Clinical Application of Three-Dimensional (3-D) Vision Systems and Virtual Reality Helmets in Video-Assisted Surgery

EMANUELE LEZOCHE, M.D., F.A.C.S., PROFESSOR OF SURGERY, DIRECTOR OF SURGICAL PATHOLOGY

ALESSANDRO M. PAGANINI, M.D., PH.D., F.A.C.S., ASSISTANT PROFESSOR OF SURGERY
UNIVERSITY OF ANCONA, ANCONA, ITALY

DAVIDE LOMANTO, M.D., PH.D.
ASSISTANT PROFESSOR OF SURGERY
UNIVERSITY OF ROME "LA SAPIENZA," ROME, ITALY

FRANCESCO CARLEI, M.D.
ASSOCIATE PROFESSOR OF SURGERY
UNIVERSITY OF L'AQUILA, L'AQUILA, ITALY

Thoraco-laparoscopic surgery presents a series of technical difficulties linked mainly to the necessity of acquiring proper motor coordination and spatial reconstruction of an operative field that is seen from a distance on a two-dimensional video monitor, in the absence of any direct tactile feedback. In an effort to improve the motor coordination of the operating surgeon and of the surgical team, many apparatuses have recently become available on the market that allow the reproduction of a 3-D image on a video monitor. Such apparatuses have technical characteristics that are substantially diverse in technology and provide significantly different end results.

MECHANISMS OF THREE-DIMENSIONAL VISION IN HUMANS

Two-dimensional vision is what one experiences, for example, when looking at a photographic print. It is the same effect given by present-day television and video monitors. Independent of the quality and resolution of the image, screens cannot provide the sense of depth that is a characteristic of human vision.

How is it possible that our eyes furnish us with a two-dimensional image of what the monitor screen displays but, at the same time, give us a 3-D image of the monitor itself?

To understand the mechanism of depth perception in higher animals, we must keep in mind that, in reality, each of the two eyes transmits a slightly different image of the same object in relation to the diverse angle of vision between the right and left eyes. To better illustrate this mechanism, consider the vision of a prism that is placed in front of us at a distance of 20 to 30 cm, and consider its two consecutive faces (the right face, which is gray, and the left face, which is white) placed in front of us so that the angle between them is aligned with our medial sagittal plane. If we look at the prism with only the right eye and then with only the left eye, what we see are two substantially diverse images of the same prism, as shown in Figure 1. In fact, when we look with only the right eye, we have a *priority* vision of the gray prism face with a foreground vision of the white prism face; in contrast, when we look at the prism with only the left eye, we experience just the opposite, that is a *foreground* vision of the gray face and a priority vision of the white face. However, by slowly moving the prism away from its original distance of 20 to 30 cm from us, the images from our left and right eyes will become more and more similar (i.e., they will become two-dimensional), with a gradual loss of 3-D resolution.

As a matter of fact, depth perception arises mainly from two spatially diverse views of the same scene. Inferior animals not possessing 3-D vision must compensate by observing a scene from different physical positions, thereby changing their angle of view. For example, a snake before attacking a target carries out a series of rapid head movements in order to obtain a picture of its target from different points of view in rapid succession, allowing it to acquire a sense of depth perception and thereby estimate the distance to its intended target.

In humans, the two diverse images coming from the right and left eyes are "fused" by the brain into a single image, reconstructing a sense of depth perception; this phenomenon was called *stereopsis* by Sir Charles Wheatstone, who discovered it in 1833. Depth perception is a complex phenomenon involving an elaboration and spatial reconstruction at the level of the central nervous system. In fact, even people who have lost an eye can recover, with time, a sense of depth perception, even to the point of comfortably driving a car, for example. It is this central nervous system reconstruction capacity that allows the expert laparoscopic surgeon to carry out complex procedures in two-dimensional (2-D) vision that would without a doubt be much easier to do with the help of 3-D vision, even by inexperienced laparoscopists.

The term 3-D is not always used properly. Such a term is also used to describe a special animation effect of some video games that gives the sensation of movement in space on the computer screen, without really corresponding to actual 3-D vision and to a sense of depth perception. In fact, there are other mechanisms that give a perceived 3-D effect such as shadow effect, lateral parallax, etc. Such an effect has been famously exploited and employed by painters of the Renaissance. As an example, note the 3-D effect that can be appreciated in this painting by Melozzo da Forlì (1438-1494) (Fig. 2), a pupil of Piero della Francesca. In this painting the 3-D effect is obtained by the parallax effect of the window frame and by the interplay of the shadows that cause the viewer to perceive that the angel is moving out of the window (San Marco Sacristy, Basilica of Loreto, Loreto, Italy).

