

# **Audiokommunikation für Telekonferenzen**

Hartmut Chodura

Fraunhofer-Institut für Graphische Datenverarbeitung, Darmstadt

## **Zusammenfassung**

Audio ist die wichtigste Komponente in interaktiven Telekonferenzen. Dieser Beitrag fasst unsere Aktivitäten im Themengebiet verteiltes Audio für Internet-Anwendungen zusammen. Zunächst werden Faktoren beschrieben, welche sich auf die Qualität einer Audioübertragung auswirken. Ein Schwerpunkt der Untersuchung liegt auf Konferenzen mit mehr als zwei Teilnehmern. Anschliessend werden zwei Eigenentwicklungen vorgestellt, die Kommunikation mit guter Unterscheidbarkeit der Teilnehmer gewährleisten. WAU ist ein Internet-Telefon, das robust gegen Qualitätsschwankungen im Netzwerk ist. Eine erweiterte Version von WAU erlaubt Mehrbenutzerbetrieb mit räumlicher Audioausgabe. Berichtet wird über die Integration von WAU in eine virtuelle Umgebung für Telekonferenzen und die Erweiterung von Microsoft-Netmeeting für mehrere Anwender. Diese Erweiterung wird im Rahmen der ADAPT Initiative im Projekt OrgTech durchgeführt, in dem unter anderem moderne Technologien zur Gruppenarbeit in kleinen und mittleren Konstruktionsbüros eingeführt werden.

## 1 Motivation

Telekonferenzen dienen dazu, miteinander zu kommunizieren bzw. gemeinsam zu arbeiten. Die wichtigste Anforderung ist dabei eine stabile und adäquate Audioverbindung. Ein Telekonferenzsystem sollte die Zusammenarbeit und Kommunikation weder behindern noch einschränken. Ideal wäre die Gegenüber akustisch so gut wahrzunehmen wie bei einem konventionellen Treffen. Aufzeichnung, Digitalisierung, Verarbeitung, Übertragung und Wiedergabe beeinflussen jedes für sich die Gesamtqualität.

## 2 Qualitätsfaktoren für Audio

Die im Folgenden beschriebenen Entwicklungen wurden durch eine umfangreiche Analyse der Qualitätsfaktoren für Audiodaten eingeleitet. Besonderes Augenmerk lag auf der Übertragung der Daten per Internet-Protokoll (IP) (Chodura, 1998a). Das Anliegen war, mindestens die Qualität des herkömmlichen Telefons zu erreichen und diese wenn möglich zu übertreffen.

### 2.1 Allgemeine Faktoren

#### 2.1.1 Digitalisierung: Abtastung und Quantisierung

Bei der Digitalisierung wird das Audiosignal abgetastet und quantisiert. Nach dem Nyquist-Theorem ist die Hälfte der Abtastrate die höchste noch aufgelöste Frequenz (Pohlmann, 1995). Die Auflösung des Telefons beträgt ca.  $4\text{ kHz}$ , was einer Abtastrate von  $8\text{ kHz}$  entspricht. Dies reicht aus, um Sprache verständlich aufzulösen. Sprache enthält jedoch noch höhere Frequenzanteile. Eine Anhebung der Abtastrate auf  $16\text{ kHz}$  oder  $22\text{ kHz}$  bringt eine deutlich verbesserte Verständlichkeit und vor allem bessere Identifizierbarkeit des Sprechers (Blauert, 1983).

Die Quantisierung bestimmt die aufgelöste Dynamik. Wird beispielsweise mit  $8\text{ Bit}$  quantisiert, erreicht man maximal  $48\text{ dB}$  Signal-Rauschabstand. Besser ist jedoch eine Quantisierung von mindestens  $16\text{ Bit}$ <sup>1</sup> ( $96\text{ dB}$  Signal-Rauschabstand), um auch Sprache gut von Hintergrundgeräuschen trennen zu können. Je mehr Bit zur Verfügung stehen, desto einfacher sind auch Übersteuerungen zu vermeiden.

---

1. Jedes zusätzlich verwendete Bit erhöht den Signal-Rauschabstand um  $6\text{ dB}$ . Die Dynamik der akustischen Wahrnehmung des Menschen liegt bei etwa  $130\text{ dB}$ .

