

(19)



(11)

**EP 3 044 787 B1**

(12)

**EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:  
**09.08.2017 Bulletin 2017/32**

(51) Int Cl.:  
**G10L 19/018** <sup>(2013.01)</sup>      **G10L 19/008** <sup>(2013.01)</sup>

(21) Application number: **14772519.6**

(86) International application number:  
**PCT/US2014/054833**

(22) Date of filing: **09.09.2014**

(87) International publication number:  
**WO 2015/038546 (19.03.2015 Gazette 2015/11)**

(54) **SELECTIVE WATERMARKING OF CHANNELS OF MULTICHANNEL AUDIO**

SELEKTIVE WASSERZEICHENMARKIERUNG VON KANÄLEN VON MEHRKANALAUDIO  
 TATOUAGE SÉLECTIF DES CANAUX D'UN AUDIO MULTICANAL

(84) Designated Contracting States:  
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

- **MEHTA, Sripal S.**  
**San Francisco, California 94103 (US)**
- **BREEBAART, Dirk Jeroen**  
**McMahons Point, New South Wales 2060 (AU)**

(30) Priority: **12.09.2013 US 201361877139 P**

(74) Representative: **Dolby International AB**  
**Patent Group Europe**  
**Apollo Building, 3E**  
**Herikerbergweg 1-35**  
**1101 CN Amsterdam Zuidoost (NL)**

(43) Date of publication of application:  
**20.07.2016 Bulletin 2016/29**

(73) Proprietor: **Dolby Laboratories Licensing Corporation**  
**San Francisco, CA 94103 (US)**

(56) References cited:  
**EP-A1- 2 562 748      US-A1- 2006 133 644**

(72) Inventors:  
 • **NURMUKHANOV, Dossym**  
**San Francisco, California 94103 (US)**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 3 044 787 B1**

**Description**CROSS REFERENCE TO RELATED APPLICATIONS

**[0001]** This application claims priority to United States Provisional Patent Application No. 61/877,139, filed on 12 September 2013.

TECHNICAL FIELD

**[0002]** The invention pertains to audio signal processing, and more particularly to watermarking of selected channels of multichannel audio programs (e.g., bit-streams indicative of object-based audio programs including at least one audio object channel and at least one speaker channel).

BACKGROUND

**[0003]** Watermarking (forensic marking) is employed in digital cinemas to prevent piracy and allow forensic tracking of illicit captures or copies of cinematic content, and is also employed in other contexts. Watermarks, which can be embedded in both audio and video signals, should be robust against legitimate and illegitimate modifications to the marked content and captures of the marked content (e.g., captures made by mobile phones or high-quality audio and video recording devices). Watermarks typically comprise information about when and where playback of the content has occurred. Thus, watermarking for theatrical use typically occurs during actual playback, and the watermarks to content played in theaters are typically indicative of theater identification data (a theater "ID") and playback time.

**[0004]** The complexity, and therefore the financial and computational cost, of watermarking audio programs increases linearly with the number of channels to be watermarked. During rendering and playback (e.g., in movie theaters) of object based audio programs, the audio content has a number of channels (e.g., object channels and speaker channels) which is typically much larger (e.g., by an order of magnitude) than the number occurring during rendering and playback of conventional speaker-channel based programs. Typically also, the speaker system used for playback includes a much larger number of speakers than the number employed for playback of conventional speaker-channel based programs.

**[0005]** It is conventional to watermark some but not all speaker channels of a multichannel audio program of the conventional type comprising speaker channels but not object channels. However, conventional watermarking of this type does not measure content of individual channels of the program to select which channels should be watermarked, and does not select which channels to watermark based on the configuration of the playback speakers (e.g., the arrangement of speakers in a room) or the audio content to be played by any of the speakers. Rather, conventional watermarking of this type typically

tries to watermark the first N channels of the program (where N is a small number consistent with the processing limitations of the watermarking system, e.g., N = 8) or all the channels if the program comprises not more than a small number of channels, but during watermarking (e.g., rendering which includes watermarking) skips randomly the watermarking of some channels depending on actually achieved processing speed (so that watermarking of some channels is skipped if otherwise, overall processing rate would fall below a threshold).

**[0006]** The inventors have recognized that watermarking (e.g., during playback in a theater) of each individual channel (or a randomly determined subset of the channels) of a multichannel audio program (or each speaker feed signal, or a randomly determined subset of the speaker feed signals, generated in response to such program) can be wasteful and inefficient. For example, watermarking of signals indicative of silent (or nearly silent) audio content will generally not contribute to an improved watermark recovery. Furthermore, watermarking of channels that are relatively quiet compared to other channels will not contribute to improved watermark recovery.

**[0007]** Although embodiments of the invention are useful for selectively watermarking channels of any multichannel audio program, many embodiments of the invention are especially useful for selectively watermarking channels of object-based audio programs having a large number of channels.

**[0008]** It is known to employ playback systems (e.g., in movie theaters) to render object based audio programs. Object based audio programs which are movie soundtracks may be indicative of many different audio objects corresponding to images on a screen, dialog, noises, and sound effects that emanate from different places on (or relative to) the screen, as well as background music and ambient effects (which may be indicated by speaker channels of the program) to create the intended overall auditory experience. Accurate playback of such programs requires that sounds be reproduced in a way that corresponds as closely as possible to what is intended by the content creator with respect to audio object size, position, intensity, movement, and depth.

**[0009]** During generation of object based audio programs, it is typically assumed that the loudspeakers to be employed for rendering are located in arbitrary locations in the playback environment; not necessarily in a (nominally) horizontal plane or in any other predetermined arrangements known at the time of program generation. Typically, metadata included in the program indicates rendering parameters for rendering at least one object of the program at an apparent spatial location or along a trajectory (in a three dimensional volume), e.g., using a three-dimensional array of speakers. For example, an object channel of the program may have corresponding metadata indicating a three-dimensional trajectory of apparent spatial positions at which the object (indicated by the object channel) is to be rendered. The trajectory may include a sequence of "floor" locations (in

the plane of a subset of speakers which are assumed to be located on the floor, or in another horizontal plane, of the playback environment), and a sequence of "above-floor" locations (each determined by driving a subset of the speakers which are assumed to be located in at least

one other horizontal plane of the playback environment). **[0010]** Object based audio programs represent a significant improvement in many contexts over traditional speaker channel-based audio programs, since speaker-channel based audio is more limited with respect to spatial playback of specific audio objects than is object channel based audio. Speaker channel-based audio programs consist of speaker channels only (not object channels), and each speaker channel typically determines a speaker feed for a specific, individual speaker in a listening environment.

**[0011]** Various methods and systems for generating and rendering object based audio programs have been proposed. During generation of an object based audio program, it is typically assumed that an arbitrary number of loudspeakers will be employed for playback of the program, and that the loudspeakers to be employed (typically, in a movie theater) for playback will be located in arbitrary locations in the playback environment; not necessarily in a (nominally) horizontal plane or in any other predetermined arrangement known at the time of program generation. Typically, object-related metadata included in the program indicates rendering parameters for rendering at least one object of the program at an apparent spatial location or along a trajectory (in a three dimensional volume), e.g., using a three-dimensional array of speakers. For example, an object channel of the program may have corresponding metadata indicating a three-dimensional trajectory of apparent spatial positions at which the object (indicated by the object channel) is to be rendered. The trajectory may include a sequence of "floor" locations (in the plane of a subset of speakers which are assumed to be located on the floor, or in another horizontal plane, of the playback environment), and a sequence of "above-floor" locations (each determined by driving a subset of the speakers which are assumed to be located in at least one other horizontal plane of the playback environment). Examples of rendering of object based audio programs are described, for example, in PCT International Application No. PCT/US2001/028783, published under International Publication No. WO 2011/119401 A2 on September 29, 2011, and assigned to the assignee of the present application.

**[0012]** An International Search Report (the "ISR") was issued in connection with the present disclosure. The ISR cited European Patent Publication number EP 2 562 748 (the "'748 document") as a "document of particular relevance". The '748 document concerns digital audio signal watermarking in real-time. Proposed in the '748 document is that the channels in a data block-based audio multi-channel signal be prioritised with respect to watermarking importance, whereby the channel priority can change for different input signal data blocks. That is, for

a current input signal block, the most important channel would be watermarked and the required processing time would be determined. If this required processing time were shorter than a predefined application-dependent threshold, the next most important channel would be marked and the additionally required processing time would be determined, and so on.

#### BRIEF DESCRIPTION OF THE INVENTION

**[0013]** The present invention provides a method for watermarking a multichannel audio program as recited in claim 1, an audio playback system as recited in claim 8 and an audio encoder as recited in claim 14; optional features are recited in the dependent claims. These and other aspects of the invention will be appreciated from the following description of example embodiments.

#### BRIEF DESCRIPTION OF THE DRAWINGS

##### **[0014]**

FIG. 1 is a block diagram of a system including an encoder, a delivery subsystem, and a decoder. The encoder and/or the decoder are configured in accordance with an embodiment of the invention.

FIG. 2 is a diagram of an embodiment of the inventive method.

FIG. 3 is a diagram of another embodiment of the inventive method.

FIG. 4 is a diagram of an embodiment of the inventive method.

FIG. 5 is a diagram of an array of speakers, some of which may be driven by watermarked signals generated in accordance with an embodiment of the inventive method.

#### Notation and Nomenclature

**[0015]** Throughout this disclosure, including in the claims, the expression performing an operation "on" a signal or data (e.g., filtering, scaling, transforming, or applying gain to, the signal or data) is used in a broad sense to denote performing the operation directly on the signal or data, or on a processed version of the signal or data (e.g., on a version of the signal that has undergone preliminary filtering or pre-processing prior to performance of the operation thereon).

**[0016]** Throughout this disclosure including in the claims, the expression "system" is used in a broad sense to denote a device, system, or subsystem. For example, a subsystem that implements a decoder may be referred to as a decoder system, and a system including such a subsystem (e.g., a system that generates X output signals in response to multiple inputs, in which the subsystem generates M of the inputs and the other X - M inputs are received from an external source) may also be referred to as a decoder system.

**[0017]** Throughout this disclosure including in the claims, the term "processor" is used in a broad sense to denote a system or device programmable or otherwise configurable (e.g., with software or firmware) to perform operations on data (e.g., audio, or video or other image data). Examples of processors include a field-programmable gate array (or other configurable integrated circuit or chip set), a digital signal processor programmed and/or otherwise configured to perform pipelined processing on audio or other sound data, a programmable general purpose processor or computer, and a programmable microprocessor chip or chip set.

**[0018]** Throughout this disclosure including in the claims, the expressions "audio processor" and "audio processing unit" are used interchangeably, and in a broad sense, to denote a system configured to process audio data. Examples of audio processing units include, but are not limited to encoders (e.g., transcoders), decoders, codecs, pre-processing systems, post-processing systems, and bitstream processing systems (sometimes referred to as bitstream processing tools).

**[0019]** Throughout this disclosure including in the claims, the expression "metadata" (e.g., as in the expression "processing state metadata") refers to separate and different data from corresponding audio data (audio content of a bitstream which also includes metadata). Metadata is associated with audio data, and indicates at least one feature or characteristic of the audio data (e.g., what type(s) of processing have already been performed, or should be performed, on the audio data, or the trajectory of an object indicated by the audio data). The association of the metadata with the audio data is time-synchronous. Thus, present (most recently received or updated) metadata may indicate that the corresponding audio data contemporaneously has an indicated feature and/or comprises the results of an indicated type of audio data processing.

