

Pázmány Péter Katolikus Egyetem
Információs Technológiai Kar
Multidiszciplináris Műszaki Tudományok Doktori Iskola
University of Bordeaux 1
Laboratoire Bordelais de Recherche en Informatique



Videó Események Detektálása és Vizuális Adat Feldolgozás Multimédiás Alkalmazásokhoz

Tézis füzet a Ph.D. Disszertációhoz

Szolgay Dániel

Tudományos Témavezetők:

Prof. Szirányi Tamás

Prof. Jenny Benois-Pineau

Prof. Tamás Roska, Akadémikus

Budapest, 2011

1. Bevezető és Probléma Felvetés

Az elmúlt 30 évben a képfeldolgozás teljes értékű mérnöki tudománnyá fejlődött és sok terület nélkülözhetetlen eszközévé vált, mint például orvosi vizualizáció, rendfenntartás, ember és gép közti kommunikáció, ipari minőség ellenőrzés és orvosi felügyelet.

A technika fejlődésével új lehetőségek nyíltak meg és az új lehetőségek legtöbbször új kihívásokat is jelentenek. A digitális képfeldolgozás kezdetén a feldolgozást mindössze néhány képre kellett elvégezni. Azonban ezredforduló előtt nem sokkal megjelentek és elterjedtek a digitális videók, miközben a folyamatosan növekvő képi adatbázisok mérete lassan meghaladta a manuálisan kezelhető méretet. Az új kihívások új módszerek megjelenését vonták maguk után: a képi keresés, a videó kódolás, a videókon események felismerése mind a szakterület részévé váltak. Ma a digitális videó kamerák mindenki számára könnyen hozzáférhető eszköznek számítanak. Ennek eredménye, hogy a videó adat mennyisége rohamosan növekszik és ezzel együtt a felvételek tartalma is bonyoltabban, nehezebben feldolgozható, hiszen legtöbbször sem a kamera, sem a használója nem profi. Így sokszor elmosódott, zajos és kiszámíthatatlan kameramozgásokat tartalmazó felvételekkel kell dolgozni. Az ilyen videók feldolgozásához az egész eljárásnak, alacsonytól a magas szintig, adaptálódnia kell a feladathoz. Munkám során alacsony és közép szintű képfeldolgozási feladatokkal foglalkoztam, melyek megoldása alap feltétele a fent említett videó adat hatékony kezelésének. A disszertáció első két részében alap képjavítási problémákkal foglalkoztam úgymint a képek elmosódásának megszüntetését segítő dekonvolúciós algoritmusok optimalizálása illetve a kép pusztán

geometriai információt tartalmazó részének előállítására a textúra elkülönítésével, míg a harmadik rész egy magasabb rendű videó értelmezési feladatra koncentrál, melynek célja mozgó objektumok elkülönítése a statikus háttértől mozgó kamerával készült felvételeken.

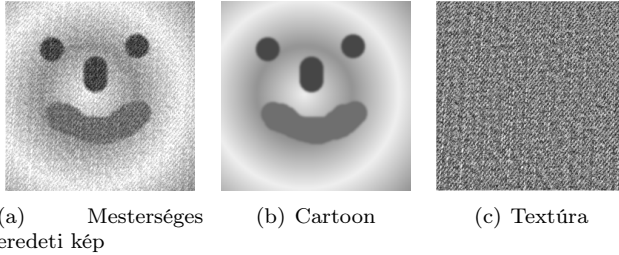
Bizonyos képfeldolgozási problémák egészen a digitális képalkotás kezdeteitől jelen vannak, újabb és egyre jobb megoldásokra várva. Ilyen feladat a képek elmosódottságának megszüntetése, mely a gyorsan mozgó kamerával készített felvételeket is jellemzi. Az elmosódottságnak – a mozgás mellett – természetesen más okai is lehetnek, mint például hibásan beállított fókusz távolság vagy a felvevő optika hibája. Ezen hibák miatt, az elkészült kép torzul, elmosódik és javítás nélkül sokszor használhatatlan. Ezt a torzulást általában konvolúcióval szokták modellezni: az eredeti ismeretlen képet konvolváljuk egy ún. pontszóródási függvényvel (PSF). A PSF egy pontszerű fényforrás a képalkotás során elszorított torzulását írja le. A cél egyértelmű: a lehető legjobb minőségben visszaállítani az eredeti képet az elmosódott kép és - bizonyos esetekben - a PSF alapján. A legtöbb eljárás úgy közelíti meg a problémát, hogy keresi azt a képet, mely konvolválva a becsült (vagy pontosan ismert) PSF-el a lehető leghasonlóbb lesz a mért, elmosódott képhez. Ez a megközelítés azonban alulhatározott problémát eredményez, mivel a keresett eredeti kép mellett sok más kép is kielégíti a fenti feltételt. Ennek hatására sok nem-regularizált iteratív dekonvolúciós módszer közös problémája, hogy előbb-utóbb zajt visznek a becsült képre (lásd az 1. ábrán). A céлом az volt, hogy találjak olyan automatikusan számolható megállási feltételt, mely az iteratív folyamatot a leghatékosabb pontban állítja le.



