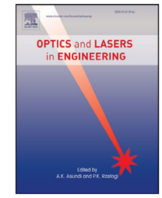
Optics and Lasers in Engineering 105 (2018) 188–197

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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



One-shot 3D scanning by combining sparse landmarks with dense gradient information

Matías Di Martino [a,*], Jorge Flores [b], José A. Ferrari [a]

[a] Facultad de Ingeniería - Universidad de la República Montevideo, J. Herrera y Reissig 565 CP 11300 Uruguay
[b] Electronic Engineering Department, University of Guadalajara Av. Revolucoón 1500, Guadalajara, Jal. CP 44840, Mexico

ARTICLE INFO

Keywords:
3D scanner
Depth retrieval
Fringe projection
Poisson equation
One-shot 3D

ABSTRACT

Scene understanding is one of the most challenging and popular problems in the field of robotics and computer vision and the estimation of 3D information is at the core of most of these applications. In order to retrieve the 3D structure of a test surface we propose a single shot approach that combines dense gradient information with sparse absolute measurements. To that end, we designed a colored pattern that codes fine horizontal and vertical fringes, with sparse corners landmarks. By measuring the deformation (bending) of horizontal and vertical fringes, we are able to estimate surface local variations (i.e. its gradient field). Then corner sparse landmarks are detected and matched to infer spare absolute information about the test surface height. Local gradient information is combined with the sparse absolute values which work as anchors to guide the integration process. We show that this can be mathematically done in a very compact and intuitive way by properly defining a Poisson-like partial differential equation. Then we address in detail how the problem can be formulated in a discrete domain and how it can be practically solved by straight forward linear numerical solvers. Finally, validation experiment are presented.

1. Introduction

Scene understanding is one of the most challenging and popular problems in the field of computer vision. To make this possible a wide variety of methods for 3D scene reconstruction have been proposed. In a first rough classification 3D scanning techniques can be divided into *Contact* and *Non-contact* methods. The last ones have become popular in the last decades and they are currently used for 3D shape reconstruction. Many commercial devices are being used [1,2] for example, in automobile industry the ATOS scanner of Gesellschaft fur Optische Messtechnik [3].

There are different kinds of non-contact methods, we will particularly focus on optical ones. They can be divided into three categories: *Time delay based* (e.g. [4]), *Image cue based* (e.g. [5,6]) and *Triangulation based* (e.g. [2,3,7,8]). The third class of methods uses -at least-two elements in order to obtain geometrical information of the scene by means of triangulation. Some methods use two or more cameras (*multi-view stereo* [9]), while other ones use a camera and control illumination sources e.g. Structured light [10–13], Photometric stereo [14], Phase Shifting [7,15–18] and Lasers Scans [2]. Methods that control or set illumination conditions are called *active* in contrast with *passive* methods in which just the relative position between the camera(s) and the objects are considered. Despite the differences that exist between Active and Passive approaches, they are based on the same geometrical principles [19,20].

We can classify existing methods into two categories: Single-Shot and Multi-Shot (see for example [21,22] and references therein). When the target object is static and the applications do not impose restriction over the acquisition time, multiple-shot approaches may be used. However, if the target is a moving scene, single shot methods must be applied. These methods have two important advantages. Firstly, they do not require any synchronization between the projector and the camera/s which leads to simpler and cheaper implementations. Secondly they can be used to scan three-dimensional dynamic scenes.

