Perceptual Consistency Improves Image Retrieval Performance

Huizhong Long, Wee Kheng Leow
School of Computing, National University of Singapore
3 Science Drive 2, Singapore 117543
longhuiz, leowwk@comp.nus.edu.sg

ABSTRACT

An ideal retrieval system should retrieve images that satisfy the user's need, and should, therefore, measure image similarity in a manner consistent with human's perception. However, existing computational similarity measures are not perceptually consistent. This paper proposes an approach of improving retrieval performance by improving the perceptual consistency of computational similarity measures for textures based on relevance feedback judgments.

1. INTRODUCTION

An ideal retrieval system should retrieve images that satisfy the user's need. It should, therefore, measure image similarity in a manner consistent with human's perception. However, existing computational similarity measures are not perceptually consistent [3, 6]. This problem leads to retrieval results that do not always meet the users' expectations [7]. To improve retrieval performance, relevance feedback technique is often used to tune computational similarity measures [4, 5, 8]. Typically, each new query resets the similarity measure back to its initial state, which is not perceptually consistent. Subsequent feedback for the query are used to adjust the weighting factors of the similarity measure to improve retrieval performance.

The main difficulty with this method is that very few users are willing to go through endless iterations of feedback in hope of retrieving the best results. A successful relevance feedback process must yield positive results within three or four iterations [1]. So, feedback methods that require many iterations to improve retrieval performance are not practically useful. Another shortcoming of this method is that previous feedback results are typically not retained in the system. Each new query starts with a similarity measure that is not perceptually consistent. The users have to go through the relevance feedback process even if the same feedback information has been given in the past. This problem is partially alleviated with user profiling.

SIGIR'01, September 9-12, 2001, New Orleans, Louisiana, USA.

This paper addresses the root of the problem of low retrieval performance: perceptual inconsistency. It proposes an approach of improving retrieval performance by improving the perceptual consistency of computational similarity measures. It uses relevance feedback judgments to incrementally transform a computational space into a perceptual space. Image similarity measured in the perceptual space would be perceptually consistent. This approach could be applied to both document and image retrieval, though this paper focuses mainly on texture image retrieval.

2. ADAPTIVE PERCEPTUAL SPACE

In [2, 3], we have demonstrated that commonly used texture features and similarity measures are not consistent with human's perception, and that mapping texture features to the perceptual texture space improves the perceptual consistency of computational features as well as their retrieval performance. This paper introduces an incremental algorithm for constructing the perceptual space using the relevance feedback judgments provided by the users.

Let $\mathcal{P}$ denote a perceptual space in which every image i is assigned a point $\mathbf{x}_i$ and the Euclidean distance between image points $\mathbf{x}_i$ and $\mathbf{x}_j$ in $\mathcal{P}$ is equal to the perceptual distance p_{ij} between images i and j. Then, the problem of incremental construction becomes one of computing the perceptual distances p_{ij} from cumulative feedback. The incremental construction algorithm can be summarized as follows:

1. Initialize image space to a computational space.
   Initialize each p_{ij} to a computational distance, e.g., the Euclidean distance between images i and j. Apply multidimensional scaling (MDS) to compute the coordinates $\mathbf{x}_i$ of each image i from the distances p_{ij}.

2. Initialize co-occurrence and feedback matrices.
   Initialize the entries c_{ij} of *co-occurrence matrix* $\mathbf{C}$ and b_{ij} of *feedback matrix* $\mathbf{B}$ to zero. Matrix $\mathbf{C}$ accumulates the relevance feedback results and $\mathbf{B}$ indicates whether feedback results have been received.

3. Update co-occurrence and feedback matrices.
   Let q denote a query image, and $\mathcal{R}$ and $\mathcal{I}$ the sets of relevant and irrelevant images marked by a user. Update the matrices as follows:

$$\begin{array}{ll} c_{qi} \leftarrow c_{qi} + 1 & \text{for each image } i \in \mathcal{R} \\ c_{qi} \leftarrow c_{qi} - 1 & \text{for each image } i \in \mathcal{I}, \text{if } c_{qi} > 1 \\ b_{qi} \leftarrow 1 & \text{for each image } i \in \mathcal{R} \cup \mathcal{I} \end{array}$$

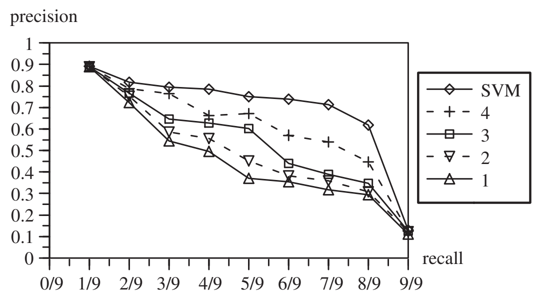

Figure 1: Precision-recall curves of the incremental space at stages 1 to 4. As information coverage increases, the curve shifts towards the upper bound achieved by direct mapping to PTS using SVM.

