Development of Prototype Autostereoscopic Imaging Systems and Applications

Spyros S. Athineos[1]

National and Kapodistrian University of Athens
Dept. of Informatics & Telecommunications
sathin@di.uoa.gr

Abstract. In this thesis, prototype 3D imaging systems are developed regarding the capturing and reproduction of 3D scenes based on Integral Photography (IP) technique, for both synthetic and real 3D scenes. Regarding the capturing of synthetic 3D scenes, an IP generator is designed based on computer simulation, by exact modeling of all the necessary optical components of a single stage IP capturing system. The simulation produces high quality photorealistic IP images with increased depth of field based on ray tracing techniques and offers the ability of reproducing the original 3D scene using an appropriate microlens array. Moreover, an extension to Integral Videography is given. Regarding the capturing of real 3D scenes, a portable IP capturing system of real 3D scenes with depth control is designed using a medium format camera along with an appropriate microlens array of high spatial sampling frequency.

Keywords: Integral Photography, Integral Imaging, Integral Videography, 3D Displays, Ray Tracing, Light Fields, Computational Photography

1 Introduction

Integral Photography (IP) or Integral Imaging (InIm), devised by Lippmann [1] in 1908, is one of the most promising methods for displaying three-dimensional (3D) images since it provides autostereoscopic viewing without eye fatigue along with full color and continuous parallax both horizontally and vertically. Today there is a revitalizing interest in IP with the evolution in micro optics, high resolution LCDs and CCDs together with the increased computational power of modern CPUs.

Nowadays it is a common practice to use computers for the generation of 3D scenes. Computer Generated Integral Photography [2] belongs to this general category and aims to the production of Integral Photography images for three-dimensional viewing. A number of software ray tracing models have been reported [3,4] for the generation of integral images. These techniques provide integral images of adequate quality but have restrictions in the complexity of the 3D scenes, while the employed ray tracing algorithms are simplified.

A similar methodology, referred in the literature as "autostereoscopic light fields"

[1] Supervisor: Nikiforos Theofanous, Professor

uses a lens array for direct viewing of light fields [5]. This technique refers to focusing and depth of field (dof) problems. However, the reconstruction stage is shortly covered with no reference to pseudoscopy elimination and the required gap between the lens array and the display panel in order to clearly differentiate between real and virtual 3D scenes.

The objective of this thesis is to develop prototype 3D autostereoscopic imaging systems using Integral Imaging, which has been selected as a most promising between all spatial display techniques. As a first step, a computer simulation has been developed, of a physically implemented single step Integral Imaging capture scheme [6], using the POV-Ray software package as the ray-tracing engine [7]. The simulation of the capturing optics has been realized by modeling the microlens array as an ordinary object of the 3D scene [8] using the ray tracer's scene description scripting language. This approach takes advantage of the optimized algorithms implemented in POV-Ray in order to produce high quality photorealistic 3D images.

Furthermore, an additional imaging lens is modeled [8] as part of our capturing setup. With this imaging lens we have the ability to capture virtual and real IP images with depth control. At the reconstruction stage, a lens array is placed on top of the IP image, and the reconstructed 3D scene is formed in space in front of the lens array or behind it.

Regarding IP capturing of real 3D scenes, relevant research has utilized a small format digital camera for the capturing of large 3D scenes at a distance from the capturing device and the reconstruction of the 3D scene using a digital projector [9]. That system exhibited increased dof and viewing angle, however, the resolution of the integrated image was poor. Furthermore, the utilized setup produced orthoscopic virtual images only.

In the relevant to Integral Imaging research area of light fields, a handheld 3D camera was presented in an attempt to capture the 4D light fields by a high-resolution 2D CCD sensor [10]. The aim of that system was not the reconstruction of the captured 3D image in space, but the presentation of different slices of the captured 3D content by computationally focusing on various depth levels of the 3D scene.

