Surface reconstruction of abdominal organs using laparoscopic structured light for augmented reality

Jeremy D. Ackerman [a], Kurtis Keller [a], and Henry Fuchs [a]
ackerman@cs.unc.edu, keller@cs.unc.edu

[a] Department of Computer Science, University of North Carolina, at Chapel Hill, Campus Box #3175, Chapel Hill, NC 27599

ABSTRACT

Creation of accurate surface models of abdominal organs is essential for many developing technologies in medicine and surgery. One application we are working towards is augmented reality (AR) visualization for laparoscopic surgery. Our current system meets some, but not all, of the requirements. We use two custom built laparoscopes, a custom built miniature projector, a standard camera, and a standard video capture and processing card to implement a laparoscopic structured light range acquisition system. We will briefly show the custom hardware but will emphasize the structured light depth extraction techniques used for the unique properties of surfaces inside the body, particularly dealing with specular reflections. In early experiments, we studied the effectiveness of our algorithm in highly specular environments by creating range images acquired from fresh animal organs. These experiments used a large projector, open abdomens, and offline image processing. We report the results of experiments using our miniature projector, and on line processing.

Keywords: structured light, depth extraction, laparoscopy

1. LAPAROSCOPIC SURFACE RECONSTRUCTION AND AUGMENTED REALITY APPLICATIONS

Laparoscopy is a surgical procedure in the abdomen performed by inserting a small telescope and elongated instruments into a created space through small incisions. A wide variety of surgical procedures can be performed laparoscopically including a cholecystectomy (removing the gall bladder), appendectomy (removing the appendix), and Nissen fundoplication (a treatment for severe reflux). Laparoscopic surgery is widely regarded as better for the patient than conventional open surgery because there is usually less post-operative pain, a shorter hospital stay, and faster recovery. Despite these benefits laparoscopic surgery is problematic. Several studies have shown that laparoscopic procedures have an increased complication rate as compared to the open version of the procedure. Also, laparoscopic procedures generally take more operating room time than do open procedures [8]. Finally the skills needed to perform basic tasks laparoscopically and the knowledge needed to understand and complete a procedure laparoscopically are considered more difficult to learn than what is needed for open surgery.

Augmented reality (AR) is a technology that allows a user's view of the real world to be combined with computer generated imagery. For medical applications this can enable a physician to see acquired imagery such as CT, MRI, or ultrasound overlaid on a patient that they are examining or operating on. An augmented reality system for assisting ultrasound guided needle breast biopsy is now in clinical trials. In order to make such a system work, the computer generated imagery must be accurately registered to the patient. In the ultrasound augmented reality system, we track the position of the calibrated ultrasound probe (and therefore the imaged tissue) relative to the physician so that the image of the tissue can be rendered accurately in space. Although the ultrasound image is two dimensional, it represents a two dimensional slice, in three-dimensional space, through the tissues being imaged.

The application of augmented reality should make laparoscopic surgery look more like the open procedure. We believe that this type of visualization may alleviate some of the disadvantages of laparoscopic surgery. A system of this type should also simplify the use of intra-operative imaging by providing an intuitive way to incorporate visual information in the context of the surgical procedure.

In order to implement an AR system to assist laparoscopic surgery a source of three-dimensional data from inside the patient is needed. Conventional laparoscopes provide two-dimensional imagery from inside the body but no information about the shape, position, or scale of the organs being viewed. These two-dimensional images do not provide enough information to accurately render the inside of the body from a new perspective.

Pre-operatively acquired and segmented three-dimensional imagery cannot be used to get the need three-dimensional information because structures in the abdomen shift significantly during the process of gaining laparoscopic access. Pre-operatively acquired imagery could be useful if added to an augmented view as a reference to the operative plan or anatomical anomalies. Any three-dimensional information gathered must be frequently updated as the procedure progresses to ensure the accuracy of the computer-generated imagery and to provide the surgeon an accurate perception of the operative site.

