

# Digitális videojel kódolók számítógépes szimulációi

Balázs Péter-Böröczky Lilla-Fazekas Kálmán  
BME Mikrohullámú Híradástechnika Tanszók

## ÖSSZEFOGLALÁS

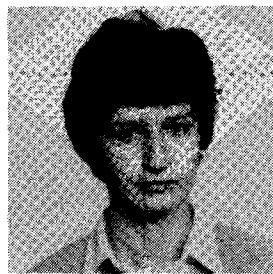
A képadat tömörítésében az alapfeladat a képek tárolásához vagy átviteléhez szükséges adatmennyiség minimalizálása. Ennek a célnak eléréséhez a jelek redundanciájának kihasználása szükséges. Az erre épülő két, széleskörben elterjedt adattömörítő eljárás a prediktív és transzformációs kódolás. E cikkben az említett két osztályba sorolható különféle forráskódolási eljárások (nem adaptív és adaptív intra-/Interframe DPCM, mozgáskompenzált DPCM, Wiener-szűrésen alapuló pel-rekurzív eljárás, transzormációs módszerek, OCT, stb.) számítógépes szimulációjával foglalkozunk. A számítógépes szimuláció a kodek méretezés nélkülözhetetlen alkotórésze, segítségével biztosítható a kitűzött célparaméterek optimális értékének megközelítése, az egzakt méretezés elérése.

## Bevezetés

Az elmúlt évtized technológiai fejlődése és annak nyomán az egyre csökkenő hardware költségek a képek adattömörítő forráskódolási eljárásait realizálhatóvá teszik. A képek digitális átvitele - elsősorban a televíziós képek műholdas és vívőhullámú átviteli alkalmazásai révén (videotelefon, konferencia TV, oktatási csúlv TV, kábel TV, TV szétosztó hálózatok, ipari alkalmazások, B-ISDN, stb.) - egyre nagyobb jelentőségűvé válik és egyre szélesebbkörű felhasználási területtel rendelkezik. Fokozódó érdeklődés kíséri mozgó képek átvitelét kifejezetten kis sebességeken is (pl. 64 kb/s, 384 kb/s, stb.). Igen jelentős kutatás folyik viszonylag olcsó, kisméretű és a különféle alkalmazásokhoz jól illeszkedő digitális kodekek kialakítására.

A képadat tömörítésében az alapfeladat a képek tárolásához vagy átviteléhez szükséges adatmennyiség minimalizálása. Ennek a célnak az eléréséhez a jelek redundanciájának kihasználása szükséges. Az erre épülő két, széles körben elterjedt adattömörítő eljárás a prediktív és a transzformációs kódolás. E cikkben az említett két osztályba sorolható különféle forráskódolási eljárások számítógépes szimulációjával foglalkozunk. A számítógépes szimuláció a kodek méretezés nélkülözhetetlen alkotórésze, segítségével biztosítható a kitűzött célparaméter/ek optimális értékének megközelítése, az egzakt méretezés elérése.

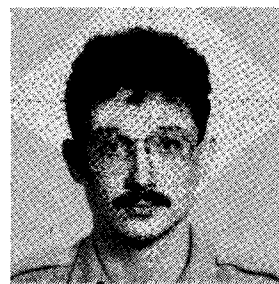
A kodek méretezés általában meglehetősen összetett feladat. A bejövő képjelek a-priori statisztikai paraméterei rendszerint nem ismertek, így az első nehéz feladat a képek adott osztályát jól leíró matematikai modell meghatározása. A megfelelően választott modell alapján határozható meg a kódolási paraméterek (kvantálási karakterisztika, prediktívus együtthatók, transformációs együttható eloszlás, illetve bit-térkép, csatornahibák hatása, illetve a hibák elleni védelem



BÖRÖCZKY LILLA

1987-ben szerzett villamosmérnöki oklevelet a Budapesti Műszaki Egyetem Villamosmérnöki Karán, Híradástechnika szakon. 3 éven keresztül digitális képkódolás témában TDK munkát végzett, szerepelt TDK Konferencián és a "Végzős Konferencia '87" rendezvényen, illetve diplomatervét is hasonló témában írta.

1987. nyarán az IAESTE szervezésében Finnországban 3 hónapos szakmai gyakorlaton vett részt. 1986 óta tagja az EURASIP-nek (European Association for Signal Processing). Jelenleg az MTA-TMB tudományos ösztöndíjasa a BME Mikrohullámú Híradástechnika Tanszékén. Kutatási témája: prediktív képkódoló eljárások vizsgálata és szimulációja.



BALÁZSPÉTER

1987-ben szerezte villamosmérnöki oklevelét a BME Villamosmérnöki Karán. Egyetemi éve alatt aktívan részt vett a tudományos diákköri munkákban. A második évben fakultatív tárgyként felvette az Elektromágneses terek gépi analízise tárgyat, harmadik évben részt vett egy mikrostrip áramköröket analizáló programcsomag kidolgozá-

sában. Negyed évben hallgatta a Videojelek digitális átvitele c. fakultatív tárgyat. Részt vett az 1987-es végzős konferencián és a HTE diplomaterv pályázatán, ahol második díjat kapott. A végzős konferencián leadott dolgozatát a Híradástechnika jelenteti meg. Jelenleg nappali szakmérnök hallgató, tématerülete a digitális jelfeldolgozás, videojel kódolás.

szükséges mértéke, várható képminőség, stb.) és végezhető el a teljes forrás-kódoló dekódoló rendszer szimulációja. A kodek méretezés végső fázisa az áramkör méretezés, amelyet ma már jórészt csak LSI és VLSI elemek felhasználásával végzünk el. Ezek az áramkör elemek vagy berendezés orientált VLSI IC-k, vagy általános célú gyorsdigitális jelfeldolgozó processzorok. Mindkét típus esetén a méretezés során "biztosra kell menni", utólagos beállításokra, javítgatásokra nincs lehetőség. Ez a tény fokozottan hangsúlyozza a szimulációs eljárások alapvető fontosságát.

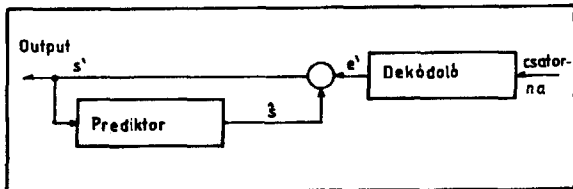
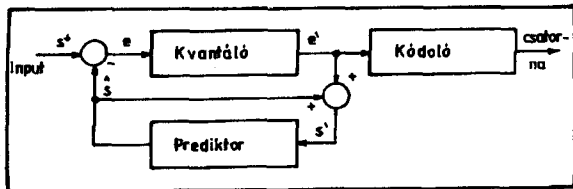
A cikk a 2., 3. és 4. fejezetben foglalja össze azokat a kódolási módszereknek az elvi alapjait, amelyeknek a szimulációjára programcsomag készül. A szimulációs eljárások az 5. és 6. fejezet-

ben kerültek ismertetésre. Míg végül az IBM-AT típusu gépen végzett szimulációk néhány szemléltető eredményét mutatja be.

