METHOD AND APPARATUS FOR CATEGORY-BASED CLUSTERING USING PHOTOGRAPHIC REGION TEMPLATES OF DIGITAL PHOTO

A clustering method and apparatus using region division templates are provided. The method includes: dividing a photo into regions by using region division templates; modeling a semantic concept included in a divided region; when it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, merging the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling; modeling a global semantic concept included in the photo by using a final local semantic concept determined after the merging; and determining one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the modeling. According to the method and apparatus, in order to more reliably extract semantic concepts included in a photo, multiple content-based feature values are extracted from region images divided by using region division templates, and the confidence degree of an input image in relation to the local semantic concept defined by using the feature values is measured. With respect to the confidence degree, the local semantic concepts of the photo are merged and a more reliable local semantic concept is extracted. By using the merged local semantic concept, the confidence degree of a global semantic concept is measured, and according to the confidence, multiple category concepts included in the input photo are extracted. By doing so, photo data can be quickly and effectively used to generate an album.
For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
Description

METHOD AND APPARATUS FOR CATEGORY-BASED CLUSTERING USING PHOTOGRAPHIC REGION TEMPLATES OF DIGITAL PHOTO

Technical Field

The present invention relates to a digital photo album, and more particularly, to a category-based photo clustering method and apparatus using region division templates.

Background Art

An ordinary digital photo album is used to transfer photos taken by a user from a digital camera or a memory card to a local storage apparatus of the user and to manage the photos in a computer. Generally, by using the photo album, users want to index many photos in a time series or in photo categories arbitrarily made by the users and browse the photos according to the index, or share the photos with other users.

In the process, a function for automatically clustering photos based on the categories included in the photos is one of major functions of a photo album. The categorization reduces the range of searching in a process of retrieving a photo desired by a user, such that the accuracy of the searching as well as the searching speed can be improved. Furthermore, by automatically classifying photos into categories desired by the user, management by the user of a large volume of photos in an album is made to be easier.

Most of the conventional categorization methods are based on text, and by using meta data specified one by one in text by a user, photos are categorized. However, the text-based method is not useful in that if the amount of photos is large, it is almost impossible for a user to specify all category information of the photos one by one, and text information is not effective in describing the semantic concepts of photos. Accordingly, a method of categorizing multimedia contents, by using content-based features, such as colors, shapes, and texture, extracted based on the contents of photos can be suggested.

To date, in order to cluster photos by using content-based features of photo images many researches have been performed. However, since a variety of semantic concepts are included in photos, automatic extraction of multiple semantic concepts is still a very difficult subject of research. As a means to solve this problem, researches of extracting major objects in a photo (image) and according to the semantic concepts of the objects, indexing or categorizing photos have been performed. Since extracting a variety of semantic concepts included in a photo is very difficult, only major semantic concepts included in the photo are extracted in this method.
Among them, the subject of the researches that have been performed mainly is extracting main subjects among semantic objects included in a photo and identifying and indexing the object, as in a method for automatic determination of main subjects in photographic images (Eastman Kodak Company) by Jiebo Luo. That is, in the categorizing of photos, research on segmentation of objects included in a photo and research on indexing or categorizing the segmented object have been performed.

However, as described above, in most cases a lot of semantic concepts are included in a photo image such that the categorization by extraction of the main subject results in loss of the other semantic concepts.

**Disclosure of Invention**

**Technical Problem**

Generally, a photo can be divided into a foreground and a background. In categorization of photo data, the semantic concept included in the foreground is important but the semantic concept included in the background is also important.

Accordingly, as a method of categorizing photo data, a method of extracting a variety of semantic concepts included in a photo by considering both the concepts of the foreground and the background, rather than the method of segmenting objects, is needed.

As a method of extracting a variety of semantic concepts from a photo, there is a method of dividing an image into smaller regions and extracting a semantic concept from each divided region. Division of an image into smaller regions has an advantage of easiness to extract a single semantic concept, but if the area of the divided image is too small, it may become difficult to extract even a single semantic concept. That is, it is not easy to determine a size by which an image is to be divided. Accordingly, an effective method of dividing an image to extract a variety of semantic concepts of a photo and a method of extracting an accurate semantic concept from a divided image are necessarily needed.

**Technical Solution**

The present invention provides a category-based clustering method and system of a digital photo album capable of extracting a variety of semantic concepts included in a photo based on content-based features of a photo and automatically classifying photos into a variety of categories.

The present invention also provides a category-based clustering method and apparatus using region division templates by which photo data is effectively divided into regions, and the semantic concept of each of the divided region is extracted, and through efficient merging of local semantic concepts in order to find the global meaning of the photo, the semantic concept included in the photo is categorized.
According to an aspect of the present invention, there is provided a clustering method of a digital photo album using region division templates, the method including: dividing a photo into regions by using region division templates; modeling a semantic concept included in a divided region; when it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, merging the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling; modeling a global semantic concept included in the photo by using a final local semantic concept determined after the merging; and determining one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the modeling.

