FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

PROJECTOS DE INVESTIGAÇÃO CIENTÍFICA E DESENVOLVIMENTO TECNOLÓGICO

# *Report on Experiments with 3D Shape Descriptors*

Autores:
**Lisha Zhang**
**Manuel João da Fonseca**
**Alfredo Ferreira**

Acrónimo: **DecorAR**

Título: **Decoração Aumentada: Combinando Realidade Aumentada e Recuperação de Objectos 3D para a Decoração de Interiores**

Referência: **POSC/EIA/59938/2004**

União Europeia - Fundos Estruturais

Governo da República Portuguesa

# Table of Contents

**Abstract:** After the review of popular 3D shape descriptors in 3D shape recognition and retrieval, we develop a prototype to implement four of the commonly used shape descriptors, as well as all the combinations of them. We use shape models in a subset of Princeton Shape Benchmark (PSB) and carry out some experiments on each of the shape descriptors, show their feature vectors and retrieved models, and compare their retrieval results. It is proved by our experiments that the shape descriptors of cord and angle histogram (CAH), shape distribution (SD) and complex function FFT (CFFFT) perform much better than complex EGI. Additionally, we concluded that most of the combined descriptors achieve better results than the single ones.

## 1. Introduction

In the previous survey [1] of 3D shape descriptors, we review some of the most popular 3D shape descriptors for 3D shape classification and retrieval and engage in the classification and comparison of these shape descriptors. In this report, we choose four appropriate 3D shape descriptors for the experiments according to the review and comparison in the survey, which are cord and angle histogram(CAH), shape distribution(SD), complex extended Gaussian image(CEGI) and complex function FFT(CFFFT).

We develop a prototype for 3D shape recognition and retrieval. It implements the four shape descriptors and all the combinations of compound descriptors. In section 2, we introduce the framework of our prototype and explain each part of it in detail. In section 3, we describe the dataset involved in our experiments and the classification of the shape models. And then we give the measures in the evaluation of the shape descriptors' performances in section 4. Section 5 analyzes the feature vectors of shape models produced by each of the shape descriptors and shows an example of various feature vectors. After that, we carry out the experiments on the single shape descriptors as well as the combined descriptors in section 6, and compare their retrieval results from the aspects of R-precision, 2R-precision and 11p-precision, and the retrieved models in top 10 and top 20 respectively. In this section, we select an example shape to further explain the effectiveness of shape descriptors. In the final section, a conclusion of experiments is given.

## 2. Prototype

Figure 1 shows the framework of our prototype. The prototype mainly contain three modules: ***file readers***(including vrml-files reader and off-files reader), ***feature extractor***, and ***shape matching***. The module of file readers reads files in formats of .wrl or .off(used in PSB database). Feature extractor module extracts the feature vectors from shape information which is acquired from vrml or off files. In the phase of training, it produces feature vectors of 3D shape models and saved them to the dataset; in the phase of recognition, it extracts feature vector from the query shape and proceeds it to the next module. The module of shape matching, after receiving a feature vector of query shape and certain feature vectors from model dataset, calculates the similarities between the query shape and the certain models and figures out similarity lists.

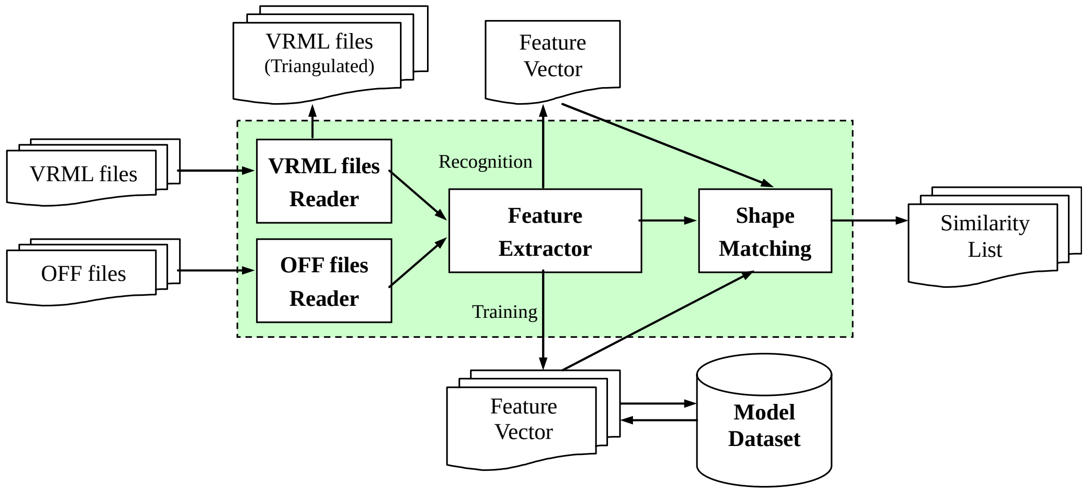



**Figure 1.** Framework of the prototype implementing 3D shape descriptors

# 3. Dataset and Measures

For the performance evaluation experiments, we used the Princeton Shape Benchmark (PSB)[2] database. PSB contains the total of 1914 models, and we used totally 400 models (Dataset1: m0-m199 & Dataset2: m800-m999) for the shape classification.

