Incremental Semi-Supervised Subspace Learning for Image Retrieval

Author: Xiaofei He

Presented by Jerry Yu

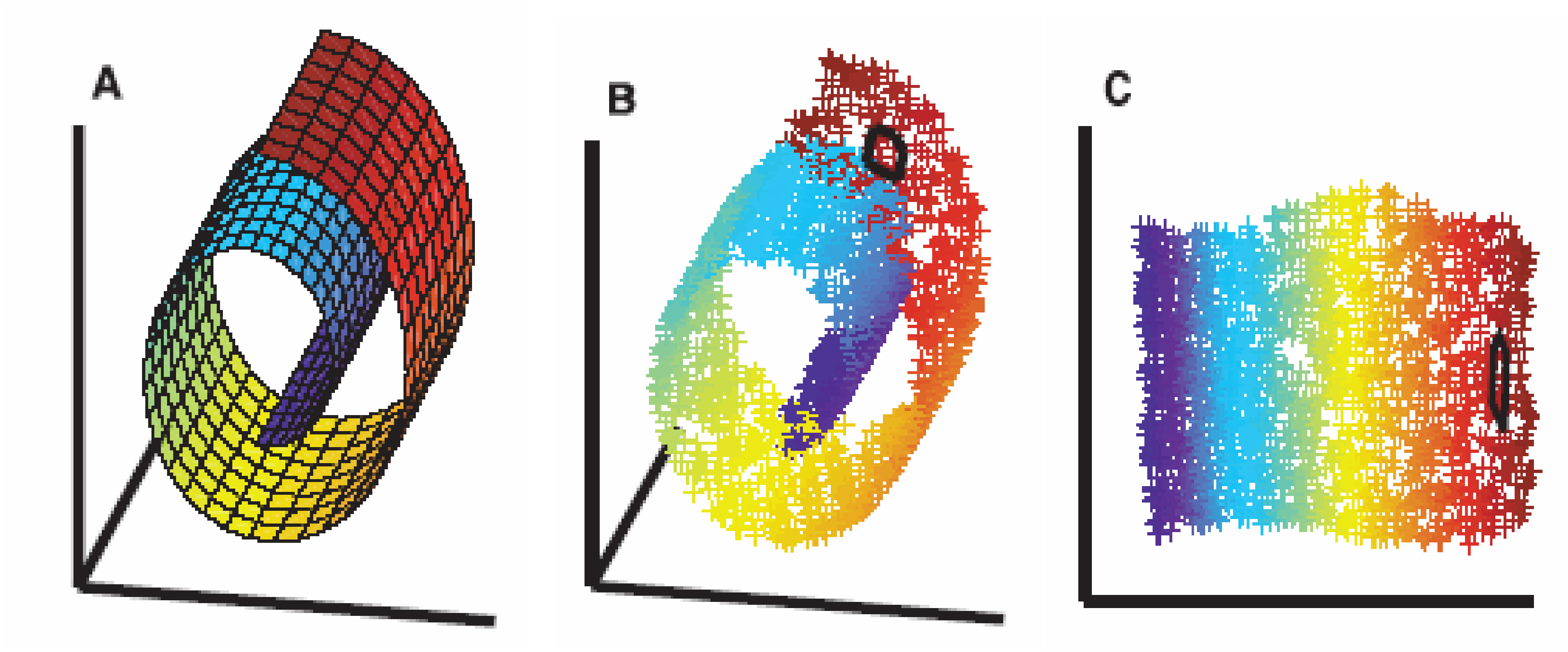
Outline

- Subspace Learning
- Image Retrieval with Relevance Feedback
- Locality Preserving Projection (LPP)
- Incremental Semi-Supervised LPP
- Experiment Results and Analysis

Subspace Learning

- In data analysis and visualization, the observed data could be in very high dimension space.
- Assuming that the data lie on a lower dimension subspace, compact representation of the data can be obtained by subspace learning.
- New approaches have been proposed:
  - ISOMAP (J. Tenenbaum and V. Silva, 2000)
  - Local Linear Embedding (S. Roweis and L. Saul, 2000)
  - Locality Preserving Projection (X. He and P. Niyogi, 2003)

