

3D Object Categorization Using Spin-Images with MPI Parallel Implementation

A. Eleliemy
ahmed_hamdy@
cis.asu.edu.eg

D. Hegazy
Computer Systems Department
Ain Shams University
Doaa.hegazy@cis.asu.edu.eg

W. Elkilani
Wail.elkilani@cis.asu.edu.eg

Abstract: Object recognition and categorization are two important key features of computer vision. Accuracy aspects represent research challenge for both object recognition and categorization techniques. High performance computing (HPC) technologies usually manage the increasing time and complexity of computations. In this paper, an approach utilizing 3D spin-images for 3D object categorization is proposed. The main contribution of this approach is that it employs the MPI techniques in a unique way to extract spin-images. The technique proposed utilizes the independence between spin-images generated at each point. Time estimation of this technique has shown dramatic decrease of the categorization time proportion to the number of workers used.

Keywords: - Object categorization, spin-image, support vector machine, MPI.

17. Introduction

3D object categorization is one of the important fields in computer vision. Its goal is making the computer (e.g. robot) able to recognize general classes of object categories, which is more human-like ability. 3D object categorization has many different applications including robot navigation and surveillance, security systems and manufacturing quality control. Such variety of applications lead to a large different sets of constraints which give each problem its flavors. In recent few years, the growing demand for on-line categorization systems has motivated a lot of researchers on this scope. So the majority of this work will target the on-line categorization systems using HPC techniques. The model proposed uses spin-image features [1] for recognizing and categorizing 3D objects. Spin images are considered an important 3D feature, used by different approaches for 3D object recognition task [1-5]. However, their high computational performance is its main disadvantage.

The proposed HPC model for 3D object categorization focuses on using MPI technique to accelerate the spin-images generation process. Moreover spin-images generated can be used for feature training of Support Vector Machine (SVM) classifier instead of recognition by surface matching. This paper assumes that the optimum system is subject to the following three constraints:

- *3D object nature constraint:* This constraint is inherited from computer vision goal where the goal is to make computers see like humans. Humans can realize the third dimension but digital images didn't have such information except in the last few years when range cameras and depth sensors have a great development. While techniques needed to infer properties of the 3D world from digital images, these devices give the depth information for the scene. Although the third dimension could be obtained by stereo images, yet development in range and depth sensor technology allows easier use

of such information [6]. The third dimension information adds more reality for the problem domain. So the system must deal with 3D world information.

- *Real time Constraint:* What is meant by real time constraint is the system ability to process at least from 10 frame per second (fps) up to 33 (fps) [7]. Such variety comes from the system domain. In many techniques the timing constraint is tied to the categorization success rate. To increase success categorization, it may increase the technique complexity which means more time. So categorization success rate will be treated as a separate constraint.
- *Success categorization rate:* This constraint is to show how system could be tolerant. Depending on application some categorization techniques could allow false categorization rate more than others.

In this paper, we present an approach for 3D object categorization using spin-image features and Support Vector Machines (SVM) classifier with a MPI parallel implementation. The proposed model tries to meet the mentioned constraints to achieve an optimum categorization performance.

The paper is organized as follows: Section II presents the related work. The proposed 3D categorization model and its parallel version are discussed in section III. Section IV presents the experimental evaluation and the conclusions are given in section V.

18. Related Work

Due to the importance of 3D object categorization, many researchers have targeted the problem which resulted in different categorization approaches such as (8). However, only few approaches addressed the on-line categorization using HPC.

As we mentioned before, spin-images are used by different categorization approaches. One of these approaches is presented in [1], where spin-images are used to recognize multiple

3D objects in a scene for 3D object retrieval. This approach is based on surface matching, by matching surfaces of already stored objects with the new two incoming captured objects. Spin-images for each known model are generated and stored. In the recognition step, spin-images are generated for each point of the incoming captured scene. Finally the similarity between the calculated new points and the spin-images from all stored models is measured. Best matching means both of recognition and localization for the object are done. Although such work doesn't discuss any timing measures, but it is clear that such technique consists of two heavy computational tasks: Spin-images Generation process and Surface Matching. Such research [1] has been extended and improved in [9]. Instead of using the original spin-image algorithm, they introduced an enhanced version called spin-image signatures. Spin image signatures used for content-based retrieval of 3D Objects and shows some important results in terms of sensitivity to model deformations. Such approach includes three steps: 1-original spin-images generation, 2-spin image signatures, 3- clustering of spin-image signatures. Total complexity is in order of $O(n^4)$, where n represents the vertices' of 3D object. Such approach violates the time constraint. In [10], the photometric information is considered. New version of the original spin-image algorithm, called textured spin-images, is introduced. Textured spin-images are used in

3D registration. Such algorithm enjoys remarkable properties, since it can give rigid motion estimates, more robust, more precise, and more resilient to noise than standard spin-images. But this approach involves the same steps to obtain spin-images of 3D object. Although it reduces complexity, it still involves intensive computation for spin-image generation and calculation.

