Large Display Light Field Estimation in a Wide Area

Pablo Roman Humanes[1,a)] Shohei Nobuhara[1,b)] Tony Tung[1,c)] Takeshi Takai[1,d)]
Takashi Matsuyama[1,e)]

Abstract: 3D reconstruction of a user in an immersive environment presents a set of challenges. In addition to the challenges of working in a low-light environment with dinamically changing illumination, the complexity of modeling a wide area in front of a display with uneven radiant intensity in different directions prevents previous methods from being applicable. We present a general method to model the light environment in front of a large display with the purpose of using it for 3D reconstruction with a single color camera by means of photometric stereo. We indirectly measure the light environment by capturing the appearance of a known object in front of the display with the same camera that will be used for reconstruction. The lighting environment is then modeled using a set of point light sources. Their positions and radiant intensities are fitted to the measurements, as is their directionality function, which accounts for existing screen gain. Our experiments verify our supposition that point light sources that account for screen gain provide a better modeling.

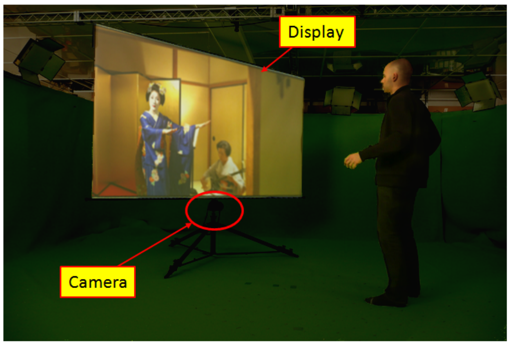

Fig. 1 Modeling light field in wide area in front of a large display with controlled content.

1. Introduction

Photometric stereo ([17],[10]) allows for very high detail 3D reconstruction of an object with a single regular camera within a controlled illumination setup ([3], [9]), but considerable results have also been obtained in uncontrolled setups ([1], [2], [8], [15]). Within the second case, the use of light coming from a display to reconstruct an object placed in front of it has seen limited attention, and previous approaches ([4], [5], [6], [7], [14]) are not designed to allow for 3D shape and reflectance (color) reconstruction of a user within a wide area in front of a large display (more than 2m diagonal). In this paper, we present a method of measuring the light field produced by such displays and a model to represent it that can be used as a first step to perform reconstruction.

[1] Kyoto University, Kyoto, 606–8501, Japan
a) pablo@vision.kuee.kyoto-u.ac.jp
b) nob@i.kyoto-u.ac.jp
c) tung@vision.kuee.kyoto-u.ac.jp
d) takesi-t@vision.kuee.kyoto-u.ac.jp
e) tm@i.kyoto-u.ac.jp

For the past few years, recent generations of game consoles have brought to the public new functionality like the creation of avatars based on the appearance of the user, usually using a camera attached to the device. Higher speed Internet connections have also brought up the possibility for the general public to transfer high definition video and detailed 3D geometry in real time. At the same time there has been an increase in the number of households that decide to install a video projector due to their cheap prices, but due to the limited luminosity of such projectors, the usual setup for these installations requires a relatively dark room. The direct implication of this is that the task of capturing images of the user and performing 3D reconstruction in such an environment becomes difficult due to the unknown and variable nature of the illumination, which mainly comes from the display.

The method presented here provides a first step to solve the problem of obtaining the 3D reconstruction and reflectance of a user in such an environment. It does so by creating a model of the light field in front of the display that can later be used to perform photometric stereo reconstruction of the user. This approach opens the possibility for said home setups and more serious virtual reality installations to support newer applications like telepresence, 3D videoconference, creation of a personalized 3D user avatar or inclusion of the real appearance of trainees and instructors in a training simulation (e.g., for emergency rescue, natural disaster contingency plan, etc.).

The photometric stereo technique consists on the estimation of the normals of the surface of an object by observing it under different lighting conditions, usually known. Its use in this scenario is justified by the requirement to obtain the real appearance (reflectance) of the user, which in this situation can only be found by separating the contribution that the reflectance and the illumination have on the appearance of the user. Its application in the presented scenario is challenging for several reasons: (1) The relatively low amounts of light in the environment result in high levels of noise in the obtained image that make accurate normal estimation difficult. (2) The use of a planar light source (e.g., a display) instead of regular lamps that can be treated as point light sources makes it harder to model the light environment. (3) Non-uniformity in the light source (projectors and LCD screens do not have a uniform radiant intensity over the whole display, and in different directions) makes it even more difficult, as does an uncontrolled environment that can present other sources of light, unknown geometry causing reflections, etc.

