

# 3D Object Retrieval Using 2D Line Drawing and Graph Based Relevance Feedback

Liangliang Cao<sup>1</sup>, Jianzhuang Liu<sup>1</sup>

<sup>1</sup>Dept. of Information Engineering  
The Chinese University of Hong Kong  
{llcao, jzliu}@ie.cuhk.edu.hk

Xiaoou Tang<sup>1,2</sup>

<sup>2</sup>Microsoft Research Asia  
Beijing, China  
xitang@microsoft.com

## ABSTRACT

This paper aims to provide a user-friendly interface for 3D object retrieval. In previous 3D retrieval systems, the user mainly uses two methods to input a query: providing an existing 3D objects, or providing partial shape information of desired objects such as text and 2D shapes. The first method fails when the user does not have a similar 3D object in hand, and the second method cannot sufficiently describe 3D shapes of objects. We believe that the best way is to have a good interface that can convert a 2D sketch drawn by the user into a 3D object as the query. A 2D line drawing is easy to be drawn and is the simplest and most direct way of illustrating a 3D object. In this paper, we develop an interface of 3D object reconstruction from line drawings, which allows the user to draw line drawings of objects with both planar and curved surfaces. In addition, in order to refine the retrieved results, we develop a relevance feedback algorithm based on a novel graph discriminant analysis. Compared with recently published relevance feedback algorithms, our algorithm achieves better retrieval performance.

**Categories and Subject Descriptors:** H.3.3 [Information Search and Retrieval]: [Query formulation, Relevance feedback]

**General Terms:** Design, Algorithms

**Keywords:** 3D object retrieval, 3D reconstruction, query generation, relevance feedback

## 1. INTRODUCTION

The recent advancement of scanning, modeling, and digitizing techniques for 3D shapes has resulted in an increasing amount of 3D models on the Internet and in corporate databases. This leads to the emergence and the flourish of a number of experimental 3D object retrieval engines in the past several years (see [13] for a review).

In most of these existing retrieval systems, the user is required to provide a 3D model as the query for retrieval. When the user does not have a desired 3D model but only a coarse shape in mind, the systems cannot help. This problem limits the use of 3D search engines in practice. To circumvent this difficulty, some systems allow the user to provide text or 2D shapes as 3D cues. However, text is ambiguous in defining 3D objects [8], and finding 3D clues from 2D shapes is time-consuming and is still a difficult problem

to solve in computer vision research. Some researchers employed Teddy [5], a sketch-to-3D tool, to help the user input 3D objects as queries [12]. However, Teddy can only construct tube- or sphere-like objects. It cannot produce many other common objects, such as the ones with planar faces.

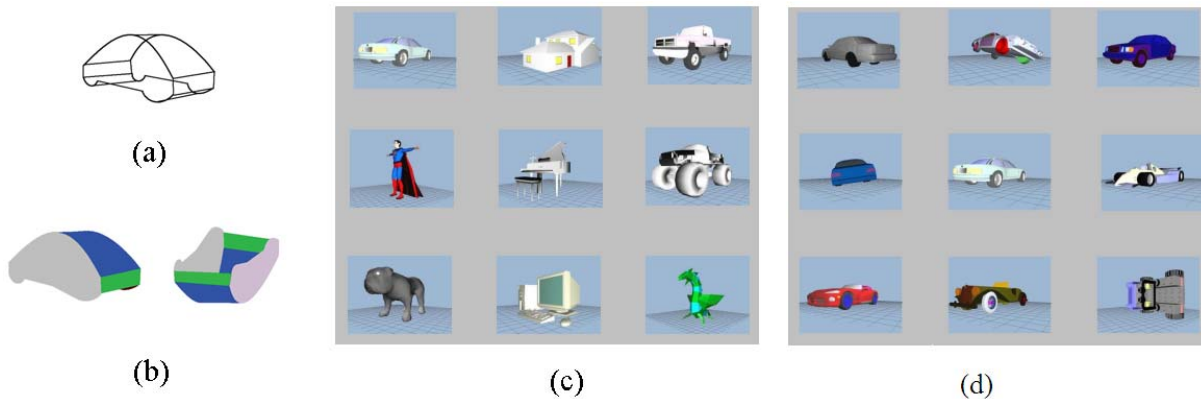
We believe that the best way of query input for 3D object retrieval is to have a user-friendly interface that can convert the user's sketches into 3D objects. A line drawing is easy to be drawn and is the simplest and most direct way of illustrating a 3D object. In this paper, we develop a novel interface which enable the user to input 3D object by drawing 2D line drawings.

