

# Képfeldolgozási módszerek

egyetemi jegyzet

Készítette: Fazekas Gábor, Hajdu András

Debreceni Egyetem,  
Informatikai Intézet,  
Debrecen, 2004.



## Köszönetnyilvánítás

A szerzők köszönetüket fejezik ki Tóth Tamás programtervező matematikus hallgatónak a hisztogram alapú technikák kidolgozásában való közreműködéséért.

# Tartalomjegyzék

<b>Előszó</b>	<b>1</b>
<b>1. Képjavítás</b>	<b>2</b>
1.1. Bevezetés	2
1.2. Hisztogram alapú eljárások	3
1.2.1. Küszöbölési technikák	4
1.2.2. Hisztogram-transzformáció	6
1.2.3. Hisztogram-kiegyenlítés	7
1.2.4. Lokális hisztogramkiegyenlítés	10
1.2.5. Multispektrális hisztogram	10
1.3. Éldetektálás	12
1.3.1. Gradiens módszerek	13
1.3.2. Laplace operátor	16
1.3.3. A második derivált tengelymetszetei	16
1.3.4. Képek élesítése	17
1.3.5. Canny éldetektálás	19
1.3.6. Parametrikus élmódellek	20
1.3.7. Élek multispektrális képeken	21
1.3.8. Élkapcsolás, élvékonyítás	21
1.3.9. Sarokdetektálás	22
1.3.10. Felüláteresztő szűrők	23
1.4. Simítás	25
1.4.1. Környezeti átlagolás	25
1.4.2. Medián szűrés	26
1.4.3. Intervallum módszer	27
1.4.4. Átlagolás inverz gradiens alapján	28
1.4.5. Nem lineáris középszűrők	28
1.4.6. Képsorozat átlagolása	29
1.4.7. Aluláteresztő szűrők	30
1.4.8. Sávszűrés	31

<b>2. Maszkolás</b>	<b>33</b>
2.1. Bevezetés . . . . .	33
2.2. Alakzatok detektálása maszkokkal . . . . .	33
2.2.1. Pontdetektálás . . . . .	34
2.2.2. Vonaldetektálás . . . . .	35
2.2.3. Éldetektálás . . . . .	35
2.2.4. Kombinált maszkrendszerek . . . . .	36
2.3. Matematikai morfológia . . . . .	39
2.4. Dilatáció és erózió . . . . .	40
2.4.1. Nyitás és zárás . . . . .	42
2.4.2. Hit-Miss transzformáció . . . . .	43
2.4.3. Kontúr kinyerése . . . . .	45
2.4.4. Lyukak kitöltése . . . . .	45
2.4.5. Összefüggő komponensek meghatározása . . . . .	45
2.4.6. Konvex burok meghatározása . . . . .	46
2.4.7. Vékonyítás . . . . .	48
2.4.8. Vastagítás . . . . .	49
2.4.9. Vázkijelölés . . . . .	50
2.4.10. Kinövések eltávolítása . . . . .	54
2.4.11. Golay ábécé . . . . .	54
<b>Irodalomjegyzék</b>	<b>55</b>

# Táblázatok jegyzéke

1.1. $64 \times 64$ -es, 8-szintű kép intenzitáshistogramjának kiegyenlítése. . . . .	10
2.1. A Golay ábécé elemei. . . . .	56

# Ábrák jegyzéke

1.1. 256-szintű (8 bites) szürkeskálás kép hisztogramja; (a) eredeti kép, (b) intenzitáshisztogram. . . . .	4
1.2. Két szintre vágás; (a) eredeti kép, (b) intenzitáshisztogram és választott küszöbérték, (c) binarizált kép. . . . .	5
1.3. Sávkimelés; (a) eredeti kép, (b) intenzitáshisztogram megfelelő küszöbértékekkel, (c) sávkimelés eredménye. . . . .	6
1.4. Szürkeárnyaltos transzformációs függvény. . . . .	7
1.5. Sűrűségfüggvény (hisztogram) kiegyenlítése folytonos esetben; (a) eredeti sűrűségfüggvény, (b) transzformációs függvény, (c) kiegyenlített sűrűségfüggvény. . . . .	9
1.6. Szürkeárnyaltos kép hisztogram-kiegyenlítése; (a) eredeti kép, (b) intenzitáshisztogram, (c) kiegyenlített kép, (d) intenzitáshisztogram. . . . .	11
1.7. Az RGB színekocka; (a) a színekocka felszíne, (b) gömbök a színekocka belsejében. . . . .	12
1.8. Multispektrális kép hisztogram-kiegyenlítése; (a) eredeti kép, (b) kiegyenlítés komponensenkénti transzformációs függvényekkel, (c) kiegyenlítés a transzformációs függvények átlagolásával. . . . .	13
1.9. Az $f(x)$ intenzitásfüggvény első és második deriváltjának viselkedése; (a) sötét-világos átmenet esetén, (b) egy lassabb világos-sötét átmenet esetén. . . . .	13
1.10. Digitális gradiens számolása; (a) Sobel gradiens, (b) Roberts gradiens. . . . .	15
1.11. A gradiensoperátor eredménye; (a) eredeti kép, (b) Sobel gradiens, (c) Roberts gradiens. . . . .	15
1.12. Kettős élék a Laplace operátor használatakor; (a) eredeti kép, (b) Laplace operátor hatása. . . . .	16
1.13. A <b>LoG</b> operátor maszkokkal való közelítései; (a) $5 \times 5$ -ös maszk, (b) $13 \times 13$ -ös maszk. . . . .	18
1.14. A <b>LoG</b> operátor hatása; (a) eredeti kép, (b) $5 \times 5$ -ös maszk esetén, (c) $13 \times 13$ -ös maszk esetén. . . . .	18
1.15. Képesítés $3 \times 3$ méretű maszkokkal; (a) enyhén élesítő maszk, (b) fokozottan élesítő maszk, (c) eredeti kép, (d) enyhe élesítés, (e) fokozott élesítés. . . . .	19
1.16. $3 \times 3$ -as maszkrendszer lokális élkapcsoláshoz. . . . .	22
1.17. $3 \times 3$ -as maszkrendszer bináris élvékonyításhoz. . . . .	22
1.18. Az ideális felüláteresztő szűrő; (a) síkkal való metszete, (b) teljes alakja. . . . .	24

1.19. Az ideális felüláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép $K = 20$ esetben, (c) szűrt kép $K = 40$ esetben. . . . .	24
1.20. A Butterworth felüláteresztő szűrő ( $t = 2$ paraméterű eset); (a) síkkal való metszete, (b) teljes alakja. . . . .	24
1.21. A Butterworth felüláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép $t = 2$ , $K = 50$ esetben, (c) szűrt kép $t = 2$ , $K = 80$ esetben. . . . .	25
1.22. Környezeti átlagolás $3 \times 3$ méretű maszkkal; (a) konvolúciós maszk, (b) eredeti kép, (c) a maszk hatása, (d) a maszk hatása 10 iterációs lépés után. . . . .	26
1.23. Környezeti átlagolás $5 \times 5$ méretű maszkkal; (a) konvolúciós maszk, (b) eredeti kép, (c) a maszk hatása, (d) a maszk hatása 10 iterációs lépés után. . . . .	26
1.24. Medián szűrés hatása; (a) eredeti (15%-os fehér zajjal terhelt) kép, (b) a medián szűrés eredménye $3 \times 3$ -as maszk esetén, (c) a szűrés eredménye 10 iterációs lépés után, (d) $5 \times 5$ -ös medián szűrés eredménye kiegészítő küszöbölést használva. . . . .	27
1.25. Intervallum módszer képsimításhoz; (a) eredeti kép, (b) a szűrés eredménye $3 \times 3$ -as maszk esetén, (c) a szűrés eredménye $5 \times 5$ -ös maszk esetén. . . . .	28
1.26. Képsorozat átlagolása; (a) - (c) (zajos) képsorozat elemei, (d) a képsorozat átlaga. . . . .	30
1.27. Az ideális aluláteresztő szűrő; (a) síkkal való metszete, (b) teljes alakja. . . . .	30
1.28. Az ideális aluláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép $K = 25$ esetben, (c) szűrt kép $K = 60$ esetben. . . . .	31
1.29. A Butterworth aluláteresztő szűrő ( $t = 2$ paraméterű eset); (a) síkkal való metszete, (b) teljes alakja. . . . .	31
1.30. A Butterworth aluláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép $t = 2$ , $K = 10$ esetben, (c) szűrt kép $t = 2$ , $K = 20$ esetben. . . . .	32
1.31. Ideális sávszűrő; (a) síkkal való metszete, (b) teljes alakja. . . . .	32
2.1. Maszkolás a képtartományban $3 \times 3$ méretű maszkkal; (a) a vizsgált képrészlet adott $x_i$ intenzitásokkal, (b) a használt maszk, (c) vektoros felírás a $\langle W^T, X \rangle$ maszkválasz meghatározásához. . . . .	34
2.2. Izolált- és határpontok meghatározása maszkkal; (a) detektáló maszk, (b) eredeti kép, (c) maszkolás eredménye. . . . .	34
2.3. Vonaldetektáló maszkrendszer a négy fő irányban. . . . .	35
2.4. Vonalak meghatározása maszkokkal; (a) vízszintes vonalak, (b) függőleges vonalak, (c)-(d) átlós vonalak. . . . .	35
2.5. Éldetektáló $3 \times 3$ -as maszkrendszer a négy fő irányban. . . . .	36
2.6. Éldetektálás $3 \times 3$ méretű maszkokkal; (a) eredeti kép, (b) vízszintes élek, (c) függőleges élek, (d) a vízszintes és függőleges élképek átlaga. . . . .	36
2.7. Laplace operátor konvolúciós maszkjai és azok hatása; (a) $L_4$ maszk, (b) $L_8$ maszk, (c) $5 \times 5$ -ös Laplace maszk, (d) eredeti kép, (e) $L_4$ maszk hatása, (f) $L_8$ maszk hatása, (g) $5 \times 5$ -ös maszk hatása. . . . .	37
2.8. $3 \times 3$ -as Laplace maszkok speciális súlyozással. . . . .	37



2.9. Kombinált maszkrendszer 3D-ben két tulajdonságaltérrel. . . . .	38
2.10. Három tulajdonság egyidejű detektálására alkalmas kilenc elemű maszkrendszer. . . . .	39
2.11. A dilatació eredménye; (a) eredeti $A$ alakzat, (b) adott $B$ struktúraelem, (c) az $A$ dilatáltja, (d) másik $B$ struktúraelem, (e) az $A$ dilatáltja erre nézve. . . . .	41
2.12. Az erózió eredménye; (a) eredeti $A$ alakzat, (b) adott $B$ struktúraelem, (c) az $A$ erodáltja, (d) másik $B$ struktúraelem, (e) az $A$ erodáltja erre nézve. . . . .	42
2.13. A morfológiai nyitás és zárás kombinálása. . . . .	44
2.14. Kontúr kinyerése morfológiai műveletekkel; (a) eredeti $A$ alakzat, (b) a felhasznált $B$ struktúraelem, (c) az $A$ halmaz $\partial A$ kontúrja. . . . .	45
2.15. Lyukak kitöltése morfológiai műveletekkel; (a) eredeti $A$ alakzat, (b) a felhasznált $B$ struktúraelem, (c) a kitöltő eljárás végeredménye, (d) az iteráció lépései. . . . .	46
2.16. Összefüggő komponensek keresése morfológiai műveletekkel; (a) eredeti $Y$ komponens kiindulóponttal, (b) a felhasznált $B$ struktúraelem, (c) az eljárás végeredménye, (d) az iteráció lépései. . . . .	47
2.17. Konvex burok meghatározása morfológiai műveletekkel; (a) eredeti $A = X_0^i$ alakzat, (b) a négy fő irányhoz tartozó struktúraelem, (c) az egyes struktúraelemek szerepe a $C(A)$ előállításában (d) a $C(A)$ konvex burok, (e) a négy irányhoz tartozó $D^1 = X_5^1, D^2 = X_1^2, D^3 = X_5^3, D^4 = X_1^4$ konvergenciahalmaz. . . . .	48
2.18. Vékonyítás morfológiai műveletekkel; (a) struktúraelemek, (b) eredeti $A$ alakzat, (c)-(j) a $B^1, \dots, B^8$ struktúraelemekkel lépésenként végzett vékonyítás (csak a változások feltüntetésével). . . . .	49
2.19. Vastagítás a komplementer vékonyításával; (a) eredeti $A$ alakzat, (b) az $A$ komplementere, (c) az $A^c$ vékonyításának végeredménye, (d) az $A^c$ vékonyításának komplementere, (e) az $A$ vastagításának eredménye az izolált részek eltávolításával. . . . .	50
2.20. A középtengely, mint a maximális sugarú körök középpontjai; (a)-(d) egyszerű alakzatok középtengelyei. . . . .	51
2.21. A vázkijelölő vékonyító eljárásban használt szomszédság. . . . .	52
2.22. Vázkijelölés vékonyítással; (a) eredeti alakzat, (b) az első iterációs lépés, (c) a második iterációs lépés, (d) végeredmény. . . . .	53
2.23. Kinövések eltávolítása morfológiai műveletekkel; (a) eredeti $A$ váz, (b) az eljárásban használt struktúraelemek, (c) az $A$ vékonyításának eredménye, (d) az $A$ vékonyításának végpontjai, (e) a végpontok dilatálása a törölt vázszegmensek visszaállításához, (f) az eljárás végeredménye. . . . .	55

# Előszó

Ez a jegyzet a "Képfeldolgozási módszerek" című tárgyhoz készült elsősorban programozó és programtervező matematikus szakos hallgatók számára. Az anyag tartalmát tekintve részben ráépül a "Képfeldolgozás matematikája" című tárgy tematikájára, így az annak keretében elsajátítható ismereteket ismertnek tételezi fel. A jegyzet fő célja a képfeldolgozási feladatokban jól használható alapvető elméleti eredmények bemutatása, valamint ezek jellegzetességeinek gyakorlati példákon keresztül való szemléltetése.

A szerzők

# 1. fejezet

## Képjavítás

### 1.1. Bevezetés

A nem ideális körülmények között készült képi felvételek gyakran zajosak, homályosak, túlzottan vagy kevéssé kontrasztosak, egyenlőtlenül megvilágítottak lehetnek. Hasonló problémák léphetnek fel a nem megfelelő digitalizálás eredményeképpen is. A képjavító eljárásokkal ezeket a képeket az emberi szemlélő számára kellemesebbé alakíthatjuk. A képfeldolgozási feladatokban a képjavításokat előfeldolgozó lépésként is végrehajthatjuk, hogy a később alkalmazott technikákkal jobb eredményeket érjünk el. A képjavító eljárásokat alapvetően aszerint osztályozhatjuk, hogy a kép- vagy a frekvenciatartományban (Fourier tartományban) alkalmazhatóak-e.

Az egyik leggyakrabban használt képjavító és képmanipuláló eljárások a szűrők. Szűrést a kép- és frekvenciatartományban egyaránt végezhetünk. A 2D-s képtartománybeli szűrés általános modellje a

$$g(x, y) = f(x, y) * h(x, y) = \int_{\mathbb{R}^2} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta$$

konvolúciós modell, ahol  $f$  az eredeti intenzitáskép,  $g$  az eredménykép,  $h$  pedig a szűrőfüggvény. Digitális képek esetén a szűrőfüggvényt ún. konvolúciós maszkokkal közelítik – az eljárást részletesen a 2. fejezetben ismertetjük. A frekvenciatartománybeli szűrésnek a jól ismert konvolúciós tétel adja az alapját, így a szűrés

$$G(u, v) = F(u, v)H(u, v)$$

módon történik, ahol  $F$ ,  $G$  és  $H$  jelöli rendre az  $f$ ,  $g$  és  $h$  függvény Fourier transzformáltját. Mivel a frekvenciatartománybeli szűrés a képtartománybeli konvolúcióval szemben a kevésbé műveletigényes szorzást használja, így a gyors és inverz gyors Fourier transzformáció segítségével sok esetben hatékonyabb annál. Amennyiben a  $h$  függvény torzításként lép fel a képen, úgy a képrekonstrukció problémaköréhez jutunk, amelyet külön fejezetben tárgyalunk részletesen.

Az alábbiakban a képjavitáshoz szükséges elemek detektálásának módszereit, valamint a legnépszerűbb képjavitó eljárások alapjait tekintjük át.

## 1.2. Hisztogram alapú eljárások

A képfeldolgozási eljárásokban a képek intenzitáshisztogramja alapvető jelentőséggel bír. Először a szürkeárnyalatos képekhez fűződő elméletet ismertetjük, ennek kiterjesztésével színes képekre pedig az 1.2.5. szakaszban foglalkozunk.

Jelöljön  $r$  egy valós értékű valószínűségi változót (az intenzitást), és tekintsünk ennek egy  $r_1, r_2, \dots, r_M$  realizációját. Mivel a hisztogram az elméleti sűrűségfüggvény diszkrét közelítése, ezért osszuk fel a számegyenest  $y_0 < y_1 < \dots < y_k$  osztópontokkal, úgy hogy  $y_0 \geq \min_{j=1}^M \{r_j\}$  és  $y_k \leq \max_{j=1}^M \{r_j\}$ . Jelölje  $n_i$  az  $[y_{i-1}, y_i)$  intervallumba eső mintaelemek számát ( $i = 1, \dots, k$ ), majd rajzoljunk az  $[y_{i-1}, y_i)$  intervallum fölé az  $n_i$ -vel arányos területű téglalapot. Az így kapott alakzatot az  $r_1, r_2, \dots, r_M$  minta  $y_0 < y_1 < \dots < y_k$  beosztáshoz tartozó hisztogramjának nevezzük.

Ha a megrajzolt téglalapok összterülete  $M$ , akkor gyakorisági hisztogramhoz jutunk. Pontosabban, a gyakorisági hisztogram az az  $f_n$  valós függvény, amelyre

$$f_n(x) = \begin{cases} \frac{\nu_i}{(y_i - y_{i-1})}, & \text{ha } x \in [y_{i-1}, y_i), \quad i = 1, \dots, k \\ 0, & \text{ha } x \notin [y_0, y_k). \end{cases}$$

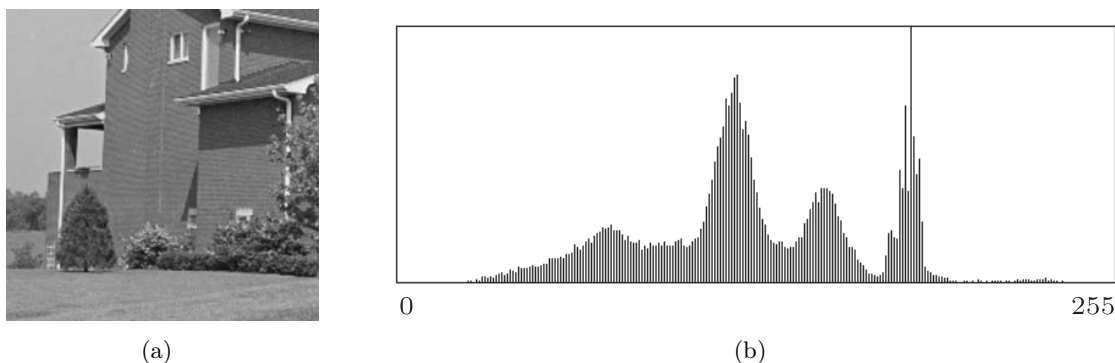
Ha a téglalapok összterülete 1, akkor a sűrűség-hisztogramot kapjuk. Ekkor az  $i$ -edik téglalap magassága:

$$\frac{\nu_i}{n(y_i - y_{i-1})}.$$

Vegyük észre, hogy a gyakorisági hisztogram és a sűrűség-hisztogram ábrázolásában az egyetlen különbség a skála beosztása. A statisztika alaptétele szerint, ha  $M, k \rightarrow \infty$  (azaz növeljük a mintaelemeszámot és finomítjuk a beosztást), akkor a sűrűség-hisztogram az elméleti sűrűségfüggvényhez tart.

A hisztogram definícióját folytonos valószínűségi változóra (analóg képre) adtuk meg. A digitális képfeldolgozás azonban diszkrét változókra épül, ám a hisztogram-fogalom bevezetése a fentiekkel megegyező módon elvégezhető. Az eljárást a könnyebb érthetőség kedvéért a jól ismert 256 szürkeárnyalatú (8 bites) intenzitásképekre mutatjuk be. Az intenzitáshisztogramok természetesen más skálázással is elkészíthetők, a színes képek hisztogramjait pedig az egyes színsatornákon külön-külön tekinthetjük. Tekintsünk tehát egy 256-szintű digitális képet, ahol a fekete és a fehér közti tartományt 256 árnyalatra bontottuk fel, úgy, hogy az  $i + 1$ -edik árnyalat az világosabb, mint az  $i$ -edik. Most a szürkeségi skálát 256 intervallumra osztjuk, ahol minden intervallumba egyetlen szürkeségi érték esik, azaz  $(y_0 < y_1 < \dots < y_{256})$ . A valószínűségi mintát maga a kép, illetve annak intenzitásértékei adják. A hisztogram  $i$ -edik oszlopának magasságát az  $[y_i, y_{i+1})$  intervallumba eső elemek száma, azaz a képen lévő

$i$ -edik szürkeárnyalatú képpontok száma adja. Ebben az esetben a hisztogram alakja attól is függ, hogy milyen nagy képet vizsgáltunk. Például egy képnek és annak négyszeresen nagyított párjának a hisztogramja nem egyezik meg. Ennek elkerüléséhez normálhatjuk a hisztogramot, leosztva az oszlopokban lévő értékeket képpontok számával. Ebben az esetben az oszlopok a megfelelő intenzitásértékek relatív gyakoriságát mutatják. A hisztogramértékek  $[0, 1]$  intervallumba normálásához pedig a maximális hisztogramértékkel kell elvégeznünk a leosztást. Az 1.1. ábrán egy szürkeárnyaltos képet és annak hisztogramját figyelhetjük meg.



1.1. ábra. 256-szintű (8 bites) szürkeskálás kép hisztogramja; (a) eredeti kép, (b) intenzitás-hisztogram.

### 1.2.1. Küszöbölési technikák

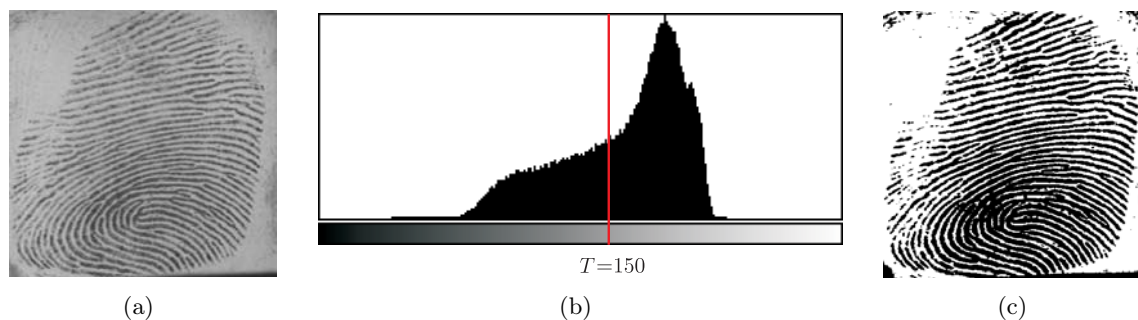
Az intenzitáshisztogram azért nagyon népszerű a képfeldolgozási eljárásokban, mert igen szemléletesen mutatja az adott kép intenzitáseloszlását, továbbá igen jelentős valószínűség-számítási és statisztikai apparátus áll a rendelkezésre a megfelelő információk (automatikus) kinyeréséhez. A hisztogram értelmezése azon a tényen alapul, hogy annak erős módusai (hegyei) általában a kép nagyobb méretű homogén intenzitású objektumaira utalnak. Ennek megfelelően, a hisztogram módusait elválasztó völgyekhez tartozó intenzitásértékek alkalmazsak lehetnek az objektumok szétválasztására vagyis a kép szegmentálására.

A kép intenzitástartományok alapján történő szétbontását küszöbölésnek nevezzük. A  $K$  szintű küszöbölés elvégzéséhez jelölje a kép teljes intenzitástartományát  $y_0 < y_1 < \dots < y_L$ , rögzítsük a  $T_1, \dots, T_K \in \mathbb{R}$  küszöbértékeket, és legyen  $T_0 = y_0$ ,  $T_{K+1} > y_L$ . A küszöbértékeknek megfelelően a képpontokat  $K + 1$  osztályba soroljuk, nevezetesen a  $j$ -ik osztályba azok a pontok kerülnek, amelyek intenzitása a  $[T_j, T_{j+1})$  intervallumba esik ( $j = 0, \dots, K$ ). A  $K = 2$  eset a kép binarizálást jelenti. A szétbontáshoz használt értékek (küszöbök) megadása általában a hisztogram lokális szélsőértékeinek (módusainak/völgyeinek) meghatározásával történik. A küszöbölés végezhető manuálisan vagy automatikusan is, ám az utóbbi megoldás többszintű küszöbölésnél lényegesen bonyolultabbá és pontatlanabbá válhat.

A küszöbölést (például binarizálást) általában előfeldolgozásként használhatjuk későbbi

képfeldolgozási műveletekhez. Az alábbiakban áttekintünk néhány egyszerű eljárást, amelyeket a hisztogram küszöböléséhez használhatunk.

