Proceeding of the 2013
RSI/ISM International Conference on Robotics and Mechatronics
February 13-15, 2013, Tehran, Iran

Unsupervised 3D Object Classification From Range Image Data By Algorithmic Information Theory

Alireza Norouzzadeh Ravari, *Student Member, IEEE,* 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 problem of unsupervised classification of 3D objects from depth information is investigated in this paper. The range images are represented efficiently as sensor observations. Considering the high-dimensionality of 3D object classification, little attention has been paid to the curse of dimensionality in the existing state-of-the-art algorithms. In order to remedy this problem, a low-dimensional representation is defined here. The sparse model of every range image is constructed from a parametric dictionary. Employing the algorithmic information theory, a universal normalized metric is used for comparison of Kolmogorov complexity based representations of sparse models. Finally, most similar objects are grouped together. Experimental results show efficiency and accuracy of the proposed method in comparison to a recently proposed method.**

I. Introduction

The problem of object detection has received a great amount of attention from computer vision and pattern recognition communities. The visual object detection is performed using feature-based representation of object images. Distinctive properties of objects extracted from images, such as shape, color and textures are employed in visual object detection. As the world is 3D in nature, the depth information should be used in the object detection algorithms. The 3D scans are made available as the observation of sensors such as stereo camera, Lidar or Microsoft Kinect.

Various descriptors are presented for representation of color, shape and depth information in the context of 3D object recognition. A shape descriptor is presented in [1], where an ensemble of angle, area and distance shape functions are employed in construction of an object descriptor. The depth information is used in [2] in order to construct a depth kernel descriptor where models the size, 3D shape and depth edges in a single framework. Another shape descriptor is expressed in [1] for classification of 3D objects observed by a Kinect camera using a database of 3D models.

Expressing 3D objects by means of descriptors, results in a high-dimensional representation which suffers from the so called curse of dimensionality problem [3]. In order to remedy this problems, in this paper, a proper 3D object classification method is presented. The proposed approach is developed such that it works in a low-dimensional space, based on range image data without training or derivation of distinctive characteristics from 3D scans. The paper is organized as follows. The next section briefly reviews some related works. Some preliminaries about sparse modeling of images based on a parametric dictionary and algorithmic information theory, are provided in section II . Section III is devoted to the description of the proposed approach. Finally, section IV is dedicated to the experimental results, which is followed by the concluding remarks.

II. Preliminaries

In order to efficiently classify 3D objects, a similarity measure from low-dimensional representation of rage image observations is necessary. Two theories are employed in the process of definition of this similarity measure in a low dimensional space. The first is the sparse representation of data based on a dictionary. The sparse modeling of 2D images, generates a parametric representation of a natural 2D image, which is unique in an over-complete dictionary [4]. This approach achieves a very compact and efficient representation of salient features of a natural image. Here we apply this method to achieve a sparse representation for each range image as a linear combination of parametric functions called atoms.

In order to accomplish the object classification task, a proper similarity measure is required. A normalized distance measure is developed in algorithmic information theory [5], which compares general objects based on the complexity of their representations. In the following sections, the sparse modeling of images and normalized distance measure are presented in more detail. Based on these theories, the proposed object classification method, is elaborated in section III.

A. Sparse Modeling

The representation of 2D images as a sparse model based on a parametric dictionary has received great amount of attention from image processing community [4]. Sparse and compact representation of an image using few atoms of an over-complete dictionary and also the flexibility of dictionary design, are some benefits of sparse modeling of images. A parametric mother function is employed for generation of dictionary atoms with a combination of 2D transformations such as translation, rotation and non-uniform scaling applied. Finally an iterative matching pursuit algorithm, [6] represents the input image approximately.The decomposition of image is expressed by a linear combination of most correlating atoms.

978-1-4673-5811-8/13/$31.00

A dictionary $\mathcal{D}$ is a set of unit norm parametric functions (atoms), constructed from a generating function denoted by $g(x,y)$. This function is defined in Hilbert space $\mathcal{H}$ with 2D transformations applied. The transformations include translation, rotation and anisotropic scaling such that the dictionary $\mathcal{D}$ spans the Hilbert space $\mathcal{H}$.

