# 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. Therefore, it should compare the similarity between images in a manner consistent with human's perception. However, existing computational similarity measurements are not necessarily consistent with human's perception. This paper proposes an approach of improving retrieval performance by improving the perceptual consistency of computational features and similarity measurements. A novel method for incrementally measuring perceptual distances and constructing perceptual space based on relevance feedback is presented. Only a small number of relevant judgments is required in each feedback. Feedbacks from multiple queries are accumulated and incrementally updates the measurements of perceptual distances between images. As more and more feedbacks are received, the constructed space becomes more consistent with human's perception. Test results show that the distances measured by this incremental method correlates very well with perceptual distances measured in psychological experiments. Moreover, the more perceptually consistent are the distance measurements, the better is the image retrieval performance.

## 1. INTRODUCTION

The end user of an image retrieval system is human. The retrieval system should retrieve images that satisfy the user's need. Therefore, it should compare the similarity between images in a manner consistent with human's perception. However, existing computational similarity measurements are not necessarily consistent with human's perception [12, 15]. This problem leads to retrieval results that do not always meet the users' expectations [19].

To improve retrieval performance, relevance feedback technique is often used to tune computational similarity measurements [6, 13, 14, 20, 22, 26]. Typically, each new query resets the similarity measurement back to its initial state, which is not perceptually consistent. Subsequent feedbacks for the query are used to adjust the weighting factors of the similarity measurement 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 hopes of retrieving the best results. A successful relevance feedback process must yield positive results within three or four iterations [9]. 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 measurement 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.

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 features and similarity measurements. This approach can be applied to both document and image retrieval, though this paper focuses mainly on image retrieval.

A direct method of improving perceptual consistency is to construct a perceptual space of images (or documents) using psychological experiments such as those in [8, 16]. The Euclidean distances measured in this perceptual space would be consistent with human's judgments. Then, images (or documents) can be mapped to the perceptual space and retrieval performed in the perceptual space would yield results that are consistent with human's judgments. This method has been explored for the retrieval of single texture images and has yielded good retrieval results [11].

The above direct approach is appropriate if the construction of the perceptual space involves a small data set, such as the 100 or so images in the Brodatz album. For general image retrieval applications, it is not feasible to construct a perceptual space using thousands of images because it is practically impossible to conduct psychological experiments involving such a large number of images.

This paper presents a novel method for incrementally mea-

This is an extended version of the paper to appear in SIGIR 2001.

suring perceptual distances and constructing perceptual space based on relevance feedback. Only a small number of relevant judgments is required in each feedback. Feedbacks from multiple queries are accumulated and incrementally updates the measurements of perceptual distances between images. If the feedbacks 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 relevant judgment remains consistent over time. Otherwise, the measurements would adapt to the changes in the user's relevant judgment.

In the remainder of this paper, existing works related to the above issues are discussed (Section 2). A precise definition of perceptual consistency is also given (Section 3). This is followed by retrieval experiments of single texture images that illustrate the fact that better perceptual consistency leads to better retrieval performance (Section 4). Next, the incremental method for computing perceptual distance and constructing perceptual space is presented (Section 5). This is followed by an application example that illustrates the use of the method (Section 5.2). Finally, this paper concludes with a summary of the results (Section 6).

## 2. RELATED WORK

The experiments performed by Payne et al. [15] on retrieving Brodatz textures show that all the computational texture features tested in their experiments usually produce retrieval results that are not always consistent with human judgments. Independent tests performed by Long and Leow [12] also confirm that computational texture features and similarity measurements are not very consistent with human's perception. Indeed, the study of human perception of image content is identified as one of the most important topics in image retrieval [21].

A common technique for improving retrieval performance is to adjust system parameters using users' feedbacks. The adjustable system parameters are weights assigned to image features (or keywords and terms, in the case of document retrieval) according to the relative importance of the features. There are three simple methods that are commonly used to adjust the weights. The first method is the so-called *tf*$\times$*idf* method [23, 25]. It sets the weight of a query feature or term according to the product of *term frequency tf* and *inverse document frequency idf*. Robertson and Sparck Jones [17] derived a variant of the *tf*$\times$*idf* formula using a probabilistic theory of relevance weighting. The third method uses Rocchio weighting formula [5, 18] which modifies a term weight in the query in proportion to the positive sum of the corresponding term weights in relevant documents and the negative sum of term weights in irrelevant documents.

