EUROGRAPHICS 2008 / G. Drettakis and R. Scopigno
(Guest Editors)

Volume 27 (2008), Number 2

Stereo Light Probe

Massimiliano Corsini, Marco Callieri and Paolo Cignoni

Visual Computing Lab, ISTI - Area della ricerca CNR
Via G. Moruzzi 1, 56124, Pisa, Italy
Email: {corsini, callieri, cignoni}@isti.cnr.it

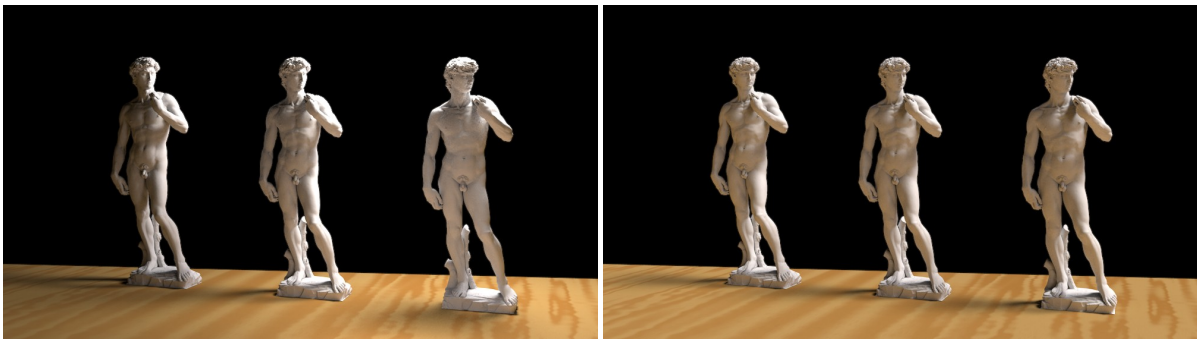

Figure 1: *(Left) Three models of the Michelangelo's David illuminated with a spatially-varying lighting environment captured with the proposed Stereo Light Probe device. (Right) The same scene with a lighting environment captured with the classical approach of single reflective ball.*

Abstract

In this paper we present a practical, simple and robust method to acquire the spatially-varying illumination of a real-world scene. The basic idea of the proposed method is to acquire the radiance distribution of the scene using high-dynamic range images of two reflective balls. The use of two light probes instead of a single one allows to estimate, not only the direction and intensity of the light sources, but also the actual position in space of the light sources. To robustly achieve this goal we first rectify the two input spherical images, then, using a region-based stereo matching algorithm, we establish correspondences and compute the position of each light. The radiance distribution so obtained can be used for augmented reality applications, photo-realistic rendering and accurate reflectance properties estimation. The accuracy and the effectiveness of the method have been tested by measuring the computed light position and rendering synthetic version of a real object in the same scene. The comparison with standard method that uses a simple spherical lighting environment is also shown.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Illumination Estimation, Light Fields, Image-Based Lighting, Reflectance and Shading I.3.7 [Three-Dimensional Graphics and Realism]:

1. Introduction

In recent years the convergence of Computer Vision and Computer Graphics is becoming more and more evident and important both from a theoretical point of view to develop new results in the field of appearance acquisition and modeling and for the development of advanced graphics applications. One of the main task of this convergence concerns the acquisition and rendering of the reflectance properties of a given scene. Only an accurate estimation of the reflectance properties of a surface make possible to produce photo-realistic faithful rendering of acquired surfaces: a critical point of the process of reflectance properties acquisition

is the precise estimation of the illumination incident on the surface of interest. Moreover the knowledge of the lighting environment is important also in many other applications, such as the integration of synthetic objects into real scenes, image-based lighting and relighting.

This task has received many effort by several researches in the past years and many methods and devices to acquire the lighting environments of a scene have been developed. Some of these methods, discussed in the next section, are purely image-based and do not require neither the knowledge of the geometry of the scene nor availability of special optical devices for the acquisition, but on the other hand their use in a standard rendering pipeline is not straightforward. Some other solutions make use of simple optical devices such as a single mirror ball, but are limited in the characteristic of the light sources estimated. Other ones are capable to capture spatially-varying lighting environments but require specific devices such as omni-directional cameras.

Here, we propose to estimate the incident light field in a compact indirect way: we estimate the position, shape and intensity of the light sources in the environment. In particular, we focus our attention on indoor environments where it is important to take into account the spatially-varying nature of the direct incident illumination, i.e. the position and shape of the light sources inside the scene. The base of our idea is to overcome the limitation of the classical spherical lighting environment approach [Deb98] and to propose an equally simple and practical tool based on a couple of mirroring balls, that act as stereo omni-directional rig to simultaneously estimate position, shape and intensity of the main light sources present in the environment.

In the next Section we describe some works related to lighting environment estimation, comparing the existing methods with the proposed one. In Section 3 we present our experimental acquisition device and the underlying algorithms used to robustly estimate the light source positions. The experimental results are shown in Section 4 and the performances and the limits of the proposed method analyzed. The conclusions are outlined in Section 5.

