

Opponensi vélemény

Dr. Bank Balázs

“Logaritmikus frekvenciafelbontású
szűrőtervezés audio alkalmazásokhoz”

című MTA doktori értekezéséről

Témaválasztás, bevezető megjegyzések

A dolgozat az audio jelfeldolgozás speciális igényeit kielégítő szűrőtervezési, modellezési, és jelút-kompenzációs módszerek kidolgozására fókuszál. A szűrőtervezési feladat megoldása egy tetszőleges specifikációt (frekvenciame-
net, impulzusválasz, model matching) legjobban közelítő szűrőt keresi adott szűrőstruktúra és számítási kapacitás (fokszám) mellett. A szerző kutatási témája a fix pólusú párhuzamos szűrő alkalmazása, vizsgálatai pedig a módszer tulajdonságainak feltérképezésére, approximációs képességeinek javítása, valamint különböző irányú kiterjesztéseire irányultak.

A téma érdekes, a problémakör fontosságát rengeteg alkalmazási terület is alátámasztja. A disszertáció témaválasztása annak ellenére korszerűnek mondható, hogy a felhasznált matematikai módszerek és kérdéfeltevések meglehetősen hosszú múltra tekintenek vissza. Számomra megdöbbentő volt azonban, hogy a dolgozatot tanulmányozva ez a meglehetősen gazdag múlt (és közelmúlt) szinte egyáltalán nem tükröződik. Mintha a szerző egy meglehetősen szűk optikán át nézné az adott területet, vagy egyáltalán nem ismerve, vagy egyszerűen ignorálva, a vonatkozó szakirodalom és az ott fellelhető eredmények egy szignifikáns részét. Ez azért is problematikus, mert a hiányolt alaperedmények – pld. approximáció, konvergencia, kvadratikus interpoláció, numerikus kérdések – relevanciája a dolgozatban megjelenő problémák szempontjából nyilvánvaló.

Lásd például: Peter S.C. Heuberger, Paul M.J. Van den Hof, Bo Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*, Springer, 2005 – a maga háromszáz körüli tételt tartalmazó hivatkozási listájával. Nem is beszélve az azóta megjelent munkákról. Nem hiszem, hogy egy Google kereséssel ne lehetett volna ezekre rábukkanni.

Általánosan megállapítható, hogy a dolgozatban felvonultatott matematikai eszközök szintje és azok használata sok kívánni valót hagy maga után és igen kedvezőtlen benyomást hagy az olvasóban. Az egyes problémakitűzések, célok, nincsenek megfelelőképpen megfogalmazva: ez nem csak a formalizált leírást érinti. Nem világos, hogy milyen adatokból dolgozunk, milyen feltételek mentén. A szükséges precizitás hiánya egyúttal értelmetlenné teszi az állítások egy jó részét: nincs kijelölve, az olvasó jóindulatára/képzelőerejére van bízva, a kontextus ami az eredményeket eredményekké teszi. Például: a dolgozatban leírtak alapján arra következtethetünk, hogy zajmentes adataink vannak és végtelen pontossággal számolunk. Ezt az általános képet legfeljebb néhány elszórt megjegyzés közelíti a valósághoz, feltéve,

ha az olvasó azokat észreveszi. Véleményem szerint az az informális stílus, ami a dolgozat sajátja, még egy gyengébb PhD dolgozatban sem engedhető meg és nagyon sokat ront a dolgozat értékén.