Since a laparoscopic camera is already present inside the abdomen, a computer vision approach to gaining three-dimensional information would seem to be a natural fit. Indeed stereo laparoscopes (a single laparoscope with two imaging paths) have been created. However, a number of conditions are present that makes passive (non-emissive) computer vision techniques difficult. The first problem is that most of the objects (organs and tissues) inside the abdomen are curved and there are few well-defined edges. The surface of tissue and organs presents two more problems. While organs do have some surface texture, it is often primarily low frequency which makes pattern matching more difficult. Tissue and organ surfaces tend to be wet and shiny which creates intense specular reflections. Specular reflections occur in locations in the image where the surface being imaged acts like a mirror, directing a large portion of the incident light directly back at the camera instead of only directing diffuse, scattered light back at the camera. Specular highlights, in comparison with the innate texture of the organ surfaces, have sharp, well-defined edges in a camera image. These specular reflections obscure the innate textures and are likely to be recognized by pattern matching algorithms as surface texture. The specular reflections appear on surfaces depending on the normal of the surface, the position of the camera, and the position of the light source, so algorithms matching specular reflections in depth from stereo methods will generate erroneous matches and therefore erroneous depth information.

The system requirements for a depth-extracting laparoscope, with consideration for our target application, have been proposed:

1. Real-time operation (range images generated at around 30Hz)
2. Operating room compatible size, sterilizability, and limited heat production
3. Ability to convert back to a conventional laparoscope if desired or in the event of device failure
4. Accuracy despite specularity and curved surfaces.

Structured light was identified as a potential method to use because it works well on curved surfaces with little or no discernable texture. The remaining challenges were to increase rate of operation and to make a device operating room compatible.

2. MINIATURIZED HIGH SPEED STRUCTURED LIGHT

We constructed a prototype depth extracting laparoscope that attempts to achieve the system requirements for AR assisted laparoscopic surgery. A photograph of the device without its driving electronics and computer system is shown in Figure 1 This device is composed of two custom designed laparoscopes, a custom built miniaturized projector, and a video camera. Customized laparoscopes were built to improve light output by the projector and light capture by the camera.

A schematic system diagram of the system is shown in Figure 2. Our initial prototype used a constant illumination source (feeding light to the projector via an unstructured fiber optic transmission cable), and synchronization of the camera shutter with the projector. We subsequently switched to using a strobe light as the light source. The strobe light is triggered with the projector using the synch in the video signal and a variable delay circuit. The strobe light is electronically isolated from the other devices using an optical isolator to prevent accidentally damaging the other electronic devices. A Matrox Genesis board (purchased in 1998) is used to control the projector, capture the camera output, and process the images.

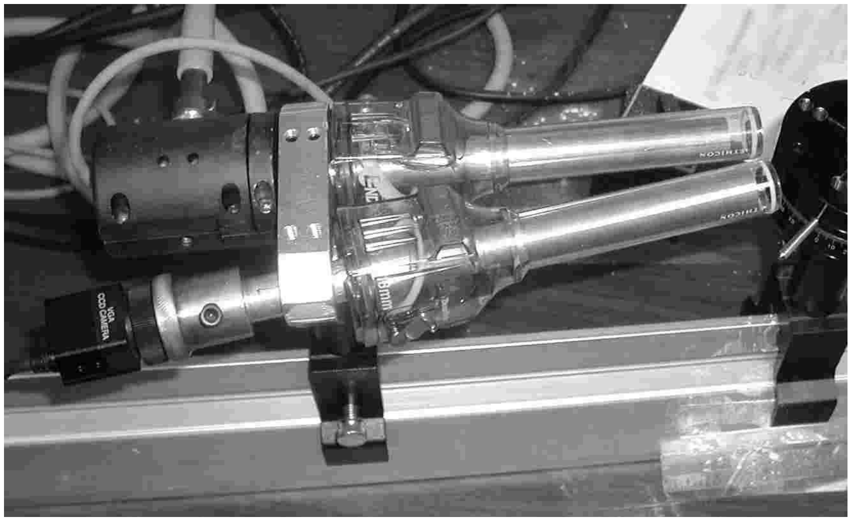

Figure 1: our structured light depth extracting laparoscope prototype shown with surgical trocars (laparoscopic access ports, Ethicon Endopath® 18mm) over the laparoscope tubes. The miniature projector is attached to the upper laparoscope, and a camera is attached to the lower laparoscope

