Building a Photo Studio for Measurement Purposes

Michael Goesele — Wolfgang Heidrich
Hendrik P. A. Lensch — Hans-Peter Seidel

Max-Planck-Institut für Informatik, Computer Graphics Group
Im Stadtwald, 66123 Saarbrücken, Germany
Email: {goesele,heidrich,lensch,hpseidel}@mpi-sb.mpg.de

Abstract

A digital photo studio was built to generate high quality real world input data for various image-based rendering and vision algorithms. Special attention was paid to carefully control the lighting conditions in order to be able to acquire exact data about the surface properties of objects using readily available digital camera technology.

This paper discusses the specific demands and requirements that arise from these goals for the equipment in the photo studio. Furthermore, we describe the typical workflow for one of our current projects, thereby illustrating the interoperation between the different devices.

1 Introduction

The need for more realistic input data for various image-based rendering and vision algorithms increases more and more. In addition to traditional multimedia applications, e-Commerce seems to be a driving force for new developments. Before buying a product online shoppers would like to get a visual impression of the product that is at least as good as the impression they can get from a paper catalogue. Once the necessary data is acquired they can look at the product from all sides by turning it on the screen using image-based rendering techniques.

Although most products are fabricated in an industrial production process, there often exists no model that can be used for this purpose. Therefore it is important to be able to acquire a realistic 3D model together with appropriate information about the object's surface properties in a fast and efficient way.

In contrast to CAD models, ray traced images, and other synthetic scenes, real world data usually contains artifacts from noise, limited sensor resolution, lens errors, or other sources. As it is often not possible to simply ignore these artifacts, new procedures have to be developed and existing algorithms must be modified in order to use data acquired from real world objects. For this reason, obtaining realistic models for geometry, reflection properties, and other physical object properties from image data has become an increasingly important area of research in the past few years.

During the lasts months we built a digital photo studio that is tailored to our specific needs in acquiring this kind of input data. It consists mainly of a digital camera, almost point shaped light sources, a low reflection (i.e. black) room, and resourceful computing equipment.

In the following sections we describe our (idealistic) requirements for each component of the photo studio and to what extend these are met by the equipment we have. Finally, we discuss the scenario for one particular project we are currently working on to illustrate the use and interoperation of the different devices.

2 The Camera

A camera is the central piece of hardware for an image-based measurement lab such as the one envisioned. Analog cameras are not practical for this kind of purpose, mostly due to the long development cycle, and the necessity to digitize the image afterwards, which is either tedious or, if done by a photo lab, out of control of the researcher.

This leaves digital cameras as the only feasible possibility. The choice of a particular model has to be made based on a variety of criteria that have to be fulfilled for different applications. These criteria, and the tradeoff between them will be described in the following.

2.1 General Requirements

When using a camera as a measurement device, we demand that it records accurate (10-12 bit per channel) and high resolution image information. But we also want to have the flexibility and user-friendliness of analog 35 mm cameras.

In addition to these criteria, an analysis of different measurements we wanted to perform in the lab showed that the camera must be capable of taking images of both still objects and living subjects, and that the ability to obtain color images was a necessity. Furthermore, the camera should be able to deal with a variety of different lighting situations, especially those that require long exposure times, which are for example necessary to generate high dynamic range images. Additionally, objects of different sizes will be captured which implies that the lenses should be interchangeable, and a wide variety of different lenses should be available. Also, since the camera will be used for measurement purposes, we would not like to introduce artifacts by using a lossy compression technique such as JPEG. These two criteria rule out most of the consumer quality cameras.

Finally, a very important criterion for simplifying the actual measurements is that the camera should be completely controllable remotely from a computer. Since measurements often require us to take series of images with different exposure, aperture, and focus settings, the need to control these parameters at the camera quickly becomes a nuisance for the researcher. Also, any kind of manual adjustment at the camera bears the danger of slight changes in camera position and orientation.