The region division templates in the modeling of the semantic concept may be expressed by the following equations:

\[
\mathcal{T}(1) = \left[ \frac{w}{4} , \frac{h}{4} , \frac{3w}{4} , \frac{3h}{4} \right], \quad \mathcal{T}(2) = \left[ 0,0, \frac{w}{2} , \frac{h}{2} \right], \quad \mathcal{T}(3) = \left[ \frac{w}{2} , 0, \frac{w}{2} , \frac{h}{2} \right], \quad \mathcal{T}(5) = \left[ \frac{w}{2} , \frac{h}{2} , \frac{w}{2} , \frac{h}{2} \right],
\]

\[
\mathcal{T}(6) = \left[ 0,0, \frac{h}{2} \right], \quad \mathcal{T}(7) = \left[ \frac{w}{2} , 0, \frac{h}{2} , \frac{h}{2} \right], \quad \mathcal{T}(8) = \left[ 0,0, \frac{w}{2} , \frac{h}{2} \right], \quad \mathcal{T}(9) = \left[ \frac{w}{2} , 0, \frac{w}{2} , \frac{h}{2} \right]
\]

\[
\mathcal{T}(10) = \left\{ \emptyset, \emptyset, \frac{w}{2} , \frac{h}{2} \right\}
\]

where \( T \) is the template of a photo, \( w \) is the length of the width of the photo, and \( h \) is the length of the height of the photo.

In the modeling of the semantic concept, the semantic concept may be modeled by extracting content-based feature values of the photo. The content-based feature values may include color, texture, and shape information of an image. In the modeling of the semantic concept the semantic concept may be composed of an item (L entity) indicating the entity of a semantic concept included in a photo and an item (Lattribute) indicating the attribute of the entity of the semantic concept. The semantic concept modeling may be modeling of the entity concept and the attribute concept of a divided region.

In the modeling of the semantic concept, modeling of the local concepts of the input photo in which regions are divided may be performed by using a support vector machine (SVM). In the merging of the semantic concepts of respective regions, the confidence degree of each local semantic concept may be measured by using one SVM for each defined local semantic concept.
In the merging of the semantic concepts of respective regions, based on the confidence degrees of local concepts allocated to 10 regions divided by using the region division templates, local concept confidence degrees of 5 basic regions may be merged according to the following equation:

\[ C'_{L}(R(1)) = \max \left\{ C_L(T) \mid T \in \{ T(1), T(10) \} \right\}, \]

\[ C'_{L}(R(2)) = \max \left\{ C_L(T) \mid T \in \{ T(2), T(6), T(8), T(10) \} \right\}, \]

\[ C'_{L}(R(3)) = \max \left\{ C_L(T) \mid T \in \{ T(3), T(6), T(9), T(10) \} \right\}, \]

\[ C'_{L}(R(4)) = \max \left\{ C_L(T) \mid T \in \{ T(4), T(7), T(8), T(10) \} \right\}, \]

\[ C'_{L}(R(5)) = \max \left\{ C_L(T) \mid T \in \{ T(5), T(7), T(9), T(10) \} \right\}. \]

where T(1), T(2), T(3), T(4), and T(5) indicate basic regions to which final local semantic concepts are allocated, and \( C' \) is the confidence degree vector of a region.

The confidence degree \( C'_{local} \) of the local concept obtained after the merging may be expressed as the following expression:

\[ C'_{local} = \left\{ C'_{local}(T(1)), C'_{local}(T(2)), C'_{local}(T(3)), C'_{local}(T(4)), C'_{local}(T(5)) \right\}. \]

where, \( C'_{local}(T) \) is the vector of a confidence degree set in relation to semantic concept \( L_{local} \) merged in divided region T. In the modeling of the global semantic concept the global concept of the input photo I in which regions are divided may be modeled by using an SVM. By using the confidence degree of a local concept as an input, the confidence degree of a global concept may be measured.

In the determining of the categories a global semantic concept having a highest confidence degree value among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept may be determined as the category of the photo. In the determining of the categories global semantic concepts having confidence degree values greater than a predetermined threshold value, among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept may be determined as the categories of the photo.

According to another aspect of the present invention, there is provided a clustering
apparatus of a digital photo album using region division templates, the apparatus including: a region division unit dividing a photo into regions by using region division templates; a local semantic concept modeling unit modeling a semantic concept included in a divided region; a local semantic concept merging unit, when it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, merging the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling; a global semantic concept modeling unit modeling a global semantic concept included in the photo by using a final local semantic concept determined after the merging; and a category determination unit determining one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the modeling.

[28] The apparatus may further include a photo input unit receiving an input of photo data for category-based clustering.

[29] The local semantic concept modeling unit may model the semantic concept by extracting content-based feature values of the photo, and the content-based feature values may include color, texture, and shape information of an image. The local semantic concept may be composed of an item (L entity) indicating the entity of a semantic concept included in a photo and an item (L attribute) indicating the attribute of the entity of the semantic concept.

[30] In the semantic concept modeling of the local semantic concept modeling unit, modeling of the local concepts of the input photo in which regions are divided may be performed by using a support vector machine (SVM). In the measuring of a confidence degree by the local semantic concept merging unit, the confidence degree of each local semantic concept may be measured by using one SVM for each defined local semantic concept.

[31] According to still another aspect of the present invention, there is provided a computer readable recording medium having embodied thereon a computer program for executing the methods.

