Duplex narrowcasting operations for multipresent groupware avatars on mobile devices

Michael Cohen*, Owen Noel Newton Fernando, Uresh Chanaka Duminduwardena and Makoto Kawaguchi

Spatial Media Group,
University of Aizu, Aizu-Wakamatsu,
Fukushima, 965-8580, Japan
Fax: (+81)(242) 37-2549
E-mail: mcohen@u-aizu.ac.jp
E-mail: newtonfernando@mixedrealitylab.org
E-mail: Uresh.Duminduwardena@dialog.lk
E-mail: kawaguchi.makoto@jri.co.jp
*Corresponding author

Abstract: The apparent paradoxes of multipresence, having avatars in multiple places simultaneously, are resolvable by an 'autofocus' feature, which uses reciprocity to project overlaid soundscapes and simulate the precedence effect to consolidate the display. We have developed an interface for narrowcasting functions for networked mobile devices deployed in a Collaborative Virtual Environment. Featuring multiple icons in a '2.5D' application, the interface can be used to control motion, sensitivity, and audibility of avatars in a teleconference. The interface is integrated with other clients through a 'servent' (server/client hybrid) HTTP–TCP/IP bridge, interoperating with a heterogeneous groupware suite including stereographic panoramic browsers and spatial audio backends and speaker arrays.

Keywords: chatspace; groupware; MMORPG; massively multiplayer online role-playing games; mobile computing; multipresence; multiuser interface; narrowcasting functions; teleconferencing.

Reference to this paper should be made as follows: Cohen, M., Fernando, O.N.N., Duminduwardena, U.C. and Kawaguchi, M. (2009) 'Duplex narrowcasting operations for multipresent groupware avatars on mobile devices', *Int. J. Wireless and Mobile Computing*, Vol. 3, No. 4, pp.280–287.

Biographical notes: Besides mobile computing, Michael Cohen's research interests include telecommunication semiotics, spatial audio, stereography, mixed and virtual reality, and telepresence. He is currently a Professor at the University of Aizu. He is on the Scientific Committee of the *Journal of Virtual Reality and Broadcasting*, a member of the ACM, Audio Engineering Society (AES), IEEE Computer Society, Institute of Electronics, Information and Communication Engineers (IEICE), Japan Mathematica Users Group (JMUG), 3D-Forum, TeX Users Group (TUG), and Virtual Reality Society of Japan (VRSJ). He is currently Vice-Chair of the Dept. of Computer and Information Systems, Graduate School of Computer Science and Engineering of the U. of Aizu.

Owen Noel Newton Fernando is a Research Fellow in the KEIO-NUS Connective Ubiquitous Technology for Embodiments (CUTE) Center of the National University of Singapore. Prior to joining Keio-NUS CUTE Center he was attached to the Mixed Reality Laboratory (MXR) of the same university. He received his BSc in Computer Science from the University of Colombo, Sri Lanka, and MSc and PhD in Computer Science and Engineering from the University of Aizu, Japan. In 2008, he completed a term of six months as a Visiting Research Scholar at the Academic Center for Computing and Media Studies of Kyoto University.

Uresh Chanaka Duminduwardena is a Manager in Software Engineering Services, Dialog Telekom PLC (Colombo, Sri Lanka). He received BSc (Hons.) in CS from University of Colombo (Sri Lanka) in 2001, and MSc in CS&E from University of Aizu (Aizu-Wakamatsu, Japan) in 2004. He was a senior solutions architect in telecommunications related software development. His research interests are mobile computing, haptic devices, and service delivery platforms (SDPs).

Makato Kawaguchi received his MS from the U. of Aizu in 2005, with interests in mobile computing and narrowcasting functions for mobile phones. He is now a System Integrator for The Japan Research Institute in Tokyo.

1 Introduction

Our group is researching CVES, realtime interactive interfaces and applications for teleëxistence and artificial reality (Fernando et al., 2002; Bolhassan et al., 2002; Cohen et al., 2002). Anticipating ubicomp networked appliances and information spaces (Okoshi et al., 2001), we are integrating various multimodal (auditory, visual, haptic) i/o devices into a virtual reality groupware suite. Such environments are characterised, in contrast to general hypermedia systems, by the explicit notion of the position (location and orientation) of the perspective presented to respective users, and often such vantage points are modelled by the standpoints and directions of icons in a virtual space. These icons might be more or less symbolic (abstract) or figurative (literal), but are representative of human users, and are therefore 'avatars' (after the Hindu notion of an earthly manifestation of a deity). Avatars reify embodied virtuality, treating abstract presence as a user interface object.

