Switched Pattern Laser Projection for Real-time Depth Extraction and Visualization through Endoscopes

Kurtis Keller [a], Jeremy Ackerman[b], and Henry Fuchs[c]

UNC Computer Science, Sitterson Hall, Chapel Hill, NC, USA
[a] keller@cs.unc.edu
[b] ackerman@cs.unc.edu
[c] fuchs@cs.unc.edu

ABSTRACT

Gathering depth information through an endoscope or laparoscope during surgical or other procedures is quite difficult. There are stereo laparoscopes but generating three-dimensional models with them is very difficult. Accurate real-time generation of three-dimensional models through a laparoscope is a needed technology to enable a wide range of surgical applications.

We have designed a miniature laparoscopic optical system consisting of a single laser whose pattern is modulated and uses the laparoscope as the optical display path into the body. Two cameras, one sensitive to the laser light and the other for full color imaging share this same tube as the laser projector but use the light from the opposite direction. The images gathered by the laser sensitive camera are used to generate a three dimensional map, and the color image is used to acquire the corresponding texture map. High-speed image processing hardware is used to generate 3D information using a structured light technique. The user can then re-render the acquired scene in 3D.

The optical system is divided into a removable upper half consisting of the cameras, laser, digital light switches and combining optics. The lower half is the laparoscope or endoscope that can be sterilized. There can be several variations in the configuration of the laparoscope optical half that tailor to different procedures.

Keywords: Depth extraction, structured light, real-time, laparoscopic

1. INTRODUCTION

During laparoscopic surgery, the surgeon is limited to the monocular viewpoint of the laparoscope. Stereo laparoscopes are available but they still are limited to their single vantage point but provide the ability for a surgeon to stereoscopically fuse the two images together. If true 3D-depth information would be included in the image, then the surgeon could have different vantage points to assist in the procedure without moving the scope. This would be especially helpful in assisting procedures in complex anatomy.

We will divide the work presented in this paper as follows:

- Application and reason for research
- Current state of the art and problems of the current systems, including our own.
- Proof of concept designs
- Medical build

We will describe earlier systems but focus more on the present research system that incorporates many, but not all, of the methods listed here. A commercial, medical version is being pursued which will combine many of the learned features here along with some that have not yet been incorporated into the prototypes or discussed here.

.

2. APPLICATIONS

As a research lab, our work is application focused. Our development focuses toward solving the problems that we encounter. Because of that, our two areas of application for the high-speed structured light are medical and teleconference. There are many more applications where 3D-depth information in digital video would be useful, but we have not investigated into them. Some of the more obvious ones are: scene reconstruction for rooms, dental cavity molding, digital studio sets for broadcast TV and other entertainment uses.

Our primary effort in real time structured light is directed at depth extraction for medical applications. Specifically, minimally invasive surgery using laparoscopes. Our area of interest is in abdominal cavity area. This particular part of the body is unique that it can be inflated with CO_2 and the organs can be viewed from the top. However, this large viewing area also causes working problems in that distances from actuator to tool end of the tools is quite great and the surgeon has difficulty perceiving the location of tool ends to body part. Having a 3D image to work from would allow the surgeon to view the area of interest from side to side, thereby greatly increasing his awareness of the task.

Figure 1 3D Laparoscope and HMD

To increase accuracy farther, a Head Mounted Display (HMD) can be worn with video see-through capabilities to augment this 3D information in a graphics computer. This augmented reality system then utilizes this 3D imaging information to the fullest and can speed up the surgical procedure times. Reduced surgical times, reduced trauma, has a strong relation to how quickly the patient recovers from a surgical procedure. Not only is a quicker procedure more cost effective but also better for the patient overall.

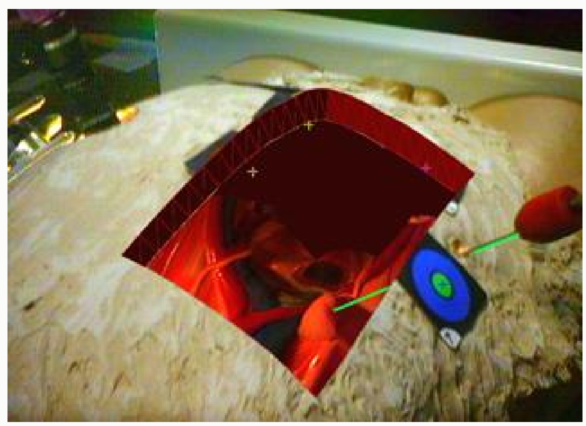

Figure 2 HMD view

Two views of a simulated medical procedure with augmented reality overlays are shown in figure 1 and 2 here. Figure 1 shows a surgeon practicing on our test simulator and the second shows the image that he sees in the HMD.

