

NEMLINEÁRISAN TORZULT OPTIKAI
HANGFELVÉTELEK HELYREÁLLÍTÁSA
REGULARIZÁLT INVERZ
KARAKTERISZTIKÁKKAL

Ph.D. értekezés tézisei

Bakó Tamás Béla
okleveles villamosmérnök

Témavezető: dr. Dabóczy Tamás
a műszaki tudomány kandidátusa

BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
MÉRÉSTECHNIKA ÉS INFORMÁCIÓS RENDSZEREK TANSZÉK

Budapest
2004.

1. Bevezetés

A professzionális hangosfilmeknél a hangot általában optikai úton rögzítik a filmszalagra. Az optikai hangrögzítés előnye, hogy a hangot a rögzített képpel együtt lehet kezelni, ugyanazokkal a technikákkal, amelyek a kép rögzítésére, másolására és vágására vonatkoznak. Éppen ezért a hangosfilm kezelése a hangrögzítés után nem igényel semmilyen további különleges eszközt a némafilmhez képest. A másik előny, hogy a hang lejátszásakor a lejátszóeszköznek nem kell hozzáérnie a filmszalaghoz és a filmszalag nem fog kopni, élettartama emiatt nem fog csökkenni. Ezen tulajdonságok miatt terjedt el az optikai hangrögzítés a hangosfilmeknél az 1920-as 30-as években és használják még napjainkban is.

Az optikai hangrögzítésnek azonban akad egy gyenge pontja is: a hang nemlineáris torzulása, ami abból fakad, hogy a rögzítéshez használt fényérzékeny anyagok ún. feketedési karakterisztikája – az adott fény expozíció hatására az előhívás után a filmen létrejövő feketedés nagysága – erőteljesen nemlineáris tulajdonságú és ez a nemlineáris tulajdonság megjelenik a lejátszott hangban is.

Az optikai hangrögzítésnek két alap módszere létezik: a változó intenzitás alapú és a változó terület alapú módszer. A hangosfilmek kezdeti időszakában főleg az intenzitás alapú hangrögzítést használták. Az intenzitás alapú hangrögzítés során a filmen a hangjel pillanatnyi értékét a filmszalag adott pontja feketedésének a mértéke határozza meg. Éppen ezért ennél a formánál a nemlineáris torzulás közvetlenül, mint a feketedési karakterisztika nemlinearitási hibája jelenik meg. Mivel a feketedés mértéke egy adott hangmintánál gyakorlatilag nem függ a korábbi illetve későbbi hangmintáktól, ezért ezt a fajta torzítást memóriamentes torzításnak tekinthetjük.

A változó terület alapú módszer kicsit később terjedt el. Ennél a módszernél a film hangsávjában ideális esetben csak teljesen fehér illetve teljesen fekete részletek vannak és a fekete és fehér részek aránya határozza meg a pillanatnyi hangértéket. A nemlinearitás ebben az esetben azért jelenik meg, mert a gyakorlatban a fekete és fehér részek közötti átmenet nem ugrásszerű

és az egyenletes átmenet a film nemlineáris feketedési karakterisztikája miatt jelentősen torzul. Ebben az esetben a torzulást jelentősen befolyásolják a korábbi, illetve későbbi hangminták, ezért a torzulás csak mint memóriával rendelkező torzulás írható le.

Lejátszás során a hang nemlineáris torzulása zavaró hatású: a torzult hang fárasztja a közönséget, akik kevésbé tudnak koncentrálni magára a filmre, ezáltal a film élvezhetősége csökken. A nemlineáris torzulás mértéke néhány esetben olyan nagy lehet, hogy a filmben elhangzó beszéd érthetősége is lecsökken. A nemlineáris torzulást azonban kompenzálni lehet.

