Cooperative Gestures: Multi-User Gestural Interactions for Co-located Groupware

Meredith Ringel Morris[*], Anqi Huang[**], Andreas Paepcke[*], Terry Winograd[*]
[*]Stanford University, Stanford, CA, USA
[**]Harvard University, Cambridge, MA, USA
merrie@cs.stanford.edu, huang3@fas.harvard.edu, {paepcke, winograd}@cs.stanford.edu

ABSTRACT
Multi-user, touch-sensing input devices create opportunities for the use of *cooperative gestures* – multi-user gestural interactions for single display groupware. Cooperative gestures are interactions where the system interprets the gestures of more than one user as contributing to a single, combined command. Cooperative gestures can be used to enhance users' sense of teamwork, increase awareness of important system events, facilitate reachability and access control on large, shared displays, or add a unique touch to an entertainment-oriented activity. This paper discusses motivating scenarios for the use of cooperative gesturing and describes some initial experiences with CollabDraw, a system for collaborative art and photo manipulation. We identify design issues relevant to cooperative gesturing interfaces, and present a preliminary design framework. We conclude by identifying directions for future research on cooperative gesturing interaction techniques.

Author Keywords
Cooperative gestures, gestures, computer-supported cooperative work, single display groupware.

ACM Classification Keywords
H5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces – computer-supported cooperative work.

INTRODUCTION
We introduce *cooperative gesturing*, a multi-user interaction technique for co-located single display groupware systems. Cooperative gestures are interactions where the system interprets the gestures of more than one user as contributing to a single, combined command. As an example, consider a system that uses a gesture (a wiping motion) to indicate deletion. If all participants simultaneously make the deletion gesture, the screen of their shared display clears entirely rather than having the normal, more local effect of only deleting material in the gesture area. The consensus and involvement of all group members is part of the definition of this sample gesture.

Cooperative gesturing trades off some performance efficiency for the benefits of enhanced collaboration, communication, awareness, and/or fun. These benefits may indirectly improve efficiency by reducing errors or miscommunications, although this possibility is not the focus of this paper. There are several motivating scenarios for the use of cooperative gesturing techniques:

Increase Participation/Collaboration: Interactions that require explicit coordination between two or more users can lead to an increased sense of group cohesion and teamwork. For many CSCW applications requiring collaboration may prove useful, with the caveat that it is important to design these systems well; naively introducing requirements can yield an application that is tedious. Educational activities are one of the most promising domains for collaborative gestures; students, especially younger children, can benefit from requiring increased group participation as a means of reducing "free rider" [16] issues. For special-needs groups, such as youngsters with Asperger's Syndrome (an autistic spectrum disorder defined by social and communicative impairments), an application that explicitly coordinates actions with others can be of therapeutic benefit [23].

Awareness of Important Events: Invocation of potentially destructive application events (e.g., quitting an application, deleting a large amount of content, etc.) can be problematic in groupware systems [21]. Requiring the coordinated effort of all group members, via the use of cooperative gestures, to invoke these important and potentially disruptive actions can help prevent accidental invocation of these commands and can increase group awareness about irreversible program actions.

Reach on Large Surfaces: In most single display groupware systems, the shared display is physically large in order to accommodate a group of users. As a result, some objects on the display are beyond a user's arm's reach. Interactions with distant objects might naturally be accomplished as cooperative gestures, with one user specifying the target and another user specifying the action, so as to avoid the

CHI 2006, April 22–28, 2006, Montréal, Québec, Canada.

need to reach into the personal territory [32] of another user. Previous work, such as the "shuffle" and "throw" gestures on the DynaWall [8], or "drag-and-pop" [2], explore single-user techniques for moving documents and icons across large displays; this paper introduces interaction techniques involving cooperation between members of a co-located group.

Implicit Access Control: Coordination and access control is a tricky issue for shared display groupware systems [21]. Although all digital documents are on a single, shared surface, some may belong to individual members of the group who may wish to restrict certain types of access by their co-workers, such as the ability to edit, copy, or even manipulate an item. Sensitive actions, such as editing a document, can be defined so as to require a cooperative gesture involving both the document's owner and the person who wishes to modify the document; in this manner, access control is implicit whenever the document's owner chooses not to participate in the cooperative gesture.

Entertainment: People engage in coordinated body movements for amusement in many social situations, such as performing "the wave" at a sporting event, or dancing in synchrony to the "YMCA" and "Macarena." Although requiring multiple people to coordinate their actions is not necessarily the most efficient interaction technique, it can lend a sociable and entertaining feel to applications for fun and creativity, such as the creation of unique forms of art that depend upon the collective input of all group members, or other game-like activities.

IMPLEMENTATION: COLLABDRAW

In order to explore the properties of cooperative gesture interaction techniques, we developed CollabDraw, which allows groups of two to four users to collaboratively create diagrams, pictures, collages, and simple animations using free-form drawing and photo collage techniques. A combination of single-user and cooperative gestural interactions controls the CollabDraw workflow.

The CollabDraw software was developed in Java, using the DiamondSpin tabletop interface toolkit [33]. The software is used by a group of two to four users seated around a DiamondTouch table [6]. The DiamondTouch can distinguish the identities of up to four simultaneous touchers by sensing capacitive coupling through special chairs. This property makes it an excellent medium for prototyping and testing cooperative gestures. However, the DiamondTouch does have limitations as a gesture-recognition device, including the coarseness and ambiguity of the input (the table has an array of 172 × 129 antennae spread over a 38'' × 31'' surface).

