Cheung Gene, Assistant Professor,

Digital Content and Media Sciences Research Division

Seeking to build interactive multimedia systems that offer outstanding compression
and data distribution performance

My area of expertise is signal processing, and I look at the entire communication
process to help optimize, design, and develop systems like free viewpoint television
(FTV). For FTV, video has to be generated, compressed, and transmitted to suit a
particular perspective selected by any of a wide range of different users. I expect
dramatic changes in the familiar TV sets currently found in every household.

## Making TV interactive

The growing presence of mobile digital devices has led to growing use of and exposure to
video data. The video data I work with is multi-view video. This video is recorded using
multiple cameras simultaneously. The camera angle seen by the viewer is selected not
by the sender, but by the viewer—interactively, via the viewer's TV. This opens up the
possibilities for future TV, as viewers will be able to enjoy video viewable from various
perspectives—like in the movie "Matrix"—and follow one player in a soccer match no
matter where the ball happens to be, or zoom in on the fingerwork of cellist Yo-Yo Ma as
he plays.

## FTV – Using a hundred cameras to create video to suit individual viewers

FTV is one of the systems currently gaining attention worldwide for next-generation TV,
and standards are beginning to take shape. Current 3D TV systems use two cameras to
create left- and right-eye images. In contrast, FTV deploys as many as 100
inter-networked cameras to allow the recreation of a near-infinite number of viewpoints.
In place of model-based analysis, video is generated and compressed by computing the
inter-relationships among pixels that make up images. This makes it suitable for any

type of scene—even indistinct images like rain or mist. Computation is relatively inexpensive, making it ideal for situations requiring real-time communication, like live broadcasts.

### An alternate reality for communication

Since FTV involves vast volumes of data, the capacity to transfer high-resolution video at high compression ratios with little quality degradation across a wide range of networks is critical.   Video conferencing participants detect and express frustration with delays as short as 100 milliseconds. System designs have to account for the entire communication process, first eliminating redundant parts of the video for compression gain, then injecting   some redundancy back into the data on purpose to compensate for potential data loss.

FTV will open up entirely new vistas in communication and allow modes of expression that transcend the bounds of reality. Examples include avatars with realistic facial expressions that replace real faces in video conferencing and technology that allows bed-ridden patients to experience the thrill of playing tennis. The list of potential applications is also growing in the area of education—for example, university lectures.

Interviewed and summarized by Rue Ikeya