K. Irie, A. E. McKinnon, K. Unsworth, I. M. Woodhead, 'Shadow Removal for Object Tracking in Complex Outdoor Scenes', *Proceedings of Image and Vision Computing New Zealand 2007*, pp. 25–30, Hamilton, New Zealand, December 2007.

Shadow Removal for Object Tracking in Complex Outdoor Scenes

Kenji Irie[1], Alan E. McKinnon[2], Keith Unsworth[2], Ian M. Woodhead[1]

[1] Lincoln Ventures Ltd, P.O. Box 133, Lincoln, Christchurch 7640, New Zealand.
[2] Applied Computing Group, Lincoln University, Canterbury, New Zealand.

Email: iriek@lvl.co.nz

Abstract

Shadows occurring in outdoor scenes can have a significant impact on the content within captured images. In object tracking applications they can grossly extend the apparent shape of an object due to the reduction in illumination intensity within cast-shadows. They have typically been removed from images using observed properties such as shape, intensity, texture retention, and shadow direction. However, these methods often fail with grey or shadow-like objects on complex backgrounds that vary in colour or texture. We present a new non-linear method of shadow removal for object tracking based on a surface's response to sunlight and skylight. The method returns full-colour images, operates on complex scenes in outdoor environments, and requires only 2 images for calibration.

Keywords: shadow removal, object tracking, image processing

1 Introduction

Object tracking in natural outdoor environments is a challenging process due to the uncontrollable lighting and high variability of textures present in outdoor scenes. Shadows are usually removed from images by use of their observed properties within the image detail, such as attachment to an object within the scene or that underlying scene texture information is retained within the umbra of the shadow. However, shadows cast in outdoor environments are difficult to differentiate from genuine objects that are added to the scene, as the complexity of the background confounds shadow detection based on these methods.

Early research into shadow removal typically focussed on identifying dark areas upon plain, flat surfaces [1]. More complex methods analysed the direction of the shadows by assessing the source illumination directly [2] or indirectly [3]. Statistical methods analyse pixel information in sequences of images [4-6] , while interactive methods provide a means to manually guide the shadow removal process [7, 8]. More recent shadow-detection research has focussed on detecting the blue colouring of shadows due to ambient skylight [7, 9-14], Most of these methods use some form of edge detection to outline shadow areas. The recent work of Finlayson et al. [14] describes a method that uses the properties of colour temperature to derive an illumination-invariant but coarsely quantised grey-scale image. They use edges in the original and illumination-invariant image to distinguish regions of shadow, and attempt to relight the shadowed region of the original image. Their method of shadow removal is dependent on edge detection of shadow and suffers from significant metamerism in the grey-scale image. Withagen et. al [13] have developed a method of shadow removal using dual illuminants and image division that models a pixel's change to illumination as a flat surface in 3-D space. The performance of their algorithm is constrained in out-door conditions as they assume flat illumination spectra in each colour-band of the camera, which does not hold for ambient skylight [15] or discharge-type illumination sources. Their method requires a large number of frames for calibration of their algorithm.

We present a new method of shadow removal for object tracking in outdoor scenes that does not rely on edge-detected shadows or flat illumination spectra, and requires only 2 images for calibration. It harnesses the colour properties of the two different illumination sources of direct sunlight and ambient skylight present in outdoor environments on sunny days. This allows our algorithm the freedom to operate on images with cluttered, complex backgrounds, as it is not dependent on the shape or position of cast shadows but on pixel response alone. First we provide the background theory of our method of shadow removal, then its application to a sequence of real-world outdoor images. We complete this paper with a discussion and conclusion.

2 Shadow Removal

Our method of shadow removal is a 3-stage process: modelling the pixel response to illumination change, shadow removal using scene change detection, and filtering of noise.

2.1 Modelling of Natural Illumination

Outdoor environments are naturally illuminated by filtered light from the sun. Daylight is typically composed of sunlight, which has a spectrum close to that of a black body radiator [16], and skylight, whose blue colour is primarily formed from the Rayleigh scattering of light particles passing through the atmosphere [17]. On sunny days the non-shadowed areas of a scene are lit predominantly by daylight, while shadows thrown by objects onto the surrounding terrain are predominantly illuminated by skylight only. It is this distinction between the illumination sources that allows for the identification of regions in cast shadow.

