



**BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
VILLAMOSMÉRNÖKI ÉS INFORMATIKAI KAR**

***Wavelet transzformáción alapuló hatékony képkódolási
eljárások***

Tézisfüzet

Írta:

Konyha Lajos
2008.

Témavezetők:

Dr. Erényi István
Dr. Fazekas Kálmán

Az értekezés tárgya és célkitűzése

A doktori értekezésben mozgóképek veszteséges tömörítésével foglalkozok, ezen belül a wavelet transzformáción alapuló képkódolási eljárásokkal kapcsolatos eredményeimet mutatom be. A veszteséges tömörítési eljárásokat két fontos jellemzővel minősítjük: a tömörítési aránnyal és a visszaállított kép minőségével. Az eljárások vizsgálatánál ezen két számszerű értéket egymás függvényében értékeljük, cél magas minőség elérése kis adatsebesség mellett. Jelenleg a gyakorlatban is használható leghatékonyabb képtömörítési eljárások a wavelet transzformáción és a wavelet együtthatókat kódoló SPIHT eljárás alapulnak. Az elért eredményeket három csoportba osztva mutatom be, melyek a képkódolás jól elkülöníthető területeit fedik le.

Az első részben a tér- és időtartománybeli redundancia kihasználásához háromdimenziós diadikus wavelet transzformációt és ehhez módosított SPIHT eljárást alkalmazok a hatékony tömörítés eléréséhez. A nagy adatmennyiség feldolgozásának gyorsítására javaslok egy újítást, mely hatékony, valósidejű képkódolást eredményez.

A második részben a kétdimenziós, tetszőleges alakú objektumok kódolásával foglalkozok. Ebben a részben az SPIHT eljárást úgy módosítottam, hogy képes legyen hatékonyan kódolni az alakhoz igazodó wavelet transzformáció eredményeként keletkezett együtthatókat.

A harmadik részben a mozgáskompensáció területén elért eredményeimet ismertetem. A tetszőleges alakú objektumok mozgásának és az összetett mozgások kezeléséhez a blokkalapú mozgáskompensáció nem alkalmazható hatékonyan. Ezekben az esetekben a hálólalapú mozgáskompensáció eredményes. Az egyik fontos feladat az objektumot lefedő háló előállítás. Erre javaslok egy új eljárást, mely a háló csomópontjait és struktúráját a mozgáskompensáció szempontjából optimálisan állítja elő.

Alkalmazott kísérleti és vizsgálati módszerek

A kutatások során az alábbi vizsgálati módszereket alkalmaztam:

- Az alkalmazott álló- és mozgóképtömörítési szabványok és az ezekben előírt, illetve javasolt eljárások vizsgálata
- Lineáris, nemlineáris, reverzibilis, irreverzibilis transzformációk vizsgálata
- Két- és háromdimenziós transzformációk vizsgálata
- Többszörös felbontású jelanalízis vizsgálata
- Egy- és többdimenziós szűrők vizsgálata
- Lineáris, nemlineáris és adaptív kvantálások vizsgálata
- Entrópia kódoló eljárások vizsgálata
- Könyvtári irodalomkutatás (folyóirat- illetve konferenci cikkek), Internetes irodalomkutatás (szakmai adatbázisokban, például IEEE Xplore)

- [6] **G. Xing, J. Li, S. Li, Y. Zhang:** Arbitrarily Shaped Video Object Coding by Wavelet, *EEE Transactions on Circuits and Systems for Video Technology, Volume: 11 Issue: 10, 2001, pp. 1135 – 1139*
- [7] **Murat Tekalp:** Digital Video Processing (Paperback), *Prentice Hall PTR, 1995*

Hivatkozási lista

Hivatkozás helye	Hivatkozott publikáció
Jurij F. Tasič, Urban Burnik: The Challenge of Personalisation and Mobility: COST #276, EUROCON 2003, Computer as a Tool , vol II. pp.411-515, Ljubljana, Slovenia, September 22-24, 2003	[S1] [S17]
Kuan-Liang Chen: Wafer ID Recognition Technology Research , Thesis Id: 093NKIT5650013, Advisor: Wei-Chih Hsu, 2004	[S12]
M.Mitrea, F.Prêteux, M.Petrescu, A.Vlad: The StirMark Watermarking Attack in the DWT Domain , <i>Proceedings 12th IEEE International Conference on Systems, Signals and Image Processing (IWSSIP'05), Halkida, Greece</i> , Vol. 2, September 2005, p. 5-9.	[S9]
Iveta Gladišová: Simple Image Indexing and Retrieval Technique Based on Dominant Colours , <i>13th International Conference on Systems, Signals and Image Processing (IWSSIP 2006)</i> , pp.345-347, Budapest, Hungary, 21-23 September 2006	[S17] [S34]
Suresh, K.V.; Kumar, G.M.; Rajagopalan, A.N.; Superresolution of License Plates in Real Traffic Videos , <i>Intelligent Transportation Systems, IEEE Transactions on</i> , Volume 8, Issue 2, June 2007 Page(s):321 - 331	[S12]
Chien-Kuo Li: Using Omni-Directional and PTZ Cameras to Implement Real-Time Tracking of Moving Objects on a DSP Board , <i>Thesis</i> , Chang Yuan Christian University, Taiwan, 2006	[S12]
Sheng-Long Lee: Using Motion Information and Multiple Classifiers in a License Plate Recognition System for Moving Vehicles on the Road , <i>Thesis</i> , Chang Yuan Christian University, Taiwan, 2004	[S12]

Irodalomjegyzék

- [1] C. Fogg, D. J. LeGall, J. L. Mitchell, W. B. Pennebaker: MPEG Video Compression Standard, *Springer, 1996*
- [2] J. M. Shapiro: Embedded Image Coding Using Zerotrees of Wavelet Coefficients, *IEEE Transactions on Signal Processing*, Vol. 41, No. 12 (1993), p. 3445-3462
- [3] A. Said, W. A. Pearlman: A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243-250, June 1996.
- [4] J. Van Overloop, W. Van De Sype, K. N. Denecker, P. De Neve, E. Sundermann, I. L. Lemahieu: Experimental Comparison of 2D and 3D Wavelet Image Compression Methods for Medical Image Sets, *Medical Imaging 1998: Image Display*, June 1998, pp. 348-358
- [5] E. Moyano, F. J. Quiles, A. Garrido, J. Duato, L. Orozco-Barbosa: Efficient 3D Wavelet Transform Decomposition for Video Compression, *Digital and Computational Video, 2001. Proceedings. Second International Workshop on 8-9 Feb. 2001 pp. 118 – 125*