## 2.1.2 Kompression

Da die zur Verfügung stehenden Bandbreiten begrenzt sind, müssen Audio- und Videosignale zumeist komprimiert werden. Nutzt man beispielsweise ISDN, stehen  $64 \text{ kBit/s}$  pro B-Kanal zur Verfügung. Ein Teil davon wird vom Protokoll beansprucht. Die Leitung teilen sich Audio- und Videoströme sowie Daten, die durch verteilte Anwendungen entstehen. Fast alle gängigen Kompressionsverfahren für Audio sind verlustbehaftet, d.h. das Audiosignal ist nach der Kompression nicht vollständig rekonstruierbar. Die Kunst besteht darin, den Verlust so wenig wahrnehmbar wie möglich zu halten. Prinzipiell kann zwischen Methoden zur Kompression speziell von Sprache, oder von Audio im allgemeinen unterscheiden. Ein weit verbreiteter Standard zur Kompression von Sprache ist GSM 06.10 (Degener, 1994). Für Sprache in hoher Qualität (engl. *wideband speech*) kommen vor allem Methoden in Frage, die auch für die Kompression von Musik verwendet werden (Gersho, 1994). Das bekannte Verfahren MP3 benutzt psychoakustische Hörmodelle, um Datenraten zu reduzieren und eignet sich sehr gut für Musik (Brandenburg et.al, 1994) / (von Leitner, 2000).

## 2.1.3 Übertragung im Internet

Das Internet wurde ursprünglich nicht zur Übertragung von Datenströmen in Echtzeit entwickelt. Es kann zu Störungen in der Verbindung kommen, insbesondere im öffentlichen Internet. Die Folge sind Datenverluste, lange Laufzeiten und Schwankungen der Laufzeiten (engl. *jitter*), die massgeblich die Qualität beeinträchtigen können.

Gehen Daten verloren, ohne dass geeignet vorgebeugt oder reagiert wird, kommt es zu Unterbrechungen im Datenstrom, die bei der Wiedergabe zumindest unangenehme Effekte wie Klicken zur Folge haben. Der Ablauf einer Telekonferenz kann gestört oder unmöglich gemacht werden. Bis zu einem gewissen Grad können Mechanismen Abhilfe schaffen, die als *Forward Error Correction* (FEC) bezeichnet werden. Unter FEC versteht man Methoden, bei denen bereits seitens des Senders versucht wird, einem möglichen Fehler durch Datenverluste vorzubeugen (Albanese et.al, 1994).

Auf Verzögerungen, die durch die Übertragung entstehen, hat man normalerweise wenig Einfluss. Die Gesamtverzögerung wird nicht nur durch die Paketlaufzeit bei der Übertragung bestimmt, sondern auch durch Verarbeitungszeit bei Aufzeichnung, Wiedergabe und Kompression. Verzögerungen über  $100 \text{ ms}$  machen sich bereits bemerkbar. Spätestens bei  $300 \text{ ms}$  ist die Interaktion stark beeinträchtigt.

#### 2.1.4 Freisprechen

Sobald die wiedergegebenen Signale vom Wiedergabesystem (Lautsprecher) zurück in das Mikrophon gelangen können, entsteht auf der Gegenseite ein Echo. Das lässt sich einerseits durch den Einsatz von Kopfhörern oder Telefonhörern unterbinden. Andererseits ist auch eine aktive Echounterdrückung möglich. Dabei gibt es mehrere technische Ansätze. Der Autor hat im Rahmen eines Projektes mit R. Storn zusammengearbeitet, der verschiedenste Verfahren verglichen und eine Implementierung auf einem DSP-System vorgenommen hat (Storn, 1996).

#### 2.1.5 Reduktion des Datenvolumens

Prinzipiell lässt sich das entstehende Datenvolumen sehr einfach reduzieren, wenn senderseitig geprüft oder gemessen wird, ob der Teilnehmer überhaupt spricht. Ist dies nicht der Fall, kann auf die Übertragung der Audiodaten verzichtet werden. Probleme bei der automatischen Überprüfung entstehen lediglich bei zu niedrigem Signalpegel oder starken Hintergrundgeräuschen.

### 2.2 Anforderungen für Mehrbenutzerbetrieb

Sollen mehrere verteilt sitzende Personen an einer Telekonferenz teilnehmen, ergeben sich zusätzliche Anforderungen an das Audiosystem:

- Jeder Teilnehmer erzeugt einen Datenstrom, d.h. das Gesamtvolumen kann sehr gross werden.
- Videobilder von verschiedenen Orten können nebeneinander dargestellt werden, Audiodaten hingegen müssen geeignet gemischt werden.

