



(12) Translation of
European patent specification

(11) NO/EP 2981958 B1

NORWAY

(19) NO
(51) Int Cl.
G10L 19/02 (2013.01)
G10L 19/032 (2013.01)
G10L 19/06 (2013.01)

Norwegian Industrial Property Office

(21)	Translation Published	2018.07.16
(80)	Date of The European Patent Office Publication of the Granted Patent	2018.03.07
(86)	European Application Nr.	14715307.6
(86)	European Filing Date	2014.04.04
(87)	The European Application's Publication Date	2016.02.10
(30)	Priority	2013.04.05, US, 201361808675 P 2013.09.09, US, 201361875553 P
(84)	Designated Contracting States:	AL ; AT ; BE ; BG ; CH ; CY ; CZ ; DE ; DK ; EE ; ES ; FI ; FR ; GB ; GR ; HR ; HU ; IE ; IS ; IT ; LI ; LT ; LU ; LV ; MC ; MK ; MT ; NL ; NO ; PL ; PT ; RO ; RS ; SE ; SI ; SK ; SM ; TR
(73)	Proprietor	Dolby International AB, Apollo Building, 3E Herikerbergweg 1-35, 1101 CN Amsterdam, NL-Nederland
(72)	Inventor	VILLEMOES, Lars, Dolby Sweden ABGävlegatan 12A, 113 30 Stockholm, SE-Sverige KLEJSA, Janusz, Dolby Sweden ABGävlegatan 12A, 113 30 Stockholm, SE-Sverige HEDELIN, Per, Dolby Sweden ABGävlegatan 12A, 113 30 Stockholm, SE-Sverige
(74)	Agent or Attorney	TANDBERG INNOVATION AS, Postboks 1570 Vika, 0118 OSLO, Norge

(54)	Title	AUDIO ENCODER AND DECODER
(56)	References Cited:	EP-A2- 0 673 014, WO-A1-2012/146757, Grant A. Davidson ET AL: "Digital Audio Coding: Dolby AC-3" In: "The Digital Signal Processing Handbook", 1 January 1999 (1999-01-01), CRC Press LLC - IEEE Press, XP055140739, ISBN: 978-0-84-938572-8 pages 41-1, section 41.2; page 6 figure 41.4 section 41.4; page 10 - page 13 page 11, line 4 - line 7 section 41.3.2; page 9 section "Frequency Banding"; page 18, HERRE J ET AL: "EXTENDING THE MPEG-4 AAC CODEC BY PERCEPTUAL NOISE SUBSTITUTION", PREPRINTS OF PAPERS PRESENTED AT THE AES CONVENTION, XX, XX, 1 January 1998 (1998-01-01), pages 1-14, XP008006769,, SCHUYLER QUACKENBUSH: "MPEG Unified Speech and Audio Coding", IEEE MULTIMEDIA, IEEE SERVICE CENTER, NEW YORK, NY, US, vol. 20, no. 2, 1 April 2013 (2013-04-01), pages 72-78, XP011515217, ISSN: 1070-986X, DOI: 10.1109/MMUL.2013.24, BRANDENBURG K: "MP3 AND AAC EXPLAINED", 15TH INTERNATIONAL CONFERENCE: AUDIO, ACOUSTICS &

SMALL SPACES, AUDIO ENGINEERING SOCIETY, US, 1 January 1999 (1999-01-01), pages 99-110, XP008004053,, RAPPORTEUR Q6/16: "Draft revised Technical Paper HSTP-MCTB (ex HSTP-MCTA (V2)) Media coding toolbox for IPTV: Audio and video codecs", 38. VCEG MEETING; 89. MPEG MEETING; 1-7-2009 - 8-7-2009; LONDON, GENEVA; (VIDEO CODING EXPERTS GROUP OF ITU-T SG.16),, no. VCEG-AL95, 11 July 2009 (2009-07-11), XP030003740, ISSN: 0000-0076

Enclosed is a translation of the patent claims in Norwegian. Please note that as per the Norwegian Patents Acts, section 66i the patent will receive protection in Norway only as far as there is agreement between the translation and the language of the application/patent granted at the EPO. In matters concerning the validity of the patent, language of the application/patent granted at the EPO will be used as the basis for the decision. The patent documents published by the EPO are available through Espacenet (<http://worldwide.espacenet.com>) or via the search engine on our website here: <https://search.patentstyret.no/>

