

Wavelet-transzformáció és JPEG2000

1. Bevezetés

1.1. Folytonos Wavelet-transzformáció

A Fourier-, a diszkrét Fourier- (DFT), és diszkrét koszinusz- (DCT) transzformációk harmonikus bázisfüggvényekre épülnek. Az állandósult állapotbeli, harmonikus bázisfüggvények tere alkalmas állandósult állapotbeli, periodikus jelek leírására, de a jeleken belüli tranziens változások, tehát a harmonikus komponensek - pl. időbeli - változásának leírására kevésbé alkalmas. Például ismeretes, hogy a DFT (FFT) frekvenciafelbontása $\Delta f = \frac{f_s}{N}$, míg az egy DFT ablakhoz tartozó keret-hossz, tehát az időfelbontás $\Delta t = N \frac{1}{f_s}$, ahol N a DFT pontszáma, f_s pedig a mintavételi frekvencia. A pontszámot növelve tehát javul a frekvenciafelbontás, de romlik az időbeli felbontás.

Mivel a Waveletek alapvető jellemzője a tranziens viselkedések jellemzésének képessége, ezért a Waveletre épülő transzformációk a harmonikus bázisfüggvényű transzformációk egy hatékony alternatíváját jelentik. Míg a folytonos Fourier transzformáció harmonikus bázisfüggvényei a $-\infty \leq t \leq \infty$ intervallumon konstans amplitúdóval értelmezettek, a wavelet bázisfüggvények gyorsan lecsengő függvények, melyek egy véges, esetleg végtelen intervallumon értelmezettek. Egy folytonos időbeli jel folytonos Fourier transzformáltja kizárólag a frekvencia egyváltozós, folytonos függvénye. Ezzel szemben egy folytonos jel folytonos Wavelet transzformáltja a *skála* (scale) és *eltolás* (shift) kétváltozós függvénye. A skála a jel frekvencia összetevőinek periódusidejével van kapcsolatban. Tehát a kis skála érték nagyfrekvenciás összetevőket, míg a nagy skála érték kisfrekvenciás összetevőket jellemez.

A Fourier transzformáció esetében a harmonikus összetevők frekvenciája nincsen kapcsolatban egymással, a spektrum folytonos $-\infty \leq f \leq \infty$ között. A Wavelet transzformáció esetében a wavelet bázisfüggvények skálázás, és eltolás (dilatáció) segítségével leszármaztathatók egy ún. anya-wavelet (prototípus) függvényből. Egy folytonos $f(t)$, $-\infty \leq t \leq \infty$, jel folytonos Wavelet transzformáltja (CWT), melyet jelöljünk $W(s, \tau)$ -vel, formailag a következőképpen definiálható:

$$W(s, \tau) = \int_{-\infty}^{\infty} f(t) \psi_{s, \tau}(t) dt, \quad (1)$$

ahol s a skála (scale), τ pedig az eltolás (shift). A $\psi_{s, \tau}(t)$ bázisfüggvények eltolással (dilatációval), és nyújtással/zsugorítással (kontrakció) származtathatók az anya-waveletből a következő módon:

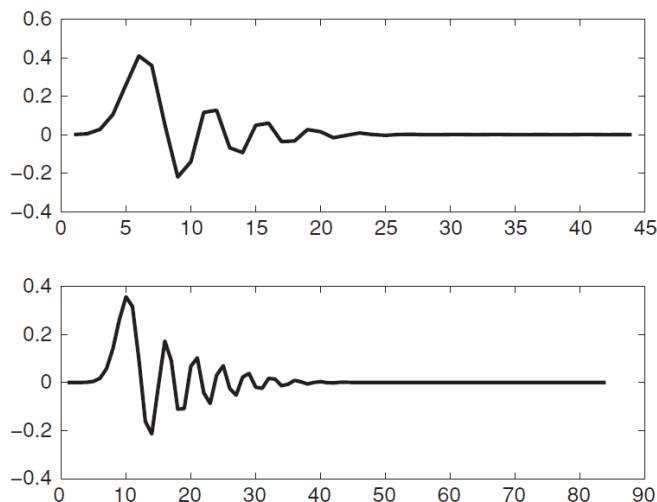
$$\psi_{s, \tau}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t - \tau}{s} \right) \quad (2)$$

A CWT esetében mind a skála, mind a shift paraméterek folytonos változók. A skála arányos a wavelet bázisfüggvény hosszával (pl. lecsengési idejével), tehát a nagy skála hosszú lecsengési idejű waveletet jelent, mely a jel kisfrekvenciás viselkedését jellemzi. A kis skálájú bázisfüggvények rövid lecsengési idejű függvények, melyek a jel nagyfrekvenciás (pl. tranziens) viselkedését jellemzik. Egyszerűen kifejezve tehát a wavelet transzformált skála és eltolás paramétere együttesen a jel adott skálával jellemezhető frekvenciakomponenseinek viselkedését írja le az eltolásnak megfelelő időpillanat környezetében. Az (2) egyenletből belátható, hogy az egy anya-waveletből leszármaztatott összes bázisfüggvény energiája egyenlő az anya-wavelet energiájával.

