Using Metaphors to Create a Natural User Interface for Microsoft Surface

Kay Hofmeester

Microsoft Surface

One Microsoft Way

Redmond, WA 98052 USA

kayh@microsoft.com

Dennis Wixon

Microsoft Surface

One Microsoft Way

Redmond, WA 98052 USA

denniswi@microsoft.com

Abstract

Creating a new model of human computer interaction is not straightforward. Only a handful of such models have been commercially successful. Those that have, such as the graphical user interface (GUI), can provide valuable lessons. When we were challenged to develop a new natural user interface design for Microsoft Surface, we drew from these lessons and from modern user research techniques. A prominent starting point resulting from this was using metaphors to develop the design. We used metaphors for two reasons: To create a user interface world that is understandable and predictable for our users, and to guide the team in creating the detailed design. We continued this practice in the user research: We focused on which metaphors worked best in the studies, learned if users understood them, and which metaphor they preferred. This case study describes the process we followed and the lessons we learned from this.

Keywords

Microsoft Surface, natural user interface, touch, metaphor, RITE

ACM Classification Keywords

H5.2 Information interfaces and presentation: User Interfaces

General Terms

Design, Human Factors

Introduction

Microsoft Surface is a hardware/software product based on a touch interface. An important characteristic that distinguishes Surface from other touch interfaces is its vision system that recognizes hand postures, gestures and objects, and is capable of handling many contacts at a time. The current software environment however, while being successful, is limited in its extensibility. This is why, when the design team was challenged to develop the next generation of Surface, we focused on innovating the user interface. Our goal was to create a new natural user interface environment that exploited the capabilities of the Surface hardware.

We started by formulating clear and simple design principles to direct our thinking and to evaluate our design throughout the process. But the design principles were not sufficient to drive the design effort forward. We needed something that would define our designs at a conceptual level, and something that would help us to creatively come up with innovative solutions. We also needed something that would integrate each of the designs and give them an internal coherence.

A user interface creates an environment for humans and computers to communicate. We imagined this environment as a "world" with its own characteristics, behaviors, and rules that a user has to learn to be able to operate in it. We needed a way to make this world understandable and predictable. We decided to use metaphors to define this user interface world. This usage was similar to the desktop metaphor guiding the

creation of the personal computer user interfaces we use today. We realized using a metaphor in the creation of a new user interface is powerful, but can be controversial [1,2,3,4,5].

These metaphors would help create a user interface world that was understandable and predictable for our users, and would guide the design team in creating an innovative and detailed user interface design. They would help us avoid replicating familiar design patterns like the GUI. We used the concept of metaphor in all phases of the design process. The design principles were always the deciding factor when selecting metaphors.

We used a design process of diverging by generating many ideas, then converging by detailing and selecting the best ones, and iterating that cycle by repeating the process with less variations and more detail [6,8].

An additional challenge was that we started with a very small core team, as the rest of the Surface organization was still working at the current version of Surface. As more people became available, we had to get them up to speed without disrupting our process. To make this easier, we were transparent about our process, communicated frequently to the rest of the organization, and involved the other Surface team members where appropriate. One of the tools we used for this was the design book. This was a digital scrapbook that contained the history as well as the latest of our thinking. We used this as a communication tool, a way to get new members up to speed, and as a tool for presentations.

We started with generating a large number of metaphors. We used the iterative diverging/converging design process to detail these into nine design concepts and narrow them down to the three most promising design proposals. We user tested these using the RITE method. Combining iterative design and user research we developed one direction representing one metaphor for the new user interface world we were developing.

Defining design principles

Early in the process we defined common design principles. The goal of these principles was to guide us in every phase of the process. They helped us in generating ideas and concepts, and they currently are helping us making decisions during the detailed design and production phase. They were also used to formulate our research program. The three most important principles we used were: Multi-user computing, designed for touch, and content is the interface.

Multi-user computing
Surface celebrates the social aspect of human interactions. It stimulates getting together, creating and sharing together.

In our definition, multi-user computing is about multiple people in the same place, at the same time. Our metaphor and our user interface would have to stimulate multi-user computing. From a research perspective multi-user computing meant that we would test people in naturally occurring pairs – for example family members.