Formai észrevételek

Az igényes formátumú angol nyelvű dolgozat 132 számozott oldal terjedelmű, beleértve az irodalomjegyzéket és egy rövid mellékletet is. A be(fel)vezetést és a nemzetközi irodalomban található előzményeket az 1.–5. fejezetek tartalmazzák. Szerintem ez nem volt egy szerencsés választás. A három, négyoldalas fejezetek olyan érzést keltenek, mintha egy monográfiának a vázlatát olvasnánk, aminek az első öt fejezetének csak a bevezetése van megírva. A pályázó saját eredményei a 6-9. fejezetekben szerepelnek. A 10. fejezet tartalmazza az összefoglalást, amelyet, saját publikációk listája, egy terjedelmes irodalomjegyzék és egy függelék követ. Az értekezés fő fejezeteit talán lehetett volna kissé arányosabban beosztani (ld. a későbbi megjegyzéseket). Az értekezésben szereplő ábrák segítik a leírtak megértését, szerkesztésük általánosságban tetszetős. Szerintem az ábrákon színek használata javította volna az olvashatóságot. Összességében megállapítható, hogy az értekezés teljesíti az MTA doktora értekezésekkel szemben támasztott formai követelményeket.

Az egyes fejezetekhez tartozó tartalmi megjegyzések és kérdések

1. fejezet ad egy motivációs háttérrel és indokolja a logaritmusos frekvenciafelbontás alkalmazását a szűrők tervezésekor. Szerencsés lett volna bevezetni és pontosan definiálni a fogalmakat, amiket használ: pld. logaritmusos frekvenciafelbontás, Kautz szűrő, stb. E nélkül eléggé nehezen értelmezhető mindaz, amit leír. Az, hogy milyen skálán (egyenletes, logaritmusos) definiálunk egy feladatot nyilván befolyásolja a megoldást. Mint ahogyan az is, hogy ehhez milyen adatok állnak rendelkezésre. Erről is lehetett volna beszélni.

Az sem teljesen világos, hogy milyen algoritmusosztályban gondolkodik a szerző: online, offline. erre csak részben ad választ a fejezet utolsó bekezdése. Mik a peremfeltételek, megszorítások aminek a tervezési eljárás meg kellene, hogy feleljen?

Csak jelzem, hogy ezen a ponton – (pp. 5), a "This work, ..., will focus on fixed-pole parallel filters, a methodology allowing the design of IIR filters

at arbitrary frequency resolution profiles, while still leading to a simple filter structure with low computational complexity." – mondat még azt is nyitva hagyja, hogy a tervezési eljárásra vagy a tervezett szűrő futására vonatkozó komplexitásról van-e szó.

Kérdés: a szerző milyen mérési adatok rendelkezésre állását tételezett fel a dolgozat során. A nem egyenletes skála adatai hogyan állnak majd elő?

2. fejezet adná meg a legfontosabb fogalmakat és jelöléseket. A szerző véges dimenziós diszkrét normákat ír fel idő illetve frekvencia tartományban. Nagyon nem szerencsés, és félrevezető, a L_p jelölés használata. Illet volna a frekvencia tartományi esetet is explicit módon felírni.

Az előző fejezetből kiindulva és a frekvenciatartományi esetre tett nem egyenletességi megjegyzés alapján nyitva marad a kérdés: mi van az időtartományban? Ott egyenletes a mintavételezés?

A 2.2 pontban tett megjegyzés csak SISO rendszerekre áll. Amúgy mi van a kauzalitással? Mi van a stabilitással?

Kérdés: Jól látom, hogy az alapfeltevés ebben a pontban az, hogy SISO kauzális és stabil LTI rendszereket vizsgál (sőt, igazából véges dimenziós SISO kauzális és stabil LTI rendszereket)?

2.2.1 pontban: "...we might also call the method equalizer design based on system identification." Amit itt felírt, az egy steril approximációs feladat (one sided model matching) lenne, ha tisztességesen meg lenne fogalmazva. Az identifikációs probléma lényege, hogy zajos adatokból dolgozunk – arról itt nem esett szó. Nem szerencsés meglevő alapfeladatokat új névvel illetni... Ami a feladat megfogalmazásából hiányzik: adott modell osztályból származó transzfer függvényeket (szűrőket) keres, ami egy rögzített frekvencia rácson produkálja a felírt funkcionál minimumát.

2.3 szakasz: Ilyen alap szintű általánosságot célszerű lett volna a mellékletbe tenni.

2.4 szakasz: Ezt a szakaszt jobb lett volna közvetlenül a 2.2 után tenni.

