Estimation of Dynamic Light Changes in Outdoor Scenes Without the use of Calibration Objects

Mikkel Sandberg Andersen Tommy Jensen Claus B. Madsen
Laboratory of Computer Vision and Media Technology
Aalborg University. DK-9220 Aalborg, Denmark
{msan01|toje01}@control.auc.dk, cbm@cvmt.dk

Abstract

The the work presented in this paper explores how dynamical light parameters in an outdoor environment can be estimated for use in a real-time augmented reality (AR) system. A method using existing inverse rendering techniques is used to acquire diffuse surface reflectances in an offline procedure. The reflectances are used in an on-line procedure to estimate the illumination parameters of an outdoor scene. The method presented reduces the light estimation problem of outdoor scenes to a modified Phong shading model with two unknown parameters, which can be determined through the use of a linear equations system. The work presented provides an elegant method for estimating dynamically changing illumination parameters without the need for a calibration object in the scene.

1. Introduction

Picture a yard where a statue has been placed, bathed by the sun, casting a shadow on the ground. There is nothing in the image pointing to the fact that the statue is not real. A cloud suddenly blocks the sun, and the lighting of the yard changes. But somehow the statue seems to be lit with greater intensity than the surroundings, and the shadow it casts on the ground has not faded into the shadow from the cloud. It becomes apparent that the statue in the centre of the yard is nothing more than a virtual object rendered on top of the video feed of a real yard.

The above scenario represents a typical problem in the implementation of a live Augmented Reality (AR) system. This raises the need for a system with the capability to adapt the lighting of the virtual objects to the dynamically changing lights of the surrounding environment. This paper proposes a method to estimate the lighting changes of an environment. The method is implementable in an augmented reality system used to augment virtual objects into a real outdoor scene and shade the object according to how the method estimates the lighting of the scene.

The outline of the paper is as follows: Section 2 will describe existing state of the art work, and conclude which progresses may be needed in the field of AR. Section 3 offers a description of how the suggested method estimates the lighting of a given scene, and represents this knowledge. And finally the paper is concluded by a presentation and discussion of the results.

2. State of the Art

Estimating scene illumination from images is the problem of estimating surface reflectance properties, because an image represents light reflected off surfaces, and which is governed by illumination and reflectances. This is why illumination estimation cannot be performed without knowledge of surface reflectance. For this reason most related work is based on placing some kind of calibration object with a priori knowledge of reflectance properties of the calibration object. Subsequently we briefly describe some of the most closely related work.

[3] suggested that the illumination distribution of a scene could be estimated by analysing shadows cast by a known calibration object onto a known surface.

[6] rendered photo-realistic architectural scenes under novel lighting conditions based on analysis of real imagery of the scene taken at different times of day.

Furthermore [1] presented a method for measuring scene radiances as a High Dynamic Range Image (HDRI) and adding virtual objects to a scene with correct lighting, using the HDRI environment map.

[2] designed an approach for automatic, real-time estimation of scene lighting for augmented reality. The approach involves placing a reflective sphere which is always in the camera's field of view. The dynamic scene illumination conditions are estimated from the environment's reflection in this special purpose sphere.

As seen from the above review, the standard approaches to determine the illumination conditions of a scene are to either have a light probe in the scene, or a calibration object. Common for both methods are that they require some sort of scene modification. The approach presented in this paper aims to develop a method that does not require scene modification, by using existing objects in the scene for calibration.

3. Estimation of scene lighting

Augmenting virtual objects into a real scene, taking into account lighting changes of the scene, presents a twofold task: Light estimation, and a rendering pipeline. This paper focuses on the light estimation part of such a system.

The illumination estimation method used to estimate the radiances of our scene has a number of constraints and assumptions for the scene, which are listed below.

- Outdoor during daytime
- Predefined diffuse surfaces
- HDRI Light probe (Environment Map)
- Known sun position
- Simple 3D model with most significant surfaces

The method assumes that the sun is the only major light source in the outdoor scene, and therefore the only direct light source needing estimation, with the sky providing secondary lighting estimated as ambient light. Furthermore the scene must contain diffuse surfaces, as these will be the sources to estimation of the scene lighting. Finally, as the light is estimated from the images recorded by a camera of the scene, the camera needs to be calibrated to fit the scene.

