ISPRS Journal of Photogrammetry and Remote Sensing 144 (2018) 105–118

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Super resolution of laser range data based on image-guided fusion and dense matching

Xu Huang[a], Rongjun Qin[a,b,*], Changlin Xiao[a], Xiaohu Lu[a]

[a] Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, USA
[b] Department of Electrical and Computer Engineering, The Ohio State University, 205 Dreese Labs, 2015 Neil Avenue, Columbus, OH 43210, USA

ARTICLE INFO

Keywords:
Super resolution
Laser range data
Image
Fusion
Matching

ABSTRACT

Low-cost/real-time laser range scanner is becoming one of the dominant tools in acquiring accurate 3D point clouds for many smart applications (e.g. automated driving), while the low point density is often the limiting factor for acquiring fine-scale information. On the other hand, stereo/multi-stereo images offer considerably higher-resolution 3D data with low cost, while the image-derived point clouds generally have a higher level of uncertainty. In order to generate accurate, dense point clouds at a low cost, this paper explores a complementary data fusion of the low-resolution-high-accuracy laser range data and the high-resolution (high-res) images, and proposes a super resolution method of laser range data through a novel dense matching framework. In general, we formulate the super resolution as maximizing a posteriori - Markov random field (MAP-MRF) problem in a constrained matching framework, where a two-step strategy is introduced to remove partial inconsistencies between laser range data and images, and the confidence of the high accuracy laser points are propagated through a uniquely designed path in the high-resolution image space, such that a global dense matching algorithm can be externally constrained to yield an accurate, dense and high-fidelity point clouds. We compared the experiment results of the proposed method with the original laser range data and other two super resolution methods of laser range data under aerial, terrestrial, and indoor scenarios. These all demonstrate that the proposed method is capable of producing sub-pixel accuracy, high-fidelity point clouds, even though the density of laser range data is considerably low (hundreds of times lower than the image resolutions).

1. Introduction

Laser range scanner or Light Detection and Ranging (LiDAR), is becoming one of the dominant tools in acquiring 3D information. By actively measuring distance to a surface through time of flight of laser beams in a high-frequency, the laser range scanner is able to collect a vast amount of individually measured high-accuracy point clouds or laser range data. Its characteristics of being robust and accurate have fueled many smart applications including automated mapping, autonomous vehicles, virtual reality, cultural heritage, robotic navigations, etc.

Though promising, the cost and payload for high quality laser range scanner are still relatively high for wide deployment in these applications. For example, the premium laser scanners (e.g. Velodyne HDL-64E, RIEGL VZ-400) being able to generate millions of points per second cost nearly $80,000 with a payload approximating 12 kg, while such a device limits their applications in civilian mobile platforms, e.g. phone, car, drones. Although there exist low-cost (hundreds to thousands of dollars) and light-weight (dozens to hundreds of grams) laser range scanners such as Canesta EP Devkit, Swiss Ranger SR4500, and RealSense SR300, such sensors are only able to collect sparse point clouds (Yang et al., 2007) and are applicable to certain applications such as robotic object avoidance, detection of moving objects and shape completion (Wu et al., 2015), while given the resolution limits (average dozens of thousands of points per second), it is not suitable for several high-level vision tasks such as extraction of sub-building level objects, e.g. cars, small architectural features.

Most of the low-cost laser range scanners are equipped with a high-resolution (high-res) camera whose image resolution or pixel density is actually dozens or even hundreds of times higher than laser range scanner. However, the high-res images along with these low-cost laser range scanners were normally used to map textures for generating low-resolution (low-res) models. A largely ignored fact is that these ready-to-use high-res images are able to form stereo pairs that can be potentially used for generating low-cost, high-density point clouds through dense image matching (DIM) techniques, thus to enhance the

* Corresponding author at: 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43210, USA.
E-mail address: qin.324@osu.edu (R. Qin).

https://doi.org/10.1016/j.isprsjprs.2018.07.001
Received 22 January 2018; Received in revised form 27 April 2018; Accepted 4 July 2018

resolution of the output for such low-cost system. On the other hand, it is generally known that DIM techniques are scene dependent and the accuracy of the 3D point clouds strongly depends on the resulting image geo-referencing, scene environment, and object types. Hence, point clouds generated from DIM techniques often contain high uncertainties. For instance, image-based point clouds through DIM are normally problematic in regions with high directional reflectivity and poor textures, such as water surfaces, snowfields. Additionally, the state-of-the-art DIM methods are generally not able to handle objects with complex structures, such as fences, wire frames or vegetation brunches. In these scenarios, a laser range scanner performs relatively well, particularly in poor-texture scenario.

Therefore, an optimal solution of generating low-cost, dense, and accurate point clouds requires the complementary data fusion of both data sources for super resolution. This, moreover, yields textured point clouds that can be used for more advanced tasks such as machine learning, object extraction, semantic recognition, and fine-scale reconstruction. This paper particularly addresses this data fusion and super resolution problem by forming it through a novel dense matching framework: we presume the availability of both low-res laser range data and geo-referenced high-resolution images and develop a laser range data super-resolution framework that is able to (1) account for partial inconsistency between both data sources, e.g. temporal changes between the acquisitions, (2) fully utilize the information of laser range data through newly devised non-local paths and most importantly (3) produce high-resolution, high-fidelity texturized point clouds without additional requirement for hardware input. The core algorithm formulates the super resolution as maximum a posteriori - Markov random field (MAP-MRF) problem, where the confidence of high accuracy laser point clouds are propagated through a uniquely designed path in the high-resolution epipolar space of stereo images, such that a global dense matching algorithm can be externally constrained to yield results that preserve both the laser range data fidelity and image data resolution. Our approach is capable of greatly enhancing the resolution of laser range data using higher-resolution images even when both of the data sources are partially inconsistent.