**[0020]** Throughout this disclosure including in the claims, the term "couples" or "coupled" is used to mean either a direct or indirect connection. Thus, if a first device couples to a second device, that connection may be through a direct connection, or through an indirect connection via other devices and connections.

**[0021]** Throughout this disclosure including in the claims, the following expressions have the following definitions:

speaker and loudspeaker are used synonymously to denote any sound-emitting transducer. This definition includes loudspeakers implemented as multiple transducers (e.g., woofer and tweeter);

speaker feed: an audio signal to be applied directly to a loudspeaker, or an audio signal that is to be applied to an amplifier and loudspeaker in series;

channel (or "audio channel"): a monophonic audio signal. Such a signal can typically be rendered in such a way as to be equivalent to application of the signal directly to a loudspeaker at a desired or nom-

inal position. The desired position can be static, as is typically the case with physical loudspeakers, or dynamic;

audio program: a set of one or more audio channels (at least one speaker channel and/or at least one object channel) and optionally also associated metadata (e.g., metadata that describes a desired spatial audio presentation);

speaker channel (or "speaker-feed channel"): an audio channel that is associated with a named loudspeaker (at a desired or nominal position), or with a named speaker zone within a defined speaker configuration. A speaker channel is rendered in such a way as to be equivalent to application of the audio signal directly to the named loudspeaker (at the desired or nominal position) or to a speaker in the named speaker zone;

object channel: an audio channel indicative of sound emitted by an audio source (sometimes referred to as an audio "object"). Typically, an object channel determines a parametric audio source description (e.g., metadata indicative of the parametric audio source description is included in or provided with the object channel). The source description may determine sound emitted by the source (as a function of time), the apparent position (e.g., 3D spatial coordinates) of the source as a function of time, and optionally at least one additional parameter (e.g., apparent source size or width) characterizing the source;

object based audio program: an audio program comprising a set of one or more object channels (and optionally also comprising at least one speaker channel) and optionally also associated metadata (e.g., metadata indicative of a trajectory of an audio object which emits sound indicated by an object channel, or metadata otherwise indicative of a desired spatial audio presentation of sound indicated by an object channel, or metadata indicative of an identification of at least one audio object which is a source of sound indicated by an object channel); and

render: the process of converting an audio program into one or more speaker feeds, or the process of converting an audio program into one or more speaker feeds and converting the speaker feed(s) to sound using one or more loudspeakers (in the latter case, the rendering is sometimes referred to herein as rendering "by" the loudspeaker(s)). An audio channel can be trivially rendered ("at" a desired position) by applying the signal directly to a physical loudspeaker at the desired position, or one or more audio channels can be rendered using one of a variety of virtualization techniques designed to be substantially equivalent (for the listener) to such trivial rendering. In this latter case, each audio channel may be converted to one or more speaker feeds to be applied to loudspeaker(s) in known locations, which are in general different from the desired position, such that

sound emitted by the loudspeaker(s) in response to the feed(s) will be perceived as emitting from the desired position. Examples of such virtualization techniques include binaural rendering via headphones (e.g., using Dolby Headphone processing which simulates up to 7.1 channels of surround sound for the headphone wearer) and wave field synthesis.

#### DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

**[0022]** Examples of embodiments of the invention will be described with reference to Figs. 1, 2, 3, 4, and 5.

**[0023]** Fig. 1 is a block diagram of an audio data processing system, in which one or more of the elements of the system are configured in accordance with an embodiment of the present invention. The Fig. 1 system includes encoder 3, delivery subsystem 5, and decoder 7, coupled together as shown. Although subsystem 7 is referred to herein as a "decoder" it should be understood that it is typically implemented as a playback system including a decoding subsystem (configured to parse and decode a bitstream indicative of an encoded multichannel audio program) and other subsystems configured to implement rendering (including watermarking) and at least some steps of playback of the decoding subsystem's output. Some embodiments of the invention are decoders (e.g., a decoder including a buffer memory of the type described herein) which are not configured to perform rendering and/or playback (and which would typically be used with a separate rendering and/or playback system). Some embodiments of the invention are playback systems (e.g., a playback system including a decoding subsystem and other subsystems configured to implement rendering (including watermarking) at least some steps of playback of the decoding subsystem's output).

**[0024]** A typical implementation of encoder 3 is configured to generate an object-based, encoded multichannel audio program in response to multiple streams of audio data and metadata provided to encoder 3 (as indicated in Fig. 1) or generated by encoder 3. A bitstream indicative of the program is output from encoder 3 to delivery subsystem 5. In other implementations, encoder 3 is configured to generate a multichannel audio program which is not an object-based encoded audio program, and to output a bitstream indicative of the program to delivery subsystem 5. The program generated by encoder 3 is delivered by delivery subsystem 5 to decoder 7, for decoding (by subsystem 8), object processing (by subsystem 9), and rendering (by system 11) for playback by playback system speakers (not shown).

**[0025]** Encoding subsystem 4 of encoder 3 is configured to encode multiple streams of audio data to generate encoded audio bitstreams indicative of audio content of each of the channels (speaker channels and typically also object channels) to be included in the program. The

encoding performed by subsystem 4 typically implements compression, so that at least some of the encoded bitstreams output from subsystem 4 are compressed audio bitstreams.

**[0026]** In typical implementations of encoder 3, a watermarking metadata generation subsystem 2 of encoder 3 is coupled and configured to generate watermarking metadata (e.g., watermark suitability values) in accordance with an embodiment of the present invention. The watermarking metadata may be generated by any of the methods described herein. For example, it may be generated by analyzing the audio data to be indicated by segments of the multichannel audio program (to be generated by encoder 3) and determining at least one watermark suitability value for each channel of each of the segments of the program. In some embodiments, the watermarking metadata for a channel of a segment is determined from the root mean square (RMS) amplitude of the channel's audio content in the segment. In some embodiments, the watermarking metadata is generated by analyzing the audio data to be indicated by segments of the program and metadata corresponding to the audio data. For example, the watermarking metadata for a channel of a segment may be determined from the RMS amplitude of the channel's audio content in the segment and from metadata corresponding to such audio content.

**[0027]** In other implementations, watermarking metadata generation subsystem 2 is omitted from encoder 3, and any watermark suitability values needed to perform an embodiment of the inventive channel-selective watermarking are generated in a playback system or decoder (e.g., in an implementation of subsystem 11 of decoder 7).

**[0028]** Formatting stage 6 of encoder 6 is coupled and configured to assemble the encoded audio bitstreams output from subsystem 4 and corresponding metadata (including watermarking metadata generated by subsystem 2) into an multichannel audio program (i.e., a bitstream indicative of such a program).

**[0029]** In a typical implementation, encoder 3 includes buffer 3A, which stores (e.g., in a non-transitory manner) at least one frame or other segment of the multichannel audio program (e.g., object based audio program) output from stage 6. The program is output from buffer 3A for delivery by subsystem 5 to decoder 7. Typically, the program is an object based audio program, and each segment (or each of some of the segments) of the program includes audio content of a bed of speaker channels, audio content of a set of object channels, and metadata. The metadata typically includes object related metadata for the object channels and watermarking metadata (e.g., watermark suitability values) for the object channels and speaker channels (in implementations in which a watermarking metadata generation subsystem 2 of encoder 3 has generated such watermarking metadata).

**[0030]** Decoder 7 of Fig. 1 includes decoding subsystem 8, object processing subsystem 9, and rendering

(and watermarking) subsystem 11, coupled together as shown. In variations on the system shown, one or more of the elements are omitted or additional audio data processing units are included. In some implementations, decoder 7 is or is included in a playback system (e.g., in a movie theater, or an end user's home theater system) which typically includes a set of playback speakers (e.g., the speakers shown in Fig. 5).

**[0031]** In some implementations, decoder 7 is configured in accordance with an embodiment of the present invention to determine watermark suitability values for channels of a multichannel audio program (e.g., an object-based, multichannel audio program) delivered by subsystem 5. In these implementations, decoder 7 is typically also configured to perform watermarking (e.g., in subsystem 11) of some channels of the program using such watermark suitability values.

**[0032]** In some implementations, decoder 7 and encoder 3 considered together are configured to perform an embodiment of the present invention. In these implementations, encoder 3 is configured to determine watermarking metadata (e.g., watermark suitability values) for channels of a multichannel audio program (e.g., an object-based, multichannel audio program) to be delivered and to include such watermarking metadata in the program, and decoder 7 is configured to identify (parse) the watermarking metadata (e.g., watermark suitability values or values determined therefrom) for the corresponding channels of the program (which has been delivered to decoder 7) and to perform watermarking of selected channels of the program using the watermark metadata.

**[0033]** Delivery subsystem 5 of Fig. 1 is configured to store and/or transmit (e.g., broadcast) the program generated by encoder 3. In some embodiments, subsystem 5 implements delivery of (e.g., transmits) a multichannel audio program (e.g., an object based audio program) over a broadcast system or a network (e.g., the internet) to decoder 7. In some other embodiments, subsystem 5 stores a multichannel audio program (e.g., an object based audio program) in a storage medium (e.g., a disk or set of disks), and decoder 7 is configured to read the program from the storage medium.

**[0034]** In typical operation, decoding subsystem 8 of decoder 7 accepts (receives or reads) the program delivered by delivery subsystem 5. In a typical implementation, subsystem 8 includes buffer 8A, which stores (e.g., in a non-transitory manner) at least one frame or other segment (typically including audio content of a bed of speaker channels, audio content of object channels, and metadata) of an object based audio program delivered to decoder 7. The metadata typically includes object related metadata for object channels of the program and may also include watermarking metadata (e.g., watermark suitability values) generated in accordance with an embodiment of the invention for object channels and speaker channels of the program. Decoding subsystem 8 reads each segment of the program from buffer 8A and decodes each such segment. Typically, subsystem 8

parses a bitstream indicative of the program to identify speaker channels (e.g., of a bed of speaker channels), object channels and metadata, decodes the speaker channels, and outputs to subsystem 9 the decoded speaker channels and metadata. Subsystem 8 also decodes (if necessary) all or some of the object channels and outputs the object channels (including any decoded object channels) to subsystem 9.

**[0035]** Object processing subsystem 9 is coupled to receive (from decoding subsystem 8) audio samples of decoded speaker channels and object channels (including any decoded object channels), and metadata of the delivered program, and to output to rendering subsystem 11 a set of object channels (e.g., a selected subset of a full set of object channels) indicated by or determined from the program, and corresponding metadata. Subsystem 9 is typically also configured to pass through unchanged (to subsystem 11) the decoded speaker channels output from subsystem 8, and metadata corresponding thereto. Subsystem 9 may be configured to process at least some of the object channels (and/or metadata) asserted thereto to generate the object channels and corresponding metadata that it asserts to subsystem 11. Subsystem 9 is typically configured to determine a set of selected object channels (e.g., all the object channels of a delivered program, or a subset of a full set of object channels of the program, where the subset is determined by default or in another manner), and to output to subsystem 11 the selected object channels and metadata corresponding thereto. The object selection may be determined by user selection (as indicated by control data asserted to subsystem 9 from a controller) and/or rules (e.g., indicative of conditions and/or constraints) which subsystem 9 has been programmed or otherwise configured to implement.