A kutatás gyakorlati része során alkalmazott kísérleti módszerek:

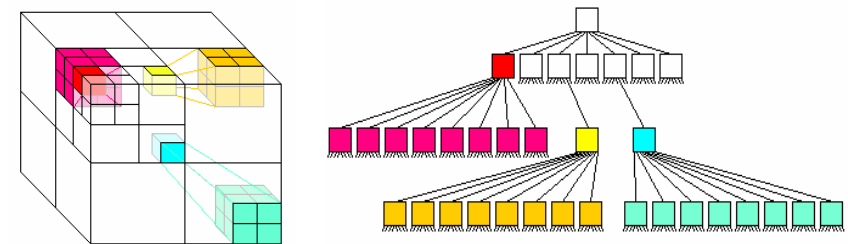
- Algoritmusok vizsgálata számítógépes szimulációval (C, C++, MATLAB)
- Képfeldolgozási eljárások számítógépes implementálása, értékelése (C, C++, MATLAB, assembly)
- Eljárások optimalizálása hardverközeli programozással
- Implementációk értékelése szabványos teszt sorozatokkal elért eredmények alapján
- Számítógépes adatfeldolgozás, megjelenítés

Az értekezés új tudományos eredményei

1. tétiscsoport: Háromdimenziós diadikus wavelet transzformáció alkalmazása valósidejű videó kódolásra.

A videó tömörítések során az elterjedt eljárásokban transzformációs kódolást alkalmaznak a térbeli redundancia kihasználásához, mely során 8x8, 16x16 vagy több szomszédos képpontot használnak föl egyszerre. Ezzel szemben az időtartománybeli redundancia kihasználásához egyszerű különbségi képeket használnak, maximum 3 képkocka alapján (aktuális képkocka, múltbeli prediktor, jövőbeli prediktor) [1].

Az emberi látórendszer (HVS) egyik tulajdonsága, hogy a hirtelen változásokhoz csak némi késleltetéssel tud alkalmazkodni, azaz az időben gyors változásokat (a térbeli gyors változásokhoz hasonlóan) gyengén érzékeli. A természetben előforduló videótartalmakban az időbeli redundancia jellemzően több egymás után következő képkockán jelentkezik (egyforma képkockák, illetve lassú, folytonos átmenet), kivéve néhány speciális esetet (jelenetváltás, mozgás következtében eltűnő vagy előtűnő részek), amikor hirtelen változás van két egymás utáni képkockán. A térbeli redundancia kihasználásához a wavelet transzformáció a leghatékonyabb eljárás, célszerű az időtartománybeli redundancia kihasználásához is ezt az eljárást használni [2, 3].



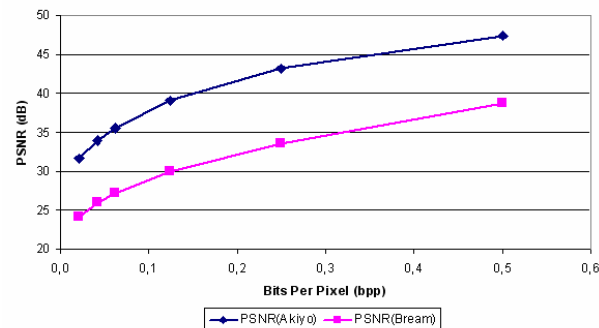
1. ábra – Háromdimenziós diadikus wavelet együtthatók (a), és az együtthatók közötti kapcsolatokat reprezentáló fa struktúra (b)

A háromdimenziós diadikus wavelet transzformáció együtthatóinak eloszlását

mutatja az 1.a ábra [4]. Az SPIHT algoritmus működése folyamán az együtthatók összeszedésének sorrendjét a wavelet együtthatók közötti összetartozást leíró térbeli orientációjú fa struktúrájú gráf határozza meg [5]. Ezt a gráfot módosítottam az 1.b ábrán látható módon, ezáltal az SPIHT algoritmust képessé tettem a keletkezett wavelet együtthatók feldolgozására.

A diadikus felbontással és SPIHT algoritmussal kapott eredményeket mutatja a 2. ábra. A felső görbe az „Akiyo” sorozattal, az alsó görbe a „Bream” sorozattal elért eredményeket mutatja. A „Bream” sorozatnál azonos tömörítési arány mellett rosszabb minőséget kapunk, mert ez sokkal dinamikusabb, összetettebb részleteket tartalmazó sorozat. Az eredményeken látszik, hogy az egyszerű anyagot tartalmazó videót 0.021-0.042bpp tömörítési arány mellett elfogadható minőségben lehet kódolni, míg a bonyolultabb struktúrákat tartalmazó képeknél körülbelül 0.1-0.2bpp információra van szükség az elfogadható minőséghez.

T1/a: Az időtartománybeli redundancia hatékony kihasználásához módosítottam az SPIHT eljárást a diadikus háromdimenziós wavelet felbontás együtthatóinak kezelésére [S14]. Az algoritmus működését meghatározó gráf struktúrájának változtatásával végzett kísérletek alapján bebizonyosodott, hogy a térbeli orientációjú gráfok által irányított működésű algoritmusok alkalmasak a háromdimenziós adatok hatékony feldolgozására. Az SPIHT eljárás szimulációs eredményei azt mutatják, hogy az elérhető tömörítési arány jelentősen jobb a korábbi megoldásokhoz képest. Az eljárás teljes működési ideje a megnövekedett adatmennyiség miatt az eredeti algoritmus működési idejének többszörösére növekedett.