Designed for touch
Surface is designed for touch from the ground up. All Surface interactions are based on touch, postures and gestures.

The goal was to make the experience simple and approachable, natural and intuitive, fast and effective. We did not want to emulate a mouse and pointer interface, but use touch as a fundamental interaction paradigm, one which was designed to work by itself and was augmented by the use of other input devices when appropriate. Thus touch was a means to an end. Our goals of simple, natural, intuitive, fast and effective were measured in the eyes of the beholder. In other words, our success in achieving these goals was measured by how users reacted to the interface.

Content is the interface
Surface focuses on content: It allows for directly interacting with the content without being hindered by controls or applications.

This was a direct departure from the GUI interface where visible controls occupy valuable screen real estate. While these controls reduce memory load (a goal of the original GUI) by making functions visible, they can overwhelm the content, which is the primary focus for the users.

Finding a metaphor

A user interface creates an environment for humans and computers to communicate. We imagined this environment as a "world" with its own characteristics, behaviors, and rules that a user has to learn to be able to operate in it. We decided to start with focusing on the metaphor that would define this user interface

world we were going to create, similar to the desktop metaphor guiding the creation of the PC user interfaces we use today.

Napkin sketches

We took the classic story of someone coming up with a million dollar idea while sitting in a bar and sketching it out on a napkin to remember it, and used it as a way of explaining the level and kind of ideas we were looking for to the larger Surface team. We organized multiple brainstorm sessions to generate a wide range of metaphors. During the sessions, we asked participants to sketch out their ideas on napkins. We used the napkin sketch as a reference to innovation and the start of something new and great. It also forced the participants to describe their idea very concisely. The generated ideas ranged from real world to abstract and from practical to esoteric. We generated about 100 metaphor ideas, examples included:

- Magnet board, a metaphor for a bulletin board, re-arrangable space, and communication hub, but also a reference to the use of physical objects for functionality or communication,
- Circle, an environment based on human physical capabilities and personal space defined by reach,
- My personal moon, a world defined as a solar system, with planets as workspaces and information centers,
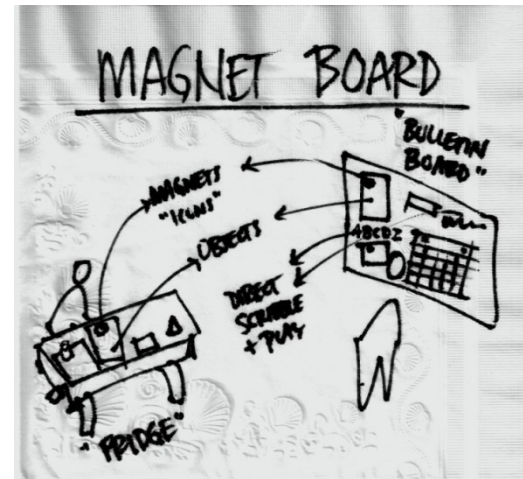- By the fireside, a metaphor of gathering around a campfire for storytelling and singing songs.
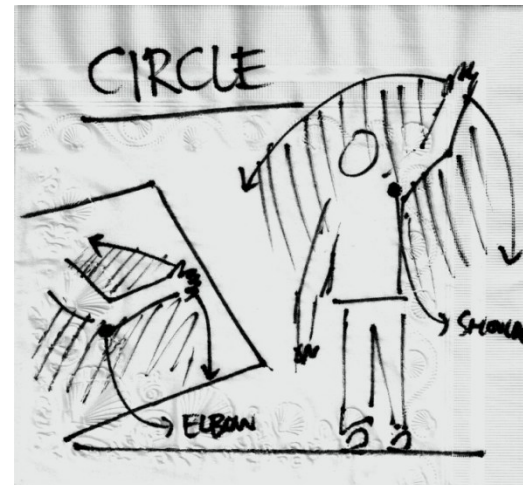