1.2. Wavelet-sorfejtés

Ha a Wavelet transzformáció skála és shift paraméterei nem folytonos, hanem diszkrét értékeket vesznek fel, akkor a folytonos $f(t)$ jel kifejezhető a Wavelet-sorfejtés segítségével. Ha a skála paraméter 2 egész hatványú értékeket veheti fel (bináris skálázás), és a shift paraméter egész értékeket vehet fel (diadikus transláció), akkor az $f(t)$ jel kifejezhető végtelen számú *skála-függvény* és *wavelet-függvény* összegeként:

$$f(t) = \sum_k a(j_0, k) \phi_{j_0, k}(t) + \sum_{j=j_0}^{\infty} \sum_{k=-\infty}^{\infty} d(j, k) \psi_{j, k}(t) \quad (3)$$



1. ábra. A Daubechies wavelet két különböző skálaparaméter értéke mellett

Az (3) kifejezés első tagja tartalmazza a $\phi_{j_0,k}$ bázisfüggvényt egy rögzített skála (j_0) mellett a shift paraméter minden lehetséges értékére. A $\phi_{j,k}$ bázisfüggvényeket *skála-függvényeknek* hívjuk, melyek egy prototípus skála-függvényből nyerhetők bináris skálázás és diadikus transláció segítségével:

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (4)$$

Ezek alapján (3) első tagja az $f(t)$ jel egy közelítése hasonlóan pl. a Fourier sor nulla frekvenciájú tagjához (átlagérték). A (3) kifejezés második tagja a *wavelet-függvények* végtelen összege, melyek szintén egy prototípus függvényből származtathatók bináris skálázás és diadikus transláció segítségével:

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (5)$$

Ebben az esetben is, a Fourier sorfejtés magasabb rendű harmonikusaihoz hasonlóan, a Wavelet-függvények az $f(t)$ jel nagyfrekvenciás, tranzien viselkedését jellemzik. A Wavelet-sorfejtésben, a (3) kifejezésben szereplő $a(j_0, k)$ és $d(j, k)$ együtthatók a folytonos Wavelet transzformáció (1) kifejezésének segítségével a következő módon határozhatók meg:

$$a(j_0, k) = 2^{j_0/2} \int_{-\infty}^{\infty} f(t) \phi_{j_0,k}(t) dt \quad (6)$$

$$d(j, k) = 2^{j/2} \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt \quad (7)$$

1.3. Diszkrét wavelet transzformáció (DWT)

Ha mind a jel, mind pedig a skála- és wavelet-függvények diszkrét idejű függvények, akkor a Wavelet-sorfejtés a Diszkrét Wavelet transzformációt (DWT) eredményezi. A Wavelet-sorfejtéshez hasonlóan a DWT két sorfejtésből áll: egy közelítő sorfejtésből, és egy részletező sorfejtésből. A DWT formális definíciója egy N hosszú, véges tartójú diszkrét jelre ($x[n]$, $0 \leq n \leq N - 1$):

$$DWT \{f(t)\} = W_\phi(j_0, k) + W_\psi(j, k), \quad (8)$$

ahol a skála (közelítő) együtthatók:

$$W_\phi(j_0, k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \phi_{j_0,k}[n], \quad (9)$$

a wavelet (részlet) együtthatók pedig:

$$W_\psi(j, k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] \psi_{j,k}[n] \quad j \geq j_0. \quad (10)$$

Az inverz DWT (rekonstrukció) pedig a következő módon írható fel:

$$x[n] = \frac{1}{\sqrt{N}} \sum_k W_\phi(j_0, k) \phi_{j_0,k}[n] + \frac{1}{\sqrt{N}} \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k) \psi_{j,k}[n] \quad (11)$$

Az (11) kifejezés második tagja egy végtelen sorösszeg, melyet a gyakorlatban véges számú tag összegeként közelítenek. A $j_0 = 0$ általában, míg N értéke 2 valamely hatványa.

1.4. Részsávós DWT kódolás

Tekintsünk egy egy dimenziós, N hosszúságú diszkrét jelet. A jel részsávós kódolása a jel többlépcsős szűrését és alulmintavételezését jelenti: az első szinten a jelet aluláteresztő, és feluláteresztő szűrőkkel szűrjük. A két szűrő kimeneteit 2:1 arányú alulmintavételezésnek vetjük alá, vagyis minden második mintát eldobunk. Tehát az első szint kimenetén két, $N/2$ hosszúságú szekvenciánk van, egyik a kisfrekvenciás, más a nagyfrekvenciás komponenseket tartalmazza. A második lépcsőben az első szint aluláteresztő szűrőjének kimenetét ismét ugyanazzal a szűrőpárossal szűrjük, és 2:1 arányban decimáljuk, tehát végeredményben a második szinten két $N/4$ hosszúságú szekvenciát kapunk. Ez a többlépcsős, részsávós szűrés-decimálás folytatható tetszőleges számú szintig, minden egyes szinten az aluláteresztő tagok kimenetének újabb részsávokra történő bontásával. Ezt a technikát *oktávósávós részsávós felbontásnak* hívják, mert minden szinten két részsávra bontjuk az előző szint kimenetét. Az eredeti jel részsávós felbontását a jelfeldolgozás *analízis szakaszának* hívjuk, az ezt megvalósító szűrőbankot pedig *analízis szűrőbanknak*. A *rekonstrukciós*, vagy más néven *szintézis szakaszban* az analízis szakasz kimeneti jeleit fordított "irányban" feldolgozva, egy ún. *szintézis szűrőbank* segítségével tudjuk az eredeti jelet, vagy annak egy közeleltését visszaállítani.

Rekonstrukció során tehát először az analízis szakasz utolsó szintjének aluláteresztő szűrőjének kimenetét 2:1 arányban interpoláljuk, majd szűrjük egy szintézis aluláteresztő szűrővel. Ugyanígy, az analízis szakasz utolsó szintjének feluláteresztő szűrőjének kimenetét 2:1 arányban interpoláljuk, majd szűrjük egy szintézis feluláteresztő szűrővel. E két szűrő kimenetét összeadjuk, majd az eljárás tovább folytatva, a megfelelő számú szintézis szakasz ismétlésével visszkapjuk az eredeti, N hosszúságú szekvenciát.

