Proceeding of the 2nd
RSI/ISM International Conference on Robotics and Mechatronics
October 15-17, 2014, Tehran, Iran

Transformation Invariant 3D Object Recognition Based On Information Complexity

Alireza Norouzzadeh Ravari and Hamid D. Taghirad, *Senior Member, IEEE*
Advanced Robotics and Automated Systems (ARAS), Industrial Control Center of Excellence (ICCE),
Faculty of Electrical and Computer Engineering, K.N. Toosi University of Technology, Tehran, Iran.
Email: a.norouzzadeh@ee.kntu.ac.ir, taghirad@kntu.ac.ir

***Abstract*—The 3D representation of objects and scenes as a point cloud or range image has been made simple by means of sensors such as Microsoft Kinect, stereo camera or laser scanner. Various tasks, such as recognition, modeling and classification can not be performed on raw measurements because of the curse of high dimensionality, computational and algorithm complexity. Non Uniform Rational Basis Splines (NURBS) are a widely used representation technique for 3D objects in various robotics and Computer Aided Design (CAD) applications. In this paper, a similarity measurement from information theory is employed in order to recognize an object sample from a set of objects. From a NURBS model fitted to the observed point cloud, a complexity based representation is derived which is transformation invariant in the sense of Kolmogorov complexity. Experimental results on a set of 3D objects grabbed by a Kinect sensor indicates the applicability of the proposed method for object recognition tasks. Furthermore, the results of the proposed method is compared to that of some state of the art algorithms.**

I. Introduction

In mobile robotics applications, the problem of 3D object recognition has received a great amount of attention from authors and researchers. This problem is tackled by two group of techniques, i.e., model based and view based [1]. In the model based category, it is assumed that a 3D model of object is available while in the view based group, one or multiple views of an object are used as its representation. Feature based representation of 3D object are also widely used [2]. Various types of feature types are used for representation of a 3D object either as a local feature or as a global one.

Since it can not be assumed that a 3D model of each object is already in hand the model based techniques are not easily applicable. Also representation of an object from various views has unsolved problems such as specification of the salient views finding the minimum number of required scans for representation. Furthermore, it has been shown that feature based representation of 3D objects suffers from the curse of dimensionality [3].

The paper is organized as follows. The next section briefly reviews some related works. Some preliminaries about NURBS curve and surface representation and algorithmic information theory, are provided in section III. Section IV is devoted to the description of the proposed approach. Finally, section V is dedicated to the experimental results, which is followed by the concluding remarks.

II. Related Works

3D object recognition is widely concerned from various communities while it has still many unsolved problems and challenges [1]. A survey of employing local features of surfaces is presented in [2], where it has been shown that local features are efficient in object recognition and more resilient to occlusion and clutter. Some 3D shape descriptors are compared in the field of object recognition in [4] where it has been concluded that local spin image descriptor outperforms the others when Kinect range images are used. Learning based method are also another type of common object recognition methods [5] where a hierarchical feature learning method is presented for object recognition.

The problems of the curse of dimensionality and lack of a normalized similarity measurement metric are rarely considered in object recognition task from either computer vision or robotic communities. Therefore, here we use a complexity based representation of NURBS models and a normalized similarity metric from information theory and provide by experimental results that the proposed approach is applicable to the problem of 3D object recognition.

III. Preliminaries

In order to have an efficient representation of 3D objects, some conditions such as the curse of dimensionality, computational complexity and storage requirements shall be considered. Non Uniform Rational Basis Spline (NURBS) are widely used for representation of curves and surfaces in solid and deformable 3D objects [6]. A 3D object can be efficiently represented by a NURBS model suing few control points and two vectors of parameters.

Another problem of conventional 3D object recognition techniques is the lack of a proper similarity measurement. Using various distance measures from Euclidean to statistical ones are used for comparison of two objects or their corresponding models. In this paper a normalized distance measure is employed for comparison of NURBS models which is defined in the information theory [7]. Later it has been shown that a complexity based representation of NURBS model can be derived which is transformation invariant. Based on these theories, the proposed method, is expressed in section IV.

