Image Presentation

13

The best usually is also the simplest.

13.1 Introduction

Image presentation is about how present the database images or retrieved images to the user in the most effective and efficient way. Given a query from a user, the list of images retrieved from the database can number from thousands to millions. Due to the limitation of both image features and image ranking, the retrieval list is usually scattered with irrelevant images. How to organize the retrieved images and present as many relevant images as possible in a very limit space is a great challenge in IR. Image presentation is part of the research on user interface and data visualization. It is a mixture of tech and arts. Many interesting image presentation methods have been developed, and they include simple browsing, category browsing, content browsing, hierarchical organization, sophisticated approaches involving user interaction, etc. In this chapter, we discuss a number of most common approaches on image presentation.

13.2 Caption Browsing

If images in the database have been labeled with captions, the simplest way to find relevant images is through caption browsing. This is typically used in personal/family photo albums stored in PCs and photo galleries on web pages, which are usually labeled, dated, and grouped by photo takers.

Caption browsing can also be used in image retrieval. If a retrieval list is not long, this method can be used to scan through the images on the list in order to find the most relevant images. Figure 13.1 shows a caption browsing given by MS Windows File Explorer.

D. Zhang, *Fundamentals of Image Data Mining*, Texts in Computer Science, https://doi.org/10.1007/978-3-030-17989-2_13

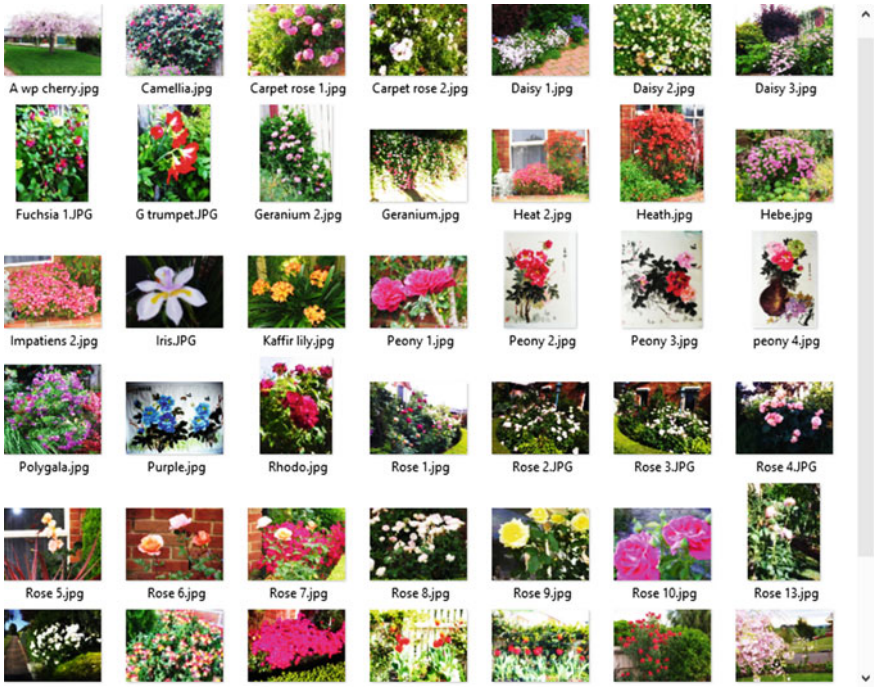

Fig. 13.1 Browsing images using image captions by Windows File Explorer

The advantage of using caption browsing is that it is effortless and it suits for universal users. Furthermore, users can search images based on image captions or date information. For small image databases or a short image retrieval list, caption browsing is the best option. A major issue with caption browsing is that the caption labels are subjective, and they are often confusing and even misleading.

13.3 Category Browsing

Category browsing is a typical approach for organizing image databases which are much larger than personal or family photo albums. Suppose images have been classified or annotated using the methods described in Part III; images can be organized into categories just like books are organized in library. Indeed, this is exactly how images are traditionally indexed in library and archives. Under each category, users can further filter the images using a simple browsing. Figure 13.2 shows an example of category browsing used by MS Windows File Explorer.

The advantage of category browsing is that very large collections of images can be broadly organized into categories every efficiently; there is no need to label individual images. This is particularly useful given there is massive amount of digital images on the Internet and most of the images are either unlabeled or mislabeled. However, images in categories of very large image collections are