Az analízis szakasz 2:1 arányú decimálása, valamint a szintézis szakasz 2:1 arányú interpolálása során visszaállíthatóság szempontjából kritikus, hogy anti-aliasing, és "image"-spektrum mentes szűrési lépéseket végezzünk. Ennek érdekében az analízis és szintézis szűrőbank szűrőinek meghatározott követelményeket kell teljesíteniük. Ez ilyen típusú szűrőket általában QMF (Quadrature Mirror Filter) szűrőbankoknk hívjuk.

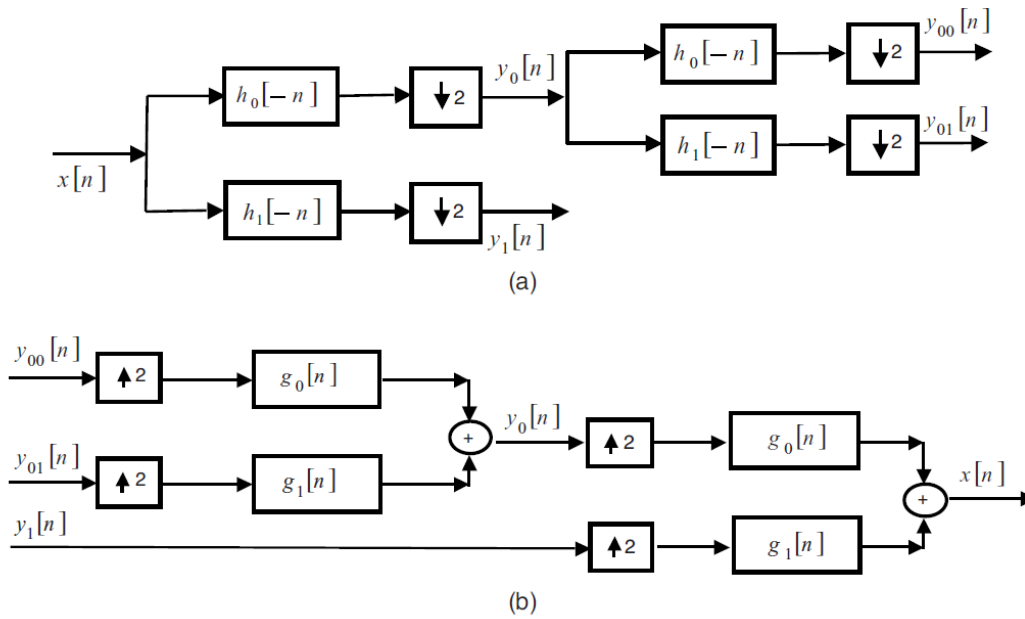
A DWT-re épülő részsávós kódolás esetében belátható, hogy a (9) és a (10) kifejezésekben található skála- és wavelet-függvények egymásba ágyazottak, vagyis egy j skála paraméterhez tartozó skála-, ill. wavelet-függvény kifejezhető a $j + 1$ skála paraméterhez tartozó skála, ill. wavelet függvények lineáris kombinációjaként:

$$\phi[2^j n - k] = \sum_l h_0[l - 2k] \sqrt{2} \phi[2^{j+1} n - l], \quad (12)$$

és

$$\psi[2^j n - k] = \sum_l h_1[l - 2k] \sqrt{2} \phi[2^{j+1} n - l], \quad (13)$$

ahol $l = m + 2k$, és $h_0[m]$ és $h_1[m]$ a lineáris kombinációk együtthatói.



2. ábra. Két szintű oktávsávós 1D DWT: (a) analízis szakasz, (b) szintézis szakasz

Ezek alapján belátható, hogy a DWT skála-, és wavelet együtthatóira is teljesül, hogy

$$W_\phi(j, k) = \sum_l h_0[l - 2k] W_\phi(j + 1, l) \quad (14)$$

$$W_\psi(j, k) = \sum_l h_1[l - 2k] W_\psi(j + 1, l) \quad (15)$$

A fenti kifejezések jelentősége az, hogy a j skála paraméterhez tartozó skála-együtthatók megkaphatók a $j + 1$ skálaparaméterhez tartozó együtthatóknak egy $h_0[-n]$ impulzusválaszú FIR szűrővel vett konvolúciójával, majd a szűrő kimenetének 2:1 arányú decimálásával, ahol $h_0[-n]$ a $h_0[n]$ impulzusválaszú FIR szűrő együtthatóinak időbeli fordítottja. Hasonlóképpen, a j skála paraméterhez tartozó wavelet-együtthatók megkaphatók a $j + 1$ skálaparaméterhez tartozó wavelet-együtthatók és egy $h_1[-n]$ impulzusválaszú FIR szűrő konvolúciójaként, majd a szűrő kimenetének 2:1 arányú decimálásával, ahol $h_1[-n]$ a $h_1[n]$ impulzusválaszú FIR szűrő együtthatóinak időbeli fordítottja. A szűrés, decimálás tetszőleges skála-szintig folytatható.

A jel visszaállítása során elindulunk a legmagasabb skála értékről, interpoláljuk 2:1 arányban a skála- és wavelet-együtthatókat, majd konvolváljuk őket a $g_0[n]$ és $g_1[n]$ impulzusválaszú FIR szűrőkkel. A két szűrő kimenetét összeadva megkapjuk a skála-együtthatókat egy alacsonyabb szinten. Az eljárást ismételve eljutunk a legalsó skála szintre, mely tulajdonképpen az eredeti jel.