## 2. Különbségi kódolás

### 2.1 DPCM kódolás

A videojelek adattömörítésére az egyik leggyakrabban alkalmazott módszer a prediktív kódolás. Az alap DPCM rendszer blokkvázlatát az 1. ábra



H484-1

1. ábra DPCM rendszer

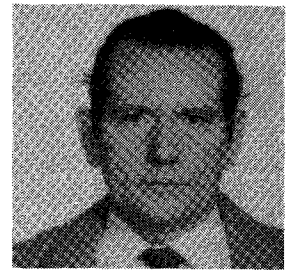
mutatja. Itt az éppen a kódolóba érkező (aktuális) képpont becslése az előzőleg átvitt és dekódolt információ alapján történik. A képpont valódi és prediktált értékének kvantált majd kódolt különbsége kerül átvitelre. Az átvitt kódszavak dekódolása után a vevő a becsült érték és kvantált különbségi jel összeadásával rekonstruálja a képpont valódi értékét. Az adó- és vevőoldali predikció megegyezik.

A DPCM kódolás lehet Intraframe és/vagy Interframe. Az első esetben a predikcióhoz csak az éppen letapogatás alatt lévő félkép rekonstruált képpontjait, míg az Interframe predikciónál az előzőleg átvitt és tárolt félkép képpontjait is felhasználjuk.

A kódolás hatásosságát fokozza a kvantáló és a prediktor adaptív vá tétele. A következő két részfejezet röviden ezek alapelveit ismerteti.

#### 2.1.1. Adaptív kvantálás

Az optimális kvantáló tervezésnél az egzakt matematikai függvények mellett figyelembe kell venni az emberi szem fiziológiai tulajdonságait is. Szubjektív tesztek segítségével elő lehet állítani előre meghatározott képekre a láthatósági különböző függvényt, az u.n. maszkolási függvényt, amely közvetetten tartalmazza a szem érzékeny-



FAZEKAS KÁLMÁN

1962-ben szerzett villamosmérnöki oklevelet a BME Villamosmérnöki Kar Híradástechnika Szakán, majd 1968-ban mérnök-tanári oklevelet. A BME Mikrohullámú Híradástechnika Tanszék adjunktusa. Több, mint tíz éven át előadója volt az Impulzustechnika című tantárgynak, jelenleg az Áramkörök c. tárgyat oktatja. A digitális képfeldolgozás és a videojelek digitális

kódolása és átvitele témakörben fakultatív előadásokat és mérnöktovábbképző tanfolyamokat tartott és tart. A digitális videojel kódolás elméleti kutatásával foglalkozik és kódoló egységek kidolgozásában vesz részt. 1962 óta tagja a HTE-nek, tagja több nemzetközi tudományos szervezetnek (IEEE, EURASIP, SPIE). Részt vesz az Interkozmosz Tanács Úrtávközlési Szakbizottságának munkájában. Virág-Pollák díjas.

ségét különböző hibákra. Ezek a szubjektívan megmért maszkolási függvények felhasználhatók az adaptív kvantálók tervezésénél is. A kép ugyanis részekre bontható aktivitás függvények segítségével, amelyeket a kép lokális statisztikai tulajdonságai determinálnak. Minden egyes részhez ezután a neki megfelelő kvantálási karakterisztikát használjuk. Az aktivitás függvény egy lehetséges definíciója

$$AMD = \max_{i,j \in DN} |d_{ij}| \quad (1)$$

ahol  $d_{ij} = S'_i - S'_j$  a szomszédos képpontok különbsége és  $DN$  az  $AMD$ -hez felhasznált képpontok indexhalmaza. A legtöbb esetben  $DN = \{1, 2, 3, 4\}$ . Az aktivitás függvény és a maszkolási függvény megfelelő összevetésével optimális kvantáló karakterisztikák tervezhetők. Az adaptív kvantálót realizálhatjuk  $N$  darab különböző karakterisztikájú kvantáló  $AMD$  aktivitás függvény szerinti kapcsolásával:

$$Q(e) = \begin{cases} Q_1(e) & AMD < a_2 \\ Q_2(e) & a_2 \leq AMD < a_3 \\ \vdots & \\ Q_n(e) & a_n \leq AMD \end{cases} \quad (2)$$

#### 2.1.2. Adaptív predikció

A képjel erősen nemstacionárius volta miatt a predikciós hiba csökkentése érhető el olyan prediktorral, amelyet adaptív vá teszünk a képjel lokális tulajdonságaihoz.

Adaptív predikciót hozhatunk létre kontur predikció, adaptív Intra/Interframe predikció, vagy

mozgáskompenzált predikció alkalmazásával. Ebben a fejezetben az adaptív Intra/Interframe predikcióval foglalkozunk, amely különösen hatékony bizonyos valós képek kódolására is.

Minden, a kódolóba érkező kép felbontható egy álló háttérre és egy változást tartalmazó területre. Nyilvánvaló, hogy az álló háttér esetében a legjobb megoldást az Interframe predikció adja, amelyben az előző félkép ugyanazon pozíciójú képpontját használjuk fel a becsléstől függően. A változó területben kisebb predikciós hibát eredményez az Intraframe predikció alkalmazása.

A predikció adaptív vá tehető aktivitás függvények segítségével. Definálható az Intraframe aktivitás függvény az (1) összefüggés szerint, míg az

$$AMD2 = \max |d_i, i+20|$$

le DN

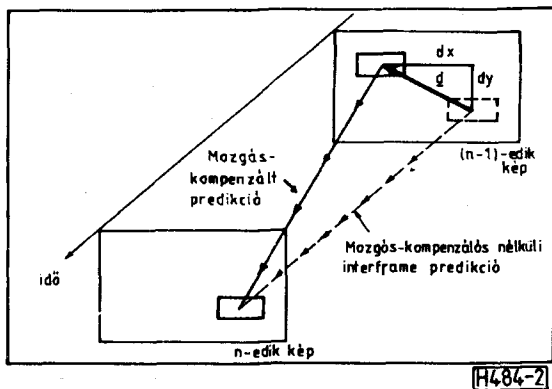
Interframe aktivitás függvény az alábbi szerint ahol  $i+20$  az előző félképre utal

A prediktorok kapcsolata az AMD és AMD2 aktivitás függvények közötti relációknak megfelelően történik. Az összehasonlításokban szereplő küszöbértékekre többféle konkrét adat található a megfelelő irodalmakban.

### 3. Mozgás-kompenzált prediktív kódolás

Nemstacionárius képanyag esetén az adaptív predikció nagyobb értékű adattömörítést képes elérni, ha figyelembe vesszük a mozgó testek kóptól-kópig (vagy félképtől-félképig) történő elmozdulását. Egy adott képszekvenciában lévő mozgás lehet a tárgyak elmozdulása, elfordulása vagy összetett mozgása a kamerához képest. A képpontok mozgás paramétereinek a képkódolásban (pl. DPCM-nél) történő hasznosítását mozgáskompenzálásnak hívják [1].

A mozgás-kompenzált predikció alapelve a következő [2]: az elmozdulás mező ismeretében végzünk becslést az aktuális kép/félkép pontjaira oly módon, hogy az előző kép/félkép mozgó képpontjait a mozgás paramétereitől függő megfelelő értékekkel eltoljuk (2. ábra).



