

An Improved Presentation Method for Relevance Feedback in a Content-based Image Retrieval System

Feng-Cheng Chang¹ and Hsueh-Ming Hang^{1,2}

¹Dept. of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan

²Dept. of Computer Sci. and Inform. Technology, National Taipei University of Technology, Taipei, Taiwan
breeze@alumni.nctu.edu.tw, hmhang@ntut.edu.tw

Abstract

Content-based image retrieval (CBIR) has been adopted as a complementary technique to the keyword-based image search. Relevance feedback (RFB) is considered as an effective means to bridge the gap between the designated features and the run-time semantics on a CBIR system. Like many other interactive system, a good user interface, which in RFB is mainly the presentation of the query results, is an important factor that affects the quality of feedbacks. By incorporating the multidimensional scaling (MDS) and the outlier detection techniques, we propose a method of presenting multiple image icons in this paper. It preserves the high-dimensional distance information and shows the selected images at proper 2D locations. Viewing the matched images with distance cues, users are able to give effective and productive feedbacks. We design and implement this presentation system and show some subjective results at the end.

1. Introduction

With the advances in consumer electronics, digital contents can be easily acquired, and thus the number of digital images increases dramatically everyday. Therefore, the image database applications, such as personal photo album and arts gallery, become more and more popular. To search a desired image in a large collection of images, the content-based image retrieval (CBIR) technique has been adopted as a technique to complement the popular keyword-based technique.

There are several factors which affect the retrieval accuracy of a CBIR system. For example, the designated image features limits the expressiveness, and the prefer-

ence deduction capability affects the precision of session-specific semantics. In many CBIR systems, feature selection is mostly done at design-time. Thus, to produce the run-time semantics using the designated features is a way to improve the retrieval accuracy. It has been known that relevance feedback (RFB)[9] is an effective technique to bridge the gap between high-level semantics and low-level features. One of the major issues of RFB is how to guide the user to give effective feedbacks to the system. The ordinary rank-list does not provide explicit clues in matching distance.

In this paper, we focus on the design of presenting the matching results. In Sec. 2, the motivation and our design goals are described. In Sec. 3, we discuss several techniques used in our improved presentation scheme. In Sec. 4, we show subjective results of our implementation, and discuss the effectiveness of our design. Finally, we conclude this work with Sec. 5.

2. Motivations and Design Goals

There are several ways to design and implement an RFB scheme. Typically, an RFB-enabled CBIR system iteratively performs the following steps: (1) analyzes the input images (and the feedback information) to derive the run-time parameters; (2) calculates the matching; (3) presents the matching results to the user; and (4) receives the user feedback information and starts another iteration. One key element in the above process is how the matching results are presented to the users so that the next iteration results would meet the user expectation faster and better.

Some of the published RFB schemes are straightforward in that the displayed instances are directly obtained from the matching results[4]. Some other approaches are complicated. For example, a few extra instances are shown together with the matching instances. These instances are more informative in terms of clustering computation. If the user happens to choose one of them, the system learns more

This work was supported in part by the NSC, Taiwan under Grants NSC 95-2221-E-009 -071 -MY3, NSC 96-2422-H-194 -002, and NSC 96-2422-H-007-003.

information from that feedback[8].

No matter how complex the RFB implementation is, the presentation of the best matched images is inevitable. A common method is to render the results as an ordered list according to their similarity to the query. The advantage of this method is simple and straightforward. It is easy to arrange all the images in a 2D display area. However, the distance (closeness) information among matched objects is discarded in this presentation. It does not give users intuitive interpretation on about how the displayed instances are related. We can improve the presentation by including the distance information. This helps the users to give more precise feedbacks that adjust the distance function (matching criterion) effectively. For example, if two subjectively similar images are displayed far apart, it hints that the distance function should be adjusted so that their distance is narrowed down. That is, the user would select them as positive feedback instances. On the other hand, two images close in display but dissimilar in content should get negative feedback so that this misrepresented distance can be corrected.

In summary, we like to design a screen presentation format that provides users clear clues in giving effective feedbacks in a CBIR system. We assume the underneath CBIR system is distance-based. Our presentation format should achieve the following goals: (1) it shows the relative distances (closeness) among all the presented instances on the screen; (2) few outliers that are away from the majority should be remapped or removed; (3) the displayed images are viewed comfortably and can be accessed easily; and (4) the computational overhead is affordable.

