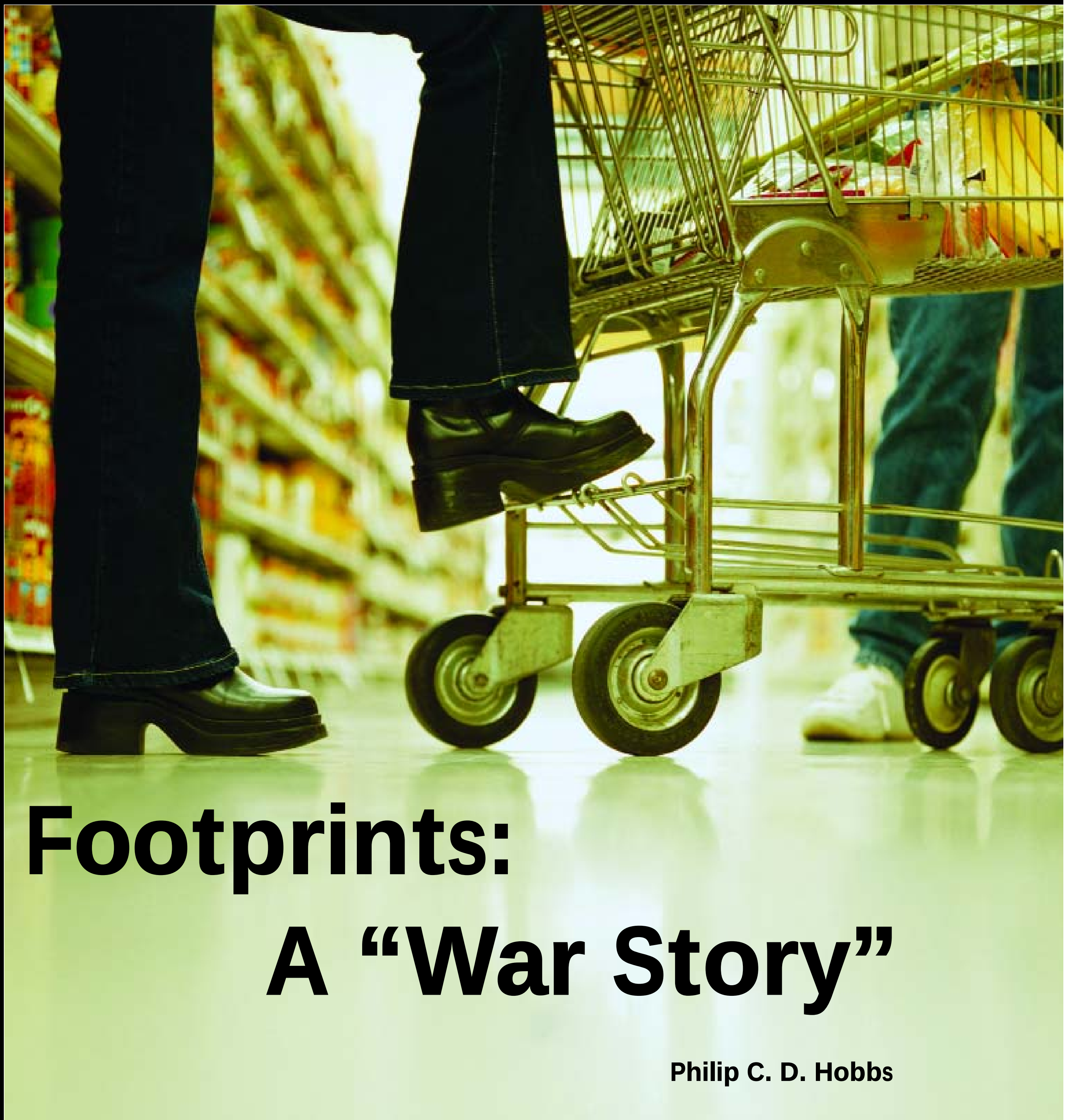
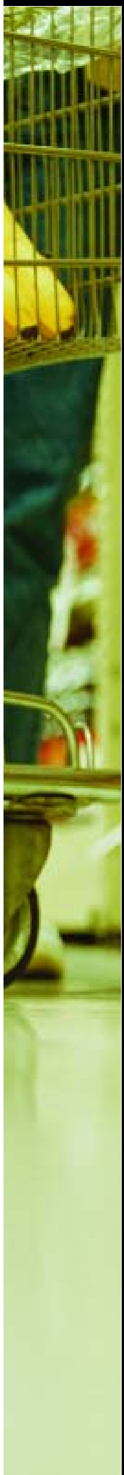


