

# Content-based 3D mesh model retrieval from hand-written sketch

Satoshi Kanai

Received: 11 September 2006 / Revised: 10 October 2006 / Accepted: 13 October 2006  
© Springer-Verlag France 2008

**Abstract** Recently, huge numbers of 3D models (solid, surface, and mesh models) have been archived in CA-X systems for industrial design, engineering, manufacturing and e-commerce. Efficient searching methods for 3D models become important for system users. For this purpose, a content-based 3D mesh model retrieval system from hand-written sketch is proposed in this paper. In the system, the user puts a query in the form of a 2D hand-written sketch to a 3D mesh model database, and the system automatically returns a set of mesh models whose 2D views are similar to the input sketch. Generic Fourier Descriptor and Local Binary Pattern which are features invariant to rotation, translation and scaling of 2D images are used to measure dissimilarities between a query sketch and images generated from 3D models. The effectiveness of the retrieval system is evaluated through case studies on industrial design process.

**Keywords** 3D Shape retrieval · Content-based retrieval · Mesh model · Generic Fourier descriptor · Local binary pattern

## 1 Introduction

With the diffusion of CAD/CAM/CAE/CG software into industrial design, engineering and manufacturing fields, a huge number of the 3D geometric models such as mesh models, surface models and solid models have been created and stored in databases in many enterprises. Moreover, advances in 3D scanners and reverse engineering software [1] could simplify generating 3D models from physical objects. On

the other hand, introduction of the Web3D [2] and MPEG-7 [3] standards enabled public to easily obtain various kinds of 3D mesh models over the internet, and these technology are increasingly introduced to 3D-product-catalog on the Web for e-commerce.

In these situations, there is a strong needs for natural, intuitive and user-friendly ways to search for a 3D model from huge model database. But, such search function is currently unavailable in CAX systems and traditional database management systems. An index-based search is usually used for this purpose where the users attach several index keywords to a 3D model at the time of archiving, and search for the models they want by inputting keywords. However, this index-based search is not necessarily satisfactory for 3D model retrieval, because it is difficult for user to find relevant keywords to express 3D shape of the model. Moreover, it is difficult to quantitatively measure the similarity between the keywords attached to the 3D models.

On the contrary, content-based information retrieval is attracting researchers attentions in the information retrieval field. It allows the user to directly input the content data as a query for database, and to retrieve a set of similar data from the database based on similarity measures [4]. So far, the content-based information retrieval systems have been intensively developed and put to practical use for the digital image and video contents [5]. Motivated by the content-based image retrieval, researches have also been taking interests in the content-based 3D model retrieval [6–8]. So far, their researches almost have been focusing on developing similarity measures between 3D models (solid models or mesh models), and on realizing model retrieval functions by inputting a query in the form of 3D model.

However, in the research, they assumed that users already have a 3D model before their query. This assumption is not practical and not user-friendly when the user of the system

S. Kanai (✉)  
Graduate School of Information Science and Technology,  
Hokkaido University, Sapporo, Japan  
e-mail: kanai@ssi.ist.hokudai.ac.jp

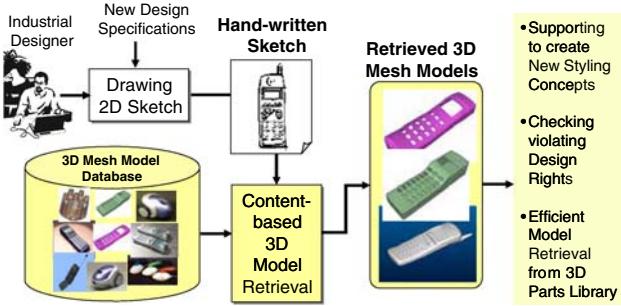


Fig. 1 An overview of proposed retrieval system

has only a dim memory or rough 2D images of the shape he/she wants to retrieve and does not have a 3D model for query yet.

To solve the drawback, a content-based 3D model retrieval system from a 2D hand-written sketch query is proposed in this research. The hand-written 2D sketch is one of simplest, most natural, intuitive and user-friendly ways to roughly express the 3D shape to be retrieved in design and engineering fields. Shown in Fig. 1, in our content-based 3D model retrieval system, the user puts a query as a 2D hand-written sketch to the 3D mesh model database, and from the database, the system returns a set of 3D models which have similar 2D views to the input sketch.

Many potential applications of this sketch-based 3D model retrieval system can be considered; aiding industrial designers to inspire new styling ideas, checking violation of existing style design rights, looking for possibility of reusing existing manufactured parts, and searching for 3D models from Web-based parts libraries. Implementing the retrieval function to the commercial 3D-CAD systems to select a CAD model file will be the other useful application. However, very few researches on sketch-based 3D model retrieval system have been studied so far.

The remaining part of the paper is organized as follows. In Sect. 2, the related works on content-based 3D model retrieval and the features of our research are discussed. In Sect. 3, an overview of our model retrieval system is provided. In Sect. 4, a rendering algorithm to generate projected images from 3D mesh model is described as a pre-processing. In Sect. 5, we describe an algorithm of the 3D model query using two kinds of image descriptors to compute dissimilarity between a hand-written sketch and a projected image of 3D model. In Sect. 6, an effectiveness of our retrieval system is evaluated by a case study on styling design of mobile phones exteriors.

## 2 Related works

There are a large number of researches on content-based information retrieval. Therefore, we only review related

works in the following two categories close to our research; (1) 3D model retrieval from 3D model query, and (2) 3D model retrieval from 2D sketch query.

The research on 3D model retrieval from 3D model query have been done in the fields of computer graphics, computer vision, pattern recognition, geometric modeling and molecular biology. Useful survey papers in this category have been published [6–8]. In [6], the content-based 3D shape retrieval methods were classified into three; feature-based, graph-based and the other one.

In the feature-based methods, shape of the 3D model is sampled, and feature or descriptor value of 3D geometry are evaluated. The method can be applied to any sort of shape model. A feature can often be expressed as a multi-dimensional vector, and the shape similarity is evaluated easily by taking a distance between two feature vectors. However, the feature-based method generally cannot discriminate detail of shapes, and partial matching cannot be done. Feature value tends to change depending on scale, orientation and position of the 3D models, and pose normalization is needed before evaluating the similarity in most case. In the feature-based methods, there are many variants depending on the feature type and similarity measures; global feature based similarity [9, 10], global feature distribution based similarity [11–13], local feature based similarity [14–16]. The feature and shape similarity from machining cost aspect for the 3D solid model was also proposed [17].

Graph-based methods attempt to extract the structure from 3D shape and to represent it as graphs indicating how shape portions are linked together. The method easily captures the topological structure of the shape. Therefore, pose normalization is not needed, and partial matching can be done. However, the matching and similarity evaluation of two graphs are more difficult than those of the feature-based approach. The pure graph based method has a limited discriminating power, because only topology is taken into account. And small changes in topology results in significant differences in similarity, therefore they are less robust than feature based methods. Three types of graph are mainly used; solid model graph based similarity [18–20], skeleton based similarity [21, 22] and Reeb graph based similarity [23].

