Haptic Support for Urgency-Based Turn-Taking

Technical Report TR-2004-14
Andrew Chan, Joanna McGrenere, Karon MacLean
Department of Computer Science, University of British Columbia
chana@cs.ubc.ca, joanna@cs.ubc.ca, maclean@cs.ubc.ca

ABSTRACT

Real-time collaboration systems that enable distributed access to a shared application often require a turn-taking protocol. Current protocols rely on the visual channel using GUI widgets, and do not support expressions of urgency. We describe a novel urgency-based turn-taking protocol that is mediated through haptics: vibrotactile signals inform users of their current role in the collaboration. For example, a control holder receives different signals according to the urgency with which collaborators request control. In an observational user study we compare three implementations of the protocol: one dominated by haptic signals, one with visual cues alone, and one balancing both modalities. Our results suggest that a modestly-sized set of well-designed haptic stimuli can be learned to a high degree of accuracy in a short time, that the inclusion of haptic stimuli can make turn-taking behavior more equitable, and that the ability to communicate urgency positively impacts collaboration dynamics.

Author Keywords

Haptic interaction, haptic icons, turn-taking, floor control, real-time collaboration, user study

ACM Classification Keywords

H.5.2 User Interfaces: Haptic I/O, Graphical user interfaces; H.5.3 Group and Organization Interfaces: Computer-supported cooperative work, synchronous interaction

INTRODUCTION

Real-time collaboration systems that enable distributed individuals to work together over the Internet offer an alternative to business travel, which can be time consuming, expensive, and exhausting. Users can share a locally running application with others who are at a distance, so that collaborators can view and interact with the application through their own computers. Many of these collaborative systems, such as Microsoft NetMeeting [21], are designed to be lightweight in that they work with existing single-user applications, and do not require the use of specialized hardware or networking infrastructure.

Typically, all participating users can view the shared application but only one can interact with it at a time, necessitating a system for taking turns. To date, these protocols invariably rely on visual feedback to convey information such as a request for control. Since computer-mediated collaboration is often oriented visually (e.g., authoring a slide presentation), visually presented requests may be distracting if too intrusive, or missed if too subtle. Current systems also lack support for users to express how urgently they wish to gain control, a need which is usually met in face-to-face collaboration.

Haptic technology offers another means of mediating collaborative turn-taking. Vibrotactile displays are already being used to transmit abstract information in other consumer applications: e.g. cell phone manufacturers like Nokia are experimenting in Japan with distinctive “vibe tones”, hoping they can transmit meanings less intrusively than their auditory counterparts; auto manufacturers are introducing haptics into vehicle cockpits (e.g. BMW’s *iDrive*). We believe that the mixed initial results in those endeavors are an illustration of the need for greater attention to the perceptual attributes of the haptic signals, as well asto their integration. As methods for designing haptic signals emerge, this kind of support should spread.

We have integrated haptic mediation into a new urgency-based turn-taking protocol: with a vibrotactile signal, we periodically inform users whether they are in control of a shared application, waiting to take control, or simply observing their collaborators’ actions. With a haptic input, users can request control with two levels of urgency; a corresponding change in the displayed haptic signal notifies the in-control user of the request and its urgency. Thus, without disrupting the in-control user’s actions, we intermittently remind him of his collaborators’ (possibly changing) intentions.

This paper builds on previous studies performed by our group [5, 6] by describing an observational study where we test our new protocol. We compare three implementations of the protocol: one dominated by haptic signals, one with visual cues alone, and one balancing both modalities. We

investigate the effectiveness of the haptic signals in conveying turn-taking information, and explore the effect of the haptic interface on the collaborative experience and task performance.

Our results suggest that a modestly-sized set of well-designed haptic stimuli can be learned to a high degree of accuracy in under 5 minutes, that the inclusion of haptic stimuli can make turn-taking behavior more equitable, that users utilize the ability to make graded requests and control-holders respond more quickly to urgent requests.

BACKGROUND AND RELATED WORK

Turn-Taking

Since Engelbart proposed the notion of "Shared Screen Conferencing" [12], there has been significant research into real-time collaboration by distributed users. One variant of such collaboration provides support for users to simultaneously *interact* with a shared artifact. Examples are the GROVE group text-editor [11] and the Tivoli electronic whiteboard [24]. However, such systems require the implementing complex distributed architectures. Despite the availability of toolkits (e.g. [29]), most applications today do not support real-time collaboration and synchronous multi-user UI design remains a research topic.

An alternate approach has been to support real-time collaborative *viewing* of a shared artifact, where users take turns in control. The two most commonly cited supporting protocols are *give* (a user who wants control must wait for the user in control to voluntarily relinquish it) and *take* (a requestor is immediately given control). Other protocols have been proposed (e.g. [13, 22]), but their effectiveness has not been extensively evaluated.

Inkpen et al. studied the effect of turn-taking protocols on pairs of co-located children [15]. While girls did not show a difference, boys shared control more equally when using *take*; the amount of time boys had control was positively correlated with their ability to complete the same task on their own. McKinlay et al. compared the effectiveness of face-to-face communication with three computer-mediated communication protocols [18]: *give*, *take*, and *free-for-all* where users could work simultaneously. Face-to-face was the clear winner; and the data suggested that the others would be ordered as *give*, *free-for-all*, and *take*.

In general, there have been few empirical comparisons of different protocols, and attention in recent years has largely shifted towards multi-user synchronous groupware.