**[0036]** If subsystem 9 is configured in accordance with a typical embodiment of the invention, the output of subsystem 9 in typical operation includes the following:

streams of audio samples indicative of a delivered program's bed of speaker channels (and optionally also corresponding metadata, e.g., watermark suitability values for the speaker channels); and streams of audio samples indicative of object channels of the program (or object channels determined from object channels of the program, e.g., by mixing) and corresponding streams of metadata (including object related metadata and optionally also watermark suitability values for the object channels).

**[0037]** Rendering subsystem 11 is configured to render the audio content determined by subsystem 9's output for playback by playback system speakers (not shown in Fig. 1). The rendering includes watermarking of selected channels of the audio content (typically using watermark suitability values received from subsystem 9 or generated by subsystem 11). Subsystem 11 is configured to map, to the available playback speaker channels,

the audio objects determined by the object channels output from subsystem 9, using rendering parameters output from subsystem 9 (e.g., object-related metadata values, which may be indicative of level and spatial position or trajectory). Typically, at least some of the rendering parameters are determined by the object related metadata output from subsystem 9. Rendering system 11 also receives the bed of speaker channels passed through by subsystem 9. Typically, subsystem 11 is an intelligent mixer, and is configured to determine speaker feeds for the available playback speakers including by mapping one or more objects (determined by the output of subsystem 9) to each of a number of individual speaker channels, and mixing the objects with "bed" audio content indicated by each corresponding speaker channel of the program.

**[0038]** In some embodiments, the speakers to be driven to render the audio are assumed to be located in arbitrary locations in the playback environment; not merely in a (nominally) horizontal plane. In some such cases, metadata included in the program indicates rendering parameters for rendering at least one object of the program at any apparent spatial location (in a three dimensional volume) using a three-dimensional array of speakers. For example, an object channel may have corresponding metadata indicating a three-dimensional trajectory of apparent spatial positions at which the object (indicated by the object channel) is to be rendered. The trajectory may include a sequence of "floor" locations (in the plane of a subset of speakers which are assumed to be located on the floor, or in another horizontal plane, of the playback environment), and a sequence of "above-floor" locations (each determined by driving a subset of the speakers which are assumed to be located in at least one other horizontal plane of the playback environment). In such cases, the rendering can be performed in accordance with the present invention so that the speakers can be driven to emit sound (determined by the relevant object channel) that will be perceived as emitting from a sequence of object locations in the three-dimensional space which includes the trajectory, mixed with sound determined by the "bed" audio content.

**[0039]** Optionally, a digital audio processing ("DAP") stage (e.g., one for each of a number of predetermined output speaker channel configurations) is coupled to the output of rendering subsystem 11 to perform post-processing on the output of the rendering subsystem. Examples of such processing include intelligent equalization or speaker virtualization processing.

**[0040]** The output of rendering subsystem 11 (or a DAP stage following subsystem 11) may be PCM bitstreams (which determine speaker feeds for the available speakers).

**[0041]** In a class of embodiments, the invention is a method for watermarking a multichannel audio program, including the steps of selecting a subset of channels of (e.g., channels determined from) at least a segment of the program for watermarking, and watermarking each

channel in the subset of channels. In some embodiments, the program is an object-based audio program (e.g., a movie soundtrack) and at least one object channel and/or at least one speaker channel of the program is watermarked. In some embodiments, a rendering system (e.g., an implementation of subsystem 11 of decoder 7 of Fig. 1) determines a set of playback speaker channels (each for playback by a different speaker of a playback system) from an object-based audio program (i.e., from at least one object channel and/or at least one speaker channel of the program), and a subset of this set of speaker channels is watermarked. In some embodiments, the selected subset is watermarked before speaker feeds are generated in response to channels of the program (e.g., by a decoder configured to receive, decode, and render the program, or during generation of the program to be delivered to a decoder for decoding and rendering). In some embodiments, the selected subset is watermarked (by a rendering system) after an encoded version of the program (e.g., an encoded bitstream indicative of the program) is decoded, but before speaker feeds are generated in response to audio content of the decoded program. In some embodiments, the selected subset is watermarked during rendering of the program (e.g., speaker feeds are generated in response to channels of the program, the speaker feeds correspond to, or are determined from, channels of the program, and a selected subset of the set of speaker feeds is watermarked).

**[0042]** Typically, the watermarking is performed in a playback system (e.g., in an implementation of decoder 7 of Fig. 1) which is coupled and configured to decode and render a multichannel audio program, and which has limited watermarking capability (i.e., the playback system does not have capability to watermark an unlimited number of audio program channels).

**[0043]** In some embodiments, a decoder (e.g., installed in a movie theater) decodes an encoded bitstream indicative of a multichannel audio program, to determine channels (speaker channels and/or object channels) of the program, or channels (speaker channels) determined from the program. A selected subset of the channels is watermarked (before or during rendering of the decoded audio), such that when the program has undergone rendering and playback, the watermark can be determined from (e.g., by processing) the sound emitted from the speaker set during playback. Thus, if the audio is recorded (e.g., illegally, by a cell phone or other device), the watermark is detectable by processing the recorded signal. The watermark may be indicative of a playback system ID (e.g., a movie theater ID) and a playback time.

**[0044]** In some embodiments, the selected subset of channels is optimized for watermark detection and recovery of information embedded in the watermark. If the channel subset selection is performed during content creation (e.g., generation of an encoded version of the program), watermarking metadata (indicative of the selected subset for each segment of a sequence of segments of the program) is typically distributed along with the audio

content of the program (e.g., the watermarking metadata is included in the program). Alternatively, the channel subset selection is performed during decoding, rendering, or playback.

**[0045]** Typical embodiments of the inventive method are expected to provide watermarking with improved watermark detectability, reduced watermarking cost, and improved quality of rendered watermarked audio (relative to that obtainable by conventional watermarking). The specific parameters of each implementation are typically determined to achieve an acceptable trade-off between robustness of watermark recovery, quality of rendered watermarked audio, and watermark information capacity.

**[0046]** In a first class of embodiments, the inventive method generates watermarking metadata (e.g., watermark suitability values) during audio program creation (e.g., in subsystem 2 of an implementation of encoder 3 of Fig. 1) including by analyzing the audio content to be included in segments of a multichannel audio program (e.g., analyzing the audio content in segments of the program each having a duration of T minutes, where the value of T is based on the watermarking algorithms to be used and amount of time required for watermark recovery) and determining at least one watermark suitability value (sometimes referred to herein as a "weight" or watermark suitability weight) for each channel of each of the segments of the program. In typical embodiments, each watermark suitability value ("WSV") is indicative of the suitability of the content of the corresponding channel (in the relevant segment of the program) for watermarking (e.g., the WSV may indicate RMS amplitude of the corresponding content, and/or recoverability of a watermark if the watermark is applied to the content). The watermark suitability values (or watermarking data determined therefrom) are included as metadata in the audio program (e.g., with each segment of each channel of the program including watermarking metadata indicative of watermark suitability of the segment of the channel or whether the segment of the channel should be watermarked). Using the watermarking metadata, a playback system can detect (typically, easily) which of the channels of each segment of the program are the most suitable for watermarking or which should be watermarked.

**[0047]** In typical embodiments in the first class, the playback system is constrained to watermark no more than a maximum number ("N") of channels of (or determined from) an audio program being decoded and rendered. For each segment of an audio program being decoded, the playback system is configured to compare the watermarking suitability values for the program's channels (e.g., for each speaker channel of a bed of speaker channels, and each object channel, of an object-based audio program), and to identify from the watermarking suitability values a subset of N of the highest-weighted (most suitable for watermarking) channels for the segment. The identified N channels of each segment are then watermarked. When the watermarking is complete

for a segment, all channels (including the N watermarked channels) to be rendered are reassembled (synchronized) and rendered (i.e., speaker feeds are generated in response to a full set of channels including the N watermarked channels).

**[0048]** Fig. 2 is a diagram of an embodiment in the first class. As indicated in Fig. 2, the process of generating the multichannel program to be watermarked and rendered (the "content creation" process, which may be performed by an implementation of encoder 3 of Fig. 1) includes steps of:

a "weighing" step (50), which includes determining watermarking suitability of each channel of a segment of the program (i.e., each speaker channel of each "bed" of speaker channels of the segment, and each object channel of the segment) from the channel's content in the segment (e.g., the RMS amplitude of the channel's audio content in the segment) and optionally also from metadata corresponding to the audio content;

a step (51) of determining a watermark suitability value ("WSV") for each channel of the segment, to be included as metadata for the corresponding audio content of each channel of the segment;

a packaging step (52), which encodes the segment as a bitstream including the samples (typically, encoded samples) of audio content of each channel of the segment packaged with the corresponding WSV (determined in step 51) and original metadata for each said channel of the segment.

**[0049]** As indicated in Fig. 2, the process of playback of the multichannel program generated in step 52 (which may be performed by an implementation of decoder 7 of Fig. 1) includes steps of:

an unpacking step (53), which includes parsing of a segment of the program into the audio content of each channel of the segment (and performing any necessary decoding of the audio samples indicative of such audio content), the WSV corresponding to the channel of the segment, and other metadata corresponding to the channel of the segment;

a step (54) of processing the WSV values for the channels of the segment to identify (select) which of the channels should be watermarked;

a step (55) of watermarking each of channels of the segment which was selected in step 54;

a step (56) of synchronizing the watermarked audio content of each watermarked channel of the segment and the non-watermarked audio content of each other channel of the segment to be rendered; and

a step (57) of rendering the synchronized watermarked and non-watermarked audio content of each channel of the segment to be rendered, thereby generating speaker feeds for each said channel of the



segment.

**[0050]** Various embodiments of the inventive method employ different methods to determine a watermark suitability value ("WSV") for each channel of a segment of a multichannel audio program, including (but not limited to) the following:

1. the WSV for a channel of the segment is determined from (e.g., is determined to be) the root mean square (RMS) amplitude of the channel's audio content in the segment;

2. the WSV for a channel of the segment is determined from the RMS amplitude of the channel's audio content in the segment and metadata (delivered with the program) corresponding to the audio content. For example, the metadata may indicate a gain (or gain increase or decrease) to be applied to the channel's audio content in the segment;

3. the segment is rendered (speaker feeds are determined for the segment from all channels of the segment) as it would be perceived in or near the center of a room (e.g., an auditorium), and the WSV for each channel of the rendered segment is determined (e.g., by an implementation of subsystem 11 of decoder 7 of Fig. 1, or by subsystem 2 of encoder 3 of Fig. 1) from the RMS amplitude of said channel of the rendered segment. For example, the segment might be rendered using zone exclusion metadata (delivered with an object-based audio program) for the segment, where the zone exclusion metadata indicates which object channels are allowed (and which object channels are not allowed) to contribute to each speaker feed for the segment (e.g., the metadata might cause audio content indicative of some objects to be played back only by speakers in specific zones of a theater). Thus, if the metadata indicates that speakers in an "exclusion" zone should not emit sound indicative of a "first" object, the speaker feeds for the speakers in the exclusion zone will not be indicative of the first object and the WSV for each corresponding channel of the rendered segment will not be indicative of RMS amplitude of audio content corresponding to the first object (although it might be indicative of RMS amplitude of audio content corresponding to objects other than the first object);

4. the WSV for a channel of the segment is determined from the energy or RMS amplitude of the channel's audio content in a limited frequency range. Watermarking algorithms often embed information in a limited frequency range only. When such watermarking is to be employed, it may be useful to compute the WSV from signal energy or RMS amplitude in the same frequency range as the frequency range to be watermarked;

5. the WSV for a channel of the segment is determined using at least one other feature (of the channel's audio content in the segment) besides RMS or

signal amplitude. For example, spread-spectrum watermarking techniques work best on wide-band audio signals and often do not perform well on narrow-band signals. The bandwidth, spectral flatness, or any other feature representative of the shape of the spectrum of the channel's audio content in the segment can be useful to estimate the robustness of the watermark detection process, and thus may be used to determine at least partially the WSV for the channel of the segment;

**[0051]** Preferably, the WSVs for the channels of a segment of a program are (or can be processed to determine) an ordered list which indicates the channels in increasing or decreasing order of suitability for watermarking. In this way, a best possible watermarking effort can be obtained which is independent of a playback system's watermarking capabilities. Because audio signals are typically time varying and dynamic in nature, the ordered list is preferably time dependent (i.e., an ordered list is determined for each segment of a program).

