

**METHOD, SYSTEM, AND NON-TRANSITORY COMPUTER-READABLE
RECORDING MEDIUM FOR RECOMMENDING CONTENTS FOR WRITING**

FIELD OF THE INVENTION

The present invention relates to a method, system, and non-transitory computer-readable recording medium for recommending contents for writing.

BACKGROUND

In general, writing (or composing) is done to fulfill a variety of goals. For example, writing a cover letter is done to fulfill a goal such as employment or college admission, and writing a thesis is done to fulfill a goal such as earning a degree or academic achievement.

In the process of the writing, it is common to consult various writing-related materials (e.g., articles, news, and reports) to improve the quality of the writing.

However, there has been no technique for automatically recommending writing-related materials in the writing process, so that the process of searching for and discovering the writing-related materials has been time and effort consuming, resulting in a problem of reduced focus on the writing.

SUMMARY OF THE INVENTION

One object of the present invention is to solve all the above-described problems in the prior art.

Another object of the invention is to automatically recommend writing-related materials (or contents) on the basis of keywords composed in the process of writing, thereby dramatically reducing the time and effort spent searching for and discovering the writing-related materials.

The representative configurations of the invention to achieve the above objects are described below.

According to one aspect of the invention, there is provided a method for recommending contents for writing, the method comprising the steps of: extracting main keywords from a user document composed of at least one block; and determining a recommended content associated with the user document with reference to at least one of a style attribute assigned to the main keywords in the user document and an attribute of a block containing the main keywords.

According to another aspect of the invention, there is provided a system for recommending contents for writing, the system comprising: a keyword extraction unit configured to extract main keywords from a user document composed of at least one block; and a content determination unit configured to determine a recommended content associated with the user document with reference to at least one of a style attribute

assigned to the main keywords in the user document and an attribute of a block containing the main keywords.

In addition, there are further provided other methods and systems to implement the invention, as well as non-transitory computer-readable recording media having stored thereon computer programs for executing the methods.

According to the invention, it is possible to automatically recommend writing-related materials (or contents) on the basis of keywords composed in the process of writing, thereby dramatically reducing the time and effort spent searching for and discovering the writing-related materials.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 schematically shows the configuration of an entire system for recommending contents for writing according to one embodiment of the invention.

FIG. 2 specifically shows the internal configuration of a content recommendation system according to one embodiment of the invention.

FIG. 3 illustratively shows a user interface screen provided by the content recommendation system according to one embodiment of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In the following detailed description of the present invention, references are made to the accompanying drawings that show, by way of illustration, specific embodiments in which the invention may be practiced. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention. It is to be understood that the various embodiments of the invention, although different from each other, are not necessarily mutually exclusive. For example, specific shapes, structures and characteristics described herein may be implemented as modified from one embodiment to another without departing from the spirit and scope of the invention. Furthermore, it shall be understood that the positions or arrangements of individual elements within each embodiment may also be modified without departing from the spirit and scope of the invention. Therefore, the following detailed description is not to be taken in a limiting sense, and the scope of the invention is to be taken as encompassing the scope of the appended claims and all equivalents thereof. In the drawings, like reference numerals refer to the same or similar elements throughout the several views.

Herein, the term "content" or "contents" encompasses digital information or individual information elements comprised of text, symbol, speech, sound, image, video, and the like. For example, such contents may comprise data such

as text, image, video, audio, and links (e.g., web links) or a combination of at least two types of such data.

Hereinafter, various preferred embodiments of the invention will be described in detail with reference to the accompanying drawings to enable those skilled in the art to easily implement the invention.

Configuration of the entire system

FIG. 1 schematically shows the configuration of the entire system for recommending contents for writing according to one embodiment of the invention.

As shown in FIG. 1, the entire system according to one embodiment of the invention may comprise a communication network 100, a content recommendation system 200, and a device 300.

First, the communication network 100 according to one embodiment of the invention may be implemented regardless of communication modality such as wired and wireless communications, and may be constructed from a variety of communication networks such as local area networks (LANs), metropolitan area networks (MANs), and wide area networks (WANs). Preferably, the communication network 100 described herein may be the Internet or the World Wide Web (WWW). However, the communication network 100 is not necessarily limited thereto, and may at least partially include known wired/wireless data communication networks, known telephone networks, or known wired/wireless television communication

networks.

