

Új lehetőségek a digitális televíziózásban

ENYEDI BALÁZS, KONYHA LAJOS, SZOMBATHY CSABA, DR. TRAN MINH SON,
DR. FAZEKAS KÁLMÁN

*Budapesti Műszaki és Gazdaságtudományi Egyetem
Szélessávú Hírközlés és Villamosságtan Tanszék, Médiatechnológia Laboratórium
{szombathy, enyedi}@mht.bme.hu*

Kulcsszavak: DVB, MPEG4, BIFS, interaktivitás, wavelet transzformáció, SPIHT

A kilencvenes években megalkotott, napjaink egyik legszélesebb körben használt képtömörítési eljárása, az MPEG2 szabvány adja a DVB-technika alapsávi jelátvitelének hátterét. A jelenlegi multimédia alkalmazások által támasztott követelmények azonban meghaladják az MPEG2 rendszerek által felkínált lehetőségeket. Világszerte ezen igények a mozgatóerői az új MPEG4 szabvány fejlesztésének. Az MPEG4 videótömörítési szabvány két legfontosabb előnye az interaktivitás lehetősége valamint az MPEG2-nél hatékonyabb tömörítési algoritmusok használata [5,6]. A digitális televíziózás az MPEG2 szabványra épült, mely nem volt felkészítve az új lehetőségek támogatására, viszont a széleskörű elterjedése miatt nem cserélhető le. Erre a problémára megoldást jelent, ha sikerül beültetni az MPEG4 tartalmat a DVB rendszerbe. Célunk az MPEG4 szabvány fent említett előnyeinek bemutatása, valamint az MPEG4 tartalomnak az MPEG2 alapú digitális műsorszórásba való beültetésének vizsgálata.

Az MPEG4 szabvány a képkocka alapú video tömörítés mellett az objektum alapú kezelésmódot is megengedi, így hang és képi anyagon kívül egyéb információkat is tárolhatunk az adatfolyamok között. Ilyen például a bináris jelenetleíró adatfolyam (BIFS), mely esemény definíciókat, vezérlő információkat is tartalmaz, így segítségével interaktivitás valósítható meg.

Az új szabvány másik nagy előnyét, a hatékony tömörítést az újabb, fejlettebb kódolási algoritmusok használatának lehetősége jelenti. Ezek közé tartozik egy alacsony bitsebességű video tömörítési eljárás, mely 3 dimenziós wavelet transzformációt és SPIHT algoritmust alkalmaz [7]. Ebben az eljárásban a DCT helyett wavelet transzformációt használunk, a hagyományos különbségi kódoláson alapuló algoritmusokkal ellentétben az időbeli redundancia kihasználására is a wavelet transzformációt alkalmazzuk. Az együtttható összedesedésére az általunk módosított SPIHT algoritmus 3 dimenziós változatát használjuk fel.

1. Bevezetés

Az MPEG4 szabvány új dimenziókat nyit az MPEG átviteli kapacitás és az általános multimédia-tömörítési technikák terén. Az MPEG4 magasabb szinten nyújt szolgáltatásokat, integrálja a tartalmat, az interaktív funkciókat és ezek megjelenítését, mint elődei. Másképpen fogalmazva, ha már a multimédia-jelenet számos összetevőt – például hang, kép stb. – foglal magában, az adattömörítésen túlmenően feltétlenül szükséges egy olyan eszköz, amely elősegíti a bonyolult jelenetek megszerkesztését és a komponensek hatékony kezelését. Összetett multimédia-jelenet kialakításához a jelenetben lévő összetevők közötti, illetve a jelenet és a felhasználók közti interaktivitás is hangsúlyt kap.