Fig. 1 Magnet board metaphor
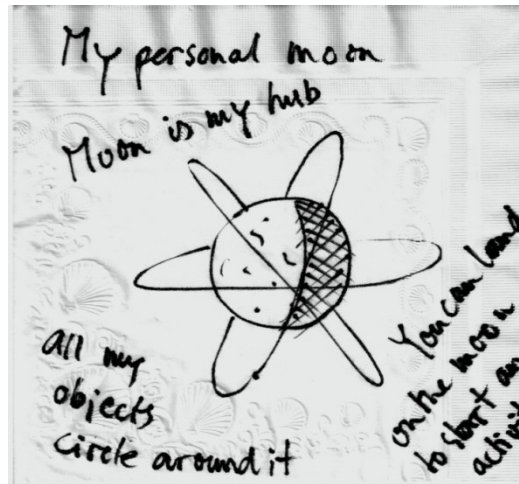

Fig. 2 Circle, personal space metaphor

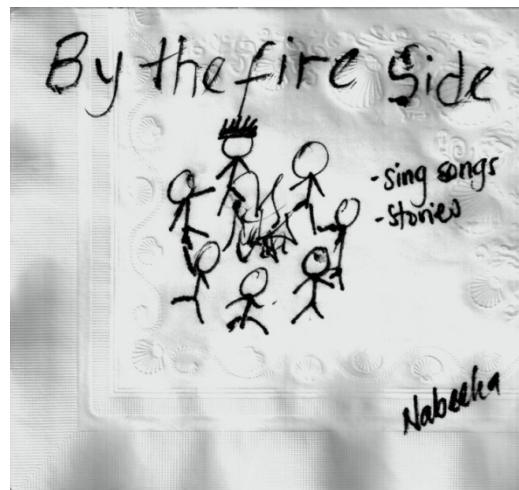**Fig. 3** My personal moon, solar system metaphor

Fig. 4 By the fireside, campfire metaphor

We grouped the ideas using the design principles, aiming for a wide range of different concepts. We made sure to have a few practical, realistic ideas, some far-out ideas, and some in-between. We used these grouping sessions as an opportunity to discuss ideas and deepen our understanding of the metaphors, and to include more members of the Surface organization in the process. At the end of this process we had selected nine groups of ideas. The next step was to organize these into nine comparable concepts.

Concepts
We started with naming the piles of napkins so we could address and discuss our concepts. Our concepts were now:

- Canvas, based on the blank canvas, drawing, and creating things together
- Garden, a metaphor of seeding, growing, tending, and community gardens
- Magazine, a book metaphor of beautiful typographic design, page-based content, and bookshelves
- Magnet board, the communication hub bulletin board
- Memory chest, a magic place of memories and discoveries
- Moon, the solar system metaphor
- Sphere, based around the idea of personal space
- Unfold, a paper and packaging metaphor of unfolding content
- Water, focused on the surface of the water, what is above and below, and the concept of sedimentation.

To make these into comparable concepts we used a format consisting of a text description of the environment, and the top three user benefits, formulated as "we-statements". We used these instead of "I-statements" to help us focus on multi-user computing. We created a collage of concept images and visual design images. Finally we added interaction sequences in the form of paper sketches and video sketches for each concept. The interaction sequence showed how to start interacting in the environment, perform an individual activity, another person joining, perform a together activity, and then end the session.

We created large poster boards for each concept. First we used these boards for a one-day brainstorm session with the user experience design and user research teams. In this session we discussed the concepts and then worked on more detailed interactions for each concept, to get a better idea of how well the concepts supported our design principles. Then we put them up in the hall and frequently presented and discussed them with members of the larger Surface team. During these discussions the concepts that fit best with the principles and had the highest potential for an innovative user interface world started to emerge.

Finally we set up a meeting with the core team and key decision makers, and decided on three concepts to detail further. Our decision was based on the three principles: Multi-user computing, designed for touch, and content is the interface. Another important factor was the range of concepts; we wanted the concepts to be distinctly different, to be able to explore different possibilities for the design. If concepts were too similar, we combined them into one. The concepts we selected were:

- Magnet (a combination of Magnet board and Garden)
- Unfold (a combination of Unfold and Magazine)
- Sphere (a combination of Moon and Sphere)

