Towards the Establishment of a Framework for Intuitive Multi-touch Interaction Design

Amy Ingram
University of North Carolina at Charlotte
9201 University City Blvd.
Charlotte, NC 28223
011-1-704-687-7741

aingra17@uncc.edu

Xiaoyu Wang
University of North Carolina at Charlotte
9201 University City Blvd.
Charlotte, NC 28223
011-1-704-687-7741

xwang25@uncc.edu

William Ribarsky
University of North Carolina at Charlotte
9201 University City Blvd.
Charlotte, NC 28223
011-1-704-687-8559

ribarsky@uncc.edu

ABSTRACT

Intuition is an important yet ill-defined factor when designing effective multi-touch interactions. Throughout the research community, there is a lack of consensus regarding both the nature of intuition and, more importantly, how to systematically incorporate it into the design of multi-touch gestural interactions. To strengthen our understanding of intuition, we surveyed various domains to determine the level of consensus among researchers, commercial developers, and the general public regarding which multi-touch gestures are intuitive, and which of these gestures intuitively lead to which interaction outcomes. We reviewed more than one hundred papers regarding multi-touch interaction, approximately thirty of which contained key findings we report herein. Based on these findings, we have constructed a framework of five factors that determine the intuition of multi-touch interactions, including direct manipulation, physics, feedback, previous knowledge, and physical motion. We further provide both design recommendations for multi-touch developers and an evaluation of research problems which remain due to the limitations of present research regarding these factors. We expect our survey and discussion of intuition will raise awareness of its importance, and lead to the active pursuit of intuitive multi-touch interaction design.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – input devices and strategies, standardization, theory and methods.

General Terms

Design, Documentation, Human Factors, Standardization.

Keywords

Multi-touch, Intuition, Interaction, Gesture

1. INTRODUCTION

Although multi-touch research has been ongoing for many years,

it is only with the recent commercialization and mass production of affordable multi-touch hardware that design questions have become more readily addressable. While multi-touch interactions are widely regarded as intuitive and easy to use, there is a lack of consistency in their implementation. The use of specific multi-touch gestures for specific interaction outcomes has yet to be standardized for the vast majority of interactions. This paper seeks to clarify the current state of multi-touch interaction with regards to intuition and encourage more thoughtful development of multi-touch interactions. While other factors should also influence multi-touch designs, the scope of this paper is limited to intuition.

Improving the intuition of multi-touch interactions is important for three reasons: improvement of the user experience among the general public, facilitation of the analysis process for analysts who use multi-touch interfaces [20], and perhaps most crucially, the cost of change. Multi-touch interactions are not yet standardized, making now the best time to ensure their efficiency and effectiveness. Interaction techniques that are standardized have a much higher cost of change; for instance, the qwerty keyboard layout is subpar and limits typing speed [19], but remains in use due to the inestimable cost to change to a more efficient layout.

As noted by Beaudouin-Lafon, the cost of change for new user interfaces is already enormous [4]. Even as factors such as computer memory and speed have increased exponentially, standards such as the WIMP (Window Icon Menu Pointer) interface and the mouse and keyboard as input devices have remained in place for more than twenty years. To be accepted, therefore, multi-touch interfaces must perform so well that their benefits outweigh twenty years of user and design experience. This is a high yet necessary standard. Pirker et al. have shown that the usability of new technology compared to previous versions is a strong determinant of whether it is adapted [21]. Also, users are not tolerant of poor design; both observational [3] and user studies [1] have demonstrated that users will give up on a multi-touch interaction altogether when they have no idea how to achieve their desired outcome. Due to the enormous cost of change involved in both transitioning to multi-touch interfaces and changing standardized interactions, it is crucial that multi-touch interactions are made as intuitive as possible now. It is our hope that the insights and questions presented within this paper will foster more discussion of this issue within the research community.

To comprehend the current state of multi-touch interaction design, we conducted a broad literature survey. The vast number of multi-touch interfaces in existence led us to choose an area of focus for our review. We found multi-touch visualizations to be appropriate due to their complexity and diversity, as discussed further in Section 5. Our survey was based on more than 100 papers found in the visualization and human-computer interaction research

fields through both IEEE and ACM. Many papers from the SIGCHI, TableTop, and UIST conference proceedings were included. The selection of this survey focused on several types of previous research, including user studies that assess the intuition and usability of multi-touch gestures, implementations of multi-touch applications, observational studies of existing multi-touch applications, and theoretical works concerning interaction theories.

The contributions of this paper are threefold. First, we examine previous research to determine a framework of factors that influence the intuition of multi-touch interaction. From these factors we provide design recommendations for multi-touch developers and evaluate the extent to which these factors have been studied. Second, we determine the level of consensus among researchers, commercial developers, and the general public regarding which multi-touch gestures are intuitive, and which of these gestures intuitively lead to which interaction outcomes. Third, we call attention to problems in intuitive multi-touch interaction design that arise concurrently with the pursuit of more robust multi-touch implementations, and propose both new research directions and possible solutions.

We conclude this introduction with a brief discussion of terminology. Throughout this paper, we make a distinction between multi-touch gestures and multi-touch interactions. We use the term multi-touch gesture to refer to the physical motion of the users' hands and fingers as they interact with the multi-touch surface. We use the term multi-touch interaction to signify the entire interaction, which includes the physical gesture, the outcome of the interaction, and the context in which the interaction is performed. The context includes visual information such as feedback, as well as the users' previous experiences and knowledge. While these terms may be interchangeable in current usage, we hope this distinction will provide language to facilitate a more in-depth consideration of the components of intuitive multi-touch interaction.

