Building Model Retrieval System for Automatic City Modeling

WEBCON4 at ACRS2014 in Nay Pyi Taw

Name: Yi-Chen Chen, Jyun-Yuan Chen

Affiliation: Department of Geomatics, National Cheng-Kung University

Country: Taiwan

Date of Birth: July 1, 1991. August 17, 1982.

E-mail: clearholder2005@gmail.com slanla@gmail.com

1. Title

Building Model Retrieval System for Automatic City Modeling

2. Design Concept and Purpose

An increasing number of three-dimensional (3D) building models are being made available on Web-based model-sharing platforms. Based on the concept of data reuse, an automatic 3D building model retrieval system is created to query the similar models by using point clouds which acquired by airborne light detection and ranging (LiDAR) systems. To encode LiDAR point clouds with sparse, noisy, and incomplete sampling, a novel encoding scheme is introduced based on a set of low-frequency spherical harmonic basis functions. These functions provide compact representation and ease the encoding difficulty coming from inherent noises of point clouds. Additionally, a data filling and resampling technique is proposed to solve the aliasing problem caused by the sparse and incomplete sampling of point clouds.

The problem of 3D building model retrieval using airborne LiDAR point clouds as input queries is addressed in this project. LiDAR is an optical scanning technique that is capable of measuring the distance to a target object. By integrating global positioning system (GPS) and inertial navigation system (INS), an airborne LiDAR gains the capability to acquire high-resolution point clouds from objects in the ground efficiently and accurately. Thus, an airborne LiDAR system can provide surveyors with the capability of digital elevation and cyber city model generation [1]. This project aims at the efficient construction of a cyber-city by encoding unorganized, noisy, and incomplete point clouds, as well as by retrieving 3D building models from model databases or from the Internet.

Recent developments on Web 2.0 technique and scanning equipment have yielded an increasing number of 3D building models in Web-based data-sharing platforms. For example, the Google 3D Warehouse is a web-based data-sharing platform which allows users upload and share their models. Based on the concept of data reuse, 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. Existing studies mainly focus on encoding and retrieving 3D polygon models using polygon models as input queries [2-7]. 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 LiDAR point clouds.

The key idea behind the proposed method is to represent noisy point clouds using a complete set of spherical harmonics (SHs). Point clouds represented by a few low-frequency SHs are insensitive to noises. In addition, SH encoding reduces data description dimensions and yields a compact shape descriptor, resulting in both storage size and search time reduction. Moreover, the inherent rotation-invariant property and multi-resolution nature of SH encoding enable the efficient matching and indexing of the model database. Besides, a data filling and resampling approach is proposed to solve encoding problems coming from the incomplete shapes of point clouds and the aliasing problems of SH coefficients attributed to the sparse sampling of point clouds.

