Image-based Airborne LiDAR Point Cloud Encoding for 3D Building Model Retrieval

Yi-Chen Chen [a, *], Chao-Hung Lin [b]

[a] Dept. of Geomatics, National Cheng Kung University, Taiwan (R.O.C.) - powcyc@gmail.com
[b] Dept. of Geomatics, National Cheng Kung University, Taiwan (R.O.C.) - linhung@mail.ncku.edu.tw

Theme Sessions, ThS2

KEY WORDS: Point Cloud Encoding, Spatial Histogram, 3D Model Retrieval, Cyber City Modeling

ABSTRACT:

With the development of Web 2.0 and cyber city modeling, an increasing number of 3D models have been available on web-based model-sharing platforms with many applications such as navigation, urban planning, and virtual reality. Based on the concept of data reuse, a 3D model retrieval system is proposed to retrieve building models similar to a user-specified query. The basic idea behind this system is to reuse these existing 3D building models instead of reconstruction from point clouds. To efficiently retrieve models, the models in databases are compactly encoded by using a shape descriptor generally. However, most of the geometric descriptors in related works are applied to polygonal models. In this study, the input query of the model retrieval system is a point cloud acquired by Light Detection and Ranging (LiDAR) systems because of the efficient scene scanning and spatial information collection. Using Point clouds with sparse, noisy, and incomplete sampling as input queries is more difficult than that by using 3D models. Because that the building roof is more informative than other parts in the airborne LiDAR point cloud, an image-based approach is proposed to encode both point clouds from input queries and 3D models in databases. The main goal of data encoding is that the models in the database and input point clouds can be consistently encoded. Firstly, top-view depth images of buildings are generated to represent the geometry surface of a building roof. Secondly, geometric features are extracted from depth images based on height, edge and plane of building. Finally, descriptors can be extracted by spatial histograms and used in 3D model retrieval system. For data retrieval, the models are retrieved by matching the encoding coefficients of point clouds and building models. In experiments, a database including about 900,000 3D models collected from the Internet is used for evaluation of data retrieval. The results of the proposed method show a clear superiority over related methods.

1. INTRODUCTION

Recent development in modeling and scanning techniques has led to an increasing number of 3D models. Many of these 3D models is free in the Internet to access. In this context, the question of ''How to generate 3D building models?'' has evolved to ''How to find them in model database and WWW?'' (Funkhouser et al., 2003). This study aims at the efficient construction of a cyber city by encoding unorganized, noisy, and incomplete building point clouds acquired by airborne LiDAR, as well as by retrieving 3D building models from model databases or from the Internet. In this scheme, a complete or semi-complete building model in databases or in the Internet is reused rather than reconstructing point cloud.

The main theme of model retrieval is accurate and efficient representation of a 3D shape. Most previous studies focus on encoding and retrieving 3D polygon models using polygon models as input queries (Funkhouser et al., 2003; Assfalg et al., 2007; Gao et al., 2011; Akgul et al., 2009; Gao et al., 2012). However, these studies do not consider model retrieval by using point clouds, which is in a great need in the topic of efficient cyber city construction with airborne LiDAR point clouds. The key idea behind the proposed method is to represent point clouds of building roofs by using top-view depth image. Building roof is the most informative part of a building in airborne LiDAR point clouds, as shown in Figure 1. In addition, the use of building roof in data encoding can avoid the difficulty coming from the insufficient sampling on the side-view of buildings. A point cloud is encoded by geometric features of its depth image,

which has the properties of rotation-invariance and noise-insensitivity. In addition, the proposed depth image encoding method reduces data description dimensions and yields a compact shape descriptor, resulting in both storage size and search time reduction.

