

Algoritmus formánsok követésére, módosítására és szintézisére

BÓHM TAMÁS, NÉMETH GÉZA

BME Távközlési és Médiainformatikai Tanszék
{bohm, nemeth}@tmit.bme.hu

Lektorált

Kulcsszavak: formánskövetés, formáns szintézis, lineáris predikció, hangkarakter módosítás

A pontos formánskövetés problémája régóta kihívást jelent a beszédtechnológia kutatóinak. Ebben a cikkben a szerzők egy olyan eljárást mutatnak be, amely alkalmas a formánsmenetek követésére és módosítására beszédjelben. Ez egy hatékony eszköz a beszéd spektrális tartalmának elemzésére és megváltoztatására, lehetőséget ad a hangminőség és a hangszínezet módosítására. A módszer a lineáris predikció modelljén alapul.

1. Bevezetés

Zöngés hangok képzésekor a hangszalagok kváziperiodikus rezgést végeznek. Az így keletkezett gerjesztőjelet a toldalékcső (garat, orrüreg és szájüreg) rezonátor-rendszerként viselkedve módosítja: a felharmonikusok egyes csoportjait felerősíti (a rezonanciahelyekenél), másokat gyengít. Ezeket a rezonanciahelyeket és környéküket formánsoknak nevezzük.

A rezonanciahely frekvenciája a formánsfrekvencia, ami a toldalékcső átviteli függvényében lokális maximumként jelenik meg. A formánst jellemzi még sávszélessége és amplitúdója. Az előbbi alatt a lokális maximum körüli, annál legfeljebb 3 dB-lel alacsonyabb értékeket tartalmazó leghosszabb frekvenciatartomány szélességét értjük. Az utóbbi a függvény értéke a csúcsonál ([1] 52. o.).

Bár a formánsok és azok változása az idő függvényében (az úgynevezett formánstrajektóriák vagy formánsmenetek) a spektrum és a spektrogram alapján az ember számára könnyen felismerhetőek, a formánsmérés és -követés automatizálása egyáltalán nem triviális feladat.

Jelentős az igény pontos formánskövetésre és a formánsmenetek módosítására, mert ez elősegíti a nyelvészeti kutatásokat és számos alkalmazás megvalósítását is lehetővé teszi. Ilyen alkalmazás például a hangkarakterének megváltoztatása (mint dialektus módosítás, beszédkorrekció vagy a hangszínezet átalakítása) és konkatenatív szövegfelolvasó rendszerek által létrehozott hullámformák formáns-trajektóriáinak simítása. Egy ilyen algoritmus arra is alkalmas lehet, hogy a beszéd személyfüggő jellemzőit megváltoztassa, így a hallgatóban egy másik beszélő érzetét keltse a nyelvi tartalom módosítása nélkül.

A fenti alkalmazások szempontjából fontos, hogy egy olyan módszert alkalmazzunk, amely képes visszaszintetizálni a beszédet a formánsstruktúra megváltoztatása után. Ezt csak egy nagy pontosságú formánsmeghatározó algoritmussal lehet elérni. A számunkra érdekes területeken alkalmazható formánsmódosítás-

sal kapcsolatban eddig nem végeztek kiterjedt kutatásokat, az irodalomban nem találtunk megfelelő megoldást.

A formánskövetést ellenben széles körben vizsgálták az elmúlt évtizedekben. A tradicionális módszerek valamilyen nem-lineáris simítással nyert spektrumon alkalmazznak egy csúcskereső algoritmust ([2] 154. o., [3]). Ide sorolhatjuk a kepsztrális spektrumszűrést is, amikor a kepsztrumból eltávolítjuk a periodicitás miatt megjelenő maximumokat (melyek a zöngének felelnek meg) és Fourier-transzformáljuk azt ([1] 241. o.). Rabiner és Schafer ezt használja fel csúcskeresővel és az „analízis szintézissel” módszerrel kombinálva [4]. Egy másik megközelítés különböző szűrőkészletek alkalmazása [5].

Néhány, elsősorban beszéd felismerésre használt módszer is alkalmazható formánskövetésre, például elterjedtek a rejtett Markov modellen (HMM) [6] és LSP-n (Line Spectrum Pair) alapuló eljárások. Ez utóbbi az LP analízis egyik megvalósítása, ami nem az idő-, hanem a frekvenciatartományban dolgozik és az előállított együtthatók a spektrum nagy amplitúdójú szakaszainak a mozgását követik (ezek a szakaszok nagyjából megfelelnek a formánsoknak).

