1.2 Surface and Tangible Computing, and the "Small" Matter of People and Design

Bill Buxton

Principal Researcher, Microsoft Research
Toronto, Ontario / Redmond Washington

1. Introduction
Hardly a day goes by that we don't see an announcement for some new product or technology that is going to make our lives easier, solve some or all of our problems, or simply make the world a better place!

However, the reality is that few of these products survive, much less deliver on their typically over-hyped promise. By virtue of their embedded microprocessors, wireless capabilities, identity tagging, and networking, these products are going to be even more difficult to get right than those that we have already produced (too often unsuccessfully) in the past.

Future products will be interactive to an unprecedented degree. Furthermore, the breadth of their form and usage will be orders of magnitude wider than what we have seen with PCs, VCRs, and microwave ovens, for example. Some will be worn or carried. Others will be embedded in the buildings that constitute our homes, schools, businesses, and cars. In ways that we are only starting to even imagine, much less understand, they will reshape who does what, where, when, why, how, with whom, for how much, and for how long.

On the other hand, as suggested by this last sentence, the extended behaviours of these products will be matched, and exceeded, by the expanded range of human behaviour and experience that they enable, encourage, and provoke - both positive and negative.

Think about the introduction of texting (more properly called "Short Messaging Service", or "SMS") into cell phones. The traditional approach would view SMS as the design of a protocol to enable text messages to be sent between cell phones, and then its implementation in hardware and software, (along with the associated model for billing for the service). Yet that description does not even *begin* to accurately characterize the real nature of SMS. This is far more accurately reflected by activities such as voting for your favourite performer in *American Idol,* or flirting with someone across the floor in a dance club. *That* is SMS, and I don't believe that you will find anyone involved in its design who would claim that they anticipated, understood, or much less considered, any of that when they were designing the feature.

Increasingly, the technologies that we design are not isolated islands – that is, they are not free-standing or complete in their own right (to the extent that they ever were, but more on that later). Rather, they are *social entities*. As with people, they have different properties and capacities when viewed as a collective, within a social, and physical context, than they have when they are viewed in isolation, independent of location or context. For example, just as I behave differently when I am alone than when with others (among other things, I talk with them, but hopefully not to myself), so will it be with our devices. When they approach other devices, or possibly people, they will become social animals. Just like you and me, their behaviour will vary, depending on whom they are with – in the same way you and I behave differently with family than we do with strangers, business colleagues, or alone.

Software will no longer resemble the all-too familiar applications that run on today's PCs, but will require a very different style, behaviour, and development process.

As we move away from standard platforms, such as the PC and the graphical user interface (GUI), the software, hardware, and business aspects of the product will not be able to exist in three different silos within an organization, much less be undertaken isolated from the larger social, cultural, and economic context within which they will function. While this may seem banal or obvious, the implications of this need for interactivity on the subtle details of the underlying circuits and components are far less so. Hence the motivation for this paper.

Regardless of which part of the pipeline we are discussing, it is the quality of the users' experience with the product that is the real object of the exercise – not the software, building, product, sign, or environment. Needless to say, any attempt to achieve this that does not intimately involve these same human users *in an appropriate way* is almost doomed to failure.

2. Back to the Classroom
Let's look back a bit in history. The example that I am going to chose relates to one of the red-hot technologies of today – big displays.

Figure 1.2.1 is from one of my favourite books, Volume II of the *Historical Atlas of Canada* (Moldofsky, Gentilcore, Measner, Matthews & Walder, 1993) [10]. It documents the introduction of the blackboard into schools in Upper Canada in the period between 1856 and 1866.

You will be excused if you are asking yourself, "Why do I care about blackboards in the 1800s, in Canada no less?" The key to the example (regardless of year or country) lies in the following question: "What preceded the blackboard?"

With just a little reflection, one will come to the obvious answer: "The slate."

But, you might ask, "Is not a slate just a little blackboard? If so, this cannot be very important."

From a technological perspective, it is hard to argue to the contrary. Slates *are* just small blackboards. They are made from the same materials ($CaCo_3$ rather than Si), and in the jargon of today, employ the same "user interface", the same "text editor", the same "operating system", and even the same "erase" operator. Moreover, you can use the same documentation for each, and if you can use one, you can use the other.

Consequently, one might quite legitimately ask, "What is the big deal?". After all, "all" that they did was make a bigger slate and put it on the wall.