1.5. Wavelet szűrőbankok

A jel Wavelet-filterbankokkal történő visszaállíthatóságának számos feltétele van. Ez első, hogy a $g_0[n]$ és $g_1[n]$ impulzusválaszú FIR szűrők átviteli karakterisztikájára teljesül, hogy

$$|G_0(f)|^2 + |G_1(f)|^2 = 2 \quad (16)$$

A gyakorlatban ezenkívül fajta DWT bír jelentőséggel:

1.5.1. Ortogonális DWT

Ortogonális DWT felbontás esetén a szűrőbankra teljesül, hogy

- A szűrő hossza L páros
- páronként teljesül $g_0[n]$ és $g_1[n]$ -re, $h_0[n]$ és $h_1[n]$ -re, $h_0[n]$ és $g_1[n]$ -re, és $g_0[n]$ és $h_1[n]$ -re a (16) feltétel
- $g_0[n]$ és $h_0[n]$ szűrők egymás időbeli fordítottjai
- $g_1[n]$ és $h_1[n]$ szűrők egymás időbeli fordítottjai
- Teljesül, hogy $h_1[n] = (-1)^{n+1}h_0[L-1-n]$
- Teljesül, hogy $g_1[n] = (-1)^{n+1}g_0[L-1-n]$
- Teljesül, hogy $\sum h_0[n] = \sum g_0[n] = 2$

A $h_0[n]$ és $g_0[n]$ szűrők aluláteresztő jellegűek, míg $h_1[n]$ és $g_1[n]$ feluláteresztő jellegűek. Ortogonális DWT szűrőbankkal nem lehet linárfázisú karakterisztikát megvalósítani. Ezért bír gyakorlati jelentőséggel a biortogonális DWT szűrőbank:

1.5.2. Biortogonális DWT

- A szűrő hossza L páros
- Mindegyik szűrő lineárfázisú (szimmetrikus együtthatókészlet)
- $g_0[n]$ és $h_0[n]$ szűrők nem egymás időbeli fordítottjai
- $g_1[n]$ és $h_1[n]$ szűrők nem egymás időbeli fordítottjai
- Teljesül, hogy $g_1[n] = (-1)^n h_0[n]$
- Teljesül, hogy $h_1[n] = (-1)^{n+1} g_0[n]$
- Teljesül, hogy $\sum h_0[n] = \sum g_0[n] = \sqrt{2}$

1.6. 2D DWT

A képtömörítési alkalmazások esetében az úgynevezett szeparábilis 2D DWT-t használják, mely esetben a 2D DWT egyszerűen és hatékonyan implementálható soronként, illetve oszloponként végzett 1D DWT-k egymásutánjaként.

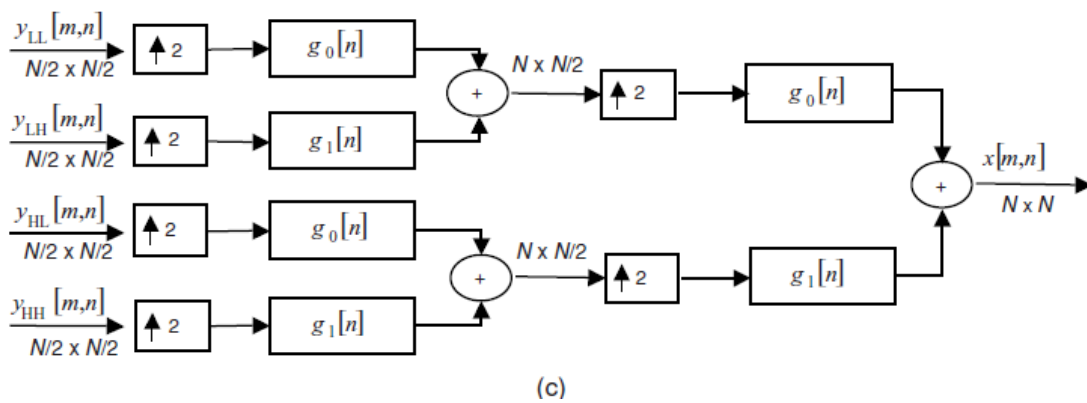
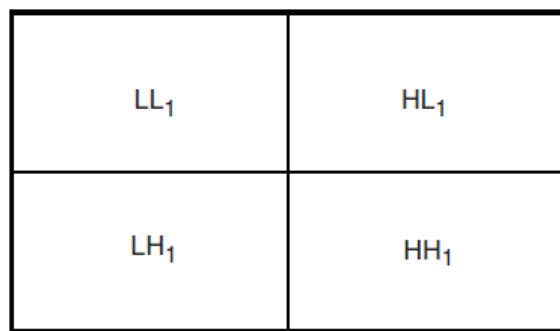
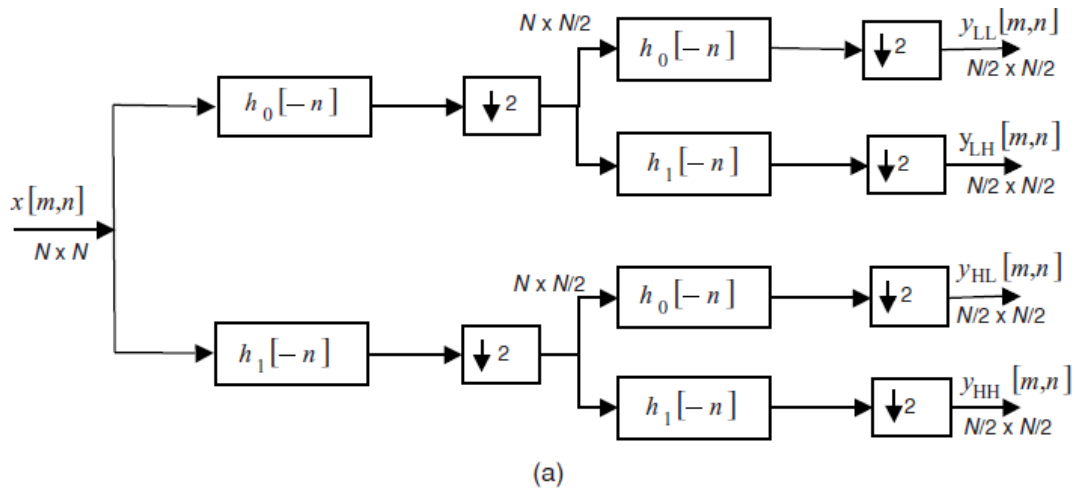
Vegyünk egy $N \times N$ -es méretű képet. Ha a kép sorait sorról sorra szűrjük a $h_0[-n]$ és $h_1[-n]$ szűrőegyütthatókkal, és minden szűrő kimenetét 2:1 arányban decimáljuk, akkor a szűrés után két $N \times N/2$ méretű együtthatóblokkot kapunk. Ezután, ha a $h_0[-n]$ szűrő kimenetének $N \times N/2$ -es együtthatóblokkját oszloponként szűrjük ismét a $h_0[-n]$ és $h_1[-n]$ szűrőegyütthatókkal, és szűrők kimenetét 2:1 arányban decimáljuk, akkor két $N/2 \times N/2$ együtthatóblokkot kapunk eredményül.