Ebben a cikkben egy olyan algoritmusról számolunk be, amely képes precíz formáns-meghatározásra, módosításra és visszaszintetizálásra. A kidolgozott algoritmus szemléltetésére egy grafikus alkalmazást hoztunk létre, amely a formánsmenetek sokrétű megjelenítésére képes. A második részben a módszer alapelvét írjuk le, míg a harmadikban az algoritmust ismertetjük. A negyedik fejezet a grafikus alkalmazást ismerteti, míg az ötödikben összefoglaljuk eredményeinket.

2. Működési elv

2.1. Spektrumszámítás lineáris predikcióval

Mint már korábban tárgyaltuk, a formáns a beszéd-spektrum helyi maximuma. A spektrum számítására alkalmazhatunk gyors Fourier-transzformációt (FFT), de

ez egy számos maximumot és minimumot tartalmazó függvényt állít elő. Egy ilyen spektrum csúcsait nehéz megbízhatóan meghatározni. Formáns-meghatározáshoz a lineáris predikción alapuló spektrumszámítás [2] az elterjedtebb, mert számos előnnyel rendelkezik:

- A predikció fokszámának (azaz az átviteli függvényben szereplő pólusok számának) változtatásával határozhatjuk meg a spektrum „felbontását”.
- A lineáris predikció a (formáns-detekció szempontjából érdekes) csúcsoknál közelíti legpontosabban a spektrumot.
- Még rövid szakaszok elemzésekor is elfogadható eredményeket ad.

A lineáris predikció együttthatóiból becsült átviteli függvény:

$$H(z) = \frac{1}{1 - \sum_{k=1}^p \alpha_k z^{-k}} \quad z = r \cdot e^{j2\pi \frac{f}{f_s}}$$

Az eredmény egy olyan spektrumközelítés, amely jóval simább, mint az FFT, de a formánsokat nem torzíja el.

2.2. Formánsmeghatározás

A formánsmeghatározás számos publikált módszere két nagyobb csoportba sorolható: a spektrum-alapú és a pólus-alapú eljárások. Míg az előbbiek az amplitúdó- és fázisspektrumot használják fel, az utóbbiak a z-tartománybeli pólusokkal számolnak.

Spektrumalapú módszer a McCandless-algoritmus, amely a spektrum abszolútértékének logaritmusában keres csúcsokat [7]. Christensen, Strong és Palmer hasonló eljárást dolgozott ki, de ők a log spektrum negatív második deriváltján alkalmazzák a csúcskeresést [8]. Yegnanarayana bebizonyította, hogy a komplex spektrum fázisának első deriváltja jelentős hasonlóságot mutat az amplitúdóspektrum alakjához [9]. Ha ezt deriváljuk kétszer a log spektrum helyett, akkor pontosabb formánsfrekvencia-bebecslést kapunk. Reddy és Swamy módszere egyszerre dolgozik a z- és az f-tartományban [10], így közeli formánsokat is képes elkülöníteni. Bár ezeket a módszereket már többen implementálták és vizsgálták, egyik sem jelent általánosan alkalmazható megoldást a problémára.

Formánskövető rendszerünk pólus-alapú formáns-meghatározást alkalmaz – hasonló ahhoz, amit Slifka és Anderson dolgozott ki beszélőmódosítás céljára [11]. A $H(z)$ átviteli függvény egy all-pole, tehát zérusokat nem tartalmazó modellt ad a toldalékcsőre. Ennek pólusai a rendszer rezonanciapontjainak, azaz az amplitúdó-spektrum helyi maximumainak felelnek meg. A pólusok az átviteli függvény nevezőjének gyökei:

$$H(z) = \frac{1}{1 - \sum_{k=1}^p \alpha_k z^{-k}} = \frac{1}{\prod_{i=1}^p (1 - p_i z^{-1})}$$

ahol az α_k -k a lineáris predikciós együttthatók.

A formánsfrekvenciákat és -sávzélességeket a pólusok $p_i = r_i \cdot e^{j\theta_i}$ alakjából határozhatjuk meg:

$$F_i = \frac{f_s}{2\pi} \varphi_i \quad B_i = \frac{f_s}{\pi} \ln\left(\frac{1}{r_i}\right)$$

Az eredmények rendkívül pontosak, de nem feltétlenül tartozik minden pólushoz formáns (például valós pólusok nem lehetnek formánsok). Ilyen pólusok például a sugárzási ellenállás vagy háttérzaj miatt jelenhetnek meg.

3. Algoritmus

Bemeneti adat a hullámforma periódushatár és hanghatár címkékkel. Pontosabb eredmények érdekében megadhatjuk a bemondás fonetikus átíratát is. A jelfeldolgozást két szakaszra bonthatjuk: analízis és szintézis. Az előbbi a formánsmenetek meghatározását jelenti és kimenetei a formánsokat leíró adatok valamint néhány segédinformáció (a lineáris predikció maradékjele és az energiaszintek). Szintézis alatt a formánsok módosítását és az eredmény visszaszintetizálását értjük.

