

UNIVERSITATEA "POLITEHNICA" DIN TIMIȘOARA

Facultatea de Electronică și Telecomunicații

Departamentul de Comunicații

DETECȚIA ȘI EXTRAGEREA FUNDALULUI

ÎN SUPRAVEGHEREA VIDEO

PRIN

TEHNICI DE ESTIMARE NONPARAMETRICE

Teză de doctorat

Conducător științific:

Prof. dr. ing. Corneliu I. Toma

Doctorand:

Ing. Codruț Ninu Ianăși

2006

SINTEZA LUCRĂRII

Această teză este strâns legată de activitatea de dezvoltare a unui sistem de videosupraveghere destinat să genereze date statistice concludente pentru analiza comportamentului consumatorilor în spații comerciale. Lucrarea prezintă doar rezultatele inovatoare din cadrul acestei activități, referitoare la estimarea fundalului prin tehnici robuste, nonparametrice. Soluțiile propuse în lucrare au permis concomitent reducerea erorilor de estimare, creșterea adaptibilității și reducerea complexității de calcul. Lucrarea este structurată în cinci capitole.

Capitolul introductiv al tezei face o prezentare a problematicii generale a supravegherii video și plasează în context estimarea și substrația fundalului, problemă asupra căreia se concentrează cea mai mare parte a contribuțiilor din teză. Capitolul este structurat pe patru paragrafe. Primul paragraf prezintă domeniile de bază ale supravegherii și analizei video inteligente, exemplificând aplicații de actualitate. Componentele unui sistem de supraveghere video general sunt descrise în paragraful 1.2. Totodată se face o prezentare a funcțiilor și metodelor de prelucrare specifice fiecăruia din blocurile componente. Paragraful 1.3 conține argumentează actualitatea temei estimării și substrației fundalului în sistemele moderne de supraveghere. Paragraful final expune și justifică structura tezei de doctorat.

Capitolul 2 este structurat pe cinci paragrafe. În primul paragraf sunt definite și discutate principalele dificultăți cu care se confruntă estimarea fundalului în aplicațiile curente. Stadiul curent al cercetării în problema estimării fundalului pentru aplicații de supraveghere video este prezentat în paragraful 2.2. Se desprind două direcții de lucru predominante. Cea tradițională, bazată pe metode de estimare parametrice și o direcție nouă, bazată pe metode de estimare nonparametrice. În paragraful 2.3 se tratează fundamentele teoretice ale metodelor de estimare parametrice, cu referire la modelul larg adoptat în literatură al sumei de gaussiene. În paragraful 2.4. sunt prezentate metodele de estimare nonparametrice. Se include o discuție referitoare la algoritmul cu translație la medie (*mean shift*) pentru detecția rapidă a modurilor funcției

densitate de probabilitate (FDP). În paragraful 2.5 sunt trecute în revistă concluziile rezultate din studiul prezentat în acest capitol.

Capitolul 3 este structurat în șase paragrafe. În primul paragraf se prezintă criteriile care stau la baza adoptării dimensiunii corecte a bufferului de cadre folosite pentru estimarea inițială a fundalului, în cazul utilizării tehnicilor de estimare nonparametrică. Paragraful 3.2 descrie proiectarea unui estimator nonparametric pentru estimarea inițială a fundalului, cu referire specifică la alegerea optimală a factorului de scală. Modalitatea de subtracție a fundalului pentru segmentarea fundal / prim-plan este fundamentată în paragraful 3.3. În paragraful 3.4 se propune un estimator nonparametric recursiv ce asigură urmărirea eficientă a schimbărilor de fundal. Rezultatele testelor de performanță ale estimatorului nonparametric recursiv sunt prezentate în paragraful 3.5. Testele evaluează atât precizia estimării în diverse condiții de funcționare cât și viteza de calcul comparativ cu soluții de referință. Concluziile capitolului sunt sintetizate în paragraful 3.6.

Capitolul 4 este structurat în șase paragrafe. În primul sunt trecute în revistă tehnicile adaptive de estimare a fundalului folosite curent. Paragraful 4.2. face o dezvoltare teoretică a ecuației de actualizare a fundalului, care este derivată din algoritmul de detecție a modului cu deplasare la medie pentru cazul unui nucleu de formă generală. Analiza efectului formei nucleului pentru estimarea funcției densitate de probabilitate asupra performanțelor statice și dinamice ale estimatorului nonparametric recursiv este subiectul paragrafului 4.3. În paragraful 4.4 este dezvoltată o variantă îmbunătățită a estimatorului recursiv pentru urmărirea mai rapidă a schimbărilor moderate de fundal. Acest nou estimator l-am denumit autoadaptiv. Performanțele noului estimator sunt evaluate comparativ cu cele ale estimatorului neadaptiv. Răspunsul estimatorului la schimbări drastice ale fundalului scenelor supravegheate este îmbunătățit semnificativ prin dezvoltarea unei versiuni noncauzale, descrise în paragraful 4.5. Concluziile capitolului sunt incluse în paragraful final.

În **capitolul 5** se face o sinteză a celor 7 contribuții principale ale doctorandului, cu caracter de noutate în domeniul detecției și extragerii fundalului în supravegherea video prin tehnici nonparametrice, prezentate sintetic în continuare, fiind însoțite de referințe bibliografice, cu referire la articolele publicate de doctorand.

NOTĂ

O parte importantă a soluțiilor și metodelor noi dezvoltate în cadrul acestei teze au fost publicate în următoarele articole:

1. C. N. Ianăși, V. Gui, C. I. Toma, D. Pescaru, "A fast algorithm for background tracking in video surveillance using nonparametric kernel density estimation", *Facta Universitatis Niš, Series Electronics and Energetics*, Vol. 18, No.1, April 2005, pp. 127-144.
2. C. N. Ianăși, C. I. Toma, V. Gui, D. Pescaru, "Kernel selection for mean shift background tracking in video surveillance", *Proceedings 4th Int. Conference on Microelectronics and Computer Science (ICMCS-05)*, Chișinău, Moldova, Vol. II, September 15-17, 2005, pp. 389-392.
3. C. N. Ianăși, V. Gui, F. Alexa, C. I. Toma, "Fast and accurate background subtraction for video surveillance, using an adaptive mode-tracking algorithm", *WSEAS Int. Conf. on Dynamical Systems and Control (WSEAS 2005)*, Venice, Italy, November 2-4, 2005, pp. 391-397.
4. C. N. Ianăși, V. Gui, F. Alexa, C. I. Toma, "Noncausal, adaptive mode-tracking estimation for background subtraction in video surveillance", *WSEAS Transactions on Signal Processing, Issue 1*, Volume 2, January 2006, pp. 52-59.

CUPRINS

SINTEZA LUCRĂRII	ii
NOTĂ	iv
CUPRINS	v
CAPITOLUL 1. MOTIVAȚIE	1
1.1 Considerații generale asupra supravegherii video	1
1.2 Componentele unui sistem de video-supraveghere	3
1.2.1 Blocul senzorilor	5
1.2.2 Blocul de preprocesare	5
1.2.3 Blocul pentru estimarea fundalului	9
1.2.4 Blocul de segmentare a mișcării	12
1.2.5 Blocul de urmărirea mișcării	14
1.2.6 Blocul de clasificare	15
1.2.7 Blocul de interpretarea informațiilor de mișcare	15
1.3 Actualitatea temei	16
1.4 Structura tezei de doctorat	17
CAPITOLUL 2. TEHNICI DE ESTIMARE A FUNDALULUI	19
2.1 Considerații generale	19
2.2 Metode de detecție a fundalului în secvențe de imagini	21
2.2.1 Modelarea fundalului la nivel de bloc	22
2.2.2 Modelarea fundalului la nivel de pixel	22
2.3 Metode Parametrice de estimare a fundalului	25
2.4 Metode Nonparametrice de estimare a fundalului	27

2.4.1	Tehnici nonparametrice de estimare a densității de probabilitate în spații multidimensionale	27
2.5	Concluzii	38

CAPITOLUL 3. ESTIMAREA NONPARAMETRICĂ RECURSIVĂ A FUNDALULUI

3.1	Adoptarea dimensiunii bufferului de cadre	39
3.2	Estimarea inițială a fundalului	41
3.3	Segmentarea fundal / prim-plan	43
3.4	Estimator nonparametric recursiv	45
3.5	Rezultate experimentale	49
3.5.1	Evaluarea calitativă a estimatorului nonparametric recursiv	49
3.5.2	Evaluarea cantitativă a estimatorului nonparametric recursiv	51
3.5.3	Evaluarea vitezei estimatorului nonparametric recursiv	53
3.6	Concluzii	56

CAPITOLUL 4. OPTIMIZĂRI ALE ESTIMATORULUI NONPARAMETRIC RECURSIV DE FUNDAL

4.1	Tehnici adaptive de estimare a fundalului	59
4.2	Ecuția de actualizare a fundalului	60
4.3	Alegerea nucleului pentru estimarea funcției densitate de probabilitate	62
4.3.1	Rezultate experimentale	63
4.4	Estimator nonparametric recursiv autoadaptiv	66
4.4.1	Adaptarea la schimbări ale fundalului de amplitudine moderată	66
4.4.2	Rezultate experimentale pentru evaluarea performanțelor estimatorului nonparametric recursiv autoadaptiv	70
4.5	Estimator nonparametric recursiv autoadaptiv noncauzal	73
4.5.1	Adaptarea la schimbări drastice ale fundalului	73
4.5.2	Evaluarea performanțelor estimatorului nonparametric noncauzal	75

4.6	Concluzii	77
CAPITOLUL 5. CONTRIBUȚII ȘI CONCLUZII		78
5.1	Contribuții teoretice	79
5.2	Contribuții aplicative	80
5.3	Considerații finale	82
INDEX FIGURI		83
BIBLIOGRAFIE		85

CAPITOLUL 1.

MOTIVAȚIE

Capitolul introductiv al tezei face o prezentare a problematicii generale a supravegherii video și plasează în context esimarea și substrația fundalului, problemă asupra căreia se concentrează cea mai mare parte a contribuțiilor din teză. Capitolul este structurat pe patru paragrafe. Primul paragraf prezintă domeniile de bază ale supravegherii și analizei video inteligente, exemplificând aplicații de actualitate. Componentele unui sistem de supraveghere video general sunt descrise în paragraful 1.2. Totodată se face o prezentare a funcțiilor și metodelor de prelucrare specifice fiecăruia din blocurile componente. Paragraful 1.3 conține argumentează actualitatea temei estimării și substrației fundalului în sistemele moderne de supraveghere. Paragraful final expune și justifică structura tezei de doctorat.

1.1 Considerații generale asupra supravegherii video

Supravegherea video s-a dezvoltat rapid în ultimul deceniu, devenind una din aplicațiile cele mai complexe în prelucrarea numerică a imaginilor. De la simpla detecție pasivă a mișcării în cadrul supravegheat în aeroporturi, bănci, spații de parcare sau amenajări cu destinație militară, care implica operatorul uman în interpretarea evenimentelor și luarea deciziilor adecvate, conceptul de videosupraveghere automată s-a extins la detecția mișcărilor anormale, urmărirea obiectelor și interpretarea evenimentelor. Două articole de sinteză reprezentative în acest sens sunt [1] și [2].

Sistemele de supraveghere tradiționale bazate pe operator uman sunt considerate depășite din perspectiva detectării în timp real a situațiilor de panică. Este binecunoscut faptul că nivelul de atenție al unei persoane scade dramatic în timp, chiar dacă se apelează la un personal special pregătit. Situația este cu atât mai dificilă cu cât

crește numărul de imagini de monitorizat (sistemele complexe multicameră pot avea peste 30 de camere de captură). Ca soluție sunt utilizate în prezent sisteme cu tehnologii de analiză video automate care ajută operatorul uman la detectarea în timp real a situațiilor amenințătoare sau în diverse investigații pe imagini înregistrate.

Direcțiile cheie de studiu în supravegherea și analiza video inteligentă sunt:

- **Detectia și urmărirea video** – o parte semnificativă din tehnologiile de detecție și urmărire video au fost dezvoltate sub un program guvernamental american numit „*Video Surveillance and Monitoring - VSAM*”: tehnologii de bază din detecție, urmarire, autocalibrare, sisteme multicameră, detecție de evenimente. [7], [25], [47], [111].
- **Identificarea video a persoanelor** – este un subiect intens cercetat, recunoașterea feței fiind o modalitate folosită curent pentru aceasta atât în cercetare cât și în sistemele industriale; și pentru acest domeniu există un program guvernamental american recent, numit „*Human ID*” [112].
- **Sisteme de supraveghere pe scară largă** – programul american „*Combat Zones That See*” explorează sistemele cu camere rapid instalabile (*rapidely deployable*) peste rețele wireless ad-hoc și transmit informații de urmărire la stații centrale în scopul monitorizării activităților sau pentru analiza pe termen lung a șabloanelor de mișcare.

Siguranța traficului rutier este pe cale de a fi asistată de sisteme de supraveghere din cele mai sofisticate, responsabile de menținerea traiectoriei în limitele de siguranță pe autostrăzi, șosele și străzi, asistență la parcare, evitarea coliziunii, inclusiv a pietonilor și a obiectelor situate în „unghiul mort”, detecția și recunoașterea semnelor de circulație pentru planificarea și urărirea rutelor, monitorizarea stării de atenție și oboseală a conducătorului auto etc. [3][4][5][6].

Monitorizarea activității umane este obiectivul sistemului de videosupraveghere W^4 dezvoltat de Haritaoglu [7]. Acesta este capabil să analizeze *ce* fac persoanele prezente în imagine, *unde* fac, *când* fac și *cine* face acțiunea în curs de desfășurare.

Numărarea persoanelor [8] este necesară pentru optimizarea comutării semafoarelor, optimizarea funcționării ascensoarelor în clădiri uriașe, de exemplu de tip zgârie-nori, sau în procesul de analiză statistică a timpilor de staționare a clienților în diverse zone ale unor magazine, pentru optimizarea serviciilor. O asemenea aplicație, pentru magazine de tip fast-food, a inspirat și o parte din soluțiile dezvoltate în cadrul tezei. Aplicația prezintă similitudini cu extragerea automată a informațiilor statistice în videosecvențe din sport, subiect de interes mai recent.

Dejucarea acțiunilor teroriste a devenit în ultimii ani o preocupare stringentă [2], ce poate fi asistată eficient prin tehnici de prelucrare numerică a semnalelor video specifice. Acestea includ detecția și urmărirea persoanelor, identificarea și recunoașterea fețelor folosind proiecții multiple, detecția unor atitudini suspecte etc.

Aplicațiile medicale ale supravegherii includ analiza mersului și a tulburărilor de somn [9][10].

