

A SIMULATOR FOR THE CAFADIS REAL TIME 3DTV CAMERA

*F. Pérez Nava⁺, J. Philip Lüke, J. G. Marichal-Hernández, F. Rosa and
J. M. Rodríguez-Ramos*

⁺Departamento de Estadística, Investigación Operativa y Computación
Departamento de Física Fundamental y Experimental, Electrónica y Sistemas
Universidad de La Laguna. 38271. Canary Islands. SPAIN.

E-mail: {fdoperez@ull.es⁺, jpluke@ull.es, jmarisher@ull.es, frosa@ull.es, jmramos@ull.es}

ABSTRACT

This paper presents a computer simulator for the CAFADIS camera. CAFADIS is a patented camera (PCT/ES2007/000046) that measures wave-front phases and 3D distances under different scenarios (from microns to kilometres), using highly specialised electronic technology, namely Graphics Processing Units (GPUs) and Field Programmable Gate Arrays (FPGAs).

Since 3DTV is one of the most important applications for future information systems it is essential to know the camera design that is capable to satisfy the demands of real time 3DTV. To study the optimal design of the CAFADIS camera for 3DTV and other applications, we have developed a computer simulator that places a virtual CAFADIS camera in a virtual world. Such simulator provides the accuracy of the camera 3D reconstruction for a specific CAFADIS camera configuration and is valuable tool for the improvement process of the device.

Index Terms— 3D TV, real time processing, rendering, GPU, FPGA, wavefront map, scene depth, stereo reconstruction.

1. INTRODUCTION

Three-dimensional television (3DTV) [1] is one of the most important future improvements of the traditional TV. The 3DTV offers a three-dimensional (3D) depth impression of the observed scene. It has been developed by the convergence of new technologies from optics, computer graphics, computer vision, and related fields. A 3DTV system is composed of several modules: image capture, 3-D scene reconstruction and representation, coding, transmission, rendering and display. The CAFADIS camera integrates the scene image capture and its 3D reconstruction. Such kind of camera will be of great importance in 3DTV systems where the techniques for 3D capture and 3D display

are decoupled from each other and where the capture device should provide the computerized representation of the 3D scene. To achieve this goal the image captured by the CAFADIS camera contains multiview information that is used to reconstruct the 3D scene. To obtain multiple perspective images, one would traditionally resort to varying configurations of sensor groups, generally two or more cameras, which determine the method for 3D distance estimation. Alternatively, the CAFADIS camera obtains multiple perspectives using micro-lenses. This multiview information is used to solve the 3D scene reconstruction problem. This is an example of an *inverse* problem: for a given a set of images, what is the 3D scene that generated them? The *forward* (or direct) problem is: given a 3D scene and the position and orientation of a camera, what is the expected image? This is the area of computer graphics known as rendering. While both problems have their own difficulties, it is widely believed that the inverse problem is considerably harder than the direct problem. In fact, the 3D multiview reconstruction problem is nowadays an active area of research [2]. In our simulator we provide a solution for both the direct (rendering) and inverse (CAFADIS output) problem. This allows us to estimate the accuracy and robustness of the distances delivered by the CAFADIS camera and provides a development environment to improve its performance. To obtain an interactive simulation we use dedicated graphics rendering devices: GPUs (Graphics Processing Units). The rapid evolution of these platforms is described in [3].

This paper is divided in five sections. In section 2 we introduce the CAFADIS camera, its optical description and its relation with other plenoptic cameras. Section 3 shows the different modules involved in the CAFADIS computer simulator, from the generation of the 3D virtual scene to the computation of the distance map and its accuracy. Section 4 contains some experimental results for a virtual scene. And finally Section 5 includes conclusions and future work.

2. THE CAFADIS CAMERA

The camera we propose for conducting tomographical 3D spatial object measurements entails one single Shack-Hartmann sensor, assembled at the image domain of a converging lens. This scheme can be used for wavefront sensing from the subdivision of the focal plane with a lenslet array [4]. The same scheme can also be used to get focused images at different scene depths [5].

The CAFADIS camera obtains enough data for the reconstruction of the 3D environment with one sole measure (i.e. one single exposure).

