US007778489B1

(12) **United States Patent**
Siegel et al.

(10) **Patent No.: US 7,778,489 B1**
(45) **Date of Patent: Aug. 17, 2010**

(54) **METHOD AND SYSTEM FOR DETERMINING THE LEGIBILITY OF TEXT IN AN IMAGE**

(75) Inventors: **Hilliard B. Siegel**, Seattle, WA (US); **David Chick**, Seattle, WA (US)

(73) Assignee: **Amazon Technologies, Inc.**, Reno, NV (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **12/266,444**

(22) Filed: **Nov. 6, 2008**

Related U.S. Application Data

(62) Division of application No. 10/790,256, filed on Mar. 1, 2004, now Pat. No. 7,466,875.

(51) **Int. Cl.**
G06K 9/36 (2006.01)

(52) **U.S. Cl.** **382/286**; 382/299; 358/1.2; 358/452

(58) **Field of Classification Search** 382/282, 382/286, 299; 358/1.2, 452
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

4,850,026	A	7/1989	Jeng	
4,944,022	A	7/1990	Yasujima	
4,994,987	A	2/1991	Baldwin	
5,014,327	A	5/1991	Potter	
5,167,016	A	11/1992	Bagley	
5,465,309	A	11/1995	Johnson	
5,956,468	A	9/1999	Ancin	
5,960,448	A	9/1999	Reicheck	
5,970,483	A	10/1999	Evans	
5,987,448	A	11/1999	Evans	
6,038,342	A	3/2000	Bernzott	
6,226,631	B1	5/2001	Evans	
6,363,179	B1	3/2002	Evans	
6,452,686	B1 *	9/2002	Svetkoff et al.	356/602
6,453,079	B1	9/2002	McInerny	
6,532,461	B2	3/2003	Evans	
6,570,974	B1 *	5/2003	Gerszberg et al.	379/218.01
6,590,608	B2 *	7/2003	Matsumoto et al.	348/231.2
6,661,919	B2	12/2003	Nicholson	
6,728,425	B1 *	4/2004	Tokuyama et al.	382/299
6,980,332	B2	12/2005	Simske	
7,199,882	B2 *	4/2007	Svetkoff et al.	356/602
2004/0199875	A1	10/2004	Samson	

* cited by examiner

Primary Examiner—Yosef Kassa
(74) *Attorney, Agent, or Firm*—Christensen O'Connor Johnson Kindness PLLC

(57) **ABSTRACT**

Legibility of text in an image of a page is determined by comparing a measure of the text in the page image with a measure of the page image itself. In one aspect, a measure of the text in the page image may be the height of a line of text, while the measure derived from the page image may be the height of the page image. A text-to-page height ratio is determined and compared to one or more thresholds for determining legibility. In another aspect of the invention, a measure of the text in a page image is obtained by measuring the word density in the page image, while the measure derived from the page image comprises compressing the page image and determining the size of the compressed image file. Legibility is then determined by comparing the measure of word density with the compressed image file size.

49 Claims, 12 Drawing Sheets