Following the categorization in (Akgul et al., 2009), 3D model retrieval methods are classified into two categories, model-based retrieval and view-based retrieval. For model-based retrieval, shape similarities are measured by using various geometric shape descriptors including shape distribution (Assfalg et al., 2007; Akgul et al., 2009), spherical harmonic function (Chen et al., 2014; Mademlis et al., 2009), shape topology (Tam and Lau, 2007), shape spectral (Jain and hang, 2007), and radon transform (Daras et al., 2006). In the topology-based method (Tam and Lau, 2007), model topologies are represented as skeletons/graphs. The methods rely on the fact that the skeleton is a compact shape descriptor, and assume that similar shapes have similar skeletons. These conditions enable a topology-based method to facilitate efficient shape matching. In the methods of shape distribution (Assfalg et al., 2007; Akgul et al., 2009), geometric features are accumulated in bins that are defined over feature spaces. A histogram of these values is then used as the signature of a 3D model. In the transformation-based methods (Chen et al., 2014; Jain and hang, 2007; Daras et al., 2006), 3D shapes are transformed to other domains, and transformation coefficients are used in shape matching and retrieval. Among them, transformation using spherical harmonic functions is the typical transformation. By utilizing the advantages of compact shape

---

* Corresponding author

description and rotation invariant, models can be efficiently retrieved.

For view-based retrieval, 3D shapes are represented as a set of 2D projections and 3D models are matched using their visual similarities rather than the geometric similarities (Gao et al., 2011; Gao et al., 2012; Chen et al., 2003; Stavropoulos et al., 2010; Papadakisa et al., 2007). Each projection is described by image descriptors. Thus, shape matching is reduced to measure similarities between views of the query object and those of the models in the database. Although methods based on projected views can yield good retrieval results, a large number of views may degrade retrieval efficiency.

Figure 1. Airbore LiDAR point cloud. The perspective view (left), top view (middle), and side view (right) of point cloud.

The related model-based and view-based methods perform very well on existing benchmarks for encoding and retrieving 3D polygon models. However, these methods cannot be applied to unorganized, noisy, sparse, and incomplete 3D point clouds. In this study, a combination of shape distribution and visual similarity is proposed to encode LiDAR point clouds with sparse, noisy, and incomplete sampling. The top-view depth image is used only for visual similarity because of the incomplete sampling on side-view of buildings. Furthermore, the distribution of geometric shape in the depth image is represented as encoded features for the matching of the input point cloud and building models in the database. These strategies enable the proposed method to consistently and accurately encode both the point cloud and polygon models, making model retrieval from airborne LiDAR point cloud and data reuse feasible. The remainder of this paper is organized as follows. Section 2 describes the methodology of point cloud encoding and building model retrieval. Section 3 discusses the experimental results, and section 4 presents the conclusions.

2. METHODOLOGY

2.1 System Overview

The system overview is as shown in Figure 2. The proposed building model encoding consists of two main components, *data encoding* and *data retrieval*. For data encoding, both the building models and point clouds, that is, the input query, are consistently encoded by a set of geometric features of depth images. To achieve consistency in encoding, an interpolation is applied to the depth image of a point cloud. This process enables building models and point clouds to have similar samplings, which facilities consistent encoding. For data retrieval, a LiDAR point cloud is selected as a query input to retrieve building models in the database. The data are retrieved by matching the encoded coefficients of point clouds and building models.
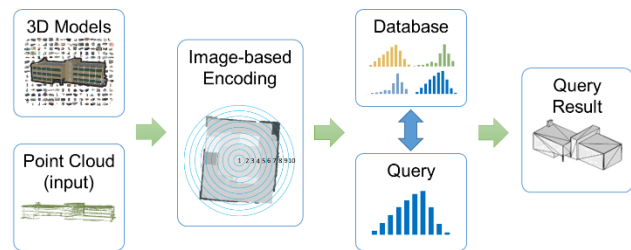

Figure 2. System overview.

2.2 Building Model Encoding

As a preprocessing, a top-view depth image is obtained by setting the projection plan to be the xy-plane, and the origins of the building models and the input point cloud are consistently set to the center of the 3D shape volume. Given the pixels in depth image $\{(x_k, y_k, d_k)\}_{k=1}^n$ where n represents the number of pixels in the depth image and d_k denotes the pixel depth, the shape origin (x_0, y_0, z_0) is defined as calculating the weighted position in the depth image, that is,

$$(x_0, y_0, z_0) = \frac{1}{n}\sum_{k=1}^n (x_k, y_k, d_k) \times d_k \qquad (1)$$

Both the point cloud and building model is moved to the shape origin (x_0, y_0, z_0) prior to the encoding. This process can reduce sensitivity to the 3D object origins in depth image generation and encoding.

Spatial histogram is used as features to represent 3D shape, in which geometric features are accumulated in spinning bins centered on the shape origin, as shown in Figure 3. In this study, three features are used, that is, height feature, edge feature, and eigen feature. These features are described as follows.

