

# Brave NUI World

*Natural user interface developments, such as Microsoft's Kinect, may indicate the beginning of the end for the mouse.*

**C**OLORADO SPRINGS, CO.-BASED independent software developer Kevin Connolly became a minor YouTube celebrity over the past several months with the demonstration of his natural user interface (NUI) hack of Microsoft's Kinect software development kit (SDK). Connolly demonstrated moving different images on a small bank of screens up and down, in and out, and sorted through three-dimensional image arrays, all through gesture alone. It was, he notes, very similar to the image manipulation featured in Steven Spielberg's 2002 futuristic film, *Minority Report*.

"I'm just some guy," Connolly says. "I made that work in a matter of hours. Imagine what we have the technology to do if one guy in his apartment can do that in a few hours."

Indeed, Microsoft's decision to release the Kinect SDK in June has garnered much attention in the technical and technology trade press and among researchers and enthusiasts. Anoop Gupta, a Microsoft distinguished scientist, says more than 100,000 individuals downloaded the SDK in the first six weeks after its release. However, the terms of the release forbid any commercial use of the SDK, and Connolly says he halted work on his nascent NUI after he got it working to his satisfaction. Yet, the release of the SDK was a signal that low-cost motion and depth-sensing technology may soon herald epochal changes in the way humans and computers interact. The Kinect hardware, for instance, is manufactured by Tel Aviv-based PrimeSense, and lists for about \$200. Researchers in numerous disciplines call such technology for such a price absolutely revolutionary.

## NUI Research

John Underkoffler's experience with NUI research goes far beyond hobbyist SDKs. In fact, Underkoffler is



**Audrey Penven created this photograph and similar ones by using Kinect's infrared structured light as a source of light.**

the man behind the *Minority Report* interface and has commercialized it via Los Angeles-based Oblong Industries, at which he is chief scientist. Underkoffler calls the Microsoft SDK release a "rhetorical event we all love." He says these events, like the 2006 release of Nintendo's Wii, "puts in the foreground for different sets of eyes—the end consumer for the Wii, the home or dorm room hacker for the Kinect SDK—the idea that it isn't going to be mouse and keyboard forever. We've seen the dialogue go from one of doubt or questioning to a kind of acceptance. Everyone now knows it isn't going to be mouse and keyboard forever, but the real question is, What is it going to be?"

Some of the next steps of NUI research are likely to be through gaming applications, which are increasingly dependent upon motion-sensing technology. Recently, four California Insti-

tute of Technology students undertook a study of motion-sensing technologies originating in gaming, and their possible uses in other computationally intensive fields. The students explored the depth camera technology in Kinect, the inertial sensors of the Wii, and electromagnetic sensing technology developed by Sixense Entertainment and Razer. One of the students, Peter Ngo, believes the idea of the NUI as a fully toolless interface is overstated, as does Amir Rubin, CEO of Sixense.

"People don't buy motion control," Rubin says. "They don't buy PCs. They don't buy consoles. They buy the experience being delivered to them. The best input device is something you don't remember is on you, and that is true even for the slogan sold by Microsoft—'You are the controller.' I agree, but if I have to alter the way I use my body in order to interact with Kinect, then it's the worst controller."

Even the most ardent NUI advocates agree with Rubin. Oblong's Underkoffler, for instance, says that writers will likely be well served by the keyboard for the foreseeable future, but that those who design ship hulls or airplane wings would be better served by three-dimensional NUIs. Moreover, he says it is vital not to graft a notion of a new interface design by simply extending two-dimensional GUI concepts onto the prototypes of three-dimensional applications. These applications will need computational capabilities along not just the flat  $x$  and  $y$  axes—an example might be a wall-sized but still two-dimensional application for designing the very three-dimensional ship's hull he mentioned—but will also need to compute the depth of the  $z$  axis. Oblong's  $g$ -speak platform, based on work Underkoffler pioneered in the 1990s at the Massachusetts Institute of Technology's Media Lab, computes this spatial environment via networked computers and screens that allow rich three-dimensional interaction. Ultimately, Underkoffler thinks a hybrid UI ecosystem will evolve.

"We're not out to replace the keyboard; let it do what it's best at," he says, "but when it comes to designing airplane wings, you do need a spatial UI. So, it's about situating the right activities in the right interaction zone."