**Advantageous Effects**

[32] According to the category-based clustering method and apparatus for a digital photo album according to the present invention, by using together user preference and content-based feature value information, such as color, texture, and shape, from the contents of photos, as well as information that can be basically obtained from photos, such as camera information and file information stored in a camera, a large volume of photos are effectively categorized such that an album can be fast and effectively generated with photo data.
Description of Drawings

[33] The above and other features and advantages of the present invention will become more apparent by describing in detail exemplary embodiments thereof with reference to the attached drawings in which:

[34] FIG. 1 is a block diagram of a structure of a photo clustering system using region division templates according to an embodiment of the present invention;

[35] FIG. 2 is a flowchart of a photo clustering method using region division templates according to an embodiment of the present invention;

[36] FIG. 3 illustrates region division templates according to an embodiment of the present invention;

[37] FIG. 4 illustrates an example of dividing a photo according to region division templates according to an embodiment of the present invention;

[38] FIG. 5 illustrates an example of entity concepts and attribute concepts of a divided region according to an embodiment of the present invention;

[39] FIG. 6 illustrates a local concept modeling in more detail according to an embodiment of the present invention;

[40] FIG. 7 illustrates an operation of grouping regions that are the objects of concept merging performed in a local semantic concept merging unit according to an embodiment of the present invention; and

[41] FIG. 8 is a relational block diagram illustrating category-based clustering process of a digital photo album according to an embodiment of the present invention.

Mode for Invention

[42] The present invention will now be described more fully with reference to the accompanying drawings, in which exemplary embodiments of the invention are shown.

[43] FIG. 1 is a block diagram of a structure of a photo clustering system using region division templates according to an embodiment of the present invention. The photo clustering system includes a region division unit 110, a local semantic concept modeling unit 120, a local semantic concept merging unit 130, a global semantic concept modeling unit 140, and a category determination unit 150. The photo clustering system may further include a photo input unit 100.

[44] The photo input unit 100 receives an input of photo data for category-based clustering. For example, a photo stream is input from an internal memory apparatus of a digital camera or a portable memory apparatus. The photo data is based on ordinary still image data. The format of the photo data includes an image data format, such as joint photographic experts group (JPEG), TIFF and RAW formats. However, the format of the photo data is not limited to these examples.

[45] The region division unit 110 divides a photo into regions by using region division
templates.

[46] The local semantic concept modeling unit 120 models a semantic concept included in the divided region and uses a local concept support vector machine (SVM) 160.

[47] When it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, the local semantic concept merging unit 130 merges the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling.

[48] The global semantic concept modeling unit 140 models a global semantic concept included in the photo by using the final local semantic concept determined through the merging, and uses a global concept SVM 170.

[49] The category determination unit 150 determines one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the global semantic concept modeling.

[50] FIG. 2 is a flowchart of a photo clustering method using region division templates according to an embodiment of the present invention. Referring to FIGS. 1 and 2, the photo clustering method using region division templates and the operation of the system for the method according to an embodiment of the present invention will now be explained.

[51] A photo stream from an internal memory apparatus of a digital camera or a portable memory apparatus is input in operation 200. The input photo is divided first, by using region division templates in operation 210. FIG. 3 illustrates region division templates according to an embodiment of the present invention. The present invention includes division of a photo with 10 base templates as shown in FIG. 3. The 10 region division base templates are expressed as the following equation 1:

\[ T = \{ T(t) \ | \ t \in \mathbb{R} \} \]

.......(1)

where \( T(t) \) is a \( t \)-th region division template.

[52] If the input photo \( I \) has dimensions of width \( w \) and length \( h \), the coordinates of each of the region division templates are expressed as the following equation 2:

\[ T(t) = \{ \text{left}(t), \text{top}(t), \text{right}(t), \text{bottom}(t) \} \]
where left(t) is the x coordinate of the left side of the t-th template, top(t) is the y coordinate of the top side of the t-th template, right (t) is the x coordinate of the right side of the t-th template, and bottom (t) is the y coordinate of the bottom side of the t-th template. According to the equation 2, the coordinates of each of the templates are as the following equations 3:

\[
\mathcal{T}(1) = \left[ \frac{\omega}{4}, \frac{\omega}{4}, \frac{3\omega}{4}, \frac{3\omega}{4} \right], \quad \mathcal{T}(2) = \left[ 0, 0, \frac{\omega}{2}, \frac{\omega}{2} \right], \quad \mathcal{T}(3) = \left[ \frac{\omega}{2}, 0, \frac{\omega}{2}, \frac{\omega}{2} \right], \quad \mathcal{T}(5) = \left[ \frac{\omega}{2}, \frac{\omega}{2}, \frac{\omega}{2}, \frac{\omega}{2} \right],
\]

\[
\mathcal{T}(6) = \left[ 0, 0, \frac{\omega}{2}, \frac{\omega}{2} \right], \quad \mathcal{T}(7) = \left[ 0, \frac{\omega}{2}, \frac{\omega}{2}, \frac{\omega}{2} \right], \quad \mathcal{T}(8) = \left[ 0, 0, \frac{\omega}{2}, \frac{\omega}{2} \right], \quad \mathcal{T}(9) = \left[ \frac{\omega}{2}, 0, \frac{\omega}{2}, \frac{\omega}{2} \right]
\]

\[
\mathcal{T}(10) = \left\{ \mathcal{T} \mid \mathcal{T} \in \mathcal{T} \right\}
\]

The input photo I divided according to the region division templates is expressed as the following equation 4:

\[
\mathcal{I} = \left\{ \mathcal{I}(\mathcal{T}) \mid \mathcal{T} \in \mathcal{T} \right\}
\]

FIG. 4 illustrates an example of dividing a photo performed in the region division unit 110. It can be seen that a local semantic concept is included in each of the divided regions. For example, in case of the first photo, it can be seen that sky is included on the top, riverside is included on the bottom left corner, and a lawn is included on the bottom right corner. That is, semantic concept information included in the photo is well expressed.