The above methods tend to be somewhat ad-hoc or provide only partial solutions [20]. Rui and Huang [20] proposed an optimization approach that optimizes for the parameters of generalized Euclidean distance, the query vector, and term weights.

In practice, it is not feasible to require a user to provide feedback for a large number of retrieved images or documents. To solve this problem, incremental relevance feedback techniques were proposed by Aalbersberg [1] and Allan [3]. These methods adjust term weights using only a few judged documents each time. Test results show that these methods can achieve good performance and is useful for handling long-standing queries with drifting notions of relevance.

Our method is similar to incremental relevance feedback techniques in that only a small number of judgments is required by the user each time. However, instead of adjusting system parameters directly, our method uses feedback information to compute perceptual distances between images. The computed distances are then used to construct a perceptual space of the images.

## 3. PERCEPTUAL CONSISTENCY

There are many ways of defining perceptual consistency. This section discusses some definitions which will be used throughout the paper. Let $p_{ij}$ denote the perceptual distance between samples $i$ and $j$, and $d_{ij}$ denote the corresponding measured or computational distance. A simple notion of perceptual consistency is that $d_{ij}$ is *proportional* to $p_{ij}$. That is, there exists a linear function $f$ such that

$$p_{ij} = f(d_{ij}), \quad \forall i, j \ . \tag{1}$$

Then, perceptual consistency can be measured in terms of the mean squared error (MSE) $e$ of linear regression:

$$e = \frac{1}{N} \sum_{i,j} (p_{ij} - f(d_{ij}))^2 \tag{2}$$

where $N$ is the number of sample pairs. The smaller the MSE, the better is the consistency. A perfect consistency has an MSE of 0.

A less stringent notion of perceptual consistency is to require that $f$ be a monotonic function which can be nonlinear. The difficulty with this definition is that it is difficult to determine the best nonlinear function to use in practice.

An alternative definition is to require that $d_{ij}$ be *statistically correlated* to $p_{ij}$. In this case, it is useful to transform the populations $\{d_{ij}\}$ and $\{p_{ij}\}$ to equivalent zero-mean unit-variance populations $\{d'_{ij}\}$ and $\{p'_{ij}\}$:

$$p'_{ij} = \frac{p_{ij} - \bar{p}}{\sigma_p}, \quad d'_{ij} = \frac{d_{ij} - \bar{d}}{\sigma_d} \tag{3}$$

where $\bar{p}$ and $\bar{d}$ are the means and $\sigma_p$ and $\sigma_d$ are the standard deviations of the populations. Then, perceptual consistency can be measured in terms of the correlation $r$:

$$r = \sum_{i,j} p'_{ij} \, d'_{ij} \ . \tag{4}$$

Substituting Eq. 3 into Eq. 4 yields the Pearson's correlation coefficient:

$$r = \frac{\displaystyle\sum_{i,j} (p_{ij} - \bar{p})(d_{ij} - \bar{d})}{\left[ \displaystyle\sum_{i,j} (p_{ij} - \bar{p})^2 \sum_{i,j} (d_{ij} - \bar{d})^2 \right]^{1/2}} \ . \tag{5}$$

The coefficient $r$ ranges from $-1$ to $+1$ (for perfect correlation). In the following sections, we will use both MSE and Pearson's correlation coefficient to measure two different aspects of perceptual consistency. Note that, with perfect consistency ($e = 0$ or $r = 1$), we obtain the following condition:

$$d_{ij} \leq d_{kl} \Rightarrow p_{ij} \leq p_{kl} \quad \text{for any textures } t_i, t_j, t_k, t_l. \tag{6}$$

That is, if perfect consistency is achieved, computational similarity would imply perceptual similarity.

# 4. PERCEPTUAL CONSISTENCY IMPROVES RETRIEVAL PERFORMANCE

This section compares the retrieval performance of various computational similarity measurements that vary in their degrees of perceptual consistency. To assess perceptual consistency, we construct a perceptual texture space (PTS) to provide the ground truth.

