Seeing Is Believing – Visualization Systems in Endoscopic Surgery

(Video, HDTV, Stereoscopy and Beyond)

Amir Szold, MD, SAGES

Edoscopic Surgery Service, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

Corespondence to:

Amir Szold, MD

Director, Endoscopic Surgery

Tel Aviv Sourasky Medical Center

6 Weizman Street, Tel Aviv 64239

Israel

amiki@tasmc.health.gov.il

Introduction

Endoscopic surgery became an important branch of surgery through processes new to the surgical community. In the learning process of new procedures and their implications, important factors were overlooked or taken for granted, mostly because they seemed intuitively very similar to the way we used to function prior to the minimal access era. However, we are gradually learning that this is not the case. Using long, limited instruments does affect our performance. Furthermore, the way we perceive reality, i.e. the fact that we use video cameras and monitors, affects our perception and performance greatly.

In a recently published seminal work (10), it has been demonstrated, although not in a strictly objective way, that severe errors made during laparoscopic procedures are not merely technical ones but rather reflect a critical misinterpretation of the video image. In reviewed cases in which a common bile duct injury occurred, the common bile duct was injured or transected not because of the surgeon's inability to perform technically well, but because of misperception of the image and wrong decision making based on false perceptual information. As the author remarks, "*These data show that errors leading to laparoscopic bile duct injuries stem principally from misperception, not errors of skill, knowledge, or judgment. The misperception was so compelling that in most cases the surgeon did not recognize a problem. Even when irregularities were identified, corrective feedback did not occur, which is characteristic of human thinking under firmly held assumptions*".

The blame, it seems, is probably on the misunderstanding most of us share; that is, we assume that our eyes are a reliable tool to interpret reality, and overlook the crucial part a third factor plays – the fact that the picture we see is a video image, that it has many limitations and may lead us to perform the wrong action based on false perceptual information.

The notion that a better video image is better is not new. Improvements in video imaging are probably one of the factors that enable us to perform much complex procedures than we used to a decade ago. But these improvements are still evolving, and several new devices are on the verge of becoming new tools to perceive the endoscopic reality.

HDTV.

Digital television (DTV) is the transmission of pure digital signals, along with the reception and display of those signals on a digital monitor. Signals can be broadcast over the air or transmitted by a cable or satellite system, where a decoder receives the signal and uses it, in digital form, to be directly displayed.

HDTV is high-resolution digital television and is the highest DTV resolution in the new set of standards. The higher resolution picture is the main selling point for HDTV. It offers 720 or 1080 lines of resolution compared to the 525 lines people are used to in the United States. The result is a very bright, detailed picture that looks "almost" three-dimensional.

There are two main formats for HDTV, interlaced and progressive. The category refers to the scanning system; in an interlaced format, the screen shows every odd line at one scan of the screen, and then follows that up with the even lines in a second scan. Since there are 30 frames shown per second, the screen shows one half of the frame every sixtieth of a second. For smaller screens, this is less noticeable. As screens get larger, the problem with interlacing is flicker.

Progressive scanning shows the whole picture, every line in one showing, every sixtieth of a second. This provides for a much smoother picture, but uses slightly more bandwidth.

There is very little evidence that there is an advantage of using HDTV over high quality 3-chip camera generated video images in endoscopic surgery. Intuitively, it would make sense that very detailed, bright images will help us better compensate for the lack of three dimensional vision, lack of haptic perception and a small field of view. However it seems this is difficult to prove objectively. A study performed in Germany by Buess et all (9) showed a trend to perform better in an inanimate model over a 3-chip camera, but the results did not reach significant statistical difference.

A serious disadvantage for routinely using HDTV in surgery is pricing. Due to component compatibility, the system has to be completely replaced at a high cost, and the image data obtained can be recorded and stored only on compatible systems. It seems that except for in Japan, where the HDTV standard is gradually accepted in the entertainment world as well as in the medical one (7), what will determine the future use of HDTV in the operating theater is the availability of compatible systems and the price.

3D / Stereoscopic video.

The way we perceive depth is complex, and based on two distinct groups of mechanisms. One is of depth clues that are monocular and very intuitive. Relative size, overlapping, shade, color, and movement analysis enable us to asses the distance of objects using "flat" information, or only one eye. This is how we can watch movies and understand them, and today this is the information we use to operate. The other group is binocular information. Our brain receives two pictures from our eyes that are not identical in angle, and the minute differences are used to process a three-dimensional picture in specialized areas in the visual cortex.

