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

Mensch-Technik-Interaktion
SmartKOM

Konvention der Dateinamen

(Datensammlung)

Version 1.2

Nicole Beringer

Ludwig–Maximilians–Universität München

Technisches Dokument Nr. 3

August 2000 (update 5 Februar 2002)

Konvention der Dateinamen

August 2000
Update 3 Juni 2001
Update 4 Oktober 2001
Update 5 Februar 2002

Nicole Beringer

IPSK LMU München
Schellingstr. 3
80799 München

Tel.: (089) 2180−5751
FAX: (089) 2800382

E−Mail: beringer@phonetik.uni−muenchen.de

Dieses Technische Dokument gehört zu Teilprojekt 1: Modalitätsspezifische Analysatoren

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 Einleitung

Zur Verwaltung der anfallenden Daten der Datensammlung ist es wichtig, bereits aus der Nomenklatur der Dateinamen Informationen über die Art der vorliegenden Aufnahmen zu erhalten.
Dabei soll zum einen die Art der vorliegenden Aufnahmen (Biometrieaufnahmen oder Wizard−Of−Oz−Aufnahmen) sowie die jeweils vorliegende Sessionnummer aus den Dateinamen hervorgehen, zum anderen sollte eine Identifikation der Domäne sowie des aufgezeichneten Hauptanwendungsfalls (vgl. [stein00]) möglich sein. Außerdem sollte man aus der Nomenklatur erkennen können, in welchem Aufnahmekanal die Aufnahme aufgezeichnet wurde.
Bei den Bezeichnungen der Daten muß man zwei Fälle unterscheiden. Zum einen wird die Dateistruktur dazu verwendet, um Abhängigkeiten der Daten untereinander wiederzugeben − also daß z.B. ein bestimmtes Audiofile zu einer bestimmten Mimikaufnahme gehört − zum anderen dient sie zur Strukturierung innerhalb der Annotationen. Die ersten sieben Zeichen des Dateinamens bilden den Namen der übergeordneten Dateistruktur (vgl. hierzu Abschnitt 8).
Da v. a. die Wizard−Of−Oz−Aufnahmen turnsegmentiert werden (vgl. [bobtk00]), muß auch die Turnnummer im Dateinamen registriert sein.
Für die Analyse der vorliegenden Daten mit den in der VP−Datenbank (vgl.[berdb00]) gespeicherten sprecher−, bzw. sprecher−aufnahmespezifischen Merkmalen wird zudem noch die Identifikation über das VP−Kürzel benötigt.
Wie in den nachfolgenden Abschnitten beschrieben ergibt sich somit ein 16stelliger Dateiname.
Die Überschreitung der für DOS− bzw. Windows−Betriebssysteme vorgeschriebenen Eingrenzung des Dateinamens auf acht Stellen entfällt, da die Erstellung des SmartKom−Systems unter Windows−NT erfolgt, was mehr als acht Stellen für den Dateinamen zuläßt.

2 Wiedergabe der Art der Aufnahme

Da sich die Aufnahmen zunächst in zwei getrennte Aufnahmearten, nämlich Biometrieaufnahmen (vgl. hierzu [berbm00]) und Wizard−Of−Oz−Aufnahmen (vgl. [stein00]) aufteilen, wird die Art der Aufnahmen als einstelliges kleinbuchstabiges Kürzel zu Beginn des Dateinamens wiedergegeben.

Dabei steht

- **b** für **B**iometrie−Aufnahmen
- **w** für **W**izard−Aufnahmen

  ferner
- **d** für **D**emodialog−Aufnahmen
- **p** für **P**robeaufnahmen
- **e** für **e**nglische Aufnahmen
- **v** für E**v**aluationsaufnahmen

3 Wiedergabe der Sessionnummer

Im Anschluß an das Kürzel zur Kennzeichnung der Art der Aufnahme wird eine 3stellige Sessionnummer ohne Abtrennung durch andere Zeichen vergeben.
Die erste Session bei den Wizardaufnahmen erhält die Sessionnummer:
001
Alle weiteren Sessions werden fortlaufend numeriert.
Achtung: Die Biometrieaufnahmen werden gemäß Absprache mit der Biometrieverarbeitung ab
000
beginnend fortlaufend numeriert.

