PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

Dolby International AB
Patent Group Europe
Apollo Building, 3E
Herikerbergweg 1-35
NL-1101 CN Amsterdam Zuidoost
PAYS-BAS

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(PCT Rule 71.1)

Date of mailing *(day/month/year)*	07.02.2017

Applicant's or agent's file reference	**IMPORTANT NOTIFICATION**
D14089WO01	

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/EP2015/075022	28.10.2015	31.10.2014

Applicant
Dolby International AB

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary report on patentability and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary report on patentability. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

The applicant's attention is drawn to Article 33(5), which provides that the criteria of novelty, inventive step and industrial applicability described in Article 33(2) to (4) merely serve the purposes of international preliminary examination and that "any Contracting State may apply additional or different criteria for the purposes of deciding whether, in that State, the claimed inventions is patentable or not" (see also Article 27(5)). Such additional criteria may relate, for example, to exemptions from patentability, requirements for enabling disclosure, clarity and support for the claims.

Name and mailing address of the international preliminary examining authority:	Authorized Officer
European Patent Office P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Fax: +31 70 340 - 3016	Bello, Marcos Tel. +31 70 340-3704

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

(Chapter II of the Patent Cooperation Treaty)

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference D14089WO01	**FOR FURTHER ACTION**	See Form PCT/IPEA/416
International application No. PCT/EP2015/075022	International filing date *(day/month/year)* 28.10.2015	Priority date *(day/month/year)* 31.10.2014
International Patent Classification (IPC) or national classification and IPC INV. G10L19/008		
Applicant Dolby International AB		

1. This report is the international preliminary examination report, established by this International Preliminary Examining Authority under Article 35 and transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 8 sheets, including this cover sheet.
3. This report is also accompanied by ANNEXES, comprising:
   a. ☒ *(sent to the applicant and to the International Bureau)* a total of 17 sheets, as follows:
      - ☒ sheets of the description, claims and/or drawings which have been amended and/or sheets containing rectifications authorized by this Authority, unless those sheets were superseded or cancelled, and any accompanying letters (see Rules 46.5, 66.8, 70.16, 91.2, and Section 607 of the Administrative Instructions).
      - ☐ sheets containing rectifications, where the decision was made by this Authority not to take them into account because they were not authorized by or notified to this Authority at the time when this Authority began to draw up this report, and any accompanying letters (Rules 66.4bis, 70.2(e), 70.16 and 91.2).
      - ☐ superseded sheets and any accompanying letters, where this Authority either considers that the superseding sheets contain an amendment that goes beyond the disclosure in the international application as filed, or the superseding sheets were not accompanied by a letter indicating the basis for the amendments in the application as filed, as indicated in item 4 of Box No. I and the Supplemental Box (see Rule 70.16(b)).

   b. ☐ *(sent to the International Bureau only)* a total of (indicate type and number of electronic carrier(s)) , containing a sequence listing, in the form of an Annex C/ST.25 text file, as indicated in the Supplemental Box Relating to Sequence Listing (see paragraph 3*ter* of Annex C of the Administrative Instructions).

4. This report contains indications relating to the following items:
   - ☒ Box No. I Basis of the report
   - ☐ Box No. II Priority
   - ☐ Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - ☐ Box No. IV Lack of unity of invention
   - ☒ Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
   - ☐ Box No. VI Certain documents cited
   - ☒ Box No. VII Certain defects in the international application
   - ☐ Box No. VIII Certain observations on the international application

Date of submission of the demand	Date of completion of this report
26.05.2016	07.02.2017
Name and mailing address of the international preliminary examining authority: European Patent Office P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk - Pays Bas Tel. +31 70 340 - 2040 Fax: +31 70 340 - 3016	Authorized officer Taddei, Hervé Telephone No. +31 70 340-3359

INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY

International application No.
PCT/EP2015/075022

Box No. I Basis of the report

1. With regard to the **language**, this report is based on

   ☒ the international application in the language in which it was filed

   ☐ a translation of the international application into , which is the language of a translation furnished for the purposes of:
   - ☐ international search (under Rules 12.3(a) and 23.1(b))
   - ☐ publication of the international application (under Rule 12.4(a))
   - ☐ international preliminary examination (under Rules 55.2(a) and/or 55.3(a) and (b))

2. With regard to the **elements*** of the international application, this report is based on *(replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report)*:

Description, Pages

1-65	as originally filed	

Claims, Numbers

1-40	filed with the letter of	14-11-2016

Drawings, Sheets

1/6-6/6	as originally filed	

☐ a sequence listing - see Supplemental Box Relating to Sequence Listing.

3. ☐ The amendments have resulted in the cancellation of:
   - ☐ the description, pages
   - ☐ the claims, Nos.
   - ☐ the drawings, sheets/figs
   - ☐ the sequence listing *(specify)*:

4. ☐ This report has been established as if (some of) the amendments annexed to this report and listed below had not been made, since either they are considered to go beyond the disclosure as filed, or they were not accompanied by a letter indicating the basis for the amendments in the application as filed, as indicated in the Supplemental Box (Rules 70.2(c) and (c-*bis*)):
   - ☐ the description, pages
   - ☐ the claims, Nos.
   - ☐ the drawings, sheets/figs
   - ☐ the sequence listing *(specify)*:

5. ☐ This report has been established:
   - ☐ taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rules 66.1(d-*bis*) and 70.2(e)).
   - ☐ without taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91(Rules 66.4*bis* and 70.2(e)).