There are 12 classes of creatures in the first part of models (m0-m199) and 8 classes of furniture in the second part of models (m800-m999), as listed in the following table:

**Table 1.** The classes of the models

| Class | Models | Class size |
|---|---|---|
| Insect | m0-m8 | 9 |
| Butterfly | m9-m15 | 7 |
| Spider | m16-m26 | 11 |
| Bird | m27-m47 | 21 |
| Duck | m48-m52 | 5 |
| Fish | m53-m82 | 30 |
| Snake | m82-m85 | 4 |
| Dog | m86-m98 | 13 |
| Pig | m99-m102 | 4 |
| Horse | m103-m108 | 6 |
| Rabbit | m109-m112 | 4 |
| Human | m113-m199 | 87 |
| Chair | m800-m828 | 29 |
| Sofa | m829-m843 | 15 |
| Shelf | m844-m869 | 26 |
| Table | m870-m938 | 69 |
| Table suit | m939-m943 | 5 |
| Bed | m944-m951 | 8 |
| Cabinet | m952-m960 | 9 |
| Decorative plant | m961-m999 | 39 |

As for the quantitative measures of performance, we used the R-precision(1R), the 2R-precision(2R), the 11 point average precision(11P) figures[3], the precisions and recalls of top10 models and top 20 models, as well as the precision-recall plot.

- R-precision: the ratio of the models retrieved from the desired class $C_k$ (i.e. the same class as the query) in the top R retrievals, in which R is the size of the class $|C_k|$.
- 2R-precision: similar to the R-precision, except that the figure is computed using the top 2R retrievals.
- 11-Point average precision: the average of precision values taken at 11 equally spaced recall values {0.0,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0}.

The 1R and 2R values favor methods having higher precision for the “near the top” retrievals. A 11P average precision value can be considered as a summary of the recall-precision plot, which emphasizes overall performance.

# 4. Feature vectors

We chose four 3D shape descriptors as mentioned in the survey. In the phase of training, feature vectors are extracted from 3D shape models with a particular shape descriptor and saved in the dataset as the representations of corresponding models; in the phase of recognition, feature vectors are also extracted with the same descriptor to represent query shapes. These shape descriptors are described in detail as follows:

(1) Cord and Angle Histogram(CAH): a cord is defined as a ray segment which joins the barycenter of the mesh with a triangle center.

- Length histogram: the length of the rays;
- Angles (xy plane, xz plane, yz plane) histograms: the angles of the rays to the three coordinate planes.

(2) Shape Distribution(SD):

- Angle histogram: angle formed by three random surface points;
- Center distance histogram: distance of a random surface point to the center of mass of the model;
- Point distance histogram: distance between two random surface points;
- Area histogram: area(square root) of the triangle defined by three random surface points;
- Volume histogram: volume(cube root) of the tetrahedron defined by four random surface points.

(3) Complex EGI(CEGI):

- Area histogram: the visible face area, composing the magnitude of complex function;
- Normal distance histogram: the normal distance of the face from the designated origin in the direction of the normal, composing the phase of complex function.

(4) Complex function FFT(CFFFT):

- Complex Fourier coefficients $\hat{r}(l,m)$: real and imaginary parts figured out as follows:

(a) express the maximal distance from the center of mass to the shape as a function of the ray-based spherical angle; (b) approximate the function using spherical harmonic basis function( $Y_{l,m}$ ); (c) figure out the complex Fourier coefficients $\hat{r}(l,m)$; and (d) form the corresponding feature vector with absolute values $\left|\hat{r}(l,m)\right|$ of dimension $\dim = (l_{\max}+1)l_{\max}/2$.

Figure 2 shows the feature vectors extracted from the 3D shape m87 with cord and angle histogram, shape distributions and complex EGI, described above, respectively.

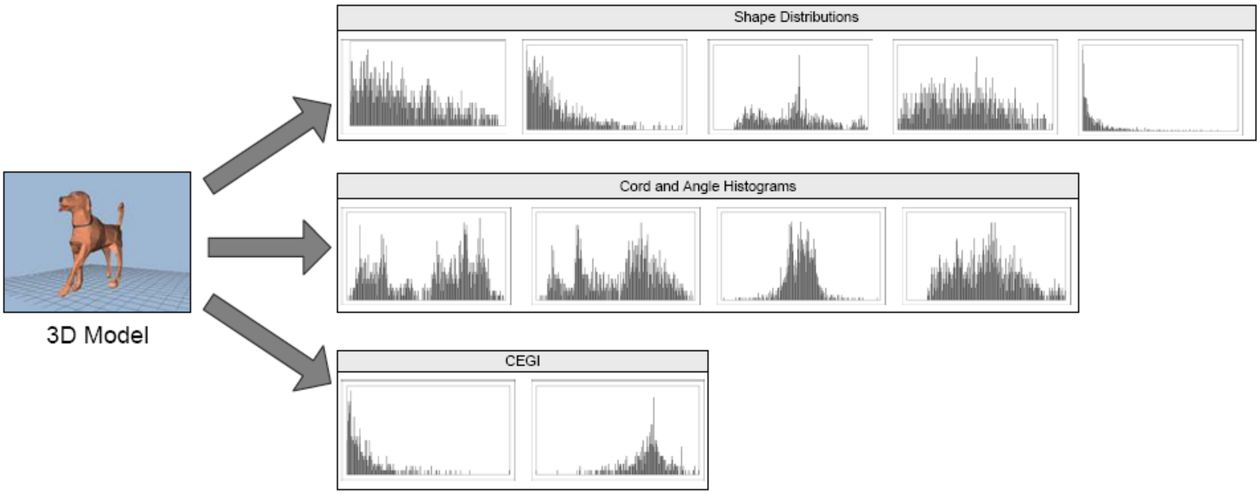



**Figure 2.** Feature vectors of different shape descriptors representing a 3D model(m87)

# 5. Retrieval results

We perform experiments on the selected PSB shape datasets with the four single 3D shape descriptors, as well as all of their combinations, and compare the recognition and retrieval results on dataset mentioned in section 3. In addition, we analyze the retrieved shapes on different model datasets to show their influences on shape descriptors.