Régebben a kompenzálás egyetlen lehetősége a filmszalag átmásolása volt, amikor is a másoláshoz használt filmanyag feketedési karakterisztikáján egy olyan munkapontot választottak ki, amelynél átmásolás után az újabb nemlineáris torzulás részben kompenzálta az eredetit. Ez a megoldás azonban rengeteg kísérletezést igényel, mivel a hang pontos torzulása nem ismert. A nemlinearitás továbbá csak részben kompenzálható, mivel a feketedési karakterisztika egyetlen részlete sem pontos inverze a karakterika más részletének. A másolás során ezenkívül a hangfelvételhez zaj adódik hozzá, ami a fényérzékeny anyag szemcsésességéből, az ebből fakadó apró fényingadozásokból és ennek a hangban való megjelenéséből adódik.

Mindezek a problémák elkerülhetők, ha a hangot digitális jelfeldolgozás segítségével állítjuk helyre. Ma már rengeteg hangrestauráló eljárás létezik, amelyek segítségével kattogásokat, sercegéseket, szélessávú zajt lehet a felvételekből eltávolítani vagy éppen a hangmagasság-ingadozást lehet lecsökkenteni, azonban viszonylag kevesen foglalkoztak a nemlineáris torzulások csökkentésével. Éppen ezért kutatásom során a régi filmek optikai úton rögzített hangjának feljavítását tűztem ki célul.

2. Vizsgálati módszerek

A régi filmek állapota az idő haladtával egyre jobban romlik, aminek oka a film hordozóanyagának öregedése. Rengeteg régi film vár felújításra, ami miatt csak kevés idő juthat az egyes filmekre. Ezért esett a választásom az intenzitás alapú hangfelvételek hangjának helyreállítására, mivel ezek a filmek készültek régebben és ezek szorulnak elsősorban megmentésre.

Az intenzitás alapú optikai hangfelvételek nemlineáris torzulása jól leírható memóriamentes nemlineáris torzulással:

$$y(t) = N(x(t)), \quad (1)$$

ahol $x(t)$ a bemenő, torzítatlan jel a t időpontban, $N()$ a rendszer nemlinearitása, $y(t)$ pedig a nemlineárisan torzult jel a t időpontban.

A nemlineáris torzulások kompenzálását három csoportra oszthatjuk:

- Amennyiben van rá lehetőségünk, átépíthetjük a nemlineáris rendszer struktúráját, amivel magát a nemlinearitást lehet megszüntetni vagy lecsökkenteni.
- Amennyiben a rendszer struktúráján nem tudunk változtatni, de hozzáférésünk van az x bemenő jelhez, akkor a bemenő jel előtorzításával kompenzálhatjuk a nemlinearitást:

$$\hat{x} = N(P(x)), \quad (2)$$

ahol $P()$ az előtorzító nemlineáris karakterisztika és \hat{x} az így előállított becslő a bemenő jelről.

- Amennyiben sem a rendszer struktúrájához, sem pedig a bemenő jelhez nem tudunk hozzáférni, akkor a kimenő jelet utólagosan is kompenzálhatjuk:

$$\hat{x} = K(y) = K(N(x)), \quad (3)$$

ahol $K()$ a kompenzáló karakterisztika.

Mivel a régi filmek esetén semmi más nem áll a rendelkezésünkre, csak a már torzult hangot tartalmazó film, emiatt az első két lehetőség nem jöhet szóba. A megoldás csak egyfajta utólagos kompenzáció lehet.

Amennyiben a nemlineáris torzulás invertálható, az utólagos kompenzáció elvileg tökéletesen megvalósítható. Azonban a gyakorlatban a jel nem állítható tökéletesen vissza, mert a felvétel és a lejátszás során megjelenő zajok, bizonytalanságok ennek határt szabnak. Ha a torzulás után a kimeneten megjelenő jelhez hozzáadódó zajt is figyelembe vesszük, akkor a (3) képlet a következőképpen módosul:

$$\hat{x} = K(y + n) = K(N(x) + n), \quad (4)$$

ahol n a kimeneti jelhez hozzáadódott, attól függetlenül tekintett nulla középértékű szélessávú zaj.

Az egyik probléma, hogy a helyreállítás erőteljes nemlineáris torzulások esetén rosszul kondicionált lehet, ami azt jelenti, hogy kis változások a kompenzálendő jelben jelentős eltéréseket okoznak a kompenzált jelben.