2. ábra - A 3D SPIHT eljárás eredményei a tömörítési arány függvényében

A háromdimenziós adatfeldolgozás velejárója a nagy adatmennyiség kezelése, ami általában erőforrásigényes. Különösen igaz ez az SPIHT eljárásra, mert ennek az algoritmusnak a leginkább időigényes része az új, jelentős együtthatók keresése, mely az 1.b ábrán látható gráf részfáinak bejárását jelenti. Ez azt jelenti, hogy egy

Magyar nyelvű kiadványban megjelent konferencia-előadás:

- [S26] Enyedi B., Konyha L., Fazekas K.: **Valós idejű mozgóképtömörítés wavelet-transzformációval**, TV2005, Televízió- és hangtechnikai konferencia és kiállítás, 2004. június 1-2., Budapest, Magyarország
- [S27] Enyedi B., Konyha L., Fazekas K.: **Megfigyelő kamerák vezérlése wavelet transzformációval, hatékony adattárolás megvalósítása**, TV 2002 konferencia, pp.149-159, 2002 május, Budapest
- [S28] Enyedi B., Konyha L., Szombathy Cs., Tran Min Son, Gschwindt A., Szokolay M., Fazekas K.: **MPEG-4 rendszerek kutatása és beintegrálása a DVB-T átviteltechnikába**, TV2005, Televízió- és hangtechnikai konferencia és kiállítás, 2005. június 1-2., Budapest, Magyarország
- [S29] Enyedi B., Konyha L., Szombathy Cs., Tran Min Son, Gschwindt A., Szokolay M., Fazekas K.: **MPEG4 alapú műsorszórás a DVB-T technikában**, HTE-BME 2005 diákkonferencia, 2005. május 19., Budapest, Magyarország

Külföldön megjelent, idegen nyelvű, lektorált folyóiratcikk:

- [S30] S. M. Tran, L. Konyha, B. Enyedi, K. Fazekas, J. Turan Jr.: **Deployment of Constrained Delaunay Mesh in VOP Shape and Texture Coding**, *Journal of Electrical Engineering*, vol.56, 07-08 (2005) pp.189-194, Slovakia, 2005
- [S31] L. Konyha, B. Enyedi, K. Fazekas: **Real Time Number Plate Localization Algorithms**, *Journal of Electrical Engineering*, vol.57, 2 (2006) pp.69-77, Slovakia, 2006
- [S32] L. Konyha, B. Enyedi, K. Fazekas: **Motion Compensated Video Compression with 3D Wavelet Transform and SPIHT**, *Radioengineering, Proceedings of Czech and Slovak Technical Universities and URSI Committees*, vol.15, n.1, pp.42-46, Slovakia, April 2006

Magyarországon megjelent, magyar nyelvű, lektorált folyóiratcikkek:

- [S33] Enyedi B., Konyha L., Szombathy Cs., Tran Min Son, Gschwindt A., Szokolay M., Fazekas K.: **MPEG-4 alapú átvitel megvalósítása a DVB-T technikában**, *Híradástechnika (LIX. 2004/7)*, pp.23-27, 2004. július, Budapest, Magyarország
- [S34] Enyedi B., Konyha L., Szombathy Cs., Tran Min Son, Gschwindt A., Fazekas K.: **Új lehetőségek a digitális televíziózásban**, *Híradástechnika (LX. 2005/9)*, pp.53-57, 2005. szeptember, Budapest, Magyarország

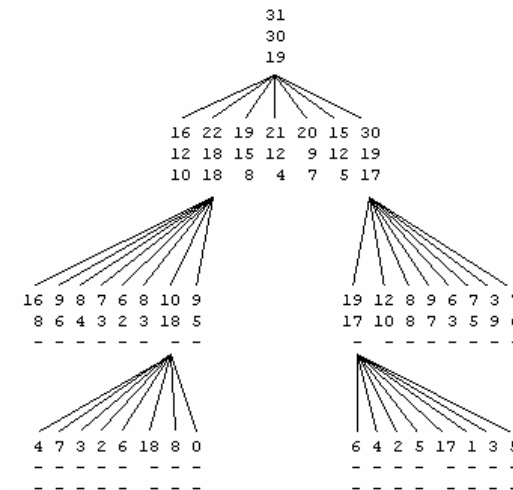
Egyéb:

- [S35] Enyedi B., Konyha L., Szombathy Cs., Tran Min Son, Gschwindt A., Szokolay M., Fazekas K.: **MPEG-4 alapú átvitel megvalósítása a DVB-T technikában**, *HTE előadás, Postamúzeum előadóterme*, 2004. szeptember 14., Budapest, Magyarország

