ICAT 2001
December 5-7, Tokyo, Japan

Mobile Control of Multimodal Groupware in a Collaborative Virtual Environment

Toshifumi Kanno, Michael Cohen, Yutaka Nagashima, and Tomohisa Hoshino

Spatial Media Group, University of Aizu; Aizu-Wakamatsu 965-8580; Japan
toshi_kanno@mac.com, mcohen@u-aizu.ac.jp, ytheater@palette.plala.or.jp, hoshino@aizu.com

Abstract

Anticipating ubicomp (for **ubi**quitous **comp**uting) networked applications and information spaces, we have integrated various multimodal (auditory, visual, haptic) I/O devices into a virtual reality groupware system. We have deployed a Java-equipped mobile phone capable of interacting with this virtual environment groupware suite, interfaced through a "servent," a **ser**ver/cli**ent** hybrid HTTP ↔ TCP/IP gateway.

Key words: mobile computing, CVE (**c**ollaborative **v**irtual **e**nvironments), groupware, CSCW (**c**omputer-**s**upported **c**ollaborative **w**ork), hand-held interface.

1. Multimodal Groupware Architecture

We have designed and implemented an architecture and framework [MSC01] [KC01b] [KCNH01] to support a **c**ollaborative **v**irtual **e**nvironment (CVE), allowing distributed users to share multimodal virtual worlds. Our CVE architecture is based upon a **c**lient/**s**erver model, and its main transaction shares the state of virtual objects and users (avatars) by N-unicast of position parameters (translation, rotation, and zoom) to client peers in a session. There is no server caching of state and changed parameters are immediately redistributed. The main features of our CVE clients are multimodal communication, platform independence, and easy network connectivity, as components are built with Java (and JMF [**J**ava **M**edia **F**ramework[1]] [GT99], including Java3D[2] [SRD97] [Pal01] [Bar01] [WG02], J2ME and Swing[3]). CVE client components implement connection to the server based on a unified protocol, and are easily linked to Java-based interfaces. The protocol has a simple interface, providing C/S communication methods like `get()` and `set()`, as shown in Fig. 1, so client developers need not concern themselves with network communication issues. Easily attachable components enable connection between various kinds of heterogeneous clients, as shown in Fig. 2, so the integrated client suite can display and manipulate various multimodal information through the internet. There are several well-developed protocols supporting VR-style groupware [SZ99] [CSM01], so the novelty of our system lies in its support for heterogeneous clients; our framework emphasizes breadth of client capability.

The CVE suite integrated by this framework includes these clients (listed in clockwise order of appearance in Fig. 2 below, starting from the bottom-left) [CM01] in various stages of development and integration:

- Internet Chair, sensing and driving azimuth of a swivel seat, with force-feedback via servomotor [Coh99] [Coh00] [KCA00] [Coh01], including display of transaural audio through nearphones, (for "near earphones") mounted straddling the chair's headrest, presenting unencumbered binaural sound with soundscape stabilization for multichannel sound image localization;
- Helical Keyboard [KC01a] [NC01b] (shown in bottom left of Fig. 3), to visualize the helical structure of a musical scale, animated in realtime by a GUI or MIDI controllers or events (sequencers);
- "Hero" mobile **he**aring tele**ro**bot,[4] autonomous or piloted, with camera and 4 microphones [HOS97] [HST+99] [YCHY01];
- Java3D models— including a robot (Fig. 4), vehicle (Fig. 5), and Internet Chair (Fig. 6)— featuring exocentric, egocentric, and stereoscopic perspective displays and control widgets;
- VR_4U_2C QTVR[5] (QuickTime Virtual Reality for panoramic photographs) browser [BMC01] (bottom right of Fig. 3), for multi-window, multi-monitor, or stereographic panoramic displays;
- PSFC (**P**ioneer **S**ound **F**ield **C**ontroller[6] in the University of Aizu Multimedia Center's Synthetic World Zone [3D Theater,[7] shown in Fig. 7]) [AMY+98] proxy [HMC01], controlling spatialization of audio sources through a DSP-driven hemispherical speaker array;
- **S**piral **S**pring (**S**wivel **S**eat) **S**oundscape-**S**tabilized GUI [CS00] (top right of Fig. 3), modeling the Internet Chair