A probléma rosszul kondicionáltsága jól megmutatható, ha felírjuk a kompenzáció képletét n kis változásaira:

$$\hat{x}|_{x=x_0} = K(N(x_0) + n) \approx K(N(x_0)) + \left. \frac{dK(x)}{dx} \right|_{x=x_0} \cdot n. \quad (5)$$

Jól látható, hogy amennyiben az adott x_0 munkapontban az inverz nemlineáris karakterisztika deriváltja nagy, akkor a zaj rendkívül felerősödik, ami használhatatlanná teheti az eredményt. A megoldást egy olyan kompenzáló karakterisztika jelentené, ami pontos eredményt szolgáltat akkor, ha a derivált értéke kicsi, és csillapítaná a zajt akkor, ha az eredeti, egzakt inverz deriváltja az adott pontban túlzottan felerősítené a zajt.

Egy másik, az előzőtől különböző probléma, hogy a helyreállítás során a kimenethez hozzáadódott zaj is nemlineárisan torzul, ami pontatlanná teheti a megfigyelt jel bizonyos paramétereinek vizsgálatát (például a helyreállított jel amplitúdója, várható értéke eltérhet a valótól). A kompenzált jel vár-

ható értékére felírhatjuk:

$$E\{\hat{x}\} = \int_{-\infty}^{\infty} f_{\hat{x}}(\hat{x})\hat{x} d\hat{x} = \int_{-\infty}^{\infty} f_n(n) \cdot K(N(x) + n) dn, \quad (6)$$

ahol $E\{a\}$ az a valószínűségi változó várható értékét jelenti, $f_a(a)$ pedig az a valószínűségi változó valószínűség-sűrűségfüggvénye. A zaj hatása a becsült jel várható értékére úgy írható, mintha a kompenzálást a zajmentes torzult jelen nem az egzakt inverzzel, hanem egy olyan függvénnyel hajtottuk volna végre, ami a kompenzáló karakterisztika és a zaj valószínűség-sűrűségfüggvényének korrelációjával írható le:

$$R_{f_n(n),K(n)}(z) = \int_{-\infty}^{\infty} f_n(n) \cdot K(z + n) dn. \quad (7)$$

Ebben az esetben a várható érték torzítatlansága érdekében egy olyan kompenzáló karakterisztikára van szükség, amely korrelálva a zaj valószínűség-sűrűségfüggvényével, az egzakt inverzet adja vissza.

A különböző területeken eddig megvalósított és használt nemlineáris kompenzálási módszerek iteratív eljárásokon alapulnak. Ezek hátránya az általában nagy számításigény, valamint, hogy sok esetben nem lehet előre megállapítani az algoritmusok konvergenciasebességét, emiatt valósidejű alkalmazásokban nehezen lehet őket használni. Céлом olyan módszer kifejlesztése volt, ami gyors és lehetőleg nem igényel iterációt, valamint minél kevesebb emberi beavatkozást igényel.

3. Új tudományos eredmények

Két új módszert adtam memóriamentes torzulások zajos környezetben való kompenzálására. Az első módszer célja, hogy a jelet úgy állítsa helyre, hogy az négyzetes értelemben minél kevésbé térjen el az eredeti jeltől. A második módszer célja, hogy a jel várható értéke torzításmentes legyen. Megvizsgáltam a módszerek konvergenciáját és javaslatokat adtam a módszerek paramétereinek optimális beállításához.

I. tézis: Memóriamentes nemlineáris torzulások utólagos kompenzálására Tyihonov-regularizáción alapuló, nem iteratív eljárást dolgoztam ki.

I/1. Megmutattam, hogy a Tyihonov-kompenzáló karakterisztika deriváltját a következő módon lehet kiszámolni (kézirat 6.4. fejezet, 50–53. o.):

$$\left. \frac{dK(y)}{dy} \right|_{y=N(x_0)} = \left. \frac{\frac{dN(x)}{dx}}{\left(\frac{dN(x)}{dx}\right)^2 + \lambda} \right|_{x=x_0}, \quad (8)$$

ahol $K(y)$ a kompenzáló karakterisztika, $N(x)$ az eredeti nemlinearitás függvény, λ pedig az ún. Tyihonov-féle regularizációs paraméter, amely segítségével kompromisszum található a túlságosan felerősödött zaj és a nagy torzítás között.