2. Related Work

In this Section we analyze several previous works that focus on illumination estimation and lighting environment acquisition.

Currently, the approach that is most commonly used in practice for lighting environment acquisition for the integration of synthetic and real objects has been proposed by Paul Debevec [Deb98]. In this work the scene of interest is subdivided in three fundamental components: the distant scene, the local scene and the synthetic object. The distant scene is modeled as a light-based model which illuminates the local scene but does not interact with it. Such light-based model is a spherical lighting environment acquired using a single mirrored ball. The reflectance properties of the local scene is estimated through inverse rendering using an approximate geometric model. The geometry and the BRDF of the synthetic scene is known.

Sato et al. [SSI99] proposed a technique to account for spatially-varying illumination by acquiring the radiance distribution of a three-dimensional scene with two omni-directional camera. The two input images are analyzed using a feature-based stereo matching algorithm. The correspondences found in the two images are used to triangulate the radiance distribution. The output of the algorithm is a triangular mesh where each vertex represents a light source of the environment. Such mesh is used to re-illuminate synthetic objects that can be inserted in the real scene, obtaining a consistent illumination. Our proposal follows a similar approach but relies on a cheaper device and overcome some robustness limitations. In another work Sato et al. [SSI03] proposed a method to estimate the lighting environment by analyzing the shadows contained in the images of the scene. The 3D model of the scene is supposed known.

More recently, Unger et al. [UWH*03] proposed two new devices for the acquisition of spatially-varying illumination. The first one, named *Mirror Sphere Array*, is composed by a planar set of mirrored balls. A set of snapshots of such spheres taken with a digital camera could be used to obtain the estimation. In particular, a specific rendering procedure presented in the same work has been developed to use directly the data obtained in the rendering phase. The second device is composed by an omni-camera mounted on a mechanical arm that allow two-axis movements to sample the lights in multiple directions and multiple spatial position. This device is able to capture a high-fidelity spatially-varying lighting environment but requires a lot of time for the acquisition.

Another device that uses multiple reflective spheres to estimate the position of a single point shaped light source in controlled laboratory settings has been developed by H. P. Lensch et al. [Len03]. This device is part of a system for the BRDF acquisition of real objects. In particular, the position of the highlight of six reflective spheres is used to build a linear system that solved provide the position of the light source used to sample the BRDF.

As mentioned previously, our acquisition method was inspired by the Sato et al. [SSI99] approach even if it presents some completely different issues and other peculiarities. Our idea is to use two HDR images of two steel balls to estimate the spatially-varying lighting environment instead of using two omni-directional cameras. In particular our work introduces a different and more robust approach for the light position estimation that does not rely on feature-based algorithm for stereo matching but adopts a region-based matching to compute correspondences on the input images. We will show in the next Section that our approach is more robust and allow the use of simple and sturdy steel balls instead of high

quality mirroring spheres or costly omnidirectional cameras. Another difference with the work of Sato et al. is that we use a different representation of the radiance distribution. While Sato et al. use a triangular mesh we opt for a set of omnidirectional point lights distributed according to the real light sources positioned inside the scene. This kind of representation allows a simpler integration with existing rendering engines and make our proposal more practical from the enduser perspective. Moreover, the many-lights rendering problems has received increasingly attention during the last years making this representation feasible also for the next-gen rendering engines [HPB07, WABG06, WFA*05].

A detailed overview of recent advance in omni-directional stereo vision is not reported here. Interested people could refer to recent papers about this topic such as [HLH06, SY05, DS02, GD03] for an in-depth examination.

3. Algorithm

Our algorithm consists of five steps. In the first step two high dynamic range images of two reflective balls are acquired by taking multiple shots at different exposure with a digital camera. Each ball is photographed separately. This two HDR images (indicated with I_1 and I_2 in the following) are the only required input of the algorithm. Then, the camera is calibrated using reference points placed around the support of each probe. In the third step the recovered extrinsic and intrinsic parameters of the camera are used to rectify the spherical images of the light reflected by the balls obtaining a rectified reflected stereo pair ($\mathcal{I}_1$ and $\mathcal{I}_2$). The fourth step consists of establishing the correspondences between regions representing the light sources in the left and right rectified images. At this point, the position of each light source can be easily calculated by intersection of the corresponding reflected rays. The absolute intensity of the lights is recovered from the values of the HDR images corrected with the recovered distance of the lights from the light probes. The intensity of the lights is given by the values of the HDR images, since such values are proportional to the radiance of the scene. At the end of these processing steps the radiance distribution of the environments is determined and stored as a set of point light sources in the scene space.

A detailed description of our acquisition device and each phase of the algorithm described are presented in the following subsections.