From an engineering perspective, this is all true. Yet, I would strongly suggest that an argument can be made that the introduction of the blackboard has had more impact on classroom education than any innovation in technology since, *including* the introduction of cheap paper, the Internet, personal computers, and perhaps all three put together!

That may seem like a bold claim! But for my point to work, you only have to concede that this argument is *plausible*. It need not be true. The gist of the example is to illustrate that a very significant impact resulted from a change in *scale*, *location,* and usage, rather than a change in technology *per se*. The change was social and organizational, not technological in the common sense of the term.

Leaping ahead, this example from the mid-1800s sheds some light on the deployment of technology in today's schools as well. In light of current work in display technologies, such as OLEDs, LEPs, e-Ink, electrowetting, etc., technologically what we are mostly doing is making a new quarry to mine a replacement for limestone. The point is what we are *not* doing, namely: making a concerted effort to distill the implications that these technologies might have on classroom education, and what implications these might hold for

 978-1-4244-2010-0/08/$25.00 ©2008 IEEE

the electrical engineer who is making the technical decisions around the circuit design of such displays.

So, how might we do things differently?

3. Other Lessons from Childhood

The blackboard and slate are not the only experiences from childhood that can help inform us in our quest. So, let me talk about two things that will help us on our path.

4. Finger-Painting

My involvement in hardware design began in the second half of the 1970s. I was involved in a group at the University of Toronto making a digital instrument for real-time musical performance. In addition to the standard organ-type keyboard, we made a number of touch-sensitive surfaces (Buxton, Hill & Rowley, 1985) [1]. These gave us a far more organic style of control than switches and buttons. One configuration of one of these tablets is shown in Fig. 1.2.2.

Around that same time, we had our first child. Very early on I realized that he already knew what I had only recently thought about. When he got into food, or any other material that he could "paint" with, he used all of his fingers and both of his hands to do so. Now, contrast what he did with your own experience with the touch pads that you have used to control a computer. You were allowed to made only one point at a time, and use only one finger at a time. My one year-old child was smarter than all of us. He wasn't educated enough to be stupid. He knew that he had hands and fingers, – instruments seemingly designed for the sole purpose of smearing everything everywhere that he wanted, whenever he could.

The good news was that at about the same time, we started to discover the same thing in our academic research (as opposed to playpen – although the difference is smaller than one might think!). Carrying on from an earlier thesis in our department (Mehta , 1982) [8], we built a tablet that was sensitive to simultaneous touches at multiple locations, and with the ability to sense the degree of each touch independently (Lee, Buxton & Smith, 1984) [7]. We stopped the work in late 1984 when I saw a much better implementation at Bell Labs – one that was transparent and mounted over a CRT. The problem was that they never released the technology, so, the whole multi-touch venture went dormant for 20 years. But, I never stopped dreaming about it. (Lesson: don't stop your research just because someone else is way ahead of you. It might be transitory, and anyhow, remember the story of the tortoise and the hare.)

5. Speaking and Hearing

Now, let's switch from finger-painting to your very first experience at making a voice-carrying telecommunications device. Unless you are considerably smarter than me, I suspect that your first experience was to explore something like the device illustrated in Fig 1.2.3.

What it shows is the kind of walkie-talkie that we used to make with string and two tin cans. These worked well, as long as we kept the string taut, and did not let it touch anything. And, if we kept at it, we would later learn that we could do the same thing with two radio speakers whose terminals were connected by two wires – with no battery or amplifier needed. Induction did the magic, and we were freed from the constraints of keeping the wire taut or away from other objects.

All this was fine. But somewhere, like with finger-painting, as we got more educated, we somehow knew less. As engineers we lost the notion that transducers can be bidirectional. Even though we knew that any speaker could also be a microphone, the evidence suggests that until very recently, we forgot to ask the seemingly obvious question: if that is true for speakers, why isn't it also true for displays?

When I asked that question around 1990 at Xerox PARC, I got the answer that I wanted to hear: they can.

You just have to anticipate the capability in how you design the circuits. Yet, it has taken fifteen years until display makers have started to understand that the exciting thing about the technology was not that you could present digital images, but that you could produce *interactive* digital images.

6. Coming Full Circle

If we return to where we started, what all of this suggests is that while we have developed the capacity to make ever-larger displays, from the perspective of blackboards, we made the board, but forgot the chalk, the erasers, and the fact that it is what we write on them ourselves, as well as what we read from them, that lets them realize their full potential.

So, what if we take that into account? What would that mean?