In order to mimic this process, several technologies to generate and display an artificial three dimensional image, or rather, a stereoscopic image, are used:

Image generation:

1. **Dual channel video**: In this technology a dual-channel optical scope is connected to two video cameras and delivers two pictures that are displayed to the viewer on a stereoscopic display. An example to this is the system used in the Intuitive© robot, the DaVinci. The advantage of this technique is that it displays a very bright, high-quality image generated from two 3-chip cameras. The disadvantages are common to many systems that implement similar techniques; the images generated by two scopes and two cameras are different not only in the picture angle, but also in brightness, color, optical distortion and sharpness. This commonly results in user side-effects such as fatigue, headache, dizziness and eye-strain. In addition, using current video and optics technology limits the size to a rather large, cumbersome camera-scope complex. Also, producing an angled scope becomes a major technical challenge, especially if the scope is rotatable.
2. **Dual chip-on-the-tip**: This closely related technology, developed and sold by Fujinon, uses dual channel video generated by two video chips that are mounted on the end of the scope. The two images created are digital, and bypass the disadvantages of the optic distortions created by the optic scope. However the problems with separated images still exist and the small distance between the chips enables to create a "weak" 3D effect. In order to understand this drawback the concept of disparity has to be explained. The disparity of a stereoscopic picture is a measure of how different from each other are the two images. This, to a certain extent, is what determines the accuracy and intensity of depth perception, and is determined mostly by the distance between the two sources of the image, which corresponds to the distance between our pupils. When the technical barriers allow a small distance, the picture generated has a weak 3D effect.

3. **Shutter mechanism**: This technology, used in the Storz 3D scope, relies on the fact that the camera is never absolutely still, and that minor angle changes occur between frames. The streaming video is thus divided by a shutter mechanism into two, slightly different steaming videos and when displayed on a stereoscopic display device a stereoscopic image is created. The advantages of this technology are that it can be generated using a single optic scope and there is no need to replace the whole system except for the camera and image processing unit. However, the small, unpredictable changes that are generated by the cameras instability create an image with low disparity, and the 3D perception is weak. In addition, the shutter mechanism creates a flicker that has user side effects similar to those that accompany dual-channel video.
4. **"Insect-eye" technology**: In this technology, developed by VisionSense, a small start-up company, a microscopic array of lenses is placed in front of a single video chip on the end of the scope, similar to the structure of an insect eye. The lens array creates many small, slightly distorted images. When fed into a powerful image-processing computer, the images are divided into "left" and "right" images using a specialized algorithm, and a streaming stereoscopic video is generated. (Figure 1.) The advantages of this technology are that the image is generated from a single CCD, thus avoiding the differences between the two eyes, and that it is the first system to truly generate an image that contains volumetric information about the observed space. This ability harbors the potential of creating hybrid images using other information, like pre-operative CT or MR or online imaginf, and manipulating the image in many ways. The disadvantage is currently lower picture intensity and system instability.

A common problem with stereoscopic systems is that they depend not only of the image-generating technology, but on the display. Ideally, the stereoscopic image should be displayed using a real 3-D display, or a multi planar image, such as a hologram. In this ideal image the viewer ia able to "look around" objects. Unfortunately the currently available hologram technology is incapable of displaying real-time, streaming video. Most other display systems are based on "stereo-pairs", i.e. the display of a slightly different picture to each eye and the creation of depth perception.

The stereoscopic display market is rapidly developing, driven mostly by the game industry. Several display technologies capable of displaying stereoscopic images are currently available:

Head mounted displays: Head-mounted displays are based on small, high resolution LCD screens that display an image to each eye. The most known system used in surgery is the Vista system. The advantages are that the user is free to move and work in a relatively ergonomic position, and that each user can have his own display system. There are several disadvantages to such systems: There is a need to wear a relatively heavy device, and some users suffer from some degree of "immersion", that is they are being cut off the real environment and lack the "feel" of what is happening around them. In addition, current LCD technology is limited in resolution and refresh-rate, or the number of times the lines appear on the screen each second. These limitations allow a medium-resolution stereoscopic image and may result in flicker and the resulting user side effects.

Polarizing screens and lenses: In this technology, two images are projected simultaneously on the screen. The image to each eye is polarized in a different angle by a polarizing screen placed in front of the projecting screen, and when the viewer uses an appropriate pair of passive polarizing lenses, each eye will see one corresponding image. The result is a stereoscopic image. The system is relatively cheap and allows adding multiple users at a low cost. The main disadvantage is that at each polarization a substantial amount of light is filtered, and by the time the image reaches the eye close to 75% of the light is filtered. The result is an image lower in intensity than usual, requiring a darker background environment.

Active shutter glasses: This stereoscopic display technology uses a screen and synchronized glasses. Both screen and glasses have a polarizing screen that rapidly alternates between two polarized images, in a synchronized, rapid rhythm. The synchronization is usually archived using infrared communication between the glasses and the screen. Each eye sees a dark screen or an image in a rapid sequence. This allows the system to display alternatively a different image to each eye, resulting in a stereoscopic picture. Again, some loss of light exists. In addition, the glasses are somewhat cumbersome and suffer from some flickering.

Autostereoscopic displays: New developments in LCD technology allowed the creation of several system that display stereoscopic images without using additional viewing devices. The system is composed of two layers: the back layer is an illumination screen divided into numerous, very thin, bright lines that correspond to pixel columns in the LCD layer in the front. Since our eyes view the screen from slightly different angles, the result is that each eye will view a different set of alternating lines. Thus, a stereoscopic image will be viewed if the screen is viewed from a certain distance and angle. This fixed viewing area is called a "sweet spot", and in a system that uses two 2D sets it is relatively small. The advantage is a high resolution and a relatively low cost. A better version divides the picture into multiple 2D sets allowing the "sweet spot" to be much bigger, and the screen viewed from a range of angles and distances. The "cost" is a lower resolution. These LCD displays are becoming available in laptop computers such as made by Sharp, and screens made by Phillips, Sanyo and DTI.