The other types of 3D model retrieval methods from 3D model query are classified to view-based similarity, volumetric error based similarity, weighted point set based similarity and deformation based similarity. 2D view-based similarities were used for 3D model retrieval from 3D model query in the methods of Macrini et al. [24] and Cyr and Kimia [25]. They selected standard viewpoints for generating the projected images, and evaluated the similarity between two shaded images. Shock graphs were used as image descriptors for evaluating similarity. The shock graph is a directed acyclic graph of a medial axis (a skeleton) for the outer boundary contours of the projected image. Instead of the shaded

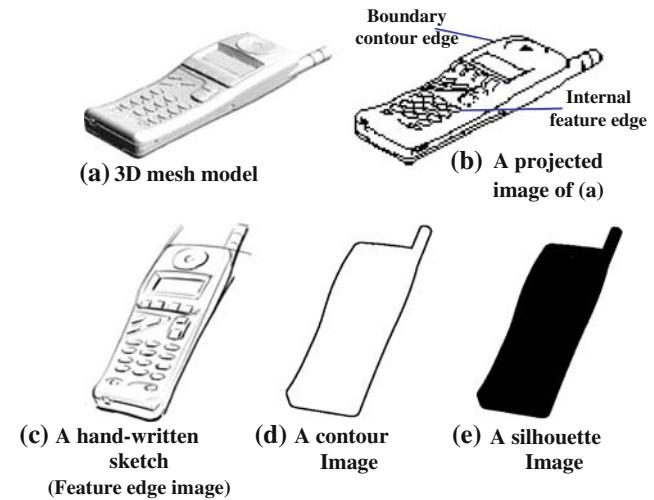
images, Obuchi evaluated the similarity between distance (2.5D) images made from 42 viewpoints around a 3D mesh model [26].

Although there are large number of the researches on 3D model retrieval from 3D model query, there is a very few on 3D model retrieval from 2D sketch query. Funkhouser et al. [27] proposed an integrated 3D model retrieval system where 3D mesh models can be retrieved both from 2D sketch query and 3D model query. In their system, the projected images of 3D model were pre-rendered from 13 view directions, and the input sketch is compared with the pre-rendered images. 2D version of spherical harmonic descriptor (SFD) was used as the 2D image descriptors for matching, and the dissimilarity between the sketch and the projected image was measured by taking an Euclidian distance between two SFDs. However, in rendering of the projected images of mesh model, they only generated images of boundary contours (silhouette curves) of the mesh model, and important feature edges on the exterior surface of the mesh model were ignored in the retrieval.

Chen et al. [28] also developed the sketch based 3D mesh model retrieval system. They pre-rendered “shaded images” for a 3D mesh model where inside region of the 2D boundary contour was completely painted black. Twenty typical viewpoints around the model were used in the rendering. Zernike moment and Fourier descriptor (FD) are used as 2D image descriptors for matching. The dissimilarity between two models were evaluated by taking minimum dissimilarity among randomly selected ten pairwise shaded images. On the other hand, Hou and Ramani [29] developed a 3D CAD model search system where the skeletal graph is used as the 3D shape features. The system is fundamentally classified into the one of 3D model retrieval from 3D model query using skeletal graph similarity, but it also allowed the user to submit a 2D skeletal sketch as a query. The user can iteratively resubmit the query sketch to narrow the search results.

Pu et al. [30] also developed a 2D sketch-based user interface for 3D CAD model retrieval. The user submit the query by sketching three orthogonal views of the 3D models. The system generates a set of three 2D orthogonal line drawings which represent shilhouettes of 3D CAD models in the database. The similarity between the query sketch and 2D line drawing in different views are measured using 2D shape distribution to retrieve the 3D CAD models. Shown in Fig. 2, a projected image of the 3D mesh generally includes both outer boundary contour edges and internal feature edges on the model surface. Both of them represents crucial visual aspects of the 3D product shape. Similarly, in the industrial design and mechanical design, feature edges on 2D sketches drawn by the designers express functionally important and very specific appearance of the 3D product shape.

However, researches on 3D model retrieval from 2D sketch query of [27, 28] only evaluated the similarity between 2D



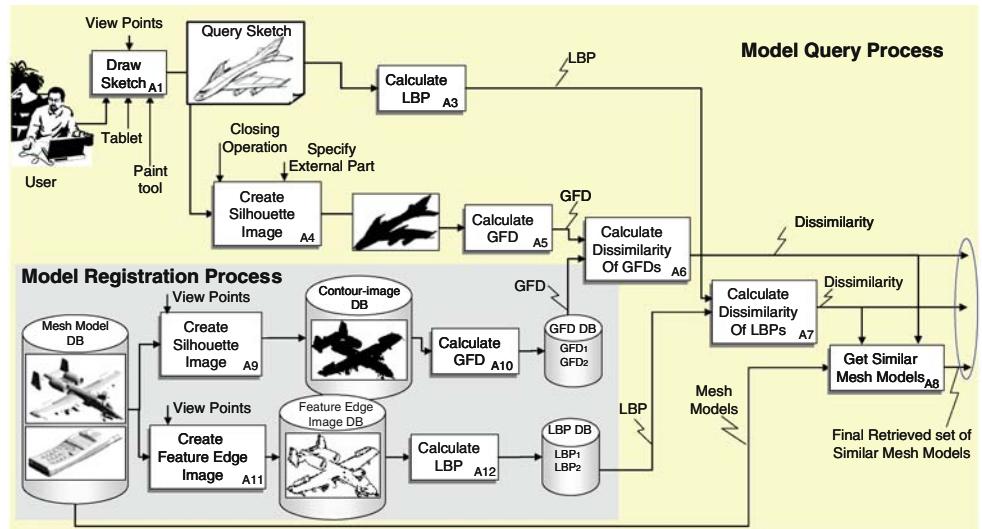
**Fig. 2** Edges and Images in projected image of 3D mesh model and hand-written sketch

contour edge images generated from the 3D mesh models. Moreover the skeletal graph approach [29] abstracted the topological structure from the 3D shape and ignored detail surface characteristic of the shape. These previous approaches did not compare feature edge structures on the projected image of 3D mesh model with the one of input sketch in their retrieval processes, and the similarities and search functions in these systems ignored the important geometric aspects in the 2D query sketch. On the other hand, the similarity and the search function in sketch-based user interface of Pu et al. [30] reflected both the contour edges and the internal feature edges on the 3D models. They limited their views to three orthogonal ones of the model, and user has to submit three orthogonal sketches to the system for query.

However, in industrial design and mechanical design, the user often expresses their shape concepts as isometric drawings for graphic explanation, because it can illustrate the 3D shape concepts with only one view drawing. Isometric drawing are often using as rough sketch in industrial design, operating instruction and users manual in manufacturing. Therefore, to make the query interface more suitable for supporting the design and engineering field, the system has to allow the user to submit a query in the form of a single view and isometric sketch including internal feature edges. Compared to the previous researches, advantages of our proposed sketch-based 3D model retrieval system are summarized as follows:

1. Our system evaluates similarities of not only outer contour edges but of internal feature edges of the query sketch in the retrieval. It also accepts a single isometric sketch as a query. This enables the user to submit a query in the form of natural and detailed 2D handwritten

**Fig. 3** Functional configuration of the proposed sketch-based 3D model retrieval system



sketch, and enables a system to obtain more relevant search results from the 3D model database.

2. Our system uses a 2D affine invariant image retrieval method by combining GFD and LBP as image descriptors. So the system realizes a search function invariant to position, orientation and scaling of 3D models and those of 2D projected images.