3. fejezet Ha jól értem, akkor az itt felsorolt simító ablakos technikák a specifikációt, vagy a használt diszkrét norma aktuális alakját érintik.

Kérdés: Hogyan biztosítható, hogy a cél (target) egy kauzális, stabil LTI rendszer legyen?

4. fejezet a frekvencia torzítás (warping) technikáját tárgyalja. Ez a fejezet másképp íródott volna meg, ha szerzőnek tágabb kitekintése van a létező eredményekre (lásd a referencia az elején). Ez a Laguerre eset, ami speciális volta miatt kissé félrevezető is lehet. Ennek oka, hogy a konstanssal kiegészített Laguerre FIR $\{D^k(z), k = 0, \dots, N + 1\}$ és a $\{\frac{z}{z-\lambda}D^k(z), k = 0, \dots, N\}$ ugyan azt az alteret fedi le. Bár Kautzra ez már nem igaz, de pólusfüggő frekvencia torzítás hasznos és definiálható – egy nemegyenletes frekvencia rács adódik, amin visszkapjuk pld. a diszkrét ortogonalitást és egy sor kedvező tulajdonságot (FFT).

Az algoritmus alapján nem egészen értem, miért kellene itt akármit is vissza transzformálni: az, hogy a kritérium és a hozzá tartozó adatok, nem egyenletesen vannak az egységkörön, nyilván befolyásolja az algoritmust, az együtthatókat ami az LS eredménye, de a transzfer függvényt nem (feltétlenül). Sok múlik a részleteken, pld. milyen adatokat használok és hogyan, ami az adott szövegrészből nem derül ki.

5. fejezet bemutatja a Takenaka-Malmquist (Kautz) rendszert. Az 5.1 szakasz utolsó paragrafusa teljesen elhibázott és talán jobb a véletlen elírások számlájára írni... Megjegyezném, hogy a TM alak (5.2) nem az egyetlen, adott pólusokkal felírt, ortonormált rendszer. Azt is szükségesnek látom megjegyezni, hogy ezek általában (pld. egyenletes rács) nem diszkrét ortonormált rendszerek.

5.2 alfejezet beszél a lehetséges pólusválasztásról. Mivel a feladat nincs precízen megfogalmazva és specifikálva, az elvárások nincsenek felsorolva, az itt leírtak nem igazán értelmezhetőek. Például: ha az impulzusválasz adott, és zaj sincs, akkor egy Ho-Kalman algoritmus meg fogja nekem adni az adott rendszer minimális állapotér-reprezentációját, és lényegében készen vagyok. Kiszámolhatom pld. az összes pólust, felírhatom mondjuk a az irányíthatósági kanonikus alakot (legalább is azzal a nagyvonalúsággal amivel a dolgozat az érdemi problémákat kezeli). Ez utóbbi alak az implementáció során műveletszám szempontjából is optimális. De ha ragaszkodom a szerző kedvenc párhuzamos alakjához, akkor egy valós modális alakra hozom a reprezentációt, megkapva a parciális tört felbontásnak megfelelő alakot (egyszeres pólusokat feltételezve az egyszerűség kedvéért). Tehát: a feladatot leíró lényegi információ nélkül nem látható miért is lesz jó (nem lesz rossz) a leírt pólusválasztás. Ennek sok oka lehet, de sajnos ezek a szövegből nem derülnek ki.

6. fejezet írja le a javasolt fix pólusú párhuzamos szűrő tervezését és tulajdonságait. A szerző saját eredményként könyveli el a parciális tört alakban történő felírást direkt módon, vagyis az együtthatóknak a megfelelő LS feladat megoldásából származó meghatározásával. A pólusrendszer rögzítésére vonatkozó megkötések nélkül, abban az általánosságban ahogy a dolgozat ezen részében szerepel, a javasolt algoritmus általában használhatatlan. Ennek magyarázatát a szerző is felírja, csak nem akarja észrevenni. A (6.25) által adott formula