3.1. Analízis

3.1.1. LP analízis és pólus-meghatározás

A lineáris predikciós analízist a zöngeperiódusokkal szinkron módon végezzük. Az analízishez Hamming-ablakot használunk, amit mindig két egymást követő periódus hosszára állítunk és periódusonként léptetjük, ezzel biztosítva a megfelelő időfelbontást. Zöngétlen hangok esetén állandó időközönként virtuális alapperiódusokat jelölünk ki.

Előbb Burg módszerrel meghatározzuk a PARCOR együttthatókat, majd azokat lineáris predikciós együttthatókká konvertáljuk, hogy kiszámíthassuk az átviteli függvényt [2].

A lineáris predikciós analízis és szintézis nem garantálja, hogy a bemenő jel energiája megegyezik a kimenő jelével. Annak érdekében, hogy az ebből eredő torzítást elkerüljük, érdemes tárolni minden egyes időszelvény energiáját. A visszaszintetizálás során ezt felhasználhatjuk az eredeti szint visszaállítására. Ugyanezt az adatot csenddetekcióra is használjuk, miután normalizáltuk az időszelvény hosszára. A veszteségmentes kódolás érdekében a lineáris predikció maradékjelét is tárolni kell.

Ezután ki kell számítanunk a rendszer pólusait, azaz az átviteli függvény nevezőjének gyökeit. Bár a nevezőben levő polinom valós együttthatós, a gyökök komplexek is lehetnek, ezért a Newton-Raphson és Brent módszerek nem alkalmasak a feladatra. Helyettük a Laguerre eljárást [12] használja az algoritmus.

Laguerre algoritmusát iteratíván alkalmazva kaphatjuk meg az összes gyököt: egyszer lefuttatva az eredmény egy valós gyök vagy egy komplex konjugált gyökképár. Miután ezekkel leosztottuk a polinomot, újra lefuttatjuk a gyökkeresést.

A komplex konjugált póluspárokhoz a 2. részben leírt formulákkal, valamint a spektrumból meghatározha-

tóak a formánsok adatai (rendre frekvencia, sávszélesség és amplitúdó). Ezeket az adatokat azonban a szintézis során nem használhatjuk fel: pontosabb a számítás, ha a formánsfrekvenciák módosítása során közvetlenül a megfelelő pólus argumentumát változtatjuk.

Az újraszintetizáláshoz tárolni kell azokat a pólusokat is, amelyekhez nem tartozik formáns.

3.1.2. Formánsokhoz tartozó pólusok kiválasztása

Nem feltétlenül tartozik minden pólushoz formáns. Annak érdekében, hogy az ilyen pólusokat ki lehessen szűrni, néhány feltételt érdemes megvizsgálni:

- A formánsfrekvencia (amit a pólus argumentumából számíthatunk) nagyobb kell, hogy legyen, mint az alapfrekvencia.
- A pólus abszolútértéke el kell, hogy érjen egy minimumértéket, azaz limitáljuk a formáns sávszélességét.
- Az időszelvény energiája meg kell, hogy haladjon egy küszöbértéket (beszéddetekció).
- Azokat a pólusokat, amelyek valós része nulla vagy nullához közeli, figyelmen kívül hagyhatjuk (ezek például alacsony frekvenciájú, keskenysávú zaj miatt jelenhettek meg).

Kizárólag azokat a pólusokat tekintjük formánsnak, amelyek eleget tesznek mind a négy feltételnek. A kiválasztott komplex konjugált póluspárokból meghatározhatóak a formánsok adatai (frekvencia, sávszélesség és amplitúdó).

A fenti feltételekkel mindössze a nyilvánvalóan téves eredményeket tudjuk kiszűrni és adhatunk egy első közelítést a formáns-pólus összerendelésre. A végleges leképezést a formánstrajektóriák folytonosságát korlátok szerint történő meghatározásával végezzük.

3.1.3. Formánstrajektóriák

A feladat tehát a formánsok trajektóriáinak követése a hanghatárok ismeretében. Rendelkezésre áll még a formáns-pólus összerendelés első közelítése. Minden egyes időszelvény formánsait le kell képeznünk a következő időszelvény formánsaira úgy, hogy azok folytonos formánsmeneteket képezzenek.

Egy formánst ahhoz a formánsmenethez rendeljük hozzá, amelyik trajektória utolsó frekvenciaértéke hozzá legközelebb van. Azokat a formánsmeneteket, amelyek nem fedik át egymást, összevonjuk. Szintén összevonjuk a párhuzamosan egymáshoz rendkívül közel futó formánstrajektóriákat. A valószínűtlenül rövid formánsmeneteket töröljük.