A derivált alapján $K(y)$ numerikus integrálással számítható. Az integrálási állandóra becslő adható a következő egyenlet megoldásával:

$$E\{N(F(y) + C)\} = E\{y\}, \quad (9)$$

ahol $F(y)$ a derivált egy előre megválasztott fix integrálási állandó mellett kiintegrált formája, C pedig a választott integrálási állandóhoz adandó korrekciós érték.

I/2. Megmutattam, hogy a Tyihonov-regularizáción alapuló eljárás $\lambda = \frac{E\{n^2\}}{E\{x^2\}}$ esetén a négyzetes értelemben optimális megoldáshoz közeli megoldást ad vissza. Ha a regularizációs paramétert más úton nem

tudjuk optimalizálni, ez a becslő jó eredményt szolgáltat, továbbá könnyen számítható (kézirat 6.4.1. fejezet, 53–60. o.).

I/3. A gyakorlatban sok esetben nem ismerjük a torzítást okozó nemlineáris függvényt. Optikai hangfelvételek helyreállítása esetén szintén felmerül ez a probléma. Ebben az esetben, a bemenő jel és a torzító rendszer bizonyos tulajdonságainak a kihasználásával vak identifikációs (blind identification) módszer dolgozható ki. A bemenő jel vizsgált részletét periodikusnak feltételezve, valamint a fényérzékeny anyag feketedési görbéje általános képletének az ismeretében kidolgoztam egy módszert, ami jó becslést ad a feketedési görbe konkrét paramétereire (kézirat 6.2. fejezet, 46–58. o.).

A módszer során a bemenő jel egy rövid szakaszát (kb. néhány-szor 100 mintát, amely rövid szakaszon a jel még periodikusnak tekinthető) véges hosszúságú Fourier-sorral modellezzük: egy alap szinuszzel és annak felharmonikusaival. A nemlinearitás modelljében a film feketedési karakterisztikájának leírására az általában használt γ függvényt alkalmazzuk:

$$y = x^\gamma. \quad (10)$$

A modellben további tulajdonságokat kell figyelembe venni, úgy mint a bemeneti offszet (O_1), a kimeneti offszet (O_2) és a kimeneti erősítés (G):

$$y = G \cdot (x + O_1)^\gamma + O_2. \quad (11)$$

A szinuszjelek fázisát, amplitúdóját, valamint a nemlinearitás modelljének paramétereit például Monte-Carlo módszer segítségével optimalizálhatjuk, hogy a modell által alkotott kimenő jel négyzetes értelemben minél kevésbé különbözzön az eredeti zajos és torz jelerészlettől. A minimális eltérésnél kapott O_1 , O_2 és G paraméterekkel alkotott nemlineáris függvény jó becslője az eredeti nemlinearitásnak.

II. tézis: Az előző tézispontban leírt kompenzálási móddal, adott értékű regularizációs paraméter mellett a hiba energiájának várható értéke a következő módon számítható:

$$E\{(x - \hat{x}(\lambda))^2\} = \int_{-\infty}^{\infty} p_x(\chi) \cdot \int_{-\infty}^{\infty} p_n(\nu) \cdot (K(N(\chi) + \nu, \lambda) - \chi)^2 d\nu d\chi, \quad (12)$$

ahol $p_x(x)$ a bemenő jel, x , valószínűség-sűrűségfüggvénye.

A regularizációs paraméter optimális értékének kiszámítása a (12) egyenlet λ szerinti minimalizálásával lehetséges. Azonban a bemenő jel valószínűség-sűrűségfüggvénye általában nem ismert, ezért az egyenlet a gyakorlatban sokszor nem oldható meg.