The rest of the paper is organized as follows: Section 2 introduces an overview of related work; Section 3 describes the methodology of the proposed method in detail; Section 4 shows the experimental results; and Section 5 draws the conclusions based on our works.

2. Related work

In recent years, there have been several attempts to fuse low-res, accurate laser range data and high-res images to produce high-res, high-fidelity point clouds. The basic idea of super resolution is to up-sample the low-res range data by projecting them to a higher-resolution (2D/3D) grid and then to estimate their actual depth or disparity (parallax in the epipolar space) in a 2D scenario or presence in the 3D space. The images can be either used as the guide for extrapolating depth values or to form stereo pairs to generate geometric data for fusion, therefore the laser range data super resolution methods can be categorized into (1) interpolation-based super resolution and (2) stereo-matching-based super resolution.

Interpolation-based super resolution techniques utilize laser points and a single registered image to estimate depths of all the points/pixels in the higher-resolution grids after up-sampling. The basic assumption of these techniques is that homogenous points must share similar depths. The homogeneity between laser points and the pixels is inferred by the intensity/color/reflectance similarities and the distances between them in the image, which is formulated as the weight values of the pixels in the interpolation. Several research works (Andreasson et al., 2006, Wang and Ferrie, 2015) locally estimate depth for each pixel through a weighted average of depths of surrounding laser points, which is simple and efficient in yielding edge-aware super resolution results, while such algorithms are generally sensitive to noises. Other researches (Diebel and Thrun, 2005, Bódis-Szomorú et al., 2015) formulated the interpolation-based super resolution as a global Markov Random Field (MRF) problem, normally regarded as more accurate and robust than the local-window-based methods (Andreasson et al., 2006). Hosseinyalamdary and Yilmaz (2015) fully utilized the geometry of laser points and the brightness changes of images to recover high-resolution surfaces, respectively, and introduced a cost function to minimize the differences between these surfaces. Their method is capable of preserving boundary and topology of surfaces, while its result may be suppressed in the presence of pseudo brightness changes (e.g. image texture). In general, these methods are capable of efficiently improving the resolution of the laser point clouds. However, given the nature that the interpolation-based method essentially only utilizes the original laser points as the major information source, its level of improvement is limited when the camera image resolution is an order of magnitude higher than laser point clouds ($>$ 10x) (Yang et al., 2007), while this is often the case of many applications such as aerial mapping (normally 25x-100x), indoor reconstruction using low-res laser range scanners ($>$ 100x).

The most relevant work is the matching-based super resolution technique, which uses at least two overlapped images to form a stereo pair and integrates the laser point clouds into an image dense matching framework. The high-fidelity of laser point clouds is propagated to the pixels in the high-res epipolar space of stereo images, such that the matching confidences of these pixels are greatly improved. Different from the interpolation-based methods, the matching-based super resolution is able to take low-res laser point clouds ($>$ 100x) as the input and significantly improve their resolution. Several early studies (Bobick and Intille, 1999, Yang, 2003) on matching-based super resolution were based on the assumption that all the laser points were inliers, and they propagated the confidence of laser points to pixels through local scanning lines. These approaches are simple and efficient in yielding high-res point clouds, while the partial inconsistency between laser points and images were not considered, nor the laser points were fully utilized due to the local propagation paths. To further utilize the reliability of laser points, some researchers used local interpolation (Rengarajan et al., 2004, Geiger et al., 2010, Wu et al., 2011, Liu et al., 2015) or global interpolation (Wang et al., 2008, Wang and Yang, 2011) to build an initial surface from discrete laser points, and then constrained the disparities of pixels to be close to the initial surfaces. This method was able to maintain the accurate geometry from the laser points to a certain extent, while the constraint may negatively impact the fine structures in stereo image matching. In addition, geometric inconsistencies between laser range data and images are an important matter of concern, and some of the existing matching-based super resolution methods are able to accommodate a small number (less than one-tenth) of laser points as inconsistencies by taking them as soft geometric constraints for confidence propagation (Lhuillier and Quan, 2002, Kim et al., 2005, Geiger et al., 2010, Wang and Yang, 2011, Huang et al., 2015, Liu et al., 2015), thus to average out these inconsistencies. However, handling significantly larger inconsistencies between the images and the laser range data (e.g. temporal changes of objects such as building demolishing or rebuilt) remains challenging.

Therefore, our proposed super resolution method for laser range data based on stereo images aims to address the aforementioned issues by (1) fully utilizing the confidence of laser points through a newly devised non-local propagation paths, (2) providing a scheme to remove large amount of inconsistent parts between laser range data and images, and finally (3) producing high-resolution, high-fidelity point clouds to recover 3D information for fine structures.

3. Methodology

In general, our proposed method aims to use the stereo images to produce pixel-wise 3D points under the constraints of the sparsely available laser range data. The stereo images serve as the guidance in