Bingham, Mark, Taylor, David, Gledhill, Duke and Xu, Zhijie

Illuminant Condition Matching in Augmented Reality: A Multi-Vision, Interest Point Based Approach

Original Citation

Bingham, Mark, Taylor, David, Gledhill, Duke and Xu, Zhijie (2009) Illuminant Condition Matching in Augmented Reality: A Multi-Vision, Interest Point Based Approach. In: Proceedings of the 2009 Sixth International Conference on Computer Graphics, Imaging and Visualization: 11-14 August 2009, Tianjin, China. IEEE Computer Society, London, UK, pp. 57-61. ISBN 9780769537894

Illuminant Condition Matching in Augmented Reality: A Multi-Vision, Interest Point Based Approach

M. Bingham, D. Taylor, D. Gledhill, Z. Xu
University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, United Kingdom
m.k.bingham@hud.ac.uk, d.taylor@hud.ac.uk, d.gledhill@hud.ac.uk, z.xu@hud.ac.uk

Abstract

For the output of an augmented reality application to appear realistic a number of issues need to be taken into consideration. The illumination correspondence between the real and virtual components should be taken into account as well as the scene level of detail and the accuracy of alignment between the two worlds. This paper focuses on matching world illumination and photometric registration methods. It introduces a new technique that aims to utilize shadow/object interest point correspondences in order to locate and virtually reproduce real-life illuminants. The technique is attractive as it makes use of natural calibration objects in the form of natural scene geometry and associated shadows. Computational complexity is kept relatively low by using an interest point based approach. Further work to be undertaken is discussed.

Keywords— **augmented reality, realism, interest points, shadow**

1 Introduction

1.1 Augmented Reality

Augmented reality (AR) is the term used to describe the concept of superimposing virtuality over images of the real world, effectively combining real and artificial environments. AR has many areas of application and in recent years the field has begun to receive interest from a number of sectors such as manufacturing, military, medical and the computer games industry [2]. One example is the Battlefield Augmented Reality System (BARS) developed by the US Naval Research Facility. BARS is a wearable device that attempts to gather intelligence from, and provide real-time information on, a soldier's surroundings using augmented reality [5]. AR gaming applications such as ARQuake [16] have been developed and allow users to interact with virtual enemies in their own every day environment. The authors of [17] have also implemented an AR system known as Tinmith which allows the user to construct AR outdoor structures via visually tracked hand movements. The Tinmith system has since been adapted for a number of applications including medical, security, entertainment, navigation, shopping, maintenance and has military potential [3]. The alignment between the real and virtual worlds must be accurate in order to achieve realistic augmentation. The process of obtaining such alignment is known as geometric registration. A number of approaches have been proposed that use either sensor data, visual cues or a hybrid combination of both. Photometric registration is the matching of light conditions between worlds. This involves detecting the pose and light qualities of any illuminant effecting the real component of the scene. A number of photometric registration methods have been explored. Researchers attempt to estimate real-world illumination conditions by gathering various metrics from the real scene. This data is then used to illuminate the artificial component. An overview of such techniques is given in section 2.

1.2 Problems and Assumptions

Despite recent advances in virtual reality systems the believable integration of real and virtual components is still a challenge. The realism of an augmented reality scene is massively dependant on robust geometric and photometric registration. The geometric problem has been mostly solved by use of either fiducial markers or interest point based tracking however photometric registration is still problematic. This is primarily due to the unpredictable nature and complexities of the real world. Existing photometric registration techniques have limitations. Such limitations include high computational complexity, the need to pre-calibrate the scene, continuous artificial object based calibration during runtime and constraints in the operational environment. Computationally complex techniques would take too long to perform the necessary calculations. Any lag time this induced would cause desynchronization between the two worlds, ultimately reducing realism and would therefore not be suitable for augmented reality application. If it took too long to recalculate an illuminant position the virtual lighting conditions may not match the real conditions for some time. Or worse, the computation may reduce the output framerate causing the scene to jitter or freeze. Techniques that require pre-calibration are often less computationally intense, however they make the as-