Kidolgoztam egy iteratív algoritmust, amellyel a regularizációs paraméter optimális értékére jó becslő adható x és $p_x(x)$ ismerete nélkül (kézirat 6.4.2. fejezet, 60–64 o.):

1. $p_x(x)_0 = p_y(y)$.
2. λ kiszámítása a (12) egyenlet λ szerinti minimalizálásával.
3. $K(y, \lambda)$, \hat{x} kiszámítása, majd ez utóbbi alapján $p_{\hat{x}}(\hat{x})_i$ kiszámítása.
4. $p_x(x)_i = p_{\hat{x}}(\hat{x})_i$.
5. Ha az iterációk száma $\geq N$ vagy az eltérés csökkenése $\leq \varepsilon$, akkor kilépés, egyébként vissza a második lépésre.

III. tézis: Megmutattam, hogy zajos, nemlineárisan torzult jel esetén – a zaj valószínűség-sűrűségfüggvénye és a nemlineáris torzulás karakterisztikája ismeretében – konstruálható egy olyan kompenzáló függvény, ami úgy állítja helyre a jelet, hogy a jel várható értéke torzításmentes legyen.

III/1. Megmutattam, hogy végtelen számú ilyen karakterisztika létezik, de ezen karakterisztikák legtöbbje a gyakorlatban alkalmatlan, mert oszcillációkat tartalmaz (kézirat 6.7. fejezet, 82–83. o.).

III/2. Megadtam egy olyan iteratív algoritmust, amelyik képes oszcilációmentes megoldást (szigorúan monoton bemenetre szigorúan monoton kimeneti választ) találni (kézirat 6.7. fejezet, 84. o.):

$$\begin{aligned} K_0(y) &= N^{-1}(y), \\ K_i(y) &= K_{i-1}(y) + \alpha \cdot (N^{-1}(y) - K_{i-1}(y) * f_n(y)), \end{aligned} \quad (13)$$

ahol $*$ korrelációt jelent, $K_i()$ a kompenzáló karakterisztikára adott i . iteráció, N^{-1} az egzakt inverz nemlinearitás függvény, $f_n(y)$ a zaj valószínűség-sűrűségfüggvénye, α pedig az iteráció sebességét és a konvergenciát befolyásoló paraméter.

III/3. Bebizonyítottam, hogy az algoritmus megfelelő paraméterbeállításal konvergens és a paraméter megfelelő értéke csak a zaj valószínűség-sűrűségfüggvénye jellegén múlik (kézirat 6.7.2. fejezet, 84–85. o.). A konvergencia feltétele:

$$\alpha \leq \frac{1}{\max(|F_n(f)|)}, \quad (14)$$

ahol $F_n(f)$ az f_n zaj valószínűség-sűrűségfüggvény Fourier-transzformáltja.

4. A gyakorlati alkalmazás lehetőségei

Az optikai úton, intenzitás módszerrel rögzített filmhang rendkívül érzékeny a felvételi és előhívási körülményekre. Rossz körülmények esetén a hang erőteljesen eltorzulhat. Ez a felvételi mód jól modellezhető egy memóriamentes torzítással rendelkező, zajos jelfeldolgozó egység modelljével.

A torz filmhang jelenleg alkalmazott egyetlen helyreállítási módja a film hangsávjának átmásolása különböző megvilágítási és előhívási paraméterekkel. Megfelelő paraméterek esetén a film feketedési karakterisztikájának olyan részére kerül a hang, amelyik kismértékben ellensúlyozni tudja az eredeti torzítást. Ezzel a módszerrel azonban nem hozható létre az optimális kompenzálás, továbbá a megfelelő paraméterek megtalálása csak „kézi” módon: emberi beavatkozással, hosszú ideig tartó próbálgatással lehetséges.

Szimulációk segítségével megmutattam, hogy az első tézispontban leírt nem iteratív kompenzálási mód segítségével zajos, torz hangjelek hatékonyan és jó minőséggel helyreállíthatóak. A Magyar Nemzeti Filmarchívumtól kapott filmfelvétel részletek segítségével megmutattam, hogy a módszer a gyakorlatban is jól alkalmazható.

Az algoritmust egy hangrestaurálással foglalkozó cég szoftveréhez implementáltam. Az algoritmus tesztelése jelenleg is folyik.

