Shadow Segmentation Using Time-of-Flight Cameras

Faisal Mufti[1] and Robert Mahony[2]

[1] Center for Advanced Studies in Engineering
faisal.mufti@ieee.org
[2] Australian National University
robert.mahony@anu.edu.au

Abstract. Time-of-flight (TOF) cameras are primarily used for range estimation by illuminating the scene through a TOF infrared source. However, additional background sources of illumination of the scene are also captured in the measurement process. This paper uses radiometric modelling of the signals emitted from the camera and a Lambertian reflectance model to develop a shadow segmentation algorithm. The proposed model is robust and is experimentally verified using real data.

Keywords: Time-of-flight, Radiometric Modelling, Shadow Segmentation, Reflectance Modelling.

1 Introduction

The presence of shadows due to lighting conditions complicates the process of shape and behaviour estimation of objects, especially where there is a significant association between the non-background points and shadow. Shadow segmentation algorithms are an important component in solutions to many computer vision scenarios such as video surveillance, people tracking and traffic monitoring [3, 19]. Background subtraction methods and colour space techniques are widely exploited for shadow segmentation [4, 8]. The drawback with such techniques is that they tend to under perform when there is insufficient colour information or when the scene is dynamic. Similarly, model based adaptive and Bayesian methods [1] suffer from same problems.

Conventional CCD camera are susceptible to dynamic range and illumination conditions. On the other hand, research in imaging devices in recent years has lead to development of range sensing cameras, especially 3D time-of-flight (TOF) [12] cameras, and is used in a number of fields; such as; for example, detection and recognition [5], 3D environment reconstruction [11] and tracking [15], etc. In general, 3D TOF cameras work on the principle of measuring time of flight of a modulated infrared light signal as phase offset after reflection from the environment and provide range and intensity data over a full image array at video frame rate [9].

This paper presents a novel algorithm for shadow detection using TOF camera technology by exploiting the additional measurement capability of a TOF camera compared to a standard CCD camera. The proposed algorithm is based on a

radiometric range model derived from TOF measurements and the background light source. We assume a statistical noise model of TOF measurements [16] and a Lambertian reflectance model to derive a radiometric range model that is independent of the coefficient of diffuse reflectivity of the environment. The radiometric range model is used to formulate a criterion to identify shadows and highlights using TOF camera data. For the purpose of this conference paper, we restrict our attention to the case of planar surfaces. More complex environment models are straight forward to develop based on the proposed approach. However, experiments show that the planar environment model is sufficient to obtain robust results for typical non-planar environments.

This paper is organised as follows: Section 2 describes TOF signal measurement, Section 3 describes Lambertian reflectance modelling from TOF perspective by exploiting TOF measurement and background light sources. In Section 4, the dependencies between measurement parameters of amplitude, range and intensity are used along with the reflectance model for sensor response to derive radiometric range model. Section 5 provides details of shadow segmentation based on radiometric range model using real data and is followed by its experimental verification.

2   Time-of-Flight Signal Measurement

Time-of-flight (TOF) sensors estimate distance to a target using the time of flight of a modulated infrared (IR) wave between the target and the camera. The sensor illuminates/irradiates the scene with a modulated signal of amplitude A (*exitance*) and receives back a signal (*radiosity*) after reflection from the scene with background signal offset I_o that includes non-modulated DC offset generated by TOF camera as well as ambient light reflected from the scene. The amplitude, intensity offset I and phase of a modulated signal can be extracted by demodulating the incoming signal $A_i = A\cos(\omega t_i + \varphi) + I$; ($t_i = i \cdot \frac{\pi}{2\omega}, i = 0, \ldots 3$) [9],

$$A := \frac{\sqrt{(A_3 - A_1)^2 + (A_0 - A_2)^2}}{2},$$

$$I := \frac{A_0 + A_1 + A_2 + A_3}{4},$$

$$\varphi := \tan^{-1}\left(\frac{A_3 - A_1}{A_0 - A_2}\right).$$

With known phase φ, modulation frequency f_{mod} and knowledge of speed of light c, it is possible to measure the un-ambiguous distance r from the camera [17].