Next, multiple content-based features are extracted from each of the divided regions and a local semantic concept is modeled in operation 220. The used multiple content-based features are expressed as the following equation 5:
\[ F = \left\{ F(f') \mid f' \in \mathcal{N}_f \right\} \]

......(5)

where \( \mathcal{N}_f \) is the number of user feature values. The present invention includes a method of extracting content-based feature values using color, texture, and shape information of an image as basic features, and basically includes a method of extracting feature values by using an MPEG-7 descriptor. However, the method of extracting the content-based feature values is not limited to the MPEG-7 descriptor.

The multiple content-based feature values extracted from a region divided by template \( T \) is expressed as the following equation 6:

\[ F_T = \left\{ F_T(f') \mid f' \in \mathcal{N}_f \right\} \]

......(6)

Based on the given region-based feature values, a local semantic concept included in each of the divided regions is modeled.

For this, first, local semantic concepts that can be included in a target category of category-based clustering are defined.

A local semantic concept, \( \mathbf{L}_{\text{local}} \), is composed of \( \mathbf{L}_{\text{entity}} \) that is an item indicating the entity of a semantic concept included in a photo and \( \mathbf{L}_{\text{attribute}} \) that is an item indicating the attribute of the entity of a semantic concept. FIG. 5 is a table showing a local concept with the entity concept of a divided region and an attribute concept expressing the attribute of the entity concept according to an embodiment of the present invention.

\( \mathbf{L}_{\text{entity}} \) that is an item indicating the entity of a semantic concept is expressed as the following equation 7:

\[ \mathbf{L}_{\text{entity}} = \left\{ L_{\text{entity}}(e) \mid e \in \mathcal{N}_e \right\} \]

......(7)

where \( \mathbf{L}_{\text{entity}} \) is an e-th entity semantic concept, and \( \mathcal{N}_e \) is the number of defined entity semantic concepts.
Lattribute that is an item indicating the attribute of a semantic concept is expressed as the following equation 8:

\[ \mathbf{L}_{\text{attribute}} = \left\{ \mathbf{L}_{\text{attribute}} (\bar{a}) \mid \bar{a} \in \mathbf{N}_a \right\} \]

......(8)

where \( \mathbf{L}_{\text{attribute}} (a) \) is an a-th attribute semantic concept, and \( \mathbf{N}_a \) is the number of defined attribute semantic concepts.

The local semantic concept \( \mathbf{L}_{\text{local}} \) is expressed as the following equation 9:

\[ \mathbf{L}_{\text{local}} = \{ \mathbf{L}_{\text{entry}}, \mathbf{L}_{\text{attribute}} \} = \{ \mathbf{L}(l) \mid l \in (\mathbf{N}_e + \mathbf{N}_a) \} \]

......(9)

where \( \mathbf{L}(l) \) is an l-th semantic concept, and can be an entity semantic concept or an attribute semantic concept.

Next, based on the local semantic concepts described as above, training sample image having respective local semantic concepts are collected. The content-based feature values are extracted from the collected images and the extracted feature values are trained by using a support vector machine (SVM).

\( \text{SVM}_{\text{local}} \) trained in relation to each of the local semantic concepts is expressed as the following equation 10:

\[ \text{SVM}_{\text{local}} = \left\{ \text{SVM}_L (\mathbf{F}) \mid \mathbf{L} \in \mathbf{L}_{\text{local}} \right\} \]

......(10)

where SVM_\( L \) is an SVM trained for semantic concept \( L \). As the input of the SVM local, the content-based feature value vector \( \mathbf{F} \) described above is input.

Next, by using the trained SVM_{local}, a local concept of the input photo I in which regions are divided is modeled. That is, the input photo I is divided into regions
according to the method described above and the divided region images are modeled by using the trained SVMlocal. The modeling of the local concept includes a process of inputting the content-based feature values extracted from the divided region images, into the SVML of semantic concept L and extracting the confidence degree of the semantic concept.

FIG. 6 illustrates the local concept modeling of the operation 220 in more detail according to an embodiment of the present invention. That is, the local concept modeling includes a local entity concept modeling in operation 600 and a local attribute concept modeling in operation 650.

The confidence degree of a semantic concept in relation to divided region T can be obtained according to the following equation 11:

\[ C_L(T) = SVM_L(F_T) \]

\[ \ldots (11) \]

where \( F_T \) is a content-based feature value vector of divided region T, and \( C_L(T) \) is the confidence degree of semantic concept L of the divided region T. The confidence degree is a measured value on how much a divided region image include a semantic concept corresponding to the region.

A confidence degree vector obtained by performing SVMs of all defined local semantic concepts is expressed according to the following equation 12:

\[ C_{local}(T) = \{ C_L(T) = SVM_L(F_T) | L \in L_{local} \} \]

\[ \ldots (12) \]

where \( C_{local}(T) \) is a confidence degree vector of each of all local semantic concepts modeled in relation to divided region T.