Although 3D objects can be reconstructed from 2D sketches drawn by the user in our 3D retrieval system, it can only provide a coarse shape description since it is not practical to ask the user to draw all the details of the desired objects. Therefore, it is often necessary to refine the retrieval results. In our work, relevance feedback is employed as a complement of the user sketching. Relevance feedback is not a new idea in 2D content-based image retrieval, and has been studied extensively as a tool to obtain better retrieval results [9]. However, few previous works study relevance feedback in 3D object retrieval. We develop a new representation called the feedback graph, based on which a new algorithm, called graph discriminant analysis (GDA), helps the user refine the search results.

Fig. 1 illustrates how our retrieval system works. Fig. 1(a) shows an example (a car) of 2D line drawings the user inputs. We can see that this is straightforward to reflect the 3D shape in mind. Fig. 1(b) gives the reconstructed 3D query in two views using our reconstruction algorithm. Fig. 1(c) shows a few retrieved objects at first. Since the query and the 3D object matching method may not result in only cars as the output, the relevance feedback step refines the retrieval results, as shown in Fig. 1(d).

## 2. RECONSTRUCTION OF A 3D QUERY FROM A 2D LINE DRAWING

The first contribution in our 3D retrieval system is the new interface which reconstructs the 3D query from line drawings provided by the user. A line drawing is defined as a 2D orthogonal projection of the edges and vertices of a 3D object in a generic view. The human visual system has the ability to interpret 2D line drawings as 3D objects without difficulty. Emulating this ability is an important research topic since the early days in machine vision [1], [6], [10], but they can only handle objects with planar faces. Our algorithm in this paper can reconstruct 3D objects with



**Figure 1:** (a) A 2D line drawing inputted by the user. (b) The reconstructed 3D object (the query). (c) Retrieved objects which are considered similar to the query. (d) The refined results after the first relevant feedback.

both planar and curved faces. It is divided into three steps: (1) transforming the line drawing of a curved object into the line drawing of a planar object, (2) reconstructing the planar object from the transformed line drawing, and (3) generating the curved faces with Bezier curves and surface patches based on the reconstructed planar object and the origin line drawing.

In the first step, we use our previous algorithm in [7] to transform the line drawing of a curved object into one representing a planar object by straightening all the curves. In the second step, we employ the method in [6] to reconstruct a 3D planar object from the transformed line drawing. Next we discuss how to generate the curved object from this planar object and the original line drawing.

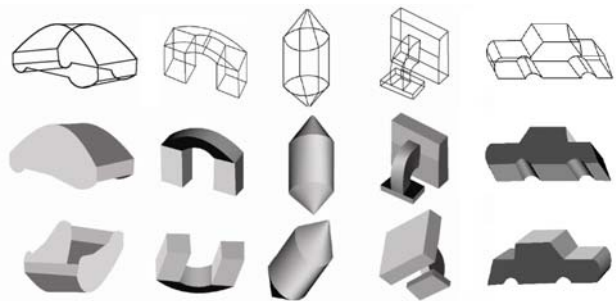
Here we assume that every curve reconstructed from a line drawing is piecewise planar in 3D space. Then we can reconstruct the curve from its projection if we know the 3D plane containing it. Since the recovered planar object contains at least one vertex of the curve in addition to its two end points, the plane containing the curve can be defined by the three non-collinear points in 3D space [7], [2]. Then the curved surfaces can be interpolated from the boundary curves.