Ezután az előző szint $h_1[-n]$ szűrőjének kimenetének $N \times N/2$ -es együtthatóblokkját is oszloponként szűrjük ismét a $h_0[-n]$ és $h_1[-n]$ szűrőegyütthatókkal, és szűrők kimenetét 2:1 arányban decimáljuk, akkor két $N/2 \times N/2$ együtthatóblokkot kapunk eredményül.

Ekkor eljutottunk a kép első szintű 2D DWT felbontásához. A négy darab $N/2 \times N/2$ méretű együtthatóblokkot jelöljük rendre $y_{LL_1}, y_{HL_1}, y_{LH_1}, y_{HH_1}$ jelölésekkel. Az első helyen álló L, illetve H betű a sorirányú aluláteresztő ($h_0[n]$), illetve feluláteresztő ($h_1[n]$) szűrést, míg a második helyen álló álló L, illetve H betű az oszlopirányú aluláteresztő, illetve feluláteresztő szűrést jelzi.

Egy második szintű DWT számításához az y_{LL_1} együtthatóblokk újabb 2D DWT transzformációjával jutunk. Ebben az esetben fa struktúrájú 2D DWT-ről beszélünk. Ha minden szinten nem csak a y_{LL} együtthatókat, hanem mind a négy együtthatóblokkot 2D DWT transzformáljuk, akkor a teljes fa struktúrájú 2D DWT felbontást kapjuk.

A kép visszaállítása az 1D DWT-hez hasonlóan, a legmagasabb szintről indulva, a szintézis szűrőbankok felhasználásával történik.



3. ábra. Egy szintű 2D DWT számítása részsávós felbontással: (a) 2D DWT, (b) együtthatók elhelyezkedése, és (c) inverz 2D DWT.

1.7. JPEG2000

A transzformációs képtömörítési eljárások célja a pixelek közötti korreláció csökkentése egy alkalmasan megválasztott transzformált tartományban, amelyben az együtthatók közötti korreláció kisebb, és a jelenergia nagy részét néhány (kis számú) együttható reprezentálja. Ha ez teljesül, akkor a transzformáció a jelenergia szempontjából kompakt. Az alapvető különbség a hagyományos JPEG, és a hasonló blokk alapú DCT transzformációt használó eljárások (pl. MPEG) és a wavelet alapú JPEG2000 tömörítés között, hogy a 2D DWT transzformációt jellemzően a kép egészére alkalmazzák, és nem a kép $N \times N$ -es blokkjaira. A 2D DWT egy szintjén tehát négy olyan együtthatóblokkot kapunk, mely az egész kép adott szintjéhez tartozó kisfrekvenciás közelítést, valamint a vízszintes, függőleges, és diagonális irányú részleteket (nagyfrekvenciás tartalmat) jellemzi,

A JPEG2000 tömörítés során a következő fő lépések végrehajtása történik:

- Offszet, és komponens transzformáció: 8 bites komponens ábrázolás esetén a komponenseket $+128$ értékű offszettel írják le, majd komponens transzformáció következik. A komponens transzformáció lehet irreverzibilis, mely általában a hagyományos YCbCr transzformációt jelenti, és kerekítési hibákat visz a képi adatba. A reverzibilis komponens transzformáció (ICT) esetén az YCbCr komponensek a következők:

$$Y = \left\lfloor \frac{R + 2G + B}{4} \right\rfloor; C_B = B - G; C_R = R - G;$$
$$G = Y - \left\lfloor \frac{C_B + C_R}{4} \right\rfloor; R = C_R + G; B = C_B + G. \quad (17)$$