978-0-7695-3789-4/09 $25.00 © 2009 IEEE
DOI 10.1109/CGIV.2009.11

sumption that lighting conditions and camera position do not change. A number of techniques only operate under certain conditions. For example, in a room of known geometry where the light sources have been manually positioned in the virtual scene. Techniques that require constant calibration at runtime require some form of calibration object. These objects are usually unnatural in appearance and therefore destroy the believability of the scene in the same way that the deployment of fiducial marker would when geometrically registering a scene. Figure 1 shows augmentations under various illumination conditions. Image A shows an unlit augmentation. In this state the augmentation does not appear three-dimensional and therefore will look out of place within the scene. Image B and C show the same object manually lit using a fiducial marker to position the light source. When lit directionally the object appears more realistic however may still look out of place in certain illumination conditions. Image D shows the object lit using similar conditions to the real surface on which it is placed. When lit in this manner artificial objects appear more congruent with the real world. Successful illuminant tracking would allow case D to be in effect continuously.

This paper focuses on photometric registration and conceptually outlines an interest point based technique that would allow for fast detection of a single illuminant using natural calibration geometry and shadows present within a scene.

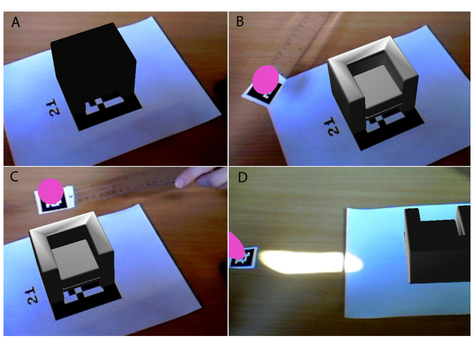

Figure 1: AR Illumination Conditions

2 Related Work

2.1 Condition Matching

Literature shows that researchers have attempted to photometrically register augmented reality worlds in a number of ways. Most existing techniques work well in constrained environments but fail if certain conditions are not met. Other techniques operate well in less constrained environments, however computational complexity is high and therefore unfeasible for real-time augmented reality processing. Feng [6] suggests a technique that makes use of spheres with Lambert surfaces as calibration objects in order to gather illumination parameters. Figure 2 shows the uniform reflectance of a Lambert surface that makes such calibration possible.