1.2 Componentele unui sistem de video-supraveghere

Sistemele moderne de supraveghere video implică tehnici de analiză în timp real a imaginilor pentru o transmisie eficientă a acestora, tehnici de analiză a imaginilor color, tehnici de focalizare a atenției bazată pe evenimente și tehnici de înțelegere a secvențelor bazată pe modele.

Un sistem de supraveghere prelucrează informații furnizate de o rețea de senzori (tipic se folosesc camere de supraveghere în circuit închis CCTV) fixe sau mobili, care funcționează continuu. O schemă bloc generală a unui sistem de videosupraveghere este reprezentată în figura 1.1.

În schema bloc prezentată, camerele statice sunt utilizate pentru acoperirea întregii scene supravegheate și furnizează o imagine globală; camerele PTZ (*Pan-Tilt Zoom*) furnizează informații de detaliu sau de scală fină despre obiectele de interes din scenă. Semnalul video de la camerele statice este folosit pentru detectarea și urmărirea mai multor obiecte din imagine, modelate în două sau trei dimensiuni. Totodată ele mai pot fi utilizate și pentru a furniza informații suplimentare grosiere despre obiecte, cum ar fi clasa obiectului (persoană, mașină, etc.) sau attribute ale obiectului (poziția capului

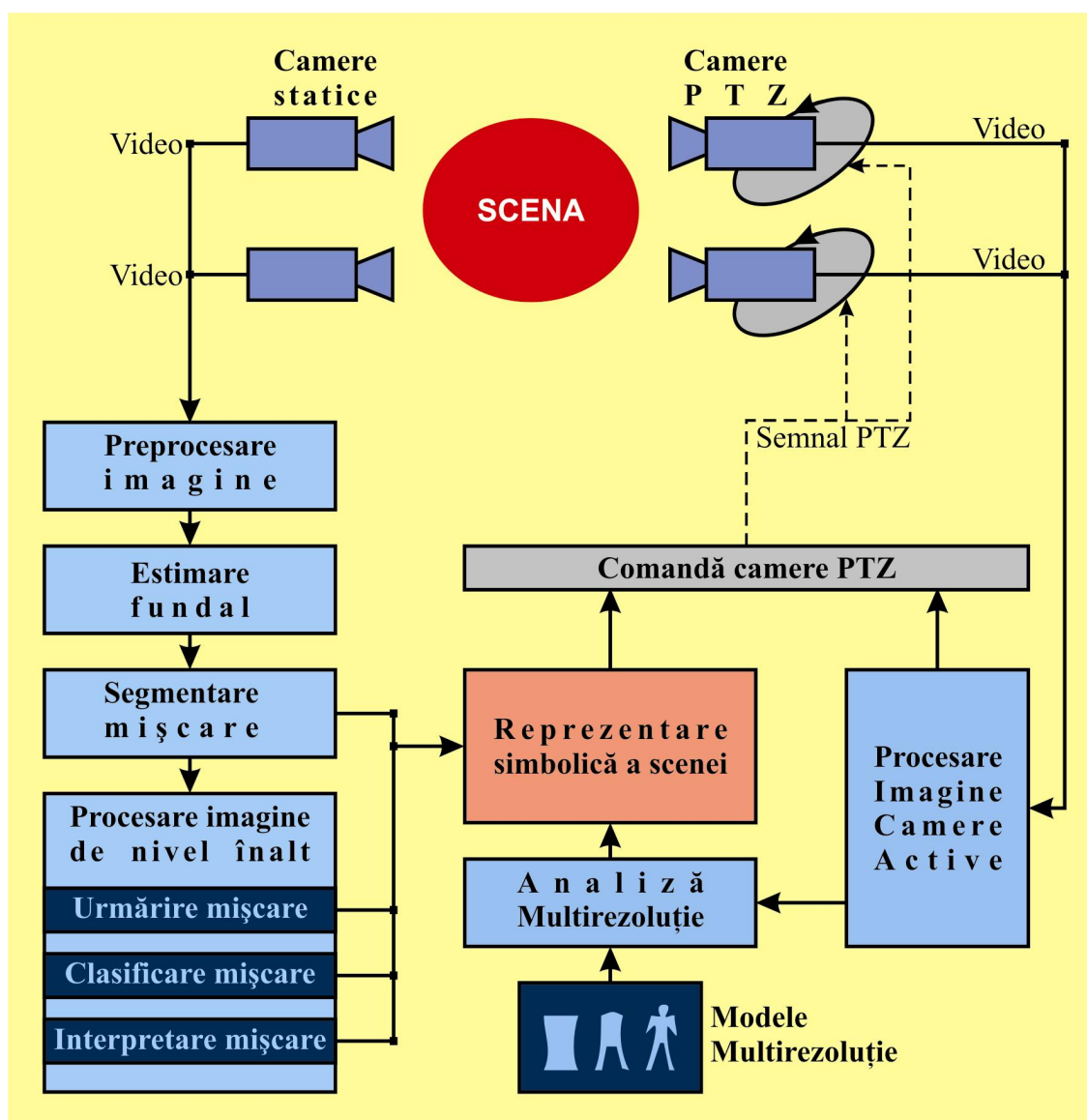


Fig. 1.1. Schema bloc generală a unui sistem de video supraveghere

persoanei, viteza de deplasare a mașinii, etc.). Informația de nivel grosier este folosită ca bază pentru a „focaliza” atenția camerelor PTZ. Imaginile furnizate de acestea sunt utilizate pentru o analiză la o scală fină. De exemplu, dacă o cameră PTZ este orientată către o persoană, analiza de scală fină poate include detecția feței. Informațiile de la analizele de scală grosiere și scală fină sunt combinate în blocul care realizează reprezentarea internă a scenei.

În cele ce urmează vom prezenta global principalele blocuri componente ale schemei bloc și câteva tehnici cheie care guvernează funcționarea lor.

1.2.1 Blocul senzorilor

La cele mai multe din aplicațiile curente în videosupraveghere, sursa principală ce furnizează datele de intrare este o cameră TV color. Sistemele cu camere multiple [8],[11][12][13] pot fi utilizate fie pentru extinderea ariei de supraveghere, fie pentru obținerea de informație de adâncime (3D), de mare utilitate în analiza imaginilor cu obiecte ocluate, sau pentru eliminarea efectelor umbrelor. Alternativ, se pot utiliza camere ultrasonice, camere sensibile în domeniul infraroșu sau, cel mai general, sisteme mixte (multimodale).

De exemplu, în [14] se combină informația de adâncime și de culoare, pentru o prelucrare mai exactă. Este cunoscut faptul că informația de adâncime nu poate fi calculată precis în zonele de imagine lipsite de textură. Inconvenientul poate fi diminuat prin utilizarea combinată a informațiilor de adâncime și de culoare prelevate de la o pereche de camere stereo color. Prelucrarea multimodală a semnalelor se referă în general la combinarea inteligentă a informațiilor referitoare la aceeași scenă, furnizate de senzori diferiți și reprezintă un domeniu de cercetare foarte activ, inclusiv în aplicații de supraveghere [15]-[19].

1.2.2 Blocul de preprocesare

Blocul de preprocesare poate avea, într-un sistem de videosupraveghere, funcții multiple. De exemplu, acest bloc poate fi responsabil de extragerea din fluxul video a imaginilor cu frecvența dorită a cadrelor și la rezoluția cerută de aplicație. Ambele operații presupun în general o re-eșantionare a imaginii [20]. De regulă este o operație de sub-eșantionare, cu toate că se pot semna și tentative de generare de imagini cu super-rezoluție, pornind de la secvențe de imagini. O aplicație interesantă în acest sens o reprezintă reconstrucția și recunoașterea numerelor matricole ale autovehiculelor sau citirea vignetei.

Pentru camerele video color, prelucrarea culorii este o operație la care se recurge frecvent în videosupraveghere. Numeroase studii argumentează faptul că sistemul RGB nu reprezintă întotdeauna cea mai bună opțiune în prelucrarea imaginilor color. Sunt preferate sistemele ce separă componenta de luminanță de cele de crominanță, cum sunt HSV, $L^*a^*b^*$ sau $L^*u^*v^*$ [21]. Conversia de la reprezentarea

RGB la una din reprezentările menționate se realizează în blocul de preprocesare. De exemplu, conversia RGB- $L^*u^*v^*$, folosită și în unele din implementările dezvoltate în cadrul tezei, se realizează prin succesiunea de transformări:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2125 & 0.7154 & 0.0721 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1.1)$$

$$L^* = \begin{cases} 1.116 \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16 & pt. Y > 0.008856Y_0 \\ 903.3 \left(\frac{Y}{Y_0} \right) & pt. Y \leq 0.008856Y_0 \end{cases} \quad (1.2)$$

unde Y_0 este luminanța maximă pentru alb,

$$u' = \begin{cases} \frac{4X}{X + 15Y + 3Z} & pt. X + 15Y + 3Z \neq 0 \\ 4 & pt. X + 15Y + 3Z = 0 \end{cases} \quad (1.3)$$

$$v' = \begin{cases} \frac{9Y}{X + 15Y + 3Z} & pt. X + 15Y + 3Z \neq 0 \\ 9/15 & pt. X + 15Y + 3Z = 0 \end{cases}$$

$$\begin{aligned} u^* &= 13L^*(u' - 0.1978) \\ v^* &= 13L^*(v' - 0.4683) \end{aligned} \quad (1.4)$$

Prima este o transformare liniară ce proiectează culoarea în spațiul de referință XYZ. Transformarea din spațiul XYZ în $L^*u^*v^*$ este neliniară. Avantajul utilizării sistemului $L^*u^*v^*$ constă în faptul că diferențele cromatice percepute de sistemul vizual uman corespund mai exact distanțelor euclidiene din spațiul $L^*u^*v^*$ decât în toate celelalte utilizate curent. De consemnat totuși că și sistemul $L^*a^*b^*$ revendică această calitate, disputa nefiind definitiv tranșată.

Una din modalitățile frecvent adoptate pentru *detecția umbrelor* constă în eliminarea completă a informației de luminanță. De exemplu, sunt folosiți doar coeficienții tricromatici [22].

$$\begin{aligned} r &= \frac{R}{R+G+B}, \\ g &= \frac{G}{R+G+B}, \\ b &= \frac{B}{R+G+B}. \end{aligned} \quad (1.5)$$

Evident,

$$r+g+b=1, \quad (1.6)$$

fiind suficientă folosirea a numai doi coeficienți. Echivalent, în [24] se folosesc componentele u/L și v/L din sistemul Luv.

Eliminarea completă a informației de luminanță în gestionarea fundalului are și unele consecințe nedorite, adesea inacceptabile: face imposibilă discriminarea obiectelor albe, negre sau gri. Mai mult, nuanțele obiectelor foarte întunecate sau foarte luminoase sunt foarte imprecis determinabile din ecuațiile de mai sus, deoarece numitorii $R+G+B$ tind la zero și în consecință zgomotul influențează drastic valorile coeficienților tricromatici. O alternativă mai bună este separarea informațiilor de intensitate și cromaticitate și tratarea lor distinctă, propusă inițial în [25]. De exemplu, se poate folosi spațiul $L^*u^*v^*$, $L^*a^*b^*$, HSV, sau chiar mai simplu, informația de intensitate se poate aproxima prin parametrul:

$$L = \frac{R+G+B}{3} \quad (1.7)$$

sau, echivalent,

$$s = R+G+B. \quad (1.8)$$

O umbră poate fi detectată pe baza unei condiții de forma [23]:

$$\alpha \leq \frac{s}{s_b} \leq \beta, \quad (1.9)$$

unde s și s_b sunt parametrii de intensitate ai pixelului curent, respectiv ai fundalului estimat la coordonatele pixelului curent. Așa cum a fost observat inițial în [26], se poate exploata faptul că umbrele au întotdeauna intensitatea diminuată în raport cu cea de intensitatea de referință a fundalului.

Extragerea muchiilor și a texturilor are numeroase aplicații în analiza imaginilor [27]-[30], putând constitui și o alternativă viabilă pentru eliminarea efectelor schimbărilor iluminării, inclusiv a efectelor umbrelor. Pozițiile liniilor de contur sunt invariante la schimbările iluminării. Un detector de contur simplu și larg utilizat datorită simplității și a rezultatelor meritorii este estimatul de gradient Sobel. Gradientii Sobel orizontali și verticali se pot calcula prin convoluții cu măștile:

$$\mathbf{H}_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad \mathbf{H}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad (1.10)$$

Contururile detectate cu ajutorul operatorilor de tip gradient necesită operații de postprocesare pentru subțiere și închidere. Operatorii pentru detecția muchiilor bazați pe derivate de ordinul 2 exploatează faptul că trecerile prin zero ale derivatei a doua coincid cu maximele derivatelor de ordinul întâi. Mai mult, tind să genereze contururi închise. Sensibilitatea mai mare la zgomot a acestor operatori poate fi redusă prin asocierea cu un operator de netezire. Se obține astfel operatorul laplacianul gaussianului (LoG) [27]:

$$\nabla^2 G(x, y) = \frac{1}{\pi\sigma^2} \left(\frac{x^2 + y^2}{2\sigma^2} - 1 \right) \exp \left\{ -\frac{x^2 + y^2}{2\sigma^2} \right\} \quad (1.11)$$

care stă la baza detectorului de muchii propus de Canny [31].

Deteția umbrelor bazată pe muchii și texturi suferă pe seama faptului că acestea sunt absente în regiunile netede. Rămâne de descoperit cea mai bună modalitate de a incorpora informația referitoare la pozițiile muchiilor în estimarea fundalului, un subiect fără îndoială promițător și insuficient explorat.

1.2.3 Blocul pentru estimarea fundalului

Conceptul de fundal joacă un rol esențial în sistemele de supraveghere contemporane. Pornind de la ipoteza că obiectivul de bază al unui sistem de supraveghere constă în dectecția și caracterizarea *activității* în imagine, că activitatea este indisolubil legată de *mişcare* și că mișcarea este, la rândul ei, indisolubil legată de *schimbare*, se poate concluziona că dectecia activității necesită *deteția schimbării* în imagine. Observând colateral că nu orice schimbare corespunde unei mișcări, schimbarea putând fi și rezultatul schimbării condițiilor de iluminare ale scenei, putem totuși afirma că mișcarea conduce la schimbare în imagine și în consecință dectecția schimbării în imagine este un obiectiv necesar (posibil insuficient) în dectecția mișcării și în ultimă instanță a activității în imagine.

Una din modalitățile cele mai simple care au fost experimentate și utilizate cu un oarecare succes în dectecția schimbării constă în diferența între perechile de cadre succesive, ilustrată în figura 2.1, pentru cazul unei regiuni de culoare constantă ce realizează o mișcare de translație – figura 2.1.(a). Rezultatul diferenței cadrelor succesive este redat în figura 2.1.(b).

