

On the Trail of the Shadow Woman: The Mystery of Motion Capture

Ben Delaney

CyberEdge
Information
Services

It was a muggy summer night in one of Santa Monica's seedier neighborhoods. I was on assignment. My job was to capture the Shadow Lady. I didn't expect it to be easy. But I guess that's why I get the big bucks.

I loitered in a dark doorway, watching and waiting. As I finished my Lucky Strike, I saw her. The Shadow Lady pulled up in her little red sports car and, without looking around, walked through an unmarked door into a big warehouse. She was wearing an overcoat and a beret. Odd, that overcoat. This July night in Santa Monica, the evening temperature was still in the 80s. I took one last drag, threw down the butt, and followed her in. It was time for some action.

I'm a motion capture technician. I handle the hard stuff—

set-up, calibration, cleaning up data—all the stuff nobody else wants to do. We need lots of room, lots of electrical power, and the freedom to work any hours we need to. That's why we work in what looks like a run-down warehouse. Inside, we have nearly a million dollars worth of gear.

The Shadow Lady is a professional dancer and actress. She's lithe, graceful, expressive, and perhaps most importantly, really patient. She gave herself the name: "All you keep is my shadow," she said at the end of a session one morning. Basically, she was right.

Catching shadows

Motion capture houses throw away the person and keep the shadow—the essence of their motion—and apply that motion to animate all kinds of characters. Perhaps you saw those dancing cars and credit cards in the Shell ad on TV. That was done with MoCap. Or you may have admired those folks strolling on the sun deck in *Titanic*. That was MoCap, too. In the past few years, as the technology has become less expensive and at the same time more accurate, MoCap has helped lessen the rigors of traditional cell animation—especially in highly cost-conscious projects like TV commercials.

A MoCap studio can use both magnetic and optical capture systems. Each has certain advantages. The magnetic system handles shoots that don't need really high accuracy, and the director wants to be able to apply the data to his character in real time. Magnetic systems generally are faster at providing data to use in animation, though optical systems are generally more accurate and can track more points. That may be changing, though. The first real-time optical system was shown at Siggraph 98 by Motion Analysis (Santa Rosa, California).

When I walked onto the capture stage, the Shadow Lady was ready to "suit up." She'd dropped the overcoat to reveal a formfitting pink leotard. The color didn't matter, of course. It was her moves we wanted. Our magnetic system included 13 receivers/sensors, each with 6 degrees of freedom. We attached them to the Shadow Lady's leotard with Velcro straps (see Figure 1). This gave us pretty good coverage—good enough for this assignment, anyway, an animated commercial.

Magnetic systems work by generating three orthogonal electro-magnetic fields from each transmitter. The

1 The Shadow Lady strikes a pose. She's wearing the Ascension MotionStar Wireless magnetic tracking system. Each small cube is a receiver. Her waist pack holds a battery, data collection unit, and radio transmitter.



Photo courtesy of Ascension Technology

host computer is aware of the timing of the signals. When the receivers pick up the signals, the host knows the distance from the transmitter by the time elapsed and the orientation of the receiver by the changes in the signal caused by tilting of the magnetic fields.

Optical systems are preferred for high-res work. An optical system can collect hundreds of data points, though you seldom need more than about thirty. It works by visually tracking small reflective markers attached to the performer at key points. Optical systems are also first choice when you have lots of action. You can't have wires getting in the way with your actors bounding around, so you need optical tracking (see Figure 2).

The more data points you track, the less extrapolation the software has to do. So for an animation of a character that doesn't look much like a human being—such as those dancing credit cards—you only need enough points to get the basic motion into a file. The animators then use that as a framework. They do plenty of work to make those points fit their character.

The object in most MoCap sessions is simple—save a data set that represents the subject's motion with the optimal level of detail and the least amount of noise. The optimal detail varies according to the project. For ergonomic study, you want fine resolution of the motion and highly repeatable measurements. For animation of a fantasy character like Moxie, the animated MTV host, you need relatively low resolution and repeatability. Sports games such as EA Sports *Madden NFL 99*, or *Knockout Kings*, where the movements are fast and accuracy is essential to realism, would fall somewhere in between, since the motion animates a relatively realistic human figure. For example, a setup to capture a martial arts sequence might use 20 to 30 sensors, and a full-body ergonomic study could use 100 or more.