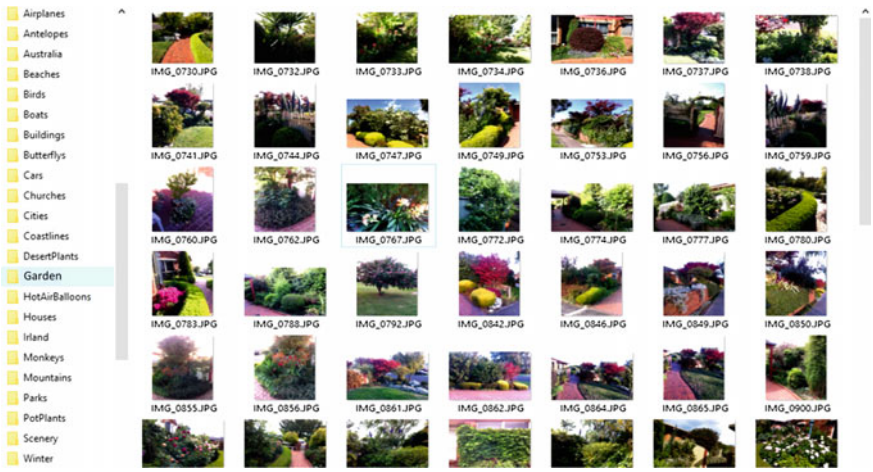

Fig. 13.2 Category browsing using MS Windows File Explorer

usually too diverse both perceptually and semantically. A common approach is to subdivide each of the categories into subcategories and a more efficient hierarchical category browsing is designed.

13.3.1 Category Browsing on the Web

More effective and efficient category browsing methods have been designed for the web where each category is initiated with one or more representative images. Figure 13.3 shows an example website using category browsing. It can be seen that the Web category browsing is more effective and efficient than a desktop file explorer, because it not only visualizes each category with one or multiple iconic images but also provides a convenient searching mechanism. Furthermore, a web browsing system also provides a user with key information about the image collection and categories, such as the number of categories, number of images, types of images, source of images, instructions on retrieving the images, how to use the images, etc. More examples of category browsing on the web can be found online by Googling for "image categories."

13.3.2 Hierarchical Category Browsing

Categories of the same concept can be merged to organize the image database into a *hierarchical* or *tree* structure. On the other hand, if there are too many images in a category, it can be subdivided into subcategories. The subjective labeling issue in image categorization can be overcome using a thesaurus such as the WordNet [1].

The design of a hierarchical category browsing system can be done either before the collection of images or during the collection of images. Due to the lack of

