

Hangvezérlő rendszer integrálása a járműbe a kényelmi és biztonsági szint emelése érdekében

Dr. Titrik Ádám*, Prof. Dr. habil Lakatos István, Hollósi János

*Széchenyi István University, 9026 Győr, Egyetem tér 1. Hungary

(Tel: 06 96 503-311; e-mail: titrika@ga.sze.hu)

Abstract: Az autonóm járművek megjelenése határán még nem célszerű a hagyományos vezető által irányított járművek fejlesztésének háttérbe szorítása - igenis van értelme ezen járművek fejlesztésének főleg úgy, hogy az egyes rendszerek/fejlesztések integrálást nyerhetnek az autonóm járművek területén is. A tanulmány célja a járművek egyes egységeinek hangvezérlési lehetőségének bemutatása, ezzel emelve a járművezető járműkényelmi és közlekedésbiztonsági szintjét. A tanulmány egy alap felhasználási lehetőséget ismertet, majd a járműbe történő integrálás folyamatát szemlélteti közlekedésbiztonság és üzembiztonsági szempontok figyelembevételével.

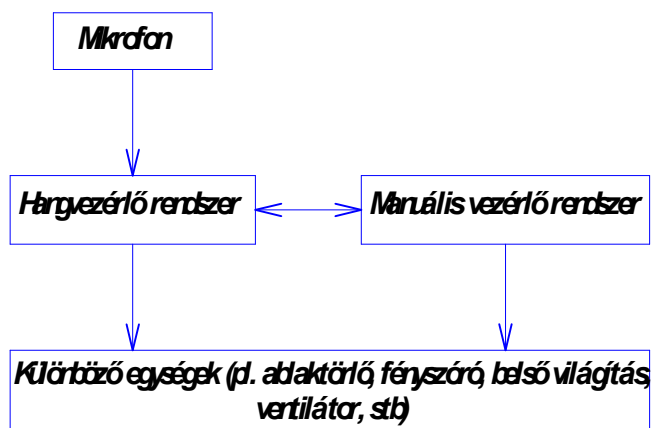
1. BEVEZETŐ

A hangvezérlő rendszert napjainkban igen elterjedtnek lehet mondani. Az egyes számítógépes applikációkat is lehetséges hangvezérelni, azonban kiemelkedő áttörést az android, IOS operációs rendszerrel működtetett okostelefonokon láthatunk. Az egyes funkciók, mint például telefonhívás kezdeményezése, vagy e-mail diktálása, vagy a böngésző keresőszavának megadása hangvezérléssel lehetséges. A hangvezérlés megvalósításához beszédhang-felismerő szoftver került felhasználásra. Több ilyen jellegű megoldás is fellelhető a piacon, melyek két fő csoportba oszthatóak: online és offline rendszerek. Online rendszer esetén kitétel, hogy a számítógép állandó internet kapcsolattal rendelkezzen, ebben az esetben a szoftver csak a kliens oldali szerepet látja el, a tényleges feldolgozás egy távoli számítógépen zajlik. Az offline megoldások esetén a tényleges számításokat végző szoftveregység is a felhasználó birtokában van, így egy számítógépen belül is megoldható a feldolgozás. Az online rendszerek előnye, hogy általában magasabb megbízhatósági fokú szoftverrendszert nyújtanak, mint az offline megoldások, ellenben az offline megoldás mellett szól az, hogy nem igényel internet kapcsolatot, míg az online rendszerhez állandó és biztos kapcsolatra van szükség. Jelen fejlesztés körülményei egy offline hangfelismerő rendszer alkalmazását igényli, hiszen a szoftver rendszer egy közúti járműbe kerül integrálásra, ahol az állandó és magas rendelkezésre állású internetkapcsolat nem kivitelezhető. Több különféle gyártótól származó offline hangfelismerő rendszer áll rendelkezésre, ezek közül több ingyenes is alkalmazható. Fontos megjegyezni, hogy a jelenlegi piacvezető mobiltelefon és táblaszámítógép operációs rendszer gyártók mind rendelkeznek saját hangfelismerő modullal, melyek fejlesztési célra ingyenesen használhatóak. Ezek a modulok jellemzően magas megbízhatósági fokkal rendelkeznek és jól optimalizáltak az adott keretrendszerre.

