

# Comments & Selected References on Gesture

Bill Buxton

First draft, May 7, 2013

Current draft, May 8, 2013

## Introduction

With interaction expanding beyond the desktop, and the accompanying emergence of new devices, applications, and contexts, “gestures” and “gesture interfaces” are becoming an ever-greater topic of interest. However, while there is certainly a new interest in the use of gesture in interaction design, and there are new technologies which greatly enhance our ability to enhance interaction through the incorporation of gesture, the topic itself is anything but new. For example, in the fields of cultural studies and human communication, there is an extensive literature on the topic of gesture. See for example, Bremmer & Roodenburg, (1991) and Kendon, (2004). Likewise, in the field of human-computer interaction, there is a literature that pre-dates the release of the Apple Macintosh Computer which presages, and in some ways is still ahead of, what we are seeing emerge today. See, in particular, Krueger (1983; 1991) and Bolt (1984).

Despite the importance of the topic to the design of emerging technologies, it is my impression is that we collectively make too little use of this prior work, work which might otherwise better inform and guide us in the real and important challenges that face us in terms of better exploiting human potential, and reducing complexity, in our designs.

This brief document is an attempt to help shift the sands, ever so slightly.

The main part consists of a brief bibliography. I make no claims for it being representative, complete, or even containing the best references in the various topics covered. Rather, it is just a list of what happens to be sitting on my personal office book-shelf. For better or worse, these are the books that I have gone to in terms of shaping my own perspective on the topic. And please note my use of the term *books*. Obviously there is a large literature consisting of articles in conference proceedings and academic journals. I make no attempt to cover that space here.<sup>1</sup>

Second, I have included a brief section which is intended to provide some inkling of how I think about such things, the importance of developing appropriate taxonomies, and the need for us collectively to be far more specific in our language when speaking about the topic. My bias, as will be quickly seen, is my unapologetic view that the term “gesture”, used alone, borders on being devoid of information.

## Background

The emerging interest in gesture should come as no surprise. Gestures – in the common usage of the term - are an integral part of human communication in general. Just watch how animated people’s hand movements can get while speaking on the phone – despite them knowing full well that the other person in the conversation can’t see them. If we can trust this example, it may be safe to conclude that at times gesture may well be as important to the speaker as to the person being spoken to. But even if true, we have still not even scratched the surface of what is there. The nature of “gesture” is far more complex than can be captured in any single example or definition. That is both the opportunity and the problem.

---

<sup>1</sup> However, [Buxton \*et al.\* \(in preparation\)](#) has an extensive bibliography, both in terms of the main [References and Bibliography](#), as well as the interim bibliographies for the chapters on [Marking Interfaces](#) and [Gesture Driven Input](#).

In fact, my sense is that the term gesture is so rich and broad, it may defy definition. Just consider the breadth of that range reflected in even a few representative examples: the notion of musical gesture (Renard, 1982), the gesture of giving roses to a loved one, giving someone the finger, waving in front of a dispenser in order to get a paper-towel, and shaking hands. How can one distill all of these meanings into a single, clean and concise definition?

Even if we restrict our definition to interaction with technology a “gesture” is required to articulate *any and every* input method that involves human motor-action. If we accept this statement as true, then even our restricted definition of the term is far too general and imprecise to be of much use. If something means almost everything, then the entropy approaches zero, and there is no information conveyed in the message. My conclusion: we need to be more specific.

Let us try to narrow things down even further, in order to increase the information conveyed. First, for the examples which follow, let us restrict ourselves to the sub-set of gestures which involve human motor-action to control some function. Let us restrict things even further, by considering the specific limbs, joints, and muscles that are primarily involved – the hand and its digits. Within this narrowed scope, then let us consider the gestures involved in terms of their dimensionality – the richness of the response resulting from the stimulus associated with the gesture.

As our example, let us contrast typing on a QWERTY keyboard and playing music on a piano. Both have keys operated by the fingers. However, no matter how hard or fast you strike the key, the QWERTY keyboard is going to enter the same thing: the character associated with the key struck. On the other hand, the response of striking a piano key is highly dependent upon how it was struck. Yes, like the QWERTY keyboard, if the associated key is struck, middle-C will sound, no matter how I strike it. But how it is struck - how hard, how fast, and when it is released - makes a world of difference in the character of the sound that results – its loudness, “bite”, and sustain. Compared to the QWERTY keyboard, the coupling between motor-action and the end result is far richer with a far higher level of dimensionality. Certainly, there are important similarities in the operation of typewriter and piano keyboards. In order to obtain a high standard of performance, both involve a high level of motor-skill involving all ten digits of the hands. But the differences are as important as the similarities, and I would argue that while both involve “gesture”, it would be wrong to lump both into some general category of “finger-operated keyboard gesture”. As the saying goes, “God is in the detail.” Understanding and insight, I would argue, comes from understanding differences as much as it does from similarities. The granularity of our analysis must be finer than this.

Therefore, let us refine things even further by considering how dependent the response to a particular stimulus is on a specific gesture. My thinking here is that there is a continuum where, at one extreme, the particular gesture used is verging on arbitrary in terms of the end result, and on the other extreme, the result could only be achieved by means of a particular specific gesture articulated in a specific way.