978-1-4799-6743-8/14/$31.00 ©2014 IEEE

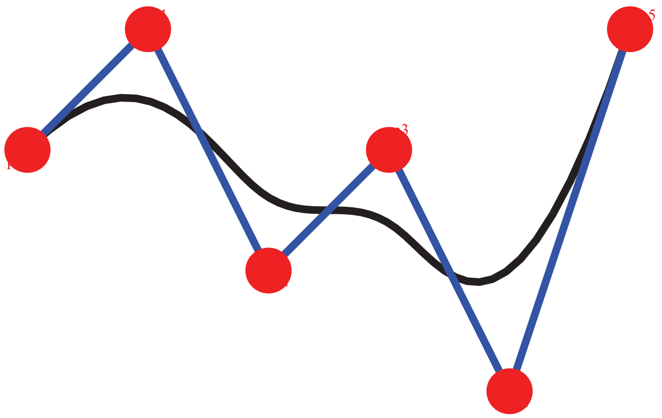

Fig. 1. A NURBS curve. $p_0, p_1, p_2, p_3, p_4, p_5$ are the control points.

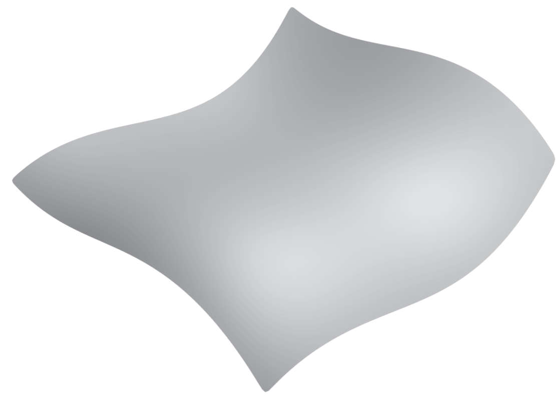

Fig. 3. A NURBS surface

A. NURBS Curves And Surfaces

In this section a brief introduction to NURBS and compressed sensing is presented. NURBS are efficient mathematical models for representation of curves and surfaces in computer aided design. The NURBS are used for modelling of automobile, ship and aircraft bodies.

Variety of 3D objects and shapes can be represented by Non Uniform Rational B-Spline (NURBS), which is a parametric geometric modelling technique. Both analytic and freeform objects can be mathematically modelled by NURBS.

A NURBS curve is defined as:

$$C(u) = \frac{\sum_{i=0}^{n} N_{i,p}(u) w_i \mathbf{P}_i}{\sum_{i=0}^{n} N_{i,p}(u) w_i} \tag{1}$$

while a NURBS surface is expressed by:

$$S(u,v) = \frac{\sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u) N_{j,q}(v) w_{i,j} \mathbf{P}_{i,j}}{\sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u) N_{j,q}(v) w_{i,j}} \tag{2}$$

where the control points are expressed by $\{\mathbf{P}_i\}$ for curves and $\{\mathbf{P}_{i,j}\}$ for surfaces. Furthermore, $\{w_i\}$ and $\{w_{i,j}\}$ are the NURBS curve and surface weights which is assumed to be real and positive.

A NURBS has m control points along u direction and n ones along v direction. The degrees of $N_{i,p}(u)$, $N_{j,q}(v)$ B-spline basis functions along parametric variables u, v is represented by p, q respectively. The basis functions are defined on the knot vectors

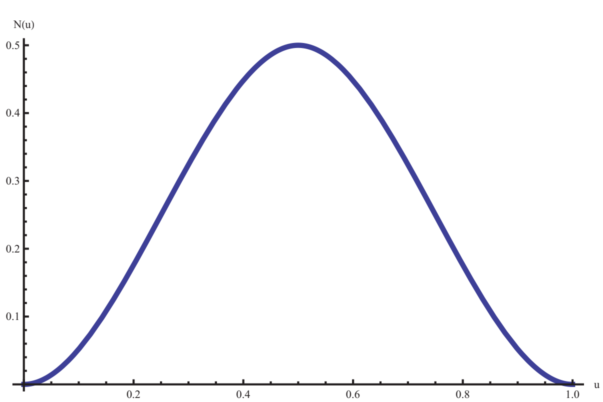

Fig. 2. A B-spline basis function

$$U = \{\underbrace{0,\ldots,0}_{p+1}, u_{p+1},\ldots,u_{r-p-1},\underbrace{1,\ldots,1}_{p+1}\} \tag{3}$$