CollabDraw's gesture recognition uses a combination of machine-learning techniques and heuristic rules. We trained the system to recognize six basic hand postures (a single finger, two fingers, three fingers, a flat palm, a single hand edge, and two hand edges), using 500 examples of each posture from each of four individuals, and regressing on this data using the SoftMax algorithm [20]. This training was sufficient to allow use of the system by new individuals who had not contributed to the training data. CollabDraw uses SoftMax's classification to recognize when one of the learned postures is performed by a user. The program then uses contextual information to determine which gesture is being performed – CollabDraw's six basic postures are used to create sixteen distinct gestural interactions. Examples of context used to further classify an identified posture are whether the hand is moving along a trajectory, whether it is near a photo, or whether another user is performing a gesture at the same time. State information about each user's past touches is maintained to increase the accuracy of these decisions. Some context (such as whether subsets of users are touching one another) is determined by exploiting special properties of the DiamondTouch – for instance, hand-holding by users on different chairs results in the table assuming that the users who sit on all of those chairs are simultaneously touching the same point whenever any one member of this "chain" touches the table's surface.

We implemented a set of cooperative gesture interactions for the CollabDraw application. Our goal in creating this initial application and gesture set was to allow experimentation with this new interaction technique in order to better understand the challenges of designing, implementing, learning, and performing cooperative gestural interactions. This set contains sixteen gestures (five single-user and eleven cooperative gestures), each of which is briefly described in the following sub-sections. The design of these gestures attempted to balance three criteria: (1) using postures and movements based on analogy to "real-world" actions when possible, (2) creating gestures distinct enough to be accurately recognized by our system given the limitations of the DiamondTouch as a recognition device, and (3) including gestures that involved several styles of cooperation (see the "Design Space" section for more discussion of this last issue).

Stroke Creation and Modification

Users can draw strokes of colored ink onto the canvas area of CollabDraw by moving the tip of a single finger on "canvas" areas of the screen (e.g., areas not occupied by photos). While drawing itself is a single-user action, the ability to modify the nature of the drawn ink is provided via a cooperative gesture. If user A places two fingers on the surface of the table while user B is drawing strokes, the width of B's stroke changes based on the distance between A's two fingers (Figure 1b). Similarly, the pressure that A applies to the surface of the table while performing this stroke-modification gesture impacts the darkness or lightness of the color drawn by B. In the event of larger groups of users (more than 2 people), the target of a stroke-modification gesture can be disambiguated by using the "partner" gesture (Figure 1a) – two users hold hands and touch the table, establishing a partnership between them. Partnerships determine which group member's strokes are

(a)

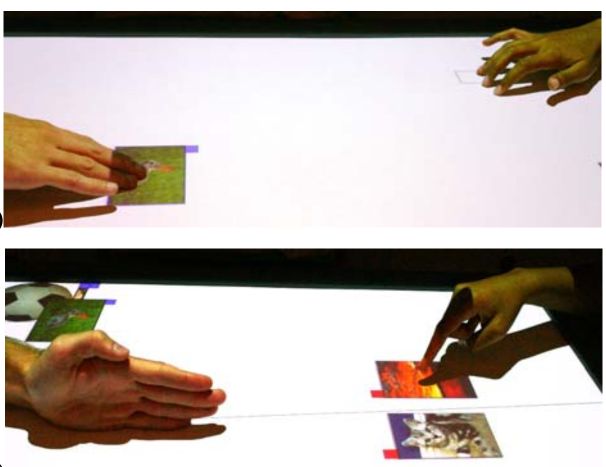
(b)

Figure 1. (a) Two users form a partnership. (b) Now, when one partner performs the “modify ink” gesture the ink drawn by her partner changes thickness.

modified by which group member’s modification gestures. Partnerships can be broken by performing the “partner” gesture a second time.

Stroke Erasure

By placing the palm of one’s hand on the surface of the table and rubbing it back and forth, a user can erase ink from the canvas. The ink immediately underneath his hand disappears. This single-user gesture has a cooperative form as well – the “clear screen” gesture. When all members of the group simultaneously perform the “erase” motion, the effect is magnified and all ink on the entire table is instantly cleared (Figure 2).



(a)

(b)

Figure 3. (a) The “throw-and-receive” gesture is one technique for passing photos long distances. (b) The “pull” action (where a partner disambiguates the target) is another option for moving photos long distances.

Figure 2. When all group members simultaneously “erase,” the “clear screen” action is invoked.

Note that for a gesture such as “clear screen” that requires the participation of *all* members of a group, the total number of group members could be determined using a variety of techniques such as pressure sensors on chairs, computer vision, or heuristics based on interaction histories during the current session. For our initial implementation of CollabDraw, the total group size was manually entered upon session start-up.

Photo Passing

The ability to manipulate digital photos as part of an artistic creation is part of the CollabDraw software. Individual users can move digital photos around the table by touching them with a single finger and dragging them to a new location. To pass photos over large distances, two cooperative gestures are available – throwing and pulling.

To throw a photo across the table, user A touches the photo with 3 fingers and makes a throwing motion while user B taps an empty location on the table with 3 fingers (Figure 3a). So long as the trajectory specified by user A’s motion is roughly aimed toward the location specified by user B, the photo will move across the table with a velocity influenced by the speed of user A’s gesture, and will snap to the endpoint specified by user B. Enhancements to this gesture could allow the receiving partner’s action to specify additional parameters, such as the orientation the thrown image should face when it arrives.