Patentkrav

1. Transformasjonsbasert talekoder (100, 170) innrettet for å kode et talesignal inn
5 i en bitstrøm; koderen (100, 170) omfattende
 - en rammebehandlingsenhet (101) innrettet for å motta et sett (132, 332) av blokker; hvor settet (132, 332) av blokker omfatter et flertall sekvensielle blokker (131) av transformasjonskoeffisienter; hvor flertallet blokker (131) angir punktprøver av talesignalet; hvor en blokk (131) av transformasjonskoeffisienter omfatter et flertall
10 transformasjonskoeffisienter for et motsvarende flertall frekvensblokker (301);
 - en omhyllingskurveestimeringsenhet (102) innrettet for å bestemme en aktuell omhyllingskurve (133) basert på flertallet sekvensielle blokker (131) av transformasjonskoeffisienter; hvor den aktuelle omhyllingskurven (133) angir et flertall spektralenergiverdier (303) for det motsvarende flertallet frekvensblokker
15 (301);
 - en omhyllingskurvekvantiseringsenhet (103) innrettet for å bestemme en kvantisert aktuell omhyllingskurve (134) ved å kvantisere den aktuelle omhyllingskurven (133);
 - en omhyllingskurveinterpolasjonsenhet (104) innrettet for å bestemme et
20 flertall interpolerte omhyllingskurver (136) henholdsvis for flertallet blokker (131) av transformasjonskoeffisienter, basert på den kvantiserte aktuelle omhyllingskurven (134) og basert på en kvantisert tidligere omhyllingskurve (134); og
 - en avflatingsenhet (108) innrettet for å bestemme et flertall blokker (140) av avflatede transformasjonskoeffisienter ved å avflate det motsvarende flertallet blokker
25 (131) av transformasjonskoeffisienter med bruk henholdsvis av det motsvarende flertallet interpolerte omhyllingskurver (136); hvor bitstrømmen blir bestemt basert på flertallet blokker (140) av avflatede transformasjonskoeffisienter.
2. Transformasjonsbasert talekoder (100, 170) ifølge krav 1, hvor
30 - den transformasjonsbaserte talekoderen (100, 170) videre omfatter en omhyllingskurveforsterkning-bestemmelsesenhet (105, 106) innrettet for å bestemme

et flertall omhyllingskurveforsterkninger (137) henholdsvis for flertallet blokker (131) av transformasjonskoeffisienter;

- den transformasjonsbaserte talekoderen (100, 170) videre omfatter en omhyllingskurveforfiningsenhet (107) innrettet for å bestemme et flertall tilpassede omhyllingskurver (139) ved å forskyve spektralenergiverdier (303) til flertallet interpolerte omhyllingskurver (136) i samsvar henholdsvis med flertallet omhyllingskurveforsterkninger (137);

- avflatingsenheten (108) er innrettet for å bestemme flertallet blokker (140) av avflatede transformasjonskoeffisienter ved å avflate det motsvarende flertallet blokker (131) av transformasjonskoeffisienter med bruk henholdsvis av det motsvarende flertallet tilpassede omhyllingskurver (139).

3. Transformasjonsbasert talekoder (100, 170) ifølge krav 2, hvor omhyllingskurveforsterkning-bestemmelsesenheden (105, 106) er innrettet for å bestemme en første omhyllingskurveforsterkning (137) for en første blokk (131) av transformasjonskoeffisienter slik at en varians til de avflatede transformasjonskoeffisientene i en motsvarende første blokk (140) av avflatede transformasjonskoeffisienter avledet med bruk av en første tilpasset omhyllingskurve (139) blir justert sammenliknet med en varians til de avflatede transformasjonskoeffisientene i en motsvarende første blokk (140) av avflatede transformasjonskoeffisienter avledet med bruk av en første interpolert omhyllingskurve (139), hvor eventuelt omhyllingskurveforsterkning-bestemmelsesenheden (105, 106) er innrettet for å bestemme den første omhyllingskurveforsterkningen (137) for den første blokken (131) av transformasjonskoeffisienter slik at variansen til de avflatede transformasjonskoeffisientene i den motsvarende første blokken (140) av avflatede transformasjonskoeffisienter avledet med bruk av den første tilpassede omhyllingskurven (139) er én.