3.1. The acquisition device

As just previously mentioned our acquisition device consists of two reflective spheres aligned on a support platform and a semi-professional digital camera (see Figure 2). More specifically, we use two steel balls with a reflection coefficient of about 0.6. Chrome balls can provide more accurate results since they are more reflective and polished, nevertheless since this is an experimental device we decide to use

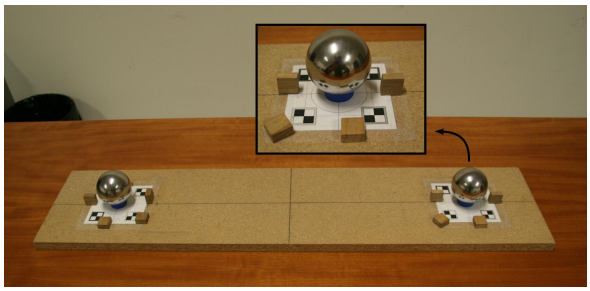

Figure 2: *The experimental acquisition device.*

steel balls that are more robust while maintaining a good reflectivity. The diameter of the steel balls used is 60 mm. The balls are positioned over a calibration pattern shown in the particular of the Figure 2. The distance between the spheres is 65 cm. The left sphere, without loss of generality, is assumed to be the reference one. The origin of the world coordinate system is assumed to be coincident with this sphere. The snapshots of each ball were captured using a Canon DS350 digital camera fixed on a tripod. A macro view is used in order to reduce lens distortion. During the acquisition phase two sets of snapshots, one for the left sphere and another one for the right one, with different exposure are taken in order to produce two high-dynamic range images; in this way we obtain values proportional to the scene radiance. The HDR images are generated using the HDRShop software [Uni]. The lighting environment we were working with did not presented a very high dynamic range like an outdoor scene; hence it was sufficient to use 5 shots to create each HDR image. The exposure settings used are: 1/40th, 1/20th, 1/10th, 1/5th, 0.4 seconds. Due to the intrinsic nature of the spherical reflection, using a single view imply a poor sampling of some portion of the surrounding space. This problem can be alleviated by taking the shots from different viewpoint and blend them together to remove the region of poor sampling. In the following examples to demonstrate the robustness of the proposed approach we have always used a single HDR image, hence a single viewpoint, for each sphere.

3.2. Camera Calibration

The position of the camera with respect to each steel ball is computed using as reference points the corners of reference symbols and small wooden boxes positioned around the support of each sphere. The intrinsic and extrinsic parameters are estimated with the popular camera calibration algorithm of Roger Tsai [Tsa87]. The calibration has to be very accurate in order to minimize problems during the rectification and the next correspondences matching phase. In fact, if the camera parameters are imprecise the mapping of the reflections on the rectified spherical stereo pairs become inexact and the correspondences cannot be calculated reliably. The

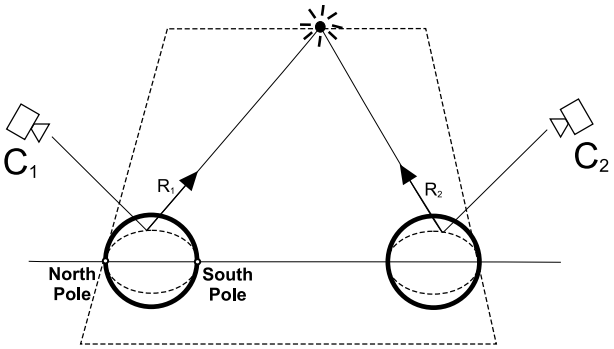

Figure 3: *How stereo light probe works. Schematization.*

use of non planar marker configuration helps to make the calibration process more robust.

After calibration we are able to map a point in world space on the image plane of the two cameras. In particular, we indicate with Π_1 and Π_2 the projection matrix ($\mathbb{R}^{3\times4}$) of the first and second camera respectively, and with M_1 and M_2 the matrix ($\mathbb{R}^{4\times4}$) to convert world coordinates into the respective camera coordinates. In the next section we use the notation $*_{1,2}$ to indicate relations that are valid for both cameras.

3.3. Rectification of Spherical Images

After the calibration phase we are able to process the input HDR images to obtain a stereo pairs from the reflections of the two steel balls. Our spherical reflection rectification algorithm follows the approach proposed in [Li06].

The rectification algorithm works in the following way: imagine you have an ideal point light p in the space. The light reflected by the reference ball reach the cameras C_1 and C_2 as schematized in Figure 3. The reflection vectors $\vec{R}_1$ and $\vec{R}_2$ are constrained on a particular plane. Such plane corresponds to the one passing along the line connecting the two spheres (the baseline) and which lies on the point p. Mapping the point light on the rectified maps is easy by considering the latitude-longitude coordinates system with the North Pole and the South Pole disposed as shown in Figure 3. Intuitively this correspond to a geographical coordinate system with the x and y axis swapped: with this reference frame the light point lies at the same longitude in both rectified maps. In particular, if we indicate with $\theta \in [0, 2\pi]$ the longitude and with $\phi \in [-\pi, \pi]$ the so-called colatitude and we associate θ to the rows of the rectified images, and ϕ to the columns of the rectified images, the search region of the correspondences is simply reduced to the horizontal axis of the rectified images.