2. TOWARDS A FRAMEWORK OF MULTI-TOUCH INTERACTION

Intuition is defined as "the power or faculty of attaining to direct knowledge or cognition without evident rational thought and inference" [10]. The goal of intuitive multi-touch interaction research is to minimize the effort required of users to learn, use, and remember the interactions so they will be as intuitive as possible given real-world limitations. In this section, we discuss five factors that influence the intuition of multi-touch interactions: direct manipulation, physics, feedback, previous experience, and physical motion.

2.1 Direct Manipulation

The most influential of these factors in the intuition of multi-touch interaction is the use of direct manipulation. Direct manipulation is defined by Tweedie [25] as a "literal representation of physical behavior in the real world." Shneiderman [23] defines direct manipulation more specifically as interaction that 1. is physical, 2. is performed on a continuously represented object, 3. results in rapid, incremental and reversible operations, and 4. has layered learning requirements, accommodating both novice and advanced users [23]. These requirements mimic the natural laws of real-world interaction; we act physically upon objects over time. Many common multi-touch interactions, such as moving, resizing, and rotating objects, make use of direct manipulation.

There are three main advantages of adapting real world interactions to a digital environment with direct manipulation, all of which improve interaction intuition. First, the ability to physically touch an object on the screen eliminates the need for intermediate interaction devices that must be learned, such as the mouse or joystick. Eliminating this barrier to interaction makes the process simpler, as demonstrated by Murata and Iwase in a study comparing pointing speed for the mouse and a finger [18]. Second, by mimicking the behavior of real-world interactions, direct manipulation makes use of knowledge so fundamental that it seems instinctual. Users have experienced real-world interactions from infancy, and can understand interactions that follow the laws of physics without conscious thought, making them intuitive. Further, when this pool of knowledge regarding the physical world is applied to multi-touch interactions, it helps users form semantic, rather than syntactic, memories [23]. Shneiderman states that the former type of memory is easier to recall due to the connections formed with previous memories and knowledge [23]. Finally, this pool of knowledge is universal. Unlike languages or cultural norms that differ, every potential user has experienced the same physical laws.

Several studies have confirmed that interactions that use direct manipulation are more intuitive than abstract interactions. For instance, Mauney et al.'s research on multi-touch interaction across different cultures, described in detail in Section 3.2, found that participant agreement was greater for direct manipulation interactions than interactions classified as symbolic [17]. Direct manipulation interactions have also been evaluated more favorably than abstract gestures in evaluations of specific interfaces. For instance, when evaluating the implementation of multi-touch interactions for navigating a large 3D environment, Fu et al. found that participants had the most difficulty using an abstract, two-handed gesture to change spatial scale, while the direct manipulation interactions—one-finger rotation and five-finger panning—users considered intuitive and easy to perform [8].

Direct manipulation is often the most intuitive solution for interactions that can make use of it. However, there are several limitations of these interactions that existing research has not fully accounted for. These include the need to distinguish between abstract and non-abstract interactions, the amount of physical effort required to perform direct manipulation actions, previous experiences that contradict direct manipulation, and the consideration of intuition for a large number of coinciding interactions rather than isolated, individual interactions.

Abstract interactions that do not mimic real-world behaviors cannot take full advantage of direct manipulation's benefits, as direct manipulation relies on the intuition of real-world behavior. Further, although Wobbrock et al. manually categorized interactions for their user study [29] and Mauney et al. reference "symbolic" gestures as a distinct gesture type [17], to the best of our knowledge, there has been no definitive work regarding what makes one interaction abstract and another not. For instance, is a finger swipe left and right for forwards and backwards motion—whether used for exploring a timeline, navigating in an internet browser, or for undoing and redoing actions—a direct manipulation interaction? Although it mimics the turning of book pages, it relies upon cultural connotations of left as past and right as future as much as any physical laws, and the motion is not likely to be continuous.

Further, direct manipulation does not take into account the costs of physical effort involved, especially if one needs to interact with scattered objects across a large surface, or use both hands to constantly rotate large environments. In isolation, the use of direct manipulation also fails to consider contradictory previous experiences, or acknowledge experience gained by other sources,

such as other multi-touch implementations or the use of other technological hardware. Finally, the use of direct manipulation is limited by the need for distinct interactions and the limited number of intuitive physical gestures, which is smaller than the number of possible interaction outcomes. This problem is discussed in further detail in Section 4.2.

Despite these limitations, direct manipulation is a powerful tool for developers. Direct manipulation allows for freeform interaction, making it useful for interactions such as annotation, selection, and movement. The intuition of direct manipulation can be further enhanced with proper use of physics and feedback, as detailed in the following two sections.

2.2 Physics
The success of direct manipulation is dependent on the user's intuitive understanding that interaction with objects on the multi-touch screen works the same way as interactions in the real world. Real world interactions are governed by the laws of physics; multi-touch interactions can make use of users' knowledge using pseudo-physics. Two aspects of physics commonly integrated into multi-touch interactions are gesture size and speed.