**Szintrevágás.** Lényegében ez a művelet jelenti a klasszikus egy- vagy többszintű küszöbölést. Az egyszintű küszöböléshez egy  $T$  küszöböt rögzítünk és a képpontokat aszerint soroljuk két osztályba, hogy az intenzitásuk  $T$  alá vagy fölé esik-e. Az eljárásra az 1.2. ábrán láthatunk példát a hisztogram és a küszöbérték feltüntetésével egy 256-szintű szürkeskálás kép esetén. Az ábrán jól megfigyelhető, hogy bár szemre a binarizálás egyszerűnek ígérkezik, a megfelelő küszöbérték meghatározása nem mindig triviális. Amennyiben a globális küszöbölés nem nyújt kielégítő eredményt célszerű lehet a küszöbölést lokálisan, a képet kisebb részekre felbontva elvégezni.



1.2. ábra. Két szintre vágás; (a) eredeti kép, (b) intenzitáshisztogram és választott küszöbérték, (c) binarizált kép.

Többszintű küszöbölés esetén az intenzitásskálán több küszöbértéket is megadunk, és a képpontok új intenzitásértékeit általában a kapott tartományok legkisebb értékeire állítjuk.

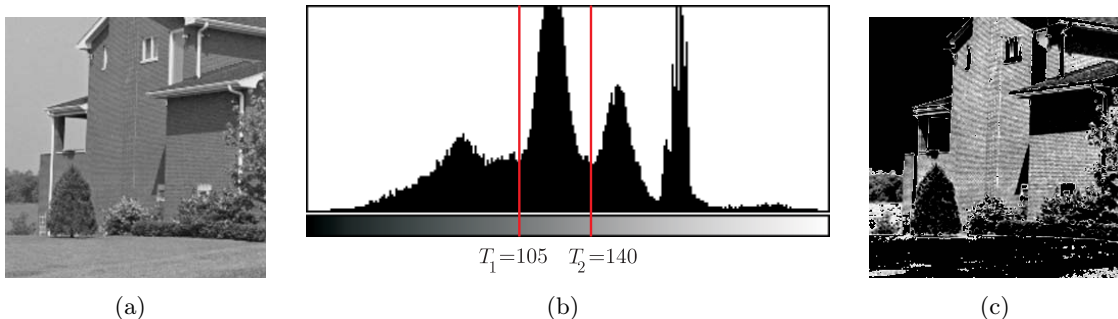
**Sávkiemelés/sávkizárás.** Az intenzitáshisztogram segítségével egyszerűen kiemelhetjük a fontosnak tartott intenzitásokat, illetve analóg módon ki is szűrhetjük ezeket a képből. Az egyes intenzitásintervallumok kiemelését sávkiemelésnek, míg azok eltávolítását sávkizárásnak nevezzük. Így például egy intenzitásáv kiemeléséhez/kizárásához adjuk meg a  $T_1$  és  $T_2$  küszöbököt, és az eredeti  $f$  képből készítsük el a megfelelő sávkiemelt

$$g(x, y) = \begin{cases} f(x, y), & \text{ha } T_1 \leq f(x, y) < T_2 \\ 0, & \text{egyébként,} \end{cases}$$

illetve sávkizárt

$$g(x, y) = \begin{cases} 0, & \text{ha } T_1 \leq f(x, y) < T_2 \\ f(x, y), & \text{egyébként} \end{cases}$$

képet. A sávkiemelés hatását az 1.3. ábrán láthatjuk egy adott küszöbölésnek megfelelően.



1.3. ábra. Sávkimelés; (a) eredeti kép, (b) intenzitáshistogram megfelelő küszöbértékekkel, (c) sávkimelés eredménye.

### 1.2.2. Hisztogram-transzformáció

Az intenzitáshistogram transzformációja előfeldolgozó lépés mellett már önmagában is jól alkalmazható bizonyos képjavítások elvégzésére. Az egyszerűbb tárgyalhatóság kedvéért használjunk most normált  $r$  intenzitásváltozót, azaz legyen  $0 \leq r \leq 1$ . Tekintsük az

$$s = T(r)$$

transzformációt, ami az eredeti kép egy  $r$  intenzitásértékéhez egy új  $s$  intenzitásértéket rendel. A transzformációval szemben az alábbi két feltételt támasztjuk:

- i)  $T(r)$  egyértékű és monoton növekvő a  $0 \leq r \leq 1$  intervallumon,
- ii)  $0 \leq T(r) \leq 1$  minden  $0 \leq r \leq 1$ -re.

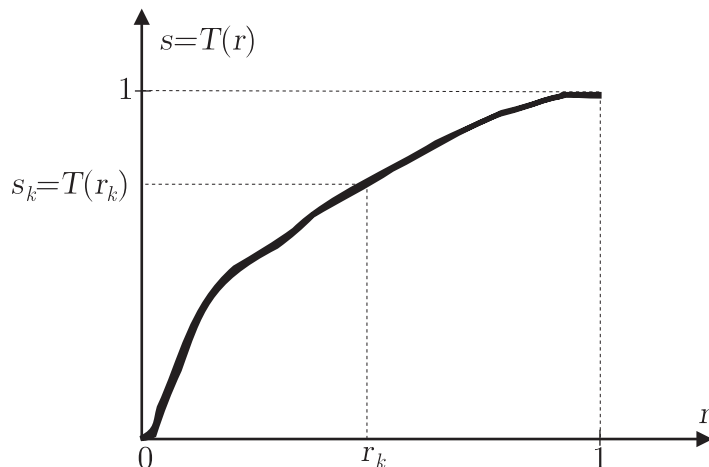
Az i) feltétel biztosítja a szürkeségi skála árnyalatai sorrendjének megőrzését a legkisebbtől a legnagyobbig, míg az ii) feltétel megtartja a használt szürkeségi skálát. Egy ilyen transzformációs függvény látható az 1.4. ábrán. Az  $s$ -et  $r$ -re visszaalakító inverz transzformáció (amennyiben létezik) a következő:

$$r = T^{-1}(s), \quad 0 \leq s \leq 1,$$

ahol a  $T^{-1}(s)$  transzformáció is teljesíti az i) és ii) feltételeket  $s$ -re nézve.

Ha az  $r$ , illetve  $s$  folytonos valószínűségi változók sűrűségfüggvénye (hisztogramja)  $p_r(r)$ , illetve  $p_s(s)$ , valamint  $s = T(r)$  ismert és  $T^{-1}(s)$  teljesíti az i) és ii) feltételeket, akkor az átalakított szürkeségi fokozatok valószínűségi eloszlása (azaz a transzformációval nyert új  $p_s(s)$  sűrűségfüggvény) a következő módon határozható meg (lásd [10]):

$$p_s(s) = \left[ p_r(r) \frac{\partial r}{\partial s} \right]_{r=T^{-1}(s)}. \quad (1.1)$$



1.4. ábra. Szürkeárnyalatos transzformációs függvény.

### 1.2.3. Hisztogram-kiegyenlítés

Azt az eljárást, ami a képet úgy transzformálja, hogy a világosság kódok egyenletes eloszlásúakká váljanak, hisztogram-kiegyenlítésnek nevezzük. Ez főként olyan képeknél hasznos, amelyek valamilyen ok miatt (például rossz megvilágítás) kevésbé kontrasztosak, azaz az intenzitásértékek az intenzitástartomány kisebb részére koncentrálnak. Az eljárás lényegében analóg módon a folytonos és diszkrét térben egyaránt elvégezhető.

#### Folytonos eset

Tekintsük most a következő speciális hisztogram-transzformációs függvényt:

$$s = T(r) = \int_0^r p_r(w) dw \quad 0 \leq r \leq 1. \quad (1.2)$$

Vegyük észre, hogy az így definiált  $T(r)$  függvény éppen az  $r$  változó eloszlásfüggvénye, ezért eleget tesz az előző szakaszban leírt i) és ii) feltételnek. Az  $s$  deriváltjára  $r$ -re vonatkozóan

$$\frac{\partial s}{\partial r} = p_r(r) \Rightarrow \frac{\partial r}{\partial s} = \frac{1}{p_r(r)}$$

teljesül. Helyettesítsük  $dr/ds$ -t az (1.1) egyenletbe:

$$p_s(s) = \left[ p_r(r) \frac{1}{p_r(r)} \right]_{r=T^{-1}(s)} = [1]_{r=T^{-1}(s)} = 1, \quad 0 \leq s \leq 1. \quad (1.3)$$

Az így kapott valószínűségi változó egyenletes eloszlású lesz a korábbi eloszlásától függetlenül.



**Példa folytonos hisztogramkiegyenlítésre.** Kövessünk végig egy konkrét példát folytonos hisztogramtranszformációra. Ehhez legyen az eredeti (analóg képhez tartozó) sűrűségfüggvény alakja (lásd még 1.5(a) ábra):

$$p_r(r) = \begin{cases} -2r + 2, & \text{ha } 0 \leq r \leq 1 \\ 0, & \text{egyébként.} \end{cases}$$

Ekkor a hisztogramkiegyenlítés az

$$s = T(r) = \int_0^r -2w + 2 \, dw = -r^2 + 2r$$

transzformációs függvény alapján történik (lásd 1.5(b) ábra). Ellenőrzésképpen vizsgáljuk meg, hogy mi lesz a kiegyenlített kép hisztogramja. Ehhez szükségünk van az  $s = T^{-1}(r)$  függvényre, azaz ki kell fejeznünk az  $r$ -et az  $s$  segítségével. Ez a  $s = -r^2 + 2r$  egyenletből most könnyen megtehető (például a másodfokú egyenlet megoldóképletét használva):

$$T^{-1}(s) = r = 1 \pm \sqrt{1-s},$$

és mivel  $r$  csak a  $[0, 1]$  intervallumon nem tűnik el, ezért csak az

$$T^{-1}(s) = r = 1 - \sqrt{1-s}$$

eset állhat fenn. Felhasználva ezt az összefüggést

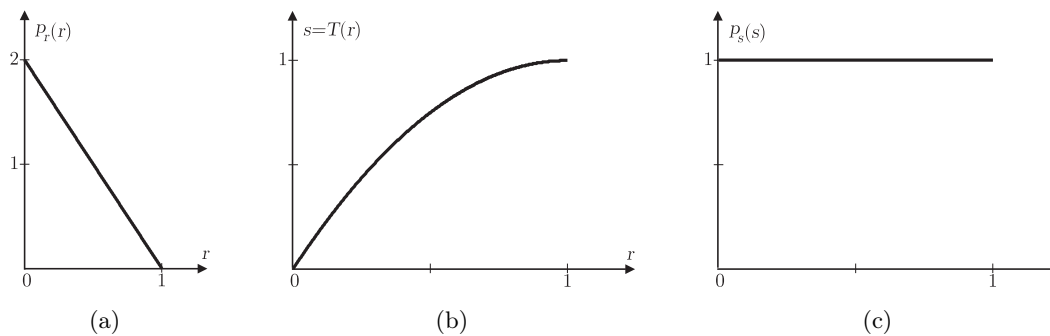
$$p_r(r) = -2r + 2 = -2(1 - \sqrt{1-s}) + 2 = 2\sqrt{1-s}$$

adódik. Ezek után a  $p_s(s)$  új sűrűségfüggvény meghatározásához (lásd 1.5(c) ábra) helyettesítsünk az (1.2) egyenletbe:

$$\begin{aligned} p_s(s) &= \left[ (-2r + 2) \frac{dr}{ds} \right]_{r=1-\sqrt{1-s}} = \left[ (2\sqrt{1-s}) \frac{d}{ds} (1 - \sqrt{1-s}) \right]_{r=1-\sqrt{1-s}} = \\ &= \left[ (2\sqrt{1-s}) \left( \frac{1}{2\sqrt{1-s}} \right) \right]_{r=1-\sqrt{1-s}} = [1]_{1-\sqrt{1-s}} = 1, \end{aligned}$$

ahol felhasználtuk, hogy

$$\frac{d}{ds} (1 - \sqrt{1-s}) = \frac{1}{2\sqrt{1-s}}.$$



1.5. ábra. Sűrűségfüggvény (hisztogram) kiegyenlítése folytonos esetben; (a) eredeti sűrűségfüggvény, (b) transzformációs függvény, (c) kiegyenlített sűrűségfüggvény.

### Diszkrét eset

A folytonos esetnél leírt módszer a diszkrét esetre való természetes átfogalmazással alkalmazható digitális képekre. Legyen adott egy  $M$  képpontból álló  $L$ -szintű digitális kép és jelölje  $r_0, \dots, r_{L-1}$  a lehetséges normált intenzitásértékeket, azaz  $0 \leq r_k \leq 1$  ( $k = 0, \dots, L-1$ ). Jelölje  $m_k$  az  $r_k$  intenzitású képpontok számát a képen, így a gyakorisághisztogram  $p_r(r) = m_k/M$  alakban írható fel.

Az (1.2) egyenlet diszkrét változata a következő:

$$s_k = T(r_k) = \sum_{j=0}^k \frac{m_j}{M} = \sum_{j=0}^k p_r(r_j), \quad 0 \leq r_k \leq 1, \quad k = 0, 1, \dots, L-1. \quad (1.4)$$

Az inverz transzformáció képlete:  $r_k = T^{-1}(s_k)$ ,  $0 \leq s_k \leq 1$ , ahol mind  $T(r_k)$ , mind  $T^{-1}(s_k)$  teljesíti az i) és ii) feltételt. Miután kiszámoltuk az  $s_k$  értékeket, már csak egy lépés marad hátra. Nevezetesen, az új intenzitásértékek nem illeszkednek pontosan a régi szűrési értékekre, ezért minden új értékhez a hozzá legközelebbi régi értéket kell hozzárendelni kerekítéssel.

**Példa diszkrét hisztogramkiegyenlítésre.** Legyen adott egy  $64 \times 64$ -es, 8-szintű kép. Az 1.2.3. táblázat oszlopai tartalmazzák rendre az intenzitásértékeket, az ezekhez tartozó képpontok számát, az intenzitásértékek relatív gyakoriságát, a transzformációs formulával kapott  $s'_k$  új intenzitásértékeket, majd ezen értékeket a régi világosságkódokra kerekítve.

Tekintsük még szemléltetésül az 1.6. ábrát is, ahol az 1.6(a) ábrán az eredeti  $411 \times 480$ -as, 256-szintű szürkeárnyalatos kép látható, alatta az 1.6(b) ábrán pedig a kép hisztogramja. A transzformációs függvényre példaként hozott 1.4. ábrán látható a hisztogramkiegyenlítés által kiszámolt  $s'_k$  értékekből adódó transzformációs függvény képe. Az 1.6(c) ábrán a kiegyenlített kép, míg az 1.6(d) ábrán annak hisztogramja látható.

1.1. táblázat.  $64 \times 64$ -es, 8-szintű kép intenzitáshistogramjának kiegyenlítése.

$r_k$	$n_k$	$p_r(r_k)$	$s'_k$	$s_k$
$r_0 = 0$	523	0,13	0,13	1/7
$r_1 = 1/7$	780	0,19	0,32	2/7
$r_2 = 2/7$	1053	0,26	0,58	4/7
$r_3 = 3/7$	818	0,20	0,78	5/7
$r_4 = 4/7$	470	0,11	0,89	6/7
$r_5 = 5/7$	222	0,05	0,94	1
$r_6 = 6/7$	164	0,04	0,98	1
$r_7 = 1$	66	0,02	1,00	1

#### 1.2.4. Lokális hisztogramkiegyenlítés

A hisztogramkiegyenlítés eredménye az eljárás globális volta miatt nem minden esetben kielégítő ([10]), mivel nem veszi figyelembe a kép lokális tulajdonságait. Ennek kiküszöbölésére használhatunk lokális eljárást, amely egy lokális szomszédságon belül javítja a képet [22]. Ennek egy lehetséges megvalósítása:

$$g(x, y) = A(x, y)[f(x, y) - m(x, y)] + m(x, y),$$

ahol

$$A(x, y) = k \frac{R}{\sigma(x, y)}, \quad 0 < k < 1,$$

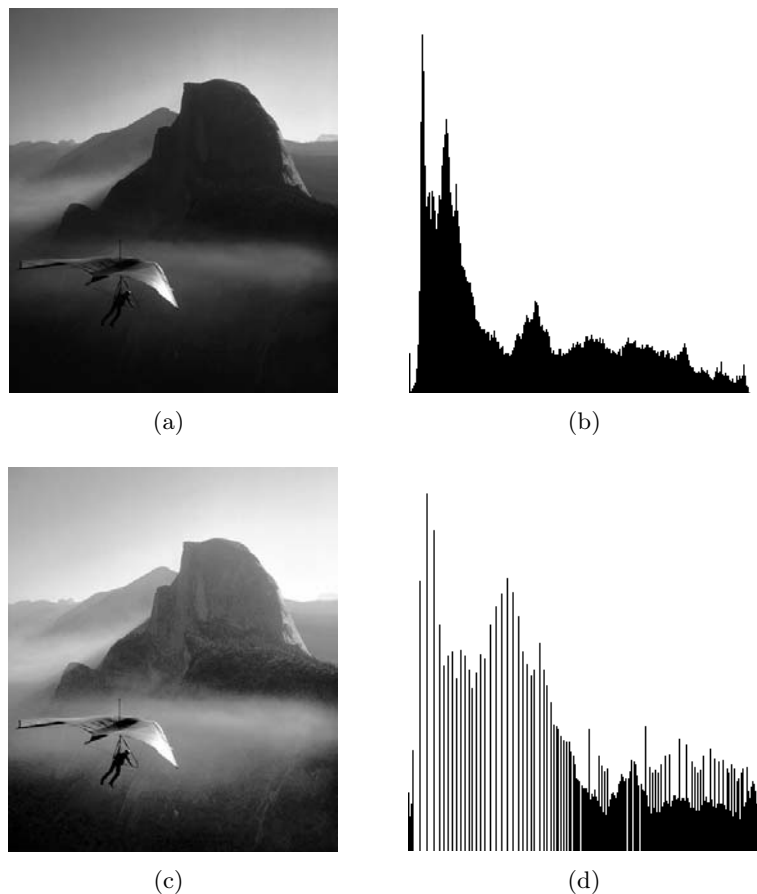
továbbá  $m(x, y)$ , illetve  $\sigma(x, y)$  az  $(x, y)$  középpontú maszk intenzitásértékeinek az átlaga, illetve szórása, míg  $R$  a kép szürkeségi szintjeinek globális átlaga.

#### 1.2.5. Multispektrális hisztogram

Természetesen nem csak szürkeárnyalatos, hanem többcsatornás képek hisztogramját is elkészíthetjük. Tekintsük például a jól ismert 24-bites (TrueColor) RGB reprezentációt, ahol három monokromatikus színkomponens (vörös,zöld,kék) aránya határozza meg a végső színt, a színkomponenseket egyenként 8 biten megadva. Ebben az esetben valójában három különböző hisztogramot készíthetünk, mindegyik komponenshez egyet-egyét.

Egy másik lehetőség közvetlenül egy 3D-s hisztogram elkészítése, amelynek alapját  $2^8 \times 2^8 \times 2^8$  színekocka adja, ahol a hisztogram  $(i, j, k)$  indexű eleme az így leírt színhez tartozó képpontok számát jelenti.

A fentiek egyszerű következménye, hogy az origó fekete, a kocka origóhoz legközelebbi három csúcsában kapnak helyet a monokromatikus színek, az origótól legtávolabbi csúcs fehér, a fehérhez legközelebbi három csúcs pedig két-két monokromatikus szín keveréke, a ciánkék (kék és zöld), a bíborvörös (piros és kék) és a sárga (piros és zöld). A koordináta-rendszer áthelyezhető másik pontba, így például a fehér színt választva origónak és az innen kiinduló élek irányvektorainak csúcsait tekintve a CMY színrendszerhez jutunk. Az 1.7(a)



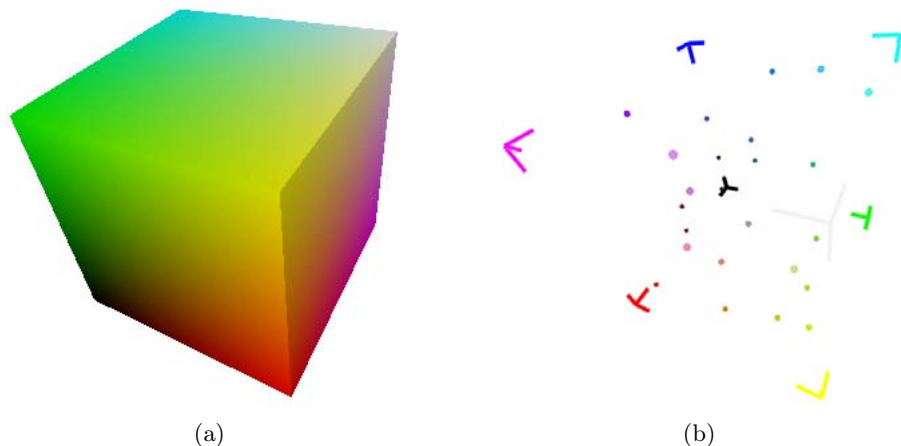
1.6. ábra. Szürkeárnyalatos kép histogram-kiegyenlítése; (a) eredeti kép, (b) intenzitáshistogram, (c) kiegyenlített kép, (d) intenzitáshistogram.

ábra a színekockát mutatja perspektivikusan, míg az 1.7(b) ábrán a kockán belül véletlenszerűen elhelyezett 24 gömb látható pozíciójuknak megfelelően színezve.

### Histogramtranszformáció multispektrális esetben

Több színcsatornával rendelkező kép esetén felmerülhet a kérdés, hogy az intenzitásképekre ismertetett histogramtranszformációt (például kiegyenlítést) hogyan célszerű végrehajtani. Egy egyszerű megközelítés, ha a képet színkomponensekre bontjuk, és azokra, mint intenzitásképekre külön-külön végrehajtjuk a histogramtranszformációt, majd a feldolgozott színcsatornákból visszaállítjuk a képet. Ha a komponensenkénti transzformációs függvények lényegesen eltérnek egymástól, akkor az ezen elven alapuló histogram-kiegyenlítés a színek torzulásához vezethet, mivel a színkomponensek változtatása függetlenül történik.

Alkalmazhatunk egy másik transzformációs módszert is, amivel kapcsolatot teremthetünk



1.7. ábra. Az RGB színekocka; (a) a színekocka felszíne, (b) gömbök a színekocka belsejében.

a színek komponensek közt. Nevezetesen, tekinthetjük az egyes színcsatornák transzformációs függvényeinek olyan súlyozott átlagát, ahol a súlyok összege 1 (például számtani közep). Az így kapott transzformációs függvényt ezután mind a három komponensre külön-külön alkalmazhatjuk. Ez az eljárás kielégíti az i) és ii) feltételeket, így valóban egy megengedett hisztogramtranszformáció. (Az ii) feltétel teljesüléséhez vegyük észre, hogy a használt súlyozás miatt az intenzitások értékkészlete megmarad, míg i) azért áll fenn, mert monoton függvények összege is monoton).

Az utóbbi eljárásnál a színek nem torzulnak, mivel minden színcsatornán ugyanazt a transzformációt végezzük el, ami egy másik színárnyalat meghatározását jelenti. Az első módszer előnye viszont, hogy amennyiben nem kell mindenképpen megakadályoznunk a szín torzulását, akkor a komponensenkénti transzformációs függvényt használva a hisztogramkiegyenlítés jobban kiemeli a kép apróbb jellemzőit. A két módszer közötti különbség jól megfigyelhető az 1.8. ábrán, ahol az 1.8(b) ábra a komponensenkénti transzformációs függvények használatával, míg az 1.8(c) ábra az egyes csatornákon számolt transzformációs függvények átlagolásával készült.

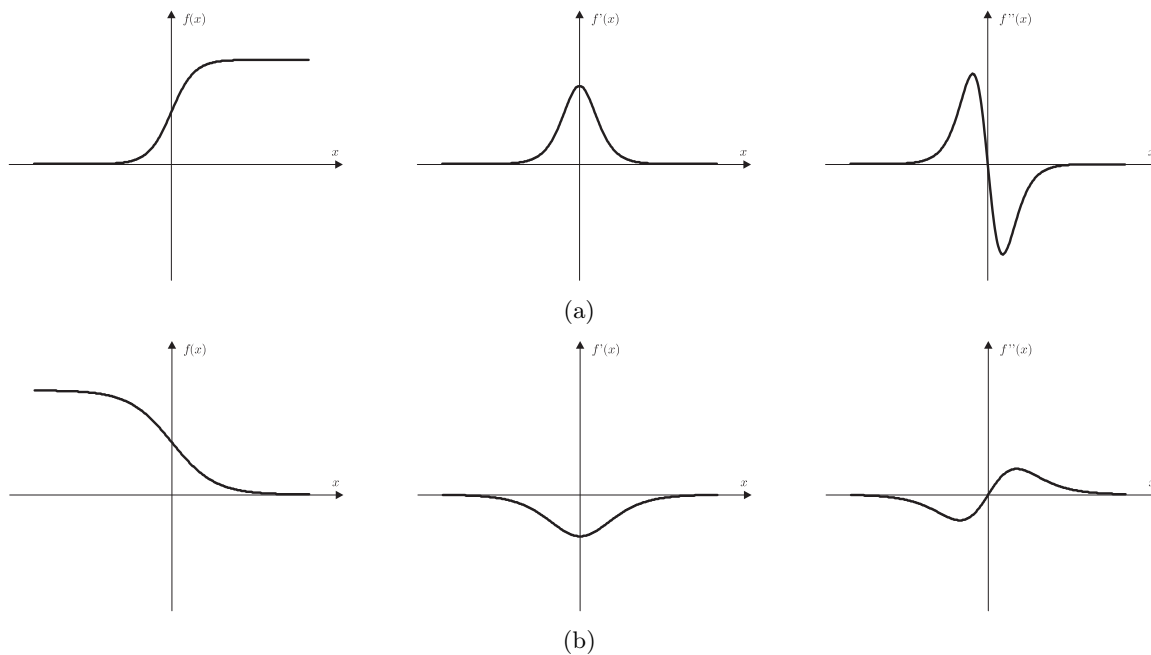
### 1.3. Éldetektálás

Az emberi szem nagyon érzékeny az intenzitásváltozásokra. Ezek a változások általában az objektumok határvonalainál jelentkeznek, amelyek érzékelése lehetőséget teremt az objektum felismerésére. A képfeldolgozásban ennek a megközelítésnek a kontúr alapú szegmentálás felel meg, melynek egy fontos lépése a kontúrponatok (élpontok) detektálása. Azoknak a helyeknek a kereséséhez pedig, ahol az intenzitásfüggvény gyorsan változik, természetes módon analitikus eszközöket, elsősorban a deriválást hívhatjuk segítségül. Az intenzitásfüggvény mind első, mind második deriváltjából hasznos információkhoz juthatunk az éllel kapcsolatban.