6. ☒ With regard to top-up searches (Rules 66.1*ter* and 70.2(f)):
   - ☒ A top-up search was carried out by this Authority on 05.09.2016 (all discovered documents are listed in the Supplemental Box Relating to Top-up Search).
     - ☐ Additional relevant documents have been discovered during the top-up search.
   - ☐ No top-up search was carried out by this Authority because it would serve no useful purpose.

7. ☐ Supplementary international search report(s) from Authority(ies) has/have been received and taken into account in establishing this report (Rule 45bis.8(b) and (c)).

* *If item 4 applies, some or all of those sheets may be marked "superseded".*

Box No. V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-40
	No:	Claims	
Inventive step (IS)	Yes:	Claims	1-32
	No:	Claims	33-40
Industrial applicability (IA)	Yes:	Claims	1-40
	No:	Claims	

2. Citations and explanations (Rule 70.7):

see separate sheet

Box No. VII Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Re Item V

Reference is made to the following documents:

D1 WO 2014/126689 A1 (DOLBY LAB LICENSING CORP [US]) 21 August 2014 (2014-08-21)

D2 HERRE J PRG A[1/4]RGEN ET AL: "MPEG Surround-The ISO/MPEG Standard for Efficient and Compatible Multichannel Audio Coding", JAES, AES, 60 EAST 42ND STREET, ROOM 2520 NEW YORK 10165-2520, USA, vol. 56, no. 11, 1 November 2008 (2008-11-01), pages 932-955, XP040508729,

D3 US 2006/165184 A1 (PURNHAGEN HEIKO [SE] ET AL) 27 July 2006 (2006-07-27)

Presence of inventive step for claims 1-32

1 D2 is regarded as being the prior art closest to the subject-matter of claim 1, and discloses

An audio decoding method comprising: receiving a two-channel downmix signal, which is associated with metadata, the metadata comprising upmix parameters for parametric reconstruction of an M-channel audio signal based on the downmix signal, where M≥4; receiving at least a portion of said metadata; (D2, Fig. 1. The downmix maybe a two channel downmix, first paragraph of section 1, page 933, left column.)

generating a decorrelated signal based on at least one channel of the downmix signal: (D2, sections 3.2.1 and 3.2.2)

determining a set of mixing coefficients based on the received metadata; and forming a K-channel output signal as a linear combination of the downmix signal and the decorrelated signal in accordance with the mixing coefficients, wherein $2 \leq K < M$, (D2, section 3.2.3 where different configurations are possible as e.g. 7.1-5.1-71.)

1.1 The subject-matter of claim 1 therefore differs from this known D2 in that:

wherein the mixing coefficients are determined such that a sum of a mixing coefficient controlling a contribution from the first channel of the downmix signal to a channel of the output signal, and a mixing coefficient controlling a contribution from the first channel of the downmix signal to another channel of the output signal, has the value 1,

wherein, if the downmix signal represents the M-channel audio signal according to a first coding format (F_1) in which:

a first (L_1) channel of the downmix signal corresponds to a certain linear combination of a first group (401) of one or more channels of the M-channel audio signal;

a second channel (L_2) of the downmix signal corresponds to a certain linear combination of a second group (402) of one or more channels of the M-channel audio signal; and

the first and second groups constitute a certain partition of the M channels of the M-channel audio signal,

then the K-channel output signal represents the M-channel audio signal according to a second coding format (F_2, F_4) in which:

each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal;

the groups corresponding to the respective channels of the output signal constitute a partition of the M channels of the M-channel audio signal Into K groups (501-502, 1301-1303) of one or more channels; and

at least two of the K groups comprise at least one channel from said first group.

The subject-matter of claim 1 is therefore new (Article 33(2) PCT)

1.2 The technical effect of the difference is to enhance the quality of the playback on a loud-speaker set-up which is different to the one foreseen on the encoder side. The solution to this problem proposed in claim 1 of the present application is considered as involving an inventive step (Article 33(3) PCT) as the application allows, when using the transmitted up-mix parameters and some computation, to avoid reconstructing the whole audio scene as encoded on the encoder side and then convert to the loudspeaker set-up. Rather by

using the downmix channels and the up-mix parameters together with some computations, channels adapted for the loudspeaker set-up are reconstructed.

2 The same reasoning applies, mutatis mutandis, to the subject-matter of the corresponding independent claim 30 which therefore are also considered not inventive.

3 Claims 2-29 and claims 31-32 are respectively dependent on claim 1 and 30 and as such also meet the requirements of the PCT with respect to novelty and inventive step.

Lack of inventive step for claims 33-40

4 The present application does not meet the criteria of Article 33(1) PCT, because the subject-matter of claims 33 is not inventive in the sense of Article 33(3) PCT.