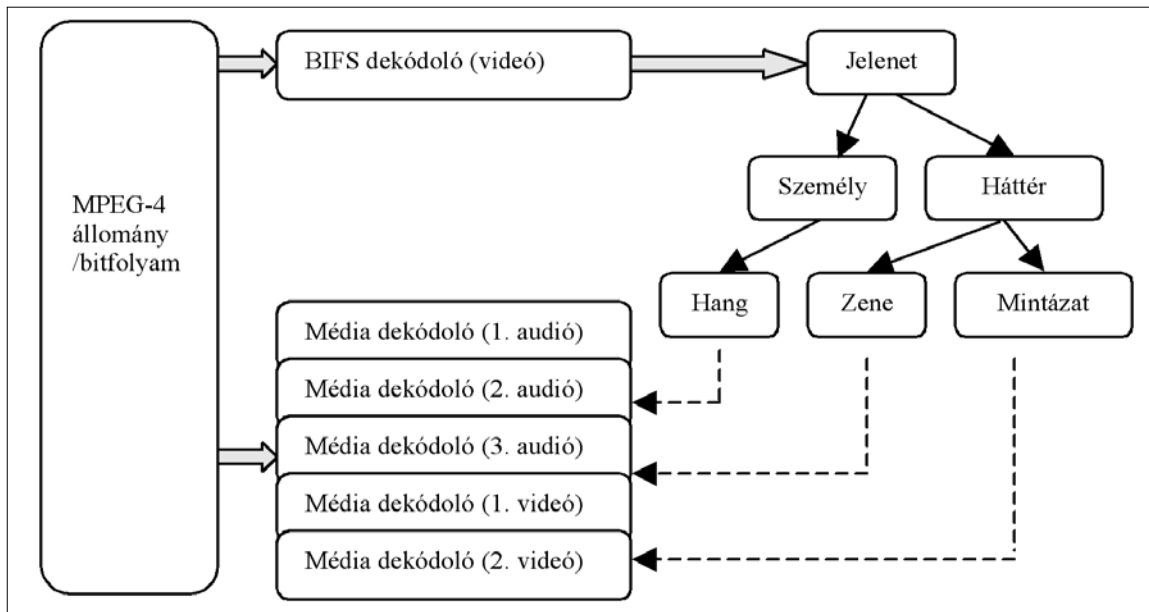
A nemzetközi MPEG4 szabvány elsőként támogatja ezt az új megoldást, nevezetesen multimédia-jelenet összeállítását és interaktivitását, egy újfajta adat – jelenet bináris leírása (BIFS) – bevezetésén keresztül. Sajnálatos módon az MPEG4 BIFS által felkínált lehetőségek még alig jelennek meg a pillanatnyi MPEG4 alapú alkalmazásokban, amelyek kizárólag az MPEG4 hatékony hang-kép tömörítési eljárásait használják ki.

1.1. Az MPEG4 BIFS struktúrája és funkciója

Az MPEG4 szabvány különböző hallható-látható objektumok kódolásával foglalkozik, például: természetes kép, hang, mintázat, 2D és 3D grafikák, szintetikus zene, hang stb. Egy bonyolultabb multimédia-jelenet rekonstruálásához nem elegendő kizárólag a tömörített hangok, képek és más összetevők továbbítása, ezért az úgynevezett BIFS-adatokat, a multimédia egy újfajta komponensét vezették be. Ezek funkciója, hogy a multimédia-jelenetet egy gráfnak megfelelő hierarchikus struktúrával írják le (1. ábra).

A gráf csomópontjai különböző objektumokat reprezentálnak, például hangot, mozgóképet, állóképet, grafikát, szöveget stb. Az MPEG4 filozófia alapján, a BIFS [4] minden lehetséges multimédia jelenetet egy hierarchikus struktúrával ír le, ahol jól látható, hogy egyszerre több hang- és képdekódoló működhet párhuzamosan. A gráf struktúrája nem feltétlenül statikus, a csomópontok változhatnak az idő múlásával és a néző beavatkozásának következtében is. Ez lehetőséget ad összetett MPEG4 alapú multimédia jelenet kialakítására is.