Previous research has tried to use the light coming from a calibrated screen to perform 3D reconstruction ([4], [5], [6], [7], [14]). The results are often satisfactory, and manage to reproduce the 3D shape and appearance of an object in front of the display. They present, though, several limitations that make unfeasible their application in our context. For [5], [6] the reconstruction volume is quite small, due to their display light model not accounting for display directionality (the reason why we can observe different radiant intensities when looking at a display from different angles), which limits their application to objects much smaller than the display size. In contrast, in the case of [7] the screen directionality function has actually been modeled. They find it by measuring the radiant intensity in different positions of the screen from a single camera position. That approach assumes the directionality function to be the same all over the display, which is mostly true in their case (using a TFT screen), and allows them to reconstruct relatively big objects. In our scenario, the use of a big projected screen means that the directionality function changes all over the display, so we cannot use the same approach. [5], [6], [7] assume that there are no reflections on the environment, and contain the whole setup — screen, object and camera — within a black containment. [14] differs considerably from the other approaches in that it allows for reconstruction without explicit calibration of the display or calculation of the lighting environment. That makes the system very easy to use, but also causes it to be limited by the need of the subject to stay motionless while several images are taken with changing lighting conditions. In our case, we measure the light over the whole volume we want to track (a typical case would be 3m×3m×3m) by sampling on the surface of a reference object, and we fit a series of point light sources (from here on PLS) to those samples to represent the whole light field. PLS have been used before when trying to reproduce the lighting environment ([12], [16]). In our case, we include a directionality function in the PLS, which allows them to better represent the directionality of the display. Our approach presents also other advantages applicable to certain scenarios. Thanks to its indirect appearance-based measuring approach, it is independent of display shape and number, and also of the geometry of the environment, without having to explicitly model either of them. Also, the calibration procedure is relatively simple, and does not require to set up additional camera positions other than the one used to capture the reconstructed object.

The proposed model assumes that the ambient light in the environment does not change between calibration and reconstruction, and that the radiant intensity of the display remains constant for the same input signal value (this just requires projectors to be warmed up before calibrating them or using them for reconstruction). Additionally, basic approaches to separate shape and albedo from multiple images with different lighting conditions require the object to be static while the display contents change. For reconstruction of a moving user in front of the display, more advanced methods of non-rigid shape matching with iterative refining of shape and reflectance are needed ([11]).

The created display illumination model can be considered satisfactory if using it we can achieve accurate estimation of the value of the radiant intensity for a specific direction and position in space within the error limits that can provide an acceptable normal estimation, and it can provide that accuracy over a big enough volume to allow for tracking of a moving user.

2. Framework Overview

In this section, we introduce our approach to model the light emitted by a display. Real life displays present certain properties that make it difficult to predict the emitted light field without direct observation. Display settings, color calibration and non-uniform radiant intensity are examples of variables that affect the light emitted from a display. We are going to present now the different properties that have to be dealt with.

- **Non-linear response function** (Figure 2): It refers to the mapping between input values to the display and actual emitted light. This relation is usually not linear, and it further depends on the display gamma settings. In our experiments, the three different color channels have different response functions, so they need to be calibrated separately. According to our measurements, after adjustment for total radiant intensity, all the pixels over the display share the same response function. That means that we only need to measure it once for each color channel.

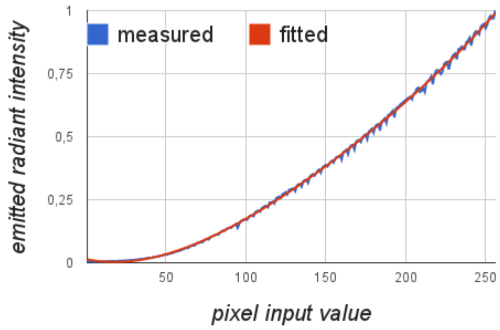

Fig. 2 Projector non-linear response function.

In our case, we separately measure the response function by directly capturing the radiance of the screen with a camera for all 8-bit intensity levels for every color channel.