Also spin-image algorithm is a base for some research at the area of face recognition. In [4], a feature based method for face recognition is introduced. Surface shape information is inferred from image brightness using a Lambertian shape- from-shading scheme. Spin-image was used to perform surface representation from the surface normal. To reduce the computation of spin-images, local spin-images are computed on the image patches using surface normal information. Although such approach could avoid some of the spin-image algorithm complexity, it calculates spin-images at some points not at each vertex which means the spin-image generation process is still as it is.

In [11], has introduced a recognition system based on a hardware accelerated implementation using Field Programmable Gate Array (FPGA). The proposed hardware is used in spin-images generation process and surface matching have. Such work enhances the overall system run time from several minutes in [1] to 4.5 seconds. However, FPGA is not the only possibility for performance enhancement different possibilities to give better performance, such as pipelining, lookup-tables for square roots and input-output stage.

Another approach is presented in [12], where the heavy task of spin-image surface matching is passed. This work introduces 3D object categorization based on spin-images and bag of features classification instead of surface matching. Such approach gives significant results in terms of recognition rate 62.5% at 100 spin-images per object. But the problem of heavy spin-images generation calculation is not avoided. There is no concern for the overall system runtime. Therefore the approach presented in [12] violates real-time constrains.

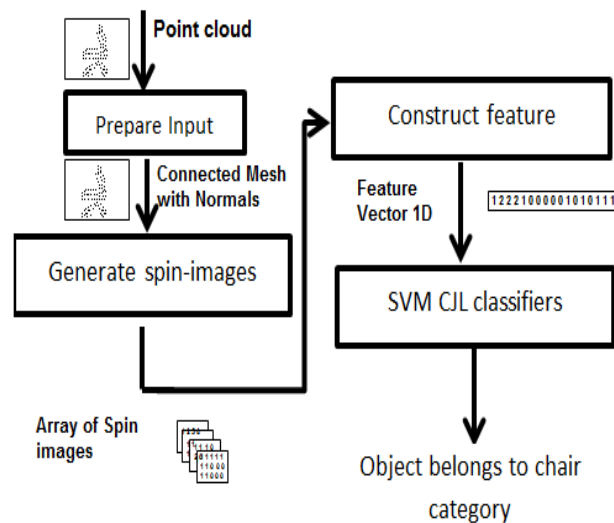


Fig. 1. Sequential 3D object categorization system

It can be noted from the previous discussion that the main drawback of the spin-image algorithms proposed is the heavy computation involved. Most authors have concentrated on elevating the recognition rate without considering system run time. Moreover, as far we have seen that neither of authors have parallelized spin-image generation algorithm using MPI. In the proposed approach, we have utilized the MPI technology to reduce dramatically the overall system runtime with maintaining the recognition rate. In fact we have enhanced slightly the recognition rate.

19. Proposed Categorization Model

A. Sequential 3D object categorization model

The proposed 3D object categorization model consists of four main steps, as shown in Figure 1. The first is the preparation step. Normally, 3D objects are obtained from different depth-image sensors in the form of point cloud. Such input needs to be converted into the form of connected mesh. Connected meshes could be obtained from point cloud by many triangulation techniques such surface reconstruction from unorganized points in [13], 3D mesh reconstruction from point cloud using elementary Vector and geometry analysis in [14] or reconstruction from point clouds in 3D in [15]. However the data gathered is synthetic 3D objects not a sensory input images as in mentioned in section IV. This means that the data is already in the form of triangulated mesh. Therefore, in the preparation step in the built model the normal vector for each point in the mesh surface is calculated. In the second step of the model, spin-images for each object are generated. Then, the generated spin-images are converted to feature vector. So, each object is represented in the model by its corresponding feature vector. Finally, the calculated feature vectors are learned and classified using SVM classifier. Now, we will give a brief description of the spin-image generation as it is an important step in the proposed model.