Zöngétlen obstruensek (zárhangok, réshangok és zár-rés hangok) előtt és után nem szükséges folytonossági korlátokat alkalmazni, mert ezen hangok képzése olyan artikulációs mozgásokkal jár, amelyek hirtelen változást idéznek elő a toldalékcső rezonanciahelyeiben. Ha a felvétel fonetikus átírata rendelkezésre áll, az algoritmus nem próbálja a formánsmeneteket ilyen hangokon keresztül összekötni. Ha nem, akkor minden hanghatáron töréspontokat iktat be – így néhány hangátmenetnél pontatlan lesz az eredmény, de

a leképezés pontossága az egész felvételre tekintve jobb lesz.

3.2. Szintézis

Az első lépésben az analízis kimenetéből előállítjuk a visszaszintetizáló algoritmus bemenetét, azaz módosíthatjuk a formánsmeneteket. A transzformáció módját az adott alkalmazás határozza meg (például trajektóriák simításakor valamilyen interpolációs eljárás). A trajektóriák módosításával közvetve a pólusokat mozgathatjuk a z-síkon.

A második lépés a visszaszintetizálás. A módosított formánsfrekvencia értékek megadják a módosított pólusokat (az analízis során alkalmazott összefüggésekkel). Ezekből összeállítjuk az átviteli függvény nevezőjét alkotó polinomot, majd ezt kifejtjük. A polinom együtthatóiként megkapjuk a lineáris predikció együtthatóit. Ezek és a külön eltárolt maradékkal felhasználásával elvégezzük a lineáris predikciós szintézist. Utolsó lépésként visszaállítjuk az egyes időszelvény energiáját (amit az analízis során eltároltunk). Az eredmény egy új hullámforma fájl, amelyben a formánsmenetek módosultak az eredetihez képest.

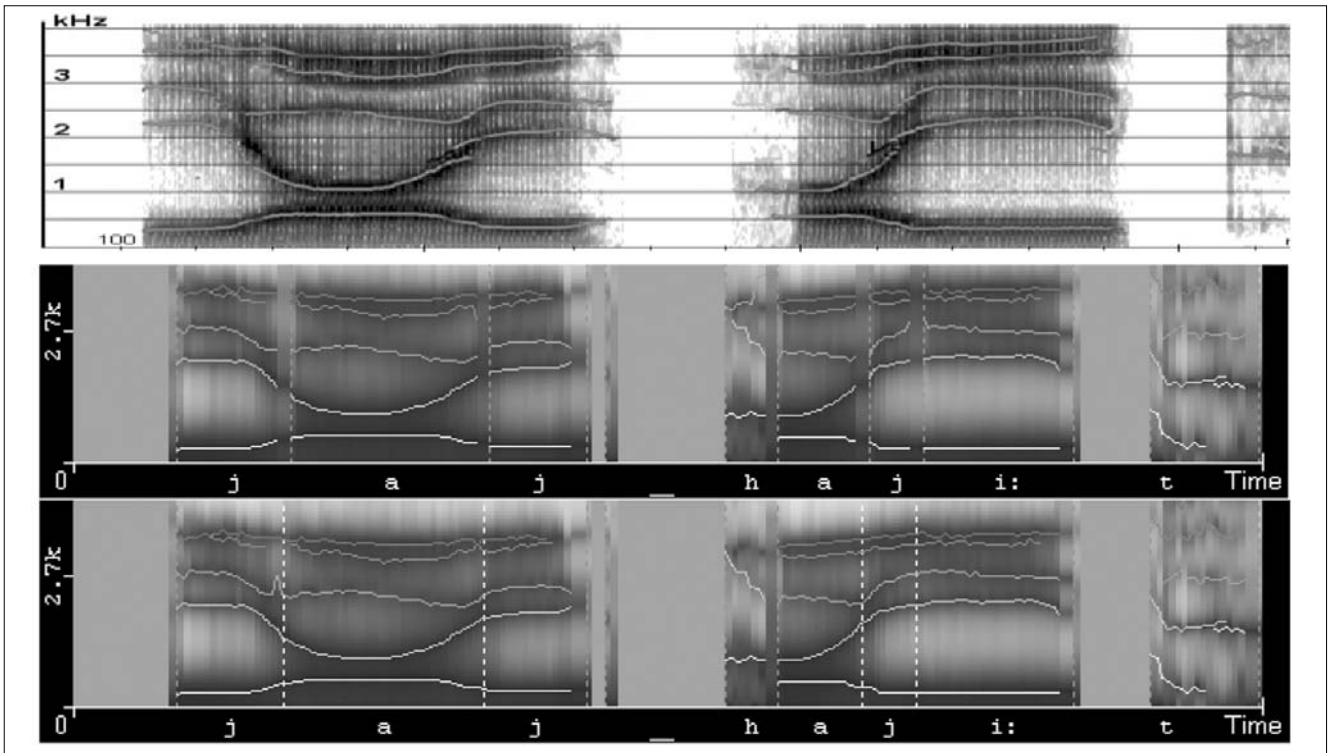
4. Eredmények

Az algoritmust implementáltuk és külön ellenőriztük az analízis és a szintézis funkciókat magyar nyelvű felvételekkel.