Currently Available Commercial Systems

Commercial web conferencing systems available today include Microsoft NetMeeting [21], WebEx [30], Macromedia Breeze [17], and Microsoft LiveMeeting [20]. All of these lightweight systems allow a user to share a single-user application with other users, and all use *give* by default, although some can also be configured to *take* instead. Except for LiveMeeting, they all include mechanisms for users to explicitly request control from the person in control. Requests are presented visually through "short-lived" dialog boxes (NetMeeting), transient tool-tips (WebEx), or persistent floating windows (Breeze). Dialog boxes can be disruptive, while tool-tips may be overlooked or conversely obscure what the user is looking at. In both NetMeeting and WebEx requests are transient – a user may be willing to relinquish control after completing his task, but the request is gone. Breeze addresses this, but the windows can still be easily missed.

Haptic Communication

There is evidence that haptic signals (haptic stimuli with which a meaning has been associated) may offer a way to nonintrusively inform users of their collaborator's desires, thus alleviating problems with current turn-taking protocols. A comprehensive understanding of our tactile psychophysical capabilities is emerging through the work of researchers such as Klatzsky and Lederman, who have, for example, documented our exquisite sensitivity to texture felt through a probe (e.g. [16]). Tan et al. measured information transfer rates of 2-3 items/second for vibrotactile stimuli independent of duration [28]: appreciable content can be conveyed through this channel.

Physically rendered virtual environment forces have also been used to aid collaboration. Sallnas had users jointly arrange cubes in a virtual environment [26], and Basdogan asked them to jointly move a ring along a wire without touching the wire [7]. Both found that haptic feedback significantly improved task performance and aided cooperation. Oakley added physical forces to telepointers in a shared editor so that users could push and pull each other [23]. Force feedback can help under workload: guiding forces have been shown to positively and nonintrusively influence behavior during an engrossing task [5, 27].

Other work has been directed towards building an abstract haptic language. One approach is to load haptic stimuli with comprehensible information, creating "haptic icons" [6]or "tactons" [9]. However, we are not aware of any studies outside of our group that specifically address the encoding of information in hierarchically clustered haptic signals.

NOVEL URGENCY-BASED TURN-TAKING SYSTEM

We describe our new protocol, the family of haptic signals designed to support it, their hardware instantiation, and the software environment in which the protocol operates.

Protocol

Conversational analysis reveals that when physically co-located, our body language communicates to our partners information such as how urgently we wish to speak [10], and that non-verbal communication is crucial for efficient information transfer [8]. We assume collaborative work has similar non-verbal communication needs; in particular, the need to express a desire for control with varying degrees of urgency. However, none of the turn-taking protocols we have observed provide flexibility in the means of requesting application control, nor do their implementations support

Figure 1: Modified *iFeel* mouse with thumb buttons.

user management of a request's intrusiveness – features we take for granted in everyday conversation.

With this in mind, we devised a protocol where a user is always in one of three states: in control of the shared application, waiting for control, or simply observing the actions of collaborators. A user obtains control in one of three ways: a Gentle Request for control, an Urgent Request for control, or by Taking control. If no one is in control at that time, any of these methods immediately result in the user obtaining control. Otherwise, requests are queued, with one queue for Gentle Requests and one for Urgent Requests. As users make requests, the user in control is made aware of the requests and their urgency. When a user releases control, the Urgent Request queue is given priority (until it is emptied) over the Gentle Request queue. If both queues are empty, no one is in control.

While this protocol is more complex than those described in the literature, we hypothesize that haptic mediation will clarify and streamline the added complexity.

Haptic Device

Logitech *iFeel* mice are used to deliver haptic feedback. These are standard optical mice with an embedded vibrotactile display, using technology licensed from Immersion Corp [14]. While the range of frequencies and amplitudes the *iFeel* supports is limited, we want to challenge the assumption that haptic technology is not ready for mainstream use by using an off-the-shelf component. This kind of display can also be placed on an arm band or similar device; our evaluation task relied heavily on the mouse, so it made sense to use the *iFeel*.

Command	Action
Gently Request Control	Press front button once
Urgently Request Control	Press front button twice; if already have Gently Requested Control, press front button once
Take Control	Hold front button for 2 seconds and release
Cancel Req. for Control	Press back button
Release Control	Press back button

Table 1: Haptic input for obtaining and releasing control.

We incorporated the ability to obtain and release control by adding two thumb buttons to the mouse. We had considered a force-sensing resistor to support squeezing the mouse to indicate more or less urgency; but this leaves unclear how much force to apply. Binary buttons do not have this problem. Commercially available mice have similar buttons to act as forward/back buttons for Internet browsing. Figure 1 shows a picture of the modified *iFeel*; button presses required for control changes are in Table 1.

Haptic signals

To support our protocol, we required 3 families of related haptic signals. The *Change in Control* family has two transient stimuli associated with (1) gaining or (2) losing control. The *In Control* family consists of three periodic, ongoing sensations delivered to the control holder: (3) no outstanding requests, or someone has made a (4) gentle or (5) strong requeste. The *Waiting for Control* family includes two sensations delivered periodically to a user waiting for control, following a (6) gentle or (7) strong request. A user simply observing the collaboration receives no haptic feedback. In total, 7 signals were needed.