Calculating Spin-images: Spin-image is a data level shape descriptor used for recognition of multiple objects. Spin- images are introduced in [1] in manner that recognition is done based on matching surfaces by matching points using the spin-image representation. Spin-image is calculated for each oriented point in the 3D mesh. Figure 5 shows the generation process of spin-images at point P. According to equations to 1 and 2, two values (α , β) are calculated with each oriented point like Point X.

$$\alpha = \sqrt{\|x - p\|^2 - (n \cdot (x - p))^2} \quad (1)$$

$$\beta = n \cdot (x - p) \quad (2)$$

In [1] the analogy to describe the image generation process, a white paper which is rotated around the normal at the point, when they touch the paper, they stick to it and form a 2D cumulative image. The generation of spin-images is affected by four important parameters: bin size, support angle, image width and support distance. Values of these parameters depend on the objects data set. Bin size determines the storage size and the descriptiveness of the spin images. The best value for bin- size is a multiple of mesh surface resolution (1). Image width

means number of rows or columns as square spin-image is used. Support angle is the maximum angle between normal of the oriented point and the normal of points that will participate in generated spin-image. Support distance equals image-width multiplied by bin-size. Estimating the best values of each parameter is done with experimental evolutions explained in section IV.

After the spin-images generation done. The resulted 2D spin-images for each 3D object are converted to be the feature vector of these object, by converting the 2D spin-image into 1D vector where all rows in 2D spin-image became at one row. The feature vectors are learned and classified using support vector machines (SVM). SVM classifiers are famous and widely used classifiers that showed robust performance in many classification tasks [16]. Due to space limitation, the reader can refer to [8] for more details. However, we provide here a short explanation of how new images are classified in the proposed approach. After training the SVM classifiers using the spin-image features, a list of SVM binary classifiers with length L is generated, where $L = \binom{n}{2}$ represents the number of categories (six categories are mentioned in section IV). This is called the classifier: Chain Judges List (CJL). Figure (2) shows how judges in the (CJL) classifier work. Assume we have 4 categories W, X, Y and Z, so judges list consist of 6 judges $J_{WX}, J_{WY}, J_{WZ}, J_{XY}, J_{XZ}$ and J_{YZ} . J_{XY} means that this SVM binary classifier is trained with the spin-image features of the images of two categories X and Y. Training each classifier in the six classifiers is performed in independent manner, each represents an object category. After training, a new previously unseen image is presented to the system for testing and the spin-image feature is calculated. The test feature vector is then presented to each binary classifier. All classifiers vote for the input image and the classifier (category) with maximum votes will win.