Non-immersive perspectives in virtual environments enable flexible paradigms of perception, especially in the context of frames-of-reference for conferencing and musical audition. Traditional mixing idioms for enabling and disabling various audio sources employ `mute` and `select` functions which selectively disable or focus on respective channels. (On many interfaces, '`mute`' and '`solo/select`' are abbreviated simply 'M' and 'S' (not to be confused with 'master/slave', 'mid/side' [as in coincident microphone techniques], etc.). Previous research (Cohen, 2000) defined sinks as symmetric duals of audio sources in virtual spaces, along with symmetric analogs of source `select` and `mute` attributes. Exocentric interfaces, which explicitly model not only sources, but also sinks, motivate the generalisation of `mute` & `select` (or `cue` or `solo`) to `exclude` and `include`, manifested for sinks as `deafen` & `attend` (`confide` and `harken`), as follows.

Narrowcasting and selection functions can be formalised in predicate calculus notation, where '¬' means 'not', 'Λ' means conjunction (logical 'and'), '∃' means 'there exists', and '♥' means 'implies'. The general expression of inclusive selection is $\texttt{active}(x) = \neg\texttt{exclude}(x) \wedge (\exists\, y\ \texttt{include}(y) \heartsuit\ \texttt{include}(x))$. So, for `mute` and `select` (`solo`), the relation is $\texttt{active}(\text{source}_x) = \neg\texttt{mute}(\text{source}_x) \wedge (\exists\, y\ \texttt{select}(\text{source}_y) \heartsuit\ \texttt{select}(\text{source}_x))$, `mute` explicitly turning off a source, and `select` disabling the collocated (same room/window) complement of the selection (in the spirit of "anything not mandatory is forbidden"). For `deafen` and `attend`, the relation is $\texttt{active}(\text{sink}_x) = \neg\texttt{deafen}(\text{sink}_x) \wedge (\exists\, y\ \texttt{attend}(\text{sink}_y) \heartsuit\ \texttt{attend}(\text{sink}_x))$.

The suite of inclusion and exclusion narrowcast commands for sources and sinks are like analogs of burning and dodging (shading) in photographic processing. The duality between source and sink operations is tight, and the semantics are identical: an object is inclusively enabled by default unless,

a it explicitly excluded (with `mute` || `deafen`)

b peers are explicitly included (with `select` [`solo`] || `attend` : `confide` or `harken`) when the respective icon is not.

Such functions, which filter stimuli by explicitly blocking out and/or concentrating on selected entities (Fernando et al., 2003) can be applied not only to other users' sinks for privacy, but also to one's own sinks for selective attendance or presence. These narrowcasting commands control superposition of soundscapes (Fernando et al., 2004). In the awareness parlance of Greenhalgh and Benford (1995), Benford et al. (1995, 1998), an aura delimited by a graphical window is like a room, sink attributes affect 'focus', and source attributes affect 'nimbus'.

A unique feature of our system is the ability of a human pilot to delegate multiple avatars simultaneously, increasing *quantity* of presence. Multiple sources are useful, for instance, in directing one's remarks to specific groups, decreasing the granularity of audibility control. Multiple sinks are useful in situations in which a common environment implies social inhibitions to rearranging shared sources like musical voices or conferences, as well as individual sessions in which spatial arrangement of sources, like the configuration of a concert orchestra, has mnemonic value.

1.1 Iαppli: I-mode with java

I-mode's Java-based iαppli service allows users of compatible terminals to download software and content from about a thousand websites. They can then use the downloaded applications and content whenever they want, without having to reconnect to the internet. Launched in January 2001, iαppli is the result of advances in wireless technology and NTT DoCoMo's cooperation in recent years with Sun Microsystems. Iαppli uses the Java platform developed by Sun for consumer electronics and built-in devices, as well as i-mode's extended library, which was developed jointly by the two companies.

Since the data processing power of mobile phones is far less than that of PCs, i-mode uses a version of Java called ‘KVM’ that runs on systems with relatively low processing power. A key feature of the KVM environment is that its security functions are tighter than those of standard Java. So, iαppli users need not worry about unauthorised viewing of their digital address books and other personal files, or the unintentional placing of calls billable to them. Although KVM iαppli cannot directly run standard Java programs, users can easily convert many Java applications for i-mode use with minimal changes in functionality. Significantly, the overall architecture of the i-mode network remains the same with the advent of iαppli. With i-mode terminals, iαppli users download Java applications from conventional HTTP internet sites just as easily as they do HTML documents, and Java programs on i-mode terminals also use the convenient HTTP format for data communications.