$$K_{i,k} = \sqrt{1 - p_k \bar{p}_k} \prod_{j=1, j \neq i}^k \frac{1}{p_i - p_j} \prod_{j=1}^{k-1} (1 - \bar{p}_j p_i) \text{ for } i \leq k$$

középső tagja okozza a problémát, ha a pólusok közel vannak. Ez amúgy klasszikus ismeret és a Vandermonde mátrixok inverzének rossz numerikus kondicionáltságával van összefüggésben. Ez a probléma akkor is fennáll, ha a szerző által javasolt LS feladatot akarjuk megoldani a megoldásban szereplő Gram mátrix rendkívül rossz numerikus kondicionáltsága miatt. Van a dolgozatban még egy hely, ahol ezt tetten lehet érni: a 9.2 táblázat a 92. oldalon.

Ezért azután a 6.4 alfejezetben levezetett, amúgy triviális, kapcsolat sem aknázható ki, általában, a kívánt reprezentáció előállítására. Még akkor sem, ha nagyvonalúan eltekintek attól, hogy a rendszer nem diszkrétan ortogonális.

Ami ebből a fejezetből hiányzik: annak igazolása (pld. hibabecslések), hogy a javasolt póluselrendezéssel és a javasolt algoritmussal (milyennel is?) valóban előállítható a feladat megoldása a feladatosztályra.

Ahhoz képest, hogy milyen fontosnak tartja a szerző, hogy a parciális tört felbontás az implementáció szempontjából előnyös, egy rövid fél oldallal elintézi a témát a 6.4.2 alfejezetben. Véleményem szerint a kérdéskör ennél jóval összetettebb és több figyelmet és magyarázatot igényel. Lásd még az idevágó kanonikus alakra alapozott példám. Arról nem is beszélve, hogy az adatmozgatás és memóriához való hozzáférés ugyan olyan tényező lehet egy adott implementáció során, mint a műveletszám.

A 6.5. alfejezet meglátásai érdekesek, hasznosak és újak. A tézis megfogalmazásakor én inkább erre a részre koncentráltam volna.

Kérdés: Mekkora modellméret és mintaszám esetén gondolja még alkalmazhatónak a javasolt eljárást?

7. fejezet új póluselrendezési eljárásokat javasol. Jó lett volna, ha az olyan utalások, mint: "the parallel filter is more advantageous from the implementation point of view in terms of computational complexity and quantization noise" alá lennének támasztva a dolgozatban. Ez ugyan is valóban tézis értékű állítás lehetne, ellentétben sok mással, ami végül is bekerült a dolgozatba.

Itt is vannak olyan részek, megfogalmazások, amit nem egészen értek: pld. 7.2 alfejezetben "It has been observed for Kautz filters that a logarithmic pole set produces logarithmic frequency resolution ..." állítás alatt mit kell érteni? Ha jól értelmezem a fogalmat a "logarithmic frequency resolution" az egy ábrázolási döntés következménye, és azt a rácsot határozza meg, amin kiértékelem/definiálom, pld., a normákat. Gondolom nem ebben az értelemben használta a szerző az adott szövegekörnyezetben. Ide kellett volna egy magyarázó utalás, hogy hogyan is kell érteni az adott mondatot.

Itt is visszaköszön az, hogy nincs pontosan megfogalmazva, hogy mit is akarok. Ezért nem világos, pld., hogy miért is baj az, ha kézzel kell valamit beállítani (pld. a pólusok sugarát).

A fejezet erénye, hogy összehasonlítja számos pólusválasztási stratégiát és ad egy intuitív képet az egyes eljárásokról. Elkerülhetetlen azonban, hogy számos, potenciálisan releváns stratégia kimaradjon a vizsgálatokból.

Kérdés: Miért is jobb a javasolt pólusválasztási stratégia a kitűzött feladatosztályra, mint más egyéb, például lásd az idézett szekvenciális stratégiát, pólusválasztási eljárás?