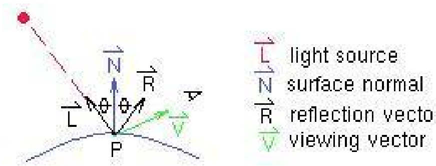

Figure 2: Lambert Illumination Model

The author claims to achieve an identical match between real and virtual components, the result being a seamless augmented reality scene. This technique operates in real-time with relatively low operational complexity but fails if multiple real light sources are present. The technique is not suited for combination with any geometric registration approach as a stationary camera is required once pre-calibration has taken place. Feng does not observe or attempt to reproduce cast shadows. Jacobs [7] presents a real-time rendering solution that simulates colour consistent virtual shadows in real-scenes. Shadow regions are estimated and then confirmed using texture information and are segmented using canny edge detection. A binary mask is then used to track which pixels contain shadow information. This information is then analyzed and the data gathered is passed to either a shadow volume or shadow map algorithm which allows for the casting of shadows from artificial objects. Shadows are cast onto both virtual and real objects. Jacobs technique out performs other shadow matching techniques as the technique correctly combines real and virtual shadows without producing an unrealistic overlap. Feng [8] identifies a number of illumination methods for augmented reality and classifies them into two categories. These are common illumination and relighting. Common illumination matching techniques attempt to simulate consistent lighting when artificial objects are inserted into a real context. Relighting techniques modify the real component in response to the insertion of a virtual object. He performs both by making use of the inverse illumination technique discussed by Patow [18]. This technique collects illumination parameters such as the Bidirectional Reflectance Distribution Function (BRDF) from the real scene for use within the virtual. The technique requires that approximate knowledge of real scene geometry be known prior to augmentation. State et al [9] propose an AR system that favors the use of shadow maps and Haller [10] suggests the use of shadow volume techniques. Both techniques allow the AR system to simulate shadows at low operational cost, after the real-world illumination data has been acquired. Yao [11] and Siala [12] present methods of locating shadows within an image, but do not perform any analysis of the data obtained. Wang [13] presents a method of detecting multiple illuminants within a scene and accurately estimating their pose. This method does not require the use of a pre-calibration object. Additionally the data collected from the technique allows for the virtual recreation of three dimensional object shapes. The illuminant detection results the technique yields are directly applicable to development of realistic augmented reality systems, however the calculations required are slow and therefore would not be capable of processing a live video stream in real-time. Wang's technique provides good results compared to a number of other techniques as it analyzes both shadows and the shading of arbitrary scene objects. The technique finds it easy to obtain multiple illuminant information from shading when specular reflections are present but finds the task difficult when observing diffuse reflections alone. An extra level of robustness exists with this technique as it is less prone to error caused by cast shadows moving outside the camera's field of view or being occluded than techniques that observe either object shading or cast-shadows exclusively. Zhang [14] presents a robust method of estimating the azimuth of a single illuminant.

Successful interest point (IP), shadow and object detection and geometric registration are a pre-requisite of the technique presented within this paper. An interest point is defined as a two-dimensional signal change; for example, where there is a corner, an edge or where the texture changes significantly [1]. Much work has been undertaken in the field of interest point detection. Many IP detection techniques have been presented by Harris and Stephens [2] and Lowe [15]. Much has been accomplished in the field of detecting and segmenting shadows and object within images and video footage. The task of geometrically registering AR by aligning the real and virtual worlds has been achieved using fiducial markers as per [4] and by tracking world interest points as detailed by State [9].

3 Multi-vision IP based Technique

The new technique aims to detect illumination conditions by studying shadow and object interest points within a scene. The technique assumes only one main illuminant is present and scene geometry is casting shadows onto planar surfaces. The technique works by observing correspondences between shadow and object interest points and estimating illuminant position in 2D for two input images then combining these results to obtain 3D coordinates. The resulting data can be used to recreate illumination conditions for use within an augmented reality scene.

The primary stages are:

1. Acquire Images
2. Identify IPs
3. Identify IP Correspondences
4. Locate illuminant direction vector from two different angles
5. Locate illuminant in 3D space
6. Perform geometric registration
7. Augment reality

Images of the real scene are obtained from two camera input devices simultaneously. The images can be taken from any angle. It is anticipated that the more acute the angle the less accurate the result. This prototype system uses two cameras observing the scene at an angle of 90 degree separation. The input images used in this paper are 3D virtual renderings as opposed to actual webcam images. This allows for the production of control images that we can eventually use to verify the accuracy of our results.

Interest points may be detected using one of the techniques discussed in section 2. The Scale Invariant Feature Transform (SIFT) method presented by Lowe [15] is preferred as it makes available additional information that is of use when detecting the correspondence between shadow interest points and object interest points. Figure 3 shows SIFT features extracted from an image. The image shows many more SIFT descriptors than would be required for the intended purpose. SIFT allows for a threshold to be applied that would reduce the number of feature extractions, limiting output to corners on objects or shadows within the scene.

Figure 3: SIFT Extracted Features

Once interest points have been obtained they need to be classified as being associated with either a cast shadow or object geometry. Correspondences between geometry IPs and shadow IPs need to be defined. We need at least

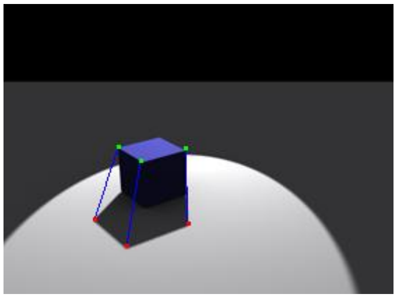

Figure 4: Shadow / Object IP Correspondences

two correspondences per image but more will offer improved accuracy. Figure 4 shows correspondences between shadow and object IPs.