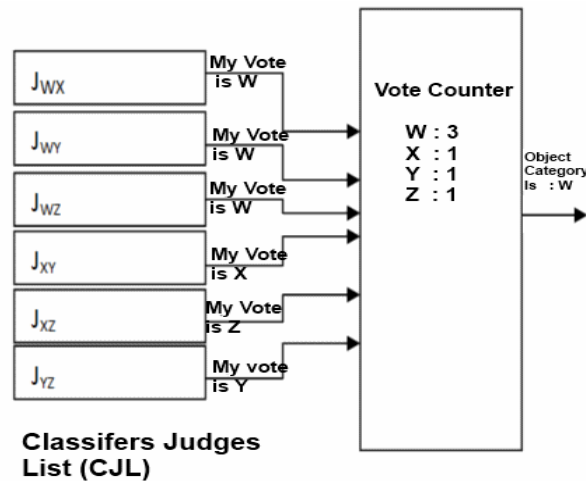


Fig. 2 Classifier judges list for different 4 categories W, X, Y and Z

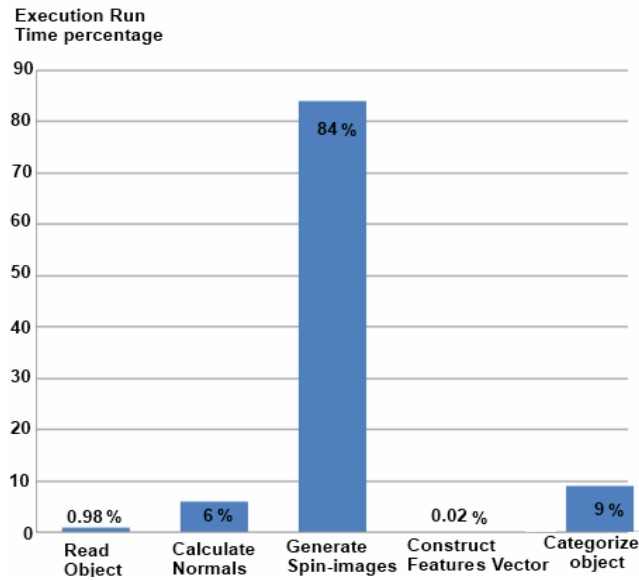


Fig. 3 System runtime analysis

B. 3Dparallelcategorizationmodel

In the sequential model the spin-image generation process consumes the majority of system time (almost 84% of system time). Figure 3 records how the system time is consumed by different system stages. According to equation 1, the task dependency between generations of spin-images is minimal. The idea is to convert the sequential spin-images generation process to parallel form. Figure 4 shows the parallel categorization model. Based on message passing interface (MPI), a group of workers cooperate to perform the 3D object categorization process.

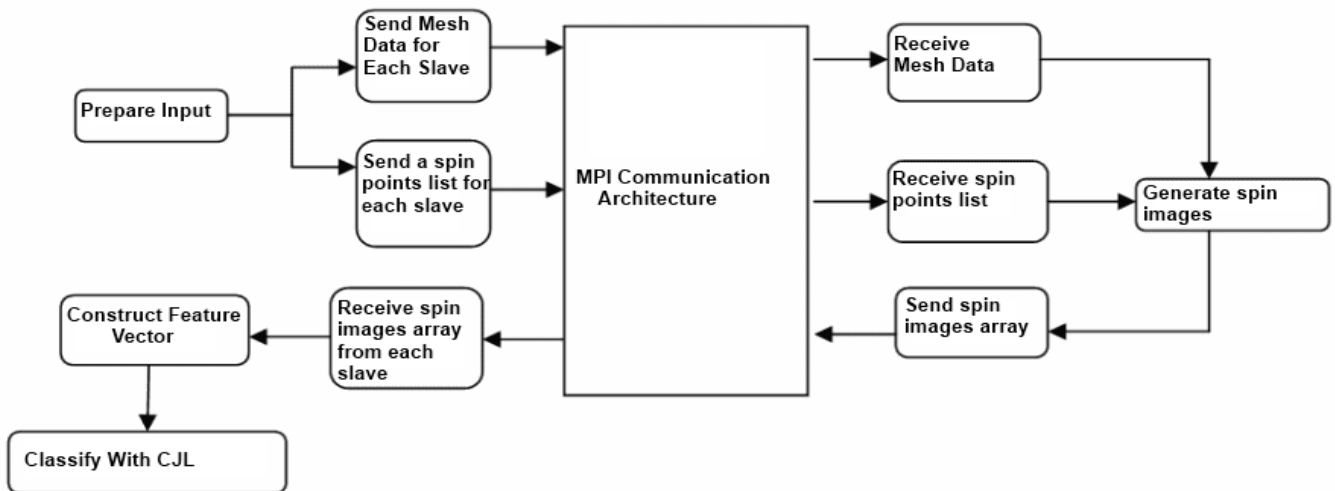


Fig. 4 Parallel 3D object categorization

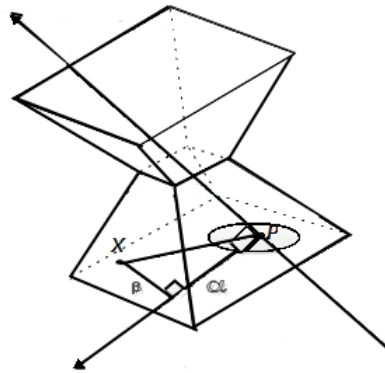


Fig. 5 The generation process of spin-image at Point P with normal n for 3D object

Based on Master/slaves model, one of these workers is the master who performs the preparation for the input 3D information. The master prepares the input to be in form of connect faces (Mesh) like in figure 5, also calculates the normal at each point. The master then sends both of prepared mesh data and list of spin-points to each slave. Each slave receives the mesh data and the list of spin-points and then calculates its spin-images. Slaves send the spin-images to the master worker which converts spin-images into feature vector (observation) to be classified according to the trained judges list as in the proposed sequential model. Category of maximum judges' votes is the decision.