3   Reflectance Model

The measurement parameters of amplitude A, intensity I, and range r are not independent but depend on the reflectance characteristics of the scene [17]. In the following discussion we consider a near-field IR point source for the camera's

active LED array, a far-field source for background illumination and ambient illumination. The primary source of illumination in TOF cameras is an IR source that produces a modulated IR signal offset and a non-modulated DC signal.
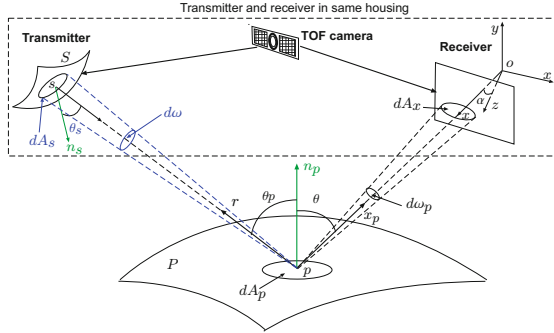

Fig. 1. Geometry of reflectance model for time-of-camera. Note that although the LED source and receiver of a physical TOF camera are co-located, it is difficult to provide a visualisation of this geometry. Here the source is shown separately to make is easier to see notation. However, in practice the directional vectors r and x_p are equal. Note that time variation (discussed in Section 2) of $A(s)$ does not need to be modelled as only the relative magnitude of $A(s)$ is of interest.

Let P be a Lambertian surface in space with n_p denoting the normal to each point $p \in P$ on the surface as shown in Figure. 1. Following the laws of radiometry [18] the amplitude of total radiance $= A(p)$ (called *radiosity*) leaving point p due to illumination by the modulated signal $A(s)$ is proportional to the *diffuse reflectance* or *albedo* $\rho_d(p)$ scaled by the cosine of arrival angle θ_p. In the present analysis, the LED point sources of the camera are part of the compact IR array of the TOF camera, and can be approximated by a single virtual modulated point source [7, p. 78] with the centre of illumination aligned with the optical axis of the camera [11]. In this case, the integration for illuminating sources can be written as a function of the exitance of a single point source at S as [7, p. 77] [17]

$$A(p) := \frac{1}{\pi}\rho_d(p)\frac{A(s)\cos\theta_p\cos\theta_s}{r^2}, \tag{1}$$

where θ_s is the angle between the normal to the source point $s \in S$ and the ray of the modulated IR signal reaching point p and r is the distance between source and the point p.

The irradiance of an image point x is obtained as

$$A(x) = \Upsilon A(p), \tag{2}$$

where $\Upsilon := \Upsilon(x)$ is the lens function [18] representing the vignetting due to aperture size and irradiance fall-off with cosine-fourth law.

The TOF camera IR source produces a DC signal from the same IR source LEDs. This signal will have the same reflectance model as has been derived for the modulated IR source (see (1)). The received signal $I_c(x)$ is given by [17]

$$I_c(x) = \Upsilon I_c(p). \tag{3}$$

The effect of this signal is an added offset to the modulated signal that provides better illumination of the scene.

For a point source $q \in Q$ that is far away compared to the area of the target surface, the exitance $I_b(q)$, does not depend on the distance from the source or the direction in which the light is emitted. Such a point source can be treated as constant [7, p. 76]. The radiosity perceived by a TOF image plane as a result of this IR source is given by [17]

$$I_b(x) = \frac{\Upsilon}{\pi} \rho_d(p) I_b(q) \cos \theta_q$$
$$= \Upsilon I_b(p). \tag{4}$$

where θ_q is the angle between normal to the surface point p.

Now consider an ambient background illumination of the scene i.e an illumination that is constant for the environment [7, p. 79] and produces a diffuse uniform lighting over the object [6, p. 273]. Let I_a be the intensity (called *exitance*) of the ambient illumination, then the received intensity $I_a(p)$ from a point p is expressed in an image plane as [17]

$$I_a(x) = \frac{\Upsilon}{\pi} \rho_a(p) I_a$$
$$= \Upsilon I_a(p), \tag{5}$$

where ρ_a is the *ambient reflection coefficient* which is often estimated empirically instead of relating it to the properties of a real material [6, p. 723]. Since it is an empirical convenience, for all practical purposes $\rho_a \approx \rho_d$.