As a result, the confidence degree of the local semantic concept in relation to 10 divided regions can be obtained and the confidence degree vector of the local semantic concept obtained in relation to the 10 divided regions is expressed as the following equation 13:
\[ C_{\text{local}} = \{ C_{\text{local}}(T|T \in T) \}_{n=1}^{10} = \{ C_{\text{local}}(T(1)), C_{\text{local}}(T(2)), C_{\text{local}}(T(3)), \ldots, C_{\text{local}}(T(10)) \}_{n=1}^{10} \]

... (13)

The defined division regions include regions spatially overlapping each other. That is, divided region T(1) overlaps T(10); T(2) overlaps T(6), T(8), and T(10); T(3) overlaps T(6), T(9), and T(10); T(4) overlaps T(7), T(8), and T(10); and finally, T(5) overlaps T(7), T(9), and T(10). Accordingly, a total of five overlapping region groups exist. In the present invention, in order to extract a more reliable local semantic concept, a process of merging the confidence degrees of the local concepts of the overlapping region groups is included in operation 230.

In the present invention, as the method of merging the semantic concepts of the overlapping region groups, included is a method by which divided regions T(1) and T(10) are merged into T(1); T(2), T(6), T(8), and T(10) are merged into T(2); T(3), T(6), T(9), and T(10) are merged into T(3); T(4), T(7), T(8), and T(10) are merged into T(4); and T(5), T(7), T(9), and T(10) are merged into T(5). The local semantic concept merging process includes a process of allocating a highest confidence degree value among semantic concepts allocated to divided regions belonging to each divided region group, to a corresponding merging region.

FIG. 7 illustrates an operation of grouping regions that are the objects of local concept merging performed in the local semantic concept merging unit 130 according to an embodiment of the present invention. The local semantic concept merging process is expressed as the following equation 14:
\[ C'_{\mathbb{L}} (T(1)) = \max \left\{ C_{\mathbb{L}} (T) \mid T \in \{ T(1), T(10) \} \right\}, \]
\[ C'_{\mathbb{L}} (T(2)) = \max \left\{ C_{\mathbb{L}} (T) \mid T \in \{ T(2), T(6), T(8), T(10) \} \right\}, \]
\[ C'_{\mathbb{L}} (T(3)) = \max \left\{ C_{\mathbb{L}} (T) \mid T \in \{ T(3), T(6), T(9), T(10) \} \right\}, \]
\[ C'_{\mathbb{L}} (T(4)) = \max \left\{ C_{\mathbb{L}} (T) \mid T \in \{ T(4), T(7), T(8), T(10) \} \right\}, \]
\[ C'_{\mathbb{L}} (T(5)) = \max \left\{ C_{\mathbb{L}} (T) \mid T \in \{ T(5), T(7), T(9), T(10) \} \right\} \]

\[ \ldots \quad (14) \]

[122] As a result, T(1), T(2), T(3), T(4), and T(5) are determined as final divided regions, and the confidence degree \( C'_{\text{local}} \) of the local semantic concept allocated to each divided region is expressed as the following equation 15:

[123]

[124]

[125]

\[ C'_{\text{local}} = \{ C'_{\text{local}} (T(1)), C'_{\text{local}} (T(2)), C'_{\text{local}} (T(3)), C'_{\text{local}} (T(4)), C'_{\text{local}} (T(5)) \} \]

\[ \ldots \quad (15) \]

[126] where, each \( C'_{\text{local}} (T) \) is the vector of a confidence degree set in relation to semantic concept \( I_{\text{local}} \) determined in divided region \( T \).

[127] Next, based on the confidence degree of the local semantic concept measured as described above, a global semantic concept, that is, a category concept, included in the input photo I is modeled in operation 240.

[128] For this, sample images of photos belonging to each category are collected, and then, from the collected sample images, the confidence degree \( C'_{\text{local}} \) of a local semantic concept is obtained through the same process as described above. Based on this confidence degree, a process of training using an SVM is performed. First, a global semantic concept, that is, a category concept, is expressed as the following equation 16:

[129]

[130]

[131]
\[ L_{\text{global}} = \{ L(g) \mid g \in M_g \} \]

......(16)

[132] where \( L(g) \) is a \( g \)-th category concept and \( N_g \) is the number of category concepts.

[133] SVM_{global} trained in relation to each category concept is expressed as the following equation 17:

\[ SVM_{\text{global}} = \{ SVM_G(C_{\text{local}}) \mid G \in L_{\text{global}} \} \]

......(17)

[137] where \( SVM_G \) is the SVM trained for category concept \( G \). As the input of \( SVM_{\text{global}} \), \( C_{\text{local}} \) that is the confidence degree set of semantic concepts extracted from the divided regions and merged is used.

[138] Next, by using the trained \( SVM_{\text{global}} \), the category concept of the input photo \( I \) is modeled in operation 240. \( C_{\text{local}} \) that is the confidence degree set of local semantic concepts of the input photo is input to an SVM for modeling each category concept, and the confidence degree of each category concept in relation to the input photo \( I \) is obtained. The confidence degree of the modeled category concept \( G \) is expressed as the following equation 18:

\[ C_G = SVM_G(C_{\text{local}}) \]

......(18)

[142] where \( C_G \) is the confidence degree of category concept \( G \). The confidence degree set \( C_{\text{global}} \) of the global category concept obtained based on the method described above is expressed as the following equation 19:
\[ C_{\text{global}} = \{ C_G \mid G \in L_{\text{global}} \} \]

......(19)

[146] The final category concept of the input photo I is determined by selecting a category having the highest confidence degree among the defined confidence degrees of the category concept \( L_{\text{global}} \). The present invention includes a method of selecting a category concept having a highest confidence degree, and a method of selecting a category concept having a confidence degree equal to or greater than a predetermined value.