Alternatively, if user A wishes to move a photo from the far end of the table toward himself, he can place the edge of his hand on the table, aimed in a line toward the target image, and can move his hand toward himself, dragging it along

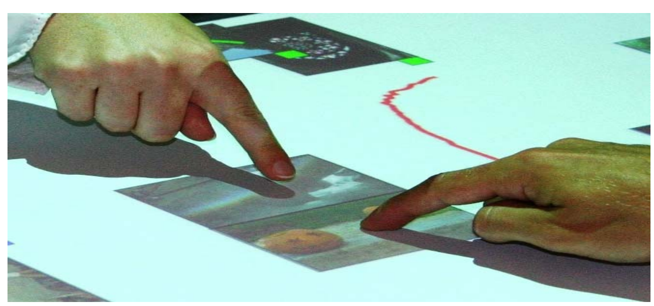


Figure 4. Two users create a collage using the “combine” gesture.

Figure 5. Four group members perform the "enlarge" gesture by simultaneously touching the corners of a single photo.

the table's surface. This action causes candidate photos along this trajectory to blink, indicating that they are potential targets of this "pull" gesture (Figure 3b). User B, who is seated near these target photos (and who may "own" them), can disambiguate A's choice (and/or grant permission for A to take a photo that B owns) by tapping one of the blinking photos with a single finger. This image then slides across the table to user A.

Combine Photos

Users may combine multiple photos to form a panorama or collage (Figure 4). To perform this action, two users each move digital photos towards each other (by dragging them with a single finger) and collide their images. When the images collide, they fuse together along the intersecting boundary, forming a single, larger image.

Enlarge Photos

Users can cooperate to increase the size of a photo to make it occupy the background of the table. To accomplish this, every member of the group must participate by touching near the corner of the target image and moving their fingers outward toward the edge of the corner (Figure 5).

Neaten Photos

Photos can be neatened into orderly piles by placing the edges of both hands on the surface of the table, and sweeping them toward each other. This causes all photos between the two hands to move into a single pile. This gesture can also be performed cooperatively, with a magnified effect: when all users simultaneously perform the "neaten" motion, the "organize table" action is invoked

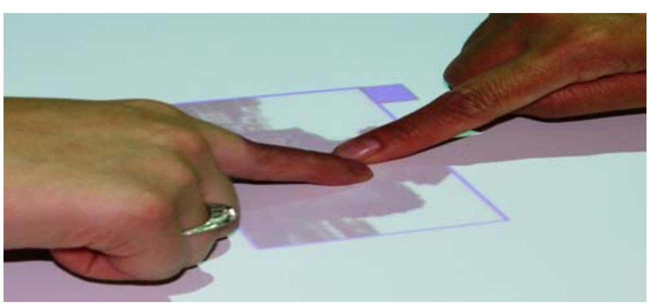


Figure 7. Two users simultaneously touch the center of a photo to transfer ownership.

Figure 6. When all four group members simultaneously perform the "neaten" gesture, the result is to organize all of the photos on the table into a single, central pile.

(Figure 6), and all photos on the entire table, regardless of whether they are between anyone's hands, move together into one single pile in the table's center, thus instantly organizing the entire work surface.

Photo Ownership and Annotation

In addition to using finger-ink to mark up CollabDraw's canvas area, individual photos can also have ink annotations added to them. These annotations remain on the photos as the photos are moved about the surface of the table. To differentiate between touching a photo to move it about the surface of the table versus touching it to draw finger-ink, a user can cover the photo with his palm. This causes a white indicator to appear above the photo, as feedback that the photo is now in annotation mode. Subsequent single-finger strokes on the image result in annotation. Covering the photo with the palm once again returns the photo to draggable mode.

Photos in CollabDraw have a notion of ownership associated with them, to allow us to explore issues related to access control that are relevant in many CSCW applications. Ownership of an image is indicated by a small colored tab above each photo. The color of this tab matches the color of the chair of the user who owns that photo. Users are only able to annotate a photo that they own. Ownership of photos can be transferred between users by performing the cooperative "exchange photo" gesture, where the two participants must simultaneously touch the center of the photo in question (and one of the participants must be the photo's owner) (Figure 7).

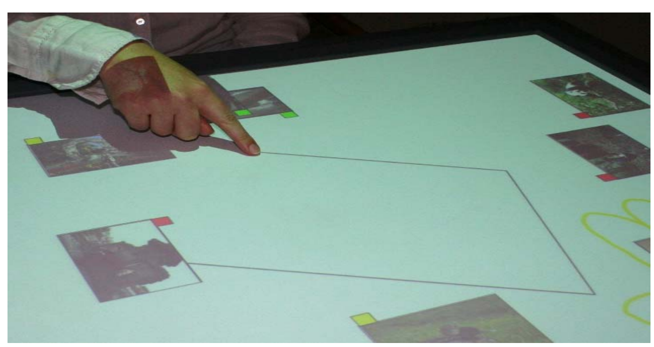


Figure 8. Group members take turns defining waypoints to animate a photo.

Photo Animation

Users can enhance their drawings with simple animations. They can cooperatively define a trajectory to be followed by target photographs. To initiate trajectory definition, a user holds the edge of her hand over an image until it begins to flash. Now, group members take turns tapping points on the table with a single finger. Each point adds to the image's trajectory, which is temporarily illustrated with black lines (Figure 8). To exit trajectory-definition mode, one user again covers the target image with her hand's edge. Now, to begin the animation, a user can mimic the "throw" gesture, pushing the target image with 3 fingers, and it will animate along the pre-defined path.

Exiting CollabDraw

Exiting CollabDraw requires the consent of all group members. To accomplish this, they must all hold hands, and then one member of the "chain" touches the table's surface with a single finger (Figure 9). This causes a menu to appear that allows the group to confirm their choice to exit the application.