We are able to vary the frequency, amplitude, and rhythm of stimuli generated by the *iFeel*. We had to ensure each signal was perceptually distinguishable from all others, yet similar to the signals in its own family. To accomplish this, we conducted a first study where we extended a technique developed by <*name removed*> to compose the set of 7

Family	State	Haptic Sensation
Change in Control	(1) User has gained control of the shared application	0.4 s, 1000-magnitude 100 Hz vibration, followed by a 0.1 s delay, followed by a 0.25 s, 8000-magnitude, 100 Hz vibration Feel: system "powering up" – weak buzz followed by a strong buzz
	(2) User has lost control of the shared application	*0.25 s, 8000-magnitude, 100 Hz vibration, followed by a 0.1 s delay, followed by a 0.4 s, 1000-magnitude 100 Hz vibration* Feel: system "powering down" – strong buzz followed by a weak buzz
In Control	(3) User is in control of the shared application	1 s, 500-magnitude, 60 Hz vibration; 1 s delay between iterations Feel: a light heartbeat
	(4) User is in control, but someone has gently requested control	1 s, 5000-magnitude, 60 Hz vibration; 1 s delay between iterations Feel: a stronger heartbeat
	(5) User is in control, but someone has strongly requested control	0.7 s, 5000-magnitude, 100 Hz vibration, followed by a 0.1 s pause, followed by a second identical vibration; 0.6 s delay between iterations Feel: a very strong heartbeat
Waiting for Control	(6) User has gently requested control	Single pulse; 1 s delay between iterations Feel: tapping a single finger on a table
	(7) User has urgently requested control	Two pulses, separated by a 0.15 s pause; 1 s delay between iterations Feel: drumming 2 fingers on a table

Table 2: Hierarchically clustered haptic signals for urgency-based protocol; 7 signals in total.

signals [4]. The resulting signals are described, including how they "feel" perceptually, in Table 2.

A second study was conducted to test whether users could easily learn the 7 signals, and whether they could detect changes in signals and accurately recall their meanings while engaged in visual and auditory distracter tasks. We found that subjects were able to learn the signals to 90% accuracy after exploring the stimuli for an average of 177 seconds; and they were able to detect changes in signals on average in 4.3 seconds while engaged in distracter tasks, with 95% identification accuracy. These performance results suggested that the haptic signals could successfully be incorporated into a turn-taking protocol. Complete details on our two initial studies can be found in [4].

We chose deliberately to always deliver periodic, low-level haptic feedback to a user in the *In Control* state, despite the amount of time they would spend there, on the premise that active (in control or queued) collaborators should always be at least peripherally aware of their state. Inactive (observing) collaborators receive no feedback. The *In Control* stimulus was designed to be subtle and not annoy; as no subject in our user study reported annoyance from that feedback, we seem to have achieved this.

Software Implementation

We modified the open-source Virtual Network Computing (VNC) [2] software to implement our protocol. VNC consists of client and server applications; using the client, a user can control the desktop of a remote computer running the server. If multiple users are viewing the same desktop, the server by default simply handles keyboard and mouse inputs in the order in which they arrive, often causing unpredictable results. No support is provided to help users decide who is in control. We modified both the client and server to use our turn-taking protocol instead.

USER STUDY

We conducted an observational exploratory user evaluation of the new haptically-supported protocol. Groups of 4 users completed furniture-layout tasks under three different conditions that differed in the amount of haptic and visual information provided.

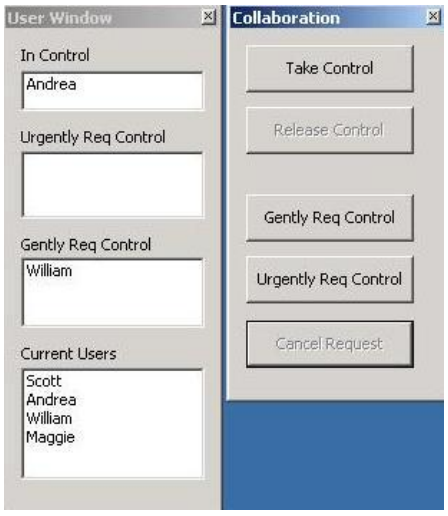

Figure 2: User window and button bar.

Our goal was to address the following research questions:

I. Can subjects learn the meanings associated with the haptic stimuli in a reasonable amount of time?

II. How will collaborative style be impacted by the different conditions?

III. How will equitability of control sharing be impacted by the different conditions?

IV. Which modality (visual, haptic, or combined) will subjects prefer for interaction information and control?

V. How will task performance be impacted by the different conditions, if at all?

Conditions

Our goal was to compare protocol mediation by the traditional visual modality (costly in attention and screen space) with the potentially less intrusive haptic channel. Thus, all three conditions in our study used the same urgency protocol and were designed as follows:

1. **Visual (V):** User Window and Button Bar. The User Window displays who is in control, who has gently and urgently requested control, and a list of the group members. The Button Bar allows users to request and release control. Both objects, shown in Figure 2, always float beside the shared application window.

2. **Haptic (H):** Haptic inputs and feedback as described on the previous page. The User Window from the V condition can also be displayed by pressing the space bar, but in this condition the window has to be dismissed before any other actions can be taken.

3. **Haptic + Visual (H+V):** Haptic input and feedback from H condition with User Window and Button Bar from V condition. Subjects can use either the Button Bar or haptic inputs to request and release control.

To minimize confounds, the same information is available in each condition, in particular persistent information about who is in control or requesting control at a given urgency. In the V condition this information is displayed continuously in the User Window, rather than through the transient tool-tips or dialog boxes used in current solutions. In the H condition we continuously transmit requests for control, and the user can invoke the User Window to identify who has made the requests.

We note that these conditions are not purely 'Visual' or 'Haptic', since both (haptic) motor activities and some visuals are necessarily involved in all cases; the conditions distinguish the primary source of information and control.

We could have made many other comparisons, for example our protocol with the more common *give* and *take*. This would have required an elaborate and lengthy study. We

felt it was more important to first investigate the protocol itself, and compare its haptic and visual instantiations.