Up to this time, our design process was relatively free-from. We followed the iterative diverging/converging cycle, and we continuously referred back to the principles. A lot of sketching in various forms, including paper, animation, clay, physical modeling, and video, was used to explore possibilities. Discussion was important to deepen our understanding, and to guide the decision making. Discussion also helped us to involve other members of the Surface organization. For the next phase however, we needed to get more specific and detailed in our designs.
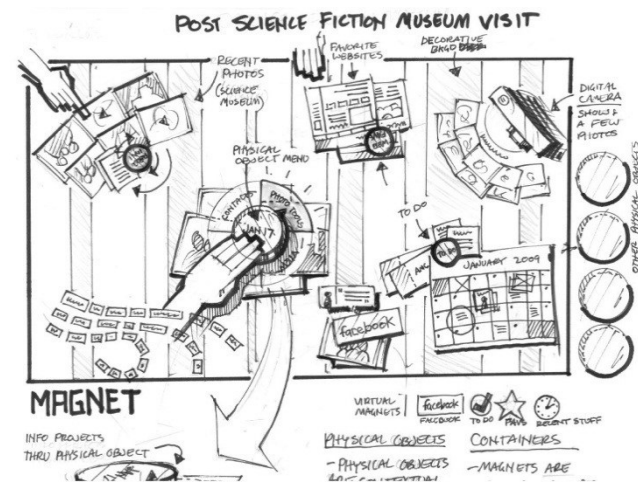

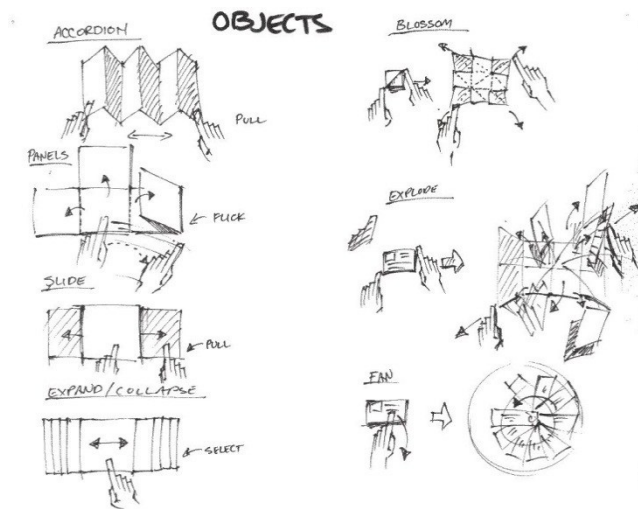
Fig. 5 Magnet interaction sketch
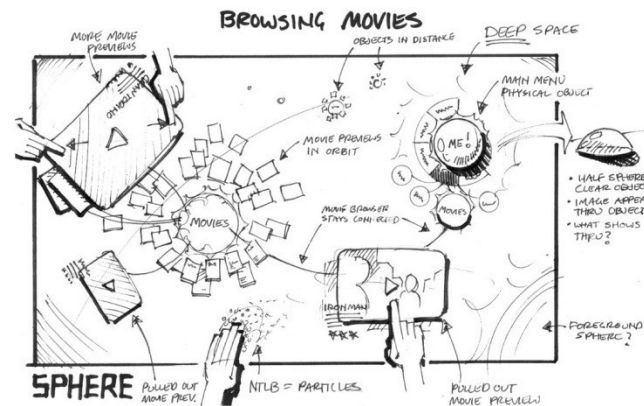
Fig. 6 Unfold interaction sketch

Fig. 7 Sphere interaction sketch

Creating prototypes

Getting our designs more specific served two purposes. Firstly, metaphors and napkin sketches would not be sufficient to build a real product. In order to do that the metaphor had to be instantiated in some form. This process both enriched and extended the metaphor and exposed inconsistencies in our concepts. Making the metaphor real gave substance to the design, clarifying exactly what needed to be built.

The second purpose was to design the instantiation so the metaphor would be clearly communicated to the user. Users can provide the most effective and detailed feedback when the presented designs are specific. In our case of a new user interface paradigm the design needed to be fully instantiated in the form of a working prototype. The working prototype could then be tested to determine if it met the design goals. We could not know if users saw the interface as natural and intuitive unless they could experience it and we could ask them. We could not know if the interface was efficient unless we measured user performance on tasks. This placed high demands on our team. Paper prototypes and partial implementations would not do; we set out to create interactive prototypes for each of the three user interface worlds. We had four weeks to go from concepts to the touch prototypes that would run on our Microsoft Surface units.

