

Fazekas Attila, Kormos János

Digitális képfeldolgozás matematikai alapjai

mobiDIÁK könyvtár

Fazekas Attila, Kormos János

Digitális képfeldolgozás matematikai alapjai

mobiDIÁK könyvtár

SOROZATSZERKESZTŐ

Fazekas István

Fazekas Attila, Kormos János

Digitális képfeldolgozás matematikai alapjai

egyetemi jegyzet
első kiadás

mobiDIÁK könyvtár
Debreceni Egyetem
Informatikai Intézet

Lektor

Fazekas Gábor

Copyright © Fazekas Attila, Kormos János, 2004

Copyright © elektronikus közlés mobiDIÁK könyvtár, 2004

mobiDIÁK könyvtár
Debreceni Egyetem
Informatikai Intézet
4010 Debrecen, Pf. 12
<http://mobidiak.inf.unideb.hu>

A mű egyéni tanulmányozás céljára szabadon letölthető. Minden egyéb felhasználás csak a szerző előzetes írásbeli engedélyével történhet.

A mű a *mobiDIÁK* önszervező mobil portál (IKTA, OMFB-00373/2003) és a *GNU Iterátor*, a legújabb generációs portál szoftver (ITEM, 50/2003) projektek keretében készült.

Tartalomjegyzék

I. Bevezetés	9
II. Alapvető fogalmak, definíciók	11
1. A digitális topológia elemei	11
2. Jól-kapcsolódó halmazok	20
III. Mintavételezés	23
1. Egydimenziós eset	23
2. Kétdimenziós eset	25
IV. Képtranszformációk	31
1. Fourier-transzformáció	31
2. Walsh-transzformáció	43
3. Hadamard-transzformáció	44
4. Diszkrét Cosinus-transzformáció	45
5. Hough-transzformáció	45
V. Képjavítások	47
1. Világosságkód-transzformáció	47
2. Szűrés	51
VI. Képrekonstrukció	55
1. Inverz szűrő	55
2. Wiener-szűrő	56
VII. Konvolúció és korreláció	59
1. Konvolúció	59
2. Korreláció	65
VIII. Szegmentálás	69
1. Gradiens módszerek	69

IX. Osztályozás	73
1. Korreláció	73
2. Statisztikus osztályozás	74
X. Lényegkiemelés	75
1. Analitikus elemzés	76
XI. Irodalomjegyzék	79

I. fejezet

Bevezetés

A vizuális információ számítógépes feldolgozásának mintegy 30 éves múltja van. A számítógéppel megoldott képfeldolgozási feladatok skálája nagyon széles, ezért ez a jegyzet nem tűzhet ki célul az olyan technikák elméleti hátterének ismertetését, amelyek optikai, analóg eszközöket használnak.

A digitális képfeldolgozás során arra törekszünk, hogy a képek elemzése révén fokozatosan megértsük, azaz helyesen értelmezzük a képben foglalt vizuális információkat; másszóval felismerjük, hogy mit ábrázol a kép.

A folyamat *fizikai szintjéhez* tartozó módszereket és eljárásokat *képátalakításoknak* nevezzük. A kiindulás valamilyen kép, amelyet először rögzíteni és digitalizálni kell. A cél az, hogy a vizuális vagy a gépi kiértékeléshez, illetve a további feldolgozáshoz az eredetinel kedvezőbb tulajdonságú képet nyerjünk. A célt a lényeges információkat megőrző átalakításokkal érjük el. Az eljárásokat globálisnak, illetve lokálisnak nevezzük attól függően, hogy a feldolgozás során egyidejűleg az összes képpont, illetve az egyes képpontoknak egy meghatározott környezetéhez tartozó képpontokat értékeljük-e ki. A képátalakítások két fő területe a következő:

- A *képkorrekciók* célja egyrészt kijavítani a leképezési hibákat, másrészt kiemelni a lényeges képi információtartalmat. A korrekciós eljárások két fő csoportba sorolhatjuk: *képjavítási* eljárások egyrészt a leképezés során óhatatlanul elszegényedő kontrasztosság fokozására és a zajok (kis területre kiterjedő képhibák) megszüntetésére irányul. A *képhelyreállítás* során olyan hibátlan kép előállítására törekszünk, amelyet egy hibátlan leképező rendszer produkált volna a kiinduló képből.
- A *szegmentálás* az értékes és a háttérpontok geometriai szétválasztását, vagyis az objektumok elkülönítését jelenti.

A folyamat *elemzési szintjén* végzett feldolgozást *képosztályozásnak* nevezzük, amelynek két fő területe az *alakfelismerés* és a *texturaelemzés*. A lényeges mozzanat mindkét területen az objektumok sajátosságai alapján végbemenő osztályozás.

- Alakfelismerésről beszélünk, ha az objektumokat a kép makrostruktúrájából származtattuk le, és az osztályozáshoz figyelembe vett sajátosságok is a makroszerkezettel kapcsolatosak.
- A texturaelemzés a kép mikroszerkezetének vizsgálatát jelenti. A texturák statisztikusan ismétlődő területelemek, amely elhanyagolhatóan kicsik ahhoz az alakzathoz képest, amelyet alkotják.

A digitális képfeldolgozás *legmagasabb szintje a képfelismerés*. Ennek során a képleírás alapján azonosítjuk az objektumot a valóságos tárgyakkal, és az egymáshoz viszonyított helyzetük, méretük alapján felismerjük, hogy mit ábrázol a vizsgált kép.

Az elmúlt években a digitális képfeldolgozás jelentősége megnövekedett. Ezt a fejlődést bizonyítja a számtalan alkalmazási lehetőség is, hiszen az ipari robottechnikától az orvosbiológiáig, az űrkutatástól a bűnüldözésig sorolhatnánk a képfeldolgozást felhasználó területek listáját.

Ez a jegyzet kilenc fejezetben ismerteti a képfeldolgozási algoritmusok elméleti hátterét, követve az immáron klasszikussá vált irodalmi felosztást.

II. fejezet

Alapvető fogalmak, definíciók

A digitális képfeldolgozás elmúlt negyedévszázados fejlődése során egyre erősödött az igény arra, hogy az egyedi gyakorlati problémák, feladatok megoldása mellett az általánosabb elméleti modellek matematikai eszközökkel történő vizsgálata is az előtérbe kerüljön. Ebben a fejezetben a digitális képfeldolgozás problémakörének ismertetéséhez szükséges alapvető fogalmakat és jelölési konvenciókat kívánjuk összegezni.

1. A digitális topológia elemei

1.1. Digitális kép

A $\mathbb{Z}^n$ -t n -dimenziós *digitális sík*nek, elemeit *pontoknak* nevezzük. Az n -dimenziós digitális sík nemüres részhalmazát n -dimenziós *digitális halmaz*nak nevezzük. Egy n -dimenziós X digitális halmaz felett értelmezett $f : X \rightarrow \{0, 1, \dots, m-1\}$ ($m \in \mathbb{N}, m \geq 2$) függvényt m -szintű *digitális kép*nek nevezzük. Az X halmazt az f *koordináta-halmazának*, a $\{0, 1, \dots, m-1\}$ halmazt az f *értékhalmazának* nevezzük. A $(p, f(p))$ rendezett párt ($p \in X$) *képpontnak* vagy *pixelnek* nevezzük, ahol p a *pixel koordinátája*, $f(p)$ a *pixel világosságkódja*.

Egy X digitális halmaz felett értelmezett 2-szintű f digitális kép esetén használjuk a *bináris kép* elnevezést is. A bináris kép *előtérén* az $F = \{p \in X \mid f(p) = 1\}$, a *hátterén* a $B = \{p \in X \mid f(p) = 0\}$ halmazt értjük. Az előtér pontjait *objektumpontoknak*, a háttér pontjait *háttérpontoknak* nevezzük.

A későbbiekben, ha egy n -dimenziós digitális halmaz p pontjának koordinátáit is fel kívánjuk tüntetni, akkor azt a $p_{(x_1, x_2, \dots, x_n)}$ alakban fogjuk megtenni. Speciálisan $n = 3$, illetve $n = 2$ esetben a $p_{(x, y, z)}$, illetve a $p_{(x, y)}$ jelölés is használatos.

Azonos X digitális halmaz felett értelmezett digitális (illetve a $\vee$ és a $\wedge$ műveletek esetén bináris) képek között a következőképpen értelmezhetünk

műveleteket:

$$\begin{aligned} f + g &= \{(p, h(p)) \mid h(p) = f(p) + g(p), p \in X\} \\ f * g &= \{(p, h(p)) \mid h(p) = f(p) \cdot g(p), p \in X\} \\ f \vee g &= \{(p, h(p)) \mid h(p) = f(p) \vee g(p), p \in X\} \\ f \wedge g &= \{(p, h(p)) \mid h(p) = f(p) \wedge g(p), p \in X\} \end{aligned}$$

Egy $f : X \rightarrow \{0, 1, \dots, m-1\}$ digitális képet *konstansnak* nevezünk, ha $f(p) = c$ ($c \in \{0, 1, \dots, m-1\}$) minden $p \in X$ -re.

Az f digitális képet *nullképnek*, illetve *egységképnek* nevezzük, ha $f(p) = 0$, illetve $f(p) = 1$ teljesül minden $p \in X$ -re.

1.2. Szomszédsági struktúrák

Legyen X egy digitális halmaz. Az (X, N) rendezett párt, ahol az N az X digitális halmazon értelmezett irreflexív, szimmetrikus binér reláció, *szomszédsági struktúrának* nevezünk.

Minden szomszédsági struktúrához hozzá lehet rendelni egy hurokmentes, többszörös éltől mentes, nem-irányított gráfot, ahol a csúcsok a koordináta-halmaz elemei. A gráf két csúcsát akkor köti össze él, ha a két pont relációban áll N -re nézve. Ezt a gráfot *szomszédsági gráfnak* nevezzük.

Az (X, N) szomszédsági struktúra minden $p \in X$ pontjára definiáljuk az $\mathcal{N}(p) = \{q \mid (p, q) \in N\}$ halmazt, amelyet a p pont (X, N) szomszédsági struktúrára vonatkozó *szomszédságának* nevezzük. A $q \in \mathcal{N}(p)$ pontot a p pont (X, N) szomszédsági struktúrára vonatkozó *szomszédjának* nevezzük.

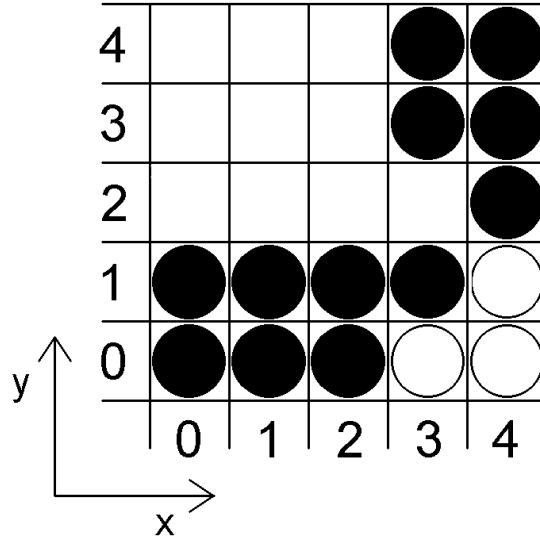
Legyen $p_{(x,y)}$ és $q_{(x',y')}$ ugyanazon kétdimenziós digitális halmaz két pontja. Definiáljuk a következő relációkat:

$$\begin{aligned} (p, q) \in N_4 &\Leftrightarrow |x - x'| + |y - y'| = 1, \\ (p, q) \in N_I &\Leftrightarrow |x - x'| = |y - y'| = 1. \end{aligned}$$

Legyen továbbá $N_8 = N_4 \cup N_I$. Ezen relációk segítségével definiálhatjuk az $\mathcal{N}_4$, az $\mathcal{N}_I$, és az $\mathcal{N}_8$ szomszédságokat, amelyeket *4-szomszédságnak*, *közvetett szomszédságnak* és *8-szomszédságnak* nevezzük. Értelmszerűen használhatjuk a *4-szomszéd*, *I-szomszéd* és a *8-szomszéd* elnevezéseket is.

Az N_4 jelölés mellett az N_D jelölés is használatos. Az $\mathcal{N}_D$ szomszédságot *közvetlen szomszédságnak*, és annak elemeit *D-szomszédoknak* nevezzük.

Az irodalomban általános módszer bináris képek szemléltetésére az alábbi grafikus reprezentáció:



1. ábra. Digitális kép grafikus reprezentációja

Ahol ○ a háttérpontokat, ● az objektumpontokat és ezek együttesen a vizsgált digitális halmaz pontjait jelölik.

Most tekintsünk egy egyszerű példát a fenti fogalmak szemléltetésére. Legyen az X digitális halmaz a következőképpen definiálva: $X = \{(0, 0), (1, 0), (2, 0), (3, 0), (4, 0), (0, 1), (1, 1), (2, 1), (3, 1), (4, 1), (4, 2), (3, 3), (4, 3), (3, 4), (4, 4)\}$. Legyen f egy X feletti bináris kép, amely rendelkezik a következő tulajdonsággal:

$$f(p_{(x,y)}) = \begin{cases} 1, & \text{ha } |x - y| \leq 2, \\ 0, & \text{egyébként.} \end{cases}$$

Ekkor az f bináris kép előtere az $F = \{(0, 0), (1, 0), (2, 0), (0, 1), (1, 1), (2, 1), (3, 1), (4, 2), (3, 3), (4, 3), (3, 4), (4, 4)\}$ digitális halmaz lesz. Nyilvánvalóan a háttér a $B = X \setminus F = \{(3, 0), (4, 0), (4, 1)\}$ összefüggéssel kaphatjuk meg. A $(3, 1)$ koordinátájú pont közvetett szomszédja a $(4, 2)$ koordinátájú pontnak. A $(0, 0)$ koordinátájú pont közvetlen szomszédja a $(0, 1)$ koordinátájú pontnak. Használatos még az irodalomban az $\mathcal{N}_4^*(p) = \mathcal{N}_4(p) \cup \{p\}$ és az

$\mathcal{N}_8^*(p) = \mathcal{N}_8(p) \cup \{p\}$ jelölés is. Például: $\mathcal{N}_4(p_{(3,1)}) = \{(4,1), (2,1), (4,1)\}$ és $\mathcal{N}_8(p_{(3,1)}) = \{(4,1), (4,2), (2,1), (2,0), (3,0), (4,0)\}$.

Legyen X és Y digitális halmaz, továbbá $(X \cup Y, N)$ egy szomszédsági struktúra. Az X és az Y egymás *szomszédjai* ezen szomszédsági struktúrára vonatkozóan, ha létezik $p \in X$ és létezik $q \in Y$ úgy, hogy p és q egymás szomszédjai az $(X \cup Y, N)$ szomszédsági struktúrára vonatkozóan.

Legyen (X, N) egy szomszédsági struktúra, p és q az X digitális halmaz tetszőleges két pontja. A $p = p_0, \dots, p_m = q$ pontsorozatot a p -t és q -t összekötő ($m \in \mathbb{N}$) m -hosszúságú, az adott szomszédsági struktúrára vonatkozó *útnak* nevezzük, ha p_i és p_{i-1} ($0 < i \leq m$) szomszédosak az adott szomszédsági struktúrára vonatkozóan.

A korábbi példánk jelöléseit felhasználva az F halmazt bontsuk fel a következőképpen: $F' = \{(0,0), (1,0), (2,0), (0,1), (1,1), (2,1), (3,1)\}$, $F'' = \{(4,2), (3,3), (4,3), (3,4), (4,4)\}$. Az F' az F'' 8-szomszédja lesz, mert $p_{(3,1)} \in F'$, $q_{(4,2)} \in F''$ és p 8-szomszédja q -nak. Könnyen ellenőrizhető, hogy F' nem 4-szomszédja F'' -nak.

Példaként keressünk egy olyan 8-utat amelyik a $(2,1)$ koordinátájú pontot és a $(4,3)$ koordinátájú pontot összeköti. Az egyik lehetséges megoldás a $(2,1), (3,1), (4,2), (4,3)$ pontsorozat. Azonban esetünkben 4-út nem létezik.

Legyen (X, N) egy szomszédsági struktúra, p és q az X digitális halmaz tetszőleges két pontja. A p és a q *kapcsolódó* az adott szomszédsági struktúrára vonatkozóan, ha létezik abban a p -t és q -t összekötő út.

Az X pontjai között a fenti definíció értelmében egy *kapcsolódási relációt* definiálhatunk a következőképpen. A p és q X -beli pontok kapcsolódóak (jelölés $p \sim q$), ha p és q között létezik út. Könnyen bizonyítható, hogy ez a reláció ekvivalencia-reláció. Ezen ekvivalencia-reláció által indukált osztályokat X -beli *maximális kapcsolódó részhalmazoknak* vagy *komponenseknek* nevezzük.

A korábbi példa alapján a $(2,1)$ és a $(4,3)$ koordinátájú pont 8-kapcsolódó, de nem 4-kapcsolódó. Fontos megjegyezni, hogy ha p és q valamely digitális halmaz pontjai, és p és q 4-kapcsolódó, akkor 8-kapcsolódó is.

Az $Y \subseteq X$ az (X, N) szomszédsági struktúra *kapcsolódó részhalmaza*, ha minden $p, q \in Y$ esetén $p \sim q$ teljesül.

Az X digitális halmazt *kapcsolódónak* nevezzük az (X, N) szomszédsági struktúrára vonatkozóan, ha pontosan egy komponense van. $K_X^N(p)$ -vel jelöljük a p -t tartalmazó X -beli maximális kapcsolódó részhalmazt.

Visszatérve a korábbi példánkhoz és annak jelöléseire a következő megállapításokat tehetjük: $K_F^8(p_{(2,1)}) = K_F^8(p_{(4,3)}) = F$, $K_F^4(p_{(2,1)}) = F'$ és $K_F^4(p_{(4,3)}) = F''$. Ezek alapján az F egyetlen 8-komponensből áll, és így 8-kapcsolódó. Azonban nem 4-kapcsolódó, mert két 4-komponense van, nevezetesen F' és F'' .

Legyen (X, N) egy tetszőleges szomszédsági struktúra és m az X elemszáma. Ekkor azt az $m \times m$ -es A mátrixot, amelynek $a_{i,j}$ eleme 1-gyel, illetve 0-val egyenlő attól függően, hogy $(p_i, p_j) \in N$, illetve $(p_i, p_j) \notin N$, *szomszédsági mátrixnak* vagy *kapcsolódási mátrixnak* nevezzük.

Az így definiált A mátrix l -edik hatványának $a_{i,j}$ eleme 1, ha létezik a p_i és a p_j pontok között l hosszúságú út az (X, N) szomszédsági struktúrában, egyébként 0-val egyenlő.

