Gefördert vom
bmb+f
Bundesministerium für
Bildung, Wissenschaft,
Forschung und Technologie

SmartKom Mensch-Technik-Interaktion

Technisches Setup der Datensammlung – SmartKom Version 1

Uli Türk

Ludwig Maximilians Universität München

Technisches Dokument Nr. 7
März 2000

März 2000

Uli Türk

Ludwig Maximilans Universität München
Schellingstr. 3
80799 München

Tel.: (089) 2180–5752
FAX: (089) 2800362

E–Mail: tuerk@phonetik.uni–muenchen.de

Dieses Technische Dokument gehört zu Teilprojekt 1: Modalitätsspezifische Analysatoren und Projektfeld II: Anwendungsszenarien

Das diesem Technischen Dokument zugrundeliegende Forschungsvorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01 IL 905 gefördert. Die Verantwortung für den Inhalt liegt bei den Autoren.

Inhaltsverzeichnis

1 Überblick

Dieses Dokument beschreibt den technischen Aufbau für die Datensammlung im SmartKom−Projekt. Neben einem Überblick über das zusammengestellte System wird auch auf die einzelnen Komponenten genauer eingangen.

Die hier vorgestellten Daten, insbesondere die geometrischen Abmessungen, wurden für die ersten Aufnahmetests festgelegt; es können sich jedoch noch Änderungen ergeben. Ein Ähnliches gilt für die eingesetzten Geräte. Auch hier kann die eine oder andere Änderung oder Ergänzung notwendig sein. Bei den Detailbeschreibungen der einzelnen Komponenten ist jeweils mitangegeben, welche Funktionalität bereits getestet wurde.

Dieses Dokuments wird laufend aktualisiert, Änderungen und Ergänzungen eingefügt. Kapitel 5 wird eine Übersicht über die erfolgten Veränderungen zeigen.

2 Allgemeines zum Aufbau

Der Aufnahmeraum, der am Institut für Phonetik eingerichtet wurde, ist durch eine Trennwand in zwei Teile gegliedert. Im vorderen, direkt von der Türe aus zugänglichen Teil ist der Aufnahmeraum, in dem die Versuchsperson mit dem SmartKom−System interagiert. Der hintere Teil, der Wizardraum bzw. Technikraum, ist durch eine verdeckte Türe in der Trennwand erreichbar. Hier arbeiten voraussichtlich zwei Wizards, die die Aufnahmen steuern und die Funktionalität des SmartKom−Systems simulieren.
Ein Versuchleiter, der die Versuchsperson betreut, wartet während der Aufnahme vor dem Raum und kann über eine Sprechverbindung mit den Wizards Kontakt halten.

Den schematischen Aufbau des simulierten SmartKom−Systems im Aufnahmeraum zeigen die Abbildungen 1 und 2. Dieser Aufbau entspricht dem Anwendungsfall von SmartKom Public: eine Benutzerin steht vor einer Projektionsfläche. Die weiteren Anwendungsfälle Home und Mobile lassen sich durch kleine Modifikationen am Aufbau (Sitzgelegenheit für die Versuchsperson bzw. Verkleinerung der Projektionsfläche) ebenfalls auf diese Weise simulieren.

In der Seitenansicht (Abb. 1) erkennt man das Hauptelement des Aufbaus, den Ständer der SIVIT−Einheit, der auf einem höhenverstellbaren Tisch montiert ist. Auf der Projektionsfläche am Fuß des Ständers liegt das Graphiktableau. Am oberen Ende des Ständers befinden sich in einer Box die SIVIT−Komponenten (Video−Beamer, Infrarot−Kamera−System) und eine zusätzliche Video−Kamera (Monitor−Kamera). Die Frontkamera nimmt das Gesicht der Versuchsperson schräg von unten auf. Die Stimmausgabe des SmartKom−Systems erfolgt über einen Frontlautsprecher unter dem Tisch.

Die Ansicht von oben (Abb. 2) zeigt die Positionierung der Seiten−Kamera. Als definierter Hintergrund für die Videoaufnahmen aus dieser Position wird eine weiße Leinwand verwendet. Bezugspunkt für die Positionierung der beiden Videokameras bildet die Projektionsfläche des Video−Beamers; die Größe dieser Fläche wurde mit 60cm x 45cm festgelegt.
Über einen Lautsprecher hinter der Versuchsperson werden Umgebungsgeräusche eingespielt.

