

Řečové technologie: od výzkumu k praxi

Datum konání: 21.05.2021

Místo konání: online (prostřednictvím MS Teams)

Název přednášky: Řečové technologie: od výzkumu k praxi

Přednášející: Doc. Ing. Jindřich Matoušek, Ph.D, Ing. Jan Švec, Ph.D

Počet účastníků: N/A

Zpracoval: Dušan Ptáček



Počítačová lingvistika (PLIN)
uvádí

**Přednáškový cyklus
externích odborníků**

doc. Ing. Jindřich Matoušek, Ph.D.
Ing. Jan Švec, Ph.D.

Katedra kybernetiky
Západočeská univerzita v Plzni

**Řečové technologie:
od výzkumu k praxi**

21.05.2021 • 13:00 • MS Teams
<https://bit.ly/3vXbvM6>

Doc. Ing. Jindřich Matoušek, Ph.D

Kybernetik a vysokoškolský pedagóg, pôsobí na Fakulte aplikovaných vied Západočeskej univerzity. Medzi jeho oblasti záujmu patria syntéza reči z textu (TTS), fonetika & fonológia, akustika či modelovanie, segmentácia a spracovanie prirodzenej reči. Je tiež spoluautorom monografie *Mluvíme s počítačem česky* a ďalších viac než 150 odborných publikácií. [1][2]

Ing. Jan Švec, Ph.D

Kybernetik a vysokoškolský pedagóg, momentálne pôsobí na Fakulte aplikovaných vied Západočeskej univerzity. Venuje sa predovšetkým rozpoznávaniu reči a hlasovým dialógovým systémom, ale všeobecnejšie jeho výskumné témy možno zaradiť pod oblasť strojového učenia a neurónových sietí. V rámci spomenutých oblastí záujmu napísal viacero odborných textov.

Rečové technológie: Od výzkumu k praxi

Prezentácia bola rozdelená na dve časti. Tej prvej sa zhostil doc. Matoušek so syntézou reči—TTS alebo Text-to-speech systém, tú druhú časť viedol dr. Švec s témou rozpoznávania a porozumenia reči.

Syntéza reči

Je to systém, ktorý umožňuje prevod písaného textu na reč bez asistencie človeka. Cieľom je skonštruovať reč z akéhokoľvek textu, pričom tento proces má niekoľko úrovní. Písaný text sa prevedie do podoby hlások a pridá sa tiež prozódia reči. Následne sa z tejto reprezentácie pomocou syntetizéra reči vytvorí samotná reč alebo rečový signál. Treba dodať, že algoritmus, ktorý číta text bez asistencie človeka, by mal vedieť prečítať aj nezmyselné výrazy.

Po ozrejmení základných princípov syntézy reči nás doc. Matoušek ďalej zoznámil s prístupmi k syntéze reči. Menovite ide o signálový prístup a modelový

či generatívny prístup.

Pri signálovom prístupe prebieha syntéza výberom jednotiek (unit selection), zatiaľ čo modelový prístup využíva viac štatistické modely. Medzi tieto modely patria skryté Markovove modely (HMM) a hlboké neurónové siete (DNN). Bez zachádzania do podrobností a pre zaujímavosť spomeňme, že WaveNet ako vokodér v DNN syntéze je v súčasnosti pomerne populárna výskumná téma.

Hodnotenie kvality syntetickej reči

Doc. Matoušek upozornil, že prakticky neexistujú také kritériá či konkrétna miera, pomocou ktorých by sme dokázali kvalitu syntetickej reči objektívne zhodnotiť. Reč je mimoriadne komplexná a rôzni poslucháči vnímajú reč odlišne. V praxi sa teda používajú posluchové testy a snahu objektivizovať hodnotenie zabezpečuje vyšší počet poslucháčov. Hodnotenie je teda akosi „sumou“ všetkých subjektívnych hodnotení jednotlivých poslucháčov na základe rôznych typov testov—testov prirodzenosti, zrozumiteľnosti alebo preferenčné testy.