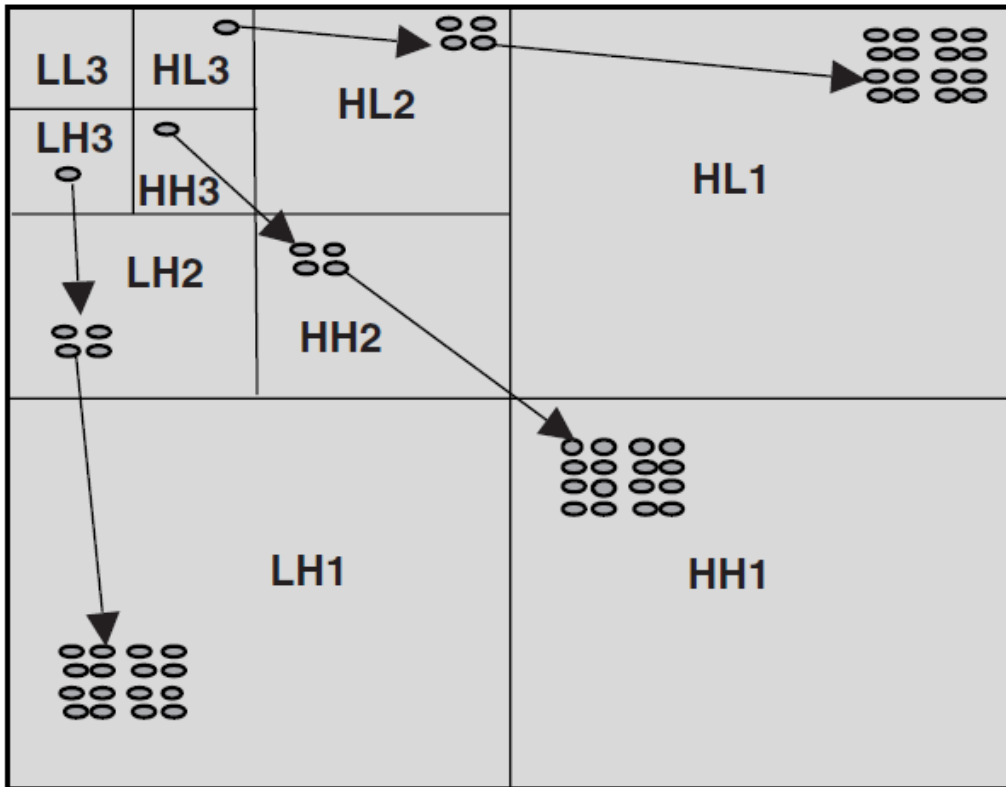
Ez a transzformáció nem jár kerekítési hibákkal, így a transzformáció reverzibilis, ugyanakkor a komponensek gyakorlatilag világosság, illetve színkülönbségi komponensek.

- Szelekekre (tile) bontás. A JPEG2000 esetén lehetőség van a kép szelekekre bontására, melyeken eltérő paraméterű tömörítés, skálázható minőség, stb. alkalmazható. A legtöbb esetben azonban az egész kép egy szeletben történő tömörítését használják
- DWT transzformáció, mely lehet irreverzibilis, vagy reverzibilis. Az irreverzibilis DWT az ún. CDF (Cohen-Daubechies-Feauveau) 9/7 -es waveletet, míg a reverzibilis DWT a CDF 5/3 kerekített, biortogonális waveletet alkalmazza. Az irreverzibilis wavelet kvantálási hibát eredményez, ezért veszteséges.
- Kvantálás
- Kvantált együtthatók leképzése (Zero-tree coding) és entrópiakódolás

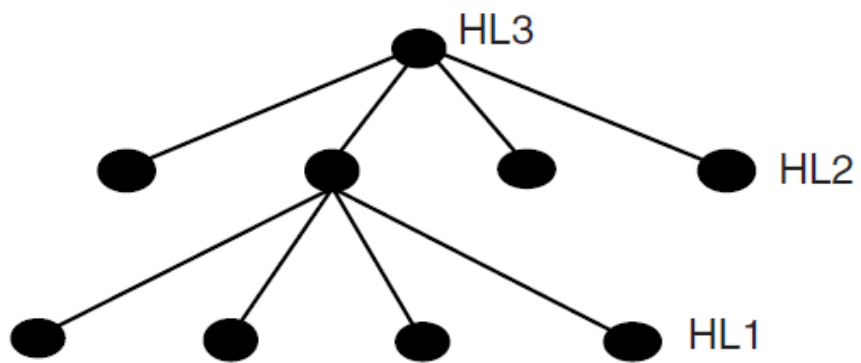
1.7.1. Zero-tree kódolás (EZW)

A "zero-tree" a nulla együtthatók fa struktúrájú elhelyezkedését jelenti. Az oktávsávós 2D DWT kódolás esetén az LL sáv minden szinten 4 részsávra osztódik, mely 4 részsáv a következő szintet alkotja. (quad-tree struktúra) Tehát, bármely szinten az LH, HL és HH részsávok bármely együtthatójához 4 együttható tartozik az eggyel alacsonyabb szinten. Ha egy LH, HL, vagy HH együttható a legmagasabb szinten nulla, vagy elhanyagolható nagyságú, akkor nagy valószínűséggel az azonos részsávhoz tartozó, alacsonyabb szinteken vele kapcsolatban álló együtthatók is nullák vagy elhanyagolhatóak lesznek.