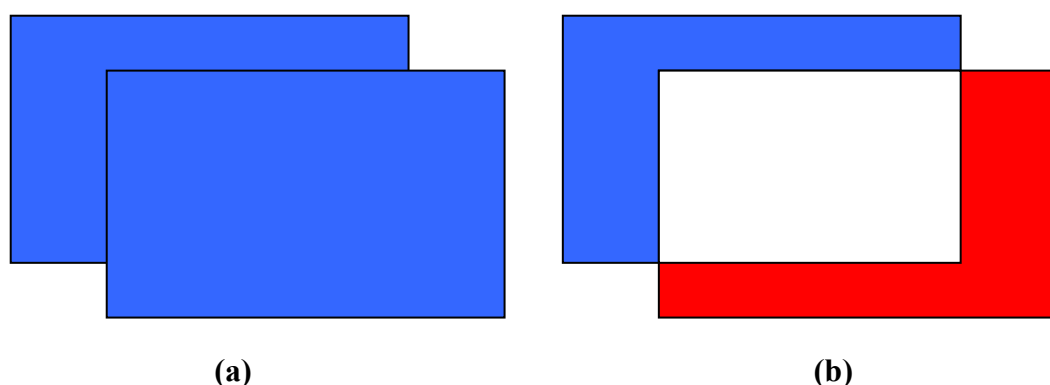


Fig. 1.2. Deteția schimbării prin diferența între cadre succesive:
a) o regiune rectangulară mobilă, de culoare constantă, în două cadre succesive;
b) rezultatul diferenței cadrelor succesive pentru imaginile din figura (a).

Este vizibil dezavantajul metodei: diferența este diferită de zero numai la margini. Mai mult, aceste margini pot fi foarte subțiri dacă viteza de deplasare este redusă, sau chiar lipsi când mișcarea încetează temporar. Dezavantajele menționate pot fi complet eliminate prin adoptarea conceptului de fundal.

Fundalul este prin definiție imaginea *obiectelor statice* din cadrul supravegheat. Deși este generat de obiecte staționare, fundalul nu este pe deplin constant. Una din cauze o constituie schimbările iluminării iar cea de-a doua este mai complicată și vizează generalizarea conceptului de obiect static de la un obiect care nu se mișcă la unul a cărui *poziție medie* este stabilă. Este exemplul tipic al unei frunze în bătaia vântului.

Menționând faptul că discuția la acest subiect este reluată în capitolele dedicate tehnicilor de estimare a fundalului, ilustrăm în figura 1.3 pricipiul detecției mișcării prin extragerea fundalului. Presupunând fundalul cunoscut (albastru deschis în figura 1.3), prin calculul diferenței între cadrul curent și fundalul estimat, obiectul mobil este detectat corect și în întregime, chiar dacă se oprește temporar din mișcare. Desigur definiția noțiunii de „temporar” se cere precizată în sens matematic, cantitativ.

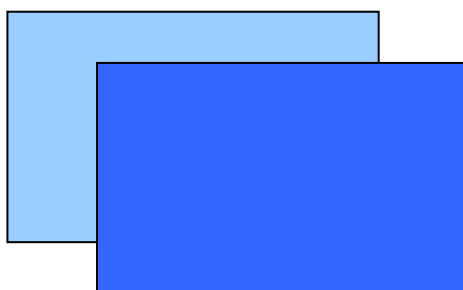


Fig. 1.3. Deplasarea unei regiuni pe un fundal cunoscut și diferit poate fi detectată corect.

Estimarea fundalului se realizează curent la nivel de pixel, folosind statistica lui pe un interval de N cadre (figura 1.4). Alternativ, se poate realiza la nivel de bloc de pixeli, ceea ce nu schimbă în esență datele problemei.

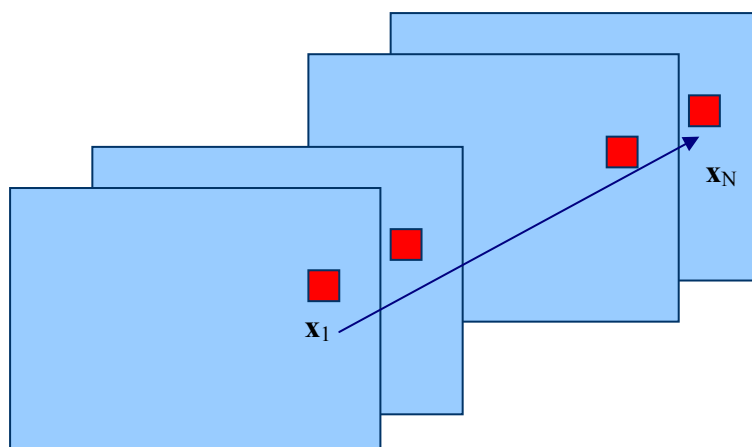


Fig. 1.4. Structura de date folosită la estimarea fundalului: buffer de N cadre, analiză la nivel de pixel (coordonate spațiale constante)

În figura 1.4, $\mathbf{x}_i$ reprezintă un vector al caracteristicilor imaginii, evaluat pentru pixelul marcat din cadrul i . De exemplu, poate fi vectorul de culoare $[R, G, B]^T$. Alegând un interval de observare adecvat, determinat de numărul de cadre N și frecvența de eșantionarea a cadrelor folosite la estimarea fundalului, este rezonabil să presupunem că vectorul caracteristicii observat *cel mai frecvent* la oricare locație din imagine corespunde vectorului caracteristicii fundalului, **b**. Observația conduce la tehnicile de detecție a fundalului bazate pe *estimarea densității de probabilitate*. **Teza prezentă este orientată spre dezvoltarea acestor tehnici.**

Problema estimării funcției (continue) a densității de probabilitate pe baza unui set finit de eșantioane observate de care se dispune este similară problemei interpolării în spații multidimensionale, fiind ilustrată 1D în figura 1.5. Prezentarea suportului teoretic aferent și stadiului curent al cercetării în estimarea densității se tratează în capitolul următor.

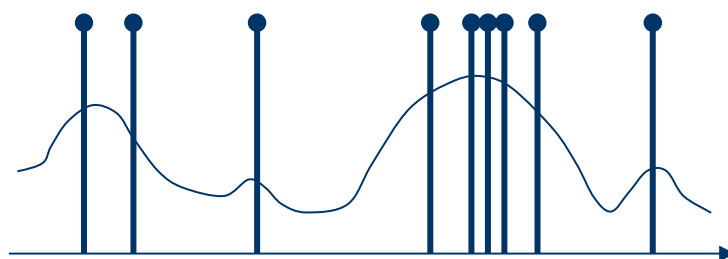


Fig. 1.5. Exemplu de estimare a funcției densitate de probabilitate pe baza unui set finit de eșantioane observate.

Revenind la schema bloc din figura 1.1, blocul de estimare a fundalului exploatează diferența între culoare, textură și mișcare în imaginile captate. În figura 1.6 sunt prezentate stadiile cheie în extragerea fundalului. Pentru detecția obiectelor în imagini aglomerate este necesară utilizarea unor tehnici de îmbunătățire a detecției. Blocul de estimare a fundalului conține un număr de mecanisme capabile să reacționeze la schimbările de ambient și la schimbările de structură ale scenei. El este în măsură să compenseze schimbarea nivelului de iluminare a scenei, să estimeze și să corecteze modificările imaginii induse de circuitele de AGC (*automatic gain control*) și de AWB (*automatic white balance*) ale camerelor CCTV. Totodată acest modul menține o hartă a regiunilor cu activitate (mișcare) intensă din scenă și actualizează încet modelul fundalului estimat doar în acele zone în care activitatea este redusă.

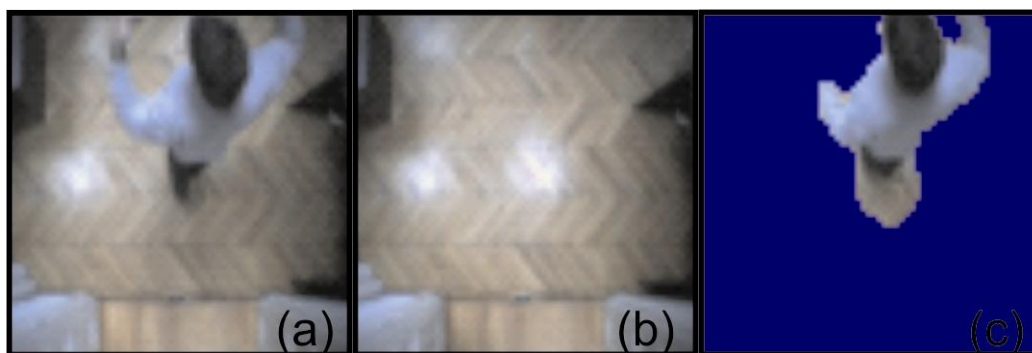


Fig. 1.6. Fazele extragerii fundalului. (a) imaginea originală; (b) fundalul extras; (c) mișcarea din imaginea originală.

În general, estimarea și extragerea fundalului este o tehnică mai potrivită pentru medii interioare, în care iluminarea generală a scenei este relativ stabilă iar mișcările deranjante sunt limitate.

1.2.4 Blocul de segmentare a mișcării

Segmentarea imaginilor sau a videosecvențelor [27], [61],[62], presupune partiționarea imaginii sau videosecvenței într-o serie de regiuni 2D sau 3D. Ideal, regiunile corespund obiectelor semantice definite. Datorită complexității și a naturii inverse a problemei, segmentarea produce de cele mai multe ori doar un set de regiuni caracterizate prin uniformitatea unor caracteristici evaluate la nivelul fiecărui pixel și grupate într-un vector al caracteristicilor. Situația este ceva mai avantajoasă în cazul

videosecvențelor, unde se pot exploata avantajos informații de mișcare ce pot fi extrase, spre deosebire de cazul segmentării imaginilor 2D.

Dacă se dispune de o estimată bună a fundalului, segmentarea mișcării se poate realiza relativ simplu, prin operația desemnată frecvent ca extragere a fundalului. Extragerea fundalului poate fi abordată teoretic ca o problemă de clasificare binară. Dat fiind un vector caracteristic $\mathbf{x}$ al unui pixel din cadrul curent prelucrat și cunoscând caracteristicile fundalului estimat la locația respectivă, de exemplu vectorul $\mathbf{b}$, este necesară o decizie cu privire la apartenența pixelului la fundal sau în caz contrar la prim-planul considerat clasa cu mișcare. Decizia pentru clasa prim-plan se face prin excluderea din clasa fundal, pentru motivul că fundalul este mult mai constant și predictibil decât prim-planul.

Un exemplu de segmentare a mișcării într-o aplicație de videosupraveghere se găsește în figura 1.7. Clasa fundal este marcată cu albastru, iar clasa prim plan cu galben. Este trasat, suplimentar, dreptunghiul de încadrare al obiectului (subiectului) mobil extras.

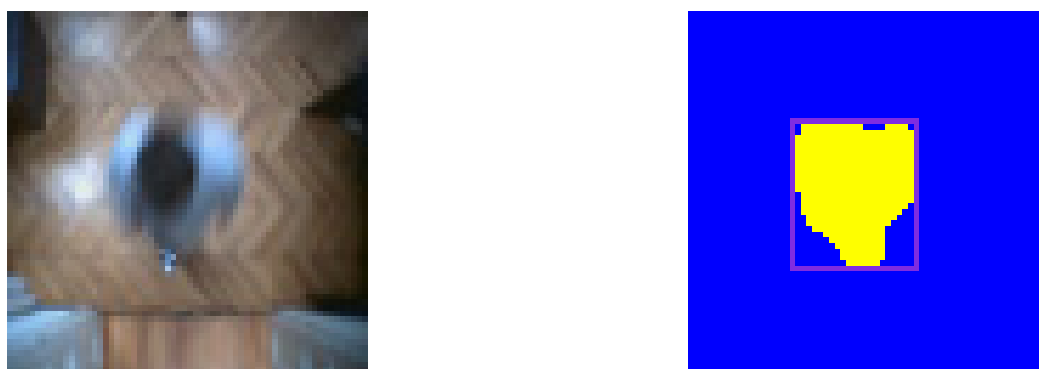


Fig. 1.7. Exemplu de segmentare a obiectului mobil (prim plan) într-o secvență de imagine, prin extracția fundalului estimat.

În general segmentarea mișcării este o abordare complementară extragerii fundalului. Problema poate fi abordată din perspectiva unei filtrări a mișcării. În figura 1.8.a este prezentată o scenă cu o persoană ce se deplasează în fața unui tufiș cu frunze ce se mișcă în bătaia vântului. O tehnică tradițională de extragere a fundalului ar avea ca rezultat imaginea din figura 1.8.b, în care practic întregul tufiș este (pe bună dreptate) clasificat ca și obiect în mișcare. În mod evident însă rezultatul obținut nu este și ceea ce se dorește de la un astfel de sistem, adică detecția persoanei ce se deplasează în fața

tufișului, deci este necesară implementarea unor tehnici de prelucrare mai complicate. O soluție este prezentată în [110] și folosește câmpul optic de mișcare (*optical flow*) în detecția mișcării dominante.

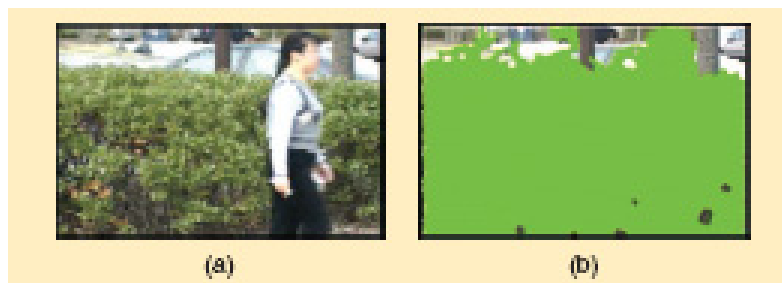


Fig. 1.8. Detecția defectuoasă a mișcării cu o tehnică standard de extragere a fundalului, într-o secvență limită (tufiș cu frunze mișcate de vânt)

1.2.5 Blocul de urmărirea mișcării

Urmărirea mișcării (*tracking*) a este un subiect de cercetare activă al ultimelor două decenii. Problema a fost abordată din perspective diverse, o prezentare sistematică și exhaustivă a subiectului fiind dificilă și în afara scopului acestui paragraf. Se presupune că obiectul urmărit a fost inițial detectat și se dispune de un prim model al acestuia. Una din primele probleme ce trebuiesc rezolvate este de a decide ce caracteristici ale obiectului vor fi utilizate în urmărire. O discuție interesantă pe acest subiect se găsește în [32]. Caracteristicile alese trebuie să fie *invariante* la modificările așteptate ale obiectului. O soluție frecvent adoptată este urmărirea pe bază de histogramă a caracteristicilor de culoare sau/și textură [33]-[35]. Histograma este invariantă la translație, rotație, chiar rescalare sau deformări moderate. Alternativ, se pot folosi contururile regiunii [36] sau forma ei [37]. Tehnicile de urmărire bazate pe regiuni segmentează obiectul urmărit și rezolvă cadru cu cadru problema corespondenței regiunilor [38]-[40]. Tehnicile bazate pe model [24], [41]-[44], exploatează informații apriorice referitoare la particularitățile obiectului sau obiectelor urmărite, concentrate într-un model al formei, siluetei etc.

