

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
19 November 2009 (19.11.2009)

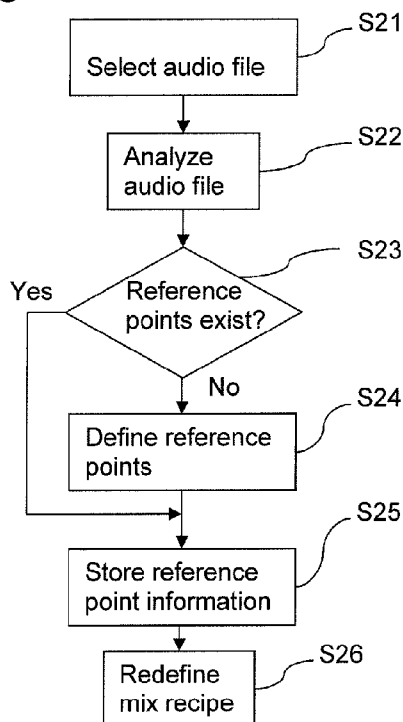
(10) International Publication Number
WO 2009/138425 A1

- (51) International Patent Classification: *G11B 27/10* (2006.01)
- (21) International Application Number: PCT/EP2009/055765
- (22) International Filing Date: 13 May 2009 (13.05.2009)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0801130-6 16 May 2008 (16.05.2008) SE
- (71) Applicant (for all designated States except US): **TONI-UM AB** [SE/SE]; P.O. Box 20075, Ölandsgatan 42, S-104 60 Stockholm (SE).
- (71) Applicants (for US only): **KLEIJN, Willem, Bastiaan** [NL/SE]; Tallåsvägen 11, S-182 75 Stocksund (SE). **EL-KHOURY, Jad** [AU/SE]; Trollesundsvägen 15, S-124 32 Bandhagen (SE).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **NORBERG, Jonas** [SE/SE]; Grindsgatan 54, S-118 57 Stockholm (SE).
- (74) Agent: **HERBJORNSEN, R.**; Albihns AB, P.O. Box 5581, Valhallavägen 117, S-114 85 Stockholm (SE).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ,

[Continued on next page]

(54) Title: AUDIO MIX INSTRUCTION FILE WITH TIMING INFORMATION REFERRING TO UNIQUE PATTERNS WITHIN AUDIO TRACKS

Fig. 2



(57) Abstract: The present invention facilitates the sharing of mix instructions files by taking into account timing differences between different files holding the same audio information. Unique reference points or segments are defined in each audio file and are used to determine the timing information related to the respective audio file within the mix instructions file.

TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

AUDIO MIX INSTRUCTION FILE WITH TIMING INFORMATION REFERRING TO UNIQUE PATTERNS WITHIN AUDIO TRACKS

Technical Field

5 The present invention relates to methods concerning the creation of mix instructions files, or mix instructions files for audio files, to computer program products and to an apparatus for playing audio files.

Background and Related Art

10 Since it has become possible to store music and other audio information in data files on hard disks or other memory, DJ systems have been developed in which the songs that are to be played are stored in a memory, for example, a data base from which they can be retrieved and played when desired. Typically these systems comprise similar functions to the traditional DJ systems in that they enable the mixing of two tracks and manipulation of each of the tracks to achieve a good mix, for example, a
15 smooth transition between two songs.

In applicant's co-pending application PCT/SE2006/050030 a hand-held DJ system comprising a data base for music tracks is proposed in which one set of controls is used to control both tracks.

20

Recently systems have become available in which the parameters used to achieve a certain mix can be stored in a mix instructions file and applied to the music files at a later time. The mix instructions file is sometimes referred to as a mix recipe or a mix recipe file. Such a system is the TRAKTOR DJ Studio 3 available from Native Instruments Software Synthesis GmbH. The content of the mix instructions file can be
25 created, for example, by registering the actions performed during a DJ session, that is, which music tracks are used and how the DJ manipulates them. This may be assisted by software that, for example, adjusts the playback speed of two tracks. For example, the mix instructions file can comprise data such as the music files' identification, the point in each music file where the playback should start from, and if the
30

playback speed should be modified, how much and at which point and/or in which interval. The system may apply the data in the mix instructions file to recreate the same mix again based on the same music files, which are stored in the system. This is a great improvement over the method previously known, which was to record the entire mix. Also, this system makes it easier to edit the mix since the parameters set in the mix instruction file can be altered instead of recording the entire mix again.

As disc jockey systems become more and more computer based, the possibility to make mixes of music tracks and share them with other users arises. Co-pending patent application PCT/SE2007/050491 discloses the possibility to create a so-called “mix instructions file” identifying each music track that is used in the mix, and setting parameters for the tracks and how to mix them, but without including the actual music tracks themselves. The mix instructions file can then be shared with other users without having to include all the music tracks. This means that a smaller file can be shared, and that copyright issues can be avoided in that each person using the mix instructions file must obtain the music files.

In this document the creator of such a recipe file will be referred to as a “Recipe Creator”, and users of such a file will be referred to as “Recipe Reader”. Note that a recipe file normally has a single Creator, whereas it can be shared by one or more Readers. Of course, any given user may be both a reader and a creator.