### 3 Overview of sketch based 3D model retrieval system

Figure 3 shows an overview of our proposed sketch-based 3D model retrieval system. The model retrieval consists of 12 activities denoted by A1–A12 in the figure. The system accepts a hand-written query sketch, and output a set of 3D mesh models which have similar 2D views to the sketch and their dissimilarity measures from 3D mesh model database. The activities in the system are mainly classified into a model registration process and a model query process. The functions of each activity are summarized as follows:

#### Model Registration Process

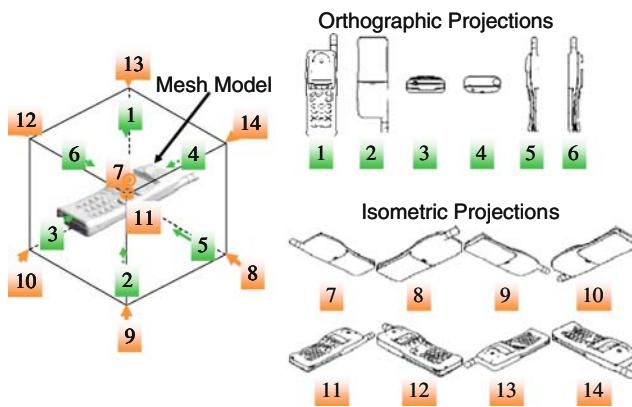
1. Feature Edge Image and Silhouette Image Generation for 3D mesh model (A9, A11): By using a non-photorealistic rendering algorithm, a set of silhouette images and feature edge images projected from several representative viewpoints are rendered from a 3D mesh model. In this paper, a “silhouette image” is the one where an inside of the contour edge is completely painted black. The generated images are stored into the silhouette image database and the feature edge image database, respectively.
2. Image descriptor Database Generation (A10, A12): Two types of image descriptors are generated. Generic Fourier descriptors (GFDs) are calculated from silhouette images, while Local Binary Patterns (LBPs) are from feature edge images. These descriptors are, respectively,

stored in image descriptor databases, and are used for image retrieval which is invariant to rotation, translation and scaling.

#### Model Query Process

3. Sketch Input (A1): The user directly draws a query sketch on the screen of Pen-tablet PC, and the image for the sketch is obtained. The sketch include contour edges and internal feature edges.
4. Silhouette Image Generation for Sketch (A4): The silhouette image is automatically generated from the image of input sketch using morphological operators.
5. Image descriptor Generation for Sketch (A3, A5): Similar to A10 and A12, GFD for a silhouette image of the sketch and LBP for a feature edge image of sketch are generated, respectively.
6. Dissimilarity Evaluation using GFD and Creation of Candidate Retrieved Set (A6, A7): A dissimilarity measures between the GFDs of the silhouette image of query sketch and ones in the silhouette image database are evaluated, and a set of 3D mesh models with less dissimilarity measures is selected to make a candidate retrieved set.
7. Dissimilarity Evaluation using LBP (A8): A dissimilarity measures between the LBP of a feature edge image of sketch and the LBPs for the 3D models in the candidate retrieved set are evaluated. And a final retrieved set of 3D mesh models is answered in order of increasing dissimilarity measures of LBP.

By using the GFD and the LBP in the search, both the dissimilarity of the contour shape and the internal edge shape of projected 3D models can be recognized in the retrieval, and effect of differences among the coordinate systems of 3D models on the search results are eliminated. In the following sections, we describe the technical details of each function.



**Fig. 4** Viewpoint settings for the rendering

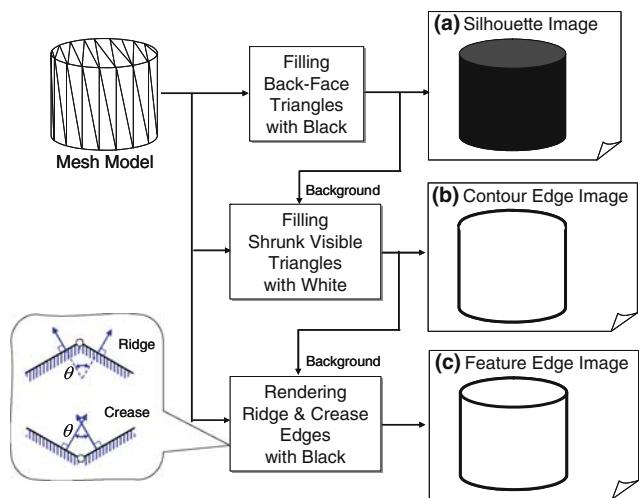
#### 4 Generation of feature edge image and silhouette image

##### 4.1 Viewpoint settings for rendering

Our retrieval method basically evaluates view-based similarity between two images; one is made from projecting a 3D mesh model, and the other from a 2D input query sketch. In the view-based evaluation, relevant selection of the projection type, the number of viewpoint positions and view directions is needed to get a proper retrieval result. If we take the larger number of viewpoints, similarity evaluation becomes more precise, but takes more processing time.

Funkhouser et al. [27] used boundary contour images of a 3D model rendered from 13 orthographic view points equally located at vertices of a tessellated sphere. Chen et al. [28] used ten shaded boundary images from 20 view points located at vertices of a dodecahedron. Macrini et al. [24] used 128 viewpoints for a 3D model, and Cyr and Kimia [25] also generated views sampled at regular ( $5^\circ$ ) intervals on a ground plane. But Pu et al. [30] only uses three orthogonal view points for retrieving engineering mechanical parts.

A user tends to draw a sketch for querying a 3D object from one of a few typical view directions which represents the outline of the object [27]. And in industrial design and mechanical design, the user often draw an rough isometric sketch for graphic explanation. Based on this fact, in our method, 14 typical viewpoints surrounding a 3D model and the 14 view directions pointing to the origin of the model are used for the rendering. Shown in Fig. 4, these viewpoints lie on eight corners and six faces on a bounding box of the model, and each of them corresponds to the isometric or the orthographic projection. This viewpoint selection provides satisfactory retrieval results in case of the industrial design which are shown in Sect. 6.

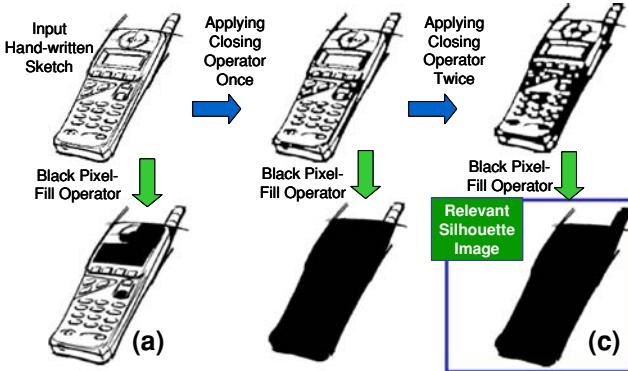


**Fig. 5** Rendering images of silhouette, contour edge and feature edge from 3D mesh model

##### 4.2 Rendering feature edge images and silhouette images for a mesh model

We assumed that a 3D mesh model consists of triangle faces, has face-edge-vertex connectivity structure, and is not necessarily closed (a mesh may have boundaries). The contour edges and feature edges on a mesh model are rendered from each of 14 viewpoints, and are stored as a silhouette image in the silhouette image databases, and a feature edge image in the silhouette image database, respectively. The rendering method we used is a simplified version of Raskers non-photo-realistic rendering method [31]. Shown in Fig. 5, a contour edge is an edge where one of the two neighbouring triangles is visible and the other one invisible from a certain viewpoint and view direction. Our system first generates a silhouette image of the model, and then generates the contour edges using the silhouette image. A silhouette image is a binary (black-and-white) image, and is easily generated by filling invisible back-face triangles with black colour shown in Fig. 5a. Then, the system overwrites slightly shrunk visible triangles with white colour on the silhouette image. As a result, contour edges of the mesh model can be easily rendered shown in Fig. 5b. After generating contour edges, then feature edges are rendered. The feature edges consist of “ridges” and “creases” on the mesh model. In the system, a ridge or a crease is rendered only by drawing a edge of the mesh model where the dihedral angle  $\theta$  of two neighbouring triangles exceeds a specified threshold angle  $Th$ . Then the system overwrites these feature edges on the contour edges to obtain the feature edge image shown in Fig. 5c.