Footprints: A "War Story"

Philip C. D. Hobbs

1047-6938/03/09/0032/6-$0015.00



Footprints, a low cost distributed imaging system, was based on low resolution, uncooled thermal infrared sensors that cost only a few dollars each. It was designed to follow the movements of people in stores so that retailers could be more responsive to customers and more efficiently organized and stocked—without invading anyone's privacy. It was a great idea, but the project turned out a little differently than expected. The story goes like this...

One of the things that makes electro-optics work so absorbing is the number of things we have to get right. Lasers, optics, electronics, software, management, finance and even marketing have to come together for a project to be the roaring success we keep hoping for. Many projects fail, and even the successful ones always seem to go through at least a couple of serious crises. Some of the problems are pure bad luck: a major customer goes out of business, the technical guru's health gives out. Others are caused by stupidity: an overbearing manager drives the whole team to quit, the system designer fails to take the second law of thermodynamics into account, the team leader gets spooked by the "time-to-market" mavens into shipping a product that isn't ready.

Most failures lie somewhere in between. They lead to sad tales that begin, "It was such a great project—it would have worked *if only* we'd. ..." Tales of the way projects *almost* fail are more fun: "Man, if so-and-so hadn't figured out a way around that one, we'd have been dead." Just about everyone who's been in science or engineering for more than a couple of years knows a few such tales.

Besides enlivening the lunch table at OSA meetings, "war stories" have great educational value.

Besides enlivening the lunch table at OSA meetings, war stories have great educational value. They represent the practical lore of the field in its chemically pure form. When we hear a war story, we can all have a good laugh and be wiser afterwards. So why are published war stories so scarce? Ego is certainly one reason: it's one thing to tell our pals how we screwed up and another thing to put it in print. Charity is an even more important reason. Most of our failures may well be our own fault (mine certainly are), but it's still hard to write about them in an amusing and truthful way without embarrassing someone else. The difficulty is even greater when the story involves mostly the failures of others. Yet these expensive lessons can save so much time, enthusiasm and money that they're well worth telling.

Footprints

In the fall of 1998, my then second-line manager, for whom I have great regard, came into my office to tell me about a new idea. A recent reorganization had put me in a group serving the retail sector and I'd had to mothball my previous project, which involved inexpensive handheld three-dimensional (3D) scanning devices. Anyway, here was my boss looking for a way to instrument retail stores to make them more competitive with Web stores in customer service. He and another guy had come up with the idea of installing a network of low-resolution cameras to watch people shop. This wasn't voyeurism; the resolution was 30 cm, or one pixel per floor tile, so people look pretty much alike. The dot-com mania was raging and traditional retailers, with their single-digit margins, were scared that Web stores would take away their best customers.