Minden csomópont bizonyos paraméterekkel rendelkezik, ezek értékeinek módosításával az adott objektum számos tulajdonsága beállítható. A csomópontok három nagy csoportba sorolhatók:



1. ábra
Multimédia-jelenet hierarchikus struktúrája

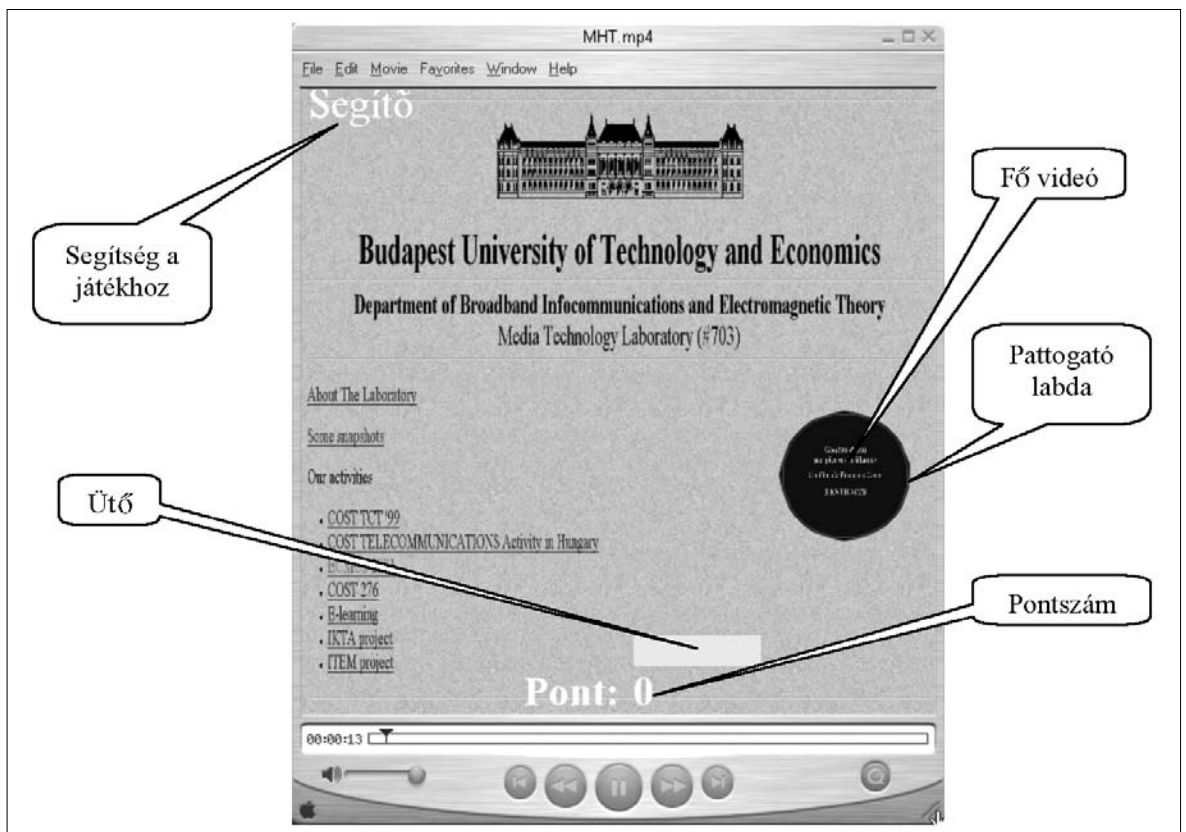
- média csomópontok, amelyek tömörített adatfolyamokon (hang, videó, állókép stb.) alapulnak;
- érzékelő csomópontok, amelyek bizonyos tulajdonsága változik külső behatásra (például nyomógomb, egérmozgás stb.);
- Script csomópont, melyek felépítését illetve viselkedését teljes mértékben programozhatjuk; a Script és az érzékelő csomópontok, valamint ezek összekötési mechanizmusa az interaktivitás alapja.

A BIFS által felkínált lehetőségek kihasználása érdekében először egy BIFS-szerkesztő struktúráját dol-

goztuk ki, majd ez alapján különböző bonyolultságú, interaktivitással rendelkező jeleneteket állítottunk össze. A 2. ábrán látható egy, az MPEG4 interaktivitásán alapuló játék (elsőként fejlesztettünk ki a BIFS lehetőségek kihasználásával MPEG4-es játékot).

