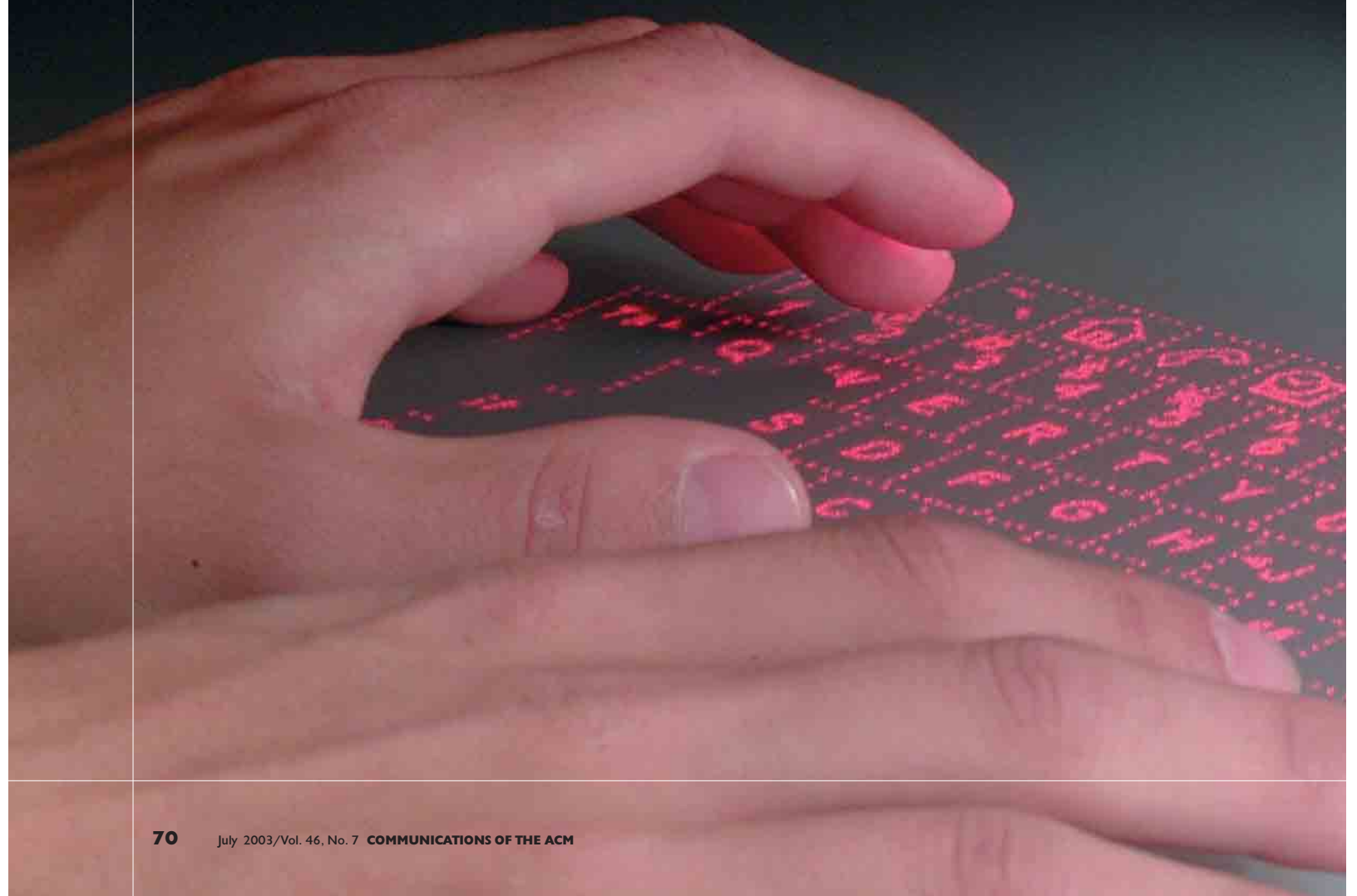


[By Carlo Tomasi, Abbas Rafii, and Ilhami Torunoglu]

FULL-SIZE PROJECTED KEYBOARD FOR HANDHELD DEVICES

A KEYBOARD MADE ENTIRELY OF LIGHT PROJECTED ONTO DESKTOPS, AIRPLANE TRAY TABLES, EVEN KITCHEN COUNTERS FUNCTIONS, FEELS, AND SOUNDS LIKE ITS MECHANICAL COUNTERPART. NEXT: PERHAPS ON LAPS AND IN THE AIR.