20. Experimental Evaluation

A.Data

The Princeton Shape Benchmark (PSB) is used in all the experimental evaluation of the proposed model. PSB is one of the public available databases that contains a variety of synthetic3D models.

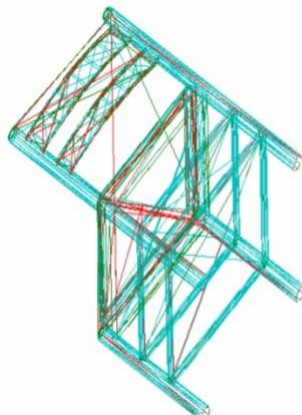


Fig. 6 Sample of Mesh object constructed at master worker

TABLE. I TABLE OF PSB FURNITURE CLASSES

Class	Number of objects
Table	69
Chair	29
Couch	15
Cabinet	9
shelf	26
Bed	8

Therefore the generated images don't suffer from noise, like in sensory captured data. Each class in PSB contains only one object [17]. The furniture objects are selected to be the focus of the proposed categorization model. Table I shows how many samples of each object are available. Also Figure 6 shows rendering of some PSB 3D Furniture objects. As mentioned

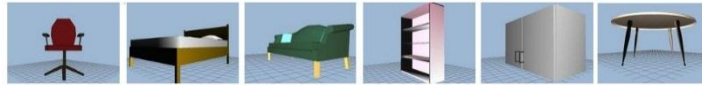


Fig. 7 Some of PSB furniture objects

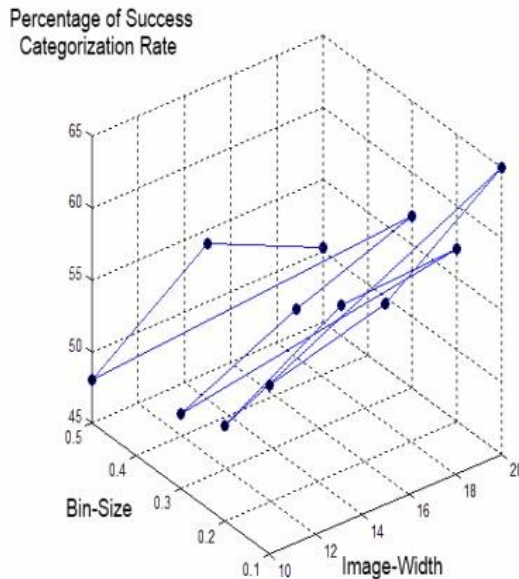


Fig. 8 Effect of bin-size parameter, support-angle and image-width on the mean of categorization rate

before, spin image generation process is done at only the oriented points. Objects in PSB are stored in a format which contains points and faces of object made by these points. We divide the gathered data into two parts: one for training and the other one is for testing. The ratio between the two parts is 80% and 20% respectively.

B.Results

As mentioned before, the quality of generated spin-images are related to the generation parameters, also they have direct effect on the categorization process. Several experiments have been already performed to select be the best parameters. Figure 8 shows the increase of bin-size for decrease of the success categorization rate. In the other side, the increase of the image width causes the increase of the success categorization rate. The best success categorization rate has been recorded at the intersection point where image-width is 20 pixels and bin-size equals 0.1.

In the proposed model, we didn't look up for minimum spin points to get highest categorization rate. Instead of that we tried different number of random spin point to identify the best categorization rate. As shown in Figure 9, the best success categorization rate has been recorded at 100 random spin points.