Zur Audio−Aufnahme stehen insgesamt 7 Mikrophone zur Verfügung: ein Richtmikrofon, das auf der Front−Kamera montiert ist, ein Mikrophon−Array aus vier Einzelmikrophonen und zwei weitere Mikrophon−Kapseln, die je nach Szenario als zwei Ansteckmikrophone oder als ein Headset−Mikrophon eingesetzt werden.

Technisches Setup zur Datensammlung – SmartKom (Version 1)

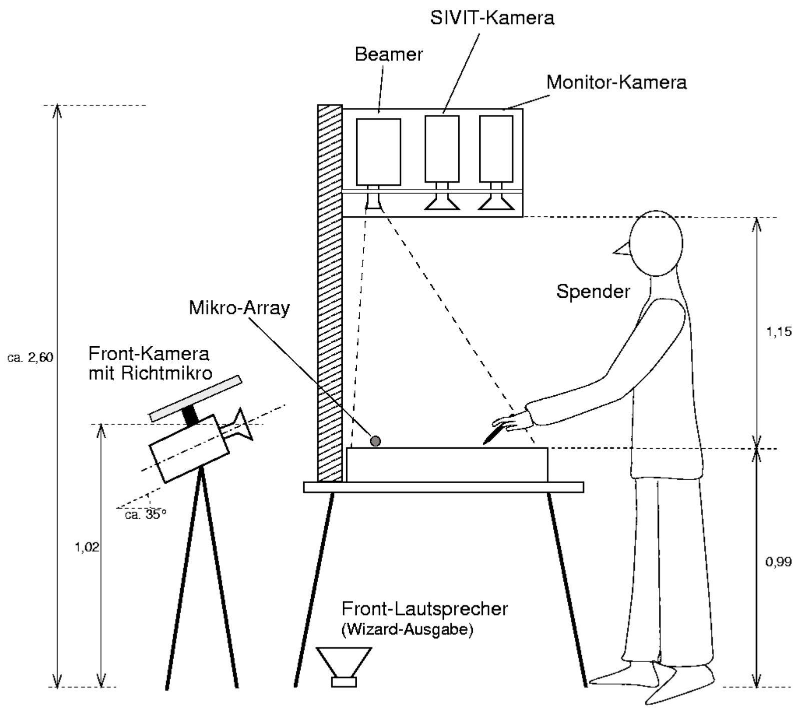

Abbildung 1 Schematischer Aufbau (Seitenansicht).
Alle Längenmaße sind in Meter angegeben.

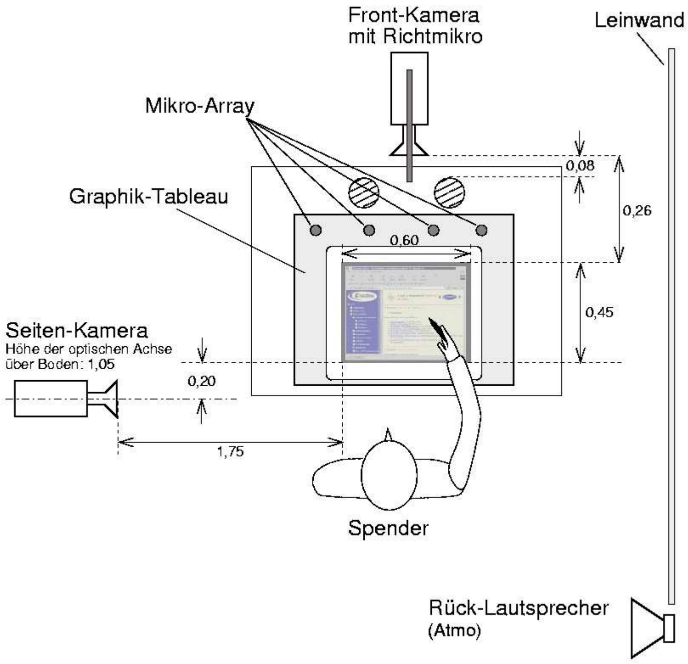

Abbildung 2 Schematischer Aufbau (Ansicht von oben).
Alle Längenmaße sind in Meter angegeben.