3. CURRENT SYSTEMS

Most endoscopic or laparoscopic devices used for surgery are simple wide field of view telescopes with a set of relay lenses or fiber bundles that focus this image on an external CCD camera. Surrounding the optics is a bundle of glass or plastic fibers that inject light inside the body from an external light source. These ubiquitous systems are the foundation of most laparoscopic surgeries. They provide high quality 2D color imaging in real time.

However, because there is but one vantage point, the imaging location is not at the surgeon's eye centric position nor is there a relation to up and down (many times a surgeon is surprised to see a cut item float "up" while viewing the procedures on a video monitor), nor is there a sense of depth or measuring capability. All of these deficiencies cause for a severe loss of presence for the surgeon making the procedures much more difficult to perform.

Attempts to reduce some of these problems have been in the form of stereo laparoscopes that try to provide a left and right eye image for the surgeon. These have had limited effectiveness. Mostly due to they only are effective at a very limited viewing distance, the stereo viewpoint is only from one particular vantage point & they don't really solve any of the other above-mentioned problems. However, in some restricted area surgeries, the stereo effect does help in gaining some presence for the surgeon.

The ability to gather real 3D imaging inside the body and to be able to manipulate it for the surgeon to view from their operating vantage point, to provide parallax and real time measurement would greatly increase the capabilities and use of laparoscopic surgery. We will describe how we are able to gather depth and range maps from inside the body and merge it with real color images in real time through a laparoscope. Following is briefly described the more common optical depth extraction methodologies for small surfaces.

3.1 Line Scan

Laser line scan cameras are normally quite accurate and can make high-resolution images in a few seconds. A secondary process or camera is required to superimpose a texture on the resultant 3D wire frame model. Laser scanners have the advantage of very distinct separation lines, not easily confused on curved surfaces and fairly high resolution. Their weak points are that they are slow compared to other depth extraction techniques which can make moving or live objects difficult to scan, fail when scanning perforated objects such as hair, light levels can be distracting to a person or animal being scanned and require a secondary camera or image to paint textures on the model. However, they do fit some live subject niches wells such as foot scanners for custom orthotics along with several very good systems for still subjects, normally using cylindrical scanning orientations.

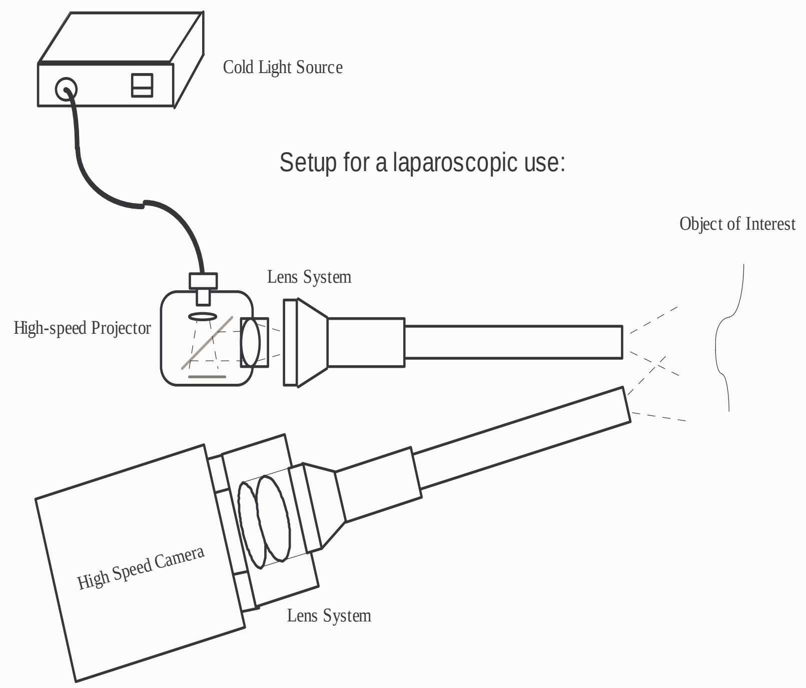

Figure 3 Structured light schematic. A single tube with parallel optical paths is used for medical versions.