In this paper a Gaussian generating function is chosen for the construction of dictionary atoms.Representing the image pixel coordinates by x and y, the Gaussian mother function is expressed as

$$g(x,y) = \frac{1}{k} e^{-(x^2+y^2)} \tag{1}$$

where k is a normalization factor such that every atom has unit norm. This Gaussian mother function is employed practically for approximate representation of natural images [4].

In order to efficiently capture 2D features of natural images, a finite set of M transformations is applied to the generating function $g(x,y)$ to construct a parametric dictionary. Translation $\mathbf{\Delta}_i$, non-uniform scaling $\mathbf{S}_i$ along image x and y axis, and rotation $\mathbf{R}_i$ are the transformations used for the construction of the parametric dictionary. These transformations can be expressed in homogeneous coordinates system as

$$\mathbf{\Delta}_i = \begin{bmatrix} 1 & 0 & \delta_{ix} \\ 0 & 1 & \delta_{iy} \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

$$\mathbf{S}_i = \begin{bmatrix} s_{ix} & 0 & 0 \\ 0 & s_{iy} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$\mathbf{R}_{\theta_i} = \begin{bmatrix} \cos\theta_i & -\sin\theta_i & 0 \\ \sin\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4}$$

The transformation parameters of each atom are represented by a 5-tuple $\gamma_i = (s_{ix}, s_{iy}, \theta_i, \delta_{ix}, \delta_{iy})$ denoting anisotropic scaling, rotation and translation along x and y axis, respectively. The corresponding affine transformation of γ_i can be expressed as

$$\mathbf{T}_a(\gamma_i) = \mathbf{\Delta}_i \mathbf{R}_{\theta_i} \mathbf{S}_i \tag{5}$$

Representing the set of M affine transformations $\mathbf{T}_a(\gamma_i)$ as Γ_a, we have

$$\Gamma_a = \{\mathbf{T}_a(\gamma_i) | \mathbf{T}_a(\gamma_i) = \mathbf{\Delta}_i \mathbf{R}_{\theta_i} \mathbf{S}_i, i = 0, \dots M-1\} \tag{6}$$

Finally, the parametric dictionary, consisting of M atoms is constructed by transforming the generating function $g(x,y)$ according to γ_i, resulting in the ith atom which is expressed by $g_{\gamma_i}(x,y)$.

$$g_{\gamma_i}(x,y) = \frac{1}{\tilde{k}} \mathbf{T}_a(\gamma_i) g(x,y) \tag{7}$$

where $\tilde{k}$ is a a normalization factor.

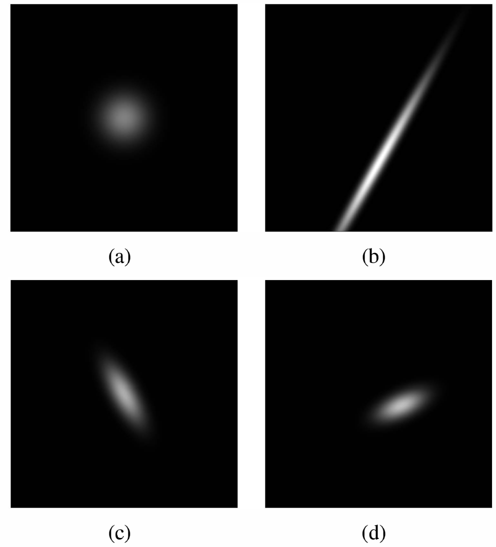


Fig. 1. Gaussian atoms with various 2D translation, rotation and scaling applied: (a) $t_x = 0, t_y = 0, \theta = 0, s_x = 1, s_y = 1$. (b) $t_x = 0, t_y = 2, \theta = \frac{\pi}{6}, s_x = 0.25, s_y = 3$. (c) $t_x = 0, t_y = 0, \theta = \frac{-\pi}{6}, s_x = 1, s_y = 2$. (d) $t_x = 1, t_y = 0, \theta = \frac{-\pi}{4}, s_x = 2, s_y = 0.5$

The over-complete parametric dictionary g_{γ_i} can be expressed as

$$\mathcal{D} = \{g_{\gamma_i} | g_{\gamma_i} = \mathbf{T}_a(\gamma_i) g, \mathbf{T}_a(\gamma_i) \in \Gamma_a\} \\ , ||g_{\gamma_i}|| = 1, i = 0 \dots M-1 \tag{8}$$

Therefore, the only limitation to design such dictionary is spanning of whole Hilbert space of input images, while the generating function should be able to capture input image structure and salient features [4]. Some dictionary atoms are depicted in Fig. 1 indicating a generating function, that has taken corresponding 2D transformations. In practice, the continuous 2D transformation space is quantized and the number of dictionary atoms M is chosen such that the spanning constraint is satisfied.