1.8. ábra. Multispektrális kép hisztogram-kiegyenlítése; (a) eredeti kép, (b) kiegyenlítés komponensenkénti transzformációs függvényekkel, (c) kiegyenlítés a transzformációs függvények átlagolásával.

Az 1.9. ábra a deriváltak viselkedését mutatja lépcsős jellegű intenzitásváltozás esetén.



1.9. ábra. Az  $f(x)$  intenzitásfüggvény első és második deriváltjának viselkedése; (a) sötét-világos átmenet esetén, (b) egy lassabb világos-sötét átmenet esetén.

### 1.3.1. Gradiens módszerek

A gradiens alapú éldetektálás intuitív háttérében az áll, hogy amíg az összegzés (átlagolás/integrálás) a képek elmosását eredményezi, addig a deriválás (mint inverz művelet) várhatóan

ellentétes hatást vált ki. Ezen a természetes gondolaton alapuló meglátás ténylegesen megállja a helyét, és a gradiens módszerekre épülő eljárások fontos szerepet játszanak az intenzitásátmenetek (élek) detektálásában, így egyúttal a képek élesítésében is.

A gradiensvektor (érintővektor) a kép első deriváltjaiból képezett

$$G[f(x, y)] = (G_x, G_y)^T$$

vektor, ahol  $G_x = \partial f / \partial x$ , illetve  $G_y = \partial f / \partial y$  a gradiensvektor két komponense. A gradiensvektor hosszát (a gradiens nagyságát), illetve irányát euklideszi módon definiáljuk, azaz

$$|G[f(x, y)]| = \sqrt{G_x^2 + G_y^2}, \quad \text{illetve} \quad \angle G = \tan^{-1} \frac{G_y}{G_x}$$

a gradiens nagysága, illetve szöge,

$$\Delta f = G_x \cos \Theta + G_y \sin \Theta$$

pedig a  $\Theta$  irányba eső változás gyorsaságát mutatja. A gradiens iránya a változás irányát, nagysága pedig a változás mértékét adja meg. Értelemszerűen, az intenzitásátmeneteknél (élpontoknál) a gradiens relatíve nagyobb, az élek pedig a gradiens irányára merőlegesen futnak. Az élpontok detektálásához általában a gradiensnek ezt a két jellemzőjét használják.

A gradiens digitális közelítésére differenciákat használhatunk. Elsőként a

$$|G[f(x, y)]| = \sqrt{(f(x, y) - f(x + 1, y))^2 + (f(x, y) - f(x, y + 1))^2}, \quad (1.5)$$

általában Sobel gradiensként ismert formula került bevezetésre. Ha jobban megfigyeljük a képletet, kiderül, hogy lényegében egy  $2 \times 2$ -es környezeten belül végzünk összehasonlítást, nyilvánvalóan az intenzitáseltéréseken alapulva. Ezt a szemléletet követve a későbbiekben az irodalomban újabb közelítő számolási módokat javasoltak, amelyek kisebb számításigénnyel bírnak, vagy egyes esetekben jobb eredményt adnak. Így az (1.5) képletnek az alábbi variánsai születtek:

$$|G[f(x, y)]| = |f(x, y) - f(x + 1, y)| + |f(x, y) - f(x, y + 1)|,$$

illetve az ú.n. Roberts gradiens

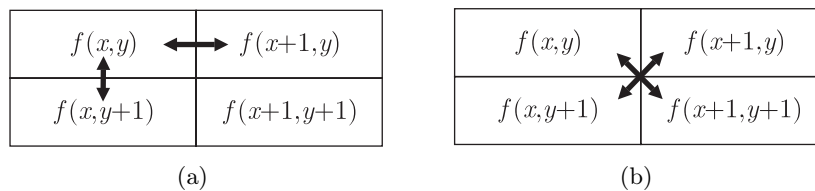
$$|G[f(x, y)]| = \sqrt{(f(x, y) - f(x + 1, y + 1))^2 + (f(x + 1, y) - f(x, y + 1))^2}, \quad (1.6)$$

$$|G[f(x, y)]| = |f(x, y) - f(x + 1, y + 1)| + |f(x + 1, y) - f(x, y + 1)|.$$

(1.5) és (1.6) alapvetően abban tér el, hogy a  $2 \times 2$ -es tartományon belül, más pontok intenzitását hasonlítjuk össze, amint ez az 1.10. ábrán is megfigyelhető.

A kiszámolt gradiensértékekből készíthető el az eredeti  $f(x, y)$  intenzitáskép  $g(x, y)$  élvagy gradiensképe. A legegyszerűbb élkép a

$$g(x, y) = |G[f(x, y)]|$$



1.10. ábra. Digitális gradiens számolása; (a) Sobel gradiens, (b) Roberts gradiens.

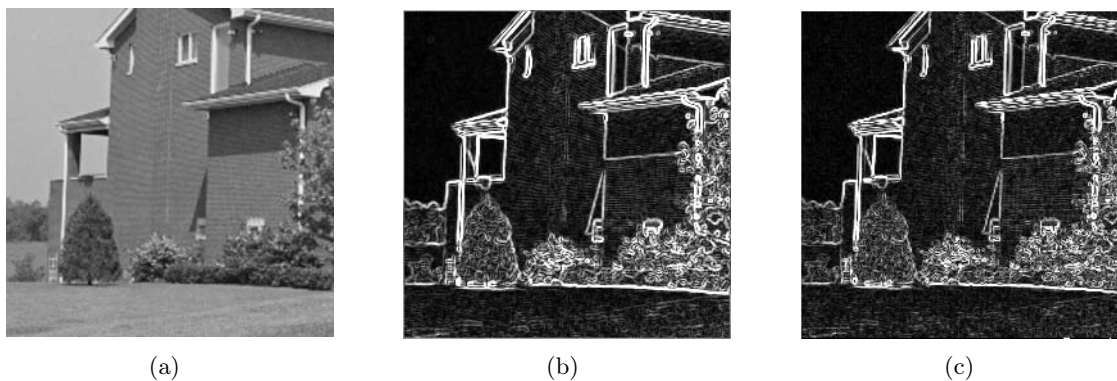
módon állítható elő. Az élkép megjelenítésekor ügyeljünk rá, hogy a gradiensértékeket kvantálnunk (kerekítenünk és normálnunk) kell a megjelenítési intenzitásokra. Mivel a kép homogén részein nincsen intenzitásátmenet, így az élkép általában sötét. Amennyiben szeretnénk megőrizni az eredeti kép homogén részeinek információt, úgy azt a következő egyszerű küszöböléssel tehetjük meg:

$$g(x,y) = \begin{cases} f(x,y), & \text{ha } |G[f(x,y)]| < K, \\ |G[f(x,y)]|, & \text{egyébként,} \end{cases}$$

ahol a  $K$  küszöb rögzített pozitív konstans. Amennyiben viszont csak a megfelelően erős élék helye a fontos, úgy a következő binarizálást célszerű választani:

$$g(x,y) = \begin{cases} 0, & \text{ha } |G[f(x,y)]| < K, \\ 1, & \text{egyébként.} \end{cases}$$

Az 1.11. ábrán a Sobel és Roberts éldetektáló eljárások eredményét láthatjuk a  $g(x,y) = |G[f(x,y)]|$  élkép definíciót használva.



1.11. ábra. A gradiensoperátor eredménye; (a) eredeti kép, (b) Sobel gradiens, (c) Roberts gradiens.

Az éldetektálás fent leírt módszerei felfoghatók egy  $2 \times 2$ -es maszkolásként is. Mivel ezek az eljárások a vizsgált képpont nagyon szűk környezetét figyelik, így természetesen zajérzékenyek. A zajérzékenység csökkenthető  $3 \times 3$ -as maszkok használatával, amelyeket a 2.2.3. szakaszban tárgyalunk bővebben.



### 1.3.2. Laplace operátor

A Laplace-operátort az  $f(x, y)$  kép (intenzitásfüggvény) második parciális deriváltjaiból származtatjuk a következő módon:

$$L[f(x, y)] = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

A Laplace-operátor eltűnik a konstans intenzitású tartományokon és az intenzitásátmeneteken, mivel az intenzitásváltozás kezdetét és végét jelzi. Ezen tulajdonságának megfelelően a Laplace operátor egyes képekre kettős éleket eredményez, amint ez az 1.12. ábrán is megfigyelhető.



(a)



(b)

1.12. ábra. Kettős élek a Laplace operátor használatakor; (a) eredeti kép, (b) Laplace operátor hatása.

Az élpontok esetén a Laplace-operátor előjeléből következtethetünk arra, hogy a vizsgált képpont az él által határolt világos, illetve sötét tartományba esik-e. Neveztesen, negatív előjel esetén az élpont a magasabb (világosabb) intenzitású tartományba, pozitív előjel esetén pedig az alacsonyabb (sötétebb) intenzitásúba esik. A Laplace-operátor digitális közelítést általában  $3 \times 3$ -as maszkolással végzik, amelynek leírását a 2.2.3. szakaszban adjuk meg.

### 1.3.3. A második derivált tengelymetszetei

Az eddig megismert éldetektáló eljárások a vizsgált képpont meglehetősen kis környezetének maszkolásán alapultak. Ennek megfelelően függenek az objektum méretétől és meglehetősen zajérzékenyek. [14]-ben megfigyelték, hogy az élek jelenlétére az intenzitásfüggvény lépcsős változása utal, így a függvény első deriváltjának szélsőértéke, míg a második deriváltjának zérushelye van ezeken a pontokon, lásd 1.9. ábra az 1D esetre. Nyilvánvaló, hogy az élatmenet annál gyorsabb, minél merőlegesebb az élre az a vágósík/egyenes amely mentén az intenzitásváltozást figyeljük. Mivel a zérushelyek detektálása általában egyszerűbben és pontosabban elvégezhető, az élek lokalizálása a második derivált zérushelyeinek keresésén is alapulhat [14]. Ekkor a legnagyobb kihívást a második derivált hatékony meghatározása jelenti.

Egy gyakori eljárás, ha zajszűréshez először simítjuk a képet, majd ezután számoljuk a második deriváltat. Ekkor a simító műveletnek két feltételt kell kielégítenie. A szűrőnek sávkorlátosnak kell lennie a frekvenciatartományban, továbbá a képtartományban a vizsgált pont szomszédos pontjai alapján kell választ adnia. Ez a két feltétel időnként ütközik, de a normális eloszlás használatával párhuzamosan optimalizálhatók. A normális (Gauss) szűrőoperátor legegyszerűbb alakja:

$$N(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}},$$

ahol  $x$  és  $y$  a képkoordináták,  $\sigma$  pedig a tekintett eloszlás szórása. Az operátor használatos normalizáló tényezővel ellátva is:

$$N(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}, \quad \text{illetve} \quad N(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

A szűrő egyetlen paramétere a  $\sigma$  szórás, ami vizsgált környezet méretével arányos. A közép-ponttól távolabb lévő képpontoknak így kisebb a befolyása, míg a  $3\sigma$ -nál messzebb lévőknek már egyáltalán nincs is.

Az éldetektálás tehát az

$$L [N(x, y, \sigma) * f(x, y)]$$

művelet (**LoG** operátor) meghatározásával történik, ahol a  $*$  konvolúció jelöli a normális simítás lépését,  $L$  pedig a második derivált (Laplace) operátort. A formula tovább egyszerűsíthető a résztvevő operátorok linearitását kihasználva:

$$L [N(x, y, \sigma)] * f(x, y).$$

Megfelelő átalakítások után a **LoG** operátor konvolúciós maszkjára

$$h(x, y) = c \left( \frac{x^2 + y^2 - \sigma^2}{\sigma^4} \right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

adódik, ahol a  $c$  tényező szerepe a maszkelemek összegének 0-ra normálása. A konvolúciós maszk digitálisan közelíthető; az 1.13. ábrán a megfelelő  $5 \times 5$  és  $13 \times 13$  méretű maszkokat láthatjuk. Formája miatt a **LoG** operátor "mexikói kalap"-ként is ismert.

A második derivált ilyentén meghatározása igen robusztus. A normális szűrés hatékonyan semlegesíti a  $3\sigma$ -nál távolabb lévő pontok hatását, a Laplace operátor pedig megbízhatóan detektálja az intenzitásváltozásokat. A megközelítés előnye, hogy a vizsgált pont nagyobb környezetét figyelembe veszi, amivel számos esetben jobb eredményt kaphatunk, mint a klasszikus, kisebb tartományokra épülő éldetektálókkal használatakor. A **LoG** operátor eredményét az 1.14. ábrán figyelhetjük meg különböző maszméretek mellett.

### 1.3.4. Képek élesítése

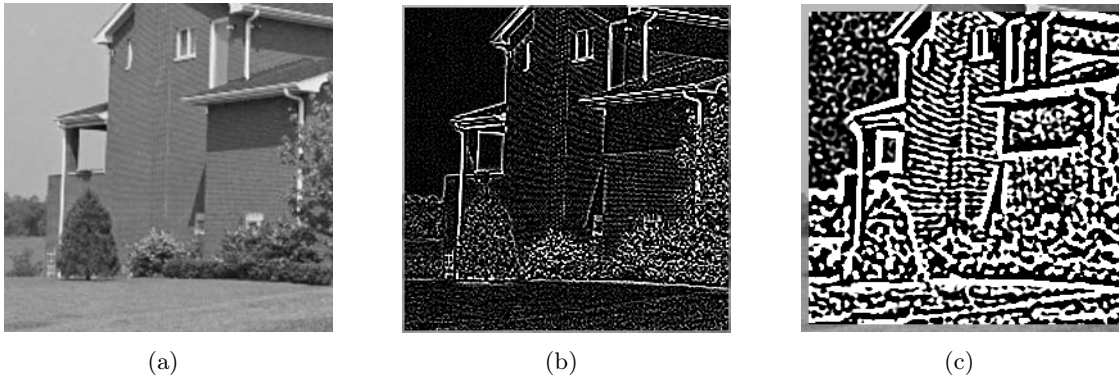
Az eddig bemutatott technikák feladata a kép intenzitásátmenetei helyének detektálása volt. A felhasználó számára sokszor nagyon hasznos lehet a képek élesítése is, amelynek eléréséhez

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

0	0	0	0	0	-1	-1	-1	0	0	0	0	0
0	0	0	-1	-1	-2	-2	-2	-1	-1	0	0	0
0	0	-2	-2	-3	-3	-4	-3	-3	-2	-2	0	0
0	-1	-2	-3	-3	-3	-2	-3	-3	-3	-2	-1	0
0	-1	-3	-3	-1	4	6	4	-1	-3	-3	-1	0
-1	-2	-3	-3	4	14	19	14	4	-3	-3	-2	-1
-1	-2	-4	-2	6	19	24	19	6	-2	-4	-2	-1
-1	-2	-3	-3	4	14	19	14	4	-3	-3	-2	-1
0	-1	-3	-3	-1	4	6	4	-1	-3	-3	-1	0
0	-1	-2	-3	-3	-3	-2	-3	-3	-3	-2	-1	0
0	0	-2	-2	-3	-3	-4	-3	-3	-2	-2	0	0
0	0	0	-1	-1	-2	-2	-2	-1	-1	0	0	0
0	0	0	0	0	-1	-1	-1	0	0	0	0	0

(a)
(b)

1.13. ábra. A **LoG** operátor maszkokkal való közelítései; (a)  $5 \times 5$ -ös maszk, (b)  $13 \times 13$ -ös maszk.

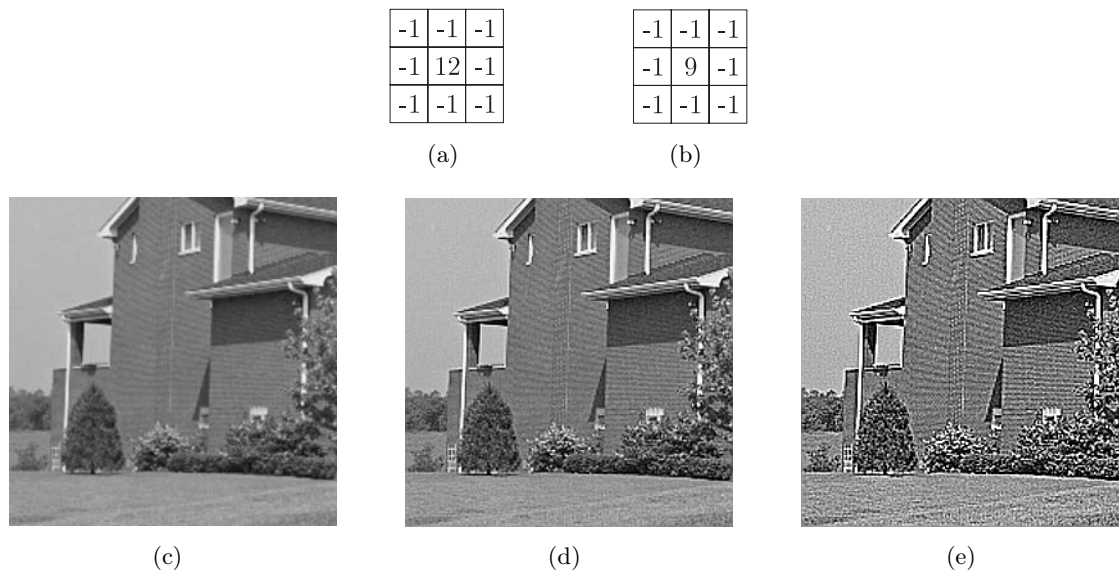


1.14. ábra. A **LoG** operátor hatása; (a) eredeti kép, (b)  $5 \times 5$ -ös maszk esetén, (c)  $13 \times 13$ -ös maszk esetén.

az éldetektálásnál használt megközelítések kínálnak jó alapot. Egy lehetséges képelesítő eljárás például a [24]-ben javasolt

$$g(x, y) = f(x, y) - cS(x, y),$$

művelet, ahol  $f$  az eredeti kép,  $g$  az eredménykép,  $c$  az élesítés fokát szabályozó pozitív konstans tényező,  $S(x, y)$  pedig az intenzitásfüggvény alakját mutató mérték, amelyet valamilyen gradiens operátorral határozhatunk meg. Nagyon gyakran a Laplace operátort használják erre a célra. Ez az elképzelés egyszerű  $3 \times 3$ -as maszkokkal is megvalósítható, ahogyan ez az 1.15. ábrán is látható. Figyeljük meg, hogy ezeken a konvolúciós maszkokon a súlyok összege nem 0, szemben a 2.2. szakaszban bemutatott valamilyen elütő tulajdonságot detektáló maszkokkal.



1.15. ábra. Képesítés  $3 \times 3$  méretű maskokkal; (a) enyhén élesítő mask, (b) fokozottan élesítő mask, (c) eredeti kép, (d) enyhe élesítés, (e) fokozott élesítés.

Képesítést a frekvenciatartományban is lehet végezni gradiensmódszerek segítségével, mivel a Fourier transzformált harmonikus függvények kombinációja. Nevezetesen, a  $\sin(nx)$  harmonikus függvény deriváltja  $n \cos(nx)$ , azaz minél magasabb a frekvencia, annál nagyobb a derivált értéke. Ez a példa is jól szemlélteti a gradiens operátorok élesítő tulajdonságát.

### 1.3.5. Canny éldetektálás

Canny [4] egy új megközelítést javasolt éldetektáláshoz, amely elsősorban fehér zajjal terhelt lépcsős élatmenetek detektálására optimális. A módszer hatékonysága miatt mindenképpen figyelmet érdemel az alkalmazásokban, de a jelen jegyzet terjedelme nem teszi lehetővé a teljes eljárás bemutatását. Itt most csak a főbb gondolatokat összegezzük a teljes leírás a [3, 4, 5, 27] publikációkban található meg.

Az éldetektáló tervezésekor az alábbi három fő kritériumot vesszük figyelembe:

- A fontos élek nem veszhetnek el, de ne legyenek valótlan élek.
- Az él detektált és tényleges helye között minimális legyen az eltérés.
- Ugyanazon él ne legyenek többszörösen detektálva. Ezt a kritériumot az első feltétel lényegében lefedi, mivel ha egy élet duplán detektálunk, az egyiket hamisnak tekinthetjük. Ez a kritérium főleg a zajos élek detektálásához fontos.

A kritériumokat kielégítő eljárás a következő fejlődési lépéseken ment keresztül:

1. Canny először 1D-s esetre készített el egy éldetektálót, amely az első két kritériumot kielégítette. Sikerült zárt alakot megadnia a szűrőfüggvényre a második momentumok segítségével.
2. Az eljárás, amely a harmadik kritériumot is figyelembe veszi, numerikus optimalizálással kapható meg. Az eredményül kapott szűrő 20%-nál kisebb hibával közelíthető normális simítószűrőkkel (lásd az 1.3.3. szakaszban leírt **LoG** operátort).
3. Sikerült a detektort 2D-re kiterjeszteni a képet és egy 2D-s normális szűrőt konvolváva, majd az eredményt a gradiens irányába deriválva.
4. Egy felső küszöböt meghaladó erős élválaszt adó pontokat élpontoknak tekintünk, a gyakran zajok által keltett gyengébb élek pedig csak akkor minősülnek tényleges éleknek, ha az erősségük meghalad egy bizonyos alsó küszöböt és erősebb élekhez kapcsolódnak.
5. A normális eloszlás  $\sigma$  szórásparaméterét használjuk a tekintett környezet méretének megadásához, nevezetesen azt a környezetet vizsgáljuk, amely esetén még megmarad egy erős válasz. Így biztosítható az él minél pontosabb lokalizálása. Canny javasolt egy szintetizáló lépést is, amelynek segítségével a kisebb  $\sigma$  paraméterű szűrésekből megjósoljuk a nagyobb  $\sigma$  esetében várható viselkedést, majd összehasonlítjuk azt a ténylegesen számolt értékekkel és csak azon új éleket fogadjuk el, amelyek esetében ez az eltérés szignifikáns volt.

### 1.3.6. Parametrikus élmodellek

A parametrikus élmodellek azon a megfigyelésen alapulnak [20], hogy a digitális képek egy folytonos (vagy szakaszonként folytonos) intenzitásfüggvény mintavételezéssel nyert zajos közelítéseként is felfoghatók. Mivel a folytonos intenzitásfüggvény nem ismert, ezért a mintából készíthetünk rá becslést, majd a kép tulajdonságait már az így kapott folytonos becslésből származtathatjuk, akár képpontméret alatti pontossággal is. A képek bonyolultsága miatt az egyetlen folytonos függvénnyel való közelítés gyakorlatilag kizárt, ezért ezt szakaszonként folytonos függvények segítségével hajthatjuk végre a képpontok környezetének, az ún. fazetták (facet-ek) segítségével. Az így kapott képpontmodellt fazettamodellnek nevezzük [11].

Az intenzitásfüggvény a képpontok környezetében különböző szintű bonyolultsággal közelíthető. A legegyszerűb a sima fazettamodell, amikor a környezetek konstans intenzitást kapnak. A lejtős fazettamodellben a környezetekben az intenzitásváltozásokat lineáris módon írjuk le, és ez a megközelítés természetesen kiterjeszthető bonyolultabb (magasabb fokú) közelítésekre, így a kvadratikus vagy harmadfokú (bi-cubic) reprezentációkra is. A fazettamodell széleskörű alkalmazhatóságára (zajszűrés, szegmentálás, statisztikailag fontos élek detektálása, gradiens alapú éldetektálás, vonal- és sarokdetektálás) [12] mutat rá.