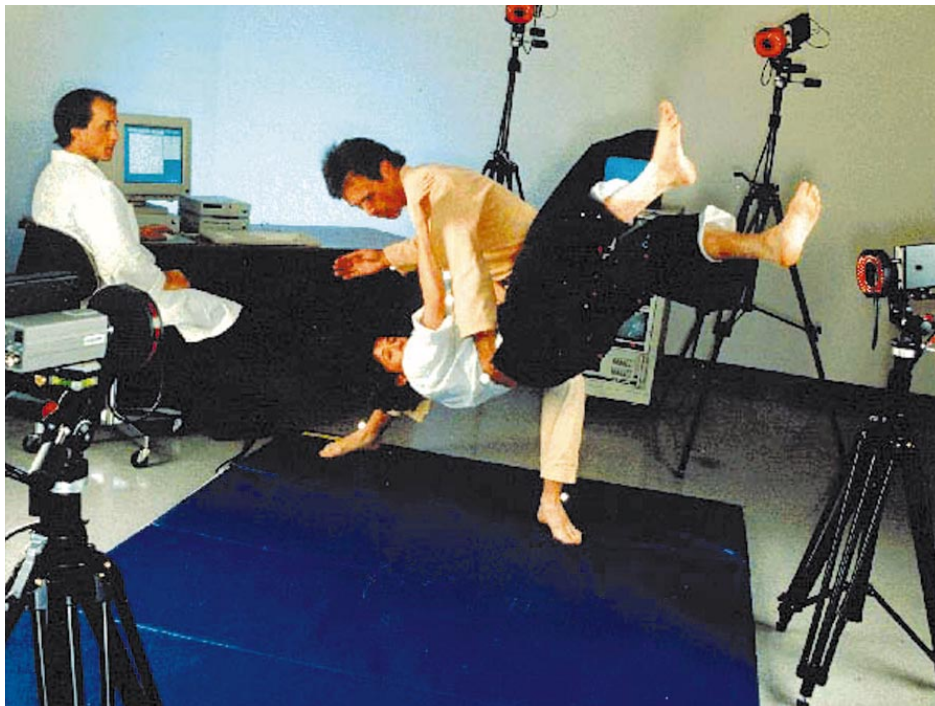
Another issue is the speed of the motion being cap-

World Wide Web URLs for MoCap

You can find out more about the following motion capture houses and equipment manufacturers by visiting their sites on the World Wide Web.

Adaptive Optics Associates: <http://www.aoainc.com>
Analogus: <http://www.analogus.com>
Ascension Technology: <http://www.ascension-tech.com>
BioVision: <http://boris.biovision.com>
Digital Domain: <http://www.d2.com>
5DT: <http://www.5dt.com>
GameTek: <http://www.gametek.com>
House of Moves: <http://www.moves.com>
MediaLab: <http://www.demon.co.uk/mlab>
Motion Analysis: <http://www.motionanalysis.com>
Performance Animation Society: <http://www.pasociety.org>
Polhemus: <http://www.polhemus.com>
Protozoa: <http://www.protozoa.com>
Richard Cray's Performance Animation Page: <http://www.pasociety.org/peranim>
Vicon Biomechanics (was Oxford Metrics): <http://www.metrics.co.uk>
Virtual Technologies: <http://www.virtex.com>

tured. Optical systems can run at 240 Hz or more—critical for discerning very fine movements, or when the motion is quick. In addition, high-speed systems typically multiplex their data capture channel among the sensors. So an optical system with 100 tracking points would need to operate about four times faster than a system tracking just 25 points to obtain the same temporal resolution. Magnetic systems operate at about 140 Hz and can support fewer sensors than optical systems. Of course, there are other considerations as well.



2 This heavy action scene uses optical tracking with a Motion Analysis system. The little balls on the actors are the reflective markers, tracked by the cameras surrounding them.

Photo courtesy of Motion Analysis

The big three

Probably the three biggest issues in MoCap for entertainment are range, interference, and wires. Let's address them one by one.

Range refers to both the distance between the performer and the capture equipment, and the size of the area in which capture can happen. These are significant and related factors. Obviously, you need to have room for the performers to move without running into the walls or equipment. This area is blocked out before the capture session, typically with the director or producer, the actor/dancer, and the capture technician. A simple dance may only need 100 square feet, while a fight scene could need 500 square feet or more. If you need to capture a performer running for some distance, treadmills sometimes can provide the required mobility. No matter how it's arranged, the working area is a big factor in setting up the session.

