

# Variációk beszédfelismerés, információkeresés és gépi fordítás használatára

Gyarmati Ágnes

Dublin City University, Centre for Digital Video Processing

Az egyre nagyobb mennyiségben tárolt, akár szabadon is hozzáférhető hang- és videófelvevételek akkor válnak igazán értékké, ha a szolgáltatók nemcsak tárolást és hozzáférést biztosítanak a digitális tartalmakhoz, hanem a felhasználóknak lehetőségük is van a videók és hangfájlok között keresni, a számukra érdekes tartalmakat könnyen megtalálni.

A hangzó szövegben való keresés hatékonyságának fejlesztését tűzte ki egyik céljával a *TREC* kutatói fórum az 1990-es évek közepén, amerikai rádiós híradásokra alapozva a kutatási irányt. Néhány év elteltével a próbálkozásokat sikeresnek, a problémát gyakorlatilag megoldottnak tekintették [1]. A rádiós hírműsorok azonban viszonylag zárt szókincsű szövegek, melyeket a bemondók többnyire szépen artikulálva olvasnak fel, így a jó minőségű, zajmentes stúdiófelvételek csak a kezdeti lépést adják élőbeszéd és információkeresés viszonyának vizsgálatában. További kutatásokra például a *CLEF* fórum beszéd-visszakereső modulja mutatott új irányt, mely az elhangzott szövegek automatikus leírásának minősége mellett a címkézés szerepét is górcső alá vette [3].

A *VideoCLEF 2009* két feladata is más megközelítésből nyújtott vizsgálati lehetőséget élőbeszéd és keresés kapcsolatához, holland nyelvű tévéműsorok felhasználásával [2]. Az egyik az automatikus címkéiosztást információkeresési módszerekkel irányozta elő megoldani, míg a másikban, az úgynevezett „Összekapcsoló feladatban” (*Linking Task*) egyes rövid, pár másodperces műsorrészletekhez kellett a Wikipédiában tartalmilag kapcsolódó oldalakat keresni. A feladat nehézségét a híryanagoknál tapasztaltaknál spontánabb beszéd és változatosabb tartalom mellett a nyelvi csavar adja: a dokumentumsorozat *holland* nyelvű, míg a linkelésnél az *angol* Wikipédia legmegfelelőbb szócikkeit kellett megtalálni. A feladattal haladó nyelvtanulót szimulálhatunk, aki már idegen nyelven is képes tévéműsorok (részleges) befogadására, de még szükségét érzi, hogy a számára különösen érdekes vagy nyelvi hiányosságai miatt magyarázatként szükséges kiegészítő információkhoz anyanyelvén, vagy legalábbis egy általa magasabb szinten beszélt (értett) harmadik nyelven juthasson hozzá (a feladatkiírásban angolul).

A feladat megoldásánál alapvetően két megközelítés kínálkozik. Az eredeti holland (beszéd-felismerővel generált) szöveg használatával a holland Wikipédiában is végezhetünk keresést, és ekkor csak utolsó lépésként, a Wikipédia saját nyelvközi linkjeit követve jutunk el az eredmény angol nyelvű cikkekhez. Másrészt a holland szöveget rögtön áttehetjük angolra, és így közvetlenül kereshetünk az angol Wikipédiában. Az eddigi eredmények azt mutatják, hogy a keresés szempontjából nem hatékony az átiratot automatikusan lefordítani, hanem érdemes a keresést inkább a forrásnyelven végezni.

## Hivatkozások

1. John S. Garofolo, Cedric G. P. Auzanne, és Ellen M. Voorhees. The TREC spoken document retrieval track: A success story. In *Text Retrieval Conference (TREC) 8*, 16–19, (2000).
2. M. Larson, E. Newman és G. J. F. Jones, Overview of VideoCLEF 2009: New Perspectives on Speech-based Multimedia Content Enrichment, In *Proceedings of the CLEF 2009: Workshop on Cross-Language Information Retrieval and Evaluation*, Korfu, Görögország, (2009). Megjelenés alatt.
3. D. W. Oard, J. Wang, G. J. F. Jones, R. W. White, P. Pecina, D. Soergel, X. Huang és I. Shafran, Overview of the CLEF-2006 Cross-Language Speech Retrieval Track, In *Proceedings of the CLEF 2006: Workshop on Cross-Language Information Retrieval and Evaluation*, Alicante, Spanyolország, Springer (2006)