### 1.3.7. Élek multispektrális képeken

Egy multispektrális kép egy pontját egy  $m$  dimenziós vektorként írhatjuk le, ahol a vektor komponenseit az  $m$  csatornán lévő intenzitásértékek adják. Több módszer is létezik a multispektrális képek éleinek detektálására, lásd [8]. Egy természetes eljárás a csatornákon külön-külön elvégezni az éldetektálást például a 1.3.1. szakaszban leírt technikákkal. Az így kapott élképeket kombinálva, a multispektrális kép pontjaihoz a csatornákon mért legerősebb élválaszt és élrányt rendelhetjük hozzá élinformációként. A maximális értékek kiválasztása mellett használatosak még a csatornákhöz tartozó élképek lineáris vagy egyéb kombinációi is [19]. Tekinthejtük még ezen kívül ugyanazon pont egyes csatornaértékeinek különbségét vagy hányadosát is [23].

Kiterjeszhető a Roberts gradiensnél (lásd 1.3.1. szakasz) ismertetett elképzelés is, amivel egy  $2 \times 2 \times m$  méretű maszkoláshoz jutunk [27], ahol az egyes képpontok befolyását mérő együttható a korrelációs együtthatóhoz hasonlít. Jelölje  $\bar{f}(x, y)$  az  $(x, y)$  pont egyes csatornákon mért intenzitásértékeinek átlagát, továbbá  $f_j$  a  $j$ -edik csatorna intenzitásfüggvényét. Az  $(x, y)$  pontban a maszkválasz az alábbi két kifejezés minimumaként definiálható:

$$\frac{\sum_{j=1}^m [f_j(x, y) - \bar{f}(x, y)] [f_j(x+1, y+1) - \bar{f}(x+1, y+1)]}{\left( \sum_{j=1}^m [f_j(x, y) - \bar{f}(x, y)]^2 \sum_{j=1}^m [f_j(x+1, y+1) - \bar{f}(x+1, y+1)]^2 \right)^{\frac{1}{2}}},$$

$$\frac{\sum_{j=1}^m [f_j(x+1, y) - \bar{f}(x+1, y)] [f_j(x, y+1) - \bar{f}(x, y+1)]}{\left( \sum_{j=1}^m [f_j(x+1, y) - \bar{f}(x+1, y)]^2 \sum_{j=1}^m [f_j(x, y+1) - \bar{f}(x, y+1)]^2 \right)^{\frac{1}{2}}}.$$

### 1.3.8. Élkapcsolás, élvékonyítás

Mivel a kép éleinek erőssége nem homogén ezért gyakran előfordul, hogy a küszöböléseken alapuló élképeken az élek megszakadnak. Az ilyen szakadások korrigálását nevezzük élkapcsolásnak.

Az egyik legegyszerűbb eljárás az élpontok egy lokális (például  $3 \times 3$  vagy  $5 \times 5$  méretű) környezetének átvizsgálása és az ott található élpontok összekapcsolása. Az  $(x, y)$  és az  $(x', y')$  élpontot akkor kapcsoljuk össze, ha az azokban számolt gradiensvektorok nagysága és iránya hasonló, azaz

$$\begin{aligned} | |G[f(x, y)]| - |G[f(x', y')]| | &< K_1, \\ | \angle G(x, y) - \angle G(x', y') | &< K_2, \end{aligned}$$

ahol  $K_1$  és  $K_2$  megfelelő pozitív küszöbök. Vegyük észre, hogy az él irányára való merőlegességük miatt a gradiensvektorok irányának hasonlósága egyben az élek irányának hasonlóságát

is jelenti. Természetes eljárásként az  $(x, y)$  és  $(x', y')$  pontokat akkor kötjük össze, ha mindkét feltételnek eleget tesznek. Használhatjuk például a [27]-ben javasolt, 1.16. ábrán látható maszkokat, amelyek, ha illeszkednek a bináris élképre, akkor a centrális képpont értékét 1-re állítják.

0	1	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	1
0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0

1.16. ábra.  $3 \times 3$ -as maszkrendszer lokális élkapcsoláshoz.

A lokális élkapcsoló eljárások mellett globális megközelítést is használhatunk, például a Hough transzformációt alkalmazva. Ilyenkor azokat a pontokat próbáljuk összekötni, amelyek egy adott egyenesre, görbére vagy valamilyen más geometria alakzatra esnek. Ehhez először kiszámoljuk a gradiensértékeket, majd elvégezzük a Hough transzformációnál tekintett (például polár) paramétersík szokásos beosztását, és kiválasztjuk a legtöbb elemet tartalmazó cellákat. Ezt követően teszteljük az egy cellába eső pontok összefüggőségét. Ha egy szakadási helyen lévő két végpont távolsága egy megadott küszöbnél kisebb, illetve a gradiensértékek ebben a két pontban hasonlóak, akkor a két végpontot összekötjük.

Amennyiben az élképen lévő élek nem egy képpont vastagságúak, akkor célszerű, hogy az élkapcsolást egy élvékonyítás előzze meg. Bináris élkép esetén az eljárás elvégezhető például az ugyancsak [27]-ben megadott és az 1.17. ábrán megfigyelhető  $3 \times 3$ -as maszkrendszerrel. A maszkok ? elemei tetszőlegesen lehetnek 0 vagy 1 értékűek, és illeszkedés esetén a centrális élpontok törlődnek.

1	?	0	?	1	1	?	1	?	?	1	?	0	0	0	0	0	?	?	?	0	0	?	?
1	1	0	0	1	?	1	1	?	?	1	0	?	1	0	0	1	1	1	1	?	?	1	1
?	0	0	0	0	0	?	?	0	0	?	0	1	1	?	0	?	1	?	1	?	?	1	?

1.17. ábra.  $3 \times 3$ -as maszkrendszer bináris élvékonyításhoz.

### 1.3.9. Sarokdetektálás

A sarokpontok meghatározása elsősorban lokális detektorokkal végezhető. Az egyik legegyszerűbb ilyen módszer a Moravec-detektor [18]

$$MO(i, j) = \frac{1}{8} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} |f(i, j) - f(k, l)|,$$

amely az  $f$  kép nagy intenzitásátmenetekenél lévő képpontjainak (sarkok, erős élek) esetében adja a maximális választ.

Jobb eredményeket kaphatunk számításigényesebb operátorok használatával, lásd [12, 29] vagy a fazettamodellt (lásd 1.3.6. szakasz) használó [13]. Ha például az intenzitásfüggvényt

$$f(i, j) = c_1 + c_2x + c_3y + c_4x^2 + c_5xy + c_6y^2 + c_7x^3 + c_8x^2y + c_9xy^2 + c_{10}y^3$$

harmadfokú módon közelítjük megfelelő  $c_k$  együtthatókkal, akkor a Zuniga-Haralick (ZH) operátor alakja:

$$\text{ZH}(i, j) = \frac{-2(c_2^2c_6 - c_2c_3c_5 - c_3^2c_4)}{(c_2^2 + c_3^2)^{\frac{1}{2}}}.$$

A tesztelesek során valamivel gyengébb eredményt adó Kitchen-Rosenfeld operátor esetén a nevezőben  $(c_2^2 + c_3^2)$  szerepel.

További megközelítések még a sarkok két fél él találkozásaként való detektálása a **LoG** operátor (lásd 1.3.3. szakasz) segítségével [17], továbbá a kontúr görbülete változásának figyelése [6].

### 1.3.10. Felüláteresztő szűrők

A nagy intenzitásúmenetek (zajok és élek) a kép Fourier transzformáltjának magasfrekvenciás összetevőiben jelennek meg. A frekvenciatartománybeli képésítés természetes ötletét tehát az adja, hogy a magas frekvenciákat megtartva, az alacsony frekvenciás összetevőket szűrjük ki. A fejezet bevezetésében bemutatott módon az  $f(x, y)$  kép frekvenciatartománybeli szűrése a  $G(u, v) = F(u, v)H(u, v)$  egyszerű szorzásként elvégezhető, ahol a  $g(x, y)$  eredménykép a  $G(u, v)$  függvény inverz Fourier transzformáltjaként állítható elő. A képésítéshez tehát olyan  $H(u, v)$  szűrőfüggvényre van szükségünk, amely gyengíti az alacsony frekvenciás összetevőket. Az ilyen szűrőket felüláteresztő szűrőknek nevezzük, amelyek legegyszerűbb változata az ún. ideális felüláteresztő szűrő. Ennek szűrőfüggvénye

$$H(u, v) = \begin{cases} 0, & \text{ha } \|(u, v)\| \leq K, \\ 1, & \text{egyébként,} \end{cases}$$

ahol  $K$  megfelelő pozitív küszöb (vágási frekvencia). Az ideális felüláteresztő szűrő síkkal való metszete és teljes alakja az 1.18. ábrán látható.

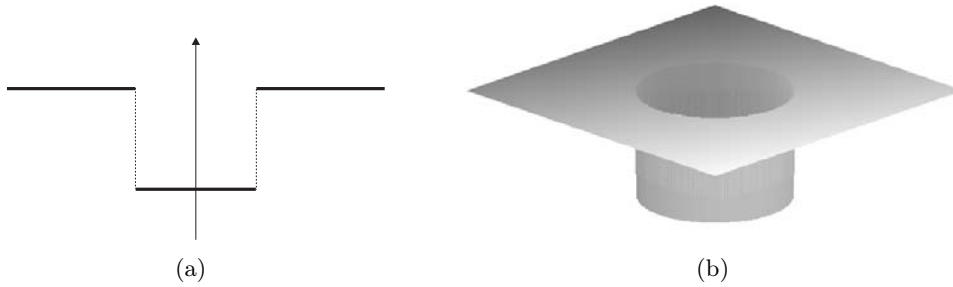
Figyeljük meg, hogy az ideális szűrő meglehetősen drasztikusan gyengíti az alacsony frekvenciás összetevőket, nevezetesen teljesen eltávolítja azokat. Bizonyos képek esetén a szűrő ezen tulajdonsága hátrányokkal is jár, amint ez az 1.19. ábrán is megfigyelhető.

Az ideális szűrő karakterisztikus jellege helyett használhatunk simább átmenetet biztosító szűrőfüggvényt is. Ilyen például a

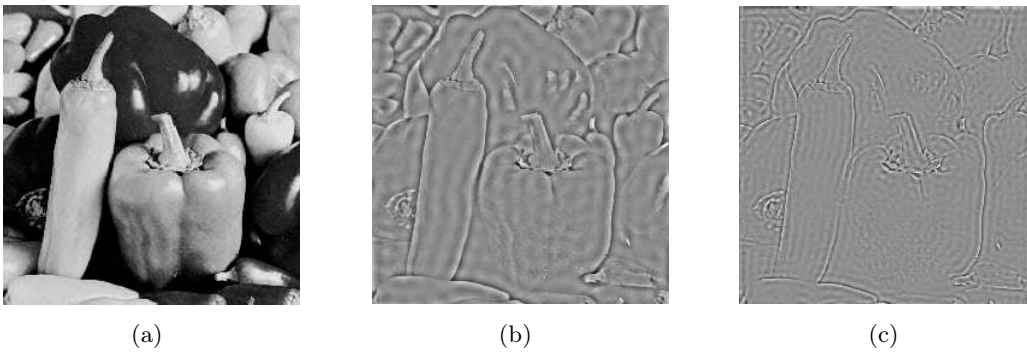
$$H(u, v) = \begin{cases} \frac{1}{1 + \left(\frac{K}{\|(u, v)\|}\right)^t}, & \text{ha } \|(u, v)\| \neq 0, \\ 0, & \text{ha } \|(u, v)\| = 0 \end{cases}$$

módon definiált a Butterworth felüláteresztő szűrő, ahol  $K$  a vágási küszöb,  $t$  pedig az átmenet sebességét szabályozó pozitív paraméter. A Butterworth felüláteresztő szűrő síkkal való



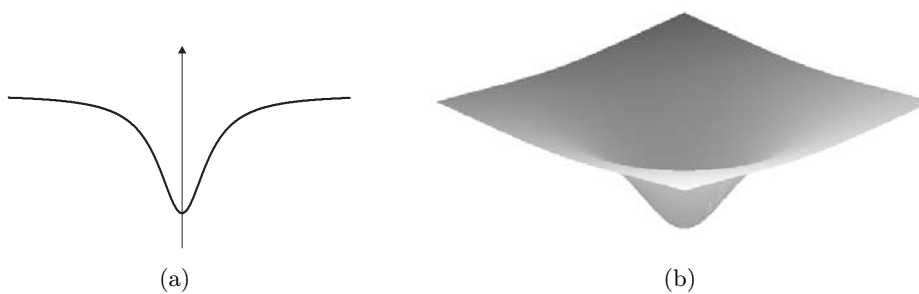


1.18. ábra. Az ideális felüláteresztő szűrő; (a) síkkal való metszete, (b) teljes alakja.

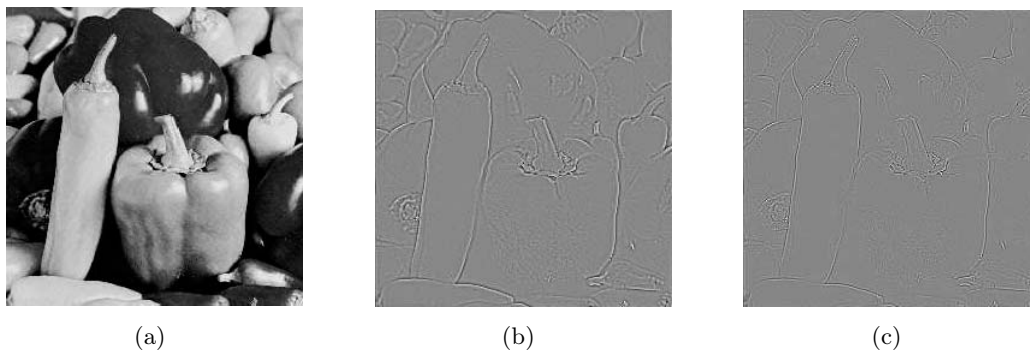


1.19. ábra. Az ideális felüláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép  $K = 20$  esetben, (c) szűrt kép  $K = 40$  esetben.

metszetét és teljes alakját mutató 1.20. ábrán jól megfigyelhető annak az ideális szűrőtől való eltérése. Az 1.21. ábrán a szűrő hatását láthatjuk.



1.20. ábra. A Butterworth felüláteresztő szűrő ( $t = 2$  paraméterű eset); (a) síkkal való metszete, (b) teljes alakja.



1.21. ábra. A Butterworth felüláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép  $t = 2$ ,  $K = 50$  esetben, (c) szűrt kép  $t = 2$ ,  $K = 80$  esetben.

## 1.4. Simítás

Az 1.3. szakaszban sok olyan technikát megismerhettünk, amelyek a képek intenzitásváltozásainak detektálásában használatosak. Az éldetektálás a képek élesítésében és a szegmentálásban játszik fontos szerepet. Ahogyan azonban erre már az 1.3.3. szakaszban is láthattunk példát, az erős intenzitásváltozások is lehetnek nemkívánatos jelenségek, különösen, ha azokat zaj okozza. Ebből következően gyakran indokolt lehet az intenzitás egyenetlenségek kiküszöbölése, azaz a képek simítása. A képsimító eljárásokat elsősorban zajsztűrésre használják, de hasznosak lehetnek a vizuális élmény kellemesebbé tételében is.

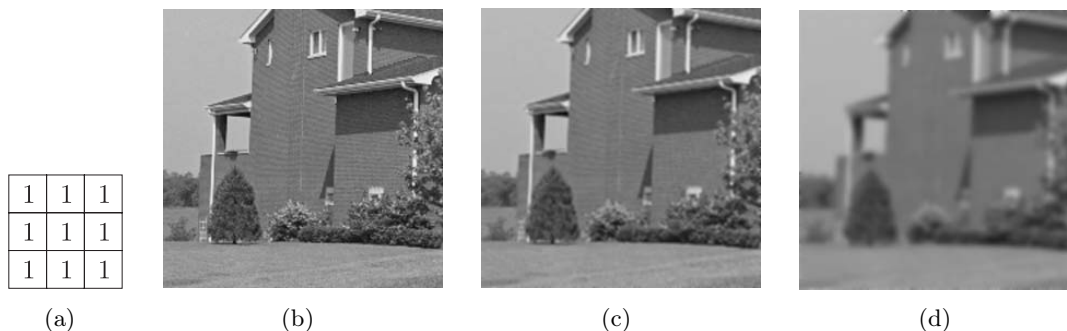
A képsimítás lényegében az éldetektálás inverz műveletének is tekinthető, így az ott közöltekhöz hasonló technikák alkalmazhatók. Ennek megfelelően a képsimító eljárásokat is alapvetően úgy csoportosítjuk, hogy a kép, vagy a frekvenciatartományban dolgoznak-e. A képtartománybeli szűrést konvolúciós maszkokkal végezzük, ami a frekvenciatartományban egy szűrőfüggvénnyel való szorzásnak felel meg a konvolúciós tétel alapján. A következőkben a képek simítására leggyakrabban használt eljárásokat tekintjük át.

### 1.4.1. Környezeti átlagolás

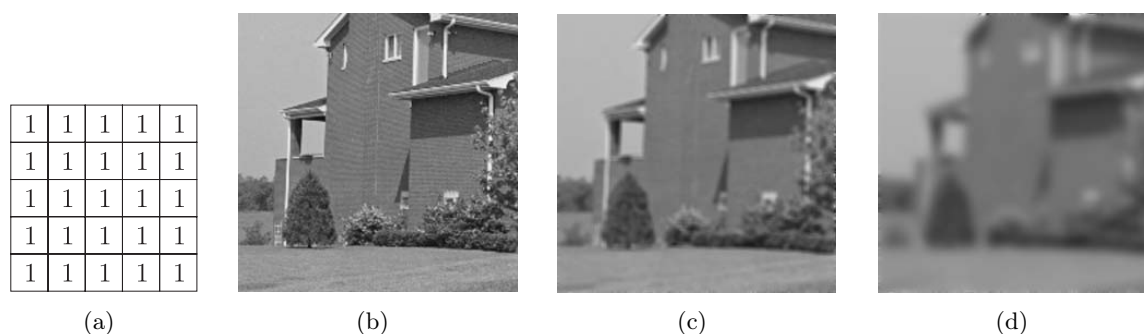
A környezeti átlagolás egy képtartományban végrehajtható konvolúciós maszkolás, amelynek  $3 \times 3$  méretű maszkja, illetve a maszkolás eredménye az 1.22. ábrán látható.

A simítást több iterációs lépésben is alkalmazhatjuk, ahogyan ez szintén megfigyelhető az 1.22. ábrán. Amennyiben kellően sokszor alkalmazzuk az iterációs lépést, úgy egy idő után elérjük a tovább már nem változó konstans intenzitású képet. Az átlagolást nagyobb maszkokkal is végezhetjük, mely esetben az 1.23. ábrán is látható erősebb simításhoz jutunk.

Vegyük észre, hogy mivel most nem egy elütő tulajdonság detektálása a célunk, a maszksúlyok összegének nem kell 0-nak lennie. A környezeti átlagolás zajsztűrésre jól használható, ám hátrányos lehet, hogy új intenzitásértékek is megjelenhetnek a képen.



1.22. ábra. Környezeti átlagolás  $3 \times 3$  méretű maszkkal; (a) konvolúciós maszk, (b) eredeti kép, (c) a maszk hatása, (d) a maszk hatása 10 iterációs lépés után.



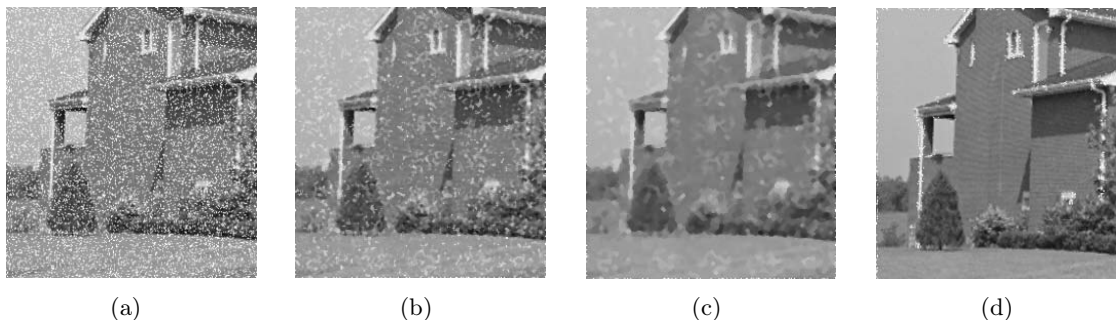
1.23. ábra. Környezeti átlagolás  $5 \times 5$  méretű maszkkal; (a) konvolúciós maszk, (b) eredeti kép, (c) a maszk hatása, (d) a maszk hatása 10 iterációs lépés után.

#### 1.4.2. Medián szűrés

Bár a zajszűrésre és képsimításra használt medián szűrő szintén képtartománybeli maszkoláson alapul, mégsem tartozik a lineáris konvolúciós szűrők közé. Ennek az az oka, hogy a maszkválaszt nem a képtartomány és a maszkkelemek konvolúciós szorzataként számoljuk.  $3 \times 3$ -as medián szűrésnél a vizsgált képtartomány pontjainak intenzitását (azaz az  $x_1, x_2, \dots, x_9$  értékeket, lásd 2.1. ábra) növekvő sorba rendezzük, majd kiválasztjuk az így kapott sorozat középső elemét (mediánját), amelyet új intenzitásértékként a centrális képponthoz rendelünk. Ahhoz, hogy valóban létezzen a sorozatnak "fizikailag" középső eleme, általában az  $x_5$  intenzitást is felhasználjuk a sorbarendezésben. Amennyiben a centrális képpontot ajánlatosabb kihagyni ebből, úgy a sorozat 8 elemet tartalmaz, amelyből értelemszerűen a 4. vagy 5. elemet tekinthetjük középsőként.

A medián szűrés alapötletét az adja, hogy a kiugró intenzitásértékek (zajok) a sorba rendezéskor a sorozat széleire szorulnak, a középső elem az adott képtartomány átlagos intenzitását mutatja. A medián szűrés népszerű módszer a zajszűrésben [10, 27], melynek egyik oka, hogy az eljárás során nem jelennek meg új intenzitásértékek (szemben például az átlag-

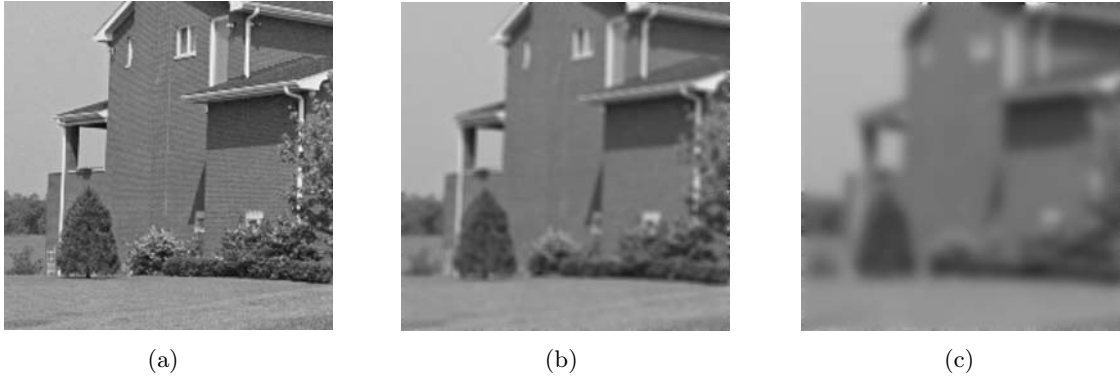
goló szűréssel). A medián szűréshez az átlagoló szűrésnél leírtakhoz hasonlóan használhatunk nagyobb (például  $5 \times 5$ -ös,  $7 \times 7$ -es) maszkokat, illetve tekinthetünk több iterációs lépést is. A medián szűrésnél az átlagoló szűréstől eltérően nem biztos, hogy megfelelő számú iterációs lépés után a konstans intenzitásképet kapjuk meg. Az eljárás eredményét az 1.24. ábrán figyelhetjük meg, klasszikus  $3 \times 3$ -as medián szűrés, több iterációs lépés, illetve egy egyszerű kiegészítő küszöbölési technika figyelembevételével (lásd [1]) után. Az utolsó ábraképből jól látszik, hogy ügyes kiegészítő technikákkal a simító maszkolásokkal igen jó zajszűrés valósítható meg. Az ábrán az is jól megfigyelhető, hogy a medián szűrés kevésbé mossa el az éleket, mint az egyszerű környezeti átlagolás.



1.24. ábra. Medián szűrés hatása; (a) eredeti (15%-os fehér zajjal terhelt) kép, (b) a medián szűrés eredménye  $3 \times 3$ -as maszk esetén, (c) a szűrés eredménye 10 iterációs lépés után, (d)  $5 \times 5$ -ös medián szűrés eredménye kiegészítő küszöbölést használva.

### 1.4.3. Intervallum módszer

Ez az eljárás szintén a képpontok egy adott méretű (például  $3 \times 3$ -as) környezetét tekinti, és a centrális képpont intenzitását akkor változtatja meg, ha annak intenzitása kívül esik a környezetében található képpontok intenzitásértékei által meghatározott intenzitásintervallumon. A centrális képpont intenzitását értelemszerűen nem vesszük figyelembe az intervallum meghatározásakor. Az intenzitásintervallum alsó végpontját a tekintett környezet minimális, felső végpontját pedig annak maximális intenzitásértéke adja. Ha a centrális képpont intenzitása kiesik ebből az intervallumból, akkor új intenzitásértékként az intervallum alsó, illetve felső végpontját kapja attól függően, hogy az eredeti érték az intervallum alá, illetve felé esett-e. A módszer sok ponton emlékeztet a medián szűréshez, így például nem eredményez a képen korábban nem létező intenzitásértékeket. A lényeges különbség, hogy a környezetétől elütő képpont új intenzitásértéke nem a "középső" környezeti érték, hanem azok alsó vagy felső határa lesz. Az intervallum módszer használható nagyobb maszkokkal is, egynél több iterációs lépést azonban nincs értelme alkalmazni. Az eljárás hatását az 1.25. ábrán figyelhetjük meg.