Study Setup

The study was conducted in a graduate student laboratory. To simulate a distributed setting, subjects were seated at workstations such that they could not easily see each other. Subjects wore Sennheiser HD280 headphones and Sony ECM-T115 lapel microphones so they could communicate with one other easily. While working together, subjects within a group used VNC clients running on Pentium III and Pentium IV computers, ranging from 733 MHz to 2 GHz, and between 256 and 512 MB of RAM. Despite the variation in hardware, task performance was similar across computers because the VNC client is not computationally intensive. Each computer had a 17" LCD display with 1280x1024 screen resolution. The clients connected via 100 Mbps LAN to the VNC server, hosted on a Pentium IV 2.67 GHz, 512 MB computer. All computers ran Windows XP.

We automated data collection in several ways. The VNC client and server were instrumented to record information in log files. The open-source screen recording program *vncrec* [3] captured groups' efforts at creating task solutions by accessing the VNC server. Another open-source program *transcode* [1] was used to create MPEG4 movies of the data. Group conversations were recorded on audio tape.

Task

Our task was designed to closely approximate real-world group collaboration. Several task characteristics were deemed important:

- Groups should share a common body of knowledge, but each individual should possess specific, specialized knowledge.
- Groups should work towards a well-defined set of goals, but there should be constraints on how the goals can be achieved.
- Group members should have conflicting interests, but collaboration should not be adversarial.

We developed a furniture-layout task that satisfied these characteristics. We created three isomorphic tasks for our three conditions, all centering around furniture layout in a typical graduate-level lab in computer science. To deepen the collaborative aspect, the three tasks shared an identical set of eight firm constraints that *had* to be observed as specific task goals were met, and an identical set of eight soft constraints that *should* be observed; perfect solutions were impossible. For example, a goal in one task was to add 5-10 workstations to an existing room. One of the firm constraints was that three-foot wide walkways had to exist to each piece of furniture in the room; a soft constraint was that noisy areas should be isolated from workstations.

Groups were given 20 minutes to formulate a solution for each task. All members knew the complete set of goals they were to achieve; however, each member was responsible for two hard and two soft constraints, which were provided in written form. To mitigate subject development of expertise in one particular aspect of the task, subjects were responsible for different sets of constraints for each of the three tasks. The tasks were designed such that creating a near-optimal solution would be very difficult in the time given, but that a sufficing solution would be possible. Tasks were completed using Microsoft Visio, a diagramming tool.

Procedure

Subjects individually completed a training phase before working together on the furniture-layout tasks. First, subjects learned how to use Microsoft Visio. We simplified its interface to make it easier to learn, hiding all functions not needed in the study. Subjects were given a brief demonstration that showed how to add, move, rotate, and remove objects and groups of objects and then used Visio to complete a brief set of exercises, ensuring that they understood its use at this level. Next, subjects were trained to identify the haptic signals we used, first reading descriptions of the stimuli and their protocol-based meanings. A Java training application was then provided: subjects were shown a screen with a button for each of the 7 signals, ordered by family, which they could play as many times as they wanted, in any order. Hints were provided as descriptions on that same screen to help subjects learn the associations (see "Feel" descriptions, Table 2). When subjects felt they had learned the signals and their meanings, they proceeded to a learning test.

In the learning test, each signal was presented three times in random order, for a total of 21 trials. Subjects had to identify 19 or more signals correctly to pass; otherwise, they returned to the initial screen for more practice before repeating the test. Identification was made by clicking on the correspondingly labeled radio button; to prevent subjects from learning positional rather than meaning associations with the stimuli, the labeled radio buttons were randomly re-ordered on each trial. No feedback was given during the test other than whether they passed or failed.

After the training phase, the group completed the three conditions. Each condition was preceded by a five-minute warm-up period where the group completed a scripted set of actions to familiarize themselves with the user interface for that condition. Groups then spent 20 minutes working on a furniture-layout task. After each condition, subjects individually completed a questionnaire and were given five-minute rest breaks. At the end of the study, subjects individually completed an overall questionnaire. They were then interviewed and debriefed as a group.

The study required one three-hour session, for which subjects were each paid $25. As an incentive, groups were told that their task solutions would be evaluated, and the top ¼ of the groups would each receive a $40 bonus.

Design and Subjects

A within-subjects design was used so that subjects could compare the modalities in the three conditions. We adopted a "2x2+1" design. The V and H conditions were

counterbalanced, as were two out of the three tasks. The remaining task was always paired with the V+H condition, and it was always the last condition presented. By placing this condition last, we were able to record and ask which modality subjects relied on, having had equal exposure to the other two approaches. Our design required four groups, an appropriate number for an exploratory study.

Subjects were recruited with the following constraints: each group had to have at least one male and one female in it, and all members of the group had to be acquainted with one another. Additionally, we required that across groups, there be diversity in technical background and academic focus. These constraints were imposed to simulate real world group composition and for the purpose of understanding how the different conditions might impact groups with various backgrounds. Subjects were not screened with respect to Visio experience because the interface to Visio was reduced to the extent that only novice behavior was permitted, reducing any advantage advanced users might have. No person with familiarity of haptic stimuli similar to those used in our study was allowed to participate.

Dependent Measures

We measured learning effort through the amount of time spent exploring the haptic stimuli and the number of attempts required to pass the learning test. We measured aspects of collaboration through the time spent in control before releasing or losing it, the time spent waiting for control after submitting a gentle or urgent request, and frequency data such as the number of requests for control. To gauge task performance, we evaluated the task solutions according to how well they satisfied the specific goals, while respecting the constraints.

We also collected qualitative data relating to all topics from questionnaires and post-study interview data. The questionnaires consisted of Likert-scale and open-ended questions, and questions where subjects ranked the modalities in order of preference.