The first step was to create detailed interaction primitives, visual design and animation design for each of the three worlds. Interaction primitives were the building blocks of our system. They consist of a visual affordance, a posture or gesture, and a resulting action. An example of a primitive is clicking a button, or resizing a photo. We created a detailed visual design

for each of the design proposals that clearly showed the metaphor of the concept, and provided the right visual affordances. And we designed animations to make the user interface operate in a smooth way, and to provide the users feedback to their actions.

We created the prototypes in WPF (Windows Presentation Foundation), which is a combination of XAML and C#. This is how the user interfaces for most Surface applications are created. At this point we were ready to test our design proposals with real users.
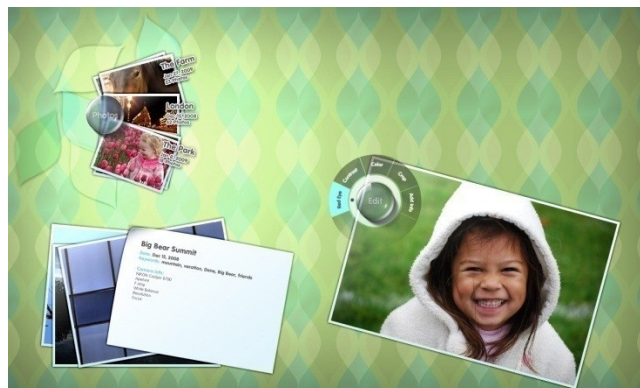

Fig. 9 Unfold prototype screenshot

Fig. 8 Magnet prototype screenshot

Fig. 10 Sphere prototype screenshot

User testing the prototypes

Rationale

Testing a new interface metaphor creates an additional challenge. We were evaluating three aspects of the design:

- Do the specific mechanics of the interface work for users?
- Do the mechanics of the interface communicate the metaphor?
- How do the users respond to that metaphor?

To address these questions in a short time frame we chose the Rapid Iterative Testing and Evaluation method – or RITE method [7]. Using this method we could rapidly fix problems with the interface during the test. This was a necessary choice because any problems we uncovered would need to be fixed so that users could perceive and evaluate the metaphor.

Test set-up

The goals of the tests were:

- Fix problems with the user interface that inhibit the study,
- Assess each of the metaphorical worlds.

Our goal of multi-user computing required that we test users in pairs. We invited 12 pairs of users to help us evaluate and iterate on our designs. The participants worked through three scenarios using the designs for the three worlds. To achieve the main goal of the study, which was to help the team refine the worlds and select one direction to pursue, we focused on the following questions:

- What words would they use to describe the system?
- Did they recognize the metaphor?
- Which world was most preferred?
- How did users respond to interaction primitives and visuals?
- After seeing all the metaphors which one did they prefer
- After seeing all the metaphors did users rank each world across the core metrics?
- How did this system compare to other systems the participants had used (i.e. GUIs)

These questions formed the basis for an open ended interview. The interviewer followed up on specifics. We also looked at usage of the metaphor.

Test experience

In accordance with the usual approach to RITE testing a cross-functional team of decision makers watched the test. This team included designers, user researchers, software developers, and program managers. A user researcher ran the tests and conducted post-test interviews. The order of presenting the different metaphors was randomized. After users had experienced all three metaphors they were asked to choose between them and describe the rationale for their choice. We were more interested in the stated reasons for the choice than a raw tally of the frequency with which each one was chosen.

Test results

Overall, Sphere was the most successful of the metaphors. While overall preference ranking for Sphere and Magnet were equal, Sphere was most accurately

described by users. It was also most preferred by those who accurately described it. The results for all three metaphors are reported in the three tables below.