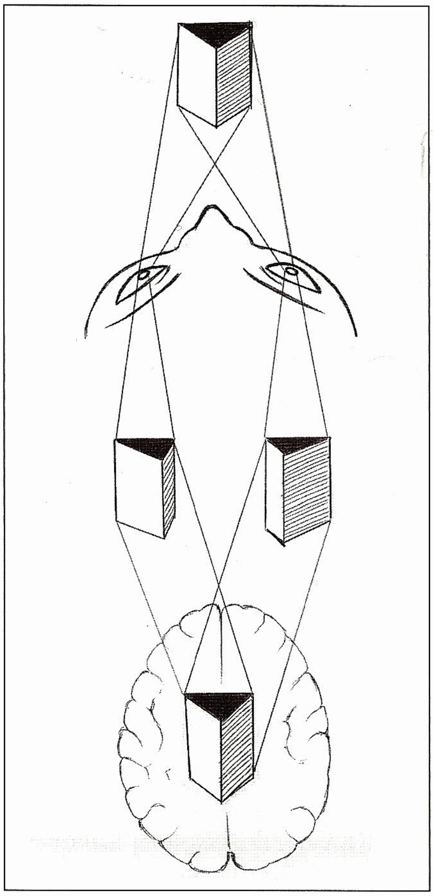

Figure 1. Depth perception in humans. Each eye receives a slightly different vision of the same object.

Figure 2. A detail from a painting by Melozzo da Forlì (1438-1494) in the San Marco Sacristy, Basilica di Loreto, Loreto, Italy.

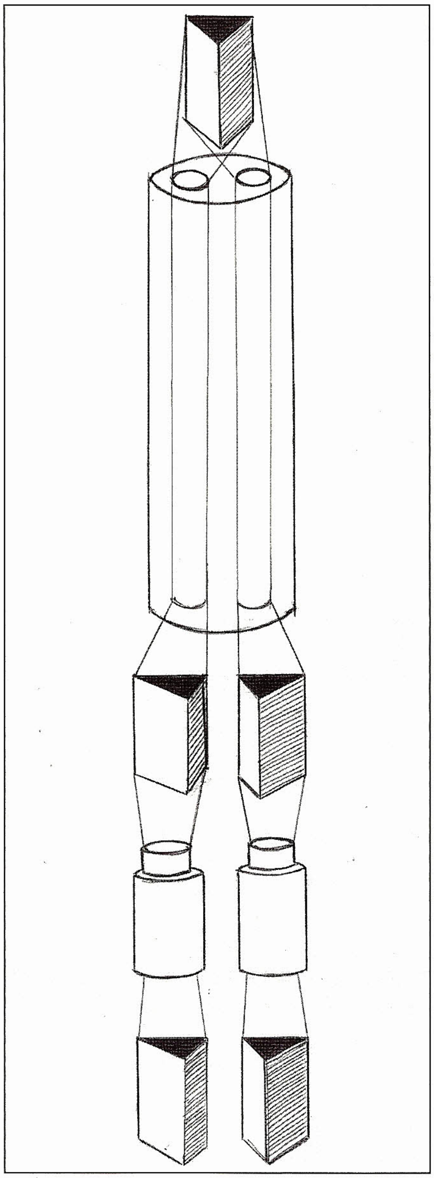

Figure 3. Double optical channel of the 3-D laparoscope.

TECHNOLOGY OF TWO-DIMENSIONAL (2-D) VIDEO SYSTEMS

In 2-D systems only one optical channel transmits the image to an image transformation cell, called CCD (Charged Coupled Device), that will convert it to electromagnetic waves. The CCD is composed of thousands of light-sensitive electronic sensors called PIXELS (PICTure ELements) that are aligned next to one another in the form of a mosaic. Light from the image activates the PIXEL sensor which in turn generates an electronic

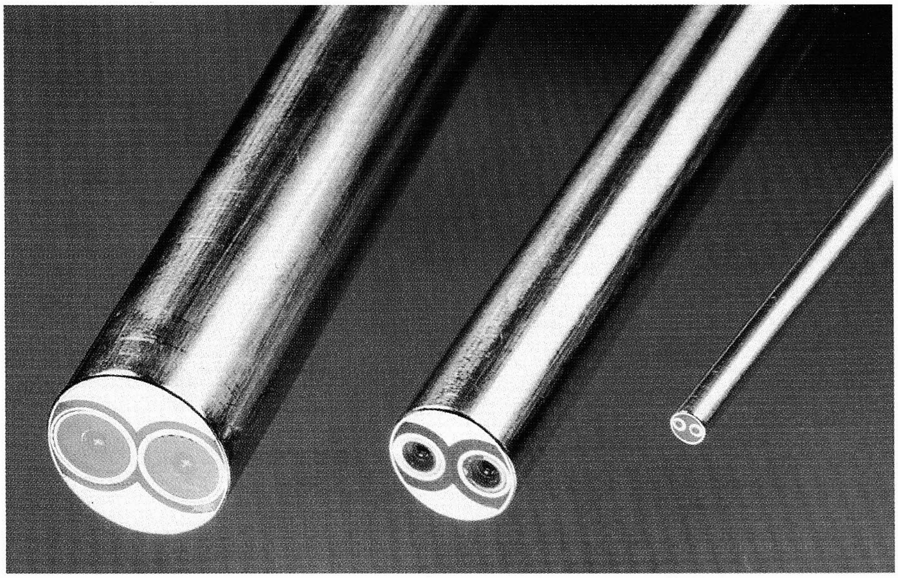

Figure 4. Double optical channels of 3-D laparoscopes of different diameters.