#### 2.2.1 Multicasting

Um Daten an alle Teilnehmer einer Konferenz zu verteilen, können die Daten vielfältig und an alle Partner getrennt versendet werden. Das ist allerdings unökonomisch, da identische Daten mehrfach über die selbe Leitung laufen.

Beim Verteilen der Daten über sogenanntes *Multicasting* wird eine andere Strategie verfolgt (Kosiur, 1997). Jeder Teilnehmer öffnet pro Datenstrom (Audio, Video, etc.) nur eine Verbindung zu einer virtuellen Netzwerkadresse. Die Daten werden an den Stellen des Netzwerks vervielfältigt, an denen es auf Grund der Netzwerktopographie notwendig ist. So müssen bei Konferenzen, an denen mehrere Rechner eines LANs angeschlossen sind, die Daten nur einmal über den Zugang zum Internet versendet bzw. empfangen werden. Für die ordnungsgemässe Zustellung der Daten sorgt ein Server, der entweder als Software auf einem Rechner innerhalb des

LANs läuft oder direkt in den Routern arbeitet (Noll et.al., 1997). Heutzutage wird *Multicasting* noch nicht im gesamten öffentlichen Internet unterstützt. Die Technologie ist jedoch für jede Form der Datenübertragung an eine grosse Gruppe von Empfängern interessant. *Broadcasting* von multimedialen Datenströmen im Internet wird immer populärer. Daher ist zu erwarten, das Multicasting bald zur Standardtechnologie wird.

### 2.2.2 Mischen der Audiodaten

Für das Mischen von Audiodaten gibt es zwei technische Ansätze:

- Die Daten werden von einem Gerät zentral gemischt oder
- die Daten werden getrennt an alle Teilnehmer verteilt und dort gemischt.

Für die erste Lösung werden unter anderem sogenannte *Media Streams Processor Units* (MPUs) verwendet, die alle Datenströme der Teilnehmer empfangen, mischen und anschliessend verteilen.

Im zweiten Fall können die Daten von jedem Teilnehmer direkt an alle Partner gesendet werden. Häufig werden die Teilnehmer unterschiedlich laut empfunden, was nicht nur an Zeitgenossen mit kraftvoller Stimme, sondern vor allem an unterschiedlichen Audiokomponenten und Pegeln liegt. Kommen die Audiodaten beim Empfänger nach Quellen getrennt an, kann die jeweilige Intensität individuell gemischt werden.

### 2.2.3 Mischen mit räumlicher Ausgabe

Ein häufiges Problem bei Mehrbenutzer-Konferenzen ist die mangelnde akustische Differenzierung der Teilnehmer (Mark et.al., 1999). Dies kann an reduziertem Dynamikumfang und eingeschränkter Frequenzauflösung der Signale liegen (vgl. Abschnitt 2.1.1). Weitere Faktoren, die zur Unterscheidbarkeit beitragen, sind individuelle Sprachcharakteristika sowie die persönliche Bekanntheit mit den Sprechern. Interessanterweise lässt sich eine bekannte Stimme auch genauer lokalisieren (Blauert, 1983).

Schlechte oder fehlende Lokalisierung ist ein weiterer Grund für schlechte Trennbarkeit der Stimmen verschiedener Sprecher. In gängigen Telekonferenzsystemen wird die Sprache verschiedener Teilnehmer üblicherweise monophon wiedergegeben und somit am selben Ort wahrgenommen. Der Mensch ist fähig, akustische Informationen zu lokalisieren und unterschiedliche Richtungen zu nutzen, um Schallquellen voneinander zu trennen. Werden die Teilnehmer räumlich, also aus verschiedenen Richtungen wiedergegeben, kann der sogenannte *Cocktail-Party-Effekt* genutzt werden (Cherry, 1953): Trotz einer relativ lauten Umgebung können leisere Signale (insbesondere Stimmen) wahrgenommen werden. Dies funktioniert

jedoch nur, wenn die einzelnen Quellen an verschiedenen Orten lokalisiert werden können.

### 3 Audiosysteme für Telekonferenzumgebungen

#### 3.1 WAU, ein Internet-Telefon

*Wavelet Audio Unit* (WAU) ist ein Internet-Telefon für Windows-Plattformen, das am Fraunhofer-IGD wurde (siehe Abbildung 1).