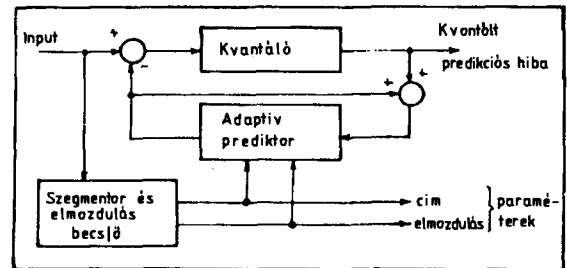
2. ábra Elmozdulás vektor

A kódolás 3 fő részből áll:

- 1) a kép/félkép elmozdulás mezejének becslése
- 2) mozgás-kompenzált predikció a becsléstől függően elmozdulás vektorok alapján
- 3) a predikciós hiba és a járulékos információk kódolása.

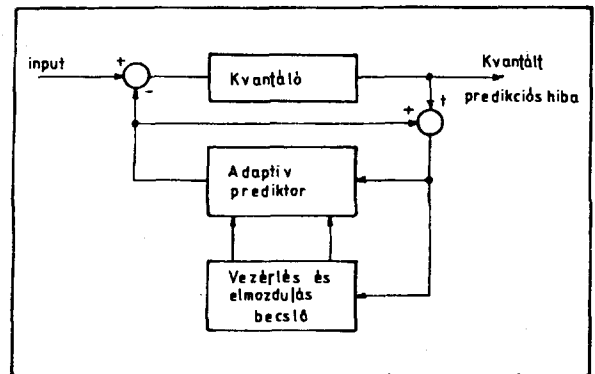
Nyilvánvaló, hogy a mozgás-kompenzált prediktív kódolás hatékonysága elsősorban a mozgás paramétereinek megfelelő becslésétől függ.

A mozgáskompenzált prediktív kódolási eljárások lényegében két csoportra oszthatók az elmozdulás vektort becslő algoritmusok alapján: blokk-illesztő és pel-rekurzív eljárásokra (3. és 4. ábra). Az egyszerűség kedvéért a továbbiakban



[H484-3]

3. ábra Blokk-illesztő mozgás-kompenzált kódoló



[H484-4]

4. ábra Pel-rekurzív mozgás-kompenzált kódoló

feltételezzük, hogy a képen belül csak egy test mozdul el és az egyenes vonalú elmozdulást végezt.

#### 3.1. Elmozdulás becslés

##### 3.1.1. Blokk-illesztő elmozdulás becslés

A blokk-illesztés az aktuális kép blokkjaihoz legjobban illeszkedő, előző képből, elmozdult blokkok keresésén alapul. Ennél az eljárásnál az első lépés, hogy a félképet blokkokra bontjuk. A blokkokra feltételezzük, hogy mindegyik képpontjuk azonos nagyságú és irányú elmozdulást végzett és így egy blokkra csak egy vektort kell meghatározni.

rozni. A blokk-illeszkedés valamilyen kritérium függvény - pl. kereszt-korrelációs függvény, négyzetes középhiba, stb. - szélső értékének keresése szerint történik. Az elmozdulás-vektort természetesen járulékosan át kell vinni (3. ábra).

A blokk-illesztő elmozdulás becslés két kritikus paramétere a blokk nagyság és a keresés lépésszáma. A lépésszám minimalizálására többféle praktikus keresési eljárást dolgoztak ki (2D-logaritmusos keresés [3], 3 lépéses keresés [3], stb.)

### 3.1.2. Pel-rekurzív elmozdulás becslés

#### 3.1.2.1. A becslés alapelve

Az alap algoritmust elsőként Netravall és Robbins publikálta 10 évvel ezelőtt [4], amelynek lényege a következő: ha egy képpont intenzitása megváltozik az egymásutáni képek/félképek között, akkor ezt a képpontot a kép mozgó területébe tartozónak tekintjük és keressük az adott  $f(x,y,t)$  intenzitás értékkel ekvivalenst az előző félkép elmozdult pozíciójában.

Legyen  $\underline{d} = (dx, dy)^T$  a pel elmozdulás vektora, ahol "T" a transzponáltat jelöli. Ha egy merev test tisztán egyenesvonalú elmozdulást végez, akkor a  $[t-\tau, t]$  intervallumra felírható:

$$f(x,y,t) = f(x-dx, y-dy, t-\tau) \quad (4)$$

ahol  $f(x,y,t)$  az aktuális pel intenzitása és  $f(x-dx, y-dy, t-\tau)$  az elmozdult képpont intenzitása az előző félképben. Defináljuk az elmozdult képkülönbség függvényt

$$dfd(x,y,dx,dy) = f(x,y,t) - f(x-dx, y-dy, t-\tau) \quad (5)$$

ahol  $\hat{dx}, \hat{dy}$  a valódi elmozdulás-vektor komponenseinek becsült értékei. Legyen  $d^{l-1} = (dx^{l-1}, dy^{l-1})$  a becslés kezdeti értéke, akkor

$$dfd(x,y,dx^{l-1}, dy^{l-1}) = f(x,y,t) - f(x-dx^{l-1}, y-dy^{l-1}, t-\tau) \quad (6)$$

Az  $f(x,y,t)$  helyébe a (4)-et írva

$$dfd(x,y,dx^{l-1}, dy^{l-1}) = f(x-dx, y-dy, t-\tau) - f(x-dx^{l-1}, y-dy^{l-1}, t-\tau) \quad (7)$$

A  $dfd$  függvényt az  $f(x-dx, y-dy, t-\tau)$  függvény  $(x-dx^{l-1}, y-dy^{l-1})$  helyen vett Taylor-sorával linearizáljuk, akkor

$$f(x-dx, y-dy, t-\tau) = f(x-dx^{l-1}, y-dy^{l-1}, t-\tau) - (d-d^{l-1}) \nabla f(x-dx^{l-1}, y-dy^{l-1}, t-\tau) + v(x,y,d^{l-1}) \quad (8)$$

ahol a  $\nabla f(\cdot)$  a 2D gradlens operator és a  $v(x,y,d^{l-1})$  a linearizálásból származó maradéktag. A (8)-t a (7)-be helyettesítve és átalakítva kapjuk a

$$dfd(x,y,dx^{l-1}, dy^{l-1}) = -\nabla^T f(x-dx^{l-1}, y-dy^{l-1}, t-\tau)(d-d^{l-1}) + v(x,y,d^{l-1}) \quad (9)$$

A  $dfd$  függvény, mint kritérium felhasználásával a pel-rekurzív algoritmus a  $d$  elmozdulás-vektor rekurzív becslésnek korrekciós tényezőjére a lehető legjobb becslést adja.

A pel-rekurzív algoritmusoknál az egyik alapvető kérdés a predikciós hiba minimalizálása mellett az, hogy az elmozdulás-vektor milyen gyorsan konvergál a valódi elmozdulás-vektorhoz. A többféle létező algoritmus közül a Wiener-szűrésen alapuló eljárás mutatja a legkedvezőbb eredményt, a továbbiakban ezt részletezzük.