Jelölje $l(p, q)$ a p és a q között az (X, N) szomszédsági struktúrában található legrövidebb út hosszát. Ha a két pont között nem létezik út, akkor az $l(p, q)$ legyen egyenlő ∞ -nel. Könnyen látható, hogy ez metrika. Az $l(p, q)$ -t a p és q *zsinórtávolságának* nevezzük.

A $d_{\max}((X, N)) = \max_{p,q \in X}(d(p, q))$ -t az (X, N) *szomszédsági struktúra* d metrikára vonatkozó *átmérőjének* nevezzük.

Egy rögzített (X, N) szomszédsági struktúra és $p \in X$ pont esetén az $e(p) = \max_{q \in X}(d(p, q))$ -t a *pont* d metrikára vonatkozó *excentritásának*, az $r((X, N)) = \min_{p \in X}(e(p))$ -t a *szomszédsági struktúra* d metrikára vonatkozó *sugarának*, a $C((X, N)) = \{q \mid e(q) = r((X, N))\}$ -t a *szomszédsági struktúra* d metrikára vonatkozó *centrumának* nevezzük.

Az (X, N) szomszédsági struktúra $Y \subseteq X$ részhalmazának p pontja *belső pont*, ha $\mathcal{N}(p) \subseteq Y$, különben *határpontnak* vagy *kontúrpontnak* nevezzük.

Az (X, N) szomszédsági struktúra $Y \subseteq X$ belső pontjainak $K(Y)$ halmazát az Y *magjának* vagy *belsejének* nevezzük. Az Y határpontjainak $R(Y)$ (vagy ∂Y) halmazát az Y *határának* vagy *kontúrjának* nevezzük.

Legyen (X, N) egy szomszédsági struktúra. Egy $p \in X$ pont *végpont*, ha $|\mathcal{N}(p)| = 1$, illetve *izolált pont*, ha $|\mathcal{N}(p)| = 0$.

Legyen p' és p'' az n -dimenziós digitális sík tetszőleges két pontja, amelyre teljesül, hogy $\text{Pr}_i(p') \leq \text{Pr}_i(p'') (1 \leq i \leq n)$, ahol Pr_i a vizsgált pont i -edik koordinátáját jelöli. A $W = \{p \mid \text{Pr}_i(p') \leq \text{Pr}_i(p) \leq \text{Pr}_i(p''), 1 \leq i \leq n\}$ digitális halmazt $(\text{Pr}_1(p'') - \text{Pr}_1(p') + 1) \times \cdots \times (\text{Pr}_n(p'') - \text{Pr}_n(p') + 1)$ -es *ablaknak* nevezzük. Az ablak *keretén* a következő formában előálló L digitális halmazt értjük:

$$L = \{p \mid \text{Pr}_i(p') = \text{Pr}_i(p) \text{ vagy } \text{Pr}_i(p'') = \text{Pr}_i(p) (1 \leq i \leq n)\}.$$

Egy X digitális halmaz által *kifeszített ablak* alatt az X -et tartalmazó összes ablak metszetét értjük.

Legyen f az X digitális halmaz felett értelmezett bináris kép. Jelölje B az f háttérét és F az előterét. *Lyukaknak* nevezzük a B azon komponenseit, amelyekhez nem lehet konstruálni olyan utat, amely a komponens egy tetszőleges elemét az X által kifeszített W ablak keretének valamely pontjával köti össze, úgy, hogy az utat alkotó pontok mindegyike a $W \setminus F$ halmaz eleme.

Most nézzünk egy-két szemléltető példát az új fogalmakra a korábbi példánk segítségével. Sem az F' , sem az F'' nem ablak. Ha például az F' halmazból a $(3, 1)$ koordinátájú pontot elhagyjuk, akkor egy ablakot kapunk. Az F'' halmaz által kifeszített ablak az $F'' \cup \{(3, 2)\}$ halmaz lesz. A példában szereplő f bináris kép nem tartalmaz lyukat (függetlenül attól, hogy 4- vagy 8-kapcsolódást választunk-e).

Legyen X és Y n -dimenziós digitális halmaz. Egy g -értékű általánosított Y -ból X -be ható *bináris t sablon* alatt egy

$$t = (g(x, y), h(x, y)) \quad x \in X, y \in Y$$

alakú függvényt értünk, ahol g minden $x \in X$ -re egy Y feletti bináris kép, amíg h minden $x \in X$ -re $Y \rightarrow \mathbb{Z}^n$ alakú injektív függvény. A h függvényt *illesztő*, a g függvényt *formázó függvénynek* nevezzük.

A korábbi jelöléseket használva a t sablont *eltolásinvariánsnak* nevezzük, ha

$$g(x, y) = g(x', y) \quad x, x' \in X\text{-re.}$$

A korábbi jelöléseket használva a t sablont *állandó alakúnak* nevezzük, ha létezik $c \in \mathbb{Z}^n$ úgy, hogy

$$h(x, y) = h(x', y) \quad x, x' \in X\text{-re.}$$

A korábbi jelöléseket használva a t sablont *alaktartónak* nevezzük, ha

$$y - y' = h(x, y) - h(x, y') \quad y, y' \in Y\text{-ra.}$$

A korábbi jelöléseket használva legyen f az X digitális halmaz feletti bináris kép. Minden t sablon esetén lehetőség van egy $p \in Y$ elem kitüntetésére, amit *centrális elemnek* nevezzük. Azt mondjuk, hogy a t sablon *pontosan*

illeszkedik a $p'' \in X$ pontban az f képre, ha minden $p' \in Y$ esetén a következő feltételek teljesülnek:

- $h(p'', p) = p''$,
- $g(p'', p') = f(h(p'', p'))$.

A korábbi jelöléseket használva azt mondjuk, hogy a t sablon *illeszkedik* a $p'' \in X$ pontban az f -re, ha t pontosan illeszkedik a p'' pontban az f' -re, ahol f' az n -dimenziós digitális sík felett értelmezett, az alábbi módon definiált bináris kép:

$$f'(x) = \begin{cases} f(x), & \text{ha } x \in X, \\ 0, & \text{ha } x \notin X. \end{cases}$$

A korábbi jelöléseket használva azt mondjuk, hogy a t sablon *túlnyúlással illeszkedik* a $p'' \in X$ pontban az f -re, ha t' illeszkedik a p'' pontban az f -re, ahol $t' = (g', h)$ alakú, ahol

$$g'(x, y) = \begin{cases} g(x, y), & \text{ha } h(x, y) \in X, \\ 0, & \text{egyébként.} \end{cases}$$

A fenti fogalmak szemléltetésére tekintsük a következő példát. Az X digitális halmaz és az f bináris kép legyen a korábbi példánkban definiált. Az $Y = \{(1, 1), (2, 1), (2, 3)\}$ és a t_1 -et meghatározó $h_1(x, y)$ -t az alábbi táblázatos megadási móddal definiáljuk:

(x, y)	$(1, 1)$	$(2, 1)$	$(2, 3)$
$(3, 1)$	$(3, 0)$	$(3, 1)$	$(4, 1)$
$(4, 1)$	$(3, 1)$	$(4, 1)$	$(4, 2)$
egyébként	$(1, 1)$	$(4, 1)$	$(4, 4)$

A $g_1(x, y)$ pedig legyen a következő!

(x, y)	$(1, 1)$	$(2, 1)$	$(2, 3)$
$(3, 1)$	0	1	0
$(4, 1)$	1	0	1
egyébként	1	0	1

Nyilvánvalóan a t_1 sablon nem eltolásinvariáns, nem állandó alakú és nem alaktartó. Legyen a t_1 centrális eleme $(2, 1)$. Ebben az esetben a t_1 sablon pontosan illeszkedik az f képre a $(3, 1)$ és a $(4, 1)$ pontokban.

Most definiáljuk a t_2 sablont a következőképpen: legyen Y a fent definiált és a centrális elem is legyen a korábban rögzített $(2, 1)$ koordinátájú pont. A $h_2(x, y) = (\text{Pr}_1(x), \text{Pr}_2(x) + \max(\text{Pr}_1(y), \text{Pr}_2(y)) - 2)$ és

$$g_2(x, y) = \begin{cases} 1, & \text{ha } y \neq (1, 1), \\ 0, & \text{egyébként.} \end{cases}$$

A t_2 sablon az f képre pontosan illeszkedik a $(4, 2)$ pontban, illeszkedik a $(4, 2), (3, 3), (0, 0), (0, 1), (0, 2)$ pontokban és túlnyúlással illeszkedik a $(4, 2), (3, 3), (0, 0), (0, 1), (0, 2), (3, 1)$ pontokban. A t_2 sablon eltolásinvariáns, állandó alakú, de nem alaktartó.

Egy f bináris képen a t sablonnal végrehajtott i -menetes sablonozás eredményét $\Psi(f, t, i)$ -vel jelöljük, amit a következőképpen definiálunk:

$$\begin{aligned} \Psi(f(p), t, 1) &= \begin{cases} f(p), & \text{ha a } t \text{ sablon illeszkedik a } p \text{ pontra,} \\ 1 - f(p), & \text{ha a } t \text{ sablon nem illeszkedik a } p \text{ pontra.} \end{cases} \\ \Psi(f(p), t, i) &= \Psi(\Psi(f(p), t, i-1), t, 1) \quad i \geq 2. \end{aligned}$$

Egy f bináris képen a t sablonnal végrehajtott sablonozás eredményét $\Psi(f, t)$ -vel jelöljük, ami $\Psi(f, t, i)$ -vel egyenlő, ahol i az a legkisebb index, amelyre teljesül, hogy $\Psi(f, t, i) = \Psi(f, t, i+1)$. Ha nem létezik ilyen i index, akkor $\Psi(f, t)$ definiálatlan.

Egy f bináris képen a $\{t_j \mid 1 \leq j \leq l\}$ sablonsorozattal végrehajtott 1-menetes sablonozás eredményét a következőképpen definiáljuk:

$$\Psi(f, \{t_j\}, 1) = \Psi(\Psi(\Psi(f, t_1, 1), t_2, 1), \dots, t_l, 1).$$

A sablonsorozattal végrehajtott i -menetes sablonozás, majd a sablonsorozattal végrehajtott sablonozás eredménye az előző definíció alapján könnyen definiálható.

Ha a $\Psi_1(f, \{t_{j_1}\}, i), \Psi_2(f, \{t_{j_2}\}, i), \dots, \Psi_m(f, \{t_{j_m}\}, i)$ sablonozások, akkor az 1-menetes $\Psi(f, \{t_j\}, 1) = \Psi_m(\dots \Psi_2(\Psi_1(f, \{t_{j_1}\}, 1), \{t_{j_2}\}, 1) \dots, \{t_{j_m}\}, 1)$ sablonozás segítségével definiálhatjuk a $\Psi(f, \{t_j\})$ sablonozást, amit m almenetes sablonozásnak nevezünk, amelynek l -edik ($1 \leq l \leq m$) almenetén a $\Psi_l(f, \{t_{j_l}\})$ sablonozást értjük. A menet, almenet helyett használhatjuk a *ciklus*, *alciklus* elnevezéseket is.

Legyen t egy g -értékű általánosított Y -ból X -be ható bináris sablon $y' \in Y$ centrális elemmel. Ha $g(x, y') = 1$ ($\forall x \in X$), akkor a t -t *törlő sablonnak*, ha $g(x, y') = 0$ ($\forall x \in X$), akkor a t -t *helyreállító sablonnak* nevezzük.

Egy g -értékű általánosított Y -ból X -be ható bináris t sablont g -értékű Y -ból X -be ható bináris sablonnak nevezzük, ha t eltolásinvariáns és alaktartó.

Egy $m_1 \times m_2 \times \cdots \times m_l$ -es W ablakból X -be ható bináris sablont $m_1 \times m_2 \times \cdots \times m_l$ -es sablonnak nevezzük.

1.3. Kereszteződési szám

Ebben az alfejezetben bináris kép alatt egy X digitális halmaz felett értelmezett f bináris képet értünk.

Legyen f az X digitális halmaz felett értelmezett bináris kép. Legyen $p_{(x,y)}$ az X egy tetszőleges pontja. Ekkor a $p_0 = p_{(x+1,y)}$, $p_1 = p_{(x+1,y+1)}$, $p_2 = p_{(x,y+1)}$, $p_3 = p_{(x-1,y+1)}$, $p_4 = p_{(x-1,y)}$, $p_5 = p_{(x-1,y-1)}$, $p_6 = p_{(x,y-1)}$ és $p_7 = p_{(x+1,y-1)}$ jelölésekkel definiáljuk a $\gamma_p(i) = f(p_i)$ függvényt. A $\gamma_p(i)$ helyett, ha nem okoz félreértést, a $\gamma(i)$ jelölést fogjuk használni.

A kereszteződési szám első ajánlója D. Rutovitz, amely megadja, hogy hányszor haladunk át objektumpontról háttérpontra és fordítva az $\mathcal{N}_8(p)$ -ben, ha a bejárési irány az óramutató járásával ellentétes.

A χ_R (Rutovitz-féle) kereszteződési számot a következőképpen definiáljuk:

$$\chi_R(p) = \sum_{k=0}^7 |\gamma_p(k+1 \bmod 8) - \gamma_p(k)|.$$

A $\chi_R(p)$ egyenlő az $\mathcal{N}_8(p)$ -ben található komponensek számának kétszeresével. A p pont törlése nem lesz hatással a kapcsolódási viszonyokra, ha $\chi_R(p) = 2$.

A $\chi_H(p)$ (Hilditch-féle) kereszteződési számot a következőképpen definiáljuk:

$$\chi_H(p) = \sum_{i=0}^3 b_i, \text{ ahol}$$

$$b_i = \begin{cases} 1, & \text{ha } \gamma(2i) = 0 \text{ és } (\gamma(2i+1) = 1 \text{ vagy } \gamma(2i+2 \bmod 8) = 1), \\ 0, & \text{egyébként.} \end{cases}$$

A $\chi_H(p)$ egyenlő a komponensek számával a $\mathcal{N}_8(p)$ -ben. Kivéve azt az esetet, amikor p mindegyik 4-szomszédja objektumpont, ebben az esetben $\chi_H(p) = 0$. Nyilvánvalóan ha $\chi_H(p) = 1$, akkor p törlése nem változtatja meg a kapcsolódási viszonyokat.

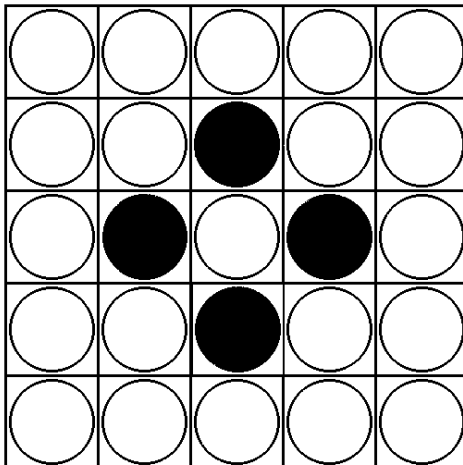
A $\chi_H(p)$ értékével karakterizálja a p pontot. Ha $\chi_H(p) = 0$, akkor belső vagy izolált pont, ha $\chi_H(p) = 1$, akkor határpont, ha $\chi_H(p) = 2$, akkor

összekötő pont, ha $\chi_H(p) = 3$, akkor *elágazó pont*, és ha $\chi_H(p) = 4$ akkor *keresztvező pontról* van szó.

Fontos különbség a $\chi_H(p)$ és $\chi_R(p)$ között a következő: a $\chi_H(p) = 1$ feltétel azt is implikálja, hogy a p pont határpont, $\chi_R(p) = 2$ nem biztosítja ugyanezt.

2. Jól-kapcsolódó halmazok

A digitális topológia kiinduló pontja az az ötlet, hogy az előtérre és a háttérre különböző kapcsolódási relációt használunk. Ezzel kerülve el a *kapcsolódási paradoxont*, amit a következő ábra mutat:



2. ábra. Kapcsolódási paradoxon

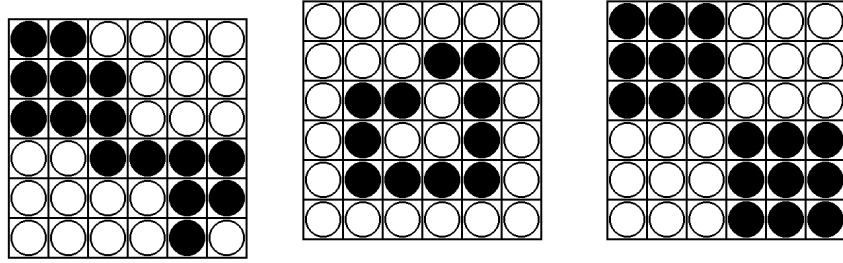
Ha 4-kapcsolódást tekintünk, akkor a nem kapcsolódó fekete pontok szétvágják a fehér képpontokat, 8-kapcsolódást feltételezve a 8-kapcsolódó fekete pontok nem vágják szét a fehér képpontokat.

A 2D-s négyzetrács esetén az ún. homogén szomszédsági struktúrák közül a 8-szomszédság és a 4-szomszédság által meghatározott kapcsolódási viszonyok lenti elvek alapján történő használatával elkerülhetjük a kapcsolódási paradoxonokat és biztosíthatjuk az Euler-karakterisztika lokális számíthatóságát. Ha az előtérre 8-, akkor a háttérre 4-kapcsolódást kell használnunk

(jelölése $(8, 4)$), illetve ha az előtérre 4-, akkor a háttérre 8-kapcsolódást (jelölése $(4, 8)$) kell használnunk. Felmerül az a kérdés, hogy vajon a lehetséges halmazok megszorításával elkerülhetjük-e a kapcsolódási paradoxonokat annak ellenére, hogy a bináris képre $(4, 4)$ - vagy $(8, 8)$ -kapcsolódást írunk elő.

Egy X digitális halmaz *majdnem jól-kapcsolódó*, ha bármely 8-komponense egyben 4-komponens is.

Egy X digitális halmaz *jól-kapcsolódó*, ha az X és annak komplementere is majdnem jól-kapcsolódó.



3. ábra. Jól-kapcsolódó, majdnem jól-kapcsolódó és egyik tulajdonsággal sem rendelkező digitális halmaz.

Egy X digitális halmaz *lokálisan 4-kapcsolódó*, ha $X \cap (\mathcal{N}_8(p))$

Tétel. 4-kapcsolódó minden $p \in X$ -re. Egy X digitális halmaz pontosan akkor jól-kapcsolódó, ha lokálisan 4-kapcsolódó.

T.Y. Kong és A. Rosenfeld megmutatta azt, hogy ha egy bináris kép esetén $(4, 4)$ - vagy $(8, 8)$ -kapcsolódást használunk és a bináris kép előtere és háttére majdnem jól-kapcsolódó, azaz a bináris kép jól-kapcsolódó, akkor az Euler-karakterisztika lokálisan számolható.

III. fejezet

Mintavételezés

Ebben a fejezetben azzal foglalkozunk, hogy egy folytonos függvény mintavételezése során kapott diszkrét függvényből milyen feltételek mellett állíthatjuk helyre a folytonos függvényt. A fejezet megértéséhez szükséges, a Fourier-transzformációval kapcsolatos ismeretek megtalálhatók a következő fejezetben.