Abbildung 1: WAU, ein Programm zum Telefonieren im Internet.

Die Qualität der Audiowiedergabe kann an die gegebene Bandbreite und Paketverlustrate angepasst werden. Je höher die effektive Bandbreite ist, desto höher kann die Qualität sein, ohne störende *drop-outs* zu erhalten. Es können sowohl analoge Modems und ISDN, als auch WANs und LANs mit höheren Übertragungsraten genutzt werden. WAU ermöglicht eine sehr gute Sprachqualität, welche die des Telefons übertrifft. Die Verzögerungszeiten sind gering gehalten, um interaktives Arbeiten zu unterstützen. Neben Standard-Kompressionsverfahren wie GSM 06.10, G.711 und ADPCM ist auch ein eigens entwickeltes Kompressionsverfahren integriert, das die skalierbare Qualität des Audiosignals gewährleistet. Je nach der zur Verfügung stehenden Bandbreite, reicht sie von Sprachqualität bis zu HiFi. Der Algorithmus basiert auf der Wavelet-Transformation (Wickerhauser, 1994). Auf Wunsch wird eine Erkennung der Sprechpausen vorgenommen, bei denen keine

Daten übertragen werden, was die im Netzwerk zur Verfügung stehenden Ressourcen schont.

In WAU ist ein einfaches aber effizientes FEC-Verfahren zur Behandlung von Paketverlusten integriert. Bis zu vier verlorene Pakete in Folge können rekonstruiert werden. Ermöglicht wird dies durch Hinzufügen redundanter Informationen. Unangenehme *drop-outs* werden vermieden. Auch bei mässiger Übertragungsqualität reisst der Audiostrom nicht ab (Chodura 1998b).

### 3.2 WAU für mehrere Benutzer

Zur Unterstützung von Konferenzen mit mehr als zwei Teilnehmern wurde von uns ein Prototyp entwickelt, der auf WAU basiert. Dieser Prototyp ist derzeit noch auf eine Bandbreite von  $88\text{ kBits/s}$  pro Audiostrom beschränkt ( $22\text{ kHz}$  Abtastrate mit ADPCM Kompression). Steht im Netzwerk Multicasting zur Verfügung, kann dies genutzt werden (siehe Abschnitt 2.2.1). Zusätzlich können die Daten räumlich wiedergegeben werden. Dies ermöglicht die freie (akustische) Positionierung der Teilnehmer einer Konferenz um den Benutzer. Hierfür wurde eigens der *Spatial Audio Server* (SAS) entwickelt. Der SAS ist eine Anwendung zur räumlichen Wiedergabe von Audioströmen oder bereits aufgezeichneten Dateien im WAV-Format. Die Audiodaten können von einer lokalen Anwendung stammen oder per IP zum Server übertragen werden. Die Position des virtuellen Zuhörers und der Teilnehmer lässt sich durch Kommandos (*events*) in Echtzeit steuern. Zusätzlich können die Klangeigenschaften der virtuellen akustischen Umgebung beeinflusst werden. Nachhall kann künstlich hinzugefügt werden. Eine höhere Nachhallzeit suggeriert einen grösseren Raum.

Der SAS kann als reine Software-Lösung oder mit zusätzlichen digitalen Signalprozessoren (DSPs) verwendet werden. In der Software-Lösung wird das Paket *Realistic Sound Experience* (RSX) verwendet, das von der Firma INTEL für Microsoft-Betriebssysteme entwickelt wurde. Mit RSX können Audiodaten binaural, d.h. zweikanalig räumlich wiedergegeben werden. Zum Einsatz kommen Filter, die aus gemessenen Aussenohrübertragungsfunktionen berechnet wurden (Begault, 1994). Zur Wiedergabe sind am besten Kopfhörer geeignet. Bei der Verwendung von zwei Lautsprechern ist die Ortung der virtuellen Schallereignisse ausserhalb des Kopfes stark eingeschränkt. Verursacht wird dies durch Übersprechen zwischen den beiden Kanälen.

Um für grössere Gruppen mit bis zu 30 Teilnehmern räumliche Audioausgabe zu ermöglichen, wurde in einem Raum unseres Instituts, der für Vorführungen genutzt wird, ein Mehrkanal-Audiosystem mit zehn Lautsprechern installiert. Für die rechenintensive räumliche Mischung von Schallquellen in Echtzeit steht eine leistungsfähige Audioworkstation zur Verfügung. Auch hier wird der SAS als

Schnittstelle zwischen der Audioworkstation und dem Internet-Telefon verwendet.