- **Non-uniform radiant intensity** (Figure 3): The majority of displays that we find in real life do not present an even radiant intensity over all their surface, even if they look like they do to the human eye. This is usually related to using a limited amount of sources of light and trying to distribute their light evenly over the whole area of the display. The actual value of minimum and maximum emitted light is also dependent on the screen brightness and contrast settings. In our case, this property is implicitly dealt with in the model.

Fig. 3 Example of non-uniform radiant intensity over the display's surface.

- **Display directionality** (Figure 4): Different display technologies exhibit different directionality characteristics. In the case of a projector screen, factors that affect it are: screen technology (front projected or back projected), screen material (which defines the reflection or transmission properties like screen gain), angle of the projector to the screen, etc. In our model, we deal with this property by using a specific representation of the light environment that is capable of dealing with directional light sources.

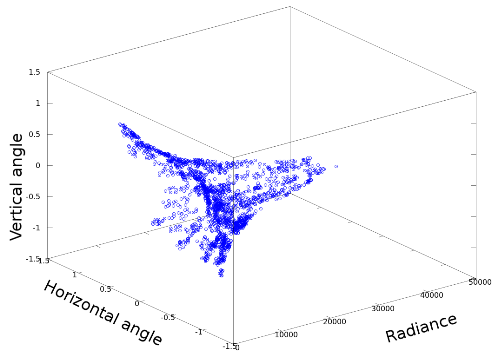

Fig. 4 Measured projector screen display directionality.

Even though in our experiments we used exclusively a projector screen, an LCD screen exhibits many similar properties. One additional property we should be careful about when calibrating an LCD screen is that cheap ones usually show color shifting and color inversion when being looked at from extreme angles.

The representation chosen for the light model uses point light sources (PLS) to show where the light is coming from and with how much radiant intensity, based on the model described in [12]. Using PLS makes it much easier to perform calculations than with more complex shapes, while still allowing for a lot of flexibility. Those PLS are used to represent both the light coming from different parts of the display, and artifacts like reflections on the environment. The position of the PLS accounts for the geometry and position of the display in relation to the environment. Additionally, we introduce a directionality function in each PLS that is designed to better represent the display directionality. These PLS are not, however, assumed to be physically correct, but rather to encode the light environment in the most accurate way. Their positions, therefore, will not necessarily exactly match the position of the display or reflecting objects. Each PLS has a maximum radiant intensity level, which accounts for the non-uniform radiant intensity property of the display, or for the albedo of a reflecting surface in case the PLS is modeling a reflection. This maximum radiant intensity value allows to map the current input values of a specific region of the display to actual light amount using the separately obtained non-linear response function. The position, radiant intensity and directionality of the PLS are determined by measuring the light emitted from different regions of the display. Both the method for measuring and the method for fitting the PLS to those measurements are explained in Section 3.

The final goal of the light model is to allow us to perform 3D reconstruction of an object within the modeled light environment. Therefore, one of the conditions that the model has to satisfy is to be able to output the normal of a specific pixel in the image when provided with that pixel's value and the contents being shown on the display. Performing reconstruction itself is out of the scope of this paper, but nevertheless we want to show the applicability of the presented model by suggesting a strategy to solve it. The goal of modeling separately the light coming from different parts of the display is to be able to consider them as if they were separate light sources. In the simplest case, we can imagine three spots on the screen with red, green and blue colors behaving as the three colored lights used in many photometric stereo techniques ([3], [9]). When performing reconstruction in a real-world scenario, with a user in front of a large display actively using the system, we do not want to constantly disturb the user by showing color patterns specifically designed to perform reconstruction of the user, but instead use the light already coming from the display according to the contents the user is visualizing, be it a game, a movie or a simulation. In practice, the contents of the display at a specific point in time may sometimes not be as informative as the three independent lights, but we expect the contents changing over time to provide us with enough information for reconstruction.

When trying to solve the value of the normal for a specific pixel, we find that there is often a big number of possible solutions. For a specific radiance value of a pixel, there can be several combinations of different values of depth and normal direction that match with it. A way of solving this ambiguity is by introducing all the possible solutions into a probabilistic framework. Those solutions are then compared with the neighboring pixels and assumptions on the continuity of depth and smoothness of normals can help us find the right solution.