1.1. Related work

The present work presents a single-shot approach that combines dense gradient information [13] with sparse absolute measurements. To that end, we designed a colored pattern that codes fine horizontal and vertical fringes, and sparse corner landmarks. By measuring the deformation (bending) of horizontal and vertical fringes, we are able to estimate surface local variations (i.e. its gradient field). This allows us to have dense information about the local shape of the test surface, however, this kind of data lacks from absolute information about surface's height. By adding sparse corner landmarks -which can be easily detected

* Corresponding author.
E-mail addresses: matiasdm@fing.edu.uy (M. Di Martino), luis.flores@cucei.udg.mx (J. Flores), jferrari@fing.edu.uy (J.A. Ferrari).

https://doi.org/10.1016/j.optlaseng.2018.01.017
Received 16 November 2017; Received in revised form 9 January 2018; Accepted 22 January 2018

e.g. using Harris corner detector-, we can obtain sparse absolute information of the surface depth map. We proved that a simple mathematical formulation can combine the dense gradient information with the sparse absolute measurements. In addition, we proved that an optimal solution (in the sense of the L2-norm) can be obtained by solving a Poisson-like partial differential equation.

This work is based on the published method "One-shot 3D gradient field scanning" [13], and substantially improves its robustness when the test surface presents large discontinuities or isolated regions as we will show in Section 3. The closest methods in the literature to the one here presented, are the methods based on Takeda's work [19], which are commonly referred as Fourier profilometry single-shot methods. These methods consist in projecting one or more fine sinusoidal patterns, which are used to infer the 3D shape of the scene by measuring the variation of the phase of these patterns. To that end, Fourier Transform properties are exploited as explained in reference [19]. Many methods based on the Fourier Transform has been published in the past decade, a review of these methods can be found in the work of Zappa et al. [23], the work of Su et al. [24] or the work of Gorthi and Rastogi [25]. For example, Tavares and Vaz [26] proposed to use orthogonal fringe patterns for the enhancement of the Fourier transform method. Other approaches improve Takeda's original method by using the Windowed Fourier transform [27], or the Wavelet Transform for extraction of the phase information (see e.g. [28]).

Important differences exist between Differential 3D (D3D) and Fourier based methods. Firstly, while Fourier methods uses Fourier transform properties to extract the surface's depth, D3D approach measures local image gradients to estimate fringes local bending which provide depth gradient information. Secondly, Fourier based methods require a final unwrapping step, while D3D method requires the numerical integration of the retrieved gradient field. Thirdly, D3D approach measures the local properties of the projected pattern, and the information extracted in a certain region is independent from the rest of the domain; in contrast, the Fourier transform is inherently a non-local operation.

The present work, takes a step forward on D3D and proposes the use of sparse absolute information to improve its robustness, specifically, when the test surface presents discontinuities or isolated regions. Similar intents were investigated to make Fourier-based approaches robust to surface discontinuities. For example, Takeda et al. [29] proposed a Fourier-based method based on spatial frequency multiplexing combined with the Gushov-Solodkin phase unwrapping algorithm. Although this work introduce a very interesting theory and inspiring new ideas, the proposed approach was susceptible to significant measurement errors. To correct them, the authors proposed a number of post-processing steps that improved the method robustness at the cost of making it more complex and difficult to implement. Another example is the work of Wei-Hung Su [30] who proposed an area-encoded fringe pattern to retrieve objects that have spatially isolated surfaces. As the previous method, a single-shot pattern is required, which combines sinusoidal fringes with fine colored stripes. Fringes are matched without ambiguity using the pattern of colored stripes.

2. Proposed method

Let us consider a two dimensional domain $\Omega \subset \mathbf{R}^2$ and a function z: $\Omega \to \mathbf{R}$. In the context of the present work Ω represents a two dimensional subset of a plane in the 3D space, and $z(x, y)$ will be associated to the height map of an arbitrary test surface (as it is illustrated in Fig. 1). Despite the previous, most of the ideas presented in this section can be extended to other physical quantities and applied in many other contexts. Let us assume that by means of some technique [12–14,31–34] we are able to retrieve empirical estimations u and v of z_x and z_y partial derivatives[1], respectively. Ideally $u(x, y) = z_x(x, y)$ and $v(x, y) = z_y(x, y)$