2. “VOICE CONTROL” RENDSZER ISMERTETÉSE

A rendszer működésének alapja offline rendszerű szoftver alkalmazása, valamint az egyes eszközök vezérlési jogosultságának megszerzése (1. ábra).

A voice control rendszer fő működési struktúrája a következő:



1. ábra: Voice control rendszer fő működési struktúrája

A kiadott hangot – hangvezérlést – mikrofonon keresztül a hangvezérlő rendszer fogadja. A hangvezérlő rendszer összekötöttesben áll a manuális vezérlő rendszerrel a parancs végrehajtási lehetőségének lekérdezése céljából. A hangvezérlő rendszer kapcsolatban áll olyan jelző rendszerrel amely jelzi az utasítást kiadó személy számára, hogy a parancs értve és teljesítve, vagy a parancs elutasításra került. A jelzés lehet vizuális, hang vagy rezgés módú.

A hangvezérlő rendszer a manuális vezérlő rendszert lekérdezve - azaz, hogy a működtetni kívánt feladat már előzőleg manuálisan elvégzésre került – és végrehajthatósági

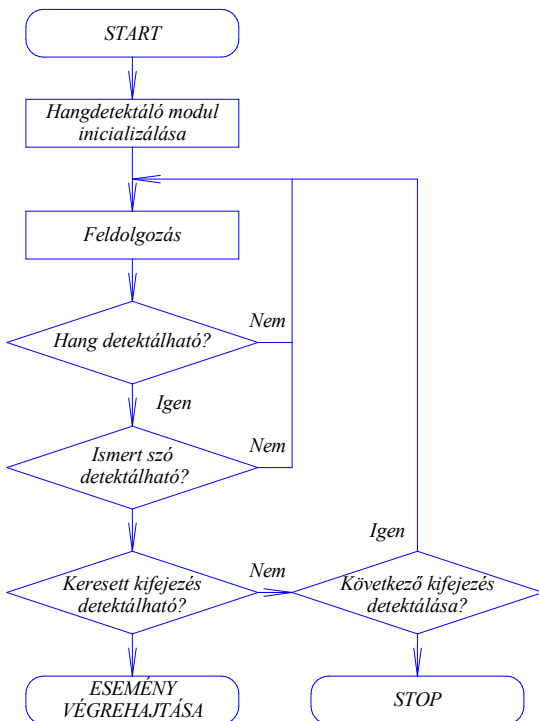
választ kapva az adott utasítást elvégzi. További lehetőség az, hogy az utasítás végrehajtása mindenképp végrehajtásra kerül függetlenül attól, hogy ez az utasítás előzőleg az adott elemnél aktiválásra került.

3. A RENDSZER ALAPJAINAK ISMERTETÉSE

A „Voice control” rendszer kiépítéséhez első sorban az alapszinten kezdtük a fejlesztést. A kutatások alapján a választás a Microsoft saját hangfelismerő rendszerére esett. A Microsoft rendszerét .NET keretrendszer alatt lehet használni, ahol a System.Speech.Recognition névtér alatt érhetjük el a szükséges eszközöket. Ebből kifolyólag a szoftver ezen részegysége .NET keretrendszer alatt, C# programozási nyelven készült. A Microsoft szoftvere alapvetően arra ad lehetőséget, hogy egy előre meghatározott felépítésű kifejezést ismerjünk fel vele, ami konstans elemekből épül fel, mely részben tartalmazhat váltakozó kifejezéseket.