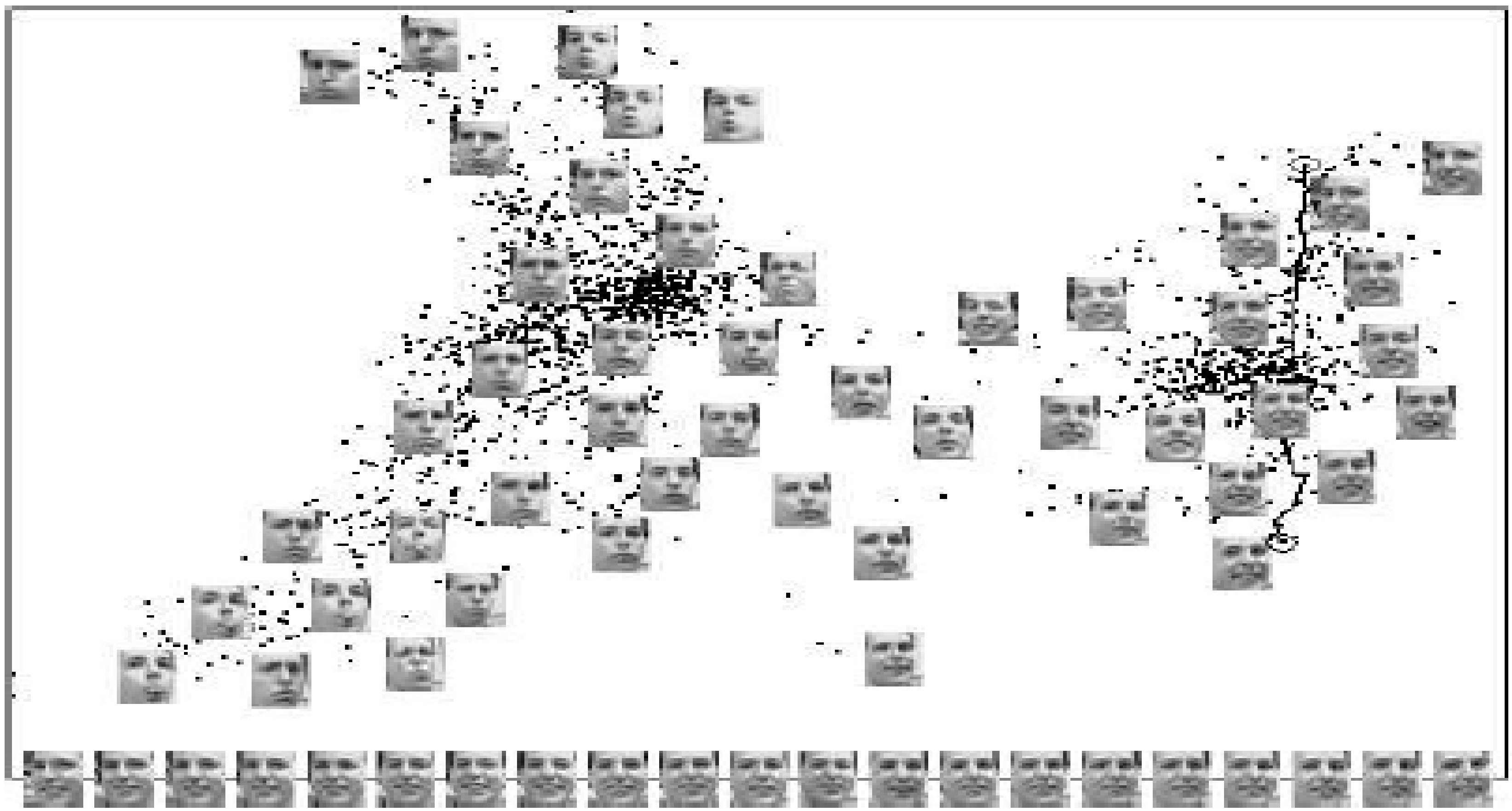
Subspace Learning: Example



Content-based Image Retrieval

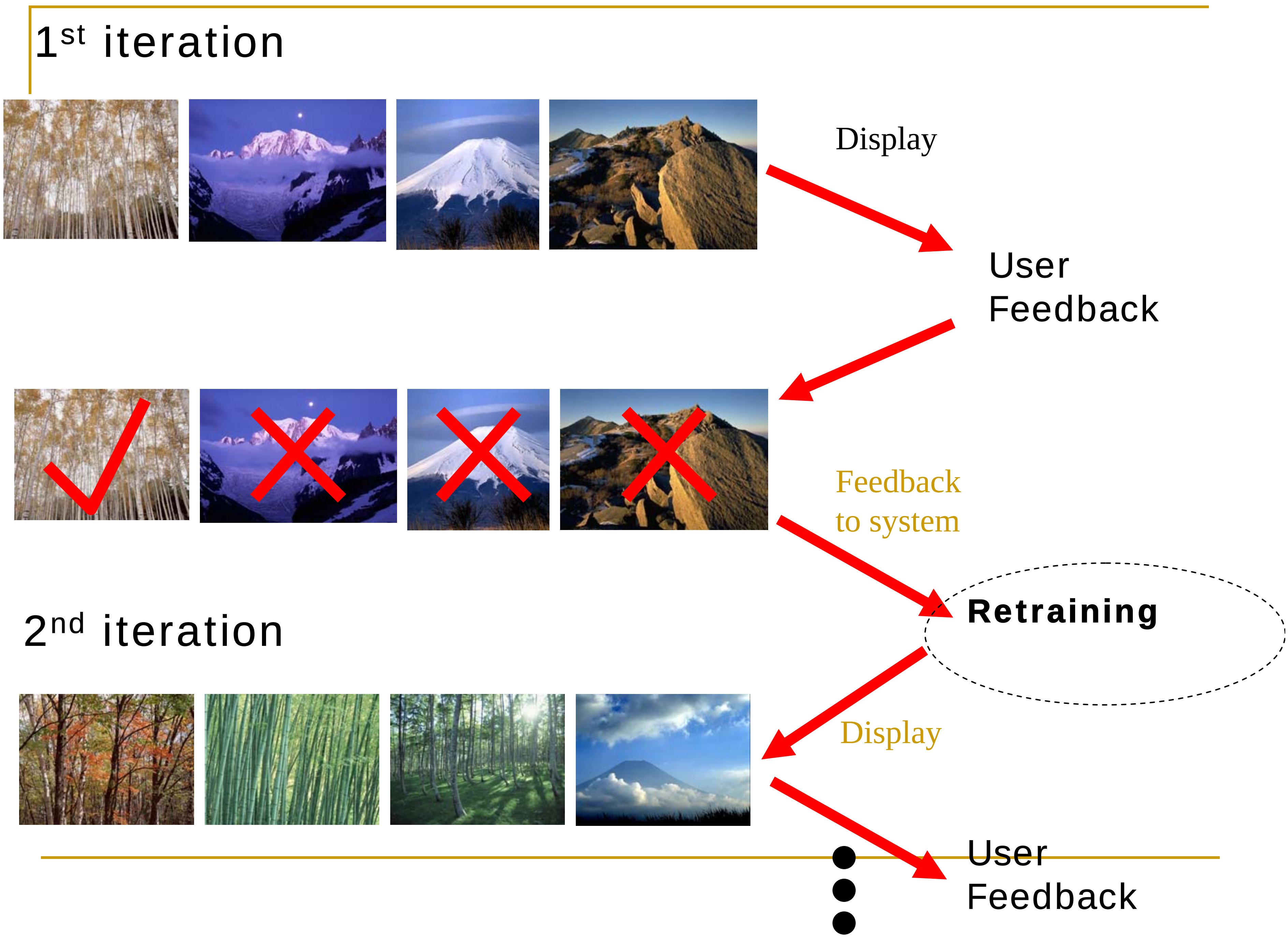
- Content-based image retrieval (CBIR): to retrieve images based on the visual content.
- One of the core problems of content-based image retrieval is image representation, that is to find compact representation for the images in low dimension space.
- Subspace learning is promising in respect to discover the structure of the nonlinear image subspace.