(a) Elmosódott kép (b) Visszaállított kép 14 iteráció után (c) Visszaállított kép 60 iteráció után

1. ábra. A példa képsorozat azt mutatja, hogyan erősítik fel a zajt a nem regularizált iteratív dekonvolúciós módszerek abban az esetben, ha az ideálisnál tovább futnak.

Videók és álló képek esetén egyaránt fontos feladat a képek felbontása geometriai és zajszerű komponensekre. Zaj eltávolítás, képtömörítés, képi jellemzők kinyerése mellett sok más feladat elvégzéséhez nyújthat segítséget, akár előfeldolgozásként is, hasonlóan az árnyék, tükröződés, füst/köd eltávolítási módszerekhez. A geometriai és textúra részek elméletileg függetlenek egymástól: az ún. cartoon kép csak geometriai információt tartalmaz, míg a textúra kép, a cartoon rész komplementereként áll elő és nem tartalmaz geometriai információt. A 2. ábrán látható egy példa szemléltetésképpen. Egy jó dekompozíciós algoritmus anélkül eltávolítja a textúrát a képről, hogy elmosná a fontos körvonalakat. A textúra kép ezután az eredeti és a cartoon kép különbségként állítható elő. Az eltér objektumok háttértől való elszeparálása egy olyan alapvető feladat, mely nagy érdeklődésre tarthat számot, hiszen ennek eredménye számos magasabb szintű algoritmus (pl.



2. ábra. Péld a cartoon/textúra dekompozícióra.

objektumok detektálása és követése) kiinduló pontja lehet. Az erős kameramozgás, a jelentős perspektíva változás és a felvételek zajossága még jobban megnehezíti a feladatot testen viselhető kamerák esetén. Általában alacsony szintű eljárásokat, úgymint zaj eltávolítást, elmosódottság megszüntetését és morfológiai javításokat használnak elő- és utófeldolgozásként a jobb eredmény elérése érdekében.

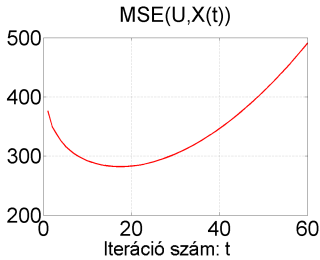
2. Kutatási Módszerek

Az utóbbi évtizedekben folyamatosan kerültek napvilágra újabb és újabb eljárások, melyek az eredeti képet próbálják visszaállítani egy elmosódott, zajos mérésből kiindulva. A disszertációban bemutatott mérésekhez egy széles körben használt, iteratív, nem regularizált dekonvolúciós algoritmust, a Richardson-Lucy [7, 8] algoritmust használtuk.