Even though the sparse approximation of images from an over-complete dictionary is an NP-hard problem [7], greedy algorithms find sub-optimal but yet efficient solutions, iteratively. One of the widely used greedy algorithms is the matching pursuit [6], in which, at every iteration the best matching atom is found by full dictionary searching. The matching pursuit converges exponentially, however, it can not find the sparsest solution [4]. Therefore, the Orthogonal Matching Pursuit (OMP) algorithm [7] is used in this paper which solves the problem of finding optimal sparse solution [8], [9].

The OMP algorithm, initially assigns the input image I_s to the residual R_0.

$$R_0 = I_s \tag{9}$$

Then iteratively at the ith step, OMP seeks the best matching atom g_{γ_i} by finding the atom which possess maximum correlation with the residual R_{i-1}.

$$\gamma_i = arg \max_{\gamma} |\langle R_{i-1}, g_\gamma \rangle| \tag{10}$$

In this relation the defined inner product in Hilbert space $\mathcal{H}$ is denoted by the operator $\langle \cdot, \cdot \rangle$.

The contribution of the selected atom is removed from the residual by orthogonal projection of R_{i-1} on to the span of selected atoms $\{g_{\gamma_i}\}$, where

$$R_i = (I - P_i)R_{i-1} \tag{11}$$

represents the orthogonal projection of $span\{g_{\gamma_i}\}$ by P_i. After N iterations, the input image I_s is expressed by a linear combination of the selected atoms.

$$I_s = \sum_{i=0}^{N-1} \langle R_i, g_{\gamma_i} \rangle g_{\gamma_i} + R_N \tag{12}$$

It is observed that the approximation error decays exponentially and the algorithm is terminated after N steps to represent the input image as a sparse model or until the norm of the residual becomes lower than a specified threshold.

After N iterations, the OMP algorithm represents the input image approximately as a linear combination of most correlating atoms. The approximate linear expansion of input image I_s expressed by

$$I_s \approx \sum_{i=0}^{N-1} \langle R_i, g_{\gamma_i} \rangle g_{\gamma_i} = \sum_{i=0}^{N-1} \zeta_i g_{\gamma_i} \tag{13}$$

is an efficient unique image representation in a low dimensional space which has application in image and video coding [10], and image transformation estimation [11]. The approximate spare model captures the salient geometrical features of input image with few atoms of a parametric dictionary.

The extracted sparse models of range images shall be compared for finding similar objects which is the main purpose of this paper. 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 III.

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 [12] and feature extraction [13].

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 [14]. 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 [15]. The NCD is mathematically expressed as

$$\text{NCD}(X, Y) = \frac{C(XY) - \min\{C(X), C(Y)\}}{\max\{C(X), C(Y)\}} \tag{14}$$

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.

III. Proposed System

In this section a similarity measurement approach is developed for 3D object classification from range images. As it is shown in the flowchart given in Fig. 2, range images are acquired sequentially as sensor observations. Then, a sparse model for each range image is constructed iteratively, from a parametric dictionary using orthogonal matching pursuit algorithm. The parametric dictionary is generated offline, containing parametric Gaussian functions known as atoms. The result of this step is the sparse model of range image, composed of a linear combination of atoms.

In order to use the NCD as a normalized similarity measure, a representation is constructed from the range image sparse model, involving its structure complexity. The Kolmogorov complexity of this representation is proportional to the number of atoms used in sparse model and their geometrical structure.

In what follows, different parts of the proposed method is explained in more details.