The other range issue is the distance between the performer and the capture equipment. The inverse square law dictates that signals get weaker in proportion to the square of the distance between the source and receiver. For magnetic trackers, that limits the range to about a 10-foot radius per transmitter. You can gang up some systems to use multiple transmitters, increasing the range and capture area. Optical trackers are virtually unlimited in range (see Figure 3), but the spatial resolution suffers because the tracking targets look smaller as they move away from the cameras. This reduces accuracy.

The second issue is *interference*. Both optical and magnetic systems suffer from interference, though the details completely differ for the two systems. Optical interference is primarily occlusion. For example, an arm moves in front of a thigh marker, or the performer turns so the camera can't see some of the markers, or one performer steps in front of another. Work-arounds—typically adding more cameras—help avoid the problem, but occlusion is a fact of life and a continuing annoyance with optical tracking.

Don't forget, though, that magnetic systems have their own problems (see Figure 4). They suffer from electrical interference caused by induced magnetic currents,

eddy patterns caused by large metallic objects nearby, or external sources of radiation, such as TVs and computer monitors. The work-arounds for these problems include sophisticated software filtering of the data.

Wires pose the third big concern. For quite a while optical systems were preferred for any sort of athletic activity because they didn't need the annoying cables that magnetic systems required. In addition, cables limit range, and when you work with more than one actor, they can get quite ridiculously tangled. It's just in the past two years that both Ascension (Burlington, Vermont) and Polhemus (Colchester, Vermont) introduced wireless magnetic systems.

I got down to work, attaching 12 sensors to the Shadow Lady's arms, legs, and shoulders. Then I had her put on a headband with one more sensor sewn on the front. I strapped the pack around her waist and plugged the sensor wires into the receiver in the pack. I used Velcro straps to hold the wires snugly to her limbs. I checked the battery again. Then I sat down at the console.

I asked the Shadow Lady to assume a series of poses and watched the stick figure on my monitor as it followed her through the moves. "Number three's a little jittery," I thought, but decided we could live with it. I dialed in zero points for each receiver and nodded to the director. "We're ready."

The director waved to the sound man, and as I hit the "Start" button, loud Caribbean music started to play. The Shadow Lady started to dance. I watched the monitor, not her. My little stick figure looked pretty good. Everything was humming. We captured about three and a half minutes, then the director yelled "Cut!" As he walked over to talk with the Shadow Lady, my mind wandered. I was thinking about where this technology came from—the roots of MoCap, so to speak.

The family tree

Motion capture is classified as "special effects," a term used to explain just about anything you see in the movies that isn't a straight camera shot. Special effects are just about as old as film itself. The term was first used in a movie credit in 1926, on the film *What Price Glory?* Back then, special effects involved some of the same mechanical tricks done in stage productions. Soon, double exposures and other in-camera tricks, like stopping the camera, moving a person or prop, and re-starting, were added to the toolkit. Early filmmakers quickly reached the limits of model making and in-camera effects, though, and started developing new tools. One of the most important was the *matte shot*.

A matte shot is created by taking a piece of film with some action on it, like the heroine running frantically from a villain, and creating a matte, or mask, that eliminates everything but her image from the film. This is

3 An optical tracking system uses two or more video cameras to find the reflective markers attached to the actors in key locations. With no wires leading from the actors, they have complete freedom of motion.

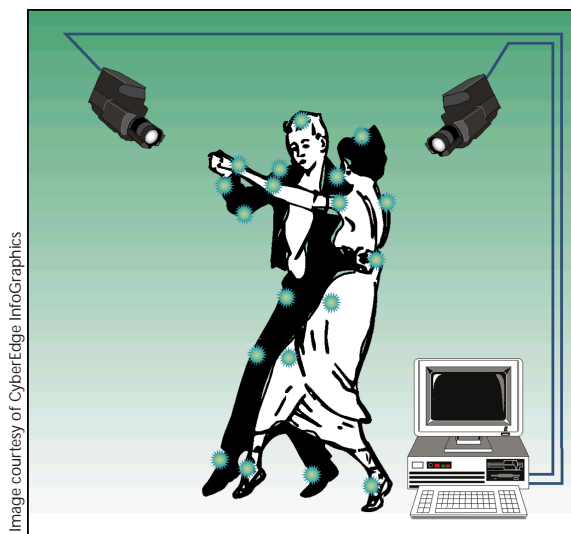


Image courtesy of CyberEdge InfoGraphics

then composited photographically with a painted background, so the shot of her running in a studio can become her running through the streets of Shanghai or through the desert. This simple trick is still used frequently, though today most of the compositing is done with computers. This technique was improved upon with the development of the *traveling matte shot*, in which the matte changes with the action. This permits combining the foreground action with another piece of film, which can include background action.