[147] The method of selecting a category concept having a highest confidence degree is expressed as the following equation 20:

[148]
[149]
[150]

\[ L_{\text{target}} = \arg \max_{G \in L_{\text{global}}} \{ C_G \} \]

......(20)

[151] where \( L_{\text{target}} \) is a category concept finally selected.

[152] The method of selecting a category concept having a confidence degree equal to or greater than a predetermined value is expressed as the following equation 21:

[153]
[154]
[155]

\[ L_{\text{target}} = \arg \{ C_G > C_{\text{th}} \} \]

......(21)

[156] where \( C_{\text{th}} \) is a threshold value of a confidence value to select a final category concept.

[157] FIG. 8 is a relational block diagram illustrating category-based clustering process of a digital photo album according to an embodiment of the present invention as described above.

[158] The present invention can also be embodied as computer readable codes on a computer readable recording medium. The computer readable recording medium is any
data storage device that can store data which can be thereafter read by a computer system. Examples of the computer readable recording medium include read-only memory (ROM), random-access memory (RAM), CD-ROMs, magnetic tapes, floppy disks, and optical data storage devices.

While the present invention has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims. The preferred embodiments should be considered in descriptive sense only and not for purposes of limitation. Therefore, the scope of the invention is defined not by the detailed description of the invention but by the appended claims, and all differences within the scope will be construed as being included in the present invention.

According to the category-based clustering method and apparatus for a digital photo album according to the present invention, by using together user preference and content-based feature value information, such as color, texture, and shape, from the contents of photos, as well as information that can be basically obtained from photos, such as camera information and file information stored in a camera, a large volume of photos are effectively categorized such that an album can be fast and effectively generated with photo data.
Claims

[1] A clustering method of a digital photo album using region division templates, the method comprising:

- dividing a photo into regions by using region division templates;
- modeling a semantic concept included in a divided region;
- when it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, merging the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling;
- modeling a global semantic concept included in the photo by using a final local semantic concept determined after the merging; and
- determining one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the modeling.

The method of claim 1, wherein the region division templates in the modeling of the semantic concept are expressed by the following equations:

\[ T(1) = \left\{ \frac{w}{4}, \frac{h}{4}, \frac{3w}{4}, \frac{3h}{4} \right\}, \quad T(2) = \left\{ 0, 0, \frac{w}{2}, \frac{h}{2} \right\}, \quad T(3) = \left\{ \frac{w}{2}, 0, \frac{w}{2}, \frac{h}{2} \right\}, \quad T(5) = \left\{ \frac{w}{2}, \frac{h}{2}, w, h \right\}, \]

\[ T(6) = \left\{ 0, w, \frac{h}{2} \right\}, \quad T(7) = \left\{ \frac{w}{2}, 0, \frac{w}{2}, h \right\}, \quad T(8) = \left\{ 0, 0, \frac{w}{2}, h \right\}, \quad T(9) = \left\{ \frac{w}{2}, 0, w, h \right\} \]

\[ T^\prime(10) = \left\{ 0, 0, w, \frac{h}{2} \right\} \]

where \( T \) is the template of a photo, \( w \) is the length of the width of the photo, and \( h \) is the length of the height of the photo.

[2] The method of claim 1, wherein in the modeling of the semantic concept, the semantic concept is modeled by extracting content-based feature values of the photo.

[3] The method of claim 3, wherein the content-based feature values comprise color, texture, and shape information of an image.

[4] The method of claim 1, wherein in the modeling of the semantic concept the semantic concept is composed of an item (\( L_{entity} \)) indicating the entity of a semantic concept included in a photo and an item (\( L_{attribute} \)) indicating the attribute of the entity of the semantic concept.

[5] The method of claim 5, wherein the semantic concept modeling is modeling of the entity concept and the attribute concept of a divided region.
The method of claim 1, wherein in the modeling of the semantic concept, modeling of the local concepts of the input photo in which regions are divided is performed by using a support vector machine (SVM).

The method of claim 7, wherein in the merging of the semantic concepts of respective regions, the confidence degree of each local semantic concept is measured by using one SVM for each defined local semantic concept.

The method of claim 2, wherein in the merging of the semantic concepts of respective regions, based on the confidence degrees of local concepts allocated to 10 regions divided by using the region division templates, local concept confidence degrees of 5 basic regions are merged according to the following equation:

\[
C'_L(T(1)) = \max \left\{ C'_L(T) \mid T \in \{ T(1), T(10) \} \right\},
\]

\[
C'_L(T(2)) = \max \left\{ C'_L(T) \mid T \in \{ T(2), T(6), T(8), T(10) \} \right\},
\]

\[
C'_L(T(3)) = \max \left\{ C'_L(T) \mid T \in \{ T(3), T(6), T(9), T(10) \} \right\},
\]

\[
C'_L(T(4)) = \max \left\{ C'_L(T) \mid T \in \{ T(4), T(7), T(8), T(10) \} \right\},
\]

\[
C'_L(T(5)) = \max \left\{ C'_L(T) \mid T \in \{ T(5), T(7), T(9), T(10) \} \right\},
\]

where T(1), T(2), T(3), T(4), and T(5) indicate basic regions to which final local semantic concepts are allocated, and \( C'_L \) is the confidence degree vector of a region.