## *5.1 Performance comparisons of descriptors*

In these experiments, we compared the performance of 3D shape descriptors on the datasets of DS1+DS2. Table 2 summarized the experiments via 1R-precision, 2R-precision and 11p-precision, both for single shape descriptors CAH, SD-A1, SD-D1, SD-D2, SD-D3, SD-D4, CEGI and CFFFT, and almost all possible combinations of them. It is shown that the most approving results out of them are produced by SD-5F(combination of the 5 shape functions), combination of CAH, SD and CFFFT, and combination of SD and CFFFT, with the highest values 42.55%, 37.97% and 36.92% in 1R respectively. In addition, the single descriptor of shape distribution-D2 and the combination of CAH and SD also yield similar results as the three mentioned above. The retrieval results of the five best combinations of descriptors are highlighted in the table.

As for single shape descriptors, SD-D2 is the best among all the single shape distribution descriptors using shape functions. If we consider SD-5f as a single descriptor integrating 5 shape functions into one descriptor, it is even better than SD-D2, therefore, the precision-recall plot of SD-5f is on the top of the others, which can be seen in Figure 3. Furthermore, all the combinations consisting of SD-5f performed better than those without SD-5f. On the contrary, it is shown in Figure 4 that CEGI descriptor produced a very poor classification result, and moreover, all combinative descriptors containing CEGI decrease the corresponding precisions of those without CEGI.

As for the combinative descriptors, it is obviously shown in the table and Figure 5 that all combinations outperformed each single descriptor and the subset combinations, except those involving CEGI.

**Table 2.** Performance comparisons of the combinations of shape descriptors

| Descriptors | | 1R-precision(%) | 2R-precision(%) | 11p-precision |
|---|---|---|---|---|
| CAH | | 33.49 | 44.98 | 0.316 |
| SD | A1 | 29.53 | 44.28 | 0.271 |
| | D1 | 32.57 | 49.82 | 0.302 |
| | D2 | 36.84 | 53.70 | 0.339 |
| | D3 | 29.57 | 44.53 | 0.265 |
| | D4 | 26.85 | 42.89 | 0.244 |
| | 5F(all) | 42.55 | 58.01 | 0.401 |
| CEGI | | 23.61 | 39.61 | 0.215 |
| CFFFT | | 31.07 | 43.95 | 0.284 |
| CAH+ SD | | 36.57 | 48.68 | 0.349 |
| CAH+ CEGI | | 30.42 | 45.70 | 0.283 |
| CAH+ CFFFT | | 35.62 | 47.29 | 0.340 |
| SD+ CEGI | | 27.10 | 44.42 | 0.248 |
| SD+ CFFFT | | 36.92 | 50.75 | 0.343 |
| CEGI+ CFFFT | | 28.93 | 45.21 | 0.265 |
| CAH+ SD+ CEGI | | 33.02 | 48.25 | 0.308 |
| CAH+ SD+ CFFFT | | 37.97 | 49.99 | 0.363 |
| CAH+ CEGI+ CFFFT | | 34.04 | 48.86 | 0.315 |
| SD+ CEGI+ CFFFT | | 31.78 | 48.20 | 0.294 |
| CAH+ SD+ CEGI+ CFFFT | | 35.95 | 51.00 | 0.335 |

## 5.2 Retrieved models

In this experiment, we choose a 3D shape model m911 from the dataset as the query shape, which belongs to the class "Table". Model m911 is shown in Figure 6. The four single shape descriptors and all their combinations are employed to retrieve m911 from 400 shape models. We observe and compare the retrieval results, which are concluded in Table 3.