For a curved face/patch, it can be parameterized by a Bezier patch  $\mathbf{S}(u, v)$ ,  $0 \leq u, v \leq 1$ , with its four Bezier boundary curves being  $\mathbf{S}(0, v)$ ,  $\mathbf{S}(1, v)$ ,  $\mathbf{S}(u, 0)$ , and  $\mathbf{S}(u, 1)$ . The Bezier patch can be approximated by:

$$\begin{aligned} \mathbf{S}^*(u, v) = & \\ & (1-u)\mathbf{S}(0, v) + u\mathbf{S}(1, v) + (1-v)\mathbf{S}(u, 0) + v\mathbf{S}(u, 1) \\ & - (1-u)v\mathbf{S}(0, 1) - u(1-v)\mathbf{S}(1, 0) \\ & - (1-u)(1-v)\mathbf{S}(0, 0) - uv\mathbf{S}(1, 1), \end{aligned} \quad (1)$$

where  $\mathbf{S}^*(u, v)$  passes through the four boundaries and the four corners exactly once. Our experimental results show that this interpolation scheme gives satisfactory 3D curved objects. More details can be found in [2].

Fig. 2 shows some examples of 3D reconstruction of both planar surfaces and curved surfaces. Our new interface provides not only a natural way for the user to sketch their ideas, but also a good approximation of more complex freeform objects. For example, a human body can be approximated by several cylinder-like and box-like objects.



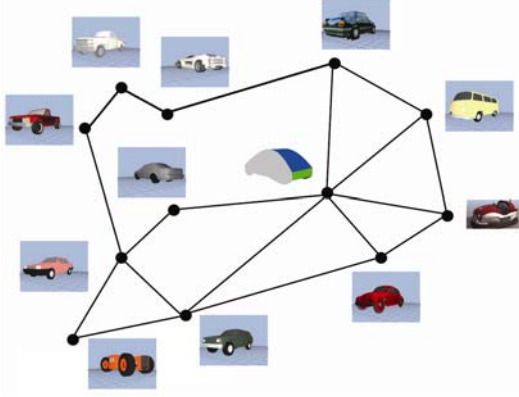
**Figure 2:** Some reconstruction examples

### 3. A GRAPH-BASED RELEVANCE FEEDBACK ALGORITHM

After a 3D object is recovered from the user's sketch input, the next step is to search a 3D object database for the objects with similar shapes. Since sketch interface does not provide detailed and concise shapes for the 3D object retrieval, a novel refining process is carried out to improve the retrieval results, which is the second contribution of this paper. This refining process is motivated by the *relevance feedback* technique first used in 2D image retrieval [9].

In image retrieval, a popular feedback learning algorithm is linear discriminant analysis (LDA), which tries to find a subspace in which the within-class instances congregate as much as possible while the between-class instances scatter as much as possible. Later, considering negative feedback instances do not belong to the same class, Zhou and Huang presented an algorithm called biased discriminant analysis (BDA), to find a subspace that clusters only positive examples while keeping negative ones away [14].

Unlike the feedback algorithm in 2D image retrieval, we consider only 3D shapes instead of colors and textures. We choose spherical harmonics for 3D shape descriptions as in [12], which enjoy many nice properties such as invariance under rotations and translations. Obviously, the feature space of 3D objects is quite different from that of 2D images. While the goal of 2D image feedback is to overcome the gap between the low-level visual features and the high-



**Figure 3: A feedback graph constructed by a group of positive feedback instances, where only part of the edges are shown.**

level semantic concepts of the image, the goal of our 3D object feedback mostly compensates the effects of inaccurate 3D queries. Thus the relevance feedback for 3D object retrieval is not a simple extension of the relevance feedback for 2D image retrieval. Next we propose a feedback graph representation of the feedbacks and develop a new learning algorithm based on it.