Instead of constructing a PTS, it is also possible to measure human's perception of texture similarity by having human subjects rank textures according to subjectively predefined properties such as directionality, coarseness, busyness, and etc. [2, 27] However, the subjects' interpretations of the meaning of these visual properties are expected to vary from one person to the other. Therefore, it is uncertain whether the individual ranking results can be combined into group ranking results that represent the perception of a typical person. Ranking along predefined properties may also bias subjects' judgments towards these properties rather than other potentially salient features [16]. Moreover, recent psychological studies show that subjective texture properties cannot be reliably associated with the dimensions of a perceptual texture space [8]. Therefore, instead of ranking textures according to predefined properties, we will construct a PTS using an experimental procedure similar to those of [8, 16].

## 4.1 Development of Perceptual Texture Space

60 subjects were asked to perform a free sorting task on 60 texture images according to their perceptual similarity. These 60 images consisted of 50 textures selected from the Brodatz album, and 10 other natural textures (such as foliage and roof tiles) not contained in the Brodatz album. Other textures in the Brodatz album were not selected because they were either scaled versions of the 50 selected textures, or their contents (e.g., flowers) were much stronger than their texture patterns. It is very difficult for the subjects to avoid classifying the images based on their contents. The images are normalized to canonical intensity, contrast, scale, and orientation [11] so as to reduce their confounding effects on human's perception of texture patterns. About half of the subjects performed the experiments twice at more than one month apart. This procedure helps to ensure the stability of the experimental data by reducing individual variations of the subjects.

After obtaining the sorting results, pairwise distances between textures were computed using Donderi's information measurement [7]:

$$p_{ij} = \sqrt{H(t_i|t_j) + H(t_j|t_i)} \tag{7}$$

where $H(t_i|t_j)$ is the *independent information* of $t_i$ after knowing the grouping of $t_j$. Independent information is the information gained by classifying two stimuli into separate groups. Let $G$ denote the set of textures used in the experiment and $G(t_i)$ denote the group in which $t_i$ belongs. The independent information $H(t_i|t_j)$ is computed as follows:

$$H(t_i|t_j) = \begin{cases} 0 & \text{if } G(t_i) = G(t_j) \\ \log_2 |G| - \log_2 |G(t_i)| & \text{if } G(t_i) \neq G(t_j) \end{cases} . \tag{8}$$

According to Eq. 7, informationally independent stimuli are dissimilar and informationally redundant stimuli are similar.

Based on the pairwise distances, multidimensional scaling was performed to derive perceptual spaces with different number of dimensions. Our study [12] showed that both 3D and 4D spaces correlate well with existing perceptual spaces of Rao and Lohse [16] and of Heaps and Handel [8]. Moreover, the 4D space fits the psychological data better than does the 3D space. Therefore, the 4D space is used in our invariant and perceptual mapping method.

## 4.2 Comparison of Retrieval Performance

An experiment was conducted to compare the retrieval performance of various computational texture models. Among the various texture features available, Gabor features are very widely used for image retrieval (e.g., [13, 10, 21]) and have been reported to best match the results of human vision study [21]. Therefore, Gabor features were used in the experiment.

The following computational models of texture similarity were compared:

- Euclidean distance:
  This is the most commonly used dissimilarity measurement.

- Fuzzy Features Contrasts Model (FFC):
  Santini and Jain developed this model [24] which has the potential of being perceptually consistent. Their similarity measurement is based on Tversky feature contrast model [28] which can account for various peculiarities of human's perceptual similarity. They have applied FFC to measure texture similarity based on Gabor features, and have obtained encouraging results. Various ranges of parameter values of FFC were tried to obtain the best retrieval performance.

- Scaled Euclidean distance:
  Scaled Euclidean distance was compared with FFC in [24].

- Perceptual Mapping:
  The most perceptually consistent way of measuring texture similarity is to first map the textures into the PTS and then measure the distances in the PTS. Support Vector Machine (SVM) was trained to perform the mapping.

