

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 August 2007 (23.08.2007)

PCT

(10) International Publication Number
WO 2007/094684 A3

(51) International Patent Classification:
G06K 9/00 (2006.01) **G06K 9/72** (2006.01)

(74) Agent: MIDTTUN, Gisle; Zacco Norway AS, P.O. Box 2003 Vika, N-0125 Oslo (NO).

(21) International Application Number:
PCT/NO2007/000053

(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.

(22) International Filing Date:
15 February 2007 (15.02.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20060787 17 February 2006 (17.02.2006) NO

(71) Applicant (for all designated States except US): LUMEX AS [NO/NO]; Øvre Vollgate 7, N-0158 Oslo (NO).

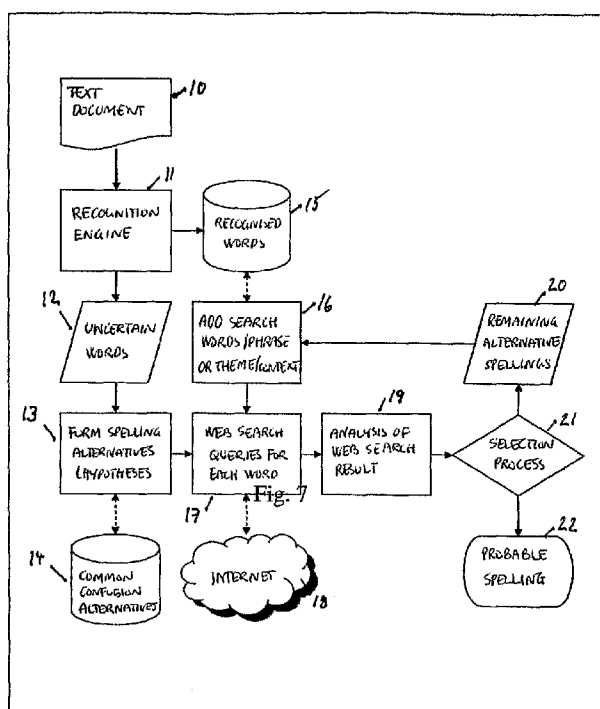
(72) Inventors; and

(75) Inventors/Applicants (for US only): MEYER, Hans, Christian [NO/NO]; Bygdøy Allé 33B, N-0262 Oslo (NO). CARLIN, Mats, Stefan [NO/NO]; Vallerveien 152 E, N-1346 Gjettem (NO). FOSSEIDE, Knut, Tharald [NO/NO]; Ingierkollveien 86, NO-1410 Kolbotn (NO).

(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).

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR VERIFICATION OF UNCERTAINLY RECOGNIZED WORDS IN AN OCR SYSTEM



(57) Abstract: The present invention provides a method and system for confirming uncertainly recognized words as reported by an Optical Character or speech Recognition process by using spelling alternatives as search arguments for an Internet search engine. The measured number of hits for each spelling alternative is used to provide a confirmation measure for the most probable spelling alternative. Whenever the confirmation measure is inconclusive, a plurality of search strategies are used to reach a measured result comprising zero hits except for one spelling alternative that is used as the correct alternative.



Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

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.

(88) Date of publication of the international search report:

13 December 2007