A SIMULATOR FOR THE CAFADIS REAL TIME 3D TV 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. Canary Islands. SPAIN.

E-mail:{ fdoperez@ull.es[+]*, jpluke@ull.es, jmariher@ull.es, frosa@ull.es, jmramos@ull.es}*

ABSTRACT

CAFADIS is a patented camera (PCT/ES2007/000046) to measure wave-front phases and distances under different scenarios (from microns to kilometres), using highly specialised electronic technology, namely Graphics Processing Units (GPUs) and Field Programmable Gate Arrays (FPGAs).

CAFADIS employs an optical system and innovative data processing techniques pertaining to the field of computer vision and image processing. The algorithms used can be parallelized, so they are ideal for real time implementation.

CAFADIS is capable of tackling any situation requiring precise metrology at high speeds. Here we present a computer simulator for the complete scene-camera process.

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

1. INTRODUCTION

Obtaining geometrical data to explain the make-up of a three-dimensional scene, in which the most ambitious goal is to obtain distance maps, remains a challenge pending a robust solution for interactive response systems. The methods most used for this purpose are based on estimating the required distances by a series of geometric calculations once a set of scene points have been coupled. To achieve multiple perspectives, one would traditionally resort to varying configurations of sensor groups, generally two or more cameras, which determine the method for distance estimation. Alternatively, it is possible to use one single camera and to obtain multiple perspective points using micro-lenses. Among the feasible combinations of sensors and optical systems are the so-called plenoptic sensors (Adelson and Wang [1]).

Moreover, it is also possible to examine the visual information generated by a 3D scene using signal theory principles. This strategy employs multi-dimensional processing techniques with high computational costs. To adhere to the requirement of interactive time-scales response it is necessary to use advanced computer resources From this point of view, the most promising specialised hardware devices are GPUs (Graphics Processing Units) and FPGAs (Field Programmable Gate Arrays). The rapid evolution of these two platforms is described in Marichal-Hernández et al. [2].

The goal of 3D scene reconstruction is to take a set of images, and estimate the positions and orientations of the cameras that produced the images, as well as a representation of the scene that was imaged. This is an example of an *inverse* problem (for example the stereo or multiview reconstruction problems). The *forward* (or direct) problem is: given a scene and the position and orientation of a set of cameras, 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 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.

This paper is divided in five sections. In section 2 we introduce the CAFADIS camera and its optical description. Section 3 shows the different modules involved in the CAFADIS simulator, from the virtual scene generator to the distance map computation. Section 4 contains some experimental results. 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. Clare and Lane [3] use this scheme for wavefront sensing from the subdivision of the focal plane with a lenslet array. Ren Ng [4] uses the same scheme to get focused images at different scene depths.

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

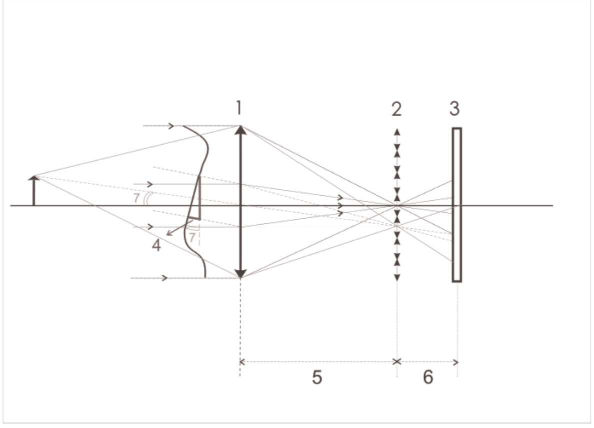

Figure 1. The CAFADIS scheme. (1) objective lens, (2) microlens array, (3) CCD, (4) wavefront, (5) microlens array position and (6) microlens focal.

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 and the distance map is obtained. This 3D map, combined with the 2D scene image, can be used as input to a 3D display [5].

In order to get the distances map as described in section 3.3, no advantage is obtained from the fact that the wavefront map can also be measured, at the same time, from the CAFADIS camera. This could also be done, as R. Ziegler *et al* [6] propose, 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 below.

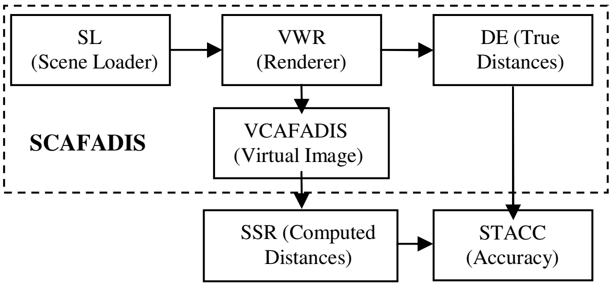

Figure 2. The SCAFADIS modules.

3.1. Scene Loader (SL)

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

3.1. 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 three-dimensional hardware acceleration. Adequately formatted information from the world, 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 for highly complex models.

3.2. Distance Extractor (DE).

What distinguishes the simulator we implemented from other virtual world viewers is the emphasis it places on acquiring the distances from the objects to the point of observation. To obtain these measurements we have 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 within complex sceneries.