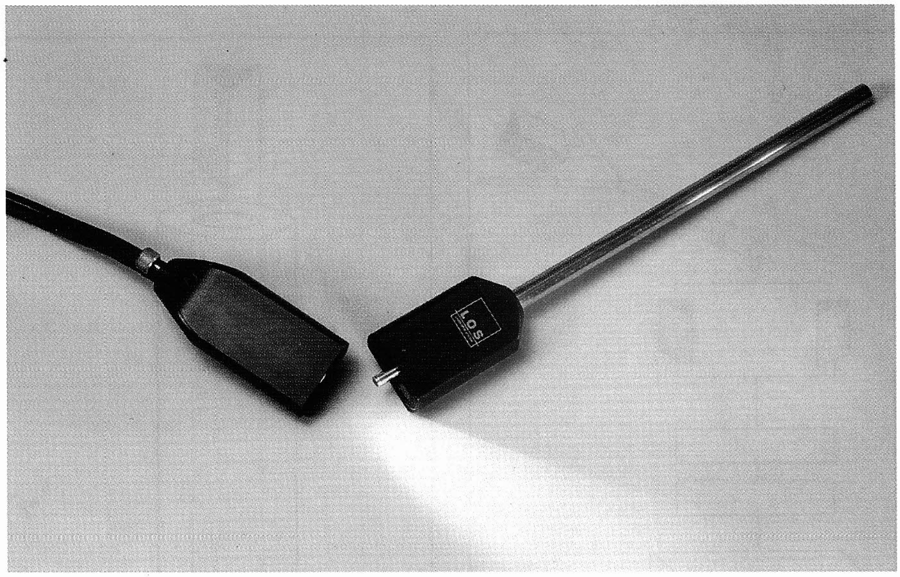

Figure 5. 0° and 30° lenses of 3-D laparoscopes.

pulse, thus allowing the image to be transmitted as an electronic signal. The result of the electronic signals generated by all the pixels per unit of time is called *frame*. Obviously, the quality of the image depends on the power of resolution of the system, which is related to the number of pixels and, therefore, to the number of lines that the pixel rows can produce. Present-day commercial videocamera CCDs consist of 300,000 to 800,000 pixels—not too many if we consider the 15,000,000 receptors present in the human eye, connected to 1,000,000 nervous fibers.

Another aspect that conditions the quality of the image is the way in which color is reproduced. Color may be produced using either a single CCD or three CCDs. In this latter case, the image passes through an optical prism that divides the image light into the three primary colors: Red, Green, and Blue (RGB). In this way, each CCD receives only one color which is processed separately by the videocamera's microprocessor, thereby reducing the risk of interference (called "image background noise"). Video cameras using such a system are called "3 CCD or 3 Chips camera." Obviously, such cameras are more expensive and cumbersome, and obviously they should not be confused with 3-D technology.

TECHNOLOGY OF THREE-DIMENSIONAL (3-D) SYSTEMS

Three-dimensional systems are based on a very complex technology. Let us examine the various components in order to evaluate the differences of the various models available on the market.

3-D laparoscope. The 3-D laparoscope (Fig. 3) is different from the 2-D model because it has a double optical channel consisting of rod lenses (first used by Hopkins in 1953). Such lenses must have a smaller diameter than that of the lenses used in a 2-D laparoscope since they must fit into a double channel of a 3-D laparoscope that must not exceed 10 mm in external diameter (Fig. 4). In addition, there must be sufficient optical fibers within this same diameter to transport light adequately to the operative field. Since the space occupied by the double-channel rod lenses is greater than in a 2-D system, it is important that particular care be taken in selecting systems that are able to transmit enough light to the operative field. Light transmission must be adequate to ensure optimal vision of the operative field, which may also include surfaces that normally absorb a large amount of light such as blood, hematomas, muscles, hepatic and splenic parenchymas.

Like 2-D laparoscopes, 3-D models are commercially available with angled lenses, but while the former have a large assortment of angled lenses (0°, 30°, 45°, and 90°), the latter have forward oblique angled lenses of only 0° and 30° angles (Fig. 5). In order for 3-D vision to produce a sense of depth perception, a frontal lens angle of not less than 30° is required.

It is also important to point out that a 2-D laparoscope allows to rotate the lens around its own axis. By doing so, a 45° forward oblique view angled lens allows to visualize the entire circumference of an anatomical structure placed in front of it. If the same maneuver is done with a 3-D laparoscope, the operator becomes disoriented due to a rotation of the spatial horizon, since the two lenses are intimately connected to the two videocameras. For this reason, it is extremely useful in 3-D systems that the lens assembly be detachable from the videocamera and that it may easily be inserted into two alternative positions, in order to allow the possibility of changing the direction of the lens looking both upwards and downwards to ensure vision from below and from above, respectively (Fig. 6).

