

# Nagy pontosságú képfeldolgozás

Csirmaz László  
Közép Európai Egyetem, Budapest

2004 január

## Absztrakt

Egy szürke skálájú képen mintegy háromszáz, nagyjából kör alakú, 10 és 30 pixel közötti átmérőjű folt középpontját és sugarát kell lehető legnagyobb pontossággal meghatározni. A feladat nehézségét az jelenti, hogy ellentétben a képfeldolgozásban szokásos pixeles pontossággal, ennél egy nagyságrenddel nagyobbra van szükség. A megoldásban különböző paraméterű szűrőkkel készített konvolúciók geometriai tulajdonságait használjuk fel; nevezetesen az illeszkedés jóságát a felület lokális maximumában mért görbülete adja meg. A cikkben az újfajta megközelítés elemeit és azok matematikai hátterét ismertetjük.

## 1. Bevezetés

Egy tipikus képfeldolgozási feladatban egy képen megadott, vagy a megadotthoz hasonló alakzatot kell keresni. Gyakran elegendő csak azt eldönteni, hogy a minta egyáltalán megtalálható-e a képen. A legtöbb ismert eljárás a minta összes előfordulását pár pixeles pontossággal meg is adja. Esetünkben képenként több mint háromszáz, nagyjából kör alakú folt sugarát és helyét kellett pixelnél nagyobb pontossággal megállapítanunk. A feladatot a következő módszer szerint kívánjuk megoldani. Elsőként definiáljuk mintáknak egy egyparaméteres sokaságát: minden szóbjövő  $r$  sugárra egy mintát, ami leírja azt, hogyan néz ki egy tipikus  $r$  sugarú folt. Ezek után egy gyors, heurisztikus módszerrel megkeressük az összes folt középpontját és sugarát pár pixel pontossággal. Minden sugárra megkeressük, hogy ehhez a sugárhoz tartozó minta hol illeszkedik a legjobban a becsült középpont közelében, majd kiválasztjuk azt a sugarat és a hozzátartozó középpontot, ahol ez az illeszkedés a lehető legjobb. A továbbiakban ennek megvalósítása során felmerülő problémákat ismertetjük azzal együtt, hogyan oldottuk meg illetve hogyan kerültük meg a problémát.

Foltok helyének közelítő meghatározása nem igényel a képfeldolgozásban ismert és használt módszereken túlmutató ötletet, így annak ismertetésétől eltekintünk. A 2. részben foglaljuk össze mindazt a szükséges előismeretet, amit a szűrőkről és a konvolúcióról – a képfeldolgozás alapvető módszereiről – fel fogunk használni. A 3. rész taglalja, hogyan kell megválasztani a különböző sugarakhoz tartozó mintákat, továbbá hogyan kereshetjük meg a foltok középpontját pixel pontossággal. A középpontokat pixelnél nagyobb pontossággal a 4. részben határozzuk meg. Az 5. rész az  $r$  sugár

meghatározásáról szól. Míg fix sugár mellett a legjobb illeszkedés helyét a konvolúciós felület lokális maximumai szolgáltatják, addig különböző sugarak mellett a lokális maximum értéke nem tükrözi vissza, hogy mennyire jó az illeszkedés. Ennek mérésére a konvolúciós felület másik jellemzőjét, a maximumban mért *görbületet* használjuk. Kiderül azonban, hogy a lokális maximum helyét megfelelő pontossággal becsülő eljárás túl nagy hibát vét a görbület esetében. A 6. részben megmutatjuk, hogyan lehet más módszerrel, megfelelő pontossággal a görbületet is számítani. Az utolsó 7. részben összefoglaljuk a feladat megoldását, és néhány nyitva maradt kérdést vetünk fel.

## 2. Szűrők

A képfeldolgozás során a leggyakrabban használt módszer *szűrők* alkalmazása [2, 6]. A képet egy  $p(x, y)$  kétváltozós valós síkon értelmezett függvénynek tekintjük; a függvény értéke a kép  $(x, y)$  pontbeli színe. Mivel fekete-fehér képről van szó, a szín egy 0 és 1 közötti valós szám, ahol 0 jelenti a feketét, 1 pedig a fehéret. A  $p(x, y)$  függvényről feltesszük, hogy a sík megfelelően nagy tartományán (vagy akár a teljes síkon) értelmezve van, és mindazon feltételeknek eleget tesz, amik biztosítják hogy a felhasznált tételek igazak legyenek (például  $p$  akárhányszor deriválható, vagy négyzetesen integrálható, stb). A feldolgozandó képen a  $p$  függvénynek az egész koordinátájú rácspontok egy téglalap alakú részén felvett értékeit találjuk valamekkora hibával. A hiba természetével és a kép minőségének javításával (zajszűrés, kontrasztjavítás) nem foglalkozunk.

Egy *szűrő* mindenütt értelmezett, sima, négyzetesen integrálható  $f$  függvény, ami rendszerint, de nem feltétlenül még centrálszimmetrikus is (vagyis  $f(x, y) = f(-x, -y)$  minden  $x, y$  valós számpárra). Az  $f$ -et eltoljuk az  $(u, v)$  pontba, majd az eltolt és tükrözött  $f$ -fel mint súllyal átlagoljuk a  $p$  képet:

$$P(u, v) = \int_{\mathbf{R}^2} p(x, y) f(u - x, v - y) dx dy. \quad (1)$$

Ez az úgy nevezett *konvolúciós integrál* a  $p(x, y)$  kétváltozós függvényhez a  $P(u, v)$  kétváltozós függvényt rendeli; szokásos még a  $P = \langle p, f \rangle$  jelölés is. Könnyű látni, hogy  $\langle p, f \rangle$  szimmetrikus:  $\langle p, f \rangle = \langle f, p \rangle$ , valamint mindkét argumentumában lineáris, például  $\langle p, c_1 f_1 + c_2 f_2 \rangle = c_1 \langle p, f_1 \rangle + c_2 \langle p, f_2 \rangle$ . Az  $f$  szűrő a konvolúció *magfüggvénye*.