3.3. Computed distances (SSR).

The distance map reconstruction in VSIDS is obtained by resolving the direct problem when in fact, for the CAFADIS camera the distance map must be obtained by resolving the inverse problem (obtaining the geometric model based on the image of the scene). This inverse problem is the stereo problem, of great importance in computer vision. To solve this problem we first compute the "focal stack" 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 [7].

3.4. Stereo Accuracy estimation: The STACC module.

Since the DE module provides us true ground data of the scene, following [8], 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.,

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

   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,

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

   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 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 3.4 are given in Table 1. A histogram of the distribution of relative errors (true distance minus estimated distance divided by z_{FOCUS}) is also shown in Figure 6.

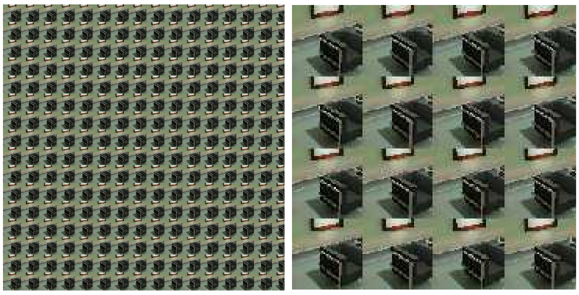

Figure 3. The complete CAFADIS image (left). Subregion of the image (left).

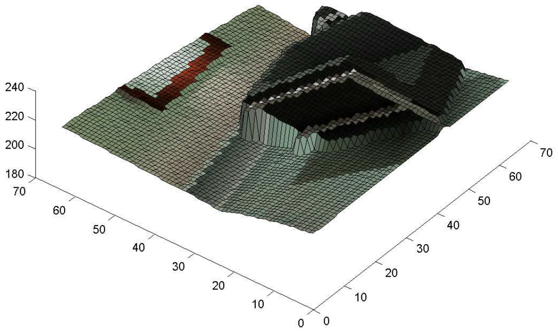

Figure 4. True distances for the "armchair" scene by rendering.

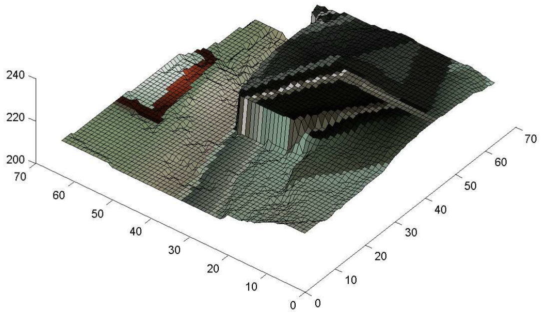

Figure 5. True distances for the "armchair" scene by CAFADIS reconstruction.

R	B
0.0463	1.00%

Table 1. Quality measures

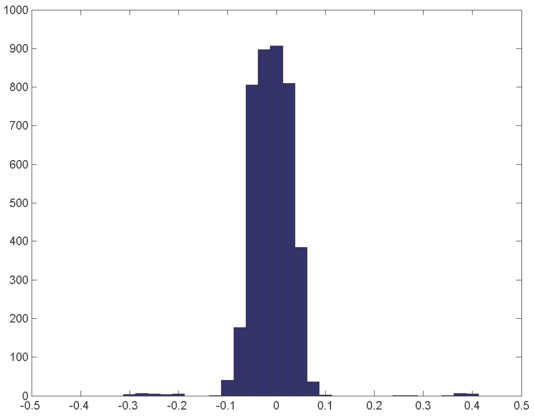

Figure 6. Relative Error (difference between rendering and CAFADIS reconstruction divided by z_{FOCUS}) histogram for the "armchair" scene.

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 real time developments and high sample rates can be successfully processed.

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) implementations, in order to increase the resolution of the real time output image and distances map. Finally, another development plan consists in adapting the obtained 3D scene images and 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] E. Adelson and J. Wang, "Single lens stereo with plenoptic camera", *IEEE transactions on pattern analysis and machine intelligence*, vol. 14, nº2, p. 99, 1992

[2] J.G. Marichal-Hernández , L.F. Rodríguez-Ramos, F. Rosa and J.M. Rodríguez-Ramos, "Atmospheric wavefront phase recovery by use of specialized hardware : graphical processing units and field-programmable gate arrays", *Applied Optics*, 44, 7587-7594, 2005.

[3] 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.

[4] Ren Ng, "Fourier Slice Photography". *Proceedings of SIGGRAPH*, Los Ángeles (U.S.A.), 2005.

[5] Philips Wowx Display Website: http://www.business-sites.philips.com/3dsolutions/Products/3DScreens/Index.html

[6] R. Ziegler, "A bidirectional Light Field – Hologram Transform", *Eurographics*, Vol. 26, 2007.

[7] V. Kolmogorov and M. Wainwright. "On optimality of tree-reweighted max-product message-passing". *Proc. Uncertainty in Artificial Intelligence*, Edinburgh 2005.

[8] 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