1.25. ábra. Intervallum módszer képsimításhoz; (a) eredeti kép, (b) a szűrés eredménye  $3 \times 3$ -as maszk esetén, (c) a szűrés eredménye  $5 \times 5$ -ös maszk esetén.

#### 1.4.4. Átlagolás inverz gradiens alapján

Az 1.3. szakaszban láthattuk, hogy a gradiensértékek kiválóan használhatók az éldetektálásban és képesítésben. Ennek megfelelően, azt a megközelítést használva, hogy az intenzitás-változás az objektumok belsejében kisebb, mint azok határán, a gradiensértékek inverzéből képsimításra alkalmas konvolúciós maszkot készíthetünk [28]. Nevezetesen, a  $\delta$  inverz gradiens függvény értéke az  $(i, j)$  pontban

$$\delta(i, j) = \begin{cases} \frac{1}{|f(x, y) - f(i, j)|}, & \text{ha } f(x, y) \neq f(i, j), \\ 2, & \text{ha } f(x, y) = f(i, j), \end{cases}$$

ahol  $f$  az eredeti kép,  $(x, y)$  pedig a tekintett maszk centrális pontjainak koordinátái. A  $\delta$  inverz gradiens a  $(0, 2]$  intervallumba esik, értéke az éleken kisebb, míg az objektumok belsejében nagyobb. A konvolúciós maszk nem centrális elemeit a

$$h(i, j) = \frac{1}{2} \frac{\delta(i, j)}{\sum_{(m, n) \in N} \delta(m, n)}$$

összefüggéssel adjuk meg, míg az  $(x, y)$  koordinátájú centrális elemre  $h(x, y) = 1/2$ . Éles intenzitásátmenetek esetén ez a maszkolás az éleket nem mossa el, csak az objektumok belsejében lévő zajokat szűri ki.

#### 1.4.5. Nem lineáris középszűrők

A nem lineáris középszűrők az átlagoló technikák kiterjesztését jelentik a következő modell alapján [27]:

$$g(x, y) = h^{-1} \left\{ \frac{\sum_{(i, j) \in N} c(i, j) h[f(i, j)]}{\sum_{(i, j) \in N} c(i, j)} \right\},$$

ahol  $f$  az eredeti kép,  $g$  az eredménykép,  $N$  az  $(x, y)$  koordinátájú képpont lokális (például  $3 \times 3$ -as) környezete,  $c$  megfelelő súlyfüggvény,  $h$  pedig egyváltozós invertálható függvény. Ha konstans  $c(i, j)$  súlyozást alkalmazunk, a szűrést homomorfnek nevezzük. Az alábbi homomorf szűréseket gyakran használják a képfeldolgozásban:

- számtani közép,  $h(f) = f$ ,
- harmonikus közép,  $h(f) = 1/f$ ,
- mértani közép,  $h(f) = \log f$ .

#### 1.4.6. Képsorozat átlagolása

Az átlagolást képsorozatok esetén is jól alkalmazhatjuk zajszűrésre. Előfordulhat, hogy a felvételekre mindig ráakódik valamilyen állandó véletlen zaj (például műholdképek esetén). Ilyenkor alkalmazhatjuk azt a technikát, hogy ugyanazon jelenetről több (zajos) felvételt készítünk, majd ezeket a felvételeket átlagoljuk. A megközelítés háttérében az áll, hogy a zajos felvételek zömén az egyes képpontok általában a valódi intenzitásértékükkel szerepelnek. Az átlagolás mellett természetesen alkalmazhatjuk az előző szakaszokban bemutatott egyéb technikákat (medián szűrés, intervallum módszer) is. A képek átlagolásának használata bizonyos feltételek mellett matematikailag is jól indokolható. Nevezetesen, tekintsük a

$$g(x, y) = f(x, y) + \mu(x, y)$$

modellt, ahol  $f$  az eredeti kép,  $\mu$  a felvétel során az arra rakódó zaj,  $g$  pedig az eredménykép. A modell realizációjaként tekintsünk  $M$  megfigyelést:

$$\begin{aligned} g_1(x, y) &= f(x, y) + \mu_1(x, y), \\ g_2(x, y) &= f(x, y) + \mu_2(x, y), \\ &\vdots \\ g_M(x, y) &= f(x, y) + \mu_M(x, y), \end{aligned}$$

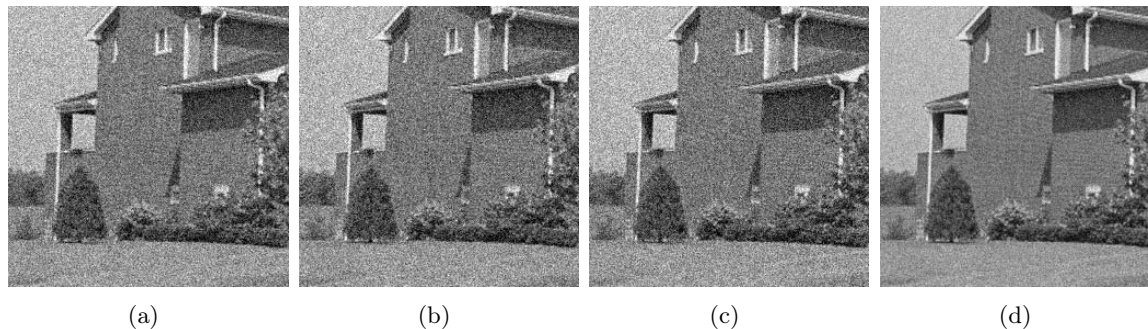
továbbá tegyük fel, hogy a zaj korrelálatlan és 0 várható értékű. Készítsük el a

$$\bar{g}(x, y) = \frac{1}{M} \sum_{i=1}^M g_i(x, y)$$

átlagképet. Ekkor  $\bar{g}$  az  $f$  torzítatlan becslése, azaz  $M \rightarrow \infty$  esetén

$$\begin{aligned} \mathbb{E}(\bar{g}(x, y)) &= f(x, y), \quad \text{és} \\ \sigma^2(\bar{g}(x, y)) &= \frac{1}{M} \sigma^2(\mu(x, y)), \end{aligned}$$

ahol  $\mathbb{E}$  egy valószínűségi változó várható értékét,  $\sigma^2$  pedig a szórásnégyzetét jelöli. Az 1.26. ábrán a fenti eljárásra láthatunk példát az eredeti felvételek 25%-os normális zajjal való terhelése mellett.



1.26. ábra. Képsorozat átlagolása; (a) - (c) (zajos) képsorozat elemei, (d) a képsorozat átlaga.

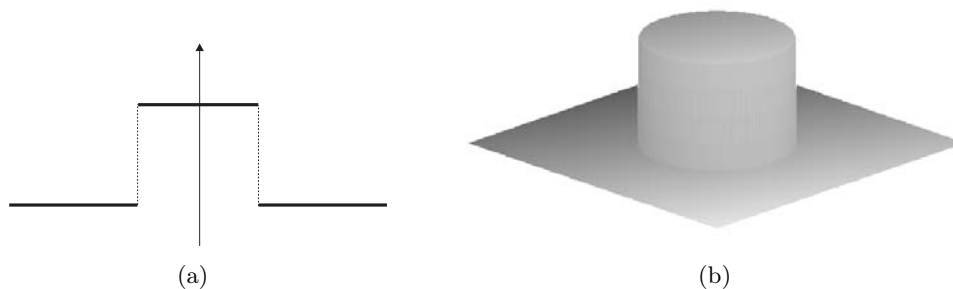
### 1.4.7. Aluláteresztő szűrők

Az aluláteresztő szűrés az 1.3.10. szakaszban tárgyalt feluláteresztő szűrésnél leírtakkal éppen ellentétes gondolon alapul. Nevezetesen, a frekvenciatartomány magas frekvenciás részeit gyengítve (vagy akár teljesen kiszűrve) a képen elmosó hatást érhetünk el. A feluláteresztő szűrésnél tárgyaltakhoz hasonlóan többféle szűrőt is alkalmazhatunk erre a célra, amelyek éppen az ellentétes hatást eredményezik, mint a megfelelő feluláteresztő szűrőpárjuk.

Az ideális aluláteresztő szűrőfüggvény:

$$H(u, v) = \begin{cases} 1, & \text{ha } \|(u, v)\| \leq K, \\ 0, & \text{egyébként,} \end{cases}$$

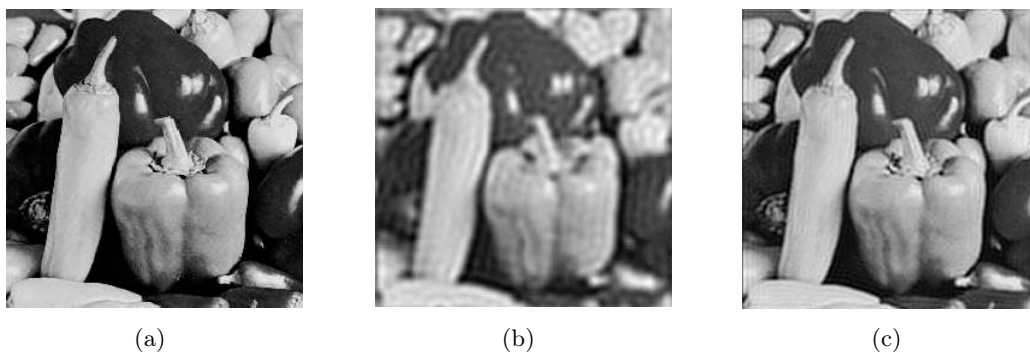
ahol  $K$  megfelelő pozitív küszöb (vágási frekvencia). Az ideális aluláteresztő szűrő síkkal való metszete és teljes alakja az 1.27. ábrán, hatása pedig az 1.28. ábrán látható.



1.27. ábra. Az ideális aluláteresztő szűrő; (a) síkkal való metszete, (b) teljes alakja.

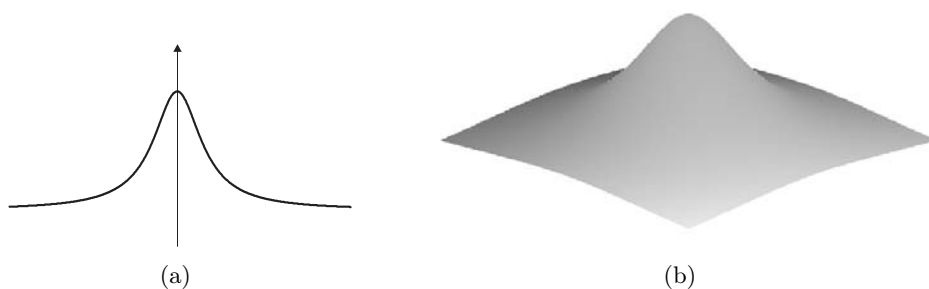
A simább átmenetet biztosító Butterworth aluláteresztő szűrő:

$$H(u, v) = \frac{1}{1 + \left(\frac{\|(u, v)\|}{K}\right)^t},$$



1.28. ábra. Az ideális aluláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép  $K = 25$  esetben, (c) szűrt kép  $K = 60$  esetben.

ahol  $K$  a vágási frekvencia,  $t$  pedig az átmenet sebességét szabályozó pozitív paraméter. A Butterworth feluláteresztő szűrő síkkal való metszetét és teljes alakját az 1.29. ábra, hatását pedig az 1.30. ábra mutatja.



1.29. ábra. A Butterworth aluláteresztő szűrő ( $t = 2$  paraméterű eset); (a) síkkal való metszete, (b) teljes alakja.

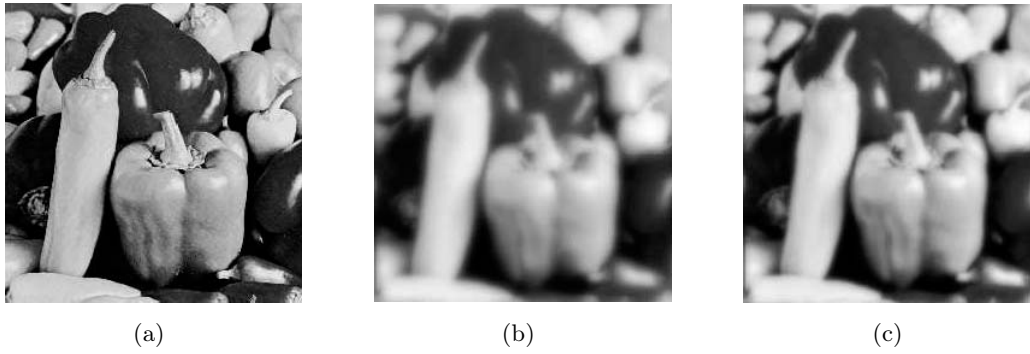
#### 1.4.8. Sávszűrés

A felül- és aluláteresztő szűrők konstrukciójának ötletét felhasználva készíthetünk speciálisabb célú sávszűrőket, amelyek feladata egy adott frekvenciasávba eső értékek gyengítése vagy erősítése. Ezzel a technikával lehetőség nyílik a kép éleinek azok erőssége alapján való feldolgozására. Példaképpen megadunk egy ideális sávkimelő szűrőfüggvényt:

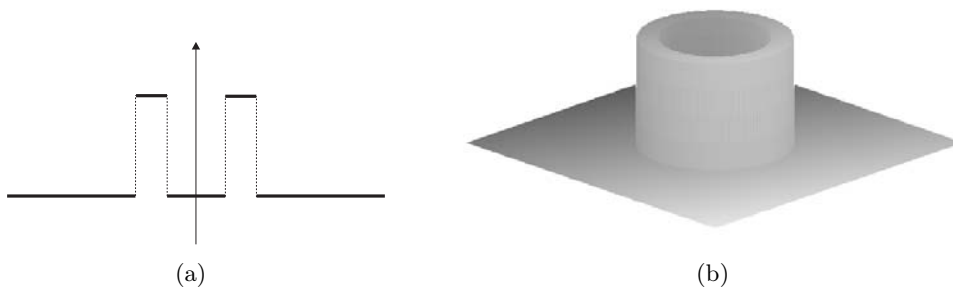
$$H(u, v) = \begin{cases} 1, & \text{ha } K_1 \leq \|(u, v)\| < K_2, \\ 0, & \text{egyébként,} \end{cases}$$

ahol  $K_1$  és  $K_2$  nemnegatív vágási frekvenciák. A szűrő síkkal való metszete és teljes térbeli alakja megfigyelhető az 1.31 ábrán.





1.30. ábra. A Butterworth aluláteresztő szűrő hatása; (a) eredeti kép, (b) szűrt kép  $t = 2$ ,  $K = 10$  esetben, (c) szűrt kép  $t = 2$ ,  $K = 20$  esetben.



1.31. ábra. Ideális sávszűrő; (a) síkkal való metszete, (b) teljes alakja.

## 2. fejezet

# Maszkolás

### 2.1. Bevezetés

A maszkolás alapötlete, hogy a végrehajtani kívánt műveletet nem egyszerre az egész képre globálisan, hanem annak kisebb részleteire lokálisan hajtjuk végre. Az ilyen jellegű eljárásokhoz maszkokat definiálunk, amelyek segítségével eldönthetjük, hogy mi történjen az egyes képrészletekkel. Maszkoknak általában kis méretű, konvex, a maszk origójára centrálszimmetrikus alakzatokat választunk, így például a négyzet rácson a leggyakrabban  $3 \times 3$ -as maszkokkal dolgozunk. Maszkolásnál a maszk elemeit megfelelő valós súlyokkal ellátva, műveletet értelmezve a képrészlet és a maszk között, a képrészlet maszk origójára illeszkedő képpontjára egy maszkválasz értéket számolunk ki, melynek segítségével eldönthetjük, hogy mi történjen az adott ponttal a továbbiakban. A maszkot a kép maszkkal megegyező méretű összes képrészletére helyezve a teljes képen elvégezhető a kívánt művelet. A maszk és a képrészlet között értelmezett műveletként nagyon gyakran az egyszerű belső szorzatot tekintjük, ilyenkor konvolúciós maszkolásról beszélünk. Az eljárás elemeit a 2.1. ábra mutatja, ahol  $X$  a vizsgált képrészletből, illetve  $W$  a használt  $3 \times 3$ -as maszkból sorfolytonosan képzett 9D-s oszlopvektor, a képrészlet  $x_5$  pontjára adott maszkválasz pedig

$$\langle W^T, X \rangle = \sum_{i=1}^9 w_i x_i.$$

Morfológiai műveletek alkalmazásakor szintén a fent leírt maszkolási technikát alkalmazzuk, ám a konvolúciós maszkolással ellentétben itt nem aritmetikai műveletet végzünk a vizsgált képrészlet és a maszk elemei között.

### 2.2. Alakzatok detektálása maszkokkal

A konvolúciós maszkolási eljárást alkalmazhatjuk olyan tulajdonságok detektálására, amelyek bizonyos alakzatok jelenlétére utalnak. Természetes elvárás, hogy amennyiben egy, a

$x_1$	$x_2$	$x_3$
$x_4$	$x_5$	$x_6$
$x_7$	$x_8$	$x_9$

$w_1$	$w_2$	$w_3$
$w_4$	$w_5$	$w_6$
$w_7$	$w_8$	$w_9$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_9 \end{pmatrix} \quad W = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_9 \end{pmatrix}$$

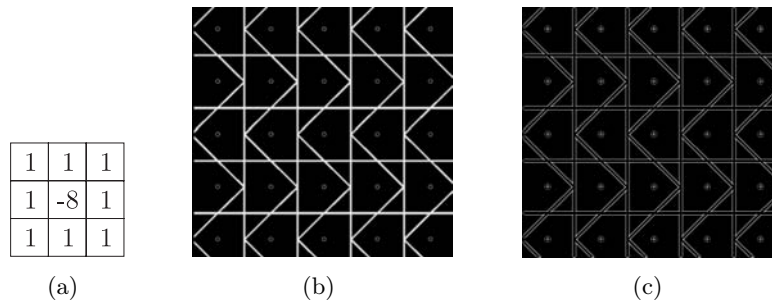
(a)
(b)
(c)

2.1. ábra. Maszkolás a képtartományban  $3 \times 3$  méretű maszkkal; (a) a vizsgált képrészlet adott  $x_i$  intenzitásokkal, (b) a használt maszk, (c) vektoros felírás a  $\langle W^T, X \rangle$  maszkválasz meghatározásához.

környezettől eltérő tulajdonságot kívánunk detektálni (például éldetektálás), akkor konstans képrészlet esetén a maszk 0 választ adjon. Más megfogalmazásban ez azt jelenti, hogy ilyen esetben alkalmazzuk a  $\sum_{i=1}^9 w_i = 0$  feltételt a maszk súlyainak megválasztásához. Ezt a feltételt az alábbiakban felhasználva, példát adunk néhány igen gyakran előforduló eset maszkokkal való detektálhatóságára.

### 2.2.1. Pontdetektálás

A pontdetektálás célja a környezetüktől erősen eltérő intenzitású (izolált vagy kontúr-) képpontok megkeresése. Az ehhez használható egyszerű maszkot és annak egy alkalmazását a 2.2. ábra mutatja.



2.2. ábra. Izolált- és határpontok meghatározása maszkkal; (a) detektáló maszk, (b) eredeti kép, (c) maszkolás eredménye.

A  $\langle W^T, X \rangle$  maszkválasz nagysága jól mutatja a vizsgált képpont izoláltságának mértékét. Így például egy megfelelő  $K > 0$  küszöböt használva,  $|\langle W^T, X \rangle| > K$  esetén a képpontot izoláltnak tekinthetjük. Vegyük észre továbbá azt is, hogy a maszksúlyok összege 0, azaz az izoláltság tulajdonsága konstans intenzitású területen a leggyengébb.

### 2.2.2. Vonaldetektálás

A vonaldetektálás célja a környezetüktől eltérő intenzitású vonalszegmensek meghatározása. Jellemzően a négy fő (vízszintes, függőleges és két átlós) irány detektálása a cél, amelyet a 2.3. ábra maszkjaival végezhetünk el.

1	1	1
-2	-2	-2
1	1	1

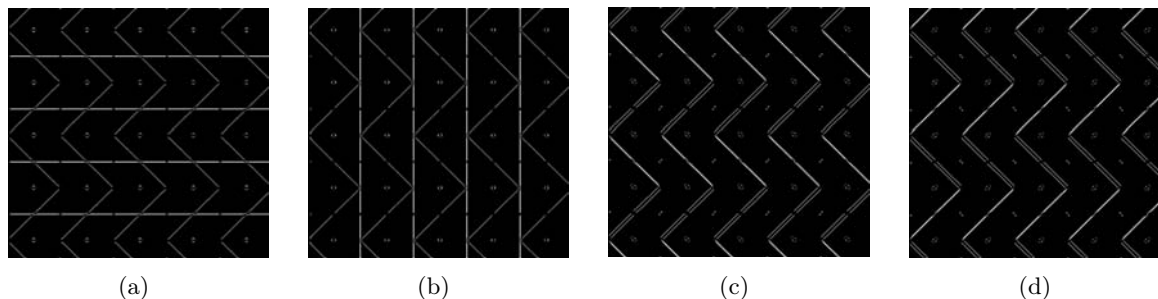
1	-2	1
1	-2	1
1	-2	1

-2	1	1
1	-2	1
1	1	-2

1	1	-2
1	-2	1
-2	1	1

2.3. ábra. Vonaldetektáló maszkrendszer a négy fő irányban.

A maszksúlyok összege most is 0, továbbá a pontdetektáláshoz hasonlóan újfent célszerű küszöbölést alkalmazni. Nevezetesen, azoknak a  $W_i$  maszkoknak ( $i = 1, 2, 3, 4$ ) a választát tekintjük relevánsnak, amelyekre  $|\langle W_i^T, X \rangle| > K$  (pozitív küszöb). Ha csak a legjellemzőbb vonalirányt keressük, értelemszerűen a  $\max_{i=1}^4 \{|\langle W_i^T, X \rangle|\}$  maszkválasszal kell dolgoznunk. A vonaldetektáló maszkok 2.2(b) képre adott eredményét a 2.4. ábrán tekinthetjük meg. Vegyük észre, hogy az egyes maszkok valóban a vizsgált irányokra adják a legerősebb választ.



2.4. ábra. Vonalak meghatározása maszkokkal; (a) vízszintes vonalak, (b) függőleges vonalak, (c)-(d) átlós vonalak.

A digitális egyenes definiálható a 8-összefüggőség megkövetelése nélkül is. Ebben az esetben az  $y = mx + b$  ( $m, b \in \mathbb{R}$ ) digitális egyenest azon  $(z_1, z_2) \in \mathbb{Z}^2$  pontok alkotják, amelyekre  $z_2 = mz_1 + b$ . Az így definiált, illetve a 8-összefüggő, de nem a négy főirányban haladó egyenesek detektálásához célszerű  $3 \times 3$  méretűnél nagyobb maszkot használni.

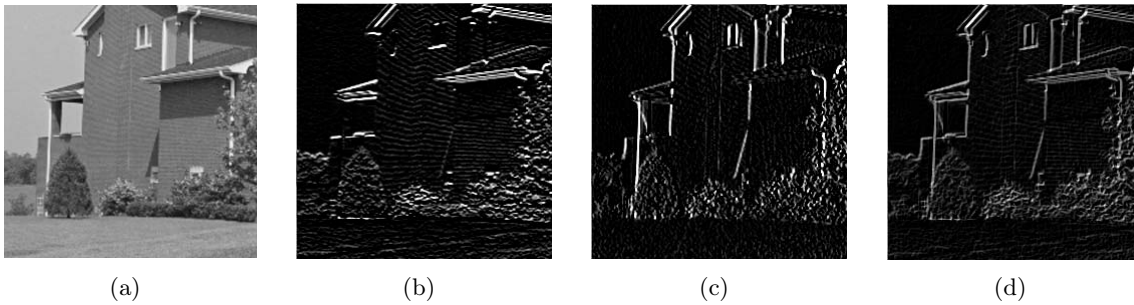
### 2.2.3. Éldetektálás

Az 1.3.1. szakaszban bemutattuk, hogyan működnek a klasszikus,  $2 \times 2$ -es környezeten alapuló éldetektáló eljárások. Ugyanerre a feladatra  $3 \times 3$ -as maszkokat is használhatunk, mint például a 2.5. ábrán látható ötelemes  $(-2, -1, 0, 1, 2)$  elemekből álló maszkrendszer, amely a négy fő irányban detektál éleket.

-1	-2	-1	-1	0	1	0	-1	-2	-2	-1	0
0	0	0	-2	0	2	1	0	-1	-1	0	1
1	2	1	-1	0	1	2	1	0	0	1	2

2.5. ábra. Éldetektáló  $3 \times 3$ -as maszkrendszer a négy fő irányban.