### Homebrewed Algorithms

Since the Microsoft SDK precludes commercial use, many early academic and enterprise projects using PrimeSense and/or stripped down Kinect hardware use either homebrewed algorithms or open source drivers and middleware released by consortia such as OpenCV or OpenNI, the natural interface forge formed in November 2010, by PrimeSense and robotics pioneer Willow Garage. OpenNI leverages the PrimeSense depth-sensing technology, which is processed in parallel by PrimeSense's system-on-a-chip processor after receiving coded near-infrared light from its partnered CMOS sensor.

OpenNI supplies a set of APIs to be implemented by the sensor devices, and a set of APIs to be implemented by the middleware components. Thus, OpenNI's API enables applications to be written and ported to operate on

**"People don't buy motion control," Amir Rubin says. "They buy the experience being delivered to them. The best input device is something you don't remember is on you."**

top of different middleware modules. It also enables middleware developers to write algorithms on top of raw data formats, regardless of which sensor device has produced them, and offers sensor manufacturers the capability to build sensors that power any OpenNI-compliant application.

In fact, one nascent healthcare industry application partially built on open source stacks by a team of surgeons and engineers at Sunnybrook Health Sciences Center in Toronto for the Kinect camera is already drawing attention.

Allowing surgeons access to medical images while not having to touch a controller—and thereby saving them the necessity to re-sterilize in order to preserve sterility around the patient—is an early enterprise triumph for the NUI concept. Computer vision specialist Jamie Tremaine says the gesture-based UI he and his colleagues developed has proven exceptionally robust and enables surgeons to view through MRI and CT scan samples that can run from 4,000 to 10,000 slides without ever having to re-sterilize.

For such an application, the hand and arm gestures recognized by the Kinect camera are suitable, but Tremaine says "a lot of the work we've done hasn't even been on the technical side as much as creating gestures in the operating room that allow very fine-grained control, but which have to be larger."

Another NUI developer, Evan Lang of Seattle-based UI design firm IdentityMine, says his work with trying to develop Kinect NUIs (on PrimeSense drivers) similar to current GUI commands revealed vexing user issues. In developing a Web button, for instance, Lang says, "I programmed it to recognize a poking gesture, where you move your hand quickly forward and quickly back. When I got some test users to try it out, and said 'Poke it or press it,' everybody had a very different idea of what that actually meant. Some did a kind of poking thing. Other people moved their hand forward but wouldn't move it back, and others, who were very cautious and deliberate about it, the machine wouldn't register as a poke."

Oblong's Underkoffler says problems such as Lang encountered are emblematic of grafting current GUI-based mechanics on an idea that needs something else.

"We believe it's not appropriate to start talking about NUIs until you have a complete solution," Underkoffler says. "If you flash back 30 years, it's like dropping an early prototype of a mouse in everybody's lap and saying, 'We have a new interface.' You don't, because you've just got a new input device. So, really it's a full loop proposition. What's the input modality? What shows up on screen? What's the analogue of the windows and scroll bars and radio buttons? Until that's not only been answered in a way to allow real work to happen, but has become kind of a standard, and more in the cognitive sense, recognizably and pervasively present, then you don't have a new interface."

Sixense's Rubin predicts the next-generation standard UI device would not be a question of which technology is most elegant, but rather, that which meets three criteria: a consumer-friendly price, an intuitive UI design, and ease of software development on top of that device.

"If you can meet the combination of those three," he says, "then you will have the next-generation standard of input devices."

### Output Perceptions

Robotics researchers such as Nicholas Roy, associate professor of aeronau-

## Allowing surgeons access to medical images while not having to touch a controller—and thereby saving them the need to re-sterilize in order to preserve sterility around the patient—is an early enterprise triumph for the NUI concept.

tics at the Massachusetts Institute of Technology, were among the first to adopt the PrimeSense sensor, and it may be their work that shows the longest-range potential for, and a new concept of, what a NUI will be.

Vehicles equipped with the three-dimensional sensors—among them a helicopter Roy and his students programmed—gain what Roy calls a “whole new sense of the human-centered environment,” and are able to sense things such as drop-offs in stairs, table legs, and so on. And now, with more researchers exploiting the extremely cheap sensor technology, it is likely that more UI work will be done exploring how a robot, or any other computer, will “think” its way to interacting with humans.

“A lot of the estimation and planning algorithms my students have developed for the helicopter, we actually reused in the context of interacting with robots,” Roy says. “If the problem is no longer how the vehicle plans to get from Point A to Point B, but the problem is how does the vehicle understand what the human wants in terms of some task, then our research programs have a lot of commonality between those two seemingly very different domains.”