When the creator shares a certain mix instructions file with many readers, each reader is expected to (1) identify the audio files that are needed in the mix, and (2) follow the included instructions that use these files in order to reproduce the expected mix result. When executing the included instructions, the timing of the instructions plays a critical role in producing the same results as expected by the creator of the mix.

This is especially critical since the reader does not have access to the exact same files used by the creator. Instead, another audio file that is identified as the same as the creator's is used.

- 5 Given that various digital representations of an audio track exist, a problem may arise when a reader uses one audio file representation that differs from that used by the creator, even though they are identified as representing the same audio track.

10 For example, the starting point of the music within the file may vary. Also, the files may have been created using different encoding and/or different sampling frequencies. Hence, although the audio contents of the files, as experienced by the user, are similar, the files may not be identical when the time-scale is considered and one file cannot necessarily be replaced for another without some adjustments.

15 Unless the difference in the timing characteristics of the different representations of the same audio track are handled by the reader of the recipe file, the mix experienced by a reader may differ from that expected of the mix creator.

Object of the Invention

It is an object of the invention to facilitate the sharing of mix instructions files.

20

Summary of the Invention

This object is achieved according to the present invention by a method of processing an audio file representing an audio track, comprising the steps of analyzing the audio file to define at least one reference point or reference segment in
25 the audio file, said reference point or segment indicating an identifiable pattern within the audio file,
said reference point or reference segment identifying a unique position within the audio track, independent of the audio file representing the audio track,
storing information related to the reference point or reference segment,
30 using this pattern to align files for use in a mix.

This first method of the invention enables the analysis of an audio track to determine its timing within the file containing it, to enable the use of the audio track with an existing mix instructions file.

5

The object is also achieved by a method of creating a mix instructions file specifying a playback mix of at least a first and a second audio file

Said method comprising the steps of:

- identifying an audio track to be used in the mix instructions file,
- 10 - storing in the mix instructions file information related to the playback of the audio track in the mix, including timing information,
- analyzing the audio file to define at least one reference point or reference segment in the audio file, said reference point or segment indicating an identifiable pattern within the audio file independent of the audio file representation,
- 15 - storing information related to the reference point or reference segment in the mix instructions file,
- defining the timing information in relation to the at least one reference point or reference segment.

20 This second method enables the creation of mix instructions files using audio files having a determined timing of the audio information within the file.

The identifiable pattern should define a position within the audio track as uniquely as possible.

25

The object is also achieved by a method of playing a mix of at least a first and a second audio track using a mix instructions file, said method comprising the following steps:

- identifying an audio file comprising an audio track used by the mix instructions
30 file,

- identifying at least one reference point or reference segment in the mix instructions file, associated with the audio track, said reference point or segment indicating an identifiable pattern within the identified audio file,
- analyzing the audio file to define at least one indication point or indication segment in the audio file matching the at least one reference point or reference segment,
- comparing the indication point or segment to at least one reference point or segment stored in the mix instructions file,
- aligning the audio file on the basis of the comparison

10

This method enables the playing of a mix instructions file generated according to the second method above, using audio files generated according to the first method above.

15

By introducing in the mix instructions file reference time points or segments for each audio file that is used in the mix, any reader can use a mix instructions file created and shared by another creator, even if the audio files to be used by the mix instructions file do not match the exact digital representations used by the creator.

20

Unique reference points or segments are defined in each audio file and are used to determine the timing information related to the respective audio file within the mix instructions file. The reference time points or segments are expected to be purely based on the audio properties of the corresponding track, and are independent of the digital representation of the audio tracks.

25

To align the two audio files sufficiently precisely to be able to replace one with the other in a mix instructions file, the reference time points or segments require a sufficiently high resolution

30

According to the invention, therefore, the reader of a mix instructions file is enabled to use a different set of tracks that are identified as similar to those used by the Mix

Creator, yet the tracks used by the Reader may differ in quality and digital representation, causing a difference in the timing properties of the tracks. So, unless this difference in timing between the tracks is taken care of, the timing of the mix instructions during playback may no longer match the Creator's timing.

5

In any of the three methods above, the reference point or segments may be defined, for example, as the point in the audio file having the highest amplitude. This will constitute an easily identifiable reference point in each audio file. Of course a defined number of amplitude maxima may be used to obtain several reference points.

10

Alternatively a beat analysis may be performed to define the reference point or segment.

Each of the methods may further comprise the following steps:

15

Dividing the second file into frames,
Performing a Fourier Transform on at least two frames,
Converting the transformed signal to periodograms to obtain corresponding spectra,
Determining the difference between the at least two frames,
Selecting at least one processed frame for use in determining the offset between a
20 first and a second audio track comprising the same audio information.

25

In the latter case, the method may further comprise the steps of determining the offset between the first and the second audio track using cross-correlation, by identifying the maximum value of the cross-correlation of a first and a second signal, each of said first and second signal being a scalar or a vector signal, the first signal representing a characteristic segment of the first track and the second signal representing at least a part of the second track, then determining the time scale to determine the offset.

Preferably, to increase the resolution, the method further comprises the steps of obtaining a first and a second cross-correlation value and interpolating between the first and second cross-correlation values.