[1] java.sun.com/products/java-media/jmf/
[2] java.sun.com/products/java-media/3D/
[3] java.sun.com/products/jfc/tsc/
[4] www.u-aizu.ac.jp/~j-huang/Robotics/robotics.html
[5] www.apple.com/quicktime/products/qt/overview/qtvr.html
[6] www.u-aizu.ac.jp/~mcohen/spatial-media/PSFC/
[7] www.mmc-aizu.pref.fukushima.jp/mmc/system/sys7.html

Client Application	
CVEClient	CVEClientIF Client application (conforms to protocol and) implements this interface (abstract superclass).
`setPosition`(x, y, z, roll, pitch, yaw) `setLocation`(x, y, z) `setOrientation`(roll, pitch, yaw) `setExtraParam`(name, value)	`getPosition`(x, y, z, roll, pitch, yaw) `getLocation`(x, y, z) `getOrientation`(roll, pitch, yaw) `getExtraParam`(name, value)

Fig. 1: Client Implementation

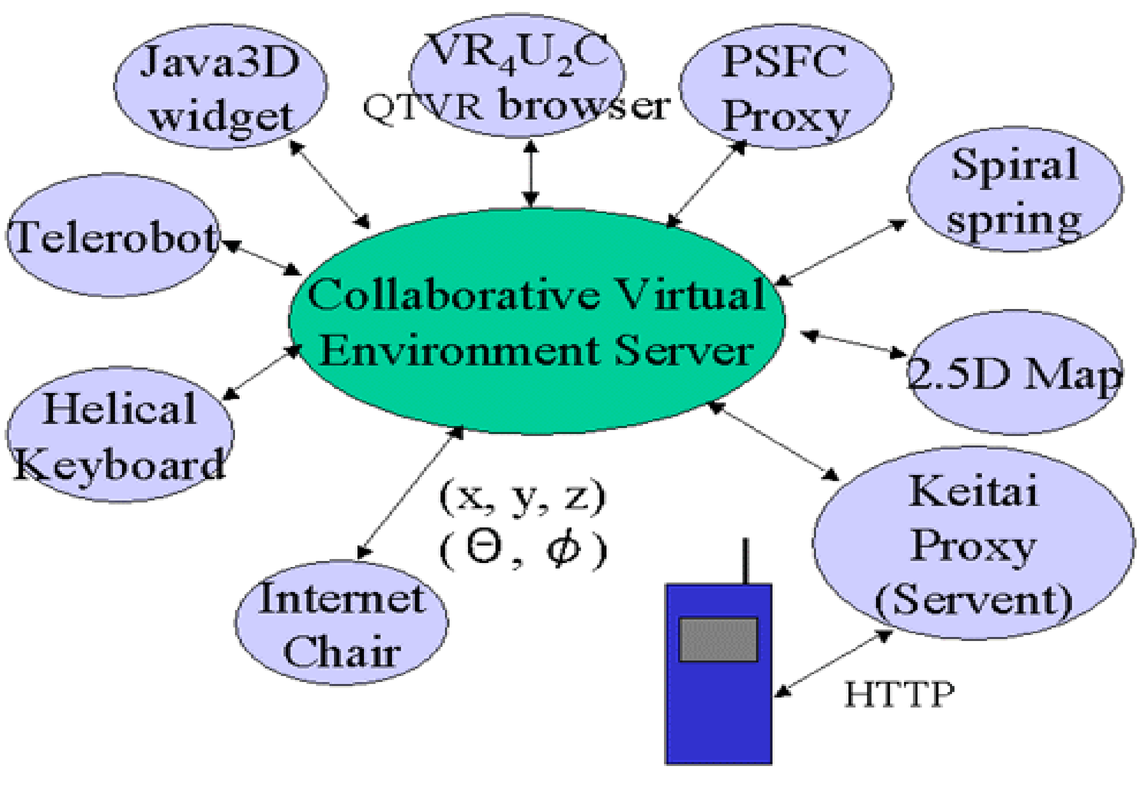

Fig. 2: CVE architecture: groupware suite

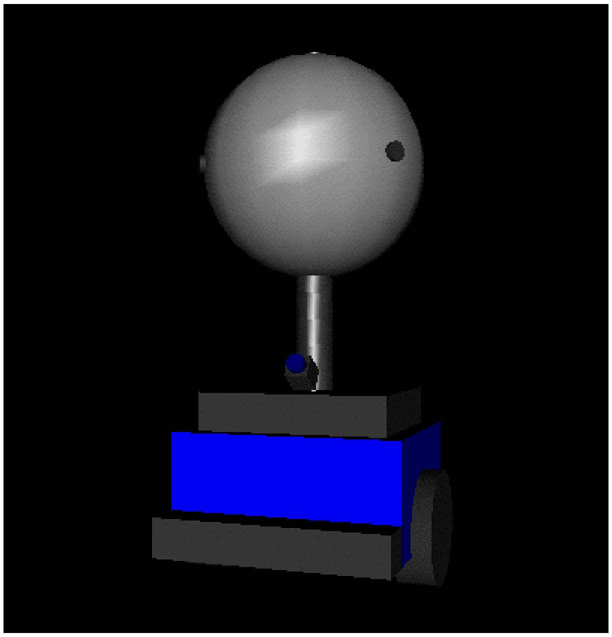

Fig. 4: Java3D Telerobot Model

with chromastereoptic visualization, and capable of directionalizing virtual sound sources, recorded or captured in realtime;

- 2.5D dynamic map [MSC01] (top left of Fig. 3), displaying and controlling iconic planar translation and rotation; and
- mobile phone (NTT DoCoMo iAppli) proxy, elaborated following.

2. Mobile Telephone Interface

We have designed and implemented a mobile telephone interface [NC01a] for use in our CVE. Programmed with J2ME (Java 2 Platform, Micro Edition,[8]) a dynamic map application runs on an (NTT DoCoMo) iAppli mobile phone ("*keitai*": 携帯 電話), as shown in Fig. 8. Featuring selectable icons with one rotational and two translational degrees of freedom, like the 2.5D map shown in Fig. 3, the interface is used to control avatars in a chatspace. The Sony model of the 503-series iAppli units features a thumb jog wheel, which can be used as a continuous controller to manipulate such icons in a teleconference. The user interface is further extended with musical and vibrational cues. We hope to eventually deploy full teleconferencing with spatial audio via such a mobile phone[9] [Rob99] for full CTI (**c**omputer-**t**elephone **i**ntegration), but unfortunately voice communication is currently disabled during such iAppli sessions, so a second phone must be used for teleconferencing.

2.1 Integrating with CVE clients

The interface is integrated with other CVE clients through a "servent" (**serv**er/cli**ent** hybrid) HTTP↔TCP/IP gate-