Soumelidis, Alexandros and Bokor, József and Schipp, Ferenc (2013) An iterative identification of pole-structure in dynamic systems based on hyperbolic metrics and Malmquist–Takenaka representation. In: Proceedings of the 52nd IEEE Conference on Decision and Control, 2013-12-10 - 2013-12-13, Florence, Italy.

Érdemes lenne összehasonlítani a javasolt "custom warping" eljárást a pólusok által meghatározott frekvenciatorzítási függvénnyel kapott megfontolásokkal (lásd β függvény az idézett cikkben, vagy lásd a monográfiát a részletekért).

8. fejezet MIMO rendszerekre terjeszti ki a javasolt eljárást. Javasolom a szerzőnek átgondolni a megtakarításokra vonatkozó állítást, ami nyilvánvaló tévedés. Ennek belátására vegyen egy diagonális rendszert – egyszerűség végett egy kétszer kettős MIMO rendszert. Triviális, hogy a komplexitás csak a rendtől függ, és nem a konkrét pólusoktól.

Az egy más kérdés, hogy más egyéb szempontból előnyös/célszerű közös pólusrendszerrel dolgozni.

9. fejezet az úgynevezett késleltetett párhuzamos szűrők tervezésével foglalkozik. Ez egy érdemi fejezet, ami a javasolt reprezentáció praktikus előállítására javasol numerikusan használható módszert. Magam részéről nem tartom meggyőzőnek a fejezetben használt érvelést. Nem látom be, hogy a késleltetés miért segítene a módszer alapproblémáján (potenciálisan rosszul kondicionált feladat). Az idő és frekvenciatartományi érvelés keveredése sem segíti a megértést.

Kérdés: Instabil szűrő esetén honnan lesz impulzus válasz információnk vagy honnan származik $A(z^{-1})$?

10. fejezet ad egy összefoglalást és egyúttal egy felsorolást is a szerző egyéb eredményeiről.

A tézisek tartalmi értékelése

1. téziscsoport foglalja össze a fix pólusú párhuzamos szűrő tervezését és tulajdonságait érintő kérdéseket. Ebben a téziscsoportban az 1.2 pontot nem tudom elfogadni tézis értékű állításnak, illetve új tudományos eredménynek.

2. téziscsoport foglalja össze az új póluselrendezési eljárásokra vonatkozó eredményeket. Ezt a tézist új tudományos eredménynek fogadom el.

3. téziscsoport MIMO rendszerek modellezésével és kompenzációjával kapcsolatos eredmények összefoglalása. Az első mondatban megfogalmazott állítástól eltekintve a tézist új tudományos eredménynek tudom elfogadni.

4. téziscsoport a késleltetett párhuzamos szűrő tervezésére vonatkozó eredményeket taglalja. Ezeket új tudományos eredménynek fogadom el.

Végezetül meg kell jegyezni, hogy a pályázó sikeres műszaki alkotással illetve beadott szabadalommal is rendelkezik, amelyben a dolgozatban ismerttetett módszereket alkalmazta.

Az értekezésben szereplő saját eredményekhez kapcsolódóan a szerző 15 lektorált első- vagy utolsó szerzős nemzetközi folyóiratcikket sorol fel. Ezek közül 6 a szakterület egyik vezető, Q1-es besorolású kiadványa. A szerző

minden egyes tézisben megfogalmazott eredményét legalább egy különálló impakt faktoros folyóiratcikkben publikálta.

Összefoglaló vélemény

Megállapítható, hogy a pályázó jelentős új tudományos eredményeket ért el a szűrőtervezés területén a PhD fokozat megszerzése óta. Kiemelendő, hogy a kidolgozott eredményeket munkatársaival számos különböző területen sikeresen alkalmazta. A kritikai észrevételek ellenére a téziseket, a megjelölt kivételekkel, a pályázó önálló, új tudományos eredményeiként elfogadom. Javaslom a nyilvános vita kitűzését és sikeres védelem esetén Bank Balázs részére az MTA doktora cím odaítélését a műszaki tudományok területén.

Budapest,
2022, Április 25



Dr. Szabó Zoltán
MTA Doktora,
egyetemi tanár