1. Egydimenziós eset

Az elemzést először az egydimenziós függvények esetén végezzük el. Tekintsünk egy $f(x)$ valós függvényt, amely a valós számok halmazán értelmezett. Tegyük fel, hogy az $f(x)$ Fourier-transzformáltja nullához tart, ha az u értéke a $[-W, W]$ intervallumon kívül esik. Ebben az esetben az $f(x)$ függvényt *sávkorlátos függvénynek* nevezzük.

A mintavételezéshez meg kell ismernünk az *impulzusfüggvény* fogalmát. A $\delta(x - x_0)$ impulzusfüggvény az alábbi módon adható meg:

$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0).$$

Ez egy olyan függvényként interpretálható, amelynél a görbe alatti terület egységnyi, és a függvény értéke az x_0 pont végtelen kicsi sugarú környezetén kívül mindenütt nulla. Ez matematikailag a következőképpen adható meg:

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = \int_{x_0^-}^{x_0^+} \delta(x - x_0) dx = 1.$$

Az $f(x)$ mintavételezett változatának előállításához $f(x)$ -et egyszerűen meg kell szorozni egy olyan $s(x)$ függvénnyel, amely egymástól Δx távolságra lévő impulzusokból áll. A konvolúciós tételből tudjuk, hogy a szorzásnak a frekvencia térben a konvolúció felel meg. Itt jegyezzük meg, hogy a transzformált függvény periodikus és a periódushossz $1/\Delta x$, továbbá az $F(u)$ egyes ismétlődései átfedhetik egymást. Az átfedések elkerülése végett Δx értékét úgy válasszuk meg, hogy a $\Delta x \leq \frac{1}{2W}$ egyenlőtlenség teljesüljön. A Δx periódushossz csökkentésével az $F(u)$ periódusait elkülöníthetjük egymástól, megszüntetve ezzel az átfedéseket. Ezen művelet fontosságát a következő lépés magyarázza meg.

Legyen a $G(u)$ függvény az alábbi módon definiált:

$$G(u) = \begin{cases} 1, & -W \leq u \leq W, \\ 0, & \text{egyébként.} \end{cases}$$

Szorozzuk meg ezzel $S(u) * F(u)$ -t. Ez lehetővé teszi számunkra az $F(u)$ teljes elkülönítését. Ezután az inverz Fourier-transzformáció az eredeti folytonos függvényt adja vissza, mivel

$$G(u) \cdot [F(u) * S(u)] = F(u) .$$

Tétel. (Whittaker-Shannon) Egy sávkorlátos függvény teljesen visszaállítható, ha a mintavételezési pontok távolságára teljesül, hogy $\Delta x \leq \frac{1}{2W}$.

Tartsuk szem előtt, hogy egy sávkorlátos függvény összes információja a frekvenciater $[-W, W]$ intervallumba esik. Ha a fenti egyenlőtlenség nem teljesül, akkor a transzformáció értékeit lerontják a szomszédos periódusok. A jelenség – amit gyakran *aliasing*nak hívnak – eleve kizárja egy alul-mintavételezett függvény teljes visszaállítását.

Az előzőekben tárgyalt függvények a valós számok teljes halmazán voltak értelmezve. Tegyük fel, hogy a vizsgált függvény mintavételezése során teljesül a Whittaker-Shannon tételben szereplő feltétel, így az aliasing most nem jelentkezik. A $[0, X]$ véges intervallumot matematikailag egy $h(x)$ függvénnyel való szorzással kapjuk meg.

$$h(x) = \begin{cases} 1, & 0 \leq x \leq X, \\ 0, & \text{egyébként.} \end{cases}$$

Ezt a függvényt szokták *ablakfüggvénynek* nevezni. Ennek Fourier-transzformáltját jelölje $H(u)$, ami az előző ábrán látható. A frekvencia térben a végeredményt a $H(u) * [S(u) * F(u)]$ konvolúció adja. Mivel $H(u)$ -nak a végtelenben is vannak frekvencia komponensei, ezen függvények konvolúciója torzulást eredményez a mintavételezett és $h(x)$ által egy véges tartományra szűkített $f(x)$ függvény frekvenciatérbeli reprezentációjában. Emiatt általános esetben azt mondhatjuk, hogy az x egy véges tartományán vett mintából az $f(x)$ függvény még akkor sem állítható vissza hibamentesen, ha a mintavételi pontok távolságára teljesül a Whittaker-Shannon-tételben szereplő egyenlőtlenség.

A fenti eset alól egyetlen kivétel van: Abban az esetben, ha az $f(x)$ sávkorlátos, periodikus és a periódushossz X -szel megegyezik, akkor a $H(u)$ -ból származó torzulások nem jelentkeznek, és ez lehetővé teszi $f(x)$ teljes visszaállítását. Megjegyzendő, hogy a visszaállított függvény a $[-\infty, +\infty]$ intervallumon értelmezett és azon kívül sem nulla.

Mint korábban már rámutattunk, a mintavételezést el lehet képzelni a vizsgált függvény és egy impulzussorozat-függvény szorzataként. Az $F(u)$ és $S(u)$ pontonkénti szorzásának a megfelelő függvények konvolúciója felel meg. A kapott függvény periódushossza $1/\Delta u$. Ha az $f(x)$ és az $F(u)$ függvényekből egyaránt N -elemű mintát veszünk, és a lépésközt úgy választjuk, hogy az N minta egy periódust fedjen le, akkor $N\Delta x = X$ és $N\Delta u = 1/\Delta x$. Ez utóbbi abból következik, hogy egy diszkrét függvény Fourier-transzformáltja periodikus és a periódushossz $1/\Delta x$. A lépésköz ilyen módon való választása azt eredményezi, hogy a képtérben a periódushossz $1/\Delta u$. Ebből következik, hogy a teljes intervallum, ahol a mintavételezést kell végezni: $1/\Delta u = N\Delta x = X$.

2. Kétdimenziós eset

Az előzőekben bevezetett fogalmak a jelölések módosítása után közvetlenül alkalmazhatók kétdimenziós függvényekre. A mintavételezés matematikailag

kétdimenziós impulzusfüggvény felhasználásával formularizálható. A $\delta(x, y)$ kétdimenziós impulzusfüggvény a következőképpen definiálható:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x - x_0, y - y_0) dx dy = f(x_0, y_0).$$

A kétdimenziós $s(x, y)$ mintavételezési függvény is impulzussorozat, ahol az impulzusok közötti intervallum Δx és Δy . Ennek Fourier-transzformáltja az $S(u, v)$, ami szintén impulzussorozat, és a periódushossz $1/\Delta x$ és $1/\Delta y$. Ha $f(x, y)$ sávkorlátos, akkor a Fourier-transzformáltja egy R véges tartományon kívül nullához tart. Legyen

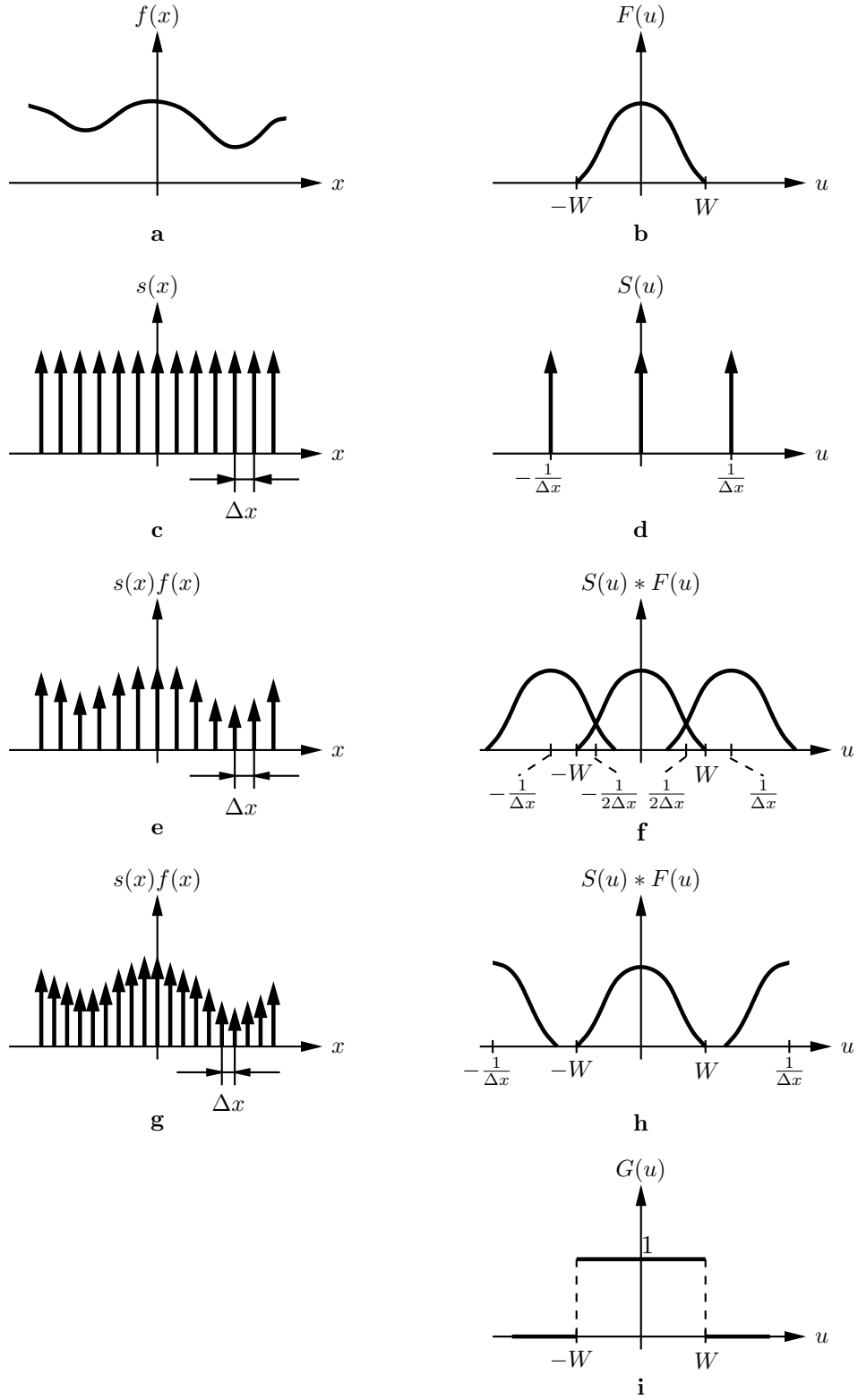
$$G(u, v) = \begin{cases} 1, & (u, v) \text{ az } R\text{-et magába foglaló téglalapon belül van,} \\ 0, & \text{egyébként.} \end{cases}$$

A $G(u, v) \cdot [S(u, v) * F(u, v)]$ inverz Fourier-transzformáltja az eredeti $f(x, y)$ függvényt adja vissza. Jelölje $2W_u$ és $2W_v$ annak a legkisebb téglalapnak az oldalait, amelyik magában foglalja az R tartományt. Ilyen jelölések mellett megfogalmazhatjuk a kétdimenziós mintavételezési törvényt:

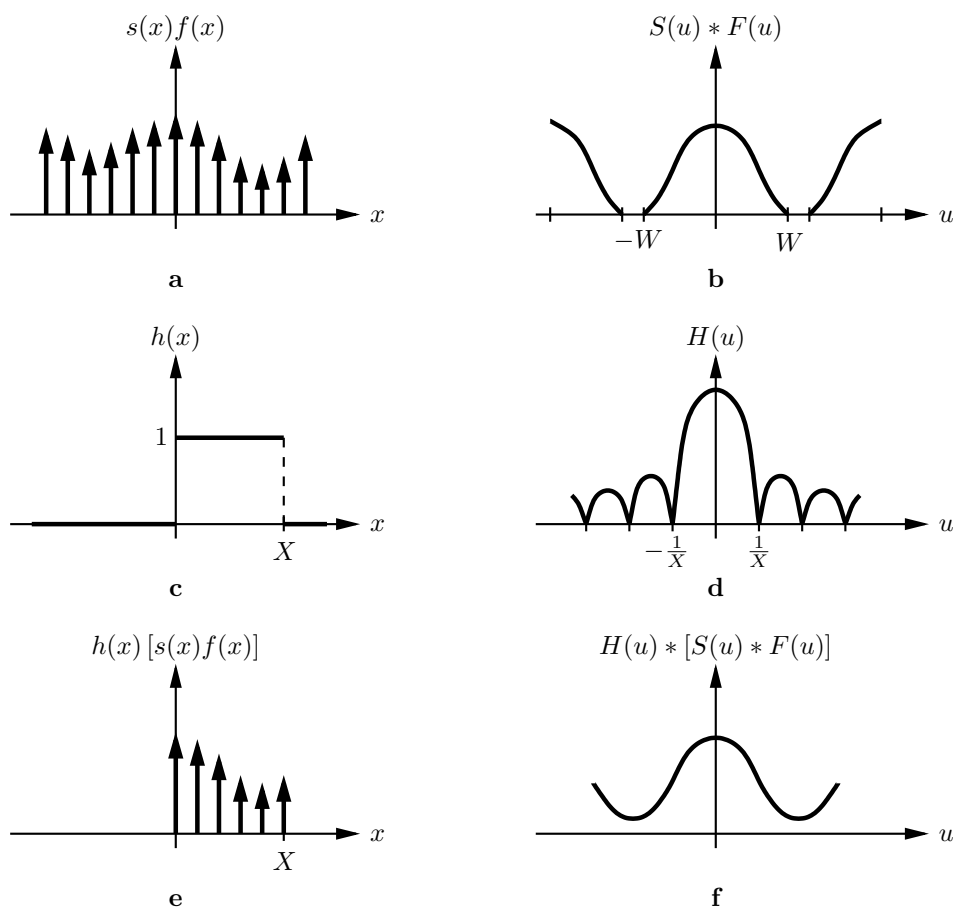
Tétel. *(Kétdimenziós mintavételezési törvény.) Egy sávkorlátos $f(x, y)$ függvény teljesen visszaállítható, ha mintavételezési periódusokra érvényes a $\Delta x \leq \frac{1}{2W_u}$, illetve a $\Delta x \leq \frac{1}{2W_u}$ egyenlőtlenség.*

Ha az $f(x, y)$ függvény egy $h(x, y)$ kétdimenziós ablakkal térben korlátozott, az egydimenziós esettel analóg módon most is jelentkezik a torzulás a $H(u, v) * [S(u, v) * F(u, v)]$ konvolúcióban. Az egydimenziós esethez hasonlóan azt kapjuk, hogy a teljes helyreállítás feltétele a periódusokra vonatkozó alábbi feltételek teljesülése:

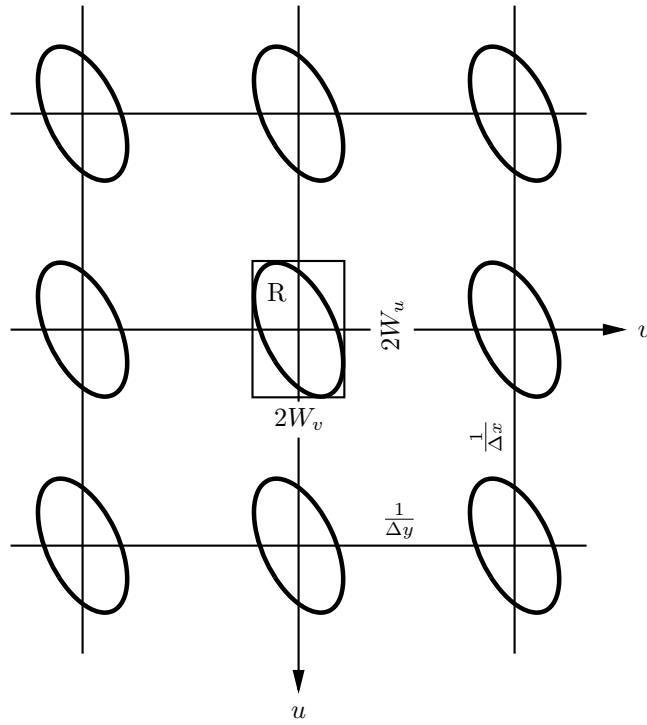
$$\Delta u = \frac{1}{N\Delta x} \quad \text{és} \quad \Delta v = \frac{1}{N\Delta y}.$$



1. ábra. Mintavételezés.



2. ábra. Whittaker-Shannon-tétel illusztrálása.



3. ábra. Kétdimenziós mintavételezés.

IV. fejezet

Képtranszformációk

A digitális képfeldolgozásban számos helyen alkalmazzák az egy-, illetve a kétdimenziós integráltranszformációkat. Ezek közül is a legfontosabb a Fourier-transzformáció, amely például a képek szűrésére használt eljárások elméleti alapjául szolgál.

1. Fourier-transzformáció

A Fourier-transzformációt elsősorban képjavításra használják. Ezt a fajta alkalmazhatóságot azon tulajdonsága teszi lehetővé, hogy a transzformáció átlagértéke arányos a kép intenzitásával, továbbá, hogy a nagyfrekvenciás komponense a képen található élek nagyságáról és helyzetéről ad információt. Ennélfogva az alacsony-, illetve a magasfrekvenciás részek vágásával a képet élesebbé, illetve simábbá tehetjük.

1.1. Folytonos eset

Ezek után térjünk rá magára a transzformációra. Legyen $f(x)$ a valós számok halmazán értelmezett folytonos függvény. Az $f(x)$ függvény *Fourier-transzformáltját* a következőképpen definiáljuk:

$$F(u) = \int_{-\infty}^{+\infty} f(x)e^{-2\pi iux} dx.$$

Adott $F(u)$ függvény esetén az $f(x)$ függvényt az *inverz Fourier-transzformáció* alkalmazásával határozhatjuk meg:

$$f(x) = \int_{-\infty}^{+\infty} F(u)e^{2\pi iux} du.$$

A fenti egyenleteket Fourier-transzformációs párnak nevezzük, létezésük folytonos és integrálható $f(x)$, valamint integrálható $F(u)$ esetén bizonyítható.

Egy valós $f(x)$ függvény $F(u)$ Fourier-transzformáltja (ha létezik) általában komplex értékű függvény, így felírható a következő alakban:

$$F(u) = R(u) + iI(u),$$

ahol $R(u)$, illetve $I(u)$ az $F(u)$ valós, illetve komplex komponense. Ezt gyakran a következő alakban fejezzük ki:

$$F(u) = |F(u)|e^{i\Psi(u)},$$

ahol $|F(u)| = |R^2(u) + I^2(u)|^{\frac{1}{2}}$ és $\Psi(u) = \tan^{-1}\left(\frac{I(u)}{R(u)}\right)$. Az $|F(u)|$ az $f(x)$ függvény *Fourier-spektruma*, $\Psi(u)$ pedig a *fázisszög-függvénye*. A $P(u) = |F(u)|^2$ az $f(x)$ *spektrum erősségének* függvénye.