2.2.1 Height Feature: The roof information is represented as pixel height in the depth image. Therefore, pixel height is used as feature in shape matching. An example of height feature histogram is shown in Figure 3.
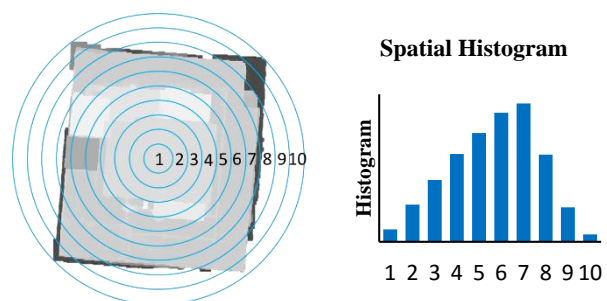

Figure 3. Result of height feature.

2.2.2 Line Feature: To extract line features of the depth image with inherent noises, the Laplacian of Gaussian (LoG) filter is adopted. In LoG filter, the Laplacian filter is second derivative filter used to find rapid changes in the depth image and the Gaussian filter is used to suppress noises. An example line feature extract and line feature histogram is shown in Figure 4.
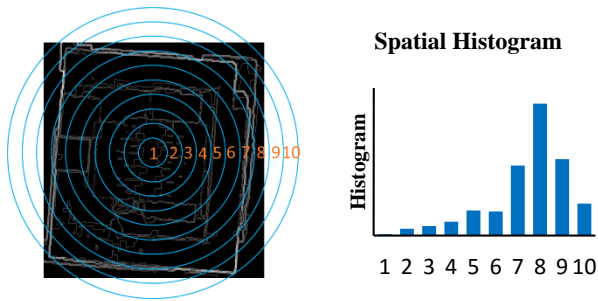
Figure 4. Result of line feature.

2.2.3 Eigen Feature: Eigen-features from the principal component analysis of a point set are useful geometric features that can describe the local geometric characteristics of a point set and indicate whether the local geometry is linear, planar, or spherical. In this study, the eigen-features of depth image is extracted and used as features for encoding. Given a 3D point set $\mathbf{P} = \{(x_i, y_i, d_i)\}_{i=1}^{n_k}$ within a circle of diameter r centered at pixel $p_c : (x_c, y_c)$ of the depth image, an efficient method to compute the principal components of the point set $\mathbf{P}$ is to diagonalize the covariance matrix of $\mathbf{P}$. In matrix form, the covariance matrix of $\mathbf{P}$ is written as

$$\mathbf{C(P)} = \sum_{p_i \in \mathbf{P}} (p_i - p_c)^T (p_i - p_c) \qquad (2)$$

The eigenvectors and eigenvalues of the covariance matrix are then computed by using matrix diagonalization technique, that is,

$$\mathbf{V^{-1}CV = D}, \qquad (3)$$

where $\mathbf{D}$ is the diagonal matrix containing the eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ of $\mathbf{C}$, and $\mathbf{V}$ contains the corresponding eigenvectors. The obtained eigenvalues are greater than or equal to zero, that is, $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$, because the covariance matrix is a symmetric semi-positive matrix. In geometry, the eigenvalues relate with an ellipsoid that represents the local geometric structure of a point set. $\lambda_1 \geq \lambda_2, \lambda_3$ represents a stick-like ellipsoid, meaning a linear structure such as building edges. $\lambda_1 \cong \lambda_2 >> \lambda_3$ indicates a flat ellipsoid, representing a planar structure. $\lambda_1 \cong \lambda_2 \cong \lambda_3$ corresponds to a volumetric structure such as corners of buildings. Some combinations of these eigenvalues provide discriminant geometric features, especially for the point clouds in urban areas. Following the definitions in (Gross and Thoennessen, 2006), the eigen-features of planarity P_λ is defined as

$$P_\lambda = (\lambda_2 - \lambda_3)/\lambda_1. \qquad (4)$$

The planarity feature has the ability to enhance planar structures. An example of eigen-feature histogram is shown in Figure 5.
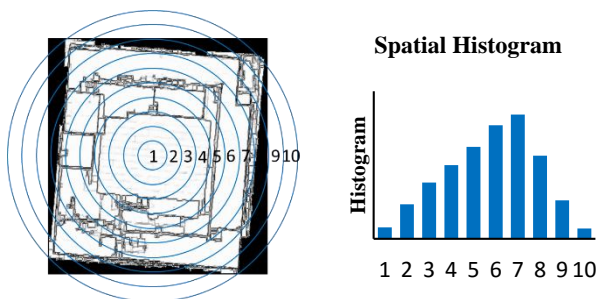

Figure 5. Result of eigen-feature.