3.2 Optical Stereo (Depth from Stereo)

Optical stereo depth extraction utilizes two or more cameras to identify features in a scene common to both cameras and use the known geometry coordinates between the cameras to interpolate the distances of item or features in the scenes. This type of system can be fast, including real time and also has the advantage that the textures are gathered at the same time by the same cameras. These systems are relatively inexpensive for close to real time use.

One large drawback with this type of depth extraction system is the difficult in extracting common features automatically, especially on curved surfaces. Automatic operation in environments where there are many planes and straight lines and other such distinctive textures are where these types of systems are most effective. For medical use, where internals of the body are all curved and surrounded in fluids which causes specularity problems, this type of depth extraction is not very effective.

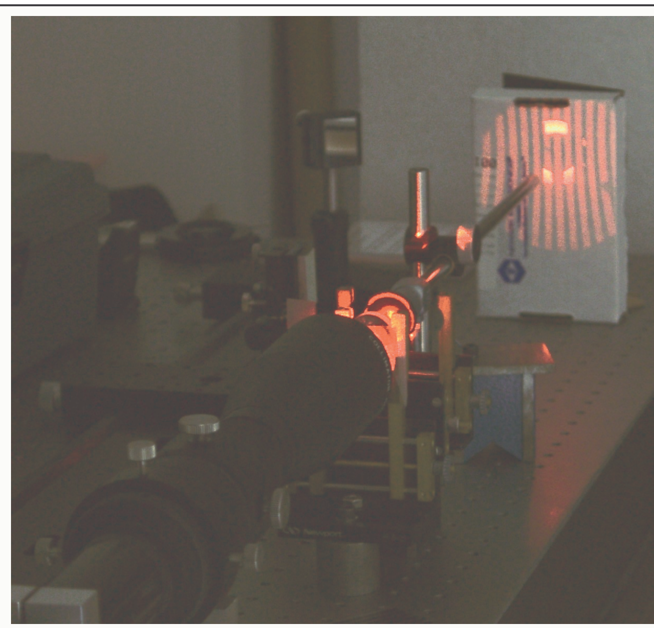

Figure 4 Optical bench with collimated laser projecting stripe pattern through laparoscope.

3.3 Structured Light

Structured light utilizes a projector to display a known pattern or patters, stripes or checkerboard designs are the most common, onto a surface. A synchronized, off-axis camera grabs these images and the coordinates of the edges of the patterns can then be identified and three-dimensional coordinates can be calculated. Progressively finer patterns are projected over this same area and greater depth detail is resulted by the increase in edges displayed. Starting with a course texture help reduce errors from strong discontinuity in features.

The number of patterns to be imaged is:

$$m \text{ stripes} = \log_2 n \text{ patterns}$$

We have hound that 5 patterns provides the best performance of speed and resolution. This type of system normally takes about 5 camera frames to grab an entire image with depth. Texture is inherent in the grabbing phase as a separate image or one inverted pattern so that all areas have been scanned. Structured light benefits medical use because of its ability, like line scan, to accurately read curved surfaces where Optical Stereo cannot. However, it has a greater lag of 5 or so frames where Optical Stereo can be as fast as one to two frames behind. We have a method that reduces this lag toward real-time implementation by continuing to update the depth map from previous frames, not just the cumination of all 5 frames all at one time.

3.3 Laser Light Source

We have utilized standard 140 watt fiber light sources as well as synchronized Xenon strobes for lighting and have found the optical efficiency to be poor for the small image generators. Utilizing a laser has some advantages. The beam is easily collimated simplifying the optics required for imaging to small image generators and displaying this image through the tiny apertures of laparoscopes. The monochromatic light also assists us in selecting the most appropriate color for feature recognition; for inner abdominal work, green light tends to penetrate tissues the least creating optimal edge contrast.