- 5 The method may also further comprise the step of converting the spectra corresponding to segment(s) to a representation that uses a perception-based frequency scale before comparing the characterizing sequences. This will result in a lower dimensionality of the vector signal, while keeping the most significant part of the audio signal, that is, the part that is in the audible range.

10

The characterizing sequence of a track may be selected as the one having the highest entropy of a number of such sequences. This will optimize the reliability of the matching.

15

The invention also relates to a computer program product characterized in that it comprises computer readable code means which, when run in an apparatus for playing audio tracks, will cause the apparatus to perform any of the above methods, and to an apparatus for playing audio tracks, comprising such a computer program product.

20

Brief Description of the Drawings

The invention will be described in more detail in the following, by way of example and with reference to the appended drawings in which:

Fig. 1 illustrates a DJ system that may be used according to the invention,

25

Fig. 2 is a flow chart of a method for creating a mix instructions file according to an embodiment of the invention,

Fig. 3 is a flow chart of a method for using a mix instructions file with new audio files,

Fig. 4 is a graphical representation of the audio content of an example audio file,

Fig. 5 is a flow chart of a method for determining reference points or segments according to a second embodiment of the invention,

Fig. 6 is a flow chart of a method of determining the offset between two files.

5 Detailed Description of Embodiments

Fig. 1 is a simplified version of Fig. 2 of co-pending application PCT/SE2007/050491. A more detailed description is given in this co-pending application. Fig. 1 illustrates a DJ system according to the invention. A computer 23
10 comprises a first data base 25 for holding audio files and a second data base 27 for holding mix instructions files similar to the ones known in the art. The computer has user input/output means represented by a keyboard 6 and a screen 8.

Mix instructions files may be created in the computer and/or retrieved from another source, for example, through the Internet, and stored in the second data base 27.

15

According to the invention the computer comprises a playback software program 33 arranged to retrieve a mix instructions file from the second data base 27 and, as prescribed in the mix instructions file, to retrieve at least one audio file comprising a music track at a time from the first data base 25 and manipulate them according to
20 the mix instructions file to create a mix of the music tracks. The computer preferably also has retrieval software 35 for retrieving, through a data network 36, mix instructions files and/or audio files from external sources. These may be used directly upon retrieval or may be stored in the second data base 27.

25

According to the invention, even if the mix instructions file is based on music tracks that are found in the first data base 25, it may be that the copies of the music tracks do not have the exact same digital representation of the tracks as the files used when the mix instructions file is created. The computer also preferably has retrieval software 35 for retrieving, through a data network 36, mix instructions files and/or mu-

sic tracks from sources such as a music track data base 37 or a mix instructions file data base 39 in the network.

The portable DJ system also comprises a mix creation unit 40 to enable the operator
5 to create mix instructions files in the portable DJ system and/or the direct retrieval
of mix instructions files and/or music tracks from the network to the portable DJ
system. From the creator's point of view, the mix instructions files are created by
user input to the computer in a way known per se. According to the invention, the
portable DJ system is arranged to analyze each audio file used in a mix instructions
10 file and define one or more reference time points or reference segments in the file to
be included in the mix instructions file. The system is also arranged to define all
timing properties of the mix instructions file in relation to the reference points or
reference segments. In this way, when the mix instructions file is used with other
audio files, the same reference time points, or segments, can be identified in these
15 other audio files and used to align the audio file as required for use in the mix. As
will be understood, this is achieved by means of software arranged in the portable
DJ system. How this can be achieved will be discussed in more detail in the follow-
ing.

20 It will be understood that each of the units 33, 35 and 40 in the computer 23 and the
playback unit 43 of the portable unit 41 are implemented as software modules stored
in the computer, or portable unit, respectively. As the skilled person will understand,
the illustration shown in Figure 2 is only a logical diagram, and the actual functions
can be implemented in software in a number of different ways.

25 As will be understood, different audio files that use different digital representations,
yet have the same audio content may differ in several respects. For example they
may have different delay times before the audio signal actually starts, or after the
audio signal ends. Also, they may have a different file format, have been created us-
30 ing a different encoding scheme and/or sampling frequency. Further, they may be

affected by noise in different ways. Hence, for example, if a reference time point has been set a number of seconds into a music track based on a file of a particular representation of the track, if a different representation of the track, obtained from another source, is used, the reference time point may be placed in the wrong position relative to the actual content of the file, if the reference time point is simply placed at the same number of seconds into the music track. According to the invention, therefore, the mix creation unit 40 of the portable DJ system is arranged, when creating a mix, to identify one or more reference points, or reference segments, which are well defined points or segments in each music track, related to the audio content of the music track. According to the invention, any time properties in the mix instructions file (such as the point in time of a cue point) are defined in relation to the reference points or reference segments of the appropriate track, making the mix instructions file independent of any digital representation of the included audio tracks. According to the invention, the mix playback unit 33 of the portable DJ system is also arranged, when reading a mix, to identify these reference points, or reference segments in order to properly align the reader's audio tracks. These reference time points or segments may be determined in a number of different ways, as will be discussed below.