A  $\sigma$  szórásiú (kétdimenziós) *Gauss-szűrő* az alábbi függvény:

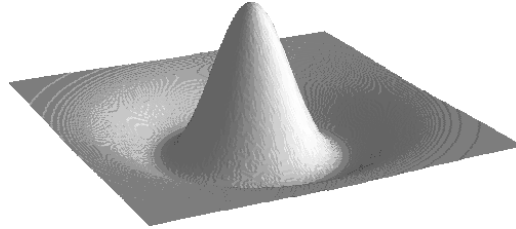
$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Szokás szerint a konstans úgy választottuk meg, hogy a szűrő teljes síkon vett integrálja (vagyis a szűrő  $L_1$  normája) éppen 1 legyen. Előszeretettel használnak kis szórásiú Gauss-szűrőt a kép hibáinak csökkentésére; nagyobb szórásiú szűrővel a képet "lágýtani" lehet, vagyis az éleket elmosni (blurring). Még nagyobb szórásiú Gauss-szűrővel vett konvolúció értéke jól méri egy képpont környezetének átlagos szürkességét. Például az élesítésnek (sharpening) hívott kontrasztjavító eljárásnál a kép  $(u, v)$

pontbeli értéket úgy módosítjuk, hogy az ottani átlagtól való eltérés  $c > 1$ -szeresére nőjön. Az átlagot egy  $\sigma_1$  szórású Gauss-szűrővel való konvolúció adja meg, az  $(u, v)$  pontbeli értéket pedig egy  $\sigma_2$  szórású Gauss-szűrő (persze  $\sigma_1$  nagyobb mint  $\sigma_2$ ). A kontrasztjavító eljárás eredménye az  $(u, v)$  pontban:

$$\langle p, G_{\sigma_1} \rangle + c \cdot (\langle p, G_{\sigma_2} \rangle - \langle p, G_{\sigma_1} \rangle) = \langle p, (1 - c)G_{\sigma_1} + cG_{\sigma_2} \rangle,$$

felhasználva a konvolúció linearitását. A javítás tehát egyetlen konvolúcióval is számolható. A két Gauss-szűrő különbsége úgy néz ki, mint egy mexikói kalap (1. ábra).



1. ábra: Két Gauss-szűrő különbsége

Gyakorlatban a számítások gyorsítása érdekében a szűrőket egészen kis méretű, egész értékű mátrixszal közelítik. A kontraszt növelésére például az alábbi  $3 \times 3$ -as mátrixot szokás használni különböző  $c$  konstansokkal:

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & c & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Visszatérve a  $\langle p, f \rangle$  konvolúcióra, ennek  $(u, v)$ -beli értékét úgy is tekinthetjük, mint annak a mértékét, hogy ebben a pontban  $p$  mennyire hasonlít az  $f$  szűrőhöz. Nézzük ugyanis  $p$  valamint  $f$  eltolt (és tükrözött) képe közötti különbség négyzetes integrálját a teljes síkon:

$$\int (p(x, y) - f(u - x, v - y))^2 dx dy = \int p^2(x, y) dx dy + \int f^2(x, y) dx dy - 2\langle p, f \rangle.$$

A jobb oldalon az első két integrál értéke független az  $(u, v)$  ponttól, tehát a négyzetes eltérés a konvolúció  $(-2)$ -szeresétől csak egy additív konstansban tér el. Minél nagyobb tehát a konvolúció, annál jobban hasonlít a  $p$  az  $f$  szűrőfüggvényre. A konvolúciónak ezt a tulajdonságát kihasználva tudunk egy képen adott mintához hasonló alakzatokat keresni. A mintát felvesszük szűrőnek; ahol a konvolúciónak (lokális) maximuma van és értéke meghalad egy bizonyos küszöböt, ott várható a minta megjelenése. Természetesen az eljárás csak az adott minta eltoltjait tudja csak megtalálni. Szerencsére esetünkben a keresendő alakzatok forgásszimmetrikusak, így nem kellett a minta elforgatásából adódó problémával küszködnünk.

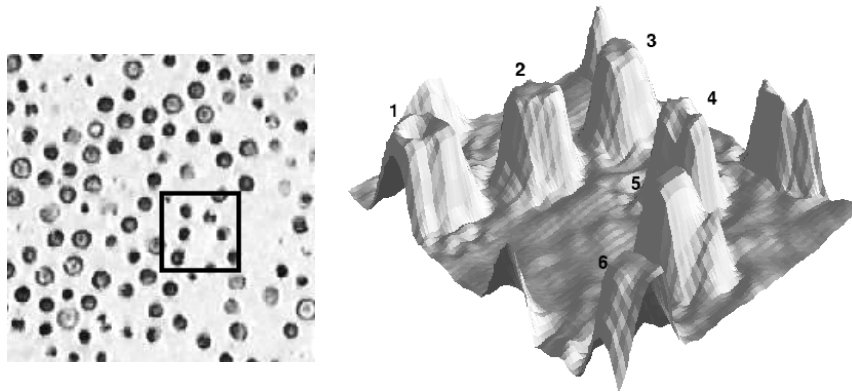
Gyakorlatban a konvolúció kiszámítása numerikus integrálást jelent. A  $p$  függvény értékét csak egész helyeken ismerjük (és ott is csak valamilyen hibával), ezért az integrált az integrálandó függvény rácspontokon felvett értékeinek összegeként közelítjük:

$$P(u, v) \approx \tilde{P}(u, v) = \sum_{i, j \in \mathbf{Z}^2} p(i, j) f(u - i, v - j).$$

Az  $f$  szűrő általában mindenütt (és nem csak a rácspontokon) van definiálva, tehát  $\tilde{P}$  értékét tetszőleges  $(u, v)$  pontban, és nem csak a rácspontokon tudjuk számítani.

