Pattern Recognition Letters 50 (2014) 15–22

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

An improved photometric stereo through distance estimation and light vector optimization from diffused maxima region

Jahanzeb Ahmad *, Jiuai Sun, Lyndon Smith, Melvyn Smith

Centre for Machine Vision, Bristol Robotics Laboratory, University of the West of England, Bristol, UK

ARTICLE INFO

Article history:
Available online 19 September 2013

Communicated by Dmitry Goldgof

Keywords:
Photometric stereo
Light vector calculation
Distance estimation

ABSTRACT

Although photometric stereo offers an attractive technique for acquiring 3D data using low-cost equipment, inherent limitations in the methodology have served to limit its practical application, particularly in measurement or metrology tasks. Here we address this issue. Traditional photometric stereo assumes that lighting directions at every pixel are the same, which is not usually the case in real applications, and especially where the size of object being observed is comparable to the working distance. Such imperfections of the illumination may make the subsequent reconstruction procedures used to obtain the 3D shape of the scene prone to low frequency geometric distortion and systematic error (bias). Also, the 3D reconstruction of the object results in a geometric shape with an unknown scale. To overcome these problems a novel method of estimating the distance of the object from the camera is developed, which employs photometric stereo images without using other additional imaging modality. The method firstly identifies Lambertian diffused maxima region to calculate the object distance from the camera, from which the corrected per-pixel light vector is able to be derived and the absolute dimensions of the object can be subsequently estimated. We also propose a new calibration process to allow a dynamic(as an object moves in the field of view) calculation of light vectors for each pixel with little additional computation cost. Experiments performed on synthetic as well as real data demonstrates that the proposed approach offers improved performance, achieving a reduction in the estimated surface normal error of up to 45% as well as mean height error of reconstructed surface of up to 6 mm. In addition, when compared to traditional photometric stereo, the proposed method reduces the mean angular and height error so that it is low, constant and independent of the position of the object placement within a normal working range.

1. Introduction

Traditional photometric stereo (PS) is used to recover the surface shape of an object or scene by using several images taken from the same view point but under different controlled lighting conditions (Barsky and Petrou, 2003; Sun et al., 2007). It was initially introduced by Woodham (1980). PS has been extensively used in many applications especially for estimating high density local surface normals in the fields of computer vision and computer graphics. It has been used for 3D modelling (Higo et al., 2009), facial expression capturing (Hern et al., 2010; Jones et al., 2011). It has also been used for medical applications (Ahmad et al., 2012; Sun et al., 2008) and in face recognition security systems (Hansen et al., 2010). Most of these applications require high accuracy reconstructed surfaces. So it is critical to estimate high accuracy surface normals in order to get accurate subsequent surface reconstruction from their integration. As we will show in the following experiment setup, 2–3 degree error in surface normal can produce up to 6 mm error in the height of reconstructed surface.

Current state-of-the-art systems normally assume that light sources are at an infinite distance from the scene so that a homogeneous and parallel incident light condition can be formed; and then the PS problem becomes solvable through a group of linear equations. In reality it is not always possible to produce parallel(collimated) incident light, especially when the object size is comparable in magnitude to the light separation and or the distance of object from light source is relatively small. Any underestimation or misalignment of the illumination may produce some error during recovery of the surface normal. For example, a 1% uncertainty in the intensity estimation will cause a 0.5–3.5 degree deviation in the calculated surface normal for a typical three-light source photometric stereo setup (Sun et al., 2007). Uncertainty in the calibration process can also lead to systemic errors when recovering surface normals and in the 3D recovered surface (Kobayashi et al., 2011; Horovitz and Kiryati, 2004).

* Corresponding author. Tel.: +44 (0)117 9656261.
E-mail addresses: jahanzeb.ahmad@uwe.ac.uk (J. Ahmad), jiuai2.sun@uwe.ac.uk (J. Sun), lyndon.smith@uwe.ac.uk (L. Smith), melvyn.smith@uwe.ac.uk (M. Smith).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.09.005

Furthermore PS gives no information concerning the absolute distance of the object from the camera. Other imaging modalities are normally required for obtaining such range data, for example laser triangulation or stereo vision techniques have been combined with the PS approach (Junyu et al., 2009; Hernández Esteban et al., 2008; Du et al., 2011; de Boer et al., 2010; Wu et al., 2011). A dense (per-pixel) surface reconstruction of a smooth and texture-less object proves to be a challenging task for many range detection imaging approaches, since they can only provide sparse surface data. In order to recover the range data at pixel resolution, we may alternatively make use of some information about the object surface itself such as convexity and smoothness.

In this paper we present a novel method to allow us to calculate the distance of an object based on the same photometric stereo imaging setup, i.e. one camera and four lights, without a requirement for any additional hardware, but with little extra computation processing cost. The object's distance from the camera is estimated by finding small patches on the object surface whose normal is pointing towards light source. This small patch is also called the diffused maxima region (DMR) and has been recently (Favaro et al., 2012) used for solving the problem of the generalized bas-relief ambiguity (GBR) (Belhumeur et al., 1999). The estimated distance is then used to calculate the light vectors at every image pixel, thereby minimizing the error associated with the assumption of a collimated light source. This approach enables the photometric stereo method to effectively work with real light sources, on Lambertian surfaces that have at least one patch with normal vectors pointing directly towards the light source, in reality this is a reasonable assumption.