Das Blockschaltbild aus Abbildung 3 zeigt alle Komponenten und ihre Verbindungen untereinander.Ausführliche Informationen zu den einzelnen Komponenten werden im nächsten Abschnitt gegeben.

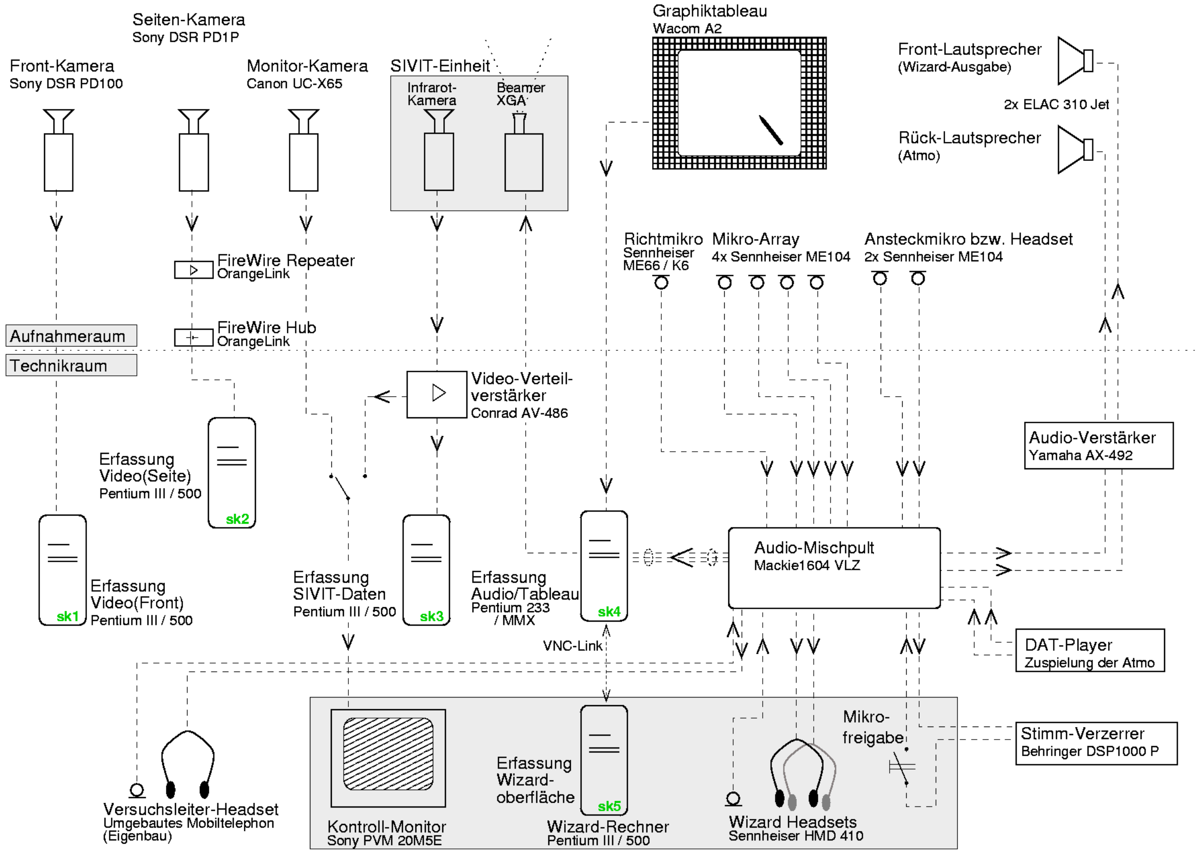

Abbildung 3   Blockschaltbild zum Aufnahme–Setup

3 Die einzelnen Komponenten

3.1 Videokomponenten

3.1.1 Videoerfassung (Front–Kamera)

Die Frontkamera (Sony DSR PD100) ist direkt über eine FireWire–Verbindung (Kabel mit 4,5m Länge) mit ihrem zugehörigen Hostrechner (Rechner [sk1], ausgerüstet mit einer MotoDV FireWire–Karte) verbunden. Diese Kamera unterstützt den geforderten Progressive–Scan–Mode.
Die Position der Kamera in Bezug auf die Projektionsfläche ist aus Abbildung 1 und 2 erkenntlich. Ihr Neigungswinkel muß vor Beginn der Aufnahmen für jede Versuchsperson neu eingestellt werden. Das Bild sollte am oberen Rand noch genügend Platz besitzen, da sich die Versuchsperson während den

Aufnahmen leicht über die Projektionsfläche beugt. Abbildung 4 zeigt ein Beispiel−Bild aus einer Testaufnahme.