EVALUATION

Fourteen paid subjects participated in a usability study to evaluate the use of cooperative gestures in CollabDraw. Six of the subjects were female, and the mean age was 25.5 years. Nine of the subjects had never used a DiamondTouch table before. Subjects completed the study in pairs of two, although CollabDraw can accommodate as many as four users. All subjects were acquainted with their partner before the study; subjects had known their partners for 2.2 years on average. Three pairs were of romantically-involved couples, while four pairs were same-sex pairs of co-workers who were not romantically involved.

The goal of this evaluation was to gauge basic aspects of the usability of cooperative gestures – would people find them intuitive or confusing? Fun or tedious? Easy or difficult to learn? The evaluation had four parts, which were all completed within a single one-hour session: (1) gesture training, (2) a gesture-performance quiz, (3) recreating a target drawing, and (4) completing a questionnaire. First, the experimenter introduced the CollabDraw application and taught each of the gestures (both single-user and

Figure 9. Group members form a chain of hands, and one user touches the table in order to exit CollabDraw.

cooperative gestures) to the participants. Because groups of size two were used, the "partner" gesture was superfluous, and was therefore not part of the evaluation. Participants could practice each gesture as many times as they wished, and could ask questions to and receive advice from the experimenter. After participants had been taught all the gestures and practiced as much as they wanted, the experimenter quizzed the subjects by naming a gesture and asking them to perform that gesture without any advice.

After the performance quiz, the subjects were provided with printouts of a target drawing, and were asked to recreate the drawing using CollabDraw without any assistance from the experimenter. The nature of the drawing required the use of several gestures (draw, annotate photos, exchange photos, modify ink, combine photos, enlarge, and animate). After completing the drawing, pairs were asked to organize the table, clear the screen of ink, and exit the application. Subjects then filled out a questionnaire asking them to rate each of the gestures along several dimensions and soliciting free-form comments. All reported ratings use a 7-point Likert scale, with a rating of 7 being positive and 1 being negative. The experimenter took notes during the sessions, and the CollabDraw software logged all user interactions with the DiamondTouch table.

Results

Overall, subjects found CollabDraw easy to use and the gestures easy to learn. Subjects took 28.8 minutes on average (stdev = 6.2 minutes) to learn all 15 gestures, and all seven pairs were able to accurately re-create the target drawing with a mean time of 8.2 minutes (stdev = 1.2 minutes). In addition, subjects made very few errors during the "quiz" portion of the session – three subjects forgot the gesture for "exchange photos," but were reminded by their partners, one subject forgot how to initiate animation and was also reminded by his partner, and one pair forgot how to clear the screen and had to be reminded by the experimenter. These results indicate that our gesture set was relatively easy for subjects to learn, remember, and use.

Overall, users found neither the single-user nor cooperative gestures confusing, as indicated by their Likert scale responses to the statements "I found the [single-user/cooperative] gestures confusing to perform" (μ=2.5 and 3.2, respectively – these ratings are not statistically different from each other, p=.20, $t(13)$=1.35). In the following sub-sections, we describe results based on observations of use and users' questionnaire ratings of the ten cooperative gestures in the CollabDraw repertoire. Although the majority of user comments were positive, in the following sections we particularly highlight some of the negative reactions since such comments are informative for improving cooperative gesture interactions.

Modify Ink

The "modify ink" gesture received poor ratings on the intuitive (μ=3.69) and fun (μ=3.21) scales, and was named by eight participants as one of their least favorite gestures. Subjects indicated that they found it confusing to need assistance to change the width of their stroke. They found the collaboration for that task to be artificial, indicated by comments such as "There's nothing inherently cooperative about ink-width changing," and "It would make more sense to modify your own ink." Users noted that the use of a cooperative gesture "seemed appropriate when the result of the gesture affected both parties involved," a rule that did not apply to "modify ink." Further, they also felt it was quite tedious and inefficient to need to interrupt their partner to ask for an ink modification since this was a task they performed frequently – as one user noted, "[my] partner had to stop what she was doing so that I could change a property of *my* picture." Performing this gesture sometimes caused unanticipated mode errors, because one partner would interrupt another to ask for an ink modification, causing his partner to forget what gesture she had been in the midst of performing, which was particularly problematic if she had been in the midst of performing a moded gesture such as photo annotation. To minimize the need to modify ink, all seven groups approached the final drawing task in a manner that required the minimum possible number of ink modifications.

Clear Screen

The "clear screen" gesture was rated as intuitive (μ=5.57), but not quite as intuitive as the corresponding single-user gesture, "erase" (μ=6.69) (p<.01, $t(13)$=4.76). "Clear screen" also received high ratings for being fun (μ=5.71). The "clear screen" gesture received mixed reactions from participants – two subjects listed it among their favorite gestures, while two subjects ranked it among their least favorite. These latter two cited the risk of accidental invocation when two people coincidentally simultaneously performed the "erase" motion. Users commented, "We had to be careful not to unintentionally affect the whole canvas when we were performing these actions." Another noted that "clear screen" was, "...too easy! I had to watch out for [my] partner erasing at the same time." This accidental invocation occurred during two of the seven test sessions.

Throw-and-Receive

The "throw-and-receive" gesture received a neutral rating on the fun scale (μ=4.5), despite the fact that during training users frequently commented that throwing photos was "cool." Five of the seven groups spontaneously used the throw gesture during unrelated portions of the training session, presumably because they found it entertaining. However, subjects commented that the throw gesture didn't seem necessary, given the small size of the DiamondTouch table (all subjects could reach the table's far end). One user commented "I'm dubious about why someone would need it [throw-and-receive] when they could just reach across the table." This apparent lack of utility might account for the low ratings – it would be interesting to see how reactions would change if larger table sizes were available.