With current camera technology, the highest possible resolution can be achieved with scanning back cameras where a single sensor line is moved across the image plane. However this technology which is comparable to a flatbed scanner makes it impossible to acquire images of living subjects.

For generating color images with digital photo cameras, three different approaches are widely used. The first one is to take several images with a monochromatic camera using different color filters (usually in the form of a filter wheel). This approach is impractical for moving objects. Alternatively, the image can be simultaneously captured by separate CCD chips, one for each color channel. Due to their complicated optical and mechanical system these cameras are difficult to produce and not very common.

The third method which yields lower quality results is to use a single CCD chip with color filters in front of the individual pixels. This way, a color image can be obtained with a single exposure, but every pixel will only have information about *one* color channel corresponding to the color of the filter in front of it. This means that for a full RGB image the individual channels have to be interpolated across the image, which reduces the spatial resolution. Despite this shortcoming, our need to deal with living subjects mandates the use of the latter solution.

A common problem of CCD cameras is high noise for large exposure times. The best available cameras for measurement purposes are used in astronomic observatories. They often contain custom built sensors with very low noise levels that can be completely controlled by a host computer [11]. They are usually cooled with liquid nitrogen and can typically be used for exposure times of up to one hour limited by external effects as cosmic rays hitting the sensor. Unfortunately, these cameras are typically only available in monochromatic versions, and the cooling systems increase size and weight of the camera.

2.2 The Choice of Camera

As a result of the above discussion, we have decided for a high-end digital camera intended for professional photography. In particular, we have chosen the Kodak DCS 560, which is a 35 mm focal plane shutter SLR camera based on a Canon EOS-1N body. It has all the usual features of current SLR cameras like an auto focus and auto exposure system and can be used with most Canon EF autofocus lenses. The 2008×3040 pixel imaging sensor is smaller than a regular 35 mm film frame leading to a lens magnification factor of 1.3. Nonetheless all the standard lenses for the Canon body can be used. Its sensitivity is comparable to a film with 80–200 ASA speed. A more detailed technical specification can be found in [4].

The camera produces images with 12 bit color depth in a Kodak proprietary TIFF format. The host computer software converts them into color or grayscale images with up to 16 bit color depth.

If the camera is used as standalone system it can be operated in the same way as an analog SLR camera. The only difference is that the images are stored on an internal 540 MB hard drive. But it is also possible to control it remotely via a serial IEEE 1394 FireWire connection. In this case most of the settings can be changes by the host computer and images can be acquired and downloaded via FireWire.

Kodak provides a SDK which allows the user to write their own programs to control the camera on a Windows or Mac platform.

3 The Lamps

3.1 Light Source Geometry

For many image-based rendering and vision algorithms the solid angle formed by the incident light rays should be well defined and as small as possible for any given surface point. This can be achieved by either using a parallel light source or a point light source.

The most common parallel light source is a collimated laser beam. It is often used in gonioreflectometers to measure the BRDF of a surface. But it is hard to illuminate a large object with a collimated laser beam as a lens with the same diameter as the object is needed. Furthermore, a laser emits light of a single wavelength which makes it impossible to record the color of the surface. Due to the coherence of laser light diffraction patterns are formed.

A point light source can be approximated by a nearly point shaped light source or by a virtual point light source created by adding an aperture, a lens system, or a reflector to a regular lamp. Placing a small aperture in front of a non-point shaped light source leads to two problems. Only a small fraction of the emitted light will pass the aperture, and the angular distribution of the light is rather inhomogeneous as the aperture forms a pin hole camera which projects an image of the light source onto the object.

For a lens based light system the effective size of the light source equals the size of the lens which is usually quite large. The same holds for a system with a reflector unless it contains an ideal point shaped light source and the light is always reflected in the direction it is coming from.

If a non-ideal light source is put at a large distance from the illuminated object the solid angle formed by the incident light rays becomes very small. Any light source can therefore be used as an approximation for distant parallel or point light sources. But as the irradiance decreases with the inverse square of the distance very powerful light sources have to be used for this approach.