1.2 *‘FPS’ and ‘MMORPG’*

High-speed, multiplayer games have yielded a new interpretation of ‘fps’: not ‘**f**rames **p**er **s**econd’, but ‘**f**irst-**p**erson **s**hooter’, as in so-called ‘twitch’ games like ‘Doom’ or ‘Quake’. Typically less violent than FPS counterparts, RPGs (**r**ole-**p**laying **g**ames) depend on coherent stories, rich graphical environments, and interaction with other players. In MMORPGs (**m**assively **m**ultiplayer **o**nline RPGs) (Kolbert, 2001) – fantasy games like Sony’s ‘EverQuest’, Cyro’s ‘Man-kind’, Origin Systems, ‘Ultima Online’, and Microsoft’s ‘Asheron’s Call’– players create characters (avatars) to explore persistent universes that exist across sessions, and ‘massive’ means on the order of thousands of users per server. Such games increasingly feature audio, including both locally-generated sound effects and distally transmitted voice channels. Advanced floor and presence control in chat-spaces and conferences spawned by such coteries is needed, like that suggested by Table 1.

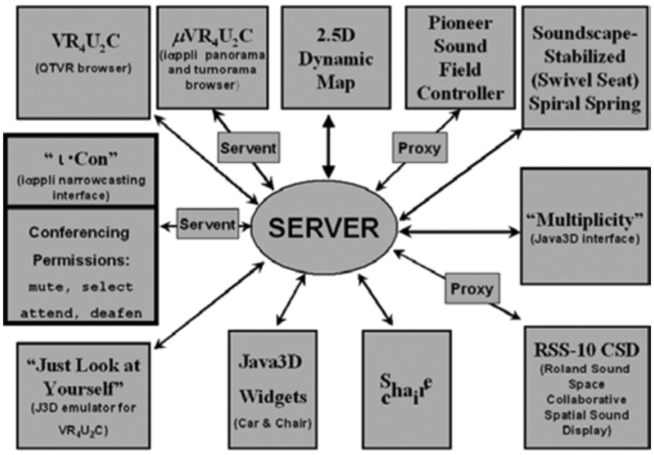
Table 1 Roles of source and sink

	Source	*Sink*
Function	Radiation/transmission	Reception
Level	Amplification	Sensitivity
Direction	OUTput	INput
Instance	Speaker	Listener
Transducer	Loudspeaker	Microphone or dummy head
Organ	Mouth	Ear
Tool	Megaphone	Ear trumpet
Include	Select (solo) or cue	Attend: confide and harken
Assert	$\overset{+}{\triangle}$	$+\triangle+$
Exclude	Mute	Deafen
Inhibit	$\overset{-}{\triangle}$	$-\triangle-$

1.3 *CVE architecture*

We have designed and implemented an architecture and framework (Kanno and Cohen, 2002) to support a CVE (Churchill et al., 2001), allowing distributed users to share multimodal virtual worlds (Cohen and Martens, 2001). Our CVE architecture (sketched in Figure 1) is based upon a **c**lient/**s**erver (C/S) model, and its main transaction shares the state of virtual objects and users (avatars) by effective multicast via replicated-unicast of position (translation, rotation) parameters and narrowcasting attributes to client peers in a session.

Figure 1 CVE architecture: groupware suite



The client/server architecture enables multimodal communication, platform independence, and easy network connectivity, as components are built with Java (and JMF (**J**ava **M**edia **F**ramework) (Gordon and Talley, 1999), QuickTime for Java (Maremaa and Stewart, 1999), Java3D (Walsh and Gehringer, 2002; Palmer, 2001; Sowizral et al., 2000), J2ME (Knudsen, 2001; Feng and Zhu, 2001), and Swing (Walrath and Campione, 1999).

There are three main classes of state in this model, ‘Location Parameter’, ‘Orientation Parameter’, and ‘Extra Parameter’. Location parameter handles translational state changes with three integer attributes x, y, and z. An orientation parameter handles the state changes of orbiter attitudes, pitch, roll, and yaw, which represent rotation of an object about the x, y, and z axes, respectively, to the desired orientation. Extra parameter handles all other state changes, especially narrowcasting states, as `<name><value>` pairs.