Research has shown that users tend to associate larger gestures with larger interaction outcomes, analogous to the use of greater force in the physical world. Koskinen et al. found that users associated larger hand motions with larger-scale outcomes, such as managing large visual objects and navigating through the environment, and associated smaller physical movements with smaller, more detail-oriented interactions, such as manipulating interface elements [12]. In another study by Kristensson et al., users interacting with a node-link diagram preferred to use one finger or a pen to complete most tasks, but used larger, two-handed interactions to manipulate the whole diagram, or move multiple objects at once [13]. Multi-touch developers should note this distinction and associate large gestures with large-scale interactions, remembering that large gestures are also physically taxing. Logically, repeatedly used interactions should not be so fatiguing to the user that they are unsuitable for extended use.

Users also expect gesture speed to influence multi-touch interactions. They "throw" objects across multi-touch tables [29], expecting that the objects' speed will increase proportionately with faster movement, like they would in the real world. Developers can take advantage of this behavior by using speed to differentiate interactions. Many multi-touch applications, such as the room-planning application presented by Wu and Balakrishnan [31], support speed differentiation for one-finger drag. This results in distinct "move" and "flick" gestures with different outcomes. In another example, the astrophysical simulation created by Fu et al. [8] allows users to differentiate between a "rotate" gesture of moderate speed, and a faster gesture that sends the rotation into auto-pilot, where the planet will continue spinning until it is stopped. However, too much subtle gesture differentiation of this sort is potentially confusing and frustrating to users. Without proper feedback, users may not understand the distinction between two unique interactions. For instance, Wobbrock et al. found that users did not naturally differentiate the number of fingers used for a gesture beyond two broad categories—touches with one to three fingers, and five finger or whole hand touches— but noted that gestures with finer distinction have been successful previously [29]. Further, users gesture at different speeds; what may be fast for one user may be slow for another. The elderly, for instance, tend to gesture more slowly than younger users, as noted by Stößel [24].

To the best of our knowledge, there has been no previous research that has quantified the effects of pseudo-physics in interaction on intuition. Despite limitations such as physical effort required, user differences and the limits of gesture differentiation, interactions that mimic the behavior of real-world physics enhance both the usability and understandability of multi-touch interfaces.

2.3 Feedback
Another crucial component of multi-touch interaction is feedback. Multi-touch screens do not provide the tactile feedback experienced in the physical world, and the presence of fingers on the display can cause occlusion if visual changes are minimal [28], resulting in confused and frustrated users who do not know whether their gestures were successful or what they accomplished.

As noted by Frisch et al., users use context to resolve gesture ambiguity [7]. When interacting on a surface with no visual feedback, study participants stated that the lack of both tangibility and visual feedback felt unnatural [12]. In another user study conducted by Wu and Balakrishnan, a user specifically requested visual feedback when selecting an object across the table with a "catch," or reverse-flick, interaction [31]. This example illustrates the confusion possible without appropriate feedback; if objects are clustered together across a table, how is the user to know which is selected without a visual indicator?

Several multi-touch implementations make good use of feedback. For instance, a tree visualization implemented by Andrews [1] features the option to move nodes closer or further apart. An immediate, continuous change to the size of the surrounding nodes indicates that the change has been made. Also, potentially abstract interactions can be made more concrete through the use of appropriate feedback. Butkiewicz and Ware make use of continuous feedback for 3D point selection, using a three finger gesture to position a continuously visible marker [5]. This allows users to continuously view the change in the marker's location as they gesture, making what might seem an abstract process more approachable. Several applications on the iPad use feedback to provide cues for future interaction; for example, the curling of a digital page when a touch approaches that corner indicates that more options are available by interacting in that space [2].

Visual feedback, while predominant, is not the only form of feedback available. Researchers have studied the use of audio as well as tactile feedback to improve the intuition of interactions on multi-touch screens. For instance, Lee and Zhai found that multiple forms of feedback, including audio, tactile, and a combination of both, improved users' performance when using touch screen buttons [15].

Multi-touch developers can make use of feedback at all stages of the interaction process: before the interaction, to provide cues that an interaction may be performed; during the interaction, to provide a visual indicator that the interaction is in progress; and afterwards, to indicate what change has occurred. Good feedback is invaluable to intuitive interaction, as it allows users to make sense of the environment.

2.4 Previous Experience
While knowledge of the physical world is the most universal, it is not the only type of previous experience that influences users' interactions. An observational study conducted by Bartoschek and Schwering [3] found that users with previous multi-touch experience were faster in learning new interactions with multi-touch tables. Several studies [9, 17] have found that a user's choice of directionality in multi-touch interactions, such as scrolling or navigating forwards and backwards, is influenced by

previous experience with multi-touch devices. For instance, Mauney et al. noted that 70% of participants who owned a multi-touch device swiped up to scroll down, compared to 50% of participants who did not own a multi-touch device [17]. Also useful to the developer are interactions familiar to users from other experiences with technology, such as double clicking and right clicking. While not innately natural the same way direct manipulations are, these interactions also build upon a pool of pre-existing knowledge. Further, their implementation may make it easier to transfer existing desktop applications to multi-touch environments.

Developers must consider the learning effect of previous multi-touch interaction experiences in their design. If interaction outcomes contradict users' expectations, they will be less intuitive; familiar or complementary outcomes will be more intuitive. Overall, the use of familiar concepts, whether derived from abstract metaphors or the use of a desktop computer, can aid in intuitive multi-touch interaction.

