HDTV Distance Learning

Carlos R. Morales
Purdue University

Abstract

While High Definition (HD) has gained market penetration in the entertainment market, it has not been implemented in any significant scale in the distance learning arena. In this paper, the author discusses a production process used to create over ninety hours of HD quality distance-learning content. The production process is noteworthy because it can achieve the production of HD content at a cost comparable to that of producing standard definition video. The author also discusses techniques for maintaining high quality in both the instructional design and the video assets.

Introduction

Purdue University offers degree-seeking programs at its main campus in West Lafayette, Indiana and at numerous smaller regional campuses. In June of 2003, members of the Computer Graphics Technology department at Purdue University's School of Technology decided to produce three full courses of instructional videos as part of a distance-learning initiative. To fully utilize the power of the video medium and the added resolution of HD, the team decided to use advanced computer graphics compositing techniques and virtual set technology.

While video is one the most powerful communication mediums, most distance-learning initiatives do not maximize the instructional power of the medium. The most effective use of instructional video comes from showing animations and moving illustrations of the concepts presented while the instructor is conducting the lecture [1]. Most distance learning initiatives, whether synchronous or asynchronous, utilize a combination of video of the instructor with synchronized presentations of PowerPoint slides. Dastbaz found that this is not very effective or efficient [2]. This ineffectiveness holds even when following recommendations for using PowerPoint in the classroom, such as those provided by Howell [3]. Creating effective distance-learning instruction carries with it the same requirements as creating high-quality instructional television shows such as those we see on NOVA or the Discovery Channel. This level of production is rarely seen in distance learning because of its high cost and lengthy production times.

In television production, the audience is typically much larger than in a distance-learning scenario. The net effect is that with television production, we can achieve a relatively low cost per unit, even if the total production cost is high. The cost of distance learning typically cannot be amortized to this extent. Television producers employ the help of editors, animators, scriptwriters, and compositors to produce content. Instructional designers, on the other hand, are primarily taught to systematically structure instruction. They are typically skilled at conducting needs assessments, defining objectives, writing assessment items, and a plethora of other instructional design related tasks in accordance with an instructional model.

Production

To produce the instructional materials at an instructional and aesthetic level that maximized instructional effectiveness, we needed to find a way to structure instruction in accordance with the cannons of instructional design, and produce the content in a manner consistent with television production at a relatively a low cost.

The total budget for the project was fifteen thousand dollars and was assigned to a group of six people. The project manager had a background in instructional design and video production. While the instructors that worked with us were experts in their field, many did not have formal instructional design experience. The project manager assisted the instructors to create an instructional map that served as a basis for the television style script that was used during the production phase of the project. Three members of the team were assigned to video compositing. Their primary responsibility was to use compositing tools such as Adobe After Effects or Discreet Combustion to create two dimensional animations.

To orchestrate the interaction of the media assets, we created a television style script. The primary purpose of this script was to guide the media creation efforts of the production team while minimizing the amount of instructor intervention. We filmed the instructor while he lectured in a traditional classroom. From this video we created a script of the instructor's lesson by using speech recognition software. The project manager identified instructional objectives, media requirements, and questions asked by the instructor which could be turned into embedded assessment items.

The script was emailed to the instructor for approval. After iterative changes were made and final approval from the instructor was received, the script was distributed to the animators and compositors, who started to create media assets. PowerPoint slides of the instructor's notes and illustrations were then created in anticipation of the production process. These slides would be used as place holders during the production phase and replaced with broadcast quality graphics during the post-production phase.

With the script and slides prepared, we proceeded to film the video of the instructor that would be used in the final product. We shot the videos in an office that we converted into a small studio. The entire room was painted or covered with blue material to allow us to replace the instructor's environment. Using the chroma-keying process, we removed the blue environment around the instructor and replaced it with computer-generated graphics. The net result was that we could make the instructor appear to be in any location that we selected. We could also make it appear that the instructor was interacting with his environment.

We shot the videos using a JVC GR-HD1, because it allowed us to capture High Definition (HD) video relatively inexpensively. A benefit of this camera is that for less than $3500.00 it can capture video at a resolution of 1280x780 pixels. Standard television is typically digitized at 720x480 or 640x480. Thus, with this camera we could achieve a considerable increase in resolution in our content at a relatively low price point.

The camera also had its disadvantages. It captures video using a customized version of MPEG-II at approximately 20 Megabits per second and sampled at 4:1:1. The chrominance portion of the video signal was getting sampled at half of the frequency of the luminance. From these specifications, we expected our chroma keying process to be more difficult than if we would have used a format such as Digital BetaCam or DVCPRO50. We found that by using software that up samples the video to 4:2:2 before producing the mattes, we were able to achieve respectable results.