A 0 súlyösszeg megtartásával definiálhatók más maszkrendszerek is, lásd [27]. Az éltulajdonság meglétét a vizsgált pontban a vonaldetektálásnál leírt küszöbölést alkalmazva ellenőrizhetjük. A 2.6(b)-(c) ábrán az ötelemes maszkkészlet vízszintes és függőleges éleket detektáló maszkjaihoz tartozó eredményeket láthatjuk. Az összes él meghatározásához tekinthetjük például ennek a két képnek az átlagát vagy maximumát (lásd 2.6(d) ábra).



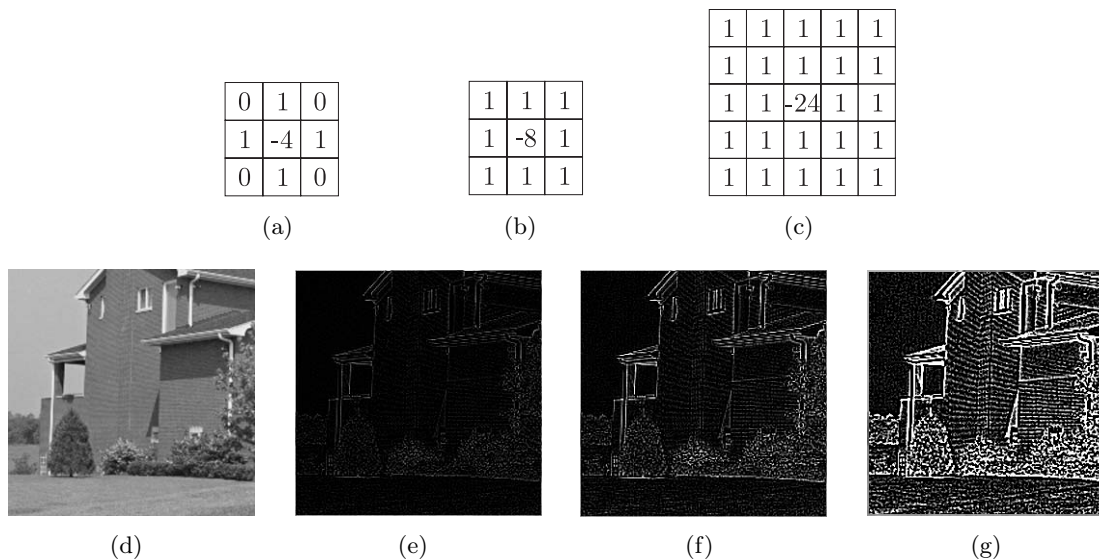
2.6. ábra. Éldetektálás  $3 \times 3$  méretű maszkokkal; (a) eredeti kép, (b) vízszintes élek, (c) függőleges élek, (d) a vízszintes és függőleges élképek átlaga.

Az első derivált közelítése mellett az éldetektálásban jól használhatók a Laplace operátort (második deriváltat) közelítő  $3 \times 3$ -as maszkok is. A két alapvető szomszédsági viszonyt megfelelően ennek két formája használatos, de léteznek nagyobb méretű maszkok is, amelyek a pont tágabb környezetét tekintik, lásd 2.7. ábra.

Vegyük észre, hogy az első deriváltat közelítő éldetektáló maszkokkal ellentétben a Laplace maszkok forgatásinvariánsak. A 2.8. ábrán látható Laplace maszkokkal változtathatjuk a centrális képpont és a környezet súlyviszonyát, ám egyúttal elveszítjük a forgatásinvarianciát.

#### 2.2.4. Kombinált maszkrendszerek

A kombinált maszkrendszerek használatával egy maszkrendszerrel több tulajdonság detektálását is végezhetjük. A maszkrendszeren belül egy tulajdonság detektálásához több maszk is tartozhat, amelynek megfelelően a maszkokat csoportosítjuk. A célunk az, hogy eldöntsük, a több vizsgált tulajdonságok közül melyik a legjellemzőbb az adott pontra. Az eljárás helyes működéséhez feltesszük, hogy a maszkvektorok ortonormált rendszert alkotnak. Ekkor az egyes csoportokba tartozó vektorok az origótól eltekintve diszjunkt tulajdonsághaltereket feszítenek fel, és egyszerűen meghatározható a vizsgált képrészlet ezen alterekkel bezárt szöge.



2.7. ábra. Laplace operátor konvolúciós maszkjai és azok hatása; (a)  $L_4$  maszk, (b)  $L_8$  maszk, (c)  $5 \times 5$ -ös Laplace maszk, (d) eredeti kép, (e)  $L_4$  maszk hatása, (f)  $L_8$  maszk hatása, (g)  $5 \times 5$ -ös maszk hatása.

2	-1	2
-1	-4	-1
2	-1	2

-1	2	-1
2	-4	2
-1	2	-1

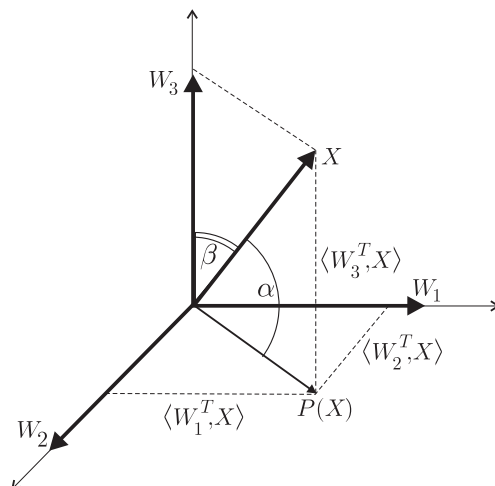
2.8. ábra.  $3 \times 3$ -as Laplace maszkok speciális súlyozással.

Természetes döntésként az adott pontra nézve azt a tulajdonságot tartjuk legvalószínűbbnek, amelynek tulajdonságaltérre a legkisebb szöveget zárja be a képrészlettel. A szög meghatározásához a maszkválasz műveletet használjuk. Az általános  $3 \times 3$  maszkok esete előtt az elvet a 2.9. ábrán is látható 3D-s esetre mutatjuk be.

Legyenek  $W_1$ ,  $W_2$  és  $W_3$  ortonormált 3D-s vektorok és  $X$  a vizsgált képrészlet. Tegyük fel továbbá, hogy két tulajdonságot vizsgálunk, ahol  $W_1$  és  $W_2$  közösen az egyik, míg  $W_3$  a másik tulajdonságot detektálja. A feladat így a felfeszített 2D-s és 1D-s tulajdonságaltér  $X$ -szel bezárt szögének meghatározása. Vegyük észre, hogy a maszkvektorok ortonormáltsága miatt a maszkválaszokra

$$\langle W_i^T, X \rangle = \|W_i^T\| \|X\| \cos \theta_i = \|X\| \cos \theta_i, \quad i = 1, 2, 3$$

adódik, ahol  $\theta_i$  az  $X$  és a  $W_i$  által bezárt szög, ami éppen az  $X$ -nek a maszkvektorok tengelyeire eső merőleges vetületének a hossza. Így az  $X$  képrészletre, illetve annak a  $W_1$  és  $W_2$  által



2.9. ábra. Kombinált maszkrendszer 3D-ben két tulajdonságtérrel.

felfeszített altérre eső  $P(X)$  vetületére

$$\begin{aligned} \|P(X)\| &= (\langle W_1^T, X \rangle^2 + \langle W_2^T, X \rangle^2)^{\frac{1}{2}}, \\ \|X\| &= (\langle W_1^T, X \rangle^2 + \langle W_2^T, X \rangle^2 + \langle W_3^T, X \rangle^2)^{\frac{1}{2}}. \end{aligned}$$

Az  $X$  képrészlet  $W_1$  és  $W_2$ , illetve a  $W_3$  által felfeszített alterekkel bezárt  $\alpha$ , illetve  $\beta$  szöge tehát  $\alpha = \arccos \frac{\|P(X)\|}{\|X\|}$ , illetve  $\beta = \arccos \frac{\langle W_3^T, X \rangle}{\|X\|}$ .

Mivel a  $3 \times 3$ -as maszkok 9D-s vektorokba rendezhetőek, ezért a páronkénti ortogonalitás legfeljebb kilenc elemű kombinált maszkrendszereket enged meg. Tegyük fel, hogy a  $W_1, W_2, \dots, W_9$  maszkok a  $C_1, C_2, \dots, C_K$  ( $K \leq 9$ ) osztályba vannak csoportosítva az egyes tulajdonságok detektálásához, így  $\sum_{i=1}^K |C_i| = 9$  adódik. Indexeljük át a maszkokat úgy, hogy a  $C_j$  maszkjainak sorszámai a  $\left[1 + \sum_{i=1}^{j-1} |C_i|, \sum_{i=1}^j |C_i|\right]$  indexintervallumba essenek. Rögzítsünk egy 9D-s  $X$  képrészletet és jelölje  $\theta_i$  ( $i = 1, \dots, K$ ) ennek a  $C_i$  osztályba tartozó maszkok

által felfeszített altérrel bezárt szögeit. Ekkor a 3D-s esethez hasonlóan megmutatható:

$$\begin{aligned}\theta_1 &= \arccos \left( \frac{\sum_{i=1}^{|C_1|} \langle W_i^T, X \rangle^2}{\sum_{i=1}^9 \langle W_i^T, X \rangle^2} \right)^{\frac{1}{2}}, \\ \theta_2 &= \arccos \left( \frac{\sum_{i=|C_1|+1}^{|C_1|+|C_2|} \langle W_i^T, X \rangle^2}{\sum_{i=1}^9 \langle W_i^T, X \rangle^2} \right)^{\frac{1}{2}}, \\ &\vdots \\ \theta_K &= \arccos \left( \frac{\sum_{i=|C_1|+\dots+|C_{K-1}|+1}^{|C_1|+\dots+|C_K|} \langle W_i^T, X \rangle^2}{\sum_{i=1}^9 \langle W_i^T, X \rangle^2} \right)^{\frac{1}{2}}.\end{aligned}$$

A  $K$  tulajdonság közül azt a  $j$ -edik tulajdonságot tekintjük a legjellemzőbbnek a vizsgált pontban, amelyre  $\theta_j = \min_{i=1}^K \{\theta_i\}$ . A 2.10. ábrán egy kilenc elemű ortogonális maszkrendszert figyelhetünk meg, amely három tulajdonság (él, vonal és átlagosság) egyidejű detektálására alkalmas. Jól megfigyelhető, hogy a szigorúbb feltételrendszer miatt a maszk súlyok kevésbé optimálisak, mint a korábbi szakaszokban ugyanezen feladatokra egymástól függetlenül megadott maszkok esetében.

1	$\sqrt{2}$	1	1	0	-1	0	-1	$\sqrt{2}$	$\sqrt{2}$	-1	0	0	1	0	-1	0	1	1	-2	1	-2	1	-2	1	1	1	1
0	0	0	$\sqrt{2}$	0	$-\sqrt{2}$	1	0	-1	-1	0	1	-1	0	-1	0	0	0	-2	4	-2	1	4	1	1	1	1	1
-1	$-\sqrt{2}$	-1	1	0	-1	$-\sqrt{2}$	1	0	0	1	$-\sqrt{2}$	0	1	0	1	0	-1	1	-2	1	-2	1	-2	1	1	1	1

2.10. ábra. Három tulajdonság egyidejű detektálására alkalmas kilenc elemű maszkrendszer.

### 2.3. Matematikai morfológia

A *morfológia* szó a biológia szakterületén az állatok és növények alakjának, struktúrájának leírását jelenti. Ennek analógiájára a képfeldolgozásban a képi alakzatok reprezentációjára és leírására a *matematikai morfológia* eszköztárát használjuk. Ide sorolható az alakzatok határvonalának, vázának és konvex burkának vizsgálata, valamint utófeldolgozásként a morfológiai szűrés, vékonyítás és a kinövések eltávolítása.

A matematikai morfológia a halmazelméleten alapul, azaz a képi objektumokat halmazokkal reprezentáljuk. A bináris képeket 2D-s (képpontok helykoordinátái), míg az intenzitásképeket 3D-s (képpontok helykoordinátái és intenzitásai) halmazokkal reprezentálhatjuk. Az analóg (folytonos) és a digitális alakzatok morfológiai elmélete szépen harmonizál, ebben



a jegyzetben az egyszerűbb érthetőség kedvéért a fogalmak és eredmények elsősorban a 2D-s digitális térre való megfogalmazását közöljük. Tekintsük át először a morfológiai műveletek definiálásához szükséges halmazelméleti alapfogalmakat.

Legyenek  $A, B \subseteq \mathbb{Z}^2$  digitális halmazok és legyen  $x = (x_1, x_2) \in \mathbb{Z}^2$  egész vektor. Ekkor

- az  $A$  halmaz  $x$  vektorral való eltoltja

$$(A)_x = \{y \in \mathbb{Z}^2 \mid y = a + x, a \in A\},$$

- az  $A$  halmaz (középpontos) tükrözöttje

$$\hat{A} = \{y \in \mathbb{Z}^2 \mid y = -a, a \in A\},$$

- az  $A$  halmaz komplementere

$$A^c = \{y \in \mathbb{Z}^2 \mid y \notin A\},$$

- az  $A$  és  $B$  halmaz különbsége

$$A \setminus B = \{y \in \mathbb{Z}^2 \mid y \in A, y \notin B\} = A \cap B^c.$$

## 2.4. Dilatáció és erózió

A morfológiai műveletek tárgyalását két olyan eljárással (a dilatációval és erózióval) kezdjük, amelyek a későbbi összetett eljárások alapelemei. Ez a két művelet duális tulajdonságú, vagyis egyik a másikból származtatható, ám az alaphalmazra gyakorolt különböző hatásuk miatt külön szokták definiálni őket.

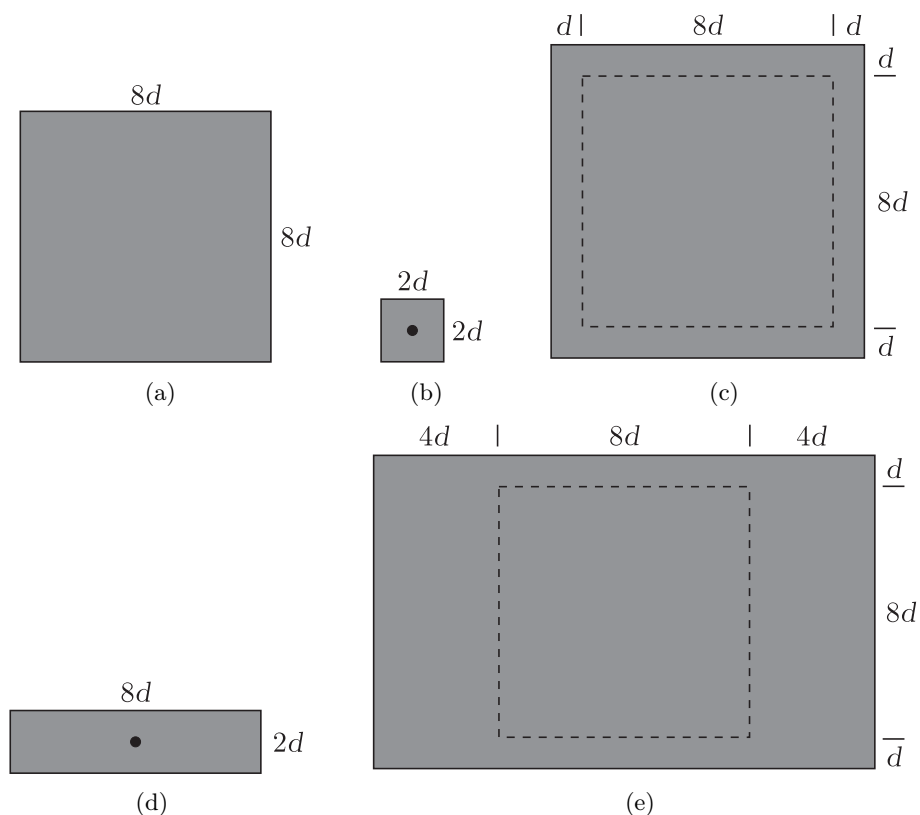
Legyen  $A, B \subseteq \mathbb{Z}^2$  és jelölje  $\emptyset$  az üres halmazt. Az  $A$  halmaz  $B$ -vel való dilatáltja alatt az

$$A \oplus B = \{x \mid [(\hat{B})_x \cap A] \neq \emptyset\}$$

halmazt értjük. A dilatáció során tehát először középpontosan tükrözzük a  $B$  halmazt, majd eltoljuk az  $x$  vektorral. Az  $A \oplus B$  dilatált meghatározásához ezután az összes olyan  $x$  eltolásvektort tekintjük, amelyekkel a  $\hat{B}$ -t eltolva annak még marad közös pontja az  $A$ -val. Ezt a gondolatmenetet felhasználva megmutatható, hogy

$$A \oplus B = \{x \mid [(\hat{B})_x \cap A] \subseteq A\}$$

is teljesül. A  $B$  halmazt, amivel az  $A$  halmazt dilatáljuk (növeljük) struktúraelemnek nevezük. Vegyük észre, hogy a  $B$  helyzete a dilatáció végeredményét csak egy eltolás mértékében befolyásolja, így feltehetjük, hogy a  $B$  tartalmazza az origót. Amennyiben a  $B$  centrálisan szimmetrikus, akkor a fenti összefüggésekben a  $\hat{B}$ -t  $B$ -re cserélhetjük. A gyakorlati alkalmazásokban használt struktúraelemnél igen gyakran a viszonylag kis méretet és a konvexitást is megkövetelik. A dilatációs eljárás hatását a 2.11. ábrán figyelhetjük meg.



2.11. ábra. A dilatació eredménye; (a) eredeti  $A$  alakzat, (b) adott  $B$  struktúraelem, (c) az  $A$  dilatáltja, (d) másik  $B$  struktúraelem, (e) az  $A$  dilatáltja erre nézve.

A dilataciónak léteznek egyéb (ekvivalens) definíciói is. Az itt közölt megközelítés azonban sok rokon vonást mutat a konvolúciós maszkok elméletével, bár a morfológia halmazműveleteken, míg a konvolúciós szűrés aritmetikai műveleteken alapul.

A dilatació duális műveletének, az eróziónak a definíciója a következő:

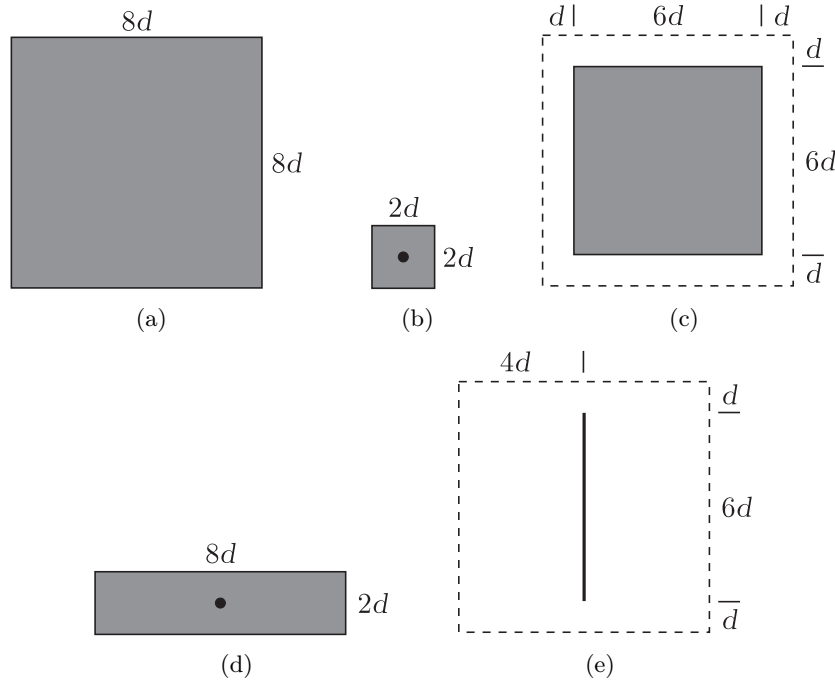
$$A \ominus B = \{x \mid (B)_x \subseteq A\}.$$

Így az  $A$  halmaz  $B$ -vel való erodáltja azon  $x$  vektorokat tartalmazza, amelyekkel  $B$ -t eltolva az még  $A$  belsejében marad. Erodáló eljárásra a 2.12. ábrán láthatunk példákat.

A dilatacióhoz hasonlóan az eróziónak is léteznek egyéb definiálási formái, ám az alkalmazásokban szintén a maszkoláshoz való hasonlósága miatt a fenti definíció a domináns. A dilatació és erózió duális viszonyára az

$$(A \ominus B)^c = A^c \oplus \hat{B}$$

összefüggés mutat rá, amit szemléletesen úgy fogalmazhatunk meg, hogy a halmaz erodáltja megegyezik a halmaz komplementere dilatáltjának komplementerével. Az állítás bizonyításá-



2.12. ábra. Az erózió eredménye; (a) eredeti  $A$  alakzat, (b) adott  $B$  struktúraelem, (c) az  $A$  erodáltja, (d) másik  $B$  struktúraelem, (e) az  $A$  erodáltja erre nézve.

ban (lásd [10]) jól megfigyelhető a két művelet közötti szoros kapcsolat. A duális viszony azért is fontos, mert segítségével könnyen beláthatók az egyik műveletre teljesülő állítások a másik műveletre nézve is.

### 2.4.1. Nyitás és zárás

Habár a dilatació és erózió duális műveletek (azaz a komplementerhalmazokon fordított a hatásuk), a két művelet nem inverze egymásnak, így a dilataciót követő erózió (ugyanazon) struktúraelemmel, illetve a műveletek fordított sorrendje nem adja vissza az alaphalmazt. Ennek belátásához tekinthetünk például egy egy pontból álló  $A$  halmazt, amelyet egy kezdő erózió teljes egészében töröl. (A fordított műveletsorrendre egy képpontnyi lyukat tekinthetünk.) A fentiek értelmében az  $A$  halmaz  $B$  struktúraelemmel történő nyitását

$$A \circ B = (A \ominus B) \oplus B,$$

míg a zárását

$$A \bullet B = (A \oplus B) \ominus B$$

módon definiáljuk. A nyitás simítja az objektum határát az azon lévő kisebb kinövések törlésével, és megszakítja a keskeny szálakat. A zárás szintén simítja a határokat, de a nyitással ellentétben az azon lévő kisebb hézagok betömésével, kitölti továbbá a kisebb lyukakat,

szakadásokat és keskeny csatornákat. Sokszor az döntheti hogy melyiküket használjuk egy alkalmazás során, hogy a nyitás csökkenti, míg a zárás növeli az alaphalmaz területét.

A nyitás és zárás nem véletlenül kapták a nevüket, mivel tulajdonságaik megfelelnek a halmaztopológiában használt nyitás és zárás fogalmának. Nevezetesen, a nyitás

$$\begin{aligned} A \circ B &\subseteq A && \text{(szűkítő),} \\ C \subseteq D &\implies C \circ B \subseteq D \circ B && \text{(monoton),} \\ (A \circ B) \circ B &= A \circ B && \text{(idempotens)} \end{aligned}$$

tulajdonságokkal, míg a zárás

$$\begin{aligned} A \bullet B &\supseteq A && \text{(bővítő),} \\ C \subseteq D &\implies C \bullet B \subseteq D \bullet B && \text{(monoton),} \\ (A \bullet B) \bullet B &= A \bullet B && \text{(idempotens)} \end{aligned}$$

tulajdonságokkal rendelkeznek.

Mivel a nyitás és zárás a dilatáción és erózióon alapul, ennek megfelelően a dualitás közöttük is fennáll, azaz a halmaz nyitása megegyezik a komplementerhalmaz zárásának komplementerével, illetve a halmaz zárása megegyezik a komplementerhalmaz nyitásának komplementerével azaz

$$(A \ominus B) \oplus B = ((A^c \oplus B) \ominus B)^c, \quad \text{illetve} \quad (A \oplus B) \ominus B = ((A^c \ominus B) \oplus B)^c,$$

ahol az egyszerűbb megfogalmazás kedvéért feltettük, hogy  $\hat{B} = B$ .

A gyakorlatban a dilatáció/erózió, illetve a nyitás/zárás műveletekben általában kis méretű ( $3 \times 3$ -as)  $B$  struktúraelemet használnak. Amennyiben ennél nagyobb változtatások szükségesek a képen (például a betömni kívánt lyukak viszonylag nagyok) akkor vagy nagyobb ( $5 \times 5$ -ös vagy  $7 \times 7$ -es) struktúraelemet kell választanunk illetve egy másik lehetőség, hogy a nyításban és zárásban a dilatációs és eróziós lépéseket egymás után többször alkalmazzuk az

$$(((A \ominus B) \ominus B) \oplus B) \oplus B, \quad \text{illetve} \quad (((A \oplus B) \oplus B) \ominus B) \ominus B$$

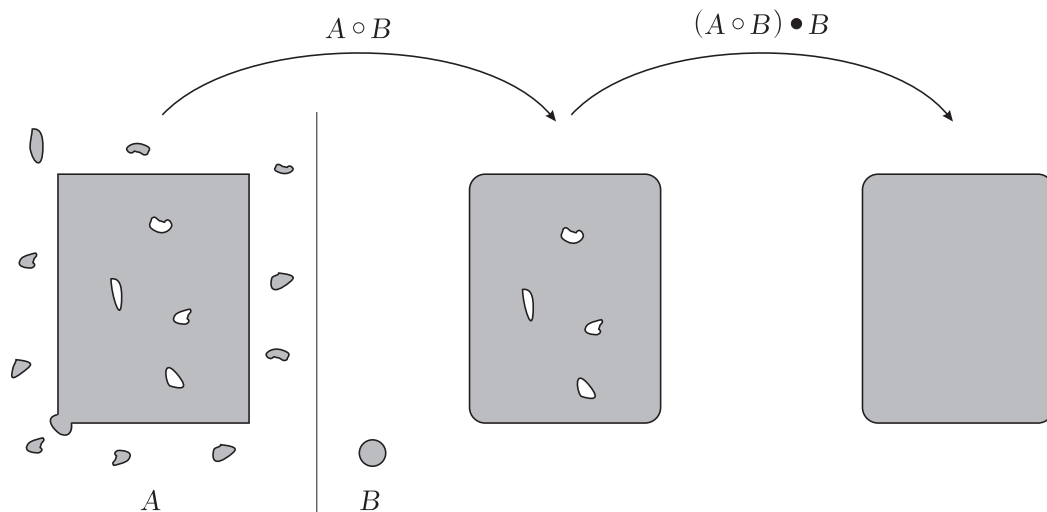
módon. Ügyeljünk arra, hogy a kiinduló halmaz méretének hozzávetőleges megőrzéséhez ilyenkor ugyanolyan számú dilatációs és eróziós lépést végezzünk. Kombiálhatjuk továbbá a nyitás és a zárás műveletét is egy lépésben

$$(A \circ B) \bullet B,$$

amivel mindkét módszer hatását érvényesíthetjük a képre. Ez utóbbi eljárást a 2.13. ábrán figyelhetjük meg.