It should also be remembered that the angled lens of the laparoscope increases light absorption. It is therefore important that the light transmission capacity of the laparoscope be quantitatively high and that the distribution of the light-transmitting

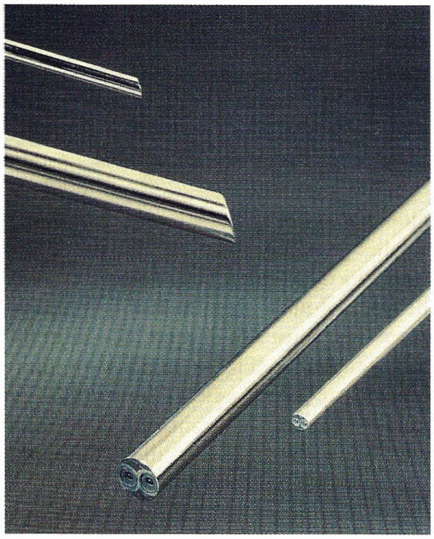

Figure 6. The lens assembly is detachable from the video camera. This allows the 30° angled laparoscope to be connected to the video camera in two alternative positions to allow vision of the operative field from below or from above.

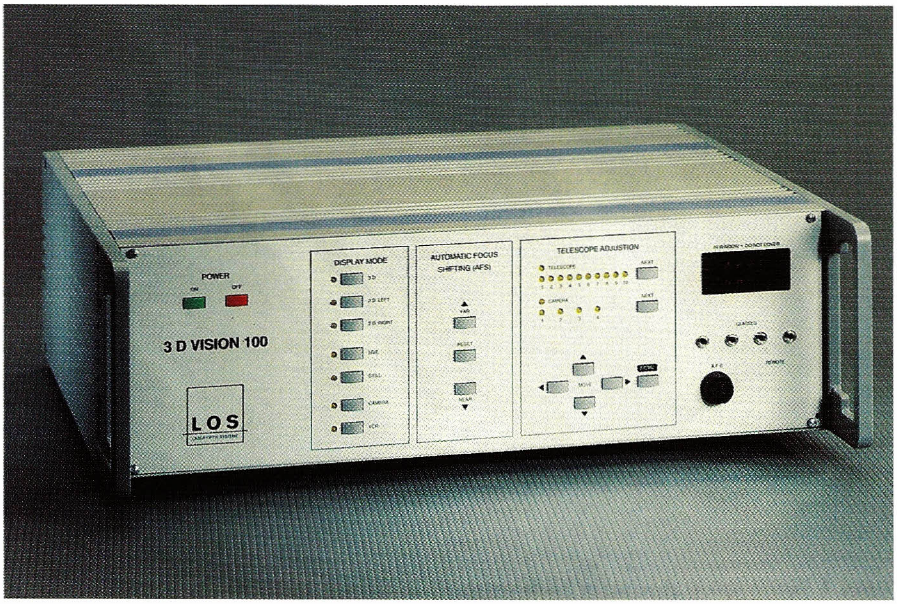

Figure 7. The 3-D electronic microprocessor.

optical glass fibers be evenly distributed around the two optical channels, to avoid a dyshomogeneous illumination of the operative field, the so-called "shadow effect."

Most 3-D systems are equipped with *two videocameras* receiving images from each of the two optical channels of the laparoscope. Each optical channel is connected to a single videocamera to ensure binocular vision of the system. At present, 3-D systems with single CCD videocameras are only available. However, 3 CCD video systems are currently being studied.

If we consider that in humans the distance of the two eyes is approximately 6.5 to 6.7 cm and that we have good depth perception from a focal distance of only a few centimeters to hundreds of meters, it is easy to understand how the laparoscope (with a distance between the two lenses of less than one-tenth of the interpupillary distance of man, i.e., about 0.5 cm) can theoretically ensure three-dimensional vision of the entire operative field.

Electronic microprocessor. An electronic microprocessor (Fig. 7) is required in 3-D systems since the two video images cannot be reproduced simultaneously on the monitor screen. In addition, the monitor's resolution may be different from that of the videocamera. In fact, present-day videocameras generally have 750 line resolution while 3-D monitors have 1,024 line resolution. A microprocessor is therefore needed to store the two images and process them in a way to make them compatible with the resolution power of the monitor, sending signals of the two images alternatively to the video screen at a