### 3. Lokális maximum keresése pixel pontossággal

Egy feldolgozandó kép tipikus részletét láthatjuk az 2. ábra bal oldalán. A jobb oldalon a kiemelt rész háromdimenziós képe látható; a foltokat a későbbi utalások kedvéért megszámoztuk.



2. ábra: A feldolgozandó kép és részének 3D képe

A keresendő alakzatok nem homogén foltok, hanem gyakran gyűrű alakúak, melyeket egy nagyon keskeny, a háttérnél világosabb csík vesz körül. Az egyes számú (bal alsó sarokban található) folt közepén is megfigyelhető egy kráter. A foltok hozzávetőleges helyét és sugarát egy gyors, heurisztikus algoritmussal megkeressük. Ezután veszünk egy  $r$  sugárhoz tartozó  $f_r$  szűrőt, és a középpont becsléstől indulva megkeressük azt a legközelebbi  $(u, v)$  rácspontot, ahol a

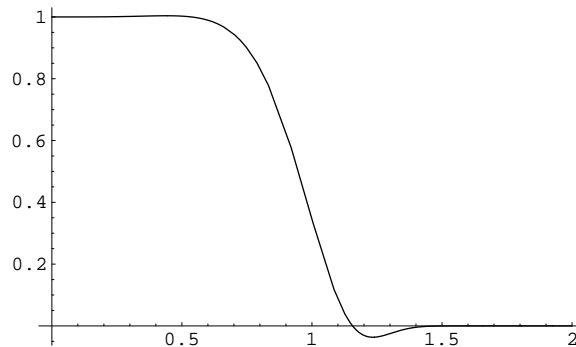
$$\tilde{P}_r(u, v) = \sum_{i, j} p(i, j) f_r(u - i, v - j)$$

közelítő összegnek lokális maximuma van. Amennyiben mind a sugarat, mind a folt helyét megfelelően jól becsültük meg, a lokális maximum helye jó közelítést ad a középpontra.

Mint az előző részben láttuk, az ideális  $f_r$  szűrő megegyezik az  $r$  sugarú folt alakjával. A foltok valamilyen fizikailag létező objektumok képei, túl nagy változatosságot

mutatnak, és nincs is rá esély, hogy valamilyen szép függvénnyel le tudjuk őket írni. Ezért olyan  $f_r$  szűrőt használunk, amely egyrészt analitikusan kezelhető, másrészt jól használható a folt középpontjának és sugarának meghatározására. Ez azt jelenti, hogy a szűrő viszonylag nagy és egyenletes meredekségű a folt szélének környékén – vagyis az origótól  $r$  távolságban –, attól távolabb és az origó környezetében pedig lényegében vízszintes kell legyen. Minél meredekebb a szűrő, annál jobban kiemelkedik az illeszkedés helye. Ugyanakkor túl meredek szűrő már túlságosan érzékeny a pontos  $r$  értékre és a folt szabályos alakjára: kicsit eltérő sugár vagy bizonytalan alakú folt esetén a lokális maximum nagyon eltolódhat az igazi középponthez képest.

További problémát okoz, hogy nem egyetlen szűrőt kell megadnunk, hanem minden szóbjövő  $r$  sugárra egyet-egyét. Különböző szűrők által adott eredményeket kell összevetnünk, ami azt jelenti, hogy a szűrőket megfelelően kell normálni. De milyen normát válasszunk? A kép átlagos szürkeségét azok a szűrők tartják meg, melyek  $L_1$  normája (a teljes síkon vett integrálja) éppen 1. Ha viszont a konvolúciót úgy tekintjük, mint ami a kép és a szűrő közötti négyzetes eltérést méri, akkor a szűrők  $L_2$  normáját kellene egyenlővé tennünk. Persze esetünkben nem szükségképpen a négyzetesen legjobban közelítő szűrő illeszkedik legjobban. Ennek oka a folt alakjában mutatkozó bizonytalanság és a folt közepén megjelenő egyenletlenség. A norma megfelelő megválasztásának kérdését megkerülhetjük, ha az  $f_r$  szűrőt  $f_1$  megfelelően felnagyított képének választjuk:  $f_r(u, v) = f_1(u/r, v/r)$ . Ekkor ugyanis  $f_r$  tetszőleges normája  $f_1$  megfelelő normájának  $r^2$ -szerese. Ennek a választásnak hátránya viszont, hogy  $f_r$ -nek sugár távolságban mért meredeksége lineárisan függ  $r$ -től: minél nagyobb a sugár, a szűrő annál kevésbé érzékeny kis elmozdulásokra.