When all the prerequisite data, listed above has been collected, the illumination estimation is able to analyse images of the scene taken by a camera. This is done using 500 randomly selected pixel samples. Assumimg that a sun model provides the direction vector to the sun, the method is able to estimate the light intensity of both direct and indirect light in the scene, if the camera has surfaces in both light and shadow within its frame. E.g. the method will estimate the RGB intensity of the sun to almost zero, when there is a heavy cloud cover, because it sees no noticeable difference between the area in direct light, and the area, that should be in shadow.

3.1. Basic Illumination Estimation

In order to estimate light, a light model is needed. There are a number of ways to represent knowledge of the lighting of an environment. In this paper the Phong shading model is the basis for the light estimation. The radiance at a point $\vec{x}$ in a scene is given by the rendering equation:

$$L_o(\vec{x},\vec{\omega_o}) = \int_{\Omega_i} f(\vec{x},\vec{\omega_i},\vec{\omega_o}) \cdot L_i(\vec{x},\vec{\omega_i}) \cdot (\vec{n}(\vec{x}) \bullet \vec{l}(\vec{x})) \quad d\vec{\omega_i} \tag{1}$$

As mentioned the illumination estimation uses diffuse surfaces to determine the lighting of a scene, hence the diffuse BRDF can be employed in the rendering equation:

$$L_o(\vec{x}) = \frac{\rho_d(\vec{x})}{\pi} \sum_{n=1}^{N} L_{i,n}(\vec{x}) \cdot (\vec{n}(\vec{x}) \bullet \vec{l}(\vec{x})) \tag{2}$$

As the illumination estimation is limited to an outdoor scene during daytime, it is assumed that the light can be divided into two parameters - direct light from the sun, and ambient light reflected off various surfaces in the scene as well as light from the sky-dome. With this assumption the rendering equation may be rewritten to:

$$L_o(\vec{x}) = \frac{\rho_d(\vec{x})}{\pi} \cdot \Big(L_a + L_d \cdot (\vec{n}(\vec{x}) \bullet \vec{l}(\vec{x})) \Big) \tag{3}$$

Equation 3 holds many unknown factors, of which some can be determined in advance.

The diffuse reflection coefficient (the albedo value) at each point in the image, $\frac{\rho_d(\vec{x})}{\pi}$, can be precalculated through an inverse rendering, from a directional HDRI environment map mapped onto the 3D model of the scene. Furthermore the normal of the surfaces in the environment, $\vec{n}(\vec{x})$, can be determined from the 3D model.

The direction vector for the incoming light ($\vec{l}(\vec{x})$), is determined by the use of a sunlight model. The sunlight vector is assumed to be the same in the entire scene, due to the distance from Earth to the Sun is of a quantity that can be considered infinite with regards to the distances within the local scene.

This leaves the radiances as unknown variables. The outgoing radiance from point $\vec{x}$, $L_o(\vec{x})$, is sampled from an input image of the the scene. The outgoing radiance from each point will be the direct radiance of the point scaled by a constant, if the camera has linear response. This scaling constant will be denoted as k, while the unit-less value provided by the camera is $P(\vec{x})$. This yields the basic rendering equation:

$$P(\vec{x}) = k \cdot \frac{\rho_d(\vec{x})}{\pi} \cdot \Big(L_a + L_d \cdot (\vec{n}(\vec{x}) \bullet \vec{l}(\vec{x})) \Big) \tag{4}$$

This leaves the two light-parameters L_a and L_d as unknown variables, both scaled by the constant k. With a suitable number of samples, the light parameters can be estimated using a linear equation system.

3.2. Specific Illumination Estimation

The ambient light, L_a, is normally a constant light contribution to the entire scene. The outcome of this simplification is that all points not lit by direct light are equally weighted, which is not consistent with reality. The light of every point in a scene is the result of an integration

(a) (b) (c) (d)

Figure 1. Frames from the test. a, b from the 100 frames virtual set c,d from the 300 frame real set.

over the surface's entire hemisphere, meaning that vertical surfaces receives less light from the sky than horizontal surfaces. This aspect must be taken into account in the rendering equation, if the estimation of the ambient light is to be more accurate than the standard ambient term.

