

DESCRIPTION

POSTED INFORMATION EXTRACTION CONTROL DEVICE
AND POSTED INFORMATION EXTRACTION CONTROL PROGRAM

Technical Field

[0001] The present disclosure relates to a posted information extraction control device and a posted information extraction control program for extracting given posted information from out of posted information (hereafter also referred to as comments) posted on a social networking service (SNS).

Background Art

[0002] Analysis of comments posted on a SNS (such as customer review analysis, SNS account analysis, posting management, attribute analysis, click count analysis, diffusion analysis, and competitive analysis) is now commonly performed.

[0003] Japanese Patent Application Laid-Open (JP-A) No. 2014-142738 describes extensively exploiting a large quantity of information posted on a SNS, and exploiting posted information that has been made more user-friendly in order to provide a more user-friendly service.

[0004] JP-A No. 2014-142738 includes a first group based on first posted information and a second group based on second posted information. A phrase extracted from the first posted information and a phrase extracted from the second posted information are compared. Based on the result of this comparison, a first keyword unique to the first posted information and a second keyword unique to the second posted information are identified. The first keyword is managed in association with the first group, and the second keyword is managed in association with the second group.

[0005] International Publication (WO) No. 2015/052785 describes a system in which an information extraction server for extracting information relating to infrastructure, a past incident DB, a space-time information DB, and a manager terminal are capable of communicating with each other. This system is capable of communicating with a client terminal for displaying information extracted by the information extraction server, and map information is employed by the client server to display received infrastructure status information using a predetermined icon at a predetermined position on a map.

SUMMARY OF INVENTION

Technical Problem

[0006] However, hitherto the utilization of comments has been limited to ascertaining trends

in comments to be exploited for marketing purposes and the like. It has therefore not been possible to evaluate the quality of comments, for example of the consistency of comment origin position information.

[0007] Hitherto, all comments posted on a SNS have been processed. These comments therefore included many in which the topic of the comment and an area (location) identified with this topic are not linked together, making evaluation of the identified area (for example a region or space designated for what is referred to as community development) impossible. Moreover, no standard has been established for evaluating such identified areas.

[0008] The present disclosure obtains a posted information extraction control device and a posted information extraction control program capable of acquiring information that is of value to a user by extracting posted information concerning a predetermined theme and having a high level of credibility from out of a large unspecified cluster of posted information posted on a SNS.

Solution to Problem

[0009] A posted information extraction control device of the present disclosure includes a selection means and an extraction means. The selection means is configured to perform natural language analysis on a large cluster of posted information that includes characters and that has been submitted over a network, in order to select from the large unspecified cluster of posted information first posted information that includes character string information indicating a position. The extraction means is configured to certify the first posted information selected by the selection means as being submitted from within a given area based on the character string information, and to extract second posted information including relevant information relating to a target present within the given area.

[0010] In the present disclosure, information with a high level of credibility is extracted from the large cluster of posted information including characters and submitted over the network.

[0011] Namely, the selection means performs natural language analysis on the large unspecified cluster of posted information. The first posted information that includes character string information identifying a position is thus selected from the large unspecified cluster of posted information.

[0012] The extraction means extracts the second posted information from the first posted information selected by the selection means based on the character string information included in the first posted information. The second posted information is information that was submitted from within the given area, and that includes relevant information relating to the target present within the given area.

[0013] The extracted second posted information is posted information that was submitted from within the given area and includes the relevant information relating to the target present within the given area, and may thus be considered to be posted information with a high level of credibility out of the large unspecified cluster of posted information.

[0014] In the present disclosure, in cases in which longitude-latitude information for a submission origin is included in the second posted information extracted by the extraction means, the longitude-latitude information is compared against the character string information indicating a position within the given area in order to determine credibility of the position indicated by the character string information.

[0015] The large unspecified cluster of posted information submitted over the network may include longitude-latitude information for submission origins. For example, if a communication terminal is provided with global positioning system (GPS) functionality and this GPS functionality is in an operational state, such position information (longitude-latitude information) for the submission origin of the posted information may be automatically embedded.

[0016] In cases in which the second posted information extracted by the extraction means includes longitude-latitude information for the submission origin, this longitude-latitude information is compared against the character string information indicating a position within the given area to determine the credibility of the position indicated by the character string information.

[0017] The position of the second posted information can thus be considered to have a higher level of credibility than in cases in which the determination of the position is based solely on the character string information.

[0018] In the present disclosure, the extraction means is configured to split the first posted information into individual word units, and to certify the first posted information as being submitted from within the given area based on respective parts of speech of the individual words.

[0019] Whether the submitter of the posted information submitted the posted information from within the given area (for example when parts of speech such as verbs in the present continuous tense are employed), or whether the submitter of the posted information submitted the posted information after already having moved out of the given area (when parts of speech such as verbs in the past tense are employed), can be determined based on the parts of speech of the individually split words.

[0020] In the present disclosure, the relevant information includes action information indicating an action with respect to the target and evaluation information indicating an

opinion about the target, and a graded evaluation is set based on a sensory expression in the character string information in the second posted information for at least the evaluation information.