Well, from a technical perspective, for example, it would mean that instead of considering displays as made up of 3 emissive pixels: R, G, & B, there should be a fourth one in the cluster, "I" for Imaging. And, if we implement things effectively, we can potentially have the makings of something that a one-year-old can appreciate.

As it turns out, there is interesting work in this direction, and my point in diving into all of this is to tease out some of what I think is important to learn from it, and how it might apply in other domains.

7. Surface Computing

Over the past couple of years, a new class of interactive device has begun to emerge, what can best be described as "surface computing". Two examples are illustrated in Fig. 1.2.4.

These typically incorporate a rear-projection display coupled with an optical system to capture touch points by detecting shadows from below. Different approaches to doing the detection have been used, but most employ some form of IR illumination coupled with IR cameras. With today's camera and signal-processing capability, reliable responsive and accurate multi-touch capabilities can be achieved.

Because they are new to most, the tendency in seeing these systems is to assume that they are all more-or-less alike. Well, in a way that is true. But on the other hand, that is perhaps no more so than to say that all ICs are more-or-less alike, since they are black plastic things with feet like centipedes which contain a bunch of transistors and other stuff. In short, the more that you know, the more you can differentiate. But even looking at the two systems in the photo, there is evidence of really significant difference.

You will be forgiven if you think that it is what is displayed on the screen. That is just your misspent youth as part of the television culture that is speaking. No, the really significant difference is that one is vertical and the other is horizontal. (Remember our description of the blackboard, where "all" it was a big slate mounted on the wall? Well, it is a bit like that.) Why is this significant? Well, this is one of those questions perhaps best answered by a child in kindergarten. They will tell you that if you put a glass of water on the vertical one, it will fall to the floor, leading to a bout of sitting in the corner. On the other hand, it is perfectly safe to put things on a table. They will stay there.

So why should we care? The key reason is that there is no reason that the only thing that can be sensed by such surfaces is one's fingers. Hence, it turns out that things like the Microsoft Surface Computer can sense more - much more - as is hinted at in Fig. 1.2.5.

Not to put too fine a spin on things, but did it ever occur to you when your meal was getting cold while you were waiting for the server to notice that your wine glass needed filling, that this was a problem that you could address through circuit design, and practically and realistically do so within the next year or so? I thought not! So what else are we missing?

To do this thread justice, let's step back for a bit of background.

8. Phycons and Graspable /Tangible Interfaces

Since the early to mid-1980s, we have become habituated to the use of graphical icons as a staple component in interacting with our computers. For example, some of the icons on my current desktop (as I write) are illustrated in Fig. 1.2.6.

The figure shows visual icons. The question that arises is: "why just for the the eyes?". What about the notion of physical or auditory icons? Well, if we think about it, we use both every day. Think about the different bells, buzzers, and assorted acoustic signals that we know the meaning of, or the shapes of knobs on our devices that we interact with. Why not incorporate these more effectively into our designs?

Well, some people have done so. For our purposes, I want to focus on physical icons, or what have alternatively come to be known as "phycons". The first time that I saw such things really developed as a concept was in Japan in 1994 while visiting a company called Wacom. There, one of the founders, Murakami-san, showed me a prototype paint program that had no menus on the screen (Fukuzaki, 1993) [4]. The entire screen real-estate was devoted to the electronic canvas, and the "icons" were held in the hands.

There was a phycon for each key function:

- A stylus: to act as your paintbrush
- An eraser: to correct mistakes
- A paint-pot: to invoke your colour picker
- A filing cabinet: to save or retrieve your work

The devices used to work with this system are illustrated in Fig. 1.2.7. What was wonderful about this system was that you could keep your eyes on your painting, and select the tool that you wanted to use eyes-free. For me, at least, this system was a revelation. It certainly stimulated research in our lab, such as Fitzmaurice, Ishii, and Buxton, (1995) [2], and Fitzmaurice (1996) [3]. This work opened up a new direction in human-computer interaction, one that has become variably known as "graspable" or "tangible" interaction.

9. Graspable Computing and the Society of Appliances.

We are going to build on four notions:

- physical icons and tangible interaction are broadly extensible concepts
- Surface as a bidirectional transducer that can sense what device is placed on it
- that the tangible devices that can be sensed on the Surface can be far more complex than credit cards, drinking glasses, erasers, pens, ink pots, etc.
- that the interactions between the Surface and the rec ognized device can go well beyond simple device and location recognition, or simple button pushing.