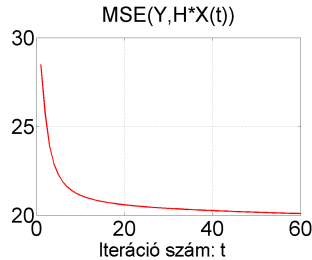
Dekonvolúciós feladatokban az eredeti kép U ismeretlen és csak a zajos elmosódott mérést Y használhatjuk, U becsléséhez. Ha $X(t)$ jelöli a dekonvolúciós algoritmus U -ra adott becslését a t -edik iteráció után, akkor elmondható, hogy a legtöbb algoritmus célfüggvénye valamilyen formában $|Y - H * X(t)|$ minimalizálásán alapul. Nyilvánvalóan a tényleges cél az $|U - X(t)|$ vagy a széleskörben használt $MSE(U, X(t))$ függvény minimalizálása és a minimum pontban a dekonvolúció leállítása lenne. Ezek a mértékek azonban direkt úton nem kiszámíthatók, hiszen U ismeretlen. Az U hiányában rendelkezésünkre álló $MSE(Y, H * X(t))$ függvény nem megfelelő helyettese a $MSE(U, X(t))$ -nek, mint ahogy ez a 3. ábrán bemutatott példa alapján is egyértelműen látszik.

Az ADE szögeltérési hibafüggvény [9] felhasználva új megál-lási kritériumot dolgoztunk ki az $X(t)$ és $X(t) - X(t - 1)$ közti függetlenség alapján. A módszer alkalmazhatóságának kísérleti is bizonyításához olyan tesztkörnyezetben dolgoztunk, ahol az U ismert és ezáltal az eredmény minősége mérhető.

A képek zajmentesítésének egy speciális esete, amikor a képet textúra- és strukturális komponensre bontjuk. A kidolgozott eljárás a nem régiben publikált BLMV [10] szűrő által készített képet használja az Anizotrop Diffúzió (AD) [11] inicializálásához. Az iteratív AD-t pedig a két komponens közti ortogonalitási feltételt felhasználva [12], automatikusan állítjuk meg, az ADE hibafüggvény (ADE) [9] segítségével. Az elkészült algoritmust összehasonlítottuk az elérhető legjobb módszerekkel mind mesterséges képeken (ahol lehetőség van a tökéletes szeparációval való összevetésre és számszerű kiértékelésre), mind valós felvételeken, melyeknek megítélése bár szubjektív mégis a legelter-



(a) Az ismeretlen $MSE(U, X(t))$



(b) A mérhető $MSE(Y, H * X(t))$

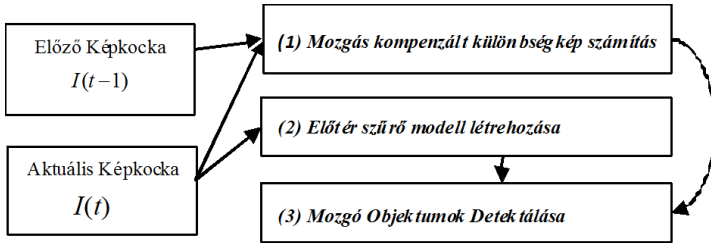
3. ábra. A mérhető négyzetes hibafüggvény $MSE(Y, H * X(t))$ nem követi az ismeretlen $MSE(U, X(t))$ függvényt, ahol U az eredeti kép, $X(t)$ a rekonstruált kép t iterációt követően és H a PSF. Látható, hogy az ideális megállási pont nem becsülhető $MSE(Y, H * X(t))$ alapján.

jedtebb validációs módszere az ilyen jellegű algoritmusoknak.

Az eddig említett algoritmusokhoz tartozó implementáció MATLAB[®] [13] környezetben készült.

A testen viselhető kamerával készített felvételeken a kamera független mozgások detektálásához létrehoztunk egy keretrendszert, mely 3 fő részből áll (lásd a 4. ábrán): (1) Különbségképek számítás mozgás kompenzált képeken, (2) Előtérszűrő modell létrehozása, (3) Mozdó objektumok detektálása.

A kamera mozgás kompenzálása nélkülözhetetlen első lépés, melyet hierarchikus blokk-illesztő algoritmus (a továbbiakban HBM) [14] és globális mozgást becselő eljárás (GME) [15] segít-



4. ábra. Az előtérben mozgó objektumok detektálásának lépései.

ségével valósítottunk meg.

Ezáltal a videó egymást követő két képkockája ábrázolhatóvá válik egy közös koordináta rendszerben és az ún. hibakép előáll a két kép különbségként. Az így kapott előtér maszkon megjelenő hamis pozitív találatok kiszűrését a háttér valószínűségi sűrűségfüggvényének becslésével végeztük. A mintapontokat az első lépés során létrehozott módosított hibaképről (MEI) kaptuk, ami egy mozgáskompenzált különbségkép kiegészítve az aktuális képkockáról származó szín információval.