Având în vedere caracterul specializat al unor asemenea soluții, se pot folosi cu succes tehnici de învățare în proiectarea sistemului de urmărire [45],[46]. Sistemul de urmărire necesită un anumit tip de filtrare, pentru a face față modificărilor obiectului urmărit, ocluziilor parțiale, sau chiar dispariției temporare prin ocluzie. Printre soluțiile

adoptate se remarcă folosirea tehnicilor bazate pe funcții nucleu [33], filtrul Kalman [24], [47],[48], sau mai modern, filtrul cu particule [49]-[53].

1.2.6 Blocul de clasificare

În unele aplicații de supraveghere video este critică determinarea tipului obiectului detectat. În literatură sunt întâlnite două abordări pentru clasificarea obiectelor: cea bazată pe imagine (*image-based*) și cea bazată pe urmărirea video (*video tracking-based*). Sistemele cu urmărire video (*video tracking-based*) utilizează statistici despre apariția, forma și mișcarea obiectelor pentru a distinge rapid între oameni, animale, vehicule, uși, pomi în bătaia vântului, etc. Clasificarea obiectelor se aplică tuturor obiectelor selectate în blocul de urmărire și le alocă acestora etichete de apartenență la clasa corespunzătoare.

1.2.7 Blocul de interpretarea informațiilor de mișcare

Rolul blocului de interpretare a informațiilor de mișcare extrase în blocul de urmărire este mai pronunțat dependent de aplicație decât al blocurilor precedente. Aici sunt sintetizate date statistice pe baza formelor, pozițiilor, vitezelor sau traiectoriilor obiectelor urmărite și se emit eventuale decizii de acțiune. O dată separate de fundal, obiectele pot fi *clasificate*. Introducerea și abandonarea unui bagaj în cadrul supravegheat (posibil atac terorist) sau scoaterea unui obiect din cadrul supravegheat (furtul unui tablou din muzeu) pot fi *detectate*. Sistemele de videosupraveghere contemporane au nu numai un rol pasiv de înregistrare de evenimente ci și unul *activ*, de interacțiune. Cel mai simplu constă în comanda camerei (zoom, tilt, pan), care devine activă (active camera) și inteligentă (smart camera). O aplicație în care interpretarea mișcării joacă un rol important și dificil este interpretarea gesturilor [54],[55]. Recunoașterea persoanelor, respectiv a tipului de activitate umană revine, de asemenea blocului de interpretare și este descrisă în [47], [56],[57].

1.3 Actualitatea temei

Estimarea și substrația fundalului reprezintă o verigă de importanță critică în aplicațiile de videosupraveghere. De succesul acestei operații depind într-o mare măsură rezultatele prelucrării la nivel înalt. Varietatea condițiilor în care trebuie să funcționeze sistemele de supraveghere introduce dificultăți majore în proiectarea sistemelor de estimare a fundalului. Un sistem performant trebuie să funcționeze precis în condițiile normale. Totodată, sistemul trebuie să reacționeze bine și la condiții neprevăzute, pe care proiectantul sistemului nu le-a avut în vedere, în sensul că este de dorit ca degradarea performanțelor să fie cât mai moderată cu putință. O asemenea calitate este denumită în literatura de specialitate *robustețe*. Preocuparea pentru *soluții robuste* în vederea artificială [58] s-a dezvoltat rapid în ultimul deceniu. Una din direcțiile prin care se obțin asemenea soluții constă în adoptarea unei modelări pe bază de *statistici robuste*, un domeniu modern al matematicii. Explorarea acestor soluții în estimarea fundalului era la debut în momentul adoptării temei de doctorat și rămâne actuală la finalizarea ei, așa cum o atestă numărul mare de articole ce continuă să fie publicate în literatură pe acest subiect.

Estimarea fundalului este una din operațiile cele mai consumatoare de timp în videosupraveghere, dat fiind faptul că prelucrarea are loc la nivelul cel mai de jos, cel de pixel. Mai mult, există un interes major pentru sisteme cu funcționare în timp real. Am menționat astfel două dintre motivele ce au impulsivat cercetări pentru algoritmi de prelucrare eficientă [59][60].

Teza este strâns legată de activitatea de dezvoltare a unui sistem de videosupraveghere destinat să genereze date statistice concludente pentru analiza

comportamentului consumatorilor în localuri de tip fast-food. Lucrarea prezintă doar rezultatele inovatoare din cadrul acestei activități, referitoare la estimarea fundalului prin tehnici robuste, nonparametrice. Soluțiile propuse în lucrare au permis concomitent reducerea erorilor de estimare, creșterea adaptibilității și reducerea complexității de calcul.

1.4 Structura tezei de doctorat

Lucrarea prezentă este structurată pe cinci capitole.

Capitolul 1 prezintă sintetic problematica sistemelor de supraveghere moderne, principalele blocuri componente ale unui sistem de videosupraveghere, din punct de vedere conceptual, fără referire la structura hardware. Este prezentat pe scurt rolul fiecăreia din componentele sistemului, între care se numără și cel pentru estimarea și subtracția fundalului. Tot în acest capitol sunt expuse actualitatea problematicii detecției și extragerii fundalului în supravegherea video și motivele care au condus la abordarea temei de cercetare ce face obiectul tezei.

Capitolul 2 prezintă stadiul curent al cercetării în problema estimării fundalului pentru aplicații de videosupraveghere. Se desprind două direcții de lucru predominante. Cea tradițională, bazată pe metode de estimare parametrice și o direcție nouă, bazată pe metode de estimare nonparametrice. În acest capitol este prezentat suportul matematic ce stă la baza tehnicilor de estimare parametrice și nonparametrice, cu referire specifică la dezvoltări mai recente privind localizarea modurilor funcției densitate de probabilitate prin algoritmul cu deplasare la medie (*mean-shift*).

Capitolul 3 are caracter de noutate și este prezentată o metodă de timp real, propusă de autor, bazată pe estimare nonparametrică, cu actualizare recursivă a fundalului estimat, ce include un test de plauzibilitate pentru accelerarea calculelor. Metoda îmbină avantajele tehnicilor de estimare nonparametrică a funcției densitate de probabilitate (FDP) bazată pe funcții nucleu, cu viteza tehnicilor bazate pe histograme. Performanțele metodei de estimare și urmărire a

fundalului propusă au fost evaluate calitativ și cantitativ, fiind comparate cu rezultatele metodelor de referință prezentate în literatură.

Capitolul 4 are, de asemenea, un caracter de noutate și prezintă soluții elaborate în cadrul cercetării curente. În prima parte se prezintă un studiu comparativ al funcțiilor nucleu asupra rezultatelor estimatorului cu urmărire pentru estimarea nonparametrică a densității, motivat de faptul că estimatorul cu urmărire are o comportare diferită atât față de estimatorii nonparametrici tradiționali cât și față de cei parametrici. În continuare sunt prezentate optimizări ale estimatorului de fundal cu urmărire la schimbări de iluminare ale scenei. Sunt avute în vedere numai soluții ce lucrează la nivel de pixel și vizează atât schimbările rapide și de amplitudine moderată ale iluminării scenei cât și schimbările drastice de conținut ale fundalului. Efectul favorabil al perfecționărilor aduse estimatorului sunt evaluate experimental și prezentate în finalul capitolului

Capitolul 5 face o sinteză a principalelor contribuții teoretice și a rezultatelor experimentale obținute în cadrul cercetării expuse în teză.

CAPITOLUL 2.

TEHNICI DE ESTIMARE A FUNDALULUI

Capitolul este structurat pe cinci paragrafe. În primul paragraf sunt definite și discutate principalele dificultăți cu care se confruntă estimarea fundalului în aplicațiile curente. Stadiul curent al cercetării în problema estimării fundalului pentru aplicații de supraveghere video este prezentat în paragraful 2.2. Se desprind două direcții de lucru predominante. Cea tradițională, bazată pe metode de estimare parametrice și o direcție nouă, bazată pe metode de estimare nonparametrice. În paragraful 2.3 se tratează fundamentele teoretice ale metodelor de estimare parametrice, cu referire la modelul larg adoptat în literatură al sumei de gaussiene . În paragraful 2.4. sunt prezentate metodele de estimare nonparametrice. Se include o discuție referitoare la algoritmul cu translație la medie (*mean shift*) pentru detecția rapidă a modurilor funcției densitate de probabilitate (FDP). În paragraful 2.5 sunt trecute în revistă concluziile rezultate din studiul prezentat în acest capitol.

2.1 Considerații generale

Detecția și substrația fundalului reprezintă un pas de importanță crucială pentru segmentarea și urmărirea automată a obiectelor mobile în aplicațiile de supraveghere video, precum și în realizarea sistemelor de interacțiune om-mașină de calcul inteligente [7], [24], [78]. De cele mai multe ori, camerele de luat vederi utilizate în aplicațiile menționate sunt statice. În consecință, este mai ușoară detecția unui fundal staționar decât a unor obiecte mobile. Cu toate acestea, detecția fundalului rămâne o problemă incomplet rezolvată în prezent. Există numeroase lucrări ce prezintă rezultate excelente în condițiile de laborator în care au fost proiectate și testate. Soluțiile propuse sunt încă perfectibile atunci când sunt confruntate cu varietatea situațiilor ce apar în aplicațiile reale.

Prin definiție, **fundalul** este o regiune mult mai stabilă decât prim-planul, aflat în mișcare. Este însă departe de a fi constant. Câteva din cele mai importante surse ale instabilității fundalului sunt examinate în cele ce urmează. Înțelegerea lor este importantă pentru adoptarea ipotezelor de lucru corecte și a modelelor matematice adecvate pentru estimarea și detecția fundalului.

Schimbări ale iluminării pot fi provocate de aprinderea sau stingerea unor surse de lumină la scene de interior, respectiv de evoluția soarelui, norilor și precipitațiilor la scene de exterior.

Umbre și reflexii pot fi provocate de obiectele mobile, afectând aspectul fundalului. La imagini cu contrast puternic, umbrele și/sau reflexiile pot afecta iremediabil informația de culoare. Obiectele umbrite pot deveni cvasi-negre, estimarea nuanței devenind extrem de sensibilă la zgomot.

Reflexii pe suprafețe lucioase pot provoca saturația senzorului de imagine și apariția unor zone apropiate de alb, cu nuanțe dificil de estimat corect din informația *RGB* și adesea dependentă mai mult de anumite limitări ale senzorilor decât de culorile inițiale din care au rezultat. Obiecte mobile luminoase pot reflecta în fundal o parte din lumina incidentă pe ele, dar asemenea modificări ale fundalului sunt mai puțin drastice și pot fi relativ ușor tolerate de metodele de detecție a fundalului utilizate curent. Detecția umbrelor cu contrast moderat este, de asemenea abordabilă cu relativ succes, deși rămâne o problemă deschisă [79].

Introducerea sau scoaterea unor obiecte din fundal determină modificarea acestuia. În primă instanță, un obiect deplasat din fundal va fi detectat ca prim-plan mobil, în timp ce regiunea descoperită va da naștere unei fantome – un fals obiect de prim-plan, static. Decizia de a modifica modelul fundalului pentru eliminarea fantomei va trebui amânată un anumit timp, pentru a nu se dizolva în fundal și obiectele ce se opresc temporar din mișcare.

Camuflajul se manifestă prin apariția în cadrul imaginii a unor obiecte mobile de culoare foarte apropiată de cea a fundalului. Există riscul ca asemenea obiecte să rămână nedetectate sau detecția să fie imprecisă și nestabilă.

Fundalul dinamic poate fi generat de obiecte de genul unui ventilator în funcțiune, afișaj digital, televizor sau obiecte flexibile aflate în bătaia vântului (ramuri

de copac, frunze etc.). Dacă fundalul real este acoperit în majoritatea timpului de obiecte mobile, în condiții de trafic extrem de aglomerat, detecția lui este semnificativ îngreunată.

Unele din problemele menționate pot fi adresate și rezolvate folosind tehnici de prelucrare de mare complexitate, ce pot fi însă prohibitive atunci când este vorba de o aplicație de timp real. La unele aplicații de supraveghere video, prelucrarea datelor off-line poate fi acceptabilă, în timp ce la altele se poate accepta o anumită elasticitate în timpul de răspuns. Pentru aplicații de interacțiune inteligentă om-calculator, reacția sistemului trebuie să fie de ordinul fracțiunilor de secundă pentru ca întârzierea să nu fie deranjantă. În linii mari, majoritatea aplicațiilor trebuie să poată ține pasul cu viteza de succesiune a cadrelor, timpul de prelucrare fiind, prin urmare, critic. Complexitatea calculelor poate fi mai redusă și șansele de detecție corectă mai mari, dacă se apelează la metode de prelucrare multimodale [14], [80], ce fac apel simultan la date de la mai multe tipuri de senzori: sisteme de stereoviziune sau multicameră, camere de luat vederi cu funcționare în infraroșu sau ultrasonice etc.

Menționând că prelucrarea multimodală a semnalelor este un domeniu nou cu dezvoltare rapidă, studiul prezent se limitează totuși la sisteme cu o cameră de supraveghere staționară.

2.2 Metode de detecție a fundalului în secvențe de imagini

Obiectivul unui sistem de supraveghere video este de a monitoriza activitatea într-o zonă specificată, situată într-o clădire sau în aer liber.

Presupunând camerele de supraveghere staționare, o metodă eficientă de detecție a obiectelor mobile constă în compararea fiecărui cadru de imagine nou cu un cadru de referință, reprezentând în cel mai fidel mod posibil fundalul. Prin eliminarea regiunilor în care cadrul curent se aseamănă suficient de bine cu cadrul de referință, se obține o segmentare rapidă a obiectelor mobile. Rezultatele acestui proces sunt desemnate în literatură prin termenul de substrație a fundalului și sunt utilizate de regulă de un modul de prelucrare la nivel mai înalt, responsabil cu urmărirea obiectelor mobile, analiza mișcării și interpretarea scenei. Este ușor de anticipat că erorile posibile

în substrația fundalului au un impact major asupra veridicității rezultatelor interpretării furnizate de modulul de prelucrare de nivel superior. În consecință, problema modelării și detecției fundalului în secvențe de imagini a fost analizată în profunzime de numeroși cercetători, existând numeroase soluții propuse, cu avantaje și dezavantaje specifice.