T ION IC ES



FULLY FUNCTIONAL,
STANDALONE PROTOTYPE OF
THE PROJECTION KEYBOARD.
THE LIGHT IS RED BECAUSE ONLY
AT THIS WAVELENGTH COULD ITS
DEVELOPERS PRODUCE A
SUFFICIENTLY INEXPENSIVE AND
POWERFUL LASER SOURCE
(CANESTA, INC., SAN JOSE, CA).

DESPITE THE SOPHISTICATION OF THE TECHNOLOGY, interacting with today's computers, cell phones, and personal digital assistants can involve painful contortions. Miniature displays and keyboards make some portable devices amazingly small, but users' hands do not shrink accordingly; neither does their eyesight sharpen to match postage-stamp-size displays. Here, we address a significant part of the problem: the keyboard. Beginning in 1999, one of us (Rafii) proposed using a single tiny sensor to observe the user's fingers, transforming motions into keystrokes. The idea was to make a keyboard out of light, projecting it onto desktops, airplane tray tables, even kitchen counters. Suddenly, there would be nothing to carry around. The keyboard would not be stowed or folded away—merely switched off. A few extra grams and a few more cubic millimeters would be integrated into users' devices of choice; typing would be as fast as their fingers allowed—on a full-size keyboard.

The projection keyboard we've been developing at Canesta over the past four years will soon be available in cell phones and PDAs. It might also replace laptops' mechanical keyboards, making the machines even thinner and lighter. It will be an unbreakable, washable keyboard or keypad, projected or etched onto flat surfaces in hospital operating rooms, automatic teller machines, pharmaceutical manufacturers' clean rooms, and space vehicles. Where projector size is not an issue, a scaled-up version can be used in interactive electronic whiteboards and wall-mounted maps and allow keyboards to be reconfigured on the fly.

In the process, we have developed sensors and software that can be used for much more than keyboards, and an array of possible applications now seems possi-

ble. The projection keyboard's sensing and interpretation system, designed by one of us (Tomasi), is innovatively simple, but the technical challenges we've faced from concept to working application have been profound. We started with clumsy emulation systems, duct tape, and a \$3,000 prototype. We've now reduced the system to three components—a projector, an infrared light source, and a sensing system—each about the size of a pencil eraser (see Figure 1). Together, they cost (to a user) less than a folding mechanical keyboard,

WE HAD TO DEVISE DYNAMIC KEYSTROKE-DETECTION ALGORITHMS THAT TRACK THE LANDING HISTORY OF FINGERS IN THE FEW MILLISECONDS BEFORE IMPACT.

and draw less power than a cell phone. The resulting system “feels” almost like a mechanical keyboard, even if users feel only the impact of their fingers on the projection surface when typing.

Technical Aspects

Early on, we realized we could not use off-the-shelf sensors and components. The requirements were

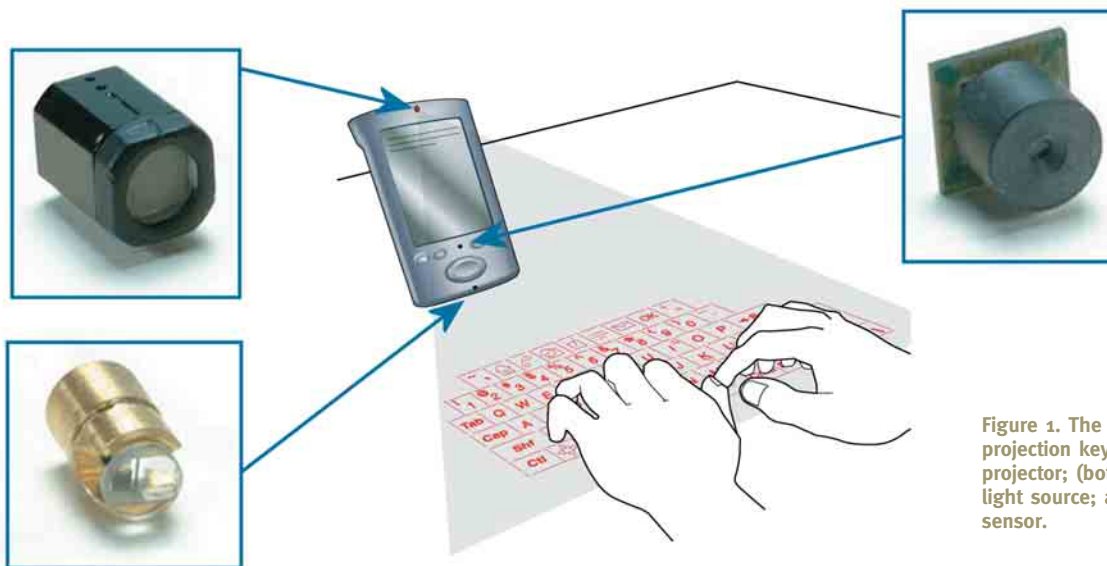


Figure 1. The components of the projection keyboard: (top left) projector; (bottom left) sensor light source; and (top right) sensor.