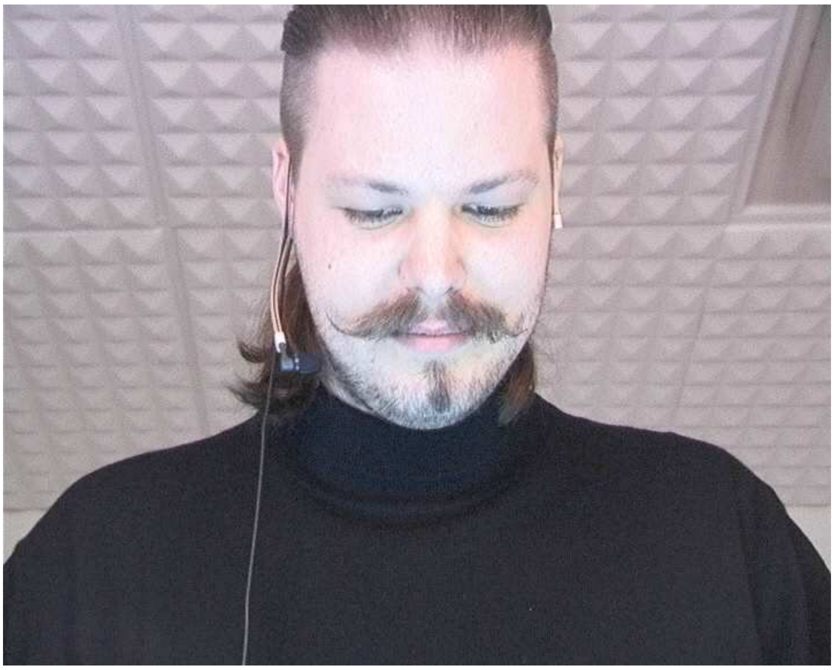

Abbildung 4   Beispiel−Bild aus der Aufnahme der Front−Kamera. Die Versuchsperson trägt hier das Headset.

3.1.2 Videoerfassung (Seiten−Kamera)

Ähnlich wie die Front−Kamera ist die Seiten−Kamera, eine Sony DSR−PD1P, über eine FireWire Verbindung an ihren Hostrechner angeschlossen. Aufgrund der größeren Kabellänge ist hier ein zusätzlicher FireWire−Repeater erforderlich, dessen Versorgungsspannung mit Hilfe eines FireWire−Hubs in den Bus eingespeist wird. Zwischen Rechner und Hub und zwischen Hub und Repeater werden zwei Kabel mit jeweils 1.8m, zwischen Repeater und Kamera ein Kabel mit 4,5m Länge verwendet.
Im Gegensatz zur Front−Kamera kann der Bildausschnitt der Seiten−Kamera einmal fest eingestellt werden.
Ein Beispiel für die Seitenansicht ist in Abbildung 5 gezeigt.

Abbildung 5 Beispiel für die Seitenansicht. Die Versuchsperson trägt hier wieder das Headset.

3.1.3 Kontrollmonitor und Monitor−Kamera

Auf dem Kontroll−Monitor im Wizardraum kann zum einem die Ausgabe der Infrarot−Kamera in der SIVIT−Einheit kontrolliert werden, zum anderen dient er den Wizards als Anzeige der Benutzer−Aktionen (die Infrarot−Kamera zeigt nur die Hände des Benutzers, nicht jedoch das projektierte Bild). Mit Hilfe der Monitor−Kamera können sie die Aktionen des Benutzers auf der Oberfläche verfolgen.

Diese Kamera, eine einfache Hi8 Video−Kamera von Canon, ist zusätzlich in die von vorne aus gesehene linke, vordere Ecke der SIVIT−Box eingebaut. Mit Hilfe eines Kugelkopf−Gelenks läßt sie sich auf die Projektionsfläche ausrichten.

Abbildung 6 zeigt die Position der Monitor−Kamera. Sie ist über eine Composite−Video−

Abbildung 6 Einbaulage der Monitor−Kamera, Ansicht von der linken Seite

Verbindung direkt an den Kontroll–Monitor angeschlossen. Das Signal der Infrarot–Kamera wird über einen Verteilverstärker zum Monitor und zum dazugehörigen Aufnahme–Rechner (Erfassung der SIVIT–Daten, [sk3]) geführt.