Pull

The "pull" gesture was voted least favorite by ten users, and received correspondingly low reviews for intuitiveness (μ=3.0), fun (μ=3.43), and comfort (μ=3.31). In addition to pointing out that the small size of the table made the pull gesture unnecessary, users also indicated that they found the specific posture involved (the use of the side of the hand) to be awkward and unnatural, commenting "In general, the edge-of-my-hand gesture is unintuitive."

Combine Photos

The "combine" gesture received generally good ratings ($\mu_{intuitive}$=5.86, μ_{fun}=5.14, $\mu_{comfortable}$=5.69), and was the source of little comment from or difficulty to users.

Exchange Photo

The "exchange photo" gesture received generally good ratings ($\mu_{intuitive}$=5.21, μ_{fun}=4.93, $\mu_{comfortable}$=5.69), although its similarity to the "enlarge photo" gesture was slightly problematic. Three subjects had to be reminded by their partners how to perform this action during the quiz. This confusion may be particular to groups of only two users, since two users are required to exchange a photo but the entire group is required to enlarge a photo. Nevertheless, users felt that the cooperative nature of this action was well-justified, as indicated by comments like, "exchange photo made sense [to make both people do the gesture]."

Organize Table

Reaction to the "organize table" gesture was similar to the response to "clear screen," the other gesture with both a single-user and whole-group interpretation. Users rated the gesture highly as being intuitive (μ=6.0) and fun (μ=6.14), but it also received a mixed response with two votes for favorite and two for least favorite gesture, with the risk of accidental invocation again being noted by its detractors.

Animate Photo

The animate gesture was named least favorite by seven users, and received correspondingly low fun (μ=4.21), comfort (μ=3.71), and intuitiveness (μ=3.86) ratings. While subjects commented that defining the actual trajectory of the animation was intuitive, they found the use of the edge of the hand to initiate and terminate this trajectory-definition phase to be unnatural. The cooperative nature of the animate gesture caused unanticipated mode errors because sometimes one user initiated it without informing their partner. Initiating this gesture put both partners in trajectory-definition mode, so if one user was unaware of the mode-switch, confusion occurred.

Enlarge Photo

Users rated this gesture as fun (μ=5.0), and had little comment on it and little difficulty in its execution, other than the aforementioned similarity between it and the "exchange" gesture.

Exit Application

The “exit” gesture received mixed reactions. Not surprisingly, couples that were romantically involved showed no reaction to the request to hold hands, but pairs of friendly co-workers found the request more unusual. One female-female co-worker pair thought the gesture was cute, smiling and saying “awwww...” when asked to hold hands, but all three male-male groups giggled or made nervous jokes. One user commented about the “exit” gesture that it was unpleasant because, “[my] partner has sweaty hands,” and another user noted, “touching was awkward.” One member of a dating couple noted, “I liked holding hands because I knew my partner, but in a work environment I would find that much more awkward,” indicating that not only how well one knew one’s partner, but also the nature of the activity would impact the acceptability of intimate cooperative gestures.

During the initial training, one male-male pair asked how the “exit” gesture worked, and the experimenter explained that by holding hands the DiamondTouch thought that both of their identities were touching the table at a single point, thereby initiating the “exit” gesture. This pair then attempted to avoid the need to hold hands during the quiz and drawing by touching their fingers very close together at one point on the table.

DISCUSSION

Lessons Learned

User feedback and observations from our evaluation of CollabDraw provided useful points to keep in mind for future iterations of cooperative gestures:

Clarity of Purpose: Users reacted most positively to cooperative gestures that served a clear purpose, commenting that they understood why actions such as exchanging photo ownership, clearing the screen, and exiting the application should require multiple users, but complaining about “unnecessary” collaboration for more mundane actions such as ink modification. We had originally envisioned this latter, “non-necessary” cooperative gesture as a possible source of amusement and creativity, but it was not viewed in this manner by users.

Accidental Invocation: Some of our cooperative gestures, such as “clear screen” and “organize table,” were based on simultaneous performance by all group members of an action that also had a valid single-user interpretation. While users indicated that these gestures were fun and intuitive, there were occasional accidental invocations of the cooperative actions when both members of a pair coincidentally simultaneously tried to perform the corresponding single-user gesture. For larger groups, accidental invocation is likely to be less frequent. Nonetheless, relying on the very small probability of accidental simultaneous action is non-optimal; interactions that avoid or mitigate this issue are desirable.

Tedium: Users complained that the “modify ink” gesture, in addition to not having a clear purpose for collaboration, was also particularly tedious because it was an action that they wanted to perform frequently, thus requiring frequent interruptions of their partner to ask for assistance. Because of the coordination overhead, cooperative gestures are probably not appropriate for frequently-used actions; rather, it may be more appropriate to add only a few cooperative gesture actions to a system, reserved for special commands that require high awareness or group consent.

Intimacy: Not surprisingly, highly intimate cooperative actions, such as the “exit” gesture that required hand-holding, were not well-received by pairs of co-workers. Even partners who were romantically involved pointed out that if the application had a business, rather than entertainment, feel to it, they might have felt awkward holding hands as well. However, gestures that required close proximity without actual skin contact, such as the “exchange photo” gesture where two users simultaneously touched near the center of a single photo, did not provoke any objections. Gestures that require skin contact might be appropriate for certain types of entertainment applications that are used among friends, but would clearly not be acceptable for more formal environments and purposes.