Az EZW kódolás a fa struktúra hatékony entrópiakódolását jelenti, melynek korszerűbb, továbbfejlesztett változatait használják a JPEG2000 tömörítésben is. A legmagasabb szint LH, HL és HH együtthatói az együttható-"családfa" szüleinek tekinthetők, és az eggyel alacsonyabb szint 4 együtthatói a "gyermek". Az EZW kódolás során az együtthatók tárolása szukcesszív approximációval ("fokozatos közelítéssel") iteratíven történik, melynek előnye, hogy tetszőleges iterációs szinten a folyamat leállítható, ha az adatok tárolására szánt bitmennyiség kimerült. Az algoritmus vázlatosan:



(a)



(b)

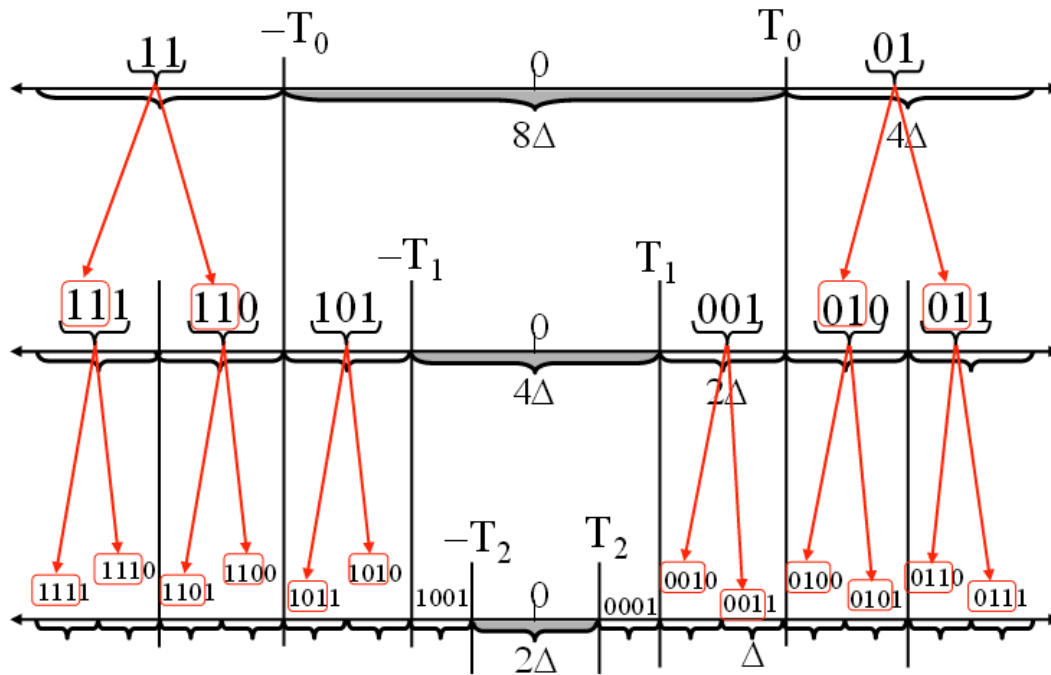
4. ábra. A 2D DWT fa-struktúrájú ábrázolása

63	-34	49	10	7	13	-12	7
-31	23	14	-13	3	4	6	-1
15	14	3	-12	5	-7	3	9
-9	-7	14	8	4	-2	3	2
-5	9	-1	47	4	6	-2	2
3	0	-3	2	3	-2	0	4
2	-3	6	-4	3	6	3	6
5	11	5	6	0	3	-4	4

5. ábra. Együtthatók leképzése, és kvantálási sorrend az EZW esetében

- a kép többszintű 2D DWT-jének kiszámítása
- A DWT együtthatók dinamikatartományának meghatározása
- T_0 küszöbszint beállítása a dinamikatartomány közepére
- Iteráció kezdete : minden következő iterációban a küszöbszintet a korábbi iteráció küszöbszintjének felére vesszük ($T_{j+1} = T_j/2$)
 - Kvantálás: T_j küszöbszint alatt ún. "dead-zone", azaz mindent nullára kvantálunk
 - Kvantálás után a kvantált együtthatókhoz szimbólumok rendelése a következők szerint (*domináns lépés*):
 - * Ha nulla, és minden belőle leszármazott együttható nulla : ZTR (zero-tree root)
 - * Ha az együttható nulla, de nem nulla (szignifikáns) leszármazottjai vannak: ISO (izolált nulla)
 - * Pozitív szignifikáns (POS)
 - * Negatív szignifikáns (NEG)
 - Két lista kezelése : *domináns* lista, és *szubordináns* lista
 - * Domináns lista: Adott iterációban még nem szignifikáns együtthatók listája
 - * Szubordináns lista: szignifikáns együtthatók kódolt értéke
 - *Szubordináns lépés*: Szignifikáns listáról az együtthatók következő legmagasabb bitsíkjának tárolása (első iterációban az MSB)
- Küszöbszint felezése, új iteráció
- Ha a felezett küszöbszint mellett a korábbi küszöbszinthez képest az együttható szignifikáns-sá válik, akkor eltávolítás a domináns listáról, és áthelyezés a szubordináns listára

Az EZW kódolás során az együtthatók vizsgálata minden iterációban azonos sorrendben (cikk-cakk) leképzésben történik. A kvantálási küszöbök felezése miatt, az új iterációban mindig csak a következő bitsík tárolása szükséges.