2.3 Encoding and Indexing

By combining the all geometric features, a point cloud or polygon model $\mathbf{S}$ is encoded as

$$F(\mathbf{S}) = \{(h_1, \cdots, h_k), (l_1, \cdots, l_k), (e_1, \cdots, e_k)\}, \qquad (5)$$

where h, l, and e represents the height, line, eigen features, respectively; k denote the number of bins in the spatial histogram. In the experiments, k is set to 20 to consider both the compact encoding and sufficient geometric description.

With the shape encoding in (5), the shape similarity measurement is formulated as the distance between the encoded coefficients of the pint cloud P and the building model M:

$$dist(\mathbf{P}, \mathbf{M}) = |F(\mathbf{P}) - F(\mathbf{M})| \qquad (6)$$

2.4 Encoding Properties

Shape retrieval based on the proposed encoding approach introduces several properties that demonstrate the potential for building model retrieval by point clouds. First, the proposed approach provides a metric in which similar shapes have small distances, whereas dissimilar ones have larger distances. Second, the proposed approach is capable of consistently encoding point clouds and polygon models with the aid of data resampling. To demonstrate this property, building models and their corresponding point clouds are tested. The encoding results in Figure 6 show that the building models and point clouds have similar coefficients, which indicates that they are consistently encoded.
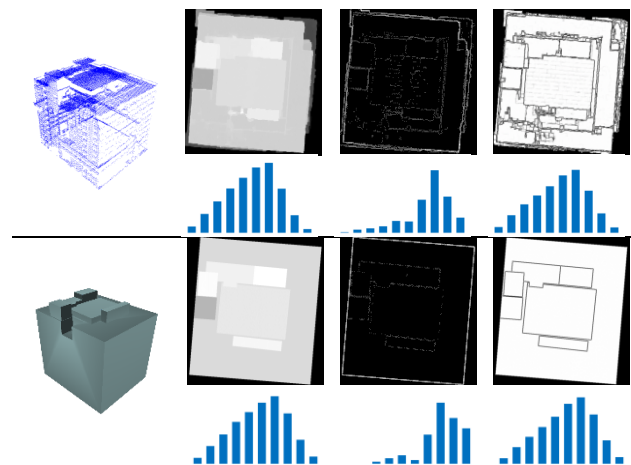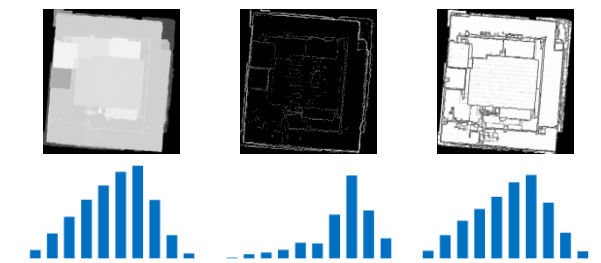

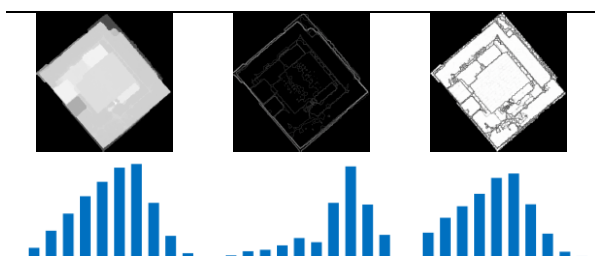Figure 6. Consistent encoding of point clouds and building models.

Figure 7. Demonstration of rotation invariance.