too daunting; for example, the keyboard projector would be unable to use photographic slides (they're optically inefficient). For sensing, ambient light is unpredictable, so the keyboard would need its own light source but would have to compete with daylight on only a few milliwatts of power. Moreover, due to the shallow, close-up viewing angle, the optics of both the projector and the sensor would be pushed to extremes. Our only option was to build everything from scratch. In the process, we designed sensors based on a variety of principles, from the special-purpose structured-light devices called "beam trangularators"¹ invented by one of us (Tomas) to methods based on the measurement of time [1] and phase¹ of returning light pulses invented by Canesta's Cyrus Bamji. As a result, we now have a battery of sensors, each suited to different work-spaces and operating conditions, though all are small, low-power, and inexpensive. For the current version of the keyboard (introduced in July 2002), we use Tomas's beam trangularators. Meanwhile, we continue experimenting with the other techniques in applications requiring more extended depth maps.

As outlined in Figure 1, the optical system projects the keyboard onto the typing surface; its infrared light source generates an invisible, fan-shaped beam grazing the surface; and its sensing system includes a processing unit. All are in fixed relative positions; no user adjustment is required.

The projector is positioned high on the host device, as shown in the figure. In the simplest sensor, a camera looks down at the typing surface at a shallow angle through a wide lens sensitive only to infrared light; the figure shows the positioning of the light source and camera. A finger striking the typing surface breaks the infrared beam, thereby becoming visible to the camera; triangulation then determines the position of the finger on the typing surface.

Even with this simple sensor, implementation of the projection keyboard under the constraints of weight, size, power, cost, reliability, and usability has involved formidable technical challenges:

Keyboard projection. Most current optical-projection systems mask light to form images at a distance. For instance, in a slide projector, the slide itself blocks part of the light hitting it; the remaining light makes it through the lens and onto the screen, forming the image. A similar concept applies to an LCD or micro-

mirror projector in which variable-transparency elements or mirrors modulate light.

All these systems are fundamentally inefficient because the light being masked away is irretrievably lost. With the image of a typical keyboard, more than 80% of the image is dark, and only 20% of the light entering a masking system reaches the typing surface. For a system running on the batteries of a portable host device, this is unacceptably inefficient. Our keyboard projector instead uses a holographic diffraction pattern to diffract a single beam of laser light directly into the pattern to be projected (for a survey of this technology, see [3]), achieving much greater efficiency.

When we first started developing the keyboard we

thought that fingers occluding the projected pattern would mean entire regions of the keyboard would disappear from the table. But as soon as we switched on our first prototype, we realized with relief that this thought was a naive

Typing speeds and error rates for users proficient in each of these keyboard methods.

Keyboard Type	Words per Minute	Error Rate
Thumb	20–30	1–4%
Projection	30–60	3–5%
Mechanical	50–80	2–3%

mistake. All users' hands block the regions behind their fingertips from their view with projected and mechanical keyboards alike. Our usability tests confirm the unimportance of these blind spots, as users have not noted projection obstruction as a significant issue.

Finger lighting and sensing. When typing, a user's useful action occurs in the thin layer of space that separates the hovering and constantly moving fingers from the surface of the keyboard. In a perfect world, an infinitesimally thin sheet of infrared light grazing the typing surface could be used as a trip switch; when a finger intersects the beam, it would also be touching the surface. In this instance, a finger that becomes visible to the camera is a finger that hits a key, and all that is left for the sensor is determining the finger's position by triangulation.

Our earliest attempts to build a projection keyboard followed this principle. However, the alignment requirements for the sheet of light proved too demanding; even in our lab we would lose calibration after a few days. The requirements were obviously unrealistic for a mass-produced device intended to work for years in unpredictable environments and on uneven surfaces. Our solution was to light a thicker, carefully shaped slab of space fanning out horizontally from the beam generator and over the typing surface to cover the sensor's wide field of view. However, though the thicker beam tolerates misalignment, the sensing and processing system would now have to sort out the complicated events occurring in this space.

¹U.S. patent pending.