Next, the content recommendation system 200 according to one embodiment of the invention may communicate with the device 300 to be described below via the communication network 100, and may function to extract main keywords from a user document composed of at least one block, and determine a recommended content associated with the user document with reference to at least one of a style attribute assigned to the main keywords in the user document and an attribute of a block containing the main keywords. Meanwhile, the content recommendation system 200 may be digital equipment having a memory means and a microprocessor for computing capabilities, and may be, for example, a server system operating on the communication network 100.

The configuration and functions of the content recommendation system 200 according to one embodiment of the invention will be discussed in detail below.

Next, the device 300 according to one embodiment of the invention is digital equipment capable of connecting to and then communicating with the content recommendation system 200, and any type of digital equipment having a memory means and a microprocessor for computing capabilities, such as a smart phone, a tablet, a smart watch, a smart band, smart glasses, a desktop computer, a notebook computer, a workstation, a personal digital assistant (PDAs), a web pad, and a mobile phone, may be adopted as the device 300

according to the invention.

Further, according to one embodiment of the invention, the device 300 may further include an application program for performing the functions according to the invention. The application may reside in the device 300 in the form of a program module. The characteristics of the program module may be generally similar to those of a keyword extraction unit 210, a content determination unit 220, a communication unit 230, and a control unit 240 of the content recommendation system 200 to be described below. Here, at least a part of the application may be replaced with a hardware device or a firmware device that may perform a substantially equal or equivalent function, as necessary.

Configuration of the content recommendation system

Hereinafter, the internal configuration of the content recommendation system 200 crucial for implementing the invention and the functions of the respective components thereof will be discussed.

FIG. 2 specifically shows the internal configuration of the content recommendation system 200 according to one embodiment of the invention.

As shown in FIG. 2, the content recommendation system 200 according to one embodiment of the invention may comprise a keyword extraction unit 210, a content determination unit 220, a communication unit 230, and a control unit 240. According to one embodiment of the

invention, at least some of the keyword extraction unit 210, the content determination unit 220, the communication unit 230, and the control unit 240 of the content recommendation system 200 may be program modules to communicate with an external system (not shown). The program modules may be included in the content recommendation system 200 in the form of operating systems, application program modules, or other program modules, while they may be physically stored in a variety of commonly known storage devices. Further, the program modules may also be stored in a remote storage device that may communicate with the content recommendation system 200. Meanwhile, such program modules may include, but are not limited to, routines, subroutines, programs, objects, components, data structures, and the like for performing specific tasks or executing specific abstract data types as will be described below in accordance with the invention.

Meanwhile, the above description is illustrative although the content recommendation system 200 has been described as above, and it will be apparent to those skilled in the art that at least a part of the components or functions of the content recommendation system 200 may be implemented in the device 300 or a server (not shown) or included in an external system (not shown), as necessary.

First, the keyword extraction unit 210 according to one embodiment of the invention may function to extract main

keywords from a user document composed of at least one block.

According to one embodiment of the invention, a user may input text (which may include characters, numbers, symbols, or any combination thereof) to the keyword extraction unit 210 (or to a server or platform that supports the functionality of the keyword extraction unit 210) to write a document (e.g., a cover letter, a thesis, or a report). (Meanwhile, the user may further input images, graphs, tables, and the like in addition to the text.) According to one embodiment of the invention, the text inputted from the user may form at least one block, and the at least one block may constitute the above document (hereinafter, "user document"). Here, the block according to one embodiment of the invention may refer to a single paragraph or a plurality of paragraphs having an extent of association not less than a predetermined level.

Next, the keyword extraction unit 210 according to one embodiment of the invention may perform preprocessing on the user document that is composed on the basis of the user input (wherein the user document may include a document that is being composed (i.e., has not been completed) by the user). For example, the keyword extraction unit 210 according to one embodiment of the invention may remove text not used in the process of recommending contents (e.g., spaces and special symbols) from the user document. As another example, the keyword extraction unit 210 according

to one embodiment of the invention may use a morphological analyzer (e.g., Komoran or Kkma) to segment the text in the user document by morphemes.

