SKOLNIK VILMOS—TÓTH LÁSZLÓ
BME Mikrohullámú Híradástechnika Tanszék

A televízió-műsorjelek DPCM kódolására vonatkozó kísérletek

ETO 621.397.334:621.376.56

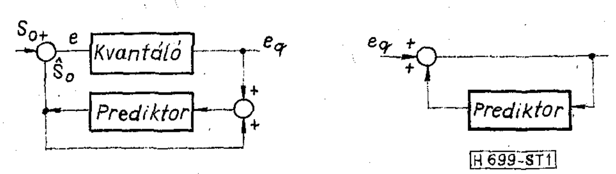
A digitális technika széles körű elterjedése a hírátvitelben lehetőséget nyújt a televíziós műsorjelek ezen elvű továbbítására a PCM hierarchia megfelelő szintjén. A távközlésben a PCM adatátviteli sebességszintek kötöttek, ezért legtöbbször a híranyagok redundanciájának megfelelő redukálásával lehet a híranyagokat diszkrét szinteken továbbítani.

A színes televíziós műsorjel és a megfelelő minőségű kísérő hangjel digitális elvű átvitele a PCM hierarchia európai III. szintjén (34 Mbit/s) lehetséges. E sebességtartomány körül működik már az ITT–SEL-rendszer, amely az átviteli útba bekapcsolta a műsorközlő műholdakat is. Jelentős kísérleti eredményeket ért már el a francia OCCITAN és a nyugatnémet BOSCH cég is.

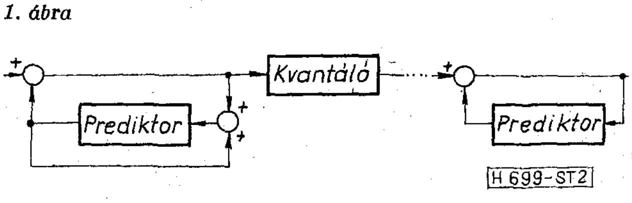
Egy ilyen jellegű, III. szintű multiplex jel felépítésének ismertetésére most nem térünk ki, de megjegyezzük, hogy tisztán a képjel továbbítására durván 30 Mbit/s áll rendelkezésre. A fennmaradó részbe vannak beültetve a hang, a különféle mérő- és adatjelek, valamint a hibajavítás is. A fenti számérték tulajdonképpen a kódolási eljárások választásának alapja.

A kódolási algoritmusok választása természetesen a videójel statisztikai jellemzőinek tükrében történhet. A redundáns információk csökkentésére több lehetőség is kínálkozik, mint például a DPCM elv, a transzformációs kódolás, illetve ezek együttes alkalmazása. A jelenlegi technikai, illetőleg technológiai feltételek alapján a DPCM elv felhasználása tűnik reálisnak. A továbbiakban egy, a fentebb említett célra alkalmas DPCM kódoló berendezés néhány tervezési szempontját vizsgáljuk meg.

A komplett DPCM kódoló analízise igen bonyolult. Erre vonatkozó eredményeket a különféle irodalmakban is csak igen speciális esetekre találhatunk. Most a kódolót alkotó két fő elem, a prediktor és a kvantáló optimalizálását külön-külön tárgyaljuk. Ez azt jelenti, hogy az 1. ábrán látható DPCM kódoló helyett a 2. ábra tömbvázlatának megfelelő kódolót vizsgáljuk.



1. ábra



2. ábra

Először a prediktort vizsgáljuk meg, amelynek megválasztását a következő szempontok határozhatják meg:

a) a videojel (most csak a világossági jelet vizsgálva), statisztikai jellemzői alapján, az aktuális képpontot megelőző mintákból a kérdéses minta lineáris, négyzetes-közép értelemben optimális becslése,

b) a videojel nem stacionárius viselkedésének és a predikciós hibák szubjektív hatásának figyelembevételével az előzőkben meghatározott predikciós algoritmusok módosítása, amely egyrészt az együtthatók módosítását, másrészt a predikció adaptívvá tételét jelentheti,

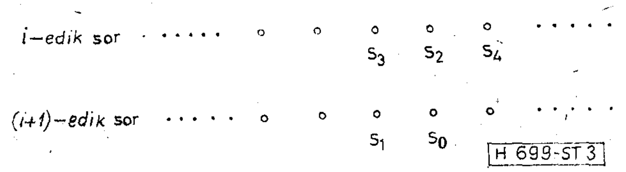
c) a dekódoló csatornahibákkal szembeni érzéketlenségének biztosítása,

d) realizálási szempontok, amelyeket a fenti módszerek mindegyikénél figyelembe kell vennünk.