## 4 Ergebnisse aus Projekten

### 4.1 Audio für *Cooperative Virtual Environments* (CVEs)

Die in Abschnitt 3.2 beschriebene Version von WAU wurde ursprünglich zur Kommunikation in dreidimensionalen Konferenzumgebungen verwendet. Das am Fraunhofer-IGD entwickelte *Cooperative Virtual Environment* (CVE) ermöglicht Treffen in einer künstlichen, dreidimensionalen Welt (Gräff et.al, 1998). Die Benutzer werden durch sogenannte Avatare repräsentiert, die für die anderen Teilnehmer durch Videobilder zu identifizieren sind (siehe Abbildung 2 links). Die Anwender

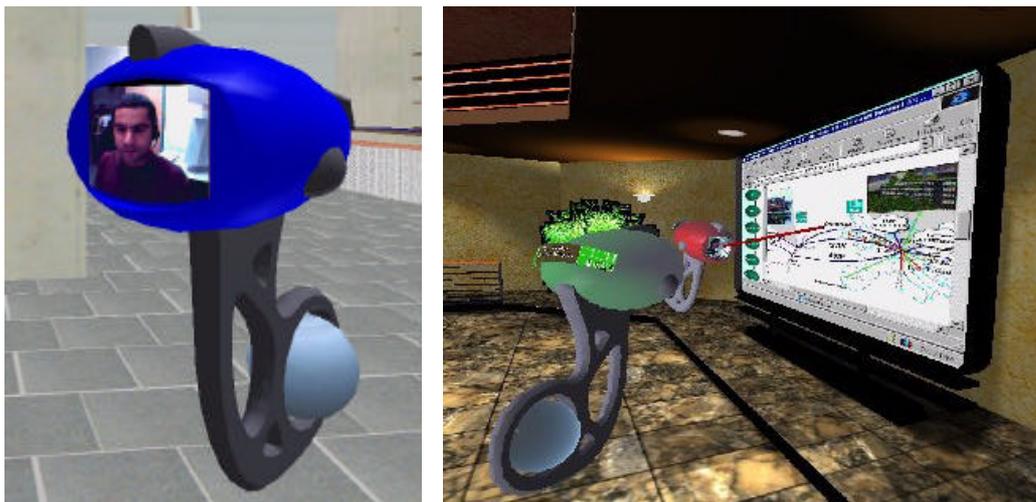


Abbildung 2: CVEs ermöglichen Treffen in virtuellen Konferenzräumen.

können ihren persönlichen Avatar durch die Szene bewegen und sehen dabei jeweils aus dessen Blickwinkel. An den Wänden des Konferenzraumes können Programme für alle sichtbar dargestellt und benutzt werden (siehe Abbildung 2 rechts). Durch den Einsatz von WAU wird im CVE eine natürliche Form der Kommunikation erreicht. Der SAS liest Position und Orientierung der Avatare aus der Szene aus und weist sie den Audioströmen der jeweiligen Benutzer zu. Die Daten werden so gemischt, dass ein Avatar aus der Richtung zu hören ist, wo er auch visuell wahrgenommen wird. Ist ein Teilnehmer ausserhalb des projizierten Blickfeldes, kann seine Position leicht akustisch gefunden werden. Je näher man sich an einem Avatar befindet, desto lauter ist sein Benutzer zu hören.

## 4.2 OrgTech - Telekommunikation für kleine und mittlere Unternehmen

Das Projekt OrgTech findet im Rahmen der Initiative ADAPT statt, die von der EU initiiert und gefördert wird. OrgTech unterstützt Konstruktionsbüros aus Nordrhein-Westfalen in den Bereichen Telekommunikation und 3D-Konstruktion. Für Telekonferenzen wird das Programm NetMeeting von Microsoft verwendet, das verteilte Anwendungen, Audio und Video ermöglicht und kostenfrei zur Verfügung steht. Ein Nachteil von NetMeeting ist die mässige Qualität der Audiokomponente (Krings et.al., 1999). Auch Mehrbenutzerbetrieb für Audio wird derzeit nicht von NetMeeting unterstützt. Dieser Umstand führte in der Vergangenheit dazu, dass NetMeeting zwar von den Anwendern genutzt wurde um Programme wie AutoCAD zu verteilen, für die Kommunikation wurde jedoch auf herkömmliches Telefon zurückgegriffen.