Our miniaturized projector is built around a 640x480 pixel ferro-reflective LCD. Input is a 60Hz video signal. The display engine was originally designed to operate at 30hz in frame-sequential color (using synchronized colored LEDs). The display shows the inverse of the displayed image for each color to prevent the liquid crystal from locking. The device is therefore capable of operating at 180hz, provided that displaying an image and its inverse is useful [5].

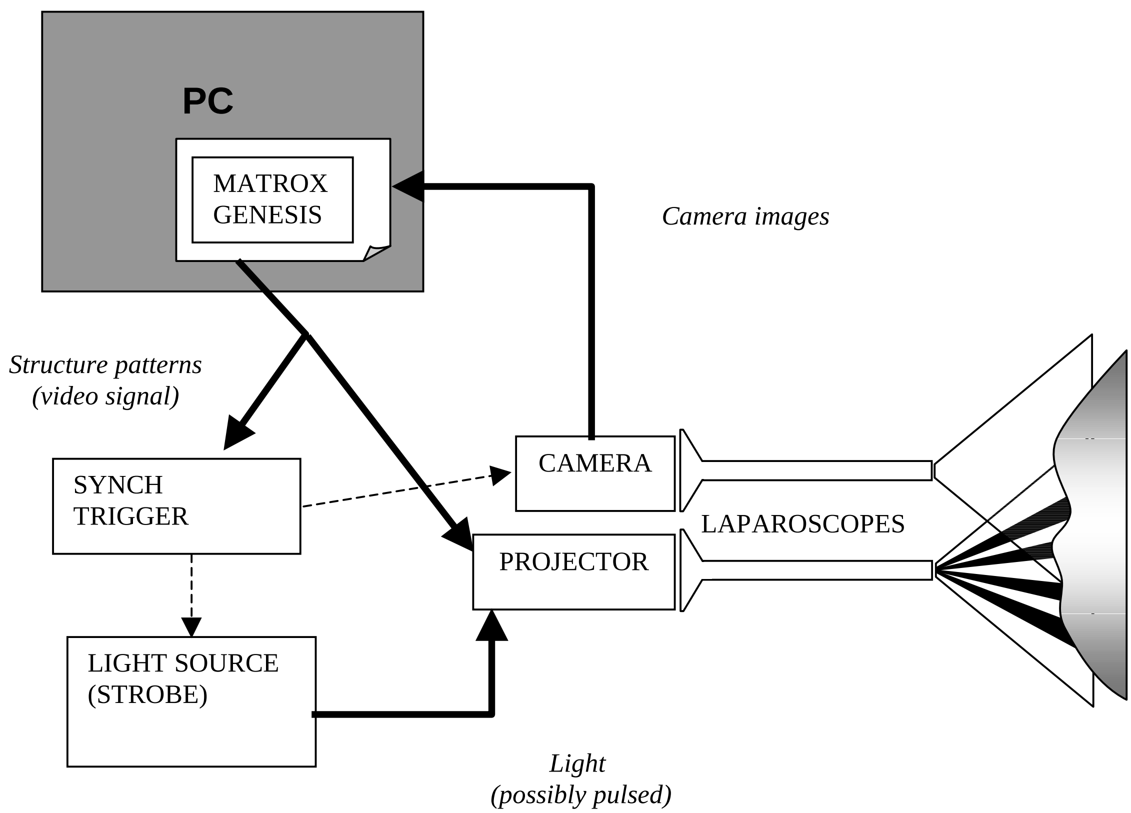

Figure 2: A schematic overview of our system for laparoscopic depth extraction. The Matrox Genesis board in the PC outputs the structured light patterns as a video signal to the miniature projector. This video signal can also be used to trigger synchronization signals for the projector's light source and/or the camera. The images acquired by the camera are captured and processed by the Matrox Genesis board.