- [S13] B. Enyedi, L. Konyha, S.M. Tran, Cs. Szombathy, K. Fazekas: **A Survey on the Interactivity Feature of MPEG-4**, 46th International Symposium ELMAR-2004 focused on Navigation, Multimedia and Marine, pp.30-38, Zadar, Croatia, June 16-18, 2004
- [S14] B. Enyedi, L. Konyha, Cs. Szombathy, K. Fazekas: **Real-time Video Compression with 3D Wavelet Transform and SPIHT**, 4th International Conference Intelligent Systems Design and Applications (ISDA 2004), Budapest, Hungary, August 26-28, 2004
- [S15] S. M. Tran, L. Konyha, B. Enyedi, Cs. Szombathy, K. Fazekas: **Experiments on Transmitting MPEG-4 Content over MPEG-2 Transport Stream**, 5th International Workshop on Image Analysis for Multimedia Interactive Services, WLAMIS 2004, Lisboa, Portugal, April, 2004
- [S16] B. Enyedi, L. Konyha, S.M. Tran, Cs. Szombathy, K. Fazekas: **A Survey on the Interactivity Feature of MPEG-4**, 11th International Workshop on Systems, Signal and Image Processing IWSSIP'04, pp.87-90, Poznan, Poland, September 13-15, 2004
- [S17] B. Enyedi, L. Konyha, K. Fazekas: **Fast Video Compression Based on 3D Wavelet Transform and SPIHT**, 7th COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication, pp.121-124, Ankara, Turkey, November 4-5, 2004
- [S18] B. Enyedi, L. Konyha, Cs. Szombathy, S.M. Tran, A. Gschwindt, K. Fazekas, M. Szokolay: **Implementation of MPEG4-Based Transmission in DVB-T Technology**, 47th International Symposium ELMAR-2005 focused on Multimedia Systems and Applications, Zadar, Croatia, June 08-10, 2005
- [S19] B. Enyedi, L. Konyha, K. Fazekas: **Threshold Procedures and Image Segmentation**, 47th International Symposium ELMAR-2005 focused on Multimedia Systems and Applications, Zadar, Croatia, June 08-10, 2005
- [S20] B. Enyedi, L. Konyha, K. Fazekas: **Fast License Plate Localization**, 6th International Scientific Conference DSP-MCOM'05, pp.10-14, Kosice, Slovakia, September 13-14, 2005
- [S21] B. Enyedi, L. Konyha, S.M. Tran, K. Fazekas: **Toward a High Abstraction Level of MPEG-4 BIFS Composition**, 6th International Scientific Conference DSP-MCOM'05, pp.50-54, Kosice, Slovakia, September 13-14, 2005
- [S22] B. Enyedi, L. Konyha, K. Fazekas: **Character Localization**, 13th International Conference on Systems, Signals and Image Processing (IWSSIP 2006), pp.317-326, Budapest, Hungary, 21-23 September 2006
- [S23] B. Enyedi, L. Konyha, K. Fazekas: **Segmentation – Region Growing Based on Adaptive Threshold Algorithms**, 13th International Conference on Systems, Signals and Image Processing (IWSSIP 2006), pp.327-332, Budapest, Hungary, 21-23 September 2006
- [S24] B. Enyedi, L. Konyha, K. Fazekas, J.Turán: **License Plate Localization and Storage Method**, 14th International Conference on Systems, Signals and Image Processing (IWSSIP 2007), pp.441-445, Maribor, Slovenia, 27-30 June 2007
- [S25] B. Enyedi, L. Konyha, K. Fazekas, J.Turán: **Small Tricks to Enhance the Accuracy of License Plate Character Recognition**, 2nd International Conference on Signals Processing and Multimedia Applications (SIGMAP 2007), pp.125-128, Barcelona, Spain, 28-31 July 2007

adott részfa gyökerétől kiindulva meg kell vizsgálni a leszármazottait (bizonyos esetekben az összes leszármazottját, máskor csak a távoli, azaz nem közvetlen leszármazottait) az új jelentős értékű csomópontok keresése során. Háromdimenziós eljárás esetén a szinteken lefelé haladva minden csomópontnak nyolc leszármazottja van (kivéve a legfelső szintet, ahol csak hét), tehát a megvizsgálandó csomópontok száma a bejárt szintek számával exponenciálisan nő. Ez a gyakorlatban előforduló képméreteknél lehetetlenné teszi a valósidejű feldolgozást.

Az eljárás felgyorsítására új változókat vezettem be a gráf csomópontjaihoz, melyek a leszármazottak illetve a távoli leszármazottak abszolút értékének maximumát tartalmazzák (3. ábra). Így a teljes részfa bejárása helyett elég az új változó értékét megvizsgálni, azaz egy részgráf bejárása helyett csak egy műveletet kell elvégezni.

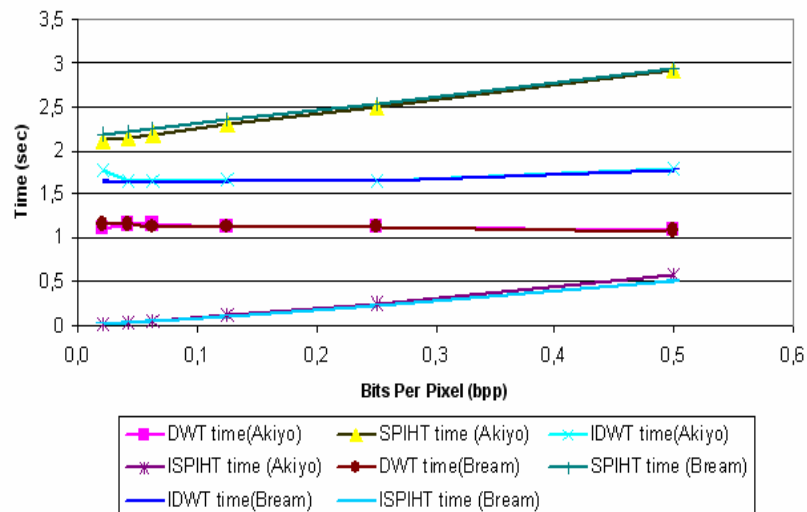


3. ábra – Az új adatok elhelyezkedése az SPIHT algoritmus grájfjában

A legelső szinten elhelyezkedő együtthatóknak nincs leszármazottjuk (az együtthatók 7/8-a a legelső szinten van), így ezekhez nem tartoznak új adatok. Az eggyel magasabb szinten található együtthatóknak nincsenek távoli leszármazottjai, így ezek maximális értékét sem kell nyilván tartani. Összességében véve az együtthatók nyolcadánál kell tárolni a leszármazottak maximumát, hatvanegyedénél a távoli leszármazottakat is, ami 14% memóriátöbbletet jelent.

A bevezetett változók segítségével a teljes kódolás és dekódolás egyes lépéseinek időigénye a tömörítési arány függvényében a 4. ábrán látható. A szimulációnál használt sorozatok 352x288 képpont méretűek, 128 képkockát tartalmaznak, így a valósidejű kódolás időkorlátja 30 másodpercenként képkockát feltételezve kicsit több, mint 4.26 másodperc. A teljes kódolás ideje a legrosszabb esetben 4 másodperc, a dekódolás ideje 2.4 másodperc, tehát mindkét folyamat

elvégezhető valós időben.



4. ábra – A gyorsított SPIHT algoritmussal végzett kódolás futási ideje

T1/b: Új adatokat vezettem be a 3D SPIHT algoritmus működésének felgyorsítására, melyek segítségével a háromdimenziós SPIHT eljárás képes valós időben feldolgozni a wavelet együtthatókat [S17, S32]. A leszármazottak és a távoli leszármazottak maximális értékének bevezetésével a jelentős csomópontokat kereső lépés (az eljárás legidőigényesebb lépése) töredékére csökkent (a részfa bejárása helyett egy művelet). Az új adatok tárolása csak 14% többletmemóriát igényel.

