

## Emerging User Interfaces

By Tim Siglin and Rosemary Roberts

Streaming media's continued maturity, with better compression and delivery tools, is paralleled at a slightly slower pace by the maturity of user interface designs. Driven by a desire to improve the immersive entertainment experience and a need to meet the customized requirements of particular verticals that are embracing the use of multiple simultaneous streams, companies are increasingly looking for intuitive ways to improve and enhance the user experience.

Three verticals and four interface types hold particular promise in 2005. Each interface type in each vertical, though, shares similar characteristics—simplicity of design, the need to simultaneously display many disparate types of information, intuitive navigation, and the ability to contextually replicate previous events.

Tomorrow's user interface also must be capable of delivering all information that streaming media capture software and hardware can provide, such as metadata about individual shots in an edited on-demand file, or global positioning information about where the live stream is being shot, which can then be tied to geographic information systems (GIS) for real-time graphical representation. This is especially important when, in the course of a lecture or keynote speech, the presenter references something locally that those physically attending the presentation can relate to, but whose meaning is lost on remote participants.

To properly assess the evolution of streaming media interfaces, one needs look no further than the closest Xbox. Today's digital gaming industry, always striving for the next level of immersion and interactivity to satisfy a demanding player base, may provide the type of "out of the box" thinking necessary to mate improved interface designs with streaming media's technology advancements. Just as a video game player has the capability of choosing the environment (type of race track or geographic location), model and color of his/her race car, power points (i.e. "turbo"), and conditions (crowded street, rain, etc.) for today's popular auto racing games, so too should near-term streaming media user interfaces provide contextually relevant environments (exhibit hall, classroom, or individual study cubical) from which to view live or on-demand content in the proper context or setting.

This article focuses on three primary markets for streaming media delivery: corporate, educational, and entertainment. Enhancements that push the interface envelope in any of these markets may in fact benefit all three, as end users in each of these verticals demand equal emphasis on immersive, passive, or interactive control of their user interfaces at any given moment within a single user session.

### **Entertainment**

As a market vertical, entertainment often is split into passive and interactive modes; as such, it provides a good platform to assess the leading edge of immersive entertainment.

While current software applications and capture hardware incorporate passive and/or interactive capabilities in limited forms—such as the current favorite “three-pane model”—tomorrow’s entertainment interfaces will look much more like a digital dashboard or military combat game screen, with 10-15 resizable panels located around the viewing screen to feature—among other things—graphics, text, and pertinent metadata about the stars, director, or production crew. Additionally, like the gaming industry’s use of multiple views in an aerial combat game, tomorrow’s streaming media interfaces will have the ability to set and recall multiple viewing preferences via a single click. These preferences would then recall particular settings if a pre-defined scenario were met: for instance, one could request a translucent window appear the first time each actress in a movie appears, stating both her name and her character’s name (which may mean, tangentially, that cameos are no longer relegated to the eagle-eyed viewer).

Sane interactive models also will come into play within future streaming media user interfaces. We say “sane” because many interactive television models require too many options to be of use in real-time or, worse, provide too little pertinent information. As an example of pertinent, timely, accessible information, let’s imagine an orange juice commercial: during the commercial, when the product appears, a verbal command of “view nutritional information” could be used to rotate the product to view an enlarged nutritional content label. Another verbal command of “resume” would return the product to its normal position and continue the commercial. The choice to view the product label was the viewer’s and did not require additional unnecessary screen real estate for a still graphic, since the information may be of interest only to a small portion of the total viewership.

The passive entertainment environment also demands a higher presentation quality in order to push the envelope toward more immersive entertainment. Some analysts expect tomorrow’s entertainment interfaces to incorporate real-time 3-D rendering, much like today’s multiplayer video games, and multiple high definition images, perhaps across several screens for enhanced viewing. Advanced digital audio delivery, through the use of 10-12 channels of audio with on-the-fly programming to accommodate differences between music, dialogue, and spatial depth already are on the horizon. Similar to the home theater experience, in which many homeowners are spending significant money to recreate—or surpass—the cineplex experience, so too will streaming video interfaces move to surpass today’s impressive gaming environments.

Much like the dominant revenue share that the VHS/DVD rental industry now holds over traditional theater-going audience revenues, downloadable movies for the home entertainment market continue to grow in popularity due to the convenience, timeliness, and quality provided by high-speed networks and advanced delivery mechanisms (see separate article on page xx dealing with the topic of next-generation delivery tools). The ability to exceed today’s DVD image quality should be achieved by mid-2005, especially now that H.264 has been approved for use in High Definition DVD players.

In addition, technologies such as round-tripping, which are found in audio products like the iPod, provide distinct advantages for streaming media service providers, including the

ability to offer services that begin within the home (downloading or buffering a portion of the movie) but continue on to a portable device such as a laptop, PDA, or vehicle screen without interruption. Flexibility of place and time becomes an added feature to the experience of convenience and quality that is not readily or easily available in today's home theater experience.

Future rich media-interface design also has the potential to inject profound changes into the media production process. At the simplest level, media production tools need to begin preserving all metadata inherent in the capture process; today more than 90% of that metadata is lost at the point of ingestion into a nonlinear editing system. At a deeper level, too, the way media is produced also will change to accommodate age-specific views. Over time, with the advent of digital cinema, movies may consistently be shot for multiple ratings, from G to R or beyond, allowing a theater owner, parent, teacher, or administrator to configure interface and content views for a specific grade, age, or access level. The cost savings and added upside in retail or rental revenues could, in turn, force media producers to use streaming and security technologies to allow omission of language or other content not suitable for a specific controlled group of viewers, without limiting the overall accessibility to the media as a whole.

