

Building immersive environments using Quicktime VR: Lessons from the real world and virtual realities

Cameron Barnes,
Dixson Library,
University of New England
cbarnes@metz.une.edu.au

Abstract

The paper discusses techniques to create immersive and compelling virtual worlds using Quicktime VR (QTVR). Attention is given to the educational potential for walk-through simulations of historic environments (including buildings and archaeological sites) through linked QTVR panoramas. This potential is discussed in the context of constructivist views of education. It is argued that QTVR panoramas can be used to stimulate situated cognition and provide rich, authentic learning environments. In addition, QTVR is ideally suited to problem-based learning approaches. However, achieving the full educational potential of QTVR requires solutions to new technological challenges as well as innovative pedagogy. Many of these solutions involve better photographic practice. Others relate to the deliberate manipulation of raw images in order to accentuate visual depth cues. It is argued that the creators of QTVR worlds can learn from the techniques developed by game designers to exploit human factors in depth perception. The paper concludes with a short consideration of future developments relating to the delivery of QTVR virtual environments over the Web.

The educational potential of QTVR

Quicktime Virtual Reality (QTVR) is a learning tool with enormous promise. The arrival of Quicktime 4.0 in 1999 provided course developers with a stable, cross-platform medium for the Web-based delivery of virtual worlds. Despite this step forward, realising the educational potential of QTVR has proven to be an elusive goal. Experience has indicated that, if QTVR is to be used to its full advantage, developers must adopt more thoughtful instructional and technical practices.

QTVR provides a means of creating and viewing photographic panoramas. Using a QTVR-equipped browser, learners can pan and zoom around a 360° virtual environment. By means of one or more hot spots within a panorama, learners can also jump to other QTVR scenes (known as nodes). Adding a series of nodes together, course developers are able to create a virtual world that can be explored over the Web. Further, a virtual QTVR world can be made with a minimum of cost. QTVR nodes can be designed and published on the Web in days, instead of the weeks or months required for the painstaking construction of computer-generated 3D buildings and landscapes.

Much of the educational importance of QTVR lies in its constructivist potential. Recent discussions of educational multimedia have stressed:

the need for open-ended exploratory authentic learning environments in which learners can develop personally meaningful and transferable knowledge and understanding
(Squires and Preece 1999: 469).

It will be argued in this paper that QTVR can be used to meet this need. However, achieving constructivist goals through QTVR requires course designers to think carefully about their motives for adopting this new technology as a learning tool.

The use of QTVR to build walk-through simulations of historic buildings and archaeological sites is often regarded almost as an end in its own sake. If pressed for a justification, course designers sometimes respond in terms of the assumed power of novelty to stimulate the learner's interest.

QTVR, it is asserted, provides learners with ‘a magic carpet’ with which to explore distant environments. Another frequent response from course designers is that QTVR provides learners with access to information difficult or impossible to obtain through video, still images or text.

These claims are not in themselves incorrect, but they ultimately sell the medium and its users short. Poorly employed, the capabilities of QTVR are educationally trivial. If the hours involved in creating QTVR environments are not to be wasted, course developers need to move away from a naïve confidence in the value of innovation for its own sake.

Evidence suggests that the initial advantages of new technologies in terms of learner motivation quickly fade (Najjar 1998: 316). As the history of the educational use of hypertext/hypermedia indicates, learner control over navigation within a computer micro-environment on its own is insufficient to ensure effective learning (Tergan 1997: 276). To make the best use of QTVR virtual worlds, learners must be provided with appropriate research tasks, guidance and opportunities to test their understanding. In addition, course designers need to appreciate the special requirements of the medium.

Constructivist educators argue that learning is (or should be) a personal process. Their ideal is a learning environment within which individuals can develop their own knowledge by constructing and progressively refining concepts about the world. Environments facilitating this form of learning are regarded as cognitively authentic. Squires and Preece observe that the notion of cognitive authenticity has led many educational theorists to identify three critical factors necessary for personally meaningful learning. These are credibility, complexity and ownership (1999: 469).