4.1. Analízis

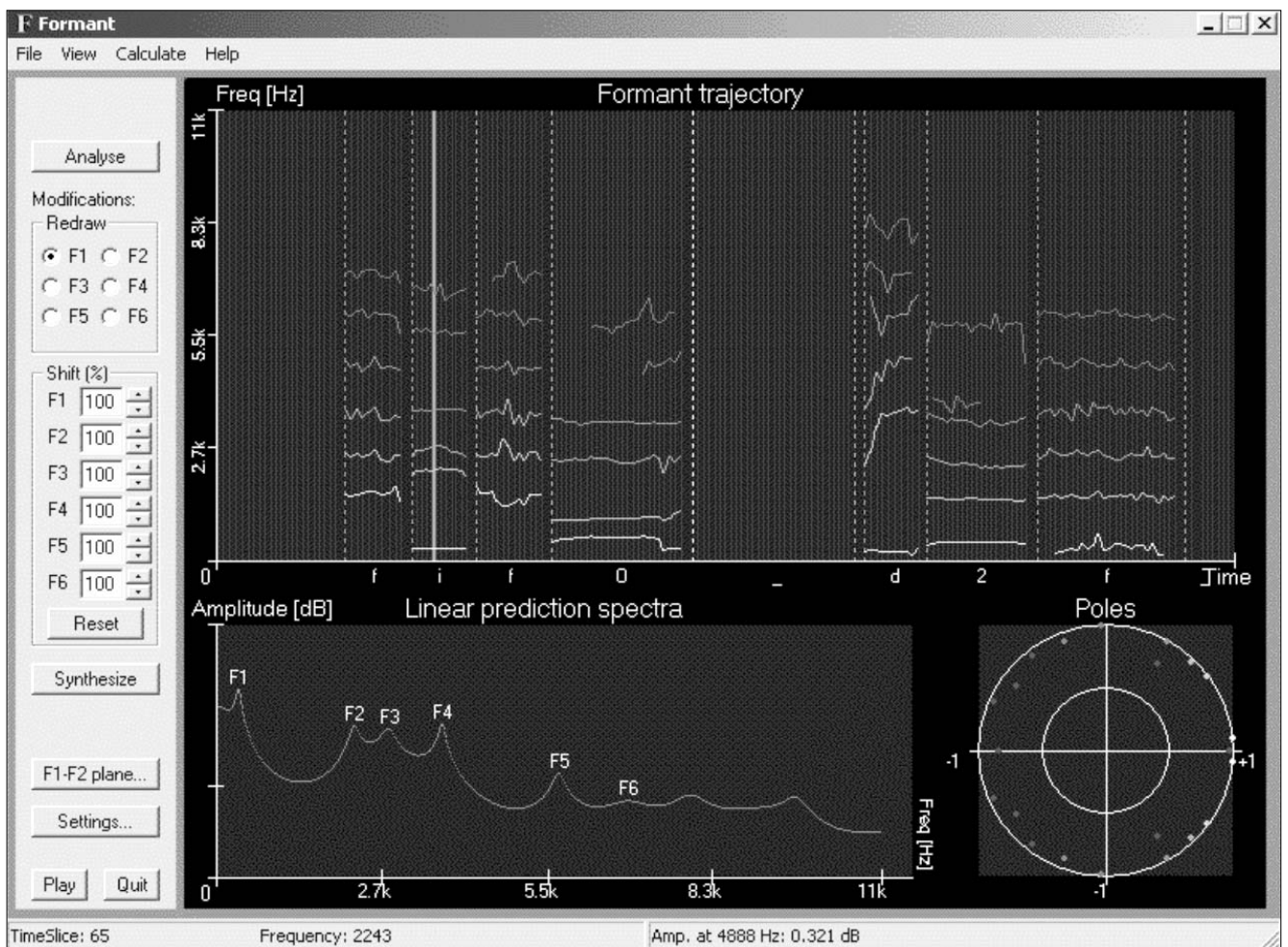
A formáns-meghatározás és -követés pontosságát három módon ellenőriztük. Először az MTA Nyelvtudományi Intézet által a rendelkezésünkre bocsátott spektrogramokkal vetettük össze az eredményeket. Ezek a referenciaként használt spektrogramok a Kay Elementrics CSL 4300B típusú digitális jelfeldolgozóval készültek. Második lépésként az egyes beszédhangok formánsfrekvenciáit összehasonlítottuk a magyar nyelvre jellemző értékekkel ([13] 40. o.).