2.5 Physical Motion

The previous sections have emphasized the minimization of effort required for users to associate a physical gesture with an interaction outcome. However, the intuition of the physical motion alone should also be taken into consideration. Gesture size and ergonomics can differentiate intuitive and non-intuitive gestures; multiple studies have demonstrated a universal preference for minimized physical effort when gesturing.

For instance, Koskinen et al. conducted a user study in which ten participants performed a set of 25 predefined gestures on a mockup touch table with no visual feedback [12]. They found that users preferred simpler gestures in every instance. Users found smaller gestures more comfortable than larger gestures, and preferred short, fast, accelerating touches. Users overwhelmingly preferred one finger to multiple finger touches, as well as one handed to two handed interactions.

Another study by Epps et al. found that one-finger pointing was by far the most common hand position, used 70.1% of the time for 36 tasks on a non-functioning table [6]. The next most common hand positions were a flat hand with separated fingers (20.0%) and a flat hand with the fingers together (11.4%) [6]. Observational studies such as Ryall et al.'s confirm that novice users usually gesture with one finger on a multi-touch table [22].

Further, the work of Frisch et al. on interactions using a multi-touch implementation of a node-link diagram, described in detail in Section 3.2, found that 60% of tasks had fewer variations when users used one hand rather than two [7]. Participants preferred one-handed over two-handed interactions for 12 of 14 tasks. Two-handed interaction was preferred for the zoom and scale tasks, larger interaction outcomes suited to larger gestures, as described in Section 2.2 above. Participants also overwhelming preferred to begin each task using one hand, or a pen and a free hand, rather than with two hands, again in every instance except scaling the entire diagram [7]. A final user study, conducted by Wobbrock et. al, found a similar preference for one-handed gestures [29]. Users preferred one-handed gestures for 25 out of 27 tasks, and were evenly divided on the remaining tasks, with two-handed gestures never preferred overall [29].

The results of these studies indicate that users are strongly inclined against overly demanding physical gestures. However, multi-touch gestures that require greater physical effort are not always unintuitive; many direct manipulation interactions require physical effort to reach for and manipulate objects. Further work needs to be conducted on the balance of the physical and mental

components of intuition, to quantify when each should take precedence for ideal intuitive multi-touch interaction.

3. THE EXTENT OF CONSENSUS

Having discussed five factors that can influence the intuition of a multi-touch interaction design, we now address presently implemented multi-touch applications, to determine what level of agreement exists regarding implementation of intuitive multi-touch interactions. Common user interactions with the mouse and keyboard became standardized over a long period of time. Interaction concepts such as "double-click" are now common across a wide spectrum of operating systems and software, and multiple keyboard shortcuts have become standardized, such as control or command-S for the operation "save".

As mentioned in the introduction, multi-touch interactions have not yet achieved this level of standardization. Wigdor et al., in the introductory paper for a CHI Special Interest Group on designing multi-touch user interfaces, lament that "with this increase of devices and interfaces, comes an unfortunate increase of non-standard design patterns and methods of user interaction" [27]. In the following section, we examine the extent of agreement regarding which multi-touch interactions are most intuitive. We approach this question of consensus with two measures: first, to what extent do multi-touch commercial developers and researchers—who presumably desire their applications to be intuitive and user-friendly—agree, and second, are the users themselves in agreement.

3.1 Consensus Among Researchers and Commercial Developers

Overviews of some recent multi-touch visualizations confirm there is a wide diversity of opinions regarding which gestures should produce which outcomes. Table 1 lists interactions used within eight multi-touch research implementations: a 3D visualization of the human body designed to aid orthopedic surgeons [16], a tree visualization [1], a room-planning interface [31], a 3D visualization of ocean currents [5], a visualization for tagged photo exploration [13], a navigable 3D astrophysical simulation [8], a tool for collaborative document analysis [11] and Google Earth [3]. A large number of interactions used unique physical gestures.

There is arguably more standardization of multi-touch gestures, if not interactions, in commercial multi-touch systems. For instance, the Apple iOS, Windows Phone 7 and Android mobile operating systems all support some variant of the gestures tap, double tap, touch and hold, drag, flick and pinch [30]. Differences remain, however: Apple's OSX differentiates between two, three and four finger dragging for scrolling, window navigation and screen navigation, while Windows 7 supports a two finger tap for centred and restored zooming, and a one finger press, one finger tap option for right clicking [30]. Microsoft Surface supports two, three and five finger pinch and spread gestures, while OSX uses a four finger pinch interaction to activate the Mission Control navigation screen [30]. Further, the majority of these gestures are available for application developers to use as they wish, leaving the actual outcome of each interaction further in question.

Although the differences are widespread, there is some level of consensus among designers. The most ubiquitous gestures are a one finger touch—which mimics the press of a button, the click of a mouse, or pointing to establish relevance—and a one finger drag, often used for movement of objects or the camera view.

These two gestures are, to the best of our knowledge, supported by all multi-touch interfaces. The interaction outcomes they produce vary depending on the context, but tend to be fundamental and common outcomes such as navigation [3, 30], selection [8, 16, 29], or movement [1, 5, 29, 31]. These gestures, and the interactions they support, are more standardized than any other multi-touch interaction.

