

PACS: beszédvezérelt POI-kereső szolgáltatás

Csáki Tibor, Vajda Péter, Vámosi János

Ygomi Europe Kft.,
4034 Debrecen, Vágóhíd utca 2.
{csaki, vajda, vamosi}@connexis.com

Kivonat: Az utóbbi években a kézisámítógépek elterjedése és a térképes alkalmazások fejlődése egyre több ember számára teszi mindennaposá a navigációs rendszerek használatát. Az autóban, vezetés közben használt rendszerek veszélye, hogy a vezető figyelmét könnyen megoszthatják. Ennek kiküszöbölésére egy olyan felhasználói felületet terveztünk, amely beszéd felismerésen alapul, így mentesíti a vezetőt a navigációs eszköz kézzel való használatától, valamint lerövidíti az információ megszerzéséhez szükséges időt. A felhasználói felület mögött komplex dialógusrendszer működik, ami szükség esetén emberi operátor beavatkozását is lehetővé teszi.

1 A PACS rendszer

A navigációs rendszerek fejlődése, a hordozható, GPS-sel felszerelt kézisámítógépek, telefonok elterjedésével egyre többen vesznek igénybe telematikai szolgáltatásokat¹. Ezzel a fejlődéssel párhuzamosan a felhasználók igényei is megnövekedtek: az útvonaltervezésen és a POI (Point of Interest) keresésen kívül ma már a helyfüggő, folyamatosan elérhető, valós idejű adatokat (pl. közlekedési információkat) kínáló szolgáltatások is mindennaposak.

Az autóvezetőknek nyújtott szolgáltatások egyik fontos kritériuma, hogy az alkalmazott eszköz a vezetők figyelmét a lehető legkisebb mértékben vonja el a forgalomtól a biztonságos közlekedés érdekében. Az autógyártó és telematikai cégek jelenleg kétféle megoldást alkalmaznak, azonban a fenti szempontból ma még ezek egyike sem kielégítő. Az egyik módszer az autóba beépített beszéd felismerő használata, amely még ma is egy drága megoldás és jelenlegi szintjén nem képes megbirkózni azokkal a nem ritkán komplex kérésekkel, amiket egy navigációs rendszer felé tehetünk. A másik megoldás az ügyfélszolgálati rendszereké, ahol a vezető egy telefonközpontossal kommunikálva jut a számára érdekes információhoz. Azonban ez a megoldás a telefonos operátorok alkalmazása miatt nem költséghatékony és a hívások időtartama, ezzel pedig költsége is nagyobb lehet, mint amit egy autóvezető elfogadhatónak tart.

A Ygomi Europe Kft.-nél megvalósított PACS (People Assisted Computer System) rendszer egy olyan helyfüggő navigációs szolgáltatás folyamatát írja le, amely a fent

¹ Projektünket a Nemzeti Kutatási és Technológiai Hivatal a Jedlik Ányos program keretében támogatta (szerződésszám: OM-00102/2007).

említett két módszer előnyeit próbálja meg ötvözni. Tipikus felhasználási módja, amikor egy autóvezető útközben egy számára érdekes helyet (POI-t) keres, és ennek adatait szeretné letölteni gépkocsija navigációs rendszerébe. A PACS rendszer jellemzője a szerveroldali beszédfelismerés, és az automatikus beszéd szintetizátorral (TTS-sel) előállított válasz, így a vezető egy rövid, géppel folytatott párbeszéd végén letöltheti autója navigációs rendszerébe a kívánt adatokat. Ugyanakkor a vezetőnek – ha szükséges – lehetősége van telefonos operátorral is felvenni a kapcsolatot, pl. ha bonyolultabb kérdést akar feltenni, vagy ha a beszédfelismerőnek nem sikerült egyértelműen felismerni a kérést. Utóbbi esetben a rendszer egy ún. Silent Agentet (néma operátort) használ, aki visszahallgatja a felvett kérést, és ha egyértelműen meg tudja válaszolni, akkor a megfelelő választ szintén TTS segítségével el tudja juttatni a vezetőhöz. Ennek a megoldásnak az előnye, hogy az operátornak nem kell beszélgetést folytatnia a klienssel, így a kérés kiszolgálása hatékonyabb és olcsóbb, mivel a „beszélgetés” időtartama rövidebb. A PACS korábbi prototípusait több európai autógyártóval is teszteltük. Ezekből a tesztekkel kitűnik, hogy a hagyományos telefonközpontos megoldáshoz képest feleannyi az átlagos párbeszéd ideje, ha Silent Agentet használunk, illetve negyedannyi, ha az operátornak nem kell közbeavatkozni, azaz ha a beszédfelismerő helyesen működik [1]. Az operátornak csak akkor kell valódi párbeszédet folytatni az autó vezetőjével, ha a felvételtől nem képes megállapítani a kérés tartalmát, vagy ha a vezetőnek az elsőre felajánlott válasz nem felel meg.

2 Megvalósítás

Az általunk bemutatott szoftver egy, a Ygomi Europe Kft.-nél elkészített PACS rendszert megvalósító demóalkalmazás, amely mobil kliensen keresztül is elérhető.

Az alkalmazás egyik fontos része a dialógusmenedzser, amely a tipikusan használt párbeszédet vezérli. A dialógus véges állapotú automatával írható le, és a rendszer konfigurálásakor megadható. A dialógusok tervezésénél figyelembe vettük, hogy a vezető figyelmét minél kevésbé vonjuk el a vezetéstől. Ennek érdekében a dialógus úgy épül fel, hogy a gyakrabban előforduló kérések esetében a felhasználó kevesebb lépésben tudjon eljutni az eredményhez, valamint a rendszer által elmondott promptok hosszát is igyekeztük minimalizálni.

A rendszer fontosabb komponensei közé tartoznak a beszédfeldolgozást megvalósító modulok, ilyen a Carnegie Mellon University-vel (CMU) közösen fejlesztett Sphinx [3] beszédfelismerő rendszer, amelyhez nyelvi modelleket fejlesztettünk ki. A modellek építéséhez felhasználtuk a saját hanganyaggyűjtésből származó tipikusan előforduló kéréseket. Beszédszintetizátorként a Nuance RealSpeak nevű termékét használjuk, szemantikus elemzőként pedig a szintén a CMU-n készült Phoenix rendszert [2]. Ezek a nyelvi modulok más alkalmazásokra is kicserélhetőek.

A rendszer több nyelvre készült el, lehetőség van a magyaron kívül angol, francia, német, spanyol és olasz nyelven is kéréseket megfogalmazni.

Hivatkozások

1. Masson, J.: Innovative Strategies For Improving Telematics Call Centre Operations. The Fully Networked Car Workshop ITU (2009)
2. Ward, W. H.: The Phoenix System: Understanding Spontaneous Speech. In: Proceedings of IEEE ICASSP (1991)
3. The CMU Sphinx Group Open Source Speech Recognition Engines
<http://cmusphinx.sourceforge.net>