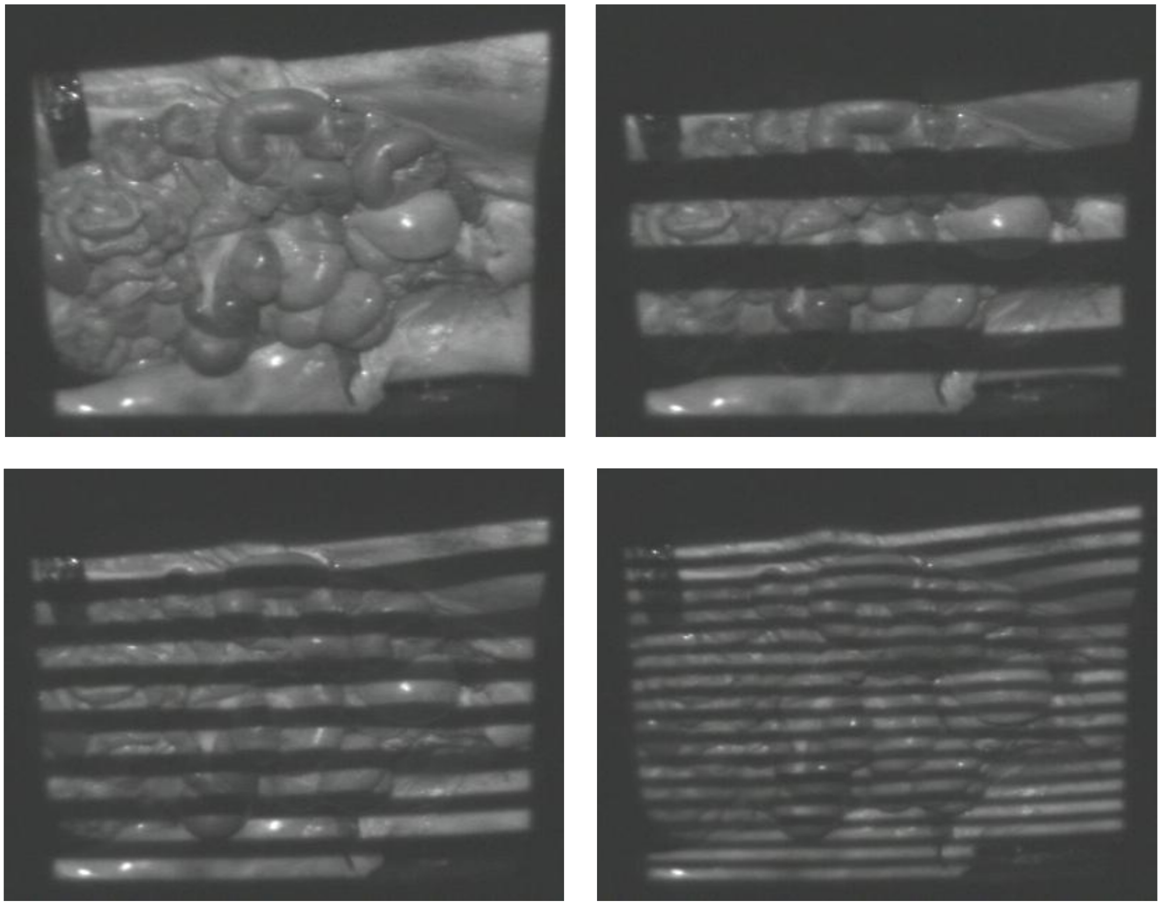

Figure 3: Sample images acquired from the abdomen of a porcine cadaver using a large projector to project stripe patterns onto the organs. The image in the upper left shows the abdomen fully illuminated. High intensity spots on the surface are the result of specular highlights.