$$V = \{\underbrace{0,\ldots,0}_{q+1}, v_{q+1},\ldots,v_{s-q-1},\underbrace{1,\ldots,1}_{q+1}\} \tag{4}$$

where $r = n+p+1$ and $s = m+q+1$. The B-spline basis functions by the CoxDe Boor formula [] is defined as:

$$N_{i,0}(u) = \begin{cases} 1 & u_i \leq u < u_{i+1} \\ 0 & otherwise \end{cases} \tag{5}$$

$$N_{i,p}(u) = \frac{u-u_i}{u_{i+p}-u_i} N_{i,p-1}(u) + \frac{u_{i+p+1}-u}{u_{i+p+1}-u_{i+1}} N_{i+1,p-1}(u) \tag{6}$$

Various methods and algorithms are available in the literature for both NURBS surface interpolation and approximation from a point cloud. In this paper we assume that the NURBS surface is fitted by a least-squares approximation.

The next section is dedicated to the Kolmogorov complexity and Normalized Compression Distance. These theories are used in development of complexity based representation of range images and are discussed in section IV.

B. Algorithmic Information Theory

The algorithmic version of information theory, estimates the information by lossless data compression which is successfully employed for content-based image retrieval [8] and feature extraction [9].

In contrast to the Shannon approach that assumes the objects are made by a known random source and represent entropy as average information, the algorithmic information theory represents objects as a symbol strings and defines the complexity in analogy to entropy. In algorithmic information theory, a string sequence X is expressed as the required input to a universal computer U which prints X on its output and stops. Also the complexity $K(X)$ is defined as the minimal length of any input for fixed U which prints X to the output.

It has been shown that the dependency of $K(X)$ to U is weak and can be ignored when $K(X)$ is sufficiently large [10]. The conditional Kolmogorov complexity is shown by $K(X|Y)$ and defined as the length of a shortest program to generate X given Y as its input.

The Kolmogorov complexity is not computable but may be approximated by a good lossless compression algorithm. Therefore, in practice the Kolmogorov complexity $K(X)$ is expressed as $C(X)$ which is the length of compressed file of X description and C is a compression algorithm. In fact, the compression algorithm estimates an upper bound for the Kolmogorov complexity. The comparison of two objects can be performed by measuring their common information. The amount of common information between two object descriptions is accomplished by the normalized compression distance metric [11]. The NCD is mathematically expressed as

$$\text{NCD}(X, Y) = \frac{C(XY) - \min\{C(X), C(Y)\}}{\max\{C(X), C(Y)\}} \quad (7)$$

where $C(XY)$ is the length of compressed file containing the concatenation of X and Y. The NCD is a metric with $\text{NCD}(X, X) = 0$ for similar string sequences and $\text{NCD}(X, Y) \leq 1$ for all pairs (X, Y). When X and Y are similar and share a great amount of information, their concatenation is compressed much more than the situation of comparing two dissimilar string sequences. Therefore, the NCD value gets close to zero. In contrast, the concatenation of two different string sequences can be compressed so much, resulting in a NCD value near to one. In order to compute NCD, any compression algorithm such as gzip, bzip2 or PPM can be used. But the block-based compression algorithms such as bzip2 fulfil the symmetry requirement of Kolmogorov complexity. The next section presents the proposed method which is constructed from complexity based representation of range images. The object classification is accomplished by comparing these representations using NCD.

IV. Proposed Method

In this section a method is developed in order to compare 3D objects by means of a normalized similarity measurement. The raw point cloud or equivalently range images are transformed into NURBS surface and then a complexity based representation is derived. The Normalized Compression Distance (NCD) from information theory is employed for measurement of the information shared by two NURBS models. Two objects are called similar, if the computed distance is lower than a threshold.

