

# COMMUNICATIONS

CACM.ACM.ORG

OF THE

# ACM

09/08 VOL.51 NO.9

## **Enterprise Information Integration** and other tools for merging data

How Do I  
Model State?

Formal Methods

Design and Code  
Reviews In the  
Internet Age

Power Management

The Puzzle of Apple

Science Policy Isn't  
Always About Science

# Ubiquitous Video

*Scalable and distributed video coding offers the promise of two-way, real-time video.*

LARISSA, A BRAZILIAN foreign-language student studying in Tokyo, gets a call on her cell phone just as she arrives at her apartment after classes. She peers at the phone's display and sees her mother sitting in the living room of the family's home in São Paulo, plus a blinking blue dot indicating the call is a live, two-way video stream. Larissa flips open her phone.

"Mama, do you like my new haircut?" Larissa asks as she lets herself into her apartment. "Is it too short?"

"No, it looks terrific," says her mother. "I have some video of your father's birthday party. Please turn on your TV."

"Okay," replies Larissa, who points her cell phone at the 50-inch, flat-panel television on her living room wall and pushes a button. The television flashes awake, picks up the video stream from the phone, and displays a high-quality video of her family celebrating her father's 49th birthday at his favorite restaurant in São Paulo.

One phone call, one stream of information. The cell phone takes only the data it needs for its two-inch display while the 50-inch television monitor takes far more data for its greater resolution—all from the same video stream.

Welcome to the future world of scalable, distributed video.

Digital video coding compresses the original data into fewer bits while achieving a prescribed picture quality, which it accomplishes largely by eliminating redundancies. Image data for a static background object, for instance, is stored just once, with subsequent frames merely pointing back to the original and registering only incremental changes.

Today's video coding paradigm exploits temporal and spatial redundancies—think of them together as repetitive elements over time—with a series of predictions, a set of represen-



tations, and a slew of cosine calculations. The goals are to remove the details the human eye can't see (whether they're too fast, dark, or small), set aesthetic rules (such as color and aspect ratio), tailor the bit and frame rates for the highest picture quality at the lowest file size, and save as much bandwidth as possible.

A video stream is broken up into pictures that are not necessarily encoded in the order in which they are played back. Encoders append such commands as "for blocks 37-214, duplicate the same blocks in the last frame," and quantize the transform coefficients to control for the limitations of human visual perception. Finally, entropy coding acts to control the statistical redundancy of the resulting coded symbols.

It's not quite instant, but in fairly short order video encoders produce a digital video file, a fraction of its original size, for an iPod, laptop, or cell phone. And with advances in scalable and distributed video coding, two-way,

real-time video, such as Larissa's conversation with her mother, is becoming a reality.

## Robust Encoding

Hybrid coding, which leverages both the temporal/predictive and frequency domains, is the basis for most current video standards. It does the hard work at the encoding step, resulting in complex encoders but just basic decoders.

A downlink model of a few encoders serving many distributed decoders serves applications for TV and cable broadcasting and on-demand Web video very well, but it makes decoder complexity its focus. Today's challenge, on the other hand, is the proliferation of wireless mobile devices—from cell phones and Internet tablets to laptops—that rely on up-links to deliver data. This requires capable device-based encoders.

In addition to robust encoding, these emerging applications require improved compression and increased

resistance to packet losses. New scalable and distributed coding solutions promise to deliver all of this—and much more.

A pair of Swiss-based standards organizations, International Organization for Standardization and International Telecommunication Union (ITU), formed a Joint Video Team in 2001 to develop a network-friendly video standard. Completed in 2003 and subsequently refined, H.264/AVC (Advanced Video Coding) attained measurably superior performance over existing standards. With the uplink model in ascendancy, there is continuing development in two promising areas: scalable video coding (SVC), an extension of the H.264/AVC standard, and distributed video coding (DVC).