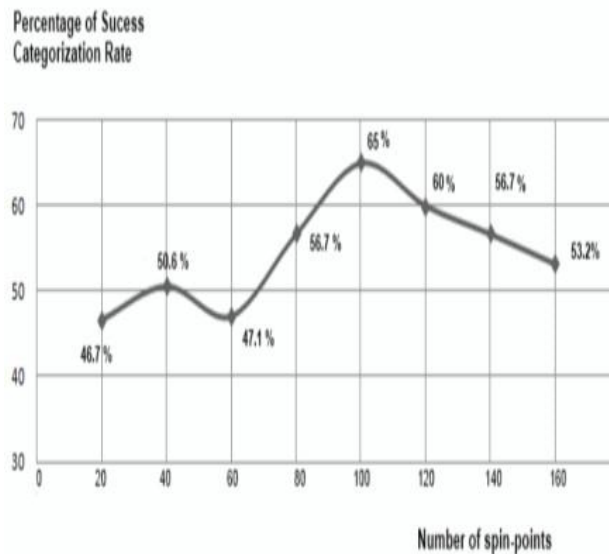


Fig. 9 Mean of success categorization rate at different number spin-point

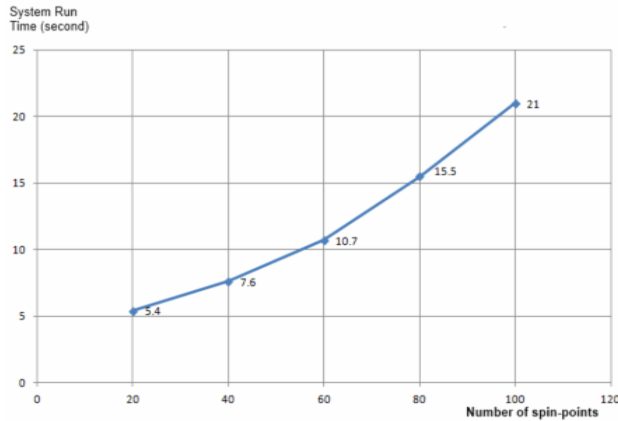


Fig. 10 Relation between system overall runtime and number of spin-points

Figure 10 shows the relation between system execution run time and number of spin points. While the number of taken spin points increases, the total amount of system execution time increases. As in figure 9, 100 spin point have best categorization rate which means 21 seconds as a total recognition time.

The confusion matrix in figure 11 shows how failure and success categorization is distributed among different system categories. For example the first row, 66% of cabinet object are correctly categorized to Cabinet, while 33% are incorrectly categorized to chair category.

Ground Truth	Predicted					
	Cabinet	Bed	Couch	Chair	Shelf	Table
Cabine	0.66	0	0	0.33	0	0
Bed	.0667	.7333	.0667	.0667	.0667	0
Couch	0	0	1	0	0	0
Chair	0	0	0	1	0	0
Shelf	0	.3333	0	.3333	.3333	0
Table	0	.1667	.3333	.1667	.1667	.1667

Fig. 11 Confusion matrix of proposed model shows the mean of success categorization rate equal to 65%

TABLE 2. TABLE OF SEQUENTIAL AND PARALLEL PARTS OF THE PROPOSED APPROACH

Number of workers	Time of Sequential Part	Time of parallel part
1	3.36	17.64 s
2	3.36	8.85 s
3	3.36	5.92 s
4	3.36	4.48 s
5	3.36	3.62 s
6	3.36	3 s

Figure 12 shows the results of the parallelized system. Increasing the number of the workers in our approach will decreased the total amount of categorization time. The proposed parallel approach has two parts: Sequential part and parallel part. Table II shows how sequential part is independent on number of workers while the parallel part is directly dependent on the number of workers. For six workers, we have reach a recognition time of 6.36 seconds. As listed previously that [11] used parallelized FPGA approach has reached 4.5 seconds. The suggested MPI approach is preferable for the well-known merits of software parallelized approaches. Obviously, that increasing number of workers will minimize categorization time runtime until the number of workers equates the number of spin points. Driving empirical formula for their relationship will be interesting future research work.

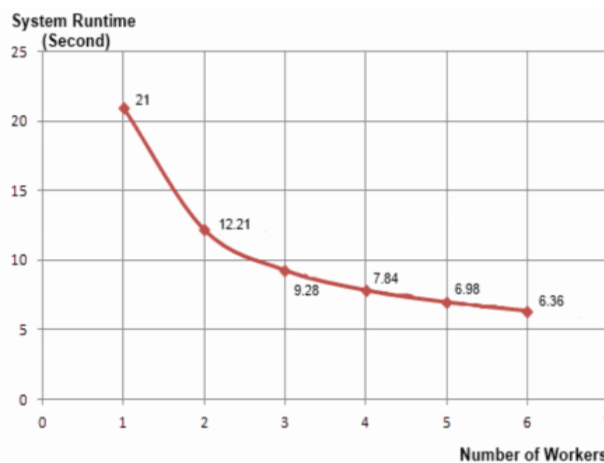


Fig. 12 Relation between number of workers and overall system run time

21. Conclusion and Future Work

Although using spin-images with SVM classifiers in 3D object categorization includes a heavy processing task, we used a distributed parallel model based on MPI to speedup the spin-image generation process which consumes the majority of the model processing power. As we can drive the maximum speed up obtained by our parallel proposed model. We recommend

such approach especially for objects with high 3D point density. Also we can parallelize the generation of single spin-image at each worker. We are planning to switch the approach proposed to use hybrid model of MPI and Open MP. We intend to use the kinect camera as sensory input and depth- image sensory database like (18).

