MIT
Technology
Review

Menu

Computing

Lying With Pixels

Seeing is no longer believing. The image you see on the evening news could well be a fake - a fabrication of fast new video-manipulation technology.

by Ivan Amato   July 1, 2000

Last year, Steven Livingston, professor of political communication at George Washington University, astonished attendees at a conference on the geopolitical pros and cons of satellite imagery. He didn't produce evidence of new military mobilizations or global pandemics. Instead, he showed a video of figure skater Katarina Witt during a 1998 skating competition.
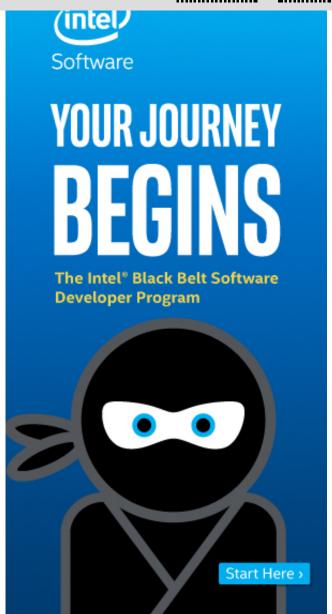
In the clip, Witt gracefully plies the ice for about 20 seconds. Then came what is perhaps one of the most unusual sports replays ever seen. The background was the same, the camera movements were the

identical to the original in all ways except for a rather important one: Witt had disappeared, along with all signs of her, such as shadows or plumes of ice flying from her skates. In their place was exactly what you would expect if Witt had never been there to begin with-the ice, the walls of the rink and the crowd.

So what's the big deal, you ask. After all, Stalin's staff routinely airbrushed persona non grata out of photos more than a half-century ago. And Woody Allen ushered a variation on reality morphing into the movies 17 years ago with Zelig, in which he inserted himself next to Adolf Hitler and Babe Ruth. In films such as Forrest Gump and Wag the Dog, reality twisting has become commonplace.

What sets the Witt demo apart-way apart-is that the technology used to "virtually delete" the skater can now be applied in real time, live, even as a camera records a scene and instantly broadcasts it to viewers. In the fraction of a second between video frames, any person or object moving in the foreground can be edited out, and objects that aren't there can be

edited in and made to look real. "Pixel plasticity," Livingston calls it. The implication for those at the satellite imagery conference was sobering: Pictures from orbit may not necessarily be what the satellite's electronic camera actually recorded.

But the ramifications of this new technology reach beyond satellite imagery. As live electronic manipulation becomes practical, the credibility of all video will become just as suspect as Soviet Cold War photos. The problem stems from the nature of modern video. Live or not, it is made of pixels, and as Livingston says, pixels can be changed.

The best-known examples of real-time video manipulation so far are "virtual insertions" in professional sports broadcasts. Last January 30, for instance, nearly one-sixth of humankind in more than 180 countries repeatedly saw an orange first-down line stretched across the gridiron as they watched the Super Bowl. New York-based Sportvision created that line and inserted it into the live feed of the broadcast. To help determine where to insert the orange pixels, several game cameras were fitted with sensors that tracked the cameras spatial positions and zoom levels. Adding to the illusion of reality was the ability of the Sportvision system to make sure that players and referees occlude the virtual line when their bodies traverse it.

Last spring and summer, as Sportvision and rivals such as Princeton Video Imaging (PVI) in Lawrenceville, N.J., were airing virtual insertion products, including simulated billboards on walls behind major league batters, a team of engineers from Sarnoff Corp. in Princeton, N.J., flew to the Coalition Allied Operations Center of NATO's Operation Allied Force in Vicenza, Italy. Their mission: transform their experimental video processing technology into an operational tool for rapidly locating and targeting Serbian military vehicles in Kosovo. The project was

dubbed TIGER, for "targeting by image georegistration." "Just dial in the coordinates and the thing goes," explains Michael Hansen, a young, caffeinated Sarnoff gadgeteer who can hardly believe he was helping fight a war last year.

Compared to PVI's job, the military's technical task was more difficult- and the stakes were much higher. Instead of altering a football broadcast, the TIGER team manipulated a live video feed from a Predator, an unmanned reconnaissance craft flying some 450 meters above Kosovo battlefields. Rather than superimposing virtual lines or ads into sports settings, the task was to overlay, in real time, "georegistered" images of Kosovo onto the corresponding scenes streaming in live from the Predator's video camera. The terrain images had been previously captured with aerial photography and digitally stored. The TIGER system, which automatically detected moving objects against the background, could almost instantly feed to the targeting officers the coordinates for any piece of Serbian hardware in the Predator's view. This was quite a technical feat, since the Predator was moving and its angle of view was constantly changing, yet those views had to be electronically aligned and registered with the stored imagery in less than one-thirtieth of a second (to match the frame rate of video recording).