The overall process of the proposed approach is depicted in Fig. 4. The point cloud representation of a 3D object is acquired by a Microsoft Kinect, stereo camera or laser scanner. Here we consider the point clouds grabbed by a Kinect sensor. A preprocessing stage is performed on the observations to remove outliers and invalid sensor observations. Having the point cloud of 3D object in hand, a least-squares NURBS fitting method is performed to construct a NURBS model [12]. Then a complexity based representation of the NURBS model is constructed which is explained in the next section in more detail. Finally, the result of 3D object recognition is determined by comparing the similarity measurement to a threshold. For similarity distances lower then the specified threshold, the object recognition is positive and in the case of higher similarity distance compared to the threshold τ, the object is reported as unknown.

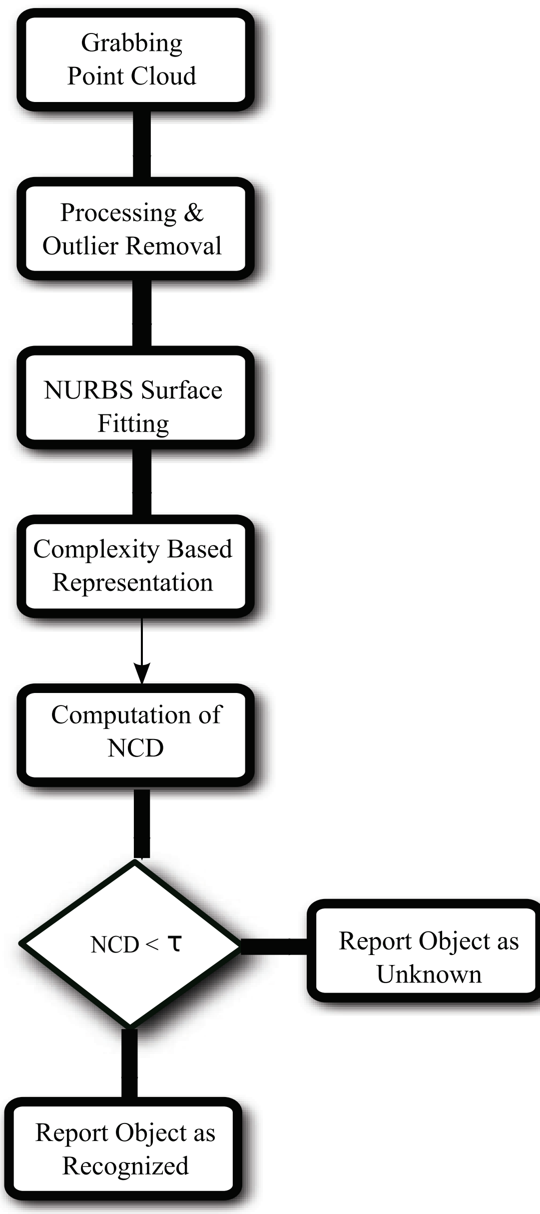

Fig. 5. The proposed method.

In what follows, the complexity based representation of NURBS is explained in more details.

A. Complexity Based Representation of NURBS Models

As mentioned before, a NURBS surface can be expressed by Eq. 2. Here we consider the case that all the NURBS surface weights are equal to one and the control points expressed in

Fig. 4. Dataset of 3D objects.

homogeneous coordinate system as $\mathbf{P}^h_{i,j}$.

$$S_n(u,v) = \sum_{i=0}^{n}\sum_{j=0}^{m} N_{i,p}(u)N_{j,q}(v)\mathbf{P}^h_{i,j} \tag{8}$$

Some properties of a NURBS surface are as follows [13].

Property 1: The NURBS surface $S_n(u,v)$ is in the convex hull of control points $\mathbf{P}_{i,j}$.

Property 2: An affine transformation is only applied to the control points when a NURBS surface is transformed.