[8] java.sun.com/j2me/

[9] java.sun.com/products/jtapi/

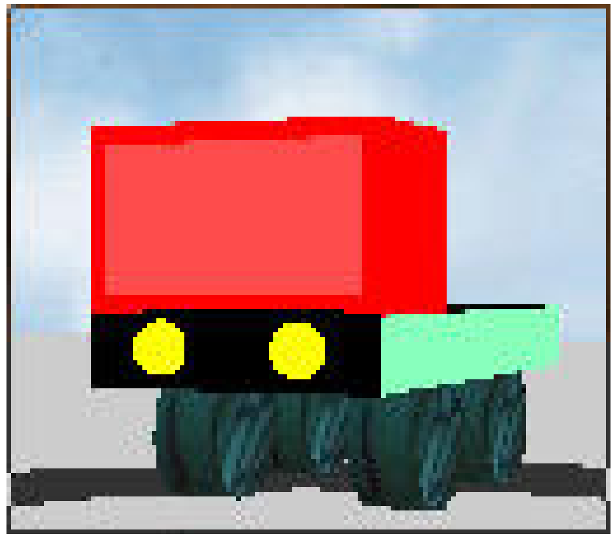

Fig. 5: Java3D Vehicle Model

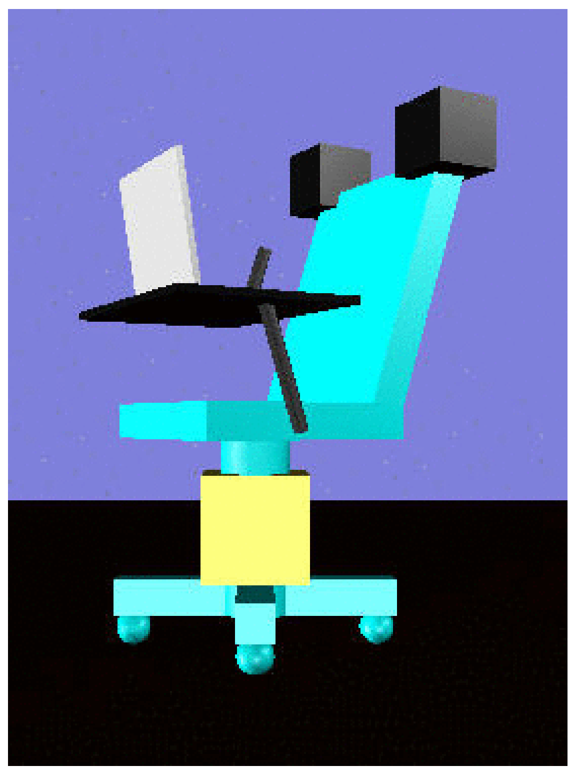

Fig. 6: Java3D Internet Chair Model

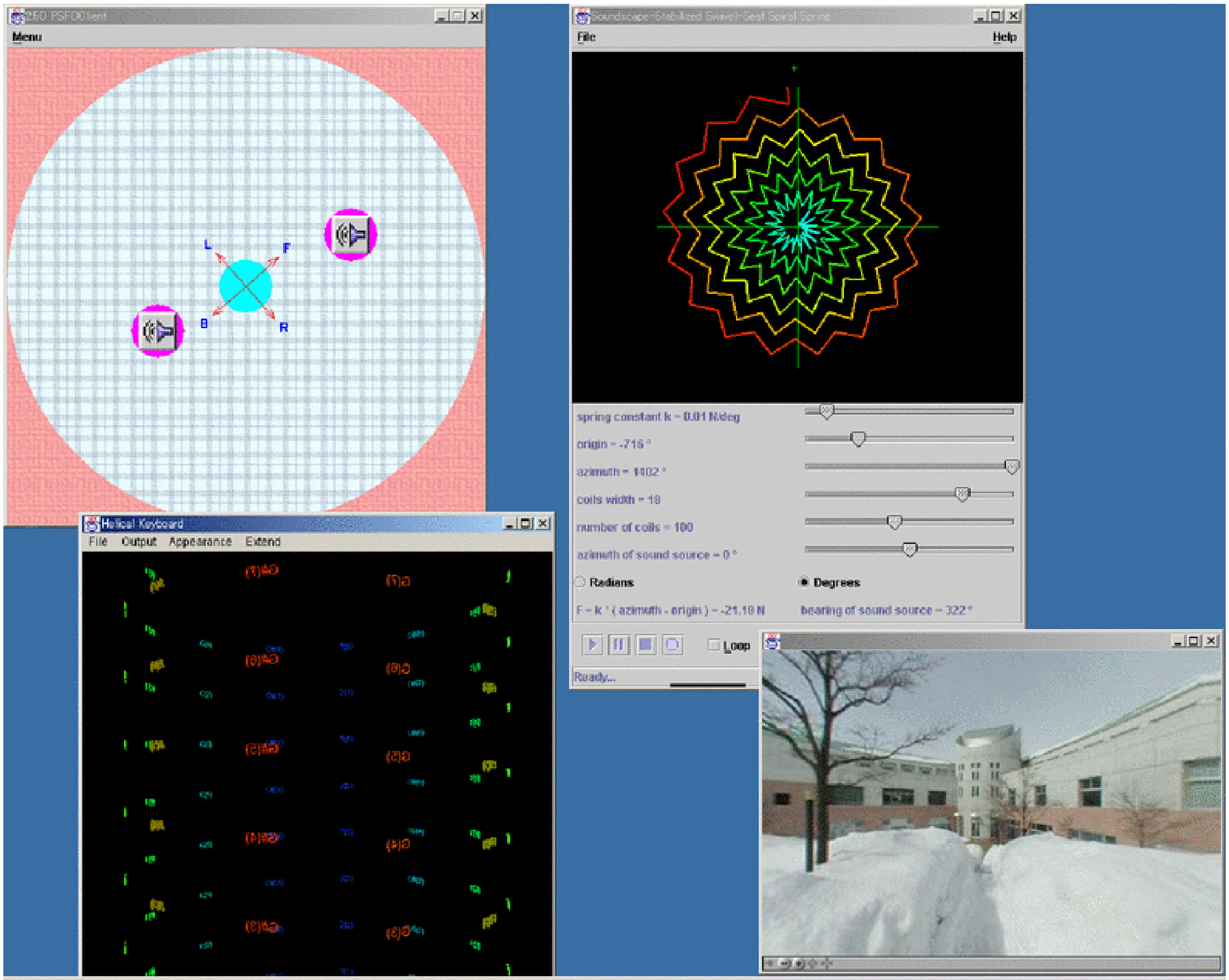

Fig. 3: Multiclient session snapshot of one user's display of **c**ollaborative **v**irtual **e**nvironment, including (counter-clockwise from top right): soundscape-stabilized (swivel-seat) spiral spring, which can directionalize (lateralize) a mixel in near-realtime from a resident file or captured audio stream; 2.5D dynamic map, in which icons can translate and rotate; Helical Keyboard, and VR_4U_2C QTVR browser, with multiwindow, multimonitor, or stereographic panoramic display. All of these clients interoperate as groupware, synchronizable locally or over the internet.