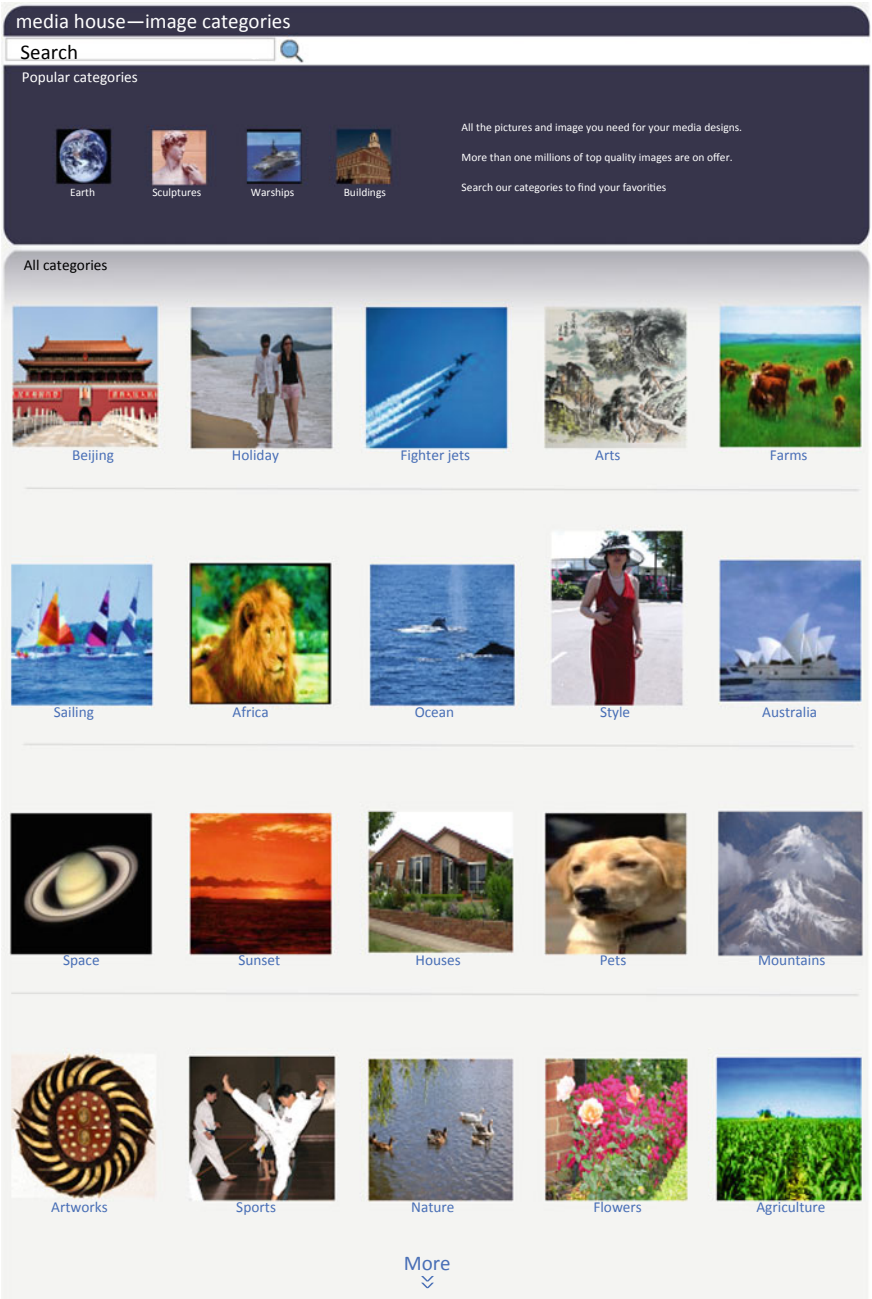

Fig. 13.3 A website of category browsing

standard image taxonomy, image categories are typically defined during the collection of images or defined "on the run". In this case, images should be put into as many categories as possible, so that higher level of categories can be created by merging similar subcategories using a thesaurus. Images can be most efficiently retrieved using a standard hierarchy like the one used to classify animals or plants.

Figure 13.4 shows an example of hierarchical organization of image categories; each top-level category is demonstrated with a number of iconic images to give an

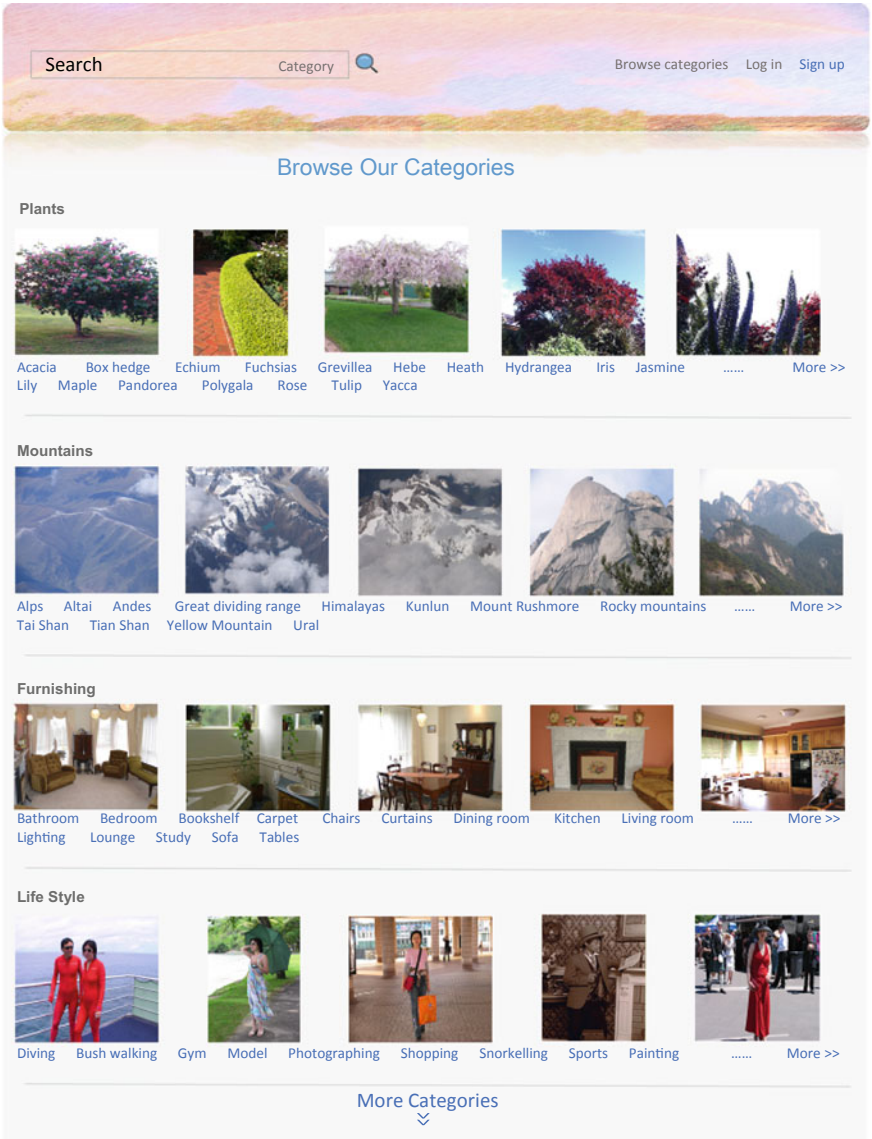

Fig. 13.4 Hierarchical organization of image categories on a website

idea on what is inside the category. It can be seen that it is a more effective and efficient presentation of image categories than that in Fig. 13.3, because more image categories can be shown to users on one web page.