Finally, 14 feature edge images and 14 silhouette images are rendered from the 14 viewpoints shown in Fig. 4 for one mesh model. They are stored in the silhouette image database and the feature edge image database, respectively.



**Fig. 6** Rendering images of silhouette, contour edge and feature edge from 3D mesh model

#### 4.3 Generating a silhouette image from an input hand written sketch

Generally, a handwritten sketch of the query has a similar 2D appearance to the feature edge image, but has a different appearance from the silhouette image. Therefore, our system has to generate a silhouette image from a handwritten sketch for evaluating dissimilarity between a sketch and a registered silhouette image.

However, in the handwritten sketch, the boundary contour edge is not necessarily closed, and pixel filling operation which fills an area inside the contour edge with black often fails as shown in Fig. 6a. To solve the problem, a several number of closing (erosion after dilation) operators [32] are applied automatically to the sketch in order to make outer boundary contour completely closed. After the closing operations, a black-pixel filling operation is applied to the sketch to obtain a relevant silhouette image shown in Fig. 6c.

### 5 Model retrieval based-on dissimilarity between projected images

#### 5.1 Overview of dissimilarity evaluation

In our system, 3D mesh model is retrieved based both on the dissimilarity between a silhouette image of the query sketch and the ones stored in the silhouette image database, and on the dissimilarity between a feature edge image of the query sketch and the ones stored in the feature edge image database. Two image descriptors are used in evaluating these dissimilarities. GFDs are used for comparing two silhouette images, while LBP are don for comparing two feature edge images. Usually, coordinate systems defining geometries of 3D mesh model differ, and the coordinate systems of their projected images also differ. Therefore, image descriptors which are invariant to the 2D similarity transformation (2D rotation,

translation and uniform scaling) are needed for relevantly evaluating the dissimilarities between two projected images.

#### 5.2 Generic Fourier descriptor

GFD is an image descriptor invariant to the 2D similarity transformation, and was originally proposed by Zhang and Lu [33] as an image content descriptor for MPEG-7. GFD is derived by applying 2D discrete Fourier transform on a polar raster sampled image, and is expressed as a multi-dimensional vector. Figure 7 shows a process of deriving the GFD. First, a given silhouette image is trimmed to an image  $I$  described in Eq. 1, whose size equals a minimum bounding box of the silhouette shape.

$$I = \{f(x, y) | 0 \leq x \leq W, 0 \leq y \leq H\} \quad (1)$$

Next, for this trimmed image  $I$ , a polar raster transform is applied to obtain a raster sampled polar image  $I_p$  expressed in Eq. 2.

$$I_p = \{f(\theta, y) | 0 \leq r \leq R, 0 \leq \theta \leq 2\pi\} \\ x = r \cos \theta + W/2, y = r \sin \theta + H/2 \quad (2)$$

where,  $R$  is an integer denoting a maximum radius of the image  $I_p$  from the image centroid ( $W/2, H/2$ ) of  $I_p$ , and is given by Eq. 3.

$$R = \lfloor \sqrt{(W/2)^2 + (H/2)^2} \rfloor \quad (3)$$

Then, a 2D discrete Fourier transform is applied to the  $I_p$  by Eq. 4.

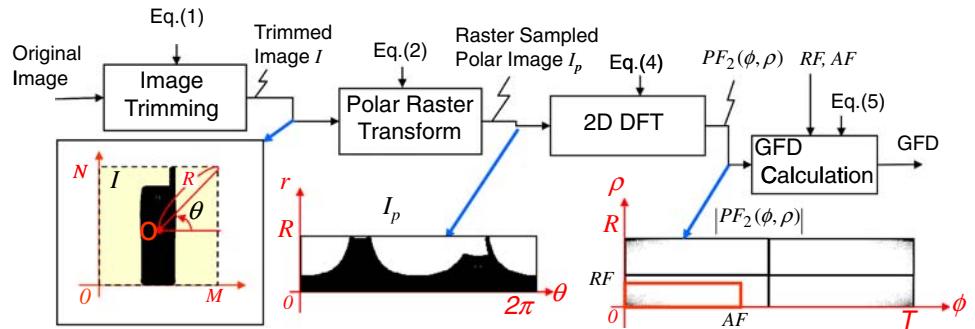
$$PF_2(\phi, \rho) = \sum_r \sum_i^{R-1} f(r, \theta_i) \exp \left[ -j2\pi \left( \frac{r}{R} \rho + \theta_i \phi \right) \right] \quad (4)$$

where  $0 \leq r \leq R$ ,  $\theta_i = 2\pi i / T$ ,  $0 \leq \rho < R$ ,  $0 \leq \phi < T$ . Physical meanings of  $\rho$  and  $\phi$  are the  $\rho$ th radial frequency and the  $\phi$ th angular frequency, respectively. GFD is a multi-dimensional vector obtained from Eq. 5.

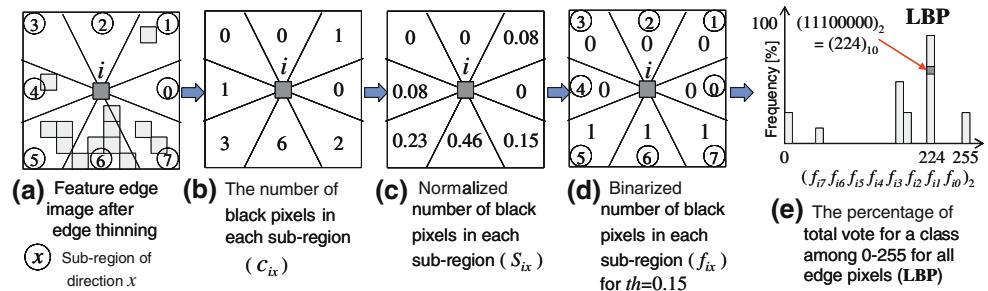
$$GFD = \left\{ \frac{|PF_2(0, 0)|}{RT}, \frac{|PF_2(0, 1)|}{|PF_2(0, 0)|}, \dots, \frac{|PF_2(AF, RF)|}{|PF_2(0, 0)|} \right\} \\ = \{GFD(0, 0), GFD(0, 1), \dots, GFD(AF, RF)\} \quad (5)$$

where,  $RF$  and  $AF$  are selected maximum radial frequency and angular frequency, respectively. By considering the symmetrical nature of power spectra  $|PF_2(\phi, \rho)|$ , the setting  $RF = R/2$  and  $AF = T/2$  is sufficient. However,  $RF$  and  $AF$  are not necessarily equals to  $R/2$  and  $T/2$  because the power of high frequencies in  $PF_2$  is relatively small and a smaller number of spectra are sufficient to capture the features of silhouette images of artificial objects. We experimentally found that  $RF = 3$  and  $AF = 12$  were sufficient for trimmed image size  $W = H = 128$  to capture the features and evaluate the dissimilarity between them.

**Fig. 7** Process of deriving generic Fourier descriptor (GFD)



**Fig. 8** Process of deriving local binary patterns (LBP)



The polar Fourier coefficients  $PF_2(\phi, \rho)$  are invariant for rotation of the original silhouette image  $I_p$ . Translation invariance is achieved by trimming operation on the  $I_p$ . Scaling invariance is achieved by the normalization  $PF_2$  in Eq. 5.

