

Inhalt

- Einführung
 - Spezifikation
- Aufbau des Tools
 - Aufnahme
 - „Stille“ entfernen
 - Speicherung
 - Kommunikation mit den anderen Teilen

Einleitung

- Dieser Teil des Projektes beschäftigt sich mit der Implementation eines Tools zur Aufnahme von Sprachdateien.

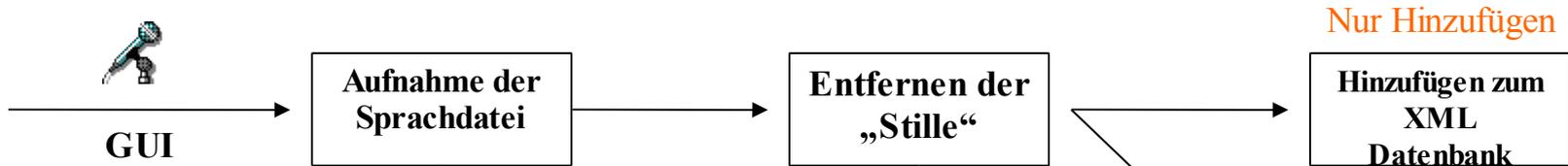
Spezifikation

- Das Programm muss Sprachdaten vom Mikrophoneingang lesen und speichern
- Die „Stille“ muss aus den Sprachdaten entfernt werden
- Die Benutzerdaten müssen in einer Datenbank abgespeichert werden
- Die Sprachdaten müssen an die anderen beiden Programmteile weitergegeben werden

Realisierung

- Zwei Teile:
 - Benutzer hinzufügen
 - 2 Sprachproben
 - Benutzer identifizieren
 - 1 Sprachprobe

Aufbau des Tools



Probleme

- Mögliche Probleme
 - Hintergrundgeräusche
 - Entfernung zwischen Sprecher und Mikrophon
 - Berührung des Mikrophons bei der Aufnahme
 - Beide Proben werden hintereinander aufgenommen

„Stille“ entfernen

- Warum muss die „Stille“ entfernt werden?
- Manuelles entfernen
- Statischer Schwellwert
 - Normalisierung
- Dynamischer Schwellwert
- Es wird eine modifizierte Version eines Programms von Tritonus.org benutzt

Manuelles entfernen

- Bestes Ergebnis
- Allerdings für Endsystem unmöglich

Statischer Schwellwert

- Es werden nur Frames aufgenommen, deren Wert über dem Schwellwert liegt
- Probleme durch unterschiedliche Lautstärken der Signale
- Probleme blieben auch bei Hinzufügen einer Normalisierung bestehen

Dynamischer Schwellwert

- Schwellwert wird dynamisch als Mittelwert der Mittelwerte der Frames berechnet
- Bessere Ergebnisse als mit statischem Schwellwert allerdings schlechtere Ergebnisse als beim manuellen Herausschneiden
- Die Ergebnisse wurden durch betrachten der Signalverläufe bewertet.

Speicherung

- XML-Datenbank ist flexibel und Plattformunabhängig
- Inhalt der Datenbank:
 - Benutzer-Identifizierungsnummer
 - Benutzername
 - Passwort
 - Individueller Schwellwert
 - Namen der beiden Sprachdateien

XML Datenbank - Beispiel

<users>

<user idUsr=„xx“>

<name>testName</name>

<password textPass=„Hallo maschine“ threshold=„x“>

<passfile version=„1“>xxtestName1.wav</passfile>

<passfile version=„2“>xxtestName2.wav </passfile>

</password>

</user>

</users>

Kommunikation mit den anderen Teilen

- WAV Dateien
- Übergeben der Dateinamen der Sprachdateien an das Tool zur Merkmalsgewinnung.
- Übergeben der Dateinamen und Schwellwerte aus der XML-Datenbank an das Tool zum Ähnlichkeitsvergleich.

Implementation

- Java SDK 1.4.2
 - Java Sound (javax.sound)
 - XML Parser (org.w3c.dom)
- Tritonus (www.tritonus.org)
 - Unabhängige Implementation der Java Sound API
 - Benutztes Programm:
„SilenceSupressingAudioRecorder.java“