If we draw a line from the shadow interest point towards the corresponding object interest point and beyond we can obtain the approximate location of the illuminant in 2D coordinate space. Intersections between two or more lines indicate a potential illuminant position. Any anomalous results are discarded and the average intersection point is recorded as the 2D illuminant position. It should be noted that the actual illuminant position may be outside of the image boundary, this does not pose a problem. Figure 5 shows the correspondence lines and detected illuminant positions for two images.

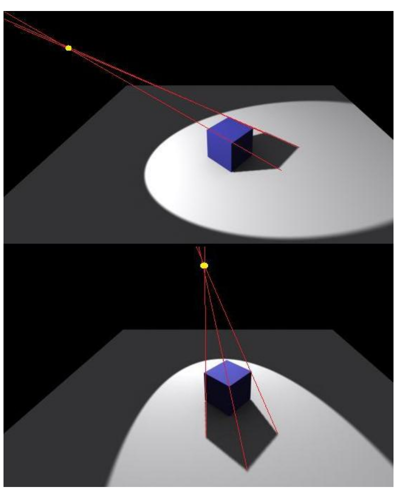

Figure 5: Illuminant Detection: 2D

Assuming we know the angle between the two input images we can position them within a three dimensional scene adjacent to each other and cast lines through the plane of each image. Slight inaccuracies can be expected therefore to ensure these two lines actually intersect an average height value is taken. The point of intersection is the approximate illuminant position in three dimensional space. Figure 6 shows the expected results. Once geometric registration has been applied this location will refer to the position of the real illuminant as shown by the sphere (at the 3D intersection point) in the image.

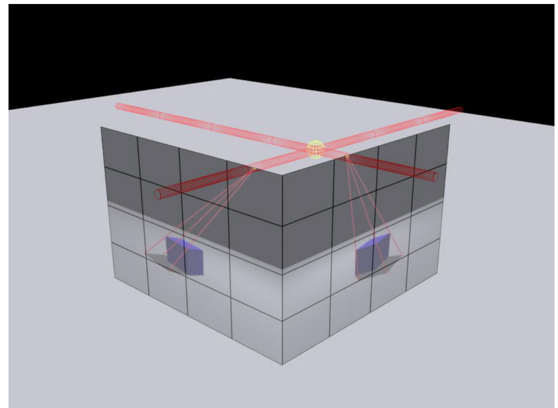

Figure 6: Illuminant Detection: 3D

Once the real light source has been located we can perform the augmentation. Geometric registration techniques should be applied to ensure the coordinate systems of the real and virtual worlds are in accurate alignment. Any augmentations can now be correctly lit and should appear to be illuminated in a similar manner as the real scene. The light source can be fed into AR shadow casting techniques such as those presented by Jacobs [7] and Haller [10]. Shadows cast by virtual objects should be at the correct angle and of the correct length. The scene can then be described as both geometrically and photometrically registered and will therefore appear more realistic in nature.

4 Conclusions and Future Work

As geometric registration techniques are now sufficiently accurate for reliable augmented reality use, the focus has recently moved onto methods of improving the realistic feel of augmented reality systems. Researchers are now investigating ways of photometrically registering worlds. A number of techniques exist, however they are often computationally complex and although they may work well for static images they are of little use for true augmented reality simulations as these generally require real time processing of video streams. Other techniques require some form of calibration. Such calibration is usually unfavorable as usability constraints are normally introduced or the realism is destroyed in the process. The technique outlined in section 3 is of low computational complexity and does not constrain the operational environment beyond the assumption that an illuminant is casting shadows somewhere within the scene.