4. Transformasjonsbasert talekoder (100, 170) ifølge ethvert foregående krav, hvor

- den aktuelle omhyllingskurven (133) angir et flertall spektralenergiverdier (303) for et motsvarende flertall frekvensbånd (302);

- et frekvensbånd (302) omfatter én eller flere frekvensblokker (301);

- omhyllingskurveestimeringsenheten (102) er innrettet for å bestemme

5 spektralenergiverdien (303) for et gitt frekvensbånd (302) basert på transformasjonskoeffisientene i flertallet sekvensielle blokker (131) for det gitte frekvensbåndet (302).

5. Transformasjonsbasert talekoder (100, 170) ifølge ethvert foregående krav, hvor

10 - en blokk (131) av transformasjonskoeffisienter omfatter MDCT-koeffisienter; og/eller

- en blokk (131) av transformasjonskoeffisienter omfatter 256 transformasjonskoeffisienter i 256 frekvensblokker (301); og/eller

15 - et sett (132, 332) av blokker omfatter fire eller flere blokker (131) av transformasjonskoeffisienter.

6. Transformasjonsbasert talekoder (100, 170) ifølge ethvert foregående krav, hvor

- den transformasjonsbaserte talekoderen (100, 170) er innrettet for å kjøre i et flertall forskjellige modi, omfattende en kortskrittmodus og en langskrittmodus;

20 - rammebehandlingsenheten (101), omhyllingskurveestimeringsenheten (102) og omhyllingskurveinterpolasjonsenheten (104) er innrettet for å behandle settet (132, 332) av blokker omfattende flertallet sekvensielle blokker (131) av transformasjonskoeffisienter når den transformasjonsbaserte talekoderen (100, 170) blir kjørt i kortskrittmodus; og

25 - rammebehandlingsenheten (101), omhyllingskurveestimeringsenheten (102) og omhyllingskurveinterpolasjonsenheten (104) er innrettet for å behandle et sett (132, 332) av blokker omfattende én enkelt blokk (131) av transformasjonskoeffisienter når den transformasjonsbaserte talekoderen (100, 170) blir kjørt i langskrittmodus, hvor eventuelt, i langskrittmodus,

- omhyllingskurveestimeringsenheten (102) er innrettet for å bestemme en aktuell omhyllingskurve (133) for den ene blokken (131) av transformasjonskoeffisienter omfattet i settet (132, 133) av blokker; og

5 - omhyllingskurveinterpolasjonsenheten (104) er innrettet for å bestemme en interpolert omhyllingskurve (136) for den ene blokken (131) av transformasjonskoeffisienter som den aktuelle omhyllingskurven (133) for den ene blokken (131) av transformasjonskoeffisienter.

7. Transformasjonsbasert taledekoder (500) innrettet for å dekode en bitstrøm for å
10 tilveiebringe et rekonstruert talesignal; dekoderen (500) omfattende

 - en omhyllingskurvedekodingsenhet (531) innrettet for å bestemme en kvantisert aktuell omhyllingskurve (134) fra omhyllingskurvedata (161) inneholdt i bitstrømmen; hvor den kvantiserte aktuelle omhyllingskurven (134) angir et flertall spektralenergiverdier (303) for et motsvarende flertall frekvensblokker (301); hvor
15 bitstrømmen omfatter data (163, 164) som angir et flertall sekvensielle blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter; hvor en blokk (148) av rekonstruerte avflatede transformasjonskoeffisienter omfatter et flertall rekonstruerte avflatede transformasjonskoeffisienter for det motsvarende flertallet frekvensblokker (301);

20 - en omhyllingskurveinterpolasjonsenhet (104) innrettet for å bestemme et flertall interpolerte omhyllingskurver (136) henholdsvis for flertallet blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter, basert på den kvantiserte aktuelle omhyllingskurven (134) og basert på en kvantisert tidligere omhyllingskurve (134); og

 - en invers-avflatingsenhet (108) innrettet for å bestemme et flertall blokker
25 (149) av rekonstruerte transformasjonskoeffisienter ved å gi det motsvarende flertallet blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter en spektralform, med bruk henholdsvis av det motsvarende flertallet interpolerte omhyllingskurver (136); hvor det rekonstruerte talesignalet blir bestemt basert på flertallet blokker (149) av rekonstruerte transformasjonskoeffisienter, hvor eventuelt den kvantiserte tidligere
30 omhyllingskurven (134) er knyttet til et flertall tidligere blokker (149) av rekonstruerte

transformasjonskoeffisienter som er umiddelbart foregående flertallet blokker (149) av rekonstruerte transformasjonskoeffisienter.