A terjedelem korlátozott volta miatt részletesebben csak az *a)* pontra vonatkozó analízissel foglalkozunk.

A videojel mintáinak lineáris négyzetes-közép becslése

A videojelet gyengén stacionárius sztochasztikus folyamatnak tekintve, feltételezzük az autokorrelációs függvény ismeretét. Ebből a stacionárius folyamatból rendre a t_0, $t_0-t_1, \ldots, t_0-t_n$ időpontokban az $S(t_0)=S_0$, $S(t_0-t_1)=S_1, \ldots S(t_0-t_n)=S_n$ mintákat vehetjük, amelyek természetesen valószínűségi változók (3. ábra).



3. ábra

A prediktor feladata, hogy egy konkrét realizáció esetén a rendelkezésre álló $S_1, S_2, \ldots S_n$ értékek alapján egy $\hat{S}_0$ becslést állítson elő az S_0 értékre. A becslés stratégiáját illetően gyakorlatilag csak a lineáris négyzetes-közép becslés jöhet szóba, ahol az S_0 érték a már realizálódott értékek egy lineáris kombinációja:

$$S_0 = a_1 S_1 + a_2 S_2 + \ldots + a_n S_n .$$

Feltételezzük, hogy a videojel várható értéke zé-

Beérkezett: 1979. X. 12.

rus. Az a_i együtthatókat úgy választjuk meg, hogy a predikció négyzetes közép hibája minimális legyen:

$$\sigma_e^2 = \mathrm{E}[(S_0-\hat{S}_0)^2] \Rightarrow \min.$$

Képezzük az a_i-k szerinti parciális deriváltakat, és azokat nullával egyenlővé téve kapjuk:

$$\frac{\partial \mathrm{E}[(S_0-\hat{S}_0)^2]}{\partial a_i} = \ldots =$$

$$= -2\mathrm{E}[(S_0-(a_1S_1+\ldots+a_nS_n))S_i]=0,$$

$$\mathrm{E}[(S_0-\hat{S}_0)S_i]=0, \quad i=1,2,\ldots,n.$$

A szélsőérték-számítás eredményeképpen az az általánosan ismert ortogonalitási elv adódott, amely szerint az optimális lineáris középbecslésnél a becslés hibája ortogonális a becslésben felhasznált adatokra.

Vezessük be az $\mathrm{E}(S_i, S_j)=R_{ij}$ jelölést, ahol az R_{ij} a mintavett sorozat i és j indexű elemeinek kovarianciája, így

$$R_{0i}=a_1R_{1i}+a_2R_{2i}+\ldots+a_nR_{ni}.$$

Ez alapján egy lineáris egyenletrendszert kapunk, amelyből az a_i együtthatók meghatározhatók. Az egyenletrendszert mátrix alakban felírhatjuk, figyelembe véve, hogy $R_{ij}=R_{ji}$, és a minták szórására bevezethetjük az $R_{ii}=\sigma^2$ jelölést.

$$\begin{bmatrix} R_{01} \\ \vdots \\ R_{0n} \end{bmatrix} = \begin{bmatrix} \sigma^2 R_{12} \ldots R_{1n} \\ \\ \sigma^2 \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{R}_0=\mathbf{R}\mathbf{a}.$$

Az $\mathbf{R}$ mátrix az $(S_1 \ldots S_n)$ sorozat autokovariancia mátrixa. Az R_{ij} értékekre azonban további megkötések származnak a folyamat stacionárius jellegéből, amelyeket konkrét esetekben, miután már megválasztottuk a becslésben szereplő pontokat, figyelembe tudunk venni.

Meghatározhatjuk a predikció négyzetes hibáját

$$\sigma_e^2=\mathrm{E}[(S_0-\hat{S}_0)^2]=\mathrm{E}[(S_0-\hat{S}_0)S_0]-\mathrm{E}[(S_0-\hat{S}_0)\hat{S}_0].$$

Az ortogonalitási elv miatt az egyenlet jobb oldalának második tagja zérus. Így

$$\mathrm{E}[S_0\cdot\hat{S}_0]=\mathrm{E}[\hat{S}_0],$$

tehát

$$\sigma_e^2=\mathrm{E}[S_0^2]-\mathrm{E}[S_0\cdot\hat{S}_0^2]=\sigma^2-\mathrm{E}[\hat{S}_0^2].$$

Azt az eredményt kaptuk, hogy a hibajel szórása kisebb, mint az eredeti jel szórása. Emellett a hibajel kevésbé korrelált mintákat tartalmaz.