13.4 Content Browsing

Images in the database or on a retrieval list can also be organized based on *numerical image features* instead of semantic labels as shown in the above sections. Numerical image features are also called low-level image features, perceptual features, or *content features*. Therefore, image retrieval based on numerical image features is called *content-based image retrieval* or CBIR for short. The idea of *content browsing* is to convert and group retrieved images as thumbnails based on their perceptual similarities such as color, texture, or shape. The organization of the thumbnails is then presented as an *image map* to guide a user to navigate to the most relevant images. This is very useful because people usually judge image similarity based on perception, and perceptual features are usually less subjective than semantic labeling which can vary widely from person to person. Figure 13.5 shows an example of content browsing.

Content browsing is a study of image *data visualization*. Key issues in content browsing include the following:

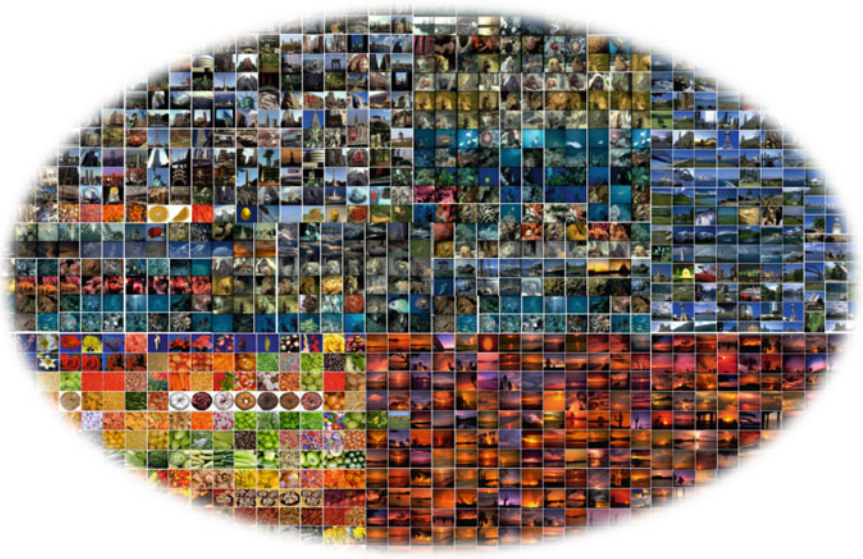

Fig. 13.5 Image presentation based on content browsing

- **Features** of images. What image feature(s) to use? Typically, images to be presented are grouped based on color, texture, shape, or combination of them, and it depends on the nature of the image data and application.
- **Similarity** measure. What similarity measure to use? Different similarity measures can lead to different perceptual experiences on the content browsing.
- **Size** of thumbnails. Should they be equal size or different sizes depending on relevance?
- **Overlap** of thumbnails. Should thumbnails be overlapped and how much?
- **Structure and layout** of thumbnails. Should they be organized in row and column, spiral, or graph?
- **Space** of presentation. Whether images are presented in 2D space or 3D space such as a cylinder or a sphere.
- **Interaction**. What type of interaction to use and how a user can interact with the thumbnails?

These factors are determined by research on how the physical world can be most effectively presented to human vision system.

13.4.1 Content Browsing in 3D Space

The limited space of a 2D screen can be extended using a 3D cylinder (Fig. 13.6) [2], or a sphere (Fig. 13.7) to present more thumbnails in an image map [3]. Images can also be presented on a curved wall as shown in Fig. 13.8 to extend the visual space.

13.4.2 Content Browsing with Focus

Content browsing with equal size thumbnails provides a global view of the visualized image data. Alternatively, a local view of the image map provides a content browsing with *focus*. For example, in Fig. 13.9, images on the top of the retrieval list are given focus and presented in the center of the screen with larger size according to their similarity to the query image, and they are then arranged in a clockwise spiral order [4].

13.4.3 Force-Directed Content Browsing

Conventional content browsing does not have overlap between images and the distance between neighboring images is uniform. In a force-directed content browsing, images in a category or a retrieval list are regarded as planets in space and the distance between images is regarded as a force between them. If two images are similar enough, the force is strong enough to attract them together. The result of a force-directed visualization is an overlapped image map as shown in Fig. 13.10a.

Fig. 13.6 Content browsing using a cylindrical image map

Fig. 13.7 Content browsing using a spherical image map

Figure 13.10b shows the force-directed image map in diminishing perspective. A bi-force-directed content browsing can be visualized using a 3D image map as shown in Fig. 13.10c. In the 3D image map, there are two types of forces

Fig. 13.8 Content browsing using a curved wall image map

representing two types of similarity between images, e.g., the horizontal force represents the color similarity and the vertical force represents the texture similarity. A close-up section of the 3D image map is shown in Fig. 13.10d.

Interactivity over an image map can also include keywords. For example, keywords can be used as a mouse-over feature of the image map. A *keyword map* and an *image map* can be created separately and put side by side, and the two maps are linked to provide a user with a joint content browsing [4].

