

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 April 2007 (26.04.2007)

PCT

(10) International Publication Number
WO 2007/045723 A1

(51) International Patent Classification:
G10L 15/14 (2006.01)

(21) International Application Number:
PCT/FI2006/050445

(22) International Filing Date: 17 October 2006 (17.10.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/252,475 17 October 2005 (17.10.2005) US

(71) Applicant (for all designated States except US): **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FI-02150 Espoo (FI).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **OLSEN, Jesper** [FI/FI]; Sturenkatu 29 D 61, FI-00510 Helsinki (FI).

(74) Agent: **TAMPEREEN PATENTTITOIMISTO OY**; Hermiankatu 12 B, FI-33720 Tampere (FI).

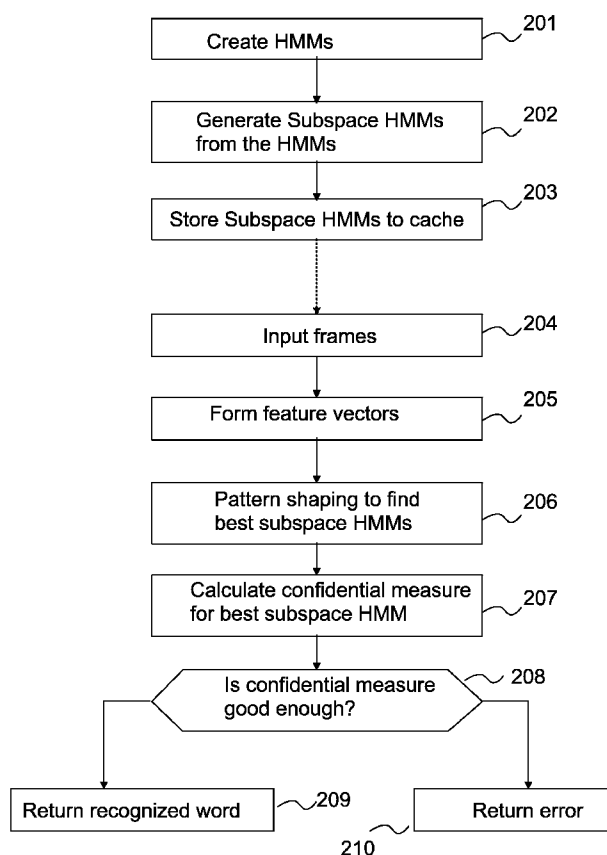
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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).

Published:
— with international search report

[Continued on next page]

(54) Title: A METHOD AND A DEVICE FOR SPEECH RECOGNITION



(57) Abstract: Method for speech recognition comprising inputting frames comprising samples of an audio signal; forming a feature vector comprising a first number of vector components for each frame; projecting the feature vector onto at least two subspaces so that the number of components of each projected feature vector is less than the first number and the total number of components of the projected feature vectors is the same as the first number; defining a set of mixture models for each projected vector which provides the highest observation probability; analyzing the set of mixture models to determine the recognition result. When the recognition result is found, the method comprises determining a confidence measure for the recognition result, the determining comprising determining a probability that the recognition result is correct; determining a normalizing term; and dividing the probability by the normalizing term.

WO 2007/045723 A1



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.