We originally attempted to use a high-speed camera (up to 220Hz) in the system, but hardware incompatibility kept us from proceeding with this camera. The working prototype uses a Toshiba IK-7XD camera operating at 30hz.

3. IMAGE PROCESSING

Structured light is a widely known method for obtaining three-dimensional information. In this paper we use bit-encoded stripe patterns. We use a simple counting method in which each projected pattern is half "on" and half "off." The n^{th} pattern (starting with pattern $n=0$) is composed of 2^{n+1} stripes alternating between "on" and "off". Figure 3 shows several stripe patterns acquired using a large projector in a porcine abdomen.

3.1. Image processing methods

For a camera capturing a series of images of a scene with a projector projecting a sequence of bit-encoded stripes and their inverses. We define the following:

$P(n)$ is the image containing the n^{th} projected bit pattern and the image
$N(n)$ is the image containing the projected inverse of the bit n^{th} pattern.
m is defined as the number of encoded bit patterns projected
b is the number of bits used to represent all calculated and captured images and $b \geq (m+1)$
$C(n)$ is a calculated image in which $P(n)$ and $N(n)$ are combined
$L(n)$ is a calculated image in which all pixels belonging to the n^{th} stripe are labeled

The combined image *C(n)* is then *P(n)-N(n)*. Practical implementation requires these images be represented by a finite number of bits and that further computation be completed using only that number of bits. Therefore, assuming representation of the images as *b*-bit unsigned integers, the following calculation for *C(n)* can be used to prevent overflow and underflow on a per pixel basis:

$$C(n) = \frac{P(n)}{2} + 2^{b-1} - \frac{N(n)}{2} \qquad (1$$

The labeled image *L(n)* can then be created in which each pixel is labeled as being part of an "on" stripe, "off" stripe, or non-identifiable. A threshold level, *r*, is used in this calculation. Increasing *r* decreases the likelihood of incorrectly labeling a pixel as belonging to a stripe, but increases the chance of labeling a pixel as non-identifiable when it actually belongs to a stripe. This calculation is again performed on a per pixel basis:

$$L(n) = \begin{cases} 0 & if \quad C(n) \le 2^{b-1} - r \\ 2^{n-1} & if \quad C(n) \ge 2^{b-1} + r \\ 2^{m} & if \quad 2^{b-1} - r < C(n) < 2^{b-1} + r \end{cases} \qquad (2$$

The labeled stripe images can conveniently be combined using a logical OR operation of the bit-masks of all the *L(n)* into a single *m+1* bit image we define as *S*. The image *S* has the characteristic that all pixels containing values less than 2^m are labeled with the number of the stripe encoded by the bit-encoding illumination patterns. All pixels with values greater than or equal to 2^m have been coded in one or more sets of image pairs ***(****P(n),N(n)****)*** as unidentifiable. The implications of a pixel not being identifiable are discussed more in the following subsection.

$$S = \bigcup_{i=1}^{n} L(i) \qquad (3$$

Several methods can then be applied to convert the stripe-encoded image *S* into range data. For the purpose of rendering a texture-mapped range image, an image of the object with little or no evidence of the projected stripe patterns can be constructed by simply adding a *P* and *N* pair together and taking appropriate precautions to prevent overflow.

The effect of this method in several scenarios is demonstrated for Table 1. The values used in this table are those representative of what might be seen in images under several different conditions but are not values we actually found in images. A "normal" pixel is one with average light reflecting properties (not highly specular and of a moderately light color). A pixel with high albido is one acquired from a surface that reflects (diffusely) light very efficiently and is usually a very bright color or white. A chalk or talc-coated surface is often used as a good example of a high albido surface. High albido surfaces often show the effects of ambient lighting quite clearly. A low albido pixel acquired from a surface is one that diffusely reflects light poorly. An example of such a surface in the abdomen is the surface of the liver which is a rich dark reddish brown, but is still capable of producing intense specular reflections.