4.1 D2 discloses:

An audio encoding method comprising: receiving an M-channel audio signal, where M ≥ 4; computing a two-channel downmix signal based on the M-channel audio signal, a first channel of the downmix signal being formed as a linear combination of a first group of one or more channels of the M-channel audio signal, and a second channel of the downmix signal being formed as a linear combination of a second group of one or more channels of the M-channel audio signal, wherein the first and second groups constitute a partition of the M channels of the M-channel audio signal; (D2, Fig. 1 and 4. In the case of e.g. K=3, M= 4 channels which are Fl, Fr, Bl, Br (Front left & right, Back left & right). The skilled person would, without involving an inventive step, downmix the M=4 channels in two signals forming partitions of the M=4 channels with e.g. the following basic formula: F=0.5x(Fl+Fr) and B=0.5x(Bl +Br))

determining upmix parameters for parametric reconstruction of the M-channel audio signal from the downmix signal, determining ~~mixing~~ parameters for obtaining, based on the downmix signal, a K-channel output signal, where 2 < K < M, wherein each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal, wherein these K groups constitute a partition of the M channels of the M-channel audio signal into K groups of one or more channels, and wherein

at least two of the K groups comprise at least one channel from said first group; and (D2, section 1, where it is disclosed that upmix parameters allow to reconstruct the e.g. 4 channels, Fl, Fr, Bl, Br, but also Fl, Fr and B. There would then be 3 groups of channels K1=Fl, K2=Fr and K3=B where two of the K groups of channels comprise at least one channel from the first group.)

outputting the downmix signal and metadata for joint storage or transmission, wherein the metadata comprises the upmix parameters and the mixing parameters. (D2, Fig. 1 where parameters are transmitted on the decoder side)

4.2 D2 is silent about determining and transmitting mixing parameters.

Nevertheless, transmitting extra parameters which are not needed as the effect of claim 33 can be achieved by using the upmix parameters would increase the bitrate without having any additional surprising effect. Such an addition of parameters corresponds to a degradation of the prior-art and does not comprise an inventive step.

5 The same reasoning applies, mutatis mutandis, to the subject-matter of the corresponding independent claims 37-39 which therefore are also considered not inventive.

6 The following dependent claims do not appear to contain any additional features which, in combination with the features of any claim to which they refer, meet the requirements of the PCT in respect of inventive step, the reasons being as follows:

Claim 34, 35: D2, Fig. 4 and 10 where the use of decorrelated for creating multichannel signal is well known. Further in D2, different parameters are transmitted for reconstructing the spatial using ICC parameters is disclosed in section 3.3.

Claim 36: D2 discloses different downmixing methods, see e.g. 4.3, transmitting the parameters adapted for each downmix and indicating which method is used does not involve an inventive step.

Claim 40: see objections related to claim 33, where mixing parameters are not needed for the decoding of the upmix parameters.

Re Item VII

7 The independent claims are not in the two-part form in accordance with Rule 6.3(b) PCT, which in the present case would be appropriate, with those features known in combination from the prior art D2 being placed in the preamble (Rule 6.3(b)(i) PCT) and the remaining features being included in the characterising part (Rule 6.3(b)(ii) PCT).

Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in D1-D3 is not mentioned in the description, nor are these documents identified therein.

8 When entering the regional phase, the applicant is reminded that for some specific region that claim 39 could be objected to as it leaves doubt about the claim's category, since it does not specify whether the program runs or is actually loaded on a computer. Therefore, it should be formulated in such a way so that its category is unambiguously determined.

DOLBY.

Dolby International AB
Patent Group Europe
Apollo Building, 3E
Herikerbergweg 1-35
1101 CN Amsterdam Zuidoost
The Netherlands
T: +31 (0) 20 6511 842
F: +31 (0) 20 6511 890
E: patentseurope@dolby.com

European Patent Office
PB 5818 Patentlaan 2
NL-2280 HV Rijswijk
Netherlands
Fax: +31 (0) 70 340 30 16

Amsterdam,	Our ref.	International Application Number
14 November 2016	D14089WO01	PCT/EP2015/075022

Dear Sirs,

This is in response to the Written Opinion of the IPEA dated 30 September 2016.

1. Amendments

In response to point 8.1 in the Written Opinion, claim 33 has been amended by replacing "wherein the groups corresponding to the respective channels of the output signal constitute a partition..." by "wherein these K groups constitute a partition...". Basis for this amendment may for example be found on page 15, line 1-6, and on page 37, lines 26-34.

Corresponding amendments have been made in claims 37 and 39.

The applicant reserves the right to file one or more divisional patent applications including any of the subject matter disclosed in the patent application as originally filed. The amendment to and/or deletion of the claims or the description should not be interpreted as abandonment of those claims or any subject matter, and is without prejudice to reinstatement of those claims and subject matter or to the filing of a divisional patent application for those claims and subject-matter.

2. Clarity

Point 8.1 of the Written Opinion has been addressed by amendments.

With regard to points 8.2 and 8.3 of the Written Opinion, it is respectfully submitted that claims 33-37 are at least sufficiently clear for a proper assessment of novelty and inventive step to be included in the International Preliminary Report on Patentability.

The feature "determining upmix parameters for parametric reconstructions of the M-channel audio signal from the downmix signal" is not unclear. At the encoder side, both the M-channel audio signal and the two-channel downmix signal are available. The skilled person knows how to determine upmix parameters suitable for parametric reconstruction of an M-channel audio signal from a two-channel downmix signal if both the downmix signal and the M-channel audio signal are available. Forcing the applicant to select a particular way of determining such upmix parameters would cause an undue limitation of the claim scope. Claim 35 is referred to in the Written Opinion as an alleged indication that special constraints are applied to the upmix parameters. The applicant respectfully disagrees. The upmix parameters may for example be the actual matrix elements to be employed when reconstructing the M-channel audio signal at the decoder side (such as the matrix elements c_i, $1 \leq i \leq 3$, and p_{ij}, $1 \leq i \leq 3$, $1 \leq j \leq 2$ on page 28, line 33 to page 29, line 5), or the upmix parameters may be a more compact form of representing such matrix elements (see page 38, lines 1-6). Claim 35 corresponds to such a compact representation. In claim 33, the metadata includes both the upmix parameters and the mixing parameters. There is no need for the upmix parameters to be in a certain format for the mixing parameters to be employed at the decoder side for obtaining the K-channel output signal.