„Tady dcera doc. Matouška”

Pre ilustráciu rozmanitého využitia syntézy reči v praxi boli pustené ukážky rôznych hlasov. Mužských i ženských, ba dokonca bol pustený hlas dcéry prednášajúceho. Veľmi zaujímavým projektom je tzv. konzervácia hlasov pacientov pred úplnou stratou hlasu. Dôležitým projektom je tiež projekt s personalizovanými hlasmi. Počuli sme napríklad hlas Václava Havla a čo bolo prekvapujúce – tento model vznikol na základe len 20 minút nahrávok z Lán.

Výskumníci pracujú aj na včlenení paralingvistických javov do syntézy reči, aby reč pôsobila prirodzenejšie. Využitie syntézy reči nachádzame trebárs i v metre v Prahe či pri automatickom dabingu.

Rozpoznávanie reči

Zjednodušene možno napísať, že rozpoznávanie reči je procesom opačným k procesu syntézy reči. Teda ide o automatický prevod hovoreného slova na text. dr. Švec spomenul na úvod svojej časti, že rozpoznávanie reči má rôzne úlohy v praxi. Menovite ide o telefónnu reč, diktovací software alebo tzv. domain-specific speech recognition.

Pri rozpoznávaní reči využívame niekoľko modelov. Ide o model zašumeného kanálu, akustický model alebo jazykový model. I v pozadí modelov rozpoznávania reči nájdeme matematické koncepty ako napríklad Bayesova veta. Tak isto neurónové siete, na ktorých sú založené jazykové modely LSTM alebo Word2vec.

Ďalším aspektom rečových technológií je porozumenie hovorenej reči. To je založené na znalostiach (napr. bezkontextové gramatiky + parsing) alebo na štatistike, prípadne na oboch prístupoch zároveň (čoho príkladom sú dialógové systémy).

Ukážky rozpoznávania reči

Keď pozeráme na živý prenos v televízii, ktorý je zároveň titulovaný, môžeme vidieť v priamom prenose i proces rozpoznávania reči, resp. výstup procesu rozpoznávania reči. Medzi ďalšie spomenuté aplikácie rozpoznávania reči patrí platforma SpeechCloud pre vývoj multimodálnych hlasových dialógov či inteligentné technológie pre zvýšenie bezpečnosti letovej prevádzky.

Záverečné zhodnotenie prednášky

Prednáška bola veľmi zaujímavá a podnetná. V mnohom naznačila, že rečové technológie budú prenikať do čoraz viac oblastí a že mnohé sci-fi predstavy sa už stali súčasťou dnešnej reality.

Na konceptuálnej úrovni bola prednáška odprezentovaná zrozumiteľne, hoci poslúchačov, ktorí sa neuspokoja s čiste konceptuálnym porozumením predneseného, čaká štúdium niektorých matematických konceptov, ktoré

prekračujú úroveň matematických predmetov určených študentom počítačovej lingvistiky.

A čo nás ako skupinu poslucháčov najviac zaujalo ?



Zdroje

[1] Matoušek, Jindřich: *Profil pracovníka*. Katedra kybernetiky, FAV ZČU, Plzeň.
[<http://www.kky.zcu.cz/cs/people/matousek-jindrich>]

[2] Encyklopedie Plzeň: Profil osobnosti - Matoušek, Jindřich.
[https://encyklopedie.plzen.eu/home-mup/?acc=profil_osobnosti&load=2807]

[3] Švec, Jan: *Profil pracovníka*. Katedra kybernetiky, FAV ZČU, Plzeň.
[<http://www.kky.zcu.cz/cs/people/svec-jan>]

[4] Švec, Jan: *Osobný profil na sociálnej sieti LinkedIn*.
[<https://cz.linkedin.com/in/jansvec>]