The technique introduced in this paper is currently being refined. New methods of automatically matching corresponding shadow and object interest points are being investigated. At present the technique relies on prior knowledge of the angle between views in order to derive 3D coordinates from the two potential illuminants in 2D space. If the technique could adapt to function with any arbitrary angle then the requirement of two input devices could be eliminated. Instead, the two input images could be taken from a video stream where the camera is in constant motion. This would provide the system with observation from different angles. A number of issues with this approach would arise. For example, objects within the scene may be repositioned between the two reference frames chosen. However an adaptive approach that would yield results of sufficient accuracy for augmented reality application could be possible. The robustness of the introduced technique is to be studied in detail. This includes studying its susceptibility to occlusions and extreme lighting conditions.

References

[1] D. Gledhill and G. Tian and D. Taylor and D. Clarke.*Panoramic Imaging - A Review.* Computers and Graphics, 2003.

[2] C. Harris and M. Stephens.*A Combined Corner and Edge Detector.* 4th Alvey Vision Conference, 1988.

[3] B. Thomas and C. Sandor.*What Wearable Augmented Reality Can Do For You* IEEE Pervasive Computing, 2009.

[4] X. Zhang and S. Fronz and N. Navab.*Visual Marker Detection and Decoding in AR.* In Proceeding International Symposium on Mixed and Augmented Reality, 2002.

[5] T. Cowper and M. Buerger.*Improving Our View of The World: Police and Augmented Reality Technology*.

[6] Y. Feng.*Estimation of Light Source Environment for Illumination consistency or Augmented Reality.* In Proceedings 2008 IEEE Congress on Image and Signal Processing, 2008.

[7] K. Jacobs and C. Angus and C.Loscos.*Automatic Generation of Consistent Shadows for Augmented Reality.* In Proceedings Graphics Interface, 2005.

[8] K. Jacobs and C. Loscos.*Classification of Illumination Methods for Mixed Reality.* In Proceedings Computer Graphics Forum, 2004.

[9] A. State and G. Hirota and C. Chen and W. Garrett and M. Livinston.*Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking.* In Proceedings of the COmputer Graphics Forum, 2006.

[10] M. Haller and S. Drab and W. Hartman.*A Realtime Shadow Approach for An Augmented Reality Application Using Shadow Volumes.* In Proceedings ACM Symposium on Virtual Reality Software and Technology, 2003.

[11] J. Yoa and Z. Zhang.*Systematic Static Shadow Detection.* In Proceedings 17th IEEE International Conference on Pattern Recognition, 2004.

[12] K. Siala and M. Chakchouk and O. Besbes and F. Chaieb.*Moving Shadow Detection with Support Vector Domain Description in the Color Ratios Space.* In Proceedings of the 17th IEEE International Conference on Pattern Recognition, 2004.

[13] Y. Wang and D. Samaras.*Estimation of Multiple Directional Light Sources for Synthesis of Augmented Reality Images.* Special Issue on Pacific Graphics, 2003.

[14] X. Zhang and S. Fronz and N. Navab.*Visual Marker Detection and Decoding in AR Systems: A Comparative Study.* In Proceedings International Symposium on Mixed and Augmented Reality, 2002.

[15] D. Lowe.*Distinctive Image Features from Scale-Invariant Keypoints.* International Journal of Computer Vision, 2004.

[16] M. Tuceryan and D. Greer and R. Whitaker and D. Breen and C. Crampton and E. Rose and K. Ahlers.*Calibration Requirements and Procedures for a Monitor Based Augmented Reality System.* IEEE Transactions on Visualization and Computer Graphics, 1995.

[17] W. Piekarski and B Thomas.*The Tinmith System: Demonstrating New Techniques for Mobile Augmented Reality Modeling.* Computuer Science Community, 2002.

[18] G. Patow and X. Pueyo.*A Survey of Inverse Rendering Problems.* In Proceedings Computer Graphics Forum, 2006.