CHI 2010 Gest: Exploring Gestural Interactions

Ankur Sardana
Sr. Interaction Designer,
Design Innovation Team,
Honeywell, India
Ankur.Sardana@Honeywell.com

Abhijit Kr. Bairagi
Industrial Designer,
Design Innovation Team,
Honeywell, India
Abhijit.Bairagi@Honeywell.com

Abstract
In this paper we describe the use of gesture based 'device interlinking' to achieve an enhanced user experience and optimize hardware utilization.

Keywords
Gesture, interaction paradigms, connect, hardware utilization, design, exploration, human factors

ACM Classification Keywords
H. Information Systems
H.5 INFORMATION INTERFACES AND PRESENTATION (e.g., HCI)
H.5.2 User Interfaces - Input devices and strategies, Theory and methods

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

General Terms
Design, Experimentation, Human factors

Introduction
Imagine returning home from a hard days work, plopping down on a favorite bean bag and tuning into ESPN (a sports channel) simply by pointing at a football lying around and then to the television.
Imagine calling up the car service station by pointing your mobile phone to your car.
Imagine pointing an mp3 player to a poster of Sting (a popular musician) on the wall to play his songs.
We imagined… and called it Gest.

A simple gesture connecting two devices opens up myriad possibilities of exciting and engrossing ways in which to interact with them.

We as humans have the natural inclination to do things physically. We relate to the physical manipulation of our environment more intimately than virtual control. Also, the natural cause and effect interactions afford the minimum amount of cognitive load. A physical

gesture is thus an apt idiom to explore for future human machine interactions.

Initiation

Current technology trend is to put all interfacial modalities into every product. Increasingly most of the devices around have visual displays, communication abilities and data storage capabilities. This is a wasteful trend. With the vision of achieving some sort of hardware usage optimization, came the thought of 'connecting' devices depending on required function. The idea was to study, analyze and discover how a user would spatially connect two devices and their expectations from the connection.

From early 1970's researchers have been working new technologies and applications based on gestural inputs. One of the earliest of them being Myron Krueger [1], who used real time image processing of live video of the user (figure 1),as an input.

Figure 1. Here the index finger is recognized and when extended, becomes a drawing tool. Shaping the hand in a fist, so that the finger is no longer extended lets the hand be moved without inking [1]

Subsequent research has yielded rich dividends and we have gestural interfaces now ready for being productized.

Figure 2. (Copyright: Canesta, Kicker Studio) Using hand gestures to control television [2]

If we could chart the progress of User Interface (UI), we could say that gesture based interactions lie somewhere in the middle of the standard click based UI and the futuristic thought based interfaces, where the user would just need to think to get the work done.

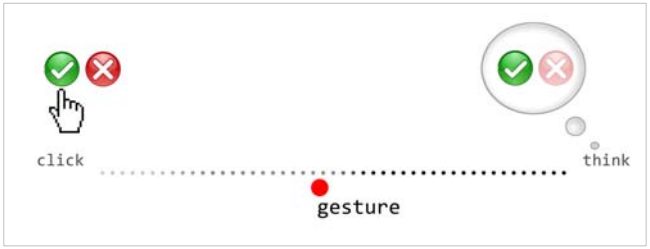


figure 3. Gesture based interfaces lying between the standard click based UI and 'thought based' interfaces

Research techniques and research

Since the answers we were trying to seek were qualitative in nature, we kept our research limited to two qualitative tests to understand how people associate gestures with functions spatially.

We conducted our research with fifteen users from a cross section of employees from our own company, the majority of them being involved in software development and lying in the age group 25-35.

We spatially simulated a home environment. This room had locations assigned for the hall, kitchen, toilet, bedroom and attic. We arranged common home appliances (using paper print outs where we could not accommodate real equipment) for the user's reference. The idea was that the user would randomly select any two accessories and then discuss

a) How would she choose to connect the two appliances?

b) What he would expect if he were to connect the two appliances?

The selection of the two devices was to be by picking a chit each from three separate boxes which contained names of three different groups of devices. The user was required to select from any two of the three.

The three groups of devices:

Group 1: *Passive objects*: Objects which allow only one way interaction with the user. For example: newspapers, books, keys, wallet, furniture etc.

Group 2: *Moderately interactive appliances*: Appliances that allow for very basic two way interaction. For example: washing machine, food processor, vacuum cleaner, microwave oven, air conditioner etc.

Group 3: *Rich interactive appliances*: Appliances which allow multimedia interaction with communication abilities and data storage. For example: personal computers, laptops, mobile phones, home theatre systems, televisions etc.

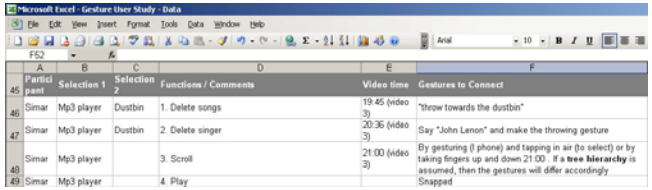
The users were asked to then think of various possible options he could have if he connected these two appliances and the various gestures he would make for the same. We recorded all the discussions with the users (with their consent) on video for later analysis.

Figure 4a. A participant interacting with the vacuum cleaner

Figure 4b. We experimented with bringing in two participants at the same time, to observe the variations in same gestures but made by different people

Analysis of Research

Functional patterns and gestures identified during the user research were further confirmed by the analysis of the collected video footage.