A two-finger pinch and spread motion for zooming or scaling objects is also widespread, although not universally used; alternative methods include holding two fingers stationary while moving one [30] and dragging multiple fingers apart or together [29]. Other multi-touch interactions are not currently implemented by developers in any standardized fashion. For instance, rotation is a common interaction outcome that can be produced with direct manipulation. However, developers have paired this interaction outcome with a wide variety of physical gestures, as shown in Table 1. This example illustrates the inconsistency common in multi-touch interactions at the present time. Overall, developers tend to show less consensus as interactions become more abstract, and more consensus with the most straightforward interaction outcomes (e.g., move this object to this location), which often make use of one finger drag and one finger touch. Both levels of consensus present unique design challenges, as discussed further in Section 4.1.

3.2 Consensus Among the General Public

Another measure of agreement regarding intuitive multi-touch interactions is the consensus among the general public. Several studies have investigated whether study participants, when told to perform a task using a multi-touch gesture, will consistently produce the same gestures for an interaction outcome.

Wobbrock et al. performed research of this sort as part of the creation of a user-defined gesture set [29]. Twenty participants were seated at a table which displayed interaction results. After viewing each result, participants were asked to perform both one and two-handed gestures they felt would best reproduce the outcome. This process was repeated for 27 interaction outcomes, such as rotate, duplicate, accept and undo. Wobbrock found a wide range of agreement levels, from 100% agreement on one-handed moving of objects, down to slightly over 10% agreement on gestures such as "next" and "cut" [29]. The mean agreement was 32% for one-handed interactions and 28% for two-handed interactions [29].

Frisch et. al performed a similar study that found a similar variation in consensus [7]. The gestures were performed on a node-link diagram. The researchers added together the number of different gestures the 17 participants used to complete 14 tasks, each performed three ways: once with a one handed gesture, once with a two handed gesture, and once with both a pen and a free hand. The outcome with the most consistent gestures was selecting a node (13 different gestures used across all interaction methods). Other outcomes with high consistency included creating different types of edges (14-15 gestures) and moving a single node (15 gestures) [7]. The outcome with the least consistency was copying a subgraph (34 gestures); other outcomes with low consistency were scaling a single node (29 gestures) and deleting nodes or edges (26 gestures) [7].

The work of Mauney et al. also found variation in the gestures chosen for different tasks [17]. This study investigated cultural differences in gestures performed by 340 participants from nine countries across Europe, Asia and North America. Participants

were asked to make gestures on a mock-up of a hand-held touch device when shown before and after images of what should occur. Mauney found that no gestures were used more frequently in one country than another; the only major difference was that participants from China used significantly more symbolic gestures than participants from other countries [17]. Overall, the mean agreement score was .35 for direct manipulation gestures and .18 for symbolic gestures [17]. The three highest agreement scores for interactions classified as direct manipulations were move (.91), rotate (.55), and stop scroll (.48) [17]. The three highest agreement scores for interactions classified as symbolic were accept/verify (.46), delete (.26) and back (.25) [17].

Overall, these studies found a wide range of variability in agreement for each interaction. For instance, Wobbrock found that duplicating an object was highly agreed upon for one-handed interaction, while in Frisch's study copying the subgraph was the least agreed upon gesture for one-handed interaction [7, 29]. However, some patterns are apparent. Users, like commercial developers, found one finger touch and drag to be the most intuitive gestures. The most agreed upon interactions overall, movement and selection, made use of one finger drag and touch respectively [6, 7, 29, 17]. Less agreed upon interactions implemented in multiple user studies include deleting (or closing) objects and zooming [6, 7, 29]. The majority of interactions were not consistently agreed upon, with few receiving even 50% agreement in any study.

Table 1. Multi-touch interactions in research implementations

Outcome	Gesture	References
selection	one finger tap	[5,8,11,13,31]
	one finger drag	[11]
pan/move camera	one finger drag	[3, 5]
	two finger drag	[16]
	five finger drag	[8]
move objects	one finger drag	[1,11,13,31]
	vertical hand position drag	[31]
	rearrange objects	[1]
throw/catch objects	one finger flick	[31]
rotate camera	one finger drag	[8, 16]
	two finger rotational drag	[16, 31]
	three finger drag	[5]
	multiple finger rotational drag	[8]
	one finger hold, one drag	[8]
	flat hand position rotates	[31]
tilt camera	two finger hold, one drag	[3]
rotate objects	one finger hold, one drags around	[1]
zoom	two finger drag apart/together	[3, 8, 16, 31]
resize objects	multiple finger drag apart/together	[11]
resize by powers of ten	two finger hold, one drag	[8]
toggle mode	one finger tap	[1, 8, 16]
show menu or details	one finger hold	[5, 11]
	horizontal hand	[31]
	tilted horizontal hand	[31]
define region	four finger touch	[16]
	two "L" shaped hands	[31]

It seems counterintuitive that multi-touch interactions, widely regarded as intuitive and easy to use, would be so seldom agreed upon. However, these user studies—which take place outside of the potentially confounding factors of a useable software program—ignore factors, such as feedback, that make implementations of these interactions easier to use. It is difficult to determine whether we can extrapolate results about what is intuitive in more complex software environments from experiments conducted in simplistic, non-realistic environments. This problem is discussed in further detail in Section 4.2.