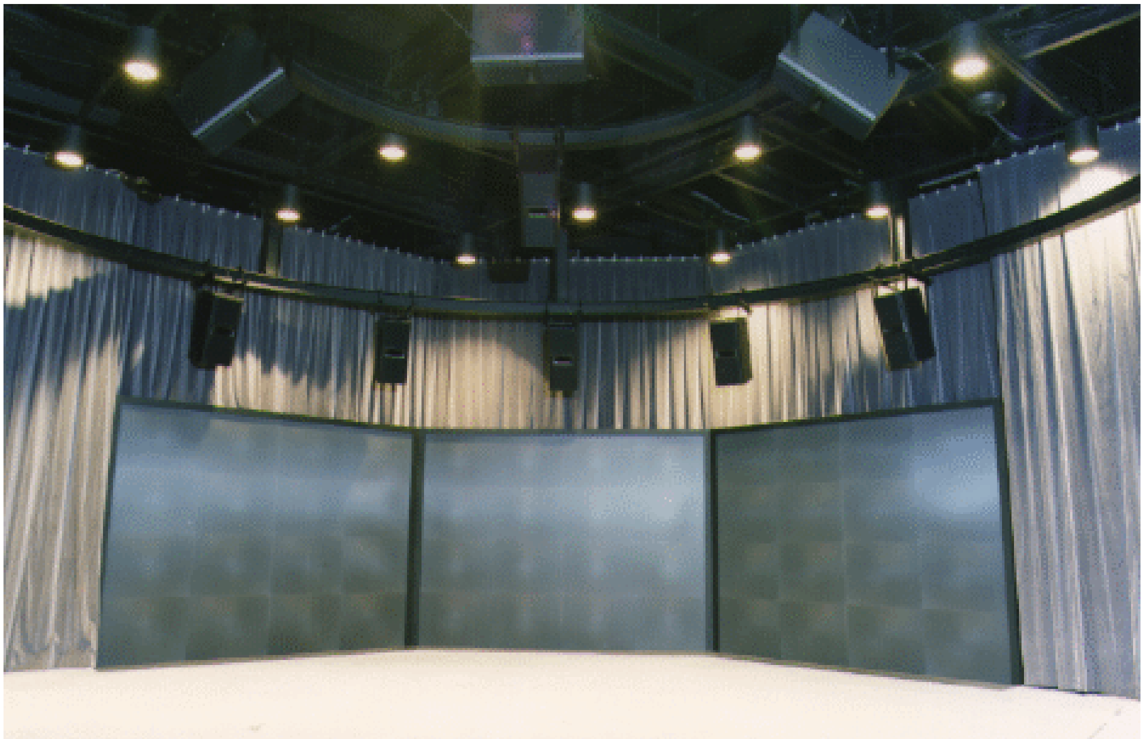

Fig. 7: Synthetic World Zone at the University of Aizu Multimedia Center, featuring PSFC

way (or "protocol analyzer") developed with Jakarta Tomcat.[10] Through the servent's servelets, via an Appache server, the mobile phone is integrated with our heterogeneous groupware suite, interacting with all the other kinds of clients.

Data communication between the "ιcon" iAppli mobile phone interface and servent is made via HTTP request. As the mobile phone is capable only of client pull (and no server push), exchanges are initiated from the mobile phone, regardless of whether the request is to send or receive updates. The servent works as client in the groupware suite, receiving data from the mobile phone 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 the mobile phone upon receiving a request from it.

Fig. 8: Jade emulator showing "ιcon" iAppli interface

3. Future Projects

Eventually mobile telephony will feature high-fidelity stereo audio capability, as suggested by Fig. 9. We also anticipate leveraging our system against emerging internet standards XML and MPEG-4[11] (for object-oriented multimedia, including audio streaming and spatialization). We are especially interested in broadband mixed reality/virtuality systems [MK94] (also known as, or related to, annotated, augmented, enhanced, hybrid, mediated,

[10] jakarta.apache.org
[11] www.cselt.it/mpeg

or virtualized reality systems), which blur sampled and synthesized data, especially realtime media streams.

Fig. 9: "Poor person's mobile stereotelephony:" two mobile phones deployed as a microphone array attached to a dummy head simultaneously calling a dual voice line (like that provided by ISDN) realizes wireless (if still low-fidelity) binaural telepresence.

4. Conclusion

We are now evaluating the usefulness and potential of our CVE, as extended by mobile interaction. Test applications will include multimedia chat spaces (social) and conferencing (business), as well as VR-style distance learning, gaming, and music. Our system is **multi**user (supporting multiple simultaneous users in realtime interactivity), **multi**media (driving graphical, auditory [stereo, transaural, and speaker-array spatial audio], musical, and video displays), **multi**modal (providing visual [WIMP/GUI plus perspective representation], auditory, haptic and force-feedback interaction), and **multi**perspective (including orthographic [2.5D map], stereographic panoramic [QTVR], and perspective [Java3D] visual displays).

5. Acknowledgments

The VR_4U_2C QTVR browser was developed by Noor Alamshah Bolhassan, the 2.5D dynamic map by Masataka Shimizu (清水 雅高), and the spiral spring client by Kenta Sasa (佐々 健太) and Shiyougo Ihara (伊原 正悟) with Takashi Wada (和田 貴志). The Helical Keyboard was developed by first author Toshifumi Kanno (菅野 才文) and Kazuhisa Nagashima (中島 一久). The Java3D models were developed by Shiyuunou Kazuki (収納 和樹) [telerobot], Shimizu Hideto (清水 英) [vehicle] and Kaneko Daisuke (金子 大輔) [Internet Chair]. The HTTP↔TCP/IP servent was developed at strategic partner "Eyes, Japan"[12] (あいづ・ジャパン). The Internet Chair is being developed in conjunction with Yamagata Daigaku (山形大学) and Mechtec[13] (メカテック), an industrial partner based in Kita-Kata (喜多方). This research has been supported by a grant from the Fukushima Prefectural Foundation for the Advancement of Science and Education.