The dissimilarity  $D_{GFD}$  between the silhouette image of a query sketch and the one stored in the silhouette image database is defined as an Euclid distance between two GFDs, and is expressed by Eq. 6.

$$D_{GFD}(GFD_Q, GFD_{Ix})$$

$$= \sqrt{\sum_{k=0}^{RF} \sum_{l=0}^{AF} \{GFD_Q(k, l) - GFD_{Ix}(k, l)\}^2} \quad (6)$$

where  $GFD_Q$  and  $GFD_{Ix}$  are GFDs of the silhouette image of query sketch and one stored in the silhouette image database, respectively. Less  $D_{GFD}$  means that one image is more similar to the other and  $D_{GFD} = 0$  means that two images are exactly identical.

### 5.3 Local binary patterns

LBP is used as an image descriptor for comparing two feature edge images. LBP itself is scale and translational invariant descriptor of edge image. The original version of LBP was proposed as an image descriptor for texture image analysis by Ojala et al. [34]. The LBP indicates statistical distribution of relative positions and orientations between two neighbouring pixels on the edge in binary image, and is independent of position and size of the edge shape.

In our system, we used a modified version of the LBP based on the Ohashi et al. [35]. Figure 8 shows the process of deriving LBP.

First, as a pre-process, the Hilditch edge thinning operation is applied to a feature edge image. Next, for each pixel  $i$  on the edge, a region around the pixel  $i$  is divided radially into eight equal sub-regions corresponding eight directions shown in Fig. 8a. Then the number of black pixels on the edge contained in each sub-region of direction  $x$  ( $x = 1, \dots, 8$ ) is counted as  $c_{ix}$  as shown in Fig. 8b. Then normalization is done where the number of edge pixels  $c_{ix}$  is divided by the total number of edge pixels  $C$  in the edge image to obtain a normalized pixel number  $S_{ix}$  by Eq. 7 shown in Fig. 8c. After the normalization,  $c_{ix}$  becomes scale invariant.

$$S_{ix} = \begin{cases} c_{ix} / C - 1 & (C \neq 1) \\ 0 & (C = 1) \end{cases} \quad (7)$$

$S_{ix}$  is then converted to a binary value by threshold. And by 8, a 8 bit binary number  $(f_{i7}, f_{i6}, f_{i5}, f_{i4}, f_{i3}, f_{i2}, f_{i1}, f_{i0})$  is obtained shown in Fig. 8d for each edge pixel where each bit corresponds to this binarized  $S_{ix}$  number by Eq. 8.

$$f_{ix} = \begin{cases} 0 & (S_{ix} \leq th) \\ 1 & (S_{ix} \geq th) \end{cases} \quad (8)$$

where  $th$  is a threshold of normalized pixel count which determines that an effective number of edge pixels exists in direction  $x$ . We experimentally set  $th = 0.15$ . Finally, for each edge pixel in the feature edge image, the system votes for a particular class whose value is equal to a 8 bit binary number  $(f_{i7}, f_{i6}, f_{i5}, f_{i4}, f_{i3}, f_{i2}, f_{i1}, f_{i0})$ . The result of the vote for all edge pixels in the image is represented as

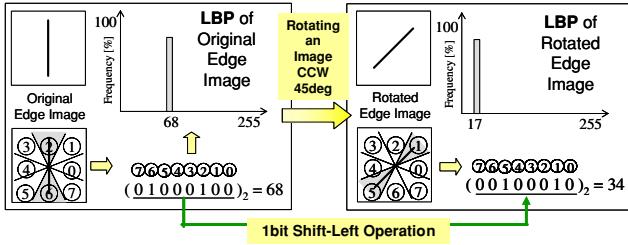


Fig. 9 Change of LBP for 45° CCW rotation of an image

a 256-dimensional vector. Each component of the vector represents the total vote in a particular direction class, and the component is then normalized to the range of 0–100 as shown in Fig. 8e. This vector is called LBP, and is expressed by Eq. 9.

$$\text{LBP} = [\text{LBP}_0, \text{LBP}_1, \dots, \text{LBP}_{255}], \quad \text{LBP}_i \in [0, 100] \quad (9)$$

As the LBP captures the relative locations between two pixels, it is translational invariant. The LBP is not invariant for rotation and mirror image transformation. But components of the LBP after the rotation and mirror image transformation can be easily calculated by applying a relevant bitwise operation to the original LBP. As shown in Fig. 9, for example, if we simply apply the 1 bit shift-left-rotation operation of a LBP, then the LBP for the other image generated by rotating 45° from the original image can be obtained.

The dissimilarity  $D_{\text{LBP}}$  between the feature edge image of the query sketch and the one in the database is defined as a minimum distance between two LBPs. This minimum distance is calculated by applying bitwise operations several times to the original LBP for the feature edge image of the query sketch and by taking minimum Euclid distance between two LBPs as Eq. 10. The bitwise operations are selected suitable for the applied rotations and mirror image transformations for the original image.

$$D_{\text{LBP}}(\text{LBP}_Q, \text{LBP}_{Ix}) = \min_{p \in \{1, 2, \dots, P\}} \sqrt{\sum_{j=0}^{255} \{l_{j'(p,j)}^Q - l_j^{Ix}\}^2} \quad (10)$$

where,  $\text{LBP}_Q$  is the LBP for a feature edge image of query sketch, and  $\text{LBP}_{Ix}$  the one of  $x$ th feature edge image in the database,  $l_{j'(p,j)}^Q$  is  $j'(p, j)$ th component of LBP for a query sketch,  $l_j^{Ix}$  is  $j$ th component of LBP for a image in the database.  $j'(p, j)$  denotes a function which maps a  $j$ th component in the LBP of database to the other component in the LBP of query sketch when applying the rotation and mirror image transformation to the query sketch.  $P$  is the number of applied bitwise operations. For example, if we apply the every 45° rotations to the sketch, then we set  $P = 8$ . While, if we only consider a horizontal mirror image transformation,

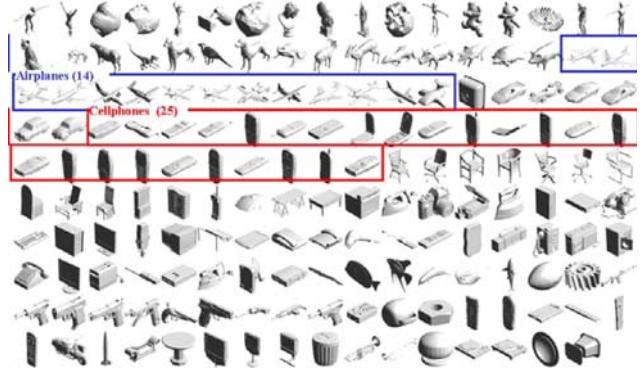


Fig. 10 An experiment set of 168 3D mesh models

then  $P = 2$ . And if we consider the every 45° rotations and horizontal mirror image transformation, then  $P = 16$ .

As a result of taking the dissimilarity of LBP as Eq. 10, the translation, rotation and scale invariant dissimilarity between a feature edge image of query sketch and one in the feature edge image database is evaluated.

#### 5.4 Retrieval by combining GFD with LBP

In the 3D model retrieval step, our system first evaluate the dissimilarity between the shilhouette image from the query sketch and the ones in the shilhouette image database, and retrieve an initial retrieval set consisting of similar 3D mesh models with less dissimilarity of shilhouette images DGFD. Then among the initial retrieval set, the system evaluates the dissimilarity between the feature edge image from the query and the ones in the feature edge image database, and rearrange the similar mesh models in order of less dissimilarities of feature edge images  $D_{\text{LBP}}$ . The final set of similar mesh models is obtained from the result of this rearrangement.