3. Proposed Presentation Method

Based on the aforementioned assumption and goals, the entire design consists of three items. The first item is how to embed the distance information into the displayed instances; the second is a method to eliminate the singular instances so that the 2D real-estate (screen) is efficiently used; and the third is a well-designed user interface.

3.1. Multidimensional Scaling

To present high-dimensional image features in a 2D display, we have to reduce the dimensionality of the feature vectors. But the high-dimensional distance should be preserved on the 2-D screen. A popular transform is the principal component analysis (PCA). In PCA, the similarity among data objects is expressed by a correlation matrix. However, the definition of the feature distance in CBIR is not always Euclidean. Therefore, a complete correlation matrix may not be always available.

Multidimensional scaling (MDS) is another technique used in some image database browsing applications[6]. MDS can be an alternative to PCA. It is flexible in that almost any kind of similarity matrix can be used including the non-metric (e.g., ordinal) information analysis. The basic concept of MDS is to re-arrange all the data objects in the low-dimensional space such that the observed (high-dimensional) similarity (distance) is best reserved. There are many approaches to find the MDS configuration of a given set of data set[7][2]. Some of them are one-shot, and some of them are iterative; some of them require complete similarity matrix, and some do not (i.e., part of the similarity values can be undefined).

In Sec. 2, we assume that a distance-based matching function is employed in this CBIR system. In other words, the matching distance between any two feature objects can be calculated. Our task is to find the 2D coordinates based on the mutual distances. To simplify the implementation, we adopt the method proposed in [2]. It is a metric MDS for transforming a distance matrix to a well-represented set of Euclidean coordinates. The operations of the MDS transformation are briefly described as follows.

Assume that we have N instances, and the distance function for instances i and j is denoted by $d(i, j)$. The first step is to prepare the $N \times N$ distance matrix \mathbf{D} . For each pair of (i, j) , we compute $d(i, j)$ and assign the value to the D_{ij} and D_{ji} components in \mathbf{D} . According to the nature of distance functions, all the diagonal components D_{kk} should be 0 (i.e., \mathbf{D} is hollow).

The second step is to prepare the $N \times 1$ mass vector \mathbf{m} . Each value m_k in \mathbf{m} represents the mass (importance) of the corresponding instance k , and the sum of all m_k should be 1. If all the instances are equally important, all $m_k = \frac{1}{N}$.

The third step is to compute the $N \times N$ centering matrix $\mathbf{\Xi}$ as

$$\mathbf{\Xi} = \mathbf{I} - \mathbf{1} \times \mathbf{m}^\top,$$

and obtain the (equivalent) cross-product matrix

$$\mathbf{S} = -\frac{1}{2}\mathbf{\Xi}\mathbf{D}\mathbf{\Xi}^\top.$$

The fourth step is to find the eigen-decomposition of \mathbf{S} , which gives

$$\mathbf{S} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top,$$

where

$$\mathbf{U}\mathbf{U}^\top = \mathbf{I} \text{ and } \mathbf{\Lambda} \text{ is diagonal matrix of eigenvalues.}$$

The final step is to derive the MDS-transformed coordinates by

$$\mathbf{F} = \mathbf{M}^{-\frac{1}{2}}\mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}, \text{ where } \mathbf{M} = \text{diag}\{\mathbf{m}\}.$$

For each row k in \mathbf{F} , (F_{k1}, F_{k2}) is the location of instance k in 2D display space.

3.2. Outlier Elimination

After mapping distances to 2D coordinates, we can view the query results as a set of image icons distributed on the screen. An example is shown in Fig. 1. However, the outputs of MDS may not always be appropriate for presentation purpose. For example, we retrieve 50 similar images from the database for a given query image. There are 48 similar ones in the database, but the top 50-list contains two dissimilar instances. In other words, there are two “singular” items in the displayed instances. The singular instances are far away from the main cluster, and thus the displayed icons are crowded and overlapped. Most of the screen is empty.