#### 2.4.2. Hit-Miss transzformáció

A Hit-Miss transzformáció egy igen általános morfológiai eszköz, amelyből például a már bemutatott erózió és dilatáció is származtatható. Az általánosabb megfogalmazás mellett a Hit-Miss transzformációt igen széles körben, például alakzatdetektálásra is használhatjuk.



2.13. ábra. A morfológiai nyitás és zárás kombinálása.

A transzformációhoz használt  $B$  struktúraelemről most feltesszük, hogy  $B = B_1 \cup B_2$  alakú, ahol a  $B_1$  és  $B_2$  halmazok diszjunktak, azaz  $B_1 \cap B_2 = \emptyset$ . Ekkor az  $A$  halmaz  $B$  halmazzal való Hit-Miss transzformálján az

$$A \circledast B = (A \ominus B_1) \cap (A^c \ominus B_2)$$

halmazzal értjük. Szemléletesen, a  $B_1$ , illetve  $B_2$  halmaz segítségével a vizsgálni kívánt objektum-, illetve háttérpontok helyét írhatjuk elő a  $B$  struktúraelemen belül. Halmazműveleteket és a dilatació/erózió dualitását használva a fenti összefüggés

$$A \circledast B = (A \ominus B_1) \setminus (A^c \oplus \hat{B}_2)$$

alakra hozható, vagyis  $A \circledast B$  olyan pontokat tartalmaz, amelyekre nézve  $B_1$  illeszkedik az  $A$ -ra,  $B_2$  pedig az  $A^c$ -re.

Az erózió is speciális Hit-Miss transzformációnak tekinthető a  $B_1 = B$ ,  $B_2 = \emptyset$  választással, mivel ekkor a definíció az

$$A \circledast B = (A \ominus B_1) = A \ominus B$$

alakra egyszerűsödik.

A Hit-Miss transzformáció egy  $X$  alakzat detektálásához a következő módon használható. Tekintsünk egy  $X$ -et tartalmazó  $W$  ablakot és legyen  $B_1 = X$ ,  $B_2 = W \setminus X$ . Ekkor az

$$A \circledast B = (A \ominus X) \cap [A^c \ominus (W \setminus X)]$$

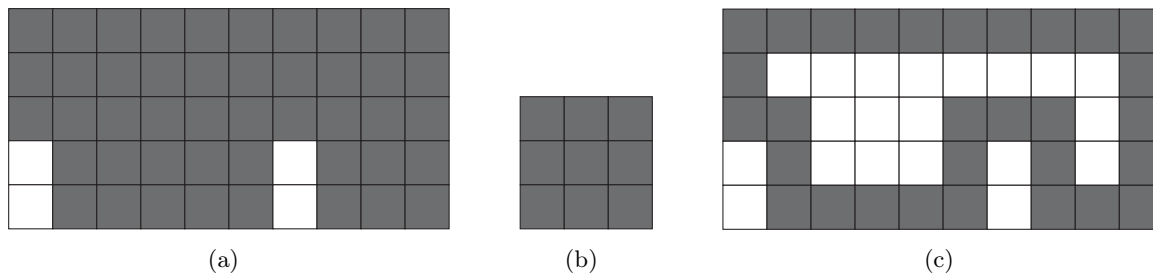
Hit-Miss transzformált éppen az  $X$  alakzat(ok) középpontjait tartja meg.

### 2.4.3. Kontúr kinyerése

Az  $A$  halmaz  $\partial A$ -val jelölt határát az  $A$  egy  $B$  struktúraelemmel való erodálásával, majd az így kapott eredményhalmaz  $A$ -ból való kivonásával kaphatjuk meg:

$$\partial A = A \setminus (A \ominus B).$$

Véges vizsgált halmaz esetén a halmaz határára helyezett struktúraelem egyes részei "lelőgnak". Ilyenkor a legegyszerűbb feltételezni, hogy ezek az elemek 0 értékűek (háttérpontok). Az eljárás eredményét a 2.14. ábrán figyelhetjük meg.



2.14. ábra. Kontúr kinyerése morfológiai műveletekkel; (a) eredeti  $A$  alakzat, (b) a felhasznált  $B$  struktúraelem, (c) az  $A$  halmaz  $\partial A$  kontúrja.

### 2.4.4. Lyukak kitöltése

Gyakran felmerülő probléma, hogy az alakzatok belseje olyan lyukakat tartalmaz, amelyeket ki szeretnénk tölteni. Az ezt megoldó iteratív eljárás első lépéseként tekintsük az  $A$  alakzatban lévő lyuk egy  $p$  pontját. Mivel az eljárás lényege, hogy a lyuk pontjait objektumpontokká alakítjuk, ezért tekintsük  $p$ -t is annak, és legyen  $X_0 = \{p\}$ . Ezt követően ismételjük az

$$X_k = (X_{k-1} \oplus B) \cap A^c$$

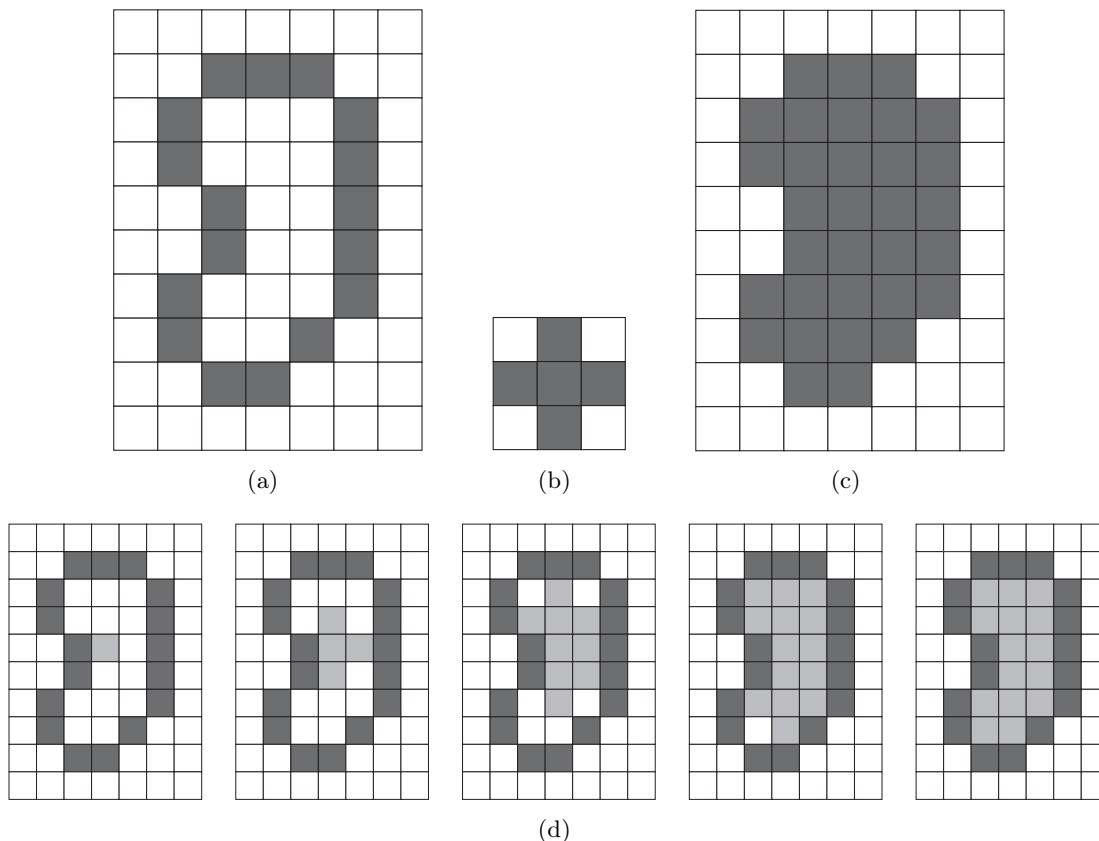
iterációs lépést a  $k = 1, 2, 3, \dots$  értékekre. Amennyiben az  $A$  alakzat véges és 8-összefüggő határral rendelkező halmaz volt, úgy a fenti iteráció véges sok lépésben betömi az abban lévő lyukakat szimmetrikus  $B$  struktúraelem használata esetén (lásd 2.15. ábra).

Mivel az iterációban használt dilataciós lépés önmagában a teljes teret kitöltené, ezért ennek megakadályozásához alkalmazzuk az  $A$  komplementerével való metszetképzést. Az ilyen jellegű morfológiai műveleteket feltételes dilatacióknak nevezzük.

### 2.4.5. Összefüggő komponensek meghatározása

Jelölje  $Y$  az  $A$  halmaz egy összefüggő komponensét, és legyen  $p \in Y$ . Ekkor a következő iterációs eljárással határozhatjuk meg  $Y$  összes elemét:

$$X_k = (X_{k-1} \oplus B) \cap A, \quad k = 1, 2, 3, \dots,$$



2.15. ábra. Lyukak kitöltése morfológiai műveletekkel; (a) eredeti  $A$  alakzat, (b) a felhasznált  $B$  struktúraelem, (c) a kitöltő eljárás végeredménye, (d) az iteráció lépései.

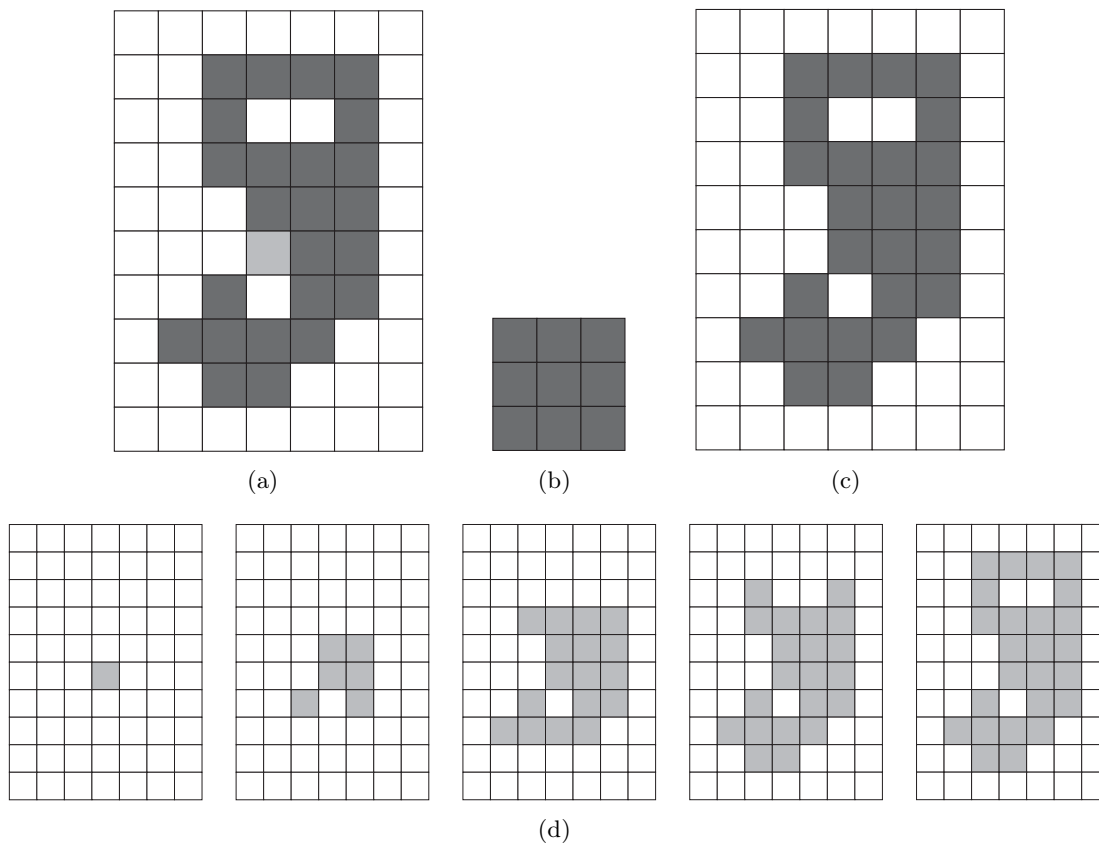
ahol  $X_0 = \{p\}$ ,  $B$  pedig egy megfelelő struktúraelem (lásd 2.16(c) ábra). Amennyiben valamely  $k$ -ra  $X_{k-1} = X_k$ , akkor a keresett komponens  $Y = X_k$ .

A lyukak betöltésére és az összefüggő komponensek meghatározására javasolt iterációs lépések nagyon hasonlóak. Az egyetlen különbség, hogy az utóbbi esetben az  $A$  halmazzal végezzük a metszetképzést annak komplementere helyett, mivel ekkor csak objektumpontokat keresünk.

#### 2.4.6. Konvex burkok meghatározása

Az alakzatok konvex burkát használhatjuk közvetlenül azok reprezentációjára, ám a konvexitás sok más eljárásban is nagyon hasznos tulajdonság. Az alábbiakban ismertetünk egy  $A$  alakzat  $C(A)$  konvex burkának előállítására használható egyszerű morfológiai algoritmust. Az eljárás az

$$X_k^i = (X_{k-1}^i \otimes B^i) \cup A, \quad i = 1, 2, 3, 4 \text{ és } k = 1, 2, 3, \dots$$



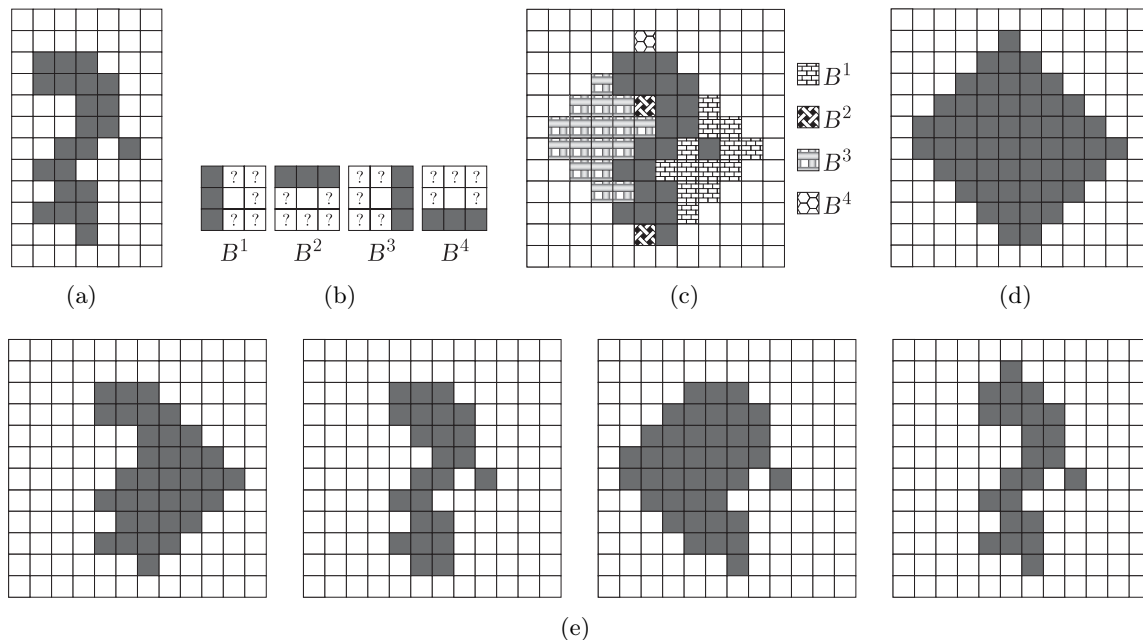
2.16. ábra. Összefüggő komponensek keresése morfológiai műveletekkel; (a) eredeti  $Y$  komponens kiindulóponttal, (b) a felhasznált  $B$  struktúraelem, (c) az eljárás végeredménye, (d) az iteráció lépései.

egyenletek implementálásából áll, ahol  $X_0^i = A$ , a  $B^i$ -k ( $i = 1, 2, 3, 4$ ) pedig a négy fő iránynak megfelelő struktúraelemek. Jelölje  $D^i$  az  $i$  irányba való konvergenciát mutató halmazt, azaz amelyre  $D^i = X_{k-1}^i = X_k^i$  (valamely  $k$  indexre). Ekkor az  $A$  konvex burkát a

$$C(A) = \bigcup_{i=1}^4 D^i$$

összefüggéssel definiáljuk. Szemléletesen, az eljárásban először a  $B^1$  struktúraelemmel végzünk Hit-Miss transzformációt a  $D^1$  konvergenciahalmaz eléréséig, majd ugyanezt elvégezzük a  $B^2$ ,  $B^3$  és  $B^4$  struktúraelemmel is és a konvex burkot a konvergenciahalmazok uniójaként állítjuk elő. A 2.17. ábrán egy példán követhetjük végig az eljárás egyes lépéseit.





2.17. ábra. Konvex burok meghatározása morfológiai műveletekkel; (a) eredeti  $A = X_0^i$  alakzat, (b) a négy fő irányhoz tartozó struktúraelem, (c) az egyes struktúraelemek szerepe a  $C(A)$  előállításában (d) a  $C(A)$  konvex burok, (e) a négy irányhoz tartozó  $D^1 = X_5^1$ ,  $D^2 = X_1^2$ ,  $D^3 = X_5^3$ ,  $D^4 = X_1^4$  konvergenciahalmaz.

### 2.4.7. Vékonyítás

Az  $A$  halmaz  $B$  struktúraelemmel való  $A \otimes B$  vékonyítását a következő műveletként definiáljuk:

$$A \otimes B = A \setminus (A \circledast B) = A \cap (A \circledast B)^c.$$

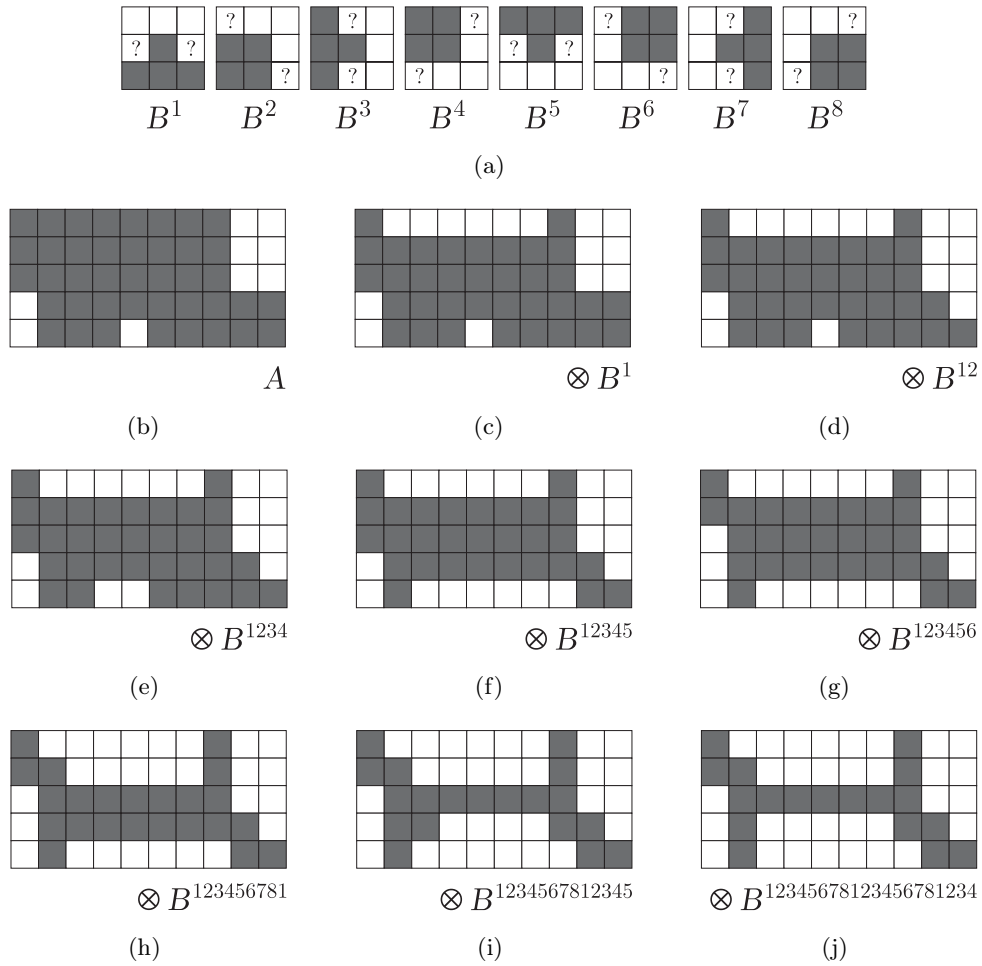
Az  $A$  szimmetrikusan (minden irányba) történő vékonyítását struktúraelemek

$$\{B\} = \{B^1, B^2, B^3, \dots, B^n\}$$

sorozatával végezhetjük el, ahol a  $B^i$  elemek egymás elforgatottjai. Ennek megfelelően a struktúraelemek sorozatával végzett vékonyítást a

$$A \otimes \{B\} = ((\dots ((A \otimes B^1) \otimes B^2) \dots) \otimes B^n)$$

módon értelmezzük. Így az  $A$  halmazon először a  $B^1$  elemmel hajtunk végre egy vékonyító lépést, ezt követően a  $B^2$ -vel, stb., majd végül a  $B^n$ -nel. A vékonyító eljárást mindaddig ismétljük, amíg változás történik. Egy vékonyításra alkalmas struktúraelem-sorozatot és az azokkal végrehajtott vékonyítás eredményét a 2.18. ábra tartalmazza. A struktúraelemek ?-lel jelzett pozícióiban objektum- és háttérpont egyaránt állhat.



2.18. ábra. Vékonyítás morfológiai műveletekkel; (a) struktúraelemek, (b) eredeti  $A$  alakzat, (c)-(j) a  $B^1, \dots, B^8$  struktúraelemekkel lépésenként végzett vékonyítás (csak a változások feltüntetésével).

### 2.4.8. Vastagítás

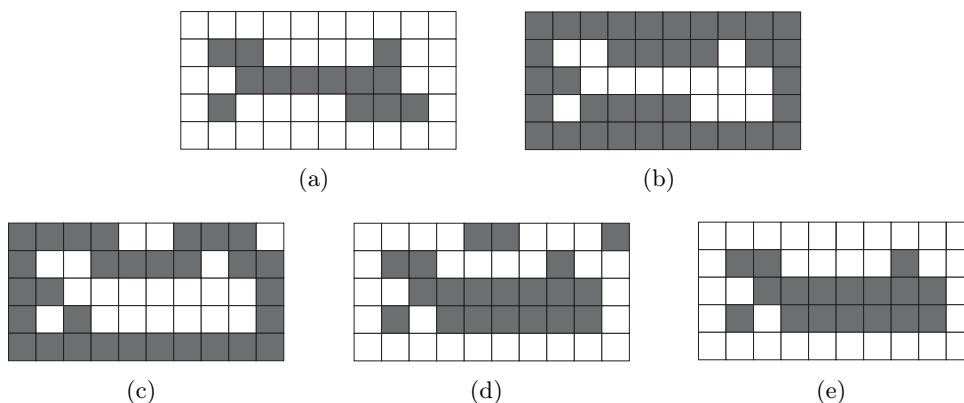
A morfológiai vastagítás a vékonyítás duális művelete, azaz a halmaz vastagítása megegyezik a komplementer halmaz vékonyításával, vagyis az  $A$  halmaz  $B$  struktúraelemmel történő vastagítása

$$A \odot B = A \cup (A \otimes B)$$

módon definiált. A vékonyításhoz hasonlóan természetesen a vastagítás is végezhető struktúraelemek

$$\{B\} = \{B^1, B^2, B^3, \dots, B^n\}$$

sorozatával. A dualitás miatt a vastagításban használt struktúraelemek megegyeznek a 2.18(a) ábrán lévőkkel, azzal a különbséggel, hogy fel kell cserélni az objektum-, illetve háttérpontok helyét. A gyakorlatban általában nem ezzel a megoldással találkozhatunk: gyakran az  $A$  halmaz komplementerét vékonyítják, majd az így kapott eredmény komplementerét tekintik. Ez az eljárás azonban az  $A$  természetétől függően izolált pontokat eredményezhet, amelyek kiszűrése általában egy utófeldolgozó lépésben valósul meg. A komplementer halmaz vékonyításán alapuló vastagítás lépéseit a 2.19. ábrán figyelhetjük meg.