4 Wiedergabe des Szenarios und des Anwendungsfalles

Es folgt ein zweistelliges Kürzel aus Kleinbuchstaben zur Identifikation des Szenarios abgetrennt durch _ .
Hier gibt es allerdings Unterschiede zwischen den Aufnahmearten, die nachfolgend beschrieben werden.

4.1 Szenarien und Anwendungsfälle in den Wizard−Of−Oz−Aufnahmen

Wie im Technischen Dokument Nr. 5 ([stein00]) detailliert beschrieben ist, teilen sich die Wizard−Aufnahmen zunächst in drei Szenarien auf:

- Public
- Mobil
- Home/Office

Daher bezeichnet die erste Stelle des zweistelligen Kürzels das Szenario wie folgt:

- **p** für **P**ublic
- **m** für **M**obil
- **h** für **H**ome/Office

Die zweite Stelle des Kürzels bezeichnet den Hauptanwendungsfall der vorliegenden Session. Etwaige Nebenanwendungsfälle werden im Rahmen der Annotationen ([bobtk00]) kommentiert.

Die Hauptanwendungsfälle werden nach Möglichkeit mit dem ersten Buchstaben des Namens gekennzeichnet. Es kann jedoch vorkommen, daß Anwendungsfälle, die ursprünglich als Hauptanwendungsfälle gedacht waren, zugunsten von anderen nur noch in Nebenanwendungen vorkommen. Diese werden ebenfalls durch eines der

nachfolgenden Kürzel gekennzeichnet, wenn die entsprechende Aufnahme hauptsächlich durch diese Anwendung gekennzeichnet ist.
Momentan werden folgende Hauptanwendungsfälle(Nebenanwendungsfälle) gekennzeichnet:

- **k** für **K**ino
- **t** für **T**ouristik
- **f** für **F**ernsehen
- **r** für **R**estaurant (nur in Nebenanwendungen)
- **n** für **N**avigation/Parkplatzsuche
- **i** für **i**nformation (nur bei englischen Aufnahmen)
- **m** für **M**usikauswahl
- **p** für **Pl**anung von Terminen
- **e** für **E**mail
- **x** für **F**ax
- **a** für **A**nrufen (=Telefonie) (nur in Nebenanwendungen)
- **h** für **H**otel
- **z** für **Z**eitplanung
- **w** für **W**etter (nur in Nebenanwendungen)
- **v** für **V**ideorekorder programmieren

Das Gesamtkürzel ergibt sich somit aus der Kombination aus Szenario und Hauptanwendungsfall, wobei die Variation des Szenarios konstant, die des Hauptanwendungsfalles noch erweiterbar ist.

4.2 Klassifikation der Biometrieaufnahmen

Die Biometrieaufnahmen lassen sich nicht wie die Wizard−Aufnahmen in Szenarien und Anwendungsfälle einteilen. Eine Klassifikation der Aufnahmen ist jedoch durchaus möglich (vgl. hierzu [berbm00]).

Konvention der Dateinamen

Sie werden wie folgt klassifiziert.

4.2.1 Handkontur

Die Aufnahme der Handkontur besteht nun mehr nur noch aus einer Geste:

1. offene Handkontur

Jede Aufnahmespezifikation kann dabei eindeutig im Szenariokürzel wiedergegeben werden, da eine Überschneidung der Aufnahmespezifikationen nicht erlaubt ist. Damit wird die Klassifizierung der Daten folgendermaßen vorgenommen:

- **lo** für die linke **o**ffene Handkontur bei der high−res−Aufnahme
- **ro** für die rechte **o**ffene Handkontur bei der high−res−Aufnahme

Wiederholungen werden über die Turnnummer geregelt. Siehe hierzu auch Abschnitt 6 bzw. die Beispiele in Abschnitt 10.