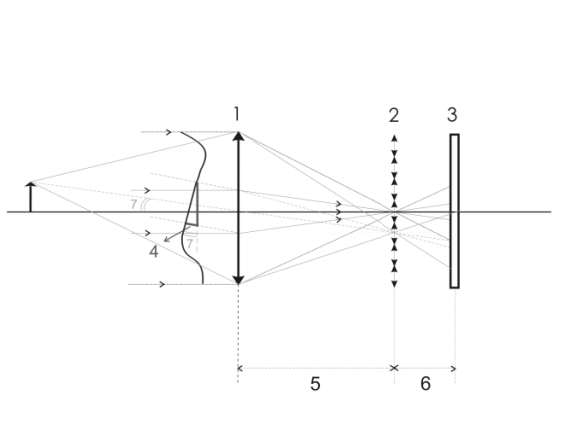


Figure 1. The CAFADIS scheme. (1) objective lens, (2) microlens array, (3) CCD, (4) wavefront, (5) microlens array position and (6) microlens focal.

The image resulting from application of the CAFADIS sensor can be seen as formed by four dimensions: two CCD co-ordinates associated to each micro-lens and a further two co-ordinates stemming from the micro-lens array (Figure 1). This image is then processed by the camera to obtain the 3D distance map. This 3D map, combined with the 2D scene image, can be used as input to a 3D display [6].

In order to recover the 3D distance map as described in Section 3.4, no advantage is obtained at the moment from the fact that the wavefront map can also be measured from the CAFADIS camera. This could also be done, as proposed in [7], using the same optical scheme to capture holograms under white light illumination.

3. THE SCAFADIS SIMULATOR

The SCAFADIS simulator as a whole consists of a system for positioning a CAFADIS virtual camera in a virtual scene. The simulator allows the camera to measure the distance and color for objects located within its field of view. SCAFADIS is composed of several modules shown in Figure 2.

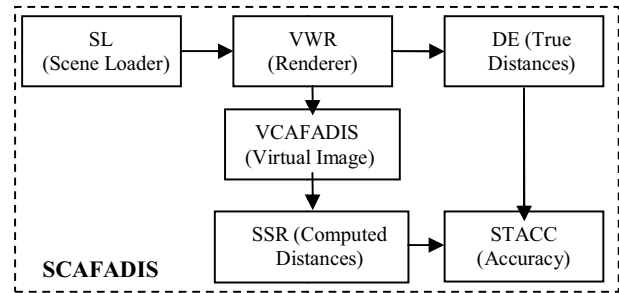


Figure 2. The SCAFADIS modules.

3.1. Scene Loader (SL)

Of fundamental importance to the SCAFADIS simulator is the versatility for representing the 3D geometry of the scene it simulates. To this end we have implemented a VRML scene loader subject to the VRML 97 norm. This allows us manipulate the geometry exported by different 3D design platforms. This module, combined with the one for hardware accelerated 3D representation described next, achieves high-resolution interactive frame rates in excess of 30 frames per second in complex scenes with one million vertices.

3.2. Virtual world rendering (VWR) and virtual image generation (VCAFADIS).

For rendering the virtual world and generating the image that provides the virtual CAFADIS camera (see Figure 3, left), we have chosen the OpenGL standard for 3D hardware acceleration. Adequately formatted information from the scene, as well as the camera's position, must be constantly relayed to the API OpenGL of the graphics card. The hardware-software platform that we have employed to generate the images consists of a nVidia Geforce 7800 GTX graphics card and an AMD 3500+ with 1 GB of memory running the GNU/Linux Debian 3.0 OS. This allows the simulator to be ready to give an interactive response even for highly complex models.

3.3. Distance Extractor (DE).

What distinguishes the simulator we implemented from other virtual world viewers is the emphasis it places on acquiring the 3D distances from the objects to the point of observation. To obtain these measurements we have also resorted to state of the art 3D graphics acceleration hardware. We use the vertex programming capability of the GPUs (Graphics Processing Units), only available on the most recent cards. This allows for higher interactivity rates with complex sceneries.