3. ábra: A szűrőt generáló  $(1 - x^8/4)e^{-x^4}$  függvény gráfja

A keresett foltok körszimmetrikusak, tehát szűrőink is azok lesznek. Az előző érvelésnek megfelelően  $f_r$ -et úgy választjuk, hogy egy valós egyváltozós  $\varphi(x)$  függvényt  $r$ -szeresére nyújtva megforgatunk az  $y$  tengely körül:

$$f_r(u, v) = \varphi\left(\frac{\sqrt{u^2 + v^2}}{r}\right).$$

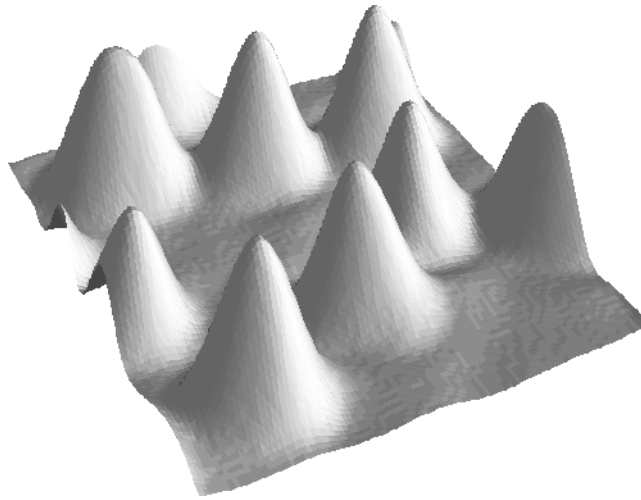
$\varphi$ -nek olyan függvényt kell választani, mely a 0 környékén nagyjából konstans, az 1-

hez közeledve meredeken csökken, majd nem sokkal 1 fölött gyorsan belesimul az  $x$  tengelybe. Az általunk választott

$$\varphi(x) = \left(1 - \frac{x^8}{4}\right)e^{-x^4}$$

függvény  $x \approx 0,7$ -ig vízszintesen halad, onnan elindul lefelé.  $x = \sqrt[4]{2} \approx 1,18$  körül átmetszi az  $x$  tengelyt, 1,5 fölött pedig már elenyészően kicsi (3. ábra). A kis negatív tartomány a folt körüli keskeny világosabb sávnak felel meg. A további elemzéseket erre függvényre tesszük; következtetéseink azonban más szűrőkre is igazak.

Alkalmazva az  $f_r$  szűrőt az  $r = 7,5$  értékkel, a kapott felület háromdimenziós képét a 4. ábra mutatja. Jól láthatók a markáns süvegek, melyek csúcsa felel meg a konvolúció lokális maximumainak, vagyis a csúcsok koordinátái adják meg a foltok középpontjait. A szűrő eltüntette az eredeti képen a foltok közepén található bemélyedéseket, a folt egyenetlenségeit és bizonytalanságokat.



4. ábra: A kép  $r = 7,5$  paraméterű szűrő alkalmazása után

A  $P = \langle p, f_r \rangle$  konvolúció lokális maximumait – vagyis a süvegek csúcsait – pixel pontossággal egyszerűen meghatározhatjuk. A maximumhoz elegendően közelről indulva rácspontokon lépkedünk. Minden lépésben a rácspontnak olyan szomszédjába megyünk, ahol a  $P$  konvolúció nagyobb értéket vesz fel. Amikor megakadunk, megtaláltuk a lokális maximumot. Az algoritmus gyorsítható ha például minden lépésben először abban az irányban próbálkozunk, amit az előző lépésben használtunk. Robosztusabbá is tehető az algoritmus, ha lépésenként az összes közvetlen (vagy másod-, harmad-) szomszéd közül választjuk ki a maximálisat.

## 4. Nagyobb pontosság felület illesztésével

A  $P_r = \langle p, f_r \rangle$  konvolúció lokális maximumainak helyét pixelnél nagyobb pontossággal is megbecsülhetjük. Tegyük fel, hogy az  $(u_0, v_0)$  rácspontban  $P_r$  értéke nagyobb, mint a környező rácspontokban. A  $P_r(u_0 + x, v_0 + y)$  függvényt egy

$$G(x, y) = A + Bx + Cy + Dxy + \frac{1}{2}Ex^2 + \frac{1}{2}Fy^2 \quad (2)$$

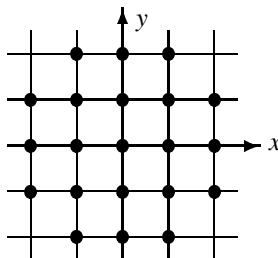
alakú másodfokú felülettel közelítjük, és  $P_r$  lokális maximumát  $(u_0 + x_0, v_0 + y_0)$ -al becsüljük, ahol  $(x_0, y_0)$  az a hely, ahol  $G(x, y)$  a lehető legnagyobb.

Szokás szerint  $G$  együtthatóit úgy határozzuk meg, hogy  $P_r(u_0 + x, v_0 + y)$  és  $G(x, y)$  eltéréseinek négyzetösszege az origó körüli néhány rácspontban a lehető legkisebb legyen. Ha  $\mathcal{A}$  ezen rácspontok halmaza, akkor a minimalizálandó kifejezés

$$\sum_{(i,j) \in \mathcal{A}} (G(i, j) - P_r(i + u_0, j + v_0))^2.$$

Ennek az  $A, B, C, D, E$  és  $F$  együtthatók szerinti parciális deriváltjai el kell tűnjenek. Ez hat lineáris egyenlet az ismeretlen együtthatókra, amiből azok meghatározhatók. Amennyiben az  $\mathcal{A}$  halmaz szimmetrikus mind a két tengelyre, az egyenletrendszer harminchat együtthatójából 24 nulla lesz, és az egyenletrendszer akár kézzel is könnyen megoldható.