A háttér sűrűségfüggvényének becslését a rendelkezésre álló minták alapján, kernel sűrűség becslés segítségével végeztük, Gaussi kernel függvényel.

A 3. lépés a mozgó objektum detektálása, ami az előzőleg elkészített sűrűségfüggvényből vont valószínűségek küszöbölését és az így kapott előtérpontok klaszterezését jelenti. Ennek meg-

valósításához a DBSCAN [16] algoritmust használtuk egy 7 dimenziós paraméter térben.

A keretrendszer C++ programozási nyelven készült, felhasználva az OpenCV könyvtár [17] függvényeit. Bár az algoritmus egy szálon fut, az elkészült munka tartalmaz egy rövid tanulmányt egy több szálon (GPU segítségével) parallel futó változat elkészítésének lehetőségeiről.

3. Új Tudományos Eredmények

1. Tézis: *A megállási feltétel meghatározása általános probléma a nem regularizált iteratív dekonvolúciós módszerek esetében. Új módszert adtam az ideális megállási pont automatikus meghatározásához, a mért jel és a jel gradiensének ortogonalitása alapján. A módszer alkalmas az eddig használt ad-hoc eljárások négyzetes hibájának alsó burkolót adni. Az elméleti megfontolást kísérletekkel támasztottam alá, melyek bizonyítják az algoritmus hatékonyságát különböző zaj modellek és jel-zaj viszony esetén.*

A tézishez kapcsolódó publikációk: [2, 3].

Új módszert dolgoztunk ki az ideális megállási pont automatikus meghatározásához nem regularizált iteratív dekonvolúciós módszerek esetén az ADE [9] ortogonalitás mértéket használva a széles körben használt négyzetes hiba mérték (MSE) helyett.

A javasolt módszer alkalmas az iteráció optimális pontban való leállítására az aktuálisan becsült jel és a jel gradiensének

függetlensége alapján, megelőzve ezzel az iterációknak egy olyan szakaszát, amely nem javítja tovább a képet, csak zajt ad hozzá. A javasolt $ADE(X_e(t), X(t))$ függvény csak mérhető értékeket tartalmaz, vagyis minden adat rendelkezésre áll a kiszámításához és elméletileg is értelmezhető. Az $ADE(X_e(t), X(t))$ függvény minimumánál a két egymást követő iterációban tett becslés közti különbség $X_e(t)$ a lehető legfüggetlenebb magától a becsült képtől ezért feltételezhetjük, hogy $X_e(t)$ nagyrészt független, zajszerű információt tartalmaz és nem a kép struktúrájára vonatkozó információt. Ezért a további iterálás inkább rontja, mint javítja a kép minőségét.

A módszerünket a széles körben ismert Richardson-Lucy [7, 8] dekonvolúciós algoritmus használatával teszteltük különböző zaj modellekkel (Gaussi és Poisson) és eltérő zaj szinttel. Az eljárás nem igényel semmilyen kalibrációt vagy manuális beállítást. A javasolt módszer ($ADE(X_e, X_{re})$) és az elméletileg legjobb megoldás ($MSE(U, X(t))$) közti korreláció 0.6726, ha az iteráció számot vesszük alapul. Míg ha a kép minőségét tekintjük, akkor a korreláció még magasabb 0.9556 lesz. A disszertációban bemutatott eredmények alapján elmondható, hogy a javasolt módszer egyértelműen jobban teljesít, mint az általában használt ad-hoc eljárások.

2. Tézis: *Új, axiomatikus módszert adtam a képen szereplő geometriai és textúra részek automatikus szétválasztására. Az eljárás alapját anizotrop diffúzió adja, melynek képtartalomtól függő, megfelelő iterációban történő leállításához az ADE ortogonalitás mértéket használtam. Kísérletekkel bizonyítottam,*

hogy a létrehozott új eljárás jobb eredménnyel választja szét a képen a textúrát és a geometriai információt, mint az utóbbi években publikált módszerek.