Third, the coefficients are inherently rotation invariant. As shown in Figure 7, the coefficients remain unchanged when rotations are applied to a building model. Fourth, the proposed encoding scheme is potentially insensitive to noise. It is because that features used in encoding are based on spatial statistics, that is, eigen-feature and spatial distribution. For example, in Figure 8, a point cloud with Gaussian noise magnitudes (the standard deviation is set to 0.1) is tested. Results show that the encoded coefficients exhibit only slight differences when noise is present in the data. Fifth, the proposed encoding facilitates the reduction of dimensionality in shape description since a small set of encoding feature is used. With the aforementioned properties, the proposed retrieval method can efficiently and accurately retrieve polygon models by point clouds.
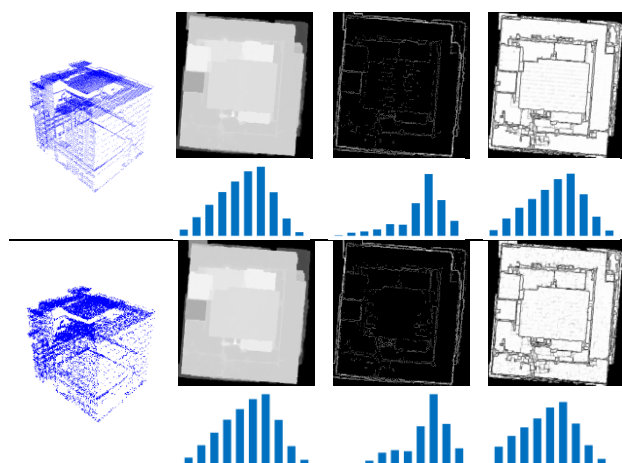

Figure 8. Demonstration of noise insensitivity.

3. EXPERIMENTAL RESULTS

A database including about 900,000 3D models collected from the Internet is used for evaluation of data retrieval. The related methods for comparison is an online system using model-based retrieval (Chen et al., 2014). The results of building model retrieval using a point cloud acquired by airborne LiDAR is shown in Figure 9 to Figure 16. The first rankings of the extracted models in each results of the proposed method are closest to input queue which indicates that the accuracy of geometric encoding and the successfulness of consistent encoding.

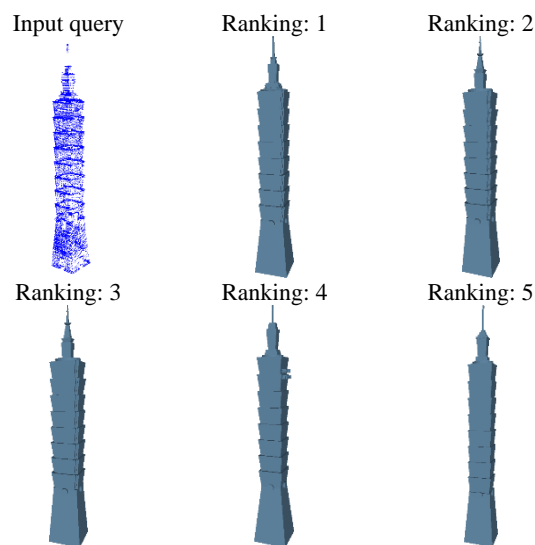Figure 9. Retrieval results of the compared method.

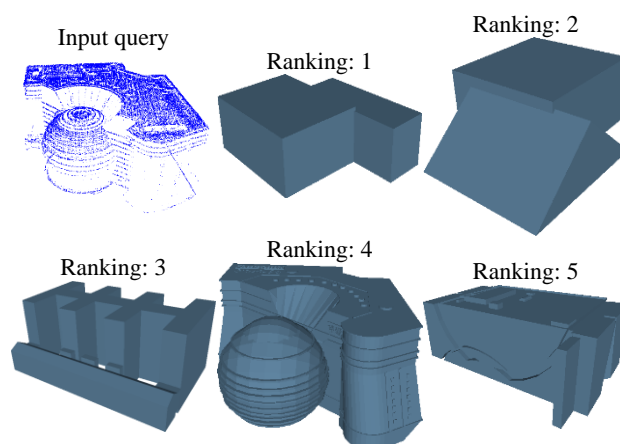Figure 10. Retrieval results of the proposed method.