3.2 Photometric Requirements

The irradiance at each surface point must be known in order to interpret the light reaching the camera from that point. It can be determined by a calibration procedure for each setup. However it is much more desirable to be able to assume that an equal amount of light is emitted from the light source within a certain area resp. a certain solid angle as it is done by a homogeneous point light source.

In order to acquire high quality color information the spectrum of the emitted light should be constant during a single photo session but also for several sessions over the course of weeks or even months. Furthermore the spectrum should be approximately constant over the visible spectrum to allow for a good color reproduction [9].

For practical reasons a point light source must be very bright as most of the light (i.e. usually more than 90%) is not emitted in direction of the object. On the other hand the power consumption should be as small as possible to avoid excessive heat production.

3.3 Real Lamps

Most of the photometric requirements in the previous section are also common for photographic and cinematographic applications. But a point light source is usually not desired as it produces very hard shadows and is rather unpleasant to look at for human subjects.

We would like to be able to record high dynamic range images [3] by combining several differently exposed images of a scene. Varying the exposure using flashes is only possible within a narrow range as the actual exposure time has no influence on the exposure. We therefore decided to use constant light sources instead of flashes.

Although tungsten lamps are often used their low color temperature leads to poor color reproduction and yellowish images which have to be corrected using white balance techniques. Due to their low efficiency these lamps have to be rather strong producing a large amount of heat.

These disadvantages can be overcome using HMI discharge lamps [10]. In these lamps a luminous arc burns between two electrodes (see Figure 1). It produces a very bright light whose spectrum is similar to daylight. Their efficiency is about 3 times larger than the efficiency of tungsten lamps.

Unfortunately HMI bulbs produce a large amount of UV radiation which must be removed as it poses a serious health hazard. Therefore HMI bulbs are usually mounted in a housing with a safety glass glass plate and a reflector rendering them unusable as point light sources.

After a long search we bought 2 Joker-Bug 800 lamps which are produced by the French company K5600. They contain a special 800 W HMI bulb with a coating that shields the UV radiation. A clear safety glass beaker encloses the bulb but still allows the lamp to be used as point light source.

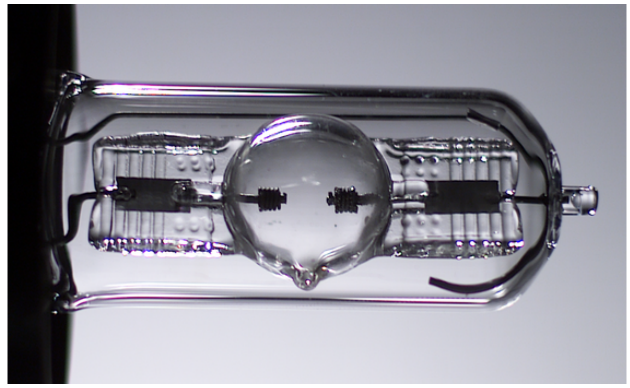

Figure 1: *A HMI bulb mounted in a lamp. An electric arc is produced between the two electrodes in the inner glass ball. The effective light source diameter is smaller than 1 cm.*

In addition we have several reflectors and diffusor boxes which can be mounted on the lamps to generate diffuse lighting situations as are commonly used by photographers.

4 The Room

Our plan was to divide the room in two parts – a small computer lab with additional storage facilities for our equipment and the actual photo studio. People should be able to work in the computer lab e.g. remote controlling the camera while the photo studio is used to take images under controlled lighting conditions.

The most important requirement for the photo studio was that no light bouncing of the walls, the floor, or the ceiling should illuminate the scene. This means that as much light as possible must be absorbed by the walls, the floor, and the ceiling of the room. Furthermore no light from the outside (i.e. the computer lab) should get inside the photo studio.