8. Transformasjonsbasert taledekoder (500) ifølge krav 7, hvor

- flertallet sekvensielle blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter omfatter en første blokk (148) av rekonstruerte avflatede transformasjonskoeffisienter ved et første mellomtidspunkt;

- omhyllingskurveinterpolasjonsenheten (104) er innrettet for å bestemme en spektralenergiverdi (303) for en gitt frekvensblokk (301) til en første interpolert omhyllingskurve (136) ved å interpolere spektralenergiverdiene (303) for den gitte frekvensblokken (301) til den kvantiserte aktuelle omhyllingskurven (135) og til den kvantiserte tidligere omhyllingskurven (134) ved det første mellomtidspunktet;

- den første interpolerte omhyllingskurven (136) er knyttet til den første blokken (148) av rekonstruerte avflatede transformasjonskoeffisienter.

9. Transformasjonsbasert taledekoder (500) ifølge krav 8, hvor

omhyllingskurveinterpolasjonsenheten (104) er innrettet for å bestemme spektralenergiverdien (303) for den gitte frekvensblokken (301) til den første interpolerte omhyllingskurven (136) ved å kvantisere interpolasjonen mellom spektralenergiverdiene (303) for den gitte frekvensblokken (301) til den kvantiserte aktuelle omhyllingskurven (135) og til den kvantiserte tidligere omhyllingskurven (134).

10. Transformasjonsbasert taledekoder (500) ifølge et hvilket som helst av kravene 8 eller 9, hvor

- flertallet sekvensielle blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter omfatter en andre blokk (148) av rekonstruerte avflatede transformasjonskoeffisienter ved et andre mellomtidspunkt;

- omhyllingskurveinterpolasjonsenheten (104) er innrettet for å bestemme en spektralenergiverdi (303) for den gitte frekvensblokken (301) til en andre interpolert omhyllingskurve (136) ved å interpolere spektralenergiverdiene (303) for den gitte

frekvensblokken (301) til den kvantiserte aktuelle omhyllingskurven (135) og til den kvantiserte tidligere omhyllingskurven (134) ved det andre mellomtidspunktet;

- den andre interpolerte omhyllingskurven (136) er knyttet til den andre blokken (148) av rekonstruerte avflatede transformasjonskoeffisienter;

5 - den andre blokken (148) av rekonstruerte avflatede transformasjonskoeffisienter etterfølger den første blokken (148) av rekonstruerte avflatede transformasjonskoeffisienter; og

10 - det andre mellomtidspunktet er senere enn det første mellomtidspunktet, hvor en differanse mellom det andre mellomtidspunktet og det første mellomtidspunktet svarer til et tidsintervall mellom den andre blokken av rekonstruerte avflatede transformasjonskoeffisienter og den første blokken (148) av rekonstruerte avflatede transformasjonskoeffisienter.

11. Transformasjonsbasert taledekoder (500) ifølge et hvilket som helst av kravene 15 7 til 10, hvor

- bitstrømmen angir et flertall omhyllingskurveforsterkninger (137) henholdsvis for flertallet blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter;

20 - den transformasjonsbaserte taledekoderen (500) videre omfatter en omhyllingskurveforfiningsenhet (107) innrettet for å bestemme et flertall tilpassede omhyllingskurver (139) ved å anvende flertallet omhyllingskurveforsterkninger (137) henholdsvis på flertallet interpolerte omhyllingskurver (136);

25 - invers-avflatingsenheten (108) er innrettet for å bestemme flertallet blokker (149) av rekonstruerte transformasjonskoeffisienter ved å gi det motsvarende flertallet blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter en spektralform, med bruk henholdsvis av det motsvarende flertallet tilpassede omhyllingskurver (139).

12. Fremgangsmåte for å kode et talesignal inn i en bitstrøm; fremgangsmåten omfattende å

30 - motta et sett (132, 332) av blokker; hvor settet (132, 332) av blokker omfatter et flertall sekvensielle blokker (131) av transformasjonskoeffisienter; hvor flertallet sekvensielle blokker (131) angir punktprøver av talesignalet; hvor en blokk (131) av

transformasjonskoeffisienter omfatter et flertall transformasjonskoeffisienter for et motsvarende flertall frekvensblokker (301); bestemme en aktuell omhyllingskurve (133) basert på flertallet sekvensielle blokker (131) av transformasjonskoeffisienter; hvor den aktuelle omhyllingskurven (133) angir et flertall spektralenergiverdier (303) for det motsvarende flertallet frekvensblokker (301);