A work was presented recently that is focused in the area of light fields but the 3D camera presented, is not based on high spatial resolution lens arrays but on optical blocks that consist of negative lenses and prisms, attached externally in front of the camera system [11]. Outcomes regarding variable focusing depth were presented. However, because of the low spatial sampling frequency, lot of post processing was needed, using computer vision techniques, such as tri-view morphing.

In this thesis, for the capturing and reproduction of real 3D scenes, a portable 3D camera has been developed, based on the configuration proposed in [10], but focused on capturing and reconstruction of real 3D scenes with increased dof. At the capturing stage, a lens array with high spatial sampling frequency is used to capture integral images presenting both real and virtual orthoscopic parts, therefore exhibiting an increased dof. At the reconstruction stage, a different MLA is used, due to the low resolution of current technology LCD displays. At an intermediate post processing stage, the captured image is pseudoscopically corrected, and scaled appropriately to match the geometrical and optical parameters of the reconstruction lens array. The produced orthoscopic 3D image is formed in space in front and/or behind the display device. As an alternative to presenting the 3D information content of the capture image, computational InIm is adopted, for the extraction of different perspectives of the original 3D scene.

2 Single Stage Synthetic IP Capturing Setup

The single stage IP capturing setup [6] that has been physically implemented with POV-Ray [7,8] is depicted in Fig. 1. In this setup, the imaging lens forms an inverted and demagnified real image of the original 3D scene. An important advantage of this capturing system is that both real and virtual integral images can be produced depending on the position of the microlens array (MLA) in the resulted image space.

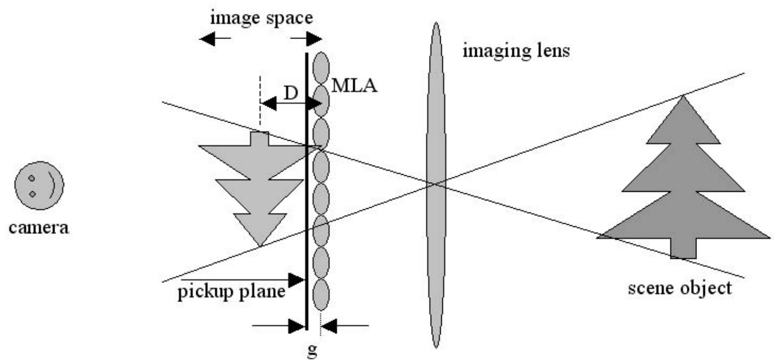


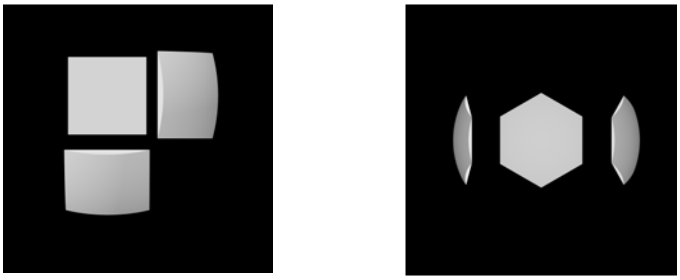
Fig.1. Single stage IP capturing setup for production of real and virtual images (distances not in scale).

The produced integral images are pseudoscopic because of the inversed depth phenomenon that is inherent in a single stage IP capturing system. These pseudoscopic images are then computationally converted to orthoscopic ones, by performing an 180º rotation of each elemental image (that is, the subimage that corresponds to each microlens) around its optical axis [12].

The relative distance D of the MLA in regard to the central plane of the image space, determines which parts of the IP image are real or virtual. When the MLA is positioned at the end of the image space towards the imaging lens as in Fig. 1, a virtual pseudoscopic integral image is produced and the pickup integral image is formed at the pickup plane at a distance g from the MLA, which is approximately equal to its back focal length. After the pseudoscopy elimination procedure, a real orthoscopic integral image is produced. At the reconstruction stage, the 3D scene floats in space in front of the MLA towards the observer. This kind of 3D reconstruction is more attractive and realistic to the observer than a virtual one [6], and for this reason it has been selected for realization in the modeled capturing setup.