A Fourier-transzformáció az $f(x, y)$ kétváltozós függvényekre is értelmezhető. A fentiekhez hasonlóan a kétdimenziós Fourier-transzformációs pár létezése bizonyítható, ha $f(x, y)$ folytonos és integrálható, valamint $F(u, v)$ integrálható:

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)e^{-2\pi i(ux+vy)} dx dy,$$

és

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v)e^{2\pi i(ux+vy)} du dv.$$

1.2. Diszkrét eset

Tegyük fel, hogy az $f'(x)$ folytonos függvényt mintavételezéssel diszkrétizáljuk. Az $f(x)$ diszkrét függvényt jelöljük a következőképpen:

$$f(x) = f'(x_0 + x\Delta x) \quad x = 0, 1, \dots, N-1,$$

ahol Δx a mintavételezési pontok távolsága. Így az $f(x)$ függvény *diszkrét Fourier-transzformáltja*

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)e^{\frac{-2\pi i u x}{N}}, \quad u = 0, 1, \dots, N-1$$

valamint az $F(u)$ függvény *inverz diszkrét Fourier-transzformáltja*:

$$f(x) = \sum_{u=0}^{N-1} F(u)e^{\frac{2\pi i u x}{N}} \quad x = 0, 1, \dots, N-1.$$

A folytonos kétdimenziós Fourier-transzformációhoz hasonlóan létezik a diszkrét kétdimenziós Fourier-transzformációs pár:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2\pi i(\frac{ux}{M} + \frac{vy}{N})},$$

valamint

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{2\pi i(\frac{ux}{M} + \frac{vy}{N})}.$$

A folytonos esettel ellentétben a diszkrét Fourier-transzformációs pár létezését szükségtelen bizonyítani, mivel $F(u)$, illetve az $F(u, v)$ függvény mindig létezik diszkrét esetben.

1.3. A Fourier-transzformáció néhány tulajdonsága

A következőkben a kétdimenziós Fourier-transzformáció néhány fontos tulajdonságát ismertetjük.

1.3.1. *Linearitás.* A Fourier-transzformáció definíciójából közvetlenül adódik, ha az $f_1(x, y), \dots, f_n(x, y)$ ($n \in \mathbb{N}$) függvényeknek létezik Fourier-transzformáltja, és azokat $F_1(u, v), \dots, F_n(u, v)$ -vel jelöljük, akkor a $c_1 \cdot f_1(x, y) + \dots + c_n \cdot f_n(x, y)$ függvénynek is létezik Fourier-transzformáltja és az $c_1 \cdot F_1(u, v) + \dots + c_n \cdot F_n(u, v)$ -vel egyezik meg.

1.3.2. *Szeperálhatóság.* A Fourier-transzformációs pár szeperábilis:

$$F(u, v) = \int_{-\infty}^{+\infty} e^{-2\pi i v y} \int_{-\infty}^{+\infty} f(x, y) e^{-2\pi i u x} dx dy,$$

és

$$f(x, y) = \int_{-\infty}^{+\infty} e^{2\pi i v y} \int_{-\infty}^{+\infty} F(u, v) e^{2\pi i u x} du dv.$$

A szeperábilítás legfontosabb előnye az, hogy az $F(u, v)$, illetve az $f(x, y)$ függvényt az egydimenziós Fourier-transzformáció, illetve inverzének kétszeri alkalmazásával kaphatjuk meg. Ez nyilvánvaló:

$$F(u, v) = \int_{-\infty}^{\infty} F(u, y) e^{-2\pi i v y} dy,$$

ahol

$$F(u, y) = \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i u x} dx.$$

Hasonlóképpen bizonyítható az inverz Fourier-transzformáció szeperábilis tulajdonsága is. Nyilván diszkrét esetben is teljesül a szeperábilítás.

1.3.3. *Eltolás.* A folytonos Fourier-transzformációs pár következő eltolási tulajdonságát egyszerű számolással ellenőrizhetjük.

$$f(x, y)e^{2\pi i(u_0x + v_0y)} \Leftrightarrow F(u - u_0, v - v_0)$$

és

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)e^{-2\pi i(ux_0 + vy_0)},$$

ahol a $\Leftrightarrow$ jel a függvény és annak Fourier-transzformáltja közötti megfeleltetést jelöli.

Szintén egyszerűen adódik diszkrét esetben a következő összefüggés:

$$f(x, y)e^{\frac{2\pi i(u_0x + v_0y)}{N}} \Leftrightarrow F(u - u_0, v - v_0)$$

és

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)e^{\frac{-2\pi i(ux_0 + vy_0)}{N}}.$$

Tekintsük most azt az esetet, amikor $u_0 = v_0 = \frac{N}{2}$. Ezt behelyettesítve az előző megfeleltetésbe a következőt kapjuk:

$$e^{\frac{2\pi i(u_0x + v_0y)}{N}} = e^{\pi i(x+y)} = (-1)^{x+y}$$

tehát

$$f(x, y)(-1)^{x+y} \Leftrightarrow F(u - \frac{N}{2}, v - \frac{N}{2}).$$

Ebből következik, hogy az $f(x, y)$ függvény Fourier-transzformáltjának origója az $N \times N$ -es frekvencia mező közepére tolódik, ha az $f(x, y)$ függvényt a $(-1)^{x+y}$ kifejezéssel megszorozzuk.

1.3.4. *Forgatás.* Vezessünk be polárkoordinátákat a következő formában:

$$x = r \cdot \cos \alpha, y = r \cdot \sin \alpha, u = \omega \cdot \cos \beta, v = \omega \cdot \sin \beta.$$

Ebben az esetben az $f(r, \alpha)$ függvény Fourier-transzformáltja az $F(\omega, \beta)$ lesz.

Könnyű megmutatni, hogy teljesül a következő kapcsolat az $f(r, \alpha)$ függvény és annak Fourier-transzformáltja között:

$$f(r, \alpha + \alpha_0) \Leftrightarrow F(\omega, \beta + \alpha_0).$$

1.3.5. *Periodicitás.* A diszkrét Fourier-transzformáció és inverze periodikus, tehát:

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N),$$

ahol N a periódus.

1.3.6. *Átlagérték.* Egy kétdimenziós függvény átlagértékének meghatározására általánosan használt az alábbi kifejezés:

$$\bar{f}(x, y) = \frac{1}{N} \frac{1}{M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y)$$

Az $u = v = 0$ értéket a (6) egyenletbe helyettesítve a

$$F(0, 0) = \frac{1}{N} \frac{1}{M} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y)$$

kifejezést kapjuk. Ebből látható, hogy az $f(x, y)$ függvény átlagértéke és Fourier-transzformáltja között érvényes az alábbi összefüggés:

$$\bar{f}(x, y) = F(0, 0).$$

1.4. Szimmetria

A Fourier-transzformáció ferdén szimmetrikus, azaz, ha az $f(x, y)$ függvény valós, akkor a Fourier-transzformáltja konjugált szimmetriát mutat:

$$F(u, v) = F^*(-u, -v).$$

1.5. Laplace-operátor

A Laplace-operátor definíciója:

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

Az $f(x, y)$ függvény Laplace-transzformáltjának Fourier-transzformáltja:

$$\mathcal{F}(\nabla^2 f(x, y)) = -(2\pi)^2(u^2 + v^2)F(u, v).$$

1.6. Megjelenítés

Egy kép Fourier transzformáltjának megjelenítése bizonyos problémákat vet fel. Általában a kép spektruma a nagyobb frekvencia értékeknél csökkenő tendenciát mutat, így ez a megjelenítés során sötét marad. Ezt kompenzálандó az $|F(u, v)|$ helyett érdemes a $\log(1 + |F(u, v)|)$ függvényt megjeleníteni.

1.7. Konvolúció

Ebben a fejezetben a függvények értelmezési tartománya (*paramétertér*) és a Fourier-transzformáltjuk értelmezési tartománya (*frekvenciatér*) közötti alapvető kapcsolatot megteremtő konvolúcióval foglalkozunk.

Az $f(x)$ és a $g(x)$ függvények $f(x) * g(x)$ -szel jelölt *konvolúcióját* a következőképpen definiáljuk:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha)g(x - \alpha)d\alpha.$$

Ha az $f(x)$ Fourier-transzformáltja $F(u)$ és a $g(x)$ Fourier-transzformáltja $G(u)$, akkor az $f(x) * g(x)$ Fourier-transzformáltja $F(u) \cdot G(u)$. Ezt az eredményt formálisan a következő formában fejezhetjük ki:

$$f(x) * g(x) \Leftrightarrow F(u) \cdot G(u).$$

Ez az összefüggés azt fejezi ki, hogy a paramétertérbeli konvolúciónak frekvenciatérbeli szorzás felel meg. Továbbá az is teljesül, hogy a paramétertérbeli szorzásnak frekvenciatérbeli konvolúció felel meg. Ezt formálisan a következőképpen fejezhetjük ki:

$$f(x) \cdot g(x) \Leftrightarrow F(u) * G(u).$$

A fent megfogalmazott tulajdonságokra *konvolúciós tételként* hivatkozunk.

1.8. Gyors Fourier-transzformáció

A Fourier-transzformált,

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)e^{-2\pi i \frac{ux}{N}}$$

kiszámításához szükséges komplex szorzások és összeadások száma N^2 -tel arányos. Ez könnyen belátható, hiszen az összegzés kifejtéséhez, az u minden N értékéhez N -szer kell komplex szorzást végrehajtani $f(x)$ és $e^{-2\pi i \frac{ux}{N}}$ között, valamint szükséges $(N - 1)$ darab összeadás is az eredmény kiszámításához. Az $e^{-2\pi i \frac{ux}{N}}$ kifejezéseket egyszerre ki lehet számítani, majd el lehet tárolni őket egy táblázatban a későbbi felhasználáshoz. Emiatt ezekben a kifejezésekben az u és x szorzását nem szokták az implementáció közvetlen részeként kezelni.

A következőkben megmutatjuk, hogy a fenti egyenlet megfelelő felbontásával a szorzó és összeadó műveletek száma $N \log_2 N$ -nel arányossá tehető. A szorzás és összeadás műveletek arányának N^2 -ről $N \log_2 N$ -re csökkentése jelentős számítási igény megtakarítását jelenti, ahogy ezt a lenti táblázat is

mutatja. Például, a transzformáció direkt implementációjának alkalmazása $N = 8192$ esetén egy olyan számítógépen, mint az IBM 7094, háromnegyed órát vesz igénybe. Ugyanaz a feladat az FFT alkalmazásával körülbelül 5 perc alatt végezhető el.

N	N^2 (Direkt FT)	$N \log_2 N$ (FFT)	Számítási nyereség ($N / \log_2 N$)
2	4	2	2,00
4	16	8	2,00
8	64	24	2,67
16	256	64	4,00
32	1 024	160	6,40
64	4 096	384	10,67
128	16 384	896	18,29
256	65 536	2 048	32,00
512	262 144	4 608	56,89
1024	1 048 576	10 240	102,40
2048	4 194 304	22 528	186,18
4096	16 777 216	49 152	341,33
8192	67 108 864	106 496	630,15

1. táblázat. Direkt FT és FFT számítási igénye.

A figyelmünket fordítsuk az egyváltozós FFT algoritmus tárgyalására. A kétdimenziós Fourier-transzformáció könnyen kiszámítható az egydimenziós transzformáció többszöri, egymás utáni alkalmazásával.

Az az FFT algoritmus, amely ebben a fejezetben tárgyalásra kerül, a „szukcesszív kettőzés” módszerén alapul. A következőkben a korábbi egyenletet az alábbi formába írhatjuk:

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) W_N^{ux},$$

ahol

$$W_N = e^{\frac{-2\pi i}{N}},$$

és N -ről feltételezzük, hogy felírható $N = 2^n$ alakban, ahol n pozitív szám. Ennek alapján N felírható $N = 2M$ alakban, ahol M is pozitív egész szám. Ezek alapján azt kapjuk, hogy

$$F(u) = \frac{1}{2M} \sum_{x=0}^{2M-1} f(x) W_{2M}^{ux} = \frac{1}{2} \left\{ \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_{2M}^{u(2x)} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_{2M}^{u(2x+1)} \right\}$$

Mivel $W_{2M}^{2ux} = W_M^{ux}$, ezért azt kapjuk, hogy

$$F(u) = \frac{1}{2} \left\{ \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux} + \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_M^{ux} W_{2M}^u \right\}.$$

Ha definiáljuk az

$$F_{\text{páros}}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x) W_M^{ux}, u = 0, 1, 2, \dots, M-1,$$

$$F_{\text{páratlan}}(u) = \frac{1}{M} \sum_{x=0}^{M-1} f(2x+1) W_M^{ux}, u = 0, 1, 2, \dots, M-1,$$

függvényeket, akkor a fent megnevezett egyenlet a következőképpen írható át:

$$F(u) = \frac{1}{2} \{ F_{\text{páros}}(u) + F_{\text{páratlan}}(u) W_{2M}^u \}.$$

Emellett, mivel $W_M^{u+M} = W_M^u$ és $W_{2M}^{u+M} = -W_{2M}^u$, ezért

$$F(u+M) = \frac{1}{2} \{ F_{\text{páros}}(u) - F_{\text{páratlan}}(u) W_{2M}^u \}.$$

Az egyenletek figyelmes vizsgálatával felismerhető az összefüggések néhány érdekes tulajdonsága. Az $F(u)$ első felének kiszámításához két $(N/2)$ -pontos transzformáció kiszámítása szükséges. Az $F_{\text{páros}}(u)$ és $F_{\text{páratlan}}(u)$ értékei kerülnek behelyettesítésre, így kapjuk meg az $F(u)$ értékét. A másik fél közvetlenül adódik az egyik egyenlet alapján, további transzformációs számítások nélkül.

A fentebb ismertetett eljárás számítási igényeinek vizsgálatához legyen $m(n)$ a szükséges komplex szorzások száma, míg $a(n)$ jelentse a szükséges összeadások számát, amelyek a módszer implementálásához kellenek. Ahogy eddig is, a minták száma legyen 2^n , ahol n pozitív szám.

Először tegyük fel, hogy $n = 1$. Egy kétpontos transzformáció kiszámításához szükségünk van $F(0)$ értékére, majd az $F(1)$ -re. Ahhoz, hogy $F(0)$ -t megkapjuk, először ki kell számítanunk $F_{\text{páros}}(0)$ és $F_{\text{páratlan}}(0)$ értékét. Ebben az esetben $M = 1$ és egy pontos transzformációkat kell számolnunk. Mivel az egyedülálló pont Fourier-transzformáltja maga a pont, ezért szorzások és összeadások sem szükségesek $F_{\text{páros}}(0)$, illetve $F_{\text{páratlan}}(0)$ kiszámításához. Az $F_{\text{páratlan}}(0)$ és W_2^0 szorzata és egy összeadás adja $F(0)$ -t. Azután $F(1)$ következik még egy összeadás segítségével (a kivonást az összeadással azonosnak tekintjük). Mivel $F_{\text{páros}}(0)W_2^0$ -t már egyszer kiszámítottuk, megkaptuk a teljes művelet igényét egy kétpontos transzformációnak: $m(1) = 1$ szorzás és $a(1) = 2$ összeadás.

A következő érték, amelyet n felvehet, a 2. A fentebb leírt gondolatmenetnek megfelelően, egy négyponthoz transzformáció két részre osztható fel. Az $F(u)$ egyik felének kiszámításához szükség van két darab, kétpontos transzformációra, ahogyan az az egyenletekben szerepel $M = 2$ esetén. Mivel egy kétpontos transzformációhoz $m(1)$ szorzás és $a(1)$ összeadás szükséges, ezért nyilvánvaló, hogy e két egyenlet értékének meghatározásához $2m(1)$ szorzás és $2a(1)$ összeadás kell. Két további szorzás illetve összeadás szükséges $F(0)$ és $F(1)$ meghatározásához. Mivel $F_{\text{páros}}(u)W_2^{2M}$ értékét már meghatároztuk az $u = \{0; 1\}$ -re, így még két összeadás megadja $F(3)$ és $F(4)$ értékét. A teljes számítási igény tehát: $m(2) = 2m(1) + 2$ és $a(2) = 2a(1) + 4$.

Ha n értéke 3-mal egyenlő, akkor $F_{\text{páros}}(u)$ és $F_{\text{páratlan}}(u)$ meghatározásához két négyponthoz transzformációval kell számoljunk. Ehhez $2m(2)$ szorzásra és $2a(2)$ összeadásra van szükség. További négy szorzás és nyolc összeadás adja az teljes transzformációt. A teljes számítási igény így $m(3) = 2m(2) + 4$ és $a(3) = 2a(2) + 8$.

Az argumentum további növelésével azt vehetjük észre, hogy bármely pozitív egész n -re, az FFT implementálásához szükséges szorzások és összeadások száma a következő rekurzív kifejezésekkel adható meg:

$$m(n) = 2m(n-1) + 2^{n-1}, n \geq 1,$$

$$a(n) = 2a(n-1) + 2^n, n \geq 1,$$

ahol $m(0) = 0$ és $a(0) = 0$, mivel egyetlen pont transzformálása nem igényel szorzásokat és összeadásokat.

A fentiek elvek implementációja jelenti a „szukcesszív kettőzés” FFT algoritmusát. A név abból ered, hogy egy kétpontos transzformáció két darab egyponthoz transzformációból számítható, egy négyponthoz transzformáció két kétpontos transzformációból, és így tovább, minden olyan N -re, amelynek értéke 2 valamely egész kitevőjű hatványa.

1.8.1. *A műveletek száma.* Ebben a részben indukció segítségével fogjuk megmutatni, hogy a fentebb megadott FFT algoritmus végrehajtásához szükséges komplex szorzások és összeadások száma

$$m(n) = \frac{1}{2}2^n \log_2 2^n = \frac{1}{2}N \log_2 N = \frac{1}{2}Nn, \quad n \geq 1$$

összefüggés által, valamint

$$a(n) = 2^n \log_2 2^n = N \log_2 N = Nn, \quad n \geq 1$$

által adott.

Először be kell látnunk, hogy a fentiek igazak $n = 1$ -re. Azt már megmutattuk, hogy $m(1) = 1$ és $a(1) = 2$.