The CVE server is a multi-threaded (Java socket) application that creates a separate thread for each client and listens for state changes. Each server can create many communication channels according to client requests and each channel has its own name. Depending on the channel ID, the session server redistributes events to relevant clients.

At least for the demonstrations we have conducted so far, server redistribution performance has not been an issue. The server can run comfortably on the same host as some of the clients without noticeable performance degradation. In our CVE, 3D image data are not transmitted on the wire, but only position updates, which are used by the respective clients for local rendering. That is, the geometric, texture maps, etc. are compiled into the application and, so, need not be distributed at run time.

2 ‘ι·Con’: Iαppli (DoJa) mobile device dynamic map

We have designed and implemented a mobile telephone interface (Nagashima and Cohen, 2001; Cohen and Kawaguchi, 2003) for use in CVEs (Kanno et al., 2001). Programmed with J2ME (Java2, Micro-Edition) (`www.sun.com/j2me`) (Ascii editing group, 2001; Yamazaki, 2001; Topley, 2002; Vacca, 2002; Kontio, 2003) our application runs on an (NTT DoCoMo) (`www.nttdocomo.com`) iαppli mobile phone, as illustrated in Figure 2. Featuring selectable icons with one rotational and two translational degrees of freedom, the ‘ι·Con’ 2.5D dynamic map interface is used to control position, sensitivity, and audibility of avatars in a groupware session. Its isosceles triangle icons are representatives of symbolic heads in an orthographic projection; its narrowcasting operations are shown in Figure 3 and Table 2. The interface is further extended with musical and vibrational cues, to signal mode changes and successful transmission/reception (which feedback is important in wireless communication, as it is much less deterministic than terrestrial systems).

Figure 2 NTT DoCoMo i-mode iαppli iJade emulator running ‘ι·Con’ application (see online version for colours)

Figure 3 Postfix grammar for keypad entry: operands are chosen by toggling icons tagged with session-unique IDs into/out of the selection set, upon which operations to change position or attributes are subsequently invoked

```
phrase := selectionToggle || operationToggle || exit
selectionToggle := channelNumber + ‘#’
operationToggle := attribute + ‘*’
attribute := (<attend> || <deafen> || <mute> || <select>) || <sink>
exit := ‘*’ + ‘*’
```

Table 2 Mnemonic initials of conferencing narrowcasting operations on the alphanumeric keypad used to toggle selection set attributes

Attributes	*Keys*
`attend`	ABC 2
`deafen`	DEF 3
`mute`	MNO 6
`select`	PQRS 7
`sink/self`	GHI 4

Current user interfaces for mobile phones cannot strictly be characterised as ‘GUI’s since, in its usual interpretation, the acronym (for ‘**g**raphical **u**ser **i**nterface’) connotes a ‘WIMP’ idiom (being itself acronymic for ‘**w**indow, **i**con, **m**enu, **p**ointer’), and the mobile phone lacks a windowing system, menus, and a cursor-style pointer. A better association might be what is sometimes called a ‘SUI’, for ‘**s**olid **u**ser **i**nterface’, as modern mobile phones feature unique interface conventions, including vibration, thumb-favoured text input, and, on some models, a jog shuttle.

In full-screen GUIs, a new selection resets the selection set, unless it is explicitly extended (typically by chorded shift + or command + <click>s, which toggle contiguous or picked objects’ membership in the selection set). For our application’s mouse- and cursor-less interface, a simple postfix grammar, shown in Figure 3, was developed for keypad entry, used to toggle avatars into and out of the selection set as a prelude for invoking some operation on them (motion: rotation or translation; attribute: set or reset). Various keys corresponding to the attribute initial invoke the narrowcasting functions.

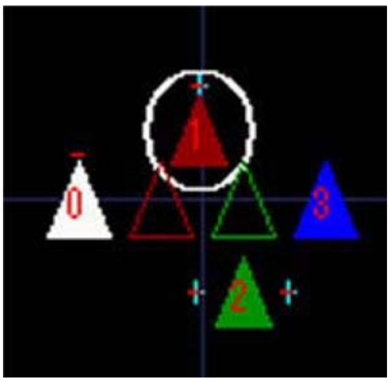
2.1 ‘ι·Con’: integrating with CVE clients

The interface is integrated with other clients through a ‘servent’ (**serv**er/cli**ent** hybrid) HTTP → TCP/IP gateway (or ‘protocol analyser’) developed with Jakarta Tomcat. Through the servent, via an Apache server, the mobile interface is integrated with all the other kinds of clients. Data communication between the ‘ι·Con’ and servent is made via HTTP request. As the DoCoMo phone is capable only of client pull (no server push), exchanges are initiated from ι·Con, regardless of whether the request is to send or receive updates. The servent performs as a client in the groupware suite, receiving data from ι·Con and passing it along to the server for multicasting to other clients. Updates which the servent receives from other clients via the server is stored (only the newest data is cached) and sent to ι·Con upon receiving a request from it.

