Virtual Reality, A Training World for Shoulder Arthroscopy.

Authors: Simon Grange*, Tim Bunker*, Jason Cooper,

**Princess Elizabeth Orthopaedic Hospital*

Exeter

EX2 4UE U.K.

Department of Computer Science

University of Exeter

Exeter, EX4 4PT U.K.

Tel: *01752 823513*
e-mail: *simon@virtualsurgery.com*

Abstract : **This study describes a potentially cost effective, new system for the use of video images as a possible resource for computer-based arthroscopic training. The feasibility of displaying images as a direct response to movements of an arthroscope made by the surgeon in training was tested.**

A 6-D locator device is used for the input of the video frames with their respective location. The problems of space and speed with respect to storage and retrieval of images; and the difficulties associated with ensuring that sufficient and appropriate images are captured is discussed. This approach provides a potentially cost-effective solution to training for minimally invasive surgery and acquisition of physical skills in triangulation and instrument manipulation.

Keywords: Simulator - Surgery - Virtual Reality - Video - Minimally Invasive Surgery - Arthroscopy - ATM - Networks - Training

1. Introduction

The last decade has seen a spectacular growth in shoulder arthroscopy (Bunker 1991). Particular skills are needed for successful arthroscopy. Using a virtual reality based surgical simulator the surgeon in training can develop such skills without placing the public at risk in a similar way to flight simulation in the flight training of Pilots. Flight simulation is an accepted system for gaining experience in a low risk environment, which is predominantly used for procedural training.

A surgeon navigating the shoulder joint through manipulation of the arthroscope has to accommodate a view of anatomy as seen magnified through the eye of an arthroscope. The surgeon performs skills such as triangulation, locating and tracking an instrument introduced through a different portal to the arthroscope. While practise on shoulder models is helpful in the acquisition of some of these skills it has drawbacks (Dumay 1995, Ota 1995, Satava 1993, 1996). The surgical model (a mechanical simulator) lacks realism (Cooper 1995);

- They are clearly artificial constructions: the visual feedback from the scope is unrealistic.
- The cost of models is appreciable and thus limits opportunities for home based training,
- Models, per se, are unable to provide any tutorial-style feedback.

Hence there is a need to develop computer based surgical simulators drawbacks (Dumay 1995, Satava 1993, 1996). An alternative to the visual feedback from the physical model is computer generated graphics feedback using state-of-the-art graphics and virtual reality techniques. This idea is being pursued by several research groups, (Richardson 1994, Wind 1986, Ziegler 1995). However, as the developers report, many compromises of visual accuracy have to be made if real time animation of shaded images is to be achieved. This technique involves high investment in equipment and lengthy development of graphical models. The visual complexity of the shoulder presently renders such an approach impractical for high-fidelity training.

An alternative that we are considering for training in shoulder arthroscopy is visual feedback from video. During arthroscopy the picture on the monitor is often recorded on video (to assist the surgeon in diagnosis, or for later review with colleagues, or indeed for tutorial-style, rather than skill-based, training purposes). Imagine sufficient video images stored to provide appropriate visual feedback to the surgeon practising on a model representing the shoulder. Further imagine that the movements of the arthroscope can be determined such that appropriate visual video images can be displayed on the monitor. The surgeon can thus practise in an environment providing a high degree of realistic manipulative and visual feedback. The model could be a simple upturned plastic jar, sitting on the other side of the world using remote access. The training would lack tactile fidelity but nevertheless provide a possibility for diagnostic and operative simulation. The essence of our approach provides for use of video in these two styles of training.

Learning arthroscopic skills requires firstly the ability to navigate, secondly, to triangulate and thirdly, pattern recognition of anatomy and pathology. This system should be excellent in all three of these modes. The potential difficulty comes when one attempts to considerably modify the environment, for example with training of specific procedures such as repairing dislocations and stitching rotator cuff muscles. This is because modelling of surgical procedures require the incorporation of a temporal dimension, i.e. there is a logical forward progression, a sequence of events. The order is vital, in that we cannot display a scene that is out of sequence and still expect the operator to "willingly suspend disbelief", since the simulation would no longer obey one aspect of the operators' reality.