4   Radiometric Range Model

From the principles of TOF camera (see Section 2) signals one knows that intensity component of TOF carries information for both, amplitude of the modulated signal and the background offset I_o [12]. The radiometric intensity of TOF camera is then

$$I := A + I_o. \tag{6}$$

The background offset I_o is composed of a DC offset I_c, due to the DC component of the illumination by the TOF camera LED array and background illumination that are modelled by an ambient illumination I_a and a background illumination I_b due to an infrared far field source present in the environment such as the Sun or other light source. Indexing the point p in the scene by the TOF receiving pixel x, one has

$$I_o(x) = I_c(x) + I_a(x) + I_b(x). \tag{7}$$

Dividing (6) by $A(x)$ and using the local shading model for IR signal and the illumination of point sources (2), (3), (4), (5), after substituting (7), one obtains

$$\frac{I(x)}{A(x)} = 1 + \frac{I_c(s)}{A(s)} + \frac{I_a r^2(x)}{A(s)\cos\theta_p\cos\theta_s} + \frac{I_b(q)\cos\theta_q r^2(x)}{A(s)\cos\theta_p\cos\theta_s}, \tag{8}$$

where $\theta_s := \theta_s(x)$ is a known function of pixel.

Define κ_a as the ratio of background ambient light I_a to modulated TOF IR source $A(s)$. Observe that κ_a does not depend upon scene or camera geometry and hence is a constant parameter over the full image array. Similarly, define κ_b as a constant ratio of far-field illumination I_b to the TOF IR source $A(s)$. Finally define κ_c as the ratio of TOF non-modulated IR source $I_c(s)$ and TOF modulated IR source $A(s)$. Since the two sources of illumination originating from the TOF camera IR LED source have the same ray geometry, then $\kappa_c(x)$ is a pixel dependent [17].

Thus, using the parameters $(\kappa_a, \kappa_b, \kappa_c, \theta_p, \theta_q)$ of sources, one obtains the radiometric relationship as

$$\boxed{\frac{I(x)}{A(x)} = 1 + \kappa_c(x) + \kappa_a \frac{r^2(x)}{\cos\theta_p\cos\theta_s} + \kappa_b \frac{\cos\theta_q r^2(x)}{\cos\theta_p\cos\theta_s}.} \tag{9}$$

Note that θ_s, is the angle measured from the camera, is stored in a look up table (projected angle of the IR beam of TOF camera with respect to each camera pixel) and angle θ_p is measurable from camera data. For any surface patch[1], it is possible to numerically compute an estimate of the angle $\theta_p := \theta_p(x)$ from the set of range measurements $r(x_i)$ associated with that patch based on an estimate of the normal vector to the surface [10].

The radiometric range model for a far-field point source of illumination depends upon the angle θ_q that is the angle between the normal to the surface and the direction of the far-field source. Hence, an estimate of the direction of background IR source (in terms of azimuth and elevation) is required and is a problem of source estimation in computer vision [2].

A specialized case of considerable practical interest is that of a planar surface. For a single planar surface it follows that θ_q is constant, while θ_p is nearly constant for a small field of TOF optical sensor over the surface, and the parameters (κ_a, κ_b) are constant over the image plane. As a result several parameters can be combined into a single constant κ_o. An approximate model is presented in [17] as

$$\kappa_o := \frac{\kappa_a}{\cos\theta_p} + \frac{\kappa_b\cos\theta_q}{\cos\theta_p} \approx \text{constant}. \tag{10}$$

As a consequence of (10), a pixel measurement of κ_o is given by

$$\kappa_o(x) = \left(\frac{I(x)}{A(x)} - \kappa_c(x) - 1\right)\frac{\cos\theta_s}{r^2(x)}. \tag{11}$$

---

[1] The surface patch must be sufficiently large to be imaged by a small window of pixels.

A detailed analysis of κ_o provides a statistical distribution of $\kappa_o(x)$ for a planar surface [17] shown in Figure
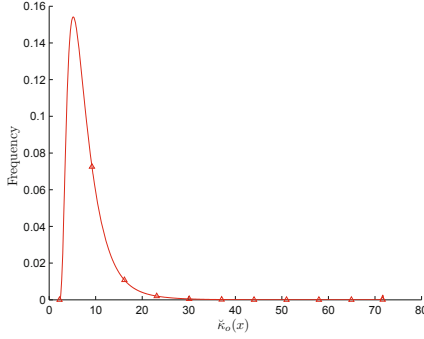

Fig. 2. Normalized histogram of $\check{\kappa}_o(x)$ of a flat surface. Pixels with $\check{\kappa}_o(x) \to \infty$ (due to amplitude and phase) have been scaled down to finite values. The heavy tail is associated with noisy data.