2.2 *Conference narrowcasting functions implemented on mobile device*

The selection functions are invoked by the key corresponding to the attribute initial, as shown in Table 2. The interface can be used to control position, sensitivity, and audibility of avatars in a groupware session. Quasi-realtime synchronisation with a CVE server motivates the use of 'ghost icons', shown as outlines, to distinguish local and session states of avatars. The teleconferencing selection attributes' graphical displays are triply encoded – by position (before the 'mouth' for `mute` and `select`, straddling the 'ears' for `deafen` and `attend`), symbol ('+' for assert and '–' for inhibit, as shown in Table 1 and Figure 4) and colour (green for assert and red for inhibit). The narrowcasting attributes are not mutually exclusive, and the encoding dimensions are orthogonal (colouring, for example, the cross bar of a plus sign red even while its vertical bar is green). Because a source or a sink is active by default, invoking `exclude` and `include` operations simultaneously on an object results in its being disabled. For instance, a sink might be first `attend`ed, perhaps as a member of some non-singleton subset of a space's sinks, then later `deafen`ed, so that both attributes are simultaneously applied. (As audibility is assumed to be a revocable privilege, such a seemingly conflicted attribute state disables the respective sink, whose attention would be restored upon resetting its `deafen` flag.) Symmetrically, a source might be `select`ed then `mute`d, akin to making a 'short list' but relegated to backup.

Figure 4 'ι·Con' J2ME dynamic map for (NTT DoCoMo) iαppli mobile phone: featuring selectable icons in a '2.5D' application, in this example, #0 is `mute`d; #1 is simultaneously `mute`d and `select`ed and also selected for rotation (as indicated by its 'halo'); and #2 is simultaneously `attend`ed and `deafen`ed (see online version for colours)



2.3 *Multipresence*

Humans are indivisible, so a person cannot be in two places or more at the same time. However, 'being able to be multipresent' enables us to overcome some fundamental constraints of this human condition. Our virtual environment interface encourages multipresence (Christein and Schulthess, 2002), by supporting sources and sinks in multiple places simultaneously, allowing, for instance, a user to monitor several spaces at once. A simple teleconferencing configuration typically consists of several icons, representing the distributed users, moving around a shared conference space. Our system allows users to have multiple avatars designated as self, effectively increasing their attendance in a conference, as humorously suggested in Figure 5. Such a feature might be used to pay close attention to multiple sources, even if those sources are not repositionable; just as in ordinary settings, social conventions might inhibit dragging someone else around a shared space.

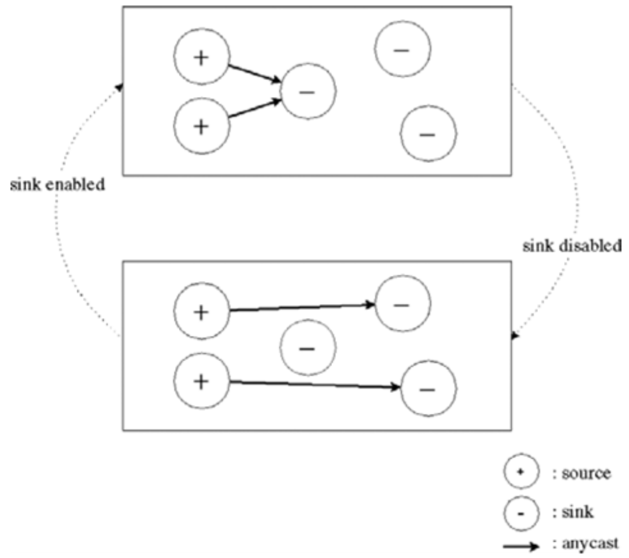
Figure 5 "I copy and paste but nothing happens!" (see online version for colours)

2.4 *Anycasting and autofocusing*

'Anycast' is a transmission between a single sender and one of possibly several receivers in a network. The term exists in contradistinction to 'multicast', transmission between a single sender and multiple receivers, and 'unicast', transmission between a single sender and a predetermined receiver. An anycasting service uses some criteria to choose a 'best' or single destination from a set of candidates. We applied the same idea, finding the best sink (closest one in a simple model) for each source in our virtual environment using an 'autofocus' technique as illustrated by Figure 6.