The texture portion utilizes a full, 3-color camera while the depth extraction portion only requires a single, monochromatic one. However, the full color camera does require a standard fiber light source to illuminate the fiber bundle in the laparoscope itself – just like a regular laparoscope. These separate cameras also allows for easy switching between 3D augmented depth extraction laparoscopy and standard 2D laparoscopy without the need for switching scopes or optical heads.

4. SWITCHED PATTERN DEPTH EXTRACTION

4.2 Structured light

Structured light is a well-known technique in computer vision to acquire depth information[1,4,6,8]. These techniques use a projector to project a known light pattern onto a scene. The scene is then viewed from a camera at a slightly different position than the projector. The distance of objects in the scene from the camera can be calculated if the position of both camera and projector are known and if the camera can recover features projected onto the scene. Distance measurements may be

Figure 5 Binary striped pattern. The inverse of the pattern displayed for the first image is displayed for the second. This ensures that all areas are illuminated.

precalculated for particular camera/projector geometry reducing the computational costs of locating features in camera images and table lookup of range values.

Different patterns can be used which are better for some types of features or extraction methods. These patterns can range from simple line, 50% on and off lines, checkerboards, concentric circles and others. We settled on the use of 50% stripes for this worked well with a single camera viewing and for inverting the lines from on to off providing for full scene illumination.

This system uses binary encoded stripe patterns. Binary encoded stripes allow use of a small number of images to derive comparable information to using many more single stripes. Simple thresholding was used in early prototypes to locate stripes in the camera images. We start with a low number of stripes that helps avoid confusion for the software and then fill the image with a higher number up to 128 lines to provide detail and measurements of closely spaced features

4.3 Animal cadaver experiments

Previous implementations of structured light fail when presented with a highly specular environment. The abdominal cavity is typically a highly specular environment due to the fluids that coat internal organs as well as the optical properties of the cells that line them. Structured light techniques also typically fail in this environment because of the difficulty associated with automatically recovering projected features.

To test whether we could recover features from images projected onto internal organs, we tested our device on a porcine cadaver. The experimental setup is shown in Figure 6.

Figure 6 Early Animal cadaver tests

An image was captured from a pattern (a series of squares) projected directly onto the porcine organs. A coating of talcum powder was then blown onto the organ to reduce the amount of specular reflection and then a second image was captured without moving the laparoscope. This approach allowed us to compare the automatically generated position of projected features on the raw organs as well as from images that we could be confident in the acquisition. Only small differences in the position of the projected features were found between the two sets of images.

5. REAL-TIME CUSTOM PROJECTOR SYSTEM

With what we learned from the structured light, digital projector depth extraction system we created the requirements for the real time system. Additionally, hardware became available that would, theoretically, let a real time system be possible.

We determined that to have an update rate of 15 frames per second with 5 x 2 (5 frames with an inverted field) and at our desired resolution of approximately 500 x 520 pixels, we would require following bandwidth just for the raw images:

15 (frames/sec) x 5 frames x 2 (fields/frame) x 500 x 520 (pixels^2) x 8 bits/pixel x 1/8 bytes/bit = 40Mbytes/second.

This is considered the minimum amount of information bandwidth. Nor the use of subsampling time where a long exposure, full frame exposure time, is used in conjunction to duplicate short frame time exposure to extend bit depth and reduce spectral reflection difficulties. With this option, necessary for medical use where the object of interest is highly specular, the bandwidth then increases to a around 80Mbytes/second. A preprocessor board would be needed.

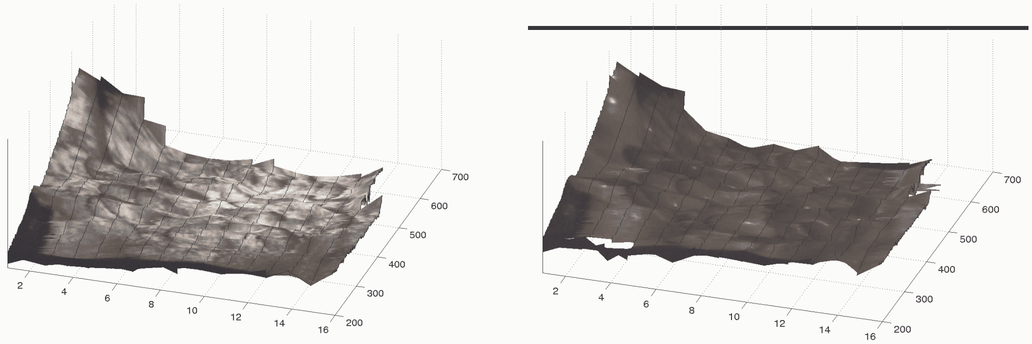

Figure 7 Depth extraction from 3D Laparoscope showing contours from inside a porcine body.