3. Point Light Source Model Definition

We measure the direction and radiant intensity of the light emitted by the display by observing the appearance of a reference object in many positions within a wide area in front of the display. The reference object has a known shape and reflectance, and its position in the image can be determined by the computer without user intervention, as to make the calibration process faster and simpler. Samples are taken over the object's surface, registering radiance value, position in space and orientation. These samples are the values that will be fitted during the calculation of the light model. Multiple images are captured of the reference object in different positions and orientations, and with varying display contents, so that the light field is sampled for the whole tracking volume for light coming from each position on the display. The positions and orientations of the reference object should be dense enough to provide a good sampling of the variation of the light field. Accordingly, areas closer to the display will require a denser sampling than areas further away.

The light intensity function, I_x, allows us to compute the radiance of a point (x) on the object's surface according to its position, orientation, and the positions (P_i) and radiant intensities (L_i) of N PLS. In our case, since we assume Lambertian reflection for the reference object, the function is quite simple. The light coming from the PLS is reduced by the square of the distance between the PLS and the object, and further reduced by the cosine of the angle (θ) between the normal of the surface (N_x) and the direction where the light is coming from ($x\vec{P}_i$). The contribution of all the PLS is then added to obtain the total radiance.

$$I_x = \sum_{i=1}^{N} \frac{L_i * cos(\theta)}{|x\vec{P}_i|^2} \tag{1}$$

This equation is used when trying to find the optimal position and radiant intensities for the PLS.

We need to also consider the relationship between the display contents and the PLS. One possibility we could think of would be to generate a different set of PLS to model the behavior of each pixel in the display. Such correspondence would be very accurate but it presents several practical inconveniences. A single pixel on the display produces a very small effect in the overall light field of the environment, therefore, when trying to capture the effect on the reference object the difference would be very small, if appreciable, and we would have quantization and noise problems. Another problem is that individually turning on every pixel in the display and registering its effect in multiple positions and angles would be prohibitively expensive, as it would require a lot of processing power for very little gain. The solution we adopted for this problem is to split the display into a limited amount of regions, or "patches", and consider all the pixels within one of them as a single entity. Each patch is thus assigned a set of PLS. When calibrating the display and estimating those PLS, all the pixels in one patch will be set to be completely white, and its effect on the reference object will be measured. The obtained radiant intensities of the PLS will therefore be a maximum value that will be applicable only when the patch is completely white. When using the model for shape reconstruction, those values need to be scaled according to the contents of the patch. To do that, the average of all the pixels in the patch needs to be calculated, but before that, the value of each of the pixels needs to be gauged according to the non-linear response function discussed in Section 2. During the calibration step, one technique we can use to reduce the capture time and/or increase the signal-noise ratio of the obtained images is multiplexed illumination ([13]).

The approach taken to calculate the positions and the radiant intensities of the PLS that model each patch is based on [12]. This approach works by generating an initial small set of candidate PLS positions. It then creates a set of equations introducing the value of the captured samples into the light intensity function (Equation 1) and solving L_i for all of them using non-negative least squares (NNLS). If we separate L_i from Equation 1, we find that the remaining elements are fixed and we can represent them by a single value K_{ij}. If we have M samples and N candidate PLS, then we can represent the previous set of equations using matrix multiplication (Equation 2). We can apply NNLS to this multiplication to obtain the values of L_i that minimize the square error. PLS whose radiant intensity is determined to be 0 or small enough are discarded, and new candidate PLS are generated around the remaining ones. This refining of the PLS positions goes on iteratively until the residual error cannot be reduced anymore. For a more detailed explanation of the algorithm refer to [12].

$$\begin{bmatrix} I(x_1) \\ I(x_2) \\ \vdots \\ I(x_M) \end{bmatrix} = \begin{bmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{M1} & K_{M2} & \cdots & K_{MN} \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_N \end{bmatrix} \tag{2}$$

The main modification we introduced to the presented procedure, besides the usage of a different reference object (see Section 4), is the a directionality function D in the PLS candidates. In the current approach, all the generated PLS have the exact same directionality function, but the direction of maximum radiant intensity is particular for each of them. For the directionality function, we chose a polynomial of the cosine of the angle (θ) to the direction of maximum radiant intensity, as expressed in equation 3. We chose this function because it presents a good ratio between the biggest and smallest radiant intensities and a relatively narrow angle of high radiant intensity (see Figure 5). Note that this function does not need to exactly correspond to the real screen gain function, since the combination of several PLS at different distances should be able to accurately represent different functions. For each step of the optimization, we generate a set of PLS in each position facing in different directions, but all of them facing roughly in the direction of the region where the samples were taken.