În cele ce urmează, se prezintă într-o succesiune progresivă, din punctul de vedere al complexității de prelucrare, principalele realizări din domeniu. Se remarcă două direcții de abordare predominante, bazate respectiv pe:

- tehnici de estimare parametrice și
- tehnici de estimare nonparametrice.

Prezentarea metodelor de estimare parametrice este inclusă în capitolul curent, în timp ce metodelor de estimare nonparametrice, de interes mai mare pentru doctorand, le este consacrat capitolul următor. Modelarea fundalului poate fi concepută la nivel de bloc sau la nivel de pixel.

2.2.1 Modelarea fundalului la nivel de bloc

Modelarea la nivel de bloc este specifică lucrărilor mai vechi, dar nu este complet abandonată [81]. Avantajele principale ale abordării la nivel de bloc sunt viteza de prelucrare și stabilitatea superioare. Ele se obțin însă în detrimentul rezoluției, ceea ce nu pentru toate aplicațiile reprezintă un compromis favorabil.

2.2.2 Modelarea fundalului la nivel de pixel

Modelul fundalului pentru un pixel se construiește pe baza unei mulțimi de vectori eșantion colectați într-un număr de cadre succesive. Componentele vectorilor reprezintă caracteristici măsurate la nivel de pixel (intensitate, culoare) sau la nivel local, într-o vecinătate a pixelului modelat (de exemplu caracteristici de textură, muchii, disparitate). Într-o situație ideală, fundalul la orice locație din cadru poate fi considerat constant, dar necunoscut. Estimarea fundalului are ca obiectiv determinarea vectorului caracteristicilor pentru fiecare locație. Vectorul caracteristicilor observat într-un cadru index k , se poate modela în forma:

$$\mathbf{x}_k = \mathbf{b} + \mathbf{n}_k \quad (2.1)$$

unde $\mathbf{b}$ este vectorul constant și necunoscut al caracteristicilor fundalului la locația observată și $\mathbf{n}_k$ este vectorul zgomotului la momentul observației, k . Notăm cu N numărul de cadre disponibil pentru estimarea fundalului, $\mathbf{b}$. Problema de estimare poate fi scrisă în forma:

$$\mathbf{b} = \arg \min_{\mathbf{y}} \{\varepsilon^2(\mathbf{y})\} \quad (2.2)$$

cu

$$\varepsilon^2(\mathbf{y}) = \sum_{k=0}^{N-1} \|\mathbf{y} - \mathbf{x}_k\|^2 = \sum_{k=0}^{N-1} \|\mathbf{n}_k\|^2 \quad (2.3)$$

care se recunoaște a fi estimatorul cu abatere pătratică minimă. El minimizează suma pătratelor distanțelor euclidiene dintre estimată și caracteristicile observate, mai precis, norma L^2 a zgomotului. Soluția se obține impunând anularea derivatei de ordinul întâi a erorii în raport cu vectorul estimat căutat:

$$\frac{\partial \varepsilon^2(\mathbf{y})}{\partial \mathbf{y}} = \frac{\partial \varepsilon^2}{\partial \mathbf{y}} \sum_{k=0}^{N-1} (\mathbf{y} - \mathbf{x}_k)^T (\mathbf{y} - \mathbf{x}_k) = 2 \sum_{k=0}^{N-1} (\mathbf{y} - \mathbf{x}_k) = 0 \quad (2.4)$$

Rezultă imediat că:

$$\mathbf{b} = \mathbf{y} = \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{x}_k \quad (2.5)$$

Prin urmare, vectorul caracteristicilor de fundal estimat este media aritmetică a eșantioanelor observate în cele N cadre de imagine utilizate pentru estimare.

Cea mai convenabilă modalitate de a pune în practică o asemenea estimare ar fi să se capteze un număr de N cadre cu scena goală (liberă de obiecte mobile). Din păcate, un asemenea scenariu arareori poate fi adoptat în aplicațiile reale, cum ar fi supravegherea unui terminal de aeroport, a unei stații de metro sau a unei autostrăzi. Mai mult, ori de câte ori fundalul s-ar schimba, procedura ar trebui reluată și scena eliberată. În mod evident, avem nevoie de o metodă de estimare capabilă să tolereze bine prezența activității în imagine și în același timp să urmărească schimbările fundalului pe parcursul derulării programului de supraveghere video.

Prezența obiectelor mobile pe durata procesului de estimare a fundalului implică prezența potențială a unor pixeli în ecuațiile (2.1) – (2.4) ce se abat drastic de la fundalul real. Denumite eșantioane aberante sau puternic deviate (outliers) în literatura referitoare la statistici robuste [58], [82], asemenea eșantioane influențează excesiv valoarea estimatei, tinzând să aibă o pondere dominantă în eroarea minimizată de estimată, datorită operației de ridicare la pătrat ce intervine în ecuația (2.3), de definiție a erorii. Efectul eșantioanelor aberante asupra estimatei poate fi redus semnificativ dacă în locul normei L^2 se folosește norma L^1 în ecuațiile (2.2) și (2.3):

$$\varepsilon(\mathbf{y}) = \sum_{k=0}^{N-1} \|\mathbf{y} - \mathbf{x}_k\| = \sum_{k=0}^{N-1} \sqrt{(\mathbf{y} - \mathbf{x}_k)^T (\mathbf{y} - \mathbf{x}_k)} \quad (2.6)$$

În acest caz, estimata minimizează suma distanțelor la eșantioanele observate și reprezintă mediana eșantioanelor utilizate la estimare. De remarcat costul de calcul considerabil mai mare în cazul medianei față de media aritmetică, în special pentru date vectoriale.

Complexitatea de calcul poate fi redusă semnificativ folosind mediana scalară pentru fiecare componentă a vectorului caracteristic, dar rămâne mult mai mare decât la media aritmetică. Calculul direct al medianei scalare conduce la complexitate de ordinul $O(N^2)$, dar există totuși soluții mult mai rapide bazate pe histogramme sau algoritmi de ordonare. Din păcate, tehnicile rapide bazate pe sortare nu pot fi extinse simplu pentru date vectoriale, datorită dificultăților legate de definirea relației de ordine.

Filtrul median are o proprietate remarcabilă, ce evidențiază gradul lui înalt de imunitate la eșantioane puternic deviate. Dacă 49 % din eșantioane sunt aberante și celelalte 51 % sunt corecte, filtrul median încă găsește valoarea corectă. În problemele de estimare a fundalului, condiția menționată este echivalentă cu aceea ca fundalul să fie vizibil cel puțin pe 51 % din numărul cadrelor folosite la estimare. Un trafic intens poate invalida condiția menționată. Un model mai realist trebuie să accepte existența unui număr mai mare de obiecte ce ocupă temporar poziții în zona pixelului pentru care se face estimarea fundalului, în intervalul celor N cadre folosite pentru estimare.

În consecință, în mulțimea eșantioanelor, se formează distribuții complexe, formate din grupuri de distribuții corespunzătoare obiectelor prezente. O ipoteză de cele

mai multe ori rezonabilă, cu condiția alegerii unui număr de cadre de estimare, N , suficient de mare este aceea că grupul de eşantioane corespunzător fundalului este cel mai numeros și mai strâns grupat în jurul modului distribuției. Echivalent, ne așteptăm ca densitatea de probabilitate a eşantioanelor în jurul modului distribuției fundalului să fie mai mare decât la orice altă distribuție, corespunzătoare unui obiect mobil. Se desprinde astfel concluzia că un model mai realist al fundalului este acela al unui vector aleatoriu, având o anumită funcție densitate de probabilitate, al cărei mod este localizat în punctul din spațiul eşantioanelor corespunzător caracteristicilor fundalului.

În unele cazuri, ce apar la aplicații de genul celor conținând un ventilator în funcțiune în fundal, poate fi necesară modelarea fundalului folosind două sau mai multe distribuții. Rămâne valabilă observația că decizia dacă un pixel având vectorul caracteristic $\mathbf{x}$ din cadrul curent aparține sau nu fundalului poate fi luată prin compararea densității de probabilitate a fundalului la locația $\mathbf{x}$, dacă aceasta este sau nu mai mare decât a oricăreia din distribuțiile existente în grupul eşantioanelor utilizate în estimarea fundalului.

2.3 Metode Parametrice de estimare a fundalului

Funcțiile densitate de probabilitate utilizate în estimarea fundalului pot fi reprezentate parametric, presupunându-se a avea forme cunoscute, parametrii urmând a fi estimați astfel încât distribuția parametrică să descrie cât mai exact eşantioanele observate. Estimarea fundalului devine astfel o problemă de estimare parametrică a densității de probabilitate, ce a fost temeinic studiată în literatura statistică. De o maximă popularitate se bucură modelul mixturii de gaussiene adoptat inițial de Stauffer și Grimson [47], [83] și urmat de mulți alții, de exemplu [80], [84], [85], [86]. Avantajul principal al modelelor parametrice este acela că pot fi folosite fără a fi necesară stocarea unui număr important de cadre. În mod tipic, trei până la cinci componente gaussiene sunt necesare în mixtură pentru modelarea adecvată a fundalului și obiectelor mobile pentru fiecare pixel. Gaussienele sunt ponderate de numărul de pixeli prezente în fiecare, prin parametri π_k , reprezentând probabilitățile apriorice ale claselor:

$$\hat{p}(\mathbf{x}) = \sum_{k=1}^K \pi_k G_k(\mathbf{x}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (2.7)$$

unde $\mathbf{x}$ este vectorul caracteristic, π_k – probabilitățile apriorice ale gaussianelor, iar $\boldsymbol{\mu}_k$ și $\boldsymbol{\Sigma}_k$ sunt vectorii medie și respectiv matricile de covarianță ale distribuțiilor componente. Pentru reducerea complexității de calcul, componentele vectorului caracteristicilor sunt uzual considerate independente. Independența caracteristicilor asigură diagonalizarea matricii de covarianță în ecuația (2.7).

Estimarea parametrilor gaussianelor se poate obține folosind algoritmul EM (expectation-maximization) introdus de Dempster [87]. Pentru aplicații de timp real, algoritmul EM original este prohibitiv din punctul de vedere al timpului de calcul. În consecință, se utilizează o aproximare, bazată pe versiunea on-line a algoritmului mediilor (K-means). Pasul E clasifică un nou vector $\mathbf{x}_t$ din cadrul curent, indice t , într-una din distribuțiile componente, prin maximizarea expectanței:

$$k = \underset{j}{\operatorname{argmax}} \{ \pi_j G_j(\mathbf{x}_t; \boldsymbol{\mu}_{j,t-1}, \boldsymbol{\Sigma}_{j,t-1}) \} \quad (2.8)$$

Formal, clasificarea se poate exprima cu ajutorul variabilelor L_k , cu $k=1,2,\dots,K$ și $L_k = 1$ dacă $\mathbf{x}_t$ aparține clasei k și $L_k = 0$ pentru toate celelalte clase. În pasul M, se actualizează parametrii mixturii, prin ecuațiile:

$$\begin{aligned} \boldsymbol{\mu}_{k,t} &= (1 - L_k \alpha) \boldsymbol{\mu}_{k,t-1} + \alpha L_k \mathbf{x}_t \\ \boldsymbol{\Sigma}_{k,t} &= (1 - L_k \alpha) \boldsymbol{\Sigma}_{k,t-1} + \alpha L_k \operatorname{diag}((\mathbf{x}_t - \boldsymbol{\mu}_{k,t-1})^T (\mathbf{x}_t - \boldsymbol{\mu}_{k,t-1})) \\ \pi_{k,t} &= (1 - L_k \alpha) \pi_{k,t-1} + \alpha L_k \end{aligned} \quad (2.9)$$

Câteva din problemele semnalate la abordarea parametrică a estimării FDP pentru substrația fundalului sunt:

- necesitatea găsirii unor inițializări de încredere;
- revenire lentă după greșeli;
- dificultăți de adaptare la schimbări rapide ale iluminării;
- dependența rezultatelor de forma reală, posibil nongaussiană, a distribuției;

- necesitatea de a se specifica în avans numărul de componente gaussiene ale mixturii.

2.4 Metode Nonparametrice de estimare a fundalului

Funcțiile densitate de probabilitate pentru fiecare locație din secvența de imagini poate fi estimată cu succes și prin metode nonparametrice, folosind estimatori cu nucleu [69], [88], așa cum o demonstrează câteva lucrări mai recente dedicate detecției fundalului în aplicații de supraveghere [23], [59].

Metodele de estimare nonparametrică pot genera estimate netede, continue, diferențiabile și precise, fără a fi nevoite să facă presupuneri privind legea de distribuție care generează datele observate. De asemenea, nu este necesară specificarea în avans a numărului de maxime (moduri) ale distribuției, iar adaptarea la datele noi este automată.

În pofida denumirii lor, metodele nonparametrice au totuși un parametru important de specificat, ce stabilește rezoluția estimatorului.

Metodele de estimare nonparametrică a FDP au fost mai puțin folosite în aplicații de vedere artificială, datorită costului de calcul substanțial mai mare decât la metodelor parametrice: $O(N^2)$, față de $O(N)$. Situația s-a mai echilibrat în urma dezvoltării unor metode de calcul rapid, bazate pe Transformarea Gauss Rapidă (TGR) [59], [60]. Legat de problema menționată, în capitolul 3 al tezei este propusă **o soluție nouă, bazată pe un calcul recursiv și teste de plauzibilitate, cu ordin de complexitate teoretic $O(N^0)$** , adică independent de lungimea bufferului de date utilizat în estimare.

2.4.1 Tehnici nonparametrice de estimare a densității de probabilitate în spații multidimensionale

Ideea de bază exploatată în majoritatea metodelor de estimare a densității de probabilitate este aceea că probabilitatea ca vectorul d -dimensional $\mathbf{x}$ să aparțină unui subdomeniu R din spațiul analizat este proporțională cu valoarea medie a densității de probabilitate pe subdomeniul respectiv, adică:

$$P = \int_R p(\mathbf{x}) d\mathbf{x} \quad (2.10)$$

Dacă subdomeniul R este suficient de mic, astfel încât densitatea de probabilitate să poată fi considerată aproximativ constantă în interiorul său, rezultă:

$$P \approx p(\mathbf{x}) \int_R d\mathbf{x} = p(\mathbf{x})V \quad (2.11)$$

unde V este volumul ocupat de subdomeniul R .