The strength of QTVR in terms of credibility is relatively uncontroversial. In most instances, the source materials are actual photographs of historic or archaeological sites. The user experience is therefore fundamentally different to that deliverable through computer-generated 3D environments. Unlike simulated worlds built using computer-generated imagery, QTVR panoramas are inherently real and believable.

QTVR virtual worlds are equally authentic in terms of complexity. The level of detail and realism included in a QTVR node far exceeds that achievable through other means (at least within the context of educational budgets). Although static in a sense that each panorama is frozen in time, the environments captured through QTVR are observably the result of complex events. These include weathering, human action (such as successive rebuildings at an historical site) and the longer term processes of growth and decay. Course developers can build on this complexity. If they choose, they can devise complex tasks, which assist learners to acquire the skills required to sensitively observe and analyse historical environments.

In terms of ownership, a similar point can be made. QTVR environments enable the learner to interact with and control the flow of information. It is the learner’s choice whether to pan to the right or the left, to look up or down, zoom in or out, or go through one door and not the other. The learner also has the freedom to go back to a visited location and re-examine features missed during an earlier excursion. There is good reason to believe that this level of control will encourage learners to process the rich information provided by QTVR worlds, and assist them to recall their new knowledge (Najjar 1998; Stemler 1997). More importantly, as the learner alone can see what he or she sees (or hears, as QTVR supports sound as well as vision), the individual learner is autonomous, engaged in a personal journey.

The concept of situated learning or situation cognition is another constructivist notion of crucial importance to the educational use of QTVR. According to theorists such as Jonassen (1994), situated learning occurs only when students work on authentic and realistic tasks, those reflecting real world contexts. Other forms of learning are said to result in decontextualised knowledge. Jonassen asserts that such decontextualised knowledge is ‘inert’ because the student has no realistic context for its employment (Jonassen, Campbell and Davidson 1994: 36).

This view provides a powerful argument for the educational use of QTVR worlds. The realistic nature of QTVR environments means that knowledge gained through their exploration is clearly situated in the real world. Learners can interrogate a QTVR world in much the same way as archaeologists or historians study a specific

location in the physical world. This advantage is lost when learners are presented, say, with a still photograph of a medieval fortification, or a schematic diagram depicting an archaeological cross section.

The potential of QTVR worlds in constructivist terms is realisable only if course developers make a sustained effort to exploit this promise. The course developer must learn to ask open-ended questions, rather than simply state the observable 'facts'. Learners must be encouraged to interrogate the virtual world in an individual manner and to take responsibility for their own learning. The complexity of a QTVR world makes it ideal for problem-based learning. This form of teaching, however, places much higher demands on the ingenuity, time and patience of course developers.

Moreover, asking a class to explore a virtual world is quite different to telling them to read a particular chapter in a textbook. In dealing with the complexity of QTVR environments, learners will often require a good deal of help. The implication is that course developers must work harder to provide mental scaffolding, anchors to prior knowledge and appropriate feedback.

In compensation, the richness of QTVR environments makes it much more likely that learners will revisit these virtual worlds. Comparatively few images in a textbook receive (or deserve) repeated study. However, the potential exists in QTVR for developers to encourage learners to retrace their steps.

Learners can be prompted to observe different features of the virtual environment, to formulate additional hypotheses based on fresh information and new questions.

Building a QTVR node

In order to examine the technical strengths and weaknesses of QTVR, it is necessary to understand the method used to build a panorama. A basic QTVR node is surprisingly simple to create. Such panoramas are typically made by taking a 35 mm film or a digital camera on location and shooting a series of pictures. After each shot, the camera is rotated a fixed number of degrees, so that a series of overlapping exposures of the scene are produced. The QTVR authoring software is then used to join the resulting photographs into a single panorama. When the QTVR movie is viewed on the Web, the user has the illusion of being within a cylindrical panorama, one which is progressively revealed as the user pans to the right or left.