$$c = \begin{cases} cos(\theta) \leq 0.5 : cos(\theta) \\ cos(\theta) > 0.5 : cos(0.5) \end{cases}$$
$$D(c) = 10.5c^3 - 16.67c^2 + 11.08c - 1.93 \tag{3}$$

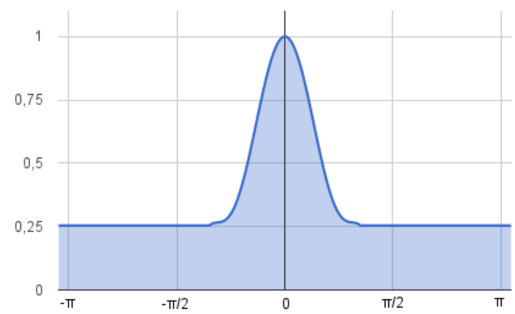

Fig. 5 Light source directionality function D(θ).

4. Experiments

4.1 Experimental setup

All the experiments have been conducted in a laboratory environment (see Figure 1). In our setup there is presence of ambient light, other than that of the display, both from artificial lighting inside the room and from the outside. The experiments have been conducted in a short amount of time to make sure that there are no big differences in background illumination. Nevertheless, each of the captured images used for calibration of the light environment is accompanied by a shot with the screen completely black to remove the effect of the background illumination. A back-projected screen (225cm/101inch diagonal) has been used as the display, together with a DLP projector, and the camera has been placed centered at the bottom of the screen, in such an angle as to get a good view of objects in front of the screen, but not of the screen itself. The camera used is a Nikon D700 digital reflex. The pictures were taken with an aperture of F/3.5, and a shutter time of 1/10.0s. The area in front of the camera consists of a smoothly colored, diffuse floor and wall, without any other visible objects around. All the samples have been taken within a roughly 3m×3m×2m volume.

The reference object used consists of a white board with a chess pattern in the center. The chess pattern is used to determine the position and orientation of the board, while the white area around the board is used for sampling. The board measures 70cm wide and 70cm tall. The pattern is 48cm wide and 36cm tall. In practice, that leaves a margin of around 9cm on the sides and around 15cm on the top and bottom of the board, since we want to avoid sampling too close to the frame of the board, which may produce shadows that interfere with the sampling.

The settings of the optimization algorithm are the following. After each round of optimization, new candidates are generated in a cube shape in 26 positions around each candidate that has not been discarded. For each round the size of the generated cube is reduced by half until a minimum size is reached, point at which the optimization ends. For each of the 26 positions, 13 light sources are generated with different directions of maximum radiant intensity in a regular shape, pointing at a maximum of $\pi/4$ degrees from the direction of the average position of the reference objects.

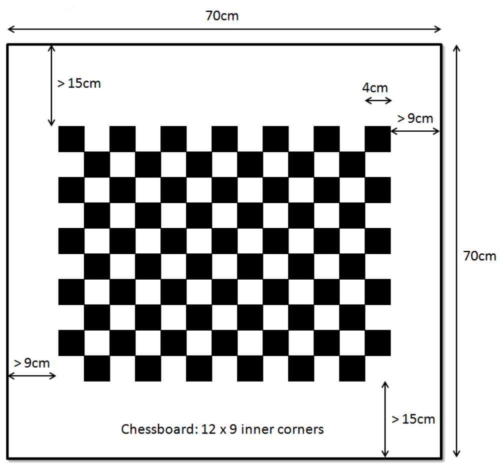

Fig. 6 Dimensions of the reference object used for system calibration.

4.2 The Experiments

We present the results of two separate experiments in this paper. The first experiment tries to establish that our model is an accurate enough representation of the light environment to be used for 3D shape reconstruction, and that such a level of accuracy can be achieved using a non-systematic non-exhaustive approach when placing the reference object for calibration. The second experiment's goal is to demonstrate that better results are achieved when adding directionality to the PLS candidates than when using isotropic ones. For both experiments we use the same set of 91 captured images, each of which shows the reference object in a different position and orientation. The same patch in the screen is filled completely in white during all the captures, while the rest of the screen is completely black. The reference object was positioned by hand by an operator with the only consideration of having a certain range of orientations around every general region of the volume visible to the camera.