Television made it possible to do matte shots in real time using a technique called *blue-screen matte*. This technique makes it seem like your local weather person is standing in front of an animated map when they're really standing in front of a blank blue screen. A mixing device combines their image with the picture of the map.

In both film and TV, matte shots just weren't satisfying enough. Directors wanted to mix animated characters and live people, and they wanted animated characters to move in a more natural fashion. Walt Disney was among the first to combine animation and live action in a feature film, and his seminal *Song of the South* was a big hit in 1946. However, viewers saw an obvious gap between the animated characters and the live actors.

Stop-action animation was another attempt to make animation more realistic and less expensive. Unlike traditional cell animation, which relies on hand drawn and colored frames that are then photographed one by one, stop-action animation uses models, which are moved in small increments and photographed after each move. Possibly the greatest practitioner of this art was Ray Harryhausen, whose films include *The Seventh Voyage of Sinbad* and *Jason and the Argonauts*. His painstaking technique included matting the stop action footage with real-life backgrounds. The effect was pretty good, but the action was obviously artificial.

Rotoscoping was developed around 1915 by Max Fleischer to help bridge the gap between natural motion and animation. This technique requires a technician to trace an actor's motion by hand on each frame of the filmed sequence. This tracing then serves as the basis for inking the actual cartoon character. You can imagine how long that takes. It works, but it's hard to do well, and the results leave something to be desired.

In the 70s, the US military began funding the development of magnetic tracking devices. They were used for following the head movement of pilots, among other things. As virtual reality appeared on the scene in the late 80s, these same trackers were adapted for use in tracking heads and hands in virtual worlds. By the mid-90s, some animators realized that they could use the same tracking systems for animation.

Optical tracking has a much longer history. Almost as

soon as moving pictures were developed (even before, if you consider Muybridge's efforts), people were tracing movements with grease pencils on negatives and analyzing those tracks. As video became inexpensive and ubiquitous, people realized they could draw on an overlay on the video and create an animation. Then computer systems were developed that could perform the process automatically. The addition of optically bright markers made the task even easier, and today optical tracking is a mainstay of performance animation.

In addition to body tracking, many studios add input gloves, such as Virtual Technology's (Palo Alto, California) CyberGlove, or 5DT's (Perseuor, South Africa) 5thGlove, to capture hand articulation. Facial expressions are captured with optical systems, such as one from Analogus (San Francisco, California), or by puppeteering, as demonstrated by MediaLab (Paris, France).

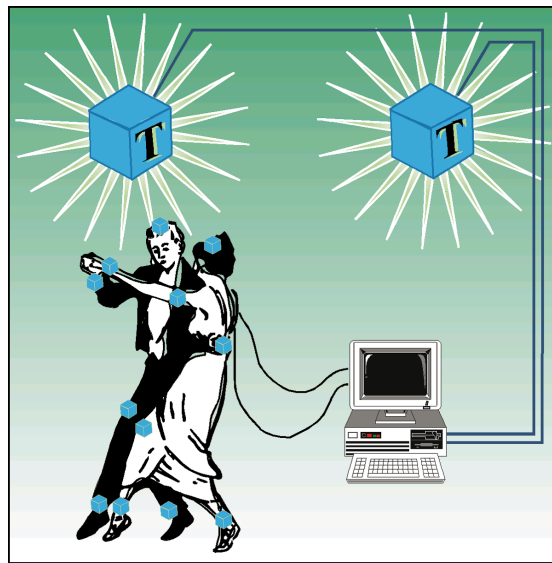


Image courtesy of CyberEdge InfoGraphics

4 A schematic of a typical magnetic tracking system setup. The two large boxes marked with T are transmitters. The small boxes on the dancers are receivers. This system is hardwired; newer units use radio transmission to eliminate wires between the actors and the host computer.

Using MoCap

Until recently, MoCap was very touchy stuff. It still demands careful attention from manufacturers to make sure everything works right. Ascension Technology offers an example.