[0021] In particular, for example, in cases in which the action information with respect to the target is the information "went to such-and-such a place", the frequency with which this action information of "went to such-and-such a place" appears can be used as grounds for determining whether or not there is an interest in performing this action.

[0022] Moreover, regarding the evaluation information with respect to the target, for example if the target is a food item, applying graded evaluations such as "really delicious", "delicious", "not bad", "disgusting", "really disgusting" can be used as grounds for determining whether or not there is a desire to eat this target.

[0023] The present disclosure further includes a notification means configured to provide notification of guidance information including a density level of the relevant information with respect to the target.

[0024] The notification means notifies the guidance information including a density level of the relevant information with respect to the target, thereby user-friendliness for the user.

[0025] The present disclosure further includes a sorting means configured to link the target with the relevant information relating to the target and sort into plural predetermined categories, and a classification means configured to request a user who uses the guidance information notified using the notification means to respond to a predetermined statement when the user registers, and to analyze the response and classify the user by plural categories. Moreover, the notification means is configured to provide notification with other guidance information corresponding to a category into which the member was placed in the classification.

[0026] The sorting means links the target with the relevant information relating to the target and sorts into the plural predetermined categories.

[0027] The classification means requests the user who uses the guidance information notified using the notification means to respond to the predetermined statement when the user registers, and performs factor analysis on the response in order to classify the member based on the plural categories.

[0028] The notification means is thus capable of notifying of guidance information corresponding to the category into which the user was placed in the classification. Since the categories classify the user according to lifestyle and personality, the notified guidance information is tailored to the user, and can thus be more valuable information than in cases in which notification is made without performing category classification.

In the present disclosure, in cases in which the large unspecified cluster of posted information includes an image, image analysis is performed including at least one out of object detection or image captioning in order to generate character string information describing the image.

The posted information may be text-only, image-only, or a combination of both text and images. In the latter cases, image analysis is performed to generate character string information describing the image to allow handling in a manner similar to that of text.

[0029] A posted information extraction control program according to the present disclosure is configured to cause a computer to operate as a posted information extraction control device.

Advantageous Effects of Invention

[0030] The present disclosure configured as described above is capable of acquiring information that is of value to a user by extracting posted information concerning a predetermined theme and having a high level of credibility from out of a large unspecified cluster of posted information posted on a SNS.

BRIEF DESCRIPTION OF DRAWINGS

[0031] Fig. 1A is a schematic view illustrating an overall system principally configured by a posted information extraction control device of an exemplary embodiment.

Fig. 1B is a block diagram to explain functionality of a posted information extraction control device according to an exemplary embodiment.

Fig. 2 is a functional block diagram illustrating functionality executed by respective sections of a posted information extraction control device according to an exemplary embodiment in block form, as arranged by control content.

Fig. 3 is a control flowchart illustrating a main routine of a posted information extraction control device according to an exemplary embodiment.

Fig. 4 is a flowchart illustrating a registree management control routine executed by a registree management function of a posted information extraction control device according to an exemplary embodiment.

Fig. 5 is a flowchart illustrating a comment analysis control routine executed by a comment analysis function of a posted information extraction control device according to an exemplary embodiment.

Fig. 6 is a flowchart illustrating an information processing control routine executed by an information provision function of a posted information extraction control device according to an exemplary embodiment.

Fig. 7 is a front view illustrating clarified survey information content stored in a

survey information storage section.

Fig. 8A is a front view illustrating a mobile terminal in the possession of a user.

Fig. 8B is a front view illustrating survey item screens sequentially displayed on a mobile terminal in the possession of a user.

Fig. 9 is a table illustrating content of survey responses to be selected by a user in a first stage of factor analysis according to a first example of an exemplary embodiment.

Fig. 10 is a table illustrating content from which items with a low level of ultimate commonality have been excluded in a second stage of factor analysis according to the first example of an exemplary embodiment.

Fig. 11 is a table illustrating a computation result of factor loading for respective factor types in a third stage of factor analysis according to the first example of an exemplary embodiment.

Fig. 12 is a correspondence table between factor types notified to survey subjects and names of the factor types in a fourth stage of image analysis according to the first example of an exemplary embodiment.

Fig. 13A is a table illustrating a list of comments in which origin position information and relevant information (evaluation information and action information) are present in a character string, the comments having been extracted from a large quantity of SNS comments by text mining processing according to a second example of an exemplary embodiment.

Fig. 13B is a table illustrating evaluation criteria for evaluation tags allocated to the extracted comments in Fig. 13A.

Fig. 14A is a front view illustrating a map image displayed on a monitor of a mobile terminal in the possession of a user, according to a third example of an exemplary embodiment.

Fig. 14B is a front view illustrating a social heat map displayed superimposed on a map image on a monitor of a mobile terminal in the possession of a user.

Fig. 15 is a block diagram to explain functionality of a comment extraction control device according to a modified example of an exemplary embodiment.