Words used to describe Sphere metaphor	Recog nized	Prefer ence
NA	No	U/M
Twister, Entertainment, Circles, Bulls Eye, Game.	Yes	U
Fluid, orbits, asteroids, suns, orbs planets universe, Copernicus, Splay outs.	Yes	S
Solar system, planets, universe, raindrops, water puddles, ripples	Yes	M
Universes, planets, molecules, atoms. Disco.	Yes	U
Control panel, home central, donuts, buttons, CDs, DVDs, like stuff is stored on disks.	Yes	S
Space, bubbles, drops	Yes	S
Eyes are bigger than my plate. To mean that there is too much on it.	No	M
Circles, bubbles, floating balloons.	Yes	S
Challenging, cluttered.	No	M
Marbles, air hockey, like hockey pucks you can throw around, tear drops or rain drops, ripple effects or records.	Yes	S
Users who recognize Sphere metaphor	**8/11 (72%)**	
"Recognizers" who prefer Sphere		**5/8 (62%)**

Table 1 Words used, recognition, and preference for Sphere. (S = Sphere, M = Magnet, U = Unfold)

In contrast only two of 11 participants recognized Magnet.

Words used to describe Magnet	Recog nized	Prefer ence
NA	No	U/M
Tactile, hands on	No	U
70's wallpaper, old school look.	No	S
Like a kid's computer game (Mama's kitchen soup). Reminds of Barbie with the bright colors.	No	M
Sci fi, Minority Report, Like Cyler on Heroes--his power and control.	No	U
Piecemeal, something that means many pieces. Components.	No	S
Chunks, pieces, game pieces, gallery	No	S
Jetsons would have this on fridge and would be like "go get milk". Mike thinks of Microsoft for new technology.	Yes	M
Like magnets on a refrigerator door. Workshop because the tools seem like a work bench.	Yes	S
Like Tom Cruise; like what he was doing in that movie. A collage but action packed collage. T.V.	No	M
Swimming, fluid, space, your space, flowing.	No	S
Users who recognize Magnet metaphor	**2/11 (18%)**	
Recognizers who prefer Magnet		**1/2 (50%)**

Table 2 Words used, recognition, and preference for Magnet. (S = Sphere, M = Magnet, U = Unfold)

Words used to describe Unfold	Recog nized	Prefer ence
Stacking	No	U/M
Visual, deck of cards,	No	U
Constructive, architect's table, deck or hand of cards, spread of cards	No	S
Scrapbook, flowers	Yes	M
Interaction like iPhone touch but you can blow things up (increase size). Like a super advanced Photoshop combining applications on a big screen.	No	U
Files, computer.	No	S
Messy, cards, flash cards, index cards, clippings, collage, journal, portal.	Yes	S
Organized, like thinking 2 steps ahead of you. Structured like a course or class. Architecture, working with blueprint, or architect's paper.	No	M
Obstacle course, dry erase. Talked about the mail flying off being like Harry Potter, Legos.	No	S
Puzzle, brain teaser, explorer, compilation, roller coaster, fun and exhilarating but also goes down.	No	M
Rolodex, credit card bills, this is keeping with the idea it is dry and regimented, so structured, does not look fun.	No	S
Users who recognize Unfold Metaphor	**2/11 (18%)**	
Recognizers who prefer Unfold		**0/2 (0%)**

Table 3 Words used, recognition, and preference for Unfold. (S = Sphere, M = Magnet, U = Unfold)

In summary users recognize *and* prefer the Sphere metaphor. In this case, recognition means users understood the metaphor in the way the designers intended it. In a sense all users had some understanding of each of the metaphors. They all could describe the metaphor in "analogical" terms, but in the case of Magnet and Unfold their understanding differed from the design intent.

After users had experienced all the metaphor, they were asked to rank them in terms of overall preference. Here we see an interesting result: Sphere and Magnet received equal ranking.
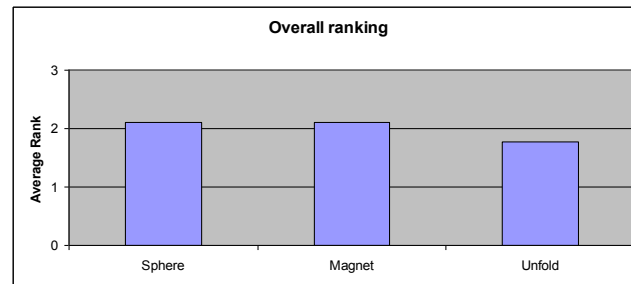

Fig 11 Rankings for each of the metaphors.