Ezután tételezzük fel, hogy a kifejezések igazak n -re. Ekkor azt kell belátnunk, hogy $n + 1$ -re is igazak. Könnyű látni, hogy

$$m(n + 1) = 2m(n) + 2^n$$

Ebből behelyettesítéssel kapjuk, hogy:

$$m(n + 1) = 2 \left(\frac{1}{2} Nn \right) + 2^n = 2 \left(\frac{1}{2} 2^n n \right) + 2^n = 2^n (n + 1) = \frac{1}{2} 2^{n+1} (n + 1).$$

Így azt kapjuk, hogy

$$a(n + 1) = 2a(n) + 2^{n+1}$$

összefüggést, amelyből

$$a(n + 1) = 2Nn + 2^{n+1} = 2(2^n n) + 2^{n+1} = 2^{n+1}(n + 1).$$

1.8.2. *Az inverz FFT.* Eleddig érintőlegesen foglalkoztunk az inverz Fourier-transzformációval. Kiderült róla, hogy bármely algoritmus, amely implementálja a diszkrét Fourier-transzformációt, az bemeneti adatok kisebb módosításával felhasználható az inverz kiszámítására is. Ennek átgondolásához vegyük az

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) e^{-2\pi i \frac{ux}{N}}$$

és az

$$f(x) = \sum_{u=0}^{N-1} F(u) e^{2\pi i \frac{ux}{N}}$$

összefüggéseket. A második egyenlet komplex konjugáltját véve, majd mindkét oldalt elosztva N -el kapjuk, hogy

$$\frac{1}{N} f^*(x) = \frac{1}{N} \sum_{u=0}^{N-1} F^*(u) e^{-2\pi i \frac{ux}{N}}.$$

Az eredményt az első egyenlettel összehasonlítva látható, hogy az utoljára kapott egyenlet jobb oldala a normál Fourier-transzformációval megegyező alakban van. Ezért, ha egy olyan algoritmus inputja $F^*(u)$, mely a normál irányú transzformáció kiszámítására készült, az eredmény $f^*(x)/N$ lesz. Ennek a komplex konjugáltját véve és megszorozva N -el megkapjuk a kívánt inverz $f(u)$ -t.

A kétdimenziós esetben véve a komplex konjugáltját a következőt kapjuk:

$$f^*(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F^*(u, v) e^{-2\pi i \frac{(ux+vy)}{N}}.$$

Ez láthatóan a kétdimenziós normál irányú transzformáció formájában van. Ebből következik, hogy ha egy olyan algoritmus inputjának választjuk $F^*(u, v)$ -t, amely a normál irányú transzformáció kiszámítására készült, akkor az eredmény $f^*(x, y)$ lesz. Ennek a komplex konjugáltját véve megkapjuk $f^*(x, y)$ -t. Abban az esetben, ha $f(x)$ és $f(x, y)$ valósak, a komplex konjugáltra nincs szükség, mivel $f(x) = f^*(x)$ és $f(x, y) \equiv f^*(x, y)$ a valós függvényekre.

1.8.3. *Implementáció.* Az ebben a fejezetben ismertetett FFT számítógépes megvalósítása egyszerű, lényegre törő. A legfontosabb dolog, amelyet észben kell tartani, hogy a bemenő adatokat a "szukcesszív" alkalmazásához megfelelő módon kell rendezni. A rendezési eljárást egy egyszerű példával illusztrálhatjuk.

Tételezzük fel, hogy az FFT kiszámításához a „szukcesszív kettőzés” módszerét akarjuk alkalmazni egy nyolcpontos függvényre, amelynek pontjai:

$$\{f(0), f(1), \dots, f(7)\}.$$

Ahogy már korábban már láttuk egyik menetben a páros argumentumokat, másik menetben a páratlan argumentumokat fogjuk használni.

Mindkét négyponos transzformáció kétpontos transzformációként kerül kiszámításra. Az első halmaz FFT-jének a kiszámításához fel kell bontanunk azt a páros részére $\{f(0), f(4)\}$ és a páratlan részére $\{f(2), f(6)\}$. Hasonlóan, a második halmazt is felbontjuk, $\{f(1), f(5)\}$ és a $\{f(3), f(7)\}$. További átcsoportosításra nincs szükség, mivel a kételemű halmazokat úgy tekintjük, hogy egy páros és egy páratlan elemből állnak. Az eredmények kombinálásával megkapjuk a bemeneti tömböt, ami $\{f(0), f(4), f(2), f(6), f(1), f(5), f(3), f(7)\}$.

Szerencsére a bemenet újrendezése általános esetben egy egyszerű bitfordító (bit-reversal) szabályt követ. Ha x az $f(x)$ bármely érvényes argumentumát jelöli, akkor a hozzá tartozó argumentumot a rendezett tömbben az x bináris reprezentációjának fordított sorrendű bitjei adják. Például, ha $N = 2^3$, akkor a hetedik elem az eredeti tömbben, $f(6)$, a negyedik elem lesz a rendezettben, mivel $6 = 110_2$, melyből $011_2 = 3$ lesz, ha a biteket megfordítjuk. Fontos, hogy ez egy balról jobbra történő megfordítást, nem szabad összekeverni a bináris komplementessel. A módszer összefoglalása látható a lenti táblázatban $N = 8$ esetén. Ha a rendezett tömböt használjuk az FFT kiszámításánál, az eredmény a Fourier-transzformáció, a megfelelő elemsorrenddel. Megfordítva, megmutatható, hogy ha a tömb a természetes sorrendben kerül felhasználásra, akkor az eredmény bit-fordítva áll elő. Ugyanez vonatkozik az inverz kiszámítására.

Eredeti argumentum	Eredeti tömb	Bitforgatott argumentum	Rendezett tömb
0 0 0	f(0)	0 0 0	f(0)
0 0 1	f(1)	1 0 0	f(4)
0 1 0	f(2)	0 1 0	f(2)
0 1 1	f(3)	1 1 0	f(6)
1 0 0	f(4)	0 0 1	f(1)
1 0 1	f(5)	1 0 1	f(5)
1 1 0	f(6)	0 1 1	f(3)
1 1 1	f(7)	1 1 1	f(7)

2. táblázat. A tömbelemek sorrendjének meghatározása.

1.8.4. *Gyors inverz Fourier-transzformáció.* Az inverz FFT kiszámításához az eredeti transzformációt is felhasználhatjuk a következő azonosság alapján. Az inverz diszkrét Fourier-transzformáció egyenletének mindkét oldalán komplex konjugáltat képezve és N -nel osztva az

$$\frac{1}{N} f^*(x) = \frac{1}{N} \sum_{u=0}^{N-1} F^*(u) e^{\frac{-2\pi i u x}{N}}$$

kifejezést kapjuk. Azt látjuk, hogy az egyenlet jobb oldala egy Fourier-transzformáció. Így ha egy Fourier-transzformáció számítását elvégző algoritmus bemeneteként az $F^*(u)$ -t adjuk meg, az eredmény az $\frac{f^*(x)}{N}$ lesz. Ennek a komplex konjugáltját képezve és N -nel szorozva az $f(x)$ függvényt kapjuk eredményül. Kétdimenziós esetben az inverz diszkrét Fourier-transzformált az

$$f^*(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F^*(u, v) e^{\frac{-2\pi i (ux+vy)}{N}}$$

alakban áll elő. Hasonlóan az egydimenziós esethez, ha a Fourier-transzformációs algoritmus bemenete az $F^*(u, v)$ függvény, úgy a $f^*(x, y)$ -t kapjuk, ennek komplex konjugáltját véve az $f(x, y)$ függvényt kapjuk eredményül. Abban az esetben, ha a kétdimenziós inverz transzformálásánál az egydimenziós algoritmust használjuk két lépésben, úgy a módszer nem az, hogy minden sor- és oszloptranszformálásnál az eredmény komplex konjugáltját képezzük, hanem az $f^*(u, v)$ -t úgy kezeljük, mint az $f(x, y)$ -t az „oda” transzformálásnál és ennek végén képezzük az eredmény komplex konjugáltját, illetve megszorozzuk N -nel.

2. Walsh-transzformáció

A Walsh-transzformáció formálisan úgy kezelhető, mint a Fourier-transzformáció, csak más ortonormált periodikus függvénybázisra épül. Azonban a Walsh transzformáció kezelése természetesebb.

Tekintsük át a transzformációt vázlatosan. Az egydimenziós diszkrét Walsh-transzformációs párt az $N = 2^n$ esetén a következő alakban definiáljuk:

$$W(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u)},$$

$$f(x) = \sum_{u=0}^{N-1} W(u) \prod_{i=0}^{n-1} (-1)^{(b_i(x)b_{n-1-i}(u))},$$

ahol $b_k(z)$ a z bináris reprezentációjának k -adik bitje. Érdekes megjegyezni, hogy a Walsh-transzformáció és az inverz Walsh-transzformáció csak egy $\frac{1}{N}$ -es együtthatóval különbözik egymástól. Így gyakorlatilag a két transzformáció során ugyanazt az algoritmust használhatjuk.

A kétdimenziós diszkrét Walsh-transzformációs pár:

$$W(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \prod_{i=0}^{n-1} (-1)^{(b_i(x)b_{n-1-i}(u) + b_i(y)b_{n-1-i}(v))}$$

és

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} W(u, v) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-1-i}(u) + b_i(y)b_{n-1-i}(v)}.$$

Ebből következik, hogy a kétdimenziós Walsh-transzformációs algoritmus változtatás nélkül használható az inverz transzformáció számítására is. A kitevőben lévő összeadást modulo-2 aritmetika szerint kell elvégezni.

A Walsh-transzformáció is szeparábilis, így a kétdimenziós Walsh-transzformáció két egymásután végrehajtott egydimenziós transzformációval számítható. Így a Walsh-transzformáció és inverze a Fourier-transzformációhoz hasonló módon számítható.

3. Hadamard-transzformáció

Az egydimenziós *Hadamard-transzformáció*s pár a következőképpen adható meg:

$$H(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)},$$

$$f(x) = \sum_{u=0}^{N-1} H(u) (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)},$$

ahol $N = 2^n$. A kitevőben lévő összeadást modulo 2 aritmetika szerint kell elvégezni. Itt is megfigyelhető, hogy az „oda” és az inverz transzformáció csak az $\frac{1}{N}$ -es együtthatóval különbözik egymástól. A kétdimenziós Hadamard transzformációs pár:

$$H(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) (-1)^{\sum_{i=0}^{n-1} (b_i(x) b_i(u) + b_i(y) b_i(v))}$$

és

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} H(u, v) (-1)^{\sum_{i=0}^{n-1} (b_i(x) b_i(u) + b_i(y) b_i(v))}.$$

A kitevőben lévő összeadást modulo-2 aritmetika szerint kell elvégezni. A Walsh-transzformáció és az Hadamard-transzformáció között szembetűnő a hasonlóság. Tanulmányozzuk a

$$wal(u, x) = \prod_{i=0}^{n-1} (-1)^{b_i(x) b_{n-1-i}(u)}$$

kifejezés alapján kiszámolt Walsh-transzformációs mátrix, illetve a

$$had(u, x) = (-1)^{\sum_{i=0}^{n-1} b_i(x) b_i(u)}$$

kifejezés alapján kiszámolt Hadamard-transzformációs mátrix közötti különbséget. Megfigyelhető, hogy a két mátrix elemei megegyeznek, csak a sorok, illetve az oszlopok sorrendje más. Valójában $N = 2^n$ esetén csak ez az egy különbség van a két transzformáció között. Mivel a képfeldolgozó algoritmusok esetén gyakran teljesül az előbb említett feltétel, így az irodalom a két transzformációt *Walsh-Hadamard-transzformáció*ként említi. Az *Hadamard-mátrix* generálására megadható egy rekurzív algoritmus. A legalacsonyabb rendű Hadamard-mátrix a következőképpen adható meg: $H_2 =$

$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Ezután a H_{2N} -nel jelölt $2N$ -edrendű mátrixot, a következőképpen állíthatjuk elő:

$$H_{2N} = \begin{pmatrix} H_N & H_N \\ H_N & -H_N \end{pmatrix},$$

feltételezve, hogy $N = 2^n$.

4. Diszkrét Cosinus-transzformáció

Az irodalomban a diszkrét Cosinus-transzformációt (DCT) többféleképpen szokták definiálni. Ezek közül a leggyakrabban használt alak kétdimenziós esetben a következő:

$$F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{2x+1}{2N}u\pi\right) \cos\left(\frac{2y+1}{2N}v\pi\right).$$

5. Hough-transzformáció

A Hough-transzformáció a digitális képfeldolgozásban általánosan használt módszer, amely egy adott kép objektumpontjainak olyan részhalmazait határozza meg, amelyekre közös egyenes illeszthető. Az elkövetkezőben vázlatosan ismertetjük a *Hough-transzformációt* folytonos esetben, majd ezek után áttérünk a diszkrét realizációjára.

Legyen F az $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ ($n \in \mathbb{N}$) objektumpontok halmaza. Egy (x_i, y_i) rögzített pont esetén az $y_i = ax_i + b$ egyenlet az (a, b) paraméterek értékétől függően egyenesek egy végtelen rendszerét határozza meg. Az alapötlet az, hogy a fenti egyenlet xy síkbeli vizsgálata helyett a $b = -x_i a + y_i$ egyenletet vizsgáljuk az ab síkban. Ennek értelmében minden F -beli (x_i, y_i) ponthoz rendeljük hozzá egy, a fenti összefüggés által meghatározott ab síkbeli egyenest. Ezek után, ha például az (x_i, y_i) és az (x_j, y_j) pontok illeszkednek egy (a, b) paraméterű egyenesre, akkor a fenti két ponthoz hozzárendelt egyenesek metszik egymást az (a, b) koordinátájú pontban. Ezen alapelvek felhasználásával k darab pont közös egyenesre való illeszkedésének az ab paramétersíkban hozzájuk rendelt k darab egyenes egy pontban történő metszése felel meg.

A diszkrét esetben az ab paramétersík $a_{min}, a_{max}, b_{min}$ és b_{max} értékek által meghatározott véges tartományát osszuk fel ekvidisztáns beosztást használva $n \cdot m$ darab részre, amely részeket celláknak fogjuk nevezni. Az $a_{min}, a_{max}, b_{min}$ és b_{max} kísérletileg meghatározott rögzített értékek. Minden cellához rendeljük hozzá egy nemnegatív egész számot, amely majd

az adott cellán áthaladó egyenesek számát jelzi. Nyilván kezdetben minden cella esetén ez az érték 0. Ezen inicializációs lépés után az F minden elemére és az összes $a_{min} \leq a \leq a_{max}$ osztáspontra határozzuk meg a b értékét, majd az (a, b) koordinátájú cellához rendelt számot eggyel növeljük meg.

A Hough-transzformáció végrehajtása után, ha az (a, b) koordinátájú cella tartalma k , akkor az (a, b) paraméterű egyenesre az F halmaz k pontja illeszkedik.

Nyilván esetünkben az F halmaz az adott pillanatban rögzített bináris kép objektumpontjainak halmaza lesz.

V. fejezet

Képjavítások

A képjavitási eljárások célja, hogy a képet a további kiértékelés vagy feldolgozás szempontjából előnyösebb formába hozzuk, és nem törekszünk arra, hogy minél jobban megközelítsük az eredetit, vagy az ideális leképezést. Éppen ellenkezőleg, esetenként egyes jellegzetességek kiemelésével, hangsúlyozásával kapjuk a számunkra kedvezőbb eredményt.

A képjavitási eljárások két nagy kategóriába sorolhatók attól függően, hogy azok a paraméterterben vagy a frekvenciaterben dolgoznak. A paraméterterben ható transzformációk közül a *világosságkód-transzformációkkal*, a frekvenciaterbeli transzformációk közül a *szűréssel* fogunk részletesebben foglalkozni.

1. Világosságkód-transzformáció

Az egyik leggyakrabban előforduló képhiba a nem megfelelő fénysűrűségből, illetve a leképező rendszerben keletkezett fényvesztéséből adódó kontraszt-szegénység. Ez szerencsés esetben az egész képen egyenletesen jelentkezik, de igen gyakran a képen belül is változik a kontraszt. A világosságkód-transzformációk célja a kontraszt növelése a világosság kódok eloszlásának megváltoztatásával.

1.1. Hisztogram-transzformáció

Az r reprezentálja a képpontok világosság kódjait. Az egyszerűség kedvéért az r -et úgy normalizáljuk, hogy a $[0, 1]$ tartományba essen. Ebben az esetben a hisztogram-transzformáció az $s = T(r)$ alakban írható fel, ahol $r \in [0, 1]$.

Tételezzük fel, hogy a hisztogram-transzformáció kielégíti az alábbi két feltételt:

- $T(r)$ a $[0, 1]$ intervallumban monoton növekvő függvény,
- $0 \leq T(r) \leq 1$, ha $0 \leq r \leq 1$.

Nyilvánvalóan az $r = T^{-1}(s)$ inverz transzformáció is kielégíti a fenti két feltételt.

A képpontok világosságkódjai valószínűségi eloszlást alkotnak, amit a hisztogrammal jellemezhetünk. Tételezzük fel egy pillanatra, hogy az eredeti, illetve a transzformált kép folytonos hisztogramját $p_r(r)$ -rel, illetve $p_s(s)$ -sel jelöljük. Abban az esetben, ha a $p_r(r)$ és a $T(r)$ ismert, akkor a $p_s(s)$ -re a következő összefüggést írhatjuk fel:

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

1.2. Hisztogram-kiegyenlítés

Az egyik legjelentősebb világosságkód-transzformáció a *hisztogram-kiegyenlítés*.

A kontraszt szegény kép hisztogramját széthúzva, kiegyenlítve a képen nem domináló szürkességi szintek is hangsúlyt kapnak. A módszernél a transzformációs függvényt az alábbi alakban írhatjuk fel:

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

Az előző egyenlet jobb oldalán az r valószínűségi változó kummulativ eloszlás függvényét ismerjük fel, így kielégíti a korábbi két feltételt. A fentiekből könnyen adódik a következő kapcsolat az r és az s között:

$$\frac{ds}{dr} = p_r(r).$$

Ezek alapján azt kapjuk, hogy

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

Most vizsgáljuk meg a diszkrét esetet. Az egyes világosság kódok előfordulási valószínűségét a

$$p_r(r_k) = \frac{n_k}{n}, \quad (0 \leq r_k \leq 1, k = 0, 1, \dots, L-1)$$

kifejezéssel közelíthetjük meg, ahol L a világosság kódok száma, n_k a k -adik világosság kódú pontok száma, n az összes képpontok száma, $p_r(r_k)$ a k -adik világosság kód előfordulási valószínűsége. A transzformációs formula diszkrét esetben a következő alakú:

$$s_k = T(r_k) = \sum_{j=0}^k \frac{n_j}{n} = \sum_{j=0}^k p_r(r_j).$$

Az inverz transzformációt a következőképpen írhatjuk fel:

$$r_k = T^{-1}(s_k) \quad 0 \leq s_n \leq 1.$$

Mind a $T(r_k)$, mind $T^{-1}(s_k)$ transzformációról feltételezzük, hogy teljesíti a korábban említett feltételeket.

A hisztogram-kiegyenlítés ezek után a következő lépésekből áll:

- Kiszámítjuk a kép hisztogramja alapján a $p_r(r_k)$ értékeket.
- A (20) alapján kiszámítjuk az s_k értékeit.
- Az s_k kiszámolt értékeit a meglévő világosság kódokkal közelítjük, amivel megkapjuk, hogy az eredeti r_k értékek most milyen világosság-kódra változnak a közelített s_k -k alapján.