Regardless, the set of multi-touch interactions that users and developers instinctively and unanimously agree upon is small, consisting of one finger touch for selection and one finger drag for movement, selection and other fundamental tasks. The lack of standardized and universally accepted interactions makes the need for well-designed multi-touch interactions even more crucial.

4. REMAINING RESEARCH PROBLEMS

The previous sections of this paper have demonstrated two major points: first, that the intuition of multi-touch interactions can be improved by considering factors such as direct manipulation, physics, feedback, previous knowledge, and physical motion, and second, that the design of multi-touch interactions has not yet become standardized, as developers do not agree on how to implement intuitive interactions, and users do not agree on which gestures are intuitive for which interactions.

As research into intuitive multi-touch interaction continues, two problems arise which have yet to be fully addressed by the research community, both of which relate to the increasing complexity of multi-touch applications. The first is the consideration of intuitive interactions within increasingly large systems, rather than in isolation. Optimizing interactions for intuitiveness becomes increasingly more complicated as the number of interactions increases. The second problem we note is the need for more effective evaluation of multi-touch interactions and implementations.

4.1 Consideration of Multiple Interactions

Intuitive multi-touch interactions must be considered within the context of other interactions, rather than in isolation, especially as multi-touch applications become more complex and start providing more interaction options to the user. As discussed in Section 3.2, the number of widely used, intuitive multi-touch gestures is small, limited to one-finger touch and drag. Designers can implement more reasonably intuitive interactions using the qualities discussed in Section 2, such as direct manipulation, physics and feedback.

The problem for designers occurs when an implementation requires either more intuitive interactions than there are simple direct manipulation gestures, or abstract interaction options. For instance, simple photo browsing applications are popular for multi-touch implementations, as they can be programmed with a minimum number of intuitive multi-touch interactions, such as moving, scaling, and rotating photos. Turning this simple application into a rigorous photo editing environment would require the addition of numerous tools, many abstract, such as free-hand selection of photo elements, annotation capabilities, photo cropping, and cut, copy and paste. As interaction complexity increases it becomes more difficult to keep the application intuitive and approachable. The designer is left to decide how many multi-touch gestures should be implemented.

One extreme approach would be to accommodate every interaction with a unique multi-touch gesture. This presents two problems: implementing more gestures than a user can learn without extensive training, and implementing less-intuitive interactions when interactions conflict. For example, users in Frisch et al.'s study found one finger drag to be the most intuitive gesture for numerous interaction outcomes, including moving nodes, creating edges, deleting objects, and copying a subgraph [7]. If these interactions were all given distinct gestures, only one could remain so intuitive.

A second extreme approach would be to limit used multi-touch gestures to only the most intuitive, primarily one-finger touch and one-finger drag. The context and speed of the gesture may differentiate interactions. Menus and buttons can also facilitate the reuse of gestures by triggering different interaction modes. For instance, an extended press on a text object in iOS triggers a menu that allows the one finger drag gesture, usually used for movement, to be used for text selection [2]. While this approach is highly intuitive in many instances, problems are present as well. Lam notes that gestures which produce different outcomes in different contexts can confuse users [14]. Further, a system that relies on menus and buttons for even the most common tasks may become frustrating, as it requires additional effort from the user for every interaction.

A combination of these approaches is beneficial when designing an intuitive multi-touch interaction system. Commonly used interactions that can make use of direct manipulation, such as navigation techniques, are good candidates for unique multi-touch gestures. Developers must weigh the effort required of the user to differentiate a gesture with the benefit derived from having the interaction readily available. Those multi-touch gestures that are deemed valuable enough to include may then be reused throughout the application. Beaudouin-Lafon introduces the concept of polymorphism—that is, the notion that tools should work in as many contexts as possible to reduce complexity [4]. This helps reduce both user frustration and required learning. One example of gesture reuse is implemented by Andrews [1], in which a double tap on a link-node display toggles the display size, while a double tap on the tree display changes the focus. A more complex interface presented by Isenberg and Fisher [11] uses one-finger touch and drag for multiple functions, including highlighting, selecting, and moving documents.

Conversely, some interactions may not be suited for multi-touch gesturing. Abstract interactions cannot make use of the benefits of direct manipulation and have some of the least-agreed upon multi-touch gestures, as discussed in Section 3.2. Implementing these interactions with menus and buttons is therefore a good way to make them more intuitive.

There has been a great deal of research into optimization of menus for a multi-touch table environment, but most such work is outside the scope of this paper. For our purposes, it is most worthwhile to note that menus and buttons can carry a great deal of the cognitive load that would otherwise burden the user. Users are used to menus and buttons from traditional computer interfaces; Mauney et al. found that users tap the touch screen for menu buttons when uncertain what to do, even when instructed to only make gestures [17]. Infrequently used interactions, and abstract interactions that cannot take advantage of direct manipulation, may be better suited for menu or button items. The user's mental effort is better spent remembering gestural interactions for frequently used interactions or those that require free-form interaction and are therefore suited for direct

manipulation. Any implementation of buttons or menus must also take into account the screen space used, as display space is both valuable and limited. A recent multi-touch visualization addresses this problem with a radial menu that only appears when the user touches the screen, making it accessible from anywhere and customizable to the situation, without taking up screen space unnecessarily [5].