5. Az értekezés témakörében készült tudományos közlemények

Külföldön megjelent idegen nyelvű folyóiratcikkek

1. Dabóczi T. and T. B. Bakó, "Inverse Filtering of Optical Images", IEEE Trans. on Instrumentation and Measurement, Vol. 50, No. 4, pp. 991-994, 2001. *lektorált, referált*
2. Tamás B. Bakó and Tamás Dabóczi, "Reconstruction of Nonlinearly Distorted Signals With Regularized Inverse Characteristics", IEEE Trans. on Instrumentation and Measurement, Vol. 51, No. 5, pp. 1019-1022, 2002. *lektorált, referált*

Nemzetközi konferencia-kiadványban megjelent idegen nyelvű előadások

1. Tamás B. Bakó, Tamás Dabóczi, "Inverse Filtering of Optical Images", IMTC 2000, Baltimore, USA, May 1-4, 2000, Proceedings of the IEEE Instrumentation and Measurement Technology Conference, No. 00CH37066, pp. 991-994. *referált*
2. Tamás B. Bakó, Tamás Dabóczi, "Reconstruction of Nonlinearly Distorted Signals with Regularized Inverse Characteristics", IMTC 2001, Budapest, Hungary, May 21-23, 2001, Proceedings of the IEEE Instrumentation and Measurement Technology Conference, No. 01CH37188, pp. 1565-1570. *referált*
3. Tamás B. Bakó, Balázs Bank, Tamás Dabóczi, "Restoration of Nonlinearly Distorted Audio with the Application to Old Motion Pictures", AES 20th International Conference on Archiving, Restoration and New Methods of Recording, Budapest, Hungary, Oct 5-7, 2001, No. 88-650002, pp. 191-198. *referált*

4. Tamás B. Bakó, Tamás Dabóczy, "Unbiased Reconstruction of Non-linear Distortions", IMTC 2002, Anchorage, USA, May 21-23, 2002. Proceedings of the IEEE Instrumentation and Measurement Technology Conference, No. 00CH37276, pp. 1099-1102. *referált*
5. Tamás B. Bakó, Tamás Dabóczy and B. A. Bell, "Automatic Compensation of Nonlinear Distortions", IMTC 2002, Anchorage, USA, May 21-23, 2002. Proceedings of the IEEE Instrumentation and Measurement Technology Conference, No. 00CH37276, pp. 1321-1325. *referált*

Magyar folyóiratcikk

1. Bakó Tamás, "Hangfelvételek digitális restaurálása", Akusztikai Szemle *megjelenés alatt, lektorált*

Magyar konferencia-kiadványban megjelent előadások

1. Tamás B. Bakó, "Deconvolution of two-dimensional signals", Proceedings of the 7th Mini-Symposium, Budapest, Jan 27-28, 2000, pp. 30-31.
2. Bakó Tamás, "Hangfelvételek digitális restaurálása", TV 2000, 9. Televízió és Hangtechnikai Konferencia és Kiállítás, Budapest, 2000. május 23-25, 131-140.
3. Bakó Tamás, "Inverz szűrés és alkalmazásai", Pro Scientia Aranyérmesek V. Tudományos Konferenciája, Sopron, 2000. november 5-7., 129-133.
4. Tamás B. Bakó, "Reconstruction of nonlinearly distorted signals", Proceedings of the 8th Mini-Symposium, Budapest, Jan 31 - Febr 1, 2001, pp. 30-31.
5. Tamás B. Bakó, "Automatic compensation of nonlinear distortions", Proceedings of the 9th Mini-Symposium, Budapest, Febr 4-5, 2002,

pp. 30-31.

6. Bakó Tamás, "Filmhang restaurálás: a nemlineáris kompenzálás egy gyakorlati alkalmazása", Pro Scientia Aranyérmesek VI. Tudományos Konferenciája, Miskolc, 2003. november 28-30., (előadás vázlatok 11-12. oldal). *megjelenés alatt*

Elektronikus publikáció

1. Bakó Tamás (BME), "Régi mozifilmek digitális hangfelújításának technikái", Nemzetközi Filmfelújítási Szeminárium, Budapest, 2001. márc. 10-15. <http://www.filmintezet.hu/magyar/filmint/filmspir/27/bako.htm>
Kereshető: Google, filmspirál 27, filmek hangjának digitális felújítása