3. Outline of the web material

3.1 Web-based User Interface

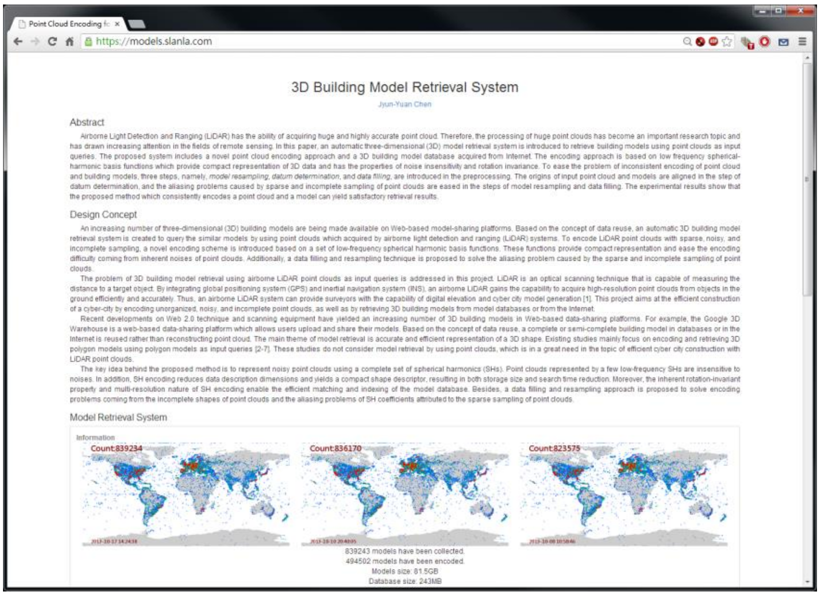

Figure 1

The proposed 3D building model retrieval system is available at https://models.slanla.com/ (Fig. 1). A document of the proposed technique is provided in the website, and the information of the model dataset is provided also. The database is managed by using SQL server and the models are automatically downloaded from Google Warehouse (http://sketchup.google.com/3dwarehouse). The dataset contains 823,575 models on Oct. 8, 2013, and database size is up to 839,234 on Oct. 17, 2013. This implies that the database is updated automatically and daily. Recently, the size of downloaded models is 81.5GB, and the size of encoded models is 243MB.

The input to this system is a building point cloud acquired by airborne LiDAR sensor. The users can input their point clouds to retrieval the similar models. The website also provides three examples of point clouds (Fig. 2).

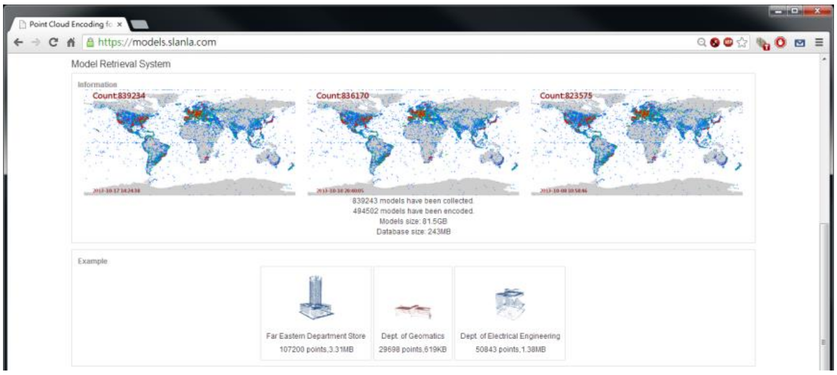

Figure 2

After input the query, the system takes a few seconds for data transmission (it depends on the local bandwidth), and 2-10 seconds for data encoding (it depends on the data size), and about 0.8 seconds for retrieval. After these processes, the system returns the retrieval models which are most similar to the query input (Fig. 3).

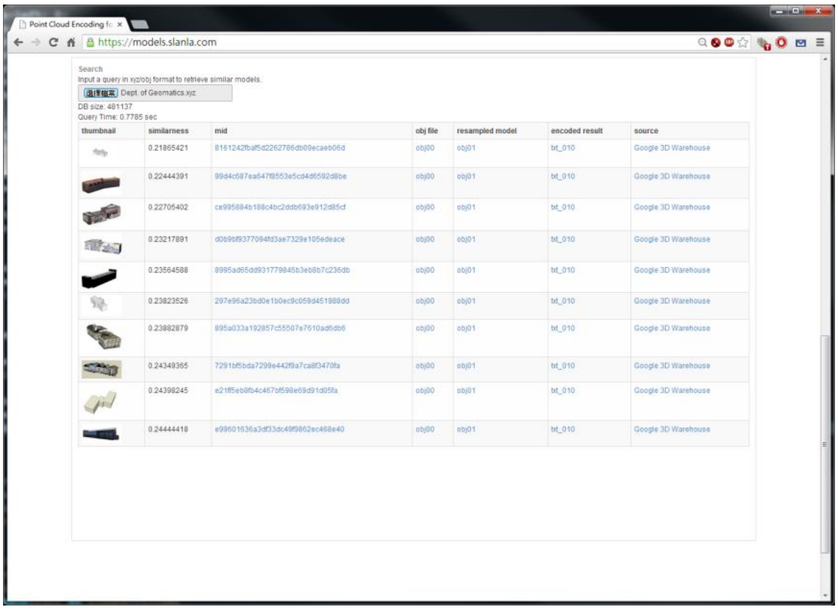

Figure 3

Each retrieval model contains several information including the encoded coefficients, the building model (in obj format), the resampled model, and the source in the Google

Warehouse. Users can get the entire information of the retrieval model from the link to Google Warehouse.