The method of claim 9, wherein the confidence degree \( C'_{\text{local}} \) of the local concept obtained after the merging is expressed as the following expression:

\[
C'_{\text{local}} = \left\{ C'_{\text{local}}(T(1)), C'_{\text{local}}(T(2)), C'_{\text{local}}(T(3)), C'_{\text{local}}(T(4)), C'_{\text{local}}(T(5)) \right\},
\]

where, \( C'_{\text{local}}(T) \) is the vector of a confidence degree set in relation to semantic concept \( L_{\text{local}} \) merged in divided region T.

The method of claim 1, wherein in the modeling of the global semantic concept the global concept of the input photo I in which regions are divided is modeled by using an SVM.

The method of claim 11, wherein by using the confidence degree of a local concept as an input, the confidence degree of a global concept is measured.
[13] The method of claim 1, wherein in the determining of the categories a global semantic concept having a highest confidence degree value among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept is determined as the category of the photo.

[14] The method of claim 1, wherein in the determining of the categories global semantic concepts having confidence degree values greater than a predetermined threshold value, among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept are determined as the categories of the photo.

[15] A clustering apparatus of a digital photo album using region division templates, the apparatus comprising:

a region division unit dividing a photo into regions by using region division templates;

a local semantic concept modeling unit modeling a semantic concept included in a divided region;

a local semantic concept merging unit, when it is assumed that a measured value indicating the degree that the image of the region includes a semantic concept corresponding to the region is a confidence degree, merging the semantic concepts of respective regions with respect to the confidence degree of the local meaning measured from the modeling;

a global semantic concept modeling unit modeling a global semantic concept included in the photo by using a final local semantic concept determined after the merging; and

a category determination unit determining one or more categories included in the input photo according to the confidence degree of the global semantic concept measured from the modeling.

[16] The apparatus of claim 15, further comprising a photo input unit receiving an input of photo data for category-based clustering.

[17] The apparatus of claim 15, wherein the local semantic concept modeling unit models the semantic concept by extracting content-based feature values of the photo, and the content-based feature values comprise color, texture, and shape information of an image.

[18] The apparatus of claim 17, wherein the local semantic concept is composed of an item \((L_{\text{entity}})\) indicating the entity of a semantic concept included in a photo and an item \((L_{\text{attribute}})\) indicating the attribute of the entity of the semantic concept.

[19] The apparatus of claim 18, wherein in the semantic concept modeling of the local semantic concept modeling unit, modeling of the local concepts of the input photo in which regions are divided is performed by using a support vector
machine (SVM).

[20] The apparatus of claim 19, wherein in the measuring of a confidence degree by the local semantic concept merging unit, the confidence degree of each local semantic concept is measured by using one SVM for each defined local semantic concept.

[21] The apparatus of claim 15, wherein in the merging of the semantic concepts of the regions, based on the confidence degrees of local concepts allocated to 10 regions divided by using the region division templates, local concept confidence degrees of 5 basic regions are merged according to the following equation:

\[ C'_{\text{local}} (T) = \max \{ C_{\text{local}} (T) | T \in \{ T(1), T(10) \} \}, \]

\[ C'_{\text{local}} (T(2)) = \max \{ C_{\text{local}} (T) | T \in \{ T(2), T(6), T(8), T(10) \} \}, \]

\[ C'_{\text{local}} (T(3)) = \max \{ C_{\text{local}} (T) | T \in \{ T(3), T(6), T(9), T(10) \} \}, \]

\[ C'_{\text{local}} (T(4)) = \max \{ C_{\text{local}} (T) | T \in \{ T(4), T(7), T(8), T(10) \} \}, \]

\[ C'_{\text{local}} (T(5)) = \max \{ C_{\text{local}} (T) | T \in \{ T(5), T(7), T(9), T(10) \} \}, \]

where \( T(1), T(2), T(3), T(4), \) and \( T(5) \) indicate basic regions to which final local semantic concepts are allocated, and \( C_{\text{local}}' \) is the confidence degree vector of a region.

[22] The apparatus of claim 21, wherein the confidence degree \( C'_{\text{local}} \) of the local concept obtained after the merging is expressed as the following expression:

\[ C'_{\text{local}} = \{ C'_{\text{local}} (T(1)), C'_{\text{local}} (T(2)), C'_{\text{local}} (T(3)), C'_{\text{local}} (T(4)), C'_{\text{local}} (T(5)) \} \]

where, \( C'_{\text{local}} (T) \) is the vector of a confidence degree set in relation to semantic concept \( T_{\text{local}} \) merged in divided region \( T \).

[23] The apparatus of claim 15, wherein the global semantic concept modeling unit models the global concept of the input photo I in which regions are divided, by using an SVM.

[24] The apparatus of claim 23, wherein in measuring of the confidence degree of a global concept by the category determination unit, by using the confidence degree of a local concept as an input, the confidence degree of the global concept is measured.
[25] The apparatus of claim 15, wherein the category determination unit determines a global semantic concept having a highest confidence degree value among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept is determined as the category of the photo.

[26] The apparatus of claim 15, wherein the category determination unit determines global semantic concepts having confidence degree values greater than a predetermined threshold value, among the confidence degrees of the global semantic concepts measured from the modeled global semantic concept are determined as the categories of the photo.