Subspace Learning and Image Representation: Example



Relevance Feedback

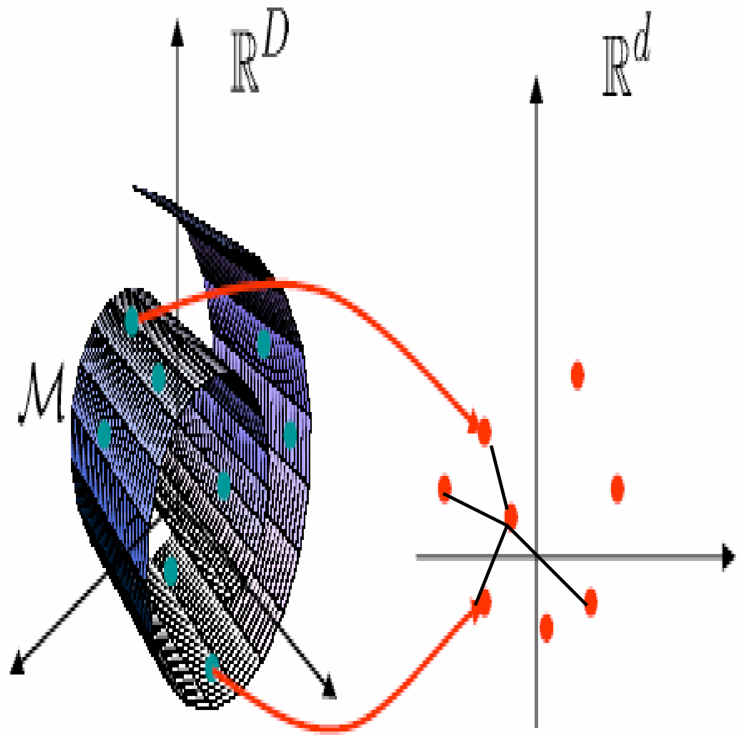
- Image understanding is another major problem to CBIR.
- Relevance feedback provides user's judgment on the subject of images and the semantic relationships between images.
- Accumulating relevance feedback would improve the understanding of the images for specific users or application.


1st iteration
Display
User Feedback
Feedback to system
Retraining
2nd iteration
Display
User Feedback

Locality Preserving Projection

- **Problem Model:**
  - Images are represented as *X* in R^D space
  - Lower dimensional space (R^d) representation *Y* of *X* is obtained by projection *A*

  $Y=A'*X$
- **Goal:** Preserving locality information by finding a projection ***A*** that minimize the sum of the squared distance from one sample to its neighborhood samples after projection.



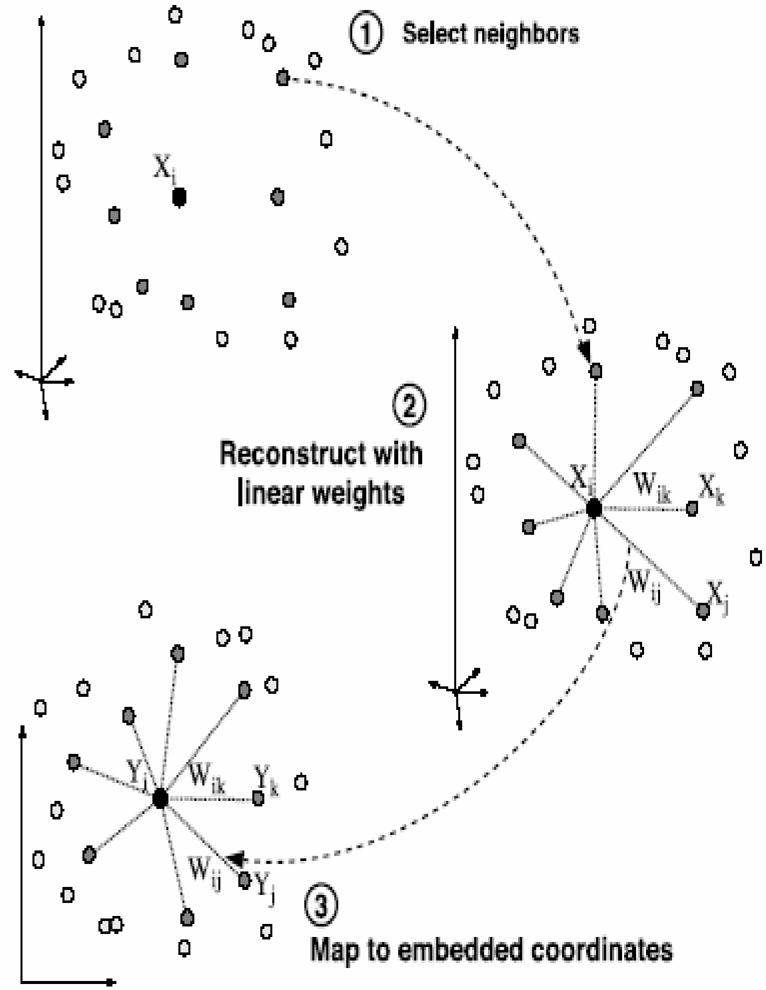
Locality Preserving Projection

- *Step 1*: Construct K-NN matrix ***S***

$$S_{ij} = \begin{cases} 1 & \text{if } \mathbf{x}_i \text{ is among the } k \text{ nearest neighbors of } \mathbf{x}_j \\ & \text{or } \mathbf{x}_j \text{ is among the } k \text{ nearest neighbors of } \mathbf{x}_i \\ 0 & \text{otherwise} \end{cases}$$