Az  $\mathcal{A}$ -ra jó választás például az a 21 elemű halmaz, ami az origó körüli 5-ször 5-ös rácsnégyzet pontjaiból áll annak négy sarkát kivéve:



A  $G(x, y)$  együtthatóinak meghatározása után megkeressük  $G$  maximumát. Ezt abban az  $(x_0, y_0)$  pontban veszi fel, ahol az  $G$ -nek  $x$  és  $y$  szerinti parciális deriváltja egyaránt eltűnik:

$$\frac{\partial G}{\partial x} = B + Dy + Ex = 0, \quad \frac{\partial G}{\partial y} = C + Dx + Fy = 0.$$

Az egyenletrendszer megoldása

$$x_0 = \frac{CD - BF}{EF - D^2}, \quad y_0 = \frac{BD - CE}{EF - D^2}. \quad (3)$$

A nevezőben álló  $d = EF - D^2$  érték a  $G$  másodfokú felület fontos jellemzője. Ennek előjele azonosítja a felület típusát a (3) szerinti  $(x_0, y_0)$  pontban. Ha  $d$  pozitív,  $G$ -nek az

$(x_0, y_0)$  pontban vagy abszolút maximuma vagy abszolút minimuma van, attól függően, hogy  $E$  (és vele együtt  $F$ ) negatív vagy pedig pozitív. Ha  $d$  nulla vagy negatív, akkor  $G$ -nek egyáltalán nincs lokális szélsőértéke, és így speciálisan  $(x_0, y_0)$  sem lehet az.

Mindezeket összerakva a  $P_r$  konvolúció lokális maximumának meghatározására az alábbi algoritmust kapjuk. Az  $(u_0, v_0)$  rácspont meghatározása után az  $\mathcal{A}$  rácspontokban négyzetesen legjobban illeszkedő  $G$  másodfokú felület (2) szerinti együtthatóit kiszámítjuk. Ha a felület  $d = EF - D^2$  diszkriminánsa nulla vagy negatív, az illesztett felületnek  $(u_0, v_0)$  környékén nincs szélsőértéke, ami az jelenti, hogy a  $P_r$  konvolúciónak sincs itt szignifikáns maximuma. Ha  $d$  pozitív, a (3) alapján számított  $(x_0, y_0)$  pontban van  $G$ -nek szélsőértéke. Még érdemes ellenőrizni, hogy  $x_0$  és  $y_0$  abszolút értékben kisebb 1-nél (tulajdonképpen az 1-hez vagy  $-1$ -hez közeli érték sem igazán elfogadható). Ha így van, az  $(u_0 + x_0, v_0 + y_0)$  pont a  $P_r$  lokális maximumának jó közelítése.

## 5. A sugár meghatározása – Gauss görbület

Amennyiben ismerjük az  $r$  sugarat, az előző pontban ismertetett eljárással a folt közép-pontját már megfelelő pontossággal meg tudjuk határozni. Feladatunk tehát  $r$  meghatározása.

Mint láttuk, a  $P_r$  konvolúció az  $f_r$  szűrő és a kép négyzetes eltérését méri. Természetesen adódik, hogy azt az  $r$ -et fogadjuk el a sugár értékének becslésére, amire  $P_r$ -nek előző pontban leírt algoritmus szerint számított lokális maximuma a lehető legnagyobb. A tapasztalat szerint azonban ez nem működik: a lokális maximumban felvett érték  $r$ -ben nagyon enyén, de monoton csökken. Ennek oka elsősorban az, hogy a foltok alakja  $r$ -rel nem lineárisan változik, míg az  $f_r$  szűrőket  $f_1$ -ből lineáris nagyítással kapjuk. Így az optimális  $r$  sugár esetén  $f_r$  már nem feltétlenül illeszkedik négyzetesen a lehető legjobban.

Megoldásként egyik lehetőség az, hogy az  $f_r$  szűrőket másképp definiáljuk. Például nagyszámú foltról statisztikát készítve próbáljuk meg azok tipikus alakját meghatározni – ehhez viszont a statisztikában részt vevő foltok pontos sugarát tudnunk kell. Ha ezt a triviálisnak egyáltalán nem tűnő problémát megoldottuk (például klaszterezéssel csoportosítjuk a nagyjából egyforma sugarú foltokat), újabb problémával kerülünk szembe: a szűrőket normálni kell. Tökéletes alakú szűrő és hiba nélküli  $p(x, y)$  értékek esetén persze a szűrők  $L_2$  normáját kell egyenlővé tenni. A mérési hibák hatását az  $L_1$  normán keresztül tudjuk kontrollálni. Mivel az  $L_1$  és  $L_2$  norma teljesen más, a normálást ismét az adatok alapján kell beállítani.

Másik lehetőségünk az, ha az illeszkedés jóságát másképpen mérjük. Ehhez vegyük jobban szemügyre a (2) felület  $(x_0, y_0)$  lokális szélsőértékét. Hogy ez a szélsőérték mennyire szignifikáns, a pozitív  $d = EF - D^2$  érték mutatja meg. Ha  $d$  kicsi, akkor  $G$  maximumát egy lapos platón alig kiemelkedve veszi fel. Minél nagyobb a  $d$ , a felület annál meredekebb, hegyesebb  $(x_0, y_0)$ -ban. Ez nem véletlen, hiszen  $d$  éppen  $G$ -nek az  $(x_0, y_0)$ -beli Gauss-görbülete, lásd [3, 4].

Húzzunk érintő síkot a  $G$  felület egy  $(x, y)$  pontjában, és az egyszerűség kedvéért tegyük fel, hogy a felület az érintési pont egy környezetében az érintő síknak ugyanarra az oldalára esik. (Ez a helyzet például, ha  $(x, y)$  lokális maximum.) Ha most az érintő síkot önmagával párhuzamosan  $\varepsilon$  távolságra eltoljuk a felület felé, akkor a felületet és a



sík metszészvonala közel ellipszis alakú lesz. A metszet ellipszis kis- és nagytengelyének félhosszát emeljük négyzetre, a négyzeteket szorozzuk össze, majd a szorzatot normaljuk le. Mivel mindkét tengely hossza  $\varepsilon^{1/2}$  nagyságrendű (érintő síkon vagyunk), a normálás  $4\varepsilon^2$ -nel való osztást jelent. (A mágikus 4-es szorzó megjelenésének okát lásd alább.) A Gauss – vagy másképpen szorzat – görbület a normált szorzat reciprokának határértéke amint  $\varepsilon$  tart a nullához.