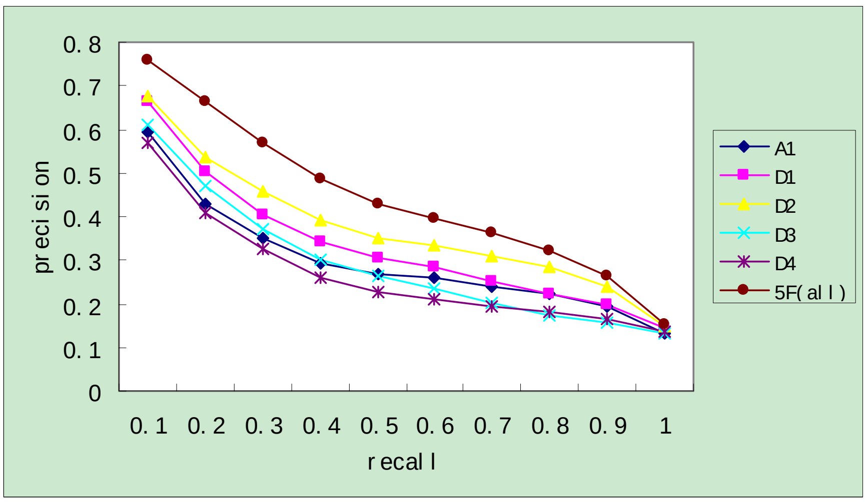



**Figure 3.** Precision-recall plots of different shape functions of shape distribution

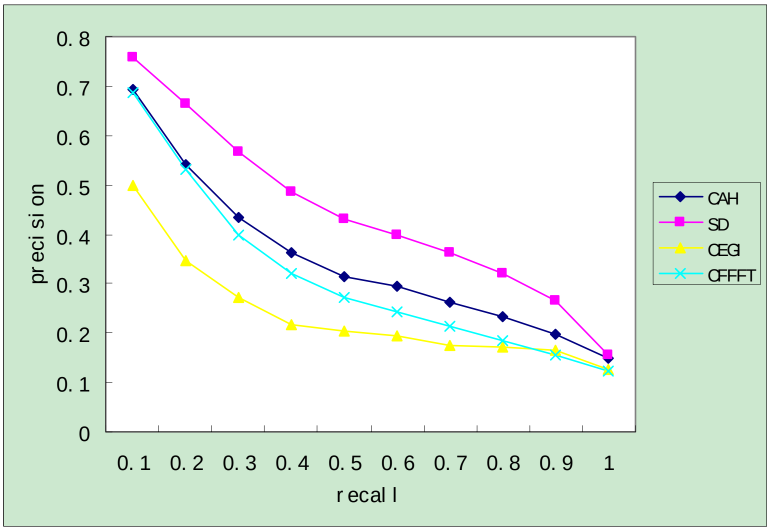



**Figure 4.** Precision-recall plots of single shape descriptors

As for single shape descriptors, it can be seen in Table 3 that CFFFT results in the best R-precision, 2R-precision and 11p-precision out of the four single shape descriptors. It retrieves 9 and 17 models of class "Table" in the top 10 and top 20 models respectively, as shown in Figure 7(c). SD and CAH behave inferior from the aspect of the three measures, while perform better in precisions and recalls of top 10 and top 20 models, as shown in Figure 7(a) and Figure 7(b). However, CEGI is poor in retrieval of this shape model and retrieve only 6 similar shape models in top 20, most of which appears in the latter part (Figure 7(d)).

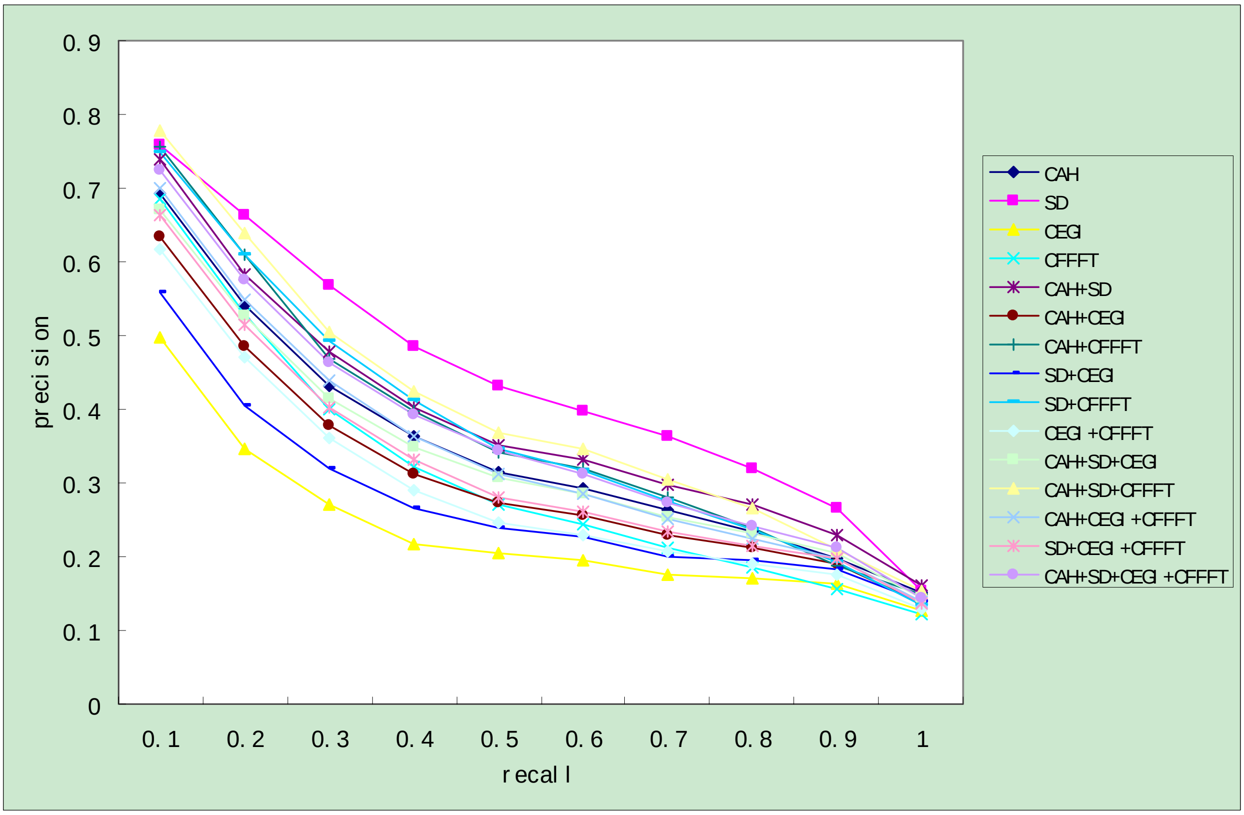



**Figure 5.** Precision-recall plots of the single and combined shape descriptors

On the other hand, it is evident that most of retrieved models in CAH and CFFFT are in similar principal coordinate axis with the query model m911, and the exceptions appear after the 12th, but those in SD descriptor present in comparatively arbitrary directions. The reason is that SD descriptor is coordinate transformation invariant, while CAH and CFFFT descriptors are more or less dependent on rotation for the employed information of angles. In addition, although information of surface normal is used in CEGI, it acts according to the surface area and basically independent on coordinate system.

