

Automatické titulkování televizních pořadů

Technologie pro lidi se SP

- především ty, které zprostředkovávají mluvenou řeč v psané formě
- manuální přepis: eScribe, titulky v televizi
- automatický přepis:
 - automatické titulkování
 - převod řeči do textu – Newton Technologies + TUL Liberec
 - Newton Dictate, MyVoice, MyDictate
 - Newton Dictate Mobile
 - využitelné i pro osoby s tělesným postižením

Titulky

- otevřené x skryté titulky
- **otevřené**
 - využívají se zejména u nedabovaných filmů v kinech
 - zpravidla jednobarevné, bez doplňkových informací o zvucích
- **skryté**
 - využívají se v televizi pro potřeby diváků se sluchovým postižením (ale nejen pro ně)

Skryté titulky



- teletext strana 888
- (téměř) doslovný přepis
- barevné odlišení mluvčích
- informace o doprovodných zvucích (klepání, štěkot psa, hra na housle)

Skryté titulky

- v ČR povinnost titulkovat určité procento pořadů
- soukromé televize: 15 % pořadů
- veřejnoprávní televize: 70 % pořadů se ST + 2 % pořadů ve znakovém jazyce

Metody titulkování

- předtočené pořady (dokumenty, filmy, ...)
 - titulky se připraví předem
- živě vysílané pořady (zprávy, přímé přenosy)
 - živé titulkování
 - titulkář
 - počítač

Automatické titulkování

- spolupráce ČT a katedry kybernetiky FAV ZČU v Plzni
- zvuková stopa putuje z Kavčích hor do Plzně, tam se zpracuje a titulky se posílají zpět do Prahy, odkud jdou do vysílání
- využití stínového řečníka – speciální trenažér a tréninkový program

Stínový řečník

- využít při titulkování většiny pořadů



- úkoly:
- poslouchá stopu pořadu
- diktuje do počítače
- vybírá řeč hlavního řečníka
- indikuje změnu řečníka
- koriguje mluvu řečníka
- zadává interpunkci
- opravuje text
- přidává nová slova do slovníku

Automatické titulkování

- první pokusy v roce 2008
 - plně automatické titulkování záznamů z parlamentu
 - úspěšnost cca 85 %
 - čtený projev, relativní klid a ticho
 - pokusy s titulkováním hokeje
 - úspěšnost přímého titulkování cca 75 %
 - se stínovým řečníkem 90 %

Aktuální situace

- od jara 2010 – titulkování záznamů z parlamentu
- od jara 2011 – titulkování Otázek Václava Moravce
- od jara 2012 – zkušební titulkování Hyde Parku



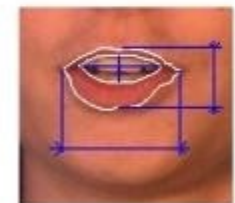
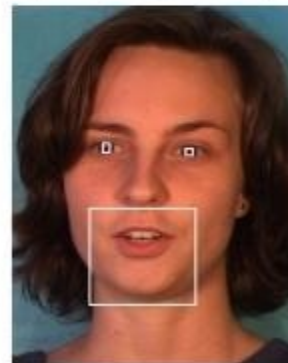
- květen 2012 – titulkování MS v hokeji
- červen 2012 – titulkování ME ve fotbale
- podzim 2012 – titulkování Star Dance

Problémy

- rozpoznávání mluvené češtiny obecně je náročné
- nutnost budování obsáhlých tematických slovníků
- zpoždění cca 5 sekund – může být matoucí
- neochota ČT financovat větší využití

Lehce mimo téma

- ZČU pracuje i na dalších projektech:
 - audiovizuální rozpoznávání a syntéza řeči
 - rozpoznávání a syntéza znakového jazyka



Děkuji za pozornost.