### 3.1 Feedback-Graph Representation

Although popular in pattern classification and dimension reduction, LDA/BDA flaws in two aspects when being applied to relevance feedback in 3D object retrieval. Theoretically, LDA is optimal only when the instances of one class are independently and identically distributed (iid) samples of a Gaussian distribution, assuming that the within-class covariances are equal for all classes. BDA suffers from the similar problem. However, 3D objects that are regarded as similar by human vision may be significantly different in geometry, such as different chairs, and persons with their arms rested or raised. In fact, the iid assumption doesn't hold for 3D object retrieval, where a geometric shape might be the deformation of another one and is probably not independent of others.

To remove the strong independence assumption made by LDA, we employ the idea in [3] which models the sampling dependence by a tree structure, and propose a new graphical representation of feedback instances. The positive feedback instances that are considered similar by the user are used to construct a feedback graph (a complete graph). A vertex of the graph denotes such an instance. An edge of the graph connects two such instances, with a weight equal to the distance of their corresponding vectors in the feature space. Fig. 3 shows an example of the feedback graph. This representation is preferred since it does not assume any distributions of instances and can avoid clustering the instances into fragmented clusters, which exists in the neighborhood-based and manifold learning algorithms when the training instances are not evenly distributed.

### 3.2 Graph-based Discriminant Analysis

Let the sets of positive feedbacks and negative feedbacks be  $PSet$  and  $NSet$ , respectively. Then the vertices of the feedback graph correspond to the instances in the  $PSet$ .

The weight of an edge between two vertices is initialized as the distance between the two feature vectors representing the two 3D objects by the spherical harmonics. The within-class compactness  $G_w$  is defined as the total weight of the minimum spanning tree ( $MST$ ) of the feedback graph in some subspace:

$$G_w = \sum_{(i,j) \in MST} \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|^2, \quad (2)$$

where  $\mathbf{x}_i, i \in PSet$  are the instances of positive feedback,  $(i, j) \in MST$  is an edge connecting  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , and  $\mathbf{W}^T$  is the transformation matrix that maps the instances into the subspace.

It is not difficult to express  $G_w$  in the form of matrix multiplication:

$$\begin{aligned} G_w &= \sum_{(i,j) \in MST} \text{Trace}(\mathbf{W}^T (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}) \\ &= \text{Trace}(\mathbf{W}^T \mathbf{X} \mathbf{C}^w \mathbf{X}^T \mathbf{W}), \end{aligned} \quad (3)$$

where  $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots]$ , and  $\mathbf{C}^w = [C_{i,j}^w]$  is called a w-connection matrix. For  $i \neq j$ ,  $C_{i,j}^w = 1$  if  $(i, j)$  is an edge in  $MST$ , otherwise  $C_{i,j}^w = 0$ ; for  $i = j$ ,  $C_{i,i}^w = d$  if vertex  $i$  is connected to  $d$  edges in  $MST$ .

The between-class compactness  $G_b$  is used to measure the distance between  $NSet$  and  $MST$  defined by

$$\begin{aligned} G_b &= \sum_{i \in NSet} \min_{j \in PSet} \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|^2 \\ &= \text{Trace}(\mathbf{W}^T \mathbf{X} \mathbf{C}^b \mathbf{X}^T \mathbf{W}), \end{aligned} \quad (4)$$

where  $\mathbf{C}^b = [C_{i,j}^b]$  is called a b-connection matrix. For  $i \neq j$ ,  $C_{i,j}^b = 1$  if  $j$  is the instance in  $MST$  closest to the negative instance  $i$ , otherwise  $C_{i,j}^b = 0$ ; for  $i = j$ ,  $C_{i,i}^b = \sum_{j \neq i} C_{i,j}^b$ .

The optimal transformation matrix  $\mathbf{W}^*$  is defined as

$$\mathbf{W}^* = \arg \max_{\mathbf{W}} \left( \frac{G_b}{G_w} \right). \quad (5)$$

$\mathbf{W}^*$  can be obtained by the eigen-decomposition of matrix [4]. We call this new discriminant *graph-based discriminant analysis* (GDA). After finding the optimal  $\mathbf{W}^*$ , we project vectors  $\mathbf{x}$  in the original feature space to vectors  $\mathbf{y}$  in the corresponding subspace by  $\mathbf{y} = \mathbf{W}^{*T} \mathbf{x}$ . Then the 3D objects which are close to the query and the positive feedbacks are selected as retrieval results. If the user is not satisfied with the current results, he/she can provide more feedbacks, and then  $\mathbf{W}^*$  and the retrieval results will be updated again.