- *Step 2*: Construct a weight matrix ***W*** that encodes neighborhood information.

$$w_{ij} = s_{ij} / \sum_j s_{ij}$$



Locality Preserving Projection

- *Step 3*: Objective function (Locality Preserving):

$$\min \sum_{i,j} w_{ij} \,|\, y_i - y_j \,|^2 = 2trace(A^T XLX^T A)$$

where $L = D - W$ and $D = diag(\sum_j w_{ij})$

- Solution: ***A*** is the d eigen vectors corresponding the d smallest eigen values of XLX^T

Incremental LPP

- LPP: Unsupervised approach can't incorporate user feedback.
- Incremental LPP: incrementally incorporate semantic information from user feedback
- *Step 1:* Start from normal LPP *(t=0)*
- *Step 2:* Update from relevance feedback

  $s_{t,ij} = 1$ if *i, jth* data from positive class

  $s_{t,ij} = 0$ if *i, jth* data from different class
- *Step 3:* Find the optimal projection after feedback

$$A_t = eig(XL_tX^T)$$

Convergence of Incremental LPP

- Suppose we finally have the category information for all images.
- Mathematical proof shows that

$$XL_{\infty}X^{T} = S_{w}$$

$$S_{w} = \sum_{i=1}^{C}\sum_{j=1}^{n_i}(x_j - m_i)(x_j - m_i)^T$$

- The incremental LPP tries to make each class cluster to its center.

Faster Incremental LPP

- Update XL_nX^T:

$$XL_nX^T = XL_{n-1}X^T + \sum_{ij}(L_n - L_{n-1})_{ij}x_ix_j^T$$

- Update 1st Eigenvector:

$$u_n^1 \approx \frac{n-1}{n}u^1{}_{n-1} + \frac{1}{n}XL_nX^T\frac{u^1{}_{n-1}}{\|u^1{}_{n-1}\|}$$

$$a_n^1 = \frac{u^1{}_n}{\|u^1{}_{n-1}\|}$$

- Calculate Other Eigenvectors with updated *X*

$$X = X - a_n^1(a_n^1)^T X$$

Experiment Setting

- COREL Database: 30 categories, 100 images in each categories
- The query image is randomly picked up from the database
- Images are preprocessed to extract 3 color features and 3 texture features.
- Length of the concatenated feature vector: 435.

Experiment Setting

- In each iteration, 4 images will be presented to user for feedback.
- Performance Evaluation Metric:

$$Accuracy = \frac{\#\ \text{of correct images}}{N}$$

- In the following experiments N=15

PCA vs LPP

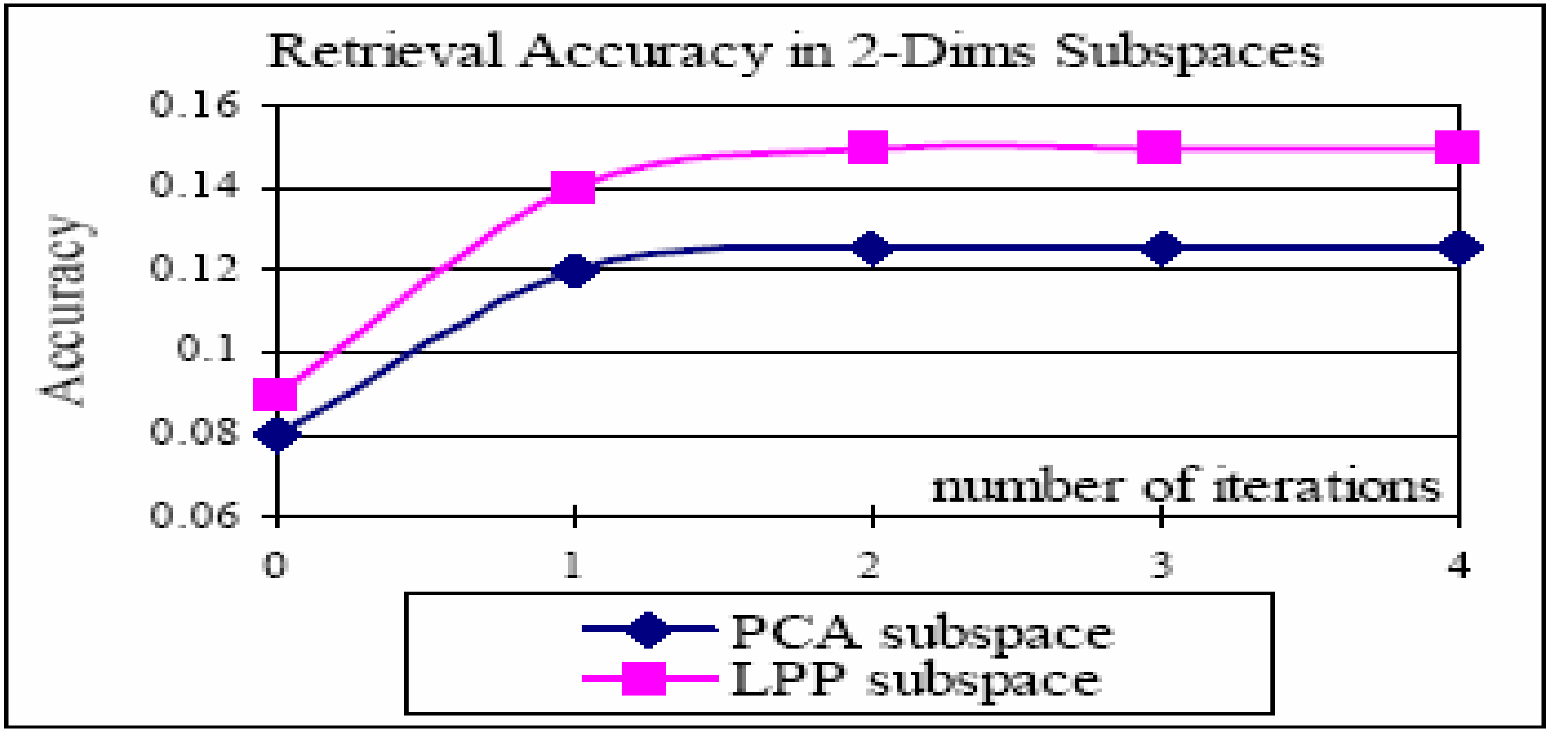


Figure 1. The image retrieval performances in PCA subspace and LPP subspace with 2 dimensions.