A NURBS surface can be uniquely described by the set of control points and two vector of parametric values known as *knots*. In order to construct the complexity based representation of a NURBS surface, all numeric values corresponding to coordinates of control points or parametric values are converted to the pseudo-random binary sequences (PRBS) [14] and stored in a file. The Kolmogorov complexity of the NURBS surface is equal to the size of compressed file. A string of PRBS related to each numeric values has the length

$$l(x) = [cx] \tag{9}$$

where the operator $[\cdot]$ represents the nearest integer value and c is a constant. This process makes the Kolmogorov complexity of the equivalent PRBS, proportional to the real value [3]. Also, it has been shown that various NURBS file formats can be classified by complexity of information [15].

In order to have a transformation invariant representation, the following steps are performed in the construction of complexity based representation. The center control points is computed as

$$C_P = \frac{1}{(m+1)(n+1)}\sum_{i=0}^{n}\sum_{j=0}^{m} P_{i,j} \tag{10}$$

The coordinates of the center of control points is stored in spherical coordinates system as $C_P(\rho_c, \phi_c, \theta_c)$. The coordinates of the control points are converted to the spherical coordinates system and stored in relation to the coordinates of C_P as $P_{i,j}(\frac{\rho}{\rho_c}, \phi - \phi_c, \theta - \theta_c)$. Then the knot parameters are stored in order to have a complete description of NURBS model. From this representation the original NURBS model can be constructed without any loss of information.

The transformation invariant property can be proved by simply computing the Kolmogorov complexity of a NURBS surface and its transformed version. According to *property 2* of NURBS surface, any transformation only affects the control points and the parameter values are unchanged. Since a NURBS complexity based representation is made of the coordinates of center of control points, coordinates of control points and parameter values, its Kolmogorov complexity can be represented as

$$K_{NURBS} = K_C + K_{CP} + K_P \tag{11}$$

where K_{NURBS} is total complexity of NURBS model, K_C is the complexity of center of control points which is a constant value, K_{CP} is the complexity of control points and finally, K_P is the complexity of parameter values.

We consider a transformation in two groups. The first is rotation and translation while the scaling is considered separately. When a rotation or translation is applied to the

(a) (b) (c)

Fig. 6. Gaussian atoms with transformation parameters: (a) Food can. (b) Apple. (c) Banana.

NURBS model, the center of control points is transformed the same as the control points. Therefore, a transformation $(\rho_T, \phi_T, \theta_T)$ is only expressed in coordinates of center of control points and in the complexity based representation, the coordinate of control points is unchanged. When a scaling s_T is applied to the NURBS, the value of ρ is scaled for both control points and their center. The normalization of coordinates of control points makes the complexity of control points unchanged. Therefore, we have

$$K_{NURBS} \approx K^T_{NURBS} \tag{12}$$

which means that the Kolmogorov complexity of NURBS model is equal to the Kolmogorov complexity of its transformed version up to an additive constant.

Finally, the NURBS model of the grabbed point cloud is denoted by series of PRBS strings with relevant Kolmogorov complexity. The complexity based representations is employed in the object recognition task. Having the more number of control points and parametric values, the representation is more complex. While few control points and parameters is related to smaller compressed file size and consequently smaller complexity. In the next section the results of experiments are presented.

V. Experimental Results

In this section the result of an experiments is presented to verify the applicability and performance of the proposed method. Also, the experimental results are compared to that of a feature-based 3D object classification [16].

In this experiment a dataset of 300 objects from 51 different categories is used [17]. We randomly selected 48 objects from the same 10 categories used in [16]. The parameters of the proposed system are shown in Table I.

The point cloud of each object is resized to 20×20 and its complexity based representation is generated. Some NURBS model fitted to the objects of DataSet are shown in Fig. 6. Table II contains the system parameters of the feature-based classification approach. Where each object is subsampled with the grid size of $1cm$. The SHOTCOLOR [18] descriptor for each 3D point is computed and the mean and standard deviation of them is achieved. The result of this experiment is shown in Table III. As it can be seen, the proposed method has a competing performance with the feature-based method, in spite of using just depth information.