4.2.2 Handgesten−Aufnahme

Die Aufnahme der Handgesten (Kombinationsgesten offen−geschlossen−offen) besteht aus einer Aufnahme mit zwei optionalen Wiederholung, die ohne Modeschmuck für jede Hand durchgeführt werden:

- **bk** Kombinationsgeste, ausgeführt von beiden Händen, je zweimal: links−rechts−links−rechts.
- **lk** optionale Wiederholung ohne Modeschmuck links, zweimal ausgeführt.
- **rk** optionale Wiederholung ohne Modeschmuck rechts, zweimal ausgeführt.

Die Turnnummer bleibt hier auf 000. Die genaue Position der Gesten wird über die .seg−bio−Files ausgedrückt; hier ist Anfang und Ende jeder Geste in der Einheit Frames angegeben.

4.2.3 Sprecherverifikation

Die Biometrieaufnahmen zur Erforschung der Sprecherverifikation basieren auf Zufallszahlenlisten bestehend aus 18 Prompts, deren Zahlen aus dem Bereich 21 bis 99 stammen, wobei alle Zehner– und Einerstellen ungefähr gleichverteilt vorkommen. Die Prompts werden automatisch von einer speziellen Siemens–Software erzeugt.

Der erste Buchstabe lautet immer „s“, der zweite Buchstabe gibt die Liste an, aus der die Prompts stammen:

- **a** Liste 1
- **b** Liste 2
- **c** Liste 3
- **d** Liste 4
- **e** Liste 6
- **f** Liste 5
- **g** Liste 7
- **h** Liste 8
- **i** Liste 9
- **j** Liste 10

Das komplette Kürzel kann folgendermaßen dargestellt werden:

- **s[a–j]** für **S**precherverifikationsdaten nach Zufallszahlenlisten **a** bis **j**

Die Klammern geben lediglich das Auswahlarray an. Sie sind im Dateinamen nicht enthalten!

Beispiele zu den kompletten Dateinamen sind im Abschnitt 10 zu finden.

4.2.4 Klassifikation der Unterschrift

Wie im TechDok Nr. 6 [berbm00] ausführlich beschrieben, muß bei der Unterschriftenerfassung nicht nur das Original sondern auch die Fälschung dokumentiert werden. Dabei werden Unterschriften sowohl auf Papier als auch auf SIVIT–Folie aufgenommen.

Konvention der Dateinamen

Es ergibt sich für die erste Stelle des Kürzels die Art der Unterschrift, also:

- **o** für die **or**iginale Unterschrift
- **f** für die ge**f**älschte Unterschrift