Subversion: We were surprised to see one subject intentionally abuse the cooperative nature of the modify ink gesture in order to ruin his partner’s drawing. This same subject also attempted to steal ownership of his partner’s photo by attempting to touch near the center of that photo at a moment when his partner happened to also be touching it, thus performing the “exchange” gesture without his partner’s conscious consent. Techniques to prevent this type of subversion are an avenue worth exploring.

Design Space

Based on our experiences designing, implementing, and evaluating an initial set of cooperative gestures, we have articulated some important axes of the design space for these interactions. By articulating this taxonomy, we hope to better understand the design possibilities for cooperative gestures, and to analyze the impact of these axes of design on their learnability, memorability, usability, and naturalness. We have excluded from our taxonomy design axes that are not unique to cooperative gestures – issues such as “naturalness” (does the gesture mimic real-world activity, or is it abstract), whether each user contributes a unimanual or bimanual action, etc. These issues are relevant to single-user gestures as well, and, while they could certainly have an impact on cooperative gesture performance, are not the focus of this paper. Based on our initial experiences with this interaction technique, we have identified seven design axes relevant to cooperative gesture interaction: symmetry, parallelism, proxemic distance, additivity, identity-awareness, number of users, and number of devices. Table 1 classifies CollabDraw’s cooperative gestures along these dimensions.

Symmetry

The "symmetry" axis refers to whether participants in a cooperative gesture perform identical actions ("symmetric") or distinct actions ("asymmetric"). In a gesture involving more than two users, it is also possibly to have a subset of users performing identical actions and another subset performing distinct actions ("partially symmetric"). Note that this differs from the use of the term "symmetry" as applied to conventional, single-user gestures, where symmetry refers to whether the two hands in a bimanual gesture perform identical actions [11].

Parallelism

"Parallelism" defines the relative timing of each contributor's actions. If all users perform their gesture simultaneously, then the collective gesture is "parallel," and if each user's gesture immediately follows another's (and yet the entire sequence accomplishes nothing unless everyone finishes their action), then it is "serial." "Partially parallel" is also possible for gestures involving more than two users, where some users perform their parts at the same time and some perform them in sequence. The level of parallelism in a cooperative gesture may impact the ability of users to conceptualize their combined actions as a single "phrase" [4] or unit.

Proxemic Distance

Proxemics [12] is the study of the distances people prefer to maintain between each other in various situations. The level of physical intimacy required to perform a cooperative gesture could impact its acceptability for different application scenarios (e.g., fun vs. business) or depending on the personal relationships among group members. For that reason, we feel that proximity is an important design consideration. We have adapted the definitions of the four canonical proxemic distances for a co-located groupware situation. "Intimate" refers to cooperative gestures in which participants must physically touch other participants. "Personal" refers to gestures in which participants must touch the same digital object (e.g., both users must touch the same image, window, text document, etc.) but their hands do not touch each other. "Social" refers to gestures in which participants must touch the same display device but can touch distant parts of the device (e.g., both users must touch the table, but each touches in the space closest to where he is seated). Lastly, "public" refers to gestures where users do not need to touch the same display (e.g., the shared display is supplemented with PDAs, and users perform their coordinated actions on these devices).

	Symmetric	Parallel	Proxemic Distance	Additive	Identity-Aware	# of Users
Partner	Y	Y	Intimate	N	N	2
Modify ink	N	Y	Social	N	N	2
Clear screen	Y	Y	Social	Y	N	All
Throw	N	Y	Social	N	N	2
Pull	N	Y	Social	N	N	2
Combine	Y	Y	Personal	N	N	2
Enlarge	Y	Y	Personal	N	N	All
Organize table	Y	Y	Social	Y	N	All
Exchange	Y	Y	Personal	N	Y	2
Animate	Y	N	Social	N	N	All
Exit	Y	Y	Intimate	N	N	All

Table 1. CollabDraw's cooperative gestures, classified.

Additivity

"Additivity" refers to a special class of symmetric, parallel gestures. An "additive" gesture is one which is meaningful when performed by a single user, but whose meaning is amplified when simultaneously performed by all members of the group. For example, in CollabDraw rubbing one's palm on the table in a back-and-forth motion erases digital ink directly under the palm. The "clear screen" action is an additive version of this gesture, invoked when all group members perform the "erase" motion simultaneously. Symmetric, parallel gestures that do not have less-powerful individual interpretations are "non-additive."

Identity-Awareness

Cooperative gestures can be "identity-aware," requiring that certain components of the action be performed by specific group members. For example, gestures whose impact is to transfer access privileges for an item from one user to another would require that the user who performs the permission-giving part of the gesture be the user who actually "owns" the object in question. Gestures with no role- or identity-specificity are "non-identity-aware."

Number of Users & Number of Devices

Cooperative gestures involve two or more users whose coordinated actions are interpreted as contributing to a single gestural interaction. The precise number of users involved is an important dimension to consider, as it could impact the complexity involved in learning and executing the gesture. The number of devices involved is also a consideration – whether users all perform their gesture on a single, shared display, or whether personal devices are involved as well. The use of a single, shared display might simplify gesture learning by increasing the visibility of group members' actions – we observed bootstrapping of this type during our evaluation of CollabDraw.