According to the invention, therefore, when creating a mix instructions file each of the audio files used in the mix instructions file should have one or more reference time points, or reference segments defined, and the definitions should be included in the mix instructions file. The reference time points or segments may be specified in the audio file, or stored in a separate data base, associated to the audio file. Examples of how to do this will be given in the following. An overall procedure for creating the mix instructions file is shown in Figure 2.

In step S21 a reference to an audio file is included in the mix instructions file. The audio file can be manipulated in any way that is common in the art, to mix it with one or more other audio files, increase its speed etc.

In step S22 the audio file is analyzed to see if it already has reference time points or segments defined for it. The reference time points or segments can be either directly embedded in the audio file, or stored in an independent data base that refers to the audio file.

5 Step S23 is a decision step. If reference points or segments exist for the file, go to step S25, if not, go to step S24.

In step S24 at least one reference point or segment is defined for the file. This can be done in a number of different ways, some of which will be discussed below.

10 In step S25 information about the reference time points, or segments is stored in connection with the audio file itself, and with the mix instructions file.

In step S26, the mix instructions file is redefined so that any time properties in the mix instructions file are defined in relation to the reference points or reference segments for the corresponding audio track.

15 Figure 3 illustrates a method of using a mix instructions file retrieved from another source, with the reader's own audio files. As discussed above, these may not be identical to the ones used when creating the mix instructions file. Co-pending application PCT/SE2007/050491 discloses a method for retrieving audio tracks to be used with the mix instructions file, if necessary.

20

In step S31 the mix instructions file is searched to identify the audio files used by it. Then for each audio file, the following steps are performed:

In step S32 the audio file is analyzed to see if it has reference time points, or segments, that match the ones defined for the audio file in the mix instructions file.

25 This means that the references should be created in the same way, so that they represent the same characteristic of the audio file.

Step S33 is a decision step: if matching references are found, go to step S35; if not, go to step S34.

In step S34 reference time points, or segments, are defined for the audio file in the same way as was done in S24 of Figure 2. Information about the references is stored in association with the audio file.

5 In step S35 the references created in step S34, or found in step S33, are compared to the reference time points for the audio file in the mix instructions file, and the result of the comparison is used to align the audio file with the one used when creating the mix instructions file.

10 The alignment of the reference time points or reference segments can be performed by determining the time difference between the reference points or reference segments of the track to be used and the reference points or reference segments of the same track as defined in the mix instructions file. This difference can then be offset for each time property in the mix instructions file, which is defined in relation to the creator's reference points or reference segments for that track. If the differences are
15 not identical for all reference time points or reference segments, then the difference can advantageously be interpolated between such points (before the first or after the last time point or segment the difference can be held constant). Linear interpolation of the differences can be used.

20 When all the audio files to be used with the mix instructions file have been aligned properly, the mix instructions file can be used together with the reader's audio files to play the mix. It would of course be possible to align the audio files as they become needed while playing the mix as well, if this could be handled fast enough. It would of course also be possible to first redefine the mix instructions file so that any
25 time properties in the mix instructions file are defined in relation to the reader's audio files, before using the mix instructions file.

30 With reference to Figure 4, some methods of determining the reference points, or segments, in the time domain will be discussed. Figure 4 shows a graphical representation in the time domain of the audio content of an audio file. As can be seen,

the audio content has clearly distinguishable features, such as amplitude peaks in a certain pattern. A simple way of determining a reference time point would be to find the maximum amplitudes in the track and determine the distance to this maximum peak from the start of the file. To create a more reliable reference, a number of the
5 highest peaks, for example three or five of the highest peaks could be used.

In another embodiment of the invention, the reference points are determined based on an analysis of certain time characteristics of the audio track such as the time positions of the audio beats in the track. An example of such an analysis is the beats
10 analysis produced by the software algorithm aufTAKT (<http://www.zplane.de/>). Such an analysis produces a vector of time positions. With the same analysis applied on different digital representation of the same audio track, different time vectors will be produced. However, given that the analysis focuses on the audio properties of the tracks, and is independent of the digital representations, the different time
15 vectors can be aligned to provide a time alignment of the different track representations.

In another embodiment of the invention, the reference points, or segments, are determined based on a sequence of periodograms, which form estimates of short-time
20 Fourier transforms, as will be discussed in more detail below.

Several methods are known for generating a hash signal. For example, WO
02/065782 discloses a method of generating a hash signal, also referred to as a signature. The hash signal may be seen as a summary of the file, and can be matched
25 with hash signals stored in a database to identify the information signal. This can be used, for example, to verify correct receipt of a large file by sending only the hash value of the file.

Fig. 5 is a flow chart of the steps performed to generate one or more sequences of periodograms, that can be used according to the invention to determine the offset between two files comprising the same music track.

In the first step S51, at least a part of the music track is divided into frames. A frame length of 20-30 ms has been found to be suitable. To increase the resolution, preferably, the frames overlap.

In step S52 a Fast Fourier Transform (FFT) is performed on each frame, to transform the frames to the frequency domain. The resulting frequency domain representation is discrete and each frequency point is referred to as a frequency "bin".