This concerns the background of spherical rectification. Now let's go to describe the exact steps of the algorithm. Starting from the input images I_1 and I_2 we compute three maps:

- The **position map** ($\mathcal{P}$) containing for each pixel the position of the sphere. The position is stored in latitude-longitude spherical coordinates (θ_S, ϕ_S).
- The **reflection map** ($\mathcal{R}$) containing for each pixel the direction where the light incident at that point is reflected. Also the direction of reflection is stored as (θ_R, ϕ_R).
- The **rectified reflection map** ($\mathcal{I}$) containing the color of the ray reflected by the point (θ_S, ϕ_S) on the surface sphere.

We arrange the values of these maps such that for each pixel of the rectified reflection map the position and reflection direction is stored in the same pixel position. We use the information of these maps in the next steps of the algorithm. More specifically, the $\mathcal{I}_1$ and $\mathcal{I}_2$ maps will be used to establish the correspondences, while $\mathcal{P}$ and $\mathcal{R}$ will be use as lookup table to determine the position on the sphere and the reflection direction of the corresponding points and calculate the light position.

The first two maps are easy to calculate, in fact it is sufficient to apply the standard reflection equation:

$$\vec{R} = \vec{I} - 2(\vec{I} \cdot \vec{N})\vec{N} \tag{1}$$

where $\vec{N}$ is the surface normal and $\vec{I}$ is the direction of the incident light. Indicating with P_1 and P_2 the center of the spheres in their respective coordinate frames and with p_1^s and p_2^s two vectors such that $P_1 + p_1^s$ and $P_2 + p_2^s$ lie on the surface of the spheres we can write:

$$\vec{I}_{1,2} = -\frac{M_{1,2}(P_{1,2} + p_{1,2}^s)}{||M_{1,2}(P_{1,2} + p_{1,2}^s)||} \tag{2}$$

$$\vec{N}_{1,2} = \frac{M p_{1,2}^s}{||M p_{1,2}^s||} \tag{3}$$

After computing the reflection vector in the relative camera coordinates frame we came back in world coordinates using the inverse transformation M_1^{-1} and M_2^{-1}. With these premises, the rectification reflection algorithm can be write:

1. Given a point on the reflective sphere in latitude-longitude coordinates (θ_S, ϕ_S):
2. $p_{1,2}^s$ is obtained by converting (θ_S, ϕ_S) in cartesian coordinates.
3. $\vec{R}_{1,2}$ is calculated using (3) and (1).
4. $\vec{R}_{1,2}$ is mapped on $\mathcal{I}_{1,2}$ by swapping its R_x and R_y coordinates and converting it in the geographical coordinate system (θ_R, ϕ_R). Then, the pixel's position on $\mathcal{I}$ is calculated as:

$$x = \text{round}\left(w\frac{\theta_R}{2\pi}\right) \tag{4}$$

$$y = \text{round}\left(h\frac{(\phi_R + \pi/2)}{\pi}\right)$$

where w is the width of the map and h is the height.

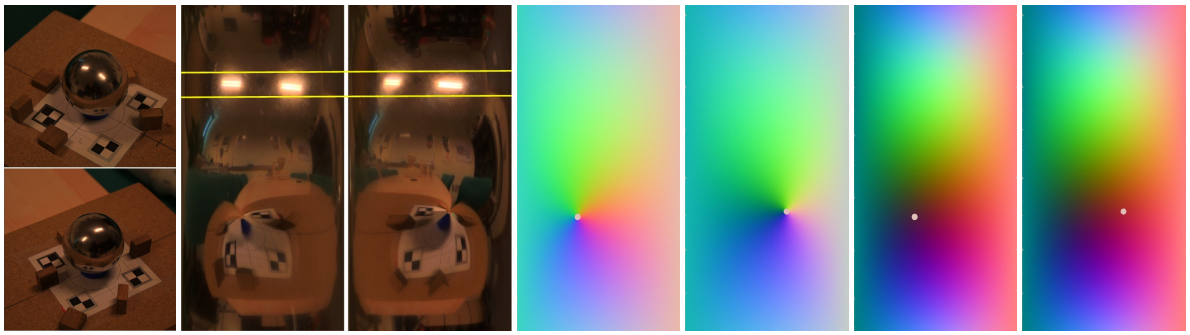

Figure 4: *(From Left-to-Right) The input images. The rectified reflection stereo pair. The RGB coded position and reflection lookup maps (for each pixel they code the position and reflection direction of the corresponding point on the sphere surface).*

5. To determine the color of the sphere point (θ_S, ϕ_S) we project on the camera image plane the point $P_{1,2} + p^s_{1,2}$. Hence, $\mathcal{I}_{1,2}(x,y) = \Pi_{1,2}(P_{1,2} + p^s_{1,2})$
6. (θ_S, ϕ_S) is stored in $\mathcal{P}_{1,2}(x,y)$.
7. $(\theta_R, \phi_R)_{1,2}$ is stored in $\mathcal{R}_{1,2}(x,y)$.