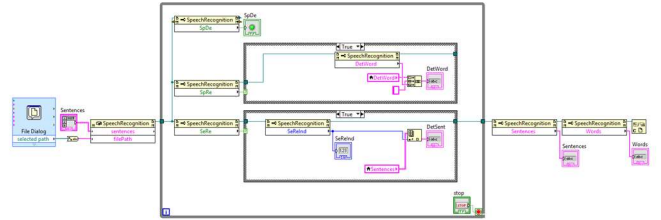
Formálisan megfogalmazva a hangfelismerő rendszer olyan kifejezések detektálását teszi lehetővé, ahol a felismert kifejezés: az „ahol” „minden” és az „üres-sztring”. A probléma ezzel a megközelítéssel, hogy nem ad lehetőséget egymástól teljesen eltérő felépítésű kifejezések felismerésére. Jelen fejlesztésnél fontos, hogy több különböző kifejezés is felismerhető legyen egyidejűleg, hiszen az egyes ablakok megjelenítéséhez más-más kifejezésre lesz szükséges.

A beszédfelismerő alkalmazás alkalmassá vált arra, hogy egyidejűleg több különböző felépítésű kifejezést is képes legyen detektálni, ami az adott fejlesztés szempontjából fontos tényező, így bármilyen kifejezés detektálható.



2. ábra: Felismerő modul blokk-diagramja

Az általunk készített és futó grafikus alkalmazás LabVIEW nyelven készült, így a .NET alapú felismerő rendszert LabVIEW környezetben kellett felhasználni. Ennek érdekében a beszédfelismerő modul úgynevezett Library-ként lett fordítva, mely lehetővé teszi, hogy más nyelvi környezetekben is felhasználható legyen, így LabVIEW alatt is. A rendszer LabVIEW környezetben történő használatát szemlélteti a 3. ábra.



3. ábra: A rendszer használata LabVIEW környezetben

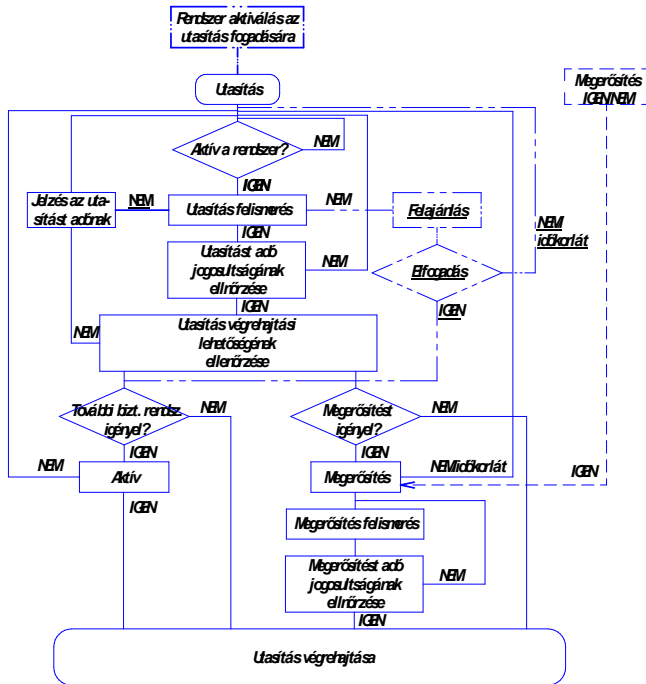
4. A “VOICE CONTROL” VEZÉRLÉS JÁRMŰBE TÖRTÉNŐ INTEGRÁLÁSI LEHETŐSÉGÉNEK ISMERTETÉSE

A gépjárműtervezés és gyártás során a járműnek több szabványnak és előírásnak is meg kell hogy feleljen a közlekedésbiztonság és járműüzembiztonság érdekében, így az ilyen jellegű rendszer integrálásának is komoly biztonsági kérdései vannak, melyek a következők:

- Csak a járművezető adhat utasítást a rendszernek?
- Képes-e a rendszer magas biztonsággal azonosítani az utasítást kiadó személy jogosultságát, hogy elfogadja és végrehajtsa az utasítást?
- Az egyéb hanghatások, mint pl. rádió, külső hanghatások befolyásolják-e a rendszer megbízhatóságát?
- Indokolt-e az egyes vezérlési utasítások megerősítése?
- A kiadott utasítás elvégezhető-e a közlekedésbiztonság veszélyeztetése nélkül?