Pixel type	*P*	*N*	*C*	*L*	Result
Normal "Off"	50	200	53	*0*	"Off"
Normal "On"	200	50	203	2^n	"On"
Shadowed (non-illuminated) pixel	50	56	125	*128*	Unidentified
Specular pixel	255	253	130	*128*	Unidentified
High albido, bright ambient "on"	250	200	153	2^n	"On"
Low albido "on"	40	20	138	2^n	"On"

Table 1: The effect of this algorithm on several different types of pixels on the resulting pixel in the label image *L*. In this example calculations are done with the number of bits per pixel *b=8*, the number of stripe patterns *m=7*, and a threshold level *r=8*.

3.2. Implications of image processing methods

Our image processing method has important ramifications on the speed of execution, the handling of motion, and the handling of specular reflections and shadows.

This algorithm is of linear order complexity with respect to the total number of pixels. The algorithm is performed completely on a per-pixel basis, possibly continuing through conversion from stripe encoding through conversion to depth information. While this means fast execution on a single processor, performance should improve linearly with an increase in the number of processors. This is ideal for image processing hardware which utilizes a large number of parallel processors. This algorithm can be pipelined in hardware because only the results from the previous step are needed for the next step.

Table 1 demonstrates the result of this method for several scenarios. In some instances specular reflections can visibly retain the structure of incident light as in the case of a mirror. In laparoscopy, however, a structured specular reflection of the light sources is unlikely due to the small aperture of the projector, columniation of light emanating from the projector, blurring effects caused by reflecting from multiple surfaces (e.g. cellular layers) within a tissue surface, and regular changes in curvature of organ surfaces. For these reasons, a specular reflection or highlight will usually appear as small region with brightly lit, often fully saturated pixels, even if the clustering of pixels is not currently being illuminated by a stripe. Indeed, even if an all black image is projected into the scene, specular reflections often still appear from background illumination coming from the projector. By acquiring a scene with a stripe pattern and its inverse, these regions can easily be identified.

Shadows provide a challenge similar to specular highlights. These are regions of the object that are visible to the camera but are not illuminated by the projector. These areas are sometimes, but not always dark in the camera image. Although these areas may be indirectly illuminated, simply identifying dark, unlit areas may be insufficient since each stripe pattern puts the same amount of light energy in the scene, these areas are usually uniformly illuminated by both patterns and their inverses.

Motion is problematic to most structured light methods. This is because small amounts of motion between each projected stripe pattern can cause pixels to be misidentified as belonging to a particular stripe. These errors most frequently occur on the edge of stripes because the object or camera/projector pair has moved slightly between acquired images. If a movement of this type occurs between acquisition of the *P* image and the corresponding *N* image, one of three possibilities occurs; an "on" pixel in *P* is illuminated in the *N* image, an "off" pixel in *P* is not illuminated in the *N* image, or a pixels illumination is not erroneously changed by the movement. In this method motion effects are labeled in the ***(P,N)*** pair as "unidentified" points. While this method will not consistently identify large motions, nor it provide a mechanism (directly) for creating range estimates as accurately as without motion, it does provide a mechanism for avoiding gross misclassification.

4. RESULTS

Previous experiments using a large projector to project structured stripe patterns on the abdominal organs of a porcine cadaver demonstrated that this image processing method works well. In these experiments we compared stripe acquisition and computed range images on raw tissue and as well as on tissue coated with a fine powder (talc) to reduce specular reflections, enhance diffuse reflection of the projected patterns, and make the albido of all tissue surfaces comparable. At high resolutions (large number of stripes) both our method labeled a large number of pixels in both the coated and uncoated abdomen as being unidentifiable. Visual inspection of the images containing these high frequency stripes confirmed that stripes were not detectable due to aliasing in the images. For lower frequency stripe patterns, our method successfully found stripe patterns more effectively than the more standard method of thresholding. Examples of calculated range images are shown in Figure 4.