In principle, the targeting step could have been hotwired to precision guided weapons. "We weren't actually doing that in Allied Force," Hansen notes. "We were just telling targeting officers exactly where Serbian targets were and then they would vector in planes to go strike the targets." That way the human decision makers could pre-empt flawed machine-made decisions. According to the Defense Advanced Research Projects Agency, TIGER technology was used extensively in

the final three weeks of the Kosovo operation, during which "80 to 90 percent of the mobile targets were hit."

So far, real-time video manipulation has been within the grasp only of technologically sophisticated organizations such as TV networks and the military. But developers of the technology say it's becoming simple and cheap enough to spread everywhere. And that has some observers wondering whether real-time video manipulation will erode public confidence in live television images, even when aired by news outlets. "Seeing may no longer be believing," says Norman Winarsky, corporate vice president for information technology at Sarnoff. "You may not know what to trust."

The Sublime to the Ridiculous

A crude form of video manipulation already is happening in the satellite imagery community. The weekly publication Space News reported earlier this year that the Indian government releases imagery from its remote-sensing satellites only after defense facilities have been "processed out." In this case, it's not real-time manipulation and it's up front, like a censor's black marker. But pixels are plastic. It is perfectly possible now to insert sets of pixels into satellite imagery data that interpreters would view as battalions of tanks, or war planes, or burial sites, or lines of refugees, or dead cows that activists claim are victims of a biotech accident.

A demo tape supplied by PVI bolsters the point in the prosaic setting of a suburban parking lot. The scene appears ordinary except for a disturbing feature: Amidst the SUVs and minivans are several parked tanks and one armored behemoth rolling incongruously along. Imagine a tape of virtual Pakistani tanks rolling over the border into India

pitched to news outlets as authentic, and you get a feel for the kind of trouble that deceptive imagery could stir up.

Commercial suppliers of virtual insertion services are too focused on new marketing opportunities to worry much about geopolitics. They have their eyes on far more lucrative markets. Suddenly those large stretches of programming between commercials-the actual show, that is-become available for billions of dollars worth of primetime advertising. PVI's demo tape, for instance, includes a scene in which a Microsoft Windows box appears-virtually, of course-on the shelf of Frasier Crane's studio. This kind of product placement could become more and more important as new video recording technologies such as TiVo and RePlayTV give viewers more power to edit out commercials.

Dennis Wilkinson, a Porsche-driving, sports-loving marketing expert who became CEO of 10-year-old PVI about a year ago, couldn't be happier about that. Wilkinson's eyes gleam when he describes a (near) future in which virtual insertion technology pushes advertisements to the personalized extreme. Combined with data-mining services by which browsers' individual likes, dislikes and purchasing patterns can be relentlessly tracked and analyzed, virtual insertion opens up the ability to shunt personally targeted advertisements over phone lines or cables to Web users and TV viewers. Say you like Pepsi but your neighbor next door likes Coke and your neighbor across the street likes Seven-Up-the kind of data harvestable from supermarket checkout records. It will become possible to tailor the soft-drink image in the broadcast signal to reach each of you with your preferred brand.

Just 15 minutes up the road from PVI, Sarnoff's Winarsky is also glowing-not so much about capturing market share as about the transforming power of the technology. Sarnoff has a distinguished

history in that regard; the company is the descendant of RCA Laboratories, which started innovating in television technology in the early 1940s and has given birth to a plethora of media technologies. The color TV picture tube, liquid crystal displays and high-definition TV all came, at least in part, from RCA qua Sarnoff, which has five technical Emmys in its lobby.

The ability to manipulate video data in real time, he says, has just as much potential as some of these forerunners. "Now that you can alter video in real time, you have changed the world," he says. That may sound inflated, but after looking at the Katarina Witt demo, Winarsky's talk of "changing the world" loses some of its air of hyperbole.

Deleting people or objects from live video, or inserting prerecorded people or objects into live scenes, is only the beginning of the deceptions becoming possible. Pretty much any piece of video that has ever been recorded is becoming clip art that producers can digitally sculpt into the story they want to tell, according to Eric Haseltine, senior vice president for R&D at Walt Disney Imagineering in Glendale, Calif. With additional video manipulation technologies, previously recorded actors can be made to say and do things they have never actually done or said. "You can have dead actors star again in entirely new movies," says Haseltine.

Contemporary shots featuring footage of dead performers have been around for several years. But the Hollywood illusion-craft that, for example, inserted John Wayne into a TV commercial required painstaking, frame-by-frame post-production work by skilled technicians. There's a big difference now, says Haseltine: "What used to take an hour [per video frame], now can be done in a sixtieth of a second." This dramatic speed-up means that manipulation can be done

in real time, on the fly, as a camera records or broadcasts. Not only can John Wayne, Fred Astaire or Saddam Hussein be virtually inserted into pre-produced ads, they could be inserted into, say, a live broadcast of The Drew Carey Show.