Figure 4

3.2 System Workflow

Fig.5 is the illustration of workflow of the proposed system, the building models and point clouds (i.e., the input query) are consistently encoded by a set of SHs (Spherical harmonics). To achieve consistency in encoding, an airborne LiDAR simulator is utilized to resample the building models. This process enables building models and point clouds to have similar samplings. In addition, a data filling process is performed to solve the aliasing problem attributed to the sparse sampling of point clouds. This process can significantly reduce aliasing errors on encoded SH coefficients, thus improving retrieval accuracy. Two major procedures, namely, preprocessing and data encoding, the preprocessing process could be satisfy the aliasing-free sampling constraint of SHs encoding, the data encoding process is encoded the 3D shape by using SH encoding. and the process reduce the transformation of rotation and transition of model and point cloud, and also reduce the data storage size to a serial SH coefficients that makes the retrieves the similar models quickly by coefficient matching.

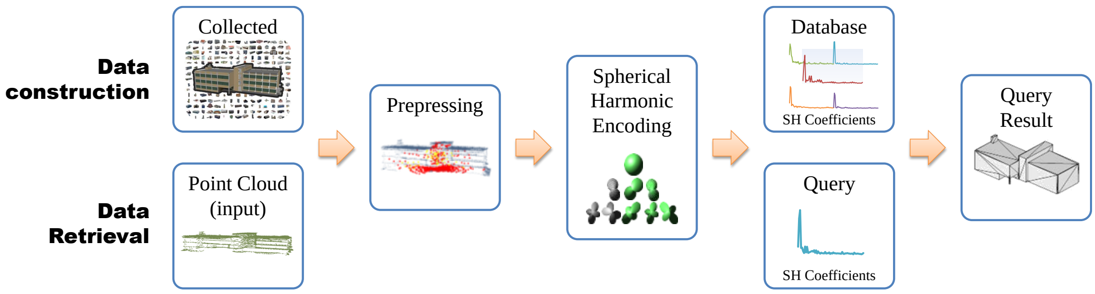

Figure 5. System workflow.

3.3 Model Encoding

A set of SH function is applied to 3D shape encoding for efficient represent point clouds and building models. The SH function is a shape descriptor, there are three properties: 1) Shape distinguishability. SH function has distinguishable outline of shape. The coefficients pair of encoding result of similar 3D building models pair is also similar, but dissimilar is not. 2) Noisy insensitive. For a similar shape pair, the SH coefficients has similar pattern in low- frequency, not influence with noisy. 3) Rotation-invariant. For a similar shape pair, but different attitude angle, the SH coefficients has same pattern not influence with attitude. These powerful properties help the retrieval system more effect and easy.

For any function $f(\theta,\phi)$ on a 3D shape can be represented by linear combination of SH basis functions:

$$f(\theta,\phi) = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} a_l^m Y_l^m(\theta,\phi), \tag{1}$$

where f is the function of shape which is model or point cloud data; $f(\theta,\varphi)$ is the vertex of shape; a_ℓ^m is unknown coefficient of the basic function Y_ℓ^m; $\ell_{\max}$ is a user-defined parameters. By utilizing the fact that the L2-norm of SH coefficients are rotation invariant, a 3D shape is encoded as follows:

$$SH(f) = \left(\left\{ \left\| a_0^m \right\| \right\}_{m=0}^{0}, \left\{ \left\| a_1^m \right\| \right\}_{m=-1}^{1}, \cdots, \left\{ \left\| a_{\ell_{\max}}^m \right\| \right\}_{m=-\ell_{\max}}^{\ell_{\max}} \right) \tag{2}$$

Preprocessing. For a pair of building model and point cloud, the concept of preprocessing is make the data encoding consistently that has resulting similar encode result. According to the property of SH, the encoded result of the similar shape pair is similar or even the same. Therefore, the basic idea is make model and point cloud looks the same before SH encoding step. To achieve this goal, there are two steps, namely, *model resampling* and *data filling*, are introduced in the preprocessing.

Model resampling: According to the aliasing-free sampling constraint Li and North, 1997 and the solved of least-square fitting, the sampling resolution must be content that:

$$\text{sampling resolution} \geq \frac{180^\circ}{l_{max}} \tag{3}$$

Most of collections of models from the internet are not enough vertices. Therefore, a LiDAR simulator is applies to resampling the collected models. The reason is that make resampled sample vertices of model are like the point cloud.