A Gauss görbületet másképpen is szokás definiálni. A felület  $(x, y)$  pontjában az érintő síkra merőleges egyenest emelünk. Ezen a normálison át fektett összes síkon meghatározzuk a felület és a sík metszetének a görbületét. A görbületek közül a minimális és a maximális két egymásra merőleges sík esetén fordul elő. A szorzatgörbület ennek a két szélsőértéknek a szorzata. Az *összeg* vagy *Minkowski* görbület pedig a minimális és maximális görbület összege. Az, hogy a Gauss görbület két definíciója ekvivalens abból következik, hogy limeszben a fenti ellipszisek kis- és nagytengelyei éppen a maximális illetve minimális görbületet adó síkokban vannak, továbbá ezek a görbületek annak a hányadosnak a határértéke, mikor  $2\varepsilon$ -t a megfelelő féltengely hosszának négyzetével osztjuk. (Innen adódik a fenti  $2\varepsilon \cdot 2\varepsilon = 4\varepsilon^2$  normáló tényező.) Ha az  $F(x, y)$  függvény által definiált felület egy pontjában az  $x$  és  $y$  szerinti parciális derivált is eltűnik, akkor abban a pontban a Gauss-görbület

$$\frac{\partial^2 F}{\partial x^2} \cdot \frac{\partial^2 F}{\partial y^2} - \left( \frac{\partial^2 F}{\partial x \partial y} \right)^2, \quad (4)$$

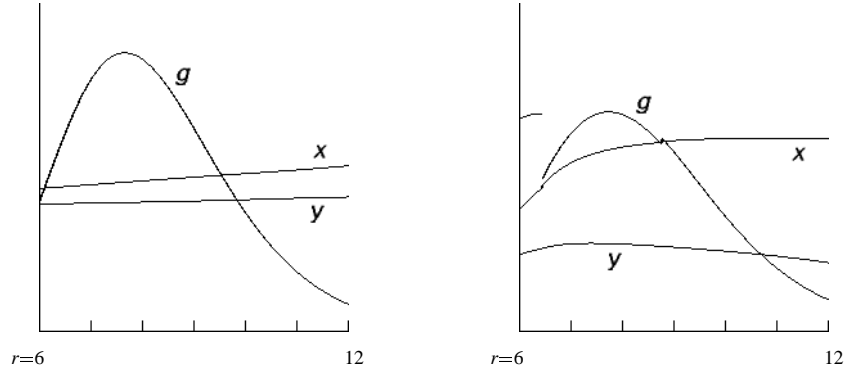
egyezsben azzal, hogy a (2) másodrendű  $G$  felület  $(x_0, y_0)$  lokális szélsőértékében ez az érték éppen  $EF - D^2$ , hiszen  $G$  második parciális deriváltjai rendre  $E$ ,  $F$ , and  $D$ . Hasonlóképpen ugyanebben a pontban a Minkowski-görbület

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}. \quad (5)$$

Mind a Gauss, mind a Minkowski görbület a felület csúcosságát, hegyességét méri. Gauss nevezetes tétele szerint a Gauss-görbület invariáns a felület hajlítgatásaival szemben, azt a felület belső geometriája meghatározza, szoros kapcsolatban van a felületre rajzolt háromszögek szögösszegével. Ugyanakkor egyik irányban hosszan elnyúló felületnél a Gauss-görbület majdnem nulla, függetlenül attól, hogy erre merőlegesen a felület mennyire meredek. Ilyenkor a Minkowski görbület jobban használható a felület görbültségének mérésére. Esetünkben a feldolgozandó foltok nagyjából kör alakúak, így  $P_r$  lokális maximumaiban a minimális és maximális görbület várhatóan közel van egymáshoz. Következésképp a kétfajta görbület ezeken a helyeken egyformán fog viselkedni.

Az illesztett  $G$  felület maximumában a görbület egyúttal becslést ad a  $P$  konvolúciós felület görbületére is. Minél nagyobb ez a görbület, annál szignifikánsabb a lokális maximum. Ez az észrevétel sugallja, hogy a különböző sugarakat annak alapján hasonlítsuk össze, hogy mekkora a lokális maximumban a görbület. Az optimális sugár a maximális görbülethez tartozik: mind kisebb mind nagyobb sugárhoz tartozó szűrő ezt a maximumot egy kissé "elkeni," és így a görbület kisebb lesz.

Ezek a megfontolások a következő eljárást sugallják egy folt helyének és sugarának becslésére. Legyen  $r_0$  valamint  $(u_0, v_0)$  az előzetesen becsült sugár illetve középpont. Az  $r_0$  körül választunk különböző  $r$  értékeket. Mindegyikkel elkészítjük a folt



5. ábra: Maximumhely és Gauss-görbület a hármas és négyes folt esetén

környékének az  $f_r$  szűrővel való  $P_r$  konvolúcióját. Az  $(u_0, v_0)$ -ból indulva megkeressük azt a rácspontot, ahol  $P_r$ -nek lokális maximuma van. Ott a 4. részben leírtak szerint kiszámítjuk a négyzetesen legjobban illeszkedő  $G$  másodfokú felületet, és azt, hogy hol veszi fel maximumát, illetve hogy ott mekkora a görbülete. Azt az  $r$ -et fogadjuk el a sugár közelítésének, amelyikre a görbület a lehető legnagyobb, és az ekkor adódó maximum helye adja meg a folt középpontját.

