

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
20. April 2006 (20.04.2006)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2006/039994 A3

(51) Internationale Patentklassifikation:
G10H 1/00 (2006.01)

(21) Internationales Aktenzeichen: PCT/EP2005/010333

(22) Internationales Anmeldedatum:
23. September 2005 (23.09.2005)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
10 2004 049 457.6
11. Oktober 2004 (11.10.2004) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **FRAUNHOFER-GESELLSCHAFT ZUR FÖRDERUNG DER ANGEWANDTEN FORSCHUNG E.V.** [DE/DE]; Hansastrasse 27c, 80686 München (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): **STREITENBERGER, Frank** [DE/DE]; Gardinistrasse 166, 81375 München (DE). **WEIS, Martin** [DE/DE]; Bergrat-Voigt-Strasse 9, 98693 Ilmenau (DE). **DERBOVEN, Claas** [DE/DE]; Schlörstrasse 3, 80634 München (DE). **CREMER, Markus** [DE/US]; 1868 San Juan Avenue, Berkeley, California 94707 (US).

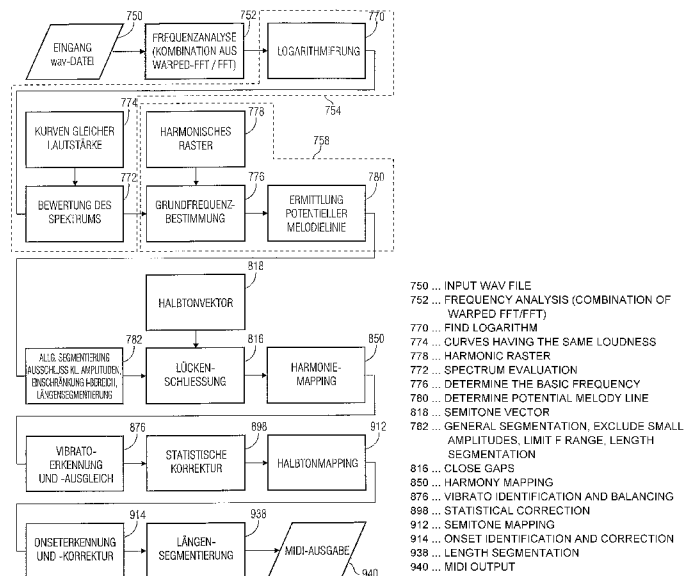
(74) Anwälte: **ZINKLER, Franz** usw.; Schoppe, Zimmermann, Stöckeler & Zinkler, Postfach 246, 82043 Pullach bei München (DE).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD AND DEVICE FOR THE EXTRACTION OF A MELODY ON WHICH AN AUDIO SIGNAL IS BASED

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR EXTRAKTION EINER EINEM AUDIOSIGNAL ZU GRUNDE LIEGENDEN MELODIE



(57) Abstract: The invention is based on the recognition that it is possible to make melody extraction or automatic transcription significantly more stable and possibly even less expensive if the assumption that the main melody is the portion of a piece of music which is perceived by humans as being the loudest and most prominent is taken into account to a sufficient degree. According to the invention, the time/spectral representation or the spectrogram of an audio signal that is of interest is scaled (772) using the curves (774) which have the same loudness and reflect the human perception of loudness in order to determine the melody of the audio signal based on the obtained perception-related time/spectral representation.

(57) Zusammenfassung: Die Erkenntnis der vorliegenden Erfindung besteht darin, dass die Melodieextraktion oder automatische Transkription deutlich stabiler und gegebenenfalls sogar unaufwendiger gestaltet werden kann, wenn die Annahme genügend Berücksichtigung findet, dass die Hauptmelodie derjenige Anteil eines Musikstückes ist, den

[Fortsetzung auf der nächsten Seite]



WO 2006/039994 A3



MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht
- vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

(88) Veröffentlichungsdatum des internationalen**Recherchenberichts:**

19. April 2007

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

der Mensch am lautesten und prägnantesten wahrnimmt. Dies aufgreifend wird gemäß der vorliegenden Erfindung die Zeit-/Spektralardarstellung bzw. das Spektrogramm eines interessierenden Audiosignals unter Verwendung der Kurven (774) gleicher Lautstärke, die die menschliche Lautstärkewahrnehmung widerspiegeln, skaliert (772), um auf der Basis der sich ergebenden wahrnehmungsbezogenen Zeit-/Spektralardarstellung die Melodie des Audiosignals zu ermitteln.

