

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
25 October 2007 (25.10.2007)

PCT

(10) International Publication Number
WO 2007/121029 A3

(51) International Patent Classification:
G06K 9/34 (2006.01)

(21) International Application Number:
PCT/US2007/064616

(22) International Filing Date: 22 March 2007 (22.03.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11/392,213 29 March 2006 (29.03.2006) US

(71) Applicant (for all designated States except US): **AMAZON TECHNOLOGIES, INC.** [US/US]; 920 Incline Way, Suite C, Incline Village, NV 89451 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **COATH, Adam, Brian** [CA/US]; 1200-12th Avenue, South, Suite 1200, Seattle, WA 98144 (US). **AKALIN, Frederick, Ziya Ramos** [US/US]; 1200-12th Avenue South, Suite 1200,

Seattle, WA 98144 (US). **GOODWIN, Robert, L.** [US/US]; 1200-12th Avenue South, Suite 1200, Seattle, WA 98144 (US). **SHAGAM, Joshua** [US/US]; 1200-12th Avenue South, Suite 1200, Seattle, WA 98144 (US).

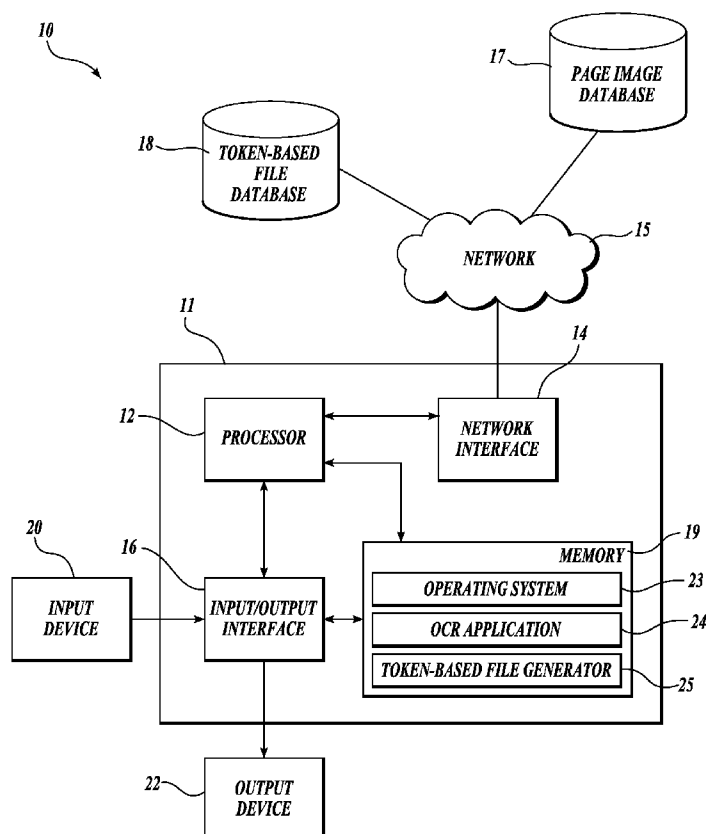
(74) Agent: **POWELL, Tracy, S.**; Christensen O'Connor Johnson Kindess PLLC, 1420 Fifth Avenue, Suite 2800, Seattle, WA 98101 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, 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, LY, MA, MD, ME, 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),

[Continued on next page]

(54) Title: CONVERTING DIGITAL IMAGES CONTAINING TEXT TO TOKEN-BASED FILES FOR RENDERING



(57) Abstract: A computer-implemented method is provided for converting a scanned-in electronic image into a token-based file. The method includes generally five steps. First, various tokens (i.e., graphical units) are identified in the electronic image. Second, the identified tokens having similar shapes are classified together to form a token group, to thereby form multiple token groups, each including one or more tokens having similar shapes. Third, in each token group, a representative token is found, which morphologically represents the shapes of tokens included in the group. Fourth, each representative token is converted into a vectorized token, which is a mathematical representation of the shape of the representative token. Fifth, each of the vectorized tokens is associated with the positions of the tokens in the electronic image represented by the vectorized token. Thus, upon rendering, the vectorized token is displayed to thereby create a page image consisting only of clean images of vectorized tokens.



European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, 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).

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

Published:

— *with international search report*

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

16 October 2008

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US07/64616

A. CLASSIFICATION OF SUBJECT MATTER

IPC: G06K 9/34(2006.01)

USPC: 382/180

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)

U.S. : 382/114, 162, 164-165, 173, 175-181, 224-225, 229-230, 306, 321

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 practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent 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

09 June 2008 (09.06.2008)

Date of mailing of the international search report

21 AUG 2008

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No. (571) 273-3201

Authorized officer

Jose L. Couso

Telephone No. (571) 272-7361