Publikációs lista

Nemzetközi részvételű konferencia kiadványában megjelent, idegen nyelvű, lektorált előadás:

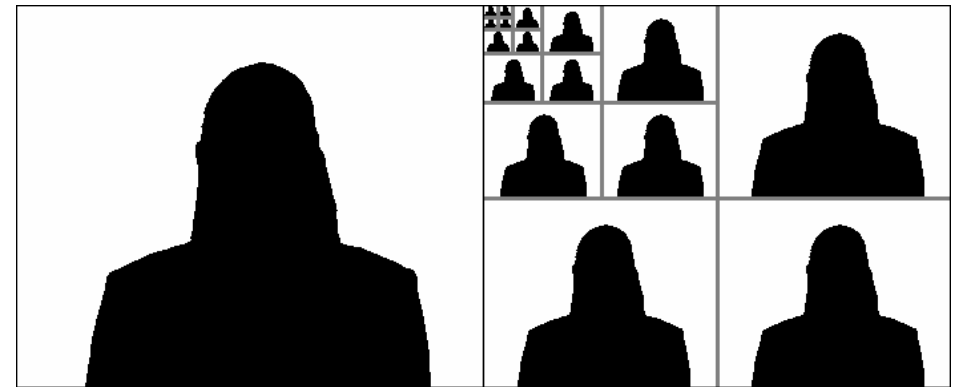
- [S1] L. Konyha, B. Enyedi, K. Fazekas: **Multimedia Distance Learning – Orthogonal Transformations**, *EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services*, pp.196-198, Budapest, Hungary, September 11-13, 2001
- [S2] L. Konyha, S. M. Tran, B. Enyedi, K. Fazekas, A. Gschwindt: **Multiresolution Image Decomposition Using Wavelet Transformation**, *Conference on Using Technology in Open and Distance Learning*, pp.49-56, Maribor, Slovenia, September 13-14, 2001
- [S3] L. Konyha, S. M. Tran, K. Fazekas, A. Gschwindt, **Deployment of wavelet transform to VOP codec scheme**, pp. 233-237, Video/Image Processing and Multimedia Communications 4th EURASIP-IEEE Region 8 International Symposium on VIPromCom, Zadar, Croatia, June 2002
- [S4] B. Enyedi, L. Konyha, K. Fazekas: **New Method for Security Systems: On-line Camera Control, Digital Data Processing**, *3rd COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication*, pp.154-160, Budapest, Hungary, October 11-12, 2002
- [S5] S. M. Tran, L. Konyha, B. Enyedi, K. Fazekas, A. Gschwindt, J. Turan Jr.: **Deployment of Constrained Delaunay Mesh in VOP Shape Coding**, *3rd COST 276 Workshop on Information and Knowledge Management for Integrated Media Communication*, pp.161-165, Budapest, Hungary, October 11-12, 2002
- [S6] S. M. Tran, L. Konyha, B. Enyedi, K. Fazekas, A. Gschwindt, J. Turan Jr.: **Deployment of Constrained Delaunay Mesh in VOP Shape Coding**, *9th International Workshop on Systems, Signals and Image Processing (IWSSIP'02)*, pp.205-211, Manchester, United Kingdom, November 7-8, 2002
- [S7] L. Konyha, B. Enyedi, S. M. Tran, K. Fazekas: **Multimedia Education – Distant Learning**, *4th European Workshop on Image Analysis for Multimedia Interactive Services*, pp.483-486, Queen Mary University, London, United Kingdom, April 9-11, 2003
- [S8] L. Konyha, B. Enyedi, S. M. Tran, K. Fazekas: **Content-Based Mesh Generation Algorithm**, *4th EURASIP Conference focused on Video / Image Processing and Multimedia Communications*, pp.175-180, Zagreb, Croatia, July 2-5, 2003
- [S9] L. Konyha, B. Enyedi, S. M. Tran, K. Fazekas: **e-Learning – Multimedia Education**, *4th cost 276 workshop, Information and Knowledge Management for Integrated Media Communication*, pp.95-97, Bordeaux, France, March 31- April 1, 2003
- [S10] S. M. Tran, L. Konyha, B. Enyedi, K. Fazekas: **Toward a High Abstraction Level of MPEG-4 BIFS Composition**, *10th International Workshop on Systems, Signals and Image Processing IWSSIP'03*, Czech Technical University in Prague, Prague, Czech Republic, September 10-11, 2003
- [S11] L. Konyha, B. Enyedi, S. M. Tran, K. Fazekas: **Paradigm of Multimedia Distance Teaching**, *10th International Workshop on Systems, Signals and Image Processing IWSSIP'03*, Czech Technical University in Prague, Prague, Czech Republic, September 10-11, 2003
- [S12] B. Enyedi, L. Konyha, Cs. Szombathy, K. Fazekas: **Strategies for Fast Licence Plate Number Localization**, *46th International Symposium ELMAR-2004 focused on Navigation, Multimedia and Marine*, pp.579-584, Zadar, Croatia, June 16-18, 2004



14. ábra - A háló struktúrája objektum alapú képkezelés esetén

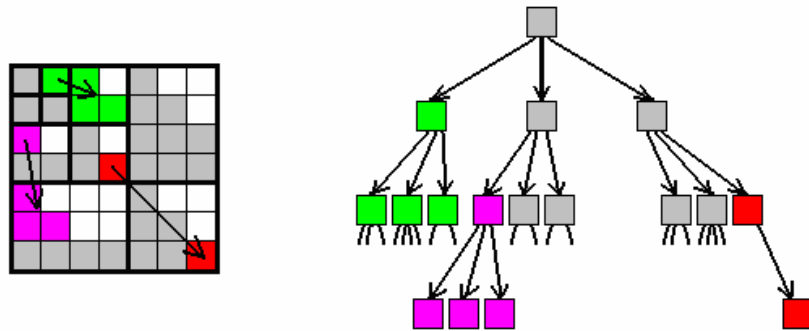
2. tézis: Tetszőleges alakú objektumok kódolása.