PCA vs LPP

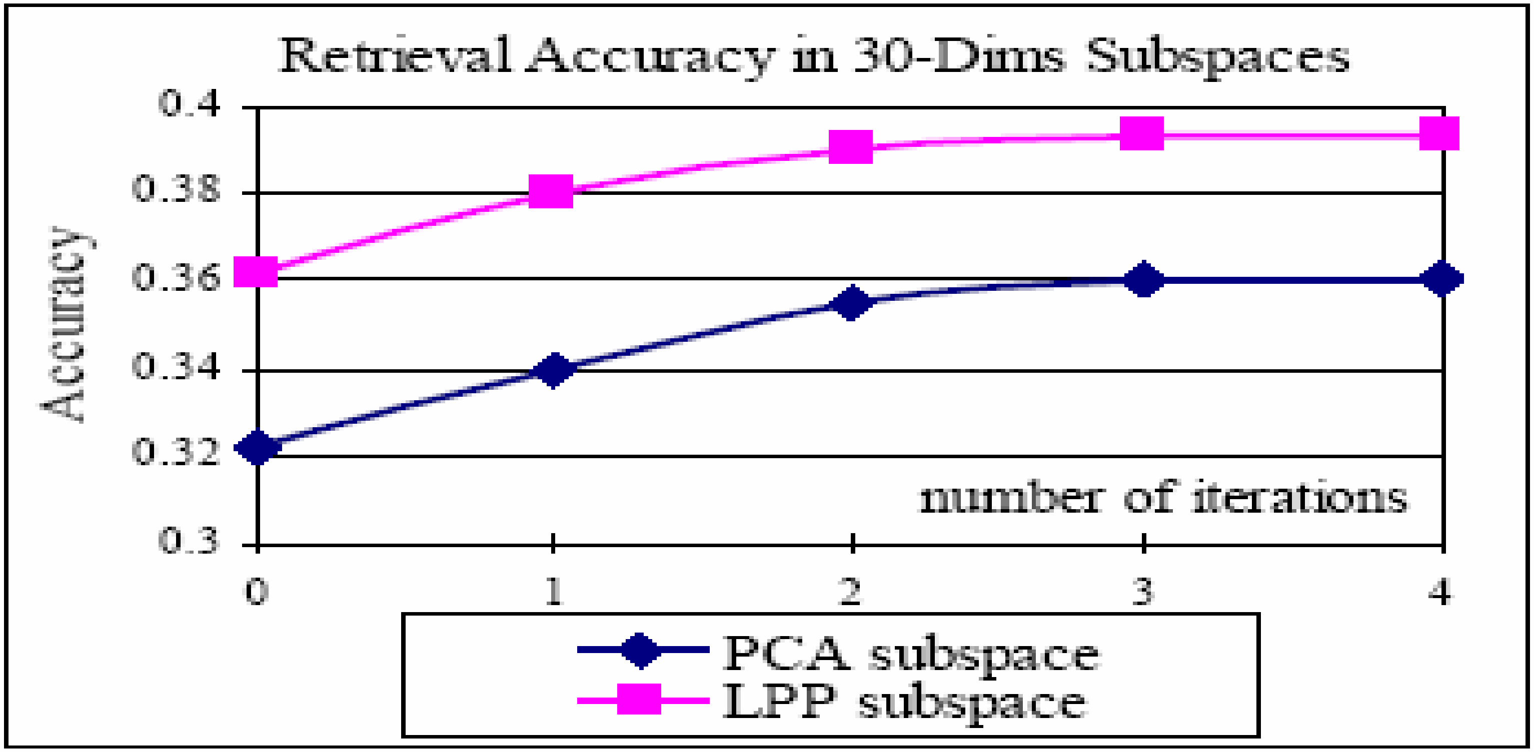


Figure 2. The image retrieval performances in PCA subspace and LPP subspace with 30 dimensions.

