(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 June 2004 (10.06.2004)

PCT

(10) International Publication Number
WO 2004/049188 A1

(51) International Patent Classification[7]: G06F 17/00

(21) International Application Number:
PCT/SG2002/000279

(22) International Filing Date:
28 November 2002 (28.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: INSTITUTE FOR INFOCOMM RESEARCH [SG/SG]; 21 Heng Mui Keng Terrace, Singapore 119613 (SG).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: XU, Changsheng [CN/SG]; Blk 248 Choa Chu Kang Avenue 2, #10-490, Singapore 680248 (SG).

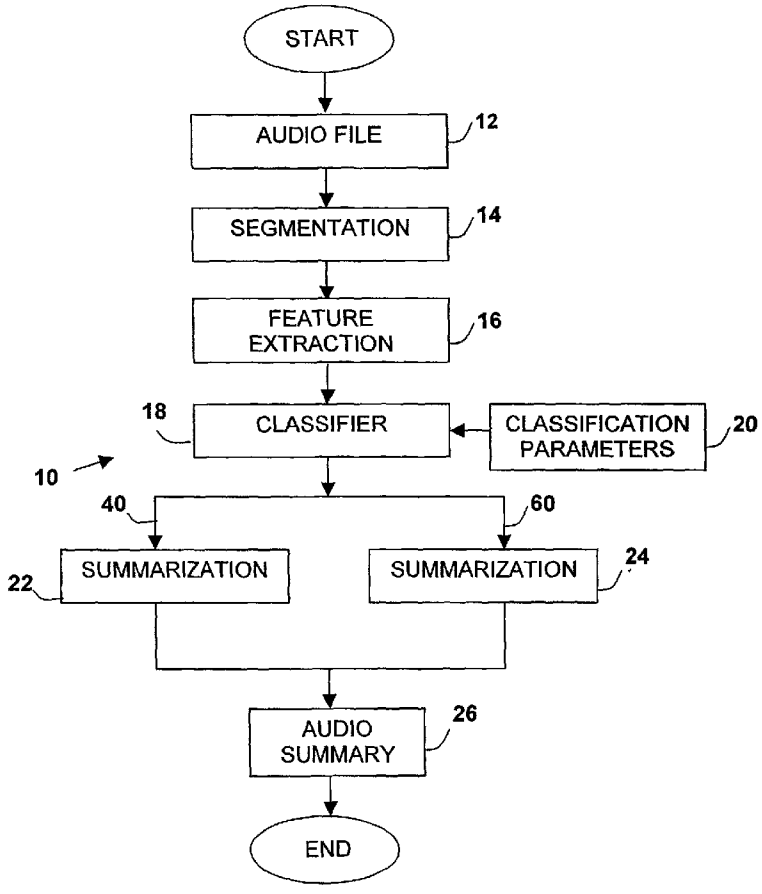
(74) Agent: ELLA CHEONG SPRUSON & FERGUSON (SINGAPORE) PTE LTD; 111 North Brige Road, #22-01 Peninsula Plaza, Singapore 179098 (SG).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

*[Continued on next page]*

(54) Title: SUMMARIZING DIGITAL AUDIO DATA




(57) Abstract: An embodiment is related to automatic summarization for digital audio raw data (12), more specifically, for identifying pure music and vocal music (40,60) from digital audio data by extracting distinctive features from music frames (73,74,75,76), designing a classifier and determining the classification parameters (20) using adaptive learning/training algorithm (36), and identifying music into pure music or vocal music according to the classifier. For pure music, temporal, spectral and cepstral features are calculated to characterise the musical content, and an adaptive clustering method is used to structure the musical content according to calculated features. The summary (22,24,26,48,52,70,72) is created according to clustered result and domain-based music knowledge (50,150). For vocal music, voice related features are extracted and used to structure the musical content, and similarly, the music summary is created in terms of structured content and heuristic rules related to music genres.

WO 2004/049188 A1

**Declarations under Rule 4.17:**

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)*

— *of inventorship (Rule 4.17(iv)) for US only*

**Published:**

— *with international search report*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*