2. MPEG4 alapú fejlesztési eredmények

Az általunk kifejlesztett, egyes jellemzőiben egyedien optimalizált MPEG4 alapú kodek az MPEG2 rendszerekben alkalmazott blokkalapú tömörítés helyett kép-



2. ábra
MPEG4 BIFS alapú multimédia játék

kocka alapú, úgynevezett wavelet transzformációt használ. A wavelet transzformáció [1,2,5,6] jelenleg a leghatékonyabban alkalmazható eljárás a képfeldolgozásban.

Előnye a korábbi algoritmusokhoz képest (pl. DCT), hogy a jelnek a térbeli és a frekvenciatartománybeli viselkedését egyszerre írja le, ráadásul mind a térben, mind a frekvenciatartományban jól igazodik az emberi látórendszer (HVS) tulajdonságaihoz; így módon kinyerhetjük a képből az emberi szem számára fontos információkat, a lényegteleneket pedig elhanyagolhatjuk. Ezek mellett a transzformáció számításiigénye kisebb, mint ami a korábbi eljárások gyorsított algoritmusának végrehajtásához szükséges. E tény igen fontos a képfeldolgozásban, bonyolult algoritmussal ugyanis nem implementálható a valósidejű jelfeldolgozás.

A wavelet transzformáció [8,9] eredményeként létrejött együtthetők tulajdonságait nagymértékben befolyásolja a transzformáció bázisfüggvénye. Az elmúlt évtizedekben sok kutatás irányult különböző alkalmazások szempontjából optimális bázisfüggvények kifejlesztésére.

A mozgókép-tartalom továbbításakor a különbségi kódolás helyett háromdimenziós wavelet transzformációt és SPIHT [8] algoritmust alkalmaztunk. Az eljárás során a hagyományos mozgáskompenzációt használó algoritmusokkal ellentétben az időbeli redundancia kihasználására is a wavelet transzformációt alkalmazzuk. Az együtthető rendezésre az általunk módosított SPIHT algoritmus háromdimenziós [9] változatát használjuk.

A háromdimenziós jelen úgy végezzük el a wavelet transzformációt, hogy először minden képkocka minden során végrehajtjuk az első transzformációs lépést (mint 1D függvény), majd ugyanezt megteesszük minden képkocka minden oszlopával. Végül minden képkocka ugyanazon koordinátájában elhelyezkedő pontjain is elvégezzük a transzformáció első lépését (időtartománybeli). Ezek után hasonlóan folytatjuk a transzformációt a következő lépésekkel.

Az első két esetben a transzformációt a térbeli koordinátatengelyek mentén végezzük el, az utolsó esetben pedig az időtengely mentén. Ezekből következik, hogy a transzformáció elvégzéséhez meg kell várni, míg az összes képkocka megérkezik. Ehhez nagyon nagy memóriaterületre van szükség, élő közvetítésben pedig megvalósíthatatlan. Ennek kiküszöbölésére a képkockákból képcsoportokat (Group of Frames, GOF) formálunk, és a transzformációt egy-egy képcsoportra végezzük el, a képcsoportokat egymástól függetlenül kezelve. A képcsoportot alkotó képkockák száma határozza meg a rendszer késleltetését, túl nagy képcsoport esetén túl nagy lesz a késleltetés, míg túl kicsi képcsoport esetén nem lesz elég hatékony a tömörítés.

A tértartománybeli wavelet transzformációhoz a Daubechies 9/3 bázisokat, míg az időtartománybelihez a Haar bázisokat használtuk. A széleken szimmetrikus kiterjesztést alkalmaztunk. Színes képeknél a transzformációt külön-külön elvégezzük mind a három komponensre (Y, U, V). Természetesen az U és a V kompo-

nensek mérete mind vízszintes, mind függőleges irányban feleakkora, mint a világosságjel mérete.