**[0052]** Such an ordered list can be split into a list of a first set of channels ("absolutely required" channels) that must be watermarked to guarantee a minimum quality of service (e.g., watermark detection robustness), and a second, ordered list which may be employed to select additional channels to be watermarked if the capabilities of the watermarking system allow for watermarking of more than just the "absolutely required" channels.

**[0053]** In a second class of embodiments, the invention is implemented by a playback system only (e.g., by an implementation of decoder 7 of Fig. 1), and does not require that an encoding system which generates the multichannel audio program (to be watermarked and rendered for playback) be configured in accordance with an embodiment of the invention (i.e., the encoding system need not identify WSVs for channels of the program). In these embodiments, the playback system determines the WSVs for channels of each segment of the program, e.g., using any of the methods described above. Fig. 3 is a diagram of such an embodiment in the second class (which may be performed by an implementation of decoder 7 of Fig. 1).

**[0054]** As indicated in Fig. 3, the process of playback of the multichannel program includes steps of:

an unpacking step (60), which includes parsing of a segment of the program into the audio content (and any corresponding metadata) of each channel of the segment (and performing any necessary decoding of the audio samples indicative of such audio content);

a "weighing" step (61), which includes generating watermarking suitability data indicative of suitability for watermarking of each channel of a segment of the program (i.e., each speaker channel of each "bed" of speaker channels of the segment, and each object channel of the segment) from the channel's

content in the segment (e.g., the RMS amplitude of the channel's audio content in the segment) and optionally also from metadata corresponding to the audio content;

a step (62) of selecting a subset of the channels of the segment using the watermarking suitability data, and watermarking each channel of the subset of the channels of the segment;

a step (63) of synchronizing the watermarked audio content of each watermarked channel of the segment and the non-watermarked audio content of each other channel of the segment to be rendered; and

a step (64) of rendering the synchronized watermarked and non-watermarked audio content of each channel of the segment to be rendered, thereby generating speaker feeds for each said channel of the segment.

**[0055]** In some embodiments in the second class, the playback system selects for watermarking a subset of a set of individual speaker channels determined from the multichannel program. For example, if the program is an object-based audio program including object channels as well as a bed of speaker channels, the playback system (e.g., an implementation of subsystem 11 of decoder 7 of Fig. 1) may determine a set of playback speaker channels (each playback speaker channel corresponding to a different speaker of a set of playback speakers) from the object channels and/or speaker channels of the program, and the playback system then selects a subset of the playback speaker channels for watermarking. The subset selection for a segment of the program may be based on RMS amplitude of each speaker channel determined from the segment of the program, or it may be based on another criterion. Fig. 4 is a diagram of such an embodiment in the second class (which may be performed by an implementation of decoder 7 of Fig. 1).

**[0056]** As indicated in Fig. 4, the process of playback of the multichannel program includes steps of:

an unpacking step (70), which includes parsing of a segment of the program into the audio content (and any corresponding metadata) of each channel of the segment (and performing any necessary decoding of the audio samples indicative of such audio content);

a step (71) of rendering audio content of the segment, thereby determining a set of playback speaker channels (each playback speaker channel corresponding to, and indicative of content to be played by, a different speaker of a set of playback speakers); a "weighing" step (72), which includes generating watermarking suitability data indicative of suitability for watermarking of each of the playback speaker channels;

a step (73) of selecting a subset of the playback speaker channels of the segment using the water-

marking suitability data, and watermarking each channel of the subset of the playback speaker channels of the segment; and

a step (74) of synchronizing the watermarked audio content of each watermarked channel of the subset of the playback speaker channels of the segment and the non-watermarked audio content of each other channel of the subset of the playback speaker channels of the segment.

**[0057]** In some embodiments in the second class, the playback system uses the configuration of the playback speakers (installed in an auditorium or other playback environment) to select the subset of channels to be watermarked, including by identifying groups (subsets) of the full set of playback speakers in distinct locations (zones) in the playback environment. These embodiments includes steps of: determining from channels of the program a set of playback speaker channels, each for playback by a different one of the playback speakers (each speaker may comprise one or more transducers), selecting a subset of the set of playback speaker channels for watermarking, and watermarking each channel in the subset of the set of playback speaker channels (thereby generating a set of watermarked channels), including by identifying groups of the playback speakers which are installed in distinct zones in the playback environment such that each of the groups consists of speakers installed in a different one of the zones, identifying suitability for watermarking of audio content for playback by each of the groups, and selecting the subset of the set of playback speaker channels in accordance with the suitability for watermarking of audio content for playback by each of at least a subset of the groups. Typically, the audio content (e.g., object channel content and speaker channel content) of the program (or a segment of the program) is rendered, thereby determining the set of playback speaker channels (each playback speaker channel corresponding to, and indicative of content to be played by, a different speaker of the set of playback speakers), and the playback system selects one playback speaker channel (or a small number of playback speaker channels) corresponding to each of the groups of speakers (e.g., a speaker channel for driving one speaker in each of the groups) or each of a subset of the groups, and watermarks each such selected playback speaker channel. This can result in watermarking of only channels that typically indicate audio content of specific type(s), and can enable recovery (with a high probability of success) of the watermarks without incurring large computation costs. These embodiments do not measure the loudness (or another characteristic) of the audio content of each channel selected for watermarking. Instead, they assume that some playback speaker channels (of a full set of playback speaker channels) are suitable for watermarking (e.g., are likely to be indicative of loud content, and/or content of specific type(s)) and should be watermarked. Typically, only playback speaker channels

that are assumed to be likely to be suitable for watermarking are watermarked, and a signal for driving a speaker from each group of the full set of speakers is watermarked. An example of such an embodiment in the second class will be described with reference to Fig. 5.

**[0058]** Fig. 5 shows an array of playback speakers in a room (e.g., a movie theater). The speakers are grouped into the following groups: front left speaker (L), front center speaker (C), front right speaker (R), left side speakers (Lss1, Lss2, Lss3, and Lss4), right side speakers (Rss1, Rss2, Rss3, and Rss4), left ceiling-mounted speakers (Lts1, Lts2, Lts3, and Lts4), right ceiling-mounted speakers (Rts1, Rts2, Rts3, and Rts4), left rear (surround) speakers (Lrs1 and Lrs2), and right rear (surround) speakers (Rrs1 and Rrs2).

**[0059]** The content to be played by the front left speaker (L), front center speaker (C), front right speaker (R), left rear speakers (Lrs1 and Lrs2), and right rear speakers (Rrs1 and Rrs2) is assumed to be suitable for watermarking, and thus the playback speaker channel corresponding to each of these speakers is watermarked (e.g., by an implementation of subsystem 11 of decoder 7). The content to be played by the left side speakers (Lss1, Lss2, Lss3, and Lss4) and right side speakers (Rss1, Rss2, Rss3, and Rss4) is assumed to be less suitable for watermarking, and thus the playback speaker channels corresponding to only two or three speakers in each of these two groups (i.e., Lss1, Lss2, Lss3, Rss1, and Rss2, as indicated in Fig. 5) is watermarked (e.g., by an implementation of subsystem 11 of decoder 7). The content to be played by the left ceiling-mounted speakers (Lts1, Lts2, Lts3, and Lts4) and right ceiling-mounted speakers (Rts1, Rts2, Rts3, and Rts4) is also assumed to be less suitable for watermarking, and thus the playback speaker channels corresponding to only two speakers in each of these two groups (i.e., Lts1, Lts2, Rts1, and Rts2, as indicated in Fig. 5) is watermarked (e.g., by an implementation of subsystem 11 of decoder 7).

**[0060]** If it is predetermined that only a maximum number ("M") of playback speaker channels will be marked (e.g., M = 16 as in Fig. 5), although rendering of a program will generate playback speaker channels for driving more than "M" playback speakers (e.g., 23 playback speaker channels for driving 23 playback speakers as in Fig. 5), the specific playback speaker channels to be watermarked may be selected as follows: one playback speaker channel for each group of speakers is selected (e.g., L, C, R, Lss1, Lrs1, Rss1, Rrs1, Lts1, and Rts1 as in Fig. 5) is selected for watermarking; then an additional playback speaker channel from each group is selected for watermarking (e.g., Lss2, Lrs2, Rss2, Rrs2, Lts2, and Rts2, as in Fig. 5) so long as the total number of channels to be watermarked does not exceed "M" (or until the total number of channels to be watermarked reaches "M"); and so on. Thus, in the Fig. 5 example, a third playback speaker channel (Lss3) from one group is selected for watermarking which brings the total number of channels to be watermarked to "M" (i.e., M = 16 in the

Fig. 5 example). Typically, the selection of the speaker channels to be marked is done once for a playback environment (e.g., an auditorium) and this selection does not change (it stays static) regardless of the content played in the environment.

**[0061]** Depending on the employed watermarking technology, watermarking can often be formulated as an additive process in which a watermark signal is added to an audio signal. The watermark signal is adjusted in terms of level and spectral properties according to the host (audio) signal. As such, the watermark can easily be faded out on one stream (channel) and faded in on another stream (channel) without creating artifacts, provided that a sufficient fade duration (typically about 10 ms or longer) is used. Thus, selection of a subset of a full set of channels for watermarking may typically be performed with a temporal granularity of on the order of tens of milliseconds (i.e., a selection is performed for each segment of the program having duration of on the order of tens of milliseconds), although it may be beneficial to perform it less frequently (i.e., to perform a selection for each segment of the program having duration of more than on the order of tens of milliseconds).

**[0062]** Content creation systems (e.g., in movie studios) typically can enable or disable audio watermarking during the content authoring process. By dynamically modify watermarking properties during content creation (i.e., by dynamically selecting different subsets of channels of content to be watermarked), the mixing engineer may influence the watermarking process to ensure that critical excerpts in the content are or are not watermarked (or are subject to watermarking which is more or less perceptible).