The pixel response P, in an RGB camera is described by the product of illumination, surface reflectance, and camera sensitivities [18] giving:

$$P_k = \int E(\lambda) S(\lambda) C_k(\lambda)\, d\lambda \,,\ k \in \{R, G, B\} \quad (1)$$

where E is the illumination spectral power distribution (SPD), S is the surface spectral reflectance function, C is the camera response, k is an RGB colour band of the camera, and λ is the wavelength.

The daylight and skylight SPDs are different for outdoor conditions, and the change in pixel response P_k can be modelled as the change in SPD from skylight to daylight:

$$P_k = \int \left(E_1(\lambda) + nE_2(\lambda)\right) S(\lambda) C_k(\lambda)\, d\lambda \,, \quad k \in \{R, G, B\},\ n \in [0,1] \quad (2)$$

where E_1 is the ambient skylight source, E_2 is the directional sunlight source, and n represents the proportion of added illumination E_2.

The use the colour band ratios (also referred to as chromaticities) $x=R/G$ and $y=B/G$ provides a space that is intensity invariant. Taking the logs of the band ratios (LBR) as in [14] distributes the colour data relatively evenly enabling easier colour segmentation. The transform of the pixel response from RGB triplets to x/y coordinates in LBR space is therefore

$$x = \log\left(\frac{\int \left(E_1(\lambda) + nE_2(\lambda)\right) S(\lambda) C_R(\lambda)\, d\lambda}{\int \left(E_1(\lambda) + nE_2(\lambda)\right) S(\lambda) C_G(\lambda)\, d\lambda} \right) \quad (3)$$

and

$$y = \log\left(\frac{\int \left(E_1(\lambda) + nE_2(\lambda)\right) S(\lambda) C_B(\lambda)\, d\lambda}{\int \left(E_1(\lambda) + nE_2(\lambda)\right) S(\lambda) C_G(\lambda)\, d\lambda} \right) \quad (4)$$

where $n \in [0,1]$. Equations (3) and (4) require the response in the green camera channel to be non-zero. This is a reasonable expectation as the bandwith of the colour filters in colour cameras is typically broad [19] and therefore the green channel would exhibit a response to almost any non-zero illumination in the visible spectrum.

2.2 Scene change detection

A change in scene information can be described by any change in pixel response that does not fall on the curve described by P in the LBR space, as illumination changes from E_1+E_2 to E_1. The addition of a region that bounds the modelled shadow curve accommodates the quantization of pixel values and addition of image noise that moves measured values of P from the modelled curve. This is shown graphically in figure 1.

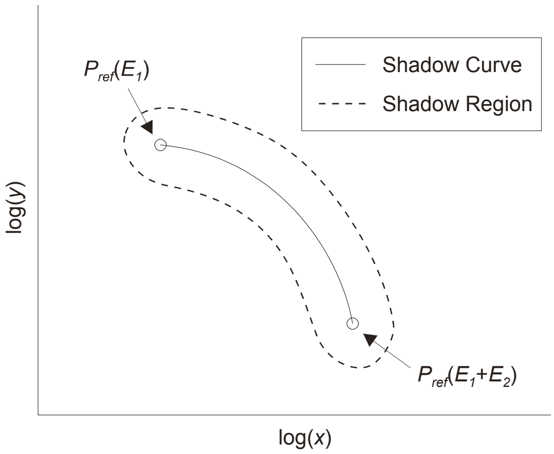

Figure 1: The modelled shadow curve for a particular surface and associated shadow region for segmenting changes in a pixel response.

If a pixel $P_{comp}(x,y)$ from an image being compared to a pixel $P_{ref}(x,y)$ from the reference image falls inside the shadow region it is then considered a shadow and $P_{comp}(R,G,B)$ is replaced by the RGB pixel from the reference image, $P_{ref}(R,G,B)$. In effect the detected cast-shadow pixel is replaced by the non-shadowed pixel from the reference image.

In practice a binary mask is created corresponding to each pixel in $P_{comp}(R,G,B)$. The mask value is zero if the pixel falls inside the shadow region shown in figure 1 otherwise it is one, indicating scene change that is not due to cast shadows.

2.3 Noise filtering

Noise within images is exhibited as temporal and spatial variations in pixels [20] and cascades into the resultant shadow-removed images if left unfiltered. Image noise is particularly detrimental to our analysis of shadows due to the colour division in the formation of the band ratios in (3) and (4), where the effect of noise increases with decreasing pixel value. A series of filters is applied to the images to improve the performance of the algorithm in the presence of noise:

- all images are convolved with a 2-D Gaussian kernel to dampen temporal and spatial variations due to noise
- small isolated areas detected as scene changes are diagnosed as erroneous and removed
- morphological opening and hole-filling is applied to the resultant areas detected as scene changes to fill small gaps left by the shadow detection algorithm.