Figure 11. Retrieval results of the compared method.
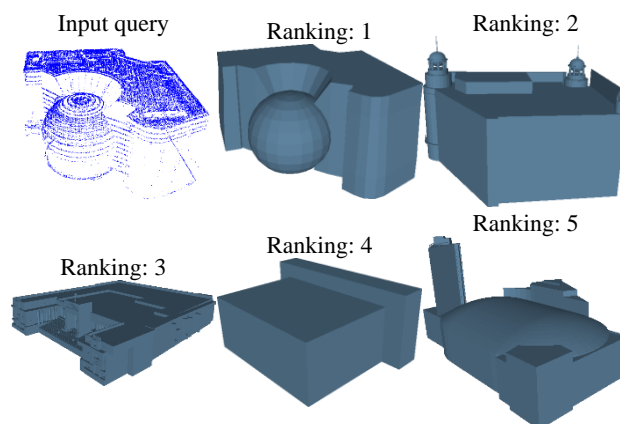
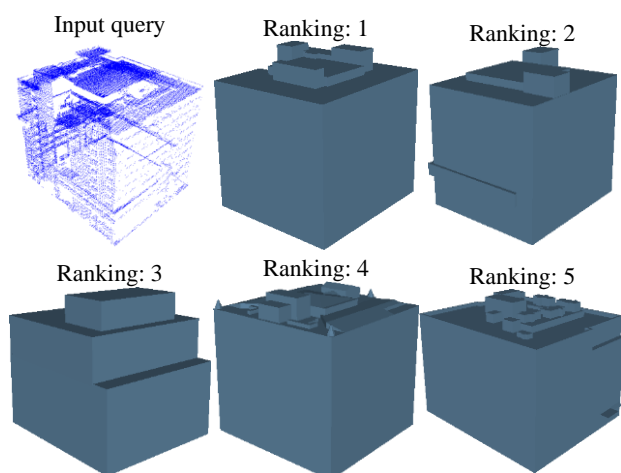Figure 12. Retrieval results of the proposed method.

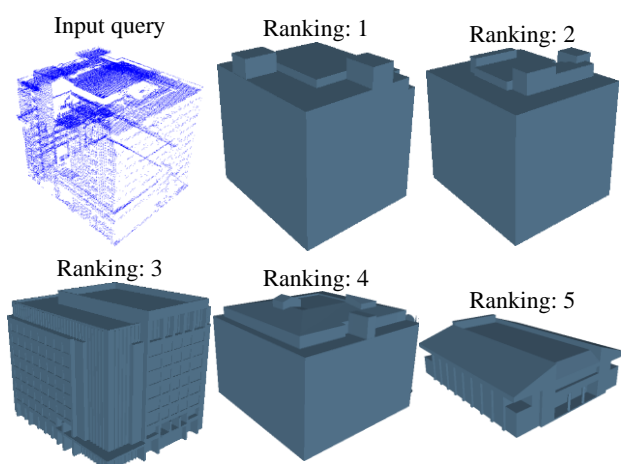Figure 13. Retrieval results of the compared method.

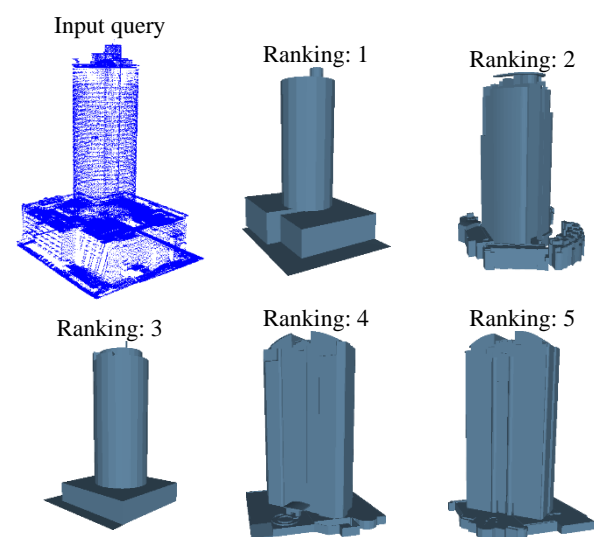
Figure 14. Retrieval results of the proposed method.

