Ergonomie: Mensch - Produkt - Arbeit - Systeme

Band 4

Hans-Jörg Zuberbühler

Quality Aspects of Multimodal Communication:

User Perception and Acceptance Thresholds

Shaker Verlag
Aachen 2003

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at http://dnb.ddb.de.

Zugl.: Zürich, ETH, Diss., 2003

Copyright Shaker Verlag 2003
All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

Printed in Germany.

ISBN 3-8322-2073-9
ISSN 1610-1898

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • eMail: info@shaker.de

Zusammenfassung

Telekommunikationsnetzwerke werden umgestellt von vermittlungsorientierten zu packetvermittelten Netzwerken. Diese Umstellung hat zur Folge, dass die Benutzer mit veränderten Netzwerkeigenschaften konfrontiert werden, wie zum Beispiel einer variablen Durchsatzrate und Übertragungsverzögerung, aber auch mit Verlusten von Datenpaketen. Diese neuen Eigenschaften stellen eine Herausforderung bezüglich Auslegung und Benutzung von packetvermittelten Netzwerken dar, da sie zu Behinderungen des gewohnten Kommunikationsprozesses führen können.

Bis anhin ist in diesem Gebiet noch wenig gesichertes Wissen vorhanden, weder darüber wie die Benutzer diese neue Situation wahrnehmen, noch darüber wie sie sich insgesamt verhalten. Dies erschwert die Konzeptionierung und Dimensionierung von Telekommunikationsnetzen, da auf fundierte Annahmen über Benutzerbedürfnisse und verlässliche Vorhersagen zur Netzwerkbelastung verzichtet werden muss. Die beschriebene Wissenslücke ist die treibende Kraft hinter der vorliegenden Arbeit, in der die Wahrnehmung und Akzeptanz der beiden Dienstqualität-Parameter *absolute* und *relative Verzögerung* untersucht werden.

Die Wahrnehmungs- und Akzeptanzschwellen der einzelnen Verzögerungsparameter werden anhand empirischer Versuche mit psychophysischer Methodik bestimmt. Dabei kommen entweder das *2-alternative forced-choice* oder das *yes-no* Paradigma sowie das adaptive psychophysische Verfahren *best-PEST* zur Anwendung. Die Experimente sind aufgeteilt in die beiden Interaktionsmodi *Mensch-Computer-Interaktion (HCI)* und *Mensch-Mensch-Interaktion (HHI)*, die beide unterschiedliche Verzögerungswahrnehmungen hervorrufen. Zur Bestimmung der HCI-Verzögerungsschwellen wird ein Versuchsaufbau eingesetzt, der Stimuluspräsentation, best-PEST Algorithmus und Datenerhebung vereint und der mit der objektorientierten Skriptsprache *Lingo* programmiert ist. Versuche im HHI Mo-

dus andererseits werden mittels einer Videokonferenzanwendung durchgeführt, die über ein emuliertes ATM-Netzwerk läuft. Dieser Versuchsaufbau besteht aus zwei oder drei Videokonferenzstationen, die über Glasfaser mit dem sogenannten *ARES*-System verbunden sind. (Das ARES-System emuliert das Echtzeit-Verhalten von ATM-Kanälen und bietet die Möglichkeit, gezielt Leistungsverschlechterungen bezüglich Verzögerung und Fehlerverhalten zu simulieren).

Im HCI Modus werden folgende Schwellwerte bestimmt:

- Relative Verzögerung zwischen auditiven Stimuli, die den visuellen vorangehen (AV).
- Relative Verzögerung zwischen visuellen Stimuli, die den auditiven vorangehen (VA).
- Absolute Verzögerung zwischen Stimmeingabe und visueller, rechnergestützter Antwort (VoiVis).
- Absolute Verzögerung zwischen Mauseingabe und visueller, rechnergestützter Antwort (MouVis).

Im HHI Modus werden folgende Schwellwerte bestimmt:

- Absolute Verzögerung bei einfacher auditiver Interaktion zweier Versuchspersonen (AudBas).
- Absolute Verzögerung bei einfacher visueller Interaktion zweier Versuchspersonen (VisBas).
- Absolute Verzögerung bei realistischer audio-visueller Interaktion zwischen drei Versuchspersonen (AudVisReal).
- Absolute Verzögerung bei realistischer auditiver Interaktion zwischen drei Versuchspersonen (AudReal).

Die Schwellwerte für relative Verzögerungen betragen 71 (±17) ms in der AV-Bedingung und 105 (±25) ms in der VA-Bedingung. Die Schwellwerte für absolute Verzögerung in HCI betragen 115 (±23) ms in der VoiVis-Bedingung und 78 (±14) ms in der MoiVis-Bedingung. In HHI betragen die Schwellwerte für absolute Verzögerung 216 (±44) ms in der AudBas-Bedingung und 237 (±92) ms in der VisBas-Bedingung. Wenn die Versuchspersonen realistische Gesprächssituationen nachzubilden haben, liegt ihre Wahrnehmungsschwelle bei 1220 ms und ihre Akzeptanzschwelle bei 2080 ms (AudVis-Real-Bedingung). In der AudReal-Bedingung betragen diese Werte 970 (±330) ms für

Wahrnehmung und 1760 (±410) ms für die Akzeptanz. Weder das Alter noch das Geschlecht der Versuchspersonen übt einen signifikanten Einfluss ($p>0.05$) auf die Schwellwerte aus.