A hisztogram kiegyenlítés jól alkalmazható minden olyan esetben, ahol az eredeti kép hisztogramja egy vékony csúcsból áll. Az eljárás ezt a csúcsot széthúzza, amivel a kép kontrasztossága jelentősen megnő.

1.3. Küszöbölés

A küszöbölési technika kiindulási alapja szintén a hisztogram. A cél a képen előforduló világosság kódok számának drasztikus csökkentése, hogy a kép vizuális információ tartalma a legkisebb mértékben változzon.

Legyenek a megadott küszöbök $k_1, k_2, \dots, k_{n-1}$. A legkisebb világosságkódot $q_{\min}$ -nel, a legnagyobb világosságkódot $q_{\max}$ -szal jelölve vezessük be a $k_0 = q_{\min}$ és a $k_n = q_{\max}$ jelöléseket. Ekkor az n -küszöbölés az

$$f'(x, y) = \begin{cases} r_1 & k_0 \leq f(x, y) \leq k_1, \\ r_2 & k_1 \leq f(x, y) \leq k_2, \\ \dots & \dots \\ r_n & k_{n-1} \leq f(x, y) \leq k_n. \end{cases}$$

hozzárendelési utasítással írhatjuk le, ahol r_i az i -edik küszöbhez előre megadott világosság kód.

Igen gyakori eset a *két szintre vágás*, amikor a képpontokat — egy küszöb megadásával — előtér, illetve háttérpontokká minősítjük.

A küszöbölések speciális esete a *sávkivágás*, amihez két küszöbre (k_1 és k_2), valamint két világosságkódra (r_1 és r_2) van szükség. Az eljárás során a $k_1 \leq f(x, y) < k_2$ világosságkódú képpontokat változatlanul hagyjuk, a többi pedig r_1 világosságkódúra színezzük át. A *sávkivágás kiemeléssel* azt jelenti, hogy a korábbi eljárásban változatlanul hagyott képpontok világosságkódját r_2 -re állítjuk.

Sávkiemelésről beszélünk abban az esetben, ha a k_1 és k_2 küszöbök közé eső világosságkódú képpontokat világosságkódját r_2 -re állítjuk be, de a többi képpont világosságkódját nem változtatjuk meg.

1.4. Környezeti átlagolás

A környezeti átlagolás a legegyszerűbb képsimító eljárás. Az $f(x, y)$ képen az (x, y) képpont világosságkódja egy előre definiált környezet átlagos világosságkódjával lesz egyenlő. Az átlagolt kép:

$$g(x, y) = \frac{1}{N} \sum_{(n, m) \in S} f(n, m)$$

ahol S a környezeti pontok koordinátáinak halmaza, N a környezeti pontok száma. A gyakorlatban a környezet alatt 3×3 és 5×5 -ös ablakokat értünk, amelynek centrális eleme az (x, y) koordinátájú pont.

A környezeti átlagolás hátránya az, hogy erősen homályosítja a képet. A homályosítási effektus csökkentése végett célszerű küszöbölni az átlagolás során. Ez azt jelenti, hogy a képpont világosságkódja csak akkor változik a környezetének átlag világosságkód értékére, ha különbségük meghalad egy adott küszöbértéket. Tehát

$$g(x, y) = \begin{cases} \frac{1}{N} \sum_{(n, m) \in S} f(n, m), & \text{ha } |f(x, y) - \frac{1}{N} \sum_{(n, m) \in S} f(n, m)| > T, \\ f(x, y), & \text{egyébként,} \end{cases}$$

ahol T a megadott küszöbérték.

1.5. Képek átlagolása

Ezt a zajszűrési módszer akkor használhatjuk eredményesen, ha ugyanarról a jelenetről több felvétel készült. Tegyük fel, hogy a képeink az eredeti $f(x, y)$ kép és valamilyen $\mu(x, y)$ zaj összegéből állnak:

$$g_i(x, y) = f(x, y) + \mu(x, y), \quad i \in I.$$

Az $\mu(x, y)$ zajról tegyük fel, hogy korrelálatlan és nulla várható értékű. Könnyű bebizonyítani, hogy több zajos kép átlagát képezve

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

az átlag várható értéke az eredeti kép lesz

$$E(\bar{g}(x, y)) = f(x, y),$$

valamint az átlag szórásnégyzete:

$$\sigma^2(\bar{g}(x, y)) = \frac{1}{N} \sigma^2(\mu(x, y)).$$

1.6. Medián szűrő

A median szűrő hatásos zajszűrő. Az (x, y) koordinátájú pont $m \times m$ -es környezetét (m páratlan) vizsgálva az (x, y) koordinátájú pont világosságkódját azon világosságkód értékkel helyettesítjük, amely – a környezetében lévő m^2 darab világosságkód értéket nagyság szerint rendezve – a rendezett sorozatban a középső helyen áll.

1.7. Box-módszer

Hasonlóan a medián szűrőhöz itt is az (x, y) koordinátájú pont $m \times m$ -es környezetét vizsgáljuk. A vizsgált képpont világosságkódját komparáljuk a környezet képpontjaival. Ha a környezetben található képpontok által meghatározott világosságkód-tartományba (minimum, maximum) beleesik a vizsgált képpont világosságkódja, akkor azt változatlanul hagyjuk. Ellenkező esetben a világosságkód-tartomány legalsó, illetve legalsó szintjére változtatjuk, attól függően, hogy melyikhez van közelebb.

2. Szűrés

A frekvenciatartománybeli szűrés alapja a konvolúció tétel. Legyen $g(x, y)$ az $f(x, y)$ kép és a $h(x, y)$ pozíció invariáns függvény konvolúciós szorzata:

$$g(x, y) = h(x, y) * f(x, y).$$

A konvolúciós tétel alapján:

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

ahol a $G(u, v)$ a $g(x, y)$, $H(u, v)$ a $h(x, y)$ és $F(u, v)$ az $f(x, y)$ függvény Fourier-transzformáltja. A $H(u, v)$ függvényt *szűrő függvénynek* nevezzük.

2.1. Aluláteresztő szűrő

A zajok és a világosságkódokban mutatkozó éles átmenetek a kép Fourier-transzformáltjának magasfrekvenciás összetevőiben jelentkeznek. Ebből következik, hogy a zajok elnyomásának egyik lehetséges módja a frekvenciatartományban végzett szűrés. Ez azt jelenti, hogy a $G(u, v) = F(u, v) \cdot$

$H(u, v)$ egyenletben szereplő $H(u, v)$ függvényt úgy kell megválasztani, hogy a kép $F(u, v)$ Fourier-transzformáltjából kiszűrje a magasfrekvenciás összetevőket, de az alacsonyabb frekvenciájú komponenseket lehetőleg változatlanul engedje át. Az ilyen szűrőket aluláteresztő szűrőknek nevezzük.

A legegyszerűbb ilyen szűrő az *ideális aluláteresztő szűrő* (ILPF) átviteli függvénye a következő:

$$H(u, v) = \begin{cases} 1, & \text{ha } \delta(u, v) \leq \delta_0, \\ 0, & \text{különben.} \end{cases}$$

ahol $\delta(u, v)$ az (u, v) pont origótól vett távolsága. A definícióból látható, hogy az ideális szűrő változatlanul átengedi a δ_0 sugarú kör belsejébe eső alacsonyfrekvenciás összetevőket, a körön kívüli magasabb frekvenciájukat pedig teljesen kiszűri. Az δ_0 -t *vágási frekvenciának* nevezzük.

Az ideális szűrő hatását tekintve bizonyos tekintetben korántsem ideális, minthogy a zajokkal együtt kiszűri a magasabb frekvenciatartományba eső éleket is, miáltal a kép igen jelentősen elhomályosodik. Ezen kívül fellép egy olyan hatás, ami a képen peridodikusan ismétlődő foltok megjelenését eredményezi. A fenti problémák kiküszöbölésére különböző, „sima” szűrőt szoktak alkalmazni.

Ilyen például az *aluláteresztő n -ed rendű Butterworth-szűrő* (BLPF), amely a következő összefüggéssel adható meg:

$$H(u, v) = \frac{1}{1 + \left(\frac{\delta(u, v)}{\delta_0}\right)^{2n}}.$$

2.2. Felüláteresztő szűrő

A torzítások gyakran úgy jelentkeznek, hogy a határátmenetek kiszélesednek, az élek elmosódnak. Sok esetben akkor is alkalmazzuk az élkiemelési eljárásokat, ha ilyen torzítás fel sem lép, mivel pszichofizikai kísérletek azt bizonyítják, hogy szubjektíve előnyösebb érzetet kelt a túlhangsúlyozott élekkel rendelkező kép, mint a valóság-hű ábrázolás.

Az eljárás célja az egyes képrészletek közötti átmeneti tartomány szűkítése, a határátmenetek hangsúlyozása, az elmosódások korrigálása. Az élkiemelést gyakran alkalmazzák a szegmentálásnál tárgyalt különböző élkitűzési eljárások előkészítő fázisaként is.

Az *ideális felüláteresztő szűrő* (IHPF) a következőképpen definiálható:

$$H(u, v) = \begin{cases} 0, & \text{ha } \delta(u, v) \leq \delta_0, \\ 1, & \text{különben.} \end{cases}$$

A *felüláteresztő n -ed rendű Butterworth-szűrő* (BHPF) a következőképpen definiálható:

$$H(u, v) = \frac{1}{1 + (\frac{\delta_0}{\delta(u, v)})^{2n}}.$$

VI. fejezet

Képrekonstrukció

Az előző fejezetben láttuk, hogy a képjavító algoritmusok legfőbb célkitűzése a kép valamilyen értelemben vett javítása. A javítási processzus gyakorlatilag passzív szűréssel valósul meg. Ezzel szemben a képrekonstrukció célja a kép minőségének konstruktív javítása, amely valamilyen módon figyelembe veszi a kép torzulásának módját. A képrekonstrukció alapproblémája: hogyan tudjuk az észlelt képből a legjobban becsülni az eredeti képet. A becslésünk jósága nagymértékben függ a torzítás fizikai okának ismeretéből, amit a-priori ismeretnek nevezünk.

A továbbiakban két alapvető képrekonstrukciós eljárást ismertetünk, ahol a torzítást a modellből eredően, ismertnek tekintjük. Mindkét eljárás az úgynevezett konstruktív szűrés elvén alapul.

1. Inverz szűrő

A legkorábbi képrekonstrukciós eljárások az inverz szűrő elvén alapultak, ami azt jelenti, hogy a torzítás átviteli függvényének inverzét alkalmazták a becsült kép előállítására.

Jelölje $f_I(x, y)$ az ideális képet, $f_0(x, y)$ a megfigyelt képet, $f'_I(x, y)$ a becsült képet, $n(x, y)$ pedig a zajt! A zajt az ideális képpel korrelálatlannak tételezzük fel. A térbeli torzítást okozó függvény átviteli függvényét jelöljük $h_D(x, y)$ -nal. Ennélfogva a megfigyelt kép:

$$f_0(x, y) = f_I(x, y) * h_D(x, y) + n(x, y)$$

ahol a $*$ a konvolúciót jelöli. A rekonstrukciós rendszer egy lineáris eltolás-invariáns szűrőből áll, amelynek átviteli függvénye $h_R(x, y)$. Így a becsült kép:

$$f'_I(x, y) = f_0(x, y) * h_R(x, y) = (f_I(x, y) * h_D(x, y) + n(x, y)) * h_R(x, y).$$

Alkalmazva az előző kifejezésre a Fourier-transzformációt:

$$F'_I(u, v) = (F_I(u, v) \cdot H_D(u, v) + N(u, v)) \cdot H_R(u, v).$$

A rekonstrukciós szűrő átviteli függvényét a

$$H_R(u, v) = \frac{1}{H_D(u, v)}$$

kifejezéssel adjuk meg, így a becsült kép Fourier-transzformáltja

$$F'_I(u, v) = F_I(u, v) + \frac{N(u, v)}{H_D(u, v)}.$$

Ez a kifejezés azt mutatja, hogy ha a zaj nulla, akkor tökéletes rekonstrukciót kapunk. A zaj jelenlétével egy olyan additív hiba jelentkezik, amely a $H_D(u, v)$ kis értékeinél egész naggyá válik. Ez azt eredményezi, hogy a kép minősége nagyfrekvenciás komponensek esetén nagy mértékben romlik.

2. Wiener-szűrő

Mint láttuk az inverz szűrő hátránya, hogy a szűrő átviteli függvényében a zaj nem szerepelt. A Wiener-szűrő ezt a problémát orvosolja.

A szűrő alkalmazásakor fel kell tételeznünk, hogy a zaj additív, korrelálatlan a képpel, ismerjük a teljesítmény-sűrűség spektrumát, valamint a várható értéke zérus. A képről az egyszerű tárgyalás kedvéért feltételezzük, hogy a várható értéke zérus.

Tekintsük a megfigyelt képet az ideális kép és a zaj összegének és ábrázoljuk vektor formában:

$$f_0 = f_I + n.$$

Alkalmazva a Fourier-transzformációt:

$$F_0 = F_I + N$$

ahol F_0 , F_I , N az f_0 , f_I , n vektorok Fourier-transzformáltja. Ha a szűrő átviteli függvényének Fourier-transzformáltja G , úgy a becsült kép Fourier-transzformáltja:

$$F'_I = GF_0,$$

ahol a G mátrix operátor mérete $M \times M$, megegyezik Fourier-transzformációs mátrix méretével. A G szűrő mátrixot úgy kell méreteznünk, hogy az minimalizálja a rekonstrukció négyzetes hibáját. A négyzetes hibát felírva:

$$E_R = ((F_I - F'_I)^T (F_I - F'_I))^T,$$

ahol T transzponáltat jelöl. Behelyettesítve, valamint felhasználva a következő kifejezéseket

$$\begin{aligned} F_I F_I^T &= C_S \\ N N^T &= C_N \\ F_I N^T &= N F_I^T = \{0\} \end{aligned}$$

azt kapjuk, hogy

$$E_R = (C_S - 2GC_S + GG(C_S + C_N))^T$$

ahol C_S az F_I jel kovariancia mátrixa, C_N az N kovariancia mátrixa. Az előző három feltétel közül az utolsó azt jelenti, hogy az ideális kép és a zaj korrelálatlan. A kapott kifejezés minimalizálásával azt kapjuk, hogy

$$G_0 = C_S(C_S + C_N)^{-1}$$

Alkalmazva az inverz Fourier-transzformációt

$$g_0 = C_S(C_S + C_N)^{-1}$$

ahol C_S az f_I jel kovariancia mátrixa, C_N a zaj kovariancia mátrixa.

A Wiener-szűrő átviteli függvényét felírhatjuk a térbeli képtorzítás átviteli függvényének segítségével is. A levezetés mellőzésével a szűrő átviteli függvénye additív zaj esetén:

$$H_R(u, v) = \frac{H_D^*(u, v)}{|H_D(u, v)|^2 + U_N(u, v)}$$

ahol $U_N(u, v)$ a zaj teljesítmény-sűrűség függvénye.

VII. fejezet

Konvolúció és korreláció

Ebben a fejezetben a képtér és a frekvenciatér közötti két alapvető kapcsolatot vizsgálunk meg. Ez a két kapcsolat – a konvolúció és a korreláció – alapvető fontosságú az olyan képfeldolgozási technikák kidolgozásánál, melyek a Fourier-transzformáción alapulnak.

A vizsgálatot az egydimenziós konvolúcióval kezdjük, folytonos paraméterek mellett. Majd a diszkrét eset, végül a kétdimenziós folytonos és diszkrét eset kerül sorra. A korreláció esetében is hasonló elveket követünk.

1. Konvolúció

Két függvény, $f(x)$ és $g(x)$ konvolúciója, melyet $f(x) * g(x)$ -el jelölünk, a következő integrálfüggvénnyel definiált:

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(\alpha) g(x - \alpha) d\alpha,$$

ahol α az integrálás mesterséges változója. Mivel a konvolúciós integrál hatását vizuálisan megjeleníteni meglehetősen nehéz, ezért két példán keresztül, grafikusán mutatjuk be a használatát a fenti egyenletnek.

Az első példa $f(x)$ és $g(x)$ konvolúcióját demonstrálja, melyek a 1a. és 1b. ábrán láthatóak.

Az integrálás elvégzése előtt határozzuk meg a $g(x - \alpha)$ -t. Ezt úgy tehetjük meg, hogy $g(\alpha)$ -t tükrözzük az origóra (1c.ábra), majd x -szel toljuk el (1d.ábra). Majd minden adott x -re, megszorozzuk $f(\alpha)$ -t $g(x - \alpha)$ -val, majd az eredményt integráljuk $-\infty$ -tól ∞ -ig. Az $f(\alpha)$ és $g(x - \alpha)$ szorzatát a besatírozott rész jelzi az 1e. ábrán (az ábra a $0 \leq x \leq 1$ esetre vonatkozik). Mivel a szorzat nulla α azon értékeire, melyek kívül esnek a $[0, x]$ tartományon, ezért azt vehetjük észre, hogy $f(x) * g(x) = \frac{x}{2}$, amely az 1e. ábra besatírozott része.

Az $[1, 2]$ intervallumhoz az 1f. ábra tartozik. Ebben az esetben $f(x) * g(x) = (1 - x/2)$. Ezért, figyelembe véve, hogy $f(\alpha)g(x - \alpha)$ nulla az olyan x -ekre, melyek kívül esnek a $[0, 2]$ intervallumon, a következőt kapjuk:

$$f(x) * g(x) = \begin{cases} \frac{x}{2}, & 0 \leq x < 1, \\ 1 - \frac{x}{2}, & 1 \leq x < 2, \\ 0, & \text{egyébként.} \end{cases}$$

Az eredmény az 1g. ábrán látható.

A konvolúciót szemléletesen a következőképp lehet értelmezni. Legyen az $f(x)$ függvény adott. Ezután vegyük a $g(x)$ függvényt, és toljuk végig a vízszintes tengely mentén. (Ez tulajdonképpen a $g(x - \alpha)$ függvény.) A konvolúció az eltolás függvényében megadja a két függvény közös része alatti terület nagyságát. Esetünkben a terület akkor lesz a legnagyobb, amikor $x = 1$. Ez látszik az 1g. ábrán is, ahol a konvolúciós függvény az $x = 1$ helyen éri el maximumát.