This steps are iterated over several values of (θ_S, ϕ_S) spanning the ranges $[0, 2\pi)$ for θ and the range $[-\pi, \pi)$ for ϕ. The sampling step is chosen in order to fill almost all the values of the pre-computed maps. This corresponds to over-sample the sphere surface.

Figure 4 shows the resulting maps. The resolution of such maps is 720×1440, encoding a good resolution in (θ, ϕ) ($\simeq$ 0.25 degrees). As it is possible to notice the light sources lie approximately at the same longitude in the rectified stereo pairs as expected.

3.4. Correspondences Estimation

At this point we have to estimate the correspondences between $\mathcal{I}_1$ and $\mathcal{I}_2$ in order to calculate the position of each irradiating element. To achieve this goal we opt for a region-based approach. This way of proceed is motivated by some preliminary experiments with multi-window stereo matching algorithm like [IB94] and [FRT97]. In some cases these correlation-based approaches are not able to compute reliable correspondences. This is caused by the fact that, depending on the position of the light sources, their shape could be severely distorted when projected on the rectified reflection maps. To accomplish for such distortions we detected blobs that represent the light sources of the scene in the rectified reflection maps and then we associate them.

The blob detection estimation is subdivided essentially in two parts. In the first part each HDR image is filtered in order to simplify the matching. Two filters are applied: a smoothing filter to reduce noise and make the image values more uniform, and a threshold filter in order to eliminate the elements of the scene that not irradiate sufficient amount of energy. Thanks to the HDR it is particularly simple to discriminate between the light sources and the other elements of the scene. The indirect illumination sources can be also identified with a careful choice of thresholds. In this first prototype we decide to set the threshold such that only the light sources are considered. More specifically the threshold is set as a percentage of the maximum radiance value (75%). After the thresholding we obtain non-uniform group of pixels that we make uniform by the application of multiple passes of a dilation filter followed by an erosion one. The kernel size of these filters depends on the dimensions of $\mathcal{I}_1$ and $\mathcal{I}_2$.

After identifying the blobs in the two images, the association is easily done by taking into account that the same blobs have to lie at the same longitude. Hence, each blobs in the first image is associated to the blob in the second image with the minimum latitude distance and within a pre-defined small range of longitude.

It is important to underline that this method to compute corresponding regions can be extended in a natural way by employing a more sophisticated segmentation algorithm for blob detection. For example by using a mean shift algorithm [CM02] to segment the elements of the scene it could be possible to associate in a robust way not only the regions that represent light sources but also those ones which irradiate a significative amount of indirect light. This segmentation could be used also to compute an approximated estimation of the geometry of the scene. In this work we do not present results in this direction leaving it as a promising future improvements of the Stereo Light Probe device.

3.5. Radiance Distribution Estimation

After having established the correspondences between brighter regions, the radiance distribution of the environment is obtained by sampling the corresponding blobs and computing the intersections between such samples.

Each blob is sampled by considering its center and its principal directions computed with standard PCA. The local coordinates are relative to the size of the blob and nor-

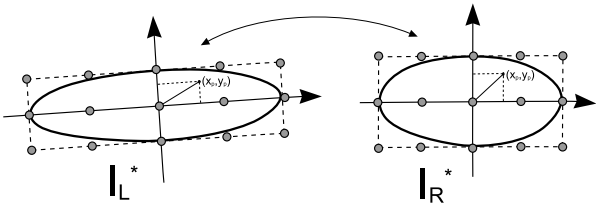

Figure 5: *Blob sampling.*

malized in the range $[-1,1]$. The samples are calculated by subdividing the oriented bounding-box of the blob in a predefined number of intervals (see figure 5).

Each sample is used to lookup $\mathcal{P}_{1,2}$ and $\mathcal{R}_{1,2}$ obtaining the position (o_1 and o_2) and the direction ($\vec{d}_1$ and $\vec{d}_2$) of the two reflected rays. We use such information to establish the position of the light sources by intersecting these rays. Since it is very difficult that these two lines intersect at an exact point we calculate the point nearest to the two reflected rays. We use the method by Goldman [Gol90] to obtain the two nearest points on the rays and then we consider the mean between these two points as the light position. This kind of linear sampling could introduce a slightly amount of distortions due to the fact that the sampling does not follow the shape of the blob but it is linear with respect to the computed oriented bounding-box. To alleviate this problem the bounding-box can be hierarchically subdivided into smaller oriented bounding-box in order to best approximate the shape of the blob.

Each light point is characterized also by the intensity and the color of the light emitted. The color of the light is determined by mapping the corresponding position o_1 on the reference sphere (I_1) and by multiplying this value according to the reflection coefficients of the steel ball. Such coefficients are obtained by taking a photograph of the steel balls reflecting a white paper and calculating the ratio between the RGB components of the paper reflected by the sphere with the components of the paper in the photo. The light intensity is corrected by two factors. The first factor takes into account the quadratic distance of the point from the reference sphere. The second factor is used to make the resulting light intensity independent of the number of samples ($N_{samples}$) used during the blob sampling. The latter scale factor is calculated according to the formula:

$$k = \frac{A}{N_{samples}} \tag{5}$$

where A is the area of the light source. More precisely the scale factor k is normalized such as the light source with the minimum area has $k = 1$. The area of the light source is computed considering the sum of the area of the reconstructed patches. Each patch is identified by four adjacent points. Figure 6 shows an example of the results of estimation of a room with two fluorescent lamps (120 samples has been used for each blob). The quality of the estimation is described in the experimental results section.