The number of 3D models included in an initial retrieval set was set to 36 in this research, which is suitable for the user identifying all models of the set at a time on a screen of computer.

## 6 Results of retrieval

### 6.1 An experiment set

We made an experimental set of 3D mesh models which contains 168 mesh models shown in Fig. 10. These were acquired from several commercial model datasets [36] and free models on the Web. The mesh model includes the major kinds of mesh model formats such as 3ds, .max, .stl, .ply, and .wrl.

In this set, there is 25 mesh models in a category of “mobile phones” and 14 mesh models in the one of “airplanes”. We

generated the silhouette image database and feature edge image database for this experimental set in advance.

## 6.2 Measure of retrieval performance

We adopted the “Precision–Recall” as a measure of retrieval performance of our system. Precision–Recall [37] has been used as a common measure of retrieval performance in many content-based image retrieval systems. Precision indicates how much the retrieved set does not contain the irrelevant answer, while Recall does how much the retrieved set contains the relevant answer without omission. If we express a set of all retrieved models as  $R$ , and a set of all relevant models as  $A$ , Precision and Recall are defined as Eq. 11.

$$\text{Recall} = |A \cap R| / |A|, \quad \text{Precision} = |A \cap R| / |R| \quad (11)$$

If we plot a graph where the Recall is a horizontal axis, and Precision is a vertical axis for the different number of the retrieved models, Precision–Recall graph can be obtained. If these plots are placed in upper right area in the graph, the system has better retrieval performance.

## 6.3 Preliminary experiments

Before the experiment on the general retrieval performance of our system, we verify the effectiveness of the GFD as a image descriptor for retrieving the silhouette image by comparing it to other two well-known 2D shape descriptors. We selected FD [38] and Complex FD [39] for the comparison. The FD is obtained from the power spectra of 1D discrete Fourier transform for the distance function between an outermost contour curve shape and its center of gravity. While CFD is obtained from the complex discrete Fourier transform for the complex expression of contour curve of the 2D shape. Both descriptors are also translational, rotational and uniform-scaling invariant. For the preliminary experiment, we selected ten silhouette images and ten feature edge images as the queries shown in Fig. 11a, b from the two databases described in 6.1. On half of the images belongs to the “airplane” category and the other half does the “mobile phone”. We retrieved 3D mesh models by using GFD, FD and CFD, respectively, and obtained the Precision–Recall graph. In this case, we defined a set of all relevant models  $A$  as the set which consists of the models belonging to the same category as the one of the query image.

The Precision–Recall graphs are shown in Fig. 12. Figure 12a indicates the graph for retrieved results for four “airplane” query images of Fig. 11a, while Fig. 12b does for “mobile phone” images of Fig. 11b. From these graphs, using GFD as the descriptor for silhouette image retrieval has Precision–Recall performance which is superior to those

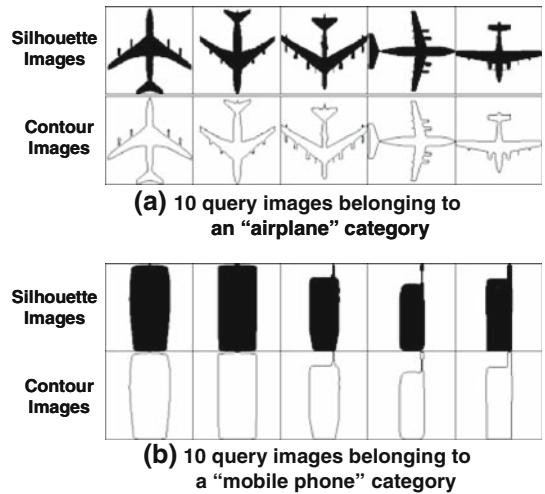


Fig. 11 Query images for preliminary experiments

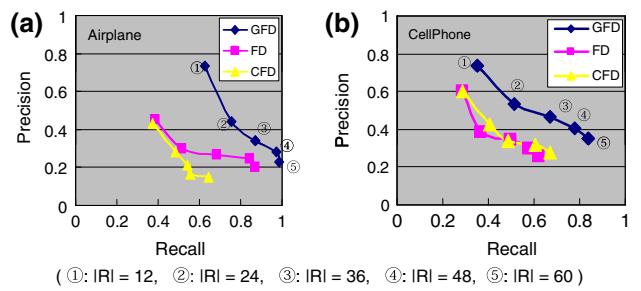


Fig. 12 Precision–Recall graphs for preliminary experiments

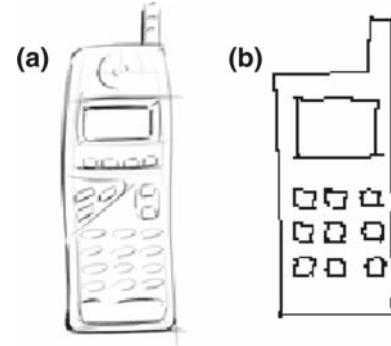


Fig. 13 Hand written sketches for query (a written by industrial designer, b written by a student)

of FD and CFD. From these results, the GFD was selected as the descriptors for evaluating the dissimilarities between silhouette images.

## 6.4 Results of model retrieval from a hand written sketch

Two hand-written sketches of mobile phones were used as queries shown in Fig. 13. Figure 13a is a rough sketch written by an industrial designer, while Fig. 13b is a scribble sketch drawn by a university student.

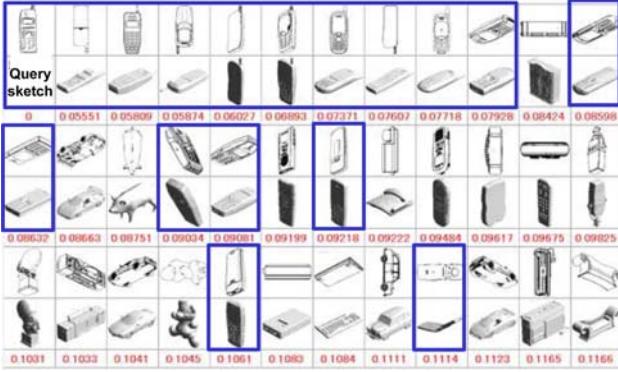


Fig. 14 A result of retrieval for Fig. 13a using GFD

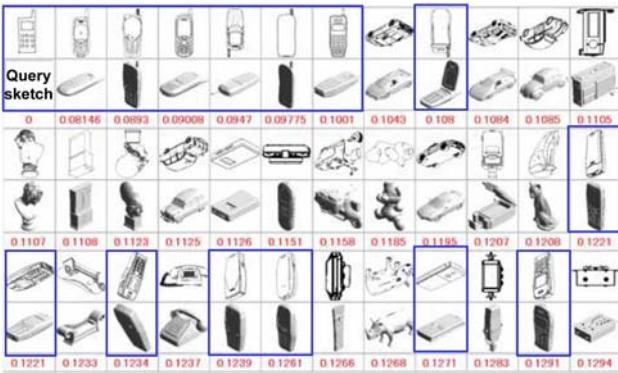


Fig. 15 A result of retrieval for Fig. 13b using GFD

The set of mesh models which was retrieved from Fig. 13a only by using the GFD is shown in Fig. 14, and those from Fig. 13b is in Fig. 15. The set of mesh models retrieved from Fig. 13a both by using the GFD and LBP described in 5.4 is shown in Fig. 13, and those from Fig. 13b is done in Fig. 17. An upper left image is the query sketch, and the figures under the image shows the dissimilarity between two silhouette images. The image bounded by the thick lines is the one which is included in the relevant categories (“mobile phones”) of the model.