Claim 34 specifies that "at least some of the mixing parameters are determined by minimizing a contribution from the decorrelated signal among such mixing parameters that cause a covariance matrix of the K channels of the output signal to coincide with a covariance matrix of the linear combinations of the one or more channels of the respective K groups of one or more channels", i.e. claim 34 specifies how mixing parameters are determined. A detailed example of how the mixing parameters may be computed is described on page 32, line 17, to page 33, line 22.

With regard to point 8.4 of the Written Opinion, it is respectfully submitted that the data represented by the computer-readable medium of claim 39 allows for parametric

reconstruction of an M-channel audio signal, as well as for provision of a K-channel output signal, where $2 \leq K < M$. This technical effect is achieved by the claimed computer-readable medium, even if the actual decoding or reconstruction may be performed by a decoder.

3. Patentability of amended claims 33, 37 and 39

In D2, a downmix signal is obtained at the encoder side by downmixing a multichannel audio signal, and spatial parameters in the form of ICC and CLD parameters are determined at the encoder side for reconstruction of the multichannel audio signal at the decoder side (see for example sections 3.2.1 and 3.2.2 in D2). No additional parameters are determined (and outputted for transmission to the decoder side) for obtaining a K-channel audio signal based on the downmix signal. In particular, no additional parameters are determined for obtaining a K-channel audio signal corresponding to a different partition of the channels of the multichannel signal. Neither section 3.2.1 nor section 4.1 in D2 (referred to in the Written Opinion) discloses determination of such parameters in addition to the ICC and CLD parameters already employed for reconstructing the multichannel audio signal.

D2 discloses use of a stereo downmix (Fig. 6, 9). The stereo downmix in D2 does not correspond to a partition of the set of original audio channels since the channels C and LFE contribute to both of the two downmix channels. Thus, claim 33 is novel over the stereo configuration disclosed in D2, because the downmix channels do not correspond to a partition of the set of original audio channels.

D2 also discloses use of a mono downmix (Fig. 4, 8). In the Written Opinion, two intermediate channels, formed in Fig. 4 in D2 after the two first layers of downmixing are read onto the two-channel downmix signal of claim 33. These two intermediate channels in D2 are combinations/mixes of the channels R, L, C, LFE and the channels Ls, Rs respectively, and are intended for further downmixing to obtain a mono downmix M_0 provided as output. Thus, the two intermediate channels are not output in the Fig. 4 configuration, but are rather downmixed into a mono signal, which is subsequently output along with the spatial parameters. However, claim 33 requires that the downmix signal (e.g., the two-channel downmix signal) be output with the upmix parameters and the mixing parameters. Thus, claim 33 is novel over the mono configuration disclosed in D2 because the mono configuration fails to output a two-channel downmix signal as required by claim 33.

While the novelty objection raised in the Written Opinion for claim 33 is based on the mono downmix embodiment from Fig. 4 in D2, and not on the stereo downmix embodiment from Fig. 6 in D2, in view of the above it is clear that claim 33 is novel over both the Fig. 4 and Fig. 6 embodiments of D2.

Section 4.1 in D2 (referred to in the Written Opinion) describes that a matrix-surround compatible stereo downmix signal may be employed, which may be obtained at the encoder side by post-processing the stereo downmix signal. It is clear from the context that the stereo downmix referred to in Section 4.1 in D2 is the stereo downmix L_0 R_0 obtained in Fig. 6 in D2, and not the two intermediate signals from Fig. 4 in D2 referred to in the novelty objection raised for claim 33 in the Written Opinion. In other words, the novelty objection in the Written Opinion is based on two different embodiments in D2 which are not compatible.

Further, it is explicitly described in section 4.1. in D2 that "the matrix-surround compatibility comes without any additional spatial information". In the Written Opinion, spatial parameters in D2 (i.e. the ICC and CLD parameters) seem to be read onto both the upmix parameters and the mixing parameters of claim 33. Even in the event that the IPEA may maintain some of its clarity objections, it is respectfully submitted that such a reading of D2 onto claim 33 is still not correct, as claim 33 requires the mixing parameters and the upmix parameters to be different sets of parameters, both included in the metadata provided as output.

Moreover, D2 does not disclose that the matrix-surround compatible downmix signal in section 4.1 in D2 may correspond to a partition of the multichannel audio signal.

Note also that there is no indication in D2 that the artistic downmix described in Section 4.3 in D2 may correspond to a partition of the multichannel audio signal. Further, the artistic downmix in Section 4.3 is employed in connection with another stereo downmix, i.e. the stereo downmix L_0 R_0 obtained in Fig. 6 in D2, and not the two intermediate signals from Fig. 4 in D2 referred to in the novelty objection raised for claim 33 in the Written Opinion.

Hence, D2 at least fails to disclose the following features of claim 33:

> "determining (340) mixing parameters for obtaining, based on the downmix signal, a K-channel output signal $(\widetilde{L_1}, ..., \widetilde{L_K})$, where $2 \leq K < M$, wherein each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal, wherein these K groups constitute a partition of the M channels of the M-channel audio signal into K groups (501–502, 1301–1303) of one or

more channels, and wherein at least two of the K groups comprise at least one channel from said first group; and

outputting (350) the downmix signal (which has two channels corresponding to a partition of the M channels of the M-channel audio signal) and metadata for joint storage or transmission, wherein the metadata comprises the upmix parameters and the mixing parameters."