RESULTS

Four groups of 4 subjects participated in the study, 16 in total (8 male, 8 female). All subjects were students at the University of <*name removed*>. Subjects ranged in age from 18 to 41, had normal tactile sensitivity, used the mouse with the right hand, and had not participated in our earlier studies. They exhibited a variety of haptics exposure: 6 reported none, 6 used game controllers, 3 had used other haptic devices, and one did not respond to the question. Each group is described below:

The Engineers (3 males, 1 female) had taken over a year of undergraduate engineering courses together, participated jointly in extracurricular activities, and kept in touch even though one subject had changed faculties.

The Long-Time Friends (2 males, 2 females) were each majoring in a different area. They had known each other since secondary school; they attended religious services, played sports, and took courses together.

The Teachers (1 male, 3 females) were completing their Education degrees, as part of a cohort of approximately 40 people who took all their classes together for one year.

The Graduate Students (2 males, 2 females) consisted of two male-female pairs from different research labs. While each pair knew each other, the pairs did not; we permitted this due to recruitment difficulty, but found it useful to have a representative on this part of the acquaintance dimension.

We now summarize the study results according to our research questions. Although we did not anticipate any statistically significant results given the small number of groups, for completeness and curiosity we did run ANOVAs on some of the dependent measures across the t conditions and, where significant, we report those results.

I. Learning and Using Haptic Stimuli

The learning component of the study was designed to ensure that subjects could achieve a threshold level of performance within a reasonable amount of time, rather than test how quickly subjects could learn the haptic signals. Therefore, subjects were encouraged to learn 'carefully' rather than 'quickly', and it is possible that learning times reported here could have been even lower.

Learning time was calculated as the time spent exploring the haptic stimuli in the Java application. The learning test itself was not included, as results would be skewed by subjects who adopted an aggressive strategy, trying the test quickly and needing multiple attempts to pass it. Indeed, we observed that 3 of the 5 subjects with the longest learning times required only one attempt to pass the learning test. Learning times ranged from 51 – 270 seconds, with a mean time of 135 seconds and a standard deviation of 64 seconds.

These results show that associations between a moderate-sized set of well-designed haptic stimuli and compatible meanings can be quickly learned to a high degree of accuracy.

II. Collaborative Style

We were interested in the impact of condition on the frequency of control transfer, the distribution of the different methods for gaining control, and the influence of method on wait times. Across all groups and conditions, the number of control changes ranged from 8 to 29. As shown in Table 3, there were nearly twice as many changes in the H and H+V conditions compared to the V condition; a repeated-measures ANOVA approached significance ($F_{2, 6} = 4.552$, $p = 0.063$, partial $\eta 2 = 0.603$). This suggests that haptics may facilitate more frequent turnover of control.

Recall there were 4 ways subjects could attempt to gain control: Take (T), Urgent Request (UR), Gentle Request (GR), or a direct Obtain Control (OC). The OC category represents a control acquisition through any of the other 3 methods *when no one is already in control*. The OC

	Visual	Haptic	H+V
Engineers	13	22*	17
Long-Time Friends	20	18*	25
Teachers	8*	20	18
Graduate Students	9*	21	29
Average	12.50	20.25	22.25
Std. Deviation	5.45	1.71	5.74

Table 3: Number of changes in control for each condition; '*' denotes condition seen first (N=16).

	Visual	Haptic	H+V
Engineers	96.3	50.1	69.1
Long-Time Friends	57.8	63.9	49.6
Teachers	153.5	55.6	70.5
Graduate Students	109.8	33.4	35.8
Col Sum/# Groups	104.3	50.8	56.3
Std. Deviation	39.5	12.9	16.6

Table 4: Control-Holder's Perspective. Average lengths of periods **in control** before releasing or losing control (sec), by group & condition. **Unweighted table average = 70.5. (N=16).**

category is distinguished because we cannot determine from the data when subjects knew someone was in control or not; for example, a T in the former situation may represent an aggressive collaborative style, whereas in the latter situation it is simply one of 3 methods to assume control when no one else has it.

Figure 3 shows the distribution of methods for gaining control used for each condition. It does not include OCs, which represent 12, 39, and 46 attempted acquisitions for V, H, and H+V respectively. The figure shows a clear preference for GRs over Ts and URs in all conditions: GRs accounted for 60-73% of the acquisition requests made when control was held by another. By contrast, the balance between the two *stronger* acquisition methods, Ts and URs, suggests their use differed by condition. In the V condition, there was a clear preference for T over UR (37%:2%), whereas in the H and H+V conditions they are used more

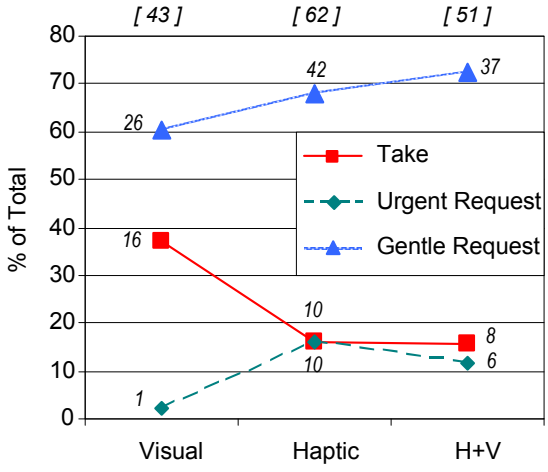

Figure 3: Distribution of methods for requesting control. [These do not include directly Obtained Controls (OCs). OC values are 12, 39, and 46 for V, H, H+V respectively.] **(N=16).**

	Visual	Haptic	H+V
Engineers	111.4	39.5	8.6
Long-Time Friends	46.8	23.8	5.9
Teachers	24.3	28.0	6.4
Graduate Students	50.0	4.3	3.0
Col Sum/# Groups	58.1	23.9	6.0
Std. Deviation	37.3	14.6	2.3