Overall, the problem of propagating intuitive interaction design from a single interaction to a large scale implementation can be alleviated in several ways: implementing a limited number of multi-touch gestures reserved for the most important and frequently used interactions; reserving the users' cognitive load by using buttons and menus for abstract and infrequently used interactions; and using buttons to toggle menu states and reuse intuitive multi-touch gestures for new interactions. While we present these guidelines extrapolated from experience and research on the intuition of individual interactions, more work needs to be conducted with multi-touch interaction systems to determine conclusively how a large number of interactions should be combined to create the most intuitive system possible.

4.2 Evaluation In Realistic Environments
A second problem requiring further consideration from the multi-touch research community is the inverse relationship between increased evaluation of multi-touch implementations and increased environment complexity. The most complicated multi-touch implementations are the least evaluated, while the most thorough evaluations are conducted in the least realistic contexts.

The lack of rigorous evaluation of multi-touch implementations manifests in several ways. Evaluations are often performed on a statistically insignificant number of users, whether experts or members of the general population, and the data collected tends to be qualitative, which is more subjective. The evaluation focus tends to be on the software implementation itself rather than the interactions, or on interactions specific to the individual implementation, rather than usability principles as a whole.

There is a recognized lack of rigorous evaluation throughout the visualization community in particular, as evidenced by the panel discussion at VisWeek 2011 on verification in visualization [26]. For instance, some multi-touch visualizations are presented without any evaluation at all [1, 5]. The multi-touch visualization presented by Isenberg and Fisher was initially evaluated by one expert and two researchers [11]; the work of Lundström et al., by a panel of five experts [16]. This is an improvement, but still not sufficient to achieve statistical significance or quantitative results about these multi-touch interactions in context. Future research must overcome the challenges inherent to such an evaluation—multiple variables, difficult to measure metrics, and noisy experimental results—to advance our knowledge regarding multi-touch interactions, and quantify existing knowledge.

The multi-touch research that is best evaluated, with larger sample sizes, controlled variables and quantitative findings, comes primarily from the HCI community, and reports findings regarding the intuition of individual gestures or interactions. However, in order to isolate intuition as a variable, these studies are primarily conducted in unrealistic environments. For instance, in a study by Koskinen participants interacted with a square on a white background [12], while Frisch et al. studied users interacting on an entirely blank table with no visual feedback [7]. This raises the question of how reliable these findings are in more realistic contexts, where there are many more factors involved.

5. LIMITATIONS
While our goal in composing this paper was to provide an overview of the current state of intuitive multi-touch interaction, this work is limited in that it does not consider the full range of multi-touch applications. As there are many multi-touch implementations found in many domains, we could not be exhaustive in our evaluation of existing multi-touch research applications. Implementations of multi-touch interaction in the field of visualization are used as examples and case studies for discussion, in part due to the diversity of applications available and the complex and varied analysis requirements of the software. However, other types of applications—for instance, games, security software, web browsers, drawing and modeling applications—have been implemented on multi-touch devices and should also be considered to ensure full ubiquity of results across fields. Also, much multi-touch research was developed commercially and the results are not always available to other researchers, making it difficult to compare with those findings. Finally, this paper is primarily concerned with findings that relate to multi-touch tables, rather than mobile devices, and findings are addressed from this perspective.

6. CONCLUSION
In this paper, we have reviewed current research and presented some design recommendations for five aspects of intuitive multi-touch interaction design. We have analyzed the limited consensus that currently exists regarding intuitive multi-touch interactions among researchers, commercial developers and the general public, as well as presented two open research areas that require more consideration as multi-touch interfaces grow more complex.

To conclude, this is an exciting time to research multi-touch interactions. Just as the best practices of interaction with the mouse and keyboard were investigated with the advent of visual interfaces in the 1980's, so are the best practices of direct manipulation on multi-touch hardware being discovered today and need to be thoroughly investigated. It is our hope that this work will inspire discussion and thoughtful pursuit of intuitive multi-touch interaction among the research community.

7. ACKNOWLEDGEMENTS

8. REFERENCES
[1] Andrews, D. 2010. MTVis: Tree Exploration Using a Multi-Touch Interface. In *Visualization and Data Analysis 2010, Proceedings of the SPIE*, 7530. 2010.

[2] Apple Inc. iPad 2, 2011. Retrieved Nov. 21, 2011, from: http://www.apple.com/ipad/.

[3] Bartoschek, T. and Schwering, A. 2011. Usability testing of the interaction of novices with a multi-touch table in semi public space. In *Proceedings of the 14th international conference on human-computer interaction: interaction techniques and environments* (HCII'11), Vol. Part II. Springer-Verlag, Berlin, Heidelberg, 71-80.

[4] Beaudouin-Lafon, M. 2004. Designing interaction, not interfaces. In *Proceedings of the working conference on advanced visual interfaces* (AVI '04). ACM, New York, NY, USA, 15-22.

[5] Butkiewicz, T. and Ware, C. 2011. Exploratory Analysis of Ocean Flow Models with Stereoscopic Multi-Touch. In *IEEE Visualization Posters*.

[6] Epps, J., Lichman, S., and Wu, M. 2006. A study of hand shape use in tabletop gesture interaction. In *CHI '06 extended abstracts on human factors in computing systems* (CHI EA '06). ACM, New York, NY, USA, 748-753.

[7] Frisch, M., Heydekorn, J., and Dachselt, R. 2009. Investigating multi-touch and pen gestures for diagram editing on interactive surfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (ITS '09). ACM, New York, NY, USA, 149-156.