As described in the application as filed (see for example page 15, lines 8-33) a technical effect of the above described distinguishing features is that the perceived playback quality at different speaker systems may be improved in a computationally efficient way. The skilled person seeking to improve the perceived playback quality at different speaker systems in a computationally efficient way would not find any teaching in D1-D3 leading to a solution within the terms of claim 33. The subject-matter of claim 33 is therefore both novel and inventive.

Claim 37 includes corresponding distinguishing features. The subject-matter of claim 37 is therefore also novel and inventive.

Claim 39 also includes corresponding distinguishing features. As described above in relation to clarity, such distinguishing features have a technical effect also for claim 39. The subject-matter of claim 39 is therefore also novel and inventive.

4. Additional distinguishing feature in dependent claim 34

Section 3.3 and Figures 4 and 10 in D2, referred to in the Written Opinion, disclose use of decorrelation and ICC parameters for obtaining a multichannel signal from a two-channel signal. However, this teaching in D2 does not in any way provide any hints towards the following feature of claim 34:

"wherein at least some of the mixing parameters are determined by minimizing a contribution from the decorrelated signal among such mixing parameters that cause a covariance matrix of the K channels of the output signal to coincide with a covariance matrix of the linear combinations of the one or more channels of the respective K groups of one or more channels."

The skilled person would find no teaching or hints in D1-D3 towards such a solution. As described above in relation to clarity, a detailed example of how the mixing parameters may be computed is described on page 32, line 17, to page 33, line 22 in the application as filed.

5. Additional distinguishing feature in dependent claim 36

Section 4.3 in D2 discloses use of a different type of downmix, referred to as artistic downmix. However, as described above, the artistic downmix is employed in connection with a spatial downmix which does not correspond to a partition of the original multichannel audio signal (see for example Fig. 6 in D2). Further, D2 fails to disclose that the artistic downmix may correspond to a different partition of the original multichannel audio signal than the spatial downmix. In particular, D2 neither discloses nor provides any hints towards the following feature of claim 36:

> "selecting one of at least two coding formats (F_1, F_2, F_3), the coding formats corresponding to respective different partitions of the channels of the M-channel audio signal into respective first and second groups associated with the channels of the downmix signal"

The skilled person would find no teaching or hints in D1-D3 towards such a solution.

6. Requests

It is believed that a favourable IPRP may be issued on the basis of the application documents presented herewith in light of the foregoing comments.

Nevertheless, we kindly request that any remaining objections be set out in a further Written Opinion, or, failing that, we request a **consultation by telephone**.

Yours sincerely,

Siegfried Soellner

For and on behalf of
Dolby International AB Patent Group Europe
(Association No. 435)

CLAIMS

1. An audio decoding method (1000) comprising:

receiving (1010) a two-channel downmix signal (L_1, L_2), which is associated with metadata, the metadata comprising upmix parameters (α_{LU}) for parametric reconstruction of an M-channel audio signal (L, LS, LB, TFL, TBL) based on the downmix signal, where $M \geq 4$;

receiving (1020) at least a portion of said metadata;

generating (1040) a decorrelated signal (D) based on at least one channel of the downmix signal;

determining (1050) a set of mixing coefficients based on the received metadata; and

forming (1060) a K-channel output signal ($\widetilde{L_1}$, ..., $\widetilde{L_K}$) as a linear combination of the downmix signal and the decorrelated signal in accordance with the mixing coefficients, wherein $2 \leq K < M$,

wherein the mixing coefficients are determined such that a sum of a mixing coefficient controlling a contribution from the first channel of the downmix signal to a channel of the output signal, and a mixing coefficient controlling a contribution from the first channel of the downmix signal to another channel of the output signal, has the value 1,

wherein, if the downmix signal represents the M-channel audio signal according to a first coding format (F_1) in which:

a first (L_1) channel of the downmix signal corresponds to a certain linear combination of a first group (401) of one or more channels of the M-channel audio signal;

a second channel (L_2) of the downmix signal corresponds to a certain linear combination of a second group (402) of one or more channels of the M-channel audio signal; and

the first and second groups constitute a certain partition of the M channels of the M-channel audio signal,

then the K-channel output signal represents the M-channel audio signal according to a second coding format (F_2, F_4) in which:

each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal;

the groups corresponding to the respective channels of the output signal constitute a partition of the M channels of the M-channel audio signal into K groups (501–502, 1301–1303) of one or more channels; and

at least two of the K groups comprise at least one channel from said first group.

2. The audio decoding method of claim 1, wherein $K = 2$.

3. The audio decoding method of claim 1, wherein $K = 3$ or $K = 4$.

4. The audio decoding method of any of the preceding claims, wherein the received metadata includes the upmix parameters and wherein the mixing coefficients are determined by processing the upmix parameters.

5. The audio decoding method of any of claims 1–3, wherein the received metadata includes mixing parameters (α_{LM}) distinct from the upmix parameters.

6. The audio decoding method of claim 5, wherein the mixing coefficients are determined independently of any values of the upmix parameters.

7. The audio decoding method of any of the preceding claims, wherein $M = 5$ or $M = 6$.

8. The audio decoding method of any of the preceding claims, wherein:

in the first coding format, each of the channels of the M-channel audio signal is associated with a non-zero gain controlling a contribution from this channel to one of the linear combinations to which the channels of the downmix signal correspond;

in the second coding format, each of the channels of the M-channel audio signal is associated with a non-zero gain controlling a contribution from this channel to one of the linear combinations approximated by the channels of the output signal; and

for each of the channels of the M-channel audio signal, the non-zero gain associated with the channel in the first coding format coincides with the non-zero gain associated with the channel in the second coding format.