5   Shadow Segmentation

Shadow segmentation based on radiometric range data is a natural outcome of the radiometric framework proposed in Section 4. An important contribution of performing shadow segmentation using TOF cameras is the ability to provide an algorithm that can be used in dynamic scenes as well as for surfaces with insufficient colour information. A close examination of (2) and (6) reveals that an object in shadow does not effect the amplitude value received by the TOF camera. However the intensity varies with the background illumination or shadowing of the object. Based on this fact, it is possible to segment the scene into shadow and highlights. The proposed shadow segmentation model is derived from the radiometric model (9) using far-field background illumination parameter κ_b. Thus, one obtains a new $\kappa_b(x)$ for this application as

$$\kappa_b(x) := \left(\frac{I(x)}{A(x)} - \kappa_c(x) - 1 \right) \frac{\cos\theta_s \cos\theta_p}{r^2(x)\cos\theta_q} - \frac{\kappa_a(x)}{\cos\theta_q}. \tag{12}$$

In terms of the measured value, one defines a measured value of $\check{\kappa}_b(x)$ as

$$\check{\kappa}_b(x) := \left(\frac{I(x)}{A(x)} - \hat{\kappa}_c(x) - 1 \right) \frac{\cos\theta_s \cos\theta_p}{\check{r}^2(x)\cos\theta_q} - \frac{\check{\kappa}_a(x)}{\cos\theta_q}, \tag{13}$$

where $\hat{\kappa}_c(x) \in \mathbb{R}^2$ is an estimate of camera based pixel parameter for an entire image, since $\kappa_c(x)$ is scene independent and can be measured offline in a set of calibration experiments. The proposed algorithm uses a k-means clustering algorithm [14] on the full set of measured $\check{\kappa}_b(x)$ obtained for a full frame. Here

the term involving $\check{\kappa}_a(x)$ is a scalar value for ambient illumination and is computed empirically assuming constant angle θ_q for a planar surface. The objective function $\Psi := \Psi(c)$ for $c = \{c_1, \cdots, c_k\}$ is defined as

$$\Psi(c) := \sum_{i=1}^{k} \sum_{x_n \in S_i} |\check{\kappa}_b(x_n) - c_i|. \tag{14}$$

The c_i values are the geometric centroids of the data points S_i, where x_n is a vector representing nth data point from radiometric range model $\check{\kappa}_b$ and $|\check{\kappa}_b(x_n) - c_i|$ is the L_1 norm. The minimisation of this problem for segmentation (partitions) of n data points into k disjoint subset S_i containing n_i data points is given by $\hat{c}$

$$\hat{c} = \arg\min_{c}(\Psi). \tag{15}$$

The proposed shadow segmentation based on $\kappa_b(x)$ value represents the ratio of background source of illumination to the camera source. The angle θ_q is assumed constant for planar cases and θ_p is relatively constant for the small field of view of a TOF camera. Under these assumption the probability distribution of $\check{\kappa}_b$ behaves like $\check{\kappa}_o$ (11).

Remark: *When shadow segmentation due to the Sun is considered then geographical knowledge of the camera pose and the Sun position can be used to estimate θ_p and θ_q given an estimate of the scene geometry [13].*

Fig. 3. CCD image of the setup taken from a 2D camera with TOF camera positioned in the lower left corner of the image

Experiments: A TOF camera was placed to capture an environment with object (flat floor) in sunlight and the remaining background appeared shadowed compared to bright light coming from the window (see Figure 3). The intensity image captured by the TOF camera is shown in Figure. 5(a). A histogram of