Table 5: Gentle Requestor's Perspective. Average time from a **gentle request** until gaining control (sec), by group & condition. **Unweighted table average 29.3. (N=16).**

	Visual	Haptic	H+V
Engineers	-	7.2	22.0
Long-Time Friends	-	1.0	3.5
Teachers	67.0	33.0	2.0
Graduate Students	-	-	-
Col Sum/# Groups	67.0	13.7	9.2
Std. Deviation	-	17.0	11.1

Table 6: Urgent Requestor's Perspective. Average time from an **urgent request** until gaining control (sec). '-': no urgent requests made. **Unweighted table average = 19.4. (N=16).**

equally (16%:16% and 16%:12% respectively). A repeated-measures ANOVA showed a significant effect of method ($F_{2, 4} = 17.401$, $p = 0.003$, partial $\eta 2 = 0.853$) with post-hoc pair-wise comparisons showing that GRs outnumbered URs ($p = 0.054$) and Ts ($p = 0.093$). The data suggests increased access of our protocol's key feature, the graded request, in conditions including H: URs represented 11% of all requests, distributed as 2/16/12% in V, H and H+V. Finally, across all conditions, control holders were more responsive to URs, releasing control in an average of 19.4 secs as opposed to 29.3 for GRs (Tables 5-6).

Several additional details aid interpretation of the numbers:

- Having only one T in the V condition may have been because request urgency was not emphasized in that condition; the requestor's name simply appeared in a different list in the User Window.
- One group, the Graduate Students, never used UR under any condition. Additionally, some subjects reported post-study that they did not feel the need to make URs because a GR was fast enough.
- In most groups, a subject typically retained control until someone else requested it. However, the Graduate Students adopted a practice in their second condition whereby they each released control as soon as they had finished their task; this accounts for the large increase in the OC counts for the H and H+V conditions.

Conditions including haptics thus seemed to facilitate increased turnover in control frequency and more even distribution of URs and Ts as compared to the V condition, indicating a generally more dynamic collaborative style. Group practices also played an important role.

III. Equitability of Sharing Control

Although control turnover was more frequent in the presence of haptics, we wanted to know whether haptics

	Visual	Haptic	H+V
Engineers	0.44	0.23	0.35
Long-Time Friends	0.39	0.25	0.24
Teachers	0.39	0.13	0.22
Graduate Students	0.47	0.25	0.35
Col Sum/# Groups	0.42	0.22	0.29
Std. Deviation	0.04	0.06	0.07

Table 7: Equitability of Control Time. Spread of percentage of time in control, between group members most and least in control (%). **Unweighted table average = 0.31. (N=16).**

promoted equitability of control amongst team members. By examining how much time subjects spent both in control and waiting for control, we can see that the collaboration dynamics changed across the conditions.

The average amount of time a subject spent in control of Visio in any one turn, before releasing or losing control, was noticeably larger in V than the other conditions (Table 4). A comparison of group means yielded a significant main effect of condition ($F_{2, 6} = 5.849$, $p = 0.039$, partial $\eta 2 = 0.661$) but Bonferroni corrected post-hoc pair-wise comparisons were not statistically significant.

This suggests that subjects in control of Visio were more responsive to requests for control in the H and H+V conditions than in the V condition. This hypothesis is supported by the average wait durations after a GR (Table 5) and a UR (Table 6) control, both of which are shorter when haptic feedback is present. For GRs, a comparison of group means showed a significant main effect of condition ($F_{2, 6} = 6.747$, $p = 0.029$, partial $\eta 2 = 0.692$) but Bonferroni corrected post-hoc pair-wise comparisons were not statistically significant. Three of the four groups did not make URs at all in the V condition, making that measure more difficult to compare across conditions.

Examination of the overall percentage of time each subject was in control under each condition revealed an interesting finding. For each group, a spread was calculated by subtracting the percentage of the subject in control the least from the percentage of the subject in control the most (Table 7). We observed that the spreads were larger in the V condition as compared to H+V and especially the H condition. Statistical analysis yielded a main effect of condition ($F_{2, 6} = 37.405$, $p < 0.001$, partial $\eta 2 = 0.926$), with post-hoc pair-wise comparisons showing that the spread was significantly larger in the V condition than the Haptic ($p = 0.011$) and H+V ($p = 0.014$) conditions.

Together these results suggest that the sharing of control is more equitable in the presence of haptics.

IV. Subject Preferences

Subjects ranked the three conditions in order of preference for obtaining control (input to the system), for displaying the turn-taking state (output/feedback from the system), and for overall preference. Table 8 shows that the H+V condition was overwhelmingly preferred overall (12 subjects) as compared to the two single-modality

	Visual	Haptic	H+V
Obtaining Control	**1** / 7 / 8	**4** / 5 / 7	**11** / 4 / 1
Conveying State	**4** / 6 / 6	**1** / 5 / 10	**11** / 5 / 0
Overall	**2** / 8 / 6	**2** / 4 / 10	**12** / 4 / 0

Table 8: Condition Preference. Number of subjects who ranked a given condition first / second / third **(N = 16).**

	Visual	Haptic	H+V
Engineers	0.69	0.28*	0.84
Long-Time Friends	0.60	0.44*	0.65
Teachers	0.22*	0.70	0.73
Graduate Students	0.41*	0.79	0.83
Col Sum/# Groups	0.48	0.55	0.76
Std. Deviation	0.21	0.23	0.09

Table 9: Normalized task scores. '*' denotes condition seen first (N=16).

conditions, which each received an equal number of overall first order rankings (2 subjects each). It is thus unsurprising that H+V was also favored for both obtaining control and displaying state (11 subjects). However, the H condition was the next preferred for obtaining control, whereas for displaying state, the V condition was next preferred.