In step S53, for each frame, the signal is then converted to a periodogram by computing the energy for each frequency bin and discarding the phase. Advantageously the square-root of the periodogram spectrum values can be taken to obtain an estimate of the short-term magnitude spectrum. Below we refer to the estimate of the magnitude spectrum or the periodogram, whichever is used, as the "spectrum".

Step S54 is an optional step in which the spectrum is converted to a perceptual scale, that is, the spectrum values within each of a set of pre-defined frequency bands are summed and scaled, where the pre-defined frequency bands and scaling factors are selected in a manner consistent with the audible frequency range. The summations can be weighted for increased accuracy. The summations result in a "perceptual" spectrum with fewer bins. The well-known mel or ERB (equivalent rectangular band) scales can be used to construct the pre-defined frequency bands.

25

Advantageously, a processing step S55 can be included, that removes the undesired sensitivity to fixed offsets in the spectrum. Such fixed offsets can be caused by stationary noise. For each time frame, step S55 has as input a spectrum corresponding to that time frame. The signal is now represented by a sequence of spectra. For each frequency bin a scalar time sequence of spectrum values exists. Thus, we can distin-

30

guish a set of scalar frequency “channels”, each channel corresponding to one frequency bin. Each channel is a time signal that has one time sample per time frame of the original signal. These channels are significantly down-sampled compared to the original audio signal. The sequence of spectra forms a vector signal with each of the channels being a component of the vector. In step S55 each frame is compared to a previous frame to determine the difference between them. That is, for each channel an output sample the previous sample is subtracted from the current sample. The resulting vector signal of time differences of the spectra describes the changes in the spectra corresponding to successive time segments. Thus, such vector signals are less sensitive to stationary additive noise than the spectra themselves.

The difference signals are sensitive to the power level of the audio file. A simple method to remove information about the power level of a signal is to consider only the sign of the signal. That is, positive signal samples are represented by +1 and negative signal samples are represented by -1. In step S56 the sign of the difference signal is determined. The end result is a binary representation of the time differences of the sequence of spectrums.

In step S57 one or more processed frames are selected for use in the method depicted in Fig. 6, for determining the offset between the file used in Fig. 5 and another file comprising the same music track. The result, as illustrated by step S57, is a characteristic representation of a part of the audio file. We refer to this characteristic representation of the part as the “characteristic vector sequence”. Some characteristic sequences will be more efficient in the sense that they can be shorter to obtain a certain alignment performance. A suitable selection criterion for selecting such a part or parts of the signal to be used for computing the characteristic vector sequence would be entropy of the characteristic vector sequence, as matching such a characteristic vector sequence in different files is likely to have features facilitating alignment. However, the parts can also be selected manually. As a third option, the parts

can be selected as the part of 500 ms centered around the middle of the file used for the original recipe.

Fig. 6 is a flow chart of how the characteristic vector sequence obtained according to Fig. 5 can be used to find the offset between two files. As explained above, each characteristic vector sequence is a vector signal segment or a set of vector signal segments. In step S61 the characteristic vector sequence is cross-correlated with the entire vector sequence representation of the other file. This can be described as sliding the characteristic vector sequence across the vector sequence of the other file and cross-correlating at different positions to find the best match. By finding the maximum in the cross-correlation, the location of the characteristic vector sequence in the other file is found, which means that the time scales of the two files can be synchronized. It is noted that the cross-correlation is performed on a vector signal that is down-sampled significantly compared to the audio signal. Thus, in step S62 an interpolation is made between the distinct cross-correlation sample values obtained, to increase the resolution of the cross-correlation curve. In step S63 the maximum value of the cross-correlation curve is found, to identify more precisely the part of the other file that matches the characteristic vector sequence. In step S64 the time scale is determined to determine the offset between the two files.

20

The offset value obtained in step S64 may be used to correct the parameters in the mix instructions file related to the processed audio file.

Claims

1. A method of analyzing an audio file representing an audio track, comprising the steps of
- 5 analyzing the audio file to define at least one reference point or reference segment in the audio file, said reference point or segment indicating an identifiable pattern within the audio file,
said reference point or reference segment identifying a unique position within the audio track, independent of the audio file representing the audio track,
10 storing information related to the reference point or reference segment,
using this pattern to align files for use in a mix
2. A method of creating a mix instructions file specifying a playback mix of at least a first and a second audio file
- 15 Said method comprising the steps of:
- identifying an audio track to be used in the mix instructions file,
 - storing in the mix instructions file information related to the playback of the audio track in the mix, including timing information,
 - analyzing the audio file to define at least one reference point or reference segment in the audio file, said reference point or segment indicating an identifiable pattern within the audio file, independent of the audio file representation,
 - storing information related to the reference point or reference segment in the mix instructions file,
 - defining the timing information in relation to the at least one reference point or reference segment.
- 20
- 25
3. A method of playing a mix of at least a first and a second audio track using a mix instructions file, said method comprising the following steps:
- identifying an audio file comprising an audio track used by the mix instructions file,
- 30

- identifying at least one reference point or reference segment in the mix instructions file, associated with the audio track, said reference point or segment indicating an identifiable pattern within the identified audio file, independent of the audio file representation
- 5 - analyzing the audio file to define at least one indication point or indication segment in the audio file matching the at least one reference point or reference segment,
- comparing the indication point or segment to at least one reference point or segment stored in the mix instructions file
- 10 - aligning the audio file on the basis of the comparison

4. A method according to any one of the previous claims, wherein the reference point is defined as the point in the audio file having the highest amplitude.