We decided to us a 24 m^2 sized area as photo studio. The walls and the ceiling were to be covered with black cloth and a black carpet was put in. Two black curtains allow for easy access to the photo studio. For the selection of the black materials the following issues were considered important.

The correct exposure for a scene is usually determined by measuring the light that is reflected from a 18% reflectance gray card. If this value is put on Zone 5 of the Zone system [1] the lowest useful density of a negative is at Zone 1 which corresponds to 1.125% reflectance. The dynamic range of a digital sensor is potentially even higher.

Unfortunately the black cloth we looked at reflects still about 2% of the incident light which means that even for an analog film the black cloth does still not look black when the gray card is correctly exposed. However we expect that this is no problem for most practical situations as we can usually manage to have a lower light level on the cloth than on the scene.

The second requirement for the materials was that the remaining light should be reflected in a diffuse way. As curtains tend to be in a curly shape arbitrary combinations of incident light and viewing direction will appear. If this leads to specular highlights the measurements will be incorrect.

Weighting these requirements we finally decided to use a thick felt to cover the walls and the ceiling. It reflects light in a very diffuse way and is dense enough to separate the computer lab from the photo studio. For the floor we choose a black needle fleece carpet. Since the curtains block any kind of air flow, and the lamps radiate a serious amount of heat, a good ventilation or air conditioning of the room is mandatory.

Figure 2 shows an image of the interior of the photo studio in a working situation. A high dynamic range image was generated from a series of images with different exposure times [3]. By applying a suitable tone-mapper (e.g. [13, 14, 15]) the dynamic range is compressed and more details can be seen.

5 Additional Infrastructure

A single compressed raw image of the DCS 560 in a Kodak proprietary TIFF format needs between 5 and 6 MB of storage. During the conversion into a readable TIFF format the size can grow up to 36 MB. As we are usually taking a large number of images, which we typically want to keep around for an indefinite period of time, a sophisticate storage solution is required.

Figure 2: *The photo studio. This image was generated from a high dynamic range image by applying a tone-mapper.*

At present we use a semi-automatic, three level hierarchical storage system. The camera is controlled by a Windows NT PC with 1 GB of main memory and several GB of temporary disk space. After the conversion and a first validation phase, all useful image data is transferred to a central file server with a 190 GB RAID 5 array via FastEthernet (a Gigabit Ethernet line between the two machines is in preparation). As third level we use a robot controlled tape archive system, from which older data can be retrieved at any time.

6 A Sample Project

Using the studio equipment as described above, the typical workflow for one of our current projects, which deals with the reconstruction of surface reflection properties from images, is as follows:

- acquisition of the 3D model from a range scanner
- calibration of the intrinsic camera parameters
- positioning and calibration of the light source
- selection of a number of camera viewpoints, and for each viewpoint
  - acquisition of image data
  - generation of high dynamic range images
  - registration of image and 3D model

- fitting of a reflection model to the samples contained in all acquired images

In the following we describe some parts of this general setting in more detail.

6.1 Acquisition of Image Data

To acquire an image of an object we are following standard procedures for object photography. However it is sometimes desirable to take images of the whole surface of the object without changing the lighting. In this case we try to illuminate the whole object with the available light sources and additional reflectors made from styrofoam.

In contrast to digital cameras used in astronomy, the DCS 560 is not a genuine measurement device. It is optimized to produce visually pleasing images instead of recording exact information about the scene. The camera and the image conversion software on the host computer perform a series of image processing steps. Some of these steps (e.g. the reconstruction of RGB data from the sensor data) are also necessary for our applications. But other steps as image sharpening or contrast enhancing lead to fake data.

At room temperature CCD sensors can produce a significant amount of noise for exposure times larger than 1 s (see Figure 3). This noise seems to be due to "hot pixels" on the chip which collect charge even when no light is hitting them. We have no means to cool the camera and therefore have to deal with this unavoidable artifact of the CCD sensor if we want to acquire images with long exposure times.