As for the combined shape descriptors, the best ones are SD_CFFFT, CAH_SD_CEGI_CFFFT, CAH_SD_CFFFT and CAH_CFFFT. It is shown in Table 3 that SD_CFFFT outperforms the other three in 1R-precison and 2R-precision, with the highest values of 65.22% and 88.41% respectively, while 11p-precision of 0.618 is a bit lower than that of CAH_SD_CFFFT, i.e. 0.614. The reason is that the retrieved models in the top R(size of class) with combined descriptor SD_CFFFT are more than those retrieved by CAH_SD_CFFFT, while the retrieval models in most of the top $R_i$($R_i$<R) are less than those of CAH_SD_CFFFT, which leads to lower average precisions, i.e. 11p-precision. Such phenomenon can be observed in Figure 8. The retrieved models by SD_CFFFT in Figure 8(b) are 10 and 17 of top 10 and top 20 respectively, and those retrieved by CAH_SD_CFFFT are respectively 10 and 20.

It is obvious that some of the combined shape descriptors outperform the single shape descriptors, but it is not absolutely right. For example, the retrieval result of CFFFT is better than CAH_CEGI_CFFFT, the combination of three single descriptors, which probably have more useful information to determine the shapes. Moreover, CFFFT performs much better than CAH_CEGI, SD_CEGI and CAH_SD_CEGI. Because of CEGI's poor retrieval results, descriptor combinations including CEGI are possibly hindered by it and not able to improve the effectiveness.

**Table 3.** Detailed information of the retrieval results of shape model m911

| Descriptor | 1R | 2R | top 10 | precision | recall | top 20 | precision | recall | 11p |
|---|---|---|---|---|---|---|---|---|---|
| CAH | 43.48% | 66.67% | 10 | 100.00% | 14.49% | 17 | 85.00% | 24.64% | 0.444 |
| SD | 53.62% | 69.57% | 10 | 100.00% | 14.49% | 19 | 95.00% | 27.54% | 0.519 |
| CEGI | 28.99% | 39.13% | 1 | 10.00% | 1.45% | 6 | 30.00% | 8.70% | 0.204 |
| CFFFT | 57.97% | 84.06% | 9 | 90.00% | 13.04% | 17 | 85.00% | 24.64% | 0.555 |
| CAH_SD | 50.72% | 69.57% | 10 | 100.00% | 14.49% | 18 | 90.00% | 26.09% | 0.502 |
| CAH_CEGI | 39.13% | 52.17% | 7 | 70.00% | 10.14% | 12 | 60.00% | 17.39% | 0.304 |
| CAH_CFFFT | 60.87% | 82.61% | 10 | 100.00% | 14.49% | 20 | 100.00% | 28.99% | 0.604 |
| SD_CEGI | 31.88% | 49.28% | 4 | 40.00% | 5.80% | 8 | 40.00% | 11.59% | 0.240 |
| SD_CFFFT | 65.22% | 88.41% | 10 | 100.00% | 14.49% | 17 | 85.00% | 24.64% | 0.608 |
| CEGI_CFFFT | 49.28% | 60.87% | 8 | 80.00% | 11.59% | 16 | 80.00% | 23.19% | 0.413 |
| CAH_SD_CEGI | 39.13% | 63.77% | 8 | 80.00% | 11.59% | 12 | 60.00% | 17.39% | 0.344 |
| CAH_SD_CFFFT | 60.87% | 82.61% | 10 | 100.00% | 14.49% | 20 | 100.00% | 28.99% | 0.614 |
| CAH_CEGI_CFFFT | 57.97% | 71.01% | 9 | 90.00% | 13.04% | 17 | 85.00% | 24.64% | 0.520 |
| SD_CEGI_CFFFT | 52.17% | 66.67% | 9 | 90.00% | 13.04% | 18 | 90.00% | 26.09% | 0.478 |
| CAH_SD_CEGI_CFFFT | 62.32% | 76.81% | 9 | 90.00% | 13.04% | 18 | 90.00% | 26.09% | 0.546 |

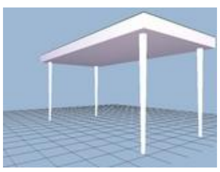


**Figure 6.** Model m911

# 6. Conclusion

This report focuses on the experiments of 3D shape recognition and retrieval with some popular shape descriptors. It introduces the framework of our prototype, describes the feature vectors of 3D shape models, and compares the retrieval results of the single descriptors and combined descriptors. It is proved that the shape descriptors of cord and angle histogram(CAH), shape distribution(SD) and complex function FFT(CFFFT) perform much better than complex EGI. Furthermore, most of the combined descriptors achieve better results than the single ones.

# References

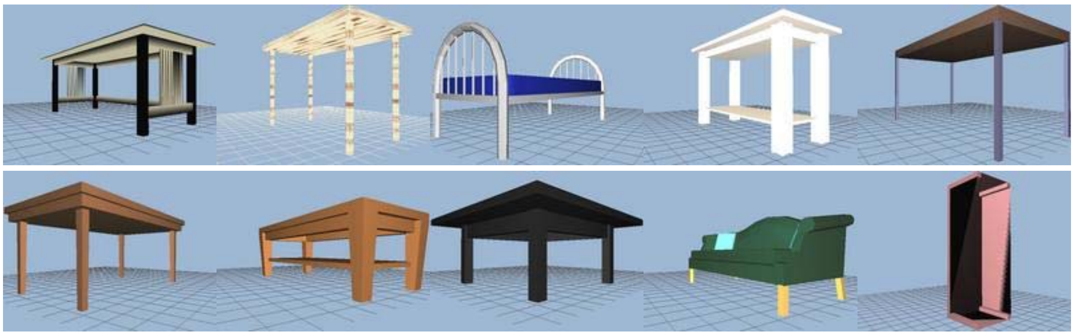
[1] Zhang Lisha, Survey on 3D Shape Descriptors. Technical report.