A tézishez kapcsolódó publikáció: [4]

Az Anizotrop Diffúzió (AD) [11] célja, hogy a képen úgy hajtson végre elmosást és ez által zajszűrést, hogy a képen szereplő erősebb éleket érintetlenül hagyja. Ehhez az összes lehetséges diffúziós irányban egy súly függvényt használ, ami meggátolja a diffúziót az adott irányba, ha ott az irányra merőleges él szerepel és megengedi a diffúziót, ha nincs ilyen él. Az AD hagyományos formájában nem alkalmas a geometriai (más néven cartoon) és a textúra információ szétválasztására, mivel a textúra is tartalmazhat erős éleket, amiket el kellene mosni, míg a cartoon is tartalmazhat gyenge éleket, amiket meg kéne őrizni. A javasolt eljárás a BLMV nem lineáris szűrő [10] által készített cartoon képet használja az AD súlyfüggvényének inicializálásához. Ezen a képen a textúrált részek már bizonyos mértékben el vannak mosva, így az AD nem fogja megőrizni őket, míg a fontosabb cartoon élek megmaradnak, így az AD súlyfüggvénye meg fogja védeni őket az elmosódástól. Az iteratív AD-t a két komponens közti ortogonalitási feltételt felhasználva, az ADE mérték segítségével automatikusan állítjuk le.

Röviden összefoglalva elmondható, hogy a javasolt eljárás elméletileg helyes megoldást kínál a geometriai és textúra részek szétválasztásához szükséges fő kérdésekre:

- Adaptív méret meghatározás a lokálisan, ADE mérték által optimalizált BLMV szűrősegítségével;

- Az adaptív BLMV által inicializált AD a textúra és a geometriaiinformáció szétválasztásához;
- Orthogonalitási kritérium szabályozza az AD leállítását és ezáltal minőségi mértéket ad a dekompozícióhoz.

Az elkészült algoritmust összehasonlítottam a ma elérhető legjobb hasonló módszerekkel (TVL1 [18], ROF [19], DPCA [20], DOSV [21], AD [11], BLMV [10]), mind mesterséges képeken numerikus kiértékelést alkalmazva, mind valós felvételeken jól meghatározott szempontokat alapján. Az eredmények valós képeken történő értékelése, nyilvánvaló szubjektivitása ellenére a ma használt legelterjedtebb módszer.

Mindkét kiértékelési megközelítés az itt bemutatott módszer egyértelmű fölényét mutatja. A javasolt módszer további előnye, hogy a többivel ellentétben nem igényel pontos manuális paraméterezést, csupán egy paraméter tartomány megadása szükséges.

3. Tézis: *Kernel sűrűségfüggvény becslésen alapuló új eljárást dolgoztam ki mozgó előtér detektálására viselhető kamerával készült felvételekhez (25 fps, 320x240 képméret), melyeket általában erős és kiszámíthatatlan kameramozgás jellemez.*

A tézishez kapcsolódó publikációk: [1, 5]

Mozgó kamerával készült felvételek feldolgozása esetén a kamera mozgás becslése és kompenzálása az első lépés, melyet hierarchikus blokk-illesztő algoritmus (a továbbiakban HBM) [14] és affin globális mozgást becslő eljárás (GME) [15] felhasználásával valósítottunk meg.

Ezáltal a videó egymást követő két képkockája ábrázolhatóvá válik egy közös koordináta rendszerben és a hibakép előáll a két kép abszolút különbségeként. Ez a hibakép ideális esetben csak előtérpontokban tartalmazna magas értékeket, míg a statikus háttér pontok a különbség képzés során kioltanak egymást. A perspektíva megváltozása, kvantálási hiba és a mozgáskompensáció kisebb pontatlansága következtében sok a hibás pozitív találat, ezért a hibakép önmagában nem alkalmas előtér modellnek. Létrehoztunk egy módosított hibaképet (MEI), ami a mozgáskompensált különbségkép kiegészítve az aktuális képkockáról származó szín információval.

A mozgó objektumok és a hiba képen jelenlévő statikus háttér elemek pixeleinek elkülönítéséhez a háttér sűrűségfüggvényének a becslését és egy valószínűség alapú döntési szabályt dolgoztunk ki.