**[0063]** Embodiments of the invention may be implemented in hardware, firmware, or software, or a combination thereof (e.g., as a programmable logic array). For example, encoder 3, or decoder 7, or subsystem 8, 9, and/or 11 of decoder 7 of Fig. 1 may be implemented in appropriately programmed (or otherwise configured) hardware or firmware, e.g., as a programmed general purpose processor, digital signal processor, or microprocessor. Unless otherwise specified, the algorithms or processes included as part of the invention are not inherently related to any particular computer or other apparatus. In particular, various general-purpose machines may be used with programs written in accordance with the teachings herein, or it may be more convenient to construct more specialized apparatus (e.g., integrated circuits) to perform the required method steps. Thus, the invention may be implemented in one or more computer programs executing on one or more programmable computer systems (e.g., a computer system which implements encoder 3, or decoder 7, or subsystem 8, 9, and/or 11 of decoder 7 of Fig. 1), each comprising at least one processor, at least one data storage system (including volatile and non-volatile memory and/or storage elements), at least one input device or port, and at least one output device or port. Program code is applied to input

data to perform the functions described herein and generate output information. The output information is applied to one or more output devices, in known fashion.

[0064] Each such program may be implemented in any desired computer language (including machine, assembly, or high level procedural, logical, or object oriented programming languages) to communicate with a computer system. In any case, the language may be a compiled or interpreted language.

[0065] For example, when implemented by computer software instruction sequences, various functions and steps of embodiments of the invention may be implemented by multithreaded software instruction sequences running in suitable digital signal processing hardware, in which case the various devices, steps, and functions of the embodiments may correspond to portions of the software instructions.

[0066] Each such computer program is preferably stored on or downloaded to a storage media or device (e.g., solid state memory or media, or magnetic or optical media) readable by a general or special purpose programmable computer, for configuring and operating the computer when the storage media or device is read by the computer system to perform the procedures described herein. The inventive system may also be implemented as a computer-readable storage medium, configured with (i.e., storing) a computer program, where the storage medium so configured causes a computer system to operate in a specific and predefined manner to perform the functions described herein.

[0067] While implementations have been described by way of example and in terms of exemplary specific embodiments, it is to be understood that implementations of the invention are not limited to the disclosed embodiments. On the contrary, it is intended to cover various modifications and similar arrangements as would be apparent to those skilled in the art, inasmuch as falling within the scope defined by the appended claims.

**Claims**

- 1. A method for watermarking a multichannel audio program, including the steps of:
  - (a) selecting a subset of channels of at least a segment of the program for watermarking, such that the selection of the subset is based on the program or on configuration of playback speakers to be employed for playback of the program; and
  - (b) watermarking (11) each channel in the subset of channels, thereby generating a set of watermarked channels; and
  - (c) analyzing (2) audio content in a segment of the program to determine values indicative of suitability for watermarking of audio content of channels of the program in the segment,

wherein step (a) includes a step of selecting the subset of channels in response to said values, and wherein step (c) includes:

- 5 a step of determining root mean square amplitude of the audio content of each of the channels in the segment, or
- 10 a step of determining energy or root mean square amplitude of the audio content in a limited frequency range of each of the channels in the segment.

- 2. The method of claim 1, also including steps of:

- 15 determining from channels of the program a set of playback speaker channels, each for playback by a different speaker of a set of speakers installed in a playback environment, where the subset of channels of the program selected in step (a) is a subset of the set of playback speaker channels, and step (a) includes steps of:

- 20 identifying groups of the speakers which are installed in distinct zones in the playback environment such that each of the groups consists of speakers installed in a different one of the zones, and identifying watermarking suitability of audio content for playback by each of the groups; and
- 25 selecting the subset of the set of playback speaker channels in accordance with the watermarking suitability of audio content for playback by each of at least a subset of the groups.

- 30 3. The method of claim 1 or claim 2, also including a step of:

- 35 after steps (a) and (b), generating speaker feeds in response to the set of watermarked channels and at least one unwatermarked channel of the program.

- 40 4. The method of any of claims 1 to 3, wherein the program includes a set of channels, said method also including a step of:

- 45 rendering the program including by generating speaker feeds in response to at least some of the channels of the program, and wherein step (a) includes a step of selecting a subset of the speaker feeds for watermarking, and step (b) includes a step of watermarking at least a segment of each speaker feed in the subset of speaker feeds.

- 50 5. The method of any of claims 1 to 4, wherein the program is an object-based audio program, and said

method includes a step of:

determining a set of playback speaker channels, each for playback by a different speaker of a playback system, from at least one object channel and/or at least one speaker channel of the program, and wherein the subset of channels selected in step (a) is a subset of the set of playback speaker channels.

6. The method of any of claims 1 to 5, wherein the program includes watermarking metadata, said method includes a step of operating a decoder to decode and render the program, and step (a) includes a step of selecting the subset of channels using the watermarking metadata, and optionally, wherein the watermarking metadata are watermark suitability values, each of the watermark suitability values of a segment of the program is indicative of suitability for watermarking of audio content of a corresponding channel of the program in the segment.

7. The method of any of claims 1 to 6, wherein the watermarking suitability value for at least one channel of the segment is at least partially determined from a number of speakers to be driven to emit content indicative of the channel during playback of the segment.

8. An audio playback system, including:

a decoding subsystem (8), coupled and configured to parse and decode an encoded bitstream to extract therefrom audio data and metadata indicative of a multichannel audio program; and a second subsystem (11), coupled and configured to select a subset of channels of at least a segment of the program for watermarking, and to watermark data indicative of each channel in the subset of channels thereby determining a set of watermarked channels, wherein the selection of the subset is based on the program or on configuration of playback speakers to be employed for playback of the program, wherein the second subsystem is configured to analyze the audio data of a segment of the program to determine values indicative of watermarking suitability of audio content of channels of the program in the segment, including by determining root mean square amplitude of the audio data of each of the channels in the segment or by determining energy or root mean square amplitude of the audio data in a limited frequency range of each of the channels in the segment, and to select the subset of channels in response to said values.

5

10

15

20

25

30

35

40

45

50

55

9. The system of claim 8, wherein the second subsystem is configured to determine from the audio data and the metadata a set of playback speaker channels, each for playback by a different speaker of a set of speakers installed in a playback environment, and to select a subset of the set of playback speaker channels as the subset of channels, including by:

identifying groups of the speakers which are installed in distinct zones in the playback environment such that each of the groups consists of speakers installed in a different one of the zones, and identifying watermarking suitability of audio content for playback by each of the groups; and selecting the subset of the set of playback speaker channels in accordance with the watermarking suitability of audio content for playback by each of at least a subset of the groups.

10. The system of claim 8 or claim 9, wherein the program includes a set of channels, and the second subsystem is configured:

to render the program including by generating speaker feeds in response to at least some of the channels of the program; and to select a subset of the speaker feeds for watermarking and watermark at least a segment of each speaker feed in the subset of speaker feeds.

11. The system of any of claims 8 to 10, wherein the program is an object-based audio program, the second subsystem is configured to determine a set of playback speaker channels, each for playback by a different speaker of a playback system, from at least one object channel and/or at least one speaker channel of the program, and to select a subset of the set of playback speaker channels as the subset of channels.

12. The system of any of claims 8 to 11, wherein the program includes watermarking metadata, the decoding subsystem is configured to extract the watermarking metadata, and the second subsystem is configured to use the watermarking metadata to select the subset of channels for watermarking, and optionally, wherein the watermarking metadata are watermark suitability values, each of the watermark suitability values of a segment of the program is indicative of suitability for watermarking of audio content of a corresponding channel of the program in the segment.

13. The system of any of claims 8 to 12, wherein the watermarking suitability value for at least one channel of the segment is at least partially determined from a number of speakers to be driven to emit con-

tent indicative of the channel during playback of the segment.

14. An audio encoder configured to generate a bitstream indicative of an encoded multichannel audio program, said encoder including:

a first subsystem (2) coupled and configured to generate watermarking metadata in response to segments of streams of audio content, wherein the watermarking metadata is indicative of suitability for watermarking of at least one segment of each of the streams, or the watermarking metadata is indicative of whether watermarking should be performed on at least one segment of each of the streams; and

a second subsystem (6) coupled and configured to generate the bitstream indicative of the encoded multichannel audio program, including by encoding at least some of the streams of audio content to generate encoded streams of audio content, and including in the bitstream each of the encoded streams of audio content, each of the streams of audio content which is not encoded, and the watermarking metadata, wherein the first subsystem is configured to analyze at least one segment of each of the streams of audio content to determine values indicative of watermarking suitability of audio content of each of the streams in the segment, including by determining root mean square amplitude of the audio content of said each of the streams in the segment or by determining energy or root mean square amplitude of the audio content in a limited frequency range of each of the channels in the segment, and to select the subset of channels in response to said values.

15. The encoder of claim 14, wherein the watermarking suitability value for at least one channel of the segment is at least partially determined from a number of speakers to be driven to emit content indicative of the channel during playback of the segment.

#### Patentansprüche

1. Verfahren für Wasserzeichenmarkieren eines Mehrkanal-Audioprogramms, die folgenden Schritte enthaltend:

(a) Auswählen einer Teilmenge von Kanälen mindestens eines Segments des Programms für Wasserzeichenmarkieren, so dass die Auswahl der Teilmenge auf dem Programm oder auf einer Konfiguration von Wiedergabelautsprechern, die für Wiedergabe des Programms einzusetzen sind, basiert; und

(b) Wasserzeichenmarkieren (11) jedes Kanals in der Teilmenge von Kanälen, wodurch eine Menge von wasserzeichenmarkierten Kanälen erzeugt wird; und

(c) Analysieren (2) von Audioinhalt in einem Segment des Programms, um Werte zu bestimmen, die eine Eignung für Wasserzeichenmarkieren von Audioinhalt von Kanälen des Programms in dem Segment angeben,

wobei Schritt (a) einen Schritt des Auswählens der Teilmenge von Kanälen als Reaktion auf die Werte enthält, und

wobei Schritt (c) enthält:

einen Schritt des Bestimmens der quadratischen mittleren Amplitude des Audioinhalts jedes der Kanäle in dem Segment, oder einen Schritt des Bestimmens der Energie oder der quadratischen mittleren Amplitude des Audioinhalts in einem begrenzten Frequenzbereich jedes der Kanäle in dem Segment.

2. Verfahren nach Anspruch 1, außerdem die folgenden Schritte enthaltend:

Bestimmen, aus Kanälen des Programms, einer Menge von Wiedergabelautsprecherkanälen, jeder für Wiedergabe durch einen verschiedenen Lautsprecher eines Satzes von Lautsprechern, die in einer Wiedergabeumgebung installiert sind, wobei die in Schritt (a) ausgewählte Teilmenge von Kanälen des Programms eine Teilmenge der Menge von Wiedergabelautsprecherkanälen ist und Schritt (a) die folgenden Schritte enthält:

Identifizieren von Gruppen der Lautsprecher, die in distinkten Zonen in der Wiedergabeumgebung installiert sind, so dass jede der Gruppen aus Lautsprechern besteht, die in einer verschiedenen einen der Zonen installiert sind, und Identifizieren von Wasserzeichenmarkierung-Eignung von Audioinhalt für Wiedergabe durch jede der Gruppen; und

Auswählen der Teilmenge der Menge von Wiedergabelautsprecherkanälen gemäß der Wasserzeichenmarkierung-Eignung von Audioinhalt für Wiedergabe durch jede der mindestens einen Teilmenge der Gruppen.