Next we generalize GDA to its kernelized version, called KGDA, as LDA and BDA to KLDA and KBDA [14]. Suppose that the feature vectors  $\mathbf{x}$  of 3D objects are linearly separable when they are mapped to a higher-dimensional feature space  $F$  by  $\phi: \mathbf{x} \mapsto \phi(\mathbf{x}) \in F$  with the kernel function  $K: K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ . Here we use the classical Gaussian kernel.

Let  $\phi(\mathbf{X}) = [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots]$ ,  $\mathbf{K} = \phi(\mathbf{X})^T \phi(\mathbf{X})$ , and  $\mathbf{W}_\phi$  be the transformation matrix that maps  $\phi(\mathbf{x})$  to  $\phi(\mathbf{y})$  in some feature space. Then the criterion similar to that in (5) can be defined as

$$\frac{G'_b}{G'_w} = \frac{\text{Trace}(\mathbf{W}_\phi^T \phi(\mathbf{X}) \mathbf{C}^b \phi(\mathbf{X})^T \mathbf{W}_\phi)}{\text{Trace}(\mathbf{W}_\phi^T \phi(\mathbf{X}) \mathbf{C}^w \phi(\mathbf{X})^T \mathbf{W}_\phi)}. \quad (6)$$

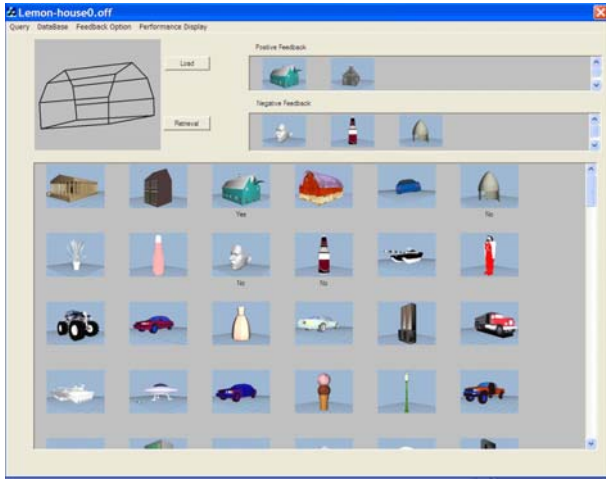


Figure 4: Our 3D retrieval tool.

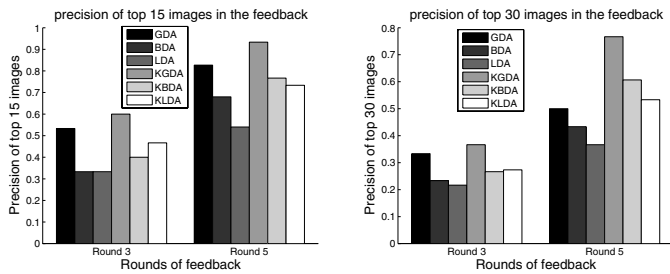


Figure 5: Comparison of the six relevance feedback algorithms.

With the kernel trick  $\mathbf{W}_\phi = \phi(\mathbf{X})\mathbf{A}$ , (6) is reduced to

$$\frac{G'_b}{G'_w} = \frac{\text{Trace}(\mathbf{A}^T \mathbf{K} \mathbf{C}^b \mathbf{K} \mathbf{A})}{\text{Trace}(\mathbf{A}^T \mathbf{K} \mathbf{C}^w \mathbf{K} \mathbf{A})}. \quad (7)$$

Since  $\phi(\mathbf{y}) = \mathbf{W}_\phi^T \phi(\mathbf{x}) = \mathbf{A}^T [K(\mathbf{x}_1, \mathbf{x}), K(\mathbf{x}_2, \mathbf{x}), \dots]$ , we need not find the optimal  $\mathbf{W}_\phi^*$ . Instead, we can directly find the optimal  $\mathbf{A}^*$  by maximizing (7).

