

# Vpliv kakovosti govora v telefoniji na samodejno razpoznavanje govorca

Robert Blatnik<sup>1,2</sup>

<sup>1</sup> Odsek za inteligentne sisteme, Institut Jožef Stefan, Ljubljana

<sup>2</sup> Mednarodna podiplomska šola Jožefa Stefana (Novi mediji in e-znanost, 2. letnik)

robert.blatnik@ijs.si

Samodejno razpoznavanje govorcev je uporaba stroja za identifikacijo ali verifikacijo identitete oseb na osnovi njihovega glasu [1], [2], [3]. Samodejno razpoznavanje govorcev spada med dosežke znanosti, ki v splošnem porastu varnostnih biometričnih metod postajajo vedno bolj aktualni v industriji in storitvah. Tehnologije za samodejno razpoznavanje govorcev se uporabljajo v telefonskem bančništvu, biometričnih sistemih vstopne kontrole, forenzičnih preiskavah kriminalnih dejanj in drugih aplikacijah, kjer obstaja potreba po razpoznavanju identitete osebe na osnovi zajetega govornega signala.

Kakovost govornega signala je ključna za uspešno razpoznavanje govorca. To je še posebej kritično v aplikacijah, kjer se govor zajema v spremenljivih pogojih, ki vplivajo na govorca ali na tehnologijo prenosa, npr. telefonska povezava. Kakovost prenosa govora, kot osnovna lastnost tradicionalne telefonske storitve, je s paketnim načinom prenosa v fiksni in mobilni telefoniji postavljena pred nove izzive, kot so potreba po hkratnih prenosih velikih količin podatkov in govora v omrežju, razvoj novih kodirnih mehanizmov, razvoj metod obnavljanja izgubljenih paketov in drugih.

Pri razvoju sistemov za razpoznavanje govorcev je potrebno analizirati vse zunanje vplive na kakovost prenosa govornega signala. V ta namen izvajamo objektivne meritve kakovosti prenosa govora in preiskujemo v kolikšni meri vpliva kakovost prenesenega govornega signala na uspešnost razpoznavanja govorca. Zgradili smo okolje za izvedbo objektivnih meritev kakovosti prenosa govora v fiksni, mobilni in IP telefoniji ter brezžičnih lokalnih omrežjih (WLAN) pri različnih pogojih delovanja, npr. različnih obremenitvah omrežja, različnih nivojih RF signala v WLAN, itd. [4]. Meritve izvajamo po standardizirani metodi za objektivno ocenjevanje kakovosti govora (ang.: Perceptual Evaluation of Speech Quality, PESQ) [5]. Algoritem metode PESQ vsebuje računalniški psihoakustični model poslušalca, s katerim na osnovi računske primerjave referenčnega (vhodnega) in testiranega (izhodnega) govornega signala objektivno izračuna oceno kakovosti. Govorni signali, ki so uporabljeni v meritvah PESQ, hkrati služijo kot osnova za izvajanje razpoznavanja govorcev in preučevanje vpliva kakovosti prenosa govora na uspešnost sistemov za razpoznavanja govorcev. Pri tem uporabljamo lastne govorne posnetke v slovenskem jeziku in posnetke iz tujih javno dostopnih govornih korpusov.

Nove metode za razpoznavanje govorcev je nujno potrebno prilagajati različnim pogojem, ki vplivajo na kakovost prenosa govora v telekomunikacijah. Preučevanje vplivov kakovosti prenosa govora na sisteme za razpoznavanje govorca je pomembno za razvoj učinkovitih in zanesljivih sistemov za biometrično razpoznavanje govorcev.

## Literatura:

- [1] P. Rose. *Forensic Speaker Identification*. London in New York: Taylor & Francis. 2005.
- [2] J. Benesty, M. M. Sondhi, Y. Huang (Eds.). *Springer Handbook of Speech Processing*. Springer-Verlag Berlin Heidelberg. 2008.
- [3] T. Šef. Razpoznavanje (identifikacija/ verifikacija) govorcev v forenzične namene. *Zbornik 10. mednarodne multikonference Informacijska družba IS 2007*. Institut "Jožef Stefan". Ljubljana. 2007.
- [4] R. Blatnik, G. Kandus, T. Javornik. VoIP/VoWLAN system performance evaluation with low cost experimental test-bed. *WSEAS trans. commun.* 2007.
- [5] A. W. Rix, J. G. Beerends, M. P. Hollier, A. P. Hekstra. Perceptual evaluation of speech quality (PESQ) - a new method for speech quality assessment of telephone networks and codecs. *IEEE ICASSP 2001*. Salt Lake City, UT, USA. 2001.