References

[AMY+98] Katsumi Amano, Fumio Matsushita, Hirofumi Yanagawa, Michael Cohen, Jens Herder, William Martens, Yoshiharu Koba, and Mikio Tohyama. A Virtual Reality Sound System Using Room-Related Transfer Functions Delivered Through a Multispeaker Array: the PSFC at the University of Aizu Multimedia Center. *TVRSJ: **T**rans. **V**irtual **R**eality **S**ociety of **J**apan*, 3(1):1–12, March 1998. ISSN 1344-011X; `www.u-aizu.ac.jp/~mcohen/welcome/publications/PSFC.ps`.

[Bar01] Jon Barrilleaux. *3D User Interfaces with Java 3D*. Manning Publications, 2001. ISBN 1-88477-790-2.

[BMC01] Noor Alamshah Bolhassan, William L.

[12] `www.aizu.com`
[13] `www.mechtec.co.jp`

Martens, and Michael Cohen. VR$_4$U$_2$C: A Multiuser Multiperspective Panoramic Browser Using QuickTime VR and Java Featuring Multimonitor and Stereographic Display. In *Proc. ICAT: **I**nt. **C**onf. **A**rtificial Reality and **T**ele-Existence*, Tokyo, December 2001.

[CM01] Michael Cohen and William L. Martens. Spatial Media Research at the University of Aizu. *JVRSJ: **J**. **V**irtual **R**eality **S**ociety of **J**apan*, 6(2):52–57, September 2001. ISSN 1342 6680; www.u-aizu.ac.jp/~mcohen/welcome/publications/spatial-media.ps.

[Coh99] Michael Cohen. The Internet Chair. In *Proc. ICAT: **I**nt. **C**onf. **A**rtificial Reality and **T**ele-Existence*, pages 29–36, Tokyo, December 1999. VRSJ. www.u-aizu.ac.jp/~mcohen/welcome/publications/ic3.ps.

[Coh00] Michael Cohen. A Design for Integrating the Internet Chair and a Telerobot. In *Proc. IS2000: Int. Conf. on Information Society in the 21st Century*, pages 276–280, Aizu-Wakamatsu, Japan, November 2000. IPSJ, IEICE, IEEE.

[Coh01] Michael Cohen. The Internet Chair. *IJHCI: Int. J. of Human-Computer Interaction*, 2001. www.u-aizu.ac.jp/~mcohen/welcome/publications/ic4.ps. In press.

[CS00] Michael Cohen and Kenta Sasa. An interface for a soundscape-stabilized spiral-spring swivel-seat. In *Proc. WESTPRAC VII: 7th **West**ern **P**acific **R**egional **A**coustics **C**onf.*, pages 321–324, Kumamoto, Japan, October 2000. ISBN 4-9980886-1-0 and 4-9980886-3-7.

[CSM01] Elizabeth F. Churchill, David N. Snowdon, and Alan J. Munro, editors. *Collaborative Virtual Environments: Digital Places and Spaces for Interaction.* Computer Supported Cooperative Work (CSCW). Springer, 2001. ISBN 1-85233-244-1.

[GT99] Rob Gordon and Stephen Talley. *Essential JMF — Java Media Framework.* Prentice Hall, 1999. ISBN 0130801046.

[HMC01] Kuniaki Honno, William L. Martens, and Michael Cohen. Psychophysically-derived control of source range for the Pioneer Sound Field Controller. In *Proc. AES: **A**udio **E**ngineering **S**ociety Conv. (110th Convention)*, Amsterdam, May 2001. Preprint #5313; www.u-aizu.ac.jp/~wlm/papers/aes110_98.ps.