- 15 5. A method according to any one of the claims 1-3, wherein a beat analysis is used to define the reference point or segment.

6. A method according to any one of the preceding claims, further comprising the following steps:

- 20 Dividing the second file into frames
Performing a Fourier Transform on at least two frames
Converting the transformed signal to periodograms to obtain corresponding spectra,
Determining the difference between the at least two frames
Selecting at least one processed frame for use in determining the offset between a
25 first and a second audio track comprising the same audio information.

7. A method according to claim 6, comprising the step of determining the offset between the first and the second audio track using cross-correlation, by identifying the maximum value of the cross-correlation of a first and a second signal, each of said
30 first and second signal being a scalar or a vector signal, the first signal representing

a characteristic segment of the first track and the second signal representing at least a part of the second track, then determining the time scale to determine the offset.

5 8. A method according to claim 7, comprising the steps of obtaining a first and a second cross-correlation value and interpolating between the first and second cross-correlation values to increase the resolution

10 9. A method according to any one of the claims 6-8, further comprising the step of converting the spectra corresponding to segment(s) to a representation that uses a perception-based frequency scale before comparing the characterizing sequences.

15 10. A method according to any one of the preceding claims wherein the characterizing sequence of a track is selected as the one having the highest entropy of a number of such sequences.

11. A computer program product characterized in that it comprises computer readable code means which, when run in an apparatus for playing audio tracks, will cause the apparatus to perform the method according to any one of the claims 1-9.

20 12. An apparatus for playing audio tracks, characterized in that it comprises a computer program product according to claim 11.