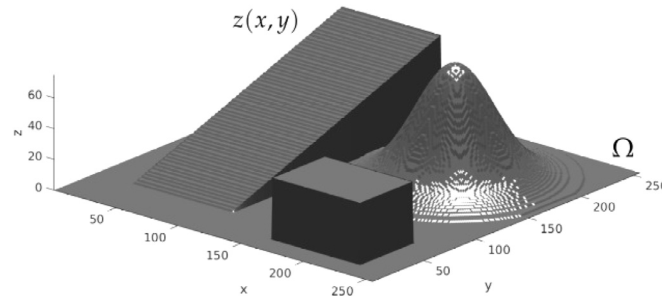


Fig. 1. Illustration of Ω and z definitions.

for all $(x, y) \in \Omega$, nevertheless, these equalities do not hold in practice due to: (a) empirical errors inherent to the method used, (b) surface discontinuities.

The estimation $\tilde{z}$ of z from the empirical data (u, v) can be formulated as the minimization problem,

$$\tilde{z} = \arg\min_{w} \int_{\Omega} \left((w_x - u)^2 + (w_y - v)^2 \right) \partial\Omega \tag{1}$$

where, of course, as the estimated gradient $(u, v) \to (z_x, z_y)$, the retrieved function $\tilde{z} \to z$. Eq. (1) can be solved by finding the solution to the Euler–Lagrange equation,

$$\begin{cases} \Delta \tilde{z} = u_x + v_y & (x, y) \in \Omega \\ \nabla \tilde{z} \cdot \hat{v} = 0 & (x, y) \in \partial\Omega \end{cases} \tag{2}$$

where $\Delta = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$ denotes the Laplacian operation, $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ the gradient vector, and $\hat{v}$ the unitary outward normal vector in the boundary set $\partial\Omega$. Eq. (2) is a well studied partial differential equation commonly referred as the Poisson equation. The two canonical formulations of the Poisson equation include Dirichlet ($\tilde{z} = 0, \quad (x, y) \in \partial\Omega$) or Neumann ($\nabla \tilde{z} \cdot \hat{v} = 0, \quad (x, y) \in \partial\Omega$) boundary conditions. In the present work Neumann boundary conditions are imposed as they are flexible enough to represent a wide range of real and complex surfaces. In order to guarantee the uniqueness of the solution of (2), one needs to impose at least one value of the solution on each connected component of the domain Ω. This is very easy to see by analyzing the minimization problem presented in (1), as only first order derivatives of the function w are involved, the functional remains invariant if we add arbitrary constants to w in any connected component of Ω. Taking into account the previous, we will consider the slightly modified Poisson equation,

$$\begin{cases} \Delta \tilde{z} = u_x + v_y & (x, y) \in \Omega \\ \nabla \tilde{z} \cdot \hat{v} = 0 & (x, y) \in \partial\Omega \\ \tilde{z} = g & (x, y) \in \Theta \end{cases} \tag{3}$$

which imposes values $g(x, y)$ to the solution at a discrete set of points $\Theta \subset \Omega$. Eq. (3) has a unique solution if $\Theta \neq \{\}$ in each connected component of the set Ω. Of course now g and the set Θ are part of the *input* data and need to be provided in addition to the gradient field (u, v).

2.1. Discrete model

In the following section we will assume that the domain of interest can be sampled as a regular rectangular grid:

$$z(x, y) = z_{ij} \quad \text{for } x = hj, \; y = hi \tag{4}$$

where h is the distance between samples (pixels) and the discrete indexes $(i, j) \in \Omega \overset{def}{=} \{1, \ldots, H\} \times \{1, \ldots, W\} \subset \mathbf{Z}^2$.

Discrete Laplacian operator: A discrete approximation for the Laplacian can be obtained through a second order Taylor expansion,

$$z_{i(j\pm1)} = z(x \pm h, y) = z(x, y) \pm z_x h + z_{xx} \frac{h^2}{2} + \mathcal{O}(h^3)$$

[1] For the sake of compactness we are adopting the notation $z_k \overset{def}{=} \frac{\partial z}{\partial k}$.