- bestemme en kvantisert aktuell omhyllingskurve (134) ved å kvantisere den aktuelle omhyllingskurven (133);

- bestemme et flertall interpolerte omhyllingskurver (136) henholdsvis for flertallet blokker (131) av transformasjonskoeffisienter, basert på den kvantiserte aktuelle omhyllingskurven (133) og basert på en kvantisert tidligere omhyllingskurve (134);

- bestemme et flertall blokker (140) av avflatede transformasjonskoeffisienter ved å avflate det motsvarende flertallet blokker (131) av transformasjonskoeffisienter med bruk henholdsvis av det motsvarende flertallet interpolerte omhyllingskurver (136); og

- bestemme bitstrømmen basert på flertallet blokker (140) av avflatede transformasjonskoeffisienter.

13. Fremgangsmåte for å dekode en bitstrøm for å tilveiebringe et rekonstruert talesignal; fremgangsmåten omfattende å

- bestemme en kvantisert aktuell omhyllingskurve (134) fra omhyllingskurvedata (161) inneholdt i bitstrømmen; hvor den kvantiserte aktuelle omhyllingskurven (134) angir et flertall spektralenergiverdier (303) for et motsvarende flertall frekvensblokker (301); hvor bitstrømmen omfatter data (163, 164) som angir et flertall sekvensielle blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter; hvor en blokk (148) av rekonstruerte avflatede transformasjonskoeffisienter omfatter et flertall rekonstruerte avflatede transformasjonskoeffisienter for det motsvarende flertallet frekvensblokker (301);

- bestemme et flertall interpolerte omhyllingskurver (136) henholdsvis for flertallet blokker (148) av rekonstruerte avflatede transformasjonskoeffisienter, basert

på den kvantiserte aktuelle omhyllingskurven (134) og basert på en kvantisert tidligere omhyllingskurve (134);

- bestemme et flertall blokker (149) av rekonstruerte

transformasjonskoeffisienter ved å gi det motsvarende flertallet blokker (148) av

5 rekonstruerte avflatede transformasjonskoeffisienter en spektralform, med bruk henholdsvis av det motsvarende flertallet interpolerte omhyllingskurver (136); og

- bestemme det rekonstruerte talesignalet basert på flertallet blokker (149) av rekonstruerte transformasjonskoeffisienter.

10 14. Fremgangsmåte for å kode et lydsignal omfattende et talesegment inn i en bitstrøm; hvor fremgangsmåten omfatter å

- identifisere talesegmentet fra lydsignalet;

- bestemme et flertall sekvensielle blokker (131) av

transformasjonskoeffisienter basert på talesegmentet, ved hjelp av en

15 transformasjonsenhet; hvor en blokk (131) av transformasjonskoeffisienter omfatter et flertall transformasjonskoeffisienter for et motsvarende flertall frekvensblokker (301);

hvor transformasjonsenheten er innrettet for å bestemme lange blokker omfattende et første antall transformasjonskoeffisienter og korte blokker omfattende et andre antall transformasjonskoeffisienter; hvor det første antallet er større enn det andre antallet;

20 hvor blokkene (131) i flertallet sekvensielle blokker (131) er korte blokker; og

- kode flertallet sekvensielle blokker (131) inn i bitstrømmen i samsvar med krav 12.

15. Fremgangsmåte for å dekode en bitstrøm som angir et lydsignal omfattende et

25 talesegment; fremgangsmåten omfattende å

- bestemme et flertall sekvensielle blokker (149) av rekonstruerte

transformasjonskoeffisienter basert på data (161, 162, 163, 164) inneholdt i

bitstrømmen i samsvar med krav 13; og

- bestemme et rekonstruert talesegment basert på flertallet sekvensielle blokker

30 (149) av rekonstruerte transformasjonskoeffisienter, ved hjelp av en

inverstransformasjonsenhet; hvor en blokk (149) av rekonstruerte

- transformasjonskoeffisienter omfatter et flertall rekonstruerte transformasjonskoeffisienter for et motsvarende flertall frekvensblokker (301); hvor inverstransformasjonsenheten (504) er innrettet for å behandle lange blokker omfattende et første antall rekonstruerte transformasjonskoeffisienter og korte blokker
- 5 omfattende et andre antall rekonstruert transformasjonskoeffisienter; hvor det første antallet er større enn det andre antallet; hvor blokkene (149) i flertallet sekvensielle blokker (149) er korte blokker.