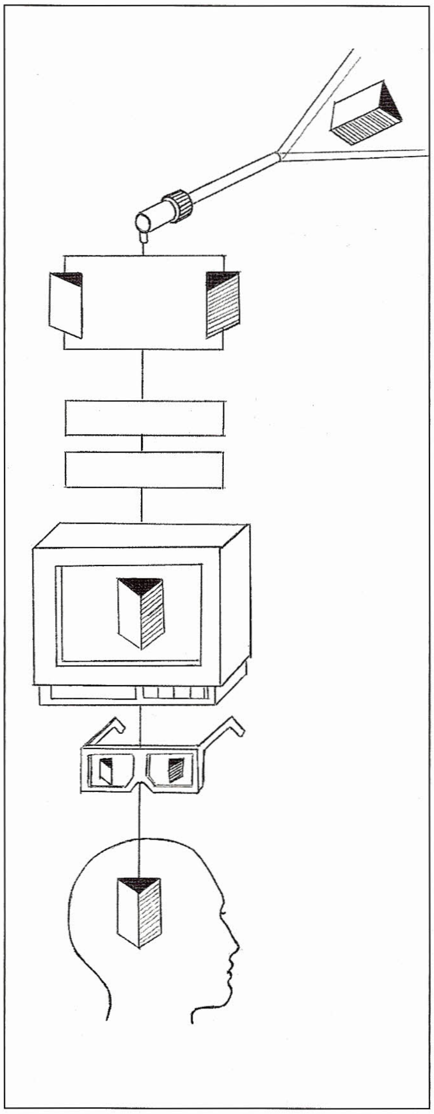

Figure 8. Reconstruction of the 3-D image by a polarized optical system.

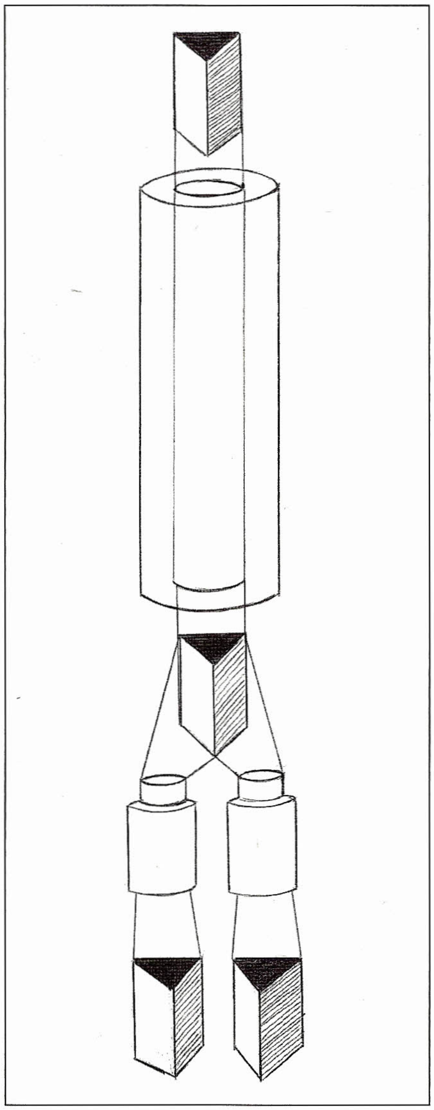

Figure 9. Optical system employing a single channel.

frequency of 120 Hz.

Currently, true *three-dimensional monitors* are not yet available; therefore, a polarized optical system is employed to reconstruct the 3-D image (Fig. 8). Such a system alternatively transmits images from the right optical channel to the right eye and images from the left channel to the left eye. The system has a high frequency monitor (120 Hz/sec) with a high resolution of 1,024 pixels. The monitor receives alternating right and left images from the microprocessor 120 times per second. If the observer does not wear *3-D glasses*, there will be an overlap of the right and left images resulting in double and confused vision. On the other hand, the use of glasses employing polarized lenses permits the left eye to see only the left image when it appears on the screen and the right eye to see only the right image. This is achieved by transmitting the images at a velocity superior to that which the human eye can perceive, taking advantage of the effect of retinal after-image, giving a continuity to both images, resulting in 3-D vision with excellent depth perception. The "opening" and "closing" of each of the lenses of the 3-D glasses is controlled by an infrared signal sent from the system and allows for perfect coordination with the image produced on the monitor screen. This mechanism of "opening and closing" is called "flick." It is similar to that used in cinematography where there is a loss of discontinuity between each film frame and the next, giving the effect of absolute image continuity (i.e., smooth motion). A screen has recently appeared on the market that is positioned in front of the monitor and that reproduces the "flick mechanism" when the observer is wearing polarized lenses. This system is better tolerated because the weight of the polarized lenses is lower.

In order to compare the technology used in this system with that used in other systems, one should note that normally employed computer screens have a resolution of about 640 horizontal pixels and 480 vertical lines, while the recently introduced High Definition screens have a resolution of more than 1,000 lines. These values are still very far from the resolution obtainable by a photographic slide or print (2,000 x 20,000 pixels). The human eye, for comparison, has an extraordinary resolution power, more than 4,000 x 40,000 pixels!