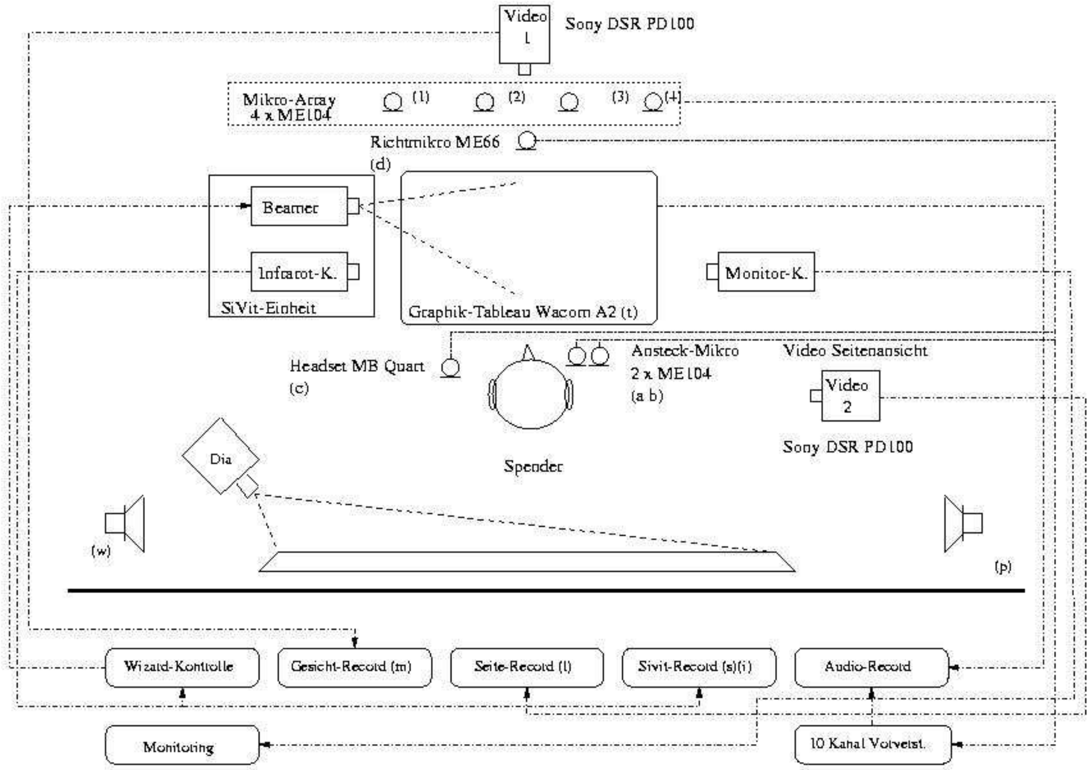
gefolgt von der Art der Schreibunterlage:

- **p** auf **P**apier
- **f** auf **F**olie

eines Sprechers XYZ. Es gilt zu beachten, daß die Fälschung der Unterschrift sich immer auf das Original beziehen muß! Es ist daher ausschlaggebend wessen Unterschrift gefälscht wurde und nicht wer die Unterschrift gefälscht hat.

5 Kennzeichnung des Aufnahmekanals

Die Kennzeichnung des Aufnahmekanals erfolgt duch ein einstelliges Kürzel. Dabei werden Mikrophon–, Video– und Koordinatendaten unterschieden.
Die nachfolgende Graphik illustriert diejenigen Aufnahmekanäle des Aufnahmesetups wie sie im weiteren Verlauf beschrieben werden.



5.1 Mikrophonkanäle

Sowohl in den Wizard–, wie auch in den Biometrieaufnahmen unterscheidet man prinzipiell zwischen vier Aufnahmearten:

- Nahbesprechungsaufnahmen
- Raumaufnahmen
- Aufnahmen mit dem Richtmikrophon
- Ausgabekanäle

5.1.1 Nahbesprechungsaufnahmen

Bei den Nahbesprechungsaufnahmen stehen die folgenden Kürzel für:

- **a** für Ansteckmikrophon 1
- **b** für Ansteckmikrophon 2
- **c** für Headset (**c**lose)

5.1.2 Raumaufnahmen

Die einzelnen Raummikrophone des Mikrophonarrays werden durchnumeriert:

- **1** für Raummikrophon **1**
- **2** für Raummikrophon **2**
- **3** für Raummikrophon **3**
- **4** für Raummikrophon **4**

5.1.3 Richtmikrophon

Die Aufnahmen mit dem Richtmikrophon (**d**irektional) wird gekennzeichnet durch:
- **d**

5.1.4 Ausgabekanäle

Zur Simulation von Umgebungsgeräuschen sowie für die Ausgabe der Sprachsynthese ist ebenfalls ein Kürzel reserviert:

- **w** für **W**izardsprachausgabe
- **p** für **P**layback Atmo vorn
- **q** für Playback Atmo hinten
- **o** für **O**perator bei den englischen Aufnahmen