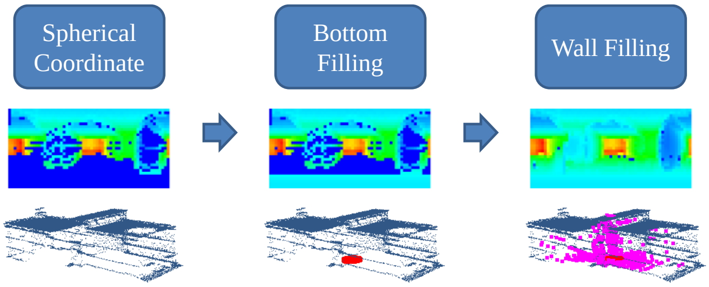

Figure 6. Processes of data filling.

Data filling: The point cloud data always is unorganized, noisy, sparse, and incomplete. Therefore, the same as model resampling, the vertices of shape must be satisfied the constraint of sampling resolution. Consider the distribution of LiDAR data, there are two procedures in the data filling process, namely, bottom filling and wall filling. According the sampling resolution, the vertices are translated into the grids which are means the latitude and longitude of spherical coordinate system. In middle of Fig.6-left, for each pixel, the color form blue to red is mean the distance of datum to vertex, but the blue pixels are mean the empty grids which are need to fills. In the bottom filling, the min z (the red points in bottom of Fig.6-middle) is uses to fills the bottom of shape. After the bottom filling, we refer the points of bottom and roof to fills the wall of building (Fig.6-right).

3.4 Experimental Results

A dataset which contain seven groups of building models shown in Fig.7 is used to verify performance of the proposed approach. In addition, a commonly used measurements precision η_s and recall η_n is adopts in the experiments. These measurements are defined as $\eta_s = \frac{TP}{TP+FP}$ and $\eta_n = \frac{TP}{TP+FN}$, where TP, FP and FN represent true positive, false positive, and false negative, respectively.

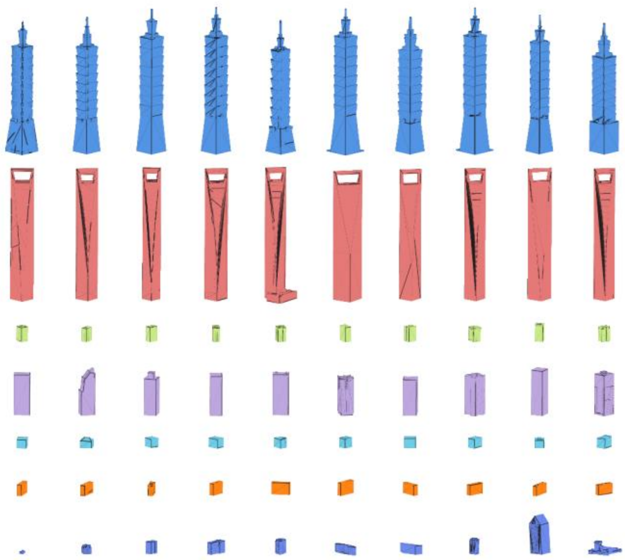

Figure 7. Test dataset.

Parameter Setting. The Maximum value of degree ℓ_{max} is a major parameter which influences the results and efficiency of SH encoding. To ensure the accuracy and efficiency of SH encoding, our approach was tested using various parameter values on simulated point clouds acquired from 3D building models. The experiment result is shows with precision-and-recall curves in Fig. 8, the performance of SH encoding is increased with a raise of ℓ_{max} from 4 to 10. However, the performance is going down when the ℓ_{max} large then 10. Although a large degree ℓ_{max} of SH encoding can be describe more detail of a 3D shape, but also lead to the result of encoding result is sensitive with tiny change between building model and point cloud. In addition, the encoding efficiency is decrease with a raise of ℓ_{max}. Therefore, the maximal degree ℓ_{max} of SH encoding is set to 10.