The idea was that stores could learn a lot about pricing, fashion, and merchandise layout by capturing people's trajectories as they went from the entrance to the checkout, and merging each trajectory with data from the point-of-sale terminal (the cash register). That way, they could compare what people could have bought with what they actually bought. Say you're a merchant who sells matching sets of accessories, including handbags, belts and shoes. Some customers bought the handbags but not the shoes: Why? Were the shoes unfashionable, expensive or just invisible? A Web store knows the difference between things you saw but didn't buy and things you never saw, but a mall store hasn't got a clue, and that's the kind of unknown that can cost a retailer real money. There seemed to be lots of these opportunities, including some that were valuable enough to persuade a big retailer to instrument every store: at 100 square feet of coverage per sensor, 50 million square feet of floor space could move a lot of iron. To figure out shopper behavior, stores pay consultants to sit with a clipboard watching time-lapse footage from security cameras: surely the world's most boring videos. Customers stand, seemingly for hours, in checkout lines that could be shortened if we could warn the store manager a few minutes in advance to open more registers. We felt that we could end both kinds of suffering, so the Footprints project was launched.

As shown in Fig. 1, the chief design constraints are ceiling height and field angle. Since stores are full of tall, narrow objects such as shelves, signs and people, to see every square foot of the selling surface the sensors would have to work at no more than a 30 degree field angle. Most stores have ceilings 14 to 20 feet high. Since most human heads and shoulders are about 5 feet off the ground, whatever sensor technology you choose, you get 70 to 200 square feet of coverage per sensor. A typical 30,000 square foot store would need a few hundred sensors, so each sensor would have to be very cheap—say $100 installed, wired and ready to go.

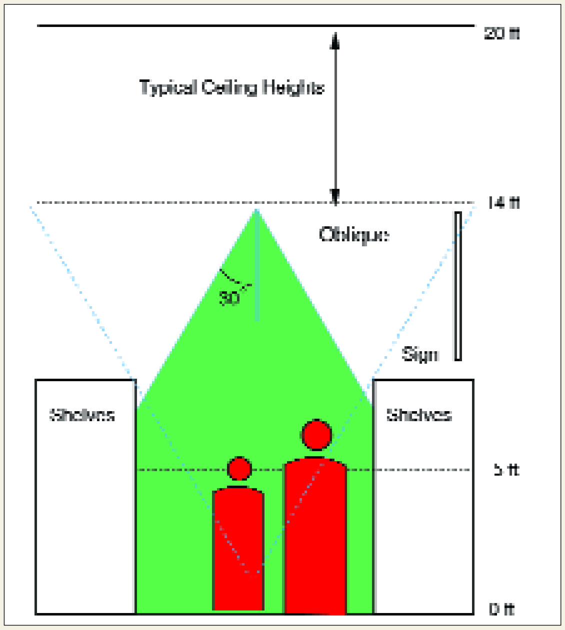


Figure 1. Typical store aisle, showing ceiling height constraint on field of view.

The appeal of thermal IR

My boss and I realized immediately that the way to go was to build low-resolution thermal infrared (IR) cameras; trying to do the job by means of large-scale machine vision would have been a nightmare. Machine vision works very well in a wide variety of situations, provided you control the lighting. Unfortunately, good machine vision lighting is more suited to an interrogation room than to Buffy's Boutique. One- or two-camera people-tracking in uncontrolled lighting has been demonstrated many times over the past 30 years, but it doesn't scale well, and even a machine vision person will tell you that the code is always a dense collection of specific hacks. In the hands of an expert it will eventually work, but for most people, it's riskier.

Thermal IR, on the other hand, fit the problem well. Not only are room-temperature objects self-luminous, floors in retail spaces are significantly cooler than human skin, so people stand out. Tracking incandescent objects against a dark background was the sort of problem I could cope with. A side benefit is that, in thermal IR, most inanimate moving objects, such as baskets and shopping carts, look just like floor and thus disappear—you could get confused by a roast chicken, but that's about it. The bad news: thermal infrared cameras are *expensive*. The cheapest ones, PZT pyroelectrics with about 256 pixels, cost $4,000 for the camera alone. That was clearly a non-starter. We weren't going to be using germanium lenses or active cooling, either.

The idea was that stores could learn a lot about pricing, fashion, and merchandise layout by capturing people's trajectories as they went from the entrance to the checkout, and merging each trajectory with data from the point-of-sale terminal.