9. The audio decoding method of any of the preceding claims, further comprising an initial step of receiving a bitstream (B) representing the downmix signal and the metadata,

wherein the downmix signal and said received metadata are extracted from the bitstream.

10. The audio decoding method of any of the preceding claims, wherein the decorrelated signal is a single-channel signal and wherein said output signal is formed by including no more than one decorrelated signal channel into said linear combination of the downmix signal and the decorrelated signal.

11. The audio decoding method of claim 10, wherein $K = 2$, and wherein forming the output signal amounts to a projection from three channels to two channels.

12. The audio decoding method of any of claims 1–9, wherein the decorrelated signal is a two-channel signal, and wherein said output signal is formed by including no more than two decorrelated signal channels into said linear combination of the downmix signal and the decorrelated signal.

13. The audio decoding method of claim 12, wherein $K = 3$, and wherein forming the output signal amounts to a projection from four channels to three channels.

14. The audio decoding method of any of the preceding claims, wherein the mixing coefficients are determined such that a pair of channels of the output signal receive contributions of equal magnitude from a channel of the decorrelated signal, the contributions from the channel of the decorrelated signal to the respective channel of the pair having opposite signs.

15. The audio decoding method of any of the preceding claims, wherein said first group consists of two or three channels.

16. The audio decoding method of any of the preceding claims, wherein the M-channel audio signal comprises three channels (L, LS, LB) representing different horizontal directions in a playback environment for the M-channel audio signal, and two channels (TFL, TBL) representing directions vertically separated from those of said three channels in said playback environment.

17. The audio decoding method of claim 16, wherein said first group consists of said three channels, and wherein said second group consists of the two channels representing directions vertically separated from those of said three channels in said playback environment.

18. The audio decoding method of claim 17, wherein $K = 2$, and wherein one of the K groups comprises both of the two channels representing directions vertically separated from those of said three channels in said playback environment.

19. The audio decoding method of claim 17, wherein the two channels representing directions vertically separated from those of said three channels in said playback environment are comprised in different groups of the K groups.

20. The audio decoding method of any of claims 1–16, wherein the M-channel audio signal comprises four channels ($LSCRN, LW, LS, LB$) representing different horizontal directions in a playback environment for the M-channel audio signal, and two channels (TFL, TBL) representing directions vertically separated from those of said four channels in said playback environment.

21. The audio decoding method of any of claims 16–17 and 19–20, wherein one of the K groups comprises both of the two channels representing directions vertically separated from those of said three channels in said playback environment.

22. The audio decoding method of any of the preceding claims, wherein the decorrelated signal is obtained by processing a linear combination of the channels of the downmix signal.

23. The audio decoding method of any of claims 1–21, wherein the decorrelated signal is obtained based on no more than one channel of the downmix signal.

24. The audio decoding method of any of claims 1–21, wherein the decorrelated signal comprises two channels, a first channel of the decorrelated signal being obtained based on the first channel of the downmix signal and a second channel of the decorrelated signal being obtained based on the second channel of the downmix signal.

25. The audio decoding method of any of claims 1–4 and 7–24, wherein said first group consists of N channels, where $N \geq 3$, wherein said first group is reconstructable as a linear combination of said first channel of the downmix signal and an $(N-1)$-channel decorrelated signal by applying dry upmix coefficients to said first channel of the downmix signal and wet upmix coefficients to channels of the $(N-1)$-channel decorrelated signal, wherein the received metadata includes wet upmix parameters and dry upmix parameters, and wherein determining the mixing coefficients comprises:

determining, based on the dry upmix parameters, the dry upmix coefficients;

populating an intermediate matrix having more elements than the number of received wet upmix parameters, based on the received wet upmix parameters and knowing that the intermediate matrix belongs to a predefined matrix class;

obtaining the wet upmix coefficients by multiplying the intermediate matrix by a predefined matrix, wherein the wet upmix coefficients corresponds to the matrix resulting from the multiplication and includes more coefficients than the number of elements in the intermediate matrix; and

processing the wet and dry upmix coefficients.

26. The audio decoding method of any of the preceding claims, further comprising:

receiving signaling (1030) indicating one of at least two coding formats (F_1, F_2, F_3) of the M-channel audio signal, the coding formats corresponding to respective different partitions of the channels of the M-channel audio signal into respective first and second groups associated with the channels of the downmix signal,

wherein the K groups are predefined, and wherein the mixing coefficients are determined such that a single partition of the M-channel audio signal into the K groups of channels, approximated by the channels of the output signal, is maintained for said at least two coding formats.

27. The audio decoding method of claim 26, wherein $K = 2$, the audio decoding method further comprising:

passing (1070) the downmix signal through as said output signal, in response to said signaling indicating a particular coding format (F_2), the particular coding format corresponding to a partition of the channels of the M-channel audio signal coinciding with a partition which the K groups define.

28. The audio decoding method of claim 26, wherein $K = 2$, the audio decoding method further comprising:

suppressing the contribution from the decorrelated signal to said output signal, in response to said signaling indicating a particular coding format, the particular coding format corresponding to a partition of the channels of the M-channel audio signal coinciding with a partition which the K groups define.