Fig. 16 is a control flowchart illustrating a main routine of a comment extraction control device according to a modified example of an exemplary embodiment.

Fig. 17 illustrates an image analysis processing control routine executed at step 122 in Fig. 16.

Fig. 18A is a front view of an image, illustrating an example of image captioning performed during image analysis.

Fig. 18B is a front view illustrating a state in which distinctive images have been extracted from a source image.

Fig. 18C is a front view illustrating image analysis text generated based on extracted distinctive images.

DESCRIPTION OF EMBODIMENTS

[0032] Fig. 1A is a schematic view of a network system principally configured by a posted information extraction control device 10 according to an exemplary embodiment. Note that the posted information of the present disclosure refers to comments (either text information or image information) posted on a SNS. Thus, posted information and comments are taken to have the same meaning.

[0033] An SNS posted comment server 14 is connected to a network 12 (such as the internet). Comments posted using plural mobile terminals 18 (such as smartphones) are accumulated in the SNS posted comment server 14 via a wireless communication device 16. Note that although comments posted using PCs and the like that are wirelessly connected to the network 12 independently of the wireless communication device 16 are not excluded, the present exemplary embodiment concerns comments relating to a predetermined theme that are submitted from a given area (location) when visiting this given area, and so explanation primarily focuses on submissions from the mobile terminals 18.

[0034] The posted information extraction control device 10 is connected to the network 12.

[0035] The posted information extraction control device 10 includes a microcomputer 20. The microcomputer 20 includes a CPU 20A, RAM 20B, ROM 20C, an input/output port (I/O) 20D, and a bus 20E such as a data bus or a control bus that connects these elements together.

[0036] A storage device 22 and a communication I/F 24 are connected to the I/O 20D.

[0037] The storage device 22 is flash memory or the like, typically in the form of a hard disc, USB memory, or SD memory, and may also be utilized as a supplementary storage space of the RAM 20B and the ROM 20C. In the posted information extraction control device 10, a comment extraction control program or the like is pre-stored in a recording region of the RAM 20B, the ROM 20C, or the storage device 22, and this program is read when appropriate to the processing situation of the CPU 20A.

[0038] Fig. 1B is a block diagram illustrating operation performed when the comment extraction control program is executed by the posted information extraction control device 10, with operation being classified according to functionality.

[0039] Namely, the posted information extraction control device 10 operates as a registree management function 26, a comment analysis function 28, and an information provision

function 30.

[0040] The registree management function 26 principally includes receiving member (user) registration, carrying out a survey (see Fig. 7, Fig. 8, and Fig. 9) performed on registration, classifying registered users into categories (see Fig. 10, Fig. 11, and Fig. 12), and learning the classified categories.

[0041] The comment analysis function 28 principally includes text mining comments acquired from the SNS posted comment server 14, determining a post origin position, identified by the text mining, and the level of credibility thereof, and evaluating targets relating to the theme (see Fig. 13). The targets include restaurants, shops, scenic spots, tourist spots, and the like. The target evaluation refers to a description of the user's feelings, such as "delicious", "disgusting", "expensive", "cheap", "beautiful", or "dirty" (collectively referred to as sensory expressions). Note that action information relating to the target, such as "went" or "saw", may also be included, and that this evaluation information and action information is collectively referred to as relevant information.

[0042] The information provision function 30 includes generating and editing social heat maps (see Figs. 14) provided to users based on the results of comment analysis, as well as managing user categories and so on. A social heat map refers to a map that is superimposed on a normally displayed map image, in which regions are displayed with different intensity or color in order to emphasize locations with a large quantity of information applicable to a user's category.

[0043] Detailed Description of Posted Information Extraction Control Device 10

[0044] Fig. 2 is a functional block diagram illustrating respective processing performed when the registree management function 26, the comment analysis function 28, and the information provision function 30 are executed by the posted information extraction control device 10 in block form, arranged by functionality. Note that the respective blocks aim to clarify the respective processing, and so are not limited to hardware configuration of the posted information extraction control device 10. Thus, some or all of the processing may be executed by what is referred to as software processing based on the comment extraction control program.

[0045] Registree Management Function 26

[0046] As illustrated in Fig. 2, a registration processing section 32 is connected to the network 12. The registration processing section 32 receives new registration and survey response information from mobile terminals 18 in the possession of users.

[0047] The registration processing section 32 is connected to a transmission section 34 and a survey analysis section 36.

[0048] Note that in cases in which the information received by the registration processing section 32 relates to a new registration, the transmission section 34 is instructed to transmit survey information.

[0049] A survey information storage section 38 is connected to the transmission section 34. As illustrated in the example in Fig. 7, survey information is stored in the survey information storage section 38.

[0050] On receiving the instruction from the registration processing section 32, the transmission section 34 reads the survey information from the survey information storage section 38 and transmits this information to the mobile terminal 18 of the user who wishes to register. Note that when transmitting the survey information, the transmission section 34 also requests that the user respond to the survey and requests that the user submit an initial posted comment.