Microsoft Excel - Gesture User Study - Data

	A	B	C	D	E	F
45	Particip ant	Selection 1	Selection 2	Functions / Comments	Video time	Gestures to Connect
46	Simar	Mp3 player	Dustbin	1. Delete songs	19:45 (video 3)	"throw towards the dustbin"
47	Simar	Mp3 player	Dustbin	2. Delete singer	20:36 (video 3)	Say "John Lenon" and make the throwing gesture
48	Simar	Mp3 player		3. Scroll	21:00 (video 3)	By gesturing (I phone) and tapping in air (to select) or by taking fingers up and down 21:00 . If a **tree hierarchy** is assumed, then the gestures will differ accordingly
49	Simar	Mp3 player		4. Play		Snapped

Figure 5: An excel sheet grab which shows how the data was captured

It was fascinating to consolidate the gathered data, to find patterns like how people meant the same things when they wanted to connect objects and sometimes had very particular expectations out of the connections.

Though the scope of this paper is not to suggest the most relevant connections, below are listed some of the results from the study.

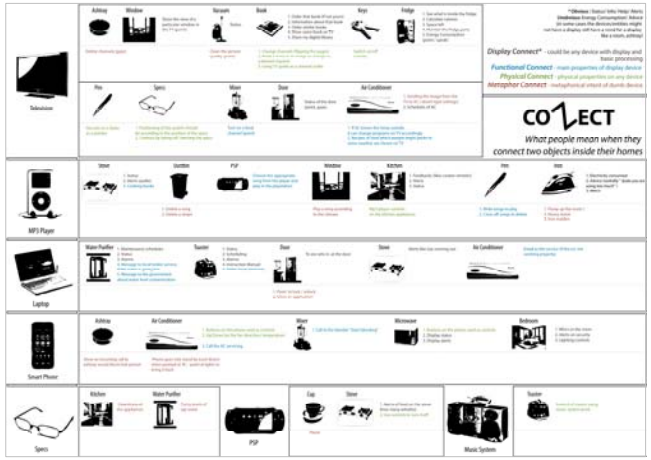


Figure 6: What people want to happen, when they connect two objects with gestures.

Similarly, there were many different ways by which users wanted to connect the two devices/objects. Some of the prominent ones being using index finger pointing for both objects; hand pointing; using a mixing gesture; using some random object (like a pen) kept around to point at both objects and using a cross like gesture with fingers.

Gest – the theory

While analyzing the responses, we could see five distinct patterns for which people wanted to connect different objects.

Display Connect – Where the user wants to use the display and processing power of one device for another device which lacks it. Example: Pointing a water purifier to a television in order to check the status of the purifier or some analyzed data of water consumption in the house on the television's screen.

Functional Connect – Where the user wants to use a property of one device to service another one. Example: Pointing gas stove to a mobile phone, to call the gas service agency

Physical Connect – Where the user wants to use the physical properties of one device. Example: a mobile phone or a pen as a pointing device

Metaphor Connect – Where the user wants to use one object as an iconic metaphor of some functionality. Example: pointing to a light bulb and a television to trigger the 'brightness' control for the television

The main gestures which were used to connect the two devices, varied according to the parameters like **functional complexity, cost, material, product design, size and spatial placement**. For example, a device which was 'respected' (due processing capacity, cost, size) was not pointed with an index finger, but with an open hand, the same rule was broken though if the device was insignificant.

Concepts and Simulation

We came up with visual scenarios based on the most prominent conclusions.

Figure 7: Scenario visualization for one of the observed patterns (Metaphor connect)

For visual depiction we needed to highlight our key findings regarding the type gestures people were most comfortable with. In addition we also wanted to portray the various functions expected. Accordingly we chose the television control scenario for depiction. The idea was to showcase the idea as simply as possible using a situation that most people can relate with. The thought also was that the video should be understandable without audio.

Feasibility

Video analytics coupled with sensing would be required for this system to work. As shown in the video, there can be multiple possibilities. Utilizing a hand worn camera to send visual inputs to the devices that need to be controlled is one possibility. Another direction of thought is having a general network of cameras in the environment that can detect the gestures. These cameras connect to a centralized server that analyses a gesture and sends commands to the intended devices which all in turn, connected to it

The need is that the cameras need to identify the objects pointed and connect to a database at a central processor with this information. The server then needs to search for the possible interactions that would be possible by connecting these two objects. If there is more than one option, then the system would either give the user options to select from among the multiple options or select one option based on the context.

For example, if the user points a vacuum cleaner to a television, he could mean that he wants to view the status/alerts etc. on the television screen or metaphorically speaking would mean that the image on the television should be cleaned (there could be more actions, these are taken for the example). The system could then analyze the context and check if the image is disturbed and then go to the other option directly without asking the user, or could give both the options to the user to select. These options could be provided in a voice based format or could be visually provided to the user.

Conclusion and next steps

Our limited research helped us in corroborating our initial hypothesis that connecting otherwise two real world objects could make sense to people. In fact it really excited people and to see them thrilled is our motivation to go ahead with prototyping the concepts.

We want to use the existing video and sensing capacities of our organization and work on prototyping the concepts for a home and office scenario. We would be exploring the UI possibilities in more details, for which a more detailed user research will be conducted.

We hope that our little foray into new gesture interaction paradigms would be one of the important milestones in HCI research in marching towards 'thought based interfaces'.

Acknowledgements

We thank the Design Innovation team's participation in the research. Also, would like to acknowledge the contribution of our team leader Parag Trivedi and our domain manager Utkarsh Biradar for their help and support in shaping up the concepts.

Example citations

[1] Buxton, W., Billinghurst , M., Guiard, Y., Sellen, A., Zhai A. *Human Input to Computer Systems: Theories, Techniques and Technology*. http://www.billbuxton.com/inputManuscript.html

[2] http://canesta.com/products-and-technology/products, http://www.kickerstudio.com/blog/2009/03/case-study-gestural-entertainment-center-for-canesta/