The ins and outs of thermal detectors

You probably have a thermal IR motion detector controlling your front porch light: it has a segmented Fresnel lens that casts a dozen or so images of everything in its field of view and a split detector wired differentially. A moving object casts a row of moving images; as they cross the split detector, a differential AC signal results. A field-effect transistor (FET**)** amplifier detects the signal and triggers the output. The split detector is made from polyvinylidene fluoride (PVDF) pyroelectric film, which is basically fluorinated Saran Wrap. Porch-light sensors are very cheap but not terribly sensitive. If we made an imaging sensor out of Saran Wrap and a molded plastic Fresnel lens, it would be cheap, but could we make it sensitive? Answering this question requires a brief excursion into the netherworld of thermal detectors.

Thermal detectors sense changes in their own temperature from which we infer changes in the scene temperature. Since we want the temperature swing to be large and fast, we use a well-insulated detector of low thermal mass and increase the numerical aperture on the image side to obtain the maximum image irradiance. Blackbodies at 300 K are not very bright. If you put two wide blackbody surfaces 5 millimeters apart in air so that each point on the surfaces sees π steradians (sr) of the other, and heat one a bit hotter than the other, radiative coupling will give you a thermal conductance of about 6 W/(m^2•K), compared with about 5 W/(m^2•K) for still air. This isn't a big number, especially since even at f/1, the detected solid angle is only 0.64 sr, not π, and the surface emissivity of the sensor is only $\approx$ 0.1 (even lower with metal films) because of its thinness. In designing a sensitive instrument, especially when it has to be cheap, a carefully-made photon budget is a necessity (neglecting it is a very common way to make a project fail). In this case, the photon budget revealed a couple of very interesting things. The first was that although minimum thermal mass was important, we could improve the signal-to-noise ratio by insulating the pixels to make the thermal time constant very long (many frame times) and speeding up the response again by digital filtering after the readout. The second was that because of the voltage-divider effect of the thermal conductances, even with 5 millimeters of air as insulation and 9-μm film, we'd get about 0.005 K ΔT on the detector per degree of scene-temperature change. With only 100 or so pixels per sensor, the pixels didn't have to be small, which helped a lot. I finally chose 3 × 5 mm pixels on a 6-mm square pitch, arranged as 12 rows of eight columns on a free-standing film in air. We got about 0.1 picocoulomb of charge per frame time (200 ms) for a 1 K scene temperature change. The choice of an oblong array was deliberate: often turning a sensor sideways will allow coverage of an irregular area with fewer sensors.

The team is up and running

By mid-1999, when the project really got going, I had three teammates: a couple of talented veterans with doctoral degrees—Sharath Pankanti and Bob Wolfe—and a young guy I'll call Reg, reputed to be a firmware whiz. For the films, I found a manufacturer in Pennsylvania, Measurement Specialties. The company mostly makes piezoelectric strain gauges and such things, but an engineer there, Mitch Thompson, was interested in working with us. We started out with 28-μm film with 400 A of nickel over 100 A of copper, because that kept the thermal mass low and could be patterned with PC board chemicals and a mask made on a laser printer. I rapidly discovered that this solution wasn't as convenient as it looked. Any time the film stretched, the metal cracked. This happened often enough to keep me very busy with silver paint. What was worse, the cracks tended to close up again when the film recovered its shape, leading to many flaky failures. Mitch suggested a change to screen-printed carbon ink, which adhered well and was so durable it would survive a hard crease, and he was right. Getting the thermal mass low enough required thinner films, so after some teething troubles due to shorts and surface leakage on the films, we eventually wound up with screen-printed electrodes on 9-μm thick PVDF. By the end of the project, the films were working well. Figure 2 is a picture of an assembled sensor, minus its lens, box and high-tech convection shield (made from a 2-mil polyethylene bag and attached with tape).