Quality problems with QTVR nodes

The relative ease with which a QTVR node can be fashioned is also one of the greatest obstacles to the effective instructional use of this technology. It is possible to take a number of pictures by hand using a low-resolution digital camera and produce a visually stunning panorama. QTVR nodes created in this manner can deliver a range of detailed visual information. However, such nodes are almost invariably unsatisfactory in terms of realism. As the user rotates his or her viewpoint, objects, buildings and landmarks appear to jump or slide into view. The underlying geometry of the virtual world is misshapen, being based on a series of flat planes rather than a smooth cylinder. In consequence, the illusion of a virtual world is lost. Instead of feeling as if he or she is present at a distant location, the user is reminded that the situation is counterfeit. In an educational context, this response is a potential distraction and threatens to interfere with the learning process.

Achieving a higher level of realism

Too often, the assumption seems to be that such geometric distortions are inevitable. In part, this is due to the multiplication of poorly created QTVR nodes on the Web. The second reason is a widespread lack of technical knowledge.

The significance of parallax error

The main cause of distortions within QTVR panoramas is parallax error. To avoid parallax error, it is essential to rotate the camera on the nodal point of the lens. The camera must therefore rotate off-centre. In addition, the film plane must be perpendicular to the plane of rotation. If not, each exposure will consist of a separate cone rather than segments of a continuous cylinder.

The significance of parallax error is sometimes difficult to grasp. The difference between the nodal point and the film plane is a few centimetres. For those without technical training, it is easy to assume that these few centimetres are relatively unimportant.

The *QuickTime VR Authoring Studio User's Manual* states that the best results are achieved when the camera is rotated over the nodal point. However, the manual fails to explain the significance of this advice (Apple Computer 1997: 80). For this reason, the warning is often missed or misunderstood.

Despite this problem, it is relatively easy to demonstrate the importance of parallax error. The following experiment indicates this point:

With one eye closed, hold your hand at arms length with your thumb pointing upward. Position your arm so that your thumb is between your eye and the vertical portion of a doorframe. If you are doing this correctly, your thumb should be obscuring part of the doorframe. Now turn your head left and right. Notice how the doorframe becomes visible as you move your head a bit left or a bit right? This is an example of image parallax due to the fact that the nodal point of your eye is not over the centre of your head rotation.

You can do this same experiment with a camera and attached lens. Using a standard 35mm camera, pivot the camera left and right over its tripod mount and you will notice the same effect of the far image becoming visible as you get a bit left or right of where you started. If you pivot the camera at a certain point under the lens, you will notice that the far image remains obscured. This special point of rotation is called the optical centre, or nodal point, of the lens (Concepts in Motion 1997).

Overcoming parallax error and other common problems

Although calculation of the nodal point is somewhat complicated, the formula is readily available on the Web (Jacobson 1997). Armed with this information, and some basic equipment (including camera leveller, camera-L bracket, tripod, click-stop pan head), it is quite possible to eliminate or minimise parallax error.

Expert technical advice can help even the amateur photographer to overcome other common difficulties. For example, a professional photographer might advise on the use of a shift lens to correct converging verticals when photographing tall buildings. Another potentially useful piece of advice includes the simple recommendation to change from the Apple recommended 15 mm lens to a 28 mm lens when shooting an outside panorama. This will avoid the situation where everything more than 3 metres away from the nodal point appears to the user as though on the edge of the observable horizon.

Creating the illusion of depth with QTVR

There is a common misconception among beginners that QTVR nodes cannot provide a convincing illusion of depth. One explanation for this mistake is that most people have little or no understanding of how they and other humans perceive visual depth. This leads to an overestimation of the obstacles to creation of more realistic QTVR worlds. Although QTVR is not a true 3D medium (at least at present), the depth information in a well-designed QTVR panorama is considerable. It is at least comparable to that in best-selling 3D computer games such as *Tomb Raider III* or *Quake Arena*. Such games have proved themselves extremely immersive.

The importance of depth cues

To appreciate the extent of the depth cues in a QTVR node (and the manner in which they can be enhanced) it is vital to consider the manner in which humans see the real world. The visual information available to our brains consists of the flat images on our retinas.

The fact that humans see their environment in depth is not solely the result of binocular depth perception (stereopsis). It is also due to the neurological processing of a range of subtle depth cues. Artists have known for centuries that careful manipulation of such cues can create a convincing illusion of depth, as witness the success of *trompe-l'oeil* paintings in the European past (Gregory 1997: 28).