Optical systems employing a single channel (Fig. 9). Besides the above described double lens system, there exists a 3-D system with a single channel optics. Such a system

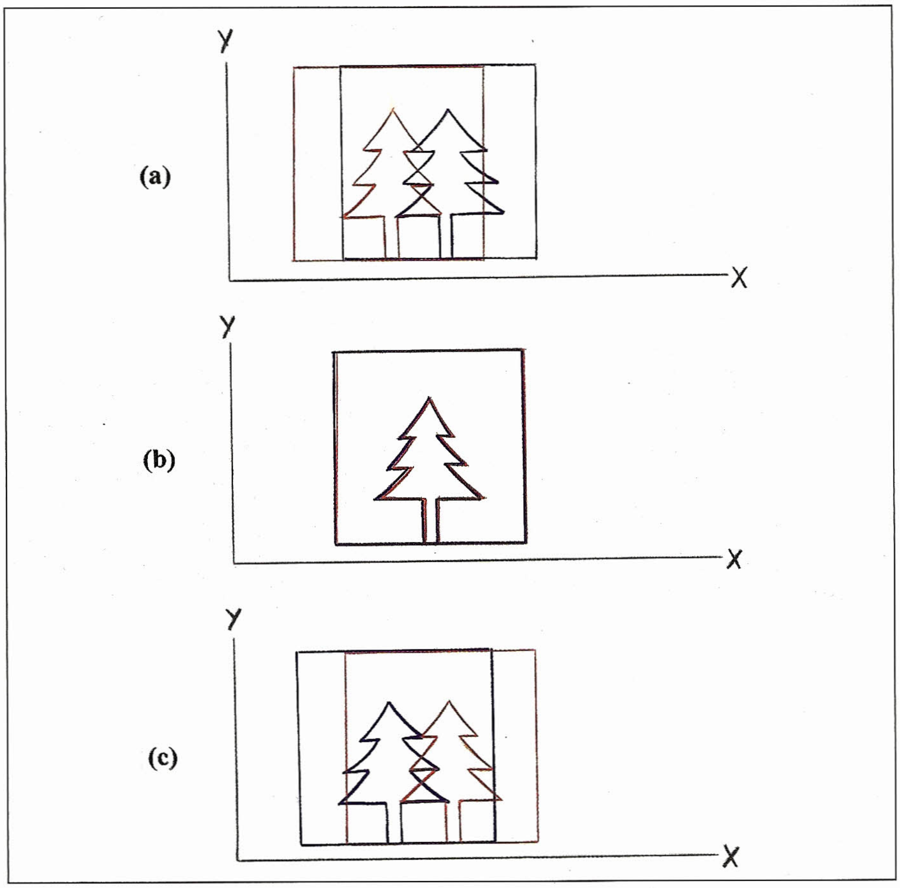

Figure 10. (a) Backwards movement; (b) overlapping of the two pictures; (c) forward movement.

"doubles" the image seen by a traditional laparoscope, alternatively transferring it to the right and left eyes, using the same technology described above. In reality, we are talking about a vision system that is significantly different from normal physiological conditions. In fact, the image that normally reaches the right eye is slightly different from that which reaches the left eye, due to the small visual angle difference between the right and left eyes. On the other hand, the two images produced by the single channel system are less different (i.e., no substantial difference in visual angles), and it is indeed this slight image difference between the right and left eyes that allows the central nervous system visual center to reproduce 3-D vision with a consequent sense of depth. Therefore, under similar conditions, this system cannot reproduce the same 3-D effect that one can obtain with a double lens system. Another limitation is that only frontal view 0° lenses can be employed.

Automatic focus shifting (AFS). All systems equipped with a double lens channel produce two diverse images. Therefore, the projection on the monitor screen of these two images translates to a certain degree one with respect to the other on the cartesian X axis. This is one of the mechanisms necessary in order to obtain a 3-D effect. However, when the two images are overly shifted and the difference overruns the ability of our eyes to compensate, we experience a sensation of visual fatigue and headache. It is also important to remember that while humans have a certain capacity to compensate for horizontal translation (along the X axis), we do not have the ability to compensate for vertical translation (along the Y axis). At present, only one apparatus is available on the international market (L.O.S., Mainz, Germany) that allows a regulation of the two axes (X and Y) of the image. This mechanism permits not only preoperative calibration but also intraoperative regulation of the horizontal translation of the two images projected on the screen in relation to the focal distance of the intended operative field (Fig. 10a). In this way, if the surgeon is working at a small focal distance (e.g., 1 cm) he or she can bring the two images on the screen to an almost complete overlap by means of a remote control (Fig. 10b). On the other hand, if the surgeon wished to work at a focal distance of 5 cm, the distance most frequently used in laparoscopic surgery, he or she would

