Functional Recognition of 3D Shapes Based on Part Contextual Relationships: Postprint

Authors: Zi Lingling, Cong Xin, Yang Pei

Date: 2018-05-20T00:00:00+00:00

Abstract

Fully leveraging the semantic information inherent in shapes for high-level analysis and understanding of three-dimensional shapes is currently a popular research topic. We propose a method that employs contextual semantic relationships among shape parts for functional recognition, which addresses the problem of automatic shape part recognition when significant variations occur in the geometric features and topological structures of 3D shapes. First, approximate convex decomposition technique is adopted to segment three-dimensional shapes into shape parts with independent semantics. Then, we propose a computational method based on contextual semantics of shape parts and utilize Support Vector Machine to achieve automatic recognition of shape parts. Experimental results demonstrate that, compared with existing methods, our approach achieves higher part matching accuracy and lower classification error rate.

Full Text

Preamble

Title: 3D Shape Function Recognition Based on Contextual Relationship of Shape Parts

Authors: Zi Lingling[1], Cong Xin[1], Yang Pei[2]
[1] School of Electronic & Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China
[2] College of Science, Liaoning Technical University, Fuxin, Liaoning 123000, China

Abstract: Leveraging semantic information embedded in shapes for high-level 3D shape analysis and understanding is currently a hot research topic. This paper proposes a function recognition method that utilizes the contextual semantic relationships of shape parts to address the automatic recognition of shape components when significant geometric and topological variations occur in 3D shapes. First, we employ approximate convexity decomposition to segment 3D shapes into parts with independent semantics. Then, we propose a computational method for shape part context semantics and implement automatic part recognition using Support Vector Machines. Experimental results demonstrate that compared to existing methods, our approach achieves higher part matching accuracy and lower classification error rates.

Keywords: contextual relationship; function recognition; 3D shapes; SVM

0 Introduction

Function knowledge, which describes the purpose of shapes, plays a pivotal role in the high-level analysis and understanding of 3D shapes [1]. Consequently, 3D shape function recognition has become a pressing research focus [2], with part-based function recognition representing one of the primary solution approaches [3]. Previous methods have achieved promising results by exploiting semantic relationships between shape parts. For instance, reference [4] proposed a part-aware similarity metric, while reference [5] introduced a connection-aware shape descriptor to express functional semantics by analyzing local shape radius variations. Reference [6] developed a high-level signature method to capture functional information of parts. However, as the volume of shareable 3D shapes grows exponentially, relying solely on semantic information compromises function recognition accuracy. To address this limitation, integrating structural information of parts becomes necessary for functional semantic extraction. Reference [7] introduced a novel pattern function transformation framework that generates shape matching pairs by comparing pattern functions and obtains compact function spaces based on each shape' s embedding basis to perform knowledge reasoning. Reference [8] constructed a component model that effectively handles shape models with similar appearances by describing human cognition regarding structure, function, and part interrelationships. Reference [9] proposed a recognition method based on sparse representation classification, while reference [10] introduced a visual element classification method; both approaches learn semantics from 3D shape models but require large-scale training datasets. Reference [11] developed a rule-based expert system to infer functional semantics from acquired knowledge. Although these methods demonstrate good recognition performance, their accuracy declines substantially when shape collections contain individuals with significant geometric and topological variations, necessitating new approaches to improve functional semantic accuracy under large-scale deformations.

This paper explores the structural characteristics and semantic relationships of shape parts in depth and proposes a Contextual-Relationship-based shape Function Recognition method (CIFR) to enhance recognition performance under large-scale deformations. Our method addresses two core problems: (a) how to decompose 3D shapes into component sets with independent semantics, and

(b) how to quantitatively compute contextual semantic relationships between parts. For the first problem, we propose a shape segmentation method based on approximate convexity decomposition, which effectively identifies components with independent semantics according to contour dissimilarity and boundary attributes of convex blocks. Approximate convexity decomposition technology provides a viable solution, and we employ it to achieve 3D shape part segmentation. For the second problem, leveraging the stability of contextual semantic relationships under large-scale deformations, we propose a semantic computation method based on part context relationships to mine potential semantic information across different shape parts, thereby improving functional recognition accuracy through contextual information.