A way to handle this aspect, is to pre-calculate the hemisphere of all points in the environment yielding an ambient occlusion map. If the visibility hemisphere for every point on the environment map is pre-calculated in such a way that surfaces with full visibility has a value of 1 and points with no light has the value 0, the visibility hemisphere aspect can be included by using this value as a weight to the ambient light [4]. The ambient occlusion is denoted as c, and the model in equation 4 can be rewritten as:

$$P(\vec{x}) = k \cdot \frac{\rho_d(\vec{x})}{\pi} \cdot \Big(c \cdot L_a + L_d \cdot (\vec{n}(\vec{x}) \bullet \vec{l}) \Big) \quad (5)$$

Another aspect that must be handled, is that certain surfaces are in shadow. This is pivotal to the success of the illumination estimation, as this determines the difference between which samples are used for estimation of ambient and direct light.

In order to determine which areas are in shadow, a binary shadow map of the scene must be constructed for every frame. To construct such a map the 3D model of the scene is used with a shadow volume algorithm to cast shadows upon itself. The sun-vector supplied by the sunlight model is used as light direction.The shadow map, $c(\vec{x})$, is incorporated into equation 5 which yields:

$$P(\vec{x}) = k \cdot \frac{\rho_d(\vec{x})}{\pi} \cdot \Big(c(\vec{x}) \cdot L_a + s(\vec{x}) \cdot L_d \cdot (\vec{n}(\vec{x}) \bullet \vec{l}) \Big) \quad (6)$$

The camera image supplied, represents a section of the albedo environment map. Assuming that the viewing direction and field of view, of the camera supplying the input images, are known, retrieving the albedo value for a given image point can be done through a simple look-up in the albedo environment map.

With a known albedo value and the pixel colour, $P(\vec{x})$, of a given image point in the input image, the unknowns of equation 6 are reduced to L_a and L_d.

To derive the light parameters of the scene several samples are taken in the input image. These samples are applied to the linear equation system in 7:

$$\begin{bmatrix} P(\vec{x_1}) \\ P(\vec{x_2}) \\ \vdots \\ P(\vec{x_M}) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\rho_d(\vec{x_1})}{\pi} \cdot c(\vec{x_1}) & \frac{\rho_d(\vec{x_1})}{\pi} \cdot s(\vec{x_1}) \cdot (\vec{n}(\vec{x_1}) \bullet \vec{l}) \\ \frac{\rho_d(\vec{x_2})}{\pi} \cdot c(\vec{x_2}) & \frac{\rho_d(\vec{x_2})}{\pi} \cdot s(\vec{x_2}) \cdot (\vec{n}(\vec{x_2}) \bullet \vec{l}) \\ \vdots & \vdots \\ \frac{\rho_d(\vec{x_M})}{\pi} \cdot c(\vec{x_M}) & \frac{\rho_d(\vec{x_M})}{\pi} \cdot s(\vec{x_M}) \cdot (\vec{n}(\vec{x_M}) \bullet \vec{l}) \end{bmatrix} \begin{bmatrix} k \cdot L_a \\ k \cdot L_d \end{bmatrix} \quad (7)$$

As seen in equation 7 two samples would suffice in order to derive an estimate of L_a and L_d to a scaling factor. Samples are only added to the equation system if it passes a goodness test.

4. Tests and Results

The method presented is based on two main processes. An *offline* part which is the collection and preparation of input data, and an *on-line* part where the direct and ambient radiances can be estimated on a per-frame basis. Please note that the term *estimated radiance* in this context actually refers to the scaled estimated radiances.

The tests have been performed on both simulated, i.e. computer rendered data, and on real images shot as time lapse sequences with a standard consumer grade digital camera.