[8] Fu, C., Goh, W., and Ng, J. 2010. Multi-touch techniques for exploring large-scale 3D astrophysical simulations. In *Proceedings of the 28th international conference on human factors in computing systems* (CHI '10). ACM, New York, NY, USA, 2213-2222.

[9] Hornecker, E. 'I don't understand it either, but it is cool' - visitor interactions with a multi-touch table in a museum. *2008 IEEE International Workshop on Horizontal Interactive Human Computer System (TABLETOP),* IEEE, 113–120.

[10] Intuition. Retrieved March 31, 2012, from Merriam-Webster: http://www.merriam-webster.com/dictionary/intuition.

[11] Isenberg, P. and Fisher, D. 2009. Collaborative brushing and linking for co-located visual analytics of document collections. In *Computer Graphics Forum*, 28, 3, 1031–1038.

[12] Koskinen, H., Laarni, J., and Honkamaa, P. 2008. Hands-on the process control: users preferences and associations on hand movements. In *CHI '08 extended abstracts on human factors in computing systems* (CHI EA '08). ACM, New York, NY, USA, 3063-3068.

[13] Kristensson, P. O., Arnell, O., Björk, A., Dahlbäck, N., Pennerup, J., Prytz, E., Wikman, J., and Åström, N. 2008. InfoTouch: an explorative multi-touch visualization interface for tagged photo collections. In *Proceedings of the 5th Nordic conference on human-computer interaction: building bridges* (NordiCHI '08). ACM, New York, NY, USA, 491-494.

[14] Lam, H. 2008. A Framework of Interaction Costs in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (November 2008), 1149-1156.

[15] Lee, S. and Zhai, S. 2009. The performance of touch screen soft buttons. In *Proceedings of the 27th international conference on Human factors in computing systems* (CHI '09). ACM, New York, NY, USA, 309-318.

[16] Lundström, C., Rydell, T., Forsell, C., Persson, A. and Ynnerman, A. 2011. Multi-Touch Table System for Medical Visualization: Application to Orthopedic Surgery Planning. *IEEE Trans. on Visualization and Computer Graphics* 17, 12 (December 2011), 1775-1784.

[17] Mauney, D., Howarth, J., Wirtanen, A., and Capra, M. 2010. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems* (CHI EA '10). ACM, New York, NY, USA, 4015-4020.

[18] Murata, A. and Iwase, H. 2005. Usability of Touch-Panel Interfaces for Older Adults. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47, 4, 767–776.

[19] Noyes, J. 1983. The QWERTY keyboard: a review. *International Journal of Man-Machine Studies,* 18, 3, 265-281.

[20] Pike, W., Stasko, J., Chang, R., and O'Connell, T. 2009. The science of interaction. *Info. Vis.* 8, 4 (Dec. 2009), 263-274.

[21] Pirker, M., Bernhaupt, R., and Mirlacher, T. 2010. Investigating usability and user experience as possible entry barriers for touch interaction in the living room. In *Proceedings of the 8th international interactive conference on Interactive TV & Video* (EuroITV '10). ACM, New York, NY, USA, 145-154.

[22] Ryall, K., Morris, M., Everitt, K., Forlines, C., and Shen, C. 2006. Experiences with and Observations of Direct-Touch Tabletops. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems* (TABLETOP '06). IEEE, Washington, DC, USA, 89-96.

[23] Shneiderman, B. 1981. Direct manipulation: A step beyond programming languages. In *Proceedings of the joint conference on easier and more productive use of computer systems. (Part II): Human interface and the user interface – Vol. 1981* (CHI '81). ACM, New York, NY, USA.

[24] Stößel, C. 2009. Familiarity as a factor in designing finger gestures for elderly users. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '09). ACM, New York, NY, USA, Article 78, 2 pages.

[25] Tweedie, L. 2007. *Proceedings of the SIGCHI conference on human factors in computing systems*. (CHI '97). New York, New York, USA, 1997, pp. 375–382.

[26] Verification in Visualization: Building a Common Culture. 2011. Retrieved December 6, 2011, from Visweek: http://visweek.org/visweek/2011.

[27] Wigdor, D., Fletcher, J., and Morrison, G. 2009. Designing user interfaces for multi-touch and gesture devices. In *Proceedings of the 27th international conference extended abstracts on human factors in computing systems* (CHI EA '09). ACM, New York, NY, USA, 2755-2758.

[28] Wigdor, D., Williams, S., Cronin, M., Levy, R., White, K., Mazeev, M., and Benko, H. 2009. Ripples: utilizing per-contact visualizations to improve user interaction with touch displays. In *Proceedings of the 22nd annual ACM symposium on user interface software and technology* (UIST '09). ACM, New York, NY, USA, 3-12.

[29] Wobbrock, J., Morris, M., and Wilson, A. 2009. User-defined gestures for surface computing. In *Proceedings of the 27th international conference on human factors in computing systems* (CHI '09). ACM, New York, NY, USA, 1083-1092.

[30] Wroblewski, L. Touch Gesture Reference Guide, 2010. Retrieved Nov. 1, 2011 from: http://www.lukew.com/touch.

[31] Wu, M. and R. Balakrishnan, R. 2003. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on user interface software and technology* (UIST '03). ACM, New York, NY, USA, 193-202.