All optical components of the capturing IP setup are modeled using the ray tracer's scene description language. The geometrical and optical characteristics of the plano-convex lenses are taken into account and Constructive Solid Geometry (CSG) techniques are used for the construction of each microlens and the microlens arrays.

Each microlens is formed as the intersection of a sphere and a parallelepiped or a hexagonal prism in order to produce square or hexagonal lenslets respectively, as depicted in Fig. 2. In either case, the formed microlenses are fully apertured and the radius *R* of the sphere corresponds to the radius of curvature of the convex surface.



(a) (b)

Fig. 2. 3D views of simulated microlenses, (a) square-based, (b) hexagonal-based

As for the imaging lens structure, a well corrected physically based camera model has been proposed for computer graphics [13], which offers a superior optical performance but uses a large number of optical elements. However, the modeling of a complex imaging lens substantially increases rendering time, since the required size of the 3D image captured must be comparable to the MLA size for optimal results. Therefore, we have used a simpler imaging lens model. The optical parameters have been specified using the ZEMAX optical design software package.

3 Experimental Results

The modeled IP capturing system has been tested using a 3D scene of increased complexity and real and virtual orthoscopic images have been produced. The detailed setup that has been modeled is depicted in Fig. 3. The imaging lens is a composition of three plano-convex large aperture lenses.

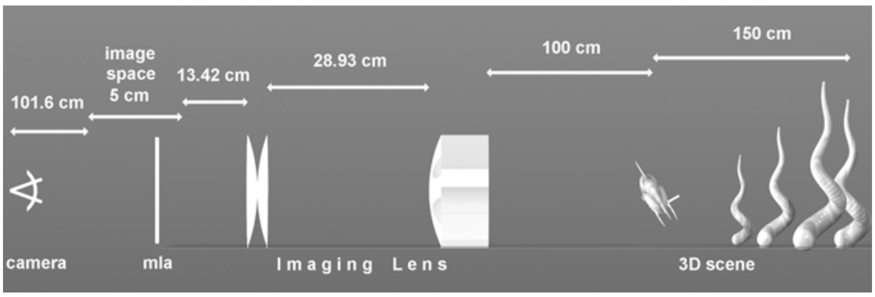


Fig. 3. Single stage IP capturing setup for generation of real and virtual orthoscopic images (distances and object lengths are not in scale)

The position of the MLA within the image space controls the type of IP images that will be produced (real or virtual). Synthetic captured IP images for four successive MLA positions relative to the central plane of the fish body are shown in Figs. 4a-4d in order to demonstrate the transition from real to virtual 3D images.