For the first experiment, we use a leave-one-out cross-validation strategy to verify the fitness of the presented model. In this case, the test observation we leave out for cross-validation is the whole set of samples belonging to a specific reference object image, and not a single sample. As a measure of goodness, we calculate the error between the samples taken from the test image and the value predicted by our model at that position and orientation. We express this error as the average and the median of the absolute error for all the samples, divided by the value of the brightest sample, to give an idea of its scale.

For the second experiment, two different models are fitted to the exact same set of samples. The first model uses simple isotropic PLS, as used in previous research [12]. The second model includes the proposed directionality function as described in Section 3 to create anisotropic PLS. The median absolute error for both cases is then compared to establish if the addition of directionality to the PLS provides any advantage. In order to make both models comparable, due to the fact that the optimization using anisotropic PLS generates several candidates for each candidate of the isotropic approach, we experimentally adjusted the parameters of each optimization to have them provide their best result.

5. Results

The capture has been conducted as detailed in Section 4. The positions of the reference object and the resulting positions and radiant intensities of the PLS after optimization can be observed in Figure 7. The figure shows the positions were the reference object was captured (grey planes), the position of the camera (small pyramid), and the position of the optimized PLS (spheres), which have been scaled in proportion to their radiant intensity, and their direction expressed by a red line. Most of the PLS are effectively found slightly above the position of the camera, that is actually where the lit-on part of the screen was situated. Additional PLS have been found further up in the volume, which we assume that model the reflections that occur on the ceiling that is made of specular plastic panels. The optimization for each of the cross-reference rounds of the experiment took on average 10 minutes to run on a computer with an Intel(R) Core(TM)2 Quad Q9400 CPU.

The results for Experiment 1 are as follows. The median abso-

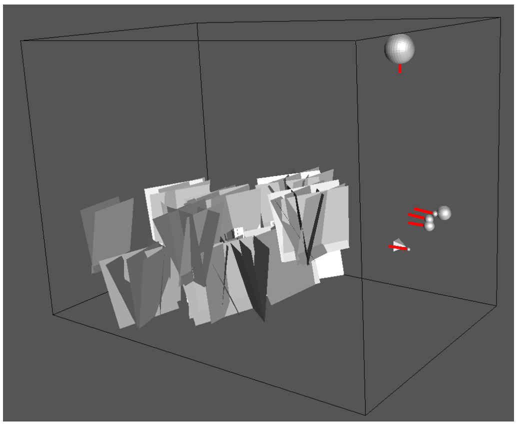

Fig. 7 Optimized positions, directions and radiant intensities of the point light sources.

lute error for all the captured reference object positions is 14.81% of the value of the brightest sample, and the average absolute error is 14.77%. Additionally we tried splitting the captured positions into two separate groups: those close to the center of the modeled volume, and those close to the edge. We did this to compare how well our approach interpolates between known data points and how well it does when it needs to extrapolate values outside of the sampled region. The result is that the median absolute error of the samples close to the center of the volume is 10.67%, compared to 26.58% for positions near the edge. The estimation is thus much better near the center of the volume, and when using the model, enough samples should be taken beyond the volume we intend to use for reconstruction.

The results for Experiment 2 are as follows. The median absolute error for isotropic PLS (without directionality) is 21.82%, and the average absolute error is 21.88%. This results compare to 14.81% and 14.77% for anisotropic PLS, as shown in the previous experiment. This shows that the addition of anisotropic PLS is advantageous and allows us to build more applicable models of the light environment than isotropic ones. In Figure 8 the difference can be appreciated. The isotropic PLS manage to reproduce the general vertical illumination of the board but fail to reproduce the more high-frequency horizontal difference. Using the directionality function, the more high-frequency darker areas at the corners can also be reproduced faithfully.