Values for parameters are determined empirically for the size of the Gaussian kernel (9x9, 2 standard deviations) and the minimum size for isolated areas (0.25% of the total image area).

2.4 The algorithm

The shadow removal algorithm used in the following experiment combines the scene-change detection properties of the illumination-invariant model with noise filtering. The following steps generate images that are free from cast shadows:

1. Gaussian filter the reference and comparison image
2. create a mask of genuine scene change from the illumination-change model
3. perform morphological opening on the mask to break up any connected areas of noise
4. size filter the remaining areas of scene change to remove patches of noise
5. fill any remaining holes within shadow areas
6. use the mask to merge the reference and comparison images to create the shadow-free image.

3 Experiment

All images were captured with a commercially available CMOS colour camera (table 1), with digital enhancements disabled.

Table 1: 1*UEYE UI1210-C* camera details

Parameter	Value
Sensor type	½" CMOS (Bayer array)
Native resolution	640 x 480
Video mode	24-bit RGB (8-bits/channel)
Interface	USB 2.0

3.1 Calibration

The model in equation (2) can be calibrated for a particular surface *S* by analysing images of the surface with illumination at the start and end points of the equation (n=0, and n=1). This can be achieved by imaging the surface in shadow where it is illuminated by skylight only and in full daylight where it is illuminated by both skylight and sunlight. A set of calibration curves were generated using a Gretag Macbeth Color Chart (GMB), which consists of 19 panels of distinct colours. The resulting calibration curves of the GMB chart are shown in figure 2. The GMB calibration curve corresponding to each pixel in the reference image $P_{ref}(x,y)$ is the curve whose n=1 value (i.e., the point at the daylight end of the curve) is nearest to the reference image pixel in LBR space. The curve thus identified is translated so that its n=1 end point coincides with the position of the pixel from the reference image in LBR space.

The shadow region was empirically set as a Euclidean threshold around the shadow curve, and set to a value of 0.06 in LBR space to provide a suitable balance between false-positive shadow and true-negative object detection.

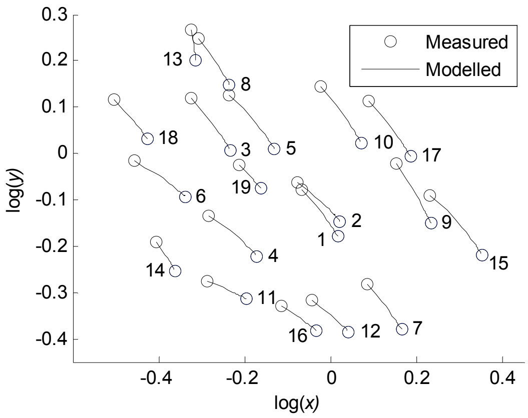

Figure 2: Model calibration of the Gretag-Macbeth Color Chart panels in LBR space using measured changes in illumination for natural outdoor lighting on a clear day.

3.2 Results

The shadow removal algorithm was applied to two image sequences of a person walking through a scene casting shadows across a complex terrain. Figure 3 shows the reference image used for the image sequence *blue man*. The terrain consists of mixed shingle packed with dirt, sporadic covering of short and long grass patches, weeds and twigs, with no uniformity in surface colour or texture.

Figure 4 illustrates the process of shadow removal for a frame of the *blue man* sequence, containing a clothed person that is significantly different in appearance to cast shadows. The top-left image contains the person wearing blue overalls and black shoes, along with a cast shadow. The raw mask from the scene-change detection algorithm strongly detects the addition of the blue overalls, as expected, with weak detection on the shiny black leather shoes. The shadowed region has speckled areas of false-positive detection. The final mask after noise filtering provides a reasonable outline of the person, with true-negative detection of shadows on the shoes. Overall, the algorithm shows good cast-shadow removal and object integrity.

Figure 5 illustrates shadow removal applied to the *grey man* sequence, which contains items of grey and black coloured clothing that more closely match the appearance of shadows than that of the *blue man* group of images. The shadow removal performance is the same as for the blue man sequence, as expected, due to the cast shadows being processed independently of the object added to the scene. The most challenging component is in the colouring of the clothing within the scene. The grey shirt and matt black pants are remarkably well conserved, considering the difficulty in visually separating them from shadow. Although the raw mask is not as solid as that in the *blue man* sequence, the final filtered mask contains a reasonably complete outline of the person.