5.2 Kanäle für die Registrierung von Koordinaten

Die Markierung der Ausgabe des Graphiktableaus bzw. die Speicherung der Positionen, die sich aus dem Sivit–System entnehmen lassen, ist wie folgt:

Konvention der Dateinamen

- **t** für **T**ableau−Daten
- **s** für **S**ivit−Koordinaten

Zu beachten ist hierbei, daß für die Biometriedatenerfassung keine Sivit−Koordinaten weitergegeben werden müssen.
Die Speicherung der Unterschriften erfolgt über den Graphiktableaukanal in einem von Siemens eigens definierten ASCII−Format.

5.3 Videokanäle

Von den insgesamt vier Videokanälen werden drei bei den Wizard−Aufnahmen gespeichert, nämlich:

- **i** für Sivit−**I**nfrarot−Output
- **m** für **M**imik
- **l** für Seitenansicht (**l**ateral)
- **o** für Beamer**o**utput
- **g** für **G**estik−Vierfach−Ansicht

Für die Biometrie ist wiederum nur ein Videokanal bei der Erfassung der Handkontur relevant:

- **i** für Sivit−**I**nfrarot−Output
- **h** für **h**igh−res−Sivit

6 Turnnummer

Nach der Registrierung von Aufnahmeart, Sessionnummer, Szenario− und Anwendungsfall sowie der Registrierung des Aufnahmekanals muß bei den Wizard−Aufnahmen die Turnnummer wiedergegeben werden.
Die 3stellige Turnnummer wird durch einen Unterstrich abgetrennt.

Konvention der Dateinamen

Die Nummerierung startet bei **001** und wird beim Turnwechsel automatisch um eins erhöht.

Man erhält folgende Notation (für den ersten Turn): **_001**

Wird die Turnaufteilung nicht verwendet, so steht als Turnnummer **unverändert**: **_000**.

7 VP−Kürzel

Ebenfalls abgetrennt durch einen Unterstrich folgt das 3stellige VP−Kürzel in Großbuchstaben:**_SPK**

Das VP−Kürzel wird automatisch durch eine zufällige Verkettung der Großbuchstaben des Alphabets generiert und ist eindeutig einem Spender zugeordnet.

Es gilt zu erwähnen, daß die Zuordnung zu schützenswerten persönlichen Daten (vgl. [berbm00]) über das Sprenderkürzel nicht möglich ist!

Es sollte ferner beachtet werden, daß sich das VP−Kürzel bei der Unterschriftenfälschung auf den Originalspender bezieht und nicht, wie ebenfalls möglich wäre, auf denjenigen, der die Unterschrift fälscht.

8 Übergeordnete Datenstruktur

Sämtliche Daten werden normalerweise im QuickTime−Format gespeichert.

Daher erhält das QT−Movie, das die Wizard−Aufnahmen referenziert, den Namen bestehend aus Sessionart (**w**), Sessionnummer (**001**), Szenariokürzel (**p, m, h**) und Hauptanwendungsfall (**k, t, f, r**).

Konvention der Dateinamen

Die Daten einer Session, sofern sie nicht im QT–Movie–Format referenziert werden (v.a. Biometrie!), erhalten als übergeordneten Dateiordnernamen die ersten 7 Zeichen des Dateinamens.

9 Markierung des Dateityps

Sowohl Aufnahme–, als auch zugehörige Annotationsdaten befolgen obige Richtlinien bei der Benennung. Welchem Dateityp ein vorliegendes File angehört, wird über die Endung wiedergegeben.