[0051] The survey is required in order to classify the user into a category (corresponding to the name of a factor as classified by factor analysis, as illustrated in the example in Fig. 12) on registration of the user.

[0052] The initial posted comment is required at least in order to increase the number of comments by the user in which a current position and an evaluation of a target at their current position are linked together. Another potential aim is to make the user aware that comments in which a current position and an evaluation are linked together are required.

[0053] In cases in which the information received by the registration processing section 32 is survey response information, the survey analysis section 36 is instructed to perform survey analysis. The survey analysis section 36 is connected to a registree category classification section 40.

[0054] The registree category classification section 40 uses the transmission section 34 to notify the user who has requested registration of their classification category (factor type), associates the registration information of the registree with the classified factor type, and registers this in a registree database 42.

[0055] The registree database 42 is connected to an AI learning processing section 44. The AI learning processing section 44 monitors changes in the preferences of registrees based on the level of credibility of comments and the contributors who post the comments on the SNS, described later, and edits (changes, adds to, deletes, or the like) the categories (factor types) of the respective registrees registered in the registree database 42.

[0056] Comment Analysis Function 28

[0057] As illustrated in Fig. 2, the comment analysis function 28 is provided with a posted comment acquisition section 46 that acquires SNS posts posted on the network 12 on a

random basis. Namely, the position of submission, level of credibility, and so on of comments are not taken into consideration when these comments are acquired.

[0058] The posted comment acquisition section 46 is connected to a contributor identification section 48, a text mining processing section 50, and a longitude-latitude information extraction section 52.

[0059] The contributor identification section 48 identifies the contributor based on the address from which the comment was submitted and so on. The identified contributor information is sent to the AI learning processing section 44 of the registree management function 26 in order to compare the contributor against registered users.

[0060] The text mining processing section 50 performs machine learning to split writing in a posted comment (i.e. written data) into individual words (such as nouns, verbs, and adjectives) in order to pick out exploitable information. A corpus 51 containing a large collection of natural language may be utilized as learning data when performing the machine learning.

[0061] The text mining processing section 50 is connected to a post origin position identification section 54 and a relevant information identification section 56.

[0062] The post origin position identification section 54 analyzes character strings rendered during the text mining processing for a character string identifying the post origin position. In cases in which a character string identifying the post origin position is present, this is sent to a linking section 58. In other words, comments in which a post origin position is not identified are eliminated at this point.

[0063] The relevant information identification section 56 identifies relevant information (namely action-related information and evaluation-related information) from the character string rendered during the text mining processing. The relevant information thus identified is sent to a relevant information evaluation section 60, which allocates the relevant information with an evaluation score. As illustrated in the example in Fig. 13B, the relevant information is allocated one out of five set graded evaluation tags (scores) (A, B, C, D, and E).

[0064] After being allocated an evaluation tag, the relevant information is sent to the linking section 58 where it is linked with the post origin position information sent from the post origin position identification section 54 (extracted comment).

[0065] This extracted comment is sent to a credibility determination section 62.

[0066] Note that the longitude-latitude information extracted by the longitude-latitude information extraction section 52 is information that can be acquired when a GPS function of the mobile terminal 18 in the possession of the user has been started up. The position of the

mobile terminal 18 (namely, the current position of the user) can be accurately identified using this longitude-latitude information.

[0067] Having extracted the longitude-latitude information, the longitude-latitude information extraction section 52 sends this longitude-latitude information to the credibility determination section 62.

[0068] On receiving the longitude-latitude information from the longitude-latitude information extraction section 52, the credibility determination section 62 compares this information against the post origin position information in the character string in the extracted comment. The level of credibility of the extracted comment is thereby determined, and extracted comments with a high level of credibility are held in a comment database 64.

[0069] In cases in which longitude-latitude information is not received from the longitude-latitude information extraction section 52, the extracted comment received from the linking section 58 is held in the comment database 64 without making such a comparison.

[0070] Note that the determination result of the credibility determination section 62 is sent to the contributor of the extracted comment, and to the AI learning processing section 44 of the registree management function 26. In cases in which the contributor is a registered user, the posting frequency and level of credibility of this user are associated and learnt, and a determination is made as to whether or not there is a need to edit the category (factor type) of this registered user.

[0071] Information Provision Function 30

[0072] An aim of the information provision function 30 is to provide a user (registree) with a social heat map generated based on extracted comments.

[0073] As illustrated in Fig. 2, an authentication section 66, a displayed map information procurement section 68, and a social heat map data transmission section 70 are connected to the network 12.

[0074] When the user operates their mobile terminal 18 to request a social heat map, this request is received by the authentication section 66. The authentication section 66 searches the registree database 42 based on discrimination information and so on contained in the request item to authenticate whether or not the user is registered.

[0075] The authentication section 66 is connected to a category identification section 72 and the displayed map information procurement section 68.

[0076] The category identification section 72 accesses the registree database 42 to identify the category of the user (registree) who has requested the social heat map, and sends this to a comment data reading section 74.