3. Verfahren nach Anspruch 1 oder Anspruch 2, außerdem einen folgenden Schritt enthaltend:

Erzeugen, nach den Schritten (a) und (b), von Lautsprecherspeisungen als Reaktion auf die

- Menge von wasserzeichenmarkierten Kanälen und mindestens einen nicht wasserzeichenmarkierten Kanal des Programms.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei das Programm eine Menge von Kanälen enthält, das Verfahren außerdem einen folgenden Schritt enthaltend:
- Wiedergeben des Programms durch Erzeugen von Lautsprecherpeisungen als Reaktion auf mindestens einige der Kanäle des Programms und wobei Schritt (a) einen Schritt des Auswählens einer Teilmenge der Lautsprecherpeisungen für Wasserzeichenmarkierung enthält und Schritt (b) einen Schritt des Wasserzeichenmarkierens mindestens eines Segments jeder Lautsprecherpeisung in der Teilmenge von Lautsprecherpeisungen enthält.
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei das Programm ein objektbasiertes Audioprogramm ist und das Verfahren einen folgenden Schritt enthält:
- Bestimmen einer Menge von Wiedergabelautsprecherkanälen, jeder für Wiedergabe durch einen verschiedenen Lautsprecher eines Wiedergabesystems, aus mindestens einem Objektkanal und/oder mindestens einem Lautsprecherkanal des Programms und wobei die in Schritt (a) ausgewählte Teilmenge von Kanälen eine Teilmenge der Menge von Wiedergabelautsprecherkanälen ist.
6. Verfahren nach einem der Ansprüche 1 bis 5, wobei das Programm Wasserzeichenmarkierung-Metadaten enthält, das Verfahren einen Schritt des Betriebens eines Decodierers zum Decodieren und Wiedergeben des Programms enthält und Schritt (a) einen Schritt des Auswählens der Teilmenge von Kanälen unter Verwendung der Wasserzeichenmarkierung-Metadaten enthält, und wobei wahlweise die Wasserzeichenmarkierung-Metadaten Wasserzeicheneignungswerte sind, jeder der Wasserzeicheneignungswerte eines Segments des Programms Eignung für Wasserzeichenmarkieren von Audioinhalt eines korrespondierenden Kanals des Programms in dem Segment angibt.
7. Verfahren nach einem der Ansprüche 1 bis 6, wobei der Wasserzeichenmarkierung-Eignungswert für mindestens einen Kanal des Segments mindestens teilweise aus einer Anzahl von Lautsprecher bestimmt wird, die anzusteuern sind, um Inhalt auszugeben, der den Kanal während der Wiedergabe des Segments angibt.
8. Audiowiedergabesystem, enthaltend:
- ein Decodierung-Untersystem (8), gekoppelt und konfiguriert zum Zergliedern und Decodieren eines codierten Bitstroms, um daraus Audiodaten und Metadaten zu extrahieren, die ein Mehrkanal-Audioprogramm angeben; und ein zweites Untersystem (11), gekoppelt und konfiguriert zum Auswählen einer Teilmenge von Kanälen mindestens eines Segments des Programms für Wasserzeichenmarkieren und zum Wasserzeichenmarkieren von Daten, die jeden Kanal in der Teilmenge von Kanälen angeben, wodurch eine Menge von wasserzeichenmarkierten Kanälen bestimmt wird, wobei die Auswahl der Teilmenge auf dem Programm oder einer Konfiguration von Wiedergabelautsprechern, die für Wiedergabe des Programms einzusetzen ist, basiert, wobei das zweite Untersystem konfiguriert ist, die Audiodaten eines Segments des Programms zu analysieren, um Werte zu bestimmen, die Wasserzeichenmarkierung-Eignung von Audioinhalt von Kanälen des Programms in dem Segment angeben, durch Bestimmen der quadratischen mittleren Amplitude der Audiodaten jedes der Kanäle in dem Segment oder durch Bestimmen der Energie oder der quadratischen mittleren Amplitude der Audiodaten in einem begrenzten Frequenzbereich jedes der Kanäle in dem Segment, und zum Auswählen der Teilmenge von Kanälen als Reaktion auf die Werte.
9. System nach Anspruch 8, wobei das zweite Untersystem konfiguriert ist, aus den Audiodaten und den Metadaten eine Menge von Wiedergabelautsprecherkanälen zu bestimmen, jeder für Wiedergabe durch einen verschiedenen Lautsprecher einer Menge von Lautsprechern, die in einer Wiedergabeumgebung installiert sind, und zum Auswählen einer Teilmenge der Menge von Wiedergabelautsprecherkanälen als die Teilmenge von Kanälen durch:
- Identifizieren von Gruppen der Lautsprecher, die in distinkten Zonen in der Wiedergabeumgebung installiert sind, so dass jede der Gruppen aus Lautsprechern besteht, die in einer verschiedenen einen der Zonen installiert sind, und Identifizieren von Wasserzeichenmarkierung-Eignung von Audioinhalt für Wiedergabe durch jede der Gruppen; und Auswählen der Teilmenge der Menge von Wiedergabelautsprecherkanälen gemäß der Wasserzeichenmarkierung-Eignung von Audioinhalt für Wiedergabe durch mindestens eine Teilmenge der Gruppen.

10. System nach Anspruch 8 oder 9, wobei das Programm eine Menge von Kanälen enthält und das zweite Untersystem konfiguriert ist:

zum Wiedergeben des Programms durch Erzeugen von Lautsprecherpeisungen als Reaktion auf mindestens einige der Kanäle des Programms; und  
zum Auswählen einer Teilmenge der Lautsprecherpeisungen für Wasserzeichenmarkieren und Wasserzeichenmarkieren mindestens eines Segments jeder Lautsprecherpeisung in der Teilmenge von Lautsprecherpeisungen.

11. System nach einem der Ansprüche 8 bis 10, wobei das Programm ein objektbasiertes Audioprogramm ist, das zweite Untersystem konfiguriert ist zum Bestimmen einer Menge von Wiedergabelautsprecherkanälen, jeder für Wiedergabe durch einen verschiedenen Lautsprecher eines Wiedergabesystems, aus mindestens einem Objektkanal und/oder mindestens einem Lautsprecherkanal des Programms und zum Auswählen einer Teilmenge der Menge von Wiedergabelautsprecherkanälen als die Teilmenge von Kanälen.

12. System nach einem der Ansprüche 8 bis 11, wobei das Programm Wasserzeichenmarkierung-Metadaten enthält, das Decodierung-Untersystem konfiguriert ist zum Extrahieren der Wasserzeichenmarkierung-Metadaten und das zweite Untersystem konfiguriert ist zum Verwenden der Wasserzeichenmarkierung-Metadaten zum Auswählen der Teilmenge von Kanälen für Wasserzeichenmarkieren, und wobei wahlweise die Wasserzeichenmarkierung-Metadaten Wasserzeicheneignungswerte sind, jeder der Wasserzeicheneignungswerte eines Segments des Programms Eignung für Wasserzeichenmarkieren von Audioinhalt eines korrespondierenden Kanals des Programms in dem Segment angibt.

13. System nach einem der Ansprüche 8 bis 12, wobei der Wasserzeichenmarkierung-Eignungswert für mindestens einen Kanal des Segments mindestens teilweise aus einer Anzahl von Lautsprecher bestimmt wird, die anzusteuern sind, um Inhalt auszugeben, der den Kanal während der Wiedergabe des Segments angibt.

14. Audiocodierer, konfiguriert zum Erzeugen eines Bitstroms, der ein codiertes Mehrkanal-Audioprogramm angibt, der Codierer enthaltend:

ein erstes Untersystem (2), gekoppelt und konfiguriert zum Erzeugen von Wasserzeichenmarkierung-Metadaten als Reaktion auf Segmente von Strömen von Audioinhalt, wobei die Wasserzeichenmarkierung-Metadaten Eignung für

Wasserzeichenmarkieren mindestens eines Segments von jedem der Ströme angeben oder die Wasserzeichenmarkierung-Metadaten angeben, ob Wasserzeichenmarkieren an mindestens einem Segment von jedem der Ströme durchgeführt werden sollte; und ein zweites Untersystem (6), gekoppelt und konfiguriert zum Erzeugen des Bitstroms, der das codierte Mehrkanal-Audioprogramm angibt, durch Codieren mindestens einiger der Ströme von Audioinhalt zum Erzeugen codierter Ströme von Audioinhalt und enthaltend in dem Bitstrom jeden der codierten Ströme von Audioinhalt, jeden der Ströme von Audioinhalt, der nicht codiert ist, und die Wasserzeichenmarkierung-Metadaten, wobei das erste Untersystem konfiguriert ist zum Analysieren mindestens eines Segments jedes der Ströme von Audioinhalt, um Werte zu bestimmen, die Wasserzeichenmarkierung-Eignung von Audioinhalt jedes der Ströme in dem Segment angeben, durch Bestimmen der quadratischen mittleren Amplitude des Audioinhalts jedes der Ströme in dem Segment oder durch Bestimmen der Energie oder der quadratischen mittleren Amplitude des Audioinhalts in einem begrenzten Frequenzbereich jedes der Kanäle in dem Segment und zum Auswählen der Teilmenge von Kanälen als Reaktion auf die Werte.

15. Codierer nach Anspruch 14, wobei der Wasserzeichenmarkierung-Eignungswert für mindestens einen Kanal des Segments mindestens teilweise aus einer Anzahl von Lautsprechern bestimmt wird, die anzusteuern sind, um Inhalt auszugeben, der den Kanal während der Wiedergabe des Segments angibt.

#### Revendications

1. Procédé de tatouage d'un programme audio multi-canaux, comportant les étapes consistant à :

(a) sélectionner un sous-ensemble de canaux d'au moins un segment du programme à tatouer, de telle sorte que la sélection du sous-ensemble soit basée sur le programme ou sur une configuration de haut-parleurs de reproduction à employer pour la reproduction du programme ; et  
(b) tatouer (11) chaque canal dans le sous-ensemble de canaux, générant ainsi un ensemble de canaux tatoués ; et  
(c) analyser (2) le contenu audio dans un segment du programme pour déterminer des valeurs indicatives de l'adéquation du tatouage du contenu audio de canaux du programme dans



le segment,

dans lequel l'étape (a) comporte une étape de sélection du sous-ensemble de canaux en réponse auxdites valeurs, et dans lequel l'étape (c) comporte :

une étape de détermination d'une amplitude efficace du contenu audio de chacun des canaux dans le segment, ou une étape de détermination de l'énergie ou de l'amplitude efficace du contenu audio dans une gamme de fréquences limitée de chacun des canaux dans le segment.

2. Procédé selon la revendication 1, comportant également les étapes consistant à :

déterminer à partir de canaux du programme un ensemble de canaux de haut-parleurs de reproduction, chacun servant à la reproduction par un haut-parleur différent d'un ensemble de haut-parleurs installés dans un environnement de reproduction, dans lequel le sous-ensemble de canaux du programme sélectionné à l'étape (a) est un sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction, et l'étape (a) comporte les étapes consistant à :

identifier des groupes des haut-parleurs qui sont installés dans des zones distinctes dans l'environnement de reproduction de telle sorte que chacun des groupes consiste en haut-parleurs installés dans une zone différente des zones, et identifier l'adéquation du tatouage du contenu audio à reproduire par chacun des groupes ; et sélectionner le sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction conformément à l'adéquation du tatouage du contenu audio à reproduire par chacun d'au moins un sous-ensemble des groupes.