3.4. Computed distances (SSR).

The distance map reconstruction in VSIDS is obtained by solving the direct problem when in fact, for the CAFADIS camera the distance map must be obtained by solving the inverse problem (obtaining the 3D geometric model based on the image of the scene). This inverse problem is the depth estimation problem, of great importance in computer vision. To solve this problem we first compute the “focal stack” from the captured image using a fast discretization of a four-dimensional generalization of the Radon transform, then we apply a focus quality operator to the focal stack and finally we compute the optimal estimated distances by describing the 3D reconstruction problem using the Markov Random Field formalism [8].

3.5. Stereo Accuracy estimation: The STACC module.

Since the DE module provides us true ground data of the scene, following [9], we compute two quality measures based on this known ground truth data:

1. Relative RMS (root-mean-squared) error between the computed distance map $z_C(i, j)$ and the ground truth map $z_T(i, j)$, i.e.,

$$R = \frac{\left(\frac{1}{N} \sum_{(i,j)} (z_C(i,j) - z_T(i,j))^2\right)^{1/2}}{z_{FOCUS}} \quad (1)$$

where N is the total number of pixels and z_{FOCUS} is the focus distance of the CAFADIS camera.

2. Percentage of bad matching pixels,

$$B = \frac{100}{N} \sum_{(i,j)} (z_C(i, j) - z_T(i, j) > \delta z_{FOCUS}) \quad (2)$$

where δ is a relative error tolerance. For the experiments in this paper we use $\delta = 0.1$.

4. EXPERIMENTAL RESULTS

In this section, we describe the experiments used to evaluate the individual modules of our system. Our main interest is to test the coherence between the direct (rendering) and inverse (CAFADIS reconstruction) estimations of the distance map.

4.1 Test data

To evaluate the system in complex environments we have selected a realistic indoor scene. The scene will be named the “armchair” scene. The CAFADIS complete image

rendered for the scene by the VCAFADIS module is shown in Figure 3 (left) and a detail from a subregion of the full image is shown in Figure 3 (right). Several difficulties arise when the 3D reconstruction algorithm tries to find the correct distances: there are many textureless regions (walls, floors) where it is very difficult to locally solve the correspondence problem, so the smoothness constraint has to be enforced. By contrast, there are abrupt discontinuities in the distance map (armchair) where the smoothness constraint fails. Note also that these abrupt changes usually coincide with occluded points.

The CAFADIS complete image is composed of 16x16 RGB subimages. Each of the subimages is composed of 64x64 pixels.

4.2 Results accuracy

The true and computed distance maps for the “armchair” scene obtained from the CAFADIS virtual image by DE and SSR modules are shown in 3D in Figures 4 and 5. Numerical results for the quality measures R and B described on Section 3.5 are given in Table 1. An histogram of the distribution of relative errors (true distance minus estimated distance divided by z_{FOCUS}) is also shown in Figure 6.

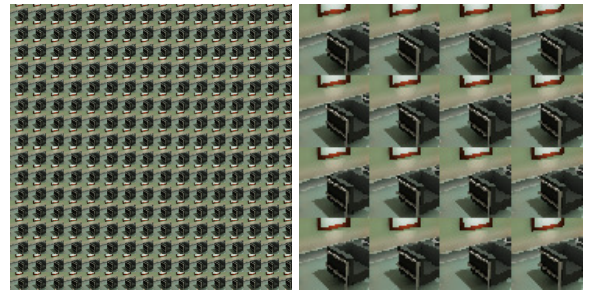


Figure 3. The complete CAFADIS image (left). Subregion of the image (left).

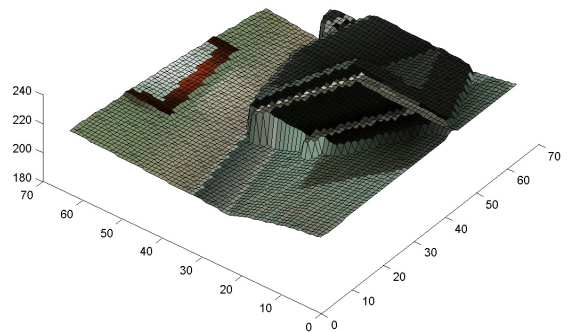


Figure 4. True distances for the “armchair” scene by rendering.