Going back to our last example, and considering a single key press, it is pretty clear that I could reliably type any particular letter on a QWERTY keyboard with my tongue. While awkward, slow, and unsanitary, the result will nevertheless be exactly the same as if I pushed the key using my finger. Likewise, I could probably make any note on the piano sound by pushing the key with my tongue. But that is where the similarity stops. In contrast to the QWERTY keyboard, I could *not* reliably replicate the result of single notes played by a skilled pianist. In fact, I couldn’t do so even if I was using the same finger that they used, much less my tongue. Touch, in musical terms, turns out to be a highly learned skill, specific to both the physiology of the gesture and the technology designed to capture it. Speaking in general terms, where this is not taken into consideration, needs to be undertaken very carefully. Let me illustrate this with a less far-fetched example.

Consider the pinch-to-zoom “gesture”. Where would you place along the continuum represented in the last QWERTY / piano keyboard example? For my part, I lean very much towards the former. Here is why. To control the zoom factor, the only thing required is to establish the position of two points on the screen and then sense the change in distance between them. If they get further apart, you zoom in

proportionally, and vice versa if they get closer together. It does not matter how those two points are established or how they are controlled. It could be a touch screen that senses the position of your thumb and index finger, therefore employing pinching. However, depending of form-factor or scaling factor, it could be just as easy, or easier, to use the index finger of each hand instead. The zooming technique – as opposed to the gesture used to articulate it – does not change in either case. This is the case even if different hardware is used. For example, one could use a mouse in one hand and a track-ball in the other, and the same technique works. Likewise, one might use a bimanual approach employing a stylus and touch surface (screen or pad) to control the zoom. These are just a few ways to articulate control of the zoom, and each involves its own set of gestures. Hence, at least in general, it seems inappropriate to characterize the technique by any single gesture. Rather, the defining nature of the interaction is a “zoom by two-point manipulation”. “Pinch-to-zoom” is just a single instance within this much broader class. It depends on the context as to whether the general class, or the means of articulation of a specific instance is more relevant to a conversation. If one does not distinguish between the two, or understand the potential significance of such distinctions, the likelihood of muddled analyses is greatly increased. It is only through making clear and explicit distinctions, and the disciplined use clear language, that we can avoid apples-and-bananas comparisons - which when they occur, get embedded in the literature, and thereby both perpetrate fuzzy thinking and retard progress.

Making sense by uncovering meaningful dimensions of differentiation is the art (or science) of creating taxonomies. Yet, in what I have written here, I have made no attempt to create such a taxonomy; rather, I have merely attempted to express the need to do so, a hint of how to approach it, and what the potential benefits might be.

Overall, this note, itself, is a gesture – one which reflects that I care enough about the topic to attempt to actively engage in its advancement. In keeping with that, comments, questions, and suggestions are welcomed.

## From My Bookshelf ...

- Bolt, R. A. (1984). *The Human Interface: Where People and Computers Meet*, Boston: Lifetime Learning Publications.
- Bremmer, J. & Roodenburg, H. (Eds.)(1991). *A Cultural History of Gesture*. Ithica: Cornell University Press.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How our hands help us think*. Cambridge MA: Belknap Press.
- Hummels, C. (2000). *Gestural Design Tools: Prototypes, Experiments and Scenarios*. PhD Thesis, Technical University of Delft.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Krueger, Myron, W. (1983). *Artificial Reality*. Reading, MA: Addison-Wesley.
- Krueger, Myron, W. (1991). *Artificial Reality II*. Reading, MA: Addison-Wesley.
- Kurtenbach, G. & Hulteen, E. (1990). Gestures in Human-Computer Communications. In B. Laurel (Ed.) *The Art of Human Computer Interface Design*. Addison-Wesley, 309-317.
- MacKenzie, C.L. & Iberall, T. (1994). *The Grasping Hand*. Amsterdam: North-Holland.
- McNeill, D (1992) *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.
- McNeill, D. (2005). *Gesture & Thought*. Chicago: University of Chicago Press.
- Morris, D. (1994). *Bodytalk: The Meaning of Human Gestures*. New York: Crown Publishers.
- Morris, D., Collett, P., Marsh, P. & O'Shaughnessy, M. (1979). *Gestures*. New York: Stein and Day
- Mulder, A. (1996). *Hand Gestures for HCI. Hand Centered Studies of Human Movement Project Technical Report 96-1 School of Kinesiology, Simon Fraser University*.
- Renard, C. (1982). *Le Geste Musical*. Paris: Hachette / Van de Velde.

- Siegel, I. (1998). *All About Bone: An Owner's Manual*. New York: Demos Medical Publishing.
- Smyth, M. & Wing, A. (Eds.)(1984).*The Psychology of Human Movement*, London: Academic Press.
- Sturman, D. & Zeltzer, D. (1994) A Survey of Glove-Based Input. *IEEE Computer Graphics and Applications*, 14, 30-39.
- Trumble, A. (2010). *The Finger: A Handbook*. New York: Farrar, Straus and Giroux.
- Wilson, F.W. (1998). *The Hand: How its use Shapes the Brain, Language, and Human Culture*. New York: Pantheon Books.