The apparent paradoxes of one's being in multiple places simultaneously are resolved by partitioning the sources across the sinks. If the sinks were in separate conference rooms, each source would be localised only with respect to the sink in the same room. If multiple sinks share a single conference room, the autofocus algorithm is employed by anticipating "the rule of the first wavefront", the tendency to perceive multiple identical simultaneous sources from different locations as a single fused source. Rather than adding and averaging the contribution of each source to possibly multiple sinks, our system renders each source with respect to only its best sink.

Figure 6 Anycast source –> sink transmissions: if an attending sink is `deafen`ed (or peers `confided` in), remaining sinks adopt orphaned sources



3 Scenarios

A simple teleconferencing configuration typically consists of several icons, representing the distributed users, moving around a shared conference space. Each of these icons represents a source, the voice of the associated user, as well as a sink, that user's ears. However, ι·Con allows users to have multiple sinks designated, effectively increasing their attendance in the conference. One could pay close attention to multiple instruments in a concert without rearranging the ensemble. One could leave a pair of ears in one conference, while sending another pair to a side caucus, even if the side caucus happens to be in the same room.

For example, in a chatspace application, a user might choose to designate two avatars as 'self', one to stand near an avatar corresponding to the user's mate, another, perhaps on the other side of the room, to stand near an avatar corresponding to the user's friend. Each of these avatars enjoys a 'local' perspective, a situation awareness encompassing where the respective conversationalists are relative to the (no longer unique) self-associated avatar, as manifested visually and audibly.

4 Complementary research

We use this interface to control multimodal groupware, including spatial audio applications integrated with panoramic stereographic browsing (Bolhassan et al., 2004). Such capability recalls the oft-aspired mission to build a "remote control for your life". We hope to eventually develop integrated teleconferencing with spatial audio (Yasuda et al., 2003) via such a mobile phone (`www.java.sun.com/product/jtapi`) with full CTI (Computer-Telephone Integration) (Roberts, 1999), but unfortunately voice communication is currently disabled during such iαppli sessions. Using our mobile networked narrowcasting interface, users will be able to control the spatialised audio (and other realtime media streams) of inevitable multiparty chatspaces, using the cocktail party effect as well as selection to make useful sense of the cacophonies, as imagined by Figure 7.

Figure 7 Social mute. (©2003 The New Yorker Collection from cartoonbank.com. All rights reserved)

"With your kind permission, I've taken the liberty of putting Marvin on 'mute.'"

Ongoing complimentary research in our group is exploring techniques for multi-windowing on mobile devices, which capability will require and amplify the multipresence-capable selection features described here, multiple avatars associated with a single human user distributed across multiple spaces. For this requirement, we can use Time-Division Multiplexing (TDM) or Space-Division Multiplexing (SDM). But the screen size of a mobile phone is too small for effective SDM, so we decided to develop TDM in our teleconferencing application for multipresence.

Anticipated windowed virtual reality mobile phone interfaces will allow teleport (`cut/paste`) and cloning (`copy/paste`) operations. For instance, a user might instantiate several avatars or 'fork presence' in spaces corresponding to music (virtual concert), home (chatspace intercom), and work or school (teleconferences), using the selection functions described here to multiplex and mix such soundscapes.

5 Conclusion

The basic goal of this research is to develop idioms for privacy and selective attention, narrowcasting for groupware applications, whether the interface is via workstation or a nomadic device like a mobile phone. This research can be considered an extension of presence technology (Vaughan-Nichols, 2003). A multi-presence feature using these idioms encourages users to install avatar representatives of themselves in several places and spaces at once. Activity or information in a space might cause the user to focus on that particular soundscape, using these narrowcasting functions. As illustrated in Figure 8, being anywhere is better than being everywhere, since it is selective; multi-presence is distilled ubiquity, narrowcasting-enabled audition (for sinks) or address (for sources) of multiple objects of regard.

Figure 8 Multi-presence is more convenient than omni-presence. (©2003 The New Yorker Collection from cartoonbank.com. All rights reserved)

"Sorry, I can't—I have to be everywhere."

References

ASCII editing group (2001) *iMode Java Programming*, ASCII, Tokyo, ISBN 4-7561-3727-X.