The positive ranking for Magnet is surprising given how few people "understood" the metaphor. In fact while users did not understand the metaphor as a whole they liked elements of its interaction. In particular they liked the immediacy of the object recognition and system response.

This positive response to Magnet is also reflected in the ranking given to all three metaphors on the scales of

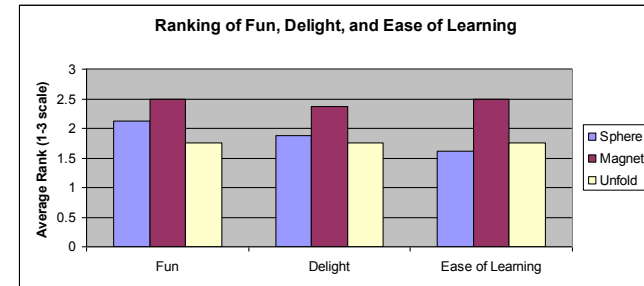fun, delightful, and easy to learn. They are shown in the figure below.

Fig 12 Ranking of Fun, Delight, and Ease of Learning

We asked participants to discuss their rankings and then tabulated the reasons given.

Reason for preference	Participants stating this reason
Objects create a menu for you at the top level like a context menu without right click	5, 7, 8, 10
Act of putting objects on the table and picking them up	5, 7, 9
Reading photos off the phone by putting it on Table	6, 8, 10, 11
Objects make sense, could tell from their shape how to use them	4, 8, 9
Objects are like toys	4, 10
Tactile, hands-on manipulation makes it easier to learn	7, 8
Colors	4, 9
Using 1 hand/two fingers to manipulate things	8, 10
Expanding photos does not degrade images	5, 8

Table 4 Reasons for preferring Magnet.

Table 4 shows that the preference for Magnet is driven by several primitive actions. In other words they liked the way the interface behaved even if they did not see the metaphor the designers intended.

Some of the overall comments by users were enlightening. In comparison to a GUI metaphor users said they felt "part of the system". For Sphere users felt that they understood the navigation model and they used the model successfully, but they could not describe it easily in words.

The final results were surprising. Many Surface team members had expected the Unfold design proposal to get the best evaluation because of its cool design, and large potential for interactions. But during the course of the study, it became clear that the users preferred either Magnet, or Sphere. They preferred Magnet for its use of tools to manipulate objects and the environment. They preferred Sphere for its clear hierarchy and edgy visual design. After evaluating the results at the end of the study we decided to combine these two concepts.

Final direction
We named this final direction Sphere, because of the strong visual appearance of the circular elements in the user interface. We also described the direction as a toolbox metaphor which combined the physical interactions from Magnet and the navigational aspects of Sphere.

This toolbox metaphor provided the excitement of choosing a tool and seeing a result. It guided us in a direction alternative to using menu structures for completing tasks. The tools could be both physical tools, using Surface's vision system to recognize the objects. Or they could be invoked using your hands as tools.

The initial presentation of the toolbox system was perceived as organic, holistic and integrated. As the user navigated to a specific task in the system, it offered individual functions appropriate to the context.

The visual direction of the system set the expectation that it was light, simple, and responsive and the activities within the system matched these expectations.
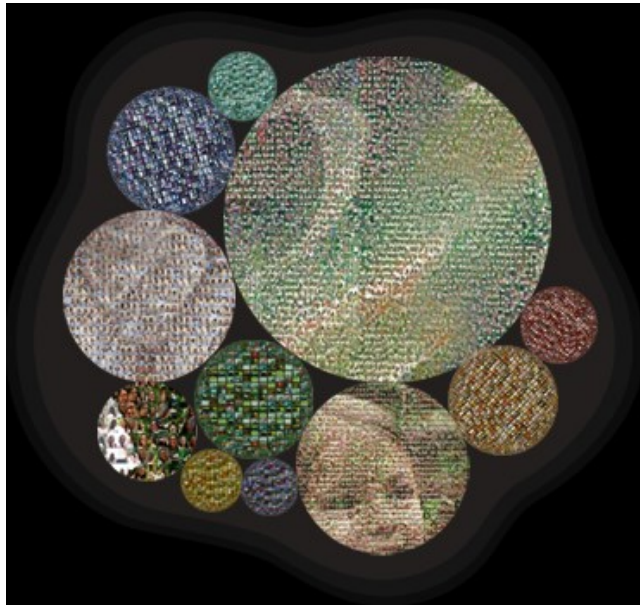