The speckled patches in the raw mask behind the person in figures 4 and 5 can be attributed to the deformation of the scene as it was walked upon. The thin extrusion from the foot in the final mask in figure 5 is from a large twig that has been moved by the foot. A montage of several original and shadow-removed images from the *blue man* and *grey man* sequences is given in figures 6 and 7. All images in both sequences demonstrated a high degree of shadow removal.

4 Discussion

At its core, the shadow removal algorithm presented is based on pixel colour and the ability to threshold any change in pixel value that does not fall into the gamut of colours described by the pixel's response under the two illumination sources of sunlight and skylight. The algorithm is prone to failure when a genuine scene change occurs where the new object's colour falls within this gamut. The segmentation performance is bound by the noise content within the image, which effectively sets the size of the threshold for the shadow region around the shadow curve.

One advantage of using a method of shadow removal based upon colour is that no assumptions are necessary for the texture or colour of the terrain the shadow is cast on. The strength of the shadow is also unimportant due to the scaling of n times the illuminant E_2 in equation (2), meaning that both hard and subtle cast-shadows can be removed and that edges are no longer required to identify shadow regions.

It is important that the imaging device is capable of capturing the range of intensities present in the image within the dynamic range of the camera. Gain and gamma can be used as long as the calibration is

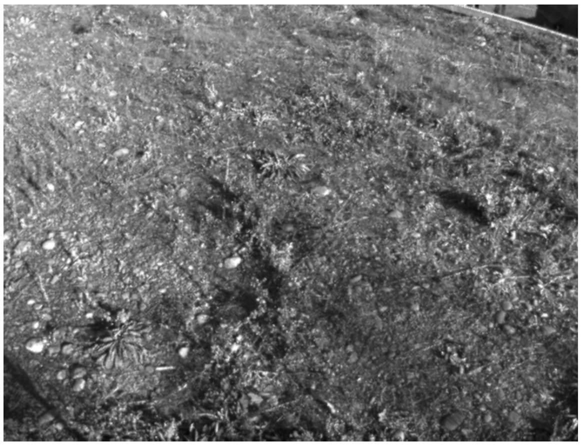

Figure 3: The intensity image of the reference image for the *blue man* image sequence.

Figure 4: The process of shadow removal from a frame of the *blue man* sequence (intensity images). (TL) original image containing a person and cast shadow. (BL) raw mask image after scene-change detection. (BR) noise-filtered mask. (TR) final shadow-removed image.

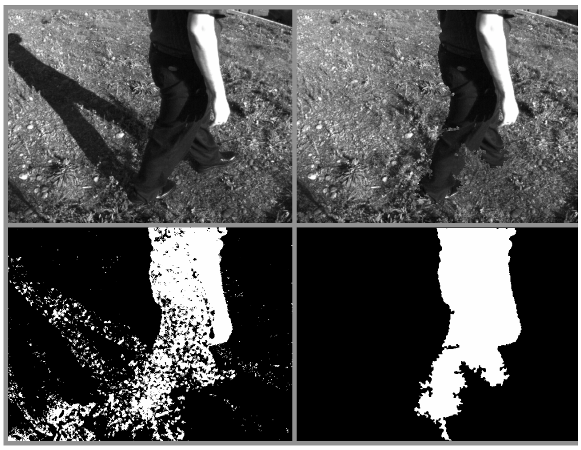

Figure 5: The process of shadow removal from a frame of the *grey man* sequence (intensity images, in the same order as figure 4).

Figure 6: Original (top) and cast shadow-removed (bottom) images from the *blue man* sequence.

Figure 7: Original (top) and cast shadow-removed (bottom) images from the *grey man* sequence.

performed with the same parameter settings that are used for subsequent image capture.

Future research will focus on a method to adaptively set the threshold in LBR space to reduce the effects of image noise.

5 Conclusion

We have developed a shadow removal algorithm [21] capable of removing shadows from complex backgrounds for use in outdoor images. It does not rely on shadow shape, or upon underlying scene texture but upon colour information only. No assumptions are made about the shape of the illumination spectra, except the expectation that the green channel response is non-zero. Results show that the algorithm performs well in a cluttered background environment, even with grey and dark coloured objects casting shadows.