A sűrűségfüggvény becslését tér-időbeli ablakból vett minták alapján kernel sűrűség becslés [22] segítségével végeztük, Gaussi kernelt alkalmazva. A tér-időbeliség arra utal, hogy a minta pontokat egy térbeli környezet különböző időpillanatokban vett értékeiből választottuk.

A kernel függvény szélességének beállításához a k legközelebbi minta pontot vettük figyelembe ahelyett, hogy csak a k . pontot használtuk volna, így csökkentve az alacsony minta számból fakadó esetleges hibákat.

A minta pontok száma a jelen feladatban erősen korlátozott és a zaj esetenként nagyon erős lehet, ezért a mintapontok választásának módja kulcsfontosságú.

Közismert eljárás a mintapont választásra rögzített kamerák esetén egy adott (x, y) koordinátájú pixel n korábbi értékének

használata [23]. Mozgó kamera esetén azonban ez a módszer nem megbízható. A mozgás kompenzáció ellenére egy valós, statikus háttérpont, ami egy adott képkockán az (x, y) koordinátájú pontnak felel meg a következő képen lehet, hogy nem pont ugyanott lesz. Ez magyarázható a mozgás kompenzáció kisebb hibáival vagy kvantálásból adódó hibával egyaránt. Ezt a hibát térben véletlenszerűnek feltételezve egy kis (x, y) középpontú térbeli ablak használatát javasoltuk.

Ezt követően az M mátrix értékei alapján egy együttes valószínűségi sűrűség függvényt számoltunk a színcsatornákra a MEI minden nem nullaértékű pontjára. Ahol az M mérési mátrix mindig az n megelőző, mozgáskompenzált képkockát tartalmazza.

Az így kapott sűrűségfüggvények alapján minden pixelt előtérnek vagy háttérnek osztályoztunk egy, a függvények karakteristikáját figyelembe vevő adaptív küszöbölés segítségével. A kapott előtér pontokat újra klasszifikáltuk a hozzájuk tartozó mozgás koordináták, szín értékek és térkoordináták alapján a DBSCAN klaszterező algoritmus segítségével.

Kísérleti úton bebizonyítottuk, hogy a bemutatott eljárás hatékonyabban működik viselhető kamerával készített felvételek esetén, mint a jól ismert Stauffer-Grimmson [24] módszer. Jelenleg az algoritmus offline működésre képes, mivel a nagy számítási igénye nem teszi lehetővé a valós idejű futtatást.

4. Új Tudományos Eredmények Lehetséges Felhasználási Területei

A viselhető kamerákkal készített videó megfigyelés rengeteg lehetőséget hordoz magában az egészségügyi, biztonságtechnikai vagy akár a szociális élet területén. Fontos kiegészítő eszköze lehet az időskori demencia diagnosztizálásának olyan esetekben, amikor a hagyományos módszerek sikertelenek, mivel a páciensek nem tudják, vagy nem akarják segíteni az orvosok munkáját. Videó logok készítésével az orvosok betekintést nyerhetnek a beteg mindennapjaiba, ami adott esetben nagy segítség lehet a helyes körkép felállításához. Biztonsági megfigyeléshez is fontos eszköz lehet a viselhető kamera olyan körülmények között, amikor a hagyományos statikus kamerák használata nem lehetséges (pl. rendőr járőrökön).

Manapság egyre népszerűbbek a blogok és az ún. life logok, melyekben a szerzőjük saját életét írja le nagyjából úgy, mint egy naplóban, kiegészítve a modern technika adta lehetőségekkel (képek, zenék, videók felöltésével). A Microsoft SenseCam [25] projektje egy viselhető kamerával a hordozója egész napját fényképekkel dokumentálja, lehetővé téve egy fényképekből álló napló könnyű létrehozását. (Ennek a szórakoztatás mellett orvosi felhasználásai is lehetnek memória zavarral küzdő betegek esetén.) Viselhető videó kamerákkal és megfelelő feldolgozó algoritmusokkal a SenseCamhez hasonló videó naplók készítése is vélhetően vonzó lenne a blog író társadalom számára.