Benford, S., Bowers, J., Fahl'en, L., Greenhalgh, C., Mariani, J. and Rodde, T. (1995) 'Networked virtual reality and cooperative work', *Presence*, Vol. 4, No. 4, pp.364–386, ISSN 1054-7460.

Benford, S., Greenhalgh, C., Reynard, G., Brown, C. and Koleva, B. (1998) 'Understanding and constructing shared spaces with mixed-reality boundaries', *ACM Trans. on Computer-Human Interaction*, Vol. 5, No. 3, September, pp.185–223.

Bolhassan, A., Cohen, M., and Martens, W.L. (2004) 'A multiuser multiperspective panorama and turnorama browser using quick-TimeVR and java featuring multimonitor and stereographic display', *TVRSJ: Trans. Virtual Reality Society of Japan*, Vol. 9, No. 1, pp.69–78, ISSN 1344-011X.

Bolhassan, N.A., Cohen, M., Fernando, O.N.N., Kamada, T., Martens, W.L., Osaka, H. and Yoshikawa, T. (2002) 'Just look at yourself!: stereographic exocentric visualization and emulation of stereographic panoramic dollying', *Proc. ICAT: Int. Conf. on Artificial Reality and Tele-Existence*, University of Tokyo, December, pp.146–153, vrsj.t.u-tokyo.ac.jp/ic-at/02146.pdf

Christein, H. and Schulthess, P. (2002) 'A general purpose model for presence awareness', *Proc. Int. Conf. on Distributed Communities on the Web*, Sydney, pp.24–34.

Churchill, E.F., Snowdon, D.N. and Munro, A.J. (Eds.) (2001) *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*, ser. Computer Supported Cooperative Work (CSCW), Springer, ISBN 1-85233-244-1.

Cohen, M. (2000) 'Exclude and include for audio sources and sinks: Analogs of mute and solo are deafen and attend', *Presence: Teleoperators and Virtual Environments*, Vol. 9, No. 1, February, pp.84–96, ISSN 1054-7460, www.u-aizu.ac.jp/~mcohen/welcome/publications/ie1.pdf

Cohen, M. and Kawaguchi, M. (2003) 'Narrowcasting operations for mobile phone CVE chatspace avatars', *Proc. ICAD: Int. Conf. on Auditory Display*, in Brazil, E. and Shinn-Cunningham, B. (Eds.), Boston, July, pp.136–139, www.icad.org/websiteV2.0/Confer-ences/ICAD2003/paper/33Cohen.pdf

Cohen, M. and Martens, W.L. (2001) 'Spatial media research at the university of Aizu', *JVRSJ: J. Virtual Reality Society of Japan*, Vol. 6, No. 2, September, pp.52–57, ISSN 1342-6680; www.u-aizu.ac.jp/~mcohen/welcome/publications/spatial-media.ps

Cohen, M., Azumi, T., Yatsuyanagi, Y., Sasaki, M., Yamaoka, S. and Takeichi, O. (2002) 'Networked speaker array streaming back to client: the world's most expensive sound spatializer?', *Proc. ICAT: Int. Conf. on Artificial Reality and Tele-Existence*, Tokyo, December, pp.162–169, vrsj.t.u-tokyo.ac.jp/ic-at/papers/02162.pdf

Feng, Y. and Zhu, J. (2001) *Wireless Java Programming with J2ME*. Sams, Beijing, ISBN 0672321351.

Fernando, O.N.N., Adachi, K. and Cohen, M. (2004) 'Phantom sources for separation of listening and viewing positions for multipresent avatars in narrowcasting collaborative virtual environments', *Proc. MNSA: Int. Wkshp. on Multimedia Network Systems and Applications (in conjunction with ICDCS: 24th Int. Conf. on Distributed Computing Systems)*, Tokyo University of Technology, Hachioji, Tokyo, March, pp.170–175, ISBN 0-7695-2087-1.

Fernando, O.N.N., Adachi, K., Duminduwardena, U.C., Kawaguchi, M. and Cohen, M. (2003) 'Audio narrowcasting for multipresent avatars on workstations and mobile phones', *Proc. ICAT: Thirteenth Int. Conf. on Artificial Reality and Tele-Existence*, Keio University, Tokyo, December, pp.106–113, ISSN 1345-1278.

Fernando, O.N.N., Cohen, M., Bolhassan, N.A., Wanasinghe, D. and Kanno, T. (2002) 'Mobile control in cyberspace of image-based and computer graphic scenes and spatial audio using stereo QTVR and Java3D', *3D Forum: J. of Three Dimensional Images*, Vol. 16, No. 4, December, pp.101–106, ISSN 1342-2189.