[0077] The displayed map information procurement section 68 also accesses the mobile

terminal 18 in the possession of the user, procures map information that is currently being displayed, and sends this to the comment data reading section 74.

[0078] The comment data reading section 74 reads comment data having origin position information from within the range of the map procured from the mobile terminal 18 of the user from the comment database 64, and sends this to a mapping processing section 76.

[0079] The mapping processing section 76 maps a heat map corresponding to the comment data onto the procured map, and uses the social heat map data transmission section 70 to send this to the mobile terminal 18 in the possession of the user. Thus, a social heat map is displayed on a monitor 18M of the mobile terminal 18 as illustrated in Fig. 14B.

[0080] Explanation follows regarding operation of the present exemplary embodiment, with reference to the flowcharts in Fig. 3 to Fig. 6.

[0081] Fig. 3 is a control flowchart illustrating a main routine of the posted information extraction control device 10 according to the present exemplary embodiment.

[0082] At step 100, determination is made as to whether or not information has been received. In cases in which determination is negative, the present routine is ended.

[0083] In cases in which determination is affirmative at step 100, processing transitions to step 102. The received information is categorized, and processing transitions to step 104.

[0084] At step 104, the processing branches according to the type of information received. Namely, in cases in which the type of information received at step 104 is determined to be registration-related, processing transitions to step 106, and registree management control (illustrated in detail in Fig. 4) is executed before ending the present routine.

[0085] In cases in which the type of information received at step 104 is determined to be a posted comment, processing transitions to step 108, and comment analysis control (illustrated in detail in Fig. 5) is executed before ending the present routine.

[0086] In cases in which the type of information received at step 104 is determined to be a social heat map request, processing transitions to step 110, and information provision control (illustrated in detail in Fig. 6) is executed before ending the present routine.

[0087] Registree Management Control

[0088] Fig. 4 is a flowchart illustrating a registree management control subroutine performed at step 106 in Fig. 3.

[0089] At step 150, the type of information received is discerned. In cases in which the type of information received is discerned at step 150 to be a new registration, processing transitions to step 152 and the survey information is read from the survey information storage section 38. Processing then transitions to step 154.

[0090] At step 154, the survey information and a response request is transmitted to the

mobile terminal 18 in the possession of the user. Next, at step 156, the user is requested to post an initial comment, and processing transitions to step 164. Note that an initial posting is not essential.

[0091] In cases in which the type of information received is discerned to be survey response information at step 150, processing transitions to step 158 and the survey response is analyzed (for example by factor analysis).

[0092] Next, at step 160, the registree is classified into a category (see Fig. 12), after which processing transitions to step 162. This category is associated with categorization information for the registree and held in the registree database, after which processing transitions to step 164.

[0093] At step 164, determination is made as to whether or not a post has been received from the registree. In cases in which determination is affirmative, processing transitions to step 166, the category of the registree is learnt based on relevant information posted by the registree, and processing returns to the main routine. In cases in which determination is negative at step 164, processing also returns to the main routine.

[0094] Comment Analysis Control

[0095] Fig. 5 is a flowchart illustrating a comment analysis control subroutine performed at step 108 in Fig. 3.

[0096] At step 200, the contributor is identified, and processing transitions to step 202, at which determination is made as to whether or not longitude-latitude information is included in the posted comment. In cases in which longitude-latitude information is determined to be included at step 202, processing transitions to step 204. Origin position information based on the longitude and latitude is temporarily held, and processing transitions to step 206. In cases in which longitude-latitude information is determined not to be included at step 202, processing transitions to step 206.

[0097] At step 206, text mining processing is executed on the posted comment based on the language stored in the corpus 51. Namely, machine learning is performed to split the writing in the posted comment (i.e. written data) into individual words (such as nouns, verbs, and adjectives) in order to pick out exploitable information. The corpus 51 containing a large collection of natural language is utilized as learning data when performing the machine learning.

[0098] Next, at step 208, determination is made as to whether or not there is a character string enabling the post origin position to be identified in the character strings obtained during the text mining processing.

[0099] In cases in which determination is negative at step 208, determination is made that

this comment is to be excluded (the comment is not worth using), and processing returns to the main routine.

[0100] In cases in which determination is affirmative at step 208, determination is made that the comment is worth using. Processing transitions to step 210, and the relevant information is identified. The relevant information is a collective term for evaluation information (terms such as "delicious" or "beautiful") and action information (terms such as "went" or "ate").

[0101] Next, at step 212, an evaluation tag is allocated to the identified relevant information (see the five types of evaluation tag in the example illustrated in Fig. 13B). Processing then transitions to step 214, at which the post origin position and the relevant information (evaluation tag) are linked together.

[0102] Next, at step 216, the temporarily held origin position information based on the longitude and latitude is read. Processing then transitions to step 218, at which the origin position information based on the longitude-latitude information is compared against the origin position as identified from the character string. Processing then transitions to step 220.