Future Work

Our initial exploration of cooperative gestures was promising. Users learned the gestures quickly, found many of them intuitive and entertaining, and provided valuable feedback on how to further improve this interaction technique. There are several interesting avenues for further research. Evaluation with larger group sizes would be informative, in order to learn whether the complexity of coordinating actions with more group members makes it more difficult to learn and execute cooperative gestures. Exploring the use of cooperative gestures in other application contexts and with other gesture sets would also be informative, since one challenge in evaluating these gestures is in determining whether our results are applicable to cooperative gestures in general or are specific to peculiarities of our particular implementation. Additionally, it would be particularly interesting to explore a set of cooperative gestures that covers combinations of axes in the design space that were not addressed by CollabDraw in order to get a better understanding of how those axes impact the usability of cooperative gestures.

RELATED WORK

Conventional, single-user gestural interactions have been explored in a variety of systems, including [1, 4, 9, 10, 17, 18, 28, 37, 39]. Our work expands upon conventional gestural interactions by exploring *cooperative* gestures, where the system interprets the input of multiple users as contributing to a single gestural command. Cooperative gestures are particularly relevant to co-located groupware systems, such as single display groupware [34]. Tabletop interfaces, in particular, are a compelling platform for this interaction technique because of the high degree of parallel activity they promote compared to shared, vertical displays [29], and because of the availability of hardware like the DiamondTouch [6], which can handle multiple simultaneous touch inputs, and can associate each input with one of four distinct user identities. Several systems [7, 22, 24, 38, 40, 41] explore gesture interactions with tabletops, but none of these systems interpret the interactions of multiple users together as a single, cooperative gesture.

Examples of cooperative gestures can be found in several prior systems. In this prior work, however, cooperative gesturing was not the focus of study, and there is no discussion of the phenomenon of cooperative gesturing per se, nor is there exploration of the design of cooperative gestures. In contrast, the focus of this paper is to define, analyze, and evaluate cooperative gesturing as an interaction technique. Nonetheless, these prior systems provide interesting examples of the use of isolated cooperative gestures:

- Several performance art and entertainment systems use a rough interpretation of a large group's motions to produce an entertaining effect [19, 26, 36] (e.g., changing the tempo of music, changing the direction of an object that moves on a large screen, etc.).
- SyncTap [25], Smart-Its Friends [15], Synchronous Gestures [13], and Stitching [14] use coordinated motion patterns (as measured by accelerometers on mobile devices) to establish ad-hoc network connections (e.g., shaking two devices at the same time or tapping them together would create a connection).
- A few systems utilize a cooperative action for facilitating participation and socialization for special user groups. The StoryTable [5] (a tabletop story-creation system for very young children) requires two children to touch certain objects to enable actions such as story playback. SIDES [23] (a cooperative, therapeutic game for adolescents with Asperger's Sydrome) requires all players to "vote" on key game actions. The ToneTable and Well of Inventions [35] are museum exhibits that allow several museum-goers to simultaneously interact with a physical simulation by manipulating trackballs; some simulation effects are possible only if multiple users interact with the display.
- Proposed interactions for voting and permission-giving in some groupware systems include examples that demonstrate cooperative gesturing. "Multi-user coordination policies" [21] include interactions for group voting. The "release" interaction for sharing digital documents on a tabletop display [27] demonstrates an access-control gesture involving two participants. iDwidgets [30] proposes identity-aware widgets for co-located groupware, and describes some potential identity-aware widgets (such as widgets requiring simultaneous input from multiple users) that could be loosely construed as supporting cooperative gesture input.

Philosophical literature on "collective intentionality" [3, 31] postulates that collective intentional behavior is a separate phenomenon from the sum of individual intentions. This philosophical underpinning provides an interesting perspective from which to appreciate the role of cooperative gesturing in an interactive system.

CONCLUSION

We have formalized the concept of cooperative gestures for co-located groupware as interactions where the system interprets the gestures of multiple group members collectively in order to invoke a single command. Judicious use of cooperative gestures can add value to applications as a means of increasing participation, drawing attention to important commands, enforcing implicit access control, facilitating reach on large surfaces, and/or enhancing social aspects of an interactive experience. We presented an initial implementation of eleven cooperative gestures in the context of CollabDraw, a tabletop art application for two to four users. Based on our evaluation of CollabDraw, we identified several issues relating to the acceptability and usability of cooperative gestures. This experience enabled us to articulate axes of a design space for this interaction technique; these axes provide a framework for future study of cooperative gesture interfaces.

ACKNOWLEDGMENTS

We thank MERL for donating a DiamondTouch table, and Heather Gentner, Ada Glucksman, Gretchen Lantz, and Dan Morris for giving permission to use video and photos of their gesturing.