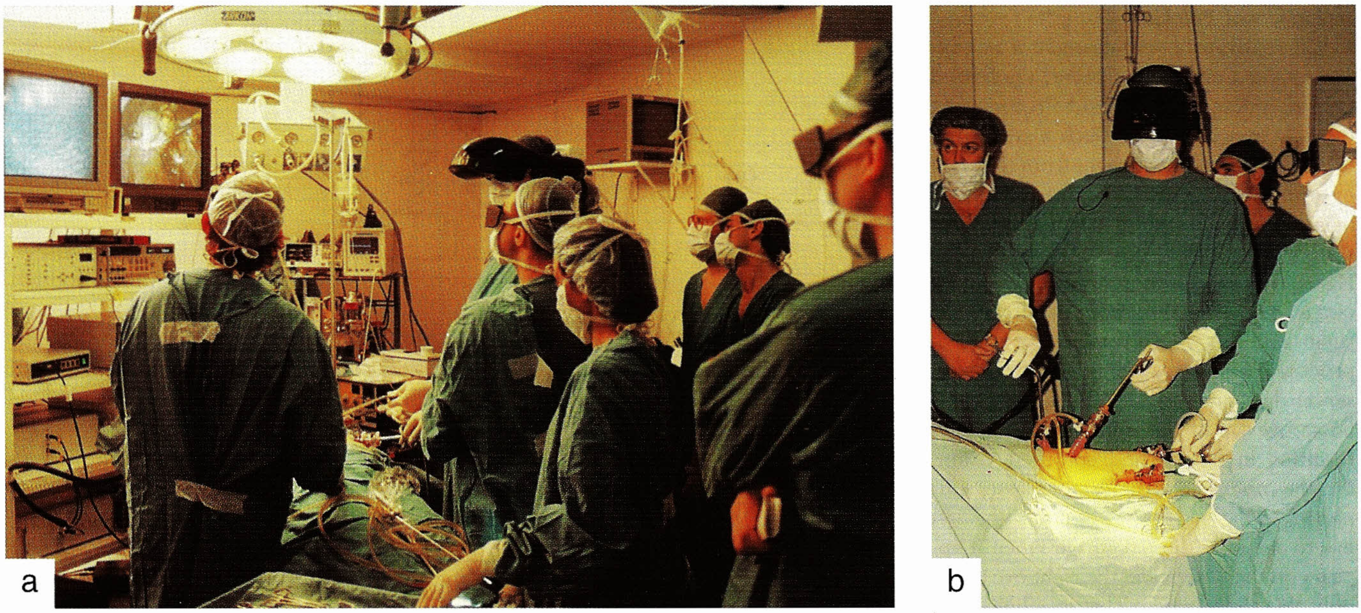

Figure 11 (a, b). Virtual reality helmets in clinical laparoscopic surgical procedures.

simply have to press a reset button and the images would return to a predetermined position with translation on the horizontal axis (Fig. 10c). If the focal distance were to exceed 5 cm, the surgeon would use the specific remote control again to increase the translation of the images in order to optimize the 3-D effect over the intended operative field. It is important to note that in all fixed focal systems there is a 3-D effect at a set distance from the lens; as one moves away from this fixed focal point, the 3-D effect is slowly diminished. This is especially relevant in advanced laparoscopic surgery (e.g., colon surgery), where the surgeon must frequently work at varying distances from the lens. These phenomena appear to be of lesser importance in instances when a small monitor screen is used. On the other hand, the image magnification of large screen monitors results also in an increased distance between the two images. Systems not equipped with Automatic Focus Shifting project images that can overrun the human eye's ability to compensate for the distance between the right and left images, resulting in visual fatigue and frequently in headache. The use of angled lenses at 30 degrees, for example, reduces the relevance of these problems because the lenses work at a less variable focal distance with respect to that of a frontal view 0° lens. This is the reason for which some companies offer 3-D systems only with a 30° forward oblique lens.

Beyond these strictly technical aspects, other aspects related to human physiopathological conditions that limit the use of these systems should be taken into account. In fact, in subjects suffering from a prevalence of one eye over the other ("lazy eye") or from latent strabismus and other more complex neurosensory conditions, the use of 3-D polarized flicking glasses may result in serious headaches or visual fatigue and nausea. An optometric check-up is indicated in such circumstances to uncover possible latent visual conditions (which are present in approximately 15% of the general population).

"Virtual Reality" Helmets

The so-called "virtual reality" helmet has been designed with the specific purpose of depriving the subject of the sense of visual perception that normally surrounds him or her, transmitting only a controlled input to his or her visual senses, which results in a total immersion of the subject into the perceived vision transmitted with the helmet.

The helmet, which is widely applied in the field of video game technology, was derived from use in the military. In fact, its first application was for air combat training and simulations. It is interesting to note that in the recent Gulf War, pilots who underwent virtual reality training performed decisively better in the real air raids than did pilots who had not undergone such training. The technology of military virtual reality training is not easily transferable to the surgical field due to high cost. A single virtual reality helmet system costs approximately $1 million.

Presently available commercial helmets have a series of technological gaps that will probably be overcome in the near future. The first companies dealing with such technology that we tested were "Liquid Image" and "Virtual Research." The former is a Canadian company and produces monoscope helmets with a 6- or 4-inch monitor screen having a resolution of 720 horizontal pixels and 280 vertical pixels, with a biconvex Fresnell focus lens. Such a helmet realizes 3-D vision by binocular vision of a liquid crystal display screen and is definitely heavier and more immersive. The Virtual Research helmet is lighter and more manageable. Virtual Research has produced a more advanced helmet with a superior resolution (1,200 x 1,024 pixels) and an innovative technology. In fact, monochromatic images of the three primary colors (red, green, and blue) are rapidly alternated (120 Hz) in repeated scans resulting in a persistence effect on the retina, which recomposes the original color image. There are many new models recently introduced on the market: "Cyber Max," "I Glasses," "d Visor," etc., and doubtless many more will appear in the near future. Recently we have been able to employ another helmet during surgery, utilizing two mini-screens that continuously project the diverse images coming from the right and left videocameras respectively.

