which denotes the adjustable balance between the main and the associated input audio signals; wherein the mixing balance input comprises scaling information which allows a deviation from a weighting of the main input audio signal and the associated input audio signal in the mixed audio signal as specified in the mixing metadata; and

a mixing unit configured to identify a dominant signal as either the main input audio signal or the associated input audio signal from the scaling information provided by the mixing metadata and from the mixing balance input, wherein the respective other input audio signal is then identified as a non-dominant signal; and wherein the dominant signal is identified by comparing the mixing balance input with the metadata scale factor for the main input audio signal;

wherein the mixing unit is further configured to scale the non-dominant signal in relation to the dominant signal; and to combine the scaled non-dominant signal with the dominant signal to yield the mixed audio signal.

12. The device of claim 11, wherein the scale factor for the non-dominant signal is determined directly from the scaling information from the mixing metadata and the mixing balance input;

the non-dominant signal is scaled using the determined scale factor; and

the scaled non-dominant signal is combined with the dominant signal into the mixed signal.

13. The device of claim 11 or 12, wherein the mixing metadata is integrated with the associated input audio signal.

14. The device of any of the claims 11-13, wherein the mixing balance input further comprises an external user input providing a value from negative values to positive values.

15. The device of any of the claims 11-14, wherein the receiver is further configured to receive multi-channel main input audio and associated input audio signals and the mixing unit is configured to mix the multi-channel main input audio and associated input audio signals.