Next, the keyword extraction unit 210 according to one embodiment of the invention may extract main keywords from the above user document (or the user document on which the preprocessing has been performed) using a natural language processing technique. For example, the keyword extraction unit 210 according to one embodiment of the invention may extract the main keywords from the user document on the basis of a result of performing a natural language processing technique such as TF-IDF (Term Frequency-Inverse Document Frequency). Here, according to one embodiment of the invention, the TF-IDF may refer to a statistical figure indicating how important a certain word is in a particular document. More specifically, the keyword extraction unit 210 according to one embodiment of the invention may determine keywords of high importance in the user document on the basis of the result of performing the TF-IDF, and may extract a predetermined number of keywords as the main keywords from among the keywords determined to be of high importance (e.g., the top five keywords may be extracted as the main keywords from among the keywords determined to be of high importance).

Next, the content determination unit 220 according to one embodiment of the invention may function to determine a

recommended content associated with the user document with reference to at least one of a style attribute assigned to the main keywords in the user document and an attribute of a block containing the main keywords.

According to one embodiment of the invention, the content determination unit 220 may assign scores to the main keywords extracted from the user document by the keyword extraction unit 210, and may apply weights to the scores.

For example, the content determination unit 220 according to one embodiment of the invention may assign scores to the main keywords extracted from the user document on the basis of ranking of importance according to the result of performing the TF-IDF. More specifically, assuming that a first keyword, a second keyword, a third keyword, a fourth keyword, a fifth keyword, and a sixth keyword are extracted as the main keywords from the user document, and their ranking of importance according to the result of performing the TF-IDF are the third, the second, the first, the fifth, and the fourth, respectively, the first keyword, the second keyword, the third keyword, the fourth keyword, and the fifth keyword may be assigned scores of 3, 4, 5, 1, and 2 points, respectively. Meanwhile, the score according to one embodiment of the invention may be set to a score out of 5, but is not necessarily limited thereto.

Further, the content determination unit 220 according

to one embodiment of the invention may apply weights to the scores assigned to the main keywords, with reference to a style attribute assigned to (or set for) the main keywords in the user document. For example, the content determination unit 220 according to one embodiment of the invention may apply weights to (e.g., double) the scores assigned to the main keywords, with reference to a style attribute assigned to the main keywords in the user document such that the main keywords are set to bold. As another example, the content determination unit 220 according to one embodiment of the invention may apply weights to the scores assigned to the main keywords, with reference to a style attribute assigned to the main keywords in the user document such that the main keywords are set to be included in a heading (or subheading) of the user document. Here, according to one embodiment of the invention, when the main keywords are set to be included in the heading (or subheading) of the user document, a greater weight may be applied as a font size of the main keywords is larger. (Here, whether the main keywords are set to be included in the heading (or subheading) of the user document may be determined with reference to the font size of the main keywords (e.g., when the font size of the main keywords is larger than the most frequent font size in the user document (or than the font size of other keywords in the user document) at or above a predetermined level, it may be

determined that the main keywords are set to be included in the heading (or subheading).) More specifically, according to one embodiment of the invention, a weight of 2 may be applied when the font size of the main keywords is 20 pt (wherein pt may represent a predetermined unit of font size), and a weight of 2.4 may be applied when the font size of the main keywords is 24 pt. That is, according to one embodiment of the invention, when the main keywords are set to be included in the heading (or subheading) of the user document, the weights applied to the scores assigned to the main keywords may be increased in proportion to the increase in the font size of the main keywords.

Further, the content determination unit 220 according to one embodiment of the invention may apply weights to the scores assigned to the main keywords with reference to an attribute of a block containing the main keywords. Here, according to one embodiment of the invention, the attribute of the block may include importance of the block, and a predetermined level may be assigned to each block on the basis of the attribute of the block. According to one embodiment of the invention, the attribute of the block may be determined (or set) with reference to a type of the user document. Specifically, according to one embodiment of the invention, the at least one block constituting the user document may have different levels of importance according to the type of the user document. The content determination

unit 220 according to one embodiment of the invention may set the importance of each of the at least one block constituting the user document, with reference to the type of the user document and a lookup table for the importance of the at least one block constituting the user document, and may assign a predetermined level to the block on the basis of the importance. For example, assuming that the type of the user document is a paper, a first block in an abstract of the user document, a second block in a body of the user document, and a third block in a conclusion of the user document may have different levels of importance, and the first block, the third block, and the second block may have higher levels of importance in that order according to the lookup table. Thus, the content determination unit 220 according to one embodiment of the invention may set the importance of the first block, the second block, and the third block constituting the user document with reference to the lookup table, respectively, and the first block, the third block, and the second block may be assigned higher levels in that order on the basis of the importance. Then, the content determination unit 220 according to one embodiment of the invention may determine which block of the user document contains the main keywords extracted from the user document, and may apply weights to the scores assigned to the main keywords with reference to the level assigned to the block.