Although theoretically it would seem that using an intuitively superior stereoscopic / 3-D vision will result in better performance, there is very little objective data to support this assumption. Several studies have shown that subjects performed the same, or slightly better using stereoscopic systems (2,3,6). The improvements observed in some of these studies were seen when errors were measured and less when performance time was a measured variable. However the studies were influenced by factors such as inferior picture quality in the stereoscopic system, inter-subject variability, and lack of standard measurement of performance.

Other studies showed a definite improvement in performance using stereoscopic systems. Studies by Taffinder (8), Birkett (1) and the author (presented at SAGES annual meeting 2002), demonstrate a clear improvement in task performance for both novices and experienced subjects. The improvement, again, was apparent when errors and efficiency of movement were measured but was also reflected in the speed of performance.

It seems that the intuitive notion that it is better to operate using two eyes rather than one, is true, but not entirely. The reasons why it was not objectively obvious in several studies come from severe limitations of the stereo technology used in the studies, including both the image generating and display technologies. In addition, the measured variables do not necessarily reflect the desired outcome in the operating room. Furthermore, as pointed out by Hofmeister (4) even when the stereoscopic image is perfect there are still limitations of scaling and viewing angles that are inherent to the use of an image instead of direct, true 3D vision.

However, the fact that in several studies the use of a stereoscopic image resulted in fewer errors of movement and judgment is a very important one. As previously said, time, or speed of performance, is a mere reflection of the true factors that should be measured – efficiency of movement and errors. As demonstrated previously, the possibility to reduce or eliminate the number of severe errors is far more important than merely reducing the time it takes to perform a procedure.

The advantage of a digital picture that contains the knowledge of the volumetric coordinates of every pixel has an enormous potential. Navigation in the abdominal cavity and inside solid organs, like the liver, may alter the way we operate. In addition, this technology enables future hybridization of several volumetric images, such as CT scan, real-time ultrasound and a video image if registration points are defined. In Japan (5) such hybrid images are already created offline in order to assist the surgeon to navigate in laparoscopic colorectal procedure.
In the future, it will be possible to incorporate the volumetric information with "smart" systems that will analyze the image based on preoperative or real time imaging and "understand" the context of the picture. This will enable the machine to warn the surgeon when working too close to "danger zones", and even performing automatic, pre-programmed tasks when coupled with an operating robot.

References

1. Birkett DH, Josephs LG, Este-McDonald J 1994) A new 3-D laparoscope in gastrointestinal surgery. Surg Endosc. 8(12):1448-51.
2. Chan AC, Chung SC, Yim AP, Lau JY, Ng EK, Li AK. Comparison of two-dimensional vs three-dimensional camera systems in laparoscopic surgery (1997) Surg Endosc. 11(5):438-40.
3. Herron DM, Lantis JC 2nd, Maykel J, Basu C, Schwaitzberg SD. The 3-D monitor and head-mounted display. A quantitative evaluation of advanced laparoscopic viewing technologies (1999) Surg Endosc 13(8):751-5.
4. Hofmeister J, Frank TG, Cuschieri A, Wade NJ. Perceptual aspects of two-dimensional and stereoscopic display techniques in endoscopic surgery: review and current problems (2001) Semin Laparosc Surg 8(1):12-24.
5. Matsuki M, Okuda J, Yoshikawa S, Tatsugami F, Masuda K, Kani H, Narabayashi I, Tanigake M, Lee SW, Nishiguchi K, Tanaka K, Toyoda M, Tanigawa N. [Clinical application of three-dimensional imaging with multislice CT for laparoscopic colorectal surgery] (2003) Nippon Igaku Hoshasen Gakkai Zasshi 63(4):154-9.
6. McDougall EM, Soble JJ, Wolf JS Jr, Nakada SY, Elashry OM, Clayman RV. Comparison of three-dimensional and two-dimensional laparoscopic video systems (1996) J Endourol 10(4):371-4.
7. Okudera H, Kobayashi S, Takemae T, Kyoshima K, Gibo H, Shibuya M, Sugita K. Introduction of high definition television system to neurosurgical documentation (1992) Neurol Res 14(5):386-8.
8. Taffinder N, Smith SG, Huber J, Russell RC, Darzi A. The effect of a second-generation 3D endoscope on the laparoscopic precision of novices and experienced surgeons (1999) Surg Endosc 13(11):1087-92.
9. van Bergen P, Kunert W, Buess GF. The effect of high-definition imaging on surgical task efficiency in minimally invasive surgery: an experimental

comparison between three-dimensional imaging and direct vision through a stereoscopic TEM rectoscope (2000) Surg Endosc 14(1):71-4.

10. Way LW, Stewart L, Gantert W, Liu K, Lee CM, Whang K, Hunter JG. Causes and prevention of laparoscopic bile duct injuries: analysis of 252 cases from a human factors and cognitive psychology perspective (2003) Ann Surg 237(4):460-9.