The advantage of such a system is the elimination of the microprocessor (cf. above paragraph on 3-D system microprocessors), with consequent cost reduction and possible improvement in image quality. Another important advantage is the elimination of the polarized lens "opening and closing" system of the 3-D glasses (flick effect), responsible for the difficulty of some people to adapt to present day 3-D technology, as previously described.

INITIAL USE OF VIRTUAL REALITY HELMETS IN CLINICAL LAPAROSCOPIC SURGICAL OPERATIONS

The first sensation that the surgeon experiences is that of complete isolation in what he or she is doing (Figs. 11a-b). Such a condition allows for a greater concentration and a more direct involvement in the problems connected with the surgical procedure. On the other hand, such isolation can also lead to problems in holding and introducing instruments through the laparoscopic access ports. Because of the helmet, the surgeon finds himself in a situation of almost total blindness regarding the surgical field outside the patient. Structural modifications were necessary to overcome this drawback and to allow a view of the patient's abdomen.

One advantage in the use of the helmet is that it always allows perfect alignment with the operating field using simple movements. As any surgeon who practices advanced laparoscopy knows, some procedures require a continuous variation of the operative field. Such is the case of anterior resection of the colo-rectum, for instance, during which the surgeon first operates on the pelvis and then operates on the left upper abdominal quadrant. Since the laparoscope is usually inserted at the umbilicus, to maintain proper alignment of the surgeon, laparoscope, operative field, and monitor, one must repeatedly change the position of the operating table and move the cart on which the video-instruments lie, with obvious disorder resulting in the operating room. On the other hand, the "virtual reality" helmet allows the surgeon to turn to the area where the operation will take place, a very simple and instinctive gesture.

Without a doubt, we are still very much in a pioneer phase. It is likely that with helmet technological improvements, the necessary improvements in resolution power and widening of the visual field will follow.

Finally, it is important to remember that the use of the "virtual reality" helmet in the operating theater may result in a problem of psychological nature. As already described in other fields, where such technology has been used, there is a risk that the subject who is totally immersed in the so-called "virtual reality" may have a desire not to return to the outside reality. This peculiar working condition will require an attempt at psychological evaluation and study.

A further application of such helmets will certainly be in the field of post-graduate teaching and training. Interactive virtual reality will undoubtedly find several applications in surgery.

CLINICAL EXPERIENCE

Since July 1994, a L.O.S. 3-D system (Mainz, Germany) has been routinely used in the authors' clinical practice to carry out both basic and advanced laparoscopic procedures.

At present the authors are conducting a study, yet to be completed, to compare a 3-D system to a 2-D system with 3 CCD videocamera in performing basic laparoscopic procedures, such as cholecystectomy. When working with the 3-D system, a 30° forward oblique view lens is usually employed, whereas the 2-D system employs a 45° forward oblique view lens. The 3-D system has been used in over 300 operations by the authors, with a maximum continuous use time of 10 hours (excluding anesthesia induction and wake-up time), without any visual side effects. Preliminary data indicate a mean time savings of about 5% to 8%. If we consider more complex specific maneuvers, requiring a greater sense of depth perception (such as cystic duct cannulation, suturing, etc.), the time saved reaches 10% to 15%.

Another relevant issue is the easier indirect recognition of some anatomical structures by the perception of their relief within the structure they are contained in. This is the case of the common bile duct in the hepatoduodenal ligament or of the vessels within the mesentery. Such an enhanced ability to recognize the anatomy is scientifically difficult to verify, given the impossibility to develop a specific study protocol.

All studies comparing 3-D with 2-D systems are biased by a personal preference of the surgeon for one system or the other, a bias that cannot be overcome. To reduce this problem, the authors are conducting a series of tests with the cooperation of medical students who, having had no previous laparoscopic experience, have also, it is assumed, not had any predetermined preference for one system or the other. However, the use of 3-D polarized glasses, which is fascinating and cumbersome at the same time, may still interfere with the student's preference for one system or the other.

Beyond the certainty that we can obtain only from scientific studies (irrefutable on the methodological level), we believe, based on our clinical experience of about 1½ years, that three-dimensional vision is a very important aid, even for the expert surgeon, in the execution of both basic laparoscopic procedures and, more importantly, in carrying out more complex, advanced laparoscopic maneuvers, such as the correct orientation of the needle in the needle-holder for more efficient suturing. In addition, it should be stressed once more that the selection and use of the best 3-D equipment is of the uppermost importance since its quality has the greatest impact on the surgeon's intraoperative comfort and well-being. Since a 3-D laparoscopic system is usually purchased as a supplemental apparatus, we feel that its selection should not depend so much on cost as on the image quality that the system is able to offer. STI