Jack Scully is vice president of Ascension Technology, manufacturers of the MotionStar and Bird tracking systems. He explained that "It used to be, five years ago, you needed somebody with a degree in electrical engineering and a programmer to get this stuff working. Now it's much easier. It's pretty much plug and play on the hardware end. We have drivers to connect our hardware to 3D Studio, Alias | Wavefront, SoftImage, all the common programs."

Scully said that Kaydara's Filmbox makes one of the best plug-ins for use with Ascension's equipment. FilmBox sits between the MotionStar and the animation software. It controls the data capture, then cleans up the data, edits it, and passes it on to the animation package. "This, more than anything else, makes the process plug and play," Scully explained.

Another step is necessary to assure that Ascension's

5 An elephant might dance like Baryshnikov, or Baryshnikov might prance like a horse. Using a large set and an optical tracking system lets this horse provide the moves for an animated character.



Photo courtesy of Motion Analysis

systems work right the first time. The company sends prospective customers a detailed survey of the equipment and software that the customer wants to use with the tracking hardware. When Ascension knows what computer, software, and environment the client intends to use, they fine-tune the system and makes recommendations to the client regarding optimal setup. The company even pays a visit with a “sniffer” that maps the electrical and magnetic characteristics of the stage. Prior to shipping a tracking system, Ascension provides recommendations regarding placement and operating conditions that will make the system more foolproof. They recommend a wooden stage, at least 18 inches away from a metal or concrete and steel floor, and isolated from power mains, large metallic objects, and RFI emitters such as monitors.

Ayes and nays

Using motion capture in animation is not without critics. Traditional animators rightfully take a lot of pride in their craft. Some of them see rotoscoping and MoCap as cheating. In a 1997 Siggraph panel on MoCap, Craig Hayes—a director, animator, and MoCap developer at Tippet Studios (Berkeley, California)—said, “Motion capture tends to be used as a crutch, or even worse, to create performances/images that could have been created with filmed, live actors.” Steph Greenberg, an independent animator who has worked at Disney and many other studios, added, “An animated character has capabilities that no human can replicate without possible injury. Characters can snap into position, their movement deliberate and uncompromising—their athletic abilities simply can’t be matched.”

Even the vice president of marketing at a leading MoCap equipment manufacturer has reservations. Chris Welch, who holds that position at Motion Analysis, said, “MoCap is a tool the animator uses. You can’t use motion capture to make Bugs Bunny walk like Bugs Bunny. But

MoCap gives the animator the time to do good work. Instead of dealing with animation one frame at a time, they can spend time on painting, backgrounds, and other stuff.”

According to Welch, the bottom line is, “If you want a character that dances like Baryshnikov but looks like an elephant, MoCap makes that look good, without spending hours on every frame. MoCap will provide the capability to put high-quality animation out again.” See Figure 5.

Gary Roberts of Centroid Studios in London, England is a fan of MoCap, which he recently used on the movie *Lost in Space*. As he explained in a press release issued by the studio,

I was tasked with capturing actor Gary Oldman for over 10 minutes of on-screen time. We elected to use the Motion Analysis system with 8 cameras . . . We used between 35 and 42 markers for the facial capture. The character Gary played had metal plates over his face, so we positioned markers to accurately replicate the movement of these plates as closely as possible. . . . We required a volume of 1.5m by 1.5m by 1m and a 200-degree field of movement in both the X- and Y-axis for Gary’s face. Using 8 cameras and a selection of 1mm to 2mm markers, we were able to achieve this without any problems.

Roberts isn’t alone in his appreciation of MoCap as a way to translate the characteristic movements of real people to animated characters. From the molten metal robot in *Terminator 2* to the strolling passengers on the deck in *Titanic*, Hollywood has taken to MoCap in a big way.

The bottom line

What really made this all practical, though, is the constant downward spiral of computer prices. It still takes a

lot of computer power to create performance animation from tracking data, especially in real time, but today that power costs a fraction of what it did even five years ago. While a good performance animation system still costs \$50,000 and up, it would have cost millions just a few years back.

Of course, filmmakers don't do motion capture just because they can. Combining captured motion data and 3D animation gives them a lot of advantages the cell animators will never have, even if cost is no object. For example, it is inconsequential to rotate a 3D character, or change the lighting, because the software treats the 3D model like a real object. When you have a 3D model, it's easy to map motion data to control points on the model. So, at a basic level, using MoCap for animation is really simple. If you need to mix live action and real-time or canned animation, 3D figures permit a greater range of interaction.