13.5 Query by Example

When come to find a required image, one of the most challenging issues for a user is how to initiate the search, which includes how to start, where to start and how to formulate a search. One of the simplest ways to start a search is a *query by example* (QBE), or search based on an example image from the user. In QBE, similar images in the database are retrieved based on the *content features* of the query image. The

Fig. 13.9 Content browsing with focus and spiral structure

justification for QBE is that often the requirement for an image cannot be described by words and a user has a sample image which is similar to what they actually want to find. The sample image can be used as equivalent keywords for a query. QBE is a useful functionality; because most of the images in the world are unlabeled, these images can be retrieved based on content features. Therefore, QBE is an essential part of an IR system.

An example of QBE is Google's "Search by image" as shown in Fig. 13.11 [5]. In the figure, the sample image was first *translated* to the keywords "Sydney opera house" which are then used as the query for the actual image retrieval. Due to the translation, Google's search by an example image can be regarded as a semi-QBE, because labels of images in the database may not be obtained from image content. This works well when a query image is properly translated; however, when a query image is incorrectly translated, it results in completely wrong retrievals. For example, Fig. 13.12 shows three example queries which are translated to "darkness", "comics", and "computer", respectively. As a result, the retrievals of the three queries are completely incorrect. Furthermore, when the hand image query is presented using two different copies, they are translated as "Naya Rivera" and

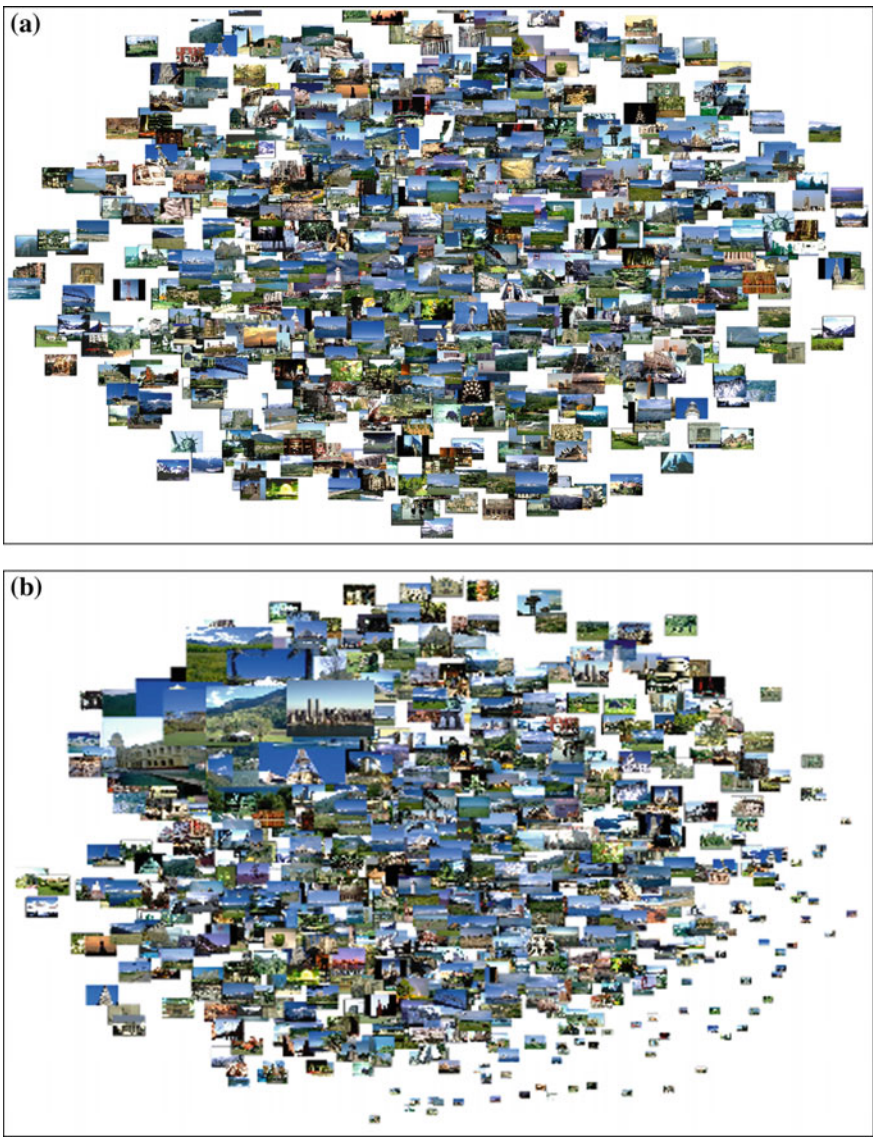

Fig. 13.10 Force-directed content browsing

"Point Cabrillo Light," respectively, which would not occur if content features are used for direct matching. This demonstrates that a CBIR-based QBE is a necessary component for an image retrieval system.

Although QBE is useful and convenient in many situations, it can be challenging to find an example image for a user. One of the ways to find example images is to use *query by keywords* or QBK which will be introduced in the following section.