3.1.4 Durchgeführte Tests

Mit beiden Digital–Videokameras (Front–Kamera und Seiten–Kamera) wurde das Speichern von Videodaten auf der Festplatte getestet. Das dafür vorgesehene Softwaretool MotoDV läuft auf einem der Rechner problemlos, führt auf dem anderen Rechner jedoch in unregelmäßigen Abständen zu Abstürzen. Hier muß noch die richtige Konfiguration des Rechners für einen sicheren Betrieb gefunden werden. Zur Sicherheit kann während den Aufnahmen zusätzlich ein Band in jeder Kamera mitlaufen, das im Fall eines Absturzes in den Rechner eingelesen werden kann. Dies wurde ebenfalls getestet.
Die aufgenommen Video–Daten können mit einer Quicktime–kompatiblen Anwendung (z.B. Adobe Premiere) problemlos geöffnet und verarbeitet werden.

3.2 Audiokomponenten

3.2.1 Aufbau

Zentrales Element der Audiokomponenten bildet das Mischpult (Mackie 1604 VLZ). Es ermöglicht die Verarbeitung mehrerer Audiosignale unabhängig voneinander:

- Verstärkung der Mikrophon–Signale und Speisung der Mikrophone mit der Phantom–Spannung
- Einspeisung der verstärkten Mikrophon–Signale in die Audio–Erfassungskarte (Terratec EWS 88 MT)
- Routing des Signals des Wizard–Mikrophons
- Routing der Hintergrundgeräusche (vom DAT–Player)
- Mischen des Kontrollsignales für die Wizards
- Kommunikation zwischen Wizards und Versuchsleiter

Aus Abbildung 7 wird der Verlauf der Audio–Signale deutlich. Die Signale der sieben Mikrophone im Aufnahmeraum werden nach dem Eingangsverstärker über die Direct–out–Ausgänge an die Audiokarte im Rechner weitergeleitet, wo sie aufgezeichnet werden. Die Sprachausgabe des Wizards an die Versuchsperson wird auf zwei Kanäle aufgeteilt: das durch den Sprach–Verzerrer verfremdete Signal wird über den Audio–Verstärker auf den Frontlautsprecher gegeben und gleichzeitig auch zur Audio–Karte weitergeleitet. Über einen Taster wird diese Sprachausgabe freigeschaltet. Das unverzerrte Sprachsignal des Wizards dient zur Kommunikation mit dem Versuchsleiter, die in den Pausen zwischen der Sprachausgabe an den Benutzer erfolgen kann. Der Versuchsleiter verwendet zur »Tarnung« ein modifiziertes Mobiltelephon, das über Kabel an das Mischpult angeschlossen ist.
Die Hintergrundgeräusche werden durch einen DAT–Player eingespielt und über den Verstärker auf den Rück–Lautsprecher ausgeben. Gleichzeitig wird dieses Signal über die Digital–Schnittstelle (S/PDIF) an die Audio–Karte weitergeleitet.
Die Wizards erhalten auf ihren Kopfhörern ein Mischsignal aus den verschiedenen Mikrophon–Signalen im Aufnahmeraum und das Sprachsignal des Versuchsleiters.

Die Aufnahme der Audiodaten im Rechner erfolgt mit einer Abtastfrequenz von 48 kHz und einer linearen Quantisierung von 16bit. Diese Abtastfrequenz wird durch den am Digital–Eingang der Karte angeschlossenen DAT–Player vorgegeben. Nach der Aufnahme müssen die Audio–Daten dann auf die geforderte Abtastfrequenz von 16kHz konvertiert werden. Dazu ist eine digitale Tiefpaß–Filterung (Bandbegrenzung auf 8 kHz)) und ein Downsampling um den Faktor drei (48 / 3 = 16) notwendig. Dies kann zwar mit der verwendeten Aufnahme–Software (Samplitude Studio) durchgeführt werden, wahrscheinlich ist jedoch weiteres Tool notwendig, das einen Batch–Processing–Betrieb ermöglicht.