#### 3.1.2.2. Wiener-szűrésen alapuló elmozdulás becslés

Blemond és társai [5] ajánlata szerint a (9)-ben a  $v(x,y,d^{l-1})$  zajtényezőt és a  $(d-d^{l-1})$  korrekciós tagot sztohasztikus folyamat mintának tekinthetjük. A cél az, hogy a "legkisebb négyzetek" módszerével a legjobb becslést adjuk a  $(d-d^{l-1})$  tagra. Ennek érdekében tekintsük a (9) szerinti  $dfd(\cdot)$  függvényt egy megfigyelésnek és a gradlens pedig egy ismert átmeneti vektornak. A becslési problémát meg lehet oldani úgy is, hogy csak az aktuális pel információt használjuk. Ha az aktuális pel egy előre definiált környezetét is figyelembe vesszük, akkor sokkal hatékonyabb becslést végezhetünk.

A Wiener-szűrésen alapuló becslési eljárás matematikailag leírása  $N$  képpontból  $(x(j), y(j))$ ,  $j=1,2,\dots,N$  származó megfigyelés alapján (9) szerint a következő:

$$\begin{aligned} dfd[x(1), y(1), d^{l-1}] &= -\nabla^T f[x(1) - dx^{l-1}, y(1) - dy^{l-1}, t-\tau] \\ & \quad (d-d^{l-1}) + v[x(1), y(1), d^{l-1}] \\ & \quad \dots \\ dfd[x(N), y(N), d^{l-1}] &= -\nabla^T f[x(N) - dx^{l-1}, y(N) - dy^{l-1}, t-\tau] \\ & \quad (d-d^{l-1}) + v[x(N), y(N), d^{l-1}] \end{aligned} \quad (10)$$

Vezessük be a gradlens vektorra a következő jelölést:

$$\nabla f[x(j) - dx^{l-1}, y(j) - dy^{l-1}, t-\tau] = [g_x(j), g_y(j)]^T \quad (11)$$

Ekkor felírhatjuk (10)-t mátrix alakban

$$\begin{bmatrix} dfd[x(1), y(1), d^{l-1}] \\ \vdots \\ dfd[x(N), y(N), d^{l-1}] \end{bmatrix} = \begin{bmatrix} g_x(1) & g_y(1) \\ \vdots & \vdots \\ g_x(N) & g_y(N) \end{bmatrix} [d - d^{l-1}] + \begin{bmatrix} v[x(1), y(1), d^{l-1}] \\ \vdots \\ v[x(N), y(N), d^{l-1}] \end{bmatrix} \quad (12)$$

\* pel = picture element

A (12) rövidebb formában

$$z = G(d - d^{l-1}) + v = Gu + v \quad (13)$$

ahol  $u$  a kezdeti becslés korrekciója. A feladat az, hogy az  $u$ -ra kell becslést ( $\hat{u}$ ) adni a "legkisebb négyzetek" módszerével, ha  $z$  és  $G$  adott.

Ahogy már az előzőekben feltételeztük, hogy a hibátag és a korrekciós tényező sztochasztikus folyamat mintái, a becslési probléma így végeredményben egy  $L$  lineáris becslő operátor keresésére vezethető vissza:

$$\hat{u} = Lz \quad (14)$$

úgy hogy

$$E\{\|u - \hat{u}\|^2\} \quad (15)$$

minimális legyen. A (15)-ben  $E\{\cdot\}$  a várható értéket,  $\|\cdot\|$  a vektor normáját jelöli.

Tegyük fel, hogy  $u$  és  $v$  ortogonálisak, akkor az ún. Wiener-szűrő alkalmazásával rendezés után a (14) kifejtve a következő:

$$\hat{u} = [G^T \cdot R_v^{-1} \cdot G + R_u^{-1}]^{-1} G^T R_v^{-1} \cdot z \quad (16)$$

ahol  $R_u$  és  $R_v$  az  $u$  és  $v$  kovariancia mátrixai.

A fentiek alapján az elmozdulás vektor egy  $\hat{u}$  becslése (13) és (16) szerint

$$d^l = d^{l-1} + \hat{u} \quad (17)$$

A Wiener-szűrően alapuló elmozdulás becslést néhány egyszerűsítő feltétel bevezetésével alkalmazhatjuk a prediktív képkódolásra. A mozgó területbe tartozó képpont  $d$  vektorának becsléséhez a szomszédos képpontok megválasztásakor lényeges szempont, hogy csak olyan képpontokban végezzük a megfigyelést, amelyek a vevőben is rendelkezésre állnak. Néhány jellegzetes konfigurációt mutat be az 5. ábra a becslésnél figyelembe vett képpontokra:

X X	X X X	X X X X	X X X X X
X O	X O	X X O	X X X O
N=4	N=5	N=7	N=9

5. ábra A becslésnél figyelembe vett képpontok

A (16) összefüggés használatához szükséges a maradéktag és a korrekciós tényező kovariancia mátrixának ismerete. Mivel ezekről a mátrixokról elég kevés ismerettel rendelkezünk tételezzük fel, hogy  $u$  és  $v$  komponensei 0 várható értékűek és függetlenek, így

$$R_u = \sigma_u^2 \cdot I(2) \quad (18.a)$$

$$R_v = \sigma_v^2 \cdot I(N) \quad (18.b)$$

ahol  $\sigma_u^2$  és  $\sigma_v^2$  az  $u$  és  $v$  szórásnégyzetek az  $I(2)$  és  $I(N)$  pedig  $2 \times 2$  illetve  $N \times N$  méretű egységmátrixok. Felhasználva a fenti egyszerűsítő feltételeket és bevezetve a

$$\mu = \sigma_v^2 / \sigma_u^2 \quad (19)$$

tényezőt, amely a maradéktag és a korrekciós tényező szórásnégyzeteinek hányadosa, a (17) becslési algoritmus a következő:

$$d^l = d^{l-1} - \begin{bmatrix} \sum_{j=1}^N g_x^2(j) + \mu & \sum_{j=1}^N g_x(j) g_y(j) \\ \sum_{j=1}^N g_x(j) g_y(j) & \sum_{j=1}^N g_y^2(j) + \mu \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_{j=1}^N g_x(j) \cdot dfd[x(j), y(j), d^{l-1}] \\ \sum_{j=1}^N g_y(j) \cdot dfd[x(j), y(j), d^{l-1}] \end{bmatrix} \quad (20)$$

Érdemes megjegyezni, hogy a többi - az előzőekben kidolgozott - pel-rekurzív elmozdulás becslés a Wiener-szűrően alapuló algoritmus speciális esetének tekinthető. Például Cafforio és Rocca algoritmusát [3] kapjuk vissza, ha  $N=1$  és  $\mu=100$ . Ha a  $\mu$ -t 0-nak vesszük és a korrekciós tényezőt  $\varepsilon = 1/2$ -el szorozzuk Walker és Rao által publikált módszerhez jutunk. A (20) egy egyszerűsített esete a Netravalli és Robbins gradiens módszerű becslése [4] is.