The algorithm was implemented on a PC with dual Pentium II 400Mhz processors and a Matrox Genesis capture and processing board in Visual C++ using the Matrox Imaging Library high-level calls for image processing functions done on the Genesis board. A median filter (3x3 kernel) was applied to camera images before processing in an effort to reduce camera noise. Average times for various parts of the algorithm are shown in Table 2. The table demonstrates that the time associated with synchronization and camera capture account for a large portion of the execution time of our working system. Conversion from labeled image to range image was not performed, but a look-up table type of conversion would add a minimal additional computational cost.

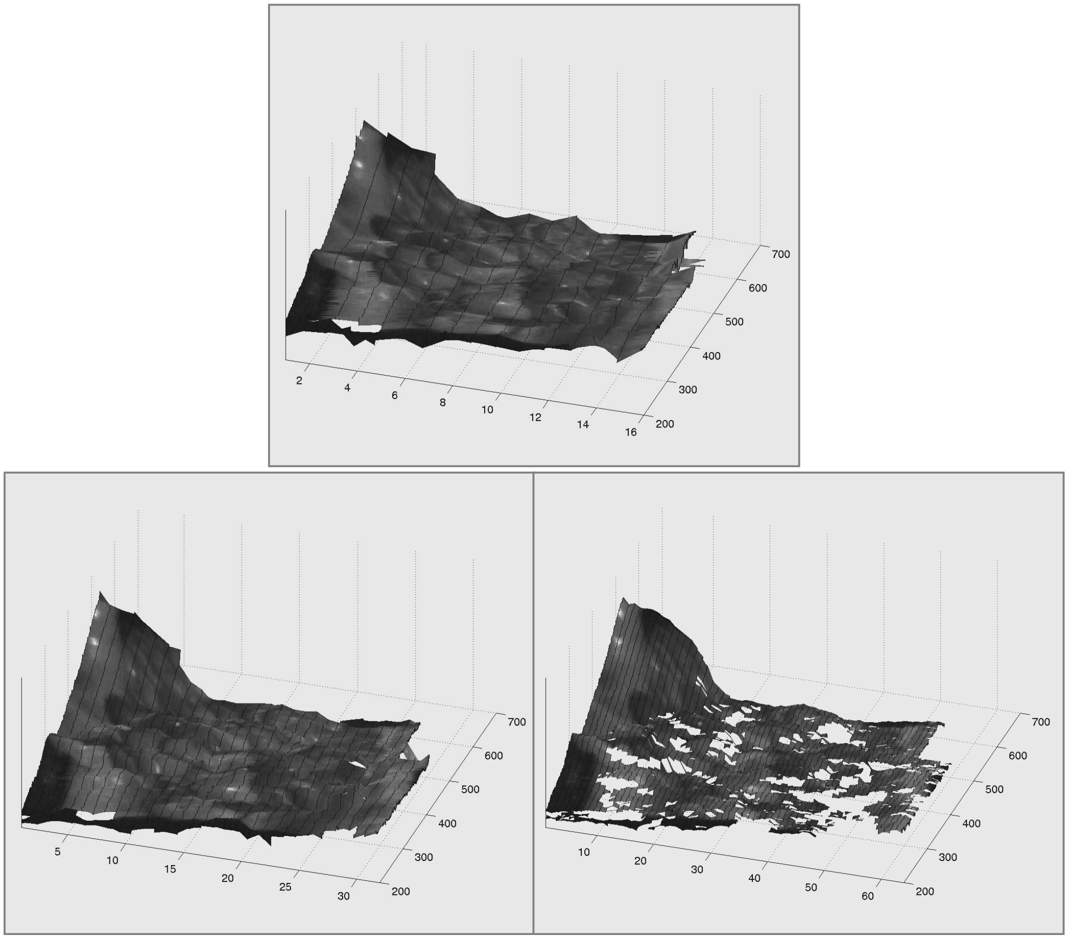

Figure 4: These three images show range images calculated from images with stripes projected onto the abdomen of a porcine cadaver. The stripes were identified using our method. These images show a progressively increasing number of stripes (top to bottom, left to right). In the image calculated with the largest number of stripes, large holes appear because aliasing of the stripes prevented their identification.