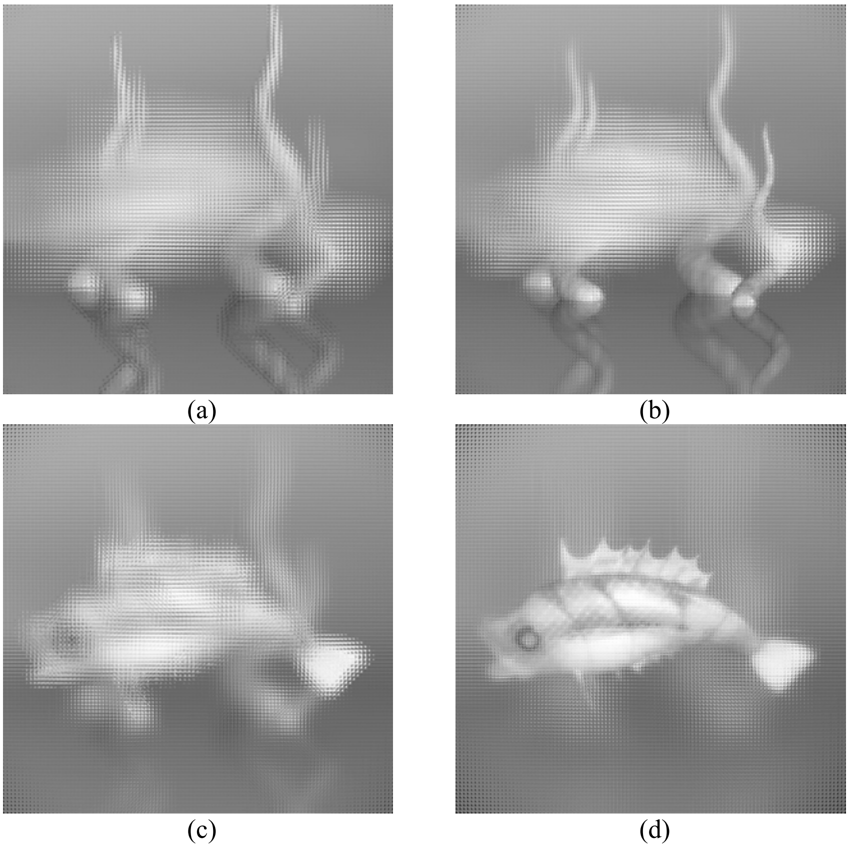


Fig. 4. IP images captured by varying the MLA position at different depths within the 3D image space. All distances are referred to the central plane of the fish image. (a) MLA at 6 cm towards the imaging lens (the MLA is in front of image space, therefore, at the reconstruction stage, the whole three-dimensional image is formed in front of the MLA - orthoscopic real image), (b) MLA at 4 cm towards the imaging lens (the fish is formed in front of the MLA, while the stems are formed just behind the MLA), (c) MLA at 1 cm towards the imaging lens (the fish is formed just in front of the MLA, while the stems are formed behind the MLA), (d) MLA at 1 cm towards the camera (the MLA is behind image space, therefore, at the reconstruction stage, the whole three-dimensional image is formed behind the MLA - orthoscopic virtual image)

The use of the same MLA in the reconstruction as the one modelled, results in a 3D scene identical to the one sampled by the MLA. A virtual 3D image was formed at a certain depth behind the display panel and exhibited smooth parallax while a real 3D image floated in space in front of the display panel. In the latter case the three-dimensional scene appeared more attractive and realistic. The resolution of the reconstructed 3D image depends strongly on its depth and image quality deteriorates as image depth increases. Therefore, in order to produce a high quality photorealistic

3D image, it is often preferable to combine a real and a virtual 3D image with a reasonable depth.

For the reconstruction stage, it is evident that a real orthoscopic 3D scene that is formed in space and in front of the lens array, cannot be easily presented using conventional 2D photography techniques. However, the 3D information contained in each IP image, as those depicted in Fig. 4, can be shown indirectly using an IP viewer that downsamples the captured IP images by appropriate spatial filtering of the corresponding pixel information under each microlens or adjacent microlenses resulting to 2D views for different viewing angles or depths. Sequences of different views of the 3D scene extracted from a single IP image were generated with this viewer and are presented in [14].

4 Integral Videography

Integral Videography (IV) is an animated extension of Integral Photography. The motivation for IV except from the obvious 3D applications has resulted from the need to have a controllable source of 3D videos for studying novel video compression techniques, which is of vital importance due to the high resolution of the IV frames needed and the associated huge volumes of data.
As in normal video an IV movie is produced as a sequence of integral images in time. However in IV, the primary parameter affecting the quality of the reconstructed 3D scene is the resolution of the display device used. For IP reconstruction, the MLA pitch determines the lateral resolution of the 3D scene produced. With maximum resolutions of 200 dpi that are currently available for LCD screens, a minimum acceptable quality can be achieved using 8x8 pixels under each microlens of 1 mm pitch size, as depicted in Fig. 5.