A separate color camera and white light fiber projection can be used to gather the color image. These can share the same optical path in the laparoscope. This also allows for standard laparoscope use when desired.

5.1 Hardware Details of High Speed Structured Light Depth Extraction System

In the current research system, we have the host computer generate the structured light image; in our case, vary width and quantities of stripes or squares. They are displayed through an ordinary VGA card at 640 x 480 resolution. Each of the three colors is considered an individual displayable frame and the images are pre-generated that way. The graphics card output goes to the custom, miniature projector. This projector grabs each color and displays a single color channel at a time in a positive black and white image and then immediately following as in inverse of that image. The inverse is to keep the LCD liquid from pooling or migrating in the display. After the image and inverse image of the red channel is displayed, then the green channel is displayed, in positive and inverse image like the red. This is repeated for the blue channel. Overall, for a single color frame sent to the miniature ferro-reflective projector, 6 output frames are produced.

Synchronized to the VGA output to the projector is the high-speed camera. We were using a high speed Dalsa CA-D6-0512 camera synchronized to the projector at 180 Hz. However, we have switched to a 60Hz Pulnix camera. Depth extraction algorithms are run on a Matrox Genesis board. This card then performs the same type of structured light depth extraction techniques that the earlier flex scope SGI based structured light system did. The depth extraction routines are now performed in real time, sending out 10 frames/second of 3D depth maps with textures from 60 synchronized image and projection fields/second.

5.2.1 High Speed Projector:

A Displaytech ferro-reflective display engine was used as the core of the projector. Because we needed much higher light levels than LED could give, but yet wanted to maintain a very small handheld package, we used a remote flex light. The particular light picked has a very efficient IR blocker which projects visible light only through the fiber onto the LCD display reducing the damage from IR heat that a normal light would impose.