Light output was a significant problem in our prototype. Multiple revisions were undertaken to improve the light output of the projector and the light gathering capability of the camera. In our most recent revision, the structure patterns appear only marginally above the background noise level in the camera when projected on a piece of white paper. Our method works reasonably well on the pair-wise images (~80%), but when L images are combined to create the S image, the successful identification of a pixel with a stripe decreases significantly (only ~30% for $n=4$).

	Convolution	*N=7*	*n=6*	*n=5*	*n=4*	*n=3*	*n=2*
Camera capture on	ON	100ms	87ms	73ms	60ms	47ms	33ms
	OFF	88ms	76ms	64ms	53ms	41ms	30ms
Camera capture off	ON	2203ms	1901ms	1602ms	1302ms	967ms	667ms
	OFF	2201ms	1901ms	1568ms	1268ms	966ms	667ms

Table 2: This table shows the actual execution time of these methods on a dual Pentium II 400Mhz PC equipped with a Matrox Genesis card. The second column describes if convolution with a 3x3 median filter was used to reduce the effects of noise in the images. These tests were conducted while capturing images and processing previously acquired images. Capturing images for processing requires leaving ample time for synchronization of the entire system which takes an average of 5 camera cycles (at 30Hz, 150ms) per image captured. The images processed were 640x480 8-bit grayscale images.

5. CONCLUSIONS

Preliminary data suggests that structured light may be a viable computer vision technique for reconstructing the surface of organs laparoscopically. However, our miniaturized high-speed structured light projector still falls short of what is needed for a medical augmented reality application.

Success will require improvements in two areas. These are light output and sensitivity, and speed.

The light output of the projector needs to be increased dramatically. While some gains could be made by using a more sensitive camara (1 lux minimum light levels rather than 3 lux currently in use), this would likely only provide marginal improvement. In one revision, we made an adjustment in the light gathering area of the light cable resulting in an approximately 30% increase in light output and an increase from 70% to 80% identification of pixels ***(P,N)*** pairs. Our current light output is not sufficient to visually recognize texture on the surface of an object in a captured image.

Our image processing methods appear to be fast enough for higher speed implementation. Approximately 10Hz output rates could be sustained with our current system. Newer processors, including on those found on the most current version of the Genesis board and newer PCs would further improve our image processing performance. Our greatest barrier to higher speed performance is capture from, and synchronization with, a high-speed camera to a high-speed projector. We plan to build a new miniature projector and laparoscope that has greater light output and will be easier to synchronize with other devices.

6. REFERENCES

1. J Ackerman, K Keller, and H Fuchs, “Real-time anatomical 3D image extraction for laparoscopic surgery”, *Medicine Meets Virtual Reality 2001*, 2001, pp. 18-22.
2. Bengt 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.
3. H Fuchs, M Livingston, R Raskar, D Colucci, K Keller, A State, J Crawford, P Rademacher, S Drake and A Meyer, “Augmented Reality Visualization for Laparoscopic Surgery”, *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention*, 1998.
4. Andreas Fountoulakis, A. Loman, S. Chadwick, “Initial Experience of 3D Video Endoscopy in General Surgery”, *Surgical Technology International*, pp. 115- 119.
5. K Keller and J Ackerman, “Real-time structured light depth extraction”, *Proc. SPIE* Vol. 3958, p. 11-18, 2000.
6. 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.
7. G.Turk and M. Levoy, “Zippered Polygon Meshes from Range Images”, *Proceedings ACM SIGGRAPH* 1994, pp. 311-318.
8. Jonsson and Niklas Zethraeus. “Costs and beneits of laparoscopic surgery {a review of the literature}”. *European Journal of Surgery, Supplement* 585:48{56, 2000.