PCA vs LPP

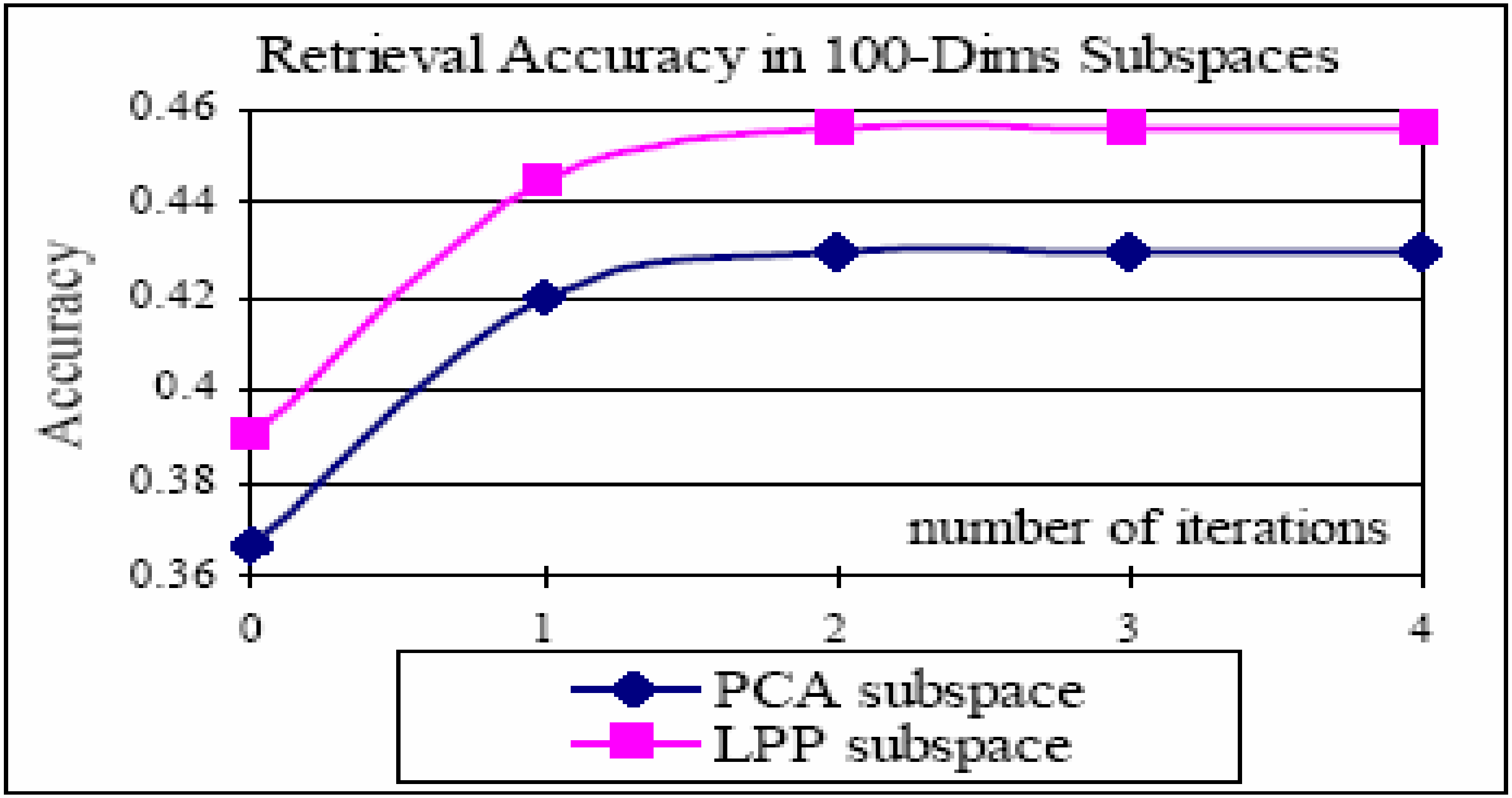


Figure 3. The image retrieval performances in PCA subspace and LPP subspace with 100 dimensions.

LPP vs Incremental LPP

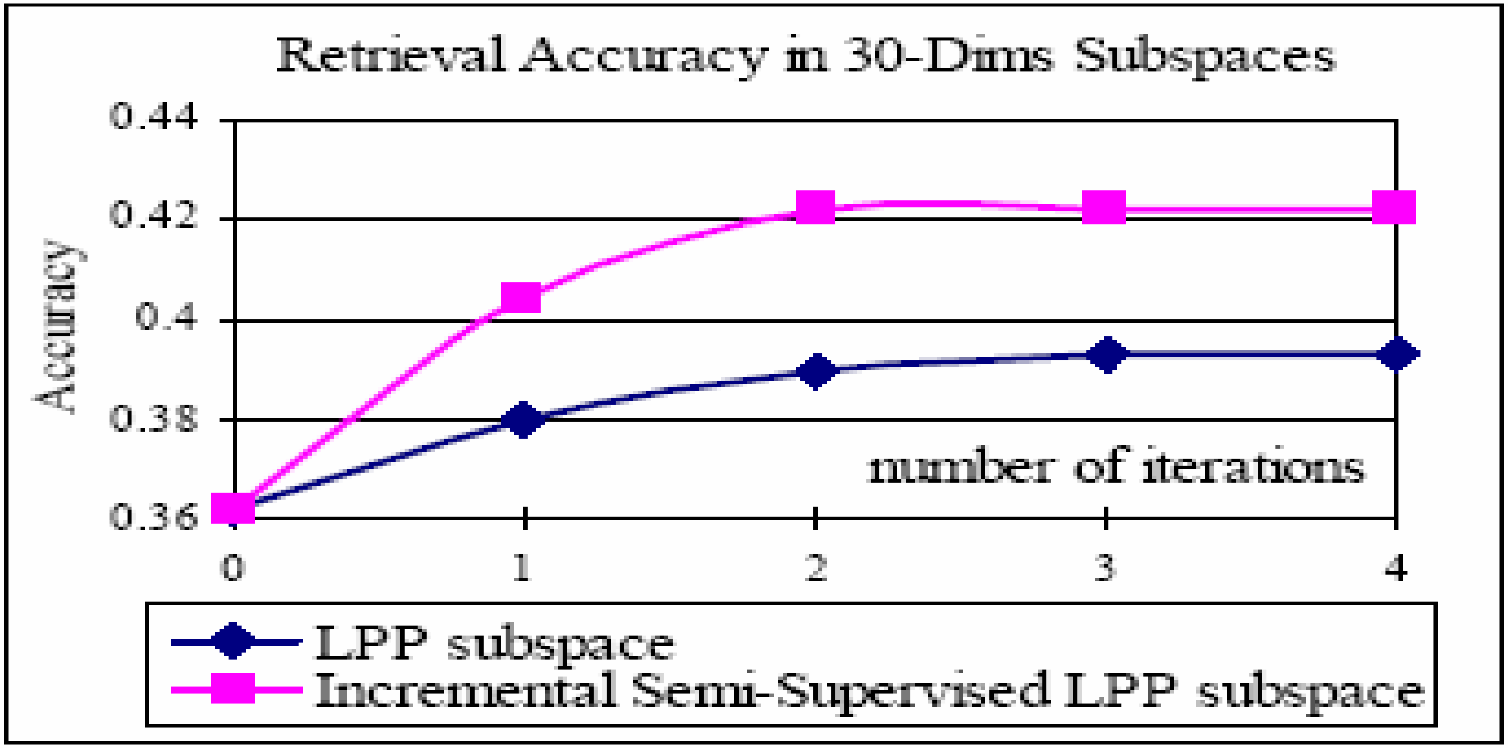


Figure 4. The image retrieval performances in LPP subspace and incremental semi-supervised LPP subspace with 30 dimensions.