$\breve{\kappa}_b(x)$ with three marked regions is shown in Figure 4. The central blue region corresponds to the highlighted area in image space. The region on the right hand side (in a red rectangle) corresponds to noisy values. Both the highlighted and the noisy regions produced significantly higher $\breve{\kappa}_b(x)$ values compared to shadowed regions (higher numerator of the right hand side of (13) either due to $I \rightarrow \infty$ or noisy signal where $A \rightarrow 0$). The values on the left side (in a red elliptical) have two peaks. These two regions are associated with the angle approximation θ_q when applied over an entire frame including non-planar regions and θ_p (sharp corners and bends in scene with respect to camera viewing direction).
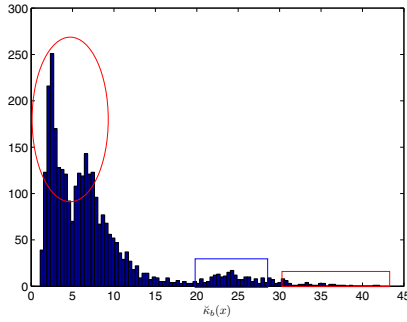

Fig. 4. Histogram of a $\breve{\kappa}_b(x)$ for a single frame

In the initial experiment the algorithm was applied to a single planar surface. The shadow segmentation was performed using the k-means clustering. The $\breve{\kappa}_b(x)$ data is segmented into two clusters of highlights and shadows and shows quite precise segmentation (see Figure 5(c)) with a precision of 99%. The proposed algorithm was also applied to an entire single frame and observed that the segmentation works reasonably well even when applied to a non-planar scene (see Figure 5(b)). However, there were certain regions of false positive due to failure of planar assumption (precision factor of 75%). These regions can either be segmented using more that two clusters or eliminated using further noise filtering [17] based on signal-to-noise ratio (SNR) estimation. Since in the case of highlight areas the estimated $\widehat{SNR}$ value is high while the noisy regions have a lower $\widehat{SNR}$ value, the image was further refined, using this two step approach as illustrated in Figure 5(d) and achieved a precision of 96%. The algorithm, when implemented in MATLAB on an Intel Core2 Duo 2.2GHz machine with 4GB RAM, performs region segmentation of a complete frame in less than a second.

A precise modelling of scene geometry (discussed in Section 4) for angle θ_q in the model (13) would result into a single region of shadow in a single step. Despite this simplification, the algorithm is effective in region segmentation based on illumination condition and unlike background subtraction and color cues, this method is independent of scene movement and works equally well on gray scale and range image data.
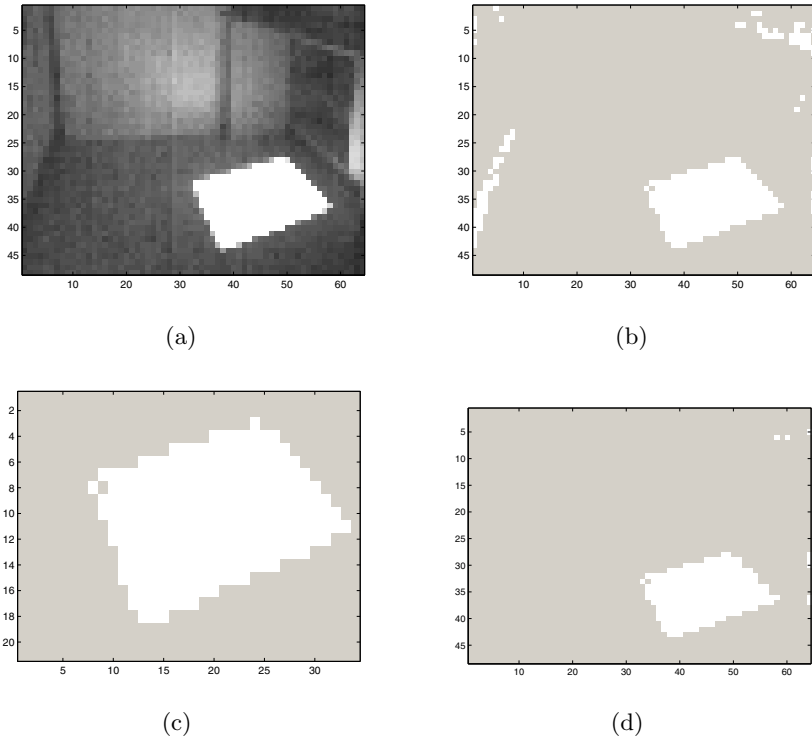
(a)

(b)

(c)

(d)

Fig. 5. (a) Intensity image observed from a TOF camera (b) segmentation of shadow and highlighted regions (non shadow) of a complete frame (c) shadow segmentation of a planar surface only. (d) Refined shadow segmentation after two step approach.