29. The audio decoding method of any of claims 26–28, wherein:

in a first coding format (F_1) of said at least two coding formats, said first group consists of three channels (L, LS, LB) representing different horizontal directions in a playback environment for the M-channel audio signal, and said second group consists of two channels (TFL, TBL) representing directions vertically separated from those of said three channels in said playback environment; and

in a second coding format (F_2) of said at least two coding formats, each of said first and second groups comprises one of said two channels representing directions vertically separated from those of said three channels in said playback environment.

30. An audio decoding system (800) comprising a decoding section (700, 1200) configured to:

receive a two-channel downmix signal (L_1, L_2), which is associated with metadata, the metadata comprising upmix parameters (α_{LU}) for parametric reconstruction of an M-channel audio signal (L, LS, LB, TFL, TBL) based on the downmix signal, where $M \geq 4$;

receive at least a portion of said metadata; and

provide a K-channel output signal ($\widetilde{L_1}$, ..., $\widetilde{L_K}$) based on the downmix signal and the received metadata, wherein $2 \leq K < M$,

the decoding section comprising:

a decorrelating section (710, 1210) configured to receive at least one channel of the downmix signal and to output, based thereon, a decorrelated signal (D); and

a mixing section (720, 1220) configured to

determine a set of mixing coefficients based on the received metadata, and

form the output signal as a linear combination of the downmix signal and the decorrelated signal in accordance with the mixing coefficients,

wherein the mixing section is configured to determine the mixing coefficients such that a sum of a mixing coefficient controlling a contribution from the first channel of the downmix signal to a channel of the output signal, and a mixing coefficient controlling a contribution from the first channel of the downmix signal to another channel of the output signal, has the value 1,

wherein, if the downmix signal represents the M-channel audio signal according to a first coding format (F_1) in which:

a first (L_1) channel of the downmix signal corresponds to a certain linear combination of a first group (401) of one or more channels of the M-channel audio signal;

a second channel (L_2) of the downmix signal corresponds to a certain linear combination of a second group (402) of one or more channels of the M-channel audio signal; and

the first and second groups constitute a certain partition of the M channels of the M-channel audio signal,

then the K-channel output signal represents the M-channel audio signal according to a second coding format (F_2, F_4) in which:

each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal;

the groups corresponding to the respective channels of the output signal constitute a partition of the M channels of the M-channel audio signal into K groups (501–502, 1301–1303) of one or more channels; and

at least two of the K groups comprise at least one channel from said first group.

31. The audio decoding system of claim 30, further comprising an additional decoding section (805) configured to:

receive an additional two-channel downmix signal (R_1, R_2), which is associated with additional metadata, the additional metadata comprising additional upmix parameters (α_{RU}) for parametric reconstruction of an additional M-channel audio signal (R, RS, RB, TFR, TBR) based on the additional downmix signal,

receive at least a portion of the additional metadata; and

provide an additional K-channel output signal $(\widetilde{R_1}, ..., \widetilde{R_K})$ based on the additional downmix signal and the additional received metadata,

the additional decoding section comprising:

an additional decorrelating section configured to receive at least one channel of the additional downmix signal and to output, based thereon, an additional decorrelated signal; and

an additional mixing section configured to:

determine a set of additional mixing coefficients based on the received additional metadata, and

form the additional output signal as a linear combination of the additional downmix signal and the additional decorrelated signal in accordance with the additional mixing coefficients,

wherein the additional mixing section is configured to determine the additional mixing coefficients such that a sum of a mixing coefficient controlling a contribution from the first channel of the additional downmix signal to a channel of the additional output signal, and a mixing coefficient controlling a contribution from the first channel of the additional downmix signal to another channel of the additional output signal, has the value 1,

wherein, if the additional downmix signal represents the additional M-channel audio signal according to a third coding format in which:

a first channel (R_1) of the additional downmix signal corresponds to a linear combination of a first group (403) of one or more channels of the additional M-channel audio signal;

a second channel (R_2) of the additional downmix signal corresponds to a linear combination of a second group (404) of one or more channels of the additional M-channel audio signal; and

the first and second groups of channels of the additional M-channel audio signal constitute a partition of the M channels of the additional M-channel audio signal,

then the additional K-channel output signal represents the additional M-channel audio signal according to a fourth coding format in which:

each of the K channels of the additional output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal;

the groups corresponding to the respective channels of the additional output signal constitute a partition of the M channels of the additional M-channel audio signal into K groups (503–504, 1304–1306) of one or more channels; and

at least two of the K groups of one or more channels of the additional M-channel audio signal comprise at least one channel from said first group of channels of the additional M-channel audio signal.

32. The decoding system of any of claims 30–31, further comprising:

a demultiplexer (801) configured to extract, from a bitstream (B), the downmix signal, said received metadata, and a discretely coded audio channel (C); and

a single-channel decoding section operable to decode said discretely coded audio channel.

33. An audio encoding method (300) comprising:

receiving (310) an M-channel audio signal (L, LS, LB, TFL, TBL), where $M \geq 4$;

computing (320) a two-channel downmix signal (L_1, L_2) based on the M-channel audio signal, a first channel (L_1) of the downmix signal being formed as a linear combination of a first group (401) of one or more channels of the M-channel audio signal, and a second channel (L_2) of the downmix signal being formed as a linear combination of a second group (402) of one or more channels of the M-channel audio signal, wherein the first and second groups constitute a partition of the M channels of the M-channel audio signal;

determining (330) upmix parameters (α_{LU}) for parametric reconstruction of the M-channel audio signal from the downmix signal,

determining (340) mixing parameters for obtaining, based on the downmix signal, a K-channel output signal ($\widetilde{L_1}, \ldots, \widetilde{L_K}$), where $2 \leq K < M$, wherein each of the K channels of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal, wherein these K groups constitute a partition of the M channels of the M-channel audio signal into K groups (501–502, 1301–1303) of one or more channels, and wherein at least two of the K groups comprise at least one channel from said first group; and

outputting (350) the downmix signal and metadata for joint storage or transmission, wherein the metadata comprises the upmix parameters and the mixing parameters.