In statistics, an observation that is not coherent to the others is called an outlier[3]. From one viewpoint, outliers can be treated as errors or noises of the data set. We simply remove them to give a better presentation. But from the other viewpoint, outliers may contain useful information for showing the property of the entire data set. We like to keep them in the case when the user wants to examine them. MDS browsing provides a good solution to the conflict issue: the modified (outlier-free) view is treated as the zoom-in version of the original (outlier-included) view.

To identify outliers, either single-variate outlier detection method or multi-variate outlier detection method can be used. However, with a few experiments, we found that the following single-variate outlier detection method is effective in our case. Firstly, the centroid of the major cluster is estimated by

$$X_c = med(x_1, x_2, \dots, x_N), Y_c = med(y_1, y_2, \dots, y_N).$$

Secondly, the distance between any instance to the centroid is computed to form a list of distances ($\{d_1, d_2, \dots, d_N\}$). Then, we apply the Hampel’s identifier[5] to detect the outlier. The median distance is defined as

$$d_m = med(\{d_1, d_2, \dots, d_N\}).$$

The median absolute deviation (MAD) is defined as

$$MAD = med(\{|d_1 - d_m|, |d_2 - d_m|, \dots, |d_N - d_m|\}).$$

As proposed in [5], for 5% outlier probability with $N = 50$, the outliers are which satisfy

$$|d_k - d_m| > 4.31 * MAD.$$

For outlier removal, there are known problems of masking effect and swamping effect. In the former case, the outlier O_2 is existent only when the outlier O_1 is removed. In the latter case, the outlier O_2 is existent only when the outlier O_1 is existent. These effects make it difficult to detect multiple outliers at once. Thus, we iteratively remove the one that has the maximum deviation in the Hampel’s identifier, until none are detected.

3.3. User Interface

Using the techniques proposed in 3.1 and 3.2, we can display image icons on a 2D area informatively (with relative distances) and efficiently (outlier removed). However, during a number of subjective tests, we found that the overlapping problem is inevitable and sometimes obscures the manipulation.

Even when we have an outlier-free distribution, some of the image icons may be co-located nearly on the same spot. There are two issues resulting from this problem. The first one is how to distinguish multiple icons stacked together. The second is how to ensure that an occluded icon can easily be seen and accessed. For the former issue, we adopt a skill saw in photo albums. We randomly rotated each icon to make it distinguishable from nearby ones. To prevent subjectively misleading, we restrict the rotation angle in the range of $\pm 60^\circ$. For the latter issue, we push the previously focused icon to the bottom, before raising the currently focused icon to the front. This ensures that a totally covered icon can be piled up and exposed, by a certain number of focus-unfocus operations on the covering icons.

3.4. Integration

With the techniques described in the above sections, the proposed presentation method for a CBIR system with interactive feedback is summarized as follows.

1. Perform the query.
2. Compute the mutual distances between each pair of the return results.
3. Compute the 2D locations for the output instances (results) using MDS (Sec. 3.1).
4. Generate the 2D presentation using the user-interface techniques (Sec. 3.3).
5. Apply the outlier detection technique (Sec. 3.2) to identify the outliers in the current view. If an outlier exists, remove the maximum-deviated one. Go to step 4 for zoom-in.
6. Stop the presentation process, and wait for the user feedback.

4. Implementation and Discussion

To examine the effectiveness of the proposed presentation method, we integrate it with a simple CBIR system. The database contains 8689 Corel images in 68 categories. The matching engine incorporates three MPEG-7 visual descriptors: color layout, scalable color, and edge histogram[1].

Since we focus on the presentation, we simply combine the three feature distances with equal weights to produce the

final distance metric. This approach implies that the similarity is not optimized to semantics, such as objects or segments. After the initial query, the top-50 results are shown in Fig. 1. These image icons are distributed with clear image feature clues. Most darker images are located toward the left side; the more colorful images are located toward the upper border; the less textured images are located on the right; and the similar images are roughly located closer to the query instance (the pink-boxed one).