#### 4. Transzformációs kódolás

Az adattömörítő eljárások másik széles körben elterjedt módszere a transzformációs kódolás. A nagyobb számítási igénye és ennek következtében komplexebb felépítése ellenére egyre kiterjedtebben és egyre hatékonyabban alkalmazzák. Gyakorlati szempontok következtében a kódolandó képet kisebb blokkokra (leggyakoribb blokk méretek  $8 \times 8$  és  $16 \times 16$  képpont) osztva végezzük el a kódolást. Így egy bejövő  $M_1 \times M_2$  képpontból álló képtömböt  $N \times N$  méretű blokkra bontunk. Az általános lineáris kétdimenziós transzformáció definiálható, mint

$$F(u, v) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j, k) A(j, k; u, v); \quad u, v = 0, 1, \dots, N-1 \quad (21)$$

ahol  $A(\cdot)$  a négydimenziós transzformációs mag. A legtöbb képtranszformációs mag szeparálható horizontális (sorirányú) és vertikális (oszlopirányú) magokra, azaz

$$A(j, k; u, v) = A_c(j, u) A_r(k, v) \quad (22)$$

Így a szeparálható kétdimenziós transzformáció két lépésben, két egydimenziós transzformációval végezhető el. A kétdimenziós szeparálható transzformáció felírható mátrix alakban is

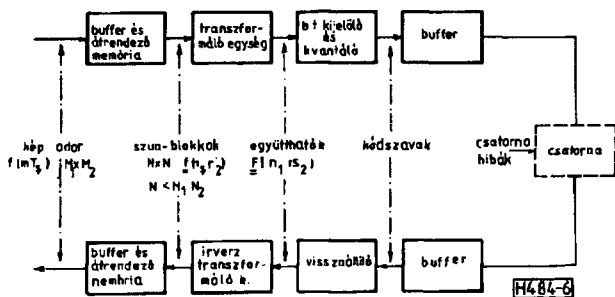
$$F = A_c \cdot f \cdot A_r \quad (23)$$

illetve két egydimenziós transzformációként

$$X = A x \quad (24)$$

ahol  $X$  és  $x$   $1 \times N$  méretű vektorok és  $A$  az  $N \times N$  méretű transzformációs mátrix.

A transzformációs kódoló általános blokkvázlatát a 6. ábra mutatja. Az adott típusú transzformáció elvégzése után a transzformációs együttható-



6. ábra Transzformációs kódoló

kat az eloszlási modelljük alapján meghatározott bit-térkép szerint kvantáljuk és a kvantált együtthatókat visszük át. A vevőoldalon az inverz műveletek fordított sorrendben követik egymást.

Mozgás-kompensált eljárás alkalmazása esetén természetesen a rendszer blokkvázlata kibővül a prediktív változathál már bemutatott elemekkel (képmemória, elmozdulás, becslő stb.).

A különféle lehetséges transzformációk közül itt most csak a diszkrét cos transzformációval (DCT) foglalkozunk. A DCT az egyik legelterjedtebben alkalmazott eljárás, mivel a legjobban közelíti az ideálisnak tekintett Karhunen-Loeve transzformációt. A 2D-DCT eljárás definíciója:

$$F(u,v) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j,k) \cos \frac{\pi u(2j+1)}{2N} \cos \frac{\pi v(2k+1)}{2N}$$

$$j, k = 0, 1, \dots, N-1; \quad u, v = 0, 1, \dots, N-1 \quad (25)$$

illetve 1D esetben

$$F(u) = \frac{1}{N} \sum_{j=0}^{N-1} f(j) \cos \frac{\pi u(2j+1)}{2N}; \quad u = 0, 1, \dots, N-1$$

$$j = 0, 1, \dots, N-1 \quad (26)$$

Hasonló módon adható meg az inverz transzformáció definíciója is.

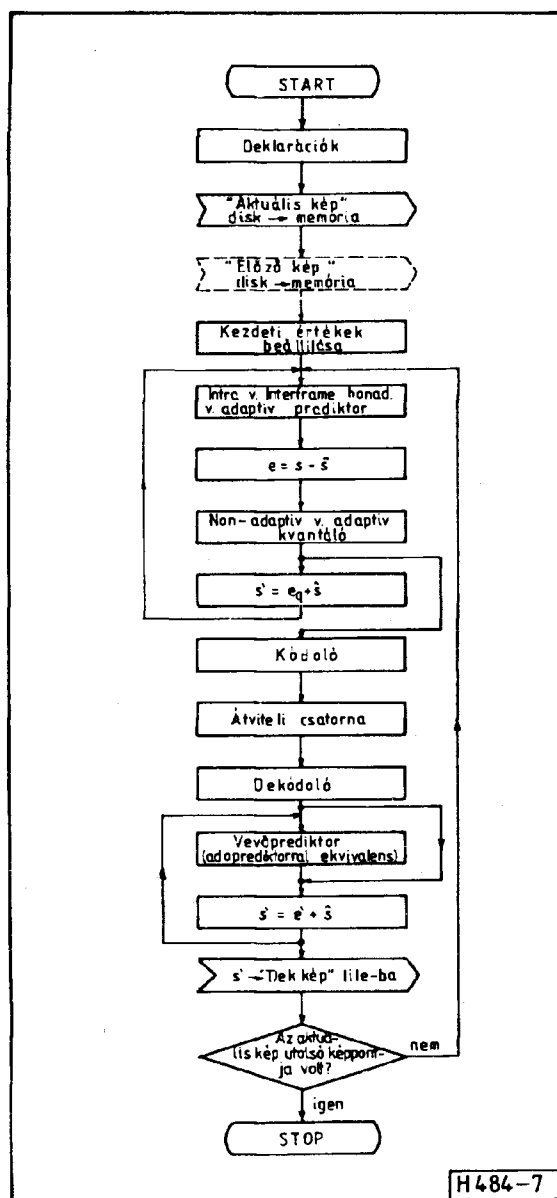
Az \$F(u,v)\$ transzformációs együtthatók eloszlása alapján határozható meg, a kvantáláshoz szükséges bit szám. A \$dc(F(0,0))\$ együttható eloszlásának modellezésére a Rayleigh valószínűség sűrűség, míg a többi együtthatóéra (\$F(u,v)\$; \$u, v = 0, 1, \dots, N-1\$ kivéve az \$u=v=0\$ esetet) a Gaussi valószínűség sűrűség alkalmazható.

A definíció alapján jól érzékelhető a meglehetősen nagy művelet szám igény. Gyors algoritmus kialakításához vezet - mátrix alakban való felírás esetén - a megfelelő mátrix faktorizáció, továbbá - a bemenő adat inverz trigonometrikus függvényvel való felírása és trigonometrikus azonosságok segítségével végzett átalakítás alapján kapott definíciós egyenletekből kiinduló - összeadó tömbökkel és memória táblázatokkal való kódoló felépítés. Mindkét esetben a valós idejű realizálás csak VLSI elemekkel hatékony. Ez lehet berendezés orientált VLSI chip, gate array-ekkel megvalósított semicustom áramkör, vagy DSP (digitális jel-

feldolgozó processzor) chipekkel felépített egység. Természetesen a szükséges szimulációk erősen függenek a megvalósítás tervezett típusától. Például DSP-kkel történő realizálás esetén jelentős, software úton végzett modellezésre, fejlesztésre is szükség van, túl magának a transzformációs eljárásnak a szimulációs vizsgálatán.

## 5. Prediktív képkódoló eljárások szimulációja

A 2. fejezetben ismertetett prediktív kópkódoló eljárások szimulációja több szempontból is lényeges. Többek között lehetővé teszi az egyes algoritmusok paramétereinek meghatározását, az adott képanyagra az optimálist közelítő eljárás ki-



7. ábra PCM kódoló folyamatábrája

választását, a real-time reallizálás feltételeinek meghatározását, a csatorna hatásának figyelembevételét, stb.