We see how this relates to Fig. 1.2.8, when we consider what happens when the physical device that I place on the Surface device is a Smartphone or two MP3 players, as is shown in.

There are a number of important things to notice in this photo. First, the albums in the unit under the user's hand are visible on the Surface, under the Zune MP3 player. Second, the track list of the current album appears up to the right. Other albums from the library or elsewhere appear scattered on the surface. Any of these can be played by touching them. But they can also be loaded in either Zune on the Surface, just by dragging, as can the albums under the user's hand be dragged from the Zune on the left to the grey one on the right.

So, what is going on here, and why is it important?

First, the nature of the interactions is extremely rich, but simple. Natural, perhaps. To get a sense of the degree to which this is true, imagine doing all of this with a conventional PC and two MP3 players. How long would it take you to teach your mother how to do it that way? How long would it take the way shown in the photo?

Second, so, what makes this so effective? To begin with, the transition in the behaviour of the Zune player, from stand-alone appliance to integration with the Surface (not to mention the other Zune) is seamless and transparent. That is, it happens automatically, is visible, and has no surprises. Things conform to expectations, even though one has never done this before.

But, it is deeper than that. The Surface and the Zunes are not really siblings. What is going on here is more of a figure-ground relationship. The Surface provides visual context that makes explicit what is going on between the two Zunes, fore example. As we go to richer transactions, one escapes the boundaries of the small displays on the Zunes, and things become visible and manipulatable.

Yet, when you pick the device up, the transition to departure is smooth and continuous. And the devices revert to their standalone behaviours.

Finally, we have only hinted at the potential richness of the potential types of interaction. We have not even touched upon integrating richer audio into the system. For example, non-speech audio cues could create a sound ecology that helps one understand what is going on. Likewise, we have not talked about speech to enhance communication with either remote people or with machines. The potential is as interesting as it is challenging.

10. Wrap-Up

So where does all of this leave us? Yes, the ideas are compelling. Yes, many of them not only work, but after up to twenty years of refinement, are being commercialized. This is all good news, and a very good example of how visions of usage can help drive technology. But what we have seen is only the tip of the proverbial iceberg in terms of realizing the real potential of technology, broadly reviewed.

Yes, we can make large interactive surfaces using rear-projection coupled with optical sensing. But that is not where we either want or need to be. For example, as well as others, we are working on thin displays, comparable to conventional laptop LCD panels, that have the bidirectional properties of our old tin-can walkie-talkie (Hodges, Izadi, Butler, Rrustemi & Buxton, 2007) [6], as is shown in Figure 9. But, even if we have inexpensive large thin multi-touch surfaces, the challenges are still substantial:

First, new sensors and capabilities in mobile devices, that can serve as tangible interfaces for interaction through Surface, is still largely open. Our preliminary work with things like MP3 players, digital cameras, and mobile phones, just scratches the surface (so to speak) and stimulate as many questions as are answered.

Wilson and Sarin (2007) [11], for example, have demonstrated the detection of mobile phones, and the use of Bluetooth to support interaction and cross-functionality between the phone and Surface, and two different phones via Surface. However, none of

 978-1-4244-2010-0/08/$25.00

the existing protocols work as one would like and this brings us to the key point – this is so because they were not designed with the needs and constraints of the types of interactions that we are working on, in mind. While brilliant, they do not do what we need.

I spoke earlier about the paradox in the speed of technology development – it goes at rocket speed, but that of a glacier as well; Simultaneously! In the perfect world, this would be ideal: we could go through several iterations of ideas so that by the time the new paradigms of interaction, such as Surface and Tangible computing are ready for prime time, everything will be in place.

But, the rapid iteration is more directed at supporting the old paradigms faster and cheaper, rather then helping shape the new ones. The reasons are not hard to understand. From the perspective of circuit design, the problems are really hard. So, one has to have one's head down working flat out to get anything done. But, there is a side of me that motivated this paper that asks, If it is so hard, then isn't it worth making sure that the things one is working on are things that are worthy of one's hard-earned skills?

I don't presume to suggest that I know better than anyone else what is worth working on. But, what I can say, and hope that I have demonstrated, is that the "brand-spanking-new revolutionary latest cool things" that are garnering attention today are ideas that have been knowable and in the works for twenty years. That is, long enough for us, and a few generations of graduate students, to ponder their implications from a circuit design perspective.