When asked to justify their overall rankings, subjects who ranked H+V first noted that the haptic feedback notified changes in state, and the User Window indicated who was in control or requesting control. One subject noted the lack of haptic feedback when no one was in control, and sometimes displayed the User Window to confirm that fact. This may be a limitation of our design that could be considered in the next iteration. Another difficulty that arose occasionally in the H condition was that the person gaining control was not who the person releasing control expected it would be, based on the audio dialog. To address this haptically, it might be necessary to add subject identification information to the stimuli; more study would be required to determine if this is perceptually workable.

Subjects clearly expressed a desire for the User Window to be present. However, the data did not show it to be necessary. After each condition, subjects were asked to indicate how much they agreed with the statement, "I constantly monitored the User Window when someone asked me for control." In the H condition, 11 subjects either disagreed or strongly disagreed with the statement, compared to 6 subjects in the H+V condition and 3 subjects in the V condition. Furthermore, the number of times the User Window was opened was measured in the H condition; on average, each subject opened the User window fewer than three times.

In summary, subjects preferred H+V overall. However, the justification for preferring features specific to V is ambiguous and we anticipate that it might diminish with familiarity of the H features.

V. Task Performance

For completeness, we checked task performance across conditions, even though significant variation was not

expected due to the short task duration. Task solutions were evaluated based on how well they satisfied the goals for the task, while following the specified hard and soft constraints. We used a points-based system: points were awarded for satisfying a specific goal, and points were deducted for failing to satisfy a constraint. The penalty for violating hard constraints was more severe than violating soft constraints. Prior to running the study, a reference solution for each task was created to estimate its maximum possible score.

Each group's task solution was scored, and the resulting score was divided by the reference solution score (Table 9). A learning effect was evident between the first and second conditions completed. On average, groups had slightly better performance in the H condition than the V condition, but the difference is not statistically significant. The better average performance in the H+V condition is almost certainly due to learning effects.

We also checked to ensure that there was no effect of task in the first two conditions, which were counterbalanced for task as well as condition order: the mean scores on those two tasks were 0.50 and 0.54 respectively.

Thus, apart from learning effects, the results suggest that the different modalities did not impact task performance. Longer exposure to the conditions, may however yield measurable differences.

DISCUSSION

Impact on Quality of Collaboration and Equitable Sharing: We found that our background-level, haptically-supplied information increased both overall turnover of control and usage of graded requests, and also seemed to promote equitability in the total amount of time each subject was in charge. To explain these positive effects, we hypothesize that users may have found this method of control exchange either less cumbersome or more informative – since no condition *compelled* users to give up control more than others, any increase in either overall frequency or use of a specific method was presumably by preference.

Thus we see consistent evidence across the available indicators that this kind of carefully designed haptic feedback can facilitate a more active, equitable distributed collaborative style, in ways that might apply to co-located collaboration as well. It will be of interest to see whether in more extensive tests, this extends beyond notions of fairness to better performance.

Stakes: URs represented 11% of all requests made when control was held by another (2% in V and 16% in H). We observe that while GRs will presumably dominate any effective interaction, the low stakes typical in an experiment suggest that this result may be conservative: despite our incentives, our subjects were conscientious and polite as well as new to the concepts involved. In a real-world (non-experiment) setting, when group members have strongly vested interests in the project outcome, the availability of choices other than rudeness or silence might allow high-stakes collaboration to be intense without becoming adversarial. This benefit could apply to both co-located and distributed situations.

Familiarity and Learning: We required our subjects to quickly learn several new concepts, including the notion of distributed collaboration, a turn-taking protocol, and identifying haptic stimuli. Newness may thus be a factor in some observations. For example, subjects liked the idea of the extra information available in conditions including V, but seldom used it. With time, would they outgrow this need for the familiar? Conversely, will they gain both the ability to process the haptic signals more automatically, and the confidence to rely on this channel?

Utilization of Information in Haptic Signals: Our learning data demonstrates our subjects' facility in learning our haptic signals, but does not tell us how easily haptic content was accessed under workload. Near the end of the study we began to follow up on a question of subjects using the haptic signals merely as binary triggers. Some claimed they used the signals only as notification; others said they were able to identify specific meanings. However, they may have been unaware of their identification ability, and of the extent to which they were utilizing meanings. Subjects in our earlier studies occasionally expressed surprise when they passed the learning test, and results demonstrated consistently accurate identification under workload. There is substantial evidence for nonconscious perception and utilization of visual stimuli [19, 25], and it is likely that this mechanism exists for haptic perception. However, longer exposure and different tests will be required to definitively answer this question.

Value of an Urgency-Based Protocol: We were not able to explicitly compare our augmentations to the presumed ideal of co-location, nor to a protocol without graded control requests. However, when given the chance to easily make graded requests, subjects did so; further, control-holders were overall more responsive to URs than to GRs. Together, these demonstrate potential impact of the urgency-communicating aspect of our protocol on both control holder and requestor behavior, and suggest that our subjects found it useful. This justifies proceeding to the next step, which is to establish whether this protocol can bring the quality of collaboration and task performance to co-located levels or even beyond.

CONCLUSIONS AND FUTURE WORK

We have described our implementation of an urgency-based turn-taking protocol in a visual collaborative task, and report observations of the behavior and verbalized reactions of a diverse group of subjects while using haptic, visual or bimodal mediation of control transfers.

Our haptic mediation appears to influence collaborative style positively relative to visual means: control turnover is higher and amount of time in control more equitable. There is also evidence that haptic mediation better facilitates our

protocol's communication of a control requestor's urgency; even though our data probably underestimates the utility of graded request urgency in real situations with higher stakes.