Gyakori az a szemlélet, melyet a későbbiekben fogunk használni, és amely az egyenletben az $f(x)$ függvénynek az impulzus függvénnyel, $\delta(x - x_0)$ -al vett konvolúcióját tekinti. Ez a következőképpen definiálható:

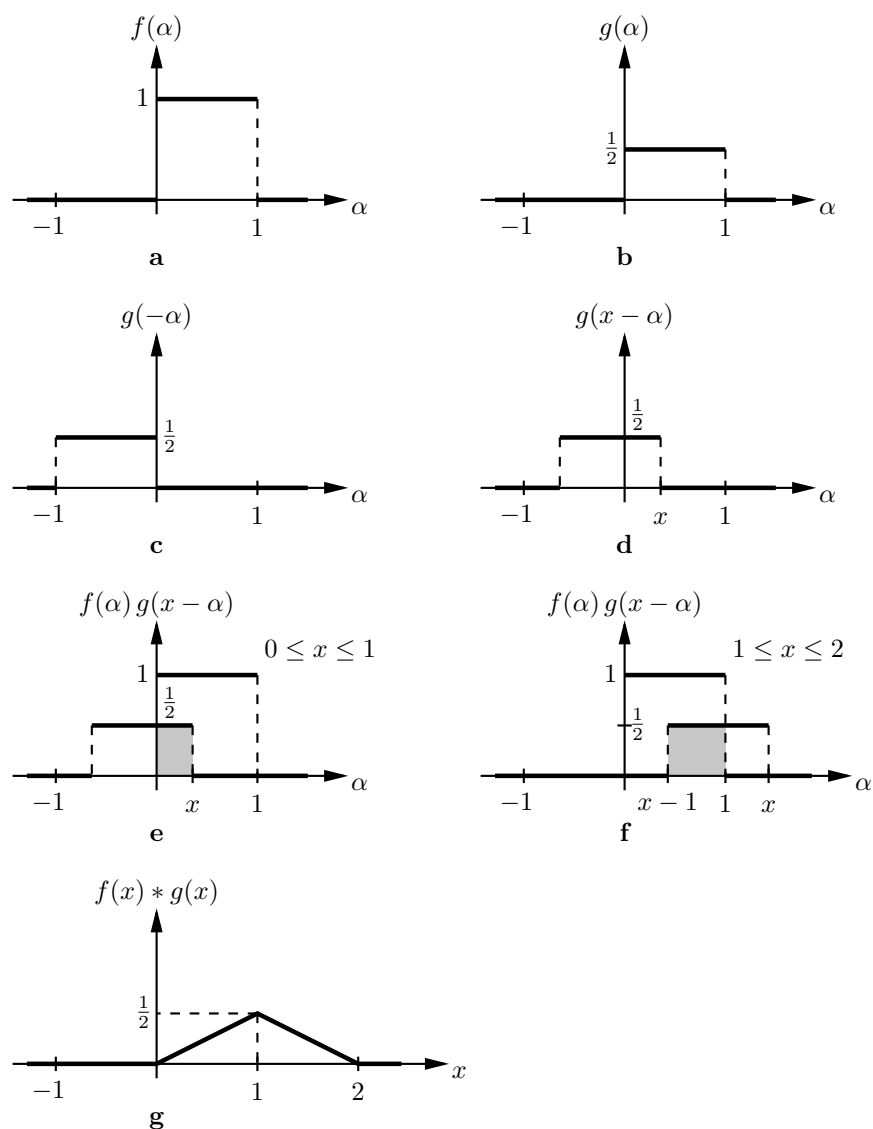
$$\int_{-\infty}^{\infty} f(x) \delta(x - x_0) dx = f(x_0).$$

A függvény x_0 kicsi környezetétől eltekintve mindenhol 0, azaz

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = \int_{x_0^-}^{x_0^+} \delta(x - x_0) dx = 1.$$

Mondhatjuk azt, hogy $\delta(x - x_0)$ az $x = x_0$ pontban van, és az impulzus erejét az $f(x)$ függvény értéke határozza meg az. Például, ha $f(x) = A$, akkor $A\delta(x - x_0)$ egy A erősségű impulzus az $x = x_0$ pontban. Általános gyakorlat, hogy grafikusan az impulzusokat egy x_0 pontban álló nyíllal reprezentálják, melynek nagysága az impulzus erejével egyezik meg. A következő ábrán látható, hogyan néz ki ez $A\delta(x - x_0)$ esetén.

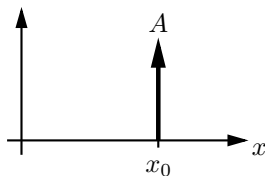
Az egyenlet használatának második illusztrációjaként tegyük fel, hogy az 1a. ábrán látható $f(x)$ függvényt konvolváljuk a $g(x) = \delta(x + T) + \delta(x) + \delta(x - T)$ függvénnyel, mely a 1b. ábrán látható. A $g(x)$ eltolásával, $f(x)$ -en mentén csúsztatva, és a korábbi két egyenlet használatával megkapjuk az eredményt, amely a 1c. ábrán látható. Megfigyelhető, hogy ebben az esetben



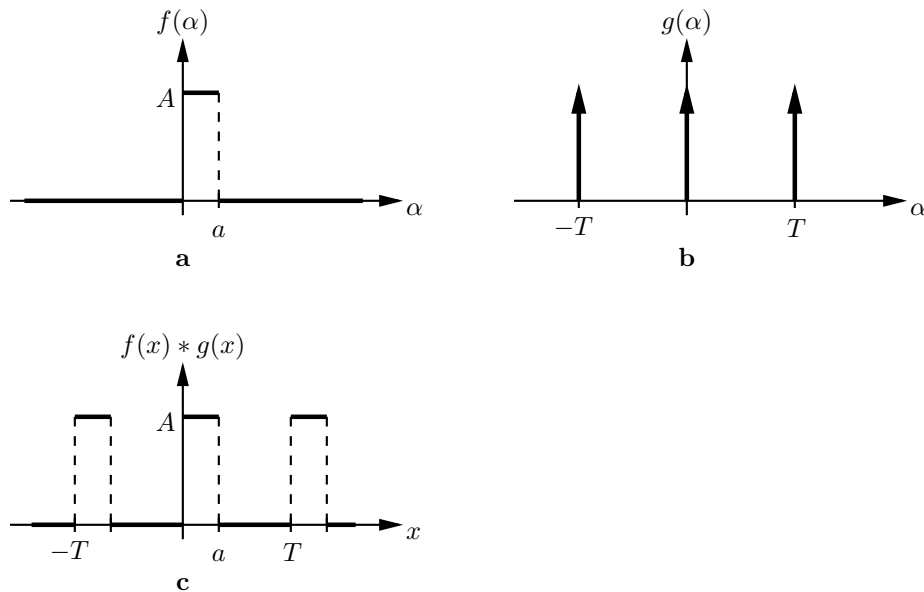
1. ábra. Konvolúció grafikus demonstrálása.

az eredmény megegyezik az $f(x)$ azokba a pozíciókba „másolásával”, ahol az impulzusok találhatók.

A konvolúció fontossága a frekvencia-tartományban végrehajtott elemzéseknél azon a tényen alapul, hogy $f(x)*g(x)$ és $F(u)G(u)$ egy Fourier-transzformációs



2. ábra. Az impulzus függvény illusztrálása.



3. ábra. Impulzus függvény konvolúciója más függvénnyel.

párt alkotnak. Más szavakkal, ha $f(x)$ Fourier-transzformáltja $F(u)$, és $g(x)$ Fourier-transzformáltja $G(u)$, akkor $f(x) * g(x)$ Fourier-transzformáltja $F(u) \cdot G(u)$. Ez az eredmény formálisan a következő formában fogalmazhatjuk meg:

$$f(x) * g(x) \iff F(u) \cdot G(u).$$

Ehhez hasonlóan kapjuk a következő összefüggést:

$$f(x) \cdot g(x) \iff F(u) * G(u).$$

Ezt a két összefüggést konvolúciós tételnek nevezzük.

Tegyük fel, hogy $f(x)$ és $g(x)$ folytonos függvények helyett diszkrétizáltak és a minták rendre A illetve B méretű tömbökben állnak rendelkezésünkre: $f(0), f(1), f(2), \dots, f(A-1)$ és $g(0), g(1), g(2), \dots, g(B-1)$. Ahogy már

korábban is láttuk, a diszkrét Fourier-transzformált és inverze periodikus függvények. Ahhoz, hogy előállítsuk a diszkrét konvolúciós tételt, amely megfelel a periodicitás követelményének, feltételezhetjük, hogy a diszkrét $f(x)$ és $g(x)$ függvények periodikusak, valamilyen M periódussal. Az eredmény szintén periodikus lesz, ugyanazzal a periódussal. A probléma az, hogy milyen értéket válasszunk M -nek. Megmutatható (Brigham [1974]), hogy ha csak nem úgy választjuk M -et, hogy

$$M \geq A + B - 1,$$

akkor a konvolúció egyes periódusai átfedők lesznek. Az átfedést gyakran burkoló hibának (*wraparound error*) nevezik. Ha $M = A + B - 1$, akkor a periódusok szomszédosak, érintkezők lesznek. Ha $M > A + B - 1$, akkor a periódusok különválnak, és a távolságuk egyenlő lesz az M és az $A + B - 1$ közötti különbséggel. Mivel feltettük, hogy a periódus nagyobb A -nál és B -nél is, ezért mindkét mintasorozat hosszát meg kell növelni M hosszúságúra. Ezt úgy is megtehetjük, hogy nullákat fűzünk hozzá az adott mintákhoz, melynek eredményei a következő kiterjesztett sorozatok lesznek:

$$f_e(x) = \begin{cases} f(x), & 0 \leq x \leq A - 1, \\ 0, & A \leq x \leq M - 1, \end{cases}$$

$$g_e(x) = \begin{cases} g(x), & 0 \leq x \leq B - 1, \\ 0, & B \leq x \leq M - 1. \end{cases}$$

Ennek alapján a következő módon definiáljuk az $f_e(x)$ és a $g_e(x)$ diszkrét konvolúcióját

$$f_e(x) * g_e(x) = \sum_{m=0}^{M-1} f_e(m) g_e(x - m),$$

minden $x = 0, 1, 2, \dots, M - 1$ -re. A konvolúciós függvény egy diszkrét, periodikus, M hosszúságú tömb $x = 0, 1, 2, \dots, M - 1$ értékekkel, melyek egy teljes periódusát leírják $f_e(x) * g_e(x)$ -nek.

A diszkrét konvolúció működése alapvetően megegyezik a folytonos konvolúcióval. A különbség csupán annyi, hogy az eltolásoknak diszkrét lépésekben való növelés foglalja el a helyét, a minták közötti távolságnak megfelelően, valamint az integrálás összegzésre cserélődik. Hasonlóan, a következő két egyenlet szintén teljesülnek diszkrét esetben, és a burkoló hiba elkerülése érdekében $f_e(x)$ -et és annak transzformáltját használjuk. Az x és u diszkrét változókról feltesszük, hogy a $0, 1, 2, \dots, M - 1$ értékeket veszik fel.

Az előbbi megfontolások az ábrán grafikusán illusztráltak, folytonos és diszkrét esetben is. A diszkrét esetben a diagramon A mintaelem szerepel a $[0, 1]$ intervallumból, mind $f(x)$ -hez, mind $g(x)$ -hez. A feltételezett periódus és így $M = A + B - 1 = 2A - 1$.

Látható, hogy a konvolúciós függvény periodikus, és mivel $M = 2A - 1$, a periódusok szomszédosak. Ha az $M > 2A - 1$ választással élünk, akkor a periódusok között nagyobb rések lesznek. Azt is fontos megjegyezni, hogy egy periódust M mintaelem teljesen leír.

A kétdimenziós konvolúció az egydimenzióshoz hasonló formájú. Tehát két függvényre, $f(x, y)$ -ra és $g(x, y)$ -ra a következő:

$$f(x, y) * g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) g(x - \alpha, y - \beta) d\alpha d\beta$$

A konvolúciós tétel kétdimenziós esetben a következő formában fogalmazható meg:

$$f(x, y) * g(x, y) \iff F(u, v) G(u, v),$$

$$f(x, y) g(x, y) \iff F(u, v) * G(u, v).$$

A burkoló hiba kiküszöbölését az

$$M \geq A + C - 1,$$

$$N \geq B + D - 1$$

választásával érhetjük el. A periodikus sorozatok az $f(x, y)$ és a $g(x, y)$ kiterjesztésével, az alábbi módon állnak elő:

$$f_e(x, y) = \begin{cases} f(x, y), & 0 \leq x \leq A - 1 \text{ és } 0 \leq y \leq B - 1, \\ 0, & A \leq x \leq M - 1 \text{ vagy } B \leq y \leq N - 1, \end{cases}$$

$$g_e(x, y) = \begin{cases} g(x, y), & 0 \leq x \leq C - 1 \text{ és } 0 \leq y \leq D - 1, \\ 0, & C \leq x \leq M - 1 \text{ vagy } D \leq y \leq N - 1. \end{cases}$$

Az $f_e(x, y)$ és $g_e(x, y)$ kétdimenziós konvolúciója a következő összefüggéssel adott:

$$f_e(x, y) * g_e(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e(m, n) g_e(x - m, y - n),$$

$x = 0, 1, \dots, M - 1$ és $y = 0, 1, 2, \dots, N - 1$.

2. Korreláció

Két folytonos függvénynek, $f(x)$ -nek és $g(x)$ -nek a korrelációját, amelyet $f(x) \circ g(x)$ -szel jelölünk, a következő összefüggéssel definiáljuk:

$$f(x) \circ g(x) = \int_{-\infty}^{\infty} f^*(\alpha) g(x + \alpha) d\alpha,$$

Az eljárás menetét a lenti ábra illusztrálja.

Az egyenlet diszkrét megfelelője a következőképpen definiálható:

$$f_e(x) \circ g_e(x) = \sum_{m=0}^{M-1} f_e^*(m) g_e(x + m),$$

$x = 0, 1, 2, \dots, M-1$. A korábban tett megjegyzések, melyek $f_e(x)$ és $g_e(x)$ periodicitására és M megválasztására vonatkoznak, szintén fennállnak most is.

Hasonló összefüggés érvényes a kétdimenziós esetben is. Ha $f(x, y)$ és $g(x, y)$ folytonos változójú függvények, akkor a korrelációjuk definíciója:

$$f(x, y) \circ g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f^*(\alpha, \beta) g(x + \alpha, y + \beta) d\alpha d\beta.$$

Diszkrét esetben pedig:

$$f_e(x, y) \circ g_e(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_e^*(m, n) g_e(x + m, y + n)$$

$x = 0, 1, 2, \dots, M-1$ és $y = 0, 1, 2, \dots, N-1$. Ahogyan a diszkrét kovolúció esetében is, $f_e(x, y)$ és $g_e(x, y)$ kiterjesztett függvények, M és N pedig a a tétel alapján kerül megválasztásra a korrelációs függvény burkoló hibáinak elkerülése végett.

Megmutatható, hogy mind a folytonos, mind a diszkrét esetre igaz a következő korrelációs tétel:

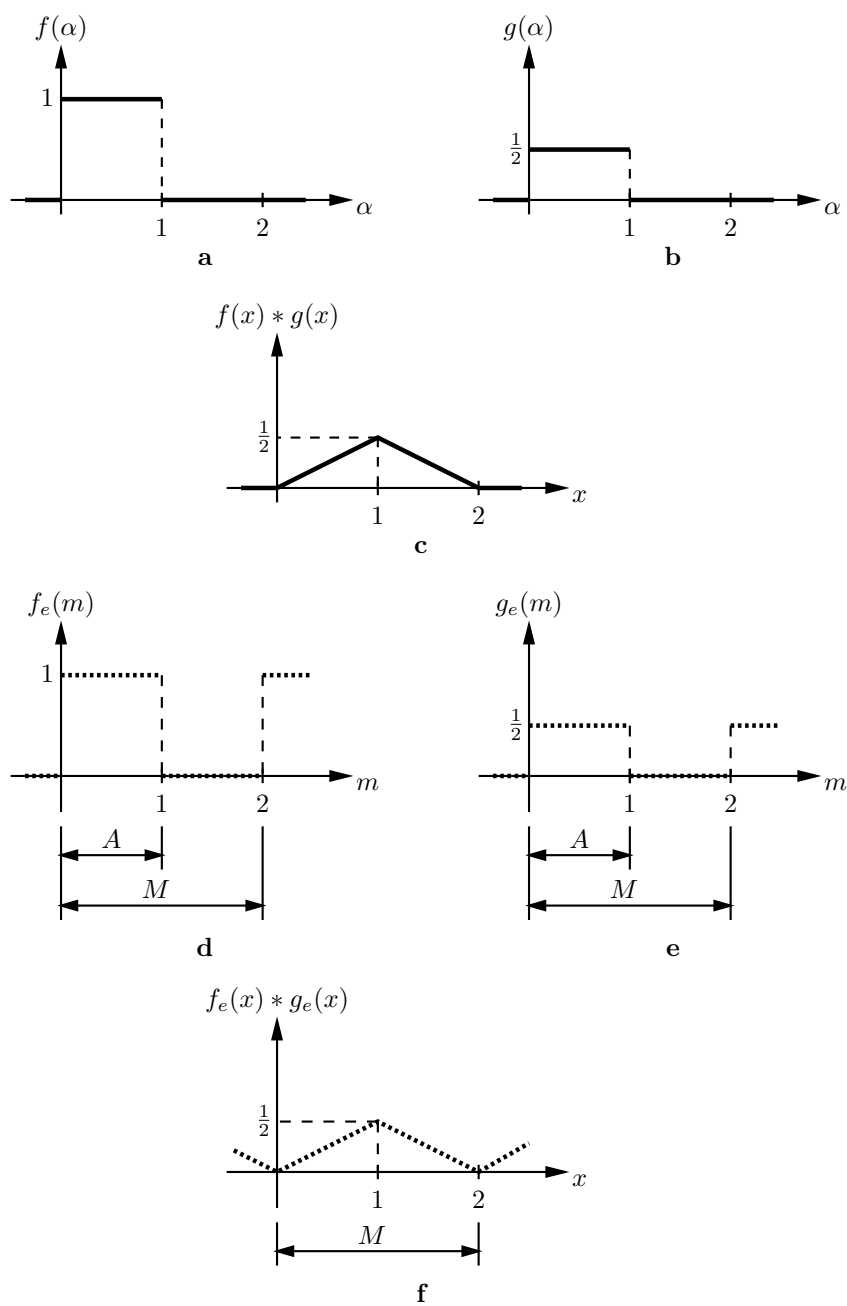
$$\begin{aligned} f(x, y) \circ g(x, y) &\Longleftrightarrow F^*(u, v) G(u, v), \\ f^*(x, y) g(x, y) &\Longleftrightarrow F(u, v) \circ G(u, v). \end{aligned}$$

Természetesen, ha a tételt diszkrét változókra értelmezzük, akkor kiterjesztett és periodikus függvényeket veszünk számításba.