A. Complexity Based Representation of Sparse Models

The normalized compression distance (NCD) is a universal metric for comparison of general object descriptions such as

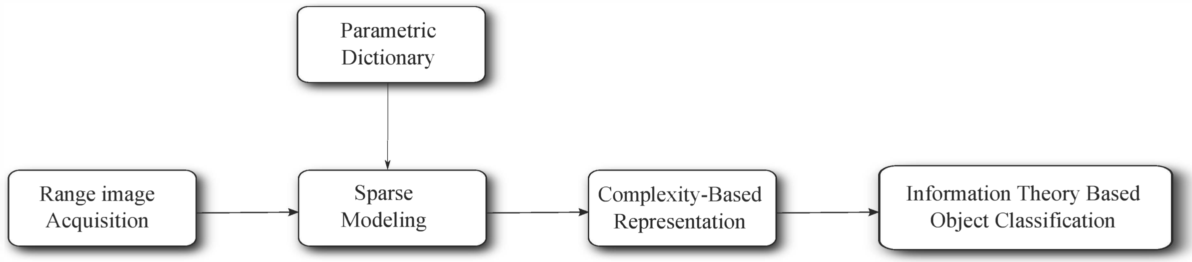


Fig. 2. The overall processing units of the proposed object classification approach.

music and genome data [16]. Unfortunately, the application of NCD in image similarity measurement is not satisfactory [17], where it has been shown experimentally that NCD can not be universally applied to the images. In order to remedy this problem, we propose to use a mapping between numbers and string patterns with proportional complexity. This is done by denoting a real value by means of a proper pseudo-random binary sequences (PRBS) [18]. Therefore the Kolmogorov complexity of the equivalent PRBS becomes proportional to the real value.

A numerical example is provided here to show why direct usage of real values for representing complexity of an object is inappropriate and how the Kolmogorov complexity of a PRBS string behaves in proportion to its length. Here a positive integer variable x is considered. In conventional methods, a string representation of x is constructed by storing its digits in a file one by one. Then the Kolmogorov complexity of x expressed by $K(x)$ is approximated by computing the length of the compressed file. The PRBS based representation of x is achieved by using it as initial value and generating a PRBS sequence with a length equal to x, denoted by x_{prbs}. In Fig. 3, the Kolmogorov complexity of a sequence of numbers and their corresponding pseudo-random binary sequence (PRBS) are depicted. The Kolmogorov complexity is approximated by bzip2 as compression algorithm. The complexity of x_{prbs} is approximately a linear function of the number itself, as it is shown in Fig. 3(a). A PRBS can not be compressed that much, due to its inherent random nature. However, a numeric representation has a low Kolmogorov complexity so it is compressed, as shown in Fig. 3(b).

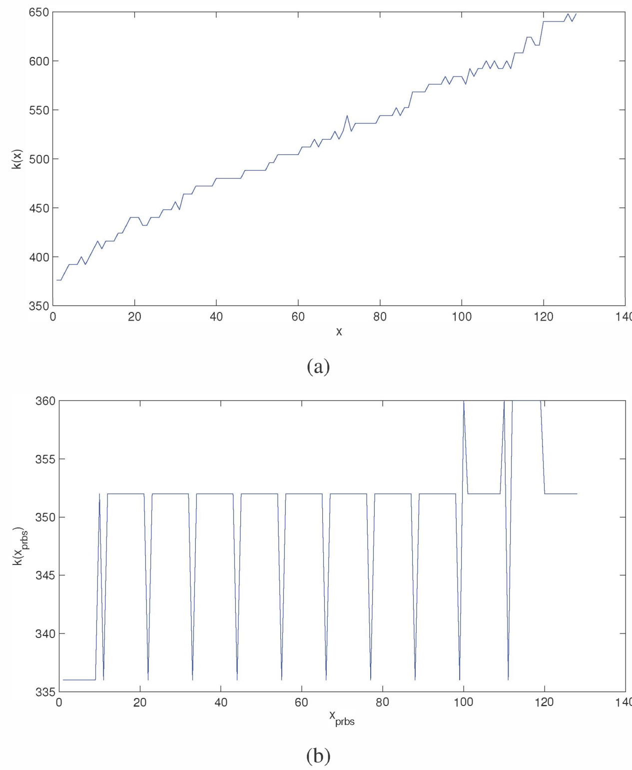


Fig. 3. Kolmogorov complexity of numbers and their corresponding PRBS: (a) The Kolmogorov complexity of a PRBS is proportional to real value denoted by x_{prbs}. (b) The Kolmogorov complexity of numbers is not proportional to the number of sequence x.