[0103] At step 220, determination is made as to whether or not the comparison result is a match. In cases in which determination is affirmative at step 220, processing transitions to step 222, and the posted comment (i.e. the linked post origin position and relevant information) is held in the comment database 64. Processing then returns to the main routine. In cases in which determination is negative at step 220, determination is made that this comment is to be excluded (i.e. the comment is not worth using), and processing returns to the main routine.

[0104] Note that configuration may be made such that comments that do not include longitude-latitude information are all held in the comment database 64, or selection/rejection may be performed by another method (for example in cases in which the origin position information in the character string is an address or the like, and can be determined to be unique).

[0105] Information Provision Control

[0106] Fig. 6 is a flowchart illustrating an information provision control subroutine performed at step 110 in Fig. 3.

[0107] At step 250, the registree database 42 is accessed, the member (registered user) is authenticated, and processing transitions to step 252.

[0108] At step 252, determination is made as to whether or not authentication has been successful. In cases in which determination is negative, processing returns to the main routine. Note that the authentication processing may be attempted a preset number of times.

[0109] In cases in which determination is affirmative at step 252, processing transitions to step 254, at which the registree database 42 is accessed, and the category of the member (registered user) is identified. Processing then transitions to step 256.

[0110] At step 256, the comment database 64 is accessed, and comment data is read for the corresponding category. Next, processing transitions to step 258, a map range being displayed for the member (registered user) is identified, and processing transitions to step 260.

[0111] At step 260, mapping processing is executed on the map based on the comment data to create a social heat map, and processing transitions to step 262.

[0112] At step 262, the social heat map data is transmitted to the member (registered user) who made access, and processing returns to the main routine.

Examples

[0113] Explanation follows regarding examples of the main processing executed by the respective functionality of the posted information extraction control device 10 (i.e. the registree management function 26, the comment analysis function 28, and the information provision function 30).

[0114] First Example

[0115] A first example is an example of a flow of processing by the registree management function 26, from a survey response request to factor analysis.

[0116] When a user requests registration, survey information is transmitted from the transmission section 34 in Fig. 2.

[0117] As illustrated in Fig. 8A, thirty survey statements (see Fig. 7) are sequentially displayed on the monitor 18M of the mobile terminal 18 in the possession of the user.

[0118] The survey is performed by making a selection from out of five responses. The five response options to each statement are "strongly disagree", "disagree", "neither agree nor disagree", "agree", and "strongly agree". The user makes a selection (by touch operation) based on their own judgement. Note that when the user touch-operates the "NEXT" button, the statement is changed (see Fig. 8B), and the user repeats the response selection for the next statement.

[0119] When the user has finished responding to the survey and transmits the survey response information, the survey analysis section 36 and the registree category classification section 40 in Fig. 2 execute factor analysis processing.

[0120] As illustrated in Fig. 9, a score is allocated to each of the survey response options. In this example, "strongly disagree" is set to -2 points, "disagree" is set to -1 point, "neither agree nor disagree" is set to 0 points, "agree" is set to +1 point, and "strongly agree" is set to +2 points, and these points are accumulated (a first stage of factor analysis). An item name

is set in advance for each statement (see Fig. 10). These item names are an important element in classifying the factor type.

[0121] As illustrated in Fig. 10, in a second stage of the factor analysis, any statements with a low level of ultimate commonality are excluded from the survey responses during the factor analysis process. In this example, statement numbers 13, 15, 25, and 30 have been excluded.

[0122] Fig. 11 illustrates factor analysis results for each set item name. In this example, factor loading has been computed for each of six factor types. These results are used to identify the factor type of the survey subject (a third stage). This identification is for example executed by performing comparisons against preset average values.

[0123] As illustrated in Fig. 12, names are allocated to each of the six factor types.

[0124] For example, the first factor type corresponds to a believer in a healthy and sustainable lifestyle, the second factor type corresponds to a selective person, the third factor type corresponds to a market watcher, the fourth factor is a passive person, the fifth factor type corresponds to a convenience-focused person, and the sixth factor is a self-bettering person. Each factor type is preferably allocated a name that instantly brings to mind the category to which it refers. Note that there is no limitation to six factor types. Moreover, there is no particular limitation to the names of the factor types, as long as each name makes the nature of the corresponding factor immediately apparent.

[0125] The regtree category classification section 40 uses the transmission section 34 to notify the survey subject of the name of their factor type. Note that the notified factor type is not necessarily a single type, and a notification such as "You have been classified into two categories: a believer in a healthy and sustainable lifestyle and a passive person" may be performed.

[0126] In the first example, factor analysis of the survey responses is performed using a principal factor method. However, other analysis methods may be employed as long as category classification can be performed.

[0127] Second Example

[0128] A second example is an example illustrating a flow of text mining processing by the comment analysis function 28.

[0129] Fig. 13A illustrates a result of subjecting a large quantity of SNS-posted comments acquired by the posted comment acquisition section 46 to text mining processing in order to extract comments (i.e. extracted comments) with character strings including both origin position information and relevant information (evaluation information and action information).

[0130] For example, the posted content in the first row in Fig. 13A is a comment including

"really delicious... department store at XXX station", which includes the evaluation information "really delicious" and the origin position information "department store at XXX station".