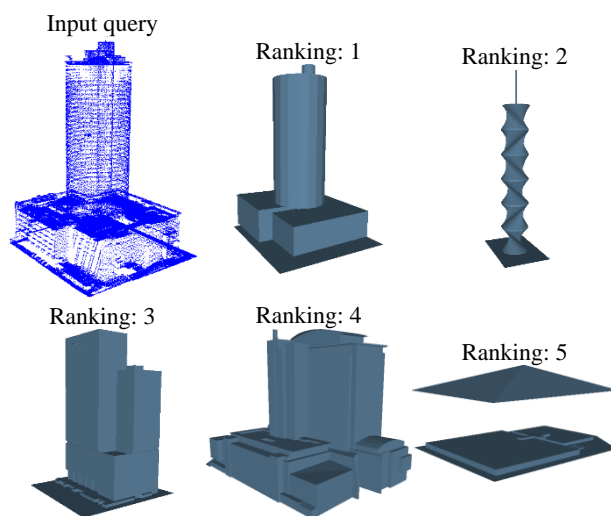Figure 15. Retrieval results of the compared method.

Figure 16. Retrieval results of the proposed method.

4. CONCLUSIONS

A building model retrieval method using a point cloud as query input was presented. With the proposed encoding approach, the building models in the database and the input point clouds can be consistently and accurately encoded. The proposed encoding approach based on geometrically spatial histogram introduces the properties of rotation invariance and noise insensitivity. The experimental results of airborne LiDAR data demonstrate the efficiency and accuracy of the proposed approach, and the qualitative and quantitative analyses show the clear superiority of the proposed method over the related methods.

REFERENCES

Funkhouser, T., Min, P., Kazhdan, M., Chen, J., Halderman, A., Dobkin, D., and Jacobs, D., 2003. A search engine for models. ACM Trans. on Graphics, 22(1), pp. 83-105.

Assfalg, J., Bertini, M., Del Bimbo, A., and Pala, P., 2007. Content-based retrieval of 3-D objects using spin image signatures. IEEE Trans. on Multimedia, 9(3), pp. 589-599.

Mademlis, A., Daras, P., Tzovaras, D., and Strintzis, M. G., 2009. Ellipsoidal harmonics for 3-D shape description and retrieval. IEEE Trans. on Multimedia, 11(8), pp. 1422-1433.

Akgul, C. B., Sankur, B., Yemez, Y. and Schmitt, F., 2009. 3D model retrieval using probability density-based shape descriptors," IEEE Trans. on Pattern Analysis and Machine Intelligence. 31(6), pp. 1117-1133.

Gao, Y., Tang, J., Hong, R., Yan, S., Dai, Q., Zhang, N., and Chua, T.-S., 2012. Camera constraint-free view-based 3D object retrieval. IEEE Trans. on Image Processing, 21(4), pp. 2269-2281.

Chen, D. Y., Tian, X. P., and Shen, Y. T., 2003. On visual similarity based 3D model retrieval. Computer Graphics Forum, 22(3), pp. 223-232.

Stavropoulos, G., Moschonas, P., Moustakas, K., Tzovaras, D., and Strintzis, M. G., 2010. 3-D model search and retrieval from range images using salient features. IEEE Trans. on Multimedia, 12(7), pp. 692-704.

Tam, K. L., and Lau, W. H., 2007. Deformable model retrieval based on topological and geometric signatures. IEEE Trans. on Visualization and Computer Graphics, 13(3), pp. 470-482.

Jain, V., and Zhang, H., 2007. A spectral approach to shape-based retrieval of articulated 3D models. Computer-Aided Design, 39(5), pp. 398V407.

Papadakisa, P., Pratikakisa, I., Perantonisa, S., and Theoharisb, T., 2007. Efficient 3D shape matching and retrieval using a concrete radialized spherical projection representation, Pattern Recognition, 40(9), pp. 2437-2452.

Gao, Y., Wang, M., Zha, Z.-J., Tian, Q., Dai, Q.-H., and Zhang, N.-Y., 2011. Less is more: efficient 3-D object retrieval with query view selection. IEEE Trans. On Multimedia, 13(5), pp. 1007-1018.

Chen, J.-Y., Lin, C.-H., Hsu, P.-C., and Chen, C.-H., 2014. Point cloud encoding for 3D building model retrieval. IEEE Trans. on Multimedia, 16(2), pp. 337-345.

Daras, P., Zarpalas, D., Tzovaras, D., and Strintzis, M. G., 2006. Efficient 3-D model search and retrieval using generalized 3-D radon transforms. IEEE Trans. Multimedia, 8(1), pp. 101V114.

Gross, H., Thoennessen, U., 2006. Extraction of lines from laser point clouds. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Calgary, Canada, pp. 86-91.