A 4. ábra az emelt szintű hangvezérlést szemlélteti, mely több elemet tartalmaz és többszintű biztonságot. A rendszer az aktiválással indul, és ezt követően adhatunk utasítást. Az utasítás elhangzásakor és a rendszert aktivált állapota után, amennyiben a rendszer nem ismeri fel az utasítást így jelez az utasítást kiadó számára különböző módon. Az első biztonsági szint az az, hogy a rendszer vizsgálja, hogy az adott utasítás kiadásához van-e jogosultsága az utasítást kiadónak. A következő biztonsági szint pedig az, hogy az utasítás végrehajtásához igényel-e további biztonsági engedélyt, pl haladás közben csomagtartó nyitási parancs elhangzása során, nem hajtódik végre az utasítás a jármű mozgása miatt.

REFERENCIÁK



4. ábra: Emelt szintű hangvezérlés folyamatábrája

Az utasítás végrehajtásának következő biztonsági szintje a megerősítés lehet, mely hangvezérlésű lehet, vagy pedig valamilyen nyomógomb aktiválásával hagyhatjuk jóvá a parancsot. A hangvezérlésű megerősítés fontos eleme az, hogy felismeri-e a rendszer a megerősítést és vizsgálja, hogy a megerősítést jogosult személy adta-e.

5. ÖSSZEFOGLALÁS

A „voice control” rendszer alapjainak definiálása megtörtént. A járműbe integrálhatóság során a közlekedésbiztonsági és üzembiztonsági kérdések vizsgálata megtörtént, mely során az egyes feladatok végrehajtásához többszintű biztonság elvégzése szükséges. Az egyszerű gyakorlati alkalmazás és a járműbe történő integrálás lehetőségének vizsgálata során, a következő fejlesztési szint az egyes eszközök és a hangvezérlő rendszer működésének összehangolása.

KÖSZÖNETNYILVÁNÍTÁS

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az Autonóm Rendszerek Nemzeti Laboratórium keretében. (RRF-2.3.1-21-2022-00002)

Zoltán Tüske, Joel Pinto, Daniel Willett, and Ralf Schlüter, “Investigation on cross-and multilingual MLP features under matched and mismatched acoustical conditions,” in Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2013.

Samuel Thomas, Sriram Ganapathy, and Hynek Hermansky, “Multilingual MLP features for low-resource LVCSR systems,” in Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2012

Thomas Niesler, “Language-dependent state clustering for multilingual acoustic modelling,” Speech Communication, vol. 49, no. 6, 2007.

Sibo Tong, Philip N. Garner, and Herv Bourlard, “An Investigation of Deep Neural Networks for Multilingual Speech Recognition Training and Adaptation,” in Proc. Interspeech, 2017

Maryam Kamvar and Shumeet Baluja. A large scale study of wireless search behavior: Google mobile search. In CHI, pages 701–709, 2006.

S. Katz. Estimation of probabilities from sparse data for the language model component of a speech recognizer. In IEEE Transactions on Acoustics, Speech and Signal Processing, volume 35, pages 400–01, March 1987

D. Povey, D. Kanevsky, B. Kingsbury, B. Ramabhadran, G. Saon, and K. Visweswariah. Boosted MMI for model and feature-space discriminative training. In Proc. of the Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), 2008

C. Van Heerden, J. Schalkwyk, and B. Strope. Language modeling for what-with-where on GOOG-411. 2009.