A hatékonysága és sokoldalúsága miatt elterjedő MPEG-4 szabvány elsőként teszi lehetővé a tetszőleges alakú objektumok kezelését a videó kódolásban. A wavelet transzformáció és az SPIHT algoritmus kiemelkedően jó hatékonysága alapján ésszerűnek látszik az eljárások módosítása úgy, hogy ne csak téglalap, hanem tetszőleges alakú képeket is képesek legyenek feldolgozni. Irodalmi eredmények alapján ismert, hogy a wavelet transzformáció alkalmassá tehető bármilyen alakú objektum kódolására [6]. Az alakhoz igazodó transzformáció egyik fontos tulajdonsága, hogy csak olyan együtthatók keletkeznek, melyekhez tartozó térbeli pozíciókat az eredeti objektum lefedi. Egy videó objektum maszkját és a keletkezett együtthatók elhelyezkedését mutatja az 5. ábra.



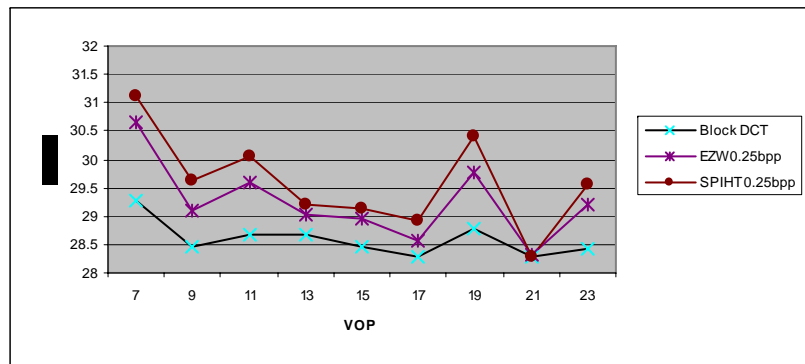
5. ábra – Videó objektum és az alakhoz igazodó wavelet együtthatók

Az SPIHT algoritmust az ilyen, módosított wavelet transzformáció eredményeinek feldolgozására kell alkalmassá tenni. Az együtthatók közötti kapcsolatokat figyelve látszik, hogy a legfelső szinten található együtthatóhoz eggyel alacsonyabb szinten három együttható tartozik, a mélyebb szinteken bármelyik együtthatónál előfordulhat, hogy 0, 1, 2, 3 vagy 4 együttható tartozik hozzá az alatta lévő szinten (kivéve a legalsó szintet). Ennek megfelelően módosítottam az SPIHT algoritmus működését irányító gráfot úgy, hogy csak azokat a csomópontokat vegye figyelembe a kódolásnál, melyekhez tartozó pozíciókban valóban keletkeztek wavelet együtthatók (6. ábra). Az így módosított eljárással az objektumok mintázata illetve maradékhibája a jelenleg elterjedt alakhoz igazodó koszinusz transzformációs és EZW eljárásoknál hatékonyabban kódolható (7. ábra).



6. ábra – Alakhoz igazodó SPIHT algoritmus gráfja

Az eljárás sebességét a javasolt módosítás két tényezője befolyásolja: egyrészt a bonyolultabb keresgélési szabályok lassítják az algoritmus működését, másrészt amiatt, hogy összességében kevesebb adatot kell feldolgozni, gyorsul az eljárás. A gyakorlatban a bonyolultabb szabályok miatt az algoritmus lassabban működik, mint az eredeti SPIHT eljárás, de a T1/a tézisben bevezetett új adatokkal a működés ugyanúgy felgyorsul. Ennek memóriaigénye azonban a leszármazottak maximumának tárolásához 25% (az együttthatók negyede nem a legalsó szinten van), a távoli leszármazottak tárolásához 6.25% (az együttthatók tizenhatoda nem az alsó két szinten van), tehát összesen 31.25%. Így a teljes alakhoz igazodó SPIHT kódolás elvégezhető valós időben.



7. ábra – Az alakhoz igazodó SPIHT összehasonlítása egyéb eljárásokkal

T2: Módosítottam az SPIHT eljárást úgy, hogy a wavelet együttthatók objektum által le nem fedett részeit a kódolás során ne vegye figyelembe, ezekre a területekre ne pazaroljon adatokat, így megnövekedett a tömörítés hatékonysága [S3]. A leszármazottak és a távoli leszármazottak maximális értékének felhasználásával ez az eljárás is jelentősen felgyorsult, a gyorsítás miatti memóriátöbblet 31.25%.

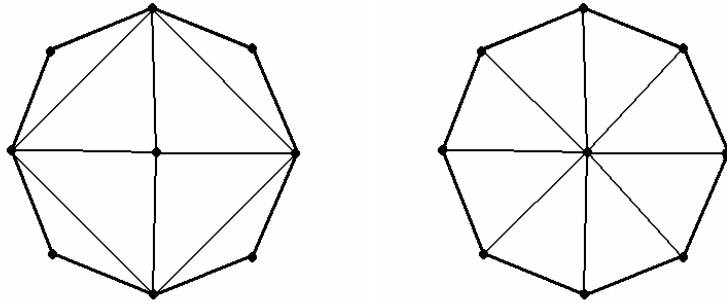
T3/c: Készítettem egy eljárást a háló összeköttetések optimalizálására, amely a pontatlan mozgásbecslésből adódó hibákkal szembeni ellenállóságot növeli [S8]. A háló optimalizálásának célja a pontatlan mozgásbecslésből eredő hibák minimalizálása, melyet a kedvezőtlen alakú (hegyesszögű) háromszögek oldalainak átrendezésével old meg.



