Information Sciences 510 (2020) 33–49

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On shortened 3D local binary descriptors

Siwen Quan[a], Jie Ma[b,*]

[a] School of Electronics and Control, Chang'an University, Xi'an 710064, China
[b] National Key Laboratory of Science and Technology on Multi-spectral Information Processing, School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China

A R T I C L E I N F O

Article history:
Received 23 March 2019
Revised 10 September 2019
Accepted 14 September 2019
Available online 14 September 2019

Keywords:
3D point cloud
Local feature description
Bit-selection
Feature matching
Binary descriptor

A B S T R A C T

The wide-spread mobile systems nowadays desire ultra lightweight local geometric features to accomplish tasks relying on correspondences. Nonetheless, most existing 3D local feature descriptors, though shown to be distinctive and robust, still are real-valued and/or high-dimensional. Accordingly, this paper conducts a comparative study on current bit-selection methods with a focus on shortening 3D local binary descriptors. By analyzing several bit-selection techniques, we develop and evaluate various approaches to obtain a shortened version of a state-of-the-art feature remaining discriminative and robust. Through extensive experiments on four standard datasets with different data modalities (e.g., LiDAR and Kinect) and application scenarios (e.g., 3D object retrieval, 3D object recognition, and point cloud registration), we show that a small subset of representative bits are sufficient to achieve promising feature matching results as the initial descriptor. Moreover, the shortened binary descriptors still hold competitive or better distinctiveness and robustness compared to several state-of-the-art real-valued descriptors, e.g., spin image, SHOT, and RoPS, albeit being dramatically more efficient to match and store. Key to the foreseen research trend of local geometric feature description is dealing with compact binary descriptors; thus, our work may pave the way for this new research direction.

1. Introduction

Local shape description for rigid data, such as 3D point clouds and meshes is a pervasive problem in the realm of computer vision, computer graphics, and robotics. The objective is to use a feature vector to fully parameterize the geometric information contained in a local shape [44]. It has been applied to numerous real-world applications, e.g., 3D object recognition, point cloud registration, simultaneous localization and mapping (SLAM), and reconstruction.

A large corpus of research on 3D local shape descriptors has been conducted in the past two decades [13,16,18,32,38,44,45,48]. The existing 3D local descriptors are either hand-crafted or learned. According to the taxonomy in [37], hand-crafted features can be categorized into three classes: histogram, signature, and hybrid. Histogram-type descriptors, e.g., fast point feature histograms (FPFH) [32] and local feature statistics histograms (LFSH) [44], represent the local shape geometry by calculating the statistical histogram of point attributes, such as normal deviation and curvatures. Signature-type descriptors, e.g., 3D shape context (3DSC) [4], mainly resort to the spatial distribution of point attributes for shape description. Hybrid-type descriptors combine the traits of the former two categories to achieve a good balance

* Corresponding author.
E-mail addresses: siwenquan@hust.edu.cn (S. Quan), majie@mail.hust.edu.cn (J. Ma).

https://doi.org/10.1016/j.ins.2019.09.028
0020-0255/© 2019 Published by Elsevier Inc.

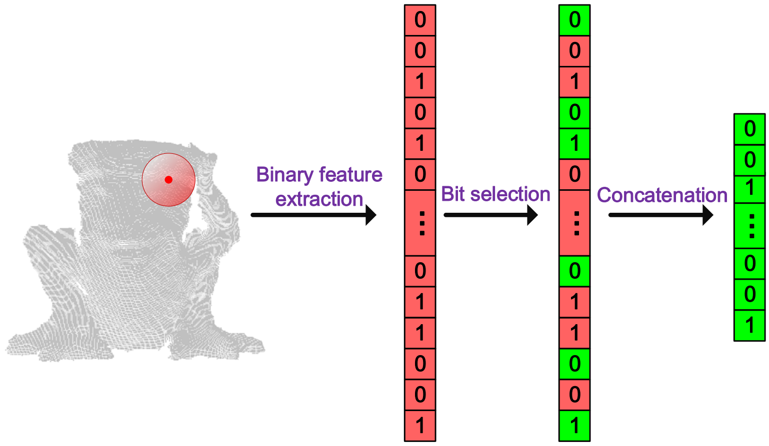

Fig. 1. Illustration of bit-selection on 3D binary descriptors. Binary descriptors are first extracted for keypoints detected on a 2.5D/3D model. After performing bit-selection on the initial bits (red), reasonable bits (green) are concatenated as a new descriptor. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