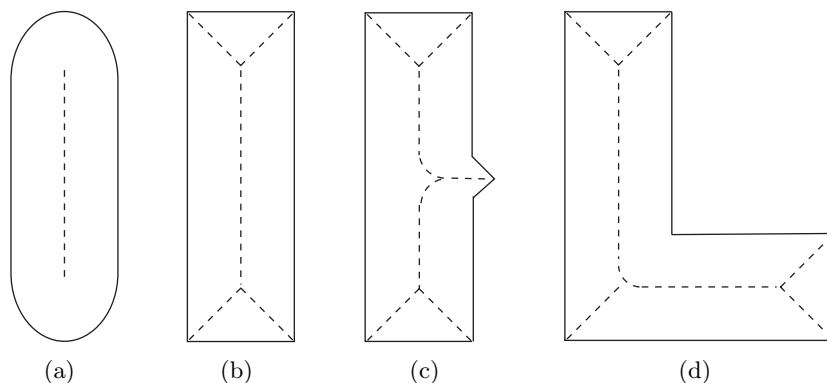


2.19. ábra. Vastagítás a komplementer vékonyításával; (a) eredeti  $A$  alakzat, (b) az  $A$  komplementere, (c) az  $A^c$  vékonyításának végeredménye, (d) az  $A^c$  vékonyításának komplementere, (e) az  $A$  vastagításának eredménye az izolált részek eltávolításával.

### 2.4.9. Vázkijelölés

A vázkijelölés sok digitális képfeldolgozási eljárás igen fontos építőköve. Fő erőssége abban rejlik, hogy segítségével az alakzatokat egyszerűbb és gazdaságosabb módon tudjuk leírni. A legfontosabb elvárások a definiált vázzal szemben, hogy egy képpont vastagságú legyen, jól kövesse az objektum alakját annak közepén haladva, és őrizze meg a topologikus viszonyokat. Az egy pont vastagságúra való zsugorítás vékonyítási technikákkal valósítható meg.

A váz elsőként a Blum [2] által 1967-ben javasolt középtengely-transzformáció (MAT) segítségével került definiálásra. Nevezetesen egy  $\partial T$  kontúrral rendelkező  $T$  tartomány minden  $p$  pontjához először meghatározzuk a  $\partial T$  legközelebbi pontját. Ha a  $\partial T$  egynél több ilyen pontot tartalmaz, akkor a  $p$  pontot a  $T$  középtengelye egyik pontjának tekintjük. Más-képp fogalmazva, a  $T$  középtengelyét azok a pontok alkotják, amelyek köré rajzolt és a  $T$  halmazban haladó maximális sugarú kör legalább két pontban érinti a  $T$  határát. Az eljárást "préritúz" hatásnak is szokták nevezni, mivel a  $\partial T$  pontjaiban "tüzeket" gyűjtva az egyes hullámfrontok éppen a középtengelyen fognak összeérni. Vegyük észre, hogy a középtengely alakja függ a távolságméréshez használt metrikától. A 2.20. ábrán néhány (folytonos térbeli) alakzatot és azok középtengelyeit láthatjuk az euklideszi távolság használata esetén.



2.20. ábra. A középtengely, mint a maximális sugarú körök középpontjai; (a)-(d) egyszerű alakzatok középtengelyei.

A középtengely meghatározása általában nem a definíció közvetlen alkalmazásával történik, mivel ez így meglehetősen költséges eljárás lenne. Ehelyett használhatunk határpontokat törölő vékonyító algoritmusokat, amelyek (1) nem törölnek végpontokat, (2) nem változtatják meg az alakzat összefüggőségét és (3) nem okozzák az alakzat túlzott erózióját.

Egy ilyen eljárás bemutatásához rendeljük az 1 értéket az alakzat pontjaihoz, illetve a 0 értéket az annak komplementeréhez tartozó pontokhoz. Az eljárás két fő lépést végez el az alakzat kontúrponthaira, vagyis azokra az 1 értékű pontokra, amelyek 8-környezetében van 0 értékű pont is. A 2.21. ábra jelöléseit használva, a  $p_1$  pontot törlésre jelöljük, ha teljesül az alábbi négy feltétel:

$$(i) \quad 2 \leq N(p_1) \leq 6;$$

$$(ii) \quad S(p_1) = 1;$$

$$(iii) \quad p_2 \cdot p_4 \cdot p_6 = 0;$$

$$(iv) \quad p_4 \cdot p_6 \cdot p_8 = 0;$$

ahol

$$N(p_1) = p_2 + p_3 + \dots + p_8 + p_9$$

a  $p_1$  pont nem 0 értékű 8-szomszédainak a száma,  $S(p_1)$  pedig a 0 – 1 átmenetek száma a  $p_2, p_3, \dots, p_8, p_9, p_2$  sorozatban. Így például az

$$\begin{array}{ccc} 1 & 0 & 1 \\ 0 & p_1 & 0 \\ 1 & 1 & 1 \end{array}$$

képrészlet esetén  $N(p_1) = 5$  és  $S(p_1) = 3$ .

Az eljárás második lépésében az (i) és (ii) feltétel ugyanaz, az (iii) és (iv) helyett azonban az

$p_9$	$p_2$	$p_3$
$p_8$	$p_1$	$p_4$
$p_7$	$p_6$	$p_5$

2.21. ábra. A vázkijelölő vékonyító eljárásban használt szomszédság.

$$(iii') \quad p_2 \cdot p_4 \cdot p_8 = 0;$$

$$(iv') \quad p_2 \cdot p_6 \cdot p_8 = 0;$$

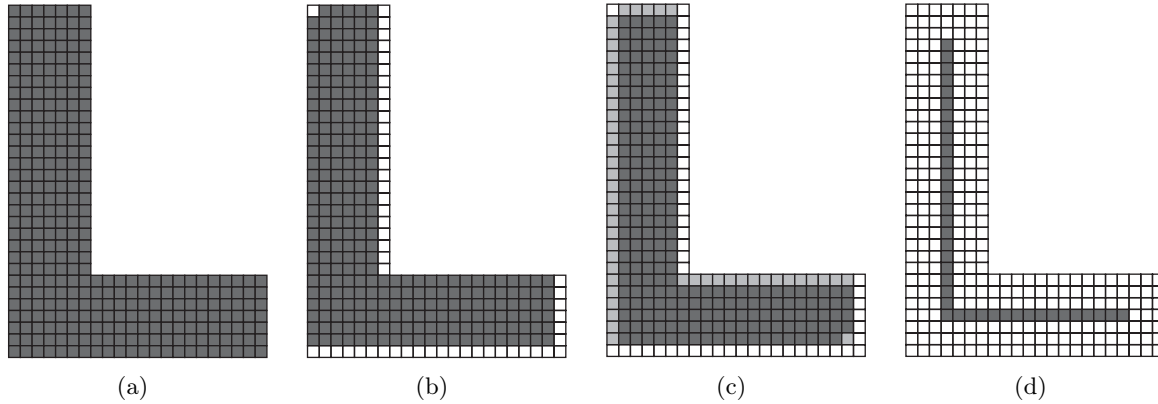
feltételeket szabjuk.

Alkalmazzuk az első lépést az eredeti alakzat minden kontúrponjtára. Ha az (i)-(iv) feltételek legalább egyike nem teljesül, akkor az aktuálisan vizsgált  $p_1$  pontot nem változtatjuk. Ha az összes feltétel teljesül, akkor a  $p_1$ -et megjelöljük törlésre, de a struktúra megőrzéséhez addig nem töröljük ténylegesen, amíg a lépést az összes kontúrpontra végre nem hajtottuk. Miután az első lépés az összes kontúrpontra lefutott, töröljük a megjelölt pontokat (vagyis az értéküket 0-ra állítjuk). Ezután a második lépést végezzük el az első lépéssel teljesen egyező módon. Összegezve, a vékonyító eljárás (1) végrehajtja az első lépést a kontúrpontra és megjelöli a törlendő pontokat, (2) törli a megjelölt pontokat, (3) végrehajtja a második lépést a megmaradt kontúrpontra és megjelöli a törlendőket, (4) törli a megjelölt pontokat. A váz kinyeréséhez az eljárást mindaddig iteratív módon folytatjuk, amíg már nem tapasztalunk változást.

Az (i) feltétel akkor sérül, ha a  $p_1$  pontnak nulla, egy vagy hét darab 8-szomszédja van. Nulla szomszéd esetén a  $p_1$  izolált, míg egy szomszéd esetén a  $p_1$  végpont, tehát a törlése nyilvánvalóan nem indokolt. A hét szomszédal rendelkező  $p_1$  pont törlése pedig az alakzat túlzott erózióját okozná. Az (ii) feltétel megóvjá az egy képpont vastagságú szegmens elhelyezkedő  $p_1$  törlését, mivel az megszüntetné a korábban meglévő összefüggőséget. Az (iii) és (iv) feltétel egyidejűleg akkor teljesül, ha ( $p_4 = 0$  vagy  $p_6 = 0$ ), illetve ( $p_2 = 0$  és  $p_8 = 0$ ). Az (i) és (ii) feltételeknek is a figyelembe vételével ez azt jelenti, hogy  $p_1$  vagy egy keleti/déli kontúrpon, vagy pedig egy észak-nyugati sarokpont, így a törlése indokolt. Hasonlóan látható, hogy a második lépés (iii') és (iv') feltétele akkor teljesül, ha ( $p_2 = 0$  vagy  $p_8 = 0$ ), illetve ( $p_4 = 0$  és  $p_6 = 0$ ), vagyis  $p_1$  vagy egy északi/nyugati kontúrpon, vagy pedig egy dél-keleti sarokpont. A 2.22. ábrán a vékonyító eljárás első két lépése és végeredménye látható.

Lantuejoul [15] (lásd még Serra [26], illetve Maragos [16]) már 1980-ban megmutatta, hogy egy  $A$  alakzat váza előállítható morfológiai eróziók és nyitások segítségével. Így az  $A$  halmaz  $S(A)$  váza előáll az

$$S(A) = \bigcup_{k=0}^K S_k(A)$$



2.22. ábra. Vázkijelölés vékonyítással; (a) eredeti alakzat, (b) az első iterációs lépés, (c) a második iterációs lépés, (d) végeredmény.

alakban, ahol

$$S_k(A) = \bigcup_{k=0}^K \{(A \ominus kB) - [(A \ominus kB) \circ B]\}$$

megfelelő  $B$  struktúraelemmel. Itt  $(A \ominus kB)$  az  $A$  halmaz  $B$ -vel  $k$  egymást követő lépésben történő erózióját jelenti:

$$(A \ominus kB) = \underbrace{((\dots (A \ominus B) \ominus B) \ominus \dots) \ominus B}_k,$$

továbbá  $K$  azt a lépésszámot, amelyre az  $A$  erodáltja még nem tűnik el, azaz

$$K = \max\{k \mid (A \ominus kB) \neq \emptyset\}.$$

Látható, hogy az  $A$  halmaz  $S(A)$  váza az  $S_k(A)$  részvázak uniójaként áll elő. Teljesül továbbá, hogy  $A$  ezekből a részvázakból az

$$A = \bigcup_{k=0}^K (S_k(A) \oplus kB)$$

módon állítható vissza, ahol  $(S_k(A) \oplus kB)$  az  $S_k(A)$  részváz  $B$ -vel  $k$  egymást követő lépésben való dilatacióját jelenti:

$$(S_k(A) \oplus kB) = \underbrace{((\dots (S_k(A) \oplus B) \oplus B) \oplus \dots) \oplus B}_k,$$

ahol a korábbiakhoz hasonlóan az összeadás legfeljebb  $K$  tagú.

### 2.4.10. Kinövések eltávolítása

A vékonyító és vázkijelölő algoritmusok végeredményeképpen előálló váz az eredeti objektum alkotóinak zajos vagy nem egyenletes volta miatt gyakran tartalmaz kisebb kinövéseket (lásd például 2.20. ábra). Ezek eltávolításához célszerű egy utófeldolgozó lépést tenni a vázkijelölést követően. Használhatunk egy vékonyításhoz hasonló eljárást, amely csak a végpontoktól visszafelé töröl. Ehhez a módszerhez inicializáló paraméterként célszerű megadnunk, hogy maximum milyen hosszúságú szegmenseket tekintünk kinövéseknek (a 2.23. ábrán mutatott példánkban a három hosszúságú kinövéseket figyeljük). A végpontoknál való vékonyításhoz tekinthetjük például a 2.23(b) ábrán látható  $B^i$  ( $i = 1, \dots, 8$ ) struktúraelemeket, így első lépésként végrehajtuk az

$$X_1 = A \otimes \{B\}$$

morfológiai vékonyítást, ahol  $B$  a  $B^i$  struktúraelemekből képzett sorozatot jelöli. A fenti vékonyításnak a 2.23(a) ábrán lévő eredeti alakzatra való háromszori szekvenciális alkalmazásnak  $X_1$  végeredményét a 2.23(c) ábrán láthatjuk. Mivel a vékonyítás nem csak a kinövéseket, hanem a fontos vázkomponenseket is rövidítette, az eljárás következő lépésének a feladata az eredeti váz helyreállítása a kinövések elhagyásával. Ehhez először elkészítjük az  $X_1$  végpontjait tartalmazó  $X_2$  halmazt (lásd 2.23(d) ábra):

$$X_2 = \bigcup_{k=1}^8 (X_1 \otimes B^k),$$

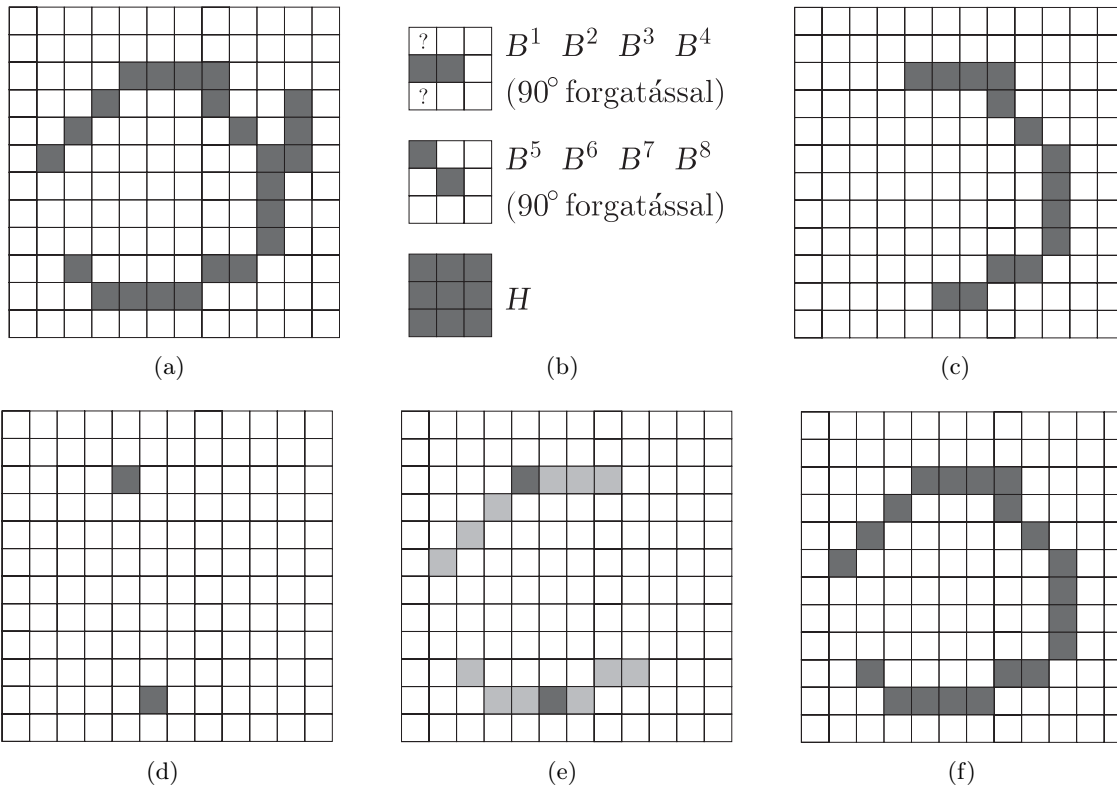
ahol a  $B^k$ -k a korábban használt struktúraelemeket jelentik. Ezután hajtsunk végre az  $X_2$ -beli végpontokra feltételes ( $A$ -val metszett) dilatációt a 2.23(b) ábrán látható szokásos  $3 \times 3$ -as  $H$  struktúraelemmel. Ez a feltételes dilatáció megakadályozza, hogy a vizsgálni kívánt tartományt elhagyjuk, amint ez a 2.23(e) ábrán is megfigyelhető. Az eljárás 2.23(f) ábrán látható végeredményét az

$$X_4 = X_1 \cup X_3$$

módon képezzük.

### 2.4.11. Golay ábécé

Ebben a szakaszban a morfológiai műveletek összefoglalásként bemutatjuk a Golay ábécé elemeit (lásd [9, 26], illetve 2.1. táblázat), amelyek a már ismert maszkok mellett újakat is tartalmaznak. A maszkolási eljárás a korábbi szakaszokban látottakhoz hasonlóan a kiindulási alakzat és a struktúraelemek közötti Hit-Miss transzformációkkal történik, ezért az új elemek használatát és az általuk szolgáltatott eljárást nem magyarázzuk el részletesen. A struktúraelemekben az 1 értékek az objektumpontokat, a 0 értékek a háttérpontokat jelölik, míg a ? helyeken akár 0, akár 1 állhat. Az ábécé elnevezést az indokolja, hogy minden morfológiai műveletnek egy betűt feleltetünk meg.



2.23. ábra. Kinövések eltávolítása morfológiai műveletekkel; (a) eredeti  $A$  váz, (b) az eljárásban használt struktúraelemek, (c) az  $A$  vékonyításának eredménye, (d) az  $A$  vékonyításának végpontjai, (e) a végpontok dilataciója a törölt vázszegmensek visszaállításához, (f) az eljárás végeredménye.

A morfológia műveletek szimmetrikus elvégzéséhez természetesen a megadott maszkok 90°-os elforgatottjait is alkalmaznunk kell. (A  $H$  és  $I$  betű esetében erre persze nincsen szükség.)

Morfológiai műveletek nem csak a négyzetrácson értelmezhetők. Így például a hatszögrács Golay ábécéjének fenti feladatokat ellátó elemeit [26]-ban megtalálhatjuk.



2.1. táblázat. A Golay ábécé elemei.

Golay betű	Struktúraelemek	Golay betű	Struktúraelemek
$H$ (erózió)	$\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$	$D$ (törlés)	$\begin{matrix} 0 & ? & ? & 0 & ? & ? \\ 0 & 1 & 1 & 0 & ? & 1 & ? & ? \\ 0 & ? & ? & ? & ? & 1 \\ & & & ? & & & & ? \end{matrix}$
$I$ (dilatáció)	$\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}$	$F$ (törlés)	$\begin{matrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & ? & 0 \\ 1 & 0 & 0 & 0 & ? & 1 \\ & & & 0 & & & & 0 \end{matrix}$
$E$ (törlés)	$\begin{matrix} ? & ? & ? \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix}$	$F2$ (törlés)	$\begin{matrix} 1 & 1 & 1 & 0 & ? & 1 \\ 0 & 1 & 0 & 0 & ? & 1 & ? & 1 \\ 0 & 0 & 0 & 0 & ? & 0 \\ & & & 0 & & & & 0 \end{matrix}$
$L$ (törlés)	$\begin{matrix} 0 & 0 & 0 & ? & ? & 0 \\ ? & 1 & ? & 1 & ? & 1 & ? & 0 \\ 1 & 1 & 1 & 1 & ? & ? \\ & & & 1 & & & & 0 \end{matrix}$	$K$ (törlés)	$\begin{matrix} 1 & 1 & 0 & ? & ? & 1 \\ ? & 1 & ? & ? & ? & 1 & ? & 0 \\ ? & ? & ? & ? & ? & ? \\ & & & ? & & & & ? \end{matrix}$
$M$ (törlés)	$\begin{matrix} 1 & ? & ? & 1 & ? & ? \\ 1 & 1 & 0 & 1 & ? & 1 & ? & ? \\ 1 & ? & ? & ? & ? & 0 \\ & & & ? & & & & ? \end{matrix}$	$C$ (növelés)	$\begin{matrix} 1 & ? & ? & 1 & ? & ? \\ 1 & 0 & ? & 1 & ? & 0 & ? & ? \\ 1 & ? & ? & ? & ? & ? \\ & & & ? & & & & ? \end{matrix}$

# Irodalomjegyzék

- [1] I. Aizenberg, C. Butakoff, D. Paliy, K. Egiazarian and J. Astola, Effective Detection and Elimination of Impulsive Noise with a Minimal Image Smoothing, In Proceedings of IEEE International Conference on Image Processing (ICIP'2003), vol.3, pages 357-360, 2003.
- [2] H. Blum, A transformation for extracting new descriptions of shape, In Wathen-Dunn, Weiant (Ed.), Symp. on Models for the Perception of Speech and Visual Form (Boston, Nov. 1964), M.I.T. Press, Cambridge, Mass.
- [3] M. Brady, Representing shape, In M. Brady, L.A. Gerhardt and H.F. Davidson, editors, Robotics and Artificial Intelligence, pages 279-300, Springer + NATO, Berlin, 1984.
- [4] J.F. Canny, Finding edges and lines in in images. Technical Report AI-TR-720, MIT, Artificial Intelligence Laboratory, Cambridge, MA, 1983.
- [5] J.F. Canny, A computational approach to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 8(6) (1986), 679-698.
- [6] J. Cooper, S. Venkatesh and L. Kitchen, Early jump-out corner detectors, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (1993), 823-828.
- [7] B. Everitt, Cluster Analysis, Heinemann Educational Books Ltd, London, 1973.
- [8] O.D. Faugeras, Three Dimensional Computer Vision: A Geometric Viewpoint. MIT Press, Cambridge, MA, 1993.
- [9] M.J.E. Golay, Hexagonal parallel pattern transformation, IEEE Transactions on Computers C-18 (1969), 733-740.
- [10] R.C. Gonzalez and R.E. Woods, Digital image processing, Addison-Wesley, Reading, MA, 1992.
- [11] R.M. Haralick, Digital step edges from zero crossing of the second directional derivatives, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (1984), 58-68.
- [12] R.M. Haralick and L.G. Shapiro, Computer and Robot Vision, Volume I, Addison-Wesley, Reading, MA, 1992.

- [13] T.S. Huang, editor, Image Sequence Processing and Dynamic Scene Analysis. Springer Verlag, Berlin, 1983.
- [14] D. Marr and E. Hildreth, Theory of edge detection, Proceedings of the Royal Society B 207 (1980), 187-217.
- [15] C. Lantuejoul, Skeletonization in quantitative metallography, In: Issues of Digital Image Processing, edited by R.M. Haralick and J.C. Simon Groningen, The Netherlands: Sijthoff and Noordhoff, 1980.
- [16] P. Maragos, Tutorials on advances in morphological image processing and analysis, Optical Engineering 26/7 (11) (1987), 623-632.
- [17] R. Mehrotra and S. Nichani, Corner detection, Pattern Recognition Letters 23 (11) (1990), 1223-1233.
- [18] H.P. Moravec, Towards automatic visual obstacle avoidance. In Proceedings of the 5th International Joint Conference on Artificial Intelligence, Carnegie-Mellon University, Pittsburgh, PA, August, 1977.
- [19] M. Nagao and T. Matsuyama, A Structural Analysis of Complex Aerial Photographs. Plenum Press, New York, 1980.
- [20] R. Nevatia, Evaluation of simplified Hueckel edge-line detector, Computer Graphics and Image Processing 6(6) (1977), 582-588.
- [21] N.R. Pal and S.K. Pal, A review on image segmentation techniques, Pattern Recognition 26 (1993), 1277-1294.
- [22] S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. Haar-Romeny, J.B. Zimmerman and K. Zuiderveld, Adaptive histogram equalization and its variations, Computer Vision, Graphics and Image Processing 39 (1987), 355-368.
- [23] W.K. Pratt, Digital Image Processing, Wiley, New York, 1978.
- [24] A. Rosenfeld and A.C. Kak, Digital Picture Processing, Academic Press, New York, 2nd edition, 1982.
- [25] A. Rosenfeld and J.L. Pfaltz, Distance functions on digital pictures, Pattern Recognition 1 (1968), 33-61.
- [26] J. Serra, Image analysis and mathematical morphology. Academic Press, New York, 1982.
- [27] M. Sonka, V. Hlavac and R. Boyle, Image processing, analysis, and machine vision. Brooks/Cole Publishing Company, Pacific Grove, CA, 1999.

- [28] D.C.C. Wang and A.H. Vagnucci, Gradient inverse weighting smoothing schema and the evaluation of its performance, *Computer Graphics and Image Processing* 15 (1981).
- [29] O. Zuniga and R.M. Haralick, Corner detection using the facet model. In *Computer Vision and Pattern Recognition*, pages 30-37, IEEE, Los Alamitos, CA, 1983.