## 4. EXPERIMENTAL RESULTS

In our experiments, we choose Princeton Benchmark [11] as our testing database, which contains 1814 models of 161 classes. A screenshot of our 3D object retrieval system is displayed in Fig 4.

Now we compare the relevance feedback performances of GDA, KGDA, LDA, KLDA, BDA, and KBDA on 3D object retrieval. We randomly select 50 objects from the database first, and then use them as the queries for testing the algorithms. After a number of feedback iterations, we compare the retrieved precision, which is measured by the ratio of the number of relevant objects to the number of total retrieved objects. At each round of feedback, we select 5 positive instances and 5 negative instances. Fig. 5 shows the comparison results of the average retrieval precision, obtained by each algorithm with the same 3D queries. The left and right figures give the results when top 15 and 30 retrieved objects are considered, respectively. Fig. 5 clearly illustrates that KGDA performs best and GDA outperforms LDA and BDA.

## 5. CONCLUSION

In this paper, we first present a new method of 3D reconstruction from 2D line drawings inputted by the user as queries. It can recover 3D objects with both planar and curved surfaces. Our method cannot reconstruct complex free form objects. However, it provides an easy-to-use interface for 3D object retrieval. We also propose a graph-based relevance feedback algorithm and its kernelized version, to refine retrieved objects. Experiments show that our approach outperforms some previous relevance feedback algorithms when they are used for 3D object retrieval.

## 6. ACKNOWLEDGMENTS

This work was supported by the Research Grants Council of the Hong Kong SAR (Project No. CUHK 414306) and the CUHK Direct Grant. The authors would like to thank Peng Du, Shuicheng Yan, Dahua Lin, and Yi Liu for their stimulating discussions and assistance in the programming.

## 7. REFERENCES

- [1] H. Barrow and J. Tenenbaum. Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence*, 17(1-3):75–116, 1981.
- [2] L. Cao. 3D object reconstruction from line drawings. *Master Thesis, Dept. Information Engineering, the Chinese University of Hong Kong*, 2005.
- [3] C. Chow and C. Liu. Approximating discrete probability distributions with dependence trees. *IEEE Trans. Information Theory*, 14(3):462–467, 1968.
- [4] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1987.
- [5] T. Igarashi, S. Matsuoka, and H. Tanaka. Teddy: a sketching interface for 3D freeform design. *SIGGRAPH*, 1999.
- [6] H. Lipson and M. Shpitalni. Optimization-based reconstruction of a 3D object from a single freehand line drawing. *CAD*, 28(8):651–663, 1996.
- [7] J. Liu, Y. Lee, and W. Cham. Identifying faces in a 2D line drawing representing a manifold object. *TPAMI*, 24(12):1579–1593, 2002.
- [8] P. Min, M. Kazhdan, and T. Funkhouser. A comparison of text and shape matching for retrieval of online 3D models. *European Conf. Digital Libraries*, 2004.
- [9] Y. Rui, T. S. Huang, and S. Mehrotra. Content-based image retrieval with relevance feedback in mars. *ICIP*, 2:815–818, 1997.
- [10] A. Shesh and B. Chen. Smartpaper: An interactive and user friendly sketching system. *EuroGraphics*, 2004.
- [11] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. The princeton shape benchmark. *Proc. Shape Modeling Applications*, 2004.
- [12] T. Funkhouser et.al. A search engine for 3D models. *ACM Trans. Graphics*, 22(1):83–105, 2003.
- [13] J. W. Tangelder and R. C. Veltkamp. A survey of content based 3D shape retrieval methods. *Proc. Shape Modeling Applications*, 2004.
- [14] X. S. Zhou and T. S. Huang. Small sample learning during multimedia retrieval using biasmap. *CVPR*, 1:11–17, 2001.