A 3D wavelet transzformáció eredményeként létrejött együtthetőkát kvantáljuk, majd összegyűjtjük. A kvantálásra az SPIHT algoritmust használtuk. A SPIHT algoritmus a következő megfigyeléseken alapszik:

- A képminőségre legjelentősebb hatással a legnagyobb helyértékű bitek vannak, ezért ezeket kell először összegyűjteni, majd sorban haladva az egyre kisebb helyértékű bitek következnek.
- Az alacsonyabb frekvenciájú részsávokban található együtthetők fontosabbak a HVS szempontjából, így ezeket kell először összeszedni.
- Az együtthetőknek az értékét és a pozícióját is tárolni kell.

Az SPIHT algoritmus akkor ér véget, ha minden bitet sikerült tárolni, vagy a keletkezett bitsorozat hossza elérte a beállított tömörítéstől függő maximális értéket, így könnyen megvalósítható mind a konstans, mind a változó bitsebességű kódolás. Az SPIHT progresszív tulajdonsága miatt a bitsebességre való skálázhatóság egyszerűen megoldható, a wavelet transzformáció kedvező tulajdonságainak köszönhetően a térbeli skálázhatóság sem okoz gondot.

3. Az MPEG4 tartalom beillesztése MPEG2 bitfolyamba

Az MPEG4 tartalom MPEG2 bitfolyamba való beillesztése már körvonalazott, az eljárás tényleges bevezetésére meg nem került sor széles felhasználói körben. A magyarázat erre egyrészt az, hogy az MPEG4 még újdonságnak tekinthető, az általa felkínált lehetőségek még nincsenek igazán kihasználva, másrészt, a szabvány hatékony tömörítő eljárásait inkább az alacsony bitsebességű csatornában, például interneten alkalmazzák először. Az IP alapú MPEG4 továbbítást néhány internetes alkalmazásban már megvalósították, például az OpenLAN videó műsorszórási VLC rendszerben. A szélessávú műsorszórás terén, azaz kábeltévé-, műholdas- illetve földi csatornán az MPEG4 jelfolyamok beiktatása még a jövő feladata.

A kutatásunk egyik célja volt, hogy áthidalja az említett rést. Ennek megfelelően egy teljes konstrukciót dolgoztunk ki az MPEG4 alapú multimédia-tartalom DVB-T rendszerben való továbbítására MPEG2 bitfolyam (TS) segítségével [3]. A szabványban leírt irányelvek alapján így konkrét (akár ipari szinten is azonnal hasznosítható) alkalmazást valósítottunk meg.

Az MPEG2 jelfolyam felépítésének kidolgozásakor biztosították a jövőbeli továbbfejlesztés lehetőségét is, ezt használtuk ki az MPEG4 adatok beillesztésekor. Az MPEG2 szabványcsomag második kiadása alapján két megoldás van az MPEG4 adatok továbbítására:

1. bitfolyam alapú:

Az MPEG4 hang-, illetve képadatokat egyszerűen hozzáadjuk az MPEG2 jelfolyamhoz, mint közönséges MPEG2 média bitfolyamokat. A PMT táblán, a társított

műsor bejegyzésén egy új jelfolyam-típus értéket vezetünk be (e paraméter a média bitfolyam jellegét határozza meg), ily módon jelezve az MPEG4 adatok jelenlétét.

Ezentúl az *MPEG4_video_descriptor* és *MPEG-4_audio_descriptor* szintaxist is kidolgozták, amelyek jellemzőkként továbbítandók a kezelt média bitfolyammal együtt. Ezen adatstruktúra segítségével a dekódoló megfelelően kezeli az MPEG4 adatokat a szükséges információ (a kódolás profilja, szintje, pontos kódoló eszköze stb.) birtokában. Noha a bitfolyam alapú módszer lehetővé teszi olyan szolgáltatások megvalósítását, amelyekben MPEG4 tömörítési eljárásokat alkalmaznak az MPEG2 helyett, nem támogatja az MPEG4 magasszintű tulajdonságait, például az interaktivitásokat. Mindezek következtében a második megoldást (lásd lejjebb) választottuk, így a bitfolyam alapú megoldással a továbbiakban nem foglalkozunk.