Végül lemértük a leképezési hibaarányt a [14] definíciója szerint. Ehhez a később ismertetendő grafikus formánskövető program segítségével összehasonlítottuk a felvételek spektrogramjait a kinyert formánsmenetekkel. A tesztalmaz 29 darab, kétszavas, magyar nyelvű felvétel volt férfi bemondóval. Mindössze két felvétel esetén (6,90%) találtunk leképezési hibát az első három formánstrajektóriában – az egyik esetben a harmadik formáns, a másik esetben mind a három formáns leképezése téves volt. [14] szerint egy névleges formánsfrekvencia értékeket használó algoritmus leképezési hibaaránya 3,62-3,99%. Ez alacsonyabb az általunk elért aránynál, de az itt bemutatott algoritmus nem használ előre definiált, tipikus formánsértékeket, így a bemondó személyétől, nemétől és nyelvétől függetlenül működik. Az említett cikk egy hasonló tulajdonságokkal rendelkező módszer hibaarányát 13,04%-ban állapítja meg.



1. ábra Egy rövid felvétel elemzése a Kay Elemetrics CSL 4300B-vel (fent), a cikkben ismertetett algoritmussal – fonetikus átírat nélkül (középen) és fonetikus átíratl (lent)

2. ábra Grafikus formánskövető és módosító program



A „jaj hajít” bemondásra adott eredmény az 1. ábrán látható. Az ábra alapján látszik, hogy zöngés hangokra pontos volt a formánskövetés, még a második szó /h/-magánhangzó átmenete után is. Ezt az esetet az irodalomban a névleges formánsfrekvenciákat nem használó algoritmusok számára problematikusként említik [14].

4.2. Szintézis

A szintézis lehetőségeit a lehetséges alkalmazások szempontjából vizsgáltuk, mert ezek alapján terveztük a módosító eljárást. A szintézis értékelése és a módszer finomítása még folyamatban van, jelenleg két egyszerű kísérletről tudunk beszámolni.

A formánsszerkezet megváltoztatásával egy magánhangzót egy másik magánhangzóvá alakíthatunk át. Például a „fésű” szóból „fásü”-t csinálhatunk, ha az első formánsmenetet magasabbra, a második formánsmenetet alacsonyabbra toljuk a frekvencia-tengelyen. Egy ilyen formánsmódosított felvételt négy magyar anyanyelvű kísérleti alannyal meghallgattunk és mindannyian egyértelműen a „fásü” értelmetlen szót hallották. Ez a technika hatékonyan bizonyult konkatenatív beszéd szintetizátor akusztikai adatbázisának kiegészítésére olyan hangokkal, amelyeket nem rögzítettünk

a bemondótól (például olyan beszédhangok, amelyek a magyarban nem fordulnak elő).