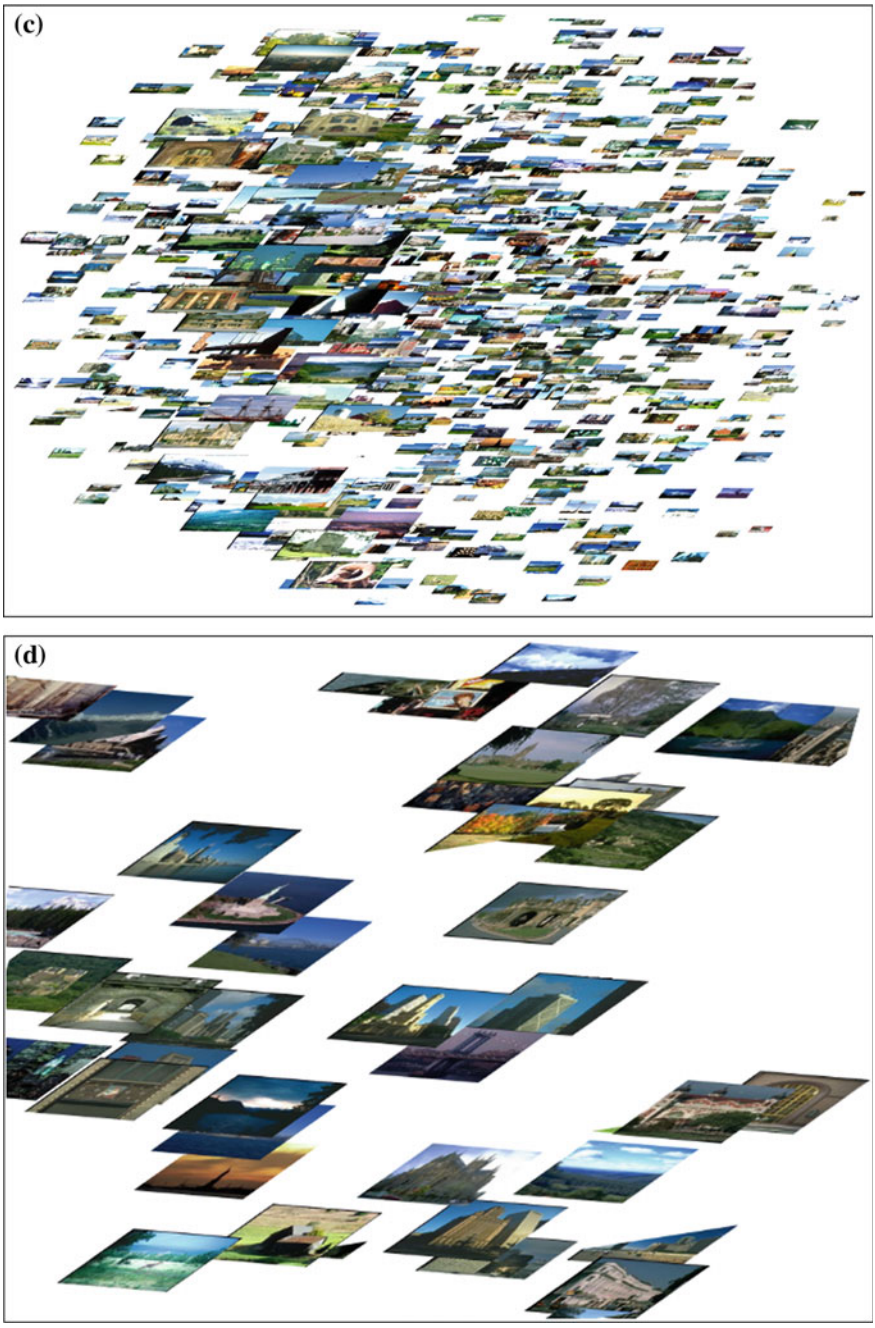

Fig. 13.10 (continued)

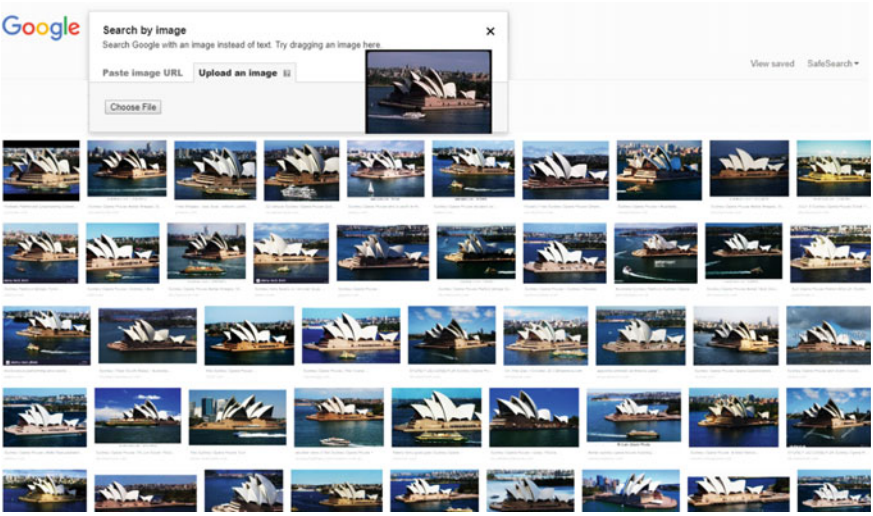

Fig. 13.11 A QBE example from Google

Fig. 13.12 Query images for Google's "Search by image"