13. ábra – A háló struktúrája különböző beállításokkal képkocka alapú képfeldolgozás esetén

A képkocka alapú képfeldolgozás esetén kapott eredmények láthatók a 13. ábrán. A bal oldali képeken az időbeli súlyozó tényező értéke 2, a jobb oldali képeken 100. A csomópontok minimális távolsága a felső képeken 1, az alsó képeken 10 képpont. A felső ábrákon a csomópontok olyan közel kerültek egymáshoz, hogy a mozgásbecslés apró hibája is a konzisztencia sérülését jelenti. Az alsó ábrákon látható hálószerkezet jól használható a mozgáskompenzációban, a háló élei általában az objektumok határával egybe esnek. Az MPEG eljárásokban használt 16x16 képpont méretű blokkalapú mozgásbecslés esetén a bemutatott 352x288 képpont méretű kép mozgását 396 mozgásvektorral lehet leírni. A bal alsó képen a háló csomópontjainak száma (a szükséges mozgásvektorok száma) 280, a jobb alsón, ahol a pontok az időben változó területekre kerültek 172.

A 14. ábrán látható objektum alapú kezelésnél a töröttvonalas közelítés a körvonal összetett részénél (fej íve) több pontot tartalmaz, az egyszerű részekenél keveset (a kép alja). A belső csomópontok itt is a kép mintázatának megfelelően helyezkednek el. Az objektumot lefedő háló összesen 122 pontból áll.

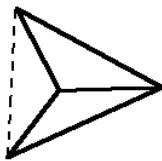


11. ábra – Háló szerkezet optimalizálás előtt (a) és után (b)

A 11. ábrán az is látható, hogy a struktúra optimalizálásához mindig két szomszédos háromszöget kell megváltoztatni úgy, hogy a közös élüket megszüntetjük, a nem közös csomópontjait összekötő új élet létrehozunk. Az optimalizálásnál a cél a torz háromszögek megszüntetése, ezért az élek cserélgetése közben arra is figyelni kell, hogy az újonnan létrejött háromszögek között ne legyen torzabb az eredeti torz háromszögnél.

Ezek alapján az eljárás a következő módon működik: Megkeresi a legtorzabb háromszöget, megpróbálja mindegyik szomszédjával felcserélni az éleket, eközben meghatározza, melyik csere eredményezi azt az elrendezést, amikor a legtorzabb háromszög alaktényezője a legkisebb (legoptimálisabb). Ha ezzel az elrendezéssel sikerül javítani a jelenlegi helyzeten, elvégzi a cserét. Ezután keresi a következő háromszöget, míg végül nem talál javítható háromszögeket.

Bizonyos esetekben a közös csúcsokat és a nem közös csúcsokat összekötő élek cseréje nem végezhető el. A 12. ábrán látható háromszögek esetén a csere után kapott háromszögek közül az egyik magába foglalja a másikat, azaz a háló konzisztenciája sérült. Ez az eset csak akkor fordulhat elő, ha a felcserélt élek nem metszik egymást, ezért az algoritmust annyiban kell módosítani, hogy csak akkor vizsgálja a felcseréléssel kapott torzultságokat, ha az élek metszik egymást.



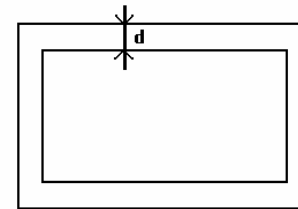
12. ábra - Nem optimalizálható háromszögek

3. Téziscsoport: Hálóalapú mozgáskompenzáció

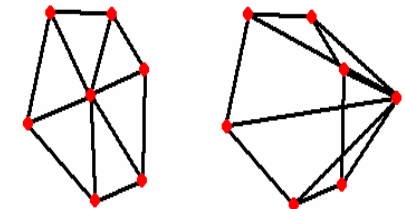
Az általánosan használt blokkalapú mozgáskompenzációnak hátránya, hogy csak az egyszerű elmozdulásokat lehet vele kezelni, összetettebb jellegű mozgások esetén (pl.: kamera közelítés, forgás, torzulós mozgások) mindig nagy lesz a kompenzáció utáni maradékhiba. Másik hátránya, hogy alapvetően téglalap alakú képkockák kezelésére fejlesztették ki, ezért nem alkalmazható tetszőleges alakú objektumok mozgásának kompenzálására. Összetettebb alakzatok és összetettebb mozgások kezelésére alkalmasak a hálóalapú mozgáskompenzációs eljárások. Ezek az eljárások a képkockára vagy a videó objektumra valamilyen sokszögekből (általában háromszögekből) álló hálót feszítenek, majd a mozgásbecslés során a csomópontok mozgásvektorait határozzák meg. A kompenzáció során a sokszögek belsejében található képpontok mozgásvektorait a sokszög csúcspontjaiból (a háló csomópontjaiból) határozzák meg valamilyen interpolációval [7]. Az egyik elterjedt hálókészítési eljárás a Delaunay eljárás, mely adott geometriai kényszereknek megfelelő hálóval fedi le az objektumot. Ennek hátránya, hogy a háló előállításánál csak a megadott geometriai kényszereket veszi figyelembe, ezért nem mindig a mozgáskompenzáció szempontjából optimális hálót készít.

Az általam javasolt eljárás a mozgáskompenzáció céljaihoz optimális hálót készít úgy, hogy az objektumok geometriai adatai mellett a mintázatukat is figyelembe veszi. Az eljárás három algoritmusból áll, ezek feladata a háló csomópontjainak kijelölése, a csomópontok összeköttetések kialakítása, a struktúra optimalizálása az összeköttetések átrendezésével.