Evolution of Incremental LPP

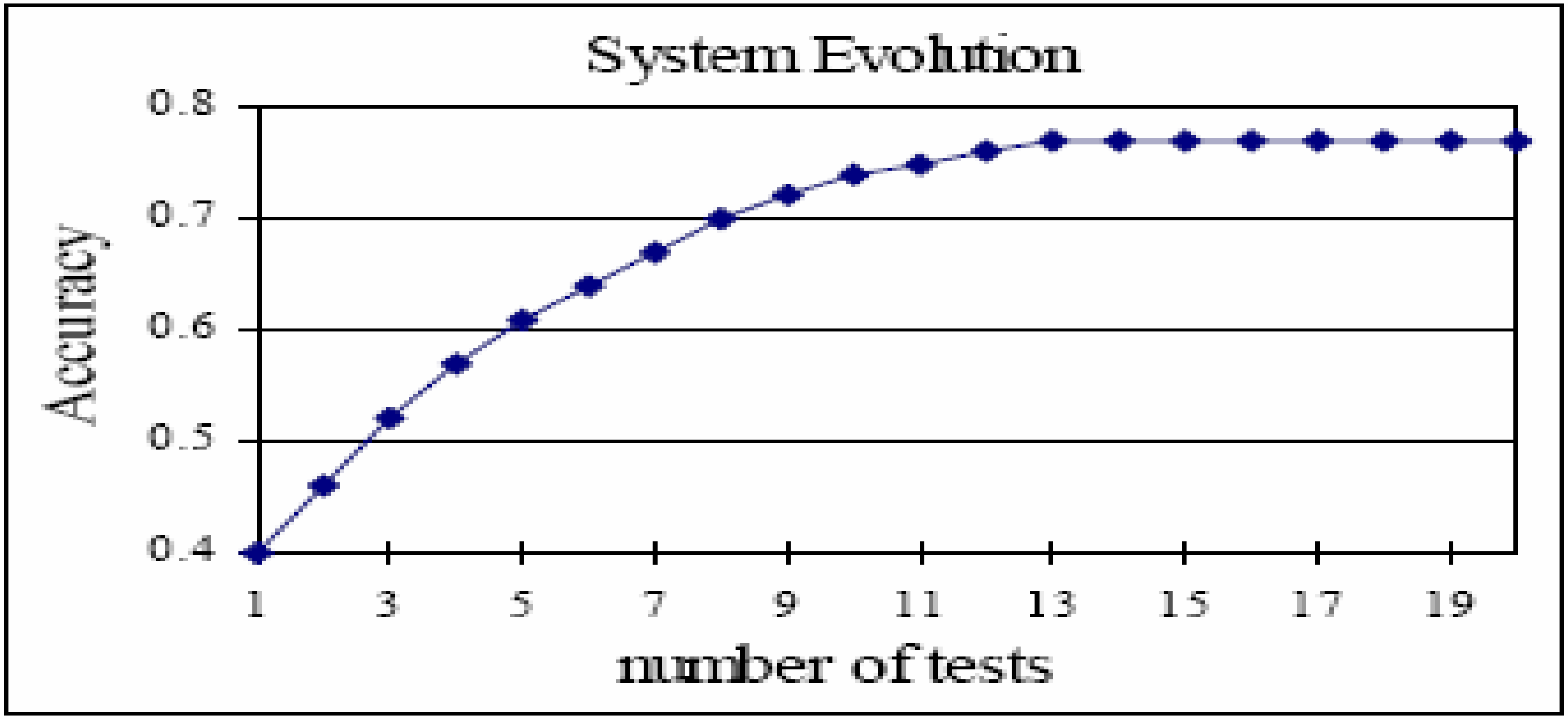


Figure 5. The retrieval accuracy of the system improves as the user's feedbacks are accumulated.