I guess that my best concluding remark would be this: There is a community that is as good at thinking about such things, as the circuit-design community is, at what it does. I wonder where we might be now if we shared our visions more clearly, much sooner, in the grand scheme of things.

While the past is water under the bridge, I can still speculate on what could happen if we do so in the future. My hope is that this paper provides a catalyst for more than a few of you to have similar thoughts. And then, collectively, we act on them.

References:
[1] W. Buxton, R. Hill, and P. Rowley, "Issues and Techniques in Touch-Sensitive Tablet Input", *Computer Graphics*, 19(3), *Proceedings of SIGGRAPH'85*, pp. 215-223, 1985.
[2] G.W. Fitzmaurice, H. Ishii, and W. Buxton, "Bricks: Laying the Foundations for Graspable User Interfaces" *Proceedings of CHI'95,* 442-449, 1995.
[3] G.W. Fitzmaurice, "User Interfaces", PhD Thesis, Department of Computer Science, 1996.
[4] Y. Fukuzaki, "Electronic Pen According to the BTRON Guideline and its Background", *Tronware*, *Vol. 4*, *Personal Media Publishers, Japan*, pp. 49-62 (in Japanese), 1993.
[5] J.Y. Han, "Low-Cost Multi-Touch Sensing through Frustrated Total Internal Reflection", *Proceedings of the 18th Annual ACM Symposium on User interface Software and Technology UIST '05*, pp. 115-118, 2005.
[6] S. Hodges, S. Izadi, A. Butler, A. Rrustemi and B. Buxton, "ThinSight: Versatile Multi-touch Sensing for Thin Form-Ractor Displays*", Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology UIST '07,* pp. 259 - 268, 2007.
[7] S. Lee, W. Buxton, and K.C. Smith, "A Multi-Touch Three Dimensional Touch-Sensitive Tablet", *Proceedings of the 1985 Conference on Human Factors in Computer Systems, CHI '85*, pp. 21-26, 1985.
[8] N. Mehta, "*A Flexible Machine Interface",* MASc. Thesis, Department of Electrical Engineering, University of Toronto, 1982.
[9] Microsoft, http://www.surface.com, 2007
[10] B.R. Moldofsky, L. Gentilcore, D. Measner, G.J. Matthews, and R.H. Walder, *"Historical Atlas of Canada II: The Land Transformed, 1800-1891"*, University of Toronto Press, 1993.
[11] A.D. Wilson and R. Sarin, "BlueTable: Connecting Wireless Mobile Devices on Interactive Surfaces Using Vision-Based Handshaking", *Graphics Interface*, pp. 119-125, 2007.

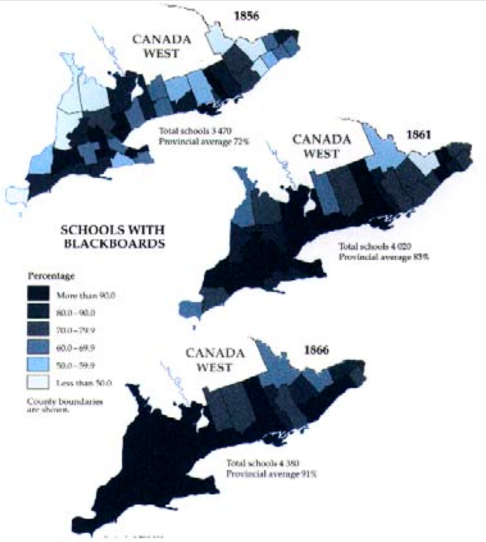

Figure 1.2.1: The introduction of blackboards into schools in Upper Canada between 1856 and 1866. From the Historical Atlas of Canada, Volume II, Plate 55 [10].

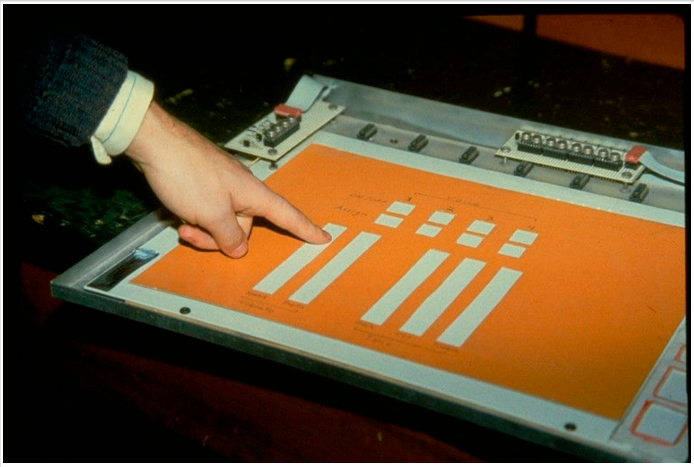

Figure 1.2.2: A Touch-Sensitive Tablet. The overlay defines a number of virtual sliders and buttons that can be used to control a digital sound synthesizer. The edges of the overlay gave tactile feedback. To enable manipulating multiple sliders at a time, we built a touch tablet capable of sensing multiple points of contact simultaneously, with pressure independent for each contact point.