### **Corporate**

The same model holds true for the corporate environment in which some portion of a presentation or recording is limited to only those with proper viewing permission. The cost of creating a single presentation with multiple viewing access levels can be more easily absorbed than today's model, which requires multiple versions of a presentation—the cost savings is especially apparent when the content of each version of the presentation requires consistent updating. In effect, the entire presentation always is available, but access levels provide one form of “editing” the content without forcing the presenter to remove information that may be pertinent for one group but superfluous or out of bounds for another.

Access and viewership levels aside, the corporate setting has its own unique set of challenges. Take, for instance, training meetings. Often, not every desired participant can attend, but those who wish to “get up to speed” by watching the streaming video later often are at a disadvantage, since the presenter may refer to content that is not available to the post-event viewer, even if the viewer is watching an “enhanced” version that contains the Powerpoint presentations prevalent in today's “advanced” three-pane interface designs. Tomorrow's interfaces also must have the flexibility of aggregating additional disparate content such as session notes or handouts provided by the presenter, so that remote or on-demand participants have access to information referenced by the presenter.

The practicalities mean that tomorrow's interfaces will act as a front end for entire content management systems, not just streaming servers, and may require links to GIS systems should a speaker reference a location near where she was presenting. This means tomorrow's capture systems must act as digital asset management (DAM) systems for all digital assets, or at least link to other DAMs, while playback interfaces must be

able to display this content in a meaningful manner at the appropriate time. Just as one might choose today to view information regarding an artist while listening to music in popular audio jukebox programs, so too will streaming video viewers be able to set preferences that allow on-the-fly reconfiguration to the viewing environment best suited for their experience at any given moment.

Another corporate sub-vertical that would benefit from immersive playback interfaces is corporate campus surveillance. Streaming interface manufacturers would be remiss in simply replicating the 4, 8, 16, or 32 window interfaces currently in vogue in campus surveillance systems. These multi-window interfaces are limited by hardware technology that stagnated years ago and have finite limitations that stifle innovative surveillance practices. Darim Corporation may have the best example of an innovative surveillance user interface—its HotActions and 3-D Spider products allow a campus to be laid out in a standard 3-D program, such as 3DS Max, with separate video streams represented by cameras and virtual monitors that grow to full screen as one navigates the 3-D environment and zooms in on a particular building or floor. This model gives the campus surveillance team the extra edge, without the confusion of trying to determine where one of 32 images displayed on a screen is geographically located.

### **Education**

In much the same way, educational institutions can use these same tools to keep their physical campuses safe. But streaming media provides many additional benefits in an educational setting. As more universities look to attract the non-traditional student—who may attend on-campus classes infrequently due to work conflicts or geographic limitations—tomorrow’s interface must provide distance education students the same immersive environment that on-campus students receive. During a live classroom lecture, distance education students using tomorrow’s user interfaces may experience Flash graphics that appear at specific, timed talking points to “invite” a viewer to mouse over and open glossary information that augments the professor’s lecture, or additional references to video and text to which the professor alludes during the course of the lecture. Take, for example, a literature professor’s comments about a particular book or author. Given work that several companies, such as Pictron, are doing to marry speech to text databases with robust search tools, tomorrow’s user interfaces may attempt to Google the author’s bibliography during the lecture and present its findings to both the professor and the distance education students.

Since tomorrow’s interfaces also might be able to tie the GIS or 3D information of the campus into a virtual world that distance education students can access to move easily from class to class, the once-popular telephone slogan, “It’s the next best thing to being there,” may morph into “It’s like being there, only better.”

Real-time exchange of information during live events also can be enhanced through incorporation of several existing technologies that have yet to be fully integrated into streaming media user interfaces:

- The ability to enter information at viewer/programming controlled intervals, such as

verifiable test points during a presentation for off-campus educational or corporate training purposes.

- Instant messaging during live streaming events on multiple levels—directly with the presenter, another individual participant, or the group as a whole.
- Speech-to-text, coupled with real-time translation of the spoken word into another language, which can then be shown as subtitles or closed-caption information for those who need this supplementary aid.
- Multiple camera angles in a given room, with only one camera streaming at a particular time, allowing a remote participant to view another single participant or sub-group of participants.

In summary, whether the desired experience fulfills an educational, corporate or entertainment goal, tomorrow's streaming media interfaces will be much more than just a replication of the television, providing viewers an immersive virtual experience that meets or exceeds the benefits of actually attending a live event. This was the premise behind a project we attempted in mid-1999 to create a virtual tradeshow. At that point, technologies such as Darim's 3-D Spider did not exist, and the project ultimately was scrapped. Still, it left an indelible impression of the potential of a robust, immersive interface that could easily be a viable business model today.

Of course, no interface can take the place of face-to-face networking—yet. But emerging streaming media interfaces could assist in quickly bridging the understanding and expectations of all parties involved, allowing for a more progressive dialog and project-specific potential.

As global commerce continues to grow, interface designers must lead the way with development of innovative and intuitive software, rather than following in technology's footsteps with substandard fixes.

Tim Siglin is an analyst and co-founder of Transitions, Inc., a consulting firm based in Kingsport, TN. He can be reached at [articles@transitioned.com](mailto:articles@transitioned.com).

Rosemary Roberts is the owner of Girl On Point, LLC, a creative services firm providing print and website content, as well as consulting for streaming media applications. She can be reached at [rosemary@girlonpoint.com](mailto:rosemary@girlonpoint.com).