A háló csomópontjainak kijelölésénél azt használtam föl, hogy a mozgásbecslés során megbízhatóbban követhetőek a nagyobb változásokat tartalmazó területek, mint a sima, homogén felületek. Emellett azt is figyelembe vettem, hogy az időben nem változó területek a mozgáskompenzáció szempontjából lényegtelenek, tehát a területek időbeli változékonyságát is fel kell használni a háló pontjainak kijelölésénél. A javasolt eljárás ezek figyelembevételével minden képpont esetén kiszámítja a változékonyságot, ami egyszerűen a térben és időben szomszédos képpontoktól való (intenzitásbeli) eltérések abszolút értékének összege. Ez a mennyiség azt tükrözi, hogy mennyire fontos, hogy az adott pozícióba csomópont kerüljön. A mozgáskompenzáció szempontjából kiemelt jelentőségű az



8. ábra – Belső csomópontoktól mentes sáv a torz háromszögek elkerüléséhez



9. ábra – A rossz mozgáskompenzáció miatt sérült háló

időtartománybeli változás, ezért az összegzésnél az időbeli különbség nagyobb súlyjal szerepelhet (ez beállítható paraméter). Az egyes képpontokra kiszámított változékonyságok összegét elosztva az előállítandó pontok számával, kiderül, hogy mennyi változékonyság jut a háló egy csomópontjára. Ezután megkeresi a legnagyobb változékonyságú csomópontot, ez lesz a háló egyik csomópontja. Ennek a pontnak a környezetében kinullázza a változékonyság értékeket, egyre távolabb haladva addig, míg a kinullázott értékek összege eléri az egy csomópontra jutó átlagos változékonyság értéket. Ezután folytatódik az eljárás a következő legnagyobb változékonyságú pont keresésével. Ezekkel a szabályokkal biztosítható, hogy a homogén területekre ne kerüljön csomópont, az egyszerű képrészeket tartalmazó területekre kevés, az összetett területekre pedig sok csomópont kerüljön.

A kép vagy objektum széléhez közel kerülő csomópontokkal mindenképp nagyon torz háromszögeket kapnánk, ezért a határvonalától adott távolságon belül az eljárás nem keres belső csomópontokat (8. ábra). A külső csomópontok felvételénél objektum alapú képkezelésnél az objektum körvonalát közelítő töröttvonal csúcspontjai lesznek a háló külső csomópontjai. Képkocka alapú képkezelésnél először a képkocka négy sarkát jelöljük ki csomópontként, majd a 8. ábrán látható biztonsági sáv változékonyság értékeit az adott részhez legközelebb fekvő oldalra összegezzük, majd az összeg változékonyságot felhasználva keresünk pontokat.

A mozgáskompensáció pontatlanságához vezethet az is, ha túl kis méretűek a háromszögek, mert ezek csúcspontjai túl közel vannak egymáshoz, és pontatlan mozgásbecslés esetén a háló konzisztenciája sérülhet (9. ábra). Ennek elkerülése érdekében az eljárás paramétere a csomópontok minimális távolsága, aminek a hatására egy pont felvétele esetén mindenképp kinullázza a megadott távolságnál közelebbi pontokhoz tartozó változékonyság értékeket.

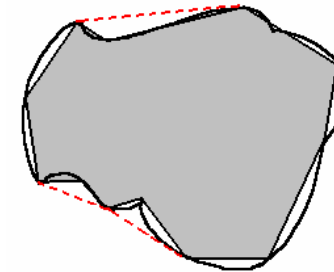
T3/a: Kifejlesztettem egy algoritmust a háló csomópontjainak felvételére, amely a mozgáskompensáció szempontjából optimális pontokat jelöl ki [S8]. A kijelölt pontok a mozgásbecslés során megbízhatóan követhetőek, emellett az összetett, részletdús területekre sok csomópont kerül, az egyszerű területekre kevés.

A következő feladat egy konzisztens háló kialakítása a csomópontok összekötésével. A háló struktúrájára a következő megkötések vannak: az összeköttetések nem keresztezhetik egymást, egy összeköttetésen két pontnak kell lennie (a végpontoknak), másik pont nem lehet rajta (a feltételek be nem tartása esetén a 9. ábrán látható struktúrát eredményezhet).

Ezek biztosításához a következő módon kell összeköttetéseket kialakítani: meg kell keresni azt a két legközelebbi pontot, melyek még nincsenek összekötve és a leendő összeköttetés nem keresztez egyetlen már meglévő összeköttetést sem. Ezáltal természetesen teljesül, hogy az élek (összeköttetések) nem kereszteznek egymást, a második feltétel pedig azért teljesül, mert a két legközelebbi pont közötti szakaszon nem lehet másik pont (ellenkező esetben az egyik végponthoz közelebb lenne a közbenső pont, mint a másik végpont). Az eddig leírtak alapján konkáv

alakú objektumok esetén előfordulhat, hogy az objektumon kívül haladó összeköttetéseket valósítunk meg. A 10. ábrán szaggatott vonallal jelölve látható, hogy az új élek a feltételeknek megfelelnek, mégis rossz struktúrát eredményeznek. Mivel ez csak akkor fordulhat elő, ha két külső pontot próbálunk összekötni, az előbbi eljárást a következőkkel kell módosítani: először összeköti sorrendben a külső pontokat, majd az új élek keresésénél csak belső csomópontból induló éleket próbál kialakítani. Ilyen elven mindig előállítható egy szabályos hálószerkezet.

T3/b: Kidolgoztam egy eljárást, amely a csomópontok alapján automatikusan létrehoz egy összeköttetési struktúrát, amely a csomópontokból háromszögalapú hálót készít [S8]. Az eljárás előnye, hogy a csomópontok koordinátáján kívül más bemenő paramétere nincs, tehát adó és vevő oldalon egyaránt alkalmazható (így nem szükséges továbbítani az összeköttetéseket).



10. ábra – Külső pontok hibás összeköttetése konkáv alakzat esetén

Az előző eljárással kapott hálózat általában sok olyan háromszöget tartalmaz, melyek a mozgáskompensáció szempontjából nem optimálisak (11.a. ábra). Az összeköttetések átrendezésével (11.b. ábra) előnyösebb, kevésbé hegyesszögű háromszögeket kaphatunk, melyek kedvezőbbek a mozgáskompensációhoz. A háromszögek torzultságának jellemzésére bevezettem a leghosszabb oldal és a másik két oldal összegének arányát (alaktényező). Nagyon hegyesszögű háromszögek esetén a hosszú oldal majdnem akkora, mint a másik kettő összege, így az alaktényező értéke majdnem 1. Ha a háromszög egyenlő oldalú, akkor a leghosszabb oldal feleakkora, mint a másik két oldal összege, ezért az alaktényező 0.5. Ezzel az egyszerű mérőszámmal minősíteni lehet az egyes háromszögeket torzultságuk alapján, azaz meg lehet találni a torz háromszögeket.