For all experiments, the number of models in initial retrieval set is limited to 35. Each retrieval could be finished within 3 s in our system even for using both of GFD and LBP. From Figs. 14, 15, 16, 17, it was shown that translation, rotation, and scaling invariant retrievals were done.

Figure 14 showed that eight similar mobile phone models with antennas could be retrieved within eighth place from the query image of Fig. 13a, and the other type of phones without antennas could be also retrieved within 35th place. However, in this retrieval, only GFD was used as image descriptors and the dissimilarity between two silhouette images was only recognized in the system. The result shows that there are several phone models in the retrieved set in a higher rank which have similar silhouettes to the query but have dissimilar

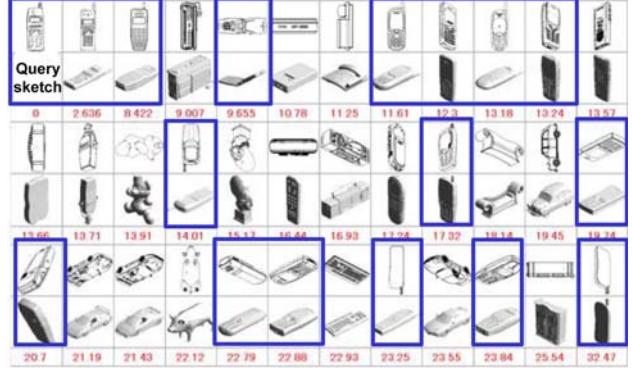


Fig. 16 A result of retrieval for Fig. 13a using GFD and rearrangement using LBP



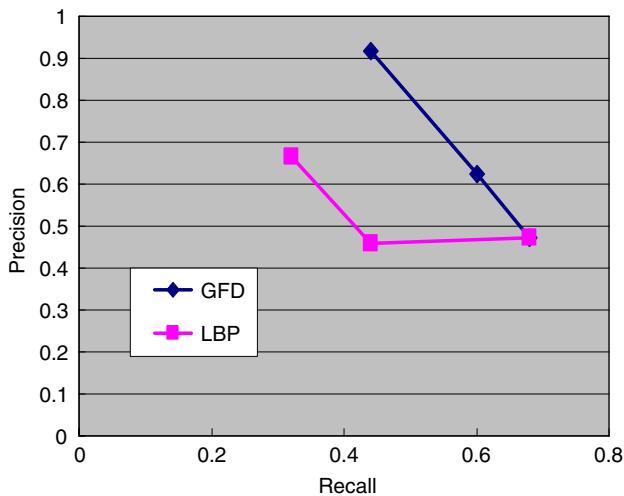
Fig. 17 A result of retrieval for Fig. 13b using GFD and rearrangement using LBP

feature edge shapes on the model (e.g. feature edges of push buttons).

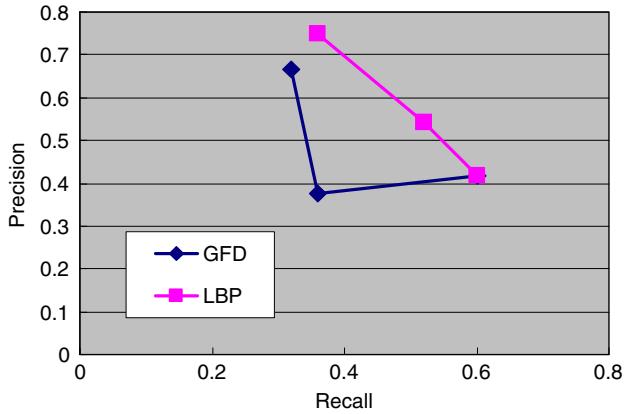
Figure 16 showed the retrieved models for Fig. 13a by using GFD and LBP. The results shows that the phone models which has similar silhouettes and similar feature edges (feature edges of buttons) to the ones of the query image appeared in higher places of the retrieval result than those of Fig. 14.

Figures 15 and 17 show the retrieved models for the scribble sketch of Fig. 13b. Figure 15 is the retrieval result only by using GFD, and Fig. 17 is the one by using GFD and LBP. Similar satisfactory results to the former ones were obtained.

Figures 18 and 19 show the precision-recall graphs corresponding to the Figs. 16 and 17. For the query of Fig. 13a, applying the LBP to the initial retrieval set caused to decrease the precision as shown in Fig. 18, because some models belonging to the different categories but resembling to the query image (keyboard, digital recorders, audio set, etc.) appeared in the higher rank after the rearrangement of ranks using the LBP. However, for the scribble query of Fig. 13b, applying the LBP to the initial retrieval set could improve the precision and recall performance of the retrieved set as shown in Fig. 19.



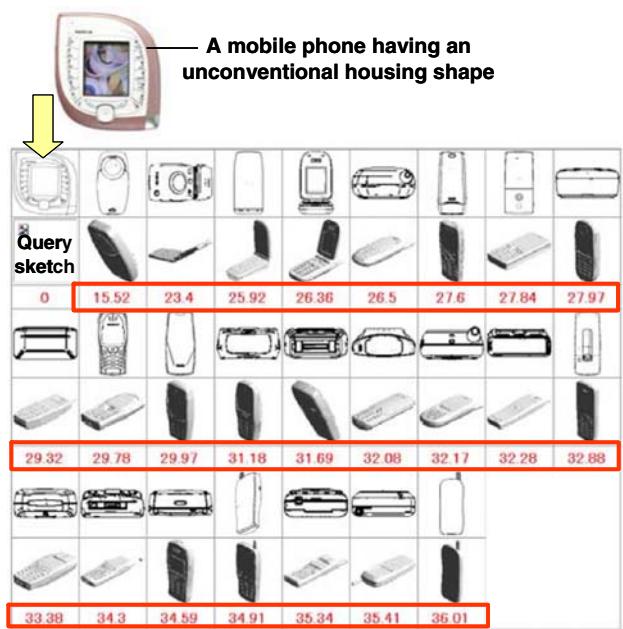
**Fig. 18** A precision–recall graph of a retrieval result of Fig. 16



**Fig. 19** A precision–recall graph of a retrieval result of Fig. 17

Figure 20 shows another retrieval result where a different handwritten sketch was inputted as a query. It was a sketch for a mobile phone having an unconventional exterior shape shown in the upper left of Fig. 20. In this case, we restricted the range of the retrieval to the mobile phone categories. The results shows that the phone models having a intuitively similar shape to the query did not appear completely, and the dissimilarities of the retrieved models increased a few hundreds as large as the ones of the former retrievals. We can interpret this results as that the mobile phone in the query has a strikingly original exteriors compared with those of usual mobile phones.

From these experimental results, an effectiveness of our proposed image based 3D mesh model retrieval was shown where GFD and LBP were used as image descriptors.



**Fig. 20** A result of retrieval for a mobile phone sketch which has an unconventional housing shape

## 7 Conclusion

A content-based 3D mesh model retrieval system from a 2D hand-written sketch query is proposed in this paper. The proposed retrieval algorithm utilizes the content based image retrieval technique for silhouette image and feature edge image. GFDs and LBP were integrated as image descriptors of the retrieval to obtain the 2D rotational, translational and line symmetric invariant retrieval. Automatic rendering of feature edge images and silhouette images for a mesh model was also shown. The mesh model dataset was developed which includes a practical engineering products. The retrieval performances were evaluated as the Precision–Recall graphs. From several experiments on the sketch-based retrieval, effectiveness and a efficiency of our retrieval system were shown in the field of industrial design.