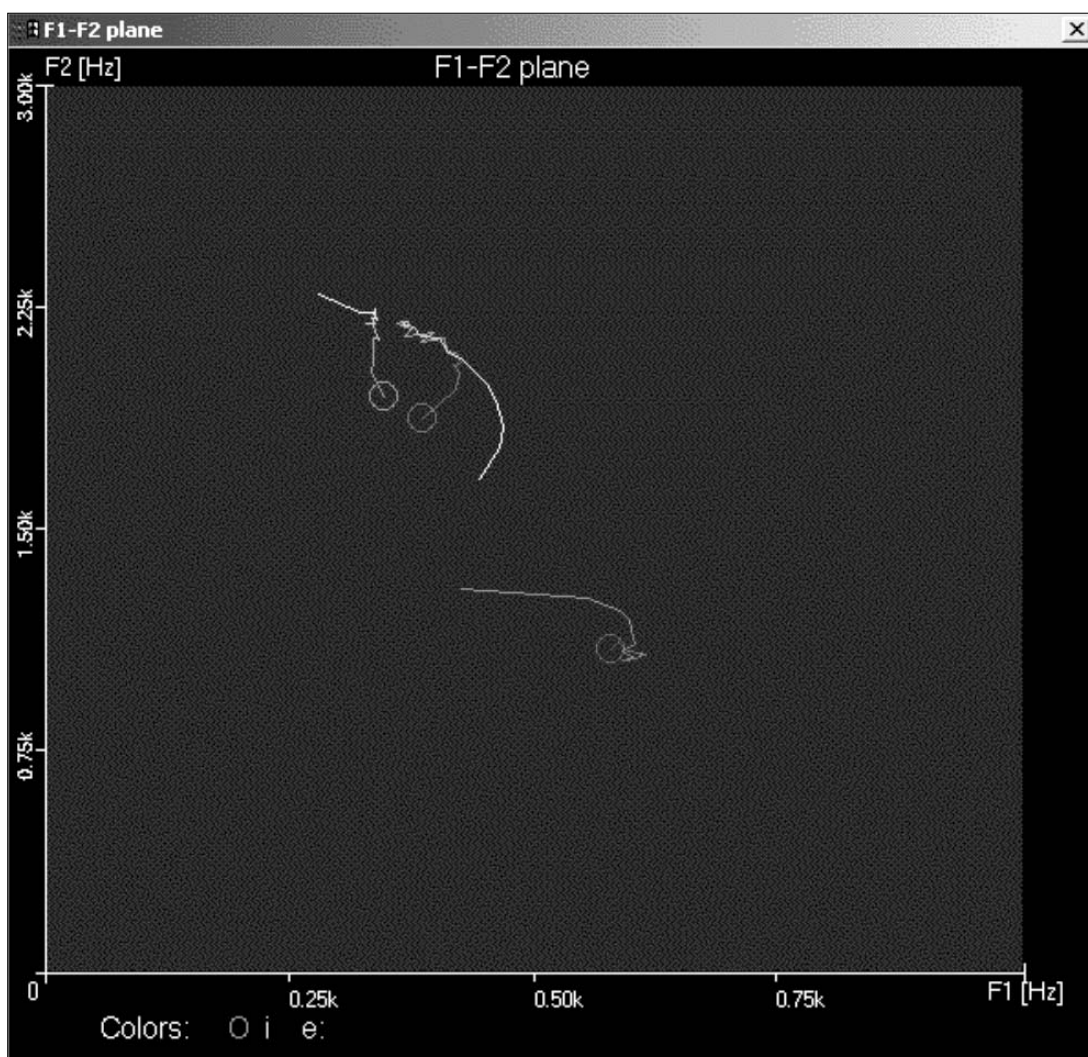
A beszédkutatók körében általánosan elfogadott nézet, hogy a magasabb formánsok a beszélőre jellemzőek. Kezdeti kísérleteket végeztünk beszélőmódosításra, amivel célunk a beszélőspecifikus jegyek megváltoztatása volt annak érdekében, hogy a beszélő személye ne legyen felismerhető. Meghallgatásos tesztek azt mutatták, hogy néhány formánsmenet megfelelő módosítása után a beszélő személyek azonosítása jóval nehezebb a hallgatók számára.

A formánsmenetek módosítása során azt tapasztaltuk, hogy nagy mértékű változtatás esetén jelentősen torzulhat a beszéd. Kisebb (legfeljebb kb. 20%) módosítások során általában jó minőségű, természetes hangzású az újrasszintetizált beszéd.

4.3. Grafikus formánskövető és módosító program

Egy grafikus alkalmazást készítettünk a módszer lehetőségeinek bemutatására és a modell ellenőrzésére (2. ábra). Ez a program oktatási és kutatási célra nyilvánosan hozzáférhető [15].

A program az analízis után az egyes formánstrajektóriákat (legfeljebb hatot) külön színekkel rajzolja ki a frekvencia-idő síkra. Az ellenőrzés érdekében a prog-



3. ábra
Magánhangzók
formánsmenetei
az F_1 - F_2 síkon

ram képes spektrogram megjelenítésre, valamint a kétféle ábrázolás egyszerre is látható. Egy időszelvet kijelölve megtekinthetjük a szelethez tartozó rövid idejű lineáris predikciós spektrumot (a program a formánsokat ezen is feltünteti) és a lineáris predikciós szűrő pólusait a komplex z-síkon (az ábra jobb alsó sarkában). Ez utóbbi egy újfajta vizualizációs megoldás, amely ugyanazt az információt hordozza, mint a lineáris predikciós spektrum.

Van lehetőség a formánsmenetek ábrázolására az F1-F2 síkon is (3. ábra). A hagyományos szórásképek helyett az első két formáns együttes mozgását folytonos görbével lehet megjeleníteni.

Az ábrán a felvétel magánhangzóinak F1-F2 menete látható. Az egyes pontok vízszintes koordinátája az első, a függőleges koordinátája a második formáns értékét fejezi ki. A görbék egyes szakaszainak fényereje az idő múlásának felel meg: a magánhangzó első alapperiódusából származó adat jelenik meg a legsötétbben, majd periódusonként világosodik a görbe és az időben utolsó szakasz a legvilágosabb. A jobb láthatóság érdekében egy kör is megjelenik a kezdőpontok körül.