2. jelenet alapú:

A videó jelenetben megszokott hang-, illetve képanyagon túlmenően ezzel a módszerrel más média-információt továbbíthatunk, például mintázatok, 2D és 3D grafikákat, szintetikus zenéket, hang effektusokat stb., ugyancsak a MPEG2 TS jelfolyamon keresztül. A különböző médiák mellett ez a módszer még biztosítja a bináris jelenet leíró (BIFS) információnak az átvitelét is.

Az MPEG2 TS jelfolyamba integrált, jelenet alapú továbbítás egy többszörös beillesztést takar. Az MPEG4 jelenet különböző összetevői (hangok, képek, grafikák stb., illetve BIFS információ) az MPEG2 program elemeinek feleltethetők meg. Az MPEG2 esetén kétszintű beillesztést ismertünk meg: az elemi bitfolyamokat először PES csomagokra bontjuk illetve szekció adatstruktúrában tároljuk el, majd mindezt TS-csomagokra bontjuk. Az MPEG4 esetén – még a PES-, illetve a szekcionálás előtt – egy sajátos összeillesztési technikát alkalmaznak (3. ábra), így a komponensekből több SL (Synchron Layer), illetve úgynevezett FlexMux bitfolyam alakul ki. A TS összeillesztése után az MPEG4

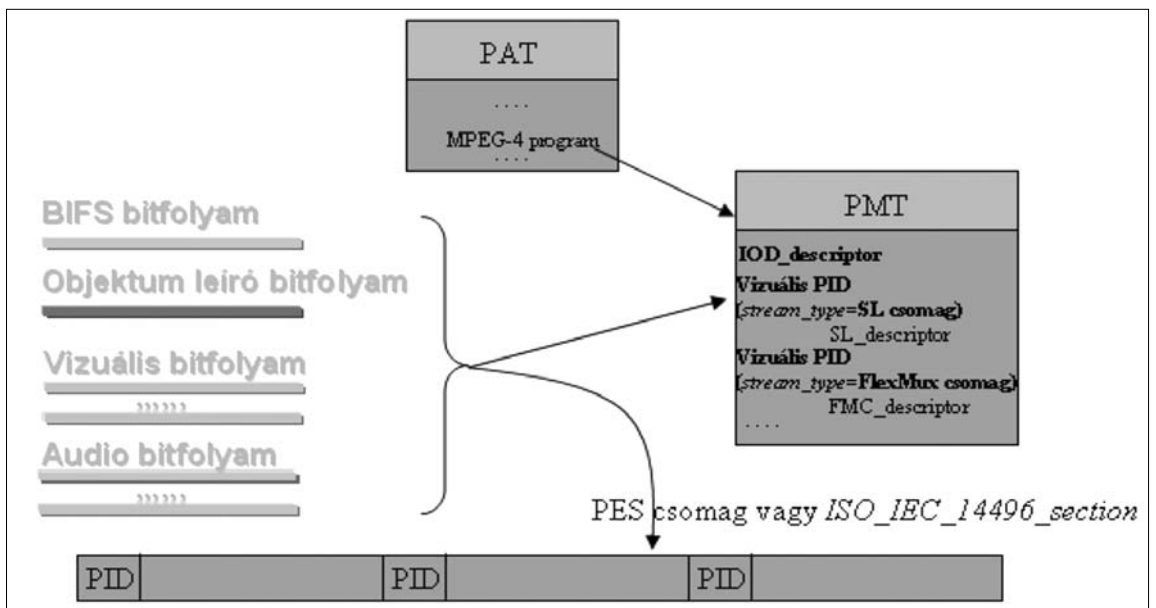
összetevők ugyanolyan bejegyzésként szerepelnek a PMT táblában, mint az MPEG2 esetén. A PMT-ben megtalálható leírók nem azonosítják a társított elemi bitfolyamok fajtáját, hanem jelezik a dekódolónak, hogy a kapcsolódó bitfolyamok legalacsonyabb „csomagolási” szintje miből (SL vagy FlexMux) áll. A dekódoló ezekből a csomagokból kiolvastva különíti el az MPEG4 értelmezési mechanizmusával a média-adatokat a BIFS illetve más rendszerinformációtól. A szétválasztásban fontos szerepet játszik az *IOD_descriptor* leíró, amely a PMT tábla része, és az MPEG4 program jellemzője. Ez a leíró különíti el a BIFS adatokat a többi elemi információtól, ezáltal a jelenet rekonstruálhatóvá válik. Ha hivatkozás van a jelenet leírásában a többi média adatra, akkor a dekódoló csak ezután használja fel őket.