22. References

- [1] A. E. Johnson and M. Hebert, “Using spin images for efficient object recognition in cluttered 3d scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 433–449, 1999.
- [2] L. Wang, J. Sun, and Q. Wang, “Modified spin- image target recognition algorithm applied to laser radar range imagery,” *Conference on Lasers and Electro- Optics/Pacific Rim 2009*, Optical Society of America, 2009.
- [3] T. Tamaki, S. Tanigawa, Y. Ueno, B. Raytchev, and K. Kaneda, “Scale matching of 3d point clouds by finding key scales with spin images,” *2010 20th International Conference on Pattern Recognition (ICPR)*, pp. 3480–3483, 2010.
- [4] M. Bock, G. Cortelazzo, C. Ferrari, and C. Guerra, “Identifying similar surface patches on proteins using a spin-image surface representation,” in *Combinatorial Pattern Matching*, ser. *Lecture Notes in Computer Science*, A. Apostolico, M. Crochemore, and K. Park, Eds. Springer Berlin Heidelberg, vol. 3537, pp. 417–428, 2005.
- [5] J. Assfalg, A. Del Bimbo, and P. Pala, “Spin images for retrieval of 3d objects by local and global similarity,” *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)*, vol. 3, , pp. 906–909, 2004.
- [6] B. Langmann, K. Hartmann, and O. Loffeld, “Depth camera technology comparison and performance evaluation,” *ICPRAM (2)*, P. L. Carmona, J. S. Sanchez, and A. L. N. Fred, Eds, SciTePress, p. 438–444, 2012.
- [7] J. Oh, G. Kim, I. Hong, J. Park, S. Lee, J.-Y. Kim, J.-H. Woo, and H.-J. Yoo, “Low-power, real-time object-recognition processors for mobile vision systems,” *Micro IEEE*, vol. 32, no. 6, pp. 38–50, 2012.
- [8] C. Cortes and V. Vapnik, “Support-vector networks”, *Machine Learning*, springer, pp. 273–297, 1995.
- [9] K. Veropoulos, G. Bebis, and M. Webster, “Investigating the impact of face categorization on recognition performance,” *Advances in Visual Computing, LNCS*, Vol. 3804, pp 207-218, 2005
- [10] N. Brusco, M. Andretto, A. Giorgi, and G. Cortelazzo, “3d registration by textured spin-images,” *Fifth International Conference on 3-D Digital Imaging and Modeling (3DIM 2005)*, pp. 262–269, 2005
- [11] J. Nylander and H. Wallin, “Improving the performance of spin-image algorithm hardware acceleration in vhdl,” *Master’s thesis, Malardalen university, Vasteras department of computer science and electronics*, 2008.
- [12] M. Baum, D. Agnes, and D. Sven, “Using spin images for 3d object classification,” *Master’s thesis, Applied Informatics Group, Bielefeld*, 2011.

- [13] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," SIGGRAPH Comput. Graph. vol. 26, no. 2, pp. 71–78, Jul. 1992
- [14] N. Wongwaen, S. Tiendee, and C. Sinthanayothin, "Method of 3d mesh reconstruction from point cloud using elementary vector and geometry analysis," 8th International Conference on Information Science and Digital Content Technology (ICIDT), vol. 1, pp. 156–159, 2012.
- [15] L. Alboul and G. Chliveros, "A system for reconstruction from point clouds in 3d: Simplification and mesh representation," 2010 11th International Conference on Control Automation Robotics Vision (ICARCV), pp. 2301–2306, 2010.
- [16] H. Song, K. K. Choi, I. Lee, L. Zhao, and D. Lamb, "Adaptive virtual support vector machine for reliability analysis of high-dimensional problems." Structural and Multidisciplinary Optimization, vol. 47, pp. 479–491, 2013.
- [17] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," Internationalin Shape Modeling, Jun. 2004.
- [18] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multiview rgb-d object dataset," 2011 IEEE International Conference on Robotics and Automation (ICRA), pp. 1817–1824, 2011.