between descriptiveness and robustness, with typical examples including signature of histograms of orientations (SHOT) [38] and rotational projection statistics (RoPS) [13]. Learned features, e.g., 3DMatch [48] and CGF [18], have achieved a clear performance gain over hand-crafted ones when trained on large datasets. These works have made great progress toward crafting discriminative and robust geometric features. However, most existing 3D local features are real-valued and high-dimensional, and incur low efficiency in both storage and feature matching. With the rise of portable and commodity range sensors, e.g., the Microsoft Kinect and Intel RealSense, there is a desperate need for ultra-light-weight feature descriptors.

Similar to the trace in 2D the image domain [6,31], designing binary descriptors recently has become a new trend in 3D the domain. Compared with float strings, bit strings greatly reduce the storage volume and time cost of feature matching. Typical examples of 3D binary features include B-SHOT [27] and local voxelized structure (LoVS) [29]. B-SHOT converts the SHOT descriptor using a quantization algorithm, and is shown to hold limited descriptiveness due to information loss [29]. LoVS achieves superior performance when benchmarked on a series of datasets against several real-valued descriptors [29]. Nonetheless, LoVS, which is composed of 729 bits, suffers from low efficiency in terms of storage and feature matching. Although 3D local descriptors have been studied since the 1990s, the development of binary descriptors began in 2015 [27] and existing binary descriptors are relatively high-dimensional [29]. We notice that in the 2D image domain, effective feature description can be achieved with as few as 32 bits, i.e., the D-BRIEF descriptor [42]. This motivates us to explore the possibility of crafting low-dimensional 3D local binary descriptors.

Dimension reduction is a common practice for achieving lower-dimensional descriptors. However, this is mostly performed on float feature descriptors. Popular dimension reduction approaches include principal component analysis (PCA) [16] and linear discriminant embedding (LDE) [15]. However they are not suitable for binary descriptors because the resulting descriptors are real-valued. There are also methods that convert real-valued descriptors to binary descriptors, to shorten a descriptor, e.g., thresholding [2] and non-linear neighborhood component analysis [40]. Yet, these methods require float descriptors as the input. For the problem on which this paper focuses, we wish to select a few reasonable bits from a initial binary descriptor to achieve the goal of shortening binary descriptors with few discriminative power loss. We notice the "binary test selection" technique in the 2D image domain [31,46,47] and borrow ideas from it. The definition of the binary test is comparing the values of two image pixels and labeling the result with '0' or '1' [31,47]. Binary test selection then refers to as the selection of a portion of all binary tests within a local image patch. Because binary test selection allows *binary in and binary out*, we apply it to shortening binary feature descriptors and dub it as *bit-selection* in our context. The distinction is that we perform treat feature bins as input instead of two image pixels.

This paper, to the best of our knowledge, presents the first attempt to compute very low-dimensional 3D local binary descriptors. Directly computing compact binary descriptors from raw data relies heavily on choosing salient pixels or feature bins, which are somewhat subjective [47]. A more prevalent solution prefers to first use a high-dimensional vector to fully characterize the information and then perform dimension reduction [15,31]. Similar to existing works that build advanced variants over existing matured features, e.g., PCA-SIFT over SIFT [17] and compressed SHOT over SHOT [22], we design distinctive, robust, and ultra-lightweight binary features based upon a state-of-the-art 3D descriptor, i.e., LoVS [29]. LoVS is a binary descriptor and exhibits competitive performance with existing real-valued features. For a given application task, we believe that not all bits provide positive contributions for feature matching. For instance, when performing 3D registration for partially overlapped data, bits extracted spatially closer to the keypoints are supposed to be more convincing, as boundary regions are usually incomplete [45]. To this end, we perform bit-selection on LoVS. To provide a better understanding of bit-selection, we illustrate its concept in Fig. 1. This work specifically studies eight bit-selection algorithms: two baselines, i.e., random sampling and uniform sampling, and six published ones, i.e., *Correlation* [31], *Entropy* [46], *Coding* [30], *Matching*