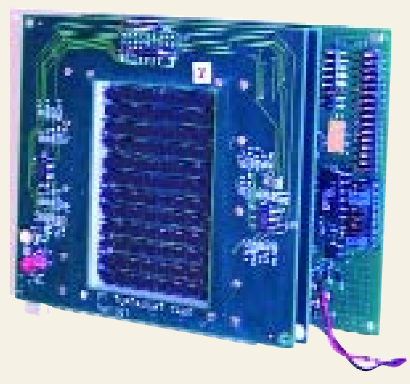


Figure 2. Assembled sensor, minus lens, box and 2-mil polyethylene convection shield. The black rectangles are the carbon-ink pixels on the free-standing film.

The code looked eerily familiar—I realized that what he was showing me was not his own work, but a reformatted version of code I had modified two days previously.

Getting the charge off the film and into the digitizer required a 96->6 multiplexer with femtoamp leakage. Integrated-circuit MOSFETs are easily good enough, but we didn't have the time or the budget for a custom multiplexer chip. Discrete MOSFETs are too expensive and don't work well: unprotected ones are too delicate, and the ones with gate-protection diodes are too leaky. I eventually used a multiplexer made out of ordinary display light-emitting diodes (LEDs), which cost only a few cents and have amazingly low leakage—I measured less than 100 fA over a bias range of -5 V to +0.5 V. The problem with diode switches is that they conduct in only one direction, whereas pyroelectric detectors produce bipolar currents. This problem was solved with a hack—by shining a little light on the switch LEDs from a group of four LEDs controlled by the processor, I could get an adjustable bias current of 0-10 pA per pixel. We thus wound up with a fundamentally minimal sensor configuration: a bit of plastic film printed with ink, plastic connectors and one LED per pixel. The sensor turned out to be pretty sensitive: about 0.13 K NEΔT.

The debugging process

We were fortunate to have a talented summer student on board in 1999 while the sensor was being brought up. Those who have had to bring up buggy hardware controlled by buggy firmware will realize that the process is much easier when the person writing the firmware isn't a beginner. Between debugging the sensor, debugging the code and spending a surprisingly long time learning what the normal behavior of the hardware was, we got a sensor running in August 1999. Meanwhile, Reg was working on a more robust version of the microcode and Bob and Sharath were writing the production communication, timekeeping, database and trajectory extraction code.

We were working with a big retailer, and we had a customer installation date of 10 p.m. Tuesday, November 23, 1999, that we couldn't slip, because the store wasn't going to let us mess around in the ceiling between the Friday after Thanksgiving and Christmas. Reg, the firmware guy, reported good progress. On Oct. 27, less than a month before the deadline, I asked Reg to give me a code walk-through the following day. Because I like to be able to put my arms around all the technology on a project, I had been learning how to program the microcontroller and

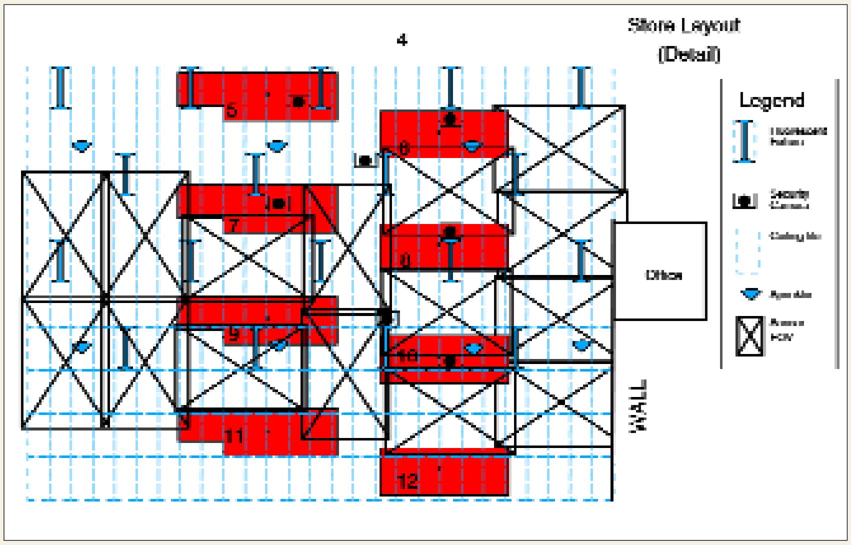


Figure 3. Layout of sensors in the store ceiling over the checkout area. Note how “busy” the ceiling is.