To solve these problems we use the raw sensor data from the DCS 560 to perform noise removal and image reconstruction using our own software.

6.2 Generation of High Dynamic Range Images

The dynamic range of analog film is approximately 10 aperture stops or roughly 1:1000. The full dynamic range of the DCS 560 is 1:4000 corresponding to 12 bit color depth. But this may still not be enough to record all information in a scene. Especially in cases where a bright light source and a shadowed area are on the same image no information can be recorded in a single image for either the bright or the dark areas. But many applications can reduce the dynamic range of the input data: Sometimes only parts of an image are viewed leaving out bright or dark areas, or a tone-mapper (e.g. [13, 14, 15]) is applied to explicitly reduce the dynamic range.

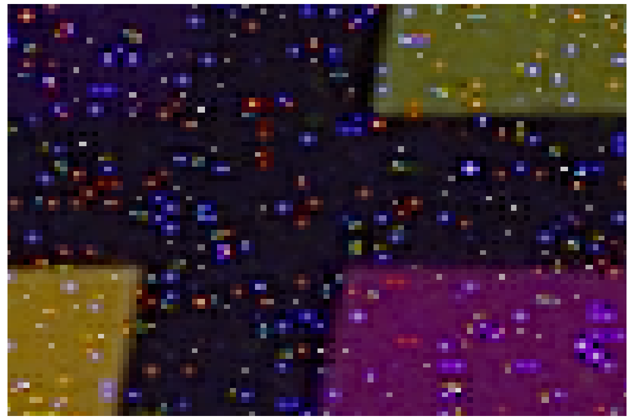

Figure 3: *A 422×281 excerpt from an image of a GretagMacBeth ColorChecker taken with the DCS 560 at a film speed of 80 ASA with 25 s exposure time. The image shows four constant colored patches separated by a black cross. The noise is mainly due to "hot pixels" on the CCD chip.*

For static scenes we therefore take usually several images with varying exposure times. By combining the values for each pixel from all suitable images an high dynamic range image can be created that contains floating point luminance values [3].

6.3 Acquisition of a 3D Model

In most cases the 3D model of an object is not available and we therefore have to acquire it ourselves. For this task we use a Steinbichler Tricolite 3D scanner. It consists from a stripe projector and a video camera mounted on a solid tripod. A binary stripe pattern is projected on the object and observed by the camera. The captured images are used to determine the position of surface points by structured light techniques. Several scans are necessary to record complete information about an object. The scans are then

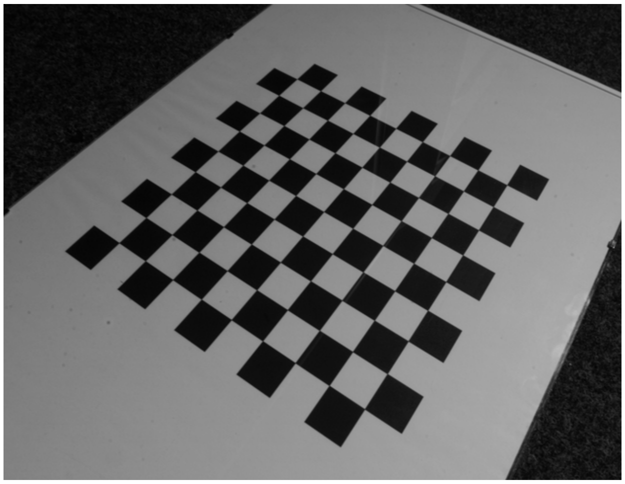

Figure 4: *Image of a calibration object taken with a 14 mm lens.*

combined to a single mesh using the Steinbichler COMETplus software.

Due to the operation principle of the scanner, concave objects are hard to scan. Each surface point must at the same time be visible from the stripe projector and from the video camera to acquire its position. Furthermore dark, patterned, or specular surfaces are hard to scan. Therefore for most objects the 3D models have to be optimized requiring user interaction.