13.6 Query by Keywords

People describe the visual world using textual language such as a blue sky, green trees, red flowers, a cheerful street, violent protests, etc. One of the natural ways to search images is to use *query by keywords* or QBK. In a QBK system, images in database are all labeled with keywords and indexed using inverted files; the search for images is just the same as search for textural documents. Unfortunately, most of the images in the world are either unlabeled or mislabeled. In a typical QBK system, image labels are obtained from image content through automatic image classification/annotation. A typical QBK system is shown in Fig. 13.13 [6, 7]. Image labels can also be obtained by analyzing textual description of images in web pages on the Internet.

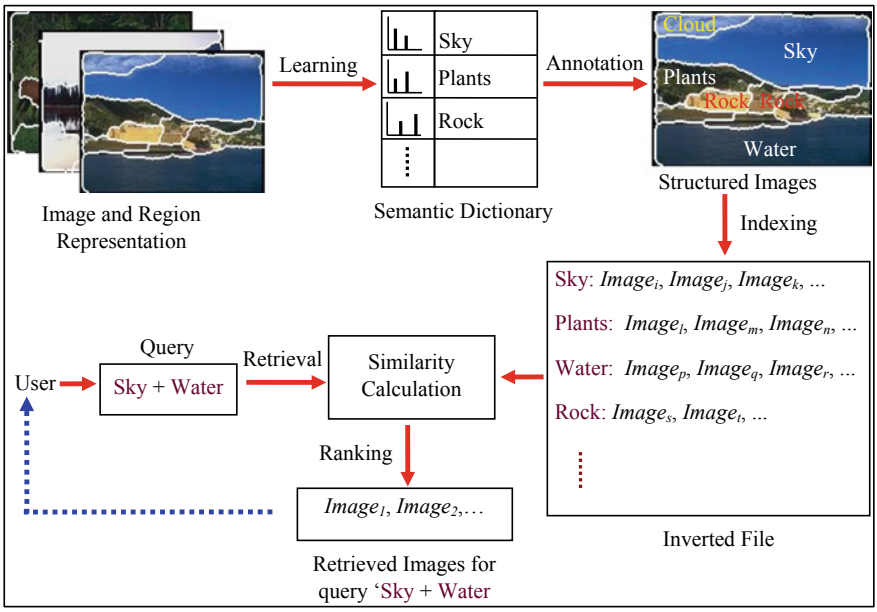

Fig. 13.13 A content-based QBK system

One of the most widely used QBK systems is Google's image search as shown in Fig. 13.14a [5]. The key features of Google's image search are summarized as follows:

- **Simplicity**. The textual input is the simplest and most intuitive query interface. It is a universal user interface.
- **Caption browsing**. Retrieved images are presented as caption browsing which is simple and intuitive.
- **Category browsing**. The categories at the top of the browsing area provide users with category browsing (Fig. 13.14b).
- **Content browsing**. QBE is included as an option for the QBK interface, shown as a camera icon at the query field (Fig. 13.14a). Some of the category browsings are also based on content features such as colors.
- **Interactivity**. An enlarged version of an image is shown as focus when a user clicks any of the images in the browsing area; the enlarged image is also linked to the image source (Fig. 13.15).
- **Hierarchical structure**. When a user clicks an image in the browsing area, it also shows the category of images similar to the clicked image at the right-hand side (Fig. 13.15). This kind of interaction can also be regarded as a QBK-supported QBE.

At this moment, Google's image search is predominantly done by QBK due to the limited capability of CBIR techniques. However, it is often a challenge for a user to formulate an image query using keywords only. In many cases, a query for