34. The audio encoding method of claim 33, wherein the mixing parameters control respective contributions from the downmix signal and from a decorrelated signal to the output signal, wherein at least some of the mixing parameters are determined by minimizing a contribution from the decorrelated signal among such mixing parameters that cause a covaraince matrix of the K channels of the output signal to coincide with a covariance matrix of the linear combinations of the one or more channels of the respective K groups of one or more channels.

35. The audio encoding method of any of claims 33–34, wherein said first group consists of N channels, where $N \geq 3$, wherein at least some of the upmix parameters are suitable for parametric reconstruction of said first group from said first channel of the downmix signal and an $(N-1)$-channel decorrelated signal determined based on said first channel of the downmix signal, wherein determining the upmix parameters includes:

determining a set of dry upmix coefficients in order to define a linear mapping of said first channel of the downmix signal approximating said first group; and

determining an intermediate matrix based on a difference between a covariance of said first group as received and a covariance of said first group as approximated by the linear mapping of said first channel of the downmix signal, wherein the intermediate matrix when multiplied by a predefined matrix corresponds to a set of wet upmix coefficients defining a linear mapping of said decorrelated signal as part of parametric reconstruction of said first group, wherein the set of wet upmix coefficients includes more coefficients than the number of elements in the intermediate matrix,

wherein said upmix parameters include dry upmix parameters, from which the set of dry upmix coefficients is derivable, and wet upmix parameters uniquely defining the intermediate matrix provided that the intermediate matrix belongs to a predefined matrix class, wherein the intermediate matrix has more elements than the number of said wet upmix parameters.

36. The audio encoding method of any of claims 33–35, further comprising:

selecting one of at least two coding formats (F_1, F_2, F_3), the coding formats corresponding to respective different partitions of the channels of the M-channel audio signal into respective first and second groups associated with the channels of the downmix signal,

wherein the first and second channels of the downmix signal are formed as linear combinations of a first and a second group of one or more channels, respectively, of the M-channel audio signal, in accordance with the selected coding format, and wherein the upmix parameters and the mixing parameters are determined based on the selected coding format,

the method further comprising:

providing signaling indicating the selected coding format.

37. An audio encoding system (200) comprising an encoding section (100) configured to encode an M-channel audio signal (L, LS, LB, TFL, TBL) as a two-channel downmix signal (L_1, L_2) and associated metadata, where $M \geq 4$, and to output the downmix signal and metadata for joint storage or transmission, the encoding section comprising:

a downmix section (110) configured to compute the downmix signal based on the M-channel audio signal, a first channel (L_1) of the downmix signal being formed as a linear combination of a first group (401) of one or more channels of the M-channel audio signal, and a second channel (L_2) of the downmix signal being formed as a linear combination of a second group (402) of one or more channels of the M-channel audio signal, wherein the first and second groups constitute a partition of the M channels of the M-channel audio signal; and

an analysis section (120) configured to determine

upmix parameters (α_{LU}) for parametric reconstruction of the M-channel audio signal from the downmix signal, and

mixing parameters (α_{LM}) for obtaining, based on the downmix signal, a K-channel output signal ($\widetilde{L_1}, \ldots, \widetilde{L_K}$), wherein $2 \leq K < M$, wherein each channel of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal, wherein these K groups constitute a partition of the M channels of the M-channel au-

dio signal into K groups (501–502, 1301–1303) of one or more channels, and wherein at least two of the K groups comprise at least one channel from said first group,

wherein the metadata comprises the upmix parameters and the mixing parameters.

38. A computer program product comprising a computer-readable medium with instructions for performing the method of any of claims 1–29 and 33-36.

39. A computer-readable medium (1100) representing:

a two-channel downmix signal (L_1, L_2);

upmix parameters (α_{LU}) allowing parametric reconstruction of an M-channel audio signal (L, LS, LB, TFL, TBL) based on the downmix signal, where $M \geq 4$, wherein a first channel (L_1) of the downmix signal corresponds to a linear combination of a first group (401) of one or more channels of the M-channel audio signal, wherein a second channel (L_2) of the downmix signal corresponds to a linear combination of a second group (402) of one or more channels of the M-channel audio signal, and wherein the first and second groups constitute a partition of the M channels of the M-channel audio signal; and

mixing parameters (α_{LM}) allowing provision of a K-channel output signal $(\widetilde{L_1}, ..., \widetilde{L_2})$ based on the downmix signal, wherein $2 \leq K < M$, wherein each channel of the output signal approximates a linear combination of a group of one or more channels of the M-channel audio signal, wherein these K groups constitute a partition of the M channels of the M-channel audio signal into K groups (501–502, 1301–1303) of one or more channels, and wherein at least two of the K groups comprise at least one channel from said first group.

40. The computer-readable medium of claim 39, wherein data represented by the computer-readable medium are arranged in time frames and are layered such that, for a given time frame, the downmix signal and associated mixing parameters for that time frame may be extracted independently of the associated upmix parameters.