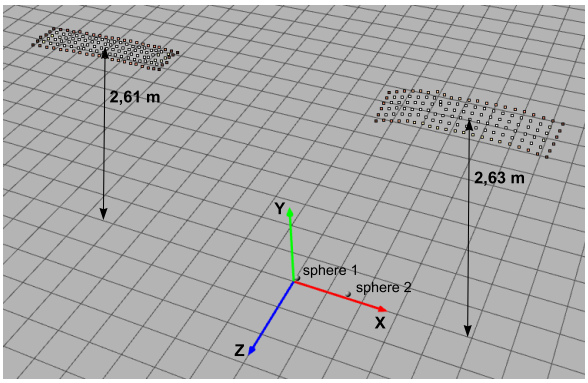

Figure 6: *Results of the estimation of a room with two fluorescent lamps.*

4. Experimental Results

We have tested the Stereo Light Probe device by acquiring the lighting environment of several real scenes. Here we present some of the acquisition results.

The aim of the three lighting environments we presents, was to measure the geometrical accuracy of the device as well as to evaluate the effectiveness of the data obtained. The recovered light position has been confronted with their measured real-world position. Then, the lighting environment has been used to re-illuminate a 3D model of an object that has been photographed under that particular real-world lighting environment. The first scenario is the one just shown during the method explanation, i.e. a room with two big fluorescent lamps, this experimented is referred as *Two Light Tubes* in the following. The second acquisition setup consists of three small light bulb with known positions. In the following we indicate this environment with *Lights*. The last lighting environment acquired (named *BigRoom*) consists of a large room (about $6 \times 5 \times 3$ m) with three fluorescent light tubes.

Even if at this prototypal development stage, the whole processing pipeline does not require skilled human intervention and could run in almost automatic way, making its use in a production environment viable. Moreover, even if the current implementation could be further optimized the overall acquisition time is quite low. The time to position the device, take the photos, build the HDR images and perform the calibration is around 1 hour. The computational time to process the input images and obtain the final estimation is also modest: the generation of the position, reflection, and rectified maps at a resolution of 720×1440 requires about 3 minutes on a standard PC.

Acquisition Name	Measure Description	Real (mm)	Estimated (mm)
Two Light Tubes	Size 1	1200×300	1281×308
Two Light Tubes	Size 2	1200×300	1301×317
Two Light Tubes	Height 1	2500	2632
Two Light Tubes	Height 2	2500	2614
Lights	Position 1	$(340, 315, 550)$	$(325, 320, 532)$
Lights	Position 2	$(-780, -600, -860)$	$(-857, -674, -956)$
Lights	Position 3	$(-2450, 1120, -1400)$	$(2814, 1378, -1644)$

Table 1: *Real vs estimate geometry of the light sources.*

Concerning the accuracy of the Stereo Light Probe device Table 1 shows a comparison between the real and the estimated geometric measures related to the *Two Light Tubes* and *Lights* environments. These results show that our method is able to correctly capture the position and the shape of the light sources. The geometric error of the estimation with respect to the distance from the light probes is relatively low. More specifically, the estimation error increases with the distance; for those lights placed in a radius of 1-1,5 m the error is about 4-5%, such error reaches about 10-11% for those lights at 3-3,5 m. The major causes of error are two. The first cause is related to small camera calibration errors, and the second one is the slight approximation in the correspondence estimation introduced by the blob sampling. Also the maps resolution influences the accuracy of the final estimation: the use of reflectance maps with higher resolution can improve the accuracy of the position determination. Finally, the use of chrome balls instead of steel balls would improve the accuracy of the system, providing more clean input images due to their more polished surface.

In order to show the reliability of the acquired data we have used the *Two Light Tubes* and the *BigRoom* lighting environment to relight 3D models and comparing the synthetic images generated with real photographs. In the first relight test (Figure 7) we compare the real photograph of a rapid-prototyped "Laurana" bust with a rendering of the digital model that has been used for the prototyping. The renderings have been generated with the basic version of NVidia® Gelato® [NVi], an hardware-accelerated high-quality rendering engine. The results of the relighting are very good from a qualitative point of view. In fact, despite the different material used (the BRDF of the material of the model is not estimated but a white plastic material has been used) the consistency of the illumination between the real and the synthetic images is very accurate. The relative absolute error between the real and the synthetic images is evaluated as:

$$\begin{aligned} \Delta(I_r, I_s) &= \frac{1}{hw} \sum_{y=1}^{h} \sum_{x=1}^{w} \frac{\Delta_R(x,y) + \Delta_G(x,y) + \Delta_B(x,y)}{3} \\ \Delta(x,y) &= \frac{I_r(x,y) - I_s(x,y)}{I_r(x,y)} \end{aligned} \tag{6}$$

where I_r is the real image, I_s is the synthetic image and the R, G and B symbols indicate the RGB components. This error is about 2.9% for the Laurana model relighted with the *Two Light Tubes* environment. Figure 7 shown a composition of the relight bust with the real images