Um aus den experimentellen Daten psychometrische Funktionen zu erhalten, werden für alle Bedingungen logistische Kurven gefittet. Der Nutzen dieser Funktionen besteht darin, dass Netzwerkplaner sowie Anbieter von Inhalten und Diensten abschätzen können, welcher Anteil Benutzer bestimmte Verzögerungswerte bemerken und/oder ablehnen wird. Diese ‚politische' Frage wird massgeblich durch ökonomische Betrachtungen beeinflusst, welches Preis-Leistungsverhältnis den Kunden angeboten werden soll.

Des weiteren werden die relativen Verzögerungsschwellwerte im Licht der neuronalen Verarbeitungsgeschwindigkeit für verschiedene Modalitäten diskutiert. Und absolute Verzögerungsschwellwerte werden bezüglich ihrer Abhängigkeit von den auszuführenden – durch verschiedene Interaktionsgrade gekennzeichnete – Aufgaben diskutiert.

Abstract

Recent trends in telecommunication networks indicate a shift away from the use of circuit-switched networks towards the use of packet-switched networks. This new networking environment will present end users with new characteristics like variations in transmission delays, and bit rates as well as potential loss of data packets. These characteristics represent a challenge in the design and use of packet-switched networks, since they may be lead to user impairments, depending on the kind of source coding and compression used in the end-systems.

It is generally agreed that very little is known about user expectations or perceptive mechanisms and user behaviour in this new situation. As a consequence, it is presently difficult to base network engineering on proper traffic forecasts and real user requirements. This lack of knowledge is the driving force behind our work, aiming to examine user perception and acceptance of the Quality of Service (QoS) parameters *absolute* and *relative delays* (also referred to as *roundtrip delay* and *synchronisation error*).

In this thesis we investigated the perception and the acceptance thresholds for particular delay parameters using psychophysical methods. I.e. threshold are obtained by means of empirical determinations applying either *2-alternative forced-choice* or *yes-no* paradigms, and using the adaptive psychophysical procedure called *best-PEST*. The experiments are conducted in the interaction modes *Human-Computer-Interaction (HCI)*, and *Human-Human-Interaction (HHI)*, which evoke different delay perceptions. HCI delay thresholds are obtained using an experimental set up that includes stimulus presentation, best-PEST algorithm, and data acquisition. It is implemented using the object-oriented scripting language *Lingo*. The experiments conducted in the HHI mode comprise threshold determinations in which the experimental subjects interact with each other over a videoconference that uses an ATM-network infrastructure. The experimental set-up consists of two

or three videoconference stations connected via fibre passing through a system called *ARES*, which emulates the behaviour of *ATM* channels in real-time with the possibility to emulate performance degradations, such as delay or errors.

In the HCI mode the following thresholds are determined:

- Relative delay between auditory stimuli preceding visual stimuli (AV).
- Relative delay between visual stimuli preceding auditory stimuli (VA).
- Absolute delay between voice input and visual computer-generated response (VoiVis).
- Absolute delay between mouse input and visual computer-generated response (MouVis).

In the HHI mode the following thresholds are determined:

- Absolute delay in basic auditory interaction between two subjects (AudBas).
- Absolute delay in basic visual interaction between two subjects (VisBas).
- Absolute delay in realistic audio-visual interaction between three subjects (AudVisReal).
- Absolute delay in realistic auditory interaction between three subjects (AudReal).

The thresholds for relative delays are 71 (±17) ms for the AV condition, and 105 (±25) ms for the VA condition. The thresholds for absolute delay in HCI are 115 (±23) ms for the VoiVis condition, and 78 (±14) ms for the MoiVis condition. In HHI the thresholds for absolute delays are 216 (±44) ms for the AudBas condition, and 237 (±92) ms for the VisBas condition. Accomplishing a realistic task the perception threshold is 1220 ms, and the acceptance threshold is 2080 ms in the AudVisReal condition. In the AudReal condition the perception threshold is 970 (±330) ms, and the acceptance threshold is 1760 (±410) ms. Age and gender of the experimental subjects have no significant effect ($p>0.05$) on these results.

To obtain psychometric functions experimental data of each condition are fitted using a logistic model. The benefit of such functions is that network planners, as well as content and service providers are delivered with a means to estimate which user percentages are expected to detect and to reject a specific delay. This 'political' question is influenced

by economical considerations, which price/performance ratio is intended to be offered to the user.

Furthermore the relative delay thresholds are discussed in the light of neural processing times for different modalities. And the absolute delay threshold is discussed regarding the task dependency represented by different degrees of interactivity.