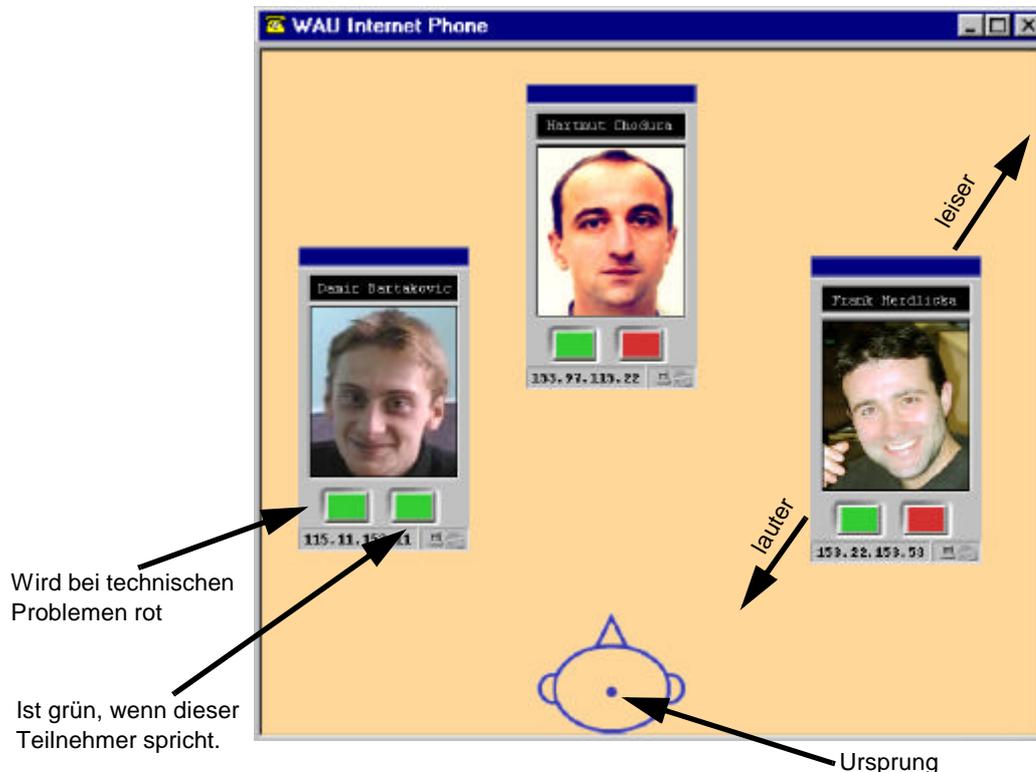


Abbildung 3: Das geplante UI für 3D-WAU. Position und Lautstärke können durch Verschieben der Symbole verändert werden.

Für dieses Einsatzgebiet werden WAU und NetMeeting in ein System integriert, das Konferenzen mit verteilten Anwendungen und Audio ermöglicht. Den Konferenzpartnern werden zunächst feste akustische (virtuelle) Positionen um den Benutzer zugewiesen. Später wird ein einfaches User-Interface die freie Wahl der Position und Lautstärke aller Konferenzpartner gestatten (siehe Abbildung 3). Wird das Symbol, das den Gesprächspartner repräsentiert nach links geschoben, ist er entsprechend von links zu hören. Durch Verschieben vom Ursprung weg wird er leiser; durch Verschieben zum Ursprung hin entsprechend lauter. Photos oder Videos auf den Symbolen werden durch die Information, wer gerade spricht bzw. wessen Verbindung wegen technischer Probleme gestört ist ergänzt. Somit die Unterscheidbarkeit zwischen Teilnehmern erhöht, sowie ein besserer Überblick über den Zustand der Konferenz ermöglicht.

## **5 Ausblick**

Durch das ständige Wachsen des Internets und die steigende Zahl von Benutzern wird auch das Interesse an Anwendungen im Internet zunehmen. Davon werden viele technische Spielereien bleiben, andere auf grössere Akzeptanz stossen. Nicht immer wird sich das technisch und qualitativ Beste durchsetzen.