A formánsmeneteket módosíthatjuk kézi átrajzolással vagy trajektóriánként egy szorzó beállításával. Az eredményt újraszintetizálhatjuk és elmenthetjük egy hullámforma fájlba.

A program széles felhasználási lehetőségekkel rendelkezik a fonetika oktatása területén. Szemléltető eszközként használható például a beszédhangok formánsstruktúráinak bemutatására vagy a magánhangzók megkülönböztető tulajdonságainak képi megjelenítésére. Alkalmas továbbá egy fonéma különböző realizációi közötti hasonlóságok és különbségek vizsgálatára és koartikulációs hatások bemutatására. Fonetikai vizsgálatok eszközeként is szolgálhat, többek között perceptív tesztek hanganyagának előállításához, nyelvjárássok vizsgálatához „analízis szintézissel” módszerrel és szemléltető ábrák készítéséhez. Jelen cikk ábráit is ezzel a programmal állítottuk elő.

5. Összefoglalás

Egy általános formánskövető és módosító algoritmus számos alkalmazás megvalósításához szükséges. Cikkünkben egy ilyen algoritmust mutattunk be és értékeltünk. Eljárásunk elfogadható eredményeket ad kevés információ alapján is és pontos eredmények érhetőek el további bemeneti adatokkal (fonetikus átirat). Az algoritmust implementáltuk és beépítettük egy grafikus alkalmazásba.

Ez a nyilvánosan elérhető program hasznos oktatási és kutatási eszköz lehet – ezt igazolja, hogy a szerzők oktatási tevékenységén kívül az ELTE Bölcsészettudományi Karán is használják. Továbbá laboratóriumunkban sikeresen alkalmaztuk a magyar magánhangzók formánsstruktúrájának kiterjedt vizsgálatára (melynek eredményei feldolgozás alatt állnak).

Irodalom

- [1] Gordos G., Takács Gy.:
Digitális beszédfeldolgozás,
Műszaki Könyvkiadó, Budapest, 1983.
- [2] Markel, J. D., Gray, A. H.:
Linear Prediction of Speech,
Springer-Verlag, Berlin, 1976.
- [3] Lobanov, B., Levkovskaya, T., Kheidorov, I.:
„Speaker and channel – Normalized set of formant parameters for telephone speech recognition”,
Proc. of Eurospeech 1999, Vol. 1., pp.331–334.
- [4] Rabiner, L. R., Schafer, R. W.:
Digital Processing of Speech Signals, Prentice-Hall,
Englewood Cliffs, 1978.
- [5] Ouni, K., Lachiri, Z., Ellouze, N.:
„Formant estimation using Gammachirp filterbank”,
Proc. of Eurospeech 2001, Vol. 4., pp.2471–2474.
- [6] Weber, K., Bengio, S., Bourlard, H.:
„HMM2 – Extraction of formant structures and their use for robust ASR”,
Proc. of Eurospeech 2001, Vol. 1, pp.607–610.
- [7] McCandless, S. S.:
„An algorithm for automatic formant extraction using linear prediction spectra”,
IEEE Trans. on Acoustics,
Speech and Signal Processing, Vol. 22, no.2, 1974.
- [8] Christensen, R. L., Strong, W. J., Palmer, E. P.:
„A comparison of three methods of extracting resonance information from predictor coefficient coded speech”,
IEEE Trans. on Acoustics,
Speech and Signal Processing, Vol. 24, no.1., 1976.
- [9] Yegnanarayana, B.:
Formant extraction from linear prediction phase spectra,
Journal of the Acoustical Society of America,
Vol. 63., p.1638., 1978.
- [10] Reddy, N. S., Swamy, M. N. S.:
„High-resolution formant extraction from linear prediction phase spectra”,
IEEE Trans. on Acoustics,
Speech and Signal Processing, Vol. 32, no.6., 1984.
- [11] Slifka, J., Anderson, T. R.:
Speaker modification with LPC pole analysis,
Proc. of ICASSP 1995, pp.644–647.
- [12] Orchard, M. T.:
„The Laguerre method for finding the zeros of polynomials”,
IEEE Transactions on Circuits and Systems, 1989.
Vol. 36, no.11, pp.1377–1381.
- [13] Olasz G.:
Elektronikus beszédelőállítás
(A magyar beszéd akusztikája és formánszintézise),
Műszaki Könyvkiadó, Budapest, 1989.
- [14] Lee, M., van Santen, J., Möbius, B., Olive, J.:
„Formant tracking using segmental phonemic information”,
Proc. of Eurospeech 1999, Vol. 6., pp.2789–2792.
- [15] <http://fonetika.nytud.hu>