This scenario generated two video files. One video file was of the instructor in the blue environment. The second file was of PowerPoint slides and software demonstrations captured by Camtasia from the instructor's computer. To facilitate editing, we shot everything in a single take. Thus, the two videos were synchronized in time. This yielded tremendous savings when it came to editing the content. The media producers could reference the live video to see the actions of the instructor while observing any drawings or illustrations created by the instructor by looking at the Camtasia video. The net result was the media creators seldom had to consult the instructor after the video session.

Compositing and Editing

The editing and compositing phase of this project turned out to be the largest challenge in terms of expenses for labor and equipment. Typically, distance-learning initiatives do not spend resources compositing computer generated images (CGI) with live video because it takes a relatively large team of people with specialized skills and equipment. We needed to find a way to use this technology to produce 90-hours of finished lectures in ninety days.

We had the correct balance of skills in the personnel involved in the project, but we lacked the equipment to allow our animators and compositors to work at a rate fast enough to meet our deadline. We had access to real-time editing packages, including Adobe Premiere Pro, Sonic Foundry's Vegas Video, and Apple's Final Cut. But, these packages did not offer the capabilities that we needed for the project. Real-Time editors excel at working with long format videos with a relatively few layers. We needed something that would allow us to stack a dozen layers of video in real-time.

Before looking at some of the hardware assisted compositors, we did a few tests to gauge our production times. Our test involved compositing a 10-minute segment in the software packages that we had available, which included Final Cut, Premiere, Vegas Video, and After Effects. The shortest production time that we could achieve was fifty hours to composite our ten minute test. This numbers were unacceptable.

The main thing that slowed our production pipeline was synchronizing the video layers with the audio of the instructor. Every time the animators tried to preview his work, the computer loaded up those layers and rendered them into RAM. This took an incredible amount of seat time for our compositors. We had no choice but to consider using a hardware assisted solution.

We first considered tying our production efforts into the University's video efforts because it allowed us to consider hardware solutions that were beyond our budgetary constraints, such as those offered by Discreet or Quantel. Both of these companies offer solutions that can handle tons of layers of at up to a resolution of 2048x1556 pixels in real-time. So how many layers is a "tons of layers?" Well, it depends on the effects applied. In something simple, 30 to 40 layers are not uncommon. On the low-end solutions in this strata cost approximately one hundred thousand dollars. On the high-end, the prices approached one million dollars. Thus, we could not proceed without external budgetary help.

Next, we considered using a hardware assisted hybrid editor/compositor from Media 100 called 844/X. This device can play eight layers of video in real-time. Using advanced render caching techniques, it can give the operator the illusion that it can render many more layers simultaneously. The final price on this solution ranged from twenty to sixty thousand dollars. Again, this option was not available to us without external budgetary assistance. If we had the funds, this would have been the solution we would have selected.

Unable to get a turn-key hardware compositor, we turned our attention to reducing the bottlenecks that plague software solutions. Video compositors are expensive because they have to move lots of data and perform image processing operations extremely fast. If we have a total of 10 layers of uncompressed video each at 720 x 480, our system would have to move 270 Megabytes per second just to load the data into RAM.

Our first idea was to find a way to deliver as much data to the CPU in our system in as little time as possible by using a very large RAM disk. Then we could use a tool like Vegas Video or After Effects and get a huge performance upgrade. We evaluated a Tyan Thunder GC-HE motherboard with 12 Gigabytes of RAM and 4 Intel Xeon processors. This motherboard can take up to 24 GB of RAM. We set up 10 GB as a RAM DISK and moved the video files onto it. The performance was astonishing. In After Effects and Combustion we almost achieved real-time results. The real boost in speed came when we used Vegas Video. We were able to throw over 50 layers at our system and it responded in real-time. Thus, for under $5000.00 we were getting the type of performance associated with a one-hundred thousand dollar system. We edited the video in eleven minute segments in order to get the video files to fit into the RAM disk.

We also considered using a real-time editing card such as those offered by Canopus and Matrox. The limitations with the cards we examined, Matrox RT.X100 and Canopus DV Storm 2 SE, is that they did not support real-time compositing effects at HD resolution and only accelerated specific software packages. Both cards accelerated Adobe Premiere and up sampled the video to 4:2:2 to perform hardware keying. This was significant because these cards could greatly reduce the cost of performing chroma keying operations, which were required in to create the illusion of a virtual set.