### Further Reading

Cohn, G., Morris, D., Patel, S.N., and Tan, D.S. Your noise is my command: Sensing

gestures using the body as an antenna, *Proceedings of the 2011 Annual Conference on human factors in computing systems*, Vancouver, British Columbia, Canada, May 7–12, 2011.

Gallo, L., Placitelli, A.P., and Ciampi, M. Controller-free exploration of medical image data: Experiencing the Kinect, *Proceedings of the 24<sup>th</sup> IEEE International Symposium on Computer-Based Medical Systems*, Los Alamitos, CA, June 27–30, 2011.

Henry, P., Krainin, M., Herbst, E., Ren, X., and Fox, D.

RGB-D mapping: Using depth cameras for dense 3D modeling of indoor environments, *Proceedings of the International Symposium on Experimental Robotics*, New Delhi, India, Dec. 18–21, 2010.

Shotton, J., et al.

Real-time human pose recognition in parts from single depth images, *Proceedings of the 24<sup>th</sup> IEEE Computer Vision and Pattern Recognition*, Colorado Springs, CO, June 20–25, 2011.

Underkoffler, J., Ullmer, B., and Ishii, H.

Emancipated pixels: Real-world graphics in the luminous room, *Proceedings of the Special Interest Group on Computer Graphics*, Los Angeles, CA, August 8–13, 1999.

Gregory Goth is an Oakville, CT-based writer who specializes in science and technology.

© 2011 ACM 0001-0782/11/12 \$10.00

## Milestones

# Katayanagi Prizes and Other CS Awards

President Obama, ACM, IEEE, Carnegie Mellon University, and Tokyo University of Technology recently honored some of the nation’s leading computer scientists.

### NATIONAL MEDALS OF SCIENCE, TECHNOLOGY, AND INNOVATION

President Obama honored Richard A. Tapia, director of the Center for Excellence and Equity in Education at Rice University, with the 2010 National Medal of Science for “his pioneering and fundamental contributions in optimization theory and numerical analysis and for his dedication and sustained efforts in fostering diversity and excellence in mathematics and science education.”

The President also presented National Medals of Technology and Innovation, recognizing scientists who have made lasting

contributions to America’s competitiveness and quality of life and helped strengthen the nation’s technological work force. The recipients were Rakesh Agrawal, Purdue University; B. Jayant Baliga, North Carolina State University; C. Donald Bateman, Honeywell; Yvonne C. Brill, RCA Astro Electronics (retired); and Michael F. Tompsett, TheraManager.

### KATAYANAGI PRIZES

Carnegie Mellon University in cooperation with the Tokyo University of Technology presented the Katayanagi Prize for Research Excellence to Barbara Liskov, Institute Professor at the Massachusetts Institute of Technology, for her record of outstanding, sustained achievement. Scott Klemmer, associate professor of computer science at Stanford University, received the Katayanagi

Emerging Leadership Prize for his demonstration of leadership promise.

### ACM AND IEEE AWARDS

Susan L. Graham, the Pehong Chen Distinguished Professor of Electrical Engineering and Computer Science Emerita of the University of California Berkeley, received the ACM-IEEE Computer Society Ken Kennedy Award “for contributions to computer programming tools that have significantly advanced software development.” Cleve Moler, chairman of MathWorks, was this year’s winner of the IEEE Computer Society Sidney Fernbach Award “for fundamental contributions to linear algebra, mathematical software, and enabling tools for computational science.” Charles Seitz, former president of Myricom, Inc., was awarded the IEEE Computer Society’s 2011 Seymour Cray

Computer Engineering Award “for innovations in high-performance message-passing architectures and networks.”

### SIGSAC AWARDS

ACM’s Special Interest Group on Security, Audit and Control (SIGSAC) presented its top honors to Virgil Gligor, co-director of Carnegie Mellon University’s CyLab, and Ravishankar Iyer, director of the Center for Reliable and High-Performance Computing at the University of Illinois. Gligor received the SIGSAC Outstanding Innovation Award for innovations in secure operating systems as well as covert channel analysis, intrusion detection, and secure wireless sensor networks. Iyer received the SIGSAC Outstanding Contributions Award for his fundamental contributions to the assessment and design of secure, dependable computing systems.