Complete simulation would be the ultimate goal, the starting point is to develop a system where we are able to simulate the whole of the 2D mapping of the 3D environment which the surgeon would normally navigate.

In order to perceive image as a moving image with no flicker, a frame reference greater than 20 per second (Hertz - Hz) is required. This can be performed in real time or put together as an animation sequence and then replayed. Obviously for minimally invasive surgery (MIS) to work real time is required. In order to achieve the quality of resolution of image and real time generation, it is necessary to look at a system which is an alternative to complete graphical generation at this stage of the evolution of affordable computer technology (Satava 1993). These are limitations of software and hardware and its ability to compute an adequate resolution of image in real time. It is from this background that we have generated the video based surgical simulator.

2. Method

In order to generate an adequate illusion of a moving image, a sequence of correct digital video images is selected by the computer and presented in relation to the observer. As the surgeon moves through the environment the sense of movement is generated by the continuing sequence of video images shown on the monitor, much as if one were looking at the monitor of a normal arthroscopy. This system works well, aiding in creating an adequate illusion, because the procedures are immersive, in that the operator is normally working from a TV monitor.

The building of a virtual world is best described in the following stages;

1. The Process of world building.
2. The Hardware
3. The Software

2.1 The Process

2.1.1 Phase 1: Recording a world.

Firstly fixed points, such as the portal of entry into the joint are recorded, and then all possible views from that portal, guided by a 3D mapping program, are recorded.

Capture Visualization

We have developed this software to show which frames, and how many angles of them, have been captured in a boxed 3D space. The capture program works in real time. The program shows a 3D grid and we have used it to capture the complete 1000 frame recording of the "desk world" a calibration environment used to test the proposed frame selection with the actual view from the camera. We have also scaled the visualization module up to deal with 60,000 points although this is easily modified for a larger or smaller numbers as is necessary. This system is referred to as the **gvo** program. A sample screen is shown in figure I. Video image frames are stored on hard disk or video tape.

2.1.2 Phase 2: Computer analysis of Videotape.

Using a modem to record co-ordinates on the LHS of audio track on VHS video relates the 3D point of the tip of the arthroscope in the virtual world with the appropriate frame in the computer's memory. This allows for recording sound on RHS should an audio reference be required for tutorials. This right hand track was not used in these experiments.

Over sampling during recording could lead to more than one frame being available for selection during playback: Thus the software allows for interpolating co-ordinates and selection of the best frame.

2.1.3 Phase 3: Playback.

Playback involves using the **Bird™** transmitter, which is moved whilst it is attached to an arthroscopic representation (a rod of wood 3mm in diameter and 240mm in length in our experiments) See Figure II.

Repeating the sequence of VR input, frame decision, frame output, generates the 'moving' image. Graphics could be integrated as an adjunct to this sequence to incorporate operative procedures in real time. The Z-co-ordinate varies least during recording, thus recording is generated in layers similar to that of an onion. This correlates with the geometry of a computer's hard disk and makes retrieving of adjacent frames a matter of moving to adjacent tracks on the disk. While enhancing the speed of playback 1.7 fold (as opposed to randomly distributed data), the search for the next best frame to display can also be localised based upon the speed of movement of the 'virtual' arthroscope. To search all frames would be too slow, thus the algorithm is generated to 'look' for the frame that is likely to be nearest, should the current path and speed of the 'virtual' arthroscope be maintained.

Grid search errors could be introduced by using a search based on the geometric partitioning of the search space, since the track of the user (i.e. **the Bird**™) will not necessarily be similar to that of the original 'frame recording', and so in effect the frame selected is the computer algorithm's 'best guess', however the impact of this is not always apparent to the user. See Figure III.

While the search locality is based upon the speed of the operator's movement and in order to minimise the time required for this in real time play back, the next step will be to pre-compute all co-ordinate choices.

2.2 Hardware - The Computer System.

2.2.1 The Flock of Birds.

This is a 6D sensor (locator) device for tracking the arthroscope. It is a serial device that can allow for 6 degrees of freedom: (Euclidean 3D, plus Roll, Pitch and Yaw). Only the first three are presently used, however the authors envisage that in the final teaching models, a 4+D system will be required in order to accommodate the rotational component of the arthroscopic technique.