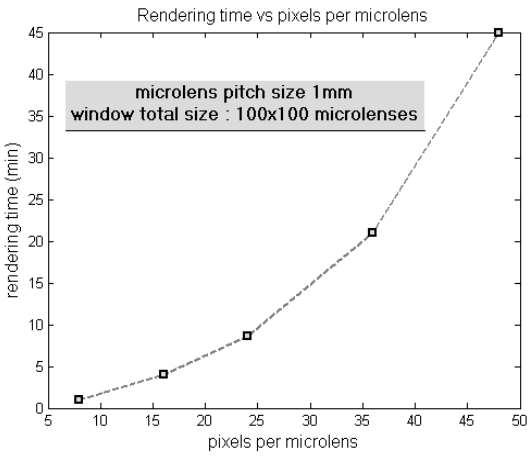


Fig. 5. Rendering time results versus number of pixels per microlens. A total of 100x100 microlenses are used for each render, with a square MLA of 1mm pitch without antialising.

Another important issue regarding IV was the use of the ray tracer antialiasing options. Rendering time was greatly affected from antialising because of the increased number of supersamples used. In Fig. 6, we depict the variation of rendering time versus antialising threshold in POV-Ray, which is a parameter inversely proportional to the number of supersamples [7]. It should be pointed out, that a trade-off exists between microimage smoothness, pseudoscopy elimination and rendering time.

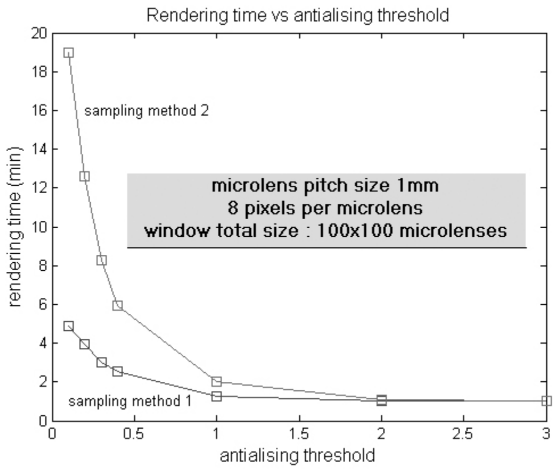


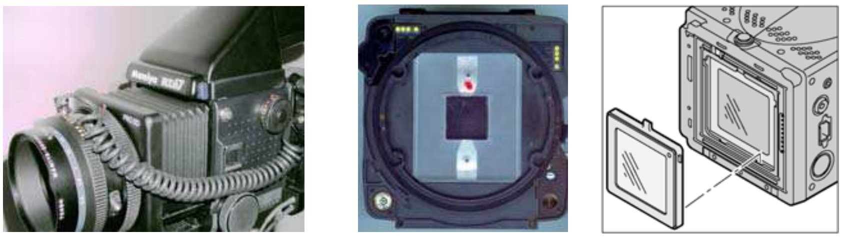
Fig. 6. Rendering time versus antialising threshold. A total of 100x100 microlenses were used for each render. A rectangular MLA was used with 1mm pitch. The sampling method 1 was an adaptive, non-recursive, super-sampling method. The sampling method 2 was an adaptive, recursive, super-sampling method with control over the maximum number of samples taken for a super-sampled pixel (we have set the corresponding depth-control parameter to 3).

Furthermore, each IV frame is substantially different in nature from a typical video frame since 3D information is embedded. Therefore, in an IV video in which the camera is still and the object is moving, the number of pixels under each microlens, the MLA size and the focusing distance of the microlenses, define the amount of 3D information enclosed in each IV frame. By downsampling the corresponding IV frame by appropriate spatial filtering of the pixel information under each microimage, we have extracted 2D views of the 3D scene for each IV frame.

5 Portable IP capturing setup of real 3D scenes

A portable IP capturing setup of real 3D scenes cannot easily be accomplished utilizing small format digital cameras, because of the increased difficulty to accurately attach a microlens array of high spatial sampling frequency in front of the sensor. Moreover, the sensor area in 35mm digital cameras is usually restricted thus limiting the captured 3D content.