[27] A computer readable recording medium having embodied thereon a computer program for executing the method of claim 1.
FIG. 2

START

INPUT PHOTO

200

DIVIDE PHOTO INTO REGIONS BY USING REGION DIVISION TEMPLATES

210

MODEL LOCAL SEMANTIC CONCEPT

220

MERGE LOCAL SEMANTIC CONCEPTS

230

MODEL GLOBAL SEMANTIC CONCEPT

240

DETERMINE CATEGORY

250

STOP

FIG. 3

SUBSTITUTE SHEET (RULE 26)
**FIG. 5**

<table>
<thead>
<tr>
<th>SEMANTIC CONCEPTS</th>
<th>LEXICONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>L_attribute (ATTRIBUTE)</strong></td>
<td><strong>ADVERB</strong></td>
</tr>
<tr>
<td><strong>ADJECTIVE</strong></td>
<td>COLOR: REDDISH, GREENISH, BLUISH, YELLOWISH, GRAYISH, WHITISH, COLORFUL (MIXED)</td>
</tr>
<tr>
<td></td>
<td>TEXTURE: SMOOTH, ROUGH, WAVE, LATTICED</td>
</tr>
<tr>
<td></td>
<td>DIRECTION (EDGE): HORIZONTAL, VERTICAL, 45 DEGREE, 135 DEGREE, NONDIRECTIONAL</td>
</tr>
<tr>
<td></td>
<td>POSITION: CENTER, LEFT-TOP, LEFT-BOTTOM, RIGHT-TOP, RIGHT-BOTTOM</td>
</tr>
<tr>
<td><strong>L_entity (ENTITY) (NOUN)</strong></td>
<td>TERRAIN: GRASS, TREE, WOOD, ROCK, SKY, CLOUD, SUN</td>
</tr>
<tr>
<td></td>
<td>WATERSIDE: RIVER, POND, SEA, FOUNTAIN, UNDERWATER, WATERFALL</td>
</tr>
<tr>
<td></td>
<td>SUNSET: SKY, SUN</td>
</tr>
<tr>
<td></td>
<td>SNOWSCAPE: SNOW</td>
</tr>
<tr>
<td></td>
<td>NIGHTSCENE: STREETLIGHT, NEON-SINE, FLASHLIGHT</td>
</tr>
<tr>
<td></td>
<td>BUILDING: SKYSCRAPER, STREET, ROAD, PAVEMENT, BRIDGE, STAIRS, BILLBOARD</td>
</tr>
</tbody>
</table>
FIG. 6

ENTITY CONCEPT MODELING

600

ATTRIBUTE CONCEPT MODELING

650

FIG. 7

SUBSTITUTE SHEET (RULE 26)
FIG. 8

LOW LEVEL FEATURES (M FEATURES)

SEMANTIC SVM IN RELATION TO LOCAL REGION
n CONCEPTS (ENTITY AND ATTRIBUTE LEXICONS)

CONCEPT MERGING FROM 10 TO 5 REGION CONCEPTS

CONFIDENCE DEGREES OF 5 REGIONS

GLOBAL SEMANTIC SVM

GLOBAL SEMANTIC CONFIDENCE

1 2 3 ... n

1 2 3 ... n

1 2 3 ... n

1 2 3 ... 5n

G₁ G₂ ... Gₖ

C₁ C₂ ... Cₖ

10 REGIONS

5 REGIONS

SUBSTITUTE SHEET (RULE 26)
A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/00(2006.01)i

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)

IPC 8 G06T 12/00; 17/30; G06T 1/00;

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Korean Utility models and applications for Utility models since 1975
Japanese Utility models and application for Utility models since 1975

Electronic database consulted during the international search (name of data base and, where practicable, search terms used)
WPI, ESPASNET, INSPECT, IEE/IEEE, PAJ, eKIPASS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>US 5,913,205 A (Ramesh Jain, et al.) 15 Jun 1999 See the abstract</td>
<td>1-27</td>
</tr>
<tr>
<td>A</td>
<td>WO 2003/090167 A2 (PAO, You-Han) 30 Oct 2003 See the abstract</td>
<td>1-27</td>
</tr>
</tbody>
</table>

Further documents are listed in the continuation of Box C.

See patent family annex.

Date of the actual completion of the international search 24 APRIL 2006 (24.04.2006)

Date of mailing of the international search report 25 APRIL 2006 (25.04.2006)

Name and mailing address of the ISA/KR
Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea
Facsimile No. 82-42-472-7140

Authorized officer
MA, Jung Youn
Telephone No. 82-42-481-5679

Form PCT/ISA/210 (second sheet) (April 2005)
<table>
<thead>
<tr>
<th>Patent document cited in search report</th>
<th>Publication date</th>
<th>Patent family member(s)</th>
<th>Publication date</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 5,913,205 A</td>
<td>15 Jun 1999</td>
<td>NONE</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 19970066989 A</td>
<td>13 Oct 1997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 01459615 A1</td>
<td>22 Sep 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 17512248 A</td>
<td>28 Apr 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20040068210 A</td>
<td>30 Jul 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WO 2003051033 A1</td>
<td>19 Jun 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CA 2480954 A1</td>
<td>30 Oct 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CN 1647109 A</td>
<td>27 Jul 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 1500052 A2</td>
<td>26 Jan 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20040101477 A</td>
<td>2 Dec 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP 01622024 A1</td>
<td>1 Feb 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>JP 16320514 A</td>
<td>11 Nov 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KR 20060010728 A</td>
<td>2 Feb 2006</td>
</tr>
</tbody>
</table>

Form PCT/ISA/210 (patent family annex) (April 2005)