A disszertációban bemutatott előtér/háttér szegmentálási eljárás csak az első lépés a videók tartalom alapú kereséséhez

vezető úton, ami az egyik legintenzívebben kutatott terület a képfeldolgozásban.

A geometriai és textúra információ szétválasztása sok egyéb algoritmus számára jelenthet jó kiinduló pontot. Tömörítési eljárásoknál a két komponens külön választásával jobb eredmény érhető el [26], ahogy azt korábbi módszerek megmutatták [27, 28]. Él kereső eljárásoknál a fontos élek megtalálásához adhat segítséget [10], képi jellemzők kinyerésére [18], valamint zajszűrésre is alkalmas abban az esetben, ha nulla középértékű véletlen zajjal van terhelve a kép [19]. Kamera mozgás becslésénél is hasznos lehet, a becslést hátráltató zaj hatásának csökkentésére.

Dekonvolúciós módszerek használata mindennapos olyan területeken, ahol digitális képeket alkalmaznak és az elmosódottság problémát jelenthet (pl.: mozgó kamerás felvételek, mikroszkópia, asztronómia). Bár manapság a regularizáció számít a fő irányvonalnak a területen, a nem regularizált módszerekkel is minőségi eredményeket lehet elérni [29]. Ezeknél a módszereknél az iterációt megállító feltétel kulcskérdés. A kidolgozott módszer elméletileg értelmezhető és effektív megoldást kínál a problémára.

Köszönetnyilvánítás

Mindenekelőtt szeretnék köszönetet mondani témavezetőimnek Dr. Szirányi Tamásnak és Dr. Jenny Benois-Pineau-nak, akik támogattak és segítettek mindenben és útmutatást adtak tanulmányaim során.

Köszönöm Roska Tamás konzulensemnek, a doktori iskola

vezetőjének tanácsai és biztatását.

Nagyon köszönöm minden kollégámnak, akik ötleteikkel, tanácsaikkal segítették munkámat.

Köszönöm a támogatást a Pázmány Péter Katolikus Egyetemnek és a Bordeaux 1 Egyetemnek, ahol a Ph.D. tanulmányaimat folytathattam. Bordeaux-i tanulmányaimat a francia állam támogatta a "Bourses Eiffel" és a "Bourses pour doctorat en cotutelle" ösztöndíjakon keresztül.

Végül, de mégis elsősorban köszönöm szüleimnek és egész családomnak hogy minden elképzelhető módon támogattak, bátorítottak és segítettek tanulmányaim során.

Publikációk

- [1] D. Szolgay, J. Benois-Pineau, R. Megret, Y. Gaestel, and J.-F. Dartigues, „Detection of moving foreground objects in videos with strong camera motion,” *Pattern Analysis and Applications*. accepted in 04.04.2011.
- [2] D. Szolgay and T. Szirányi, „Orthogonality based stopping condition for iterative image deconvolution methods,” in *Computer Vision ACCV 2010*, vol. 6495 of *Lecture Notes in Computer Science*, pp. 321–332, Springer Berlin / Heidelberg, 2011.
- [3] D. Szolgay and T. Sziranyi, „Optimal stopping condition for iterative image deconvolution by new orthogonality criterion,” *Electronics Letters*, vol. 47, no. 7, pp. 442–444, 2011.

- [4] D. Szolgay and T. Sziranyi, „Adaptive image decomposition into cartoon and texture parts optimized by the orthogonality criterion,” *IEEE Transactions on Image Processing*. Submitted in May 2011.
- [5] R. Megret, D. Szolgay, J. Benois-Pineau, P. Joly, J. Pinquier, J.-F. Dartigues, and C. Helmer, „Wearable video monitoring of people with age dementia : Video indexing at the service of health care,” in *International Workshop on Content-Based Multimedia Indexing, 2008.*, pp. 101 – 108, june 2008.
- [6] D. Szolgay, C. Benedek, and T. Sziranyi, „Fast template matching for measuring visit frequencies of dynamic web advertisements,” *Proceedings of VISAPP 2008, Third International Conference Computer on Vision Theory and Applications*, pp. 228–233, 2008.