2-D Visualization of LPP

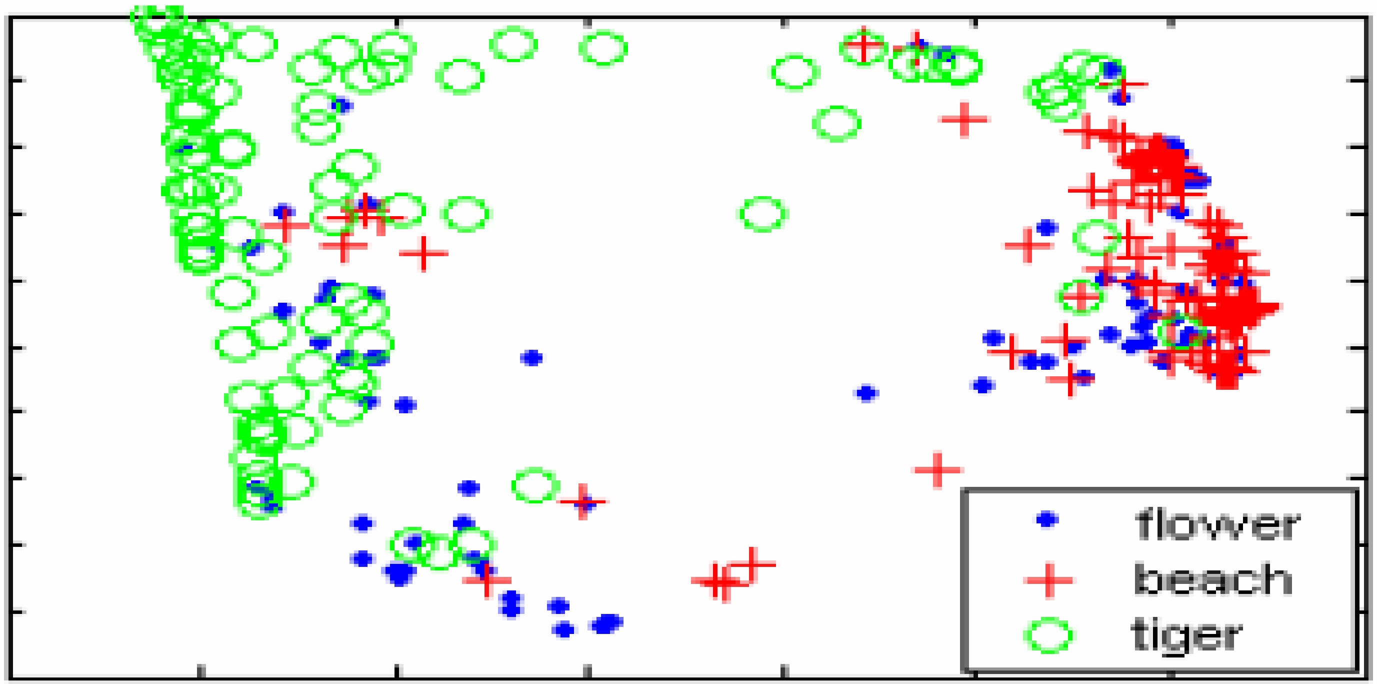


Figure 6. 2-D visualization of three image classes using the LPP algorithm.

2-D Visualization of Incremental LPP

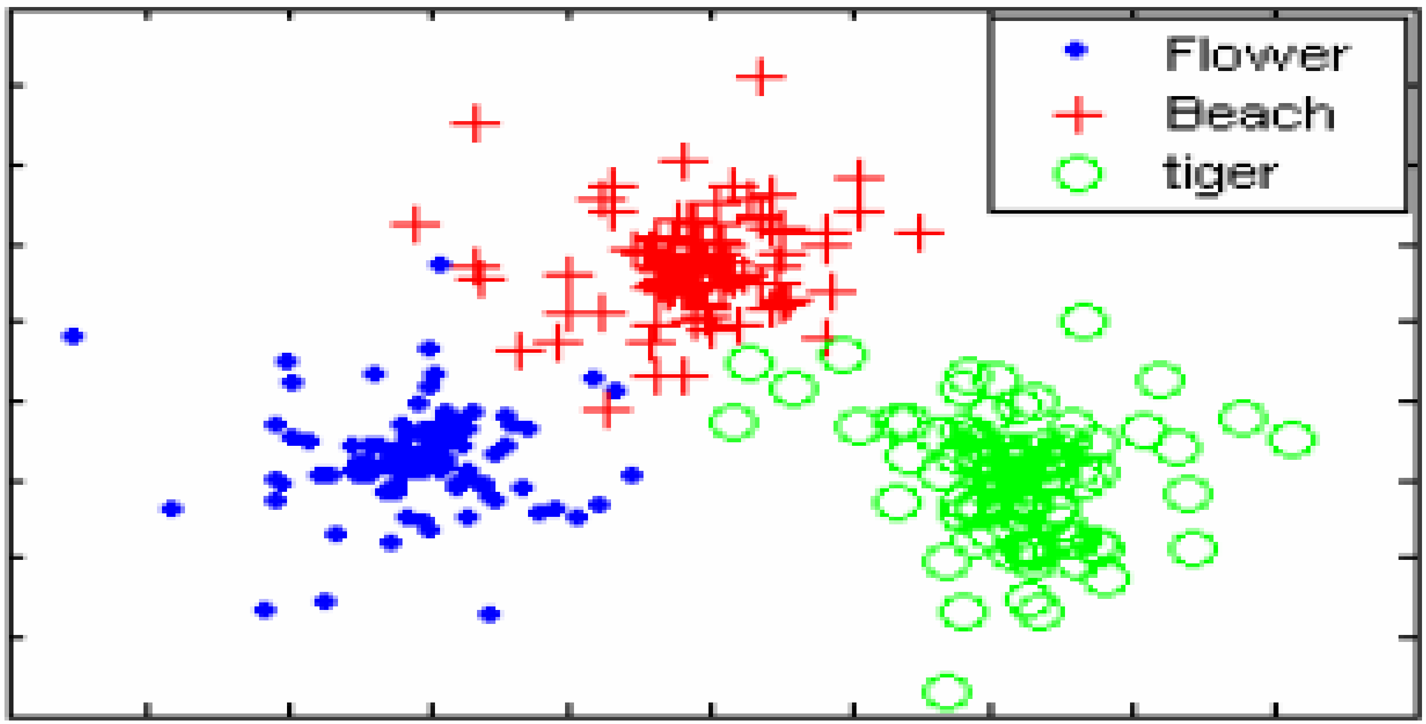


Figure 7. 2-D visualization of three image classes using incremental semi-supervised LPP algorithm.

Conclusion

- LPP preserves local structure which might be more important to learn image subspace.
- Incremental LPP is proposed to accumulate long-term retrieval experience.
- Theoretical study shows that Incremental LLP converges to minimize S_w.

Future Work

- Is user feedback reliable? (Noise problem)
- Is the local structure (geometrical structure of image subspace) consistent to human perception?

Reference Paper

- [1] Incremental Semi-Supervised Subspace Learning for Image Retrieval, Xiaofei He, *ACM conference on Multimedia*, 2004
- [2] Locality Preserving Projections, Xiaofei He, and Partha Niyogi, *Advances in Neural Information Processing Systems*, 2003
- [3] Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering, M. Belkin and P. Niyogi, *Advances in Neural Information Processing Systems*, 2002
- [4]Nonlinear Dimension Reduction by Local Linear Embedding, S. Rowis and L. Saul, *Science*, 2000
- [5] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford, “A Global Geometric Framework for Nonlinear Dimensionality Reduction,” *Science*, 2000.