To the best of our knowledge we are the first to use the DMR in this way, i.e. to enhance the PS method by reducing the well know problem of distortion in the recovered 3D surface by improving the light vector direction estimation and adding range data by using the convexity and smoothness of real objects and without using other additional imaging modalities. Paper is organized as following, in next section we will discuss related work after that photometric stereo technique is discussed. In Section 4 proposed method is discussed followed by experiments and results in Section 5. Finally in Section 6 paper is concluded.

2. Related work

The common low cost approach to produce collimated light is to use convex lenses or concave mirrors; but even in these cases, only a narrow parallel light beam with similar physical size to that of the lens or mirror can be obtained. To produce a collimated light source for a larger scene area, a possible solution is to develop a custom optical system with an array of specially aligned individual light units. Unfortunately this results in a high hardware and setup cost (Smith et al., 2005). Another practical solution is to set the light sources far away from the object (Ashdown, 1995), so that the light can be approximated as a distant radiation point source. This strategy may help to provide evenly distributed radiance across the object surface, but it sacrifices the majority of the illumination intensity, and correspondingly decreases the signal/noise ratio of the whole system. In addition, such a distant lighting setup usually means a large impractical working space is required. So this approach is only suitable for those light sources able to produce high levels of energy and those applications where a large redundant space is available. In terms of the availability and flexibility of current commercial illumination, the distant illumination solution is often not an optimal choice.

A nearby light source model has been considered as an alternative by Kim and Burger (1991) and Iwahori et al. (1990) to reduce the photometric stereo problem to find a local depth solution using a single non-linear equation. By distributed the light sources symmetrically in a plane perpendicular to camera optical axis, they were able to get a unique solution of non-linear equations. However, selection of initial values for the optimisation process and limitations in the speed for solving non-linear equation are the main problems with this method.

A moving point light source based solution has been proposed by Clark et al. (1992) termed "Active Photometric Stereo". By moving a point light along a known path close to the object surface a linear solution can be formulated to solve the photometric stereo problem. However, the range of motion of light must be closely controlled in order to guarantee the efficiency of the solution.

Kozera and Noakes introduced an iterative 2D Leap-Frog algorithm able to solve the noisy and non-distant illumination issue for three light-source photometric stereo (Kozera and Noakes, 2006). Because distributed illuminators are commercially available, Smith et al. approximated two symmetrically distributed nearby point sources as one virtual distant point light source for their dynamic photometric stereo method (Smith and Smith, 2005). Unfortunately, none of these methods lend themselves to a generalized approach.

Varnavas et al. (2010) implemented parallel CUDA based architecture and computed light vectors at each pixel by manually placing shiny sphere at the four corners of the field of view and assuming a flat plane at that distance, so that a changing light direction was taken into account. However in practice the whole surface of the object is not flat and is not necessarily at the same distance from the light source, especially when the size of the object is comparable to the distance of the light source.

3. Photometric stereo

According to the Lambertian reflectance model the intensity I of light reflected from an object's surface is dependent on the surface albedo ρ and the cosine of the angle of the incident light as described in Eq. 1. The cosine of the incident angle can also be referred as dot product of the unit vector of the surface normal $\overrightarrow{N}$ and the unit vector of light source direction $\overrightarrow{L}$, as shown in Eq. 2 .

$$I = \rho\cos(\phi_i) \tag{1}$$

$$I = \rho(\overrightarrow{L}.\overrightarrow{N}) \tag{2}$$

When more than two images (four images are used in the following work) from same view point are available under different lighting conditions, we have a linear set of Eqs. 1 and 2 and this can be represented in vector form as shown in Eq. 3.

$$\overrightarrow{I}(x,y) = \rho(x,y)[L]\overrightarrow{N}(x,y) \tag{3}$$

$\overrightarrow{I}$ is the vector formed by the four pixels $((I^1(x,y), I^2(x,y), I^3(x,y), I^4(x,y))^T$ from four images, $[L]$ is the matrix composed by the light vectors $(\overrightarrow{L^1};\overrightarrow{L^2};\overrightarrow{L^3};\overrightarrow{L^4})$. Where, 1, 2, 3 and 4 is the number with respect to the individual light source direction. $[L]$ is not a square and so not invertible, but the least square method can be used to compute Pseudo-Inverse and local surface gradients $p(x,y)$ and $q(x,y)$, and the local surface normal $N(x,y)$ can be calculated from the Pseudo-Inverse using Eqs. (4)–(6) where $\overrightarrow{M}(x,y) = (m_1(x,y), m_2(x,y), m_3(x,y))$.

$$\overrightarrow{M}(x,y) = \rho(x,y)N(x,y) = ([L]^T[L])^{-1}[L]^T\overrightarrow{I}(x,y) \tag{4}$$

$$p(x,y) = \frac{m_1(x,y)}{m_3(x,y)}, q(x,y) = \frac{m_2(x,y)}{m_3(x,y)} \tag{5}$$