In a similar test conducted by Ma and Manjunath [13], human subjects were asked to classify similar Brodatz textures into the same group. During retrieval, textures in the same

**Table 1: Correlation of various computational texture models with PTS measured in Pearson's correlation coefficient and mean squared error (MSE) of linear regression.**

| | Euclidean | Scaled Euclidean | FFC | SVM |
|---|---|---|---|---|
| Pearson | 0.273 | 0.282 | 0.430 | 0.859 |
| MSE | 0.131 | 0.121 | 0.098 | 0.019 |

group as the query were regarded as relevant. In our tests, similar grouping is employed except that the groups are derived from $k$-means clustering of the textures in PTS. Out of the 60 textures in PTS, 20 groups were derived, giving an average of 3 textures per group.

For the retrieval experiment, 9 non-overlapping samples were cropped from each of the 60 texture images used in constructing the PTS. These samples constituted a set of 540 texture images. Retrieval was performed by presenting each of all the images in turn as a query texture. Retrieval performance was measured by computing the precision-recall curve averaged over all the queries. A nine-point average precision at the recall levels of $i/9$, $i = 1, \ldots, 9$, was computed. Recall levels were set at fractions of 9 because the number of relevant images for each query was a multiple of 9.

Perceptual consistency of the four computational models was assessed by comparing the distances measured by these models with those measured in PTS. Both Pearson's correlation coefficient (Eq. 5) and the mean squared error (MSE) of linear regression (Eq. 2) were used as the measurement of consistency. The larger the Pearson's coefficient, or the smaller the MSE, the more consistent computational features are to PTS (Section 2).

Table 1 summarizes the results, and shows that perceptual mapping with SVM is most consistent with PTS. FFC is less perceptually consistent than does perceptual mapping. But, compared to Euclidean and scaled Euclidean distances, FFC is more perceptually consistent. That is, FFC indeed captures some aspects of human perception.

Figure 1 shows the precision-recall curves of the four computational texture models. Perceptual mapping with SVM achieves the best performance. It's precision remains very high ($> 0.7$) until recall level is raised to 8/9 and above. As expected, Euclidean and scaled Euclidean distances perform poorly while FFC achieves moderate performance.

Figure 2 shows a plot of precision vs. perceptual consistency at various recall levels. Perceptual consistency is measured as Pearson's correlation coefficient. It is evident from Fig. 2 that perceptual consistency is highly correlated to precision and recall levels. In particular, there is a perfect correlation between perceptual consistency and precision at the recall level of 6/9. Therefore, perceptual consistency can be used, in addition to precision and recall, as an indicator of the overall retrieval performance.

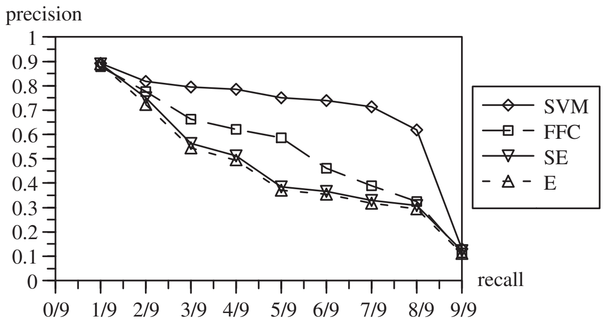


**Figure 1: Precision-recall curves of various computational texture models. E: Euclidean, SE: scaled Euclidean.**

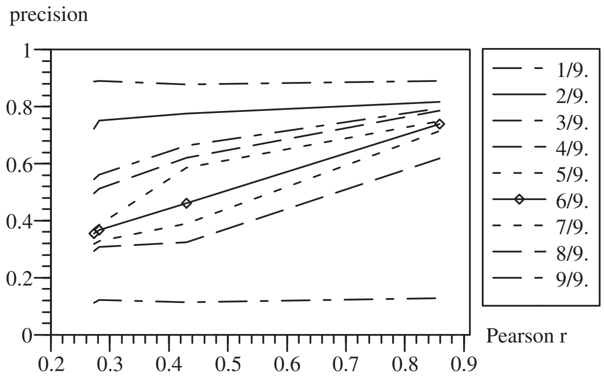


**Figure 2: A plot of precision vs. Pearson's correlation coefficient at various recall levels. Precision and recall are highly correlated with Pearson's coefficient, especially at the recall level of 6/9.**

# 5. ADAPTIVE PERCEPTUAL SPACE

The previous section demonstrated a simple method of achieving a high level of perceptual consistency, i.e., mapping texture features to the perceptual space. In practical applications, however, it may not be possible to construct a perceptual space a priori because of the large amounts and variations of images. This section presents a method for incrementally computing perceptual distances and constructing perceptual space using relevance feedback.