Az eljárás eredményét a hármas és négyes számú foltokra az 5. ábra mutatja. A vízszintes tengelyen az  $r$  sugár fut 6-tól 12-ig. A  $g$ -vel jelölt görbe a Gauss-görbület, az  $x$  illetve  $y$  pedig a lokális maximum helye. Mindkét esetben jól látható a görbület maximuma  $r = 7,6$  illetve  $r = 7,7$  esetén, tehát ezek az értékek adják meg a foltok sugarát. A hármas foltnál  $g$ ,  $x$  és  $y$  várakozásainknak megfelelően viselkedik:  $x$  és  $y$  nagyjából egyenletesen nő: miközben a sugár  $r = 6$ -ról  $r = 12$ -re nő,  $x$  mindegy kétharmad,  $y$  pedig egynegyed pixelnyit mozdul el. Eközben a  $g$  görbület először meredeken nő, majd a maximum elérése után meredeken zuhan. A négyes foltnál  $x$  és  $y$  nagyjából folytonosan követi  $r$ -et,  $x$  összesen másfél pixelnyit,  $y$  nagyjából egyharmad pixelnyit mozog; a görbületnek viszont  $r = 6,5$ -nél egy nagyobb,  $r = 8,8$  körül pedig egy kisebb ugrása van. Bár az ugrások ellenére a görbületnek jól meghatározott maximuma van, ami persze a keresett sugarat is megadja, általános esetben a szakadások miatt az eljárás egyáltalán nem, vagy csak igen körülményesen alkalmazható.

## 6. A végső eljárás

A görbületnél adódó szakadások okát vizsgálva nézzük meg egy kicsit pontosabban a fenti eljárást. Először is kiválasztjuk az  $r$  sugárhoz tartozó  $f_r$  szűrőt, majd az előzetesen becsült rácspontból indulva keresünk egy közeli  $(u_r, v_r)$  rácspontot helyet, ahol a konvolúciós integrál  $\tilde{P}_r$  közelítő értékének lokális maximuma van. Ezután meghatározzuk azt a  $G_r$  másodfokú felületet, amely az  $(u_r, v_r)$  rácspontnak egy (21 rácspontot tartalmazó) környezetében négyzetesen a legjobban illeszkedik  $\tilde{P}_r$ -re. Végül kiszámítjuk  $G_r$ -nek a maximumát valamint azt, hogy a maximumban mennyi  $G_r$  görbülete. A

görcbületben akkor tapasztalunk szakadást, mikor az  $(u_r, v_r)$ -beli lokális maximum épen átugrik egyik rácspontról egy másikba. Az illesztett  $G_r$  felület, míg helyesen ad számot arról, hogy hol is található  $\tilde{P}_r$  lokális maximuma, már nem elég jó ahhoz, hogy  $\tilde{P}_r$ -nek a maximumbeli görcbületét is jól megbecsülje.  $\tilde{P}_r$  görcbületét tehát közvetlenül kell kiszámítani. A (4) és (5) képletek szerint ehhez elegendő  $\tilde{P}_r$  második parciális deriváltjainak ismerete.

Láttuk a 2. rész végén, hogy ha a szűrő mindenütt értelmezve van (ami esetünkben fennáll), akkor  $\tilde{P}_r$  értékét nem csak rácspontokban, hanem tetszőleges  $(x, y)$  pontban tudjuk számítani a következőképpen:

$$\tilde{P}_r(x, y) = \sum_{i, j \in \mathbb{Z}^2} p(i, j) f_r(x - i, y - j).$$

Ennek alapján  $\tilde{P}_r$  parciális deriváltjai a jobb oldal tagonként deriválásával adódik. Így  $\tilde{P}_r$  egy parciális deriváltja annak a konvolúciónak az értéke, melyben a magfüggvény az  $f_r$  szűrő megfelelő parciális deriváltja. Így például

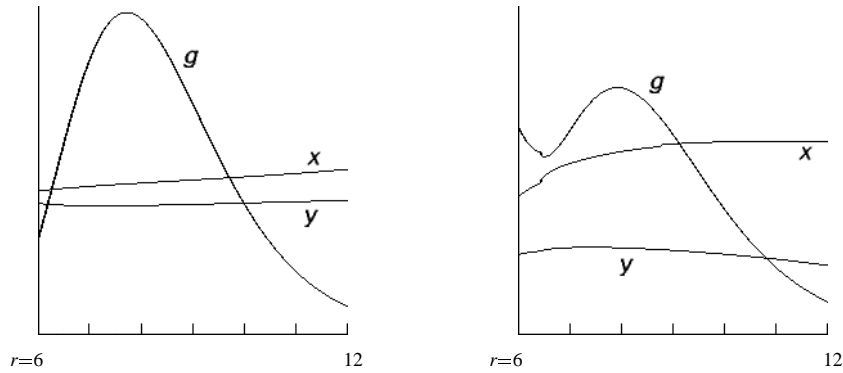
$$\frac{\partial^2}{\partial x^2} P_r = \langle p, \frac{\partial^2}{\partial x^2} f_r \rangle \quad \text{illetve} \quad \frac{\partial^2}{\partial y^2} P_r = \langle p, \frac{\partial^2}{\partial y^2} f_r \rangle.$$

Ezeknek a konvolúcióknak az értéke pedig tetszőleges pontban számítható.

Tekintsük a  $\tilde{P}_r$  függvénynek az  $(u, v)$  pont körüli Taylor sorát, és vegyük abban a legfeljebb másodrendű tagokat:

$$\tilde{P}_r(u + x, v + y) \approx \tilde{P}_r(u, v) + \frac{\partial \tilde{P}_r}{\partial x} x + \frac{\partial \tilde{P}_r}{\partial y} y + \frac{\partial^2 \tilde{P}_r}{\partial x \partial y} xy + \frac{1}{2} \frac{\partial^2 \tilde{P}_r}{\partial x^2} x^2 + \frac{1}{2} \frac{\partial^2 \tilde{P}_r}{\partial y^2} y^2, \quad (6)$$

$\tilde{P}_r$ -nek ebben a másodrendű közelítésében az összes együtthatót tetszőleges  $(u, v)$  pontban ki tudjuk számítani.