Eine wichtige Anforderung an Technologie ist, dass sie einfach zu benutzen und einzusetzen ist. Im Projekt OrgTech ist ein Themenschwerpunkt, den Anwendern neue Technologien näher zu bringen und deren Vorteil gegenüber konventionellen Methoden herauszustellen. Arbeitsabläufe müssen gegebenenfalls erweitert oder überarbeitet werden, was den Prozess der Einführung häufig verlangsamt. Auch bei dem um Mehrbenutzer-Audio erweiterten NetMeeting werden wir mit Geduld und Aufmerksamkeit verfolgen, wie unsere Anwender das neue System annehmen.

## **6 Danksagung**

Besonderer Dank gilt den Mitarbeitern und Studenten in der Abteilung Kommunikation & Kooperation des Fraunhofer-IGD, die bei der Entwicklung von WAU beteiligt waren und wertvolle Anregungen gegeben haben, allen voran Kornel Knöpfle, Andreas Gräff, Damir Bartakovic, Volker Fitzek, Norbert Schiffner und unserem Abteilungsleiter Stefan Noll.

Bedanken möchten wir uns auch bei den Kollegen aus dem Projekt OrgTech für ihre Unterstützung bei der Integration unserer Audiotechnologien.

## 7 Literaturliste

- Albanese, A.; Blömer, J.; Edmonds, J.; Luby, M.; Sudan, M. (1994):  
*Priority Encoding Transmission*, Proceedings of 35th Annual Symposium on Foundations of Computer Sciences, Santa Fe, S. 604 - 612.
- Begault, D. R. (1994): *3-D Sound for Virtual Reality and Multimedia*, Academic Press, S. 230.
- Blauert, J. (1983): *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, S. 37 - 50.
- Brandenburg K.; Stoll G. (1994): *ISO-MPEG-1 Audio: A generic Standard for Coding of High-Quality Digital Audio*, Journal of the Audio Engineering Society, Vol. 42, Nr. 10, S. 780 - 791.
- Cherry, E. C. (1953): *Some experiments on the recognition of speech with one and two ears*.  
Journal of the Acoustical Society of America, Vol. 24, S. 975 - 979.
- Chodura, H. (1998a): *High-Quality Digital Audio for Network Applications*,  
Technischer Bericht, Fraunhofer-Institut für Graphische Datenverarbeitung, Darmstadt.
- Chodura, H. (1998b): *WAU-An Internet Telephone*, Topics - Reports on Computer Graphics,  
Haus der Graphischen Datenverarbeitung, Darmstadt, Vol. 10, Nr. 3 / 98, S. 13 - 14.
- Degener, J. (1994): *Digital Speech Compression*, Dr. Dobb's Journal, December 1994, S. 30 - 34.
- Gersho, A. (1994): *Advances in Speech and Audio Compression*, Proceedings of the IEEE,  
Vol. 82, Nr. 6, S. 900 - 918.
- Gräff, A.; Noll, S.; Schiffner, N. (1998): *Collaborative Virtual Environments*, Proceedings  
European Multimedia, Microprocessor Systems and Electronic Commerce Conference and  
Exhibition, EMMSEC 1998, Bordeaux.
- Kosiur, D. (1998): *IP Multicasting - The Complete Guide to Interactive Corporate Networks*,  
John Wiley & Sons.
- Krings, M.; Stiemerling, O.; Wulf, V. (1999): "There's a whole lot of sharin' going on" – *On the Versatility of Application Sharing*, Technischer Bericht, Universität Bonn.
- von Leitner, F. (2000): *Die Kunst des Weglassens, Grundlagen der Audio-Kompression*, c't,  
Heise Verlag, Nr. 3 / 2000, S. 130 - 137.
- Mark, G.; Grudin, J.; Poltrock, S. E. (1999): *Meeting at the Desktop: An Empirical Study of Virtually Collocated Teams*. Proceedings of the Sixth European Conference on Computer-Supported Cooperative Work, Kluwer Academic Publishers, S. 159 - 178.
- Noll, S.; Macedonia, M. R. (1997): *A transatlantic research and development environment*, IEEE  
Computer Graphics and Applications, Vol. 17, Nr. 2 / 1997, S. 76 - 82.
- Pohlmann, K. C. (1995): *Principles of Digital Audio*, 3. Auflage. McGraw-Hill, S. 23 - 28.
- Storn, R. M. (1996): *Echo Cancellation Techniques for Multimedia Applications - A Survey*,  
Technischer Bericht Nr. 96-46, International Computer Science Institute, Berkeley.
- Wickerhauser, M. V. (1994): *Adapted Wavelet Analysis from Theory to Software*, A. K. Peters Ltd.