6 Acknowledgement

This work was supported by the New Zealand Foundation for Research, Science and Technology programme LVLX0401. We also wish to acknowledge the assistance from reviewer #2 in particular, who drew our attention to the very recent work of Withagen et. al [13].

7 References

[1] C. Jiang and M. O. Ward, "Shadow identification," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1992, pp. 606-612.

[2] G. Funka-Lea and R. Bajcsy, "Combining color and geometry for the active, visual recognition of shadows," in *Proc. IEEE 5th Intl. Conf. Computer Vision*, 1995, pp. 203-209.

[3] J. M. Wang, Y. C. Chung, C. L. Chang, and S. W. Chen, "Shadow detection and removal for traffic images," in *Proc. IEEE Int. Conf. Networking, Sensing and Control*, vol. 1, 2004, pp. 649-654.

[4] F. Porikli and J. Thornton, "Shadow flow: a recursive method to learn moving cast shadows," in *Proc. IEEE 10th Int. Conf. on Computer Vision*, vol. 1, 2005, pp. 891-898.

[5] O. Javed and M. Shah, "Tracking and Object Classification for Automated Surveillance," in *Proc. 7th European Conf. Computer Vision-Part IV*: Springer-Verlag, 2002, pp. 343-357.

[6] M. M. Ibrahim and A. Rajagopal, "Shadow Detection in Images." US: Honeywell International Inc., 2007.

[7] G. Finlayson, S. D. Hordley, and M. Drew, "Removing Shadows From Images using Retinex," in *Proc. 10th Color Imaging Conf.:*

Color Science and Engineering Systems, Technologies, Applications. Scottsdale, Arizona, 2002.

[8] Y. Lu, H. Xin, J. Kong, B. Li, and Y. Wang, "Shadow Removal Based on Shadow Direction and Shadow Attributes," in *Proc. Int. Conf. Computational Intelligence for Modelling, Control and Automation, and Intelligent Agents, Web Technologies and Internet Commerce*, 2006, pp. 37.

[9] J. A. Marchant and C. M. Onyango, "Shadow-invariant classification for scenes illuminated by daylight," *Journal of the Optical Society of America A*, vol. 17, pp. 1952-1961, 2000.

[10] H. Jiang and M. S. Drew, "Shadow-resistant tracking in video," in *Proc. IEEE Int. Conf. Multimedia and Expo*, vol. 3: IEEE Computer Society, 2003, pp. 77-80.

[11] S. G. Erbou, H. B. D. Sorensen, and B. Stage, "Detection of Cast Shadows in Surveillance Applications," in *Den 14. Danske Konference i Monstergenkendelse og Billedanalyse*: Informatics and Mathematical Modelling, Technical University of Denmark, 2005.

[12] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, pp. 1079-1087, 2004.

[13] P. J. Withagen, F. C. A. Groen, and K. Schutte, "IAS technical report IAS-UVA-07-02 Shadow detection using a physical basis," Intelligent Autonomous Systems, University of Amsterdam 2007.

[14] G. D. Finlayson, S. D. Hordley, C. Lu, and M. S. Drew, "On the removal of shadows from images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, pp. 59-68, 2006.

[15] J. Hernandez-Andres, R. L. Lee, and J. Romero, "Calculating Correlated Color Temperatures Across the Entire Gamut of Daylight and Skylight Chromaticities," *Applied Optics*, vol. 38, pp. 5703-5709, 1999.

[16] S. T. Henderson, *Daylight and its spectrum*, 2 ed. Bristol: Adam Hilger Ltd, 1977.

[17] D. K. Lynch and W. Livingston, *Color and Light in Nature*. Cambridge: Press Syndicate of the University of Cambridge, 1995.

[18] R. Gershon, A. D. Jepson, and J. K. Tsotsos, "From [R, G, B] to Surface Reflectance: Computing Color Constant Descriptors in Images," in *Proc. Int. Joint Conf. Artificial Intelligence*. Milan, 1987, pp. 755-758.

[19] J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*: CRC Press, 2006.

[20] K. Irie, A. E. McKinnon, K. Unsworth, and I. M. Woodhead, "A technique for evaluation of CCD video-camera noise (Accepted for publication)," *IEEE Trans. Circuits and Systems for Video Technology, to be published*, 2007.

[21] K. Irie, "Shadow removal method." NZ Pat Appln No 551684: Lincoln Ventures Ltd, 2007.