We believe that the haptic mediation was more usable both perceptually and socially because by occupying a different sensory modality than the primary task, it was less intrusive. At least some individuals seemed able to access haptic signal content under workload despite the short learning time, and earlier data suggests that the content may be acted upon without the user's awareness. Socially, we expect to see 'shy' users more comfortable requesting control, and dominating users more aware of others' needs.

The potential utility of lightweight haptic signals and our urgency-based protocol extend to other categories of distributed collaboration and shared control over resources or a group's attention – including co-located settings, where it can be awkward to get the attention of a dominating participant: familiar examples are meeting floor control, team document editing, design reviews. In more specialized examples, air traffic controllers need to smoothly transfer responsibility and information as airplanes cross jurisdictions, as do emergency personnel when coordinating a response to a massive forest fire. The use of haptic signaling might unintrusively facilitate other kinds of group communication: for example, students may learn better when they can confidentially ask instructors to slow down.

We plan to continue this investigation by considering factors such as longer exposure to signals and our protocol, as well as scenarios with higher stakes; obtaining more extensive longitudinal data from which performance distinctions as well as patterns of interaction may emerge; and comparing our protocol with others.

REFERENCES

[1] Transcode. Accessed 2004. http://zebra.fh-weingarten.de/~transcode/.

[2] Virtual Network Computing. Accessed 2004. http://www.realvnc.com/.

[3] Vncrec. Accessed September. http://www.sodan.org/~penny/vncrec/.

[4] Anonymous-1 (Reference removed for anonymity).

[5] Anonymous-2 (Reference removed for anonymity).

[6] Anonymous-3 (Reference removed for anonymity).

[7] Basdogan, C., Ho, C., Srinivasan, M., and Slater, M., "An Experimental Study on the Role of Touch in Shared Virtual Environments," ACM TOCHI, vol. 7:4, pp. 443-460, 2000.

[8] Boyle, E., Anderson, A., and Newlands, A., "The Effects of Visibility on Dialogue and Performance in a Cooperative Problem Solving Task," Language & Speech, vol. 37:1, pp. 1-20, 1994.

[9] Brewster, S., and Brown, L., "Non-Visual Information Display Using Tactons," in Proc. of CHI 2004, Extended Abstracts, pp. 787-788, 2004.

[10] Duncan, S., and Niederehe, G., "On Signalling That It's Your Turn to Speak," Journal of Experimental Social Psychology, vol. 10, pp. 234-247, 1974.

[11] Ellis, C., Gibbs, S., and Rein, G., "Groupware: Some Issues and Experiences," Comm. of the ACM, vol. 34:1, pp. 39-58, 1991.

[12] Engelbart, D., Toward High Performance Knowledge Workers. Accessed 2004. http://www.bootstrap.org.

[13] Greenberg, S., "Sharing Views and Interactions with Single-User Applications," in Proc. of ACM/IEEE Conf. on Office Information Systems, pp. 227-237, 1990.

[14] Immersion Corp. Accessed 2004. http://www.immersion.com/.

[15] Inkpen, K., McGrenere, J., Booth, K., and Klawe, M, "Turn-Taking Protocols for Mouse-Driven Collaborative Environments," in Proc. of Graphics Interface '97, pp. 138-145, 1997.

[16] Klatzky, R., Lederman, S., et al., "Feeling Textures through a Probe: Effects of Probe and Surface Geometry and Exploratory Factors," Perception & Psychophysics, vol. 65:4, pp. 613-631, 2003.

[17] Macromedia Breeze. Accessed 2004. http://www.macromedia.com/software/breeze/.

[18] McKinlay, A., Arnott, J., et al., "A Study of Turn-Taking in a Computer-Supported Group Task," in Proc. of HCI '93: People and Computers VIII, pp. 383-394, 1993.

[19] Merikle, P., Smilek, D. and Eastwood, J.D., "Perception without Awareness: Perspectives from Cognitive Psychology," Cognition, vol. 79, pp. 115-134, 2001.

[20] Microsoft LiveMeeting. Accessed 2004. http://main.placeware.com/.

[21] Microsoft Windows Technologies, Windows Netmeeting. Accessed 2004. http://www.microsoft.com/windows/netmeeting/.

[22] Myers, B., Chuang, Y., Tjandra, M., Chen, M., and Lee, C., Floor Control in a Highly Collaborative Co-Located Task. Accessed 2004. http://www-2.cs.cmu.edu/~pebbles.

[23] Oakley, I., Brewster, S., and Gray, P., "Can You Feel the Force? An Investigation of Haptic Collaboration in Shared Editors," in Proc. of Eurohaptics, 2001.

[24] Pedersen, E., McCall, K., Moran, T., and Halasz, F, "Tivoli: An Electronic Whiteboard for Informal Workgroup Meetings," in Proc. of InterCHI '93, pp. 391-398, 1993.

[25] Rensink, R.A., "Visual Sensing without Seeing," Psychological Science, vol. 15, pp. 27-32.

[26] Sallnas, E., Rassmus-Grohn, K., and Sjostrom, C., "Supporting Presence in Collaborative Environments by Haptic Force Feedback," ACM TOCHI, vol. 7:4, pp. 461-476, 2000.

[27] Steele, M.G., B., "Shared Control between Human and Machines," in Proc. of Human Factors and Ergonomics Soc., 2001.

[28] Tan, H., Durlach, N., Reed, C., and Rabinowitz, W, "Information Transmission with a Multifinger Tactual Display," Perception & Psychophysics, vol. 61:6, pp. 993-1008, 1999.

[29] Tse, E., and Greenberg S, "Sdgtoolkit: A Toolkit for Rapidly Prototyping Single Display Groupware," in Proc. of ACM CSCW, 2002.

[30] WebEx Communications Inc. Accessed 2004. http://www.webex.com/.