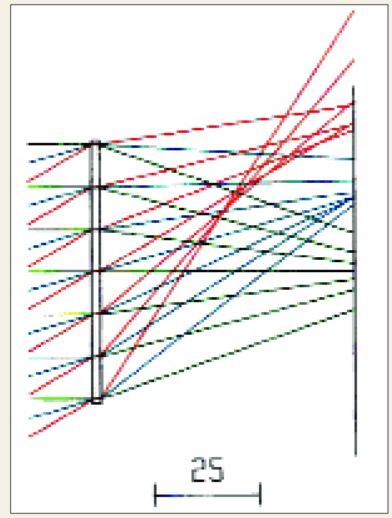


Figure 4. Performance of an optimized aspheric Fresnel lens vs. field angle. [Courtesy of R. Budd.]

had made successively modified versions of the summer student’s microcode.

When Reg and I sat down the next day, he started running me through his source files. The code looked eerily familiar—I realized that what he was showing me was not his own work, but a reformatted version of code I had modified two days previously. Contrary to his rosy progress reports, he had no microcode of his own to show for months of work, and we were three weeks away from customer installation. The summer student’s code had serious design problems that made it crash several times a day, so we couldn’t use it for the installation.

A change of venue

I had to write the complete microcode myself in three weeks. This would have been impossible if we hadn’t had the summer student’s code for debugging the sensor (even though none of that code wound up in the installed version). I didn’t sleep much for those three weeks, but two days before the installation, I had six sensors with microcode; the code was stripped to the minimum, but it didn’t crash. Bob had been doing a lot of the project management. He had lined up an installation crew, rented a ceiling lift and figured out what we were going to need to do the job. (Store ceilings are complicated, as shown in Fig. 3, which shows the sensor fields of view superimposed on the store layout as seen from above the ceiling.) Then, the day before the scheduled installation, the retailer cancelled. No reason was given, but because the parent company insisted that we move the installation to another store, we suspected that that particular store manager had objected to the idea. As frustrating as this was, it wasn’t altogether a bad thing since a couple of bugs surfaced in that version of the microcode. The compiler was also hopelessly buggy, and porting to another was painful because of the very different libraries, interrupt handling and startup code.I eventually had to bite the bullet, and changed compilers the following spring. Life got better.

During 2000, we were funded under an interdivisional joint program called First-of-a-Kind or FOAK. A FOAK is intended to fund early development of solutions that IBM can then sell. It runs for a year, no more and no less, and it requires an outside partner—sort of a customer who doesn’t pay anything. Determined behind-the-scenes work from my managers had lined up some powerful champions in IBM’s Retail Store Systems Division, who understood the vision and the high level of customer interest in our project. This was important because in a big organization like IBM, building something useful is not enough: you have to have a channel to the relevant project division.

The retailer’s objectives were somewhat different from ours. We wanted to build a trajectory-following system that would eventually cover the whole store, whereas the chain was primarily interested in measuring the length of checkout queues. This looked ok from our point of view because we didn’t possess enough sensors to outfit a significant fraction of the floor space anyway. Unfortunately, the store we got moved to had higher ceilings, 18 feet rather than the 13.5 feet we’d designed for. Since the relevant area was at shoulder height, our pixel size had just doubled—to two floor tiles per pixel—and we were looking at the area in which people were most crowded together. We suddenly had a resolution problem.