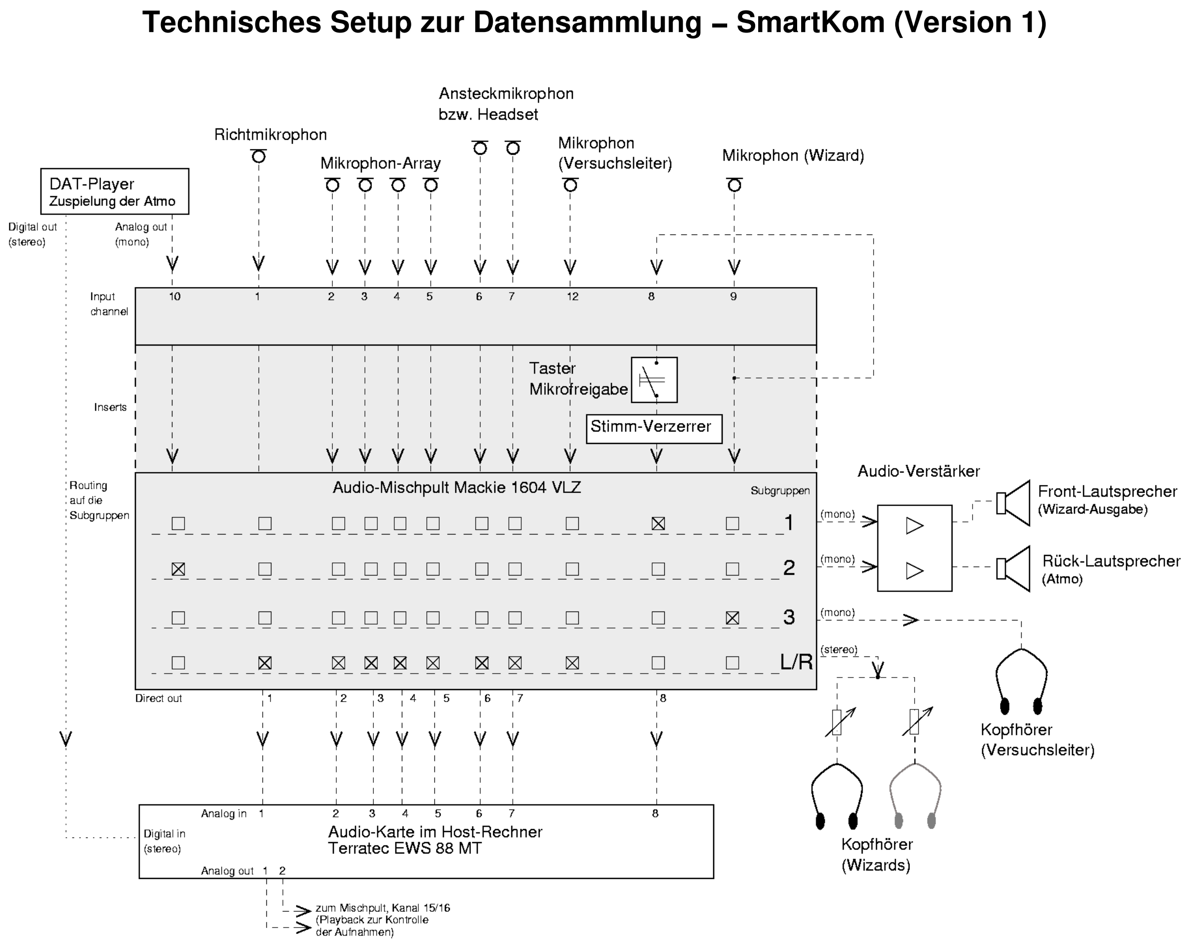

Abbildung 7   Routing der Audio–Signale

3.2.2 Durchgeführte Tests

Die einzelnen Audiokomponenten wurden erfolgreich getestet. Die Aufnahme der Audio–Daten auf Festplatte funktioniert mit den angegebenen Software–Treibern problemlos. Bislang konnte nur ein Test mit allen acht Analog–Eingängen durchgeführt werden. Ein Probelauf mit der Aufnahme von allen zehn Eingängen (acht analog, zwei digital) steht noch aus.
Noch nicht geklärt ist, ob der verwendete Rechner noch Resourcen für die Auswertung der Tableau–Daten und den Betrieb der VNC–Verbindung (siehe 3.5.1) besitzt.