REFERENCES

1. Baudel, T. and Beaudoin-Lafon, M. Charade: Remote Control of Objects Using Free-Hand Gestures. *Communications of the ACM*. 36(7), 28-35.
2. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. Drag-and-Pop and Drag-and-Pick: Techniques for Accessing Remote Screen Content on Touch- and Pen-operated Systems. *Interact 2003*.
3. Bratman, M. 1999. Faces of Intention. Cambridge, MA: Cambridge University Press.
4. Buxton, W. Chunking and Phrasing and the Design of Human-Computer Dialogues. *IFIP Conference*, 1986, 475-480.
5. Cappelletti, A., Gelmini, G., Pianesi, F., Rossi, F., and Zancanaro, M. Enforcing Cooperative Storytelling: First Studies. *ICALT 2004*.
6. Dietz, P. and Leigh, D. DiamondTouch: A Multi-User Touch Technology. *UIST 2001*, 219-226.
7. Everitt, K., Shen, C., Ryall, K., and Forlines, F. Modal Spaces: Spatial Multiplexing to Mediate Direct-Touch Input on Large Displays. *CHI 2005 Extended Abstracts*, 1359-1362.
8. GeiBler, J. Shuffle, throw, or take it! Working Efficiently with an Interactive Wall. *CHI 1998 Extended Abstracts*, 265 -266.
9. Grossman, T., Wigdor, D., and Balakrishnan, R. Multi-Finger Gestural Interaction with 3D Volumetric Displays. *UIST 2004*.
10. Guimbretiere, F., Stone, M., and Winograd, T. Fluid Interaction with High-Resolution Wall-Size Displays. *UIST 2001*, 21-30.
11. Guiard, Y. Asymmetric division of labor in human skilled bimanual action: the kinematic chain as a model. *Journal of Motor Behaviour*. 19 (4), 1987, 486-517.
12. Hall, E.T. The Hidden Dimension. New York: Anchor Books, 1966.
13. Hinckley, K. Synchronous Gestures for Multiple Persons and Computers. *UIST 2003*, 149-158.
14. Hinckley, K., Ramos, G., Guimbretiere, F., Baudisch, P., and Smith, M. Stitching: Pen Gestures that Span Multiple Displays. *AVI 2004*, 23-31.
15. Holmquist, L., Mattern, F., Schiele, B., Alahuhta, P., Beigl, M., and Gellersen, H. Smart-Its Friends: A Technique for Users to Easily Establish Connections between Smart Artefacts. *UbiComp 2001*, 116-122.
16. Joyce, W.B. On the Free-Rider Problem in Cooperative Learning. *Journal of Educational Business*. 74(5), 271-274, May-June 1999.
17. Krueger, M.W., Gionfriddo, T., and Hinrichsen, K. VIDEOPLACE - An Artificial Reality. *CHI 1985*, 35-40.
18. Malik, S., Ranjan, A., Balakrishnan, R. Interacting with Large Displays from a Distance with Vision-Tracked Multi-Finger Gestural Input. *UIST 2005*, 43-52.
19. Maynes-Aminzade, D., Pausch, R., and Seitz, S. Techniques for Interactive Audience Participation. *ICMI 2002*, 15 - 20.
20. McCullagh, P. and Nelder, J.A. Generalized Linear Models. Chapman and Hall, 2nd edition, 1989.
21. Morris, M.R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Beyond "Social Protocols": Multi-User Coordination Polices for Co-located Groupware. *CSCW 2004*, 262-265.
22. Oka, K., Sato, Y., and Koike, H. Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems. *IEEE International Conference on Automatic Face and Gesture Recognition*, 2002, 429-434.
23. Piper, A.M., O'Brien, E., Morris, M.R., and Winograd, T. SIDES: A Collaborative Tabletop Computer Game for Social Skills Development. *Stanford Univ. Technical Report*, 2005.
24. Rekimoto, J. SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. *CHI 2002*, 113-120.
25. Rekimoto, J. SyncTap: Synchronous User Operation for Spontaneous Network Connection. *Personal and Ubiquitous Computing*, 8(2), May 2004, 126-134.
26. Reynolds, M., Schoner, B., Richards, J., Dobson, K., and Gershenfeld, N. An Immersive, Multi-User, Musical Stage Environment. *SIGGRAPH 2001*, 553-560.
27. Ringel, M., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-User Interactive Tables. *CHI 2004 Extended Abstracts*, 1441-1444.
28. Ringel, M., Berg, H., Jin, Y. and Winograd, T. Barehands: Implement-Free Interaction with a Wall-Mounted Display. *CHI 2001 Extended Abstracts*, 367-368.
29. Rogers, Y. and Lindley, S. Collaborating Around Large Interactive Displays: Which Way is Best to Meet? *Interacting with Computers*, 2004.
30. Ryall, K., Esenther, A., Everitt, K., Forlines, C., Morris, M.R., Shen, C., Shipman, S. and Vernier, F. iDwidgets: Parameterizing Widgets by User Identity. *Interact 2005*.
31. Searle, J.R. Collective intentions and actions. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, chapter 19, pages 401-416. MIT Press, Camridge, MA, 1990.
32. Scott, S.D., Carpendale, M.S.T., and Inkpen, K. Territoriality in Collaborative Tabletop Workspaces. *CSCW 2004*, 294-303.
33. Shen, C., Vernier, F., Forlines, C., and Ringel, M. DiamondSpin: An Extensible Toolkit for Around-the-Table Interaction. *CHI 2004*, 167-174.
34. Stewart, J., Bederson, B., and Druin, A. Single Display Groupware: A Model for Co-present Collaboration. *CHI 1999*.
35. Taxén, G., Hellström, S.-O., Tobiasson, H., Back, M., and Bowers, J. The Well of Inventions – Learning, Interaction and Participatory Design in Museum Installations. *ICHIM 2003*.
36. Ulyate, R. and Bianciardi, D. The Interactive Dance Club: Avoiding Chaos in a Multi Participant Environment. *CHI 2001 Workshop on New Interfaces for Musical Expression*.
37. Vogel, D. and Balakrishnan, R. Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users. *UIST 2004*.
38. Wellner, P. Interacting with Paper on the DigitalDesk. *Communications of the ACM*, 36(7), 1993, 87-96.
39. Wolf, C. and Rhyne, J. Gesturing with Shared Drawing Tools. *CHI 1993 Extended Abstracts*, 137-138.
40. Wu, M. and Balakrishnan, R. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. *UIST 2003*, 193-202.
41. Wu, M., Shen, C., Ryall, K., Forlines, C., and Balakrishnan, R. Gesture Registration, Relaxation, and Reuse for Multi-Point Direct-Touch Surfaces. *IEEE Tabletop 2006*, 183-190.