Kutatásunk keretében egy teljes beillesztő konstrukciót dolgoztunk ki, amelynek segítségével az MPEG4 multimédia tartalmat a szabványos DVB-T rendszerben továbbíthatjuk. A 4. ábra az elemi egységeket és azok kapcsolatait szemlélteti.

A teljes rendszer integrálása sikerrel járt. Továbbítottunk és megjelenítettünk mind normál videó anyagot, mind a fentiekben leírt, BIFS szolgáltatáson alapuló MPEG4 alapú interaktív alkalmazást. Az ábrán látható rendszerben a DVB-T adót követően mindkét megoldás alkalmazható, ezek segítségével sikerült komplett valós idejű, MPEG4 alapú, DVB-T átviteltechnikába ágyazott digitális műsorszóró rendszert megvalósítanunk.

4. Összefoglalás

Noha napjaink földi digitális műsorszóró rendszerei elsősorban még az MPEG2 kódolásra épülnek, a jövőben egyértelműen az MPEG4 rendszerek előtérbe kerülése várható. Külön kiemelendő, hogy az MPEG4 rendszerek DVB-T technikába való sikeres beintegrálása ipari szinten is nagy jelentőséggel bír: amennyiben a szolgáltatók a jövőben úgy döntenek, hogy áttérnek



3. ábra Jelenet alapú beillesztés

a jelenlegi MPEG2 rendszerekről MPEG4-re, nem szükséges az igen drága adástechnikai eszközöket (jelfolyam-generátorokat, analizátorokat, modulátorokat stb.) kicserélni, minimális beruházással megoldható a váltás.

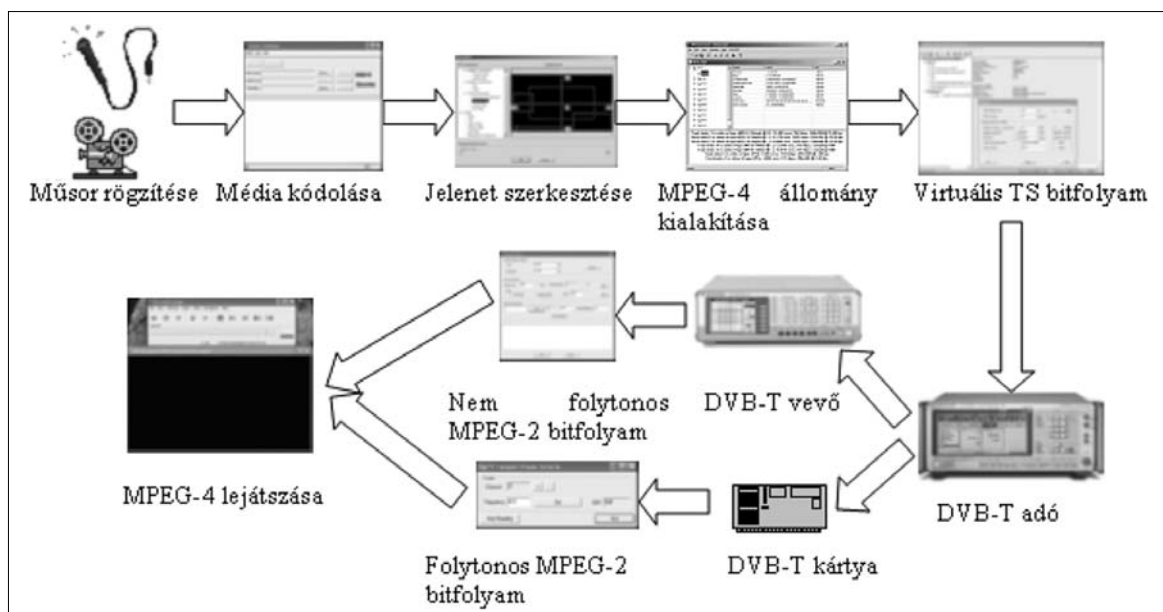
Az áttérés során a legnagyobb problémát a már piacon lévő készülékek lecserélése jelentheti, ugyanis senki sem akar kidobni egy viszonylag újonnan vásárolt vevőkészüléket az új kódolási eljárások alkalmazása miatt. E probléma feloldására két kézenfekvő megoldás létezik. Az egyik, hogy ugyanazon műsort mind a régi, MPEG2 alapú kódolással, mind az új megoldással is sugározzák. Az új adások alacsonyabb sávszélesség-igénye miatt a szükséges sávszélesség nem duplázódna, de így is gazdaságtalan ez az eljárás.