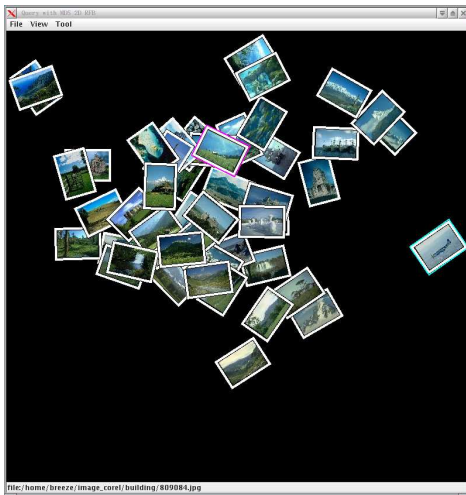


Figure 1. Initial query results.

However, one may notice that the center is crowded due to an outlier laid near the right border. After applying the Hampel's identifier, the outlier (the rightmost image) is detected, and the outlier-free view is shown in Fig. 2. It is obvious that the distributions along the vertical and horizontal directions are now more balanced. The icon-rotation technique makes an icon rarely fully-overlapped. It is not difficult to identify that several images are very close to each other.

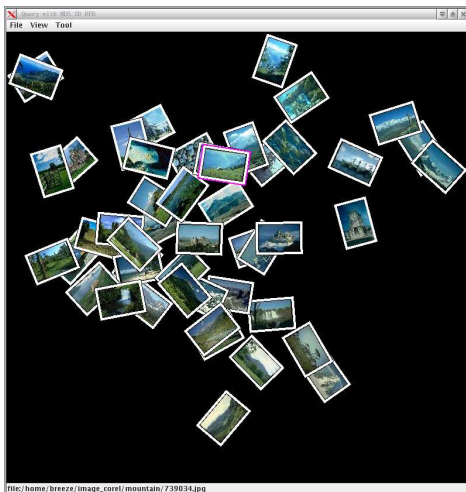


Figure 2. Results with outliers removed.

nique makes an icon rarely fully-overlapped. It is not difficult to identify that several images are very close to each other.

For the computation complexity, this method is more complicated than the rank-list based presentation. However, since the number of instances for presentation is not large, the overhead of the computation is affordable on an ordinary PC. In our experiments, the matching takes about 25 seconds, while the presentation takes about 1 second.

5. Conclusions

In this paper, we proposed an improved presentation method for relevance feedback in a CBIR system. The assumption is that a good presentation of the query results can guide the user to give effective feedbacks. The MDS technique is incorporated to transform the high-dimensional mutual distances to 2D locations. The outlier detection technique is adopted to improve the use of the valuable 2D space. The rotated icons with push-to-bottom rule guarantee that the overlapped icons can be viewed and accessed easily. We carefully adjusted these schemes and integrated them together with a simple CBIR system. The subjective results show that the design goals are achieved and this approach effectively improves the presentation, comparing to the conventional rank-list approach.

References

- [1] ISO/IEC JTC1/SC29/WG11, FDIS N4203. MPEG Committee, Jul. 2001.
- [2] H. Abdi. Metric multidimensional scaling (mds): Analyzing distance matrices. In N. J. Salkind, editor, *Encyclopedia of Measurement and Statistics*. SAGE, 2007.
- [3] I. E. Ben-Gal. Outlier detection. In O. Maimon and L. Rokach, editors, *The Data Mining and Knowledge Discovery Handbook*, pages 131–146. Springer, 2005.
- [4] F.-C. Chang and H.-M. Hang. A relevance feedback image retrieval scheme using multi-instance and pseudo image concepts. *IEICE Trans. on Information and Systems*, E89-D(5):1720–1731, May 2006.
- [5] L. Davies and U. Gather. The identification of multiple outliers. *Journal of the American Statistical Association*, 88(423):782–792, Sep. 1993.
- [6] Y. Rubner, C. Tomasi, and L. J. Guibas. A metric for distributions with applications to image databases. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 59, Washington, DC, USA, 1998. IEEE Computer Society.
- [7] F. W. Young. Multidimensional scaling. In Kotz-Johnson, editor, *Encyclopedia of Statistical Science*, volume 5. John Wiley & Sons Inc., 1985.
- [8] C. Zhang and T. Chen. An active learning framework for content-based information retrieval. *IEEE Trans. Multimedia*, 4(2):260–268, Jun. 2002.
- [9] X. S. Zhou and T. S. Huang. Relevance feedback in image retrieval: A comprehensive review. *Multimedia Systems*, 8(6):536–544, April 2003.