Therefore, PRBS strings are employed to construct a complexity based representation from sparse model of range images. In order to construct this representation, a range image sparse model with N atoms is considered. This sparse model can be uniquely expressed by describing its atoms with their corresponding gain values. In the complexity based representation, each atom is described by a proper PRBS string which indicates the atom gain and its transformation parameters. Since a large value is mapped to a corresponding large Kolmogorov complexity value, the structure complexity of range image sparse model is encoded in this representation.

Here, mapping of real values to a PRBS string is accomplished by generating a PRBS string with a length defined as

$$l(x) = [cx] \tag{15}$$

where the operator $[\cdot]$ represents the nearest integer value and c is a constant.

Therefore, the string representation of range image sparse model is converted to the Kolmogorov complexity based representation by substitution of every real value by its corresponding PRBS string.

Each PRBS is constructed by using the real value $r \in \mathcal{R}$, which can be an atom gain or a transformation parameter, as initial value and generating a sequence of $l(r)$ binary bits with the help of linear feedback shift register (LFSR) method [18].

This representation encodes all information required for construction of range image from its sparse model into patterns with proportional complexity.

Here we present an example that better illustrates the main idea of the proposed method. The progressive reconstruction of a range image is depicted in Fig. 4. The input range image is shown in Fig. 4(a), where its related color image can be seen in Fig. 4(b). The OMP algorithm iteratively finds the best matching atom from dictionary and construct an approximation of the input image. An approximate representation of input range image with 20 atoms is shown in Fig. 4(c). This representation can be expressed by a set of parameters λ_i where $i = 1, \ldots, 20$. Replacing each parameter value with its corresponding PRBS, the complexity based representation is achieved and stored in a file. The Kolmogorov complexity of this PRBS-based representation is practically computed by measuring the size of the compressed file. Here the bzip2 compressor is used and the Kolmogorov complexity of the PRBS-based representation is computed as $k(I_s) = 812$, where the file size is expressed in bytes. The same process is repeated in Fig. 4(d) up to Fig. 4(g). It can be seen that with increasing the number of atoms, more detailed reconstruction of range image is achieved. Also, increasing the number of atoms, results in a more complicated representation and therefore, the Kolmogorov complexity of representation increases. Here representation of the input range image by 80, 140, 200 and 280 atoms, results in Kolmogorov complexity of 2663, 3925, 5213 and 6551, respectively.

Therefore, the sparse model of each range image can be expressed by a PRBS based representation with appropriate Kolmogorov complexity. Finally, the NCD can be efficiently applied on the constructed complexity based representations in order to perform classification task. The next section presents the experimental results of 3D object classification using the proposed approach.

IV. 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 [19].

In this experiment a dataset of 300 objects from 51 different categories in used [20]. We randomly selected 48 objects from the same 10 categories used in [19]. In order to make a fair comparison, no training is accomplished and objects are classified based on minimum distance. The parameters of the proposed system are shown in Table I. The range images of each object is resized to 100×100 and its complexity based representation is generated. 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 [21] 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
Number of scales	3
Number of rotations	5
Maximum residual norm	0.1
Range image resolution	100×100
PRBS length magnification constant	1

TABLE II
Feature-Based Algorithm Parameters

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

V. Conclusions

In this paper a new approach is presented for 3D object classification from range images, based on the sparse modeling of images and algorithmic information theory. While the state-of-the-art algorithms use high-dimensional feature-based representations, here we perform the object classification in low-dimensional space. A sparse representation for every captured range image is constructed using a parametric dictionary. Then a complexity based representation is generated from the range image sparse model by means of pseudo-random binary sequences. From this representation the sparse model of range image 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 classified in the same group. Experimental results show efficiency and accuracy of the proposed method in comparison to one recently proposed method.

References

[1] W. Wohlkinger and M. Vincze, "Ensemble of shape functions for 3d object classification," in *Robotics and Biomimetics (ROBIO), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2987–2992.