6. ábra: Maximumhely és Gauss-görcbület iterálás után

Ha  $(u, v)$ -nek az a pontot választjuk, amelyet a 5. részben ismertett eljárás eredményez, akkor  $\tilde{P}_r$  lokális maximumának még jobb közelítését adja a (6) másodrendű

felület maximuma. Ebben a maximumban (6) Gauss-görbülete

$$\frac{\partial^2 \tilde{P}_r}{\partial x^2} \cdot \frac{\partial^2 \tilde{P}_r}{\partial y^2} - \left( \frac{\partial^2 \tilde{P}_r}{\partial x \partial y} \right)^2 \quad (7)$$

jobb közelítést ad  $\tilde{P}_r$  görbületére. A 6. ábra mutatja az ennek alapján számított maximumhelyet és Gauss-görbületet. A maximum helye pár század pixellel mozdult el a négyzetesen legjobban illeszkedő felület maximumától, míg a görbület értéke jól láthatóan kisimult.  $\tilde{P}_r$  görbületének ez a közelítése tehát már használható a folt sugarának megállapítására.

A (6) felület hat együtthatója hat konvolúció kiszámítását jelenti, melyekben a magfüggvények értékeit helyben kell kiszámítani. Amennyiben elfogadjuk, hogy  $(u, v)$  megfelelően jól közelíti a lokális maximumot, akkor csak a görbületet kell számítanunk. A Gauss-görbület esetén (7) szerint ez három konvolúciót jelent, míg a Minkowski görbület esetén – a konvolúció linearitása miatt – ez egyetlen konvolúcióval is számítható:

$$\frac{\partial^2 \tilde{P}_r}{\partial x^2} + \frac{\partial^2 \tilde{P}_r}{\partial y^2} = \sum_{i,j \in \mathbb{Z}^2} p(i,j) \left( \frac{\partial^2 f_r}{\partial x^2} + \frac{\partial^2 f_r}{\partial y^2} \right) (x-i, y-j). \quad (8)$$

## 7. Összefoglalás

A kitzűzött feladat, miszerint határozzuk meg a képen található foltok helyét és nagyságát pixelnél nagyobb pontossággal, a bevetésben vázolt módszerrel megoldható. A 3. részben tárgyaltuk, hogy az  $f_r$  szűrőknek milyen feltételeket kell kielégíteniük. Fix  $r$  érték mellett a legjobb illeszkedést a kép és az  $f_r$  szűrő konvolúciójának maximuma adja. A maximum helyét pixelnél nagyobb pontossággal megkaphatjuk, ha a konvolúciós felületet egy négyzetesen legjobban illeszkedő másodrendű felülettel helyettesítjük. Természetesen a négyzetesen legjobban illeszkedő felület helyett használhatnánk a 6. részben bemutatottak mintájára azt a másodfokú polinomot, melynek együtthatói a rácspontban kiszámított parciális deriváltak. A maximumban mind a Gauss, mind a Minkowski görbület jó mérőszáma annak, hogy mennyire jó az illeszkedés. Azt az  $r$  sugarat kell választanunk, amire a görbület maximális.

A görbület rosszul viselkedhet abban az esetben, mikor az  $r$  sugár pár pixellel kisebb a folt valódi sugaránál, és a folt belsejében egyenetlenségek vannak. Ezért a követendő eljárás az, hogy  $r$ -et nagyobbra választjuk, majd addig csökkentjük, amíg a görbület nő.

Érdekes lenne megvizsgálni, hogy a  $p$  függvény, vagyis a kép hibái hogyan mutatkoznak a lokális maximum illetve a görbület értékében. Szokás szerint a hibáról feltesszük, hogy pixelenként független, kis szórású, nulla várható értékű normális eloszlásból származik. Elképzelhető, hogy a négyzetesen illeszkedő felület kisebb hibával adja meg a lokális maximumot, mint a harmadrendű tagoknál levágott hatványsor, különösen ha a hiba nagy lehet.

A másodendű közelítést arra használtuk, hogy a konvolúciós felület lokális maximumát megkeressük, vagyis olyan  $(u, v)$  pontot, ahol a felület mindkét parciális de-

riváltja nulla. Elképzelhető, hogy erre egy közvetlen, gyors algoritmus is adható, mely ezeknek a parciális deriváltaknak explicit alakját használja.

Egy folt sugarának megállapításához a görbület függvény maximumát kell megkeresni. Ha ehhez a (8) szerinti Minkowski görbületet használjuk, akkor (8) analitikusan deriválható, és a derivált zéróhelyét kell keresni. Egy egyváltozós függvény zéróhelyét numerikusan jóval egyszerűbb meghatározni, mint a maximumát. Lehet-e ezt az észrevételt egy használható algoritmussá fejleszteni?

## Irodalomjegyzék

- [1] Winkler, G(1996), *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Heidelberg, Springer
- [2] R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, Addison Wesley, 1992
- [3] Hajós György: *Differenciálgeometria*, egyetemi jegyzet, Tankönyvkiadó, Budapest, 1971
- [4] Lánzos Kornél: *A geometriai térfogalom fejlődése*, Gondolat, Budapest, 1976
- [5] J. Polzehl, V. Spokoiny, *Image denoising: pointwise adaptive approach*, preprint, Weierstass Institute, 2001
- [6] R. Fisher, S. Perkins, A. Walker, E. Wolfart, *Hypermedia image processing reference*, <http://www.dai.ed.ac.uk/HIPR2>