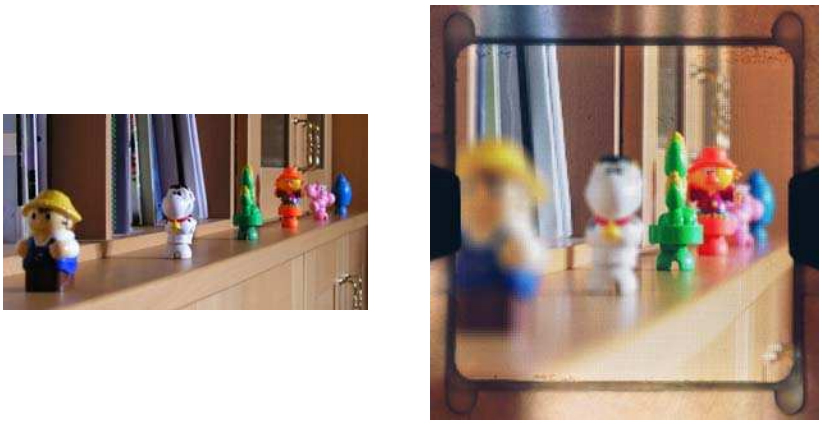
Therefore, a medium format camera, as depicted in Figure 7a, was utilized which could be easily disassembled in parts and accepts either film or a digital back with a high-resolution CCD sensor.



(a) (b) (c)

Figure 7. (a) medium format camera Mamiya RZ67 Pro IID, accepting either film or a digital back. (b) film holder with MLA attached (c) digital back Kodak DCS Pro Back Plus (the CCD sensor and a removable IR filter are depicted).

The digital back Kodak DCS Pro Back Plus is equipped with the Kodak full frame color CCD sensor KAF 16802CE with dimensions 3.7cm x 3.7cm and pixel size 9μm. This sensor has 4080x4080 active pixels (16Mpixels) offering the required high resolution for InIm capturing. An MLA with typical pitch size of 100um or 125um must be placed in front of this sensor, to give 11x11 or 13x13 pixels under each microlens, respectively. Depending on the f/# of the MLA used, the 3D content of the captured image is adequate for the extraction of depth information and multiple perspectives of the 3D scene. Alternatively, a film holder can be used in place of the digital back, with the MLA accurately positioned in front of the film, as depicted in Fig. 7b. This setup, results in high quality IP images as the one depicted in Fig. 8b.



(a) (b)

Figure 8. Captured IP image of real 3D scene, (a) 3D scene captured by a conventional 2D camera, (b) captured IP image exhibiting both real and virtual parts

The MLA used in the capturing process, has a high spatial sampling resolution. Therefore, this MLA cannot be matched with an LCD display for the reconstruction of the 3D image, due to the low resolution of current LCD display technology. A possible solution to this problem is the appropriate scaling of the captured image [15], in order to match the geometrical and optical parameters of an MLA suitable for use with a high-resolution LCD display.

6. Conclusions & Future Work

In this thesis, a novel way for producing high quality, photorealistic integral images of complex synthetic or real 3D scenes has been proposed.

Capturing of synthetic 3D scenes has been accomplished using an advanced general-purpose ray tracing software package. With this approach all necessary optics are modeled like ordinary objects of the 3D scene. This methodology constitutes a source of IP images and IVs with controllable 3D content for developing new compression techniques for 3D still images and videos and studying the reconstruction stage concerning viewing angle and depth.

The proposed methodology offers full depth control and positioning of the reconstructed 3D scene. Besides, the modeling of an MLA using real world parameters, further ensures that the reconstructed 3D scene has optimum quality. In addition, the proposed technique has the advantage of allowing the combination of real and virtual IP images for autostereoscopic viewing of complex photorealistic 3D scenes exhibiting mixed depth in front and behind the display device. The methodology presented can be easily extended to Integral Videography, producing high-quality 3D videos along with depth control.

Currently, raster graphics are the dominating technology used for computer graphics, but the rendered images can hardly reach the photorealism achieved with ray tracing techniques especially for more advanced 3D scenes [16]. Ray tracing has increased computational cost compared to raster graphics. However, as the complexity of the 3D scene increases, the ray tracing approach takes advantage over raster graphics concerning computational requirements, thus it is expected that hardware accelerated ray traces will prevail in the future in computer graphics [17, 18]. In this context, the proposed methodology is expected to be of significant importance for computer generated three-dimensional display techniques. However, more work should be done in modeling physically realizable, well-corrected lens systems of increased complexity [13] especially in case of modeling MLAs with sizes comparable to CCDs. In addition, an important drawback of the proposed ray tracing approach is that the rendering time is far from considered real time, thus hardware-accelerated ray tracing techniques should be considered [17, 18].