[HOS97] Jie Huang, Noboru Ohnishi, and Noboru Sugie. Building ears for robots: Sound localization and separation. *Artificial Life and Robotics (Springer-Verlag)*, 1(4):157–163, 1997.

[HST+99] J. Huang, T. Supaongprapa, I. Terakura, F. Wang, N. Ohnishi, and N. Sugie. A model based sound localization system and its application to robot navigation. *Robotics and Autonomous Systems (Elsevier Science)*, 27(4):199–209, 1999.

[KC01a] Toshifumi Kanno and Michael Cohen. A Helical Keyboard Client. In *Proc. CIT: 2nd Int. Conf. on **C**omputer and **I**nformation **T**echnology*, pages 163–165, Shanghai, September 2001. ISSN 1007-6417.

[KC01b] Toshifumi Kanno and Michael Cohen. An architecture for collaborative virtual environments. In *Proc. HC2001: 4th Int. Conf. on **H**uman and **C**omputer*, pages 31–38, Aizu-Wakamatsu, Japan, September 2001.

[KCA00] Nobuo Koizumi, Michael Cohen, and Shigeaki Aoki. Japanese patent #3042731: Voice reproduction system, March 2000.

[KCNH01] Toshifumi Kanno, Michael Cohen, Yutaka Nagashima, and Tomohisa Hoshino. Mobile control of multimodal groupware in a distributed virtual environment. In William L. Martens, editor, *Proc. IWSM: **I**nt. **W**ksp. on **S**patial **M**edia*, pages 107–111, Aizu-Wakamatsu, October 2001.

[MK94] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual dispays. IEICE *Trans. Inf. Sys.*, E77-D(12):1321–1329, December 1994.

[Mor00] Ralph Morelli. *Java, Java, Java: Object-Oriented Problem Solving.* Prentice Hall, 2000. ISBN 0130113328.

[MSC01] Takashi Mikuriya, Masataka Shimizu, and Michael Cohen. A collaborative virtual environment featuring multimodal information controlled by a dynamic map. *3D Forum: J. of Three Dimensional Images*, 15(1):133–136, 3 2001. ISSN 1342-2189.

[NC01a] Yutaka Nagashima and Michael Cohen. Distributed virtual environment interface for a mobile phone. In *Proc. HC2001: 4th Int. Conf. on **H**uman and **C**omputer*, pages 43–46, Aizu-Wakamatsu, Japan, September 2001.

[NC01b] Kazuhisa Nakashima and Michael Cohen. Animated extensions to a helical keyboard client: Chord-chords, chord-kites, and intelligent spatialization. In *Proc. HC2001: 4th Int. Conf. on **H**uman and **C**omputer*, pages 39–41, Aizu-Wakamatsu, Japan, September 2001.

[Pal01] Ian Palmer. *Essential Java 3D fast: developing 3D graphics applications.* Springer, 2001. ISBN 1-85233-394-4.

[Rob99] Spencer Roberts. *Essential* JTAPI*: Java telephony.* Prentice Hall, 1999. ISBN 0-13-080360-X.

[SRD97] Henry Sowizral, Kevin Rushforth, and Michael Deering. *The Java 3D API Specification.* Addison-Wesley, 1997. ISBN 0201325764.

[SZ99] Sandeep Sighal and Michael Zyda. *Networked Virtual Environments: Design and Implementation.* Addison-Wesley: ACM Press, 1999. ISBN 0-201-32557-8.

[WG02] Aaron E. Walsh and Doug Gehringer. *Java 3D API jump-start.* Prentice-Hall, 2002. ISBN 0-13-034076-6.

[YCHY01] Yasuhiro Yamazaki, Michael Cohen, Jie Huang, and Tomohide Yanagi. Augmented audio reality: compositing mobile telerobotic and virtual spatial audio. In *Proc. ICAT:* **I***nt.* **C***onf.* **A***rtificial Reality and* **T***ele-Existence*, Tokyo, December 2001.