Although six degrees of freedom are required to completely specify the location (3D) and orientation (roll, pitch, and yaw) of an orientable surgical implement held in Euclidean space, 4 (four) degrees of freedom are adequate to describe a 'virtual arthroscope' entering or rotating via a fixed portal. **The Bird**™ is used to identify an object's location in 3D space (normally described as *x, y, z* Cartesian co-ordinates) and its angle rotation against these planes (*x', y', z'*). It does so through transmitting and receiving a signal using low energy Electromagnetic radiation . Its effective range (radius) is approximately 1 metre. The angle of rotation against the x plane, is called *pitch*, the rotation against the y plane *yaw*, and the rotation against the z plane, *roll*. See Figure IV. This acts as a generator of co-ordinates, which acts as a reference to each stored frame.

2.2.2 SGI Indy/Indigo. Silicon Graphics Workstation.

The computer is equipped to the following specifications;

An ATM card, which is essential to allow communication, and although 96 Mbytes RAM is available, which could be used to cache recent frames, currently only 10 Mbytes of RAM are in use. Stripped 1OGb hard disks are used for storage, with a 10 kbyte frame size allowing for 1 million frames on disk. Using this system, there is NO use of the computer's graphics hardware.

2.2.3 The ATM: CellStack™.

This is another vital part of the technology, being necessary as a bidirectional TV to ATM converter. It can convert Digital Video, (jpeg compressed) 768x524 pixels, to High Definition TV, with CD/Dat stereo audio, and with RS232 serial I/O, a function which would not presently be possible using the Ethernet card system.

2.2.4 ATM: Network.

A 16 port Switch system is used for *Guaranteed Bandwidth* information transfer at a speed of 155 Mbit both ways, receiving inputs from Fibre or UTP cables, allowing for good integration with **IP** and even the option for use as a meter, recording the volume of data transfer for possible charging. ATM programming consists of relatively simple streams, like TCP, but detailed description is beyond the scope of this paper. Client/Server Video Worlds is a key concept here, in that by allowing for remote access, distance learning techniques become possible, and in addition, Remote Collaboration with Video Conferencing techniques using the same technology are possible. It will be essential to arrange multi-role applications to ensure a shared load and cost for the multimedia laboratories that are likely to be the initial venues for this resource.

2.2.5 Hardware - Biological

For the experiments, freshly prepared lamb or pork shoulders were used. These were mounted in a vice, so that access to the capsule over the portal was readily available. This acts as the fulcrum about which the arthroscope moves and thus is a potential source of error for the bird co-ordinates, which are reciprocated from the other end of the arthroscope. The specimens can be positioned to allow some distraction of the joint space, mimicking the clinical procedure, otherwise the architecture is unaltered, and irrigation of the joint cavity (necessary clinically) was not used in the experiments.

3. Discussion:

If we relate the temporal component to these modalities, then we produce the resultant table which considers the types of sequenced images used to create virtual realities;

Input (Recording)	*Sequence of data collection vital*	*Real Time Image Resolution*	*Frame Dependant*	*Interactivity*
Film (Traditional, ie.pre-production)	No	Excellent	Yes	None
Cartoon	No	Good	Yes	None
Video World	No	Good	Yes	Limited
Graphic World	No	Poor	No	Good
Output (Replay)	*Sequence of data collection vital*	*Real Time Image Resolution*	*Frame Dependant*	*Interactivity*
Film (Traditional, ie.pre-production)	Yes	Excellent	Yes	Passive
Cartoon	Yes	Good	Yes	Passive
Video World	No	Good	Yes	Limited
Graphic World	No	Poor	No	Good

It must be emphasised that **film** is considered in the traditional sense, and development of image manipulation in the editing of film sequences will certainly have an impact upon frame dependency, as individual frames can be modified (The Dinosaurs in Jurassic park weren't real !). Though this is prohibitively expensive for the video surgical simulator situation at present, it is significant that a major market for the high speed networking equipment (ATM) is in fact, the post-production film industry. The relationship between the technology and the application is perhaps best illustrated in the table below; when we consider the practicality of the tasks needed in a purely video world system for simulation;