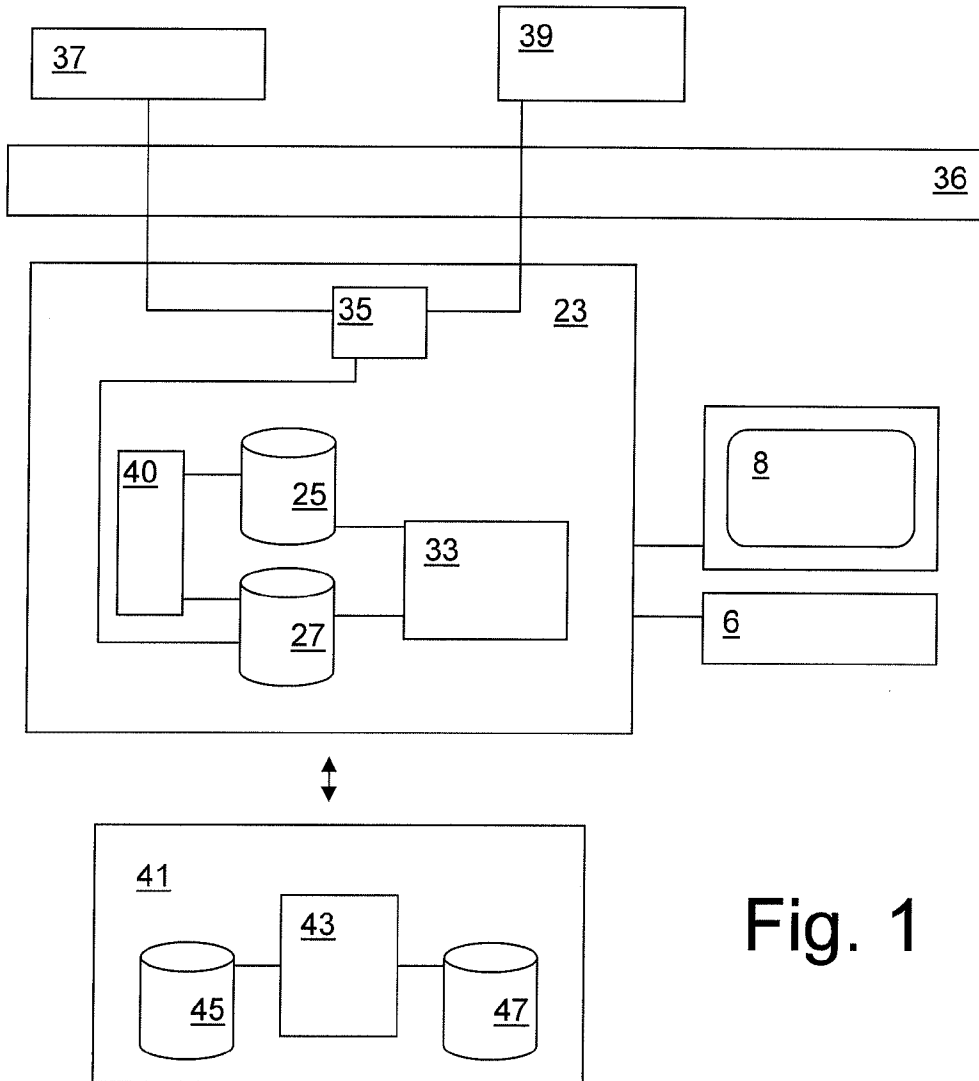


Fig. 1

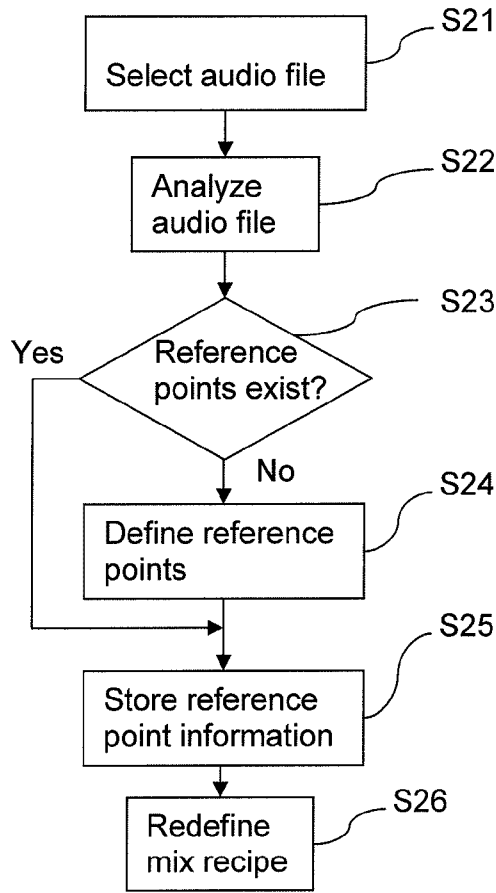


Fig. 2

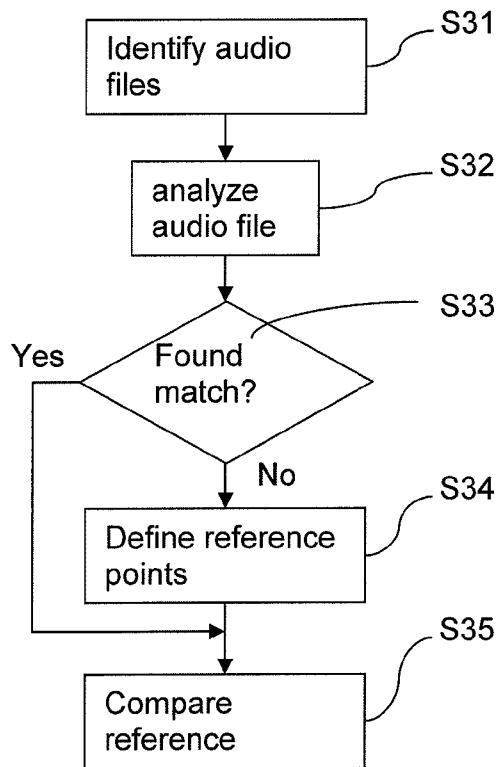


Fig. 3

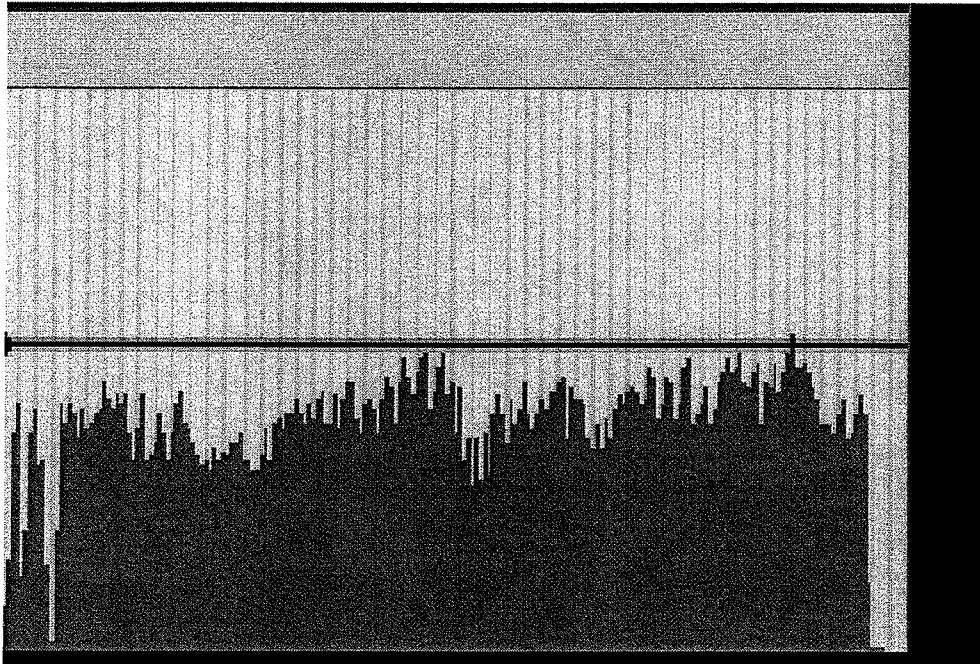


Fig. 4

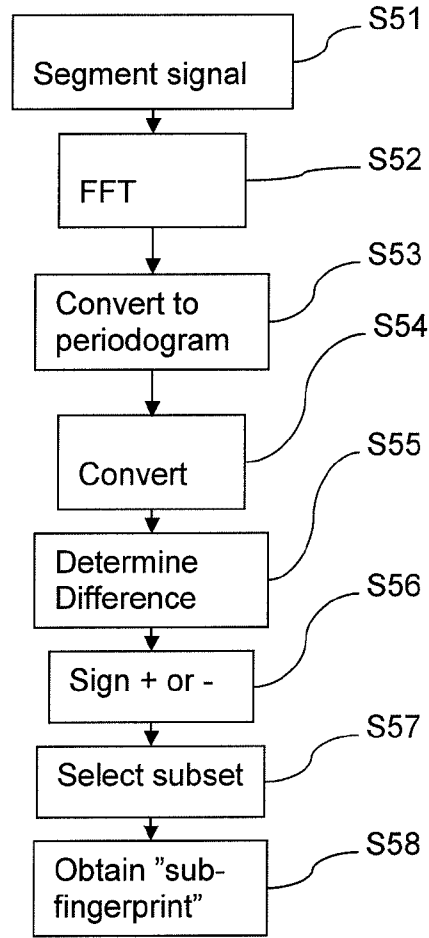


Fig. 5

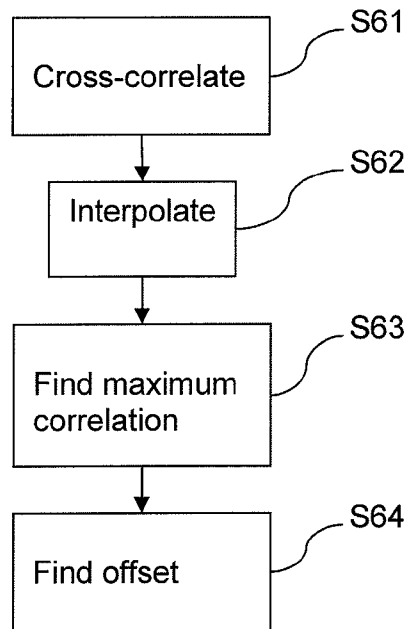


Fig. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2009/055765

A. CLASSIFICATION OF SUBJECT MATTER
INV. G11B27/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G11B G10H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2007/066813 A (SONY CORP [JP]; KOMORI AKIHIRO [JP]; YAMASHITA KOSEI [JP]; MIYAJIMA YA) 14 June 2007 (2007-06-14)	1-3, 11, 12
Y	paragraphs [0019], [0031], [0032], [0041] - [0043]; figures 5, 6, 11 - & EP 1 959 427 A (SONY CORP [JP]) 20 August 2008 (2008-08-20)	4-10
Y	US 2003/160944 A1 (FOOTE JONATHAN [US] ET AL FOOTE JONATHAN [US] ET AL) 28 August 2003 (2003-08-28) paragraphs [0010], [0037] - [0040], [0066], [0067]; figure 6	4-10
	----- -/--	

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents :

A document defining the general state of the art which is not considered to be of particular relevance	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
E earlier document but published on or after the international filing date	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
O document referring to an oral disclosure, use, exhibition or other means	* & * document member of the same patent family
P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 31 August 2009	Date of mailing of the international search report 04/09/2009