[0131] Fig. 13B illustrates criteria employed when allocating evaluation tags to evaluation information in character strings from extracted comments.

[0132] There are five grades of evaluation tag, these being A, B, C, D, and E.

[0133] The evaluation tag A is allocated to extracted comments that include both a descriptor and a modifier, such as "really delicious" or "the most beautiful ever", and that correspond to a positive evaluation. The extracted comment in the first row has the evaluation tag A.

[0134] The evaluation tag B is allocated to extracted comments that include a descriptor, such as "delicious" or "beautiful", and that correspond to a positive evaluation.

[0135] The evaluation tag C is allocated to extracted comments that are not evaluation comments, but do include action information that reports on a situation, such as "went" or "bought".

[0136] The evaluation tag D is allocated to extracted comments that include a descriptor, such as "disgusting" or "dirty", and that correspond to a negative evaluation.

[0137] The evaluation tag E is allocated to extracted comments that include both a descriptor and a modifier, such as "absolutely disgusting" or "really dirty", and that correspond to a negative evaluation.

[0138] The evaluation tags may be based on numbers (points) instead of letters.

[0139] An evaluation tag is allocated to each of the extracted comments in Fig. 13A based on the evaluation criteria in Fig. 13B, and the extracted comments are held in the comment database 64 in this evaluation tag-allocated state. Note that the longitude-latitude information is what is referred to as stealth data, and so the presence or absence of longitude-latitude information is not recorded in Fig. 13A. However, in cases in which longitude-latitude information is present, this information may be compared against origin position information identified in the text mining processing so as to determine whether or not to hold this information in the comment database 64.

[0140] Third Example

[0141] A third example is an example illustrating a display format of a social heat map by the information provision function 30.

[0142] Fig. 14A is a front view illustrating a mobile terminal 18 in the possession of a user who has requested a social heat map.

[0143] For example, by starting up a map application (including a GPS function) installed in the mobile terminal 18, a map image 18A of the surroundings of the current position is

displayed at a preset scale, and a marker 18B indicating the current position is also displayed on the monitor 18M.

[0144] When, for example, a social heat map is requested using a different window while in this state, regions 18C with different intensities or colors are displayed superimposed on the map image 18A that is currently being displayed. Namely, the regions 18C are displayed at locations where there is a large quantity of information applicable to the user's category.

[0145] In this case, the more applicable an object is to the relevant category, the greater the intensity of display. Note that with the intensity, the display color may be changed from red, through yellow, and then to green in a sequence corresponding to the sequence from high intensity to low intensity.

[0146] As illustrated in Fig. 14B, in cases in which a location or the like that is particularly recommended for the user's category is present, a special marker 18D (a star in this case) may be displayed.

[0147] As previously described, the present exemplary embodiment may be utilized for the purpose of urban development, in order to ascertain the potential of a particular area using a social heat map and select candidates for future business development, or to select locations to introduce measures to boost existing redevelopment districts.

[0148] Alternatively, an architectural designer may utilize the present exemplary embodiment to ascertain the potential of a project site using a social heat map, and to reflect this in a traffic flow plan or in building exterior planning in order to maximize the impact in a particular area.

[0149] Alternatively, an architectural designer may utilize the present exemplary embodiment to ascertain the potential of a project site using a social heat map, and to reflect this in a traffic flow plan or in building exterior planning in order to minimize the effect on a particular area.

[0150] Alternatively, an information service may develop an application enabling deployment of a social heat map of a particular area on mobile terminals 18 (smartphones and the like) in order to encourage members to explore this particular area. So doing may provide indirect health benefits.

Modified Example

In the present exemplary embodiment, text mining developed specifically for text is performed on the posted information (comment) to identify the post origin position, determine the level of credibility, evaluate targets, and so on. However, the posted information may include an image, or may be configured solely of an image.

Although some images are not relevant, as a modified example, an image may be

analyzed to generate text describing the image, which may then be handled similarly to a comment in the present exemplary embodiment.

Explanation follows regarding a modified example of the present exemplary embodiment, with reference to Fig. 15 to Fig. 17. Note that configuration portions that are the same as those in the present exemplary embodiment are appended with the same reference numerals, and explanation of the configurations thereof is omitted.

Fig. 15 is a block diagram to explain functionality of a comment extraction control device according to the modified example of the present exemplary embodiment.

As illustrated in Fig. 15, an image analysis function 29 is connected to the comment analysis function 28.

The image analysis function 29 primarily executes object detection and conversion to text using image captioning.

In the object detection, an image (region) of an object with a distinguishing feature is extracted from the image, and the extracted image is categorized (vehicle, animal, building, etc.) to generate character-based information.

Image captioning is used to express a situation in (the content of) an image as a character string.

Namely, the object detection and image captioning both correspond to functionality to convert the image to text, and a character string (comment) created by the text conversion is analyzed by the comment analysis function.