	Training Procedure	*Video Worlds - (complexity rating *=simple, *****= complex due to increased number of required frames)*
1	Steering in an 'unfamiliar' environment	*
2	Navigation	*
3	Pattern Recognition	**
4	Triangulation	**
5	Pathology Simulation	***
6	Procedural training	*****

This suggests that a hybrid Video graphics world may provide the optimum system in order to allow for interactivity as well as reducing computational demands, see figure V. Ultimately, when the computer resources become affordable, then this system should still have a valid role on the 'input' side of the equation, in that it allows for the rapid collection of video data with Cartesian co-ordinates in a 3D environment, that could then be converted to pixel data with Artificial Intelligence techniques.

4. Conclusion

Capturing useful video worlds is possible, but further development of the capture system is likely to require a degree of automation. Further evaluation is needed.

The current system can be run at 20 frames/sec., thus the illusion of REAL TIME playback is possible. Using ATM + the K-Net video codec increased throughput 10 fold, and it is likely that this will be the standard vehicle for the technological process. Future work will involve taking the technique into the clinical environment and the development of integrated computer graphics for the development of a more interactive system with pathology incorporated.

It is essential that such a technology finds a niche within the rapidly developing framework of the Computer based distance learning strategies of the Medical Educational Establishment (Pinciroli 1995), and thus development alongside the current medical educational infrastructure reforms ((Solomon 1992, Thompson1994) will be necessary. Integration with other input and output systems in both the 'world building' development stage and training application stage will be necessary, e.g. VR headsets, 2 channel stereo vision, tactile feedback systems and of course a larger VR database of worlds, must be developed.

Possible remote access systems will be explored, including H.261 on Superjanet based systems. It should be emphasised that at present this system is faster and cheaper, than graphical based systems, producing images of a higher fidelity. It has the potential to be integrated with graphical systems to greatly enhance their performance, and in the future could have a significant role to play not only as a 'stand alone' system but also being used for accurate rendering of graphics based worlds.

5. References:

1. Bunker TD. Wallace WA, editor.Shoulder Arthroscopy. London: Dunitz; 1991;Shoulder Arthroscopy.

2. Dumay A.C.M., Jense G.J. Endoscopic Surgery Simulation in a virtual environment. Computers in Biology and Medicine 1995;2(25):139-48.

3. Cooper J., Ford L., Watson G. A Training potential for Non-invasive surgery. [Abstract] ftp://ftp dcs ex ac uk/pub/usr/lindsey/garth ps z 1995;1-21.

4. Ota D., Loftin B., Saito T., Lea R., Keller J. Virtual reality in surgical education. Computers in Biology and Medicine 1995;2(25):127-37.

5. Pinciroli Francesco and Valenza Paolo. An inventory of computer resources for the Medical application of virtual reality. Computers in Biology and Medicine 1995;25(2):115-25.

6. Richardson M. On the leading edge, Texas researchers are developing the next generation of medical technology. Texas Medicine 1994;8(90):12-6.

7. Satava RM. Virtual reality Surgical simulator. The first steps. Surgical Endoscopy 1993;3(7):203-5.

8. Satava RM. Emerging medical applications of virtual reality:a surgeon's perspective. Artificial Intelligence in Medicine 1996;4(6):281-8.

9. Solomon DJ, Osuch JR., Anderson K., Babel J., Gruenberg J., Kisala J., Milroy M., Stawski W. A pilot study of the relationship between experts's ratings and scores generated by the NBME's Computer based examinationsystem. Academic medicine 1992;2(67):130-2.

10. Thompson AR., Wilton PB., Scott-Conner CE, Hall TJ., Anglin BL., Muakkassa FF., Poole GV. The integration of laparoscopy into a surgical residency and its implications for the training environment. Surgical Endoscopy 1994;9(8):1054-7.

11. Wind G., Dvorak VK., Dvorak JA. Computer graphic modelling in surgery. Orthopaedic Clinics of North America 1986;4(17):657-68.

12. Ziegler Rolf, Fischer Georg, Muller Wolfgang, and Gobel Martin. Virtual Reality Arthroscopy training simulator. Computers in Biology and Medicine 1995;25(2):193-203.