Es werden folgende Dateitypen in der Endung unterschieden:

qt	Quicktime–Movie (enthält sämtliche Aufnahmespuren)
spr	Spenderprotokolldaten
rpr	Recordingprotokolldaten
rpr–bio	Recordingprotokolldaten für die Biometrieaufnahmen
trl	Audioannotationsdaten
ush	**U**ser**S**tate **h**olistisch
usm	**U**ser**S**tate **m**imisch
usp	**U**ser**S**tate **p**rosodisch
ges	Gestikannotationsdaten
par	Partiturdaten
mar	Markerfile
mar–bio	Markerfile für Audio–Biometrieaufnahmen (Zahlenreihen)
seg–bio	Markerfile für Gesten–Video–Aufnahmen

Dabei verwenden alle Dateitypen die in Abschnitt 8 definierte übergeordnete Datenstruktur mit einer Ausnahme: Die Partiturdaten werden turnsegmentiert abgespeichert.

10 Beispiele für Dateinamen

Nachfolgend wird die in den vorherigen Abschnitten dargestellte Vergabe der einzelnen Kürzel durch Beispiele näher erläutert.

Wie im vorigen Abschnitt beschrieben, werden die übergeordneten Dateiordner bzw. das QuickTime–File durch die ersten sieben Zeichen des Dateinamens repräsentiert.
Demnach heißt das QT–File für die erste Wizardsession im Szenario Public mit Anwendungsfall Kino:
w001_pk.qt .
Beispiele für Dateinamen in den Biometriesessions wären:
b000_lki_001_SDF
Die erste Session der ersten Biometrieaufnahme der Kombination der Handkontur links als Sivit–Infrarot–Daten der VP SDF.
b001_fpt_001_SDF
wäre analog die erste Fälschung der Unterschrift der VP SDF auf Papier in der zweiten Biometriesession aufgezeichnet mit dem Graphiktableau.

Für die Wizard–Aufnahmen werden die Dateinamen in den einzelnen Spuren des QT–Files wiedergegeben. Es gilt zu beachten, daß nicht zu jedem Turn sämtliche Aufnahmekanäle registriert sein müssen. Denn nicht jede Spracheingabe ist z.B. begleitet von einem Gesten–Event.
Die Daten des ersten Raummikrophons der ersten Wizardsession im Szenario Public mit Hauptanwendungsfall Kino, erster Turn der VP ABC sähen folgendermaßen aus:
w001_pk1_001_ABC .
Andere Aufnahmekanäle im ersten Turn der ersten Wizardsession im Szenario Public mit Hauptanwendungsfall Kino der VP ABC könnten analog lauten:
w001_pks_001_ABC
Für die Sivit–Koordinaten,

Konvention der Dateinamen

w001_pkm_001_ABC

für die Mimikaufnahme oder

w001_pki_001_ABC

für die Infrarot−Aufnahme des Sivit−Systems.

Die Annotationen wären demnach abgespeichert in:

w001_pk.trl (spontanprachliche Transliteration der Session w001_pk)

w001_pk.ges (Gestikannotation der Session w001_pk)

w001_pk.mim (Mimikannotation der Session w001_pk)

w001_pk.spr (spenderspezifisches Protokoll der Session w001_pk)

w001_pk1_001_ABC.par (Partitur der ersten Wizard−Session im Szenario Public mit Hauptanwendungsfall Kino, erster Raumkanal, erster Turn, VP ABC)

11 Literatur

[bobtk00] Nicole Beringer, Daniela Oppermann, Susanne Burger: Transliteration spontanprachlicher Daten − Lexikon der Transliterationskonventionen − SmartKom(Version 1). SmartKom Technisches Dokument Nr. 2, Februar 2000.

[stein00] Silke Steininger: Aufnahmeplan Wizard−Of−Oz Aufnahmen. SmartKom Technisches Dokument Nr. 5, April 2000.

[berdb00] Nicole Beringer: VP−Datenbank. SmartKom Technisches Dokument Nr. 4, Mai 2000.

[berbm00] Nicole Beringer: Aufnahmeplan Biometrieaufnahmen. SmartKom Technisches Dokument Nr. 6, April 2000.