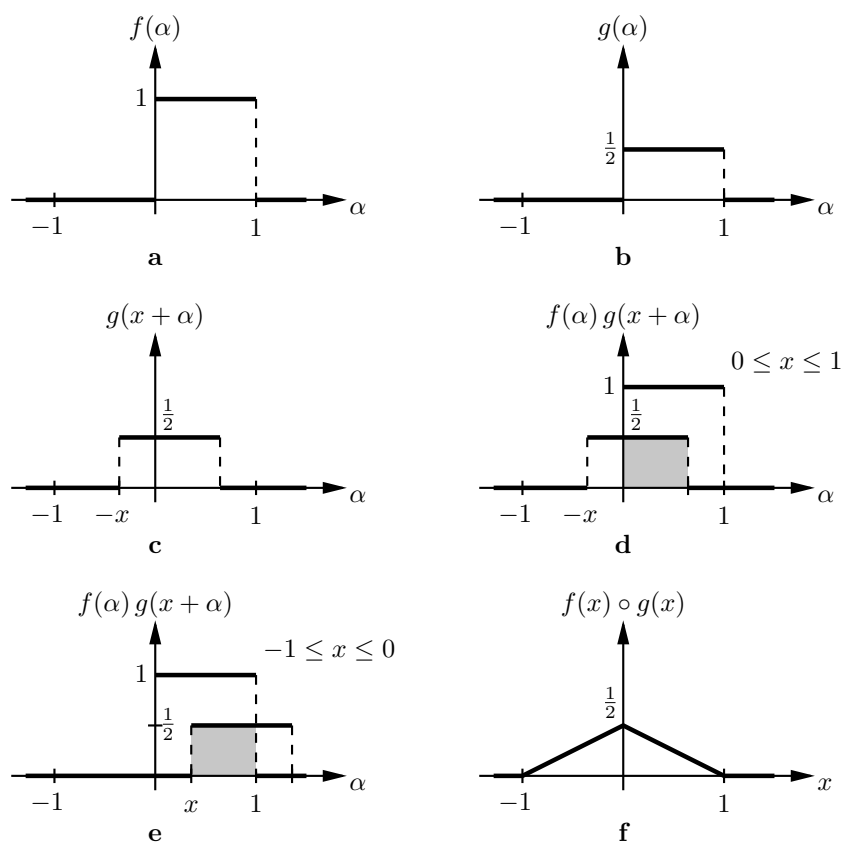
A korreláció egyik legfontosabb alkalmazása a képfeldolgozás területén a sablon vagy a prototípusillesztés, ahol az a feladat, hogy több ismert kép közül megtaláljuk azt az egyet, amelyik a legjobban illeszkedik egy adott,

de ismeretlen képre. Az egyik megközelítés szerint az ismeretlen képnek ki kell számítani a korrelációját az ismert képek mindegyikével. A legjobb illeszkedést az a kép adja, amelyre a korrelációs függvény a legnagyobb. Mivel az eredmény korrelációk kétdimenziós függvények, emiatt a függvények legnagyobb amplitúdóját kell megkeresni. Ahogyan a diszkrét konvolúció esetében is, $f_e(x, y) \circ g_e(x, y)$ kiszámítása sokkal hatékonyabb a frekvenciatarományban, ahol FFT algoritmus segítségével kaphatjuk meg a normál, illetve az inverz transzformáltakat.

Ha összehasonlítjuk a diszkrét és a folytonos korrelációt vagy konvolúciót, meg kell jegyezzük, hogy az a mód, ahogyan a diszkrét eseteket definiáltuk, egyenértékű a folytonos formulák négyzetes integrálásával. Ezért ha szeretnénk a diszkrét és folytonos eseteket összehasonlítani egy abszolút bázis segítségével, akkor a korábbiak alapján az egyik egyenletet meg kell szorozni Δx -szel, míg a másik egyenleteket $\Delta x \Delta y$ -nal, ahol a Δx és a Δy a mintavételezési pontok távolsága.



4. ábra. A periódus alkalmas választása.



5. ábra. Kapcsolat a konvolúció és a korreláció között.

VIII. fejezet

Szegmentálás

Egy digitális kép szegmentálásán a képpontok valamilyen sajátosságvektor(ok) szerinti osztályozását, majd a kapott osztályozásra nézve összefüggő képponthalmazok, azaz tartományok megkeresését értjük.

A szegmentálással foltokhoz, illetve élekhez jutunk, attól függően, hogy a figyelembe vett sajátosságok hasonlósági, illetve különbözőségi jellemzőket mérnek-e. Ideális esetben ezek egybeesnek a képen található objektumok határoló felületeivel, illetve kontúrvonalaikkal. Mindkét feladat megoldására nagyon sok módszer létezik, amelyek sokszor még céljukat tekintve is erősen különböznek. A továbbiakban kizárólag az éldetektálással fogunk foglalkozni.

1. Gradiens módszerek

Arra való tekintettel, hogy az éleket a képen a világosságkód változások jelzik, ezért célszerű a képfüggvény első deriváltját képezni, mivel a derivált képnek az élék helyén lesz szélsőértéke. Erre használhatók a képfüggvény első parciális deriváltjai.

1.1. Digitális gradiens

Digitális képeken a deriváltak helyett természetesen differenciákat használunk:

$$\Delta_x f(i, j) = f(i, j) - f(i - 1, j)$$

$$\Delta_y f(i, j) = f(i, j) - f(i, j - 1)$$

ennélfogva a kép a θ irányba eső „változási gyorsasága”:

$$\Delta f(i, j) = \Delta_x f(i, j) \cos \theta + \Delta_y f(i, j) \sin \theta$$

A digitális gradiens nagysága:

$$a_g(i, j) = ((\Delta_x f(i, j))^2 + (\Delta_y f(i, j))^2)^{\frac{1}{2}}$$

amit gyakran közelítenek az

$$a_g(i, j) \simeq |\Delta_x f(i, j)| + |\Delta_y f(i, j)|$$

kifejezéssel.

1.1.1. *Roberts gradiens.* A Roberts gradiens a digitális gradiens közelítése, amely függőleges és vízszintes éleket detektál. A gradiens nagysága:

$$a_R = |f(i, j) - f(i + 1, j + 1)| + |f(i, j + 1) - f(i + 1, j)|$$

Mivel a gradiens operátorok 4 pontra számolják a gradiens nagyságát, rendkívül érzékenyek a zajokra, illetve az intenzitás egyenletlenségeire.

1.1.2. *Laplace-operátor.* A magasabb rendű deriváltak közül a Laplace-operátort szokták éldetektálásra használni. A másodrendű Laplace-operátor az élek mentén kettős vonalat eredményez, mivel a másodrendű deriváltak ott van szélsőértéke, ahol a világosságkódok változása elkezdődik, és ott ahol befejeződik. A másodrendű Laplace-operátor:

$$\Delta^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2},$$

amelynek digitális közelítése függ az alakzatra meghatározott szomszédságtól is. Tekintsük most a Laplace-operátor 4-szomszédos digitális közelítését:

$$\Delta_4^2(i, j) = |f(i + 1, j) + f(i - 1, j) + f(i, j + 1) + f(i, j - 1) - 4f(i, j)|,$$

illetve a 8-szomszédos közelítését:

$$\begin{aligned} \Delta_8^2 f(i, j) = & |f(i + 1, j) + f(i - 1, j) + f(i, j + 1) + f(i, j - 1) + f(i - 1, j - 1) + \\ & + f(i + 1, j - 1) + f(i - 1, j + 1) + f(i + 1, j + 1) - 8f(i, j)|. \end{aligned}$$

1.2. Maszkos éldetektorok

Az általunk bemutatandó éldetektor diszkrét élmazkokat alkalmaz annak eldöntésére, hogy az adott kép tartalmaz-e él elemet, és ha igen, milyen irányút. Az adott kép domináns él-orientációja:

$$\max\{B \oplus T_i\}$$

ahol T_i a maszk-készlet egyes irányba mutató elemei. Ha a fenti kifejezés által visszaadott érték meghalad egy adott küszöbértéket, úgy az általa meghatározott irányú élt elfogadjuk.

A küszöbérték meghatározásakor két probléma merülhet fel. Az első az élaaktivitás problémája. Ugyanis a gradiens képen lévő élek nagysága igen változó, vannak kis gradiens értékű finom élek, illetve erősebbek. Ez a küszöbölésnél gondot okozhat, mivel a túl alacsony küszöbérték zajossá teszi a képet, a túl magas küszöb esetében pedig hasznos információ vesz

el. Ezt hivatott kiküszöbölni az élektivitási index (EAI). Amennyiben a maszkkészlet 8 irányára kiszámolt gradiens érték rendre $y_0, y_1, \dots, y_7$, úgy az EAI-t a következőképpen definiáljuk:

$$EAI = \frac{\max_{k=0,1,\dots,7}\{|y_k|\}}{(\frac{1}{8} \sum_{k=0}^7 y_k^2)^{\frac{1}{2}} - 1}.$$

Az EAI értéke 0, ha az adott képpontban nincs elfogadható él, ami azt mutatja, hogy ott nincs élektivitás. A másik probléma a küszöbölés globális volta. Ezt a lokálisan adaptív küszöböléssel lehet kiküszöbölni.

IX. fejezet

Osztályozás

Az osztályozási feladat során az alakzatok, vagy alakzatcsoportok hasonlóságát, illetve különbözőségét célszerű valamilyen számszerű mértékkel megadni. Ezt a számszerű mértéket nevezzük hasonlósági, vagy különbözőségi mértéknek. A különbözőségi mértéket általában távolságnak nevezzük. Az alábbiakban bemutatott hasonlósági mértékek elsősorban a statisztikus osztályozásnál használatosak.

1. Korreláció

Gyakran adódhat az a feladat, hogy egy adott $f(x, y)$ $M \times N$ -es, digitális képen egy $g(x, y)$ $M' \times N'$ -méretű részképhez hasonlót keresünk. Általában teljesülnek az alábbi feltételek: $M' < M$ és $N' < N$. A megoldást az $f(x, y)$ és a $g(x, y)$ képek közötti korreláció számítása adja. Ennek legegyszerűbb formája:

$$R(m, n) = \sum_x \sum_y f(x, y)g(x - m, y - n),$$

ahol $m = 0, 1, \dots, M - 1$, $n = 0, 1, \dots, N - 1$, és az összegzést a $g(x, y)$ által definiált részképen kell elvégezni.

Az $R(m, n)$ függvény maximuma azt a helyet mutatja, ahol az $f(x, y)$ képre legjobban illeszkedik a $g(x, y)$. Meg kell jegyeznünk, hogy az $f(x, y)$ kép széleihez közel eső (m, n) értékekre az illeszkedés mértéke kevésbé lesz pontos.

Az előző kifejezésnél valamivel bonyolultabb az alábbi:

$$R(m, n) = \frac{\sum_x \sum_y f(x, y)g(x - m, y - n)}{\sqrt{\sum_x \sum_y f^2(x, y)}}.$$

A nevezőbeli normáló tényezőt szintén a $g(x, y)$ által meghatározott részképen történő összegzéssel kell kiszámítani.

2. Statisztikus osztályozás

Mint láttuk, az osztályozás célja az, hogy a megfigyelt alakzatot véges számú diszjunkt osztály valamelyikébe be tudjuk sorolni. Statisztikus osztályozásról akkor beszélhetünk, ha az alakzatot az $\vec{x}$ valószínűségi vektor változó (alapvektor) írja le, amelynek egy realizációja $\vec{x} = (x_1, x_2, \dots, x_r) \in \Omega_x$, ahol $r = 1, \dots, N$ a jellemzők sorszáma, Ω_x az r -dimenziós mintatér. Jelölje $C_1, C_2, \dots, C_s$ a felismerendő osztályokat, és tegyük fel, hogy a döntést (osztályba sorolást) az alakzatot leíró $\vec{x}$ alapvektor alapján kell megtenni. Az osztályba sorolást az alább ismertetésre kerülő döntési szabályok segítségével végezhetjük el. Az Ω_x mintatérre s darab $D_1, D_2, \dots, D_s$ tartományra osztjuk fel, és $x \in D_i$ esetén elfogadjuk a C_i hipotézist, vagyis azt, hogy x a C_i osztályba tartozik. Ezt a döntési szabályt az alábbi döntésfüggvénnyel is megfogalmazhatjuk: Legyen $\delta(x)$ az Ω_x halmazon értelmezett olyan függvény, amelynek értékkészlete a $\{1, 2, \dots, s\}$ halmaz. A $\delta(x) = i$ esetén elfogadjuk el a C_i hipotézist. Két osztály esetében a döntési függvényt általában a $\delta(x) = \frac{3}{2} + \frac{1}{2}\text{sign}D(x)$ alakban adjuk meg, ahol $D(x)$ olyan valós értékű függvény, amely a D_1 döntési tartományon negatív, a D_2 döntési tartományon pedig nemnegatív értékű. A $D(x)$ függvényt szokás *szeparáló függvénynek* nevezni. A $D(x)$ függvényt úgy kell megválasztani, hogy a $D(x) = 0$ előfordulási valószínűsége 0 legyen. A döntési szabály jóságát a hibás osztályozás valószínűségével mérhetjük, minél kisebb hibás osztályozás valószínűsége, annál jobb a döntési szabályunk. Hibás osztályozás akkor lép fel, ha például a C_i hipotézis igaz ugyan, de $x \in D_i$, vagyis másikat fogadunk el.

X. fejezet

Lényegkiemelés

Az alapvektor geometriai módszerrel történő meghatározására az objektum és annak konvex burkának mérhető paraméterei szolgálnak. Ezek pl. a következők lehetnek: terület kerület, átmérő, vetített magasság, vetített szélesség, hosszúság, orientáció, Feret-féle átmérő, konvex burok ekvivalens területe, konvex burok ekvivalens kerülete, konvex burok átmérője. De ezeken kívül más jellemzőket is meghatározhatunk, pl. az Euler-féle számot, amely az objektum topológikus tulajdonságait tükrözi.

A továbbiakhoz definiáljuk az alábbi 2×2 -es mintákat, amelyeket és $k \cdot 90^\circ$ -os elforgatottjaikat bit-négyeseknek nevezünk:

A bit-négyesekkel igen egyszerű módon meg tudjuk határozni az Euler-számot, megkülönböztetve a négy, illetve nyolc szomszédságra:

$$E_4 = \frac{1}{4} [n(Q_1) - n(Q_3) + 2n(Q_D)],$$

illetve

$$E_8 = \frac{1}{4} [n(Q_1) - n(Q_3) - 2n(Q_D)],$$

ahol $n(Q_i)$ a képen található Q_i bit-négyesek száma.

A területet, illetve a kerületet a bit-négyesek segítségével is kifejezhetjük:

$$T = \frac{1}{4} [n(Q_1) + 2n(Q_2) + 3n(Q_3) + 4n(Q_4) + 2n(Q_D)],$$

$$K = n(Q_1) + n(Q_2) + n(Q_3) + 2n(Q_D).$$

<table><tr><td>0</td><td>0</td></tr><tr><td>0</td><td>0</td></tr></table>	0	0	0	0	<table><tr><td>1</td><td>0</td></tr><tr><td>0</td><td>0</td></tr></table>	1	0	0	0	<table><tr><td>1</td><td>0</td></tr><tr><td>1</td><td>0</td></tr></table>	1	0	1	0	<table><tr><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td></tr></table>	1	0	1	1	<table><tr><td>1</td><td>1</td></tr><tr><td>1</td><td>1</td></tr></table>	1	1	1	1	<table><tr><td>1</td><td>0</td></tr><tr><td>0</td><td>1</td></tr></table>	1	0	0	1
0	0																												
0	0																												
1	0																												
0	0																												
1	0																												
1	0																												
1	0																												
1	1																												
1	1																												
1	1																												
1	0																												
0	1																												
Q_0	Q_1	Q_2	Q_3	Q_4	Q_D																								

1. ábra. Bit-négyesek.

1. Analitikus elemzés

Az egyedi alakzatok analitikus elemzésére két módszert mutatunk be. Az egyik módszer az alakzat kontúrját leíró függvény Fourier-sorát képezi, a másik pedig momentumok segítségével szolgálja az alakzat leírását.

1.1. A kontúr Fourier analízise

A módszer lényege a következő. Képezzük az alakzat kontúrját leíró függvényt. Ez az ún. alakfüggvény (az ívhossz függvényében a görbület) megadását jelenti, amely a kerületre nézve periodikus. A görbület-függvény Fourier-sorát képezve megkapjuk az alakzatot tömören leíró vektort, mivel adott alakzatot már kellőképpen jellemez az első néhány Fourier-komponense. Ábrázoljuk az alakzatot komplex számsíkon úgy, hogy az egyes kontúrpontok abcisszái a valós tengelyre, ordinátái a képzetes tengelyre essenek. Legyen az alakfüggvény $B = B(s)$, amely periodikus, létezik Fourier-sora:

$$B(s) = \sum_{i=-\infty}^{\infty} F_f e^{fis},$$

ahol

$$F_f = \frac{1}{2\pi} \int_0^{2\pi} B(s) e^{-fis} ds.$$

A bináris képen az alakzatfüggvény szakaszonként lineáris, tehát egy N szögű sokszögnek tekinthető. Ennek köszönhetően az egyes Fourier-együtthatók a levezetést mellőzve:

$$F_f = \frac{N}{(2\pi f)^2} \sum_{k=1}^N (p_{k-1} - p_k) e^{-fik \frac{2\pi}{N}}$$

Válasszuk az $F_0 = 0$ -t, ami azt jelenti, hogy a Fourier együtthatók alapján visszaállított alakzat súlypontját az origóba toljuk. A p_i -k az egyes csúcsok koordinátái.

1.2. Alakjellemzés momentumok segítségével

Az alakzatokat nemcsak kontúrpontjaival, hanem belső pontjai alapján is leírhatjuk. Ezt a momentumok segítségével tehetjük meg.

Az $f(x, y)$ 2-dimenziós folytonos függvény (p, q) -adik momentuma a valószínűség elméletéből ismert definíció alapján:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

A centrális momentum:

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy$$

ahol $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$.

A centrális momentum diszkrét esetben:

$$\mu_{pq} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (x - \bar{x})^p (y - \bar{y})^q f(x, y).$$

A momentumokból számos olyan mérték állítható elő, melyek kielégítően jellemzik az objektumot, és az objektumon végrehajtott lineáris transzformációkkal szemben invariánsak.

XI. fejezet

Irodalomjegyzék

1. Azriel Rosenfeld and Avinash C. Kak: *Digital Picture Processing I-II.*, Academic Press, 1982.
2. Ghymes Balázs: *Digitális képfeldolgozó algoritmusok I-IV.*, Kézirat, 1982.
3. Álló Géza, Hegedűs Gy. Csaba, Kelemen Dezső és Szabó József: *A digitális képfeldolgozás alapproblémái*, Akadémiai Kiadó, 1989. Tankönyvkiadó, Budapest, 1988.
4. Piero Zamperoni: *Methoden der digitalen Bildsignalverarbeitung*, Friedr. Vieweg & Sohn, 1989.
5. Bernd Jähne, Horst Haußecker, and Peter Geißler, *Handbook of Computer Vision and Applications*, Academic Press, 1999.
6. Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, John Wiley & Sons, 2001.
7. Sankar K. Pal and Amita Pal, *Pattern Recognition. From Classical to Modern Approaches*, World Scientific, 2001.