Gordon, R. and Talley, S. (1999) *Essential JMF – Java Media Framework*, Prentice-Hall, Upper Saddle River, NJ, USA, ISBN 0130801046.

Greenhalgh, C. and Benford, S. (1995) 'Massive: A collaborative virtual environment for teleconferencing', *ACM Trans. on Computer-Human Interaction*, Vol. 2, No. 3, September, pp. 239–261.

Kanno, T. and Cohen, M. (2002) 'An architecture for collaborative virtual environments', *3D Forum: J. Three Dimensional Images*, Vol. 16, No. 1, March, pp.166–174, ISSN 1342-2189.

Kanno, T., Cohen, M., Nagashima, Y. and Hoshino, T. (2001) 'Mobile control of multimodal groupware in a distributed virtual environment', in Tachi, S., Hirose, M., Nakatsu, R. and Takemura, H. (Eds.): *Proc. ICAT: Int. Conf. on Artificial Reality and Tele-Existence*, University of Tokyo, Tokyo, December, pp.147–154, ISSN 1345-1278 www.pi.titech.ac.jp/~hase/ICATPHP/upload/39camera.pdf; vrsj.t.u-tokyo.ac.jp/ic-at/papers/01147.pdf

Knudsen, J. (2001) *Wireless Java: Developing with Java 2, Micro Edition*, Apress, Berkeley, CA, USA, ISBN 1-893115-50-X.

Kolbert, E. (2001) 'Pimps and dragons: How an online world survived a social breakdown', *The New Yorker*, 28 May, pp.88–98.

Kontio, M. (2003) *Mobile Java with J2ME*, IT Press, Finland, ISBN 951-826-554-2.

Maremaa, T. and Stewart, W. (1999) *QuickTime for Java: A Developer Reference*, Academic Press, Morgan Kaufman, ISBN 0-12-305440-0.

Nagashima, Y. and Cohen, M. (2001) 'Distributed virtual environment interface for a mobile phone', *3D Forum: J. of Three Dimensional Images*, Vol. 15, No. 4, December, pp.102–106, ISSN 1342-2189.

Okoshi, T., Wakayama, S., Sugita, Y., Iwamoto, T., Nakazawa, J., Nagata, T., Furusaka, D., Iwai, M., Kusumoto, A., Harshima, N., Yura, J., Nishio, N., Tobe, Y., Ikeda, Y. and Tokuda, H. (2001) 'Smart space laboratory project: toward the next generation computing environment', *Proc. IEEE Third Int. Workshop on Networked Appliances*, March, Singapore, pp.115–121.

Palmer, I. (2001) *Essential Java 3D Fast: Developing 3D Graphics Applications*, Springer, London, ISBN 1-85233-394-4.

Roberts, S. (1999) *Essential* JTAPI*: Java Telephony*, Prentice-Hall, Upper Saddle River, NJ, USA, ISBN 0-13-080360-X.

Sowizral, H., Rushforth, K. and Deering, M. (2000) *The Java 3D API Specifi-cation*, 2nd ed., Addison-Wesley, Reading, MA, USA, ISBN 0-201-71041-2.

Topley, K. (2002) *J2ME in a Nutshell*, O'Reilly, Sebastopol, CA, USA, ISBN 0-596-00253-X.

Vacca, J.R. (2002) *I-Mode Crash Course*, McGraw-Hill, New York, ISBN 0-07-138187-2.

Vaughan-Nichols, S.J. (2003) 'Presence technology: More than just instant messaging', *(IEEE) Computer*, Vol. 36, No. 10, October, pp.11–13, ISSN 0018-9162.

Walrath, K. and Campione, M. (1999) *The JFC Swing Tutorial*. Addison-Wesley, Reading, MA, USA, ISBN 0-201-43321-4.

Walsh, A.E. and Gehringer, D. (2002) *Java 3D API Jump-start*, Prentice-Hall, Upper Saddle River, NJ, USA, ISBN 0-13-034076-6.

Yamazaki, Y. (2001) *How to Make An iAppli*, SOFTBANK, ISBN 4-7973-1573-3.

Yasuda, Y., Ohya, T., McGrath, D. and Flanagan, P. (2003) '3-D audio communications services for future mobile networks', *AES 23rd Int. Conf.*, May, Helsingør, Denmark.