1 CIFR Method Overview

Let $\{S_1, S_2, \dots, S_n\} \in \mathcal{S}$ denote a set of 3D shapes. Each shape $S_i \in \mathcal{S}$ is decomposed into independent part sets $\{p_1, p_2, \dots, p_m\} \in \mathcal{P}$, with $\{c_1, c_2, \dots, c_m\} \in \mathcal{C}$ representing functional category labels. The function recognition task retrieves parts from test shape S_t whose part p_t shares the same functional category. We decompose this problem into two sub-problems: (1) given a 3D shape, and (2) for any two shape parts, how to quantitatively compute semantic similarity between p_i and p_j and apply it to an SVM classifier for functional category recognition. Based on this, we propose the CIFR method, whose framework is illustrated in [Figure 1: see original paper]. The CIFR method consists of three main stages: shape segmentation, which divides 3D shapes into parts with independent semantics; semantic computation for the obtained shape parts; and finally, SVM-based functional recognition of shape parts.

2.1 Shape Segmentation

Drawing upon human visual perception characteristics for complex shapes [12, 13], we decompose shapes based on minimum principles. The process involves three key steps. First, we over-segment S_i to obtain an approximate convex block set and construct a shape convex graph from the extracted blocks, using graph-based spectral clustering to acquire the initial block collection $\mathcal{B} = \{B_1, B_2, \dots, B_n\}$.

Second, we compute the convex rank $R(B_i)$ for each initial block $B_i \in \mathcal{B}$ using:

$$R(B_i) = \frac{|VIS(B_i)|}{|B_i|}$$

where $VIS(B_i)$ represents the total number of mutually visible point pairs in B_i, i.e., $(i, j) \in VIS(B_i)$ with $i \in B_i$ and $j \in B_i$, and points i and j are mutually visible. Sorting $R(B_i)$ in descending order yields the weakly convex component set $\mathcal{D} = \{D_1, D_2, \dots, D_m\}$. Components with convex rank greater than threshold α are merged with their adjacent component D_{i+1}, where $\alpha = 0.7$. To integrate weakly convex components with similar visibility attributes, we repeat this merging operation three times to obtain a new set $\mathcal{D} = \{D_1, D_2, \dots, D_{m'}\}$.

Finally, we characterize geometric features of components using contour dissimilarity and boundary attributes to merge weakly convex components with similar geometric characteristics. Contour dissimilarity is computed using the Earth Mover' s Distance between weakly convex components:

$$dist(D_i, D_j) = EMD(h_i, h_j)$$

where h_i and h_j are histograms of the shape diameter function for D_i and D_j, respectively.

Boundary attributes are characterized by the convexity of seam lines:

$$VS_{i,j} = \{(\omega_p, \omega_q) : \omega_p \in D_i, \omega_q \in D_j, \theta(n_p, n_q) \leq \zeta\}$$
$$NS_{i,j} = \{(\omega_p, \omega_q) : \omega_p \in D_i, \omega_q \in D_j, \theta(n_p, n_q) > \pi - \zeta\}$$

where $VS_{i,j}$ and $NS_{i,j}$ represent convex and concave seam lines between D_i and D_j, respectively, ω_p is a sampling point in D_i with direction $-n_p$, $\theta(n_p, n_q)$ is the angle between n_p and n_q, ζ is the normal adjustment threshold, and C_k is the neighbor graph formed by seam lines.

Weakly convex components D_i and D_j are merged if they satisfy:

$$\frac{||NS_{i,j}||}{||VS_{i,j}||} \geq \eta \quad \text{and} \quad dist(D_i, D_j) \leq \sigma$$

where η and σ are merging thresholds, yielding the independent part set $\mathcal{P}$.

2.2 Semantic Computation and Function Recognition

When significant geometric and topological changes occur in 3D shapes, the contextual semantic relationships between shape parts remain relatively stable. We leverage this property for functional classification. Let graph $G(V, E)$ represent the 3D shape part relationship, where nodes V correspond to shape parts and edges E represent different types of relationships between nodes. Problem 2 transforms into: given relation graphs G_p and G_q for shapes S_p and S_q, how to quantitatively compute a graph kernel Ψ that measures functional semantic similarity between shape parts, such that Ψ can compare subgraphs $G'_p \in G_p$ and $G'_q \in G_q$.