6.4 Camera Calibration and Image Rectification

Most algorithms assume that images are acquired using a perspective projection. But this is only true for pin hole cameras but not for cameras with a lens system.

Thus the actual transformation of the camera's lens system must be determined for which a checker board calibration object is used (see Figure 4). Several images are taken with the same lens settings (i.e. object distance, focal length) that are used to capture the actual images of the object. An implementation of the Harris detector [5] included in Bouguet's camera calibration toolbox [2] is used to extract the feature points from the calibration images.

As the geometry of the calibration object is known, it can be used to extract the intrinsic data, especially radial and tangential distortion coefficients using standard camera calibration techniques [6, 12, 16]. With this information pure perspective projection images are generated.

6.5 Registration of Images and 3D Model

To generate a complete model of the object the image data and the 3D model have to be registered against each other.

For objects with a distinct silhouette the images can be registered with the 3D model using a silhouette matching algorithm [7, 8]. The input image is segmented to extract the silhouette of the object. Afterwards, a non-linear optimization is used to match the extracted silhouette with the silhouette of the object.

In each optimization step the 3D model is rendered white in front of a black background. Then the segmented image is combined using a per pixel XOR-operation. The number of pixels that differ can be calculated by evaluating the histogram of the image. Since all these steps can be executed in hardware the non-linear optimization can be significantly accelerated. Figure 5 shows an example of an image that is registered with the corresponding 3D model.

6.6 Generation of Reflection Models

Once we have acquired several different view of an object we can use them to further analyze the reflection properties of the object surface.

7 Conclusion

The main challenge in building the photo studio was to find equipment that suites our needs. As our requirement are in many cases different from the requirements of traditional photographers we sometimes had to search for a long time until we found a good solution.

8 Acknowledgments

We would like to thank all people that were involved in building our photo studio, especially

Figure 5: *Image of a wooden elk. After registering it with the 3D model of the elk a novel view is rendered. Parts of the model for which no image information is available are shown in black.*

the Max-Planck-Institut's technical and computing resource staff, Gernot Worm at Bremaphot, Marc Galerne at K5600, and Kodak's Developer Relations Group.

References

[1] Ansel Adams. *The Negative*. Little, Brown and Company, 6th paperback printing edition, 1999.

[2] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab. See http://www.vision.caltech.edu/bouguetj.

[3] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. *Proceedings of SIGGRAPH 97*, pages 369–378, August 1997.

[4] Eastman Kodak Company. Kodak Professional DCS 500 Series Digital Cameras – User's Guide, 1998.

[5] C. Harris and M.J. Stephens. A combined corner and edge detector. In *Alvey88*, pages 147–152, 1988.

[6] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. In *CVPR97*, 1997.

[7] Hendrik Lensch, Wolfgang Heidrich, and Hans-Peter Seidel. Hardware-accelerated silhouette matching. In *SIGGRAPH Sketches*, 2000.

[8] Hendrik P.A. Lensch, Wolfgang Heidrich, and Hans-Peter Seidel. Automated texture registration and stitching for real world objects. Submitted.

[9] Osram GmbH. Halogen-Metalldampflampen Foto Optik.

[10] Osram GmbH. Lighting Program '98/99 Photo Optics.

[11] A. G. Davis Philip, Kenneth A. Janes, and Arthur R. Upgren, editors. *New Developments in Array Technology and Applications. Proceedings of the 167th Symposium of the International Astronomical Union*, 1994.

[12] R.Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *RA*, 3(4):323–344, 1987.

[13] Jack Tumblin and Holly Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, 1993.

[14] Greg Ward. A contrast-based scalefactor for luminance display. In P.S. Heckbert, editor, *Graphics Gems IV*, pages 415–421. Academic Press, 1994.

[15] Greg Ward-Larson, Holly Rushmeier, and C. Piatko. A visbility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):291–306, 1997.

[16] Zhengyou Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research, 1999. Updated version of March 25, 1999.