Ecuția (2.11) sugerează posibilitatea de estimare a densității de probabilitate în centrul domeniului R prin:

$$\hat{p}(\mathbf{x}) = \frac{P}{V} = \frac{\int_R p(\mathbf{y}) d\mathbf{y}}{\int_R d\mathbf{y}} \quad (2.12)$$

Deoarece în practică, numărul eșantioanelor disponibile pentru estimarea densității de probabilitate este finit, alegerea volumului V este problematică. Dacă volumul este foarte mic, așa cum o cere ecuația (2.11), există riscul să conțină prea puține eșantioane sau să nu conțină nici un eșantion. La limita inferioară, dacă volumul tinde la zero, se obțin densități infinite în punctele spațiului corespunzătoare eșantioanelor existente pentru estimare și zero în rest. Dacă volumul este foarte mare, se obține o estimată plată, excesiv netezită prin medierea operată de integrare. La limita superioară, dacă domeniul se extinde în tot spațiul disponibil, rezultă o densitate de probabilitate constantă în jurul oricărui punct. Niciuna din situațiile evocate nu este dorită.

O soluție ingenioasă pentru evitarea extremelor menționate constă în impunerea condiției ca volumul domeniului R să reprezinte un procent precizat din volumul total, dependent de numărul de eșantioane disponibile. Considerând un spațiu normalizat, cu volumul total unitar, se poate impune $V_n = 1/\sqrt{n}$.

Este soluția adoptată de estimatorul Parzen [88]. O soluție alternativă este să se impună ca domeniul R să conțină un procent precizat din eșantioanele disponibile,

respectiv k din totalul de n eşantioane. Metoda este cunoscută sub denumirea de estimator kNN (de la denumirea englezească „k-nearest-neighbor”) [88]. Estimatorul Parzen foloseşte un domeniu hipercubic de latură h , cu volumul

$$V = h^d \quad (2.13)$$

Definim funcţia fereastră:

$$K(\mathbf{u}) = \begin{cases} 1, & \text{pentru } |u_i| \leq 1/2, i = 1, 2, \dots, d \\ 0 & \text{in rest} \end{cases} \quad (2.14)$$

Funcţia $K(\mathbf{u})$ defineşte un hipercub de volum unitar, centrat în originea spaţiului d -dimensional. Numărul eşantioanelor dintr-un hipercub de latură h , centrat în $\mathbf{x}$, este:

$$k = \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (2.15)$$

Deoarece probabilitatea ca un eşantion să aparţină hipercubului considerat este:

$$P = \frac{k}{n} \quad (2.16)$$

densitatea de probabilitate a eşantioanelor în punctul $\mathbf{x}$ devine:

$$\hat{p}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h^d} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) \quad (2.17)$$

Ecuţia (2.17) sugerează o abordare mai generală a problemei estimării densităţii de probabilitate. În esenţă, funcţia fereastră este folosită această ecuaţie ca o funcţie de interpolare, fiecare eşantion contribuind la estimata în punctul $\mathbf{x}$ în funcţie de distanţa până la $\mathbf{x}$. Fereastra hipercubică centrată pe $\mathbf{x}$ ia valoarea 1 pentru eşantioanele aflate în interiorul ei şi zero în rest, corespunzător unei interpolări de ordinul zero. Este firesc să ne punem întrebarea dacă putem folosi şi alte funcţii de interpolare şi ce condiţii trebuie să satisfacă aceste funcţii pentru a se obţine estimate legitime.

Condiția:

$$K(\mathbf{u}) \geq 0 \quad (2.18)$$

este necesară pentru a se garanta non-negativitatea estimatei, în timp ce menținerea volumului elementar pentru estimare conform ecuației (2.13) și necesitatea ca $p(\mathbf{x})$ să aibă integrala unitară conduc la:

$$\int_R K(\mathbf{u}) d\mathbf{u} = 1 \quad (2.19)$$

O asemenea condiție este îndeplinită de o largă varietate de funcții. De regulă acestea se aleg simetrice și monoton descrescătoare. Indiferent de forma particulară, parametrul h are un efect de scalare extrem de important. Dacă h tinde spre zero, volumul elementar tinde spre zero și contribuțiile eșantioanelor scad extrem de rapid cu distanța față de punctul pentru care se face estimarea densității. Pe măsură ce h crește, estimata în fiecare punct tinde să includă contribuțiile unui număr mai mare de eșantioane și estimata este mai netedă.

2.4.1.1 Convergența mediei

Ne propunem să calculăm media estimatei densității de probabilitate conform ecuației (2.17). O notăm cu $\bar{p}(\mathbf{x})$ și o considerăm o variabilă aleatoare. Conform definiției,

$$\begin{aligned} \bar{p}(\mathbf{x}) &= E\{\hat{p}(\mathbf{x})\} = E\left\{\frac{1}{n} \sum_{i=1}^n \frac{1}{V_n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)\right\} = \\ &= \frac{1}{n} \sum_{i=1}^n E\left\{\frac{1}{V_n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)\right\} = \frac{1}{n} \sum_{i=1}^n \int \frac{1}{V_n} K\left(\frac{\mathbf{x} - \mathbf{u}}{h}\right) p(\mathbf{u}) d\mathbf{u} = \\ &= \int \frac{1}{V_n} K\left(\frac{\mathbf{x} - \mathbf{u}}{h}\right) p(\mathbf{u}) d\mathbf{u}. \end{aligned} \quad (2.20)$$

Ecuația obținută ne arată că media estimatei este rezultatul convoluției densității de probabilitate reale și necunoscute cu funcția fereastră interpolatoare. Estimata este o versiune netezită a densității reale.

Considerând $V_n = 1/\sqrt{n}$, pe măsură ce n tinde la infinit, parametrul h și volumul V_n tind spre zero, iar funcția $K(\mathbf{u}/h)/V_n$ tinde spre forma Dirac.

În consecință, dacă funcția densitate de probabilitate, $p(\mathbf{x})$, este continuă, când numărul eșantioanelor tinde spre infinit și volumul V_n , respectiv rezoluția ferestrei, tind spre zero, media estimatei densității de probabilitate converge spre valoarea reală a funcției densitate de probabilitate.

2.4.1.2 Convergența dispersiei

Ecuția (2.20) arată că *media* estimatei poate fi făcută să convergă spre densitatea reală chiar și pentru un număr de eșantioane, n , finit, dacă volumul V_n tinde la zero. Desigur, pentru diferite mulțimi de n eșantioane vor rezulta estimate diferite, motiv pentru care este firesc să căutăm o caracterizare a variațiilor acestor rezultate, prin dispersia (varianța) estimatei. Deoarece $\hat{p}(\mathbf{x})$ este suma unor funcții de variabile aleatoare statistic independente, dispersia ei este suma dispersiilor componentelor sumei:

$$\begin{aligned}\sigma_n^2 &= E\{(\hat{p}(\mathbf{x}) - \bar{p}(\mathbf{x}))^2\} = E\left\{\sum_{i=1}^n \left(\frac{1}{n} \frac{1}{V_n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right) - \frac{1}{n} \bar{p}(\mathbf{x})\right)^2\right\} = \\ &= nE\left\{\frac{1}{n^2} \frac{1}{V_n^2} K^2\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)\right\} - \frac{1}{n} \bar{p}^2(\mathbf{x}) \leq nE\left\{\frac{1}{n^2} \frac{1}{V_n^2} K^2\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right)\right\} = \\ &= \frac{1}{nV_n} \int \frac{1}{V_n} K^2\left(\frac{\mathbf{x} - \mathbf{u}}{h}\right) p(\mathbf{u}) d\mathbf{u} \leq \frac{\sup(K(\cdot))}{nV_n} \bar{p}(\mathbf{x}).\end{aligned}\tag{2.21}$$

Se observă că pentru a se obține o dispersie redusă a estimatei, este necesară utilizarea unui volum V_n mare, respectiv o scară h grosieră. În mod evident, ne aflăm în fața cunoscutei dileme din domeniul estimării: medie exactă sau dispersie redusă?

Totuși este important să evidențiem faptul că alegând, de exemplu, $V_n = 1/\sqrt{n}$ sau $V_n = V_1 / \log n$, putem asigura ca dispersia să tindă la zero chiar și când V_n tinde la zero, cu n tinzând la infinit. Acest rezultat teoretic important ne arată posibilitatea obținerii unei estimări asimptotice foarte bune a densității de probabilitate, când dispunem de eșantioane suficient de multe. Din păcate, nu dă nici un indiciu privind alegerea funcției fereastră și a volumului V_n sau a parametrului de rezoluție h pentru

situația practică din prelucrarea imaginilor, în care numărul eşantioanelor disponibile, n , este limitat drastic.

2.4.1.3 Funcții nucleu

Calitatea estimatei funcției densitate de probabilitate depinde de alura funcției fereastră, numită și *funcție nucleu a operatorului de estimare*, și de parametrul de scară, h . Mai general decât în ecuația (2.17), funcția densitate de probabilitate poate fi estimată prin:

$$\hat{p}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) \quad (2.22)$$

unde

$$\hat{p}(\mathbf{x}) = K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K_{\mathbf{H}}(\mathbf{H}^{-1/2} \mathbf{x}) \quad (2.23)$$

cu $\mathbf{H}$ o matrice $d \times d$ simetrică și pozitiv definită, ce stabilește lărgimea de bandă a estimatorului. Din considerente de complexitate, uzual $\mathbf{H}$ are forma diagonală. Mai mult, pentru spații omogene, este preferată forma proporțională cu matricea unitate, $\mathbf{H} = h^2 \mathbf{I}$, cu avantajul de a avea un singur parametru de scară, ceea ce conduce la un estimator de forma definită în ecuația (2.17).

Pentru spații multidimensionale, funcția nucleu poate fi definită în mod convenabil pornind de la o fereastră unidimensională simetrică, folosind una din modalitățile următoare:

$$K^P(\mathbf{x}) = \prod_{i=1}^d K_1(x_i) \quad (2.24)$$

sau

$$K^R(\mathbf{x}) = \alpha K_1(\|\mathbf{x}\|) \quad (2.25)$$

unde α este o constantă de proporționalitate strict pozitivă ce se alege astfel încât estimata densității de probabilitate să se integreze la valoarea unitară:

$$\alpha = \frac{1}{\int_{R^d} K_1(\|\mathbf{x}\|) d\mathbf{x}} \quad (2.26)$$

În prima variantă, funcția nucleu multidimensională se obține ca produs al funcțiilor nucleu unidimensionale. Este și cazul ferestrei hipercubice ce poate fi generată ca produs al ferestrelor unidimensionale rectangulare:

$$rect(x) = \begin{cases} 1, & |x| \leq \frac{1}{2} \\ 0, & \text{in rest} \end{cases} \quad (2.27)$$

În cel de-al doilea caz, funcția nucleu multidimensională se obține rotind fereastra unidimensională în spațiul caracteristicilor. Se obțin astfel nuclee cu simetrie radială.

Restrângându-ne atenția la funcții nucleu de forma:

$$K^R(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2) \quad (2.28)$$

unde $c_{k,d}$ este o constantă ce asigură din nou integrarea la 1 a funcției densitate de probabilitate (FDP), putem defini estimatorul cu ajutorul *profilului* $k(x)$, o funcție unidimensională definită doar pentru valorile pozitive ale variabilei scalare x .

Performanțele estimatorului pot fi apreciate cu ajutorul abaterii pătratice dintre estimată și valoarea exactă, integrată pe întregul domeniu de variație al vectorului $\mathbf{x}$. În practică se poate obține doar o aproximare asimptotică a acestei erori, minimizată de nucleul Epanechnikov [89], având profilul

$$k_E(x) = \begin{cases} 1-x, & 0 \leq x \leq 1 \\ 0, & x > 1 \end{cases} \quad (2.29)$$

cu nucleul corespunzător,

$$K_E(\mathbf{x}) = \begin{cases} \frac{d+2}{2V_d}(1-\|\mathbf{x}\|^2), & \|\mathbf{x}\| \leq 1 \\ 0, & \|\mathbf{x}\| > 1 \end{cases} \quad (2.30)$$

unde V_d este volumul sferei de rază unitară în spațiul d -dimensional.

Profilul:

$$k_N(x) = \exp\left(-\frac{1}{2}x\right), \quad x \geq 0 \quad (2.31)$$

corespunde nucleului normal,

$$K_N(\mathbf{x}) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right) \quad (2.32)$$

ce poate fi generat identic prin oricare din ecuațiile (2.24) sau (2.25).

În practică, nucleul normal se trunchiază adesea, pentru a fi redus la un suport finit.

2.4.1.4 Detecția maximelor locale ale funcției densitate de probabilitate

La estimarea fundalului în videosecvențe, estimarea completă a FDP nu este necesară. Spre deosebire de prim-plan, fundalul corespunde unui vector ce maximizează local FDP. În consecință, facem observația că integrala abaterii pătratică de estimare, folosită curent în literatură pentru evaluarea calității estimatorilor FDP, nu reprezintă în mod necesar cel mai bun criteriu de performanță în problema estimării fundalului. Ceea ce ne interesează în ultimă instanță este localizarea precisă a maximelor FDP în spațiul caracteristicilor. Valoarea FDP în punctele de maxim nu trebuie cunoscută cu precizie deosebită, fiind interesantă doar pentru eliminarea maximelor inconsistente.

Pornind de la observația că maximele FDP sunt puncte în care gradientul de densitate este nul, se poate construi un algoritm de tip gradient ascendent, pentru localizarea maximelor. Gradientul funcției densitate de probabilitate poate fi estimat și el direct, pe baza eșantioanelor disponibile, pornind de la ecuația (2.22), pe care o rescriem folosind profilul nucleului în forma:

$$\hat{p}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (2.33)$$

Definind estimata gradientului FDP prin gradientul estimatei FDP și folosind liniaritatea operatorului gradient, putem scrie [90], [91]:

$$\hat{\nabla} p_{h,K}(\mathbf{x}) \equiv \nabla \hat{p}_{h,K}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x} - \mathbf{x}_i) k'\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (2.34)$$

Presupunând că derivata profilului nucleului există pentru toate numerele reale pozitive, exceptând un număr finit de puncte, definim funcția derivată a profilului,

$$g(x) = -k'(x) \quad (2.35)$$

Folosind funcția $g(x)$ ca profil, obținem nucleul corespunzător:

$$G(\mathbf{x}) = c_{g,d} g(\|\mathbf{x}\|^2) \quad (2.36)$$

unde $c_{g,d}$ este o constantă de normalizare. Nucleul $K(\mathbf{x})$ este denumit nucleu umbră al lui $G(\mathbf{x})$. Astfel, nucleul Epanechnikov este nucleul umbră al nucleului uniform în interiorul sferei de rază unitate, în timp ce umbra nucleului gaussian este tot un nucleu gaussian. Introducând $g(x)$ în ecuația (2.34), se obține:

$$\begin{aligned} \hat{\nabla} p_{h,K}(\mathbf{x}) &= \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{x}) g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) = \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \right] \left[\frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \right]. \end{aligned} \quad (2.37)$$

Ambii factori ai produsului de mai sus au semnificații importante. Primul,

$$p_{h,G}(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right) \quad (2.38)$$

este estimata FDP folosind nucleul $G(\mathbf{x})$. Al doilea este *translația mediei*:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{x} \quad (2.39)$$

adică diferența dintre media ponderată a eșantioanelor în jurul lui $\mathbf{x}$, folosind nucleul $G(\mathbf{x})$ și $\mathbf{x}$.