Based on geometric features of 3D shapes and node context relationships, we compute the graph kernel using:

$$\Psi(G_p, G_q) = \sum_{l=0}^{\infty} \sum_{p_A \in G_p, p_B \in G_q} \psi_l(p_A, p_B)$$

where the recursive function ψ_l is defined as:

$$\psi_l(p_A,p_B)=\begin{cases}\Psi_N(p_A,p_B)\times\sum_{\substack{p_S\in N_{G_p}(p_A)\\ p'_S\in N_{G_q}(p_B)}}\Psi_E(e_{p_Ap_S},e_{p_Bp'_S})\cdot\psi_{l-1}(p_S,p'_S) & \text{if } l>0\\ \Psi_N(p_A,p_B) & \text{if } l=0\end{cases}$$

Here, l is the path length, $p_A \in G_p$, $p_B \in G_q$, $N_G(x)$ denotes all neighboring nodes of x in graph G, and e and e' are edges connecting p_A to p_S and p_B to p'_S, respectively. Ψ_N is the node kernel reflecting geometric feature similarity between two parts, characterized by attributes including shape descriptors based on Euler distance, shape dimensions based on unit spheres, and principal direction features based on PCA. Ψ_E is the edge kernel representing the similarity matrix of different node context relationships, where for any edges $e_p \in E_p$ and $e_q \in E_q$, $\Psi_E(e_p,e_q)=\delta(e_p,e_q)$ with δ being the Kronecker delta function (value 1 if e_p and e_q share a context relationship, 0 otherwise).

The contextual semantic relationships between parts are summarized in , comprising basic and derived relationships. Basic relationships include containment, symmetry, horizontal support, side contact, and adjacency. Derived relationships consist of three semantic inference rules designed to capture potential contextual semantic information based on basic relationships. In the table, p_i represents any part in G with $i \in \{1,2,3\}$.

Table 1: Node Contextual Relationships

Relationship Type	Condition	Expression
Containment	If more than 50% of p_1' s bounding box lies within p_2' s bounding box	$EN(p_1,p_2)=1$
Symmetry	If p_1 and p_2 are rotationally, translationally, or reflectively symmetric	$SY(p_1,p_2)=1$
Horizontal Support	If p_1' s contact surface is perpendicular to p_2' s gravity vector	$HS(p_1,p_2)=1$
Side Contact	If the normal of p_1' s contact region is perpendicular to p_2' s principal symmetry axis	$SC(p_1,p_2)=1$

Relationship Type	Condition	Expression
Adjacency	If p_1 and p_2 share common vertices or their bounding boxes overlap	$AD(p_1, p_2) = 1$
Derived Relationships		
Symmetry Transitivity	If $SY(p_1, p_2) = 1$ and $SY(p_2, p_3) = 1$	Then $SY(p_1, p_3) = 1$
Support Transitivity	If $HS(p_1, p_2) = 1$ and $SY(p_2, p_3) = 1$	Then $HS(p_1, p_3) = 1$
Contact Transitivity	If $SC(p_1, p_2) = 1$ and $SY(p_2, p_3) = 1$	Then $SC(p_1, p_3) = 1$

Given SVM' s superior performance in controlling overfitting, computational efficiency, and classification accuracy compared to many traditional methods [14], we employ SVM to build a functional semantic classifier. During training, we select Ψ as the kernel function to compute support vectors and optimal parameters. In the testing phase, we first segment the test model to obtain semantic parts, then identify functional categories of semantic parts using the optimal parameter set.

3 Experiments and Analysis

The experimental dataset comprises the Princeton Shape Benchmark [15] and the COSEG dataset [16], containing six categories: airplane, ant, chair, goblet, candelabrum, and bird. Sample shape models and functional categories are shown in . Experiments were conducted on a 3GHz CPU with 3GB RAM using Matlab R2010b. For each shape category, half of the dataset was randomly selected as training set and the other half as test set. The normal adjustment parameter ζ was set to 0.05π, and the weakly convex component merging thresholds η and σ were set to 0.8 and 0.15, respectively.

Table 2: Experimental Data

Category	Functional Parts
Airplane	fuselage, wing, tail
Ant	antenna, head, body, leg
Chair	chair leg, chair back, seat
Goblet	cup body, stem, base
Candelabrum	container, handle, flame, candle
Bird	body, tail, wing

Since functional recognition relies on high-level semantics of shape categories, we validate functional recognition accuracy using classification error rate:

$$Error_i = \frac{N_e}{N_e + N_c} \times 100\%$$

where N_e is the number of incorrectly classified functional semantics and N_c is the number of correctly classified ones. Ground-truth functional categories were manually annotated for each shape. Lower error rates indicate higher functional recognition accuracy.

[Figure 2: see original paper] compares the functional semantic error rates between our method and the semantic correspondence method GCFR [4]. The results demonstrate that CIFR achieves lower classification error rates for every shape category, primarily due to our proposed semantic computation method based on part contextual relationships that enhances functional semantic classification accuracy.