Capturing of real 3D scenes has been accomplished using a medium format camera along with an MLA of high spatial sampling resolution. However, much work has to be done in the field of digital cameras for the development of the next generation 3D cameras that incorporate 3D capturing systems as well as the required processing algorithms.

References

1. G. Lippmann, "La Photographie Integrale", Comptes-Rendus Academie des Sciences 146, pp 446-451 (1908).
2. B. Lee, S. -W. Min, S. Jung, and J.-H. Park "A three-dimensional display system based on computer-generated integral photography", J. Soc. 3D Broadcast. Imag. 1(1), pp 78–82 (2000).
3. Y. Igarashi, H. Murata, M. Ueda "3D Display System Using a Computer Generated Integral Photograph", Japan J.Appl.Phys. 17, pp 1683-1684 (1978).

4. Graham Milnthorpe, Malcolm McCormick, Neil Davies "Computer Modeling of Lens Arrays for Integral Image Rendering", IEEE Computer Society, Proc. of EGUK'02, pp 136-141 (2002).
5. A. Isaksen, L. McMillan, and S.J. Gortler, "Dynamically reparameterized light fields", SIGGRAPH '00 Proceedings, pp 297-306 (2000).
6. Ju-Seog Jang, Bahram Javidi "Formation of orthoscopic three-dimensional real images in direct pickup one-step integral imaging", Optical Engineering 42(7), pp1869-1870 (2003).
7. povray: http://www.povray.org
8. S. Athineos, N. Sgouros, P. Papageorgas, D. Maroulis, M. Sangriotis, N. Theofanous "Photorealistic Integral Photography using a Ray Traced Model of the Capturing Optics", in Journal of Electronic Imaging, 15(4), Oct/Dec 2006.
9. Ju-Seog Jang, Bahram Javidi, "Depth and lateral size control of threedimensional images in projection integral imaging", Opt. Express 12(16), pp. 3778-3790 (2004).
10. Stanford technical report, light field camera: http://graphics.stanford.edu/papers/lfcamera
11. Todor Georgeiv, Ke Colin Zheng, Brian Curless, David Salesin, Shree Nayar, and Chintan Intwala1, "Spatio-Angular Resolution Tradeoff in Integral Photography, " Tomas Akenine-Möller and Wolfgang Heidrich (Editors), Eurographics Symposium on Rendering (2006).
12. N.Sgouros, S.Athineos, M.Sangriotis, P.Papageorgas and N.Theofanous, "Accurate lattice extraction in integral images, " in Opt. Express 14(22), pp.10403-10409 (Oct. 2006).
13. C. Kolb, D. Mitchell, and P. Hanrahan, "A Realistic Camera Model for Computer Graphics", Computer Graphics SIGGRAPH '95 Proceedings, pp 317-324 (1995).
14. imaging results with 2D viewer: http://imaging.di.uoa.gr
15. Jae-Hyeung Park, Heejin Choi, Yunhee Kim, Joohwan Kim and Byoungho Lee, "Scaling of Three-Dimensional Integral Imaging", Japanese Journal of Applied Physics 44 (1A), pp 216–224 (2005).
16. Jim Hurley, "Ray Tracing Goes Mainstream", Intel Technology Journal 09(02), pp 99-107 (2005), http://developer.intel.com/technology/itj/index.htm
17. Timothy J. Purcell, *Ray Tracing on a Stream Processor*, PhD thesis, Stanford University (2004).
18. Sven Woop, Jörg Schmittler , Philipp Slusallek , "RPU: a programmable ray processing unit for realtime ray tracing", ACM Transactions on Graphics 24(3), pp 434 – 444, (2005).