



(43) International Publication Date
7 April 2011 (07.04.2011)

(10) International Publication Number
WO 2011/041795 A1

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/US2010/051357

(22) International Filing Date:
4 October 2010 (04.10.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/247,973 2 October 2009 (02.10.2009) US

(72) Inventor; and

(71) Applicant : **MUSULURI, Aravind** [IN/US]; 237 Paradise Lake Drive, Birmingham, AL 35244 (US).

(74) Agents: **DENNEN, Ann, I.** et al.; Lanier Ford Shaver & Payne P.C., P.O. Box 2087, Huntsville, AL 35804-2087 (US).

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, — with international search report (Art. 21(3))

(54) Title: SYSTEM AND METHOD FOR BLOCK SEGMENTING, IDENTIFYING AND INDEXING VISUAL ELEMENTS, AND SEARCHING DOCUMENTS

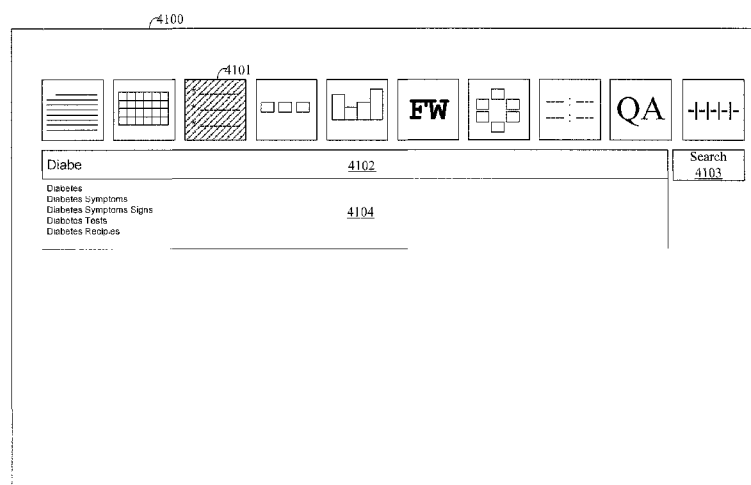


FIG. 41

(57) Abstract: A method for segmenting, identifying and indexing visual elements, and searching documents comprises for each document generating metadata, segmenting the document into blocks using the metadata, performing block operations on the identified blocks, identifying and indexing inline visual elements using data and metadata rules, identifying and indexing block visual elements using profiles, and searching for documents containing visual elements.