[2] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, D. Jacobs, A search engine for 3D models, ACM TOG, 22(1), pp.83-105, 2003.

[3] Ryuatrou Ohbuchi, Yushin Hata. Combining Multiresolution Shape Descriptors for 3D Model Retrieval. Conference proceedings ISBN 80-86943-03-8. WSCG’2006, Plzen, Czech Republic, 2006.

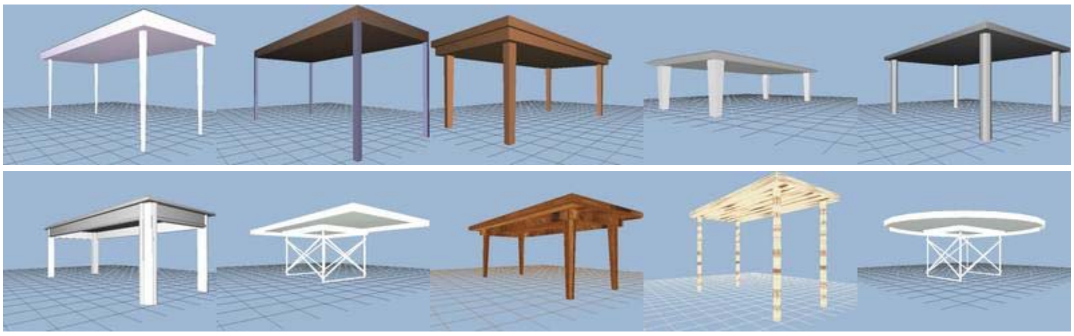
# Annex – Retrieval Results for query *m911*