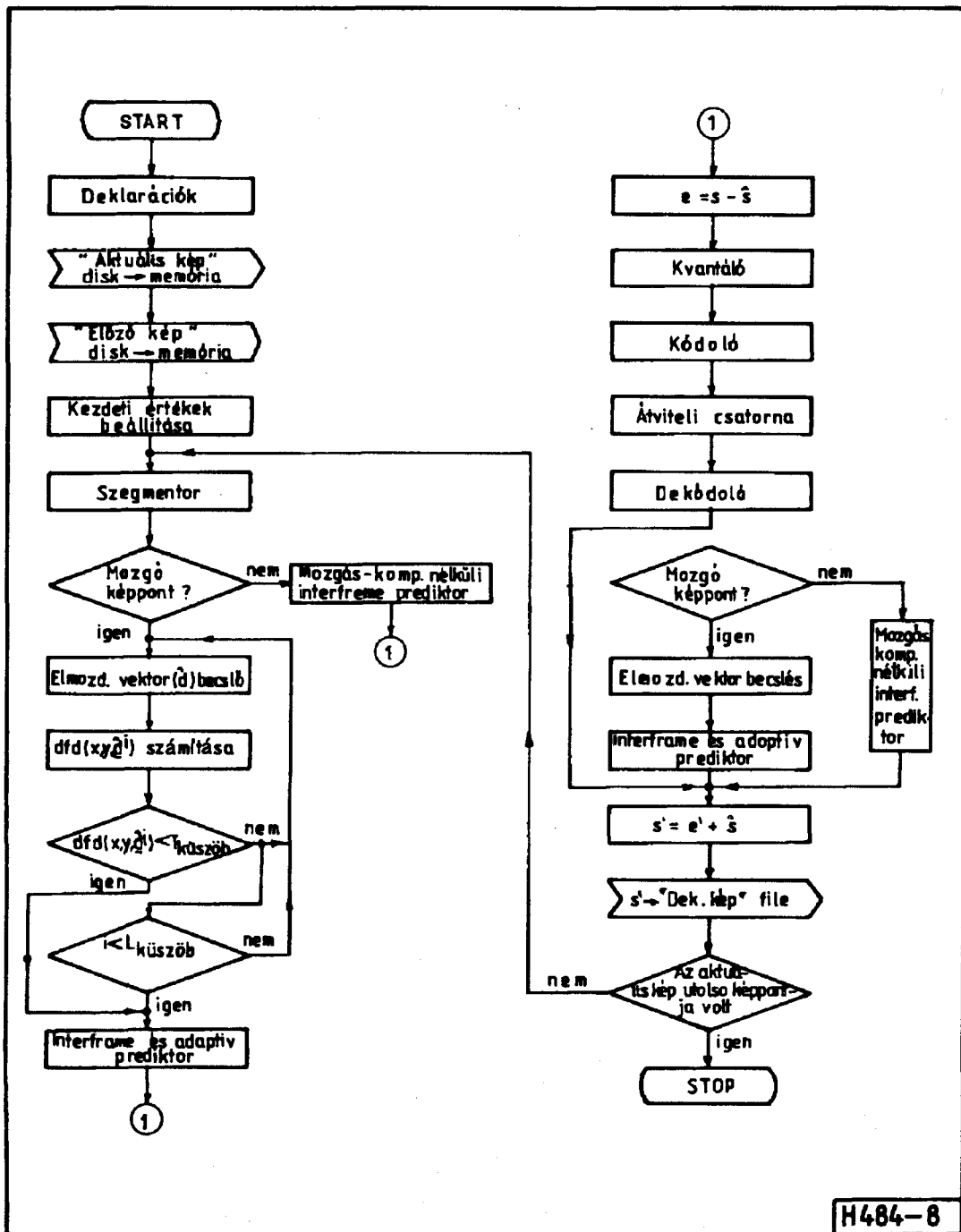
Az előzőekben vázolt célokra egy szimulációs programcsomag kifejlesztésében dolgozunk, amely két fő csoportból áll:

- mozgás-kompenzáció nélküli DPCM kódolás
- mozgás-kompenzált prediktív kódolás

A programok C nyelven íródtak, amely nagyon sok előnnyel rendelkezik a vázolt szimulációk esetében.

### 5.1 Mozgáskompenzáció nélküli DPCM kódolás

A programcsomag ezen része tartalmazza Intraframe és Interframe DPCM kódoló-dekódoló egységek szimulációját, amelynek nem adaptív és adaptív (adaptív kvantálás és/vagy adaptív predikció) változatokból épülnek fel. A programok felépítését a 7. ábrán lévő folyamatdiagram illusztrálja.



8. ábra Wiener-szűrésen alapuló mozgás-kompenzált kódoló folyamatábrája

Megjegyezzük, hogy ezek a programok jó referenciaként szolgálnak a mozgás-kompenzált predektív eljárások vizsgálatánál.

## 5.2. Mozcás-kompenzált predektív kódolás

Ahogy már az elméleti összefoglalóban is említettük a kópszekvenclákban a testek mozgása 3-féle lehet: transláció, rotáció vagy összetett mozgás. E szerint a programcsomag mozgás-kompenzált interframe kódolással foglalkozó része is 3 egységre osztható. Eddig elkészültek az egy képen belül csak translációs mozgást végző testet tartalmazó képanyagra használható pel-rekurzív eljárások szimulációs programjai. A Wiener-szűrősen alapuló elmozdulás becslést használó predektív kodek szimulációs programjának folyamatábráját a 8. ábra mutatja.

Ez a program többek között lehetővé teszi, hogy az eljárás elméleti vizsgálatában felmerülő kérdésekre is választ kapjunk, például a  $d$  kezdeti értékének a megválasztására, a gradiens és a  $dfd(.)$  függvény számításának lehetséges alternatívái közül a legmegfelelőbb alkalmazására, a translációs mozgás nagyságának hatására a dekódolt kép minőségére, stb. Ehhez az egységhez tartozik a speciális eseteknek tekinthető algoritmusok, a Walker-Rao, a Cafforlo-Rocca, és a gradiens módszer eljárások szimulációs programja.

A másik két fajta mozgás típusra a mozgás paramétereinek meghatározása lényegesen bonyolultabb, ezek meghaladják a cikk keretét.