## 5.1 Incremental Construction Algorithm

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 feedbacks.

The incremental construction algorithm is as follows:

1. Initialize perceptual 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 $p_{ij}$. That

is, the space is initialized to a computational space.

2. Initialize co-occurrence matrix.
Initialize the entries $c_{ij}$ of matrix $\mathbf{C}$ to zero. Matrix $\mathbf{C}$ is called the *co-occurrence matrix* which is used to accumulate the relevance feedback results. Also initialize the entries $b_{ij}$ of matrix $\mathbf{B}$ to zero. Matrix $\mathbf{B}$ is called the *feedback matrix* which indicates whether relevance feedback results have been received.

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

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

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 those of $p_{ij}$. That is, the distances $d_{ij}$ whose $b_{ij} = 1$ are rescaled to a canonical mean and variance. The distances $p_{ij}$ are also rescaled to the same mean and variance. After rescaling, $p_{ij}$ 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 when new feedback results are received. Initially, the space is a computational space. As more and more feedbacks are received, the distances measured in the space become closer and closer to perceptual distances, and the space gradually transforms from a computational space to a perceptual space. Eventually, when sufficient feedback results are received, the space becomes a perceptual space.

For a matrix $\mathbf{C}$ that is very large (say, greater than 100×100), standard MDS algorithm may have problem with large matrix manipulations. In this case, incremental multidimensional scaling [4] can be used instead. In practical applications in which many entries of $\mathbf{B}$ are zero (i.e., many feedback information is not available), sparse matrix representations of the matrices can be used to reduce memory requirements.

## 5.2 Retrieval Tests

An experiment was conducted to test the effectiveness of the incremental construction algorithm. For the incremental method to be effective, it is necessary that each feedback iteration provides new relevance information that is not already captured in the previous feedbacks. A feedback that does not provide new information is not useful for the algorithm and does not yield any improvement in the retrieval performance. This constraint is also true for any incremental relevance feedback method, such as those in [1, 3].

This retrieval experiment used the same images described in Section 4.1. Initially, the constructed space was initialized to a Euclidean space, and a support vector machine (SVM) was trained to map the Gabor features of the textures to the coordinates in the constructed space. For each retrieval trial, an image is randomly chosen as the query image, and the trained SVM was used to retrieve and rank the images according to the distances measured in the constructed space. Then, relevance feedback was given to the top 20 retrieved images so that there was a high likelihood that each feedback provided some new relevance information that was not already included in the preceding feedbacks. The incremental algorithm was applied to update the perceptual space after collecting 10 feedbacks, each containing non-redundant information, and the SVM was re-trained. After re-training, the SVM was used to retrieve images and the precision and recall of the retrieval were computed. This test procedure was repeated several times.

Recall that the images used in the experiment consist of 9 non-overlapping samples cropped from each of the 60 texture images. During relevance feedback, any information (relevant or irrelevant) associated with any one of the samples is as good as that associated with the other samples. That is, there is redundancy in the feedback information associated with the samples of a texture type. On the other hand, feedbacks associated with different texture types are not redundant. We use the term *information coverage* to refer to the amount of non-redundant information covered by a set of feedback results.

Figure 3 plots the perceptual consistency of the system at various stages. 10 feedback iterations, each containing non-redundant information, are performed between each stage. As more and more feedback results are received, the information coverage increases. Perceptual consistency of the incrementally constructed space was checked against the perceptual texture space (PTS) constructed using psychological experiments (Section 4.1). The normalized precision and recall were averaged over 10 random retrieval tests. Figure 3 shows that, as information coverage increases, the perceptual consistency of the incrementally constructed space approaches the upper bound achieved by direct mapping to PTS using SVM.

Figure 4 plots the precision-recall curves at various stages. In this case, all the images served as query images one at a time, and precision and recall were averaged over all the queries. The precision-recall curve of the SVM model is also displayed to indicate the best performance achievable. The curve for SVM corresponds to the case of 100% information coverage. That is, all the distances measured in the constructed space are perceptual distances. Figures 3 and 4 show that the incremental method indeed improves retrieval performance over time. Moreover, the constructed space becomes more and more consistent with actual perceptual texture space as more and more feedback results are received.