To further validate recognition performance, we measure part matching accuracy. For any shape category, we construct a manually annotated part correspondence matrix $M = [m_{ij}]$ $(i, j = 1 \ldots N)$, where N is the total number of parts in that category. $m_{ij} = 1$ if parts i and j share the same functional category, and 0 otherwise. The average part matching accuracy for a shape category is defined as:

$$\phi(i) = \frac{1}{N_p} \sum_{p \in P} \frac{\sum_{q \in Q} m_{pq} \cdot AC(p, q)}{\max(AC(p, q))}$$

where p and q represent two shapes from any category. For part i in shape p, $\phi(i)$ identifies the part in shape q with the same functional category as i, and N_p is the total number of parts in P. The value is 1 if all parts in p match functional categories in q, and 0 if no parts are correctly classified. [Figure 3: see original paper] shows the average part matching accuracy compared with GCFR [4], demonstrating that CIFR achieves higher average matching accuracy across different shape models.

4 Conclusion

The diversity of geometric features and topological structures across different 3D shapes makes recognizing functionally similar parts with different appearances a challenging problem. Semantic relationships between shape parts contain latent structural information that provides valuable observations for statistical methods. Therefore, we formulate recognition as a prediction problem and propose a shape function recognition method based on part contextual relationships. The key innovations are: (1) a shape segmentation method based on approximate convexity decomposition that effectively decomposes 3D shapes into multiple part sets with different semantics; (2) a semantic computation method based on part contextual relationships that constructs an SVM classifier using functional semantic similarity as the kernel function to effectively predict functional semantics of different shape parts. Experimental results demonstrate that our method achieves excellent recognition performance and high accuracy.

References

[1] Fan G, Zhang X, Ding M. Gaussian process for human motion modeling: a comparative study [C]// Proc of IEEE International Workshop on Machine Learning for Signal Processing. [S.l.]: IEEE Signal Processing Society, 2011: 1-6.

[2] Yu T, Wang R. Scene parsing using graph matching on street-view data [J]. Computer Vision and Image Understanding, 2016, 145: 70-80.

[3] Kim V G, Chaudhuri S, Guibas L, et al. Shape2pose: human-centric shape analysis [J]. ACM Trans on Graphics, 2014, 33(4): Article 120.

[4] Laga H, Mortara M, Spagnuolo M. Geometry and context for semantic correspondences and functionality recognition in man-made 3D shapes [J]. ACM Trans on Graphics, 2013, 32(5): Article 150.

[5] Feng J, Liu Y, Gong L. Junction-aware shape descriptor for 3D articulated models using local shape-radius variation [J]. Signal Processing, 2015, 112: 14-23.

[6] Léon V, Bonneel N, Lavoué G, et al. Continuous semantic description of 3D meshes [J]. Computers & Graphics, 2016, 54: 47-56.

[7] Kuang Z, Li Z, Qian L, et al. Modal function transformation for isometric 3D shape representation [J]. Computers & Graphics, 2015, 46: 209-220.

[8] Wang J, Yuille A. Semantic part segmentation using compositional model combining shape and appearance [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2015: 1788-1797.

[9] Shu Z, Wang P, Yu X, et al. 3D model recognition algorithm based on local sparse representation [J]. Journal of Computer-Aided Design & Computer Graphics, 2014, 26(11): 1938-1947.

[10] Aubry M, Maturana D, Efros A A, et al. Seeing 3D chairs: exemplar part-based 2D-3D alignment using a large dataset of CAD models [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2014: 3762-3769.

[11] Daniela X, Berta C, Rubén F. A rule-based expert system for inferring functional annotation [J]. Applied Soft Computing, 2015, 35: 373-385.

[12] Asafi S, Goren A, Cohen-Or D. Weak convex decomposition by lines-of-sight [J]. Computer Graphics Forum, 2013, 32(5): 23-31.

[13] Kaick O V, Fish N, Kleiman Y, et al. Shape segmentation by approximate convexity analysis [J]. ACM Trans on Graphics, 2014, 34(1): 1-11.

[14] Li P, Dong L, Xiao H, et al. A cloud image detection method based on SVM vector machine [J]. Neurocomputing, 2015, 169: 34-42.

[15] Chen X, Golovinskiy A, Funkhouser T. A benchmark for 3D mesh segmentation [J]. ACM Trans on Graphics, 2009, 28(3): 341-352.

[16] COS. 2012. The shape coseg dataset. http://web.siat.ac.cn/yunhai/ssl/ssd.htm [EB/OL].

Note: Figure translations are in progress. See original paper for figures.