In the final relight test we took a photograph of a small-scale model of the David of Michelangelo. This dense polymer model (height about 40 cm) has been cast in a mold generated from the data of the Digital Michelangelo's Project. The model has been placed in a large room and its position with respect to the Stereo Light Probe device manually measured. Then, the lighting environment of this room, the *BigRoom* environment, has been used to relight the 3D model. Figure 8 shows a detail of the final rendering ($\Delta(I_r, I_s) \simeq 3.2\%$). As it is possible to see even in the presence of small estimation errors (the fluorescent light tubes have a distance of about 5 m from the device), the illumination is still consistent with the real one. This lighting environment has been also used to produce the images in the teaser (Figure 1). These two images are generated considering three version of the David scaled to be height 1.8 m and placed at about 1.2 m of distance each other. The first image has been illuminated with the environment acquired, the second image as we use a single ball for the lighting environment estimation, i.e. only the direction of the light sources has been considered. The resulting images emphasizes the importance of the recovered position and shape of the light sources.

As it is possible to notice from the experimental results here presented, even if the characterization of the light sources is done in terms of omnidirectional point light sources, the quality of the final relighting is very good. In fact, most of the indoor light sources are well represented by omnidirectional sources, since they are constructed to produce a well distributed illumination in their environment. Nevertheless, it is important to underline that, in some cases, this omnidirectional representation can produce results with lower quality, for example, a window in a room. Since the light coming from an external source (mainly the sun) does not radiate from the window geometry in every direction, but along a favorite path. This could cause problems in a rendering using the output of the device. Another important considerations about the results that it is possible to obtain with the device concern the occlusion of the light sources. The

device is design to work with non-occluded light sources: the reconstruction can fails when one of the two reflective ball can "see" one light sources but the other cannot due to occlusion. This problem can be round on by choosing appropriately the position of the device or by making multiple acquisition.

5. Conclusions

We proposed a new method to acquire spatially-varying lighting environments. The acquisition is done with a simple stereo-omni device composed by two reflective balls and a digital camera. The experimental results demonstrate the effectiveness of the method. More specifically, even if the final lights position estimation is affected by small errors, such estimation is sufficiently accurate to be used in several applications such as augmented reality and reflectance properties estimation.

The major benefit of the proposed method is the good trade-off between the accuracy of the estimation and the simplicity and cheapness of the device required to obtain them. Additionally, the processing algorithm is not computationally expensive. The main critical point of the whole pipeline is the calibration phase that requires a good accuracy in order to obtain reliable results, but the use of non planar marker configurations allows to overcome these issues.

Future works regard the building of an hand-machine device and the improvement of the correspondences estimation by taking into account the nature of distortions on the rectified reflected stereo pairs in order to avoid the approximation given by the current blob sampling. Another interesting directions for future investigations is to extend the capability of the device by employing a more sophisticated segmentation algorithm in order to simultaneously acquire the light sources and the geometry of the scene.

6. Acknowledgment

This work is partially supported by EU IST NoE 507832 "EPOCH". The David 3D model is courtesy of the Digital Michelangelo Project, Stanford University. We would like to thank Marco Tarini for useful discussions about this work.

References

[CM02] COMANICIU D., MEER P.: Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell. 24*, 5 (2002), 603–619.

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1998), ACM Press, pp. 189–198.

[DS02] DOUBEK P., SVOBODA T.: Reliable 3d reconstruction from a few catadioptric images. In *OMNIVIS'02: Proceedings of the Third Workshop on Omnidirectional Vision* (Washington DC, USA, 2002), IEEE Computer Society, pp. 71–78.

[FRT97] FUSIELLO A., ROBERTO V., TRUCCO E.: Efficient stereo with multiple windowing. In *CVPR '97: Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)* (Washington, DC, USA, 1997), IEEE Computer Society, p. 858.

[GD03] GEYER C., DANIILIDIS K.: Conformal rectification of omnidirectional stereo pairs. *cvprw 07* (2003), 73.

[Gol90] GOLDMAN R.: Intersection of two lines in three-space. In *Graphics Gems I*, Glassner A. S., (Ed.). Academic Press, 1990, p. 304.

[HLH06] HWANG Y., LEE J., HONG H.: Omnidirectional camera calibration and 3d reconstruction by contour matching. In *Advances in Visual Computing*, vol. 4291 of *Lecture Notes in Computer Science*. Springer, 2006, pp. 881–890.

[HPB07] HAŠAN M., PELLACINI F., BALA K.: Matrix row-column sampling for the many-light problem. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM Press, p. 26.

[IB94] INTILLE S. S., BOBICK A. F.: Disparity-space images and large occlusion stereo. In *ECCV '94: Proceedings of the third European conference on Computer Vision (Vol. II)* (Secaucus, NJ, USA, 1994), Springer-Verlag New York, Inc., pp. 179–186.