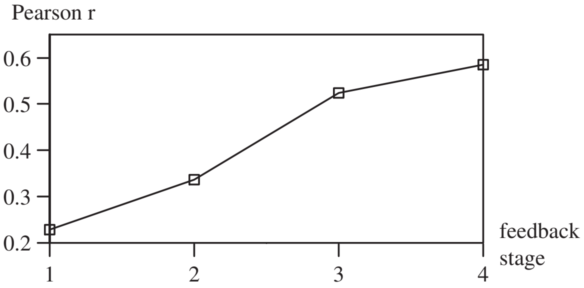


**Figure 3: Perceptual consistency, measured as Pearson's coefficient, at each stage of the incremental construction. As information coverage increases, the perceptual consistency of the space also increases.**

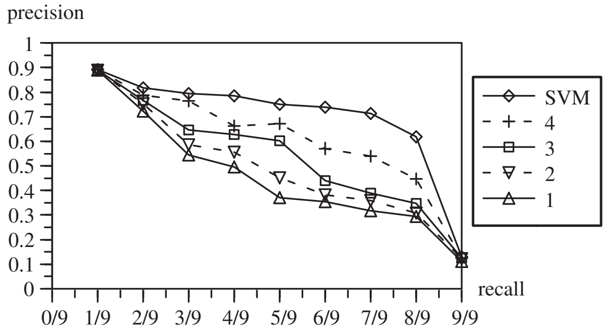


**Figure 4: Precision-recall curves of the incremental space at stages 1 to 4. As information coverage increases, the precision-recall curve of the space shifts towards the upper bound achieved by direct mapping to PTS using SVM.**

## 5.3 Phase Shifts of Incremental Space

The incremental construction method shows an interesting phenomenon called *phase shifting*. Initially, the constructed space is just a computational space. As more and more feedback results are received, the space begins to shift towards a perceptual space. To further explore this phenomenon, a simulation test was conducted. In this test, the distances $p_{ij}$ were initialized to the Euclidean distances between images $i$ and $j$. Then, a fraction of the distances $p_{ij}$ were randomly selected and replaced by the true perceptual distances obtained from the psychological experiment for constructing the PTS (Section 4.1). Then, the correlation between the constructed space and PTS was measured. As the amount of true perceptual distances contained in the space increases, the amount of information coverage also increases.

Figure 5 plots the perceptual consistency of the constructed space at various percentage of coverage. In addition, the operating points of the four computational texture models discussed in Section 4.2 are marked in the figure according to their perceptual consistency. It can be seen that, below 5% coverage, the space is a Euclidean space. At 20% coverage, the space shifts to an FFC space. It becomes a highly perceptually consistent space at 80% coverage. In between about 30% and 70% coverage, the space behaves as a mixture of computational and perceptual spaces. Therefore, the constructed space (marked as squares in Fig. 5) undergoes phase shifts from computational towards perceptual as more and more distances are replaced by true perceptual distances. This phase shifting property offers another advantage in addition to making the constructed space perceptually consistent. As a user's relevance judgment changes over time, the space will also change accordingly, thus adapting to the user's changing need.

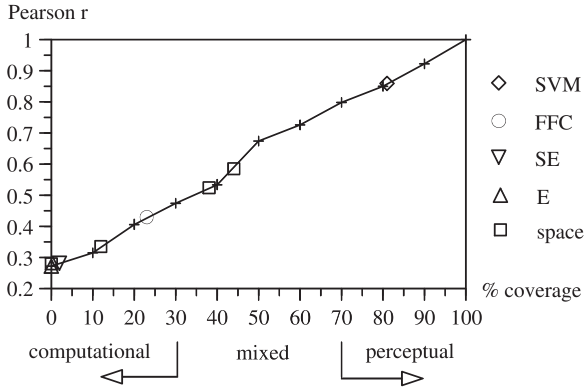


**Figure 5: Phase shifts of incremental space. The line indicates the degree of perceptual consistency at various percentage of information coverage. The labels along the line mark the positions of the four texture models (E: Euclidean, SE: scaled Euclidean, FFC, and SVM mapping) according to their perceptual consistency. At low coverage, the incrementally constructed space behaves as a computational space. It shifts towards a mixed space at moderate coverage, and a perceptual space at high coverage.**

# 6. CONCLUSION

This paper shows that perceptual consistency of computational texture models is highly correlated with their retrieval performance. Perceptual consistency can, therefore, be used as an indicator of retrieval performance in addition to precision and recall. A straightforward method of improving retrieval performance is to construct a perceptual space based on psychological experiments of human perception. Computational features of images can then be mapped to the perceptual space so that the measurement of image similarity in the space would be consistent with human's perception.