3.3 SIVIT–Einheit

3.3.1 Aufbau

Die SIVIT–Einheit wurde für den Einsatz in diesem Test–System an zwei Punkten modifiziert. Die Geräte–Box, die den Video–Beamer und das Infrarot–System enthält, wurde weiter nach oben gesetzt, um ein ausreichend großes Projektionsbild zu erhalten. Der Abstand zwischen der Oberfläche des Graphiktableaus und der Unterkante der Geräte–Box beträgt jetzt 1,15 m (siehe Abb. 1). Dieser Umbau erfordert eine Neukalibrierung mit Hilfe der SIVIT–Treiber–Software.
Die zweite Änderung betrifft den Hostrechner. Der zusammen mit der SIVIT–Einheit gelieferte Siemens PC wird, wie bereits erwähnt, zur Aufnahme der Audio–Daten verwendet. Die Falcon–

Grabber−Karte und die SIVIT−Software sind stattdessen in dem angegebenen AK−Office PC installiert.
Für einen störungsfreien Betrieb ist eine Antireflexions−Folie (Fa. Siemens) notwendig, die auf die Projektionsfläche aufgeklebt werden muß. In unserem Fall ist das Graphiktableau ganzflächig mit dieser Folie beklebt.

3.3.2 Durchgeführte Tests

Die SIVIT−Einheit funktioniert problemlos. Das Protokollieren der Finger−Koordinaten konnte noch nicht getestet werden, da die Treiber−Software noch nicht dafür ausgestattet ist.

3.4 Graphiktableau

3.4.1 Aufbau

Zum Ausgleich der geneigten Fläche am Fuß des SIVIT−Ständers wird das Graphiktableau um 180° gedreht verwendet (eigentliche Oberkante zeigt zur Versuchsperson). Damit ergibt sich eine leicht zur Benutzerin geneigte Projektionsfläche mit einer Neigung von ca. 2°. Wie bereits erwähnt, ist die gesamte Oberseite des Graphiktableaus mit der Antireflexions−Folie beklebt.

3.4.2 Durchgeführte Tests

Das Graphiktableau funktioniert mit den mitgelieferten Anschlußkabel und Treibern. Es fehlen jedoch noch die notwendige Verlängerung der Anschlußleitung und der Test dieses Gesamtaufbaus.

3.5 Präsentation und Steuerung der Benutzeroberfläche

3.5.1 Konzept

Da die Versuchsperson auf der Benutzeroberfläche keinen Hinweis auf die Tatsache erhalten soll, daß das System von den Wizards simuliert wird, die Wizards aber selbst auf die Steuerung der Oberfläche per Maus angewiesen sind, ist folgendes Konzept entworfen worden:
Die Wizards steuern die Oberfläche an einem Rechner, der zu Protokollzwecken seine eigene Bildschirmausgabe mit einer niedrigen Wiederholrate (z.B. 4 fps) in eine Datei schreibt. Der gesamte Bildschirm wird über das Software−Protokoll VNC an einen zweiten Rechner (Audio−Rechner) übertragen, der diesen Inhalt in einem gesonderten Fenster darstellt. Die Bildschirmausgabe dieses Rechners erfolgt über den Video−Beamer.
Das Icon des Mauszeigers kann hier durch ein leeres Sprite ersetzt werden, sodaß er für die Versuchsperson unsichtbar wird.

3.5.2 Durchgeführte Tests

Eine VNC−Verbindung zwischen zwei Rechnern läßt sich problemlos aufbauen. Die Darstellung des Mauszeigers kann in den Systemeinstellung zwischen sichtbar und unsichtbar umgeschaltet werden.
Das Mitprotokollieren der Bildschirmausgabe ist prinzipiell möglich, führt aber selbst bei niedrigen Wiederholraten zu starker Auslastung des Rechners.
An einem Test der gesamten Konfiguration wird noch gearbeitet.