Studies of human visual perception have identified a range of pictorial depth cues, which can be used to generate the appearance of depth on a flat plane. Textbook authors such as Coren, Ward and Enns (1994: 329–37) frequently provide lists of these. The range of known depth cues includes the following:

- *interposition* or *occlusion* (objects which obscure others are assumed to be closer);
- *attached shadows* (these are taken to indicate the relative distance of one object from another);
- *cast shadows* (these are used as a guide to the intrinsic shape of an object);
- *relative size* (smaller objects are understood as more distant);
- *linear perspective* (the convergence of receding parallel lines and planes);
- *relative brightness* (the brighter of two objects is interpreted as closer);
- *relative height* (proximity to the horizontal plane is perceived as an indication of greater distance);
- *texture gradients* (the smaller and more tightly packed the elements of a texture appear, the more distant they are perceived to be); and
- *attenuation* (the distance of objects from a central light source is inferred from their relative lightness or darkness).

Lessons from computer game design

In attempting to develop effective virtual worlds using QTVR, educators can learn a good deal from the computer game industry. The designers of successful 3D computer games have achieved the creation of immersive experiences through the systematic manipulation of pictorial depth cues.

Game developers place great importance on the achievement of a consistent surface geometry in their virtual worlds. Experience has shown that such geometries need not correspond exactly with those of the real world. ID Software's *Doom* provides an excellent example. Although the developers incorporated cues such as perspective, relative size and relative height into their design, the geometry of the game world was never intended to be more than approximate. *Doom* is crude by today's standards, but at the time of its release it was tremendously effective and popular.

Rich, detailed surfaces are often used by game developers to enhance the illusion of depth and realism. This is despite the fact that most game textures are no more than two-dimensional 'skins' or 'bump maps' stretched over irregular polygons.

The lesson for the developer of QTVR nodes is that it is essential to keep the level of detail as high as possible. Higher resolutions will mean larger file sizes, but they will also help to retain subtle depth cues. These cues include the effect of perspective on smaller surface features and texture gradients.

Game design has shown that the creators of QTVR panoramas should not be afraid to improve on nature. Programs such as Adobe Photoshop are likely to prove essential tools for the development of effective QTVR panoramas. Photoshop can be used to extend the range of tones and colours in a photograph, to sharpen shadows and to accentuate highlights. When used cautiously, these changes can heighten the effectiveness of common depth cues such as shadows and attenuation.

Another important lesson from the world of computer game design is the value of including floors and ceilings in a QTVR panorama. The ability to look up and down (as well as from side to side) has greatly enhanced the realism of the latest 3D games. Achieving this goal in the creation of a QTVR node is not difficult. All that is needed is to fix the camera vertically rather than horizontally and to use a wide-angle lens when taking the source photographs. The inclusion of lines and planes on floors and ceilings within the field of view will also help to exaggerate linear perspective, accentuating the illusion of depth.

Moreover, fixing the camera vertically allows the QTVR designer to make greater use of relative height cues. Using a low viewpoint, including a foreground and placing a subject (such as a standing stone) high within the panorama will provide the learner with an enhanced impression of depth.

Employment of ambient sound will assist in the production of more realistic QTVR nodes. Background sounds have proven an effective tool in the establishment of mood or atmosphere within a game. When shooting

photographs for a Quicktime node, it is possible to make a stereo recording of the ambient noise. After editing and mixing, this sound can be added to the finished node to enhance the illusion of reality. Furthermore, sound can be used in other ways. The transition from an outside to an inside node can be made more credible if outside sounds are replaced by the muffled sounds common to indoor environments.

Another lesson from game design is that QTVR creators should carefully plan their nodes. Few factors disrupt a game player's absorption as much as a sudden inconsistency, which reminds him or her of the fictional nature of the game environment. Poor programming means that it is common for players to experience mounting frustration as they attempt to reach inexplicably unattainable locations (such as ledges or wall niches) which appear easily accessible on their screen.