6. Conclusions and Future Work

We have presented in this paper a novel approach for the modeling of light emitted from a large display over a wide area with applications in 3D shape reconstruction. Our model uses a reference object to measure the light field within a specific volume and creates a set of point light sources (PLS) that approximates that light field. In contrast with conventional work, we introduced a new search strategy and added anisotropic PLS. We have shown that we can generate a model that we believe is capable of approximating the light environment accurately enough to be used for photometric stereo 3D reconstruction, and that the use of anisotropic PLS allows us to better model the light coming from a display. The presented model is also very general and applicable to any number of displays independent of their shape, without any explicit calibration of their geometry. We expect these interesting qualities to open the door for new algorithms that allow to obtain the 3D shape and reflectance of an object in contexts where it was not possible until now.

The next logical step for this research is to go beyond theory and use the presented model to perform actual shape reconstruction. Initially, a static object with changing screen contents can be used. Obtaining its shape and reflectance can be a step before progressing into the reconstruction of a moving user. To do that, a non-rigid shape model of the user needs to be fitted to each of the captured frames, and then an iterative process of refinement of both shape and reflectance can be applied.

Additional steps can also be taken to try to improve the current model. Different directionality functions for the PLS may give better results, and its worth checking what directionality functions work best for different kinds of displays. It may also be worth trying other search strategies for the PLS, which may give better results in certain cases. Also, a more complex reference object that includes several different normals may reduce the amount of images that need to be captured to model the light environment.

One possible future research path we consider especially interesting is the optimization of the screen contents to help improve the accuracy of the reconstruction. Substantially changing the contents of the screen for the sole purpose of reconstruction can be a hindrance for the user, but learning how to change the contents in subtle ways (brightness, color tonality, etc.) that do not bother the user and are in accordance with the estimated light model can be an interesting challenge.

Acknowledgments This study is partially supported by JSPS Ayame project.

References

[1] Ackermann, J., Langguth, F., Fuhrmann, S. and Goesele, M.: Photometric stereo for outdoor webcams, *CVPR - Computer Vision and Patern Recognition - 2012*, pp. 262–269 (2012).

[2] Basri, R., Jacobs, D. and Kemelmacher, I.: Photometric Stereo with General, Unknown Lighting, *Int. J. Comput. Vision*, Vol. 72, No. 3, pp. 239–257 (online), DOI: 10.1007/s11263-006-8815-7 (2007).

[3] Brostow, G. and Vogiatzis, G.: Video normals from colored lights, *Pattern Analysis*, Vol. 33, No. 10, pp. 2104–2114 (2011).

[4] Clark, J. J.: Photometric stereo with nearby planar distributed illuminants, *Computer and Robot Vision, 2006. The 3rd Canadian Conference on*, pp. 16–16 (2006).

[5] Clark, J. J.: Photometric stereo using LCD displays, *Image and Vision Computing*, Vol. 28, No. 4, pp. 704–714 (2010).

[6] Francken, Y., Hermans, C., Cuypers, T. and Bekaert, P.: Fast Normal Map Acquisition Using an LCD Screen Emitting Gradient Patterns, *2008 Canadian Conference on Computer and Robot Vision*, pp. 189–195 (2008).

[7] Funk, N. and Yang, Y.-H.: Using a Raster Display for Photometric Stereo, *Fourth Canadian Conference on Computer and Robot Vision (CRV '07)*, pp. 201–207 (2007).

[8] Hayakawa, H.: Photometric stereo under a light source with arbitrary motion, *J. Opt. Soc. Am. A*, Vol. 11, No. 11, pp. 3079–3089 (online), DOI: 10.1364/JOSAA.11.003079 (1994).

[9] Hernández, C., Vogiatzis, G., Brostow, G. J., Stenger, B. and Cipolla, R.: Non-rigid photometric stereo with colored lights, *Vision, 2007. ICCV* (2007).

[10] Horn, B. K. P. and Brooks, M. J.(eds.): *Shape from shading*, MIT Press, Cambridge, MA, USA (1989).

[11] Letouzey, A. and Boyer, E.: Progressive Shape Models, *CVPR - Computer Vision and Patern Recognition - 2012*, (online), available from ⟨http://hal.inria.fr/hal-00677506⟩ (2012).

[12] Matsuyama, T., Nobuhara, S., Takai, T. and Tung, T.: Chapter 6: Estimation of 3d dynamic lighting environment with reference objects, *3D Video and its Applications*, Springer, pp. 203–238 (2012).