A fentiek alapján, elegendő számú képre vonatkozó mérések birtokában, a prediktor-együtthatók közel optimális megválasztását el lehet végezni. Az együtthatók meghatározása amúgy sem csak matematikai optimalizálási kérdés, hiszen a már korábban leírt egyéb szempontok is befolyásolják a rendszer tervezését. A másik igen lényeges probléma az, hogy a videojelre tett kiindulási feltételezéseink nem teljesülnek, így eredményeink csak közelítő jellegűek, amelyeket konkrét mérésekkel kell ellenőrizni.

Tapasztalatok, szimulációk alapján a következőkben összefoglaljuk a különböző predikciós eljárásokat (a 3. ábra szerint).

Egypontos predikció:	1. $\hat{S}_0=S_1$
Kétpontos predikció:	2. $\hat{S}_0=1/2.S_2$
	3. $\hat{S}_0=1/2.S_1+1/2.S_4$
Hárompontos predikció:	4. $\hat{S}_0=S_1+S_2-S_3$
	5. $\hat{S}_0=S_1+1/2.(S_2-S_3)$

A vizsgálatok eredménye gyakorlatilag az, hogy a három képponton (S_1, S_2, S_3) alapuló becslés már nem javítható lényegesen a további képpontok figyelembevételével.

Kvantálás

A kvantálás lehetséges módszereinek áttekintését eleve a számunkra érdekes területre szűkítjük. Figyelembe véve, hogy fix szóhosszúságú, 3 vagy 4 bites kódolás jöhet szóba mind az Y, mind a színkülönbségi jelekre, nem térünk ki az entrópiakódolási módszerekre.

Így a kvantáló tervezésének alapja mindenképpen a kvantálás hatását jellemző torzítás minimalizálása adott számú döntési szint esetén.

A kvantáló tervezése két alapvető feladatból áll. Az első a kvantálási zajok szubjektív hatását is figyelembe vevő torzítási mértéknek a meghatározása, mely az optimalizálás alapja. A második feladat megfelelő numerikus módszerek alkalmazásával, a már adott optimalizálási kritériumok alapján, az optimális kvantáló reprezentációs és döntési szintjeinek a meghatározása.

A következőkben csak a szóba jövő optimalizálási kritériumokat tekintjük át. A következő jelöléseket fogjuk használni:

x_i döntési szintek,

y_i reprezentációs szintek.

Az indexelést úgy választhatjuk meg, hogy a kvantáló az x_i és x_{i+1} döntési szintek közötti bemenőjelhez az y_i reprezentálási szintet rendeli.

N a reprezentációs szintek száma,

$f(x)$ a hibajel sűrűségfüggvénye.

A torzítás legkézenfekvőbb mértéke az átlagos négyzetes torzítás (MSE — mean square error):

$$D_{\mathrm{MSE}}=\sum_{i=1}^{N}\int_{x_i}^{x_{i+1}}(x-y_i)^2 f(x)\,\mathrm{d}x.$$

Az így definiált torzítás minimalizálására Max dolgozott ki algoritmust. Fix N esetére, különböző $f(x)$ eloszlásokra meghatározta az optimális reprezentálási, illetve döntési szinteket.

Rendelkezésre állnak a kiszámított adatok pl. a Laplace-eloszlásra is, amellyel jól modellezhető a DPCM kódolóban előálló különbségi jel. Az eredmények azt mutatják, hogy az MSE kvantáló igen erősen kompressziós jellegű, kicsi a maximális reprezentálási szintje, így a nagy ugrások átvitele torzított lesz, míg a kis amplitúdójú tartományban a bontása finomabb, mint kellene. A másik probléma az, hogy a kvantálók MS hibája és a nyert szubjektív képminőség nem arányos egymással.

Az MSE kvantálók kedvezőtlen szubjektív hatása felveti annak a szükségességét, hogy szubjektív vizsgálatokon alapuló súlyozófüggvényeket keressünk, amelyeket a torzítás kifejezésében az $f(x)$ helyére írva olyan torzítási kritériumot nyerhetünk, amely jobban kifejezi a képminőséget. Ez alapján bevezethető az átlagos négyzetes szubjektív torzítás (MSSE):

$$D_{\mathrm{MSSE}}=\sum_{i=1}^{N}\int_{x_i}^{x_{i+1}}(x-y_i)^2\cdot V(x)\cdot \mathrm{d}x.$$

A fenti kifejezést először csak egypontos predikcióra vizsgáljuk. Ilyenkor az előálló különbségi jel közvetlenül arányos az eredeti jel meredekségével. A $V(x)$ függvény az ún. láthatósági függvény. Ez annak a zajnak a teljesítményével arányos, amit a jelhez adva, mindannyiszor, amikor a jel változási sebessége adott korlátokon belül megközelíti x-et, ugyanazt a szubjektív képminőséget kapjuk, mint ami egységnyi teljesítményű fehér zajnak a teljes jelhez való hozzáadásával adódik. A láthatósági függvények képről képre változnak, és alapvetően két összetevő határozza meg a jellegüket. Egyrészt a statisztikai jellemzők, másrészt pszichofizikai jellemzők. Ez a hatás általánosságban úgy fogalmazható meg, hogy bizonyos ingerek maszkolni képesek az emberi szem érzékenységét egyéb ingerekkel szemben. Így a láthatósági függvény felírható:

$$V(x)=F\{\mathrm{p}(x),\,\mathrm{m}(x)\},$$

ahol $\mathrm{m}(x)$ a jel meredekségétől függő maszkfüggvény. Ez írja le a szem kis fényváltozásokkal szembeni érzékenységét a jel meredekségének a függvényében. Limb és Rubinstein javasolták a láthatósági függvény következő felbontását:

$$V(x)=\frac{p^{\alpha}(x)}{m(x)}.$$

Az $\mathrm{m}(x)$ maszkfüggvény igen tág határok között független a képstatisztikától, ezt mérések is igazolják. Az α tényező értékére is adtak becsléseket, és mérések alapján meghatározták $\mathrm{m}(x)$ jellegét. A $\mathrm{p}(x)$ sűrűségfüggvény jellegére is adhatunk konkrét feltételeket (Laplace-eloszlású), így a kvantáló méretezése elvégezhető. A fő probléma az, hogy több pontos predikció esetén az előálló különbségi jel nem fejezi ki közvetlenül a fényességváltozás meredekségét. Így a D_{MSSE} kifejezésben még egy tényezőt figyelembe kell venni, amely a hibajel és a jel változási meredekség kapcsolatát írja le.

Az előző módon konstruált MSSE kvantáló általában kevésbé komprimált karakterisztikát szolgáltat, mint az MSE. Emellett mérések igazolták, hogy az MSSE széles határok között arányosnak tekinthető a szubjektív képminőséggel. Jó minőségű műsorjel átvitelére az utóbbi típusú kvantálóval felépített DPCM húrokról nem állnak rendelkezésünkre adatok, így csak kísérletileg dönthető el a módszer hasznossága. Valószínűnek látszik, hogy megfelelő eredményeket csak adaptív kvantálóval érhetünk el. Ez a gyakorlatban úgy realizálható, hogy több kvantáló karakterisztikát alakítunk ki a különböző dinamikájú képrészletek számára, amelyeket célszerűen a prediktor által előállított jel vezérel (pl. a SEL rendszernél is). Ebben az esetben a kapcsolójel a dekódolóban is előállítható, így a kvantáló átkapcsolását végző jelet nem kell külön átvinni. Adaptív rendszerben a vezérelt kvantáló esetén mind az MSE, mind az MSSE kvantálók megfelelő eredményt adhatnak.

A DPCM kódoló két alapvető elemének optimalizálási szempontjait külön-külön tekintettük át. A teljes húrok optimalizálása is lehetséges. Az eljárás lényege, hogy a különbségi jelből kiindulva tervezhető meg a kvantáló. Ezután a kvantálót a húrokba téve felvehetjük a különbségi jel hisztogramját. Ez alapján új kvantáló tervezhető. Ezen eljárás ismétlésével, mérések alapján néhány lépésben előállítható az optimális kvantáló.

IRODALOM

[1] O'*Neal:* Predictive Quantization Systems. BSTJ 1966/5—6.

[2] *Papaulis:* Probabllity, random variables and stochastic process. McGraw—Hill, New York, 1965.

[3] *Pratt:* Digital Image Processing. Wiley-Interscience, New York, 1978.

[4] *Zschunke:* DPCM picture coding with adaptive prediction. IEEE Com. 1977. Nov.

[5] *Limb—Rubinstein:* On the Design of Quantisers for DPCM Coders. IEEE Com. 1978. May.

[6] Jones*:* Minimum Distortion Quantisers. NASA TND 8384; AD—A040—033 1977. March