

Sketch-based 3D shape retrieval

Mathias Eitz¹, Kristian Hildebrand¹, Tamy Boubekeur² and Marc Alexa¹
¹TU Berlin, ²Telecom ParisTech & LTCI CNRS

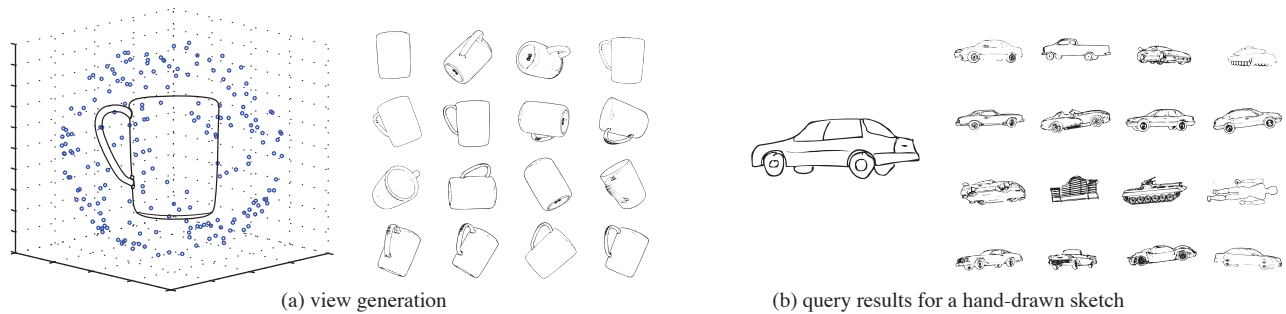


Figure 1: 3D model search from simple user sketches: a) random view generation; b) retrieval results given a sketch of a car.

1 Introduction

As large collections of 3D models are starting to become as common as public image collections, the need arises to *quickly* locate models in such collections. Models are often insufficiently annotated such that a keyword based search is not promising. Our approach for content based searching of 3D models relies entirely on visual analysis and is based on the observation that a large part of our perception of shapes stems from their salient features, usually captured by dominant lines in their display. Recent research on such feature lines has shown that 1) people mostly draw the same lines when asked to depict a certain model and 2) the shape of an object is well represented by the set of feature lines generated by recent NPR line drawing algorithms [Cole et al. 2009]. Consequently, we suggest an image based approach for 3D shape retrieval, exploiting the similarity of human sketches and the results of current line drawing algorithms. Our search engine takes a sketch of the desired model drawn by a user as the input and compares this sketch to a set of line drawings automatically generated for each of the models in the collection.

2 Feature Lines based Search

We build our engine upon a *bag-of-features* search [Squire et al. 1999], which has become the method of choice for affine invariant image retrieval during the last years. This allows us to map the difficult 3D retrieval problem to a simpler image retrieval problem.

In a bag-of-features approach, each image is represented by a large set of small local features that do *not* carry any spatial information. Those features are quantized to form a “visual vocabulary” in a training step and each image is then represented by its specific distribution of “visual words”. To query a collection, the distribution of visual words of the query image is computed and the collection images with most similar distributions are returned. We accelerate this process by using inverted indices. As a result, only images that have at least one visual word in common with the query need to be compared and the query is extremely quick. Consequently, we need to define two specific components for our scenario: 1) generating a complete set of representative line drawings for each model and 2) new line-based local descriptors, which provide a compact translation-invariant representation of these lines. Note that image-based descriptors cannot be used since sketches are usually rough and texture-less.

We generate the images by sampling a set of camera position uniformly from the bounding sphere of the model, see Figure 1 a) (we used 50 views per model in this work). The camera position is used as input for view-dependent line drawing algorithms, such as Apparent Ridges [Judd et al. 2007]. This results in a set of curves in the plane, which are the input for computing the descriptors. Note that the combination of random views and descriptors based on lines naturally supports retrieval based on parts of the shape – a task that is usually difficult in systems based on global shape descriptors.

As descriptors for the projected lines we propose to use a histogram of oriented gradients extracted from small local regions. Note that the queries generated by users are binary sketches and thus only contain information about the direction of a line – so information about the length of the gradient is missing.

3 Results and Discussion

Since our proposed shape retrieval engine is based on a *bag-of-features* search it comes with several desirable properties compared to existing methods [Funkhouser et al. 2003]. Since we use small *local* features, users do not need to sketch the full shape, part-based retrieval is naturally supported. The retrieval is translation invariant – we are currently working on integrating full invariance to affine transformations. Rotation invariance can be easily achieved by computing the local histograms with respect to a reference frame define by the image content. A query is very fast and we believe that this approach can be readily scaled to much larger shape collections. We have built our initial system on the 2,000 models contained in the Princeton shape database; a query on this collection takes only about 10 ms. Finally, as most 3D models do not carry texture information, we can use our engine to search them from pictures as well, by using recent image contour extraction methods.

References

- COLE, F., SANIK, K., DECARLO, D., FINKELSTEIN, A., FUNKHOUSER, T., RUSINKIEWICZ, S., AND SINGH, M. 2009. How well do line drawings depict shape?
- FUNKHOUSER, T., MIN, P., KAZHDAN, M., CHEN, J., HALDERMAN, A., DOBKIN, D., AND JACOBS, D. 2003. A search engine for 3D models. *ACM Transactions on Graphics* 22, 1.
- JUDD, T., DURAND, F., AND ADELSON, E. H. 2007. Apparent ridges for line drawing. *ACM Transactions on Graphics* 26, 3.
- SQUIRE, D., MUELLER, W., MUELLER, H., AND RAKI, J. 1999. Content-based query of image databases. In *Scand. Conf. on Image Analysis*.