6   Summary

Time-of-flight (TOF) cameras are primarily used for range estimation by illuminating the scene through a TOF IR source. Unlike conventional cameras where only a single parameter is measured as intensity, TOF camera's additional measurement facilitate in deriving a radiometric range model for the TOF camera. The model comprising light sources and scene geometry is independent of reflectivity for all practical purposes. Despite the fact that the radiometric model is based on much simpler assumption, the framework proved robust and effective. Experimental results prove the effectiveness of this model for TOF cameras in vision based algorithms.

References

1. Benedek, C., Sziranyi, T.: Bayesian foreground and shadow detection in uncertain frame rate surveillance videos. IEEE Trans. Image Process. 17(4), 608–621 (2008)
2. Cao, X., Shah, M.: Camera calibration and light source estimation from images with shadows. In: Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2005, vol. 2, pp. 918–923 (2005)

3. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell. 25(10), 1337–1342 (2003)
4. Elgammal, A., Duraiswami, R., Harwood, D., Davis, L.: Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proc. IEEE 90(7), 1151–1163 (2002)
5. Fardi, B., Dousa, J., Wanielik, G., Elias, B., Barke, A.: Obstacle detection and pedestrian recognition using a 3D PMD camera. In: Proc. IEEE Intell. Vehicles Symp., pp. 225–230 (2006)
6. Foley, J.D., Da, A.V., Feiner, S.K., Hughes, J.F.: Computer Graphics: Principles and Practices. Addison-Wesley Publishing Company, Inc., Reading (1997)
7. Forsyth, D.A., Ponce, J.: Computer Vision: A Modern Approach. Prentice-Hall, Englewood Cliffs (2003)
8. Jacques, J., Jung, C., Musse, S.: Background subtraction and shadow detection in grayscale video sequences. In: Proc. 18th Brazilian Symposium on Computer Graphics and Image Processing SIBGRAPI 2005, pp. 189–196 (2005)
9. Kahlmann, T., Remondino, F., Guillaume, S.: Range imaging technology: new developments and applications for people identification and tracking. In: Proc. SPIE-IS&T Electronic Imaging, San Jose, CA, USA, vol. 6491 (January 2007)
10. Klasing, K., Althoff, D., Wollherr, D., Buss, M.: Comparison of surface normal estimation methods for range sensing applications. In: Proc. IEEE International Conference on Robotics and Automation, ICRA 2009, May 12-17, pp. 3206–3211 (2009)
11. Kuhnert, K.D., Stommel, M.: Fusion of stereo-camera and PMD-camera data for real-time suited precise 3D environment reconstruction. In: Proc. IEEE/RSJ Int. Conf. Intell. Robot. Systs. (2006)
12. Lange, R., Seitz, P.: Solid-state time-of-flight range camera. IEEE J. Quantum Electron. 37, 390–397 (2001)
13. Leroy, M., Roujean, J.L.: Sun and view angle corrections on reflectances derived from NOAA/AVHRR data. IEEE Trans. Geosci. Remote Sens. 32(3), 684–697 (1994)
14. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., vol. 1, pp. 281–297. Univ. of Calif. Press (1967)
15. Meier, E., Ade, F.: Tracking cars in range image sequnces. In: Proc. IEEE Int. Conf. Intell. Trans. Systs., pp. 105–110 (1997)
16. Mufti, F., Mahony, R.: Statistical analysis of measurement processes for time-of-flight cameras. In: Proc. SPIE Videometrics, Range Imaging, and Applications X, vol. 7447-21 (2009)
17. Mufti, F., Mahony, R.: Radiometric range image filtering for time-of-flight cameras. In: Proc. Int. Conf. on Computer Vision Theory and Applications (VISAPP 2010), Angers, France, vol. 1, pp. 143–152 (May 2010)
18. Sillion, F.X., Puech, C.: Radiosity and Global Illumination. Morgan Kaufmann, San Francisco (1994)
19. Tsai, V.: A comparative study on shadow compensation of color aerial images in invariant color models. IEEE Trans. Geosci. Remote Sens. 44(6), 1661–1671 (2006)