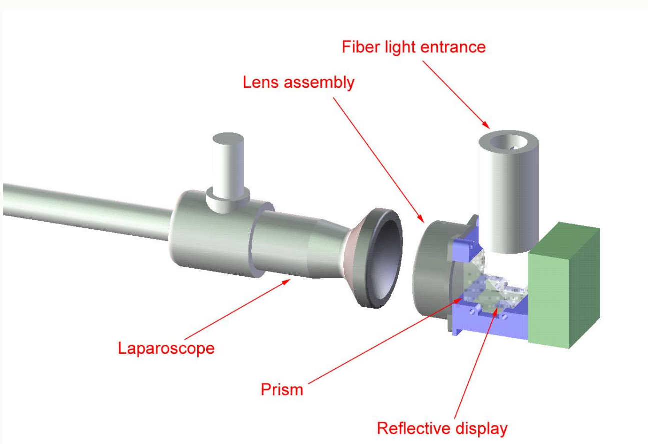

Figure 8 Laparoscope projector

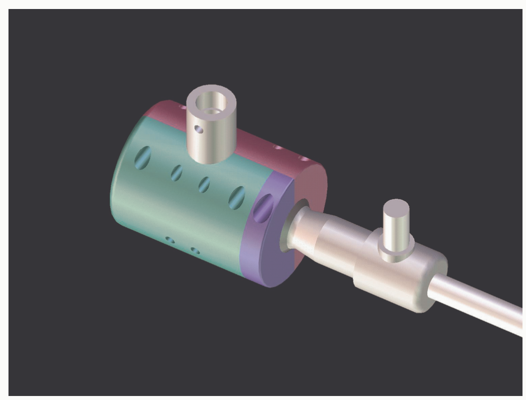

Figure 9a Projector Assembled

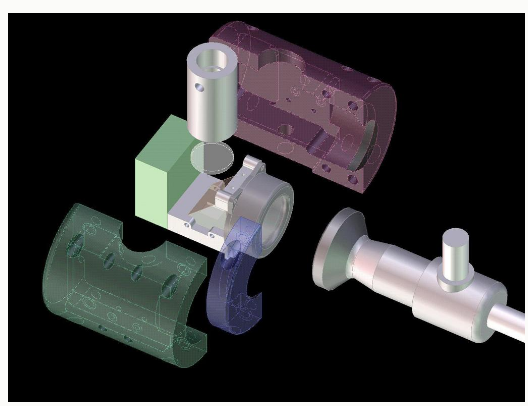

Figure 9b Exploded view of Projector

Collimation lenses were used at the end of the fiber bundle to concentrate the light evenly onto the display. New lenses were also installed on the exit side of the imager to allow the image to be projected through a standard endoscope eye aperture. They were focused to project about the same distance that a laparoscope is set for. A special quick-change mount was also integrated to allow us to switch out different endoscopes and laparoscopes quickly. Overall the projector is a scant 50mm in diameter by 67mm long. The images to the right show an assembled projector attached to a standard laparoscope extending from the right. The bottom image is an exploded view showing the fiber light entrance (top) with collimator lens; the LCD display engine and cube (center); and the exit lenses and mount (toward laparoscope).

There is now a manufacture that is producing a high-speed ferro-reflective display that has a reduced crystal migration effect. By reducing the number of inverse images one has to make, required because of the crystal migration, one can speed up the overall update rate of the system.

5.2.2 Cameras

We used two different cameras to gather high-speed images from this system. The Dalsa DA-0512 and Matrox Genesis LVDS system theoretically were a matched pair; however, our models of both components were from the first month of introduction and they both were, unfortunately, very unstable together. Because of this, we switched back to a Pulnix progressive scan camera grabbing 60 frames per second. Other high speed cameras can be used.

5.2.3 Processor board

The Matrox Genesis board with LVDS option was the preferred graphics card for the real-time nature of these experiments. The onboard processor was able to convert the image sets as fast as we expected it to. However, as mentioned above, interfacing to the matched camera via the high-speed LVDS was not easy and this particular arrangement is not recommended. A more robust and high-speed digital linkage is needed between camera and image grabber/processor board to make this system more than a laboratory system. Because of the Matrox card and Dalsa flakiness together, we could not use this system on subjects.

On the Matrox board, we grab six frames in two separate sets. From each of these sets, three sets of depth information is extracted. Each set having finer and finer resolutions off axis, same as the first system running on the Silicon Graphics workstation. These sets are then sent out from the Genesis card to the PC processor where a texture map with lookup tables of depth and light intensity are updated. This data is then grabbed as a 3D skin and the corresponding texture is overlaid. The display of this information is controlled from an external input device, normally a tracker on a HMD or surgical headband and then displayed to either a monitor or HMD for the surgeon to view.

5.3 Future additions

There are additions that we were not able to include during these initial experiments that we feel would greatly improve the usefulness of the system:

Because the inside of the body is wet, it is highly specular. This causes problems for projection devices blinding the camera from various places on the surfaces. To reduce this blinding problem on the imager, increased digital depth from 8 bits to 12 or more are needed. Currently, there are not any high resolution, high-speed cameras with greater than 8 bits output of sensitivity. For now, we propose a secondary scanning of the image with the camera's shutter open for 1/8 the normal duration. Combining this data to the full intensity data will result in 12 bits of light intensity information per pixel.

Another method to reduce the specularity problem is by various polarizing schemes between both the camera and the projector. Since the reflective LCD is already polarized, unwanted reflections by adding a polarizer on the camera side can be reduced for some angles.

We are also replacing the strobe light with an onboard laser based projector. This easily collimated light source has shown to be very bright, controllable and easy to image through the small aperture of a standard laparoscope. Figure 10 shows an abbreviated schematic of what is in the DMD/laser based depth extraction engine.