6. ábra. Szukcesszív approximációs kvantálás, és a kvantálási intervallumok alakulása az EZW iterációkban

1.8. DCP - Digital Cinema Package

A digitális mozis disztribúciós formátuma, a DCP a JPEG2000-re épülő kizárólag intra-frame tömörítést használja a következő okok miatt: A DCT-re épülő blokk alapú kódolást alkalmazó MPEG kodek család (MPEG-1, MPEG-2, H264/AVC, HEVC, stb.) blokkosodási jelenségeket eredményez a blokkok határán, mely a blokkosodást gátló szűrők alkalmazása mellett is (H264, HEVC) láthatóvá válhat a képen - ez a moziban nem engedhető meg. Ezenkívül az inter-frame, tehát képek közötti korrelációt is csökkentő prediktív kódolás (P és B képek) a hirtelen, gyors mozgásoknál, képváltásoknál a mozgás leképzés romlását eredményezheti, illetve ismét blokkosodást eredményezhet, mely a moziban szintén nem engedhető meg.

Az SMPTE által elfogadott digitális film disztribúciós csomag (DCP) formátumának főbb paraméterei:

- JPEG2000 alapú intra-frame tömörítés
- 24, 25, 30, 48, 50, és 60 Hz képfrekvencia támogatás 2K felbontás mellett
- 24, 25, és 30 Hz képfrekvencia támogatás 4K felbontás mellett
- 12 bit/komponens ábrázolás
- Tisztán exponenciális 2.6 gamma értékű gamma-korrekció
- DCI-P3, vagy CIE XYZ színtér
- Maximum bitráta: 250 Mbps
- Felbontások:
 - Flat (1998x1080 vagy 3996x2160), 1.85:1
 - Scope (2048x858 vagy 4096x1716), 2.39:1
 - (U)HDTV (1920x1080 vagy 3840x2160), 16:9
 - Full (2048x1080 vagy 4096x2160) 1.9:1

2. Ellenőrző kérdések

- Mi a kapcsolat a DFT frekvencia és időfelbontása között ? (képlet)
- Milyen két paramétere van egy wavelet bázisfüggvénynek, ezek mit fejeznek ki (szemléletesen) ?
- Rajzoljuk le vázlatosan az oktávsávós részsávós kódolás analízis, illetve rekonstrukciós szakaszát.
- Mi az alapvető különbség a JPEG és JPEG2000 transzformációs kódolása között ?

3. Feladatok

3.1. 1D DWT

Vonatkozó MATLAB script: `dwt_1D_example.m`

A script segítségével egy harmonikus, valamint tranziens komponensekből álló 1D diszkrét jel 2 szintű DWT-jét vizsgálhatjuk.

- Plotoljuk ki a jel harmonikus (x1) és tranziens (x2) komponenseit.
- Vizsgáljuk meg az analízis és szintézis FIR szűrőket mind időtartományban, mind frekvenciatartományban
- Vizsgáljuk meg az első-, valamint a második szintű skála (approximation), és wavelet (detail) együtthatókat. Figyeljük meg a jel tranzienseinek megjelenési formáját az együtthatókban. Hol, és hogyan jelennek meg, és miért ott ?
- A kódban szerepelő db8 ortogonális vagy biortogonális wavelet ?

3.2. 2D DWT

Vonatkozó MATLAB script: `dwt_2d_example.m`

A script segítségével a cameraman ábra két szintű 2D DWT felbontását tekinthetjük meg.

- Az LH,HL és HH részsávok közül mely felelős a vízszintes, a függőleges, és az átlós irányú részletekért ?

3.3. MATLAB WAVEMENU

A wavemenu begépelésével egy szerteágazó grafikus toolbox segítségével vizsgálhatjuk a wavelet transzformációt, illetve az arra épülő tömörítési eljárásokat.

- A *One Dimensional/Wavelet-1D* gomb segítségével különböző jelekre vizsgálhatunk többszintű wavelet felbontás eredményeit:
 - A File/Example analysis/Basic signals alatt "előregyártott" példa jelekre vizsgálhatjuk az 1D DWT felbontást
 - Vizsgáljuk meg a Triangle+Sine és az Electrical consumption jelet db Wavelet 2,4, és 8 szintű felbontása mellett
 - Értelmezzük a Display mode/Full decomposition at Levels ... szintek állítása esetén az együtthatókat
 - Vizsgáljuk meg az együtthatók hisztogramjait

- Értelmezzük az eredményeket
- *A Specialized Tools/True compression 2D* gomb használatával vizsgálhatjuk pl. az EZW tömörítés működését. Töltsünk be egy képet, értelmezzük a grafikus felületet.
 - Mi a hatása és szerepe a Nb. of encoding loops beállításnak ?
 - Milyen hatása van annak, hogy hány szintű DWT felbontás történik ?
 - Használjuk az Inspect Wavelet tree gombot. Próbáljuk meg értelmezni a látottakat.

3.4. JPEG és JPEG2000 összehasonlítása

Vonatkozó MATLAB script : `JPEG_vs_JPEG2000.m`

A kód segítségével előre beállított minőségi lépcsőket definiálhatunk: a minőségi szintet a JPEG esetében a `Q_JPEG` definiálja 1 és 100 között, a JPEG2000 esetében a tömörítési arányt a `CR_JPEG2000`, melynek értéke egynél nagyobb (pl. `CR_JPEG2000=10` esetén a cél fileméret 10%-a az eredetinek.) A JPEG esetében a minőségi szint nem jelent közvetlenül bitsebességvezérlést.

A kód a beállított lépcsőkkel elvégzi a bemeneti kép kódolását, és az eredményezett fileméret függvényében megjeleníti a PSNR, valamint az SSIM értékeket. Szintén információt kapunk a kódolási, illetve dekódolási időkről.

Ezen kívül lehetőség van a forrásfájl egy adott, megközelítőleg azonos fileméretet eredményező JPEG és JPEG2000 tömörítéssel kódolt változatainak összehasonlítására.