3. Procédé selon la revendication 1 ou la revendication 2, comportant également une étape consistant à :

après les étapes (a) et (b), générer des signaux d'alimentation des haut-parleurs en réponse à l'ensemble de canaux tatoués et d'au moins un canal non tatoué du programme.

4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel le programme comporte un ensemble de canaux, ledit procédé comportant également une étape consistant à :

rendre le programme y compris en générant des

signaux d'alimentation de haut-parleurs en réponse à au moins certains des canaux du programme, et dans lequel l'étape (a) comporte une étape de sélection d'un sous-ensemble des signaux d'alimentation de haut-parleurs à tatouer, et l'étape (b) comporte une étape de tatouage d'au moins un segment de chaque signal d'alimentation de haut-parleur dans le sous-ensemble de signaux d'alimentation de haut-parleurs.

5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel le programme est un programme audio à base d'objet, et ledit procédé comporte une étape consistant à :

déterminer un ensemble de canaux de haut-parleurs de reproduction, chacun pour une reproduction par un haut-parleur différent d'un système de reproduction, à partir d'au moins un canal objet et/ou d'au moins un canal de haut-parleur du programme, et dans lequel le sous-ensemble de canaux sélectionnés à l'étape (a) est un sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction.

6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel le programme comporte des métadonnées de tatouage, ledit procédé comporte une étape d'exploitation d'un décodeur pour décoder et rendre le programme, et l'étape (a) comporte une étape de sélection du sous-ensemble de canaux en utilisant les métadonnées de tatouage, et facultativement, dans lequel les métadonnées de tatouage sont des valeurs d'adéquation de tatouage, chacune des valeurs d'adéquation de tatouage d'un segment du programme indique l'adéquation du tatouage du contenu audio d'un canal correspondant du programme dans le segment.

7. Procédé selon l'une quelconque des revendications 1 à 6, dans lequel la valeur d'adéquation du tatouage d'au moins un canal du segment est déterminée au moins partiellement à partir d'un nombre de haut-parleurs à commander pour émettre un contenu indicatif du canal durant la reproduction du segment.

8. Système de reproduction audio, comportant :

un sous-système de décodage (8), couplé et configuré pour analyser syntaxiquement et décoder un flux binaire codé afin d'en extraire des données audio et métadonnées indicatives d'un programme audio multicanal ; et un second sous-système (11), couplé et configuré pour sélectionner un sous-ensemble de canaux d'au moins un segment du programme à tatouer, et tatouer des données indicatives de

- chaque canal dans le sous-ensemble de canaux déterminant ainsi un ensemble de canaux tatoués, dans lequel la sélection du sous-ensemble est basée sur le programme ou sur une configuration de haut-parleurs de reproduction à employer pour la reproduction du programme, dans lequel le second sous-système est configuré pour analyser les données audio d'un segment du programme pour déterminer des valeurs indicatives de l'adéquation du tatouage du contenu audio de canaux du programme dans le segment, y compris en déterminant une amplitude efficace des données audio de chacun des canaux dans le segment, ou en déterminant l'énergie ou l'amplitude efficace des données audio dans une gamme de fréquences limitée de chacun des canaux dans le segment, et sélectionner le sous-ensemble de canaux en réponse auxdites valeurs.
9. Système selon la revendication 8, dans lequel le second sous-système est configuré pour déterminer à partir des données audio et des métadonnées un ensemble de canaux de haut-parleurs de reproduction, chacun servant à la reproduction par un haut-parleur différent d'un ensemble de haut-parleurs installés dans un environnement de reproduction, et sélectionner un sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction comme sous-ensemble de canaux, y compris en identifiant des groupes des haut-parleurs qui sont installés dans des zones distinctes dans l'environnement de reproduction de telle sorte que chacun des groupes consiste en haut-parleurs installés dans une zone différente des zones, et identifiant l'adéquation du tatouage du contenu audio à reproduire par chacun des groupes ; et sélectionner le sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction conformément à l'adéquation du tatouage du contenu audio à reproduire par chacun d'au moins un sous-ensemble des groupes.
10. Système selon la revendication 8 ou la revendication 9, dans lequel le programme comporte un ensemble de canaux, et le second sous-système est configuré pour :
- rendre le programme y compris en générant des signaux d'alimentation de haut-parleurs en réponse à au moins certains des canaux du programme ; et sélectionner un sous-ensemble des signaux d'alimentation de haut-parleurs à tatouer, et tatouer au moins un segment de chaque signal d'alimentation de haut-parleur dans le sous-ensemble de signaux d'alimentation de haut-parleurs.
11. Système selon l'une quelconque des revendications 8 à 10, dans lequel le programme est un programme audio à base d'objet, le second sous-système est configuré pour déterminer un ensemble de canaux de haut-parleurs de reproduction, chacun pour une reproduction par un haut-parleur différent d'un système de reproduction, à partir d'au moins un canal objet et/ou d'au moins un canal de haut-parleur du programme, et sélectionner un sous-ensemble de l'ensemble de canaux de haut-parleurs de reproduction comme sous-ensemble de canaux.
12. Système selon l'une quelconque des revendications 8 à 11, dans lequel le programme comporte des métadonnées de tatouage, le sous-système de décodage est configuré pour extraire les métadonnées de tatouage, et le second sous-système est configuré pour utiliser les métadonnées de tatouage afin de sélectionner le sous-ensemble de canaux à tatouer, et facultativement, dans lequel les métadonnées de tatouage sont des valeurs d'adéquation de tatouage, chacune des valeurs d'adéquation de tatouage d'un segment du programme est indicative de l'adéquation du tatouage du contenu audio d'un canal correspondant du programme dans le segment.
13. Système selon l'une quelconque des revendications 8 à 12, dans lequel la valeur d'adéquation du tatouage d'au moins un canal du segment est déterminée au moins partiellement à partir d'un nombre de haut-parleurs à commander pour émettre un contenu indicatif du canal durant la reproduction du segment.
14. Codeur audio configuré pour générer un flux binaire indicatif d'un programme audio multicanaux codé, ledit codeur comportant :
- un premier sous-système (2) couplé et configuré pour générer des métadonnées de tatouage en réponse à des segments de flux de contenu audio, dans lequel les métadonnées de tatouage sont indicatives de l'adéquation du tatouage d'au moins un segment de chacun des flux, ou les métadonnées de tatouage sont indicatives qu'un tatouage doit être exécuté ou non sur au moins un segment de chacun des flux ; et un second sous-système (6) couplé et configuré pour générer le flux binaire indicatif du programme audio multicanaux codé, y compris en codant au moins certains des flux de contenu audio pour générer des flux de contenu audio codés, et en incluant dans le flux binaire chacun des flux de contenu audio codés, chacun des flux de contenu audio qui n'est pas codé, et les métadonnées de tatouage, dans lequel le premier sous-système est configuré pour analyser au moins un segment de

chacun des flux de contenu audio pour déterminer des valeurs indicatives de l'adéquation du tatouage du contenu audio de chacun des flux dans le segment, y compris en déterminant une amplitude efficace du contenu audio de chaque dit flux dans le segment, ou en déterminant l'énergie ou l'amplitude efficace du contenu audio dans une gamme de fréquences limitée de chacun des canaux dans le segment, et sélectionner le sous-ensemble de canaux en réponse auxdites valeurs.

5

10

15. Codeur selon la revendication 14, dans lequel la valeur d'adéquation du tatouage d'au moins un canal du segment est déterminée au moins partiellement à partir d'un nombre de haut-parleurs à commander pour émettre un contenu indicatif du canal durant la reproduction du segment.