Fig. 16 is a control flowchart illustrating a main routine of the comment extraction control device according to the modified example of the present exemplary embodiment. Note that processing that is the same as that in the main routine of the comment extraction control device of the present exemplary embodiment is suffixed with the reference numeral A.

At step 100A, determination is made as to whether or not information has been received. In cases in which determination is negative, the present routine is ended.

In cases in which determination is affirmative at step 100A, processing transitions to step 102A. The received information is categorized, and processing transitions to step 104A.

At step 104A, the processing branches according to the type of information received. Namely, in cases in which the type of information received at step 104A is determined to be registration-related, processing transitions to step 106A, and registree management control (illustrated in detail in Fig. 4) is executed before ending the present routine.

In cases in which the type of information received at step 104A is determined to be a posted comment, processing transitions to step 120, and determination is made as to whether or not the posted information includes an image.

In cases in which determination is affirmative at step 120, processing transitions to step 122, image analysis processing is executed, and processing transitions to step 108A.

At step 108A, comment analysis control (illustrated in detail in Fig. 5) is executed, and the present routine is ended.

In cases in which the type of information received at step 104A is determined to be a social heat map request, processing transitions to step 110A, and information provision control (illustrated in detail in Fig. 6) is executed before ending the present routine.

Fig. 17 illustrates an image analysis processing control routine executed at step 122 in Fig. 16.

At step 124, image analysis (i.e. at least one out of object detection or image captioning) is executed, after which processing transitions to step 126, and text is generated based on the image analysis. Note that the text generated based on the image is referred to as image analysis text in order to differentiate from text that is present initially.

Next, at step 128, determination is made as to whether the posted information is an image with text, or just an image.

In cases in which the posted information is determined to be an image with text at step 128, processing transitions to step 130. The image analysis text is merged with the initial text, and then processing transitions to step 132. The merged texts are sent to the comment analysis function 28 as a comment, the present routine is ended, and processing transitions to step 108A in Fig. 16.

In cases in which the posted information is determined to be just an image at step 128, processing transitions to step 132. The image analysis text is sent to the comment analysis function 28 as a comment, the present routine is ended, and processing transitions to step 108A in Fig. 16.

Figs. 18 illustrate an example of an image caption created during image analysis.

Fig. 18A illustrates a source image 80. As illustrated in Fig. 18B, distinctive images are extracted from the source image 80. In this example, an image of a cat (within a dotted line frame 80A) and an image of a ball (within a single-dotted dashed line frame 80B) have been extracted.

As illustrated in Fig. 18C, image analysis text is generated based on the extracted distinctive images. In this example, the image analysis text "A cat laying on the chair with a ball" is generated.

[0151] The entire content of the disclosure of Japanese Patent Application No. 2018-116210 filed on June 19, 2018 is incorporated by reference in the present specification. All cited documents, patent applications, and technical standards mentioned in the present

specification are incorporated by reference in the present specification to the same extent as if each individual cited document, patent application, or technical standard was specifically and individually indicated to be incorporated by reference.

CLAIMS

1. A posted information extraction control device comprising:

a selection means configured to perform natural language analysis on a large unspecified cluster of posted information that includes at least one of characters or an image and that has been submitted over a network, in order to select, from the large unspecified cluster of posted information, first posted information that includes character string information indicating a position; and

an extraction means configured to certify the first posted information selected by the selection means as being submitted from within a given area based on the character string information, and to extract second posted information including relevant information relating to a target present within the given area.

2. The posted information extraction control device of claim 1, wherein:

in cases in which longitude-latitude information for a submission origin is included in the second posted information extracted by the extraction means, the longitude-latitude information is compared against the character string information indicating a position within the given area in order to determine credibility of the position indicated by the character string information.

3. The posted information extraction control device of claim 1 or claim 2, wherein:

the extraction means is configured to split the first posted information into individual word units, and to certify the first posted information as being submitted from within the given area based on respective parts of speech of the individual words.

4. The posted information extraction control device of any one of claim 1 to claim 3, wherein:

the relevant information includes action information indicating an action with respect to the target and evaluation information indicating an opinion about the target; and

a graded evaluation is set based on a sensory expression in the character string information in the second posted information for at least the evaluation information.

5. The posted information extraction control device of any one of claim 1 to claim 4, further comprising:

a notification means configured to provide notification of guidance information

including a density level of the relevant information with respect to the target.

6. The posted information extraction control device of claim 5, further comprising:

a sorting means configured to link the target with the relevant information relating to the target and sort into a plurality of predetermined categories; and

a classification means configured to request a user who uses the guidance information notified using the notification means to respond to a predetermined statement when the user registers, and to analyze the response and classify the user by a plurality of categories;

the notification means being configured to provide notification with other guidance information corresponding to a category into which the user was placed in the classification.

7. The posted information extraction control device of any one of claim 1 to claim 6, wherein in cases in which the large unspecified cluster of posted information includes an image, image analysis is performed including at least one of object detection or image captioning, in order to generate character string information describing the image.

8. A posted information extraction control program configured to cause a computer to operate as the posted information extraction control device of any one of claim 1 to claim 7.