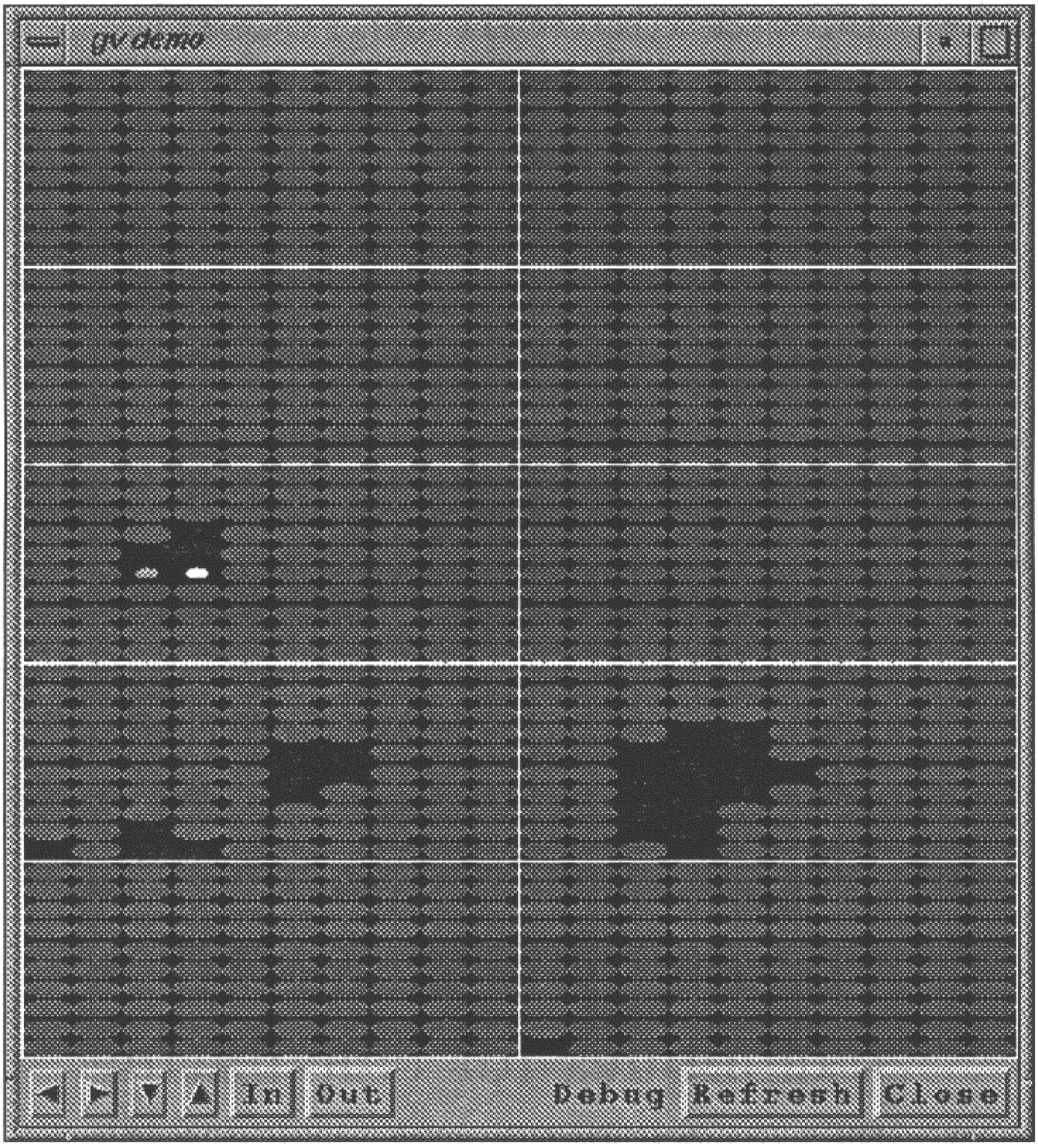

Figure I.