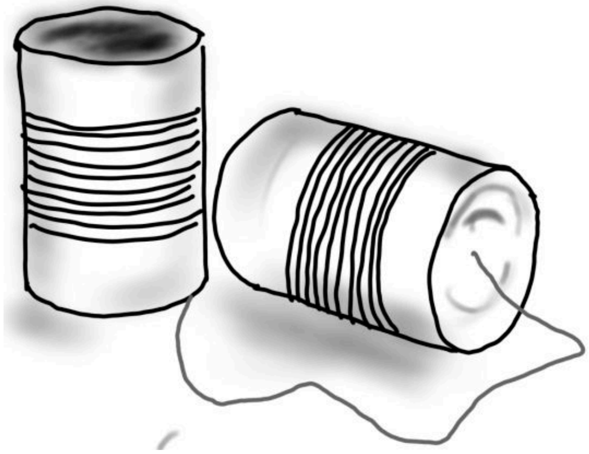

Figure 1.2.3: A Tin-Can Walkie-Talkie. As long as you kept the string taut, and it didn't touch anything, you had your very own home-made voice-communication device. The important underlying message of significance that most of us missed was the existence proof of symmetrical bi-directional transducers. That is, every loudspeaker can also be a microphone.

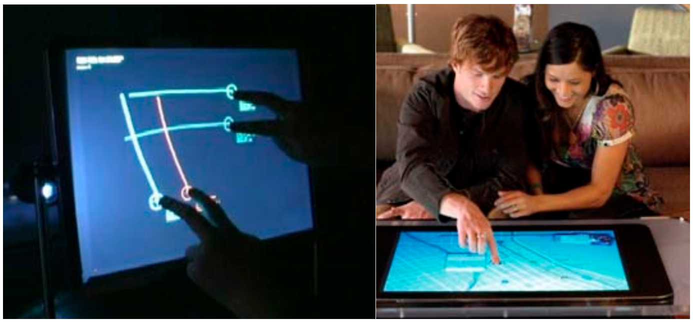

Figure 1.2.4: Surface Computing: The left image shows multi-touch interaction on a vertical rear-projection display. This is finger painting worthy of any kindergarten student. The work is by Jefferson Han of NYU (Han, 2005) [5]. The right image is a coffee-table-like horizontal Surface from Microsoft (2007) [9]. Likewise, it uses a rear-projection multi-touch sensing display.

Figure 1.2.5: Recognizing a Credit Card and a Drink on the Surface. Objects can be identified by techniques such as shape recognition, or tagging using approaches such as optical bar codes or RFID tags. Based on context, the surface can then enable certain relevant transactions. Microsoft (2007) [9].

Figure 1.2.6: Graphical Icons on my Desktop Using Microsoft's Vista Operating System.

 978-1-4244-2010-0/08/$25.00 ©2008 IEEE

Figure 1.2.7: The Wacom Paint Phycons (Fukuzaki, 1993) [4]. The physical icons can be seen on the graphical tablet: the filing cabinet, ink-pot, eraser, and stylus. The system worked using electro-magnetic resonance. Each phycon had a coil in it with a unique resonant frequency, determined by its unique coil length. Hence, each device could be identified, as well as location determined, by the tablet circuitry. No batteries or wires to the devices were required. (Photo compliments of Wacom).

Figure 1.2.8: Two Microsoft Zune MP3 Players on the Surface Computer. What is significant here is that we see that Surface is more than just a big interactive screen. It is a means to make explicit and visible the functions, context, and inter-relations and transactions amongst the devices on it, as well as its own resources.

Figure 1.2.9: ThinSight: An Integrated Optical Multi-Touch- Sensing Thin-Form-Factor Display (Hodges, et al., 2007 [6]). The fingers are illuminated by IR from behind the LCD, and the fingers are detected by the IR light reflected back through the display. The black background rectangles show the finger shadows.