Top 10

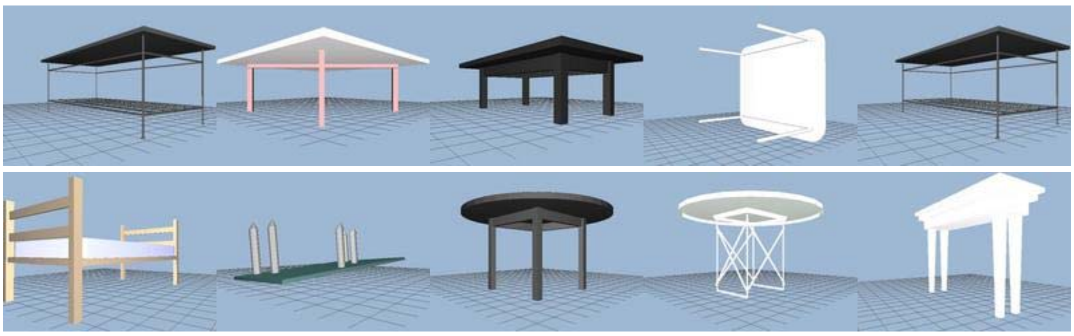


Top 11-20

**Figure 7(a).** Top 20 of retrieval results with Cord and angle histogram(CAH)



Top 10



Top 11-20

**Figure 7(b).** Top 20 of retrieval results with Shape distribution(SD)

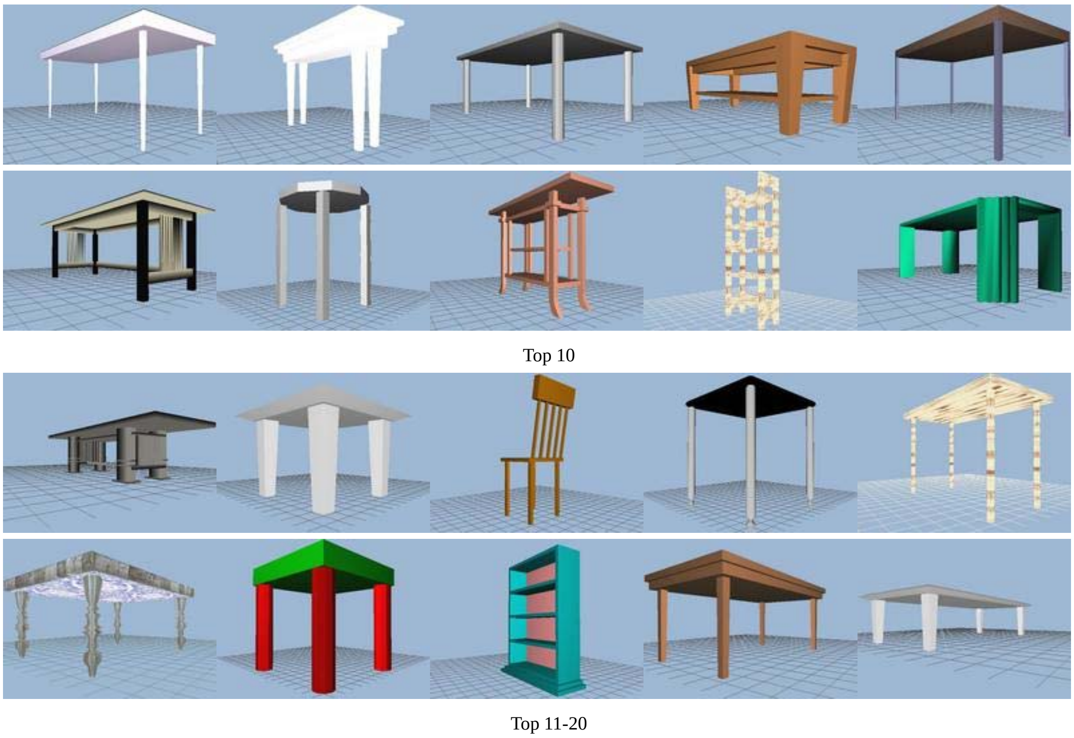



**Figure 7(c).** Top 20 of retrieval results with Complex Function FFT(CFFFT)

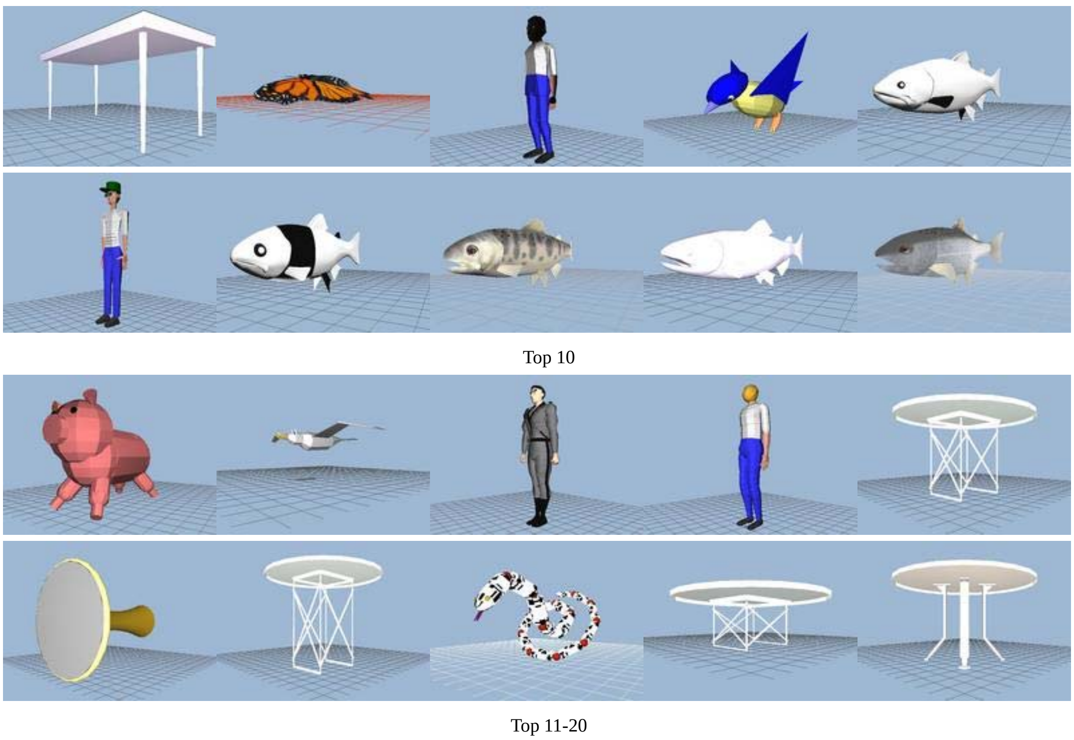
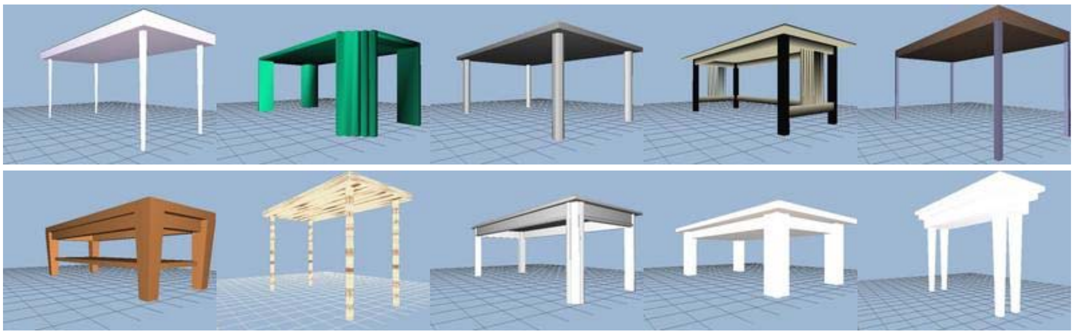
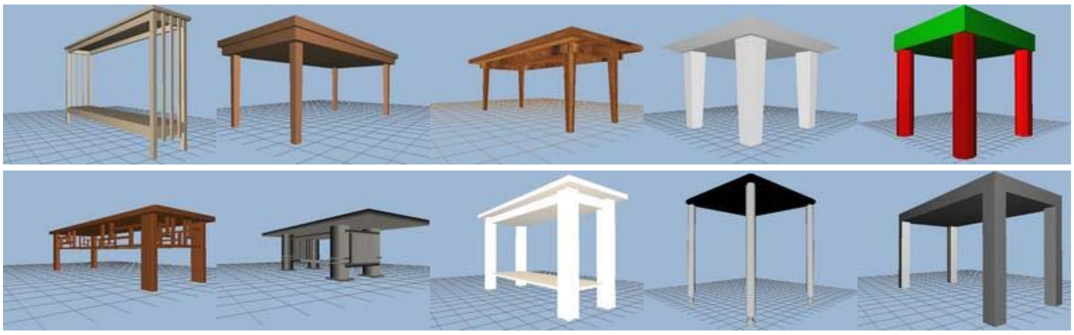