4 Geräteliste SmartKom−Datensammlung

4.1 Audio−Komponenten

Hardware

- Sennheiser NB2 Nackenbügel
- 6x Sennheiser Mikrofonkapseln ME 104 (Nierencharakteristik)
- 4x Sennheiser KA 100 SP, Stahlkabel (3m), gerader Mikrostecker auf XLR
- 2x Sennheiser KA 100 P XLR, (3m), Winkel−Mikrostecker auf XLR
- Sennheiser Richtmikrophon, ME 66
- Sennheiser Korpus f. Richtmikrophon, K6
- Mackie Audiomixer, 1604 VLZ Pro
- Terratec Audio−Karte EWS 88 MT
- 2x ELAC CL 310 Jet, Monitor−Boxen
- Yamaha, Verstärker AX−492, 2x120 W
- Sennheiser Headset, HMD410
- Standard Kopfhörer

Software

- SEK'D Samplitude Studio, Version 5.56, für Win
- Terratec Audio−Treiber für Win NT 4, Version 4.02.12 Build 75

Kabel

- 6x XLR−Verlängerung, 10m
- XLR−Anschlußkabel, 7,5m
- 14x Cinch auf Klinke, 6,3mm, Mono

4.2 Video−Komponenten

Hardware

- Sony DiviCam DSR−PD 100AP
  Firewire
- Sony DiviCam DSR PD 1
  Firewire
- Canon HI−8 VideoCam UCX 65
  22/440−Zoom, 470000 Pixel
- 2x Sachtler DV4 Kamera−Stativ
- 2x Digital Origin MotoDV,
  FireWire PCI−Karte inkl. Premiere LE, Photoshop LE, Capture− und Foto−Software
- 2x IDS Falcon, PCI−Grabberkarte, monochrom
- Sony PVM 20M5E, Studiomonitor
  500 Linien Auflösung, Composite A/B, YC A/B, 2xAudio in.
- Conrad AV−486, Audio−Video−Verteilverstärker
- OrangeLink FireWire Hub
- OrangeLink FireWire Repeater

Software

- MotoDV Capture Utility

Kabel

- 2x FireWire−Kabel, 4,5m, 6 pol. auf 4 pol.
- 2x FireWire−Kabel, 1,8m, 6 pol. auf 6 pol.
- Videokabelverlängerung, 10m

4.3 SIVIT–Einheit

- SIVIT StarterKit mit Siemens PC (Pro M5, siehe unter Hostrechner), OMB und XGA–Beamer

4.4 Graphiktableau

Hardware

- Wacom UltraPad A2, seriell für PC

Software

- Wacom–Treiber–Set, Version 4.40

4.5 Hostrechner

Hardware

- Video–Daten (Gesicht) [sk1]
  AK Office Pentium III/500, 128 MB RAM, 2x18GB UW–SCSI Cheetah, 8 MB ATI–XPert, 48x–CD, 10/100 LAN
- Video–Daten (Seite) [sk2]
  AK Office Pentium III/500, 128 MB RAM, 1x18GB UW–SCSI Cheetah, 8 MB ATI–XPert, 48x–CD, 10/100 LAN
- Video–Daten (SIVIT) [sk3]
  AK Office Pentium III/500, 128 MB RAM, 18GB UW–SCSI Cheetah, 8 MB ATI–XPert, 48x–CD, 10/100 LAN
- Wizard–Rechner, Frame–Grabber für Wizard–Oberfläche [sk5]
  AK Office Pentium III/500, 128 MB RAM, 13GB UW–SCSI Cheetah, 8 MB ATI–XPert, 48x–CD, 10/100 LAN
- Audio–Daten [sk4]
  Siemens Pro M5, Pentium II/233 MMX, 32MB RAM

Software

- Windows NT 4, ServicePack 6

4.6 Schnittrechner

Hardware

- Audio–Video–Schnittrechner[sk6]
  AK Office Pentium III/500, 256 MB RAM, 2x18GB UW–SCSI Cheetah, 8 MB ATI–XPert, 48x–CD, 10/100 LAN
- Monitor EIZO F77S, 21 Zoll

Software

- Windows NT 4, ServicePack 6
- Adobe Premiere 5.1

4.7 Sonstiges

Netzwerk–Komponenten

- Techworks Switch LSW 10/100 16–H–2

Mechanischer Aufbau

- Tisch für den Aufbau, 75 x 120cm, 84cm hoch

Leinwand

- Roll–Leinwand, 2m x 2m

Sprechverbindung für Versuchsleiter

- modifiziertes Mobilteil eines schnurlosen Telephons, Hörerkapsel und Mikrokapsel über Kabel (10m) mit Mischpult verbunden.

5 Versionen dieses Dokuments

Version 1, 23.3.2000

Erste Ausgabe