In practical application, it may not be feasible to construct such as space due to the large amounts and variations of images. To solve this problem, this paper presented an incremental method for measuring perceptual distances and constructing the perceptual space using relevance feedback results. This method requires only a small amount of relevance judgments at each relevance feedback iteration. It accumulates the feedback information over several queries, and use the information to construct the perceptual space.

The constructed space exhibits an interesting phase shifting property. Initially, the space is initialized to a computational space. As more and more feedback results are received, the space shifts from a computational space towards a perceptual space. This property confirms that the method can indeed construct a perceptual space incrementally. In other words, the incremental construction method provides a means of transforming a computational space to a perceptual space in a seamless manner. It also has the additional advantage of allowing a retrieval system to gradually adapt to the user's changing need.

# 7. REFERENCES


[1] I. J. Aalbersberg. Incremental relevance feedback. In *Proc. SIGIR '92*, pages 11–22, 1992.

[2] N. Abbadeni, D. Ziou, and S. Wang. Computational measures corresponding to perceptual textural features. In *Proc. ICIP*, 2000.

[3] J. Allan. Incremental relevance feedback for information filtering. In *Proc. SIGIR '96*, pages 270–278, 1996.

[4] W. Basalaj. Incremental multidimensional scaling method for database visualization. In *Proc. SPIE*, volume 3643, pages 149–158, 1999.

[5] C. Buckley and G. Salton. Optimization of relevance feedback weights. In *Proc. SIGIR '95*, pages 351–357, 1995.

[6] I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos. PicHunter: Bayesian relevance feedback for image retrieval. In *Proc. ICPR*, 1996.

[7] D. C. Donderi. Information measurement of distinctiveness and similarity. *Perception and Psychophysics*, 44(6):576–584, 1988.

[8] C. Heaps and S. Handel. Similarity and features of natural textures. *J. Expt. Psycho.: Human Perception and Performance*, 25(2):299–320, 1999.

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

[10] W. K. Leow and S. Y. Lai. Scale and orientation-invariant texture matching for image retrieval. In M. K. Pietikäinen, editor, *Texture Analysis in Machine Vision*. World Scientific, 2000.

[11] H.-Z. Long and W. K. Leow. Perceptual texture space for content-based image retrieval. In *Proc. Int. Conf. on Multimedia Modeling (MMM 2000)*, pages 167–180, 2000.

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

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

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

[15] 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.

[16] A. R. Rao and G. L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55(3):218–233, 1993.

[17] S. E. Robertson and K. S. Jones. Relevance weighting of search terms. *Journal of American Society for Information Science*, 27(3):129–146, 1976.

[18] J. J. Rocchio. Relevance feedback in information retrieval. In G. Salton, editor, *The SMART Retrieval System—Experiments in Automatic Document Processing.* Prentice Hall, 1971.

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

[20] Y. Rui and T. S. Huang. A novel relevance feedback technique in image retrieval. In *Proc. ACM MM '99*, pages 67–70, 1999.

[21] Y. Rui, T. S. Huang, and S.-F. Chang. Image retrieval: Current techniques, promising directions, and open issues. *Journal of Visual Comm. and Image Rep.*, 10:39–62, 1999.

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

[23] G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval.* McGraw-Hill, 1983.

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

[25] W. M. Shaw. Term-relevance computations and perfect retrieval performance. *Information Processing and Management*, 31(4):491–498, 1995.

[26] J. R. Smith and S.-F. Chang. Multi-stage classification of images from features and related text. In *Proc. 4th Europe EDLOS Workshop*, 1997.

[27] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. SMC*, 8(6):460–47, 1978.

[28] A. Tversky. Features of similarity. *Psychological Review*, 84(4):327–352, 1977.