The combination of real-time, virtual insertion with existing and emerging post-production techniques opens up a world of manipulative opportunity. Consider Video Rewrite technology, which its developers at the Interval Corp. and the University of California, Berkeley first demonstrated publicly three years ago. With just a few minutes of video of someone talking, their system captures and stores a set of video snapshots of the way that a person's mouth-area looks and moves when saying different sets of sounds. Drawing from the resulting library of "visemes" makes it possible to depict the person seeming to say anything the producers dream up-including utterances that the subject wouldn't be caught dead saying.

In one test application, computer scientist Christoph Bregler, now of Stanford University, and colleagues digitized two minutes of public-domain footage of President John F. Kennedy speaking during the Cuban missile crisis in 1962. Using the resulting viseme library, the researchers created "animations" of Kennedy's mouth saying things he never said, among them, "I never met Forrest Gump." With technology like this, near-future political activists conceivably will be able to orchestrate webcasts of their opponents saying things that might make Howard Stern sound like a mensch.

Haseltine believes video manipulation techniques will quickly be carried to their logical extreme: "I can predict with absolute certainty," he says, "that one person sitting at a computer will be able to write a script, design characters, do the lighting and wardrobe, do all of the acting and

dialog, and post production, distribute it on a broadband network, do all of this on a laptop-and viewers won't know the difference."

The End of Authenticity

So far, the widely witnessed applications of real-time video manipulation have been in benign arenas like sports and entertainment. Already last year, however, the technology began diffusing beyond these venues into applications that raised eyebrows. Last fall, for instance, CBS hired PVI to virtually insert the network's familiar logo all over New York City-on buildings, billboards, fountains and other places-during broadcasts of the network's The Early Show. The New York Times ran a front-page story in January raising questions about the journalistic ethics of altering the appearance of what is really there.

The combination of real-time virtual insertion, cyber-puppeteering, video rewriting and other video manipulation technologies with a mass-media infrastructure that instantly delivers news video worldwide has some analysts worried. "Imagine you are the government of a hypothetical country that wants more international financial assistance," says George Washington University's Livingston. "You might send video of a remote area with people starving to death and it may never have happened," he says.

Haseltine agrees. "I'm amazed that we have not seen phony video," he says, before backpedaling a bit: "Maybe we have. Who would know?"

It's just the sort of scenario played out in the 1998 movie Wag the Dog, in which top presidential aides conspire with a Hollywood producer to televise a virtually crafted war between the United States and Albania to deflect attention from a budding Presidential scandal. Haseltine and

others wonder when reality will imitate art imitating reality.

The importance of the issue will only intensify as the technology becomes more accessible. What now typically requires an $80,000 box of electronics the size of a small refrigerator should soon be doable with a palm-sized card (and ultimately a single chip) that fits inside a commercial video recorder, according to Winarsky. "This will be available to people in Circuit City," he says. Consumer gear for virtual video insertion is likely to require a camcorder with a specialized image-processing card or chip. This hardware will take signals from the camera's electronic image sensors and convert them into a form that can be analyzed and manipulated in a computer using appropriate software-much as photo editors at newspapers use Adobe Photoshop and other programs to "clean up" digital image files. A home user might, for instance, insert absent family members into the latest reunion tape or remove strangers they would prefer not to be in the scene-bringing Soviet-style historical revisions right into the family den.

Combine the potential erosion of faith in video authenticity with the so-called "CNN effect" and the stage is set for deception to move the world in new ways. Livingston describes the CNN effect as the ability of mass media to go beyond merely reporting what is happening to actually influencing decision-makers as they consider military, international assistance and other national and international issues. "The CNN effect is real," says James Currie, professor of political science at the National Defense University at Fort McNair in Washington. "Every office you go into at the Pentagon has CNN on." And that means, he says, that a government, terrorist or advocacy group could set geopolitical events in motion on the strength of a few hours' worth of credibility achieved by distributing a snippet of well-doctored video.

With experience as an army reservist, as a staffer with a top-secret clearance on the Senate's Intelligence Committee, and as a legislative liaison for the Secretary of the Army, Currie has seen governmental decision-making and politicking up close. He is convinced that real-time video manipulation will be, or already is, in the hands of the military and intelligence communities. And while he has no evidence yet that any government or nongovernment organization has deployed video manipulation techniques, real-time or not, for political or military purposes, he has no problem conjuring up disinformation scenarios. For example, he says, consider the impact of a fabricated video that seemed to show Saddam Hussein "pouring himself a Scotch and taking a big drink of it. You could run it on Middle Eastern television and it would totally undermine his credibility with Islamic audiences."

For all the heavy breathing, however, some experts remain unconvinced that real-time video manipulation poses a real threat, no matter how good the technology gets. John Pike, an analyst of the intelligence community for the Federation of American Scientists in Washington, D.C., says the credibility risks are simply too great for governments or serious organizations to get caught attempting to spoof the public. And for the organizations that would be willing to risk it, says Pike, the news folks-knowing just what the technology can do-will become increasingly vigilant.