After evaluating these solutions, we decided to make some changes to our production pipeline to ease the requirements on our hardware. We also decided to employ both the RAM disk and a hardware accelerated card. By limiting the number of simultaneous layers on screen to a total of six, we were able to use a Matrox X.100 for most of the editing for our standard definition version of the lessons. For the HD version, we purchased Aspect HD. This product allowed us to manipulate up to 4 layers of HD quality video in real-time.

We did not go with the Tyan motherboard to save on some of our expenses. We found a Super Micro motherboard at one third of the cost which allowed us use up to 6 GB of RAM. Any effects that required more than six layers were composited in After Effects and Combustion on the machine with the 6GB of RAM. Those assets were then passed over to the editor working on the machine with the Matrox RTX.100 or the editor working with Aspect HD, but required rendering at the end of the process.

In the end, this solution allowed us to achieve the type real-time response often associated with systems costing hundreds of thousands of dollars for less than six thousand dollars. Figure 1 shows a sample frame from a lesson that utilizes virtual set technology, live video, and computer generated images to teach linear algebra concepts.

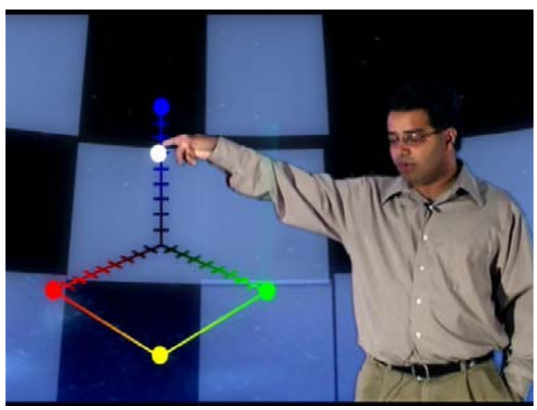

Figure 1. Sample Frame

Storage

Our network and storage infrastructure was not adequate for the project. For each hour of finished video, we produced approximately 30 GB, which was extremely slow to transfer over our 100-megabit network. We alleviated the problem by placing gigabit network cards into our production machines and assembling a homebrewed Network Attached Storage (NAS) device composed of twelve 250GB drives.

Typically our IT department does not use any home-brewed hardware because the cost of maintenance is typically higher than with turnkey solutions. In this case, we made a compelling enough argument to secure permission to build our own homebrewed NAS. Compared the cost of a turnkey 3 TB NAS, we saved over ten thousand dollars without sacrificing performance or ease of maintenance. The combination of the gigabit Ethernet solution and the homebrewed NAS yielded a reduction in idle for our animators by a factor of eleven.

Conclusion

This project demonstrates that using instructional design practices and high-end computer graphics compositing we can produce instructional content that is not only pedagogically effective, but also maximizes the power of the video medium. The cost of producing graphics-rich video can be reduced by utilizing a structured production pipeline and strategically utilizing the latest technological advances.

Based on our experiences, we would recommend the following to anyone undertaking a similar project: (1) Use a process that minimizes the time requirement from the instructor. Instructors will be more willing to participate if they do not have to crate media assets or have to spend time guiding the creation of those assets; (2) implement a process that insures an instructionally effective script. In our case, this involved videotaping the instructor twice and having an instructional designer review the script; (3) Use multi-layer video instead of three-dimensional graphics whenever possible. This will reduce the amount of production time; (4) strategically select and use contemporary video and computer peripherals. Many times current consumer versions of video gear can outperform and cost less than professional video gear produced a year or two ago.

In conclusion, by utilizing these techniques, it is possible to generate distance-learning content that is instructionally effective and maximizes the effectiveness of the video medium.

References

[1] Wetzel, C. D., Radtke, P.H., Stern, H., W. (1994). *Instructional effectiveness of video media*, (pp. 152-210). Mahwah: Lawrence Erlbaum Associates.

[2] Dastbaz, M. (2000). Hypermedia aided learning (HAL): a viewpoint on delivering education in the new millennium. *Proceedings of the 9th Annual Meeting of Information Visualization*, *9*, 44-47.

[3] Howell, D., Howell, D., Morrow, J. (2002). *Using powerpoint in the classroom*, (pp. 34-76). Thousand Oaks: Corwin Press.

[4] Johnson, D.F. (1996). Will my visuals work on my video distance learning system. *Journal of Interactive Instruction Development*, *9*, 22-25.

CARLOS R. MORALES

Carlos R. Morales is an assistant professor of computer graphics at Purdue University at Kokomo. He holds a BA in Telecommunications and an MS Ed. in Curriculum and Instruction. Prior to working at Purdue University, Carlos worked as a Technical Director. Some of his clients have included Microsoft, Chicago Bulls Organization, First Alert and Brach's Candies. His research interest includes distance learning, animation, and multimedia development. He can be reached at crmorales@tech.purdue.edu