TABLE I
Proposed System Parameters

Parameter	Value
Similarity threshold	0.5
Point cloud size	20×20
PRBS length magnification constant	1

TABLE II
Feature-Based Algorithm Parameters

Parameter	Value
Feature type	SHOTCOLOR
Sampling method	Sub-sampling 1 cm

TABLE III
Experiment Result

Method	Accuracy
Proposed method	92.17%
Feature-based method	76.02%

VI. Conclusions

In this paper a new approach is presented for 3D object recognition from NURBS surface models, based on the algorithmic information theory. While the state-of-the-art algorithms use high-dimensional feature-based representations, here we perform the object recognition in low-dimensional space. A complexity based representation is constructed for each NURBS model. This transformation invariant representation is expressed by means of pseudo-random binary sequences. From this representation the NURBS model can be constructed without any information loss. Then from the information theory a normalized compression distance metric is employed for similarity measurement. A normalized difference matrix is generated by pairwise comparison of complexity based representations. The objects with minimum complexity based distance are recognized if the similarity distance is lower than a specified threshold. Experimental results show efficiency and accuracy of the proposed method in comparison to one recently proposed method.

References

[1] Y. Gao and Q. Dai, "View-based 3-d object retrieval: Challenges and approaches," *MultiMedia, IEEE*, vol. PP, no. 99, pp. 1–1, 2014.

[2] Y. Guo, M. Bennamoun, F. Sohel, M. Lu, and J. Wan, "3d object recognition in cluttered scenes with local surface features: A survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2014.

[3] A. Norouzzadeh Ravari and H. D. Taghirad, "Unsupervised 3d object classification from range image data by algorithmic information theory," in *Robotics and Mechatronics (ICRoM), 2013 First RSI/ISM International Conference on*. IEEE, 2013, pp. 319–324.

[4] M. As' ari, U. U. Sheikh, and E. Supriyanto, "3d shape descriptor for object recognition based on kinect-like depth image," *Image and Vision Computing*, vol. 32, no. 4, pp. 260–269, 2014.

[5] L. Bo, X. Ren, and D. Fox, "Unsupervised feature learning for RGB-D based object recognition," in *Experimental Robotics*. Springer, 2013, pp. 387–402.

[6] P. Palmer, A. Mir, and M. González-Hidalgo, "Deformable objects representation," in *Deformation Models*. Springer, 2013, pp. 3–47.

[7] T. Cover and J. Thomas, *Elements of information theory*. Wiley-Interscience, 2006.

[8] D. Cerra and M. Datcu, "A fast compression-based similarity measure with applications to content-based image retrieval," *Journal of Visual Communication and Image Representation*, 2011.

[9] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 8, pp. 1226–1238, 2005.

[10] M. Li and P. Vitanyi, *An introduction to Kolmogorov complexity and its applications*. Springer-Verlag New York Inc, 2008.

[11] M. Li, X. Chen, X. Li, B. Ma, and P. Vitányi, "The similarity metric," *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3250–3264, 2004.

[12] D. Brujic, I. Ainsworth, and M. Ristic, "Fast and accurate nurbs fitting for reverse engineering," *The International Journal of Advanced Manufacturing Technology*, vol. 54, no. 5-8, pp. 691–700, 2011.

[13] L. Piegl and W. Tiller, "The nurbs book. 2nd," 1997.

[14] A. Tan and K. Godfrey, "The generation of binary and near-binary pseudorandom signals: an overview," *Instrumentation and Measurement, IEEE Transactions on*, vol. 51, no. 4, pp. 583–588, 2002.

[15] V. A. Cicirello and W. C. Regli, "A flexible and extensible approach to automated cad/cam format classification," *Computers & Graphics*, vol. 37, no. 5, pp. 484–495, 2013.

[16] L. Alexandre, "3d descriptors for object and category recognition: a comparative evaluation," in *Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vilamoura, Portugal*, 2012.

[17] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multi-view rgb-d object dataset," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1817–1824.

[18] F. Tombari, S. Salti, and L. Di Stefano, "A combined texture-shape descriptor for enhanced 3d feature matching," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 809–812.