*The **gvo** program displaying a 10x10x10 three dimensional grid of the recording area. As the surgeon records, circles are eliminated from the display, indicating that adequate frames have been recorded in the region.*

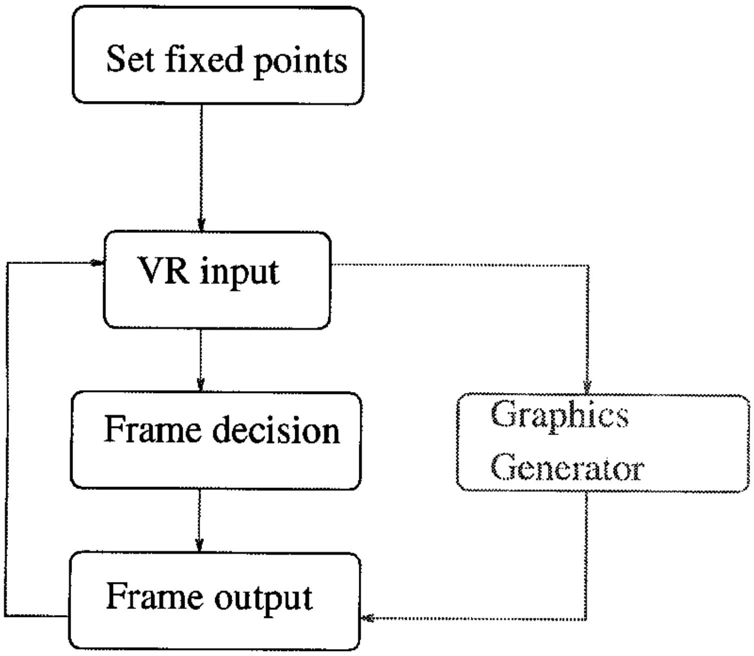

Figure II.

During playback, the operation loop "VR input", "Frame decision", "Frame output" is repeatedly cycled. This is begun by correlating fixed points and could be implemented with graphics generation running in parallel.

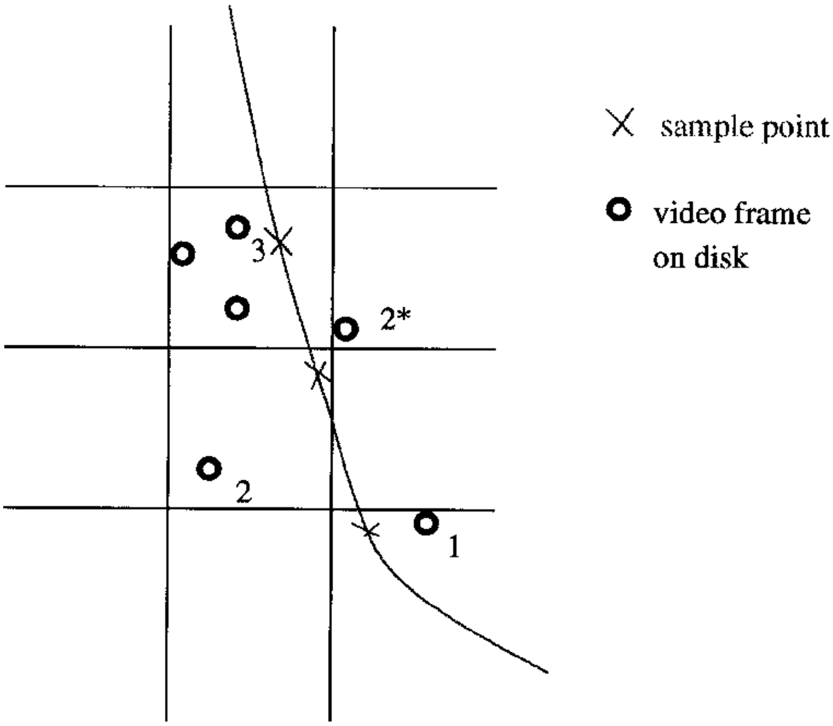

Figure III.