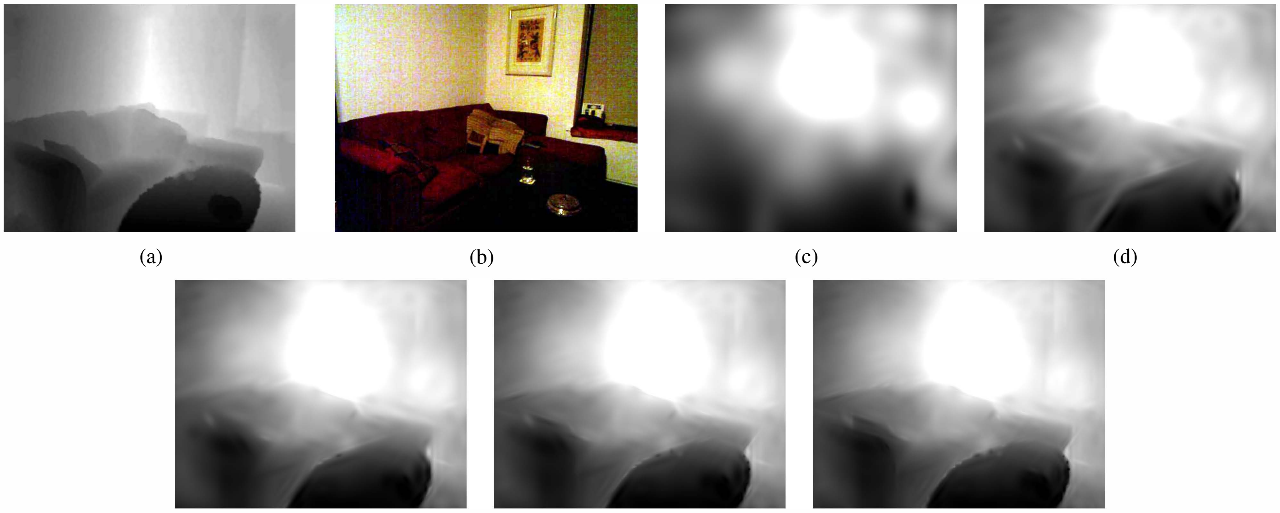
[2] L. Bo, X. Ren, and D. Fox, "Depth kernel descriptors for object recognition," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 821–826.

[3] M. Radovanović, A. Nanopoulos, and M. Ivanović, "Hubs in space: Popular nearest neighbors in high-dimensional data," *The Journal of Machine Learning Research*, vol. 9999, pp. 2487–2531, 2010.

[4] R. Figueras i Ventura, P. Vandergheynst, and P. Frossard, "Low-rate and flexible image coding with redundant representations," *Image Processing, IEEE Transactions on*, vol. 15, no. 3, pp. 726–739, 2006.

TABLE III
Experiment Result

Method	Accuracy
Proposed method	87.36%
Feature-based method	76.02%



(a) (b) (c) (d) (e) (f) (g)

Fig. 4. Progressive reconstruction of a range image. (a) Original range image (I_s). (b) Colour image. Reconstructed range images with (c) 20 atoms, $k(I_s) = 812$, (d) 80 atoms, $k(I_s) = 2663$, (e) 140 atoms, $k(I_s) = 3925$, (f) 200 atoms, $k(I_s) = 5213$ and (g) 280 atoms, $k(I_s) = 6551$.

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

[6] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3397–3415, 1993.

[7] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constructive approximation*, vol. 13, no. 1, pp. 57–98, 1997.

[8] E. Livshitz, "On the optimality of the orthogonal greedy algorithm for μ-coherent dictionaries," *Journal of Approximation Theory*, vol. 164, no. 5, pp. 668–681, 2012.

[9] M. Davenport and M. Wakin, "Analysis of orthogonal matching pursuit using the restricted isometry property," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4395–4401, 2010.

[10] A. Rahmoune, P. Vandergheynst, and P. Frossard, "Sparse approximation using m-term pursuit and application in image and video coding," *Image Processing, IEEE Transactions on*, vol. 21, no. 4, pp. 1950 –1962, april 2012.

[11] E. Kokiopoulou, D. Kressner, and P. Frossard, "Optimal image alignment with random projections of manifolds: Algorithm and geometric analysis," *Image Processing, IEEE Transactions on*, vol. 20, no. 6, pp. 1543 –1557, june 2011.

[12] 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.

[13] 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.

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

[15] 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.

[16] R. Cilibrasi and P. Vitanyi, "Clustering by compression," *Information Theory, IEEE Transactions on*, vol. 51, no. 4, pp. 1523 – 1545, april 2005.

[17] P. Vázquez and J. Marco, "Using normalized compression distance for image similarity measurement: an experimental study," *The Visual Computer*, pp. 1–22, 2011.

[18] 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.

[19] 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.

[20] 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.

[21] 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.