Meanwhile, the content determination unit 220 according to one embodiment of the invention may determine a recommended content associated with the user document with reference to the scores assigned to the main keywords and the weights applied to the scores. Here, the content according to one embodiment of the invention may include a variety of contents such as text (e.g., news, papers, and reports), images, and videos.

For example, the content determination unit 220 according to one embodiment of the invention may determine a content including at least one of the main keywords extracted from the user document as the recommended content associated with the user document. Then, the content determination unit 220 according to one embodiment of the invention may determine the recommended content that is to be preferentially recommended, with reference to the number of main keywords included in the recommended content determined as above. Here, when each of a plurality of recommended contents includes the same number of main keywords, the content determination unit 220 according to one embodiment of the invention may determine the recommended content to be preferentially recommended with reference to a score obtained by summing the scores (more specifically, weighted scores) assigned to the main keywords included in each of the recommended contents.

More specifically, according to one embodiment of the

invention, it may be assumed that a first keyword, a second keyword, a third keyword, a fourth keyword, and a fifth keyword are extracted as the main keywords from the user document, and scores (more specifically, weighted scores) assigned to the keywords are 10, 12, 20, 8, and 15 points, respectively. The content determination unit 220 according to one embodiment of the invention may determine a first content including the first and second keywords, a second content including the first, second, and third keywords, a third content including the first, second, third, and fourth keywords, and a fourth content including the first, second, third, and fifth keywords as recommended contents associated with the user document. The content determination unit 220 according to one embodiment of the invention may determine a recommended content to be preferentially recommended, with reference to the number of main keywords included in the recommended contents determined as above, wherein with respect to the third content and the fourth content including the same number (four) of main keywords, the recommended content to be preferentially recommended may be determined by summing the scores assigned to the main keywords included in each of the recommended contents. That is, according to one embodiment of the invention, a score obtained by summing the scores assigned to the main keywords included in the third content (i.e., the first, second, third, and fourth keywords) is 50 points, and a score

obtained by summing the scores assigned to the main keywords included in the fourth content (i.e., the first, second, third, and fifth keywords) is 57 points, so that the fourth content may be recommended in preference to the third content. Finally, according to one embodiment of the invention, the recommended contents may be provided to the user in the order of the fourth content, the third content, the second content, and the first content.

Meanwhile, the content determination unit 220 according to one embodiment of the invention may cause the recommended content to be dynamically determined on the basis of changes to the main keywords extracted from the user document. For example, the content determination unit 220 according to one embodiment of the invention may cause the recommended content to be provided to the user in real time via a user interface screen while the user composes the document, and may cause a content different from a previously recommended content to be provided to the user to reflect changes that the user makes to the content of the document (e.g., modification or addition of text).

Next, the communication unit 230 according to one embodiment of the invention may function to enable data transmission/reception from/to the keyword extraction unit 210 and the content determination unit 220.

Lastly, the control unit 240 according to one embodiment of the invention may function to control data

flow among of the keyword extraction unit 210, the content determination unit 220, and the communication unit 230. That is, the control unit 240 according to the invention may control data flow into/out of the content recommendation system 200 or data flow among the respective components of the content recommendation system 200, such that the keyword extraction unit 210, the content determination unit 220, and the communication unit 230 may carry out their particular functions, respectively.

FIG. 3 illustratively shows a user interface screen provided by the content recommendation system 200 according to one embodiment of the invention.

According to one embodiment of the invention, a block associated with a user input may be displayed in a first space 310. Here, according to one embodiment of the invention, the block associated with the user input may be a block in which text is entered by the user input. The content recommendation system 200 (more specifically, the content determination unit 220) according to one embodiment of the invention may cause the block associated with the user input to be displayed separately from other blocks, with reference to the occurrence of the user input to the user document. For example, the content recommendation system 200 according to one embodiment of the invention may cause only the block associated with the user input to be displayed in the first space 310 without displaying other

blocks of the user document. Further, the content recommendation system 200 according to one embodiment of the invention may cause at least a part of the block associated with the user input to be highlighted on the basis of a visual element, with reference to the occurrence of the user input to the user document. For example, the content recommendation system 200 according to one embodiment of the invention may cause a title (e.g., heading, subheading, or tentative heading) of the block associated with the user input to be highlighted on the basis of at least one of color and boldness, with reference to the occurrence of the user input to the user document (see 311 in FIG. 3).