The frame decision algorithm reported in {Cooper J. 1995} could lead to errors when frame 2 is incorrectly chosen instead of frame 2.*

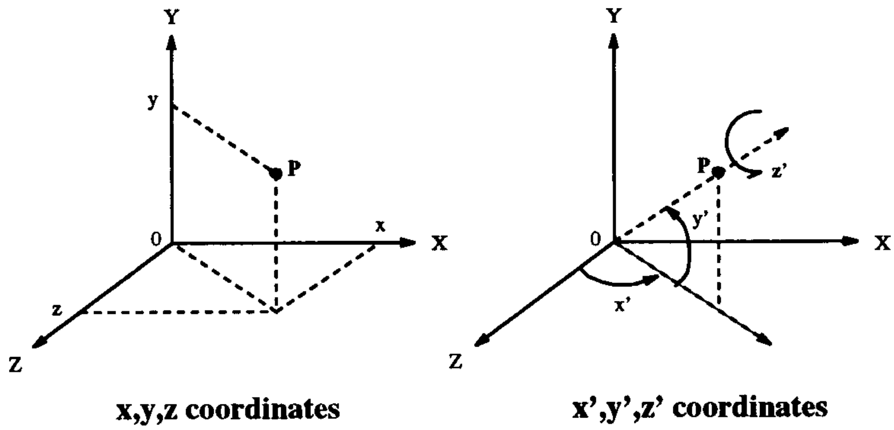

Figure IV.

The two types of co-ordinates give 6 degrees of freedom for a surgical implement.

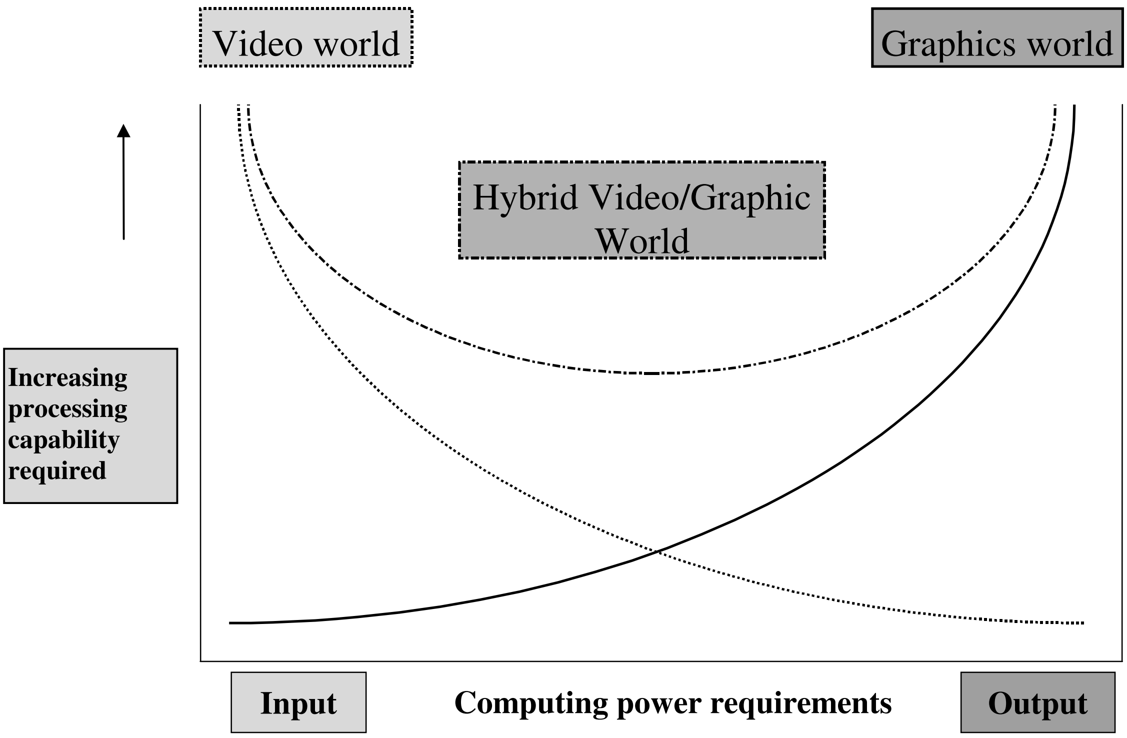

Figure V.

Graph to visualise the concept of a hybrid video and graphics combined 'world' and the anticipated saving of computing resources.