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2005/010333

A. CLASSIFICATION OF SUBJECT MATTER INV. G10H1/00				
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) G10H G10G				
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, WPI Data, INSPEC				
C. DOCUMENTS CONSIDERED TO BE RELEVANT				
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
A	GIULIANO MONTI, MARK SANDLER: "Automatic Polyphonic Piano Note Extraction Using Fuzzy Logic in a Blackboard System" PROC. OF THE 5TH INT. CONFERENCE ON DIGITAL AUDIO EFFECTS (DAFX-02) SEPTEMBER 26-28, 2002. SESSION 2, 26 September 2002 (2002-09-26), pages DAFX-39-DAFX-44, XP002364043 HAMBURG, GERMANY 2. FFT Front-End and Psychoacoustic Masking figure 1 <div style="text-align: center;"> ----- -/-- </div>	1,33,34		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.				
* Special categories of cited documents :				
<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none; vertical-align: top;"> "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "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) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed </td> <td style="width: 50%; border: none; vertical-align: top;"> "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 "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 "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. "&" document member of the same patent family </td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "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) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"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 "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 "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. "&" document member of the same patent family
"A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "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) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed	"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 "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 "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. "&" document member of the same patent family			
Date of the actual completion of the international search <div style="text-align: center; font-size: 1.2em;">23 January 2006</div>		Date of mailing of the international search report <div style="text-align: center; font-size: 1.2em;">08/02/2007</div>		
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer <div style="text-align: center; font-size: 1.2em;">Lecoite, M</div>		

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2005/010333

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>KARJALAINEN M ET AL: "Multi-pitch and periodicity analysis model for sound separation and auditory scene analysis" ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1999. PROCEEDINGS., 1999 IEEE INTERNATIONAL CONFERENCE ON PHOENIX, AZ, USA 15-19 MARCH 1999, PISCATAWAY, NJ, USA, IEEE, US, vol. 2, 15 March 1999 (1999-03-15), pages 929-932, XP010328512 ISBN: 0-7803-5041-3 2. Reduced Complexity Model of Auditory Pitch Perception figures 1,2</p>	1,33,34
A	<p>HARVEY FLETCHER AND W.A. MUNSON: "Loudness, Its Definition, Measurement and Calculation" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, vol. 5, October 1933 (1933-10), XP002364044 Definitions figure 4</p>	1,33,34

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen
PCT/EP2005/010333

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES INV. G10H1/00		
Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC		
B. RECHERCHIERTE GEBIETE		
Recherchierte Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) G10H G10G		
Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen		
Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) EPO-Internal, WPI Data, INSPEC		
C. ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	GIULIANO MONTI, MARK SANDLER: "Automatic Polyphonic Piano Note Extraction Using Fuzzy Logic in a Blackboard System" PROC. OF THE 5TH INT. CONFERENCE ON DIGITAL AUDIO EFFECTS (DAFX-02) SEPTEMBER 26-28,2002. SESSION 2, 26. September 2002 (2002-09-26), Seiten DAFX-39-DAFX-44, XP002364043 HAMBURG, GERMANY 2. FFT Front-End and Psychoacoustic Masking Abbildung 1 <div style="text-align: center; margin-top: 20px;"> ----- -/-- </div>	1,33,34
<input checked="" type="checkbox"/> Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen <input type="checkbox"/> Siehe Anhang Patentfamilie		
* Besondere Kategorien von angegebenen Veröffentlichungen : "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist "E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist "L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt) "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist "T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist "&" Veröffentlichung, die Mitglied derselben Patentfamilie ist		
Datum des Abschlusses der internationalen Recherche	Absenddatum des internationalen Recherchenberichts	
23. Januar 2006	08/02/2007	
Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Bevollmächtigter Bediensteter Lecointe, M	

INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP2005/010333

C. (Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	<p>KARJALAINEN M ET AL: "Multi-pitch and periodicity analysis model for sound separation and auditory scene analysis" ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1999. PROCEEDINGS., 1999 IEEE INTERNATIONAL CONFERENCE ON PHOENIX, AZ, USA 15-19 MARCH 1999, PISCATAWAY, NJ, USA, IEEE, US, Bd. 2, 15. März 1999 (1999-03-15), Seiten 929-932, XP010328512 ISBN: 0-7803-5041-3 2. Reduced Complexity Model of Auditory Pitch Perception Abbildungen 1,2</p>	1,33,34
A	<p>HARVEY FLETCHER AND W.A. MUNSON: "Loudness, Its Definition, Measurement and Calculation" JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, Bd. 5, Oktober 1933 (1933-10), XP002364044 Definitions Abbildung 4</p>	1,33,34