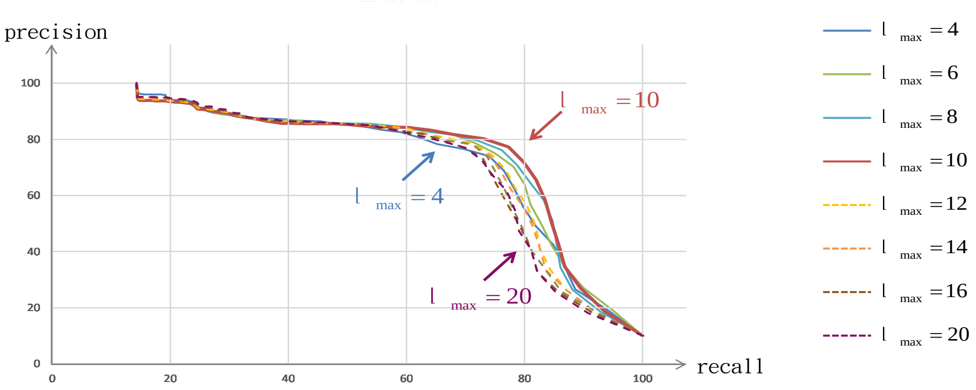

Figure 8. Testing for parameter ℓ_{max}.

Retrieval evaluation. The data preprocessing is proposed to makes the sample points similar to the airborne LiDAR point Cloud. To test the encoding performance of the proposed retrieval approach, the related methods, typical SH encoding approach [2,8] and shape distribution [3,6,9], have used to compare with our approach. The experimental result shows with precision-and-recall curves in Fig. 9, the proposed encoding result (blue line) has better performance with typical SH encoding approach (green line), and shape distribution (red line). The major reason is the proposed approach is resamples the model with the LiDAR simulator and fills the model and point cloud in the same condition. This implies that the similar pair of building model and airborne point cloud has similar outline and also has similar encoded coefficients.

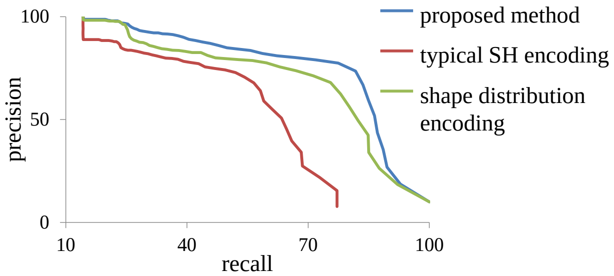

Figure 9. Retrieval evaluation.

4. Website

Our system is construction on the Internet, the website's URL is http://models.slanla.com. This system allow user upload a point cloud data of building, the data formation is XYZ ANSI-formation. The uploaded file must be a clean data not contain outlier data, like the point cloud of tree, car, person...etc.

References

[1] J. P. Wilson, ”Digital terrain modeling”, Geomorphology, vol. 137, issue 1, pp. 107-121, 2012.

[2] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, ”A search engine for models”, ACM Trans. on Graphics, vol. 22, no. 1, pp. 83-105, 2003.

[3] J¨urgen Assfalg, Marco Bertini, Alberto Del Bimbo, and Pietro Pala, ”Content-based retrieval of 3-D objects using spin image signatures”, IEEE Trans. on Multimedia, vol. 9, no. 3, pp. 589-599, 2007.

[4] A. Mademlis, P. Daras, D. Tzovaras, and M. G. Strintzis, ”Ellipsoidal harmonics for 3-D shape description and retrieval”, IEEE Trans. on Multimedia, vol. 11, issue 8, pp. 1422-1433, 2009.

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

[6] C. B. Akgul, B. Sankur, Y. Yemez, and F. Schmitt, ”3D model retrieval using probability density-based shape descriptors”, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 31, no. 6, pp. 1117-1133, 2009.

[7] Y. Gao, J. Tang, R. Hong, S. Yan, Q. Dai, N. Zhang, and T.-S. Chua, ”Camera constraint-free view-based 3D object retrieval”, IEEE Trans. on Image Processing, vol.21, no.4, pp. 2269-2281, 2012.

[8] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, ”Rotation invariant spherical harmonic representation of shape descriptors”, Proc. Eurographics/ACM SIGGRAPH Symp. Geometry Processing, pp. 156-164, 2003.

[9] D. Y. Chen, X. P. Tian, and Y. T. Shen, ”On visual similarity based 3D model retrieval”, Computer Graphics Forum , vol. 22, no. 3, pp. 223-232, 2003.