A polyethylene Fresnel lens was the natural choice for the optical system. Fresnel lenses may produce crummy images, but with only 96 pixels of 3×5 mm size, how

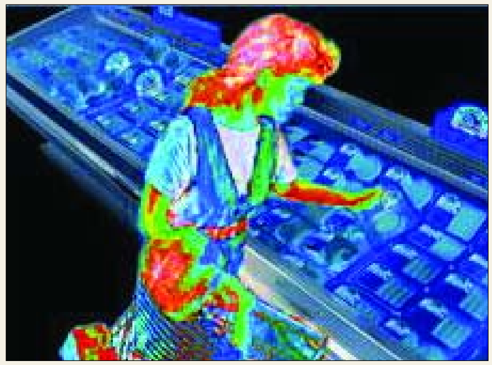


Figure 5. What shoppers don't look like in the IR. [From Beyond2000.com.]

Figure 6. One frame of Footprints data (six-sensor mosaic), showing four people walking around. A mild filter was used to reduce visual artifacts caused by rectangles but the individual pixels are clearly visible. [Courtesy of R. H. Wolfe.]

good a lens do you need? Actually, you need something a bit better than that—when my colleague Russell Budd simulated the lens we were using, it turned out that our 60-mm $f/1$ Fresnel lens running at a 30 degree field angle had a focused spot size of almost 1 cm, if you call that a focus. (Figure 4 shows the results calculated for minimum worst-case spot size). He also produced a much better design by moving the aperture stop forward and using a much larger Fresnel lens, making the optical system nearly telecentric on the image side. By then it was too late to fix it; we would have needed a diamond-turned mold with a three-month lead time and a completely redesigned mechanical package.

We arranged an installation date in the new store in mid-May 2000. This was in the middle of my battle with the metal films and silver paint, but we still managed to scrape together six sensors that worked (at 200 square feet per sensor, we didn't need as many). We were all ready, with the crew and the ceiling lift lined up, when the retailer cancelled the day before, *again*. At this point, it was becoming harder to explain the long hours and deadline pressure to our families.

Media attention

We had also started to get media attention from *The New York Times*, National Public Radio and the Discovery Channel in Canada, plus a number of online reports. Some of the coverage was silly and sensational. One online article had a sort of "Big-Blue-Is-Watching-You" slant, with a false-color image of a woman shopping for frozen food, purporting to be what Footprints would see. Of course, the photo was taken from the side, and resolution was 5 mm instead of 300 mm, but how alarmed are readers going to get by a 3-pixel monochrome image of someone's head taken from the top? A dead giveaway of the nature of this picture, shown in Fig. 5, was that the freezer was the same color as part of the woman's face, indicating origins in Photoshop rather than in thermal IR. We had a demo system running to show customers who visited IBM Research, and we were invited to stage a demo at a conference at Euro Disney put on by our chairman for the chief information officers of big customers, but we had to turn the opportunity down because we just weren't ready.

One benefit of the delays was that when we finally installed sensors in the last store in late June 2000, we were able to use solid firmware with good time-keeping and data handling, relatively advanced signal processing and the first carbon-ink sensor films. We installed four at first; about six weeks later, we replaced them with six others with the final version of the films, which were highly reliable and sensitive. Those sensors took data until January without a single failure and without the need to reboot. We finished the FOAK at the end of 2000, having done a good but not stellar job at counting people in checkout queues.

Of course it didn't help us much. The queue-counting problem with the too-big pixels took resources away from the trajectory-extraction job, which never really got done. Inventing on a schedule, Reg, the retailer's prevarication and my own inexperienced project management forced us to work in emergency mode almost the entire time. Our main champion in the product division changed jobs, so the path to market went away. To cap it all off, we had two management changes in Research, so our local champions went away too, and being the only hardware effort in an all-software department made us an anomaly. Footprints was officially cancelled on March 20, 2001. RIP.

Philip C. D. Hobbs (pcdh@us.ibm.com) is with IBM T. J. Watson Research Center, Yorktown Heights, N.Y.

References

1. Interested readers can find the full gory details in a SPIE paper, "A $10 thermal imager," Proc. SPIE 4563, or at www.pergamos.net.