A témához kapcsolódó irodalom

- [7] W. Richardson, „Bayesian-based iterative method of image restoration,” *JOSA*, vol. 62, pp. 55–59, 1972.
- [8] L. Lucy, „An iterative technique for rectification of observed distributions,” *The Astronomical Journal*, vol. 79, pp. 745–765, 1974.
- [9] L. Kovacs and T. Sziranyi, „Focus area extraction by blind deconvolution for defining regions of interest,” *IEEE Tr. Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 1080–1085, 2007.

- [10] A. Buades, T. Le, J.-M. Morel, and L. Vese, „Fast cartoon + texture image filters,” *IEEE Transactions on Image Processing*, vol. 19, no. 8, pp. 1978–1986, 2010.
- [11] P. Perona and J. Malik, „Scale-space and edge detection using anisotropic diffusion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 629–639, 1990.
- [12] J.-F. Aujol and G. Gilboa, „Constrained and snr-based solutions for tv-hilbert space image denoising,” *J. Math. Imaging Vis.*, vol. 26, pp. 217–237, November 2006.
- [13] MATLAB, *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc., 2010.
- [14] M. Bierling, „Displacement estimation by hierarchical block matching,” pp. 942–951, 1988.
- [15] M. Durik and J. Benois-Pineau, „Robust motion characterisation for video indexing based on mpeg2 opticalflow,” *In Proc. of the International Workshop on Content-Based Multimedia Indexing*, pp. 57–64, 2001.
- [16] M. Ester, H. Peter Kriegel, J. S. and X. Xu, „A density-based algorithm for discovering clusters in large spatial databases with noise,” pp. 226–231, AAAI Press, 1996.
- [17] G. Bradski and V. Pisarevsky, „Intel’s computer vision library: applications in calibration, stereo segmentation, tracking, gesture, face and object recognition,” in *IEEE*

Conference on Computer Vision and Pattern Recognition, 2000. Proceedings, vol. 2, pp. 796–797, 2000.

- [18] W. Yin, D. Goldfarb, and S. Osher, „Image cartoon-texture decomposition and feature selection using the total variation regularized L1 functional,” in *Variational, Geometric, and Level Set Methods in Computer Vision*, pp. 73–84, Springer, 2005.
- [19] L. I. Rudin, S. Osher, and E. Fatemi, „Nonlinear total variation based noise removal algorithms,” *Phys. D*, vol. 60, pp. 259–268, November 1992.
- [20] F. Zhang, X. Ye, and W. Liu, „Image decomposition and texture segmentation via sparse representation,” *Signal Processing Letters, IEEE*, vol. 15, pp. 641–644, 2008.
- [21] R. Shahidi and C. Moloney, „Decorrelating the structure and texture components of a variational decomposition model,” *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 299–309, 2009.
- [22] E. Parzen, „On estimation of a probability density function and mode,” *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.
- [23] A. Mittal and N. Paragios, „Motion-based background subtraction using adaptive kernel density estimation,” vol. 2, pp. 302–309, june-july 2004.
- [24] C. Stauffer and W. Grimson, „Learning patterns of activity using real-time tracking,” *IEEE Transactions on Pattern*

- Analysis and Machine Intelligence*, vol. 22, pp. 747–757, aug 2000.
- [25] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood, „Sensecam: a retrospective memory aid,” *International Conference on Ubiquitous Computing, LNCS 4206*, pp. 177–193, 2006.
- [26] N. Sprljan, M. Mrak, and E. Izquierdo, „Image compression using a cartoon-texture decomposition technique,” *Proc. Int. Work. on Image Analysis for Multimedia Interactive Services (WIAMIS)*, p. 91, 2004.
- [27] M. Kunt, A. Ikonomopoulos, and M. Kocher, „Second-generation image-coding techniques,” *Proceedings of the IEEE*, vol. 73, no. 4, pp. 549–574, 1985.
- [28] D. Barba and J.-F. Bertrand, „Optimization of a monochrome picture coding scheme based on a two-component model,” in *9th International Conference on Pattern Recognition, 1988.*, pp. 618–622 vol.1, nov 1988.
- [29] S. C. L. Zou, H. Zhou and C. He, „Dual range deringing for non-blind image deconvolution,” *International Conference on Image Processing*, pp. 1701–1704, 2010.