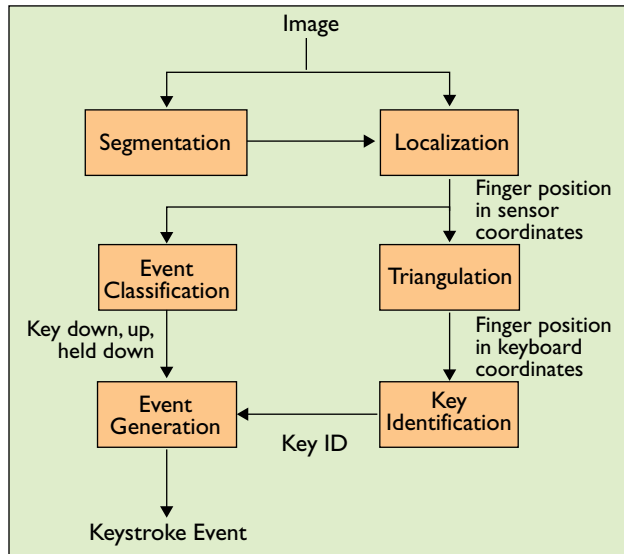


Figure 2. Flow diagram of the keystroke interpretation algorithm.

Here, up to 10 fingers are hovering, typing, or just lingering, often moving in sympathy with the fingers doing the actual typing. As a result, we had to make our sensors faster, devising dynamic keystroke-detection algorithms to track the landing history of fingers in the few milliseconds before impact.

Another fundamental difficulty was soon evident in the darkness of our lab. In the real world, the dim light of the infrared beam would have to compete like David against the Goliath of ambient light. We therefore combined several variations on the theme of matched filtering. The sensor lens, made of plastic, is colored with a dye that blocks all but a very narrow spectral band around the frequency of the beam-generator light, or spectral matched-filtering. In addition, the sensor's electronic shutter is opened in synchrony with short and powerful pulses from the beam generator, thereby realizing a temporal matched filter. Finally, the embedded software processes the images with state-of-the-art signal-detection techniques [6] to extract from a noisy background the dynamics of dim blobs. Dynamics include matched filtering in shape, signal intensity, and trajectory space. The combination of these techniques, along with relentless efforts by one of us (Torunoglu) to optimize the code, now allows a modest amount of power from the host device to prevail over the flood of ambient light. After months of darkness, we were able to open the windows and type in sunlight.

Interpretation. A finger that strikes a surface moves at velocities close to 40 cm/sec., traversing the entire thickness of the infrared beam in less than 10 millisec. Other fingers on the same hand move in sympathy with the main active finger, often stopping just one or two millimeters from the typing surface. Due to fil-

tering, the camera does not see the surface but infers finger-to-surface contact (or lack thereof) from the motion of the fingers, as well as more specifically from the history of the height of each blob in the image over several frames. This inference, in turn, requires that blob identity be maintained over time, accomplished through a matching and state-estimation algorithm based on dynamic programming.

Fingers occluding one another complicate the picture, even if the vertical placement of the camera and beam generator minimizes the problem. When two keys are held down simultaneously (such as Shift-q to obtain a capital Q), the finger closer to the camera may hide the one behind it.

Life just above a keyboard is clearly very complicated, so one-finger studies (such as [3]), even while achieving acceptable error rates for an input device, are only the beginning of what needs to be done to make a projection keyboard really work. An additional software layer of reasoning is required. Moreover, some keys must be “sticky” in certain circumstances.

The candidate fingers are identified and segmented from the other possible background objects in the image, as outlined in Figure 2. Localization determines the position and time of a keystroke. Event classification determines the type of action: landing, hold, move, and takeoff. Triangulation transforms image points into keyboard positions a table then maps to the identity of the key associated with that position. Key identity and event type determine the appropriate keyboard event.

Mapping from coordinates to keyboard events is implemented through reconfigurable tables, enabling applications in which the keyboard layout changes dynamically. To facilitate working with different layouts, we have developed an interactive designer software tool for creating new layouts and downloading them into a device.

Careful power management is essential for a system that runs on the limited-capacity batteries of cell phones and similar host systems. The projection keyboard is turned off when not in use for a long time and dimmed during shorter periods of inactivity. Sensing can be turned off much more frequently and abruptly without the user noticing. When the sensor is in power-save mode, it passively monitors its environment, using minimal computational resources. As soon as it detects typing activity, the sensing system springs back to life to interpret keyboard events.

Usability

With a mechanical keyboard, a key being pressed delivers a variety of haptic sensations to the typist. A descending finger senses contact with the key, over-

comes a counteracting spring force, travels further downward beyond that point of resistance, then stops when the key reaches the bottom of its stroke. In addition, the sharp edge of a key discreetly informs the finger when it hits the key away from its center, as well as its direction. Familiar sounds of physical typing accompany all these events, and their nature and the users' reactions to them are well understood [2]. With a projection keyboard, sound feedback is still possible, and contact with a key carries the information of impact with an inert surface. Despite the projection keyboard having less feedback, our usability tests were a pleasant surprise.