Further, according to one embodiment of the invention, a recommended content associated with the user document (e.g., a heading of the recommended content) may be displayed in a second space 320. Here, according to one embodiment of the invention, a first content 321 may be recommended in preference to a second content 322 on the basis of ranking of preferential recommendation. For example, the first content 321 may be exposed above the second content 322 in the second space 320. Further, according to one embodiment of the invention, when the user selects (or clicks on) a tab for the first content 321 or the second content 322 (or a graphical element representing the content), a detailed page for the content may be displayed in the second space 320. For example, when the

user selects a tab for the first content 321, a summary of the first content 321, a link to view the original of the first content 321, and the like may be displayed in the second space 320.

The embodiments according to the invention as described above may be implemented in the form of program instructions that can be executed by various computer components, and may be stored on a computer-readable recording medium. The computer-readable recording medium may include program instructions, data files, and data structures, separately or in combination. The program instructions stored on the computer-readable recording medium may be specially designed and configured for the present invention, or may also be known and available to those skilled in the computer software field. Examples of the computer-readable recording medium include the following: magnetic media such as hard disks, floppy disks and magnetic tapes; optical media such as compact disk-read only memory (CD-ROM) and digital versatile disks (DVDs); magneto-optical media such as floptical disks; and hardware devices such as read-only memory (ROM), random access memory (RAM) and flash memory, which are specially configured to store and execute program instructions. Examples of the program instructions include not only machine language codes created by a compiler, but also high-level language codes that can be executed by a computer using an interpreter.

The above hardware devices may be changed to one or more software modules to perform the processes of the present invention, and vice versa.

Although the present invention has been described above in terms of specific items such as detailed elements as well as the limited embodiments and the drawings, they are only provided to help more general understanding of the invention, and the present invention is not limited to the above embodiments. It will be appreciated by those skilled in the art to which the present invention pertains that various modifications and changes may be made from the above description.

Therefore, the spirit of the present invention shall not be limited to the above-described embodiments, and the entire scope of the appended claims and their equivalents will fall within the scope and spirit of the invention.

WHAT IS CLAIMED IS:

1. A method for recommending contents for writing, the method comprising the steps of:

extracting main keywords from a user document composed of at least one block; and

determining a recommended content associated with the user document with reference to at least one of a style attribute assigned to the main keywords in the user document and an attribute of a block containing the main keywords.

2. The method of Claim 1, wherein in the determining step, scores assigned to the main keywords are weighted with reference to at least one of the style attribute and the attribute of the block, and the recommended content is determined with reference to the weighted scores.

3. The method of Claim 2, wherein the attribute of the block is determined with reference to a type of the user document.

4. The method of Claim 1, wherein in the determining step, the recommended content is dynamically determined on the basis of changes to the main keywords extracted from the user document.

5. The method of Claim 1, wherein with reference to occurrence of a user input to the user document, a block associated with the user input in the user document is displayed separately from other blocks.

6. The method of Claim 5, wherein at least a part of the block associated with the user input is highlighted on the basis of a visual element.

7. A non-transitory computer-readable recording medium having stored thereon a computer program for executing the method of Claim 1.

8. A system for recommending contents for writing, the system comprising:

- a keyword extraction unit configured to extract main keywords from a user document composed of at least one block; and

- a content determination unit configured to determine a recommended content associated with the user document with reference to at least one of a style attribute assigned to the main keywords in the user document and an attribute of a block containing the main keywords.

9. The system of Claim 8, wherein the content determination unit is configured to weight scores assigned

to the main keywords with reference to at least one of the style attribute and the attribute of the block, and determine the recommended content with reference to the weighted scores.

10. The system of Claim 9, wherein the attribute of the block is determined with reference to a type of the user document.

11. The system of Claim 8, wherein the content determination unit is configured to cause the recommended content to be dynamically determined on the basis of changes to the main keywords extracted from the user document.

12. The system of Claim 8, wherein with reference to occurrence of a user input to the user document, a block associated with the user input in the user document is displayed separately from other blocks.

13. The system of Claim 12, wherein at least a part of the block associated with the user input is highlighted on the basis of a visual element.