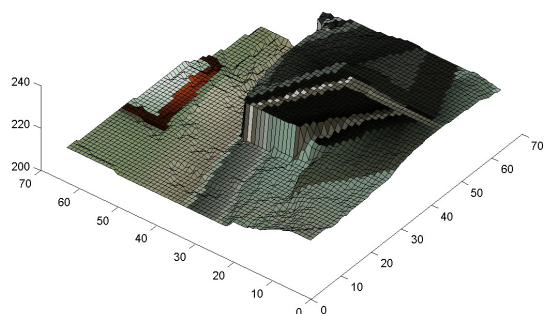


Figure 5. True distances for the “armchair” scene by CAFADIS reconstruction.

R	B
0.0463	1.00%

Table 1.Quality measures

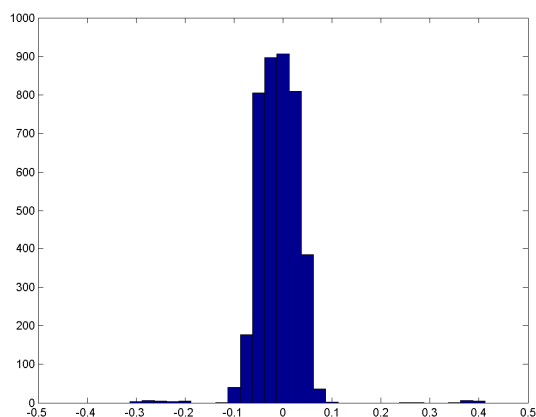


Figure 6. Relative Error (difference between rendering and CAFADIS reconstruction divided by z_{FOCUS}) histogram for the “armchair” scene.

Results show that a reliable 3D reconstruction is obtained with the CAFADIS camera.

5. CONCLUSIONS AND FUTURE WORK

We have implemented a robust tool for testing the CAFADIS camera on different scenarios running on powerful electronic hardware, namely, a GPU where high sample rates can be successfully obtained.

Future improvements will consist in porting the simulator environment to the CUDA [9] (Compute Unified Device Architecture) platform. We are also planning to port the simulator to FPGA (Field Programmable Gates Arrays) and VHDL [10] (Very High Speed Integrated Circuit Hardware Description Language), in order to increase the resolution of the output image and 3D distance map. Finally, another development plan consists in adapting the obtained scene images and 3D distance maps to drive various kinds of 3DTV displays.

6. ACKNOWLEDGMENTS

This work has been funded by “Programa Nacional I+D+i” (Project DPI 2006-09726) of the “Ministerio de Ciencia y Tecnología”, and by the “European Regional Development Fund” (ERDF).

7. REFERENCES

- [1] L. Onural: "Television in 3-D: What Are the Prospects?", Proceedings of the IEEE vol 95, n° 6, pp. 1143 – 1145, 2007.
- [2] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," CVPR , pp. 519-528, 2006.
- [3] J.G. Marichal-Hernández , L.F. Rodríguez-Ramos, F. Rosa and J.M. Rodríguez-Ramos, "Atmospheric wavefront phase recovery by use of specialized hardware : GPU's and FPGA's", Applied Optics, 44, pp. 7587-7594, 2005.
- [4] R. M. Clare and R. G. Lane, "Wave-front sensing from subdivision of the focal plane with a lenslet array," JOSA A, Vol. 22, pp. 117-125, 2005.
- [5] R. Ng, "Fourier Slice Photography". *Proceedings of SIGGRAPH*, 2005.
- [6] Philips Wowx Display Website: <http://www.business-sites.philips.com/3dsolutions/Products/3DScreens/Index.html>
- [7] R. Ziegler, "A bidirectional Light Field – Hologram Transform", *Eurographics*, vol. 26, 2007.
- [8] V. Kolmogorov and M. Wainwright. "On optimality of tree-reweighted max-product message-passing". *Proc. Uncertainty in Artificial Intelligence*, 2005.
- [9] D. Scharstein and R. Szeliski. "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms". *IJCV* 47: pp. 7-42, 2002
- [9] Cuda website: http://www.nvidia.com/object/cuda_home.html
- [10] VHDL cookbook Peter J. Ashenden: [tams-www.informatik.uni-hamburg.de/vhdl/doc/cookbook/VHDL-Cookbook.pdf](http://www.informatik.uni-hamburg.de/vhdl/doc/cookbook/VHDL-Cookbook.pdf)