This principle is clearly relevant to the designers of QTVR simulations of historic buildings. If there are two doors out of a room, there should be two hot spots linking the learner to new nodes. If windows are open, users should be able to look out through these windows. If it is not feasible to create additional nodes, then doors and windows should be closed or otherwise indicated as out of bounds. Roping off a doorway is a well-established convention in museums and historic sites. It should cause learners no problems in the context of say, a QTVR simulation of an historic building.

When the first text-based adventures were developed in the 1970s, players had to endure an annoying phenomenon. Whenever they entered a room in a dungeon or cave complex, they received the same stereotyped description of the room's contents. This was despite the fact that most rooms had different entrances and exits and were visited many times during a single game.

The lesson that future QTVR developers should draw from the experience of past players is that careful attention should be given to the point at which a node begins to play. When a learner enters a node from one direction, the node should play from the appropriate point. It is undesirable for a learner to be presented with the view from the southern end of a room if he or she has entered from a door to the north.

In contrast to the latest 3D games, QTVR nodes constrain the user to a single point. This fact need not interfere with the creation of an immersive experience. The success of the game *Myst* has shown that users are quite capable of inferring movement from a succession of still images. This principle should be applicable to the QTVR environment, where the user has a wide degree of control over his or her visual field.

The future of QTVR on the Web

At present, the potential for the delivery of QTVR virtual worlds over the Web is constrained by a number of technical factors. These include playback speeds, bandwidth limits and the slower processors of older computers. In addition, existing video compression/decompression schemes (codecs) generally require severe trade-offs between playback speed and visual quality (Niederst 1999: 350–51). These conditions inhibit the wider adoption of QTVR as an educational tool for the creation of Web-based virtual worlds.

The existing situation can only be temporary. Desktop processors are doubling in speed every 18 months or so. The price of Internet access is falling in real terms while bandwidth increases.

New codecs continue to be released. On past evidence, these factors will inevitably lead to improvements in visual quality and better compression ratios. Hence, there is every reason to believe that the creation and online delivery of immersive worlds through QTVR will soon become far more common. The challenge for educators is to ensure that the best use will be made of this extremely powerful learning tool.

References

- Apple Computer, 1997 *QuickTime VR Authoring Studio: User's Manual*, Apple Computer, Cupertino, California.
- Concepts in Motion, 1997 (accessed 22 October 1999), 'Tips and techniques', *Outside the Lines*, <http://www.outsidethelines.com/Tips.html>.
- Coren, S., Ward, L.M. and Enns, J.T. 1994, *Sensation and Perception*, 4th edn, Harcourt Brace, Forth Worth, Texas.

- Gregory, R. 1997, *Mirrors in Mind*, Penguin, London.
- Jacobson, D. 1997 (accessed 22 October 1999), 'Lens tutorial', *photo.net*, <http://photo.net/photo/optics/lensTutorial.html>.
- Jonassen, D.H. 1994, 'Thinking technology: Towards a Constructivist design model', *Educational Technology*, vol. 33, no. 4, pp. 34–37.
- Jonassen, D.H., Campbell, J.P. and Davidson, M.E. 1994, 'Learning with media: Restructing the debate', *Educational Technology, Research and Development*, vol. 42, no. 2, pp. 31–39.
- Najjar, L.J. 1998, 'Principles of educational multimedia user interface design', *Human Factors*, vol. 40, no. 2, pp. 511–523.
- Niederst, J. 1999, *Web Design in a Nutshell*, O'Reilly, Sebastopol, California.
- Squires, D. and Preece, J. 1999 'Predicting quality in educational software: Evaluating for learning, usability and the synergy between them', *Interacting with Computers*, vol. 11, pp. 467–483.
- Stemler, L. K. 1997, 'Educational characteristics of multimedia: A literature review', *Journal of Educational Multimedia and Hypermedia*, vol. 6, nos. 3–4, pp. 339–359.
- Tergan, S. 1997, 'Multiple view, contexts and symbol systems in learning with hypertext/hypermedia: A critical view of research', *Educational Technology*, July–August, pp. 5–8.