[Len03] LENSCH H. P. A.: *Efficient, Image-Based Appearance Acquisition of Real-World Objects*. PhD thesis, Max-Planck-Institut für Informatik, Saarbrücken, Germany, 2003.

[Li06] LI S.: Real-time spherical stereo. In *ICPR06* (2006), vol. III, pp. 1046–1049.

[NVi] NVIDIA: Nvidia® Gelato®. More info on: `http://www.nvidia.com/page/gz_home.html`.

[SSI99] SATO I., SATO Y., IKEUCHI K.: Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE Transactions on Visualization and Computer Graphics 5*, 1 (/1999), 1–12.

[SSI03] SATO I., SATO Y., IKEUCHI K.: Illumination from shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence 25*, 3 (2003), 290–300.

[SY05] SATO T., YOKOYA N.: Omni-directional multi-baseline stereo without similarity measures. In *OMNIVIS2005: Proceedings of the 6th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras* (Oct 2005), pp. 193–200.

[Tsa87] TSAI R.: A versatile camera calibration technique

Figure 7: Two Light Tubes *environment: (Left) Real image. (Right) Composition of the real image and the synthetic image obtained with the acquired spatially-varying illumination.*

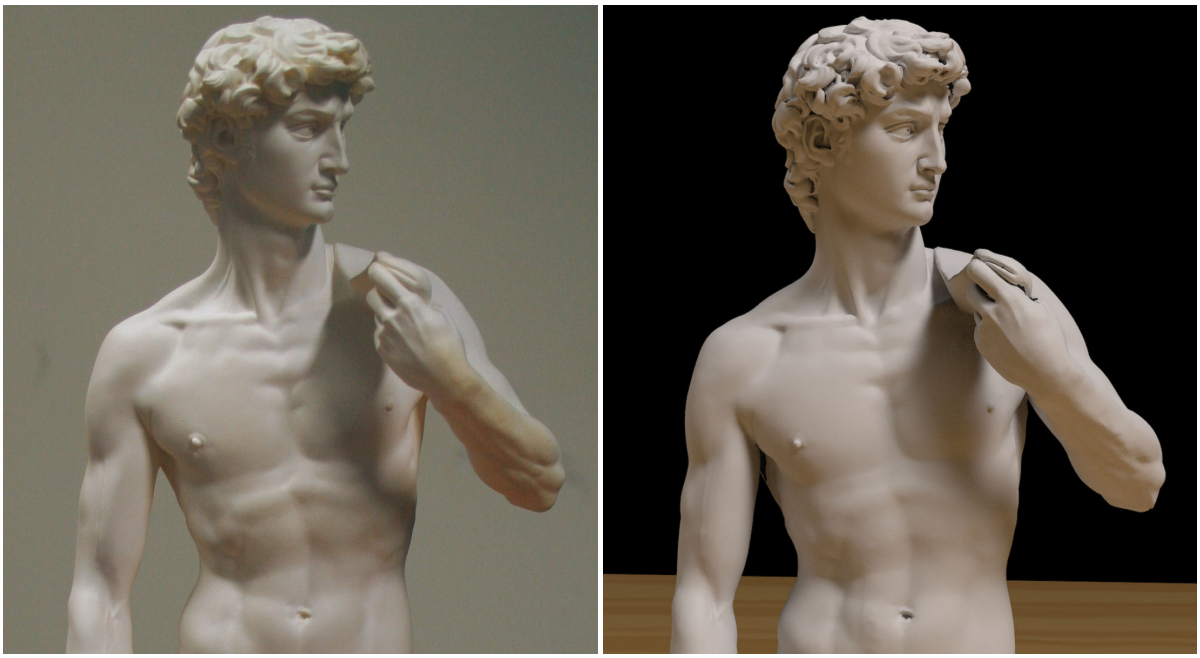

Figure 8: BigRoom *environment: (Left) A photograph of a small-scale model of the Michelangelo's David. (Right) A synthetic image generated using the Digital Michelangelo's Project data.*

for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation RA-3*, 4 (Aug. 1987).

[Uni] UNIV. OF SOUTHERN CALIFORNIA: HDRShop®. More info on: `http://www.hdrshop.com`.

[UWH*03] UNGER J., WENGER A., HAWKINS T., GARDNER A., DEBEVEC P.: Capturing and rendering with incident light fields. In *EGRW '03: Proceedings of the 14th Eurographics workshop on Rendering* (Aire-la-Ville, Switzerland, Switzerland, 2003), Eurographics Association, pp. 141–149.

[WABG06] WALTER B., ARBREE A., BALA K., GREENBERG D. P.: Multidimensional lightcuts. In *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM Press, pp. 1081–1088.

[WFA*05] WALTER B., FERNANDEZ S., ARBREE A., BALA K., DONIKIAN M., GREENBERG D. P.: Lightcuts: a scalable approach to illumination. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), ACM Press, pp. 1098–1107.