Fig. 13 Final direction sketch

Next steps

With this final direction we entered the next phase of the design process: To create a detailed interaction and visual design for the Surface user interface. This user interface was based on the guiding principles of the metaphor, and we regularly tested with our users if it was understandable and predictable.

Some of the details changed in this process, but the navigational principles and the toolbox metaphor were maintained. At several points in this process we reviewed our designs based on these principles, and corrected design flaws. Our understanding of the metaphor enabled us to better follow the chosen design direction.

Conclusion

Defining a new user interface environment is not an easy task. Starting with design principles and using the route of the metaphor proved successful in our case. The design principles helped us to stay on course during the multiple iterations, and the diverging and converging phases. The metaphor provided the beacon to focus on, and guided the creation of detailed design and primitives.

User research provided valuable insights in how multi-user computing was experienced by users, which metaphors worked well, and which did not.

Lessons learned
Our process of iterating on diverging and converging design cycles proved to work. It helped us to both be creative, and move towards a single design. It was also very useful in explaining to non-designers in the organization what our process was, and how we planned to arrive at our goal.

The prototypes were very effective both as a design tool, and as a user testing tool. Designing for a prototype forced the team to consider all aspects of the design and it created a deep understanding of the design. It made the designs specific enough for users to evaluate.

User testing of the prototypes was unique. We found that users understood and preferred one metaphor: Sphere. We also found that users liked many of the interaction primitives of Magnet. For the design team the test produced surprising results. This showcased how important it is to balance expert opinions with user testing.

RITE testing worked well in this case. Because of the short timeframes, we needed to arrive at a working prototype very quickly. The rapid iteration in the RITE testing sessions proved useful in arriving at a prototype that generated user feedback.

Because of the very short given timeframes it proved very important to work together in a team and divide up tasks. Building trust in the team through communication was essential for this. The process could sometimes get messy, but was always very creative. As long as the goal was clear, and there was trust, this worked well.

Documenting the results during the process in a design book worked well as a way of collecting all the information, providing an overview of where we were at any point in time, and for informal presentations. To finalize the design book in a formal deliverable proved to be very time intensive though.

Metaphors have been discussed extensively in the HCI community. Surprisingly tests of alternative metaphors with users have been relatively rare. We found that user testing was very effective in both choosing a metaphor and understanding where to compromise the metaphor with elements derived from other metaphors.

Logically users need to understand a metaphor in order to realize many of its benefits. Practically users need to appreciate a metaphor in order to adopt a new system.

References
1. Barr, P., Khaled, R., Noble, J. and Biddle, R. (2005) A Taxonomic Ananlysis of User-Interface Metaphors in the Microsoft Office Project Gallery. Australasian User Interface Conference.

2. Blackwell, A. (2006) The Reification of Metaphor as a Design Tool. CHI 2006, ACM Press.

3. Halasz, F. and Moran, T. (1981) Analogy Considered Harmful. CHI 1981, ACM Press.

4. Khoury, G. and Simoff, S. (2004). Elastic Metaphors: Expanding the Philosophy of Interface Design. Computing and Philosophy Conference.

5. Lakoff, G. and Johnson, M. (1980) Metaphors we live by. Chicago, University of Chicago Press.

6. Laseau, P. (1980). Graphic Thinking for Architects and Designers. New York: Van Nostrand Reinhold Company.

7. Medlock, M., Wixon, D., McGee, M. and Welsh, D. (2005) The Rapid Iterative Test and Evaluation Method: Better Products in Less Time. Cost Justifying Usability, An Update for the Internet Age. Bias, R. and Mayhew, D. (eds). Boston, Morgan Kaufmann.

8. Pugh, S. (1990). Total Design: Integrated Methods for Successful Product Engineering. Reading MA: Addison-Wesley.