**Figure 7(d).** Top 20 of retrieval results with Complex EGI(CEGI) from DS1+DS2

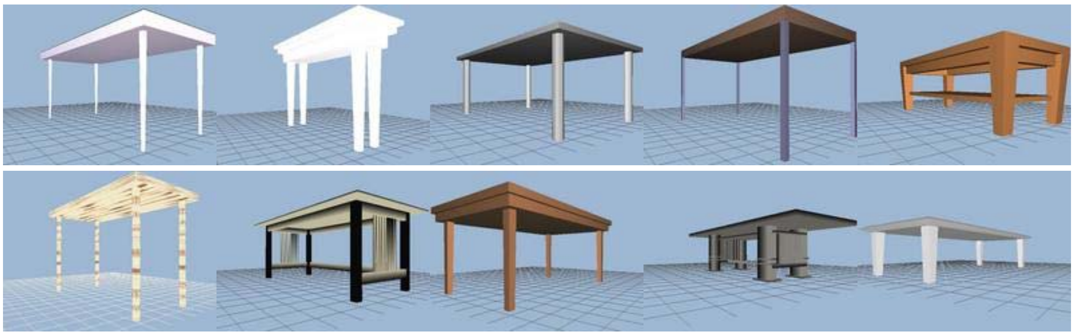


Top 10



Top 11-20

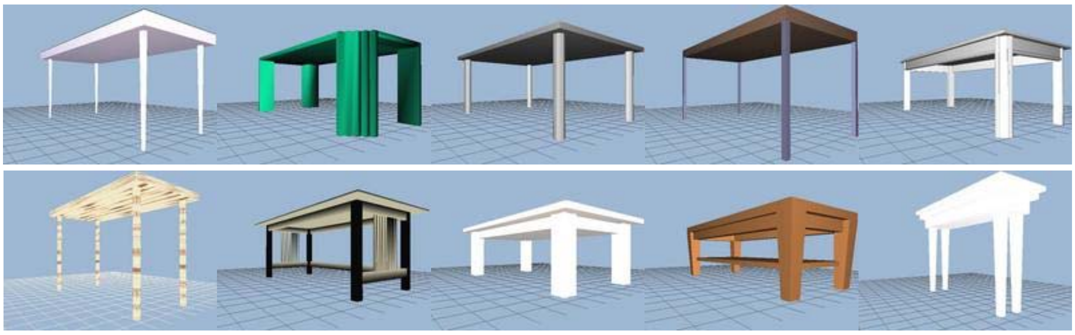
**Figure 8(a).** Top 20 of retrieval results with the combination of CAH and CFFFT



Top 10

Top 11-20

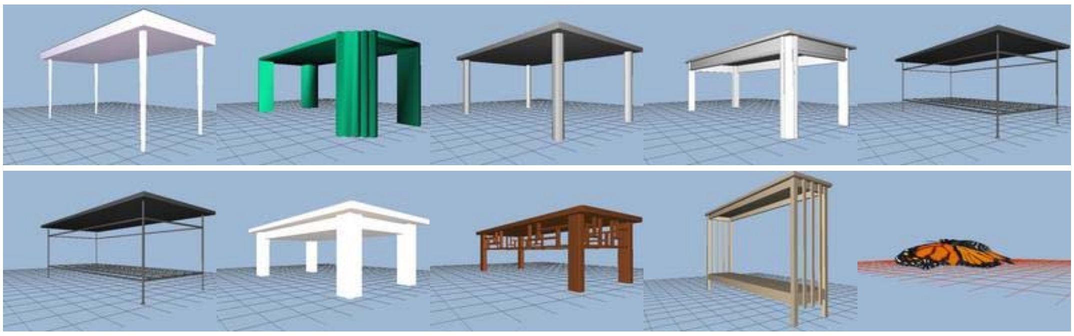
**Figure 8(b).** Top 20 of retrieval results with the combination of SD and CFFFT



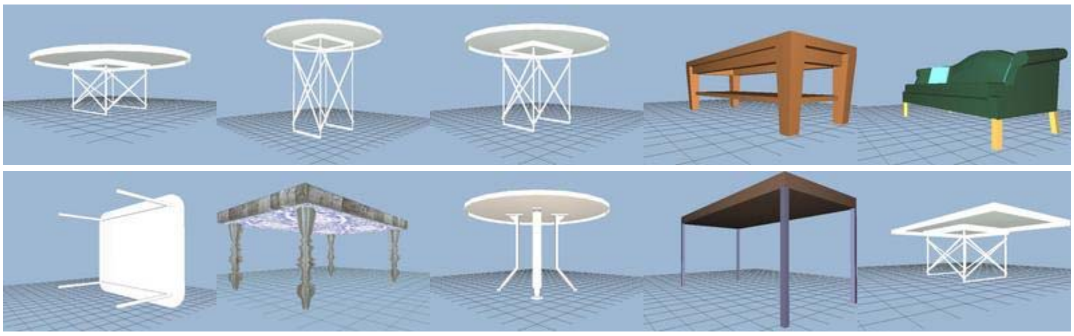
Top 10

Top 11-20

**Figure 8(c).** Top 20 of retrieval results with the combination of CAH, SD and CFFFT



Top 10



Top 11-20

**Figure 8(d).** Top 20 of retrieval results with the combination of CAH, SD, CEGI and CFFFT