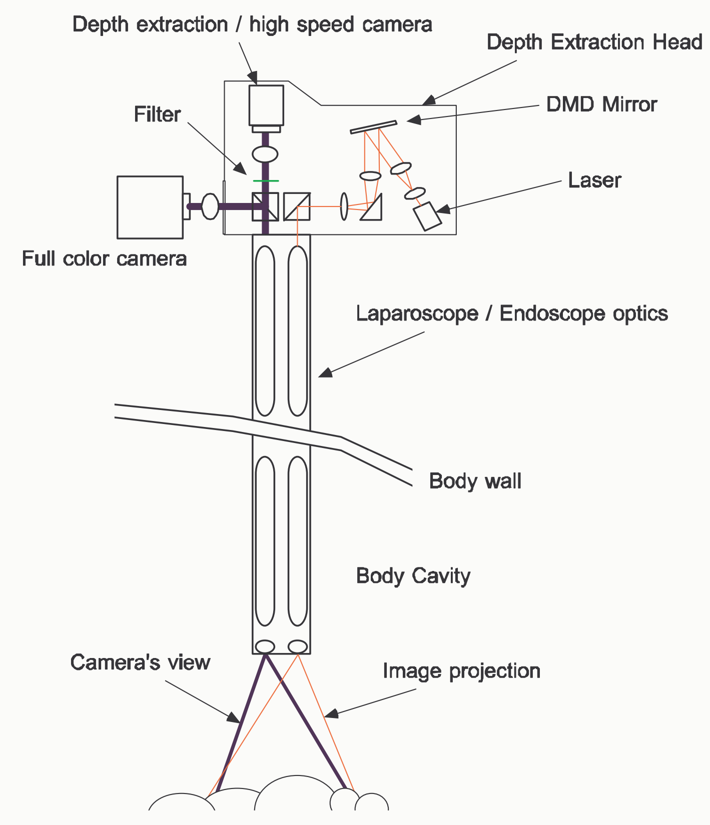

Figure 10 Laser based 3D depth extraction

6. SUMMARY

Taking a standard depth extraction method like structured light into a unique, highly specular environment such as inside the human body and then shrinking the size of the equipment and use requirements into that required for a surgical environment, puts many strains on the initial system. The best approach to solving these problems is by dividing up the various challenges and prototyping the different sub components separately. Combining them all at once limits ones ability to easily change a component without affecting the other components.

With this design philosophy, we have been able to combine several sub systems into a small, usable and effective depth extraction laparoscope which should be suitable for challenges found in the surgical environment.

7. ACKNOWLEDGMENTS

We want to thank the following for all the help and knowledge that they lent us for this research: Michael Rosenthal, and Andrei State.

8. REFERENCES

1. R.C. Dailey, L. G. Hasselbrook, S. Tungate, J. M. Reisig, T. A. Reed, B. K. Williams, J. S. Daughterys, and M. Bond, "Topolgraphical Analysis with Time Modulated Structured Light", SPIE Proceedings 2488, 1995, pp. 396-407.

2. Andreas Fountoulakis, A. Loman, S. Chadwick, "Initial Experience of 3D Video Endoscopy in General Surgery", *Surgical Technology International*, pp. 115- 119.

3. Henry Fuchs, Mark A. Livingston, Ramesh Raskar, D'nardo Colucci, Kurtis Keller, Andrei State, Jessica R. Crawford, Paul Rademacher, Samuel H. Drake, and Anthony A. Meyer, "Augmented Reality Visualization for Laparoscopic Surgery", MICCAI (Proceedings of First International Conference on Medical Image Computing and Computer-Assisted Intervention) 1998.

4. P. Levoie, D. Ionescue, & E.M. Petriu, " 3-D Object Model Recovery from 2-D Images Using Structured Light", IEEE Instrument Measurement Technology Conference, pp. 377-382.

5. J. Marco, Mark A. Livingston, and Andrei State. "Managing Latency in Complex Augmented Reality Systems." Proceedings on Interactive 3D Graphics Series, ACM SIGGRAPH, pp. 49-54, 1997.

6. R Raskar, G Welch, W Chen, "Tabletop Spatially Augmented Reality: Bringing Physical Models to Life using Projected Imagery", Second Int Workshop on Augmented Reality (IWAR'99), October 1999, San Francisco, CA.

7. Andrei State, Gentaro Hirota, David T. Chen, William F. Garrett, and Mark A. Livingston, "Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking," Proceedings of SIGGRAPH '96, Computer Graphics, Annual Conference Series, 1996, pp. 429-438.

8. G.Turk and M. Levoy, "Zippered Polygon Meshes from Range Images", Proceedings ACM SIGGRAPH 1994, pp. 311-318.

9. Ackerman, Jeremy, Kurtis Keller, "Real-time Antatomical 3D Depth Extraction For Laparoscopic Surgery", Medicine Meets Virtual Reality, 2001.