Testing on simulated data has the apparent advantage of giving a completely controlled environment. This means that we are able to conclusively determine the reflection coefficients and the actual values of direct and ambient radiances present. These can be compared to the estimated radiances produced by the proposed method for a basic verification of functionality. Shown in Fig. 3.1.a and Fig. 3.1.b are two images from a simulated test sequence of 100 frames. A visual comparison of a graph of the original radiances from the rendered input sequence, and the per-frame estimated radiances from the proposed

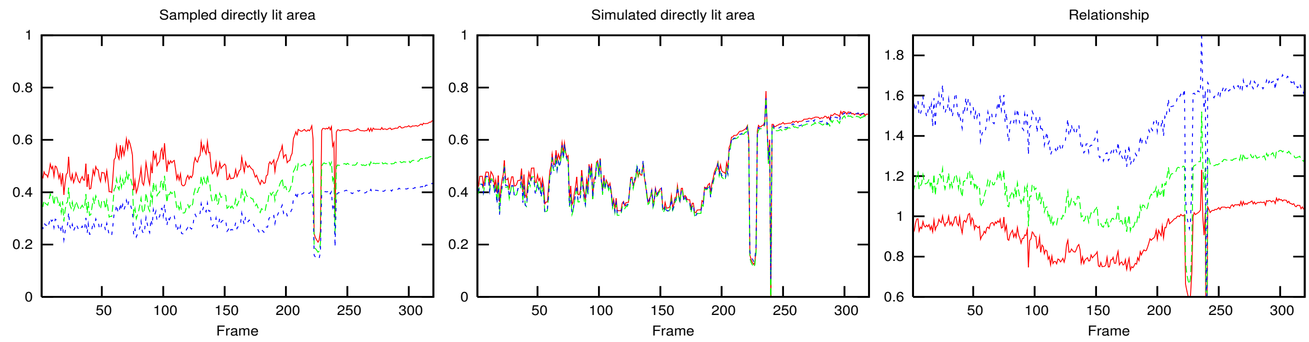

Figure 2. Left graph shows average pixel intensities sampled in a directly lit area. Middle graph shows simulated pixel intensities in a directly lit area. Right graphs is the relationship between the two curves. All data are normalised.

method clearly shows a correlation between the known and the estimated radiances.

Testing on real data involves a slightly more complicated process to obtain ground truth about the actual radiances present in the photographed images. A 300 frame sequence, of which 2 frames are shown in Fig. 3.1.c and 3.1.d, was used. To verify the results of a test in real data, the test involves two steps, first of which is to estimate the radiances. Step two is to verify the quality of the estimates. This is done by sampling the average intensity of a selected surface in the original image sequence. The results are compared with a "virtual surface" inserted at the same location and lit by the chosen light model and the estimated radiances. Fig. 2 shows the relationship between the average intensity of the original surface, and the "virtual surface".

In the ideal situation, the intensities would exhibit a constant relationship. Fig. 2 shows that this is almost the case except for the last 40 frames. This deviation is most probably caused by the quality of the test material. The two main problems are that the stone surface on the ground is assumed to be diffuse and have a vertical surface normal. Inaccuracies in these two parameters affects the estimates especially towards the end of the sequence, since the direct light estimation relies almost solely on that surface.

5. Discussion

When moving towards photo-realism within the field of Augmented Reality (AR), ne must keep in mind, that AR, as opposed Virtual Reality (VR), co-exists with the real world. Consequently, to maintain the illusion that virtual objects are a part of the real world, phenomenas inherent to the real world must be considered. This paper has focused on one major factor in supporting this illusion; illumination. The initial goal, to estimate light parameters in an outdoor environment with both predictable and unpredictable dynamic changes, has been researched through this work.

The method presented in this paper is able to use existing geometry to update parameters of a light model, leaving the scene unmodified. The principle of handling dynamic light changes in an Augmented Reality (AR) system by updating light parameters as presented, is not seen in existing research.

Two movie sequences of the method used in an OpenGL-based rendering system can be downloaded from `http://www.control.aau.dk/~toje01-nobackup`.

References

[1] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. 1998.

[2] M. Kanbara and N. Yokoya. Real-time estimation of light source environment for photorealistic augmented reality. In *Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, United Kingdom*, pages 911–914, August 2004.

[3] I. Sato, Y. Sato, and K. Ikeuchi. Illumination distribution from shadows. In *Proceedings: CVPR99*, pages 306–312, 1999.

[4] A. Whitehurst. Depth map based ambient occlusion lighting, 2001. `http://www.andrew-whitehurst.net/amb\_occlude.html`.

[5] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: recovering reflectance models of real scenes from photographs. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 215–224, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[6] Y. Yu and J. Malik. Recovering photometric properties of architectural scenes from photographs. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 207–217, New York, NY, USA, 1998. ACM Press.