Másik megoldás egy fokozatos áttérés lehetősége, ahol a fő szerep a készülégyártóknak jut. Minden új vevőben könnyen megteremthető az a technikai háttér, hogy a dekódoló algoritmusok lecserélhetők, frissíthetők legyenek. Olyan jelfeldolgozó áramkört kell választani, melyben több kódolási algoritmust lehet tárolni. Így egy adott adás esetén azt kell tudni, hogy az adott műsor dekódolásához melyik eljárást válassza ki a készülék. Ezt az információt egyszerűen hozzá lehet adni a hasznos adatokhoz vagy szervízinformációkhoz. A kérdés már csak annyi, hogy milyen módon kaphatja meg a készülék az új dekódoló rutint. A válasz triviális, a műsorokkal együtt az algoritmusokat is ki kell sugározni, melyet vesz a készülék és elraktároz. Ezen kívül a készülék internetes csatlakoztatásával is megoldható lehet a letöltés és frissítés, tehát az újabb készülékeket alkalmassá lehet tenni bármilyen módon kódolt multimédia adat megjelenítésére.

Irodalom

[1] L. Konyha, B. Enyedi, K. Fazekas, „Multimedia Distance Learning – Orthogonal Transformations”, EURASIP Conference on Budapest, Sept. 2001.

- [2] B. Enyedi, L. Konyha, K. Fazekas, „Using Wavelet Transform for Guiding Observation Cameras and Efficient Data Storage”, 3rd COST #276 Workshop on Information and Knowledge Management for Integrated Media Comm., Budapest, Oct. 2002.
- [3] S. M. Tran, L. Konyha, B. Enyedi, Cs. Szombathy, K. Fazekas, „Experiments on Transmitting MPEG-4 Content over MPEG-2 Transport Stream”, WIAMIS 5th International Workshop on Image Analysis for Multimedia Interactive Services, Lisboa, Portugal, 2004.
- [4] S. M. Tran, K. Lajos, E. Balazs, K. Fazekas, Sz. Csaba, „A Survey on the Interactivity Features of MPEG-4”, 46th International Symposium ELMAR-2004 focused on Navigation, Multimedia and Marine, Zadar, Croatia, 2004.
- [5] Turan, J., Bojkovic, Z., Filo, P., Samcovic, A., Ovsenic, L., „Signal Processing with Continuous Kernel Hough Transform”, FACTA UNIVERSITATIS (NIS), Ser. Elec. Eng. Vol.18, 2005, pp.113–126.
- [6] Bojkovic, Z., Turan, J., Samcovic, A., Ovsenic, L., „Coding, Streaming and Watermarking – Some Principles in Multimedia Signal Processing”, Acta El. et Inf., Vol.4, No.3, 2004, pp.15–20.
- [7] S. G. Mallat, „A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”
- [8] Amir Said, William A. Pearlman, „A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees”, IEEE Transaction on Circuit and Systems for Video Technology, Vol.6, June 1996.
- [9] V. Bottreau, M. Bénétière, B. Felts, B. Pesquet-Popescu, „A Fully Scalable 3D Subband Video Codec”



4. ábra
MPEG4
tartalom
beillesztése
DVB-T
rendszerbe