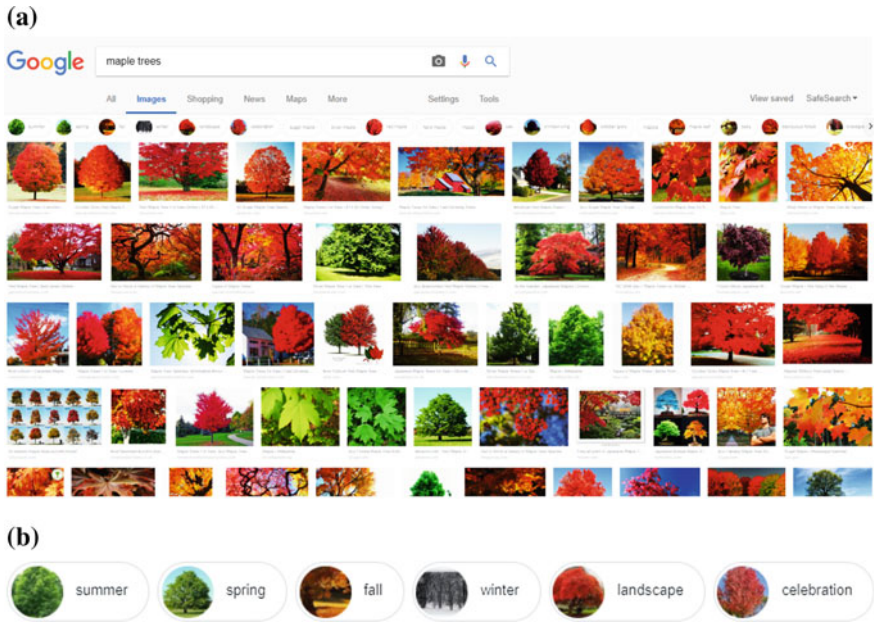

Fig. 13.14 A QBK example from Google. **a** An example from Google's Image search; **b** category browsing from Google's Image search

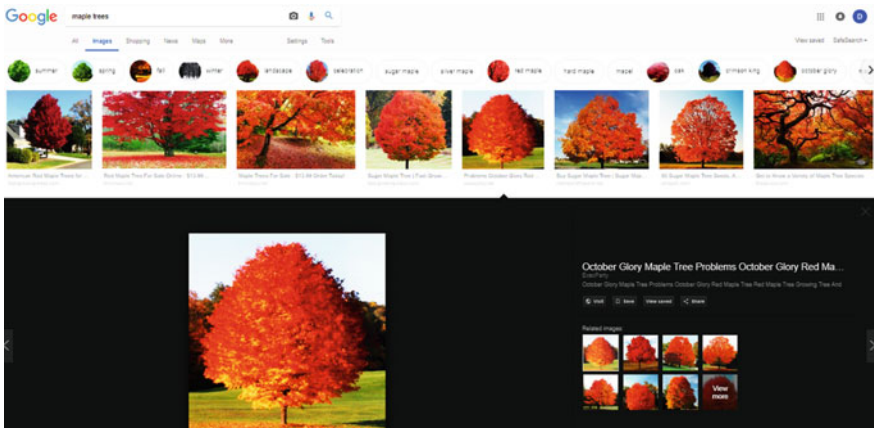

Fig. 13.15 A hierarchical QBK in Google. An enlarged image at the black strip and its relevant category at the right-hand side of the enlarged image

an image can only be described by a visual form or QBE. As CBIR techniques become more mature, it is expected that both QBE and QBK will play equal roles in image search.

An effective *relevance feedback* or RF mechanism is an essential feature of an image retrieval system. This is because an initial retrieval result is likely inaccurate; the RF provides a user a chance to interfere and refine the retrieval result. RF works

by letting a user to select a number of relevant images from a retrieval list; the IR system then uses these relevant images as a collective query (instead of a single query) to refine or update the retrieval list so that it meets the user's expectation. RF is particularly useful for mass scale image search such as Google, because the chance of identifying relevant images from the top of a retrieval list is high.

13.7 Summary

In this chapter, different image presentation techniques have been discussed. Once images in an image database have been classified, indexed, and ranked, the next step is how to present the images to users. We have introduced a number of common methods of image presentation including caption browsing, category browsing, content browsing, QBE, QBK, etc. Each of them has its applications and limitation. As image databases become larger and larger, an image retrieval system needs to integrate all of these individual image presentation techniques so that images can be found both effectively and efficiently.

An example of such kind of integrated image retrieval system has been shown using Google image search application. It is shown that Google image search has integrated caption browsing, category browsing, hierarchical category browsing, QBE, and QBK into a single IR system. The success of Google's search engine shows that in terms of design a retrieval system, the simplest is the best. The chapter also demonstrates that powerful image visualization techniques such as a Google earth like image map can be utilized to enhance user experience on an IR system.

References

1. WordNet. https://wordnet.princeton.edu/. Accessed Feb 2019
2. Schöffmann K, Ahlström D (2011) Similarity-based visualization for image browsing revisited. In: IEEE international symposium on multimedia
3. Quadrianto N, Kersting K, Tuytelaars T, Buntine WL (2010) Beyond 2D-grids: a dependence maximization view on image browsing. In: Proceedings of the international conference on multimedia information retrieval, pp 339–348,
4. Wang C et al (2015) Similarity-based visualization of large image collections. Inf Vis 14 (3):183–203
5. Google.com. Accessed Oct 2018
6. Islam M (2009) SIRBOT—semantic image retrieval based on object translation. PhD thesis, Monash University
7. Zhang D, Islam M, Lu G (2013) Structural image retrieval using automatic image annotation and region based inverted file. J Vis Commun Image Represent 24(7):1087–1098