An example of video scalability is when a “server has this 20Mbps coded video and you have a connection that can deliver 10Mbps,” says Gary Sullivan, a video architect with Microsoft and chair of the ITU-T Video Coding Experts Group. “If the video is encoded in a scalable way, the server can take just the subset of the data that represents the lower quality and give you that.”

Video data is delivered in packets, and if the video is not coded in a scalable manner, there’s basically very little a person can do other than decode all of them, notes Sullivan. However, if the video is encoded in a scalable way, then some packets belong to the base layer and some packets belong to the enhancement layer. Sullivan muses that it’s possible to create a bitstream with 10 layers, covering a wide range of decoders. “It’s a nice concept, but

**“We know where [distributed video coding] may arrive from a theoretical point of view, but we still don’t know how to arrive there in practice,” says Fernando Pereira.**

has been difficult to achieve,” he says.

#### Charting a New Course

A professor at the Electrical and Computers Engineering Department at Portugal’s Instituto Superior Técnico and the chair of many ad hoc video standards groups, Fernando Pereira is trying to chart video’s course from scalable to distributed. Not only will there be the multiple layers from SVC, but the new distributed video encoding will dynamically divvy up the work between encoders and decoders.

Pereira likens progress in the field of video coding to paleontologist Stephen Jay Gould’s description of “punctuated equilibrium” in evolution during which periods of stasis are interrupted by flurries of “creative destruction” and rapid change.

Video coding’s state of the art in

the early 1970s was represented by the Slepian-Wolf theorem that describes lossless coding—a way to reduce file sizes without losing any bits—with rather small compression factors. By 1976, Abraham Wyner and Jacob Ziv had derived the Wyner-Ziv theorem that essentially defines the conditions under which the picture quality can be achieved even when the coding process is not lossless.

Because it does not delete irrelevant information, the Slepian-Wolf theorem by itself has little practical application in video compression today. However, the Slepian-Wolf and Wyner-Ziv theorems together suggest the potential to compress two signals in a distributed way, with two separate encoders supplying a single joint decoder, says Pereira. He is confident this approach can achieve “a coding efficiency close to that of the predictive, joint encoding and decoding schemes” now in widespread use.

As opposed to conventional coding, in DVC the task of motion estimation is performed only on the decoder side to generate motion-compensated predictions for each input frame. The coding efficiency of a DVC scheme is judged to a great degree on the quality of these predictions.

The new DVC model promises substantial advantages for existing and emerging applications. They include flexible resources (DVC allocates varying amounts of encoder complexity to the decoder, which results in low encoder complexity and low battery consumption), improved resilience (DVC codecs do not rely on repetitive prediction loops, so channel in-

#### Workplace Technology

## Keeping Focus Amid the Distractions

Today’s cubicle dweller switches gears every three minutes, moving from one task to *anything* else bombarding his or her attention. Indeed, the office environment has become such a breeding ground for interruptive technologies like emailing, cybersurfing, and IMing, that disruptions are now consuming as much as 28% of the average U.S. worker’s day and

sap productivity by as much as \$650 billion a year, according to Manhattan-based business research firm Basex.

In response, a growing number of tools and technologies are emerging to help keep workers on task. Microsoft is working on some remedies for attention disruption disorder that include AI systems that observe humans at work and

build models that predict the cost and benefit of interrupting someone, *BusinessWeek* reports. A prototype of an email triage program called Priorities ranks messages in order of perceived importance. The Outlook Mobile Manager enables Outlook to recognize urgent messages. And Bounded Deferral holds messages until a recipient is ready for a “cognitive break.”

IBM is also on the attention management track, now testing a prototype IM answering machine known as IMSavvy that can “sense” when a worker is too busy to answer calls or messages and will relay that sentiment to would-be interrupters. The system also offers a whisper option, flickering text on a worker’s screen even if the worker has instructed the system to withhold messages.

terference errors do not propagate over time), multiview independence (when used in a multiview video context, DVC encoders do not jointly process multiple views and thus do not need inter-camera, inter-encoder communication, saving energy), and codec-independent scalability (in current scalable codecs, a prediction approach from lower to upper layers requires the encoder to know the coding solutions for each layer, and the DVC approach allows each layer to use a discrete codec, unknown to the encoder, as knowledge of every layer is no longer necessary).