Folosind ultimele două ecuații, estimata gradientului FDP evaluate la scara h cu nucleul $K(\mathbf{x})$ devine:

$$\hat{\nabla} p_{h,K}(\mathbf{x}) = \hat{p}_{h,G}(\mathbf{x}) \frac{2c_{k,d}}{h^2 c_{g,d}} \mathbf{m}_{h,G}(\mathbf{x}) \quad (2.40)$$

ecuație ce se poate rescrie în forma:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\hat{\nabla} p_{h,K}(\mathbf{x})}{\hat{p}_{h,G}(\mathbf{x})} \quad (2.41)$$

Rezultă că vectorul de translație a mediei (VTM), calculat cu nucleul $G(\mathbf{x})$, este proporțional cu *gradientul normalizat* al FDP, estimate cu nucleul $K(\mathbf{x})$. Normalizarea este în raport cu estimata calculată cu nucleul $G(\mathbf{x})$. Prin urmare, VTM (engl. *mean shift vector*) este orientat în *direcția cu cea mai rapidă creștere a densității*.

Ecuația (2.41) este intuitivă: media locală este deplasată spre regiunea în care densitatea de eșantioane este mai mare. Pornind de la această observație, se poate construi un algoritm simplu pentru detecția maximelor locale ale FDP. Pașii de bază ai algoritmului sunt:

- calculul vectorului de translație a mediei (VTM), $\mathbf{m}_{h,G}(\mathbf{x})$,

- translatarea cu $\mathbf{m}_{h,G}(\mathbf{x})$ a centrului ferestrei (nucleului) $G(\mathbf{x})$.

Acești pași se repetă până la atingerea unui punct staționar de maximă densitate. Pozițiile succesive ale centrului ferestrei corespund unor locații cu densități din ce în ce mai mari. În punctul de maxim, gradientul este nul, deci VTM este nul și algoritmul converge. În realitate, simplul fapt că translațiile succesive ale centrului ferestrei au loc pe direcțiile estimatelor gradientului de densitate nu garantează convergența algoritmului, decât dacă deplasările sunt infinitesimale. Este totuși de remarcat o proprietate favorabilă a VTM: modulul lui este invers proporțional cu densitatea estimată. În consecință, deplasările scad în amplitudine, pe măsură ce fereastra se apropie de poziția corespunzătoare densității maxime.

2.4.1.5 Algoritmul cu translație la medie (ATM)

Fie $\{\mathbf{y}_j\}_{j=1,2,\dots}$ secvența locațiilor succesive ale centrelor nucleului $G()$,

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}, \quad j = 1, 2, \dots$$

(2.42)

unde $\mathbf{y}_1 = \mathbf{x}$ este punctul de start al algoritmului. Se observă că $\mathbf{y}_{j+1}$ este media ponderată a eșantioanelor în jurul locației precedente, $\mathbf{y}_j$, cu ponderile definite de nucleul $G()$, respectiv profilul corespunzător, $g()$. Secvența corespunzătoare a estimatelor succesive ale densității, folosind nucleul K , este:

$$\{\hat{p}_{h,K}(j)\}_{j=1,2,\dots}$$

$$\hat{p}_{h,K}(j) = \hat{p}_{h,K}(\mathbf{y}_j), \quad j = 1, 2, \dots$$

(2.43)

S-a demonstrat recent [91] următoarea teoremă:

Teorema 1. *Dacă nucleul $K()$ are un profil convex și monoton descrescător, secvențele*

$\{\mathbf{y}_j\}_{j=1,2,\dots}$ și $\{\hat{p}_{h,K}(\mathbf{y}_j)\}_{j=1,2,\dots}$ sunt convergente și secvența $\{\hat{p}_{h,K}(\mathbf{y}_j)\}_{j=1,2,\dots}$ este monoton crescătoare.

De menționat că proprietatea de secvență monoton crescătoare și valoarea finită a densității estimate cu nuclee cu suport finit asigură convergența, în timp ce convergența nu implică monotonia. Caracterul monoton crescător asigură convergența spre maxim. Mulțimea punctelor ce converg în același maxim local formează un bazin de atracție al maximului. Pe această bază, este posibilă segmentarea datelor prin identificarea maximelor FDP (modurilor) folosind algoritmul cu translație la medie.

Algoritmul cu translație la medie permite identificarea modului FDP asociat oricărui punct din spațiul caracteristicilor. El poate fi utilizat ca un instrument eficient atât pentru filtrarea cât și pentru segmentarea datelor. În filtrare, fiecare din vectorii de intrare, $\mathbf{x}$, este înlocuit cu vectorul $\mathbf{y}_c$ la care algoritmul inițializat cu $\mathbf{x}$ converge. În segmentare, vectorii de date asociați aceluiași mod sunt grupați într-o clasă comună.

2.5 Concluzii

După o prezentare succintă a principalelor dificultăți ce survin în estimarea fundalului, s-au prezentat argumentele pentru modelarea statistică a distribuțiilor multidimensionale ale caracteristicilor de fundal. S-au prezentat comparativ avantajele și dezavantajele modelelor parametrice și nonparametrice și s-au introdus bazele teoretice ale modelării parametrice și nonparametrice ale fundalului.

O extindere considerabil mai mare a avut-o modelarea nonparametrică, metodă ce a fost adoptată pe parcursul tezei. Este tratat teoretic, în detaliu și algoritmul cu deplasare la medie (mean-shift), un instrument modern pentru localizarea modurilor funcțiilor densitate de probabilitate, ce va fi exploatat într-o manieră nouă în capitolele următoare.

CAPITOLUL 3.

ESTIMAREA NONPARAMETRICĂ RECURSIVĂ A FUNDALULUI

Pornind de la o analiză atentă a modului în care tehnicile de estimare nonparametrică a densității de probabilitate cu funcții nucleu pot fi utilizate în estimarea și substrația fundalului pentru aplicații de supraveghere video, în acest capitol se propune **o nouă soluție pentru estimarea nonparametrică a fundalului**, cunoscută sub denumirea de **metodă de estimare nonparametrică recursivă a fundalului** [105]. Noua metoda propusă s-a dovedit a fi în același timp rapidă și exactă.

Capitolul este structurat în șase paragrafe. În primul paragraf se prezintă criteriile care stau la baza adoptării dimensiunii corecte a bufferului de cadre folosite pentru estimarea inițială a fundalului, în cazul utilizării tehnicilor de estimare nonparametrică. Paragraful 3.2 descrie proiectarea unui estimator nonparametric pentru estimarea inițială a fundalului, cu referire specifică la alegerea optimă a factorului de scală. Modalitatea de substrație a fundalului pentru segmentarea fundal / prim-plan este fundamentată în paragraful 3.3. În paragraful 3.4 se propune un estimator nonparametric recursiv ce asigură urmărirea eficientă a schimbărilor de fundal. Rezultatele testelor de performanță ale estimatorului nonparametric recursiv sunt prezentate în paragraful 3.5. Testele evaluează atât precizia estimării în diverse condiții de funcționare cât și viteza de calcul comparativ cu soluții de referință. Concluziile capitolului sunt sintetizate în paragraful 3.6.

3.1 Adoptarea dimensiunii bufferului de cadre

O ipoteză de lucru simplificatoare ce se adoptă curent la modul implicit în estimarea nonparametrică a fundalului [23] este aceea că, în mulțimea celor N cadre

utilizate la estimarea fundalului, vectorul caracteristic al fundalului este observat cel puțin o dată în interiorul marjei de eroare dorite. Considerăm în cele ce urmează această ipoteză validă, cu toate că metoda de estimare pe care o propunem poate da rezultate exacte chiar și când ipoteza menționată este infirmată.

Din punct de vedere teoretic, pentru un fundal constant, lungimea bufferului de cadre, N , trebuie să fie cât mai mare. Așa cum se va arăta, în cadrul metodei propuse, valoarea lui N nu afectează complexitatea de calcul, exceptând estimarea inițială, la inițializarea sistemului. Lungimea bufferului de cadre utilizat în estimarea și substrația fundalului trebuie totuși menținută în limite rezonabile din considerente de spațiu de memorie consumat. De asemenea, valori excesiv de mari ale lui N pot afecta negativ capacitatea estimatorului de a se adapta la schimbări rapide ale nivelului de iluminare al scenei. Experimentele noastre au confirmat că lungimi de ordinul sutelor de cadre sunt suficiente pentru o gamă largă de situații practice. În aceste condiții, pentru a se acoperi un timp de observație suficient de mare, este preferabilă subeșantionarea cadrelor la estimarea fundalului, ceea ce are și un efect benefic asupra reducerii timpului mediu de calcul afectat estimării fundalului.

O soluție posibilă pentru reducerea spațiului de memorie ocupat de cadrele utilizate în estimarea și urmărirea fundalului o constituie stocarea datelor într-o structură de date de tip histogramă. Dacă numărul de coloane al histogramei, M , este mai mic decât N , se obține reducerea spațiului de memorie și în același timp sunt create premisele unei prelucrări mai rapide. Pentru imagini monocromatice, soluția este atractivă. În schimb, pentru imagini color cuantizate pe M niveluri, rezultă histogramă cu un număr de M^3 celule, ceea ce face soluția mult mai puțin atractivă. De exemplu, pentru $M = 64$, se ajunge la $2^{18} = 256$ Mega celule. Valori mai reduse ale lui M pot afecta nepermis de mult acuratețea estimării.

Cu toate că nu am ales soluția de a stoca în forma de histogramă cadrele pentru estimarea fundalului, în lucrarea prezentă folosim ideea de prelucrare rapidă bazată pe o histogramă grosieră, cu număr de celule moderat, pentru construcția unui test rapid de plauzibilitate de apartenență la fundal, ce elimină de la calculul exact un procent important din datele folosite la estimare.

3.2 Estimarea inițială a fundalului

Procesul de estimare a fundalului propus de doctorand este compus din două etape de prelucrare distincte:

- Prima etapă, prezentată în acest paragraf, este de inițializare și este derulată o singură dată, la pornirea sistemului, fiind, în esență, similară cu metoda propusă de Elgamal [23].
- Etapa a doua, prezentată în paragraful următor, actualizează recursiv estimarea inițială și este mult mai rapidă.

Se presupune că dispunem de un set de N cadre de imagine pentru estimare. Nu se presupune absența activității în aceste cadre. Vectorii caracteristicilor reprezintă triplete de culoare, RGB :

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3.1)$$

La fiecare locație spațială, funcția densitate de probabilitate (FDP) este estimată folosind ecuația:

$$\hat{p}(\mathbf{x}_k) = \frac{1}{N} \sum_{i=1}^N K_h(\mathbf{x}_k - \mathbf{x}_i), \quad k = 1, 2, \dots, N \quad (3.2)$$

cu

$$K_h(\mathbf{x}_k - \mathbf{x}_i) = \prod_{c=1}^3 \text{rect}\left(\frac{x_{kc} - x_{ic}}{h_c}\right) \quad (3.3)$$

și

$$\text{rect}(u) = \begin{cases} 1, & \text{pentru } |u| \leq \frac{1}{2} \\ 0 & \text{altfel} \end{cases} \quad (3.4)$$

Pentru simplificarea notațiilor, coordonatele spațiale ale vectorilor de date au fost omise. Factorul de scară al estimatorului este calculat adaptiv, în funcție de dispersia datelor de intrare, distinct pentru fiecare locație spațială, conform ecuației:

$$h_c = \alpha \underset{i}{\text{median}} |x_{ic} - x_{(i-1)c}| \quad (3.5)$$

unde i este indexul cadrului și α o constantă.

Se observă că factorul de scară este ales proporțional cu mediana deviațiilor absolute, la fel ca în [23]. Mediana deviațiilor absolute dintre cadre consecutive reprezintă un estimator robust al varianței intraclasă a datelor, fiind foarte puțin afectat de salturile de mare amplitudine relativ infrecvente, ce apar când muchii diferite ale unui obiect mobil sunt proiectate la o anumită locație.

Pentru calculul estimativ rapid al medianei, se folosește o soluție recursivă:

$$\text{median}(i) = \text{median}(i-1) + \eta \text{sign}(x_{ic} - x_{(i-1)c}) \quad (3.6)$$

unde η este o constantă subunitară, având semnificația unei rate de învățare.

Soluția poate fi înțeleasă pe baza proprietății medianei de a se situa la mijlocul șirului ordonat al eșantioanelor folosite în filtrare. Într-o formulare alternativă, proprietatea cere egalitatea dintre numărul de eșantioane mai mici decât mediana și numărul de eșantioane mai mari decât aceasta. Prin ecuația (3.6) mediana estimată este modificată incremental de fiecare eșantion de date. Cât timp mediana nu este situată la mijlocul șirului ordonat, estimata este modificată asimetric, deoarece vor exista eșantioane mai mici și respectiv mai mari în proporții diferite. În consecință estimata tinde să se stabilizeze la valoarea mediana a eșantioanelor de date.

Detecția prim-planului mobil este posibilă prin discriminare cu prag în domeniul densității de probabilitate estimate:

$$\mathbf{x}_k \in \mathbf{F} \Leftrightarrow \hat{p}(\mathbf{x}_k) < Th, \quad (3.7)$$

unde Th este un prag ales convenabil.

Culoarea fundalului este dată de

$$\mathbf{b} = \arg \max_{\mathbf{x}_k} \{\hat{p}(\mathbf{x}_k)\} \quad (3.8)$$

și poate fi, de asemenea, utilizată pentru discriminare între fundal și prim-plan.

3.3 Segmentarea fundal / prim-plan

Principial, segmentarea fundal/prim-plan poate fi obținută pe baza ecuației (3.7), fără extragerea modului distribuției fundalului, $\mathbf{b}$.

Cu toate acestea, în lucrarea prezentă s-a adoptat o soluție bazată pe calculul explicit al fundalului, $\mathbf{b}$, pentru că permite evitarea estimării densității pentru fiecare vector de intrare și efectuarea acestei estimări la o rată substanțial mai mică. În plus, $\mathbf{b}$ poate fi util și la detecția umbrelor. În consecință, un pixel nou, $\mathbf{x}$, este clasificat ca mobil, sau aparținând de prim-plan, dacă:

$$d(\mathbf{x}, \mathbf{b}) > Th \quad (3.9)$$

unde $d(\mathbf{x}, \mathbf{b})$ este o măsură adecvată a similarității dintre doi vectori de culoare și Th un prag de decizie.