The literature, including [4], has established that sound can substitute for or reinforce haptic feedback. We found that a faint click generated electronically upon recognition of a keystroke is an enormous help and markedly increases typing speed. We also experimented with modulating the quality of the sound as a function of where a key is hit. Although this trick has hardly any effect on typing speed, users learn from it and tend to type more reliably (their fingers drift less), even if they do not look at the keyboard.

We've observed that users get accustomed to typing on the projection keyboard in a few minutes, reaching their steady-state speed after an average of 15 minutes of practice. Typing is not quite as rapid as on a mechanical keyboard but easily beats thumb-operated keyboards; the table here compares the typing speeds and error rates for users proficient in each of the keyboard methods listed.

Practicing on a projection keyboard takes place on a familiar-looking interface, so instructions are not necessary. Though users' typing is slower in their first few minutes on the device, they do nevertheless type from the start, employing the keyboard for useful tasks with their first keystrokes. Moreover, users experience less fatigue than with mechanical keyboards; only a light touch is required to activate a key. However, on the projection keyboard today, idle fingers must hover, a possible source of fatigue for touch typists. We are experimenting with keystroke-detection algorithms that allow fingers to rest.

Next Steps

We face at least two fundamental physical-interaction design challenges: enabling users to type on arbitrary surfaces, say, directly on their laps while sitting on chairs in waiting rooms; and enabling users to type in midair, thereby obviating the need for a typing surface altogether. Not only do they involve technology but the user's perception of the naturalness of using the system as well.

In another direction, how can we predict which

technology will enable us to deliver a practical, fully dynamic projection technology? Special eye-mounted miniature displays are a promising option. However, a system based on projecting an interface onto a surface may be less awkward to use, at least for text, as it would obviate the need to wear special contraptions.

Small, inexpensive, power-thrifty sensors like the ones we are building are just beginning to open a world of electronic perception technology to exploration. The projection keyboard is only the first of many embedded applications to come from this technology. Other applications could include: automobile safety systems that detect the size and position of passengers so airbags are optimally deployed in crashes; video games and remote controls that interact with the user through gesture or body movement; and facial-recognition systems using 3D shapes to identify their subjects more accurately.

More than a decade ago, Mark Weiser of Xerox PARC, said, "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" [7]. This fundamental unobtrusiveness is the main metric of validity for the fledgling field of electronic perception technology. **C**

REFERENCES

1. Bamji, C. *CMOS-Compatible Three-Dimensional Image Sensor IC*. U.S. Patent 6,323,942, Nov. 2001.
2. Beringer, D. and Peterson, J. Underlying behavioral parameters of the operation of touch-input devices: Biases, models, and feedback. *Human Factors* 27, 4 (Apr. 1985), 445–458.
3. Mantyjarvi, J., Koivumaki, J., and Vuori, P. Keystroke recognition for virtual keyboards. In *Proceedings of the IEEE International Conference on Multimedia & Expo (ICME'02)* (Naples, Italy, July 3–5). IEEE, Piscataway, NJ, 2002, 429–432.
4. Richard, P., Birebent, G., Coiffet, P., Burdea, G., Gomez, D., and Langrana, N. Effect of frame rate and force feedback on virtual object manipulation. *Presence* 5, 1 (Jan. 1996), 95–108.
5. Turunen, J., Noponen, E., Vasara, A., Miller, J., and Taghizadeh, M. Electromagnetic theory of diffractive optics. In *Proceedings of the Workshop on Digital Holography*, F. Wyrowski, Ed. (Prague, Czechoslovakia, May 19–21). SPIE, Bellingham, WA, 1992, 90–99.
6. Van Trees, H. *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, New York, 1968.
7. Weiser, M. The computer for the 21st century. *Sci. Am.* 262, 9 (Sept. 1991), 66–75.

CARLO TOMASI (tomasi@cs.duke.edu) is an associate professor of computer science at Duke University, Durham, NC, and a software architect at Canesta, Inc., San Jose, CA.

ABBAS RAFII (ilhami@canesta.com) is a co-founder and executive vice president of Canesta, Inc., San Jose, CA.

ILHAMI TORUNOGLU (arafii@canesta.com) is a co-founder and executive vice president of Canesta, Inc., San Jose, CA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

© 2003 ACM 0002-0782/03/0700 \$5.00