However, if cost is an object, MoCap animation overwhelms traditional cell work. Traditional animation can cost \$20,000 per minute of finished footage, or even more. A good MoCap house, working with journeymen animators, can cut that cost to as little as \$500 per minute. For certain types of work, the savings could be even greater. For example, if you wanted to include an animated Michael Jackson in a piece using traditional cell animation, someone would need to become an expert in animating Jackson's moon walk, gestures, and physical style. It would be obvious to anyone familiar with Jackson if the moves were off. With MoCap, this issue just doesn't exist. Simply wire Michael, as he was for his *Ghosts* music video, and have him do his thing. When you see that footage, you immediately know that it is Michael Jackson dancing, not an imitation.

These economies and fidelities drive the MoCap business today. A recent study conducted by CyberEdge Information Services (Sausalito, California) found that about 30 to 40 MoCap houses provide motion capture services for hire. Those service bureaus bring in, on average, around \$675,600 each (in 1998). Many more studios are divisions of movie studios, special effects houses, and game publishers. They contribute savings, rather than revenue, to the bottom line. Business is very good for the service bureaus—the study respondents anticipated average growth of more than 75 percent in 1999.

That growth rate, if accurate, bodes well for equipment manufacturers. MoCap users spend, on average, more than \$90,000 per system. Ascension Technology and Polhemus, who both make magnetic systems, control the lion's share of the MoCap market, with around 40 percent between them. Optical system manufacturers Vicon (Oxford, UK) and Motion Analysis (Santa Rosa, California) follow them in market share, and a flock of smaller companies divvy up the rest of the market.

Magnetic MoCap systems cost less than optical systems, starting at around \$20,000 and rising quickly as you add sensors and range. Optical systems require an initial investment of more than \$100,000. Of course, these costs do not include the computers, software, and miscellany required to actually do any work.

Whither MoCap?

Though MoCap's future looks assured, a few concerns still exist. The CyberEdge study revealed that while most users of MoCap systems are pretty satisfied with their equipment, several areas bother them. The single biggest issue for them is what they categorized as the general difficulty of using the gear, plus its lack of robustness. This is still new technology, and it shows. Wires drag about everywhere, connectors break, and interference pops up constantly. Magnetic systems can be difficult to calibrate, and optical systems often require very precise set-ups. Also, MoCap users want greater range, fewer wires, and greater accuracy in the measurements provided. While price isn't a major issue—end customers say they're willing to pay what it takes—users, especially users of optical systems, want systems that cost less.

Systems will certainly continue to improve, and the problems will be solved. MoCap will get less expensive, easier to use, and more common. The big question is, where will MoCap show up next?

Within a few years, and perhaps even sooner, we will see the first non-cartoon virtual actors. Non-cartoon means that you may have to look twice to know if you're seeing a live person or a digital duplicate. Libraries of MoCap data will make it possible for these synthetic thespians to talk the talk as they walk the walk—Groucho's goofy prance, John Wayne's swagger, or Marilyn Monroe's seductive swirl. Michael Jackson may dance on for a hundred years, now that his moves are recorded. Perhaps you've seen the dancing baby that opens the "Ally McBeal" TV show. In a few years, the whole cast may have originated in 3D Studio.

With virtual actors will come interesting new legal discussions. What makes up an actor's personality? If Arnold Schwarzenegger has his face and body scanned, his motion captured, and his voice recorded, does a synthetic Arnold have the same rights as the flesh and blood model? Only time (at \$500 and up per hour of legal fees) will tell.

The real performers left, such as those reading the news on TV, will largely work in virtual sets. While not a MoCap application, virtual sets use very similar technology for tracking camera movement. With a virtual set, the camera angle is a vital factor in rendering the proper view of the set. Cameras are tracked with the same systems used to track actors. We'll undoubtedly see virtual actors, virtual sets, and live actors all in the same show, and probably soon.

Dawn was breaking over the mountains when I finally left the studio. Our session had been a success. I had bagged the prize and, as usual, had earned my money. As I waited for the Red Car, I wanted nothing more than a tall, cold carrot juice, an aspirin, a shower, and a bed. And maybe some scrambled eggs. And sausage. With home fries. I definitely needed some rest. I knew that next week I'd be seeing the Shadow Lady every time I turned on my TV—as a dancing banana. ■

Delaney may be captured at ben@cyberedge.com.

Contact Potel at potel@wildcrest.com.