15

20

25

30

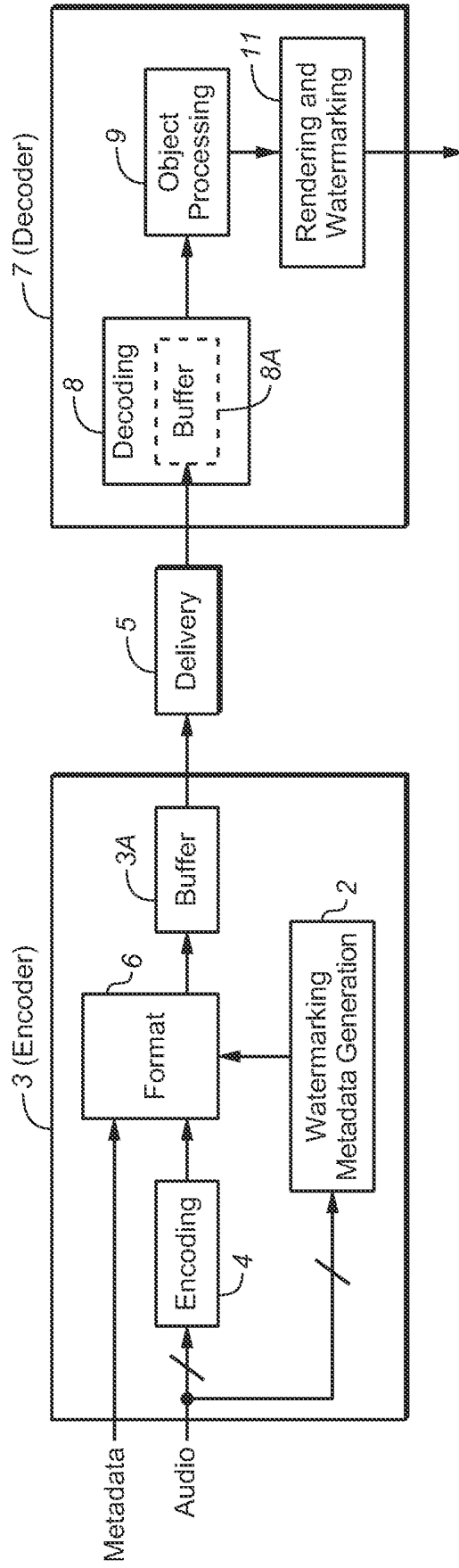
35

40

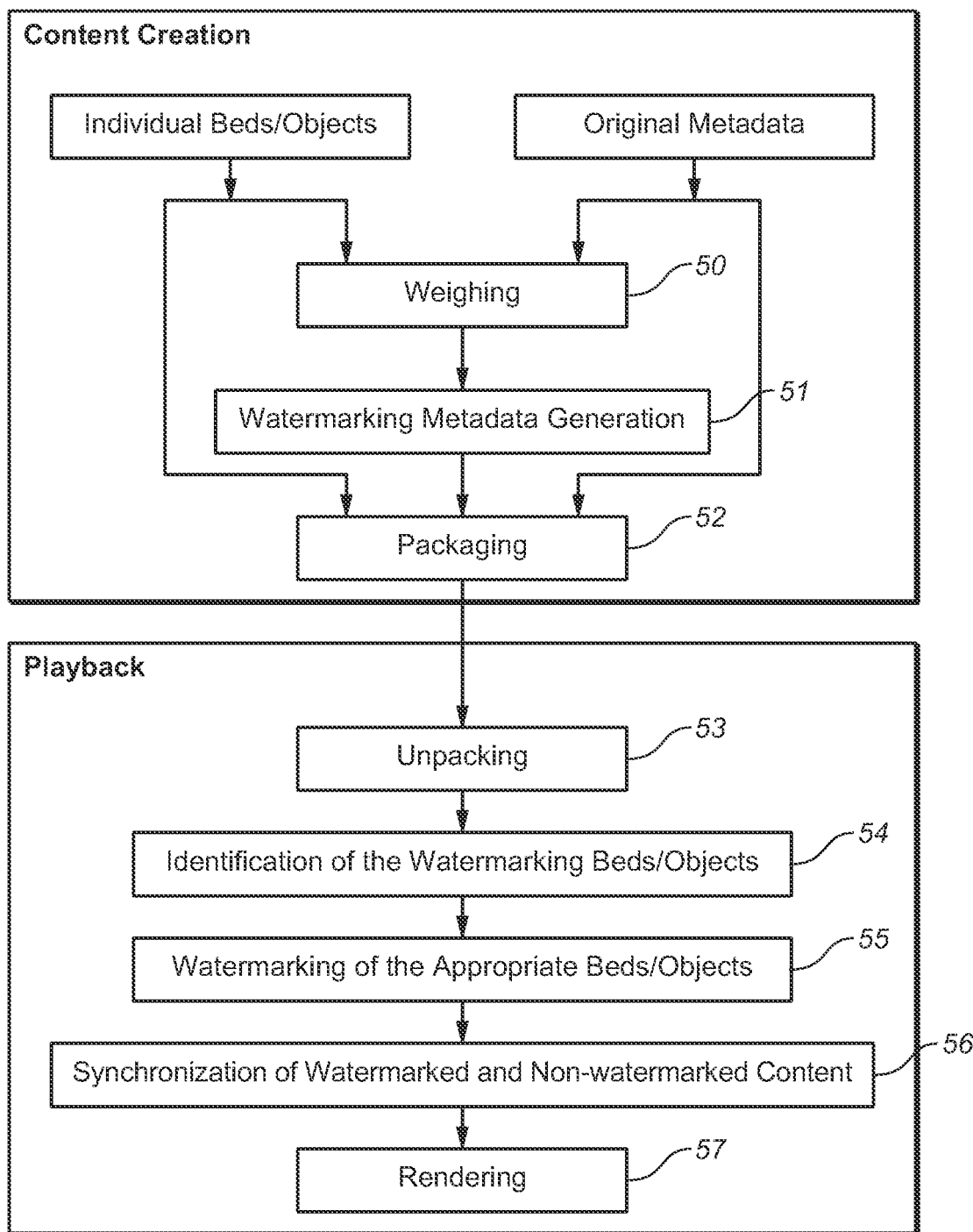
45

50

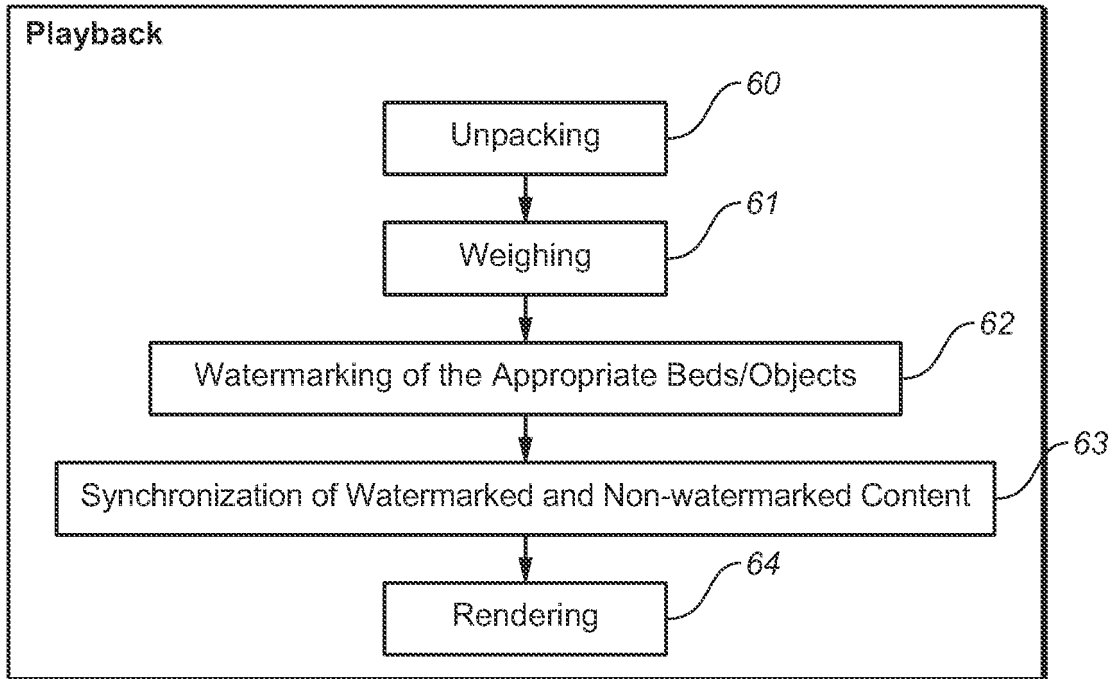
55



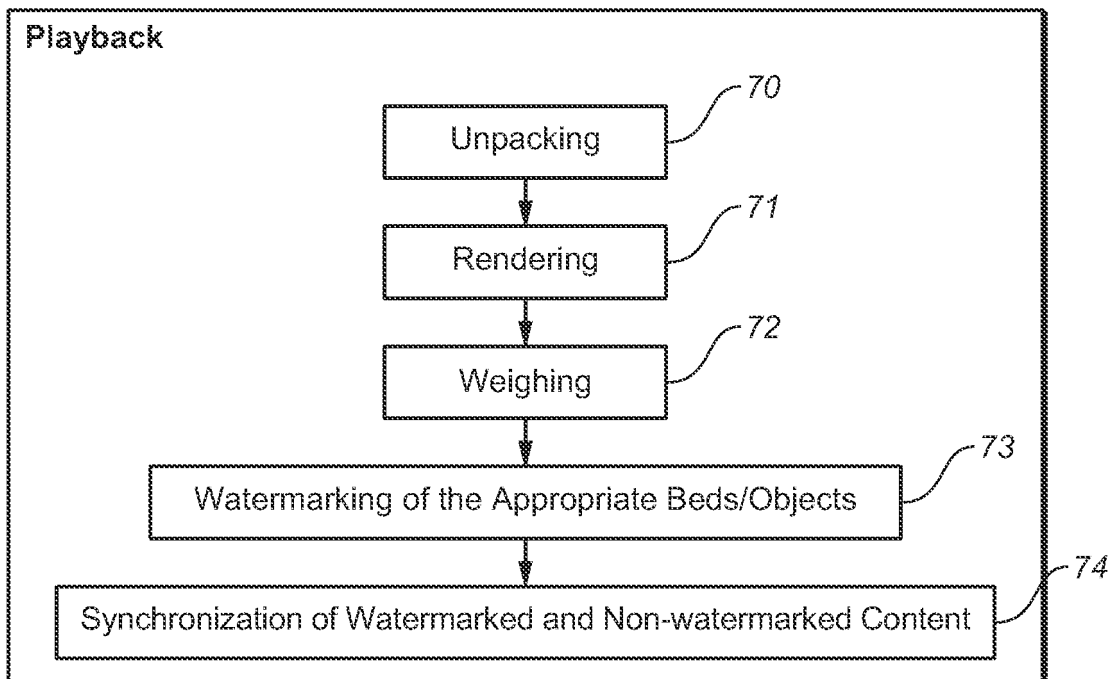
**FIG. 1**



**FIG. 2**

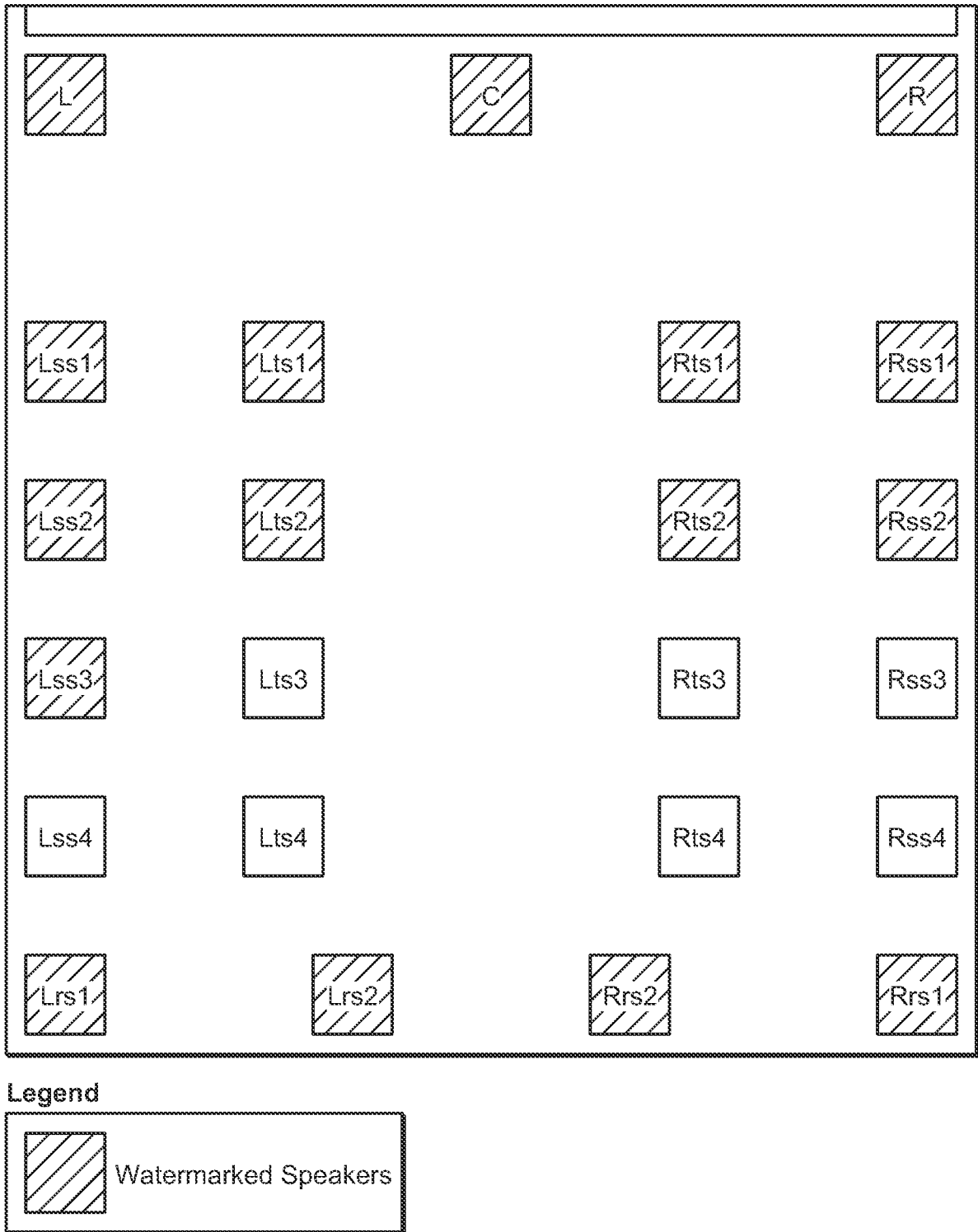


**FIG. 3**



**FIG. 4**

Watermarked Speakers in an Auditorium



**FIG. 5**

**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- US 61877139 B [0001]
- US 2001028783 W [0011]
- WO 2011119401 A2 [0011]
- EP 2562748 A [0012]