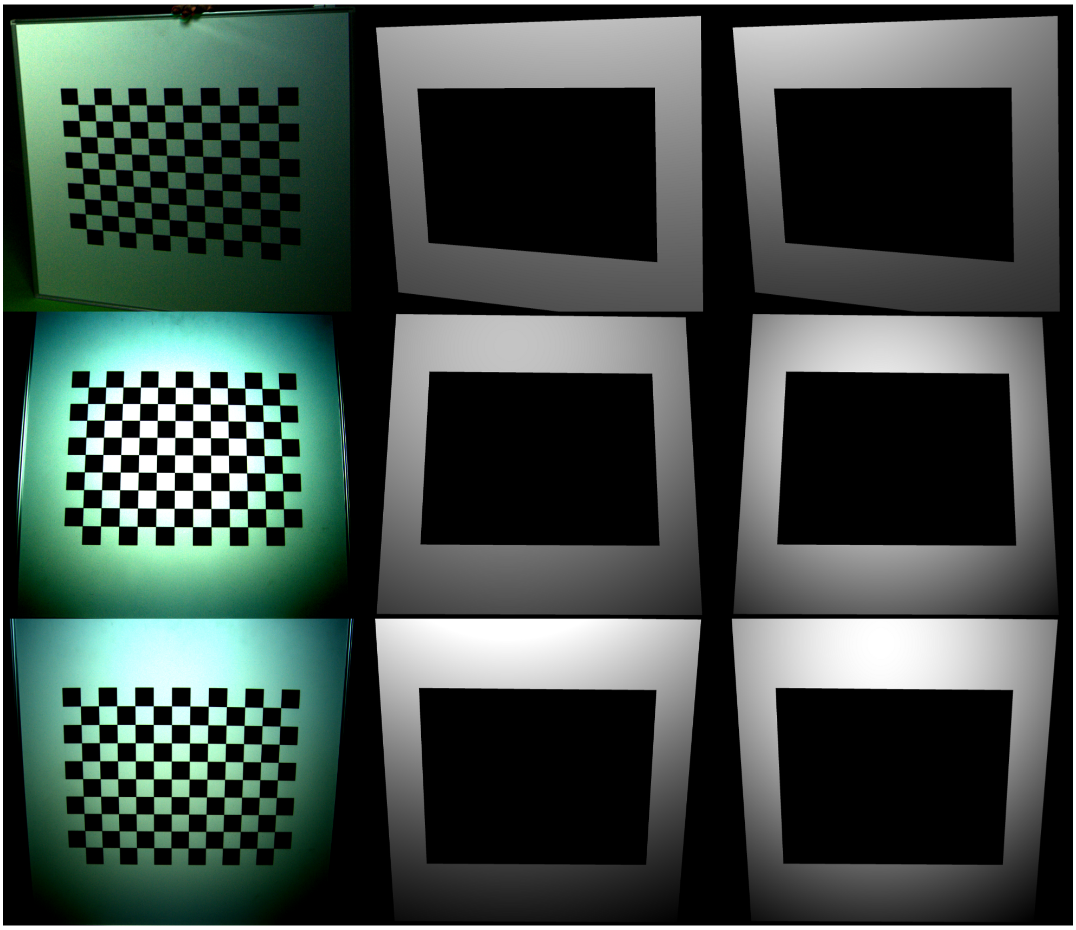

Fig. 8 Comparison between original image (left), optimization without directionality (center) and with directionality (right).

[13] Schechner, Y., Nayar, S. and Belhumeur, P.: A theory of multiplexed illumination, *Proceedings Ninth IEEE International Conference on Computer Vision*, pp. 808–815 vol.2 (2003).

[14] Schindler, G.: Photometric stereo via computer screen lighting for real-time surface reconstruction, *Proceedings of 3DPVT-the 4th Internation Symposium on 3D Data Processing, Visualisation and Transmission*, Vol. 3 (2008).

[15] Shi, B., Matsushita, Y., Wei, Y., Xu, C. and Tan, P.: Self-calibrating photometric stereo, *IN: PROC. IEEE CONF. COMPUTER VISION AND PATTERN RECOGNITION* (2010).

[16] Weber, M. and Cipolla, R.: A Practical Method for Estimation of Point Light-Sources, *PROCEEDINGS BMVC 2001*, Vol. 2, pp. 471–480 (2001).

[17] Woodham, R.: Photometric method for determining surface orientation from multiple images, *Optical engineering* (1980).