These benefits will positively impact video-related applications such as mobile videoconferencing and video email. "The future will tell us in which application domain the dis-

tributed source coding principles will find success," says Pereira.

### Video Everywhere


Although Pereira sees important roles for academia and industry, "DVC is still very much an academic exercise with very few companies involved," he says. "MPEG [the family of standards used for coding audiovisual information] is not involved at all because it is too early to think about any standardization, and we still don't know what the best solution may be."

"With the continuing convergence of Internet, cable-based technologies, and wireless, bandwidth should also increase and we'll be seeing more on-demand and live video applications very soon," says Kevin Bee, CEO of Uptime Video, a video encoding firm based in Thousand Oaks, CA. This growing convergence has already led Adobe to include H.264 compatibility in its Flash Player 9, a move that has exponentially extended the codec's reach.

"We know where DVC may arrive from a theoretical point of view, but we still don't know how to arrive there in practice," says Pereira.

Sullivan concurs. "H.264 itself gets easier to implement over time, but it will take a lot of work to make a better compression-capable codec," he says. "We're not there yet, and won't be for several years at least."

One major area of scientific research is human cognition. "Audio people had to enter this area earlier and deeper because the amount of redundancy in audio is much lower than in video, and they had to deal with irrelevancy in a more efficient way," says Pereira. Clearly, he concludes, a better understanding of visual perception and the manner in which the human visual system responds to compression are among the most important next steps.

"The bottom line is that it is time for research and hard work," says Pereira. "We should not go too fast in terms of making products so as to avoid 'killing the goose that laid the golden egg.' But, honestly, I don't even know if there's a goose yet." 

Logan Kugler is a Los Angeles-based freelance writer who writes about business and technology.

## Information Technology

# Goodbye, Computer Mouse

The days of the computer mouse are nearing their end, Gartner analyst Steve Prentice has told the BBC News. Prentice expects the functionality of the mouse to be gradually replaced during the next three to five years by emerging alternative user interfaces that rely on facial recognition, movement, and gestures.

Prentice says the mouse has staying power in the desktop computer environment, but believes that "for home entertainment or working on a notebook, it's over."

He notes that Apple, Intel, and Microsoft are now promoting gestural interfaces for future computer use and that NEC, Panasonic, and Sony are demonstrating applications that use facial and movement recognition.

"With the [Nintendo] Wii you point and shake and it vibrates back at you so you have a two-way relationship there," says Prentice. "The new generation of smart phones like the iPhone all have tilting mechanisms or you can shake the device to do one or more things. Even the multi-touch interface is so much more powerful and flexible than in the past allowing you to zoom in, scroll quickly, or contract things."

Of course, not everyone agrees with Prentice. "The death of the computer mouse is greatly exaggerated," says Rory Dooley, senior vice president and general manager of Logitech's control devices unit, who notes that much of the developing world has still to get online. "There are around one billion people online, but the world's population is over five billion.... The mouse will be even more popular than it is today as a result."

Invented by Douglas Engelbart of the Stanford Research Institute, the computer mouse will celebrate its 40th anniversary later this year. Engelbart, recipient of the 1997 ACM A.M. Turing Award, never received any royalties for his invention, in part due to its patent expiring in 1987 before the widespread popularity of personal computers.

## Coming Next Month in COMMUNICATIONS

*An editorial debate on e-voting*

*Natural computing*

*Topology of dark networks*

*Code spelunking*

*A conversation with Pat Selinger*

*GPU architectures*

*Scene completion using millions of photographs*

*Geometry, flows, and graph partitioning algorithms*

*The enrollment crisis in Canada*

*Q&A with Daphne Koller, winner of the first ACM-Infosys Foundation Award*

**And the latest news on E2E systems and cryptography, green computing, and the Deep Web**