Problema găsirii unei măsuri adecvate a similarității dintre două culori a fost îndelung studiată în literatură, fiind de interes major nu numai în vederea artificială ci și în compresia imaginilor, filtrarea imaginilor sau colorimetrie. Soluția cea mai simplă constă în utilizarea normei Euclidiene a vectorului diferență de culoare în spațiul *RGB*. Este cunoscut că diferențele de culoare în acest spațiu, adoptat din considerente legate de tehnologiile de captare și reproducere a imaginilor în televiziune, nu corespund fidel evaluărilor subiective, bazate pe percepția vizuală.

O corespondență mai bună cu aceste evaluări se obține pe baza distanțelor Euclidiene măsurate în spațiul Luv sau în spațiul Lab. Un avantaj suplimentar al acestor spații este accesul direct la informația de luminanță, utilă în detecția umbrelor. Accesul direct la informația de luminanță se regăsește și în spațiul *HSV*, cu avantajul că pentru

conversia coordonatelor color din spațiul *RGB* în spațiul *HSV* este nevoie doar de o transformare liniară simplă și rapidă:

$$\begin{aligned}s &= (R + G + B) / 3, \\ r &= R / s, \\ g &= G / s.\end{aligned}\tag{3.10}$$

Aparent, informația de luminanță, s , poate fi ignorată, rezultând o metodă invariantă la luminanță și în consecință imună și la efectele de umbrire. Totuși, o soluție bazată exclusiv pe componentele cromatice r și g ar conduce la imposibilitatea distingerii obiectelor incolore (alb, negru, gri). Mai mult, culorile obiectelor foarte întunecate sunt incert definite, în timp ce obiectele foarte luminoase pot provoca saturarea senzorilor de imagine ai camerei TV.

Cu mențiunea că problema detecției umbrelor rămâne un subiect deschis și de interes major în prelucrarea imaginilor, în lucrarea prezentă am adoptat spațiul color *HSV* cu norma L^1 ponderată, cu rezultate pe care le consider mulțumitoare:

$$d(c_1, c_2) = |s_1 - s_2| + M|r_1 - r_2| + M|g_1 - g_2|\tag{3.11}$$

unde M este valoarea maximă a semnalelor R, G sau B .

Scalarea este esențială pentru compensarea paletii valorice foarte diferite între componenta de luminanță și cele cromatice în spațiul *HSV*.

Măștile de segmentare ale obiectelor obținute prin discriminarea cu prag a diferențelor de culoare, evaluate cu ajutorul ecuației (3.11), sunt afectate de numeroase surse de eroare. Unele, de genul celor induse de prezența umbrelor sau camuflaj sunt dificil de eliminat și pot necesita prelucrare multimodală. Altele, induse de zgomot și constând în mici grupuri izolate de pixeli, corespunzând unor obiecte false sau găuri false în obiecte reale, pot fi corectate eficient prin tehnici de filtrare spațială. Soluția de filtrare adoptată de majoritatea autorilor este filtrarea morfologică, datorită în primul rând a existenței metodelor de filtrare morfologică binară rapidă.

În teză am folosit o soluție de filtrare care exploatează informația reziduală conținută de imaginea diferență dintre cadrul curent și fundalul estimat, în locul filtrării

morfologice a măștilor de segmentare, obținute prin binarizarea imaginii diferență, folosită curent de majoritatea cercetătorilor.

Soluția adoptată constă în filtrarea liniară a imaginii diferență cu un operator trece-jos, urmată de binarizarea propriu-zisă. Ansamblul celor două operații reprezintă o convoluție cu prag. Am folosit un filtru binomial cu masca:

$$\mathbf{H}_B = \frac{1}{65536} \begin{bmatrix} 1 & 8 & 28 & 56 & 70 & 56 & 28 & 8 & 1 \\ 8 & 64 & 224 & 448 & 560 & 448 & 224 & 64 & 8 \\ 28 & 224 & 784 & 1568 & 1960 & 1560 & 784 & 224 & 28 \\ 56 & 448 & 1568 & 3136 & 3920 & 3136 & 1568 & 448 & 56 \\ 70 & 560 & 1960 & 3920 & 4900 & 3920 & 1960 & 560 & 70 \\ 56 & 448 & 1568 & 3136 & 3920 & 3136 & 1568 & 448 & 56 \\ 28 & 224 & 784 & 1568 & 1960 & 1560 & 784 & 224 & 28 \\ 8 & 64 & 224 & 448 & 560 & 448 & 224 & 64 & 8 \\ 1 & 8 & 28 & 56 & 70 & 56 & 224 & 64 & 8 \end{bmatrix}. \quad (3.12)$$

Pentru a menține viteza de prelucrare ridicată, am exploatat separabilitatea filtrului binomial și am factorizat pe biți expresia convoluției 1D.

3.4 Estimator nonparametric recursiv

Se propune o metodă rapidă pentru urmărirea schimbărilor de fundal, ce îmbină avantajele tehnicilor de estimare nonparametrică a funcției densitate de probabilitate (FDP) bazată pe funcții nucleu, cu viteza tehnicilor bazate pe histogramme.

Implementarea directă a ecuațiilor (3.2) și (3.8) presupune evaluarea densității pentru fiecare vector de intrare, conducând la un număr de N^2 operații. O analiză atentă relevă faptul că, după epuizarea primelor N cadre, densitățile pentru pixelii cadrelor noi se pot obține printr-o procedură recursivă, pe baza rezultatelor preexistente și câteva observații simple.

La recepționarea unui cadru de estimare nou, un vector de date nou înlocuiește în bufferul de memorie-cadre de lungime N vectorul culoare cel mai vechi. Pentru toți cei $N - 1$ vectori de date neschimbați, noile densități se pot calcula în funcție de cele

vechi, prin adăugarea contribuției vectorului de date nou intrat și scăderea contribuției vectorului scos din memorie:

$$\hat{p}_{new}(\mathbf{x}) = \hat{p}_{old}(\mathbf{x}) + \frac{1}{N} K_h(\mathbf{x} - \mathbf{x}_{new}) - \frac{1}{N} K_h(\mathbf{x} - \mathbf{x}_{old})$$

(3.13)

Calculul ecuației de mai sus consumă numai două operații pe pixel (la nucleul uniform, numai două adunări), adică $2(N - 1)$ operații pentru evaluarea ecuației în toate punctele pre-existente în bufferul de memorie-cadre. Se obține astfel reducerea complexității de calcul de la $O(N^2)$ la $O(2N)$.

În realitate, se poate obține o reducere a complexității de calcul și mai drastică, pornind de la observația că evaluarea ecuației (3.13) pentru toate datele din memoria de cadre la o locație dată este inutilă. Numai dacă densitatea este suficient de mare pentru a fi șanse de a corespunde fundalului, operația își atinge scopul. Pentru pixelii ce nu aparțin fundalului curent, o evaluare grosieră și rapidă a FDP, cu ajutorul unei histograme 3D este suficientă pentru verificarea plauzibilității ipotezei ca aceștia să maximizeze FDP. Actualizarea histogramei 3D pentru un vector color necesită numai 3 operații de incrementare (decrementare, dacă acesta iese din bufferul de date).

În lucrarea prezentă am folosit o histogramă color 3D de format $16 \times 16 \times 16$.

Presupunând că fundalul este modelat cu o culoare unică la fiecare locație, definită prin proprietatea de a maximiza funcția densitate de probabilitate (FDP), ceea ce trebuie verificat la fiecare pixel nou este dacă el modifică sau nu maximum FDP. Un pixel nou poate fi inclus într-una din următoarele două categorii:

- aparține fundalului curent;
- nu aparține fundalului curent.

În prima situație, pixelul nou se situează în interiorul ferestrei de estimare centrate pe fundalul curent estimat, $\mathbf{b}$, în timp ce în situația a doua este în afara acestei ferestre.

Dacă pixelul aparține fundalului, este necesară actualizarea exactă a densității fundalului, conform ecuației (3.13), pentru $\mathbf{x} = \mathbf{b}$. Dacă pixelul nu aparține fundalului curent, $\mathbf{b}$, există încă o șansă ca el să maximizeze FDP și astfel să schimbe radical

fundalul estimat. Este cazul relativ infrecvent al obiectelor introduse sau scoase din fundal sau cazul unei schimbări foarte drastice a iluminării. Dacă FDP la pixelul nou, evaluată cu ajutorul histogramei este superioară unui prag, pixelul nou trece testul de plauzibilitate și densitatea lui este re-evaluată exact, pe baza ecuației (3.2). Câștigul de viteză se obține datorită faptului că acest calcul exact trebuie efectuat extrem de rar, așa cum au dovedit-o și testele noastre experimentale extinse.

O descriere concisă a algoritmului de urmărire a fundalului propusă de doctorand este redată în figura 3.1.

```

if(  $K_h(\mathbf{x}_{new} - \mathbf{b}) \neq 0$  )
    update(  $\mathbf{b}$  and  $\hat{p}(\mathbf{b})$  );
else if(  $\text{Hist}(\mathbf{x}_{new}) > \text{threshold}$  )
    if(  $\hat{p}(\mathbf{x}_{new}) > \hat{p}(\mathbf{b})$  )
         $\mathbf{b} =: \mathbf{x}_{new}$ ;
    
```

Fig. 3.1. Descriere concisă a algoritmului rapid de urmărire a fundalului.

Actualizarea funcției densitate de probabilitate (FDP) pentru pixelii noi de fundal se realizează conform ecuației (3.13), în timp ce culoarea fundalului se actualizează folosind ecuația:

$$\mathbf{b}_{new} = (1 - \alpha)\mathbf{b}_{old} + \alpha\mathbf{x}_{new} = \mathbf{b}_{old} + \alpha(\mathbf{x}_{new} - \mathbf{b}_{old}) \quad (3.14)$$

Formula de actualizare este similară celei folosite în actualizarea mediilor (modurilor) la estimarea parametrică a FDP prin mixturi de gaussiene. În contextul estimării nonparametrice adoptate ca suport teoretic de bază în lucrarea prezentă, motivația este legată, de fapt, de algoritmul cu translație la medie, descris în capitolul precedent.

Regula corespunde unei singure iterații ale algoritmului ATM, corespunzătoare estimării FDP cu nucleul Epanechnikov. Deoarece punctul de start este deja apropiat de mod, includerea unui număr mai mare de iterații nu ar aduce schimbări importante. Așa

cum am mai menționat, uzual ATM converge extrem de rapid, în numai doi-trei pași. De remarcat că fundalul estimat prin ATM este un vector continuu, nediscretizat.

Ecuția de actualizare (3.14) este echivalentă cu (3.13) pentru nucleul Epanechnikov, dacă se alege:

$$\alpha = \frac{1}{1 + N\hat{p}(\mathbf{b}_{old})} \quad (3.15)$$

În practică, am folosit valori ale lui α constante, în intervalul $(0,05 \div 0,1)$. Acestea s-au dovedit a nu influența esențial rezultatele obținute.

Dimensiunile celulelor de cuantizare ale histogramei trebuie alese comparabile cu parametrul de scară ale ferestrei de estimare, h_c , preferabil mai mari, în scopul evitării deciziilor false negative. Pragul de decizie pentru densitatea estimată pe baza histogramei s-a stabilit la o fracțiune, β , din densitatea maximă, estimată pentru fundal.

În toate experimentele descrise mai jos, s-a folosit $\beta = 0,5$, ceea ce asigură o funcționare corectă chiar pentru o marjă de eroare de 50 % pentru estimata densității bazată pe histogramă. Valoarea nu este critică și nu influențează prea mult nici timpii de prelucrare, deoarece pentru marea majoritate a pixelilor de prim-plan, densitatea este cu cel puțin un ordin de mărime mai mică decât cea a fundalului. Parametrul de scară, h_c , este actualizat la fiecare cadru. La fel și valoarea fundalului, $\mathbf{b}$, dar numai pentru un pixel dintr-un bloc de dimensiunea 4×4 . Este nevoie de un număr de 16 cadre pentru 4×4 actualizarea întregului bloc.

Prelucrarea secvențială a pixelilor pe blocuri are două avantaje:

- pe de o parte, permite subeșantionarea cadrelor la estimarea fundalului, fără a introduce salturi importante ale timpilor de prelucrare;
- pe de altă parte, imperfecțiuni ale procesului de estimare a culorii fundalului generează după substracție pixeli izolați, ce sunt eliminați cu ușurință prin procedurile de post-filtrare descrise la paragraful 3.3. În cazul unui întreg bloc eronat, erorile rezultate ar fi corelate și mult mai dificil de eliminat prin filtrare.

3.5 Rezultate experimentale

Performanțele estimatorului nonparametric recursiv propus în paragraful precedent au fost evaluate calitativ și cantitativ, fiind comparate cu rezultatele obținute prin metoda nonparametrică introdusă de Elgamal [23], bazată pe estimarea FDP în cele N puncte stocate în memoria de cadre pentru fiecare pixel.

3.5.1 Evaluarea calitativă a estimatorului nonparametric recursiv

Imaginea din figura 3.2.(a) a fost obținută prin aplicarea estimatorului nonparametric de referință Elgamal, folosind o funcție nucleu rectangulară, conform ecuațiilor (3.3) și (3.4) și factorul $\alpha = 1$ în ecuația (3.5) care intervine în calculul factorului de scală. Imaginea din figura 3.2.(b) este obținută cu aceeași funcție nucleu, folosind estimatorul nonparametric recursiv de fundal.



Fig. 3.2. Performanțele privind estimarea și urmărirea fundalului:
a) fundal obținut prin metoda Elgamal;
b) fundal obținut prin urmărire, utilizând algoritmul cu translație la medie, propus de doctorand.

Cu toate că în 128 de imagini folosite la estimarea fundalului prezintă probleme severe de reflexie, umbre și trafic intens, ambele imagini de fundal extrase sunt de o calitate bună, adecvată substracției. Imaginile sunt asimptotic identice, cu observația unui efect favorabil de netezire cu păstrarea conturilor prezent la fundalul obținut prin metoda propusă în teză.

Granularitatea mai mare a fundalului extras prin metoda propusă de Elgamal se explică prin natura discretă a estimatei la implementarea lor, datorită faptului că FDP este estimată numai pentru cele N eşantioane color din bufferul de cadre, în timp de algoritmul cu translație la medie caută maximul FDP pe baza unei estimate continue, fundalul estimat, $\mathbf{b}$, nefiind constrâns să coincidă cu nici un vector eşantion. Fundalul estimat prin metoda propusă de doctorand reprezintă o medie condiționată a eşantioanelor din proximitatea maximului, ceea ce explică aspectul puțin mai neted al imaginii. Rezultatele estimării fundalului pentru alte două cadre din aceeași secvență de imagini sunt prezentate în figura 3.3.