The future works of this research includes the improvement of the image descriptors for feature edge image and implementation of the user feedback function in the system to obtain more subjectively relevant retrieval results.

**Acknowledgments** This work was financially supported by the grant-in-aid of Intelligent Cluster Project (“Sapporo IT Carrozzeria”) founded by Japanese MEXT.

## References

1. Várady, T., Martin, R.R., Cox, J.: Special issue: reverse engineering of geometric models. *Comput. Aided Des.* **29**(4), 253–254 (1997)

2. Bennis, F., Chedmail, P., Hélaré, O.: Representation of design activities using neural networks: application to fuzzy dimensioning. Montréal-Québec, IDMME'2000 (2000)
3. Manjunath, B.S., Salembier, P., Sikora, T.: Introduction to MPEG-7: Multimedia Content Description Interface. Wiley, Hoboken (2002)
4. Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(12), 1349–1380 (2000)
5. Hanjalic, A., Langelaar, G.C., van Roosmalen, P.M.B., Biemond, J., Lagendijk, R.L.: Image and video databases: restoration, watermarking and retrieval. *Adv. Image Commun.* **8**, 313–434 (2000)
6. Tangelder, J.W.H., Veltkamp, R.C.: A survey of content based 3d shape retrieval methods. In: Shape Modeling International, pp. 145–156 (2004)
7. Corney, J., Rea, H., Clark, D.: Coarse filters for shape matching. *IEEE Comput. Graph. Appl.* **22**(3), 65–74 (2002)
8. Iyer, N., Jayanti, S., Lou, K., Kalyanaraman, Y., Ramani, K.: Three dimensional shape searching: state-of-the-art review and future trends. *Comput. Aided Des.* **37**(5), 509–530 (2005)
9. Kazhdan, M., Chazelle, B., Dobkin, D., Funkhouser, T., Rusinkiewicz, T.: A reflective symmetry descriptor for 3d models. *Algorithmica* **38**(1), 201–225 (2004)
10. Zhang, C., Chen, T.: Indexing and retrieval of 3d models aided by active learning. *ACM Multimedia*, pp. 615–616 (2001)
11. Osada, R., Funkhouser, T., Chazelle, B., Dobkin, D.: Shape distributions. *ACM Trans. Graph.* **21**(4), 807–832 (2002)
12. Ip, C.Y., Lapadat, D., Sieger, L., Regli, W.C.: Using shape distributions to compare solid models. In: Proc. ACM Solid Modeling '02, pp. 273–280 (2002)
13. Ohbuchi, R., Otagiri, T., Ibato, M., Takei, T. (2002) Shape-similarity search of three-dimensional models using parameterized statistics. In: Proc. Pacific Graphics 2002, pp. 265–274
14. Shum, H.-Y., Hebert, M., Ikeuchi, K.: On 3d shape similarity. In: Proc. IEEE Computer Vision and Pattern Recognition, pp. 526–531 (1996)
15. Johnson, A.E., Hebert, M.: Using spin images for efficient object recognition in cluttered 3d scenes. *PAMI* **21**(5), 635–651 (1999)
16. Chua, S.J., Jarvis, R.: Point signatures: a new representation for 3d object recognition. *Int. J. Comput. Vis.* **25**, 63–5 (1997)
17. Cardone, A., Gupta, S.K.: Identifying similar parts for assisting cost estimation of prismatic machined parts. In: Proceedings of 2004 ASME/DETC, number 57761, ASME/DETC (2004)
18. El-Mehalawi, M., Miller, R.A.: A database system of mechanical components based on geometric and topological similarity. Part i: representation. *J. Comput. Aided Des.* **35**(1), 83–94 (2003)
19. El-Mehalawi, M., Miller, R.A.: A database system of mechanical components based on geometric and topological similarity. part ii: indexing, retrieval, matching and similarity assessment. *J. Comput. Aided Des.* **35**(1), 95–105 (2003)
20. McWherter, D., Peabody, M., Shokoufandeh, A., Regli, W.C.: Database techniques for archival of solid models. *Solid Modeling '01*, pp. 78–87 (2001)
21. Sundar, H., Silver, D., Gagvani, N., Dickenson, S.: Skeleton based shape matching and retrieval. In: Proc. of Shape Modeling International 2003, pp. 130–139 (2003)
22. Iyer, N., Lou, K., Jayanti, S., Kalyanaraman, Y., Ramani, K.: Shape based searching for product lifecycle applications. *Comput. Aided Des.* **37**, 1435–1446 (2005)
23. Hilaga, M., Shinagawa, Y., Kohmura, T., Kunii, T.L.: Topology matching for fully automatic similarity estimation of 3d shapes. In: ACM SIGGRAPH 2001, pp. 12–17, August 2001
24. Macrini, D., Shokoufandeh, A., Dickenson, S., Siddiqi, K., Zucker, S.: Viewbased 3-d object recognition using shock graphs. *ICPR* 2002 (2002)
25. Cyr, C.M., Kimia, B.: 3d object recognition using shape similarity-based aspect graph. *ICCV01*, pp. I:254–I:261 (2001)
26. Ohbuchi, R., Nakazawa, M., Takei, T.: Retrieving 3d shapes based on their appearance. In: Proc. 5th ACM SIGMM Workshop on Multimedia Information Retrieval (MIR 003), (MIR 003) (2003)
27. Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., Dobkin, D.: A search engine for 3D models. *ACM Trans. Graph.* **22**(1), 83–105 (2003)
28. Chen, D.-Y., Tian, X.-P., Shen, Y.-T., Ouhyoung, M.: On visual similarity based 3d model retrieval. In: Computer Graphics Forum (EG 2003 Proceedings), number 22(3) (2003)
29. Hou, S., Ramani, K.: Dynamic query interface for 3d shape search. In: Proceedings of 2004 ASME/DETC, number 57687, DETC2004, (2004)
30. Pu, J.T., Lou, K., Ramani, K.: A 2d sketch-based user interface for 3d cad model retrieval. *Comput. Aided Des. Appl.* **2**(6), 717–725 (2005)
31. Raskar, R.: Hardware support for non-photorealistic rendering. In: Proc. of SIGGRAPH/ Eurographics Workshop on Graphics Hardware (HWWS), pp. 41–47, August 2001
32. Pierre, S.: Morphological Image Analysis. Springer, Berlin (2004)
33. Zhang, D., Lu, G.: Shape-based image retrieval using generic fourier descriptor. *Signal Process. Image Commun.* **17**, 825–848 (2002)
34. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recognit.* **29**(1), 51–59 (1996)
35. Ohashi, G., Nagashima, Y., Mochizuki, K., Shimodaira, Y.: Edge-based image retrieval using a rough sketch. *J. Inst. Image Inf. Televis. Eng. (in Japanese)* **56**(4), 653–658 (2002)
36. De espóna 3d enciclopedia. <http://www.deespona.com>
37. Müller, H., Müller, W., Squire, D.M., Marchand, S., Pun, T.: Performance evaluation in content-based image retrieval: overview and proposals. *Pattern Recognit. Lett.* **22**, 593–601 (2001)
38. Zhang, D., Lu, G.: A comparative study of curvature scale space and fourier descriptors for shape-based image retrieval. *J. Vis. Commun. Image Present.* **14**, 41–60 (2003)
39. Glunlund, G.H.: Fourier preprocessing for hand paint character recognition. *IEEE Trans. Computer* **C21**(2), 195–201 (1972)