This accumulation step can be executed several times before going on to the next step.

4. Update perceptual distances.
This step involves merging the new distance measurements $d_{ij} = n - c_{ij}$ with the old perceptual distances p_{ij}. The variable n denote the total number of feedback iterations executed so far. To merge them successfully, d_{ij} and p_{ij} must be normalized to the same scale. This is achieved by equalizing the population means and variances of d_{ij} and p_{ij}. After equalization, p_{ij} is replaced by d_{ij} if $b_{ij} = 1$.

5. Update perceptual space.
Re-compute the coordinates $\mathbf{x}_i$ of each image i from the distances p_{ij} using MDS.

Steps 3 to 5 are repeated for new feedback results.

2.1 Retrieval Tests

Experiments were conducted to test the algorithm's performance. Figure 1 plots the precision-recall curves at various relevance feedback stages. Ten feedback iterations, each containing non-redundant information, were performed between each stage. As more and more feedback results were received, information coverage (i.e., information conveyed by the feedback judgments) increased and retrieval performance approached the upper bound achieved by directly mapping computational features to the perceptual texture space using support vector machine (SVM) [2, 3]. The perceptual texture space was constructed using human judgment data collected in psychological experiments [2, 3].

The incremental construction method shows an interesting phenomenon called *phase shifting*. Initially, the perceptual consistency of the texture space was as low as that of a computational space (e.g., Euclidean and scaled Euclidean [9] spaces). As more feedback results were received, the space shifted towards a perceptual space (Fig. 2). This phase shifting property offers another useful advantage. As a user's relevance judgment changes over time, the space will also change accordingly, thus adapting to the user's changing need.

3. CONCLUSION

This paper presented an incremental method of measuring perceptual distances and constructing a perceptual space using relevance feedback judgements. The method requires only a small amount of relevance judgments at each feedback iteration. It accumulates the feedback information over different queries to construct the perceptual space. The constructed space exhibits phase shifting from a purely computational space towards a perceptual space as more and more feedback results are received. This property confirms that the method can indeed construct a perceptual space incrementally. If the feedback judgments are provided by the same user, then the perceptual distances measured would be consistent with a single user's perception. Otherwise, the measurements would reflect the average perception of typical users. In the case of a single user, the measurements would eventually stabilize if the user's judgment remains consistent over time. Otherwise, the measurements would adapt to the changes in the user's judgment.

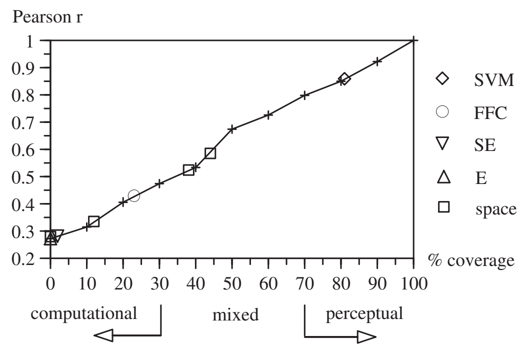

Figure 2: Phase shifts of texture space. The line indicates the perceptual consistency at various percentage of information coverage. The labels mark the positions of various similarity measures (E: Euclidean, SE: scaled Euclidean, FFC [9], and SVM mapping [3]) according to perceptual consistency.

4. REFERENCES

[1] R. R. Korfhage. *Information Storage and Retrieval*. John Wiley & Sons, 1997.

[2] H.-Z. Long and W. K. Leow. Invariant and perceptually consistent texture mapping for content-based image retrieval. In *Proc. ICIP*, 2001.

[3] H.-Z. Long and W. K. Leow. Perceptual texture space improves perceptual consistency of computational features. In *Proc. IJCAI*, 2001.

[4] W. Y. Ma and B. S. Manjunath. Texture features and learning similarity. In *Proc. IEEE CVPR '96*, pages 425–430, 1996.

[5] T. P. Minka and R. W. Picard. Interactive learning using a "society of models". In *Proc. IEEE CVPR '96*, pages 447–452, 1996.

[6] J. S. Payne, L. Hepplewhite, and T. J. Stonham. Perceptually based metrics for the evaluation of textural image retrieval methods. In *Proc. IEEE ICMCS '99*, pages 793–797, 1999.

[7] Y. Rui and T. Huang. Optimizing learning image retrieval. In *Proc. CVPR*, 2000.

[8] Y. Rui, T. S. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in MARS. In *Proc. ICIP '97*, 1997.

[9] S. Santini and R. Jain. Similarity measures. *IEEE Trans. PAMI*, 21(9):871–883, 1999.