--	---

Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Authorized officer Seeger, Alexander
--	--

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2009/055765

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2003/065517 A1 (MIYASHITA MASAHICO [JP] ET AL) 3 April 2003 (2003-04-03) paragraphs [0005] - [0008], [0060] - [0062], [0075], [0084] - [0105], [0129] - [0140], [0154]; figures 1-3,8,9 -----	1,5,11, 12
X	US 2007/227337 A1 (YOSHIKAWA HAJIME [JP] ET AL) 4 October 2007 (2007-10-04) paragraphs [0010], [0050] - [0058], [0118] - [0120]; figures 1-3 -----	1,5,11, 12
X	EP 1 876 583 A (SONY CORP [JP]) 9 January 2008 (2008-01-09) paragraphs [0001], [0050] - [0069], [0083] - [0106]; figures 1,9,17,18,21,26 -----	1,5,11, 12

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2009/055765

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2007066813 A	14-06-2007	CN 101322179 A	10-12-2008
		EP 1959427 A1	20-08-2008
		KR 20080074975 A	13-08-2008
		US 2009044689 A1	19-02-2009
EP 1959427 A	20-08-2008	CN 101322179 A	10-12-2008
		WO 2007066813 A1	14-06-2007
		KR 20080074975 A	13-08-2008
		US 2009044689 A1	19-02-2009
US 2003160944 A1	28-08-2003	JP 2003259302 A	12-09-2003
US 2003065517 A1	03-04-2003	EP 1315143 A2	28-05-2003
		JP 2003108132 A	11-04-2003
US 2007227337 A1	04-10-2007	WO 2005104088 A1	03-11-2005
EP 1876583 A	09-01-2008	CN 101160615 A	09-04-2008
		JP 2006301471 A	02-11-2006
		WO 2006114998 A1	02-11-2006
		KR 20070122452 A	31-12-2007
		US 2009088877 A1	02-04-2009