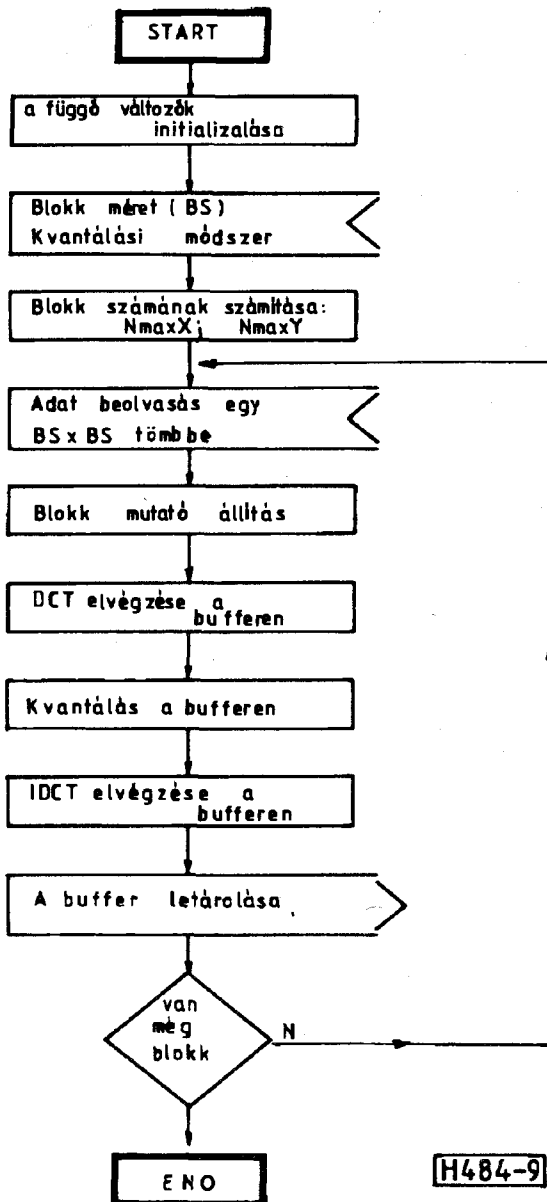
## 6. Transzformációs kódoló eljárások szimulációja

Az igen hatékony, de meglehetősen bonyolult, nagy mennyiségű számítási műveletet igénylő transzformációs kódoló rendszer jellemzőinek a meghatározása számítógépes szimulációkkal történik. A legfontosabbakat említve a teljesség igénye nélkül

- a blokk nagyság meghatározása ( $N$  értékének a megválasztása),
- a transzformációs együtthatók eloszlásának vizsgálata,
- bit-kijelölés (bit-térkép)
- kvantálási karakterisztika,
- csatorna hibák hatásának analízise
- képminőség meghatározása (négyzetes hiba illetve súlyozott négyzetes hiba számítása),
- mozgás-kompenzált esetben az elmozdulások paramétereinek számítása,
- a különböző közbenső fokozatok számítási pontosságának analízise,
- a várható műveletidők becslése.

Szemléltetésül a DCT kódolás folyamat diagramját adtuk meg a 9. ábrán a legegyszerűbb változatot feltételezve.

Külön kiegészítő megfontolások szükségesek a DSP-k alkalmazása esetén:



9. ábra 2D-DCT folyamatábrája

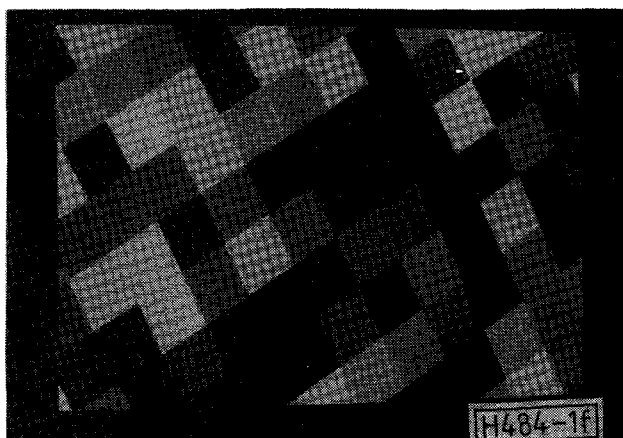
- A DSP típusának megválasztása. Itt egymástól független, különféle szempontok érvényesülhetnek
- a DSP fix szószélessége (bitszélessége) és a blokk méret egyeztetése, a fix bitszélességből adódó számítási korlátok figyelembevétele
- a software fejlesztés követelményeinek figyelembevétele
- a műveleti idők becslése és összevetése a DSP jellemzőivel.

## 7. Néhány szimulációs eredmény

A szimulációs programcsomag tagjai az előzőekben említett kódolási eljárások részletes analízisét teszik lehetővé. A konkrét vizsgálatokhoz bemenőjelként a sztohasztikus képmódel [8] alap-

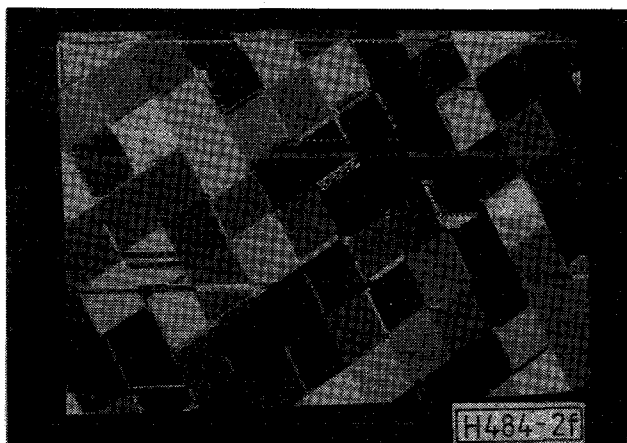


ján ugyancsak a számítógéppel generált jelet (a képméret: 199 sor x 319 képpont) használjuk első lépésként (1. foto). Ez a mesterséges kép megfele-



1. foto Bemelő mesterséges kép

lő élstruktúrával rendelkezik ahhoz, hogy a kódolási eljárások szempontjából kritikus jelenségeket vizsgálni tudjuk. Mozgás szimulációjához a memória tartalom eltolását használtuk, azaz két egymástáni kép egy képsorozatban elmozdulás esetén a tároláskor a memóriában egymáshoz képest eltolva jelenik meg. A 2. foto Interframe



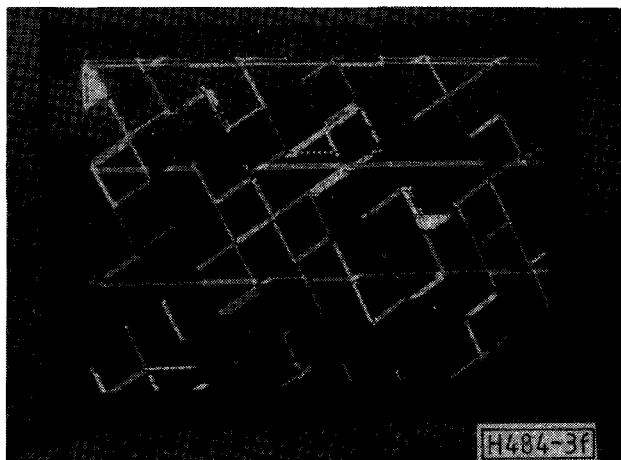
2. foto Interframe DPCM kodek kimenő képe

DPCM-kodek dekódolt képét mutatja, amikor a két kép között 3-sornyi függőleges irányú elmozdulást realizálhatunk. Az Interframe eljárásban fix predikciót valósíthatunk meg

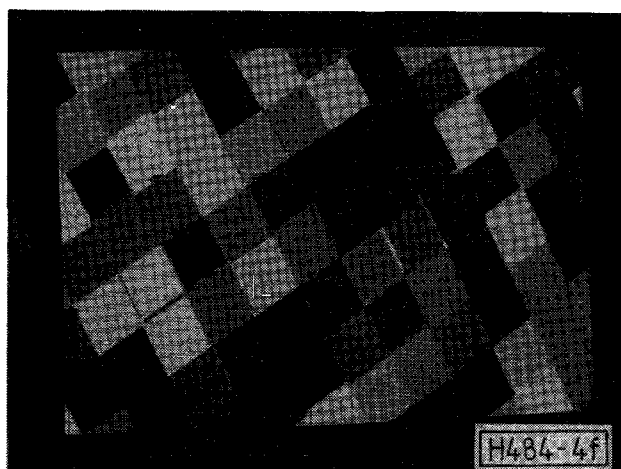
$$S_{i,j,t} = a_1 S_{i,j-1,t} + a_2 S_{i-1,j,t} + a_{20} S_{i,j,t-1}$$

ahol  $i$  a sorszám,  $j$  a kóppontszám,  $t$  az idő és  $a_1 = 0,25$ ,  $a_2 = 0,125$ ,  $a_{20} = 0,625$ . A kvantálás adaptív. Jól látható, hogy ez az eljárás mozgásra elég érzékeny, sok hiba keletkezik. A 3. foto mutatja a bemelő és dekódolt kép különbségét.

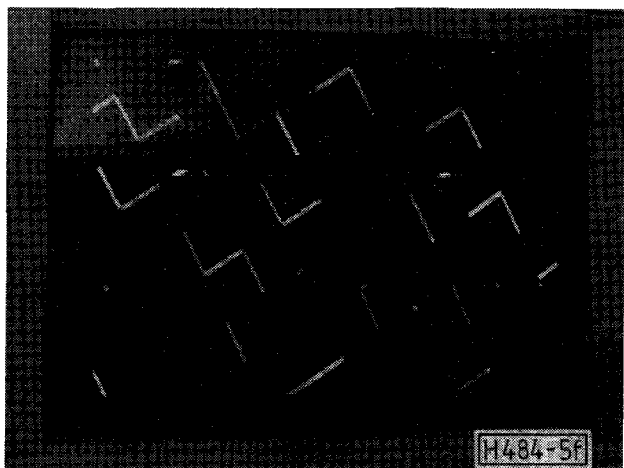
A 4. foto Wiener-szűrésen alapuló pel-rekurzív mozgás-kompenzált predektív kodek kimenő képe, még ebben az esetben a különbségi képet az 5. foto mutatja. Jól érzékelhető a javulás az előző



3. foto Különbségi kép Interframe DPCM esetén

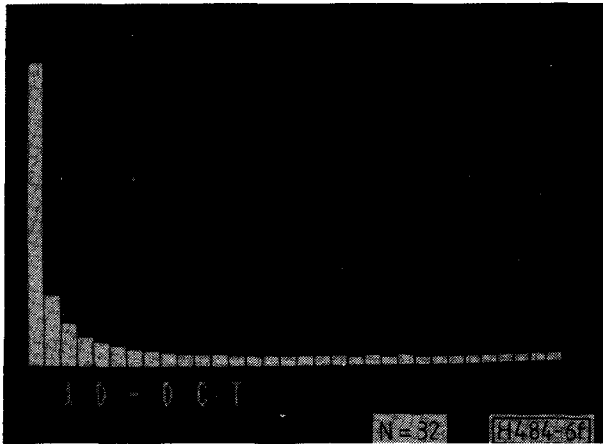


4. foto Wiener-szűrésen alapuló pel-rekurzív mozgás-kompenzált kodek kimenő képe

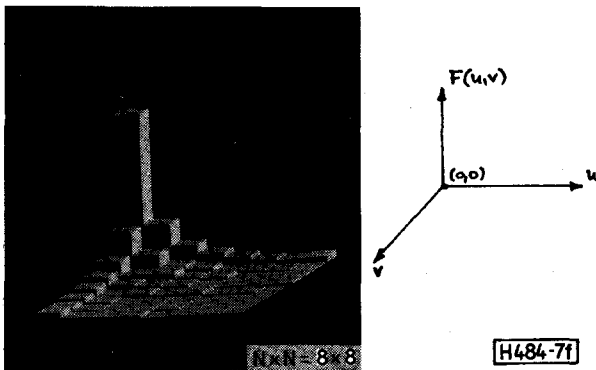


5. foto Különbségi kép a pel-rekurzív mozgás-kompenzált kodek esetén

módszerhez képest. Normál felbontás (nagyobb sorszám és nagyobb pontszám) és természetes képek esetén ezek a hibák alig észrevehetőek. Itt az erős geometrikus struktúra miatt a hibák szubjektíven erősebben érzékelhetők.



6. foto Transzformációs együtthatók eloszlása 1D-DCT esetén



7. foto Transzformációs együtthatók eloszlása 2D-DCT esetén blokkméret 8x8.

A transzformációs módszerek közül a DCT együtthatók eloszlásának szimulációjáról adunk még két fotót (6. és 7.). A módszerek kiterjedtsége e cikk keretében nem teszi lehetővé a további részletes szemléltetést.

## 8. Következtetések

A szimulációs programcsomag flexbilis, hatékony lehetőséget biztosít a különféle predektív és transzformációs kódolási eljárások analízisére. A szimulációs eljárás a kodek méretezések meghatározó alapeleme, enélkül optimális vagy legalábbis az optimálist jól közelítő rendszer kialakítása lehetetlen. A konkrét méretezési igények ki-elégítésén túl, az elméleti vizsgálatoknak is fontos eszköze a szimuláció.

A szimulációs programcsomagot jól egészítik ki a sztohasztikus képmoell alapján készült különféle paraméterű mesterséges képeket generáló programok, valamint a képek statisztikai paraméterrelt számító programok.

## IRODALOM

- [1] Böröczky Lilla; "Mozgás-kompensált kódoló tervezése színes videojelre", Diplomater, BME Mikrohullámú Hírad. Tsz., 1987. jun.
- [2] Lilla Böröczky, "Theory of Motion Compensated Predictive Coding", Newsletter, Technical University of Budapest, Vol. 6. No. 1, 1988.
- [3] H. G. Musmann, P. Pirsch and H. J. Grallert, "Advances In Picture Coding", Proc. of the IEEE, Vol. 73, No. 4, April 1985, pp. 523-548.
- [4] A. N. Netravalli and J. D. Robbins, "Motion-Compensated Television Coding: Part I", BSTJ., Vol. 58, Mar. 1979, pp. 631-670.
- [5] J. Blemond, L. Looljenga, D. E. Boekee and R. H. J. M. Plompen, "A Pel-Recursive Wiener-Based Displacement Estimation Algorithm", Signal Processing, Vol. 13, No. 4, Dec. 1987, pp. 399-412.
- [6] W. C. Wong and R. Stede, "Adaptive discrete cosine transformation of pictures using an energy distribution logarithmic model. The Radio and Electronic Engineer, Vol. 51, No 11/12. pp. 571-578 November/December 1981.
- [7] Fazekas Kálmán, Böröczky Lilla, "Mozgáskompensált kódolás realizálási kérdése", Interkozmosz Tudományos Szeminárium Kiadványa (megjelenés alatt)
- [8] Fazekas Kálmán, Balázs Péter, "Képjelek modellezése", Interkozmosz Tudományos Szeminárium Kiadványa (megjelenés alatt)
- [9] Balázs Péter, "DCT kódoló realizálása DSP-kkel", Diplomater, BME Mikrohullámú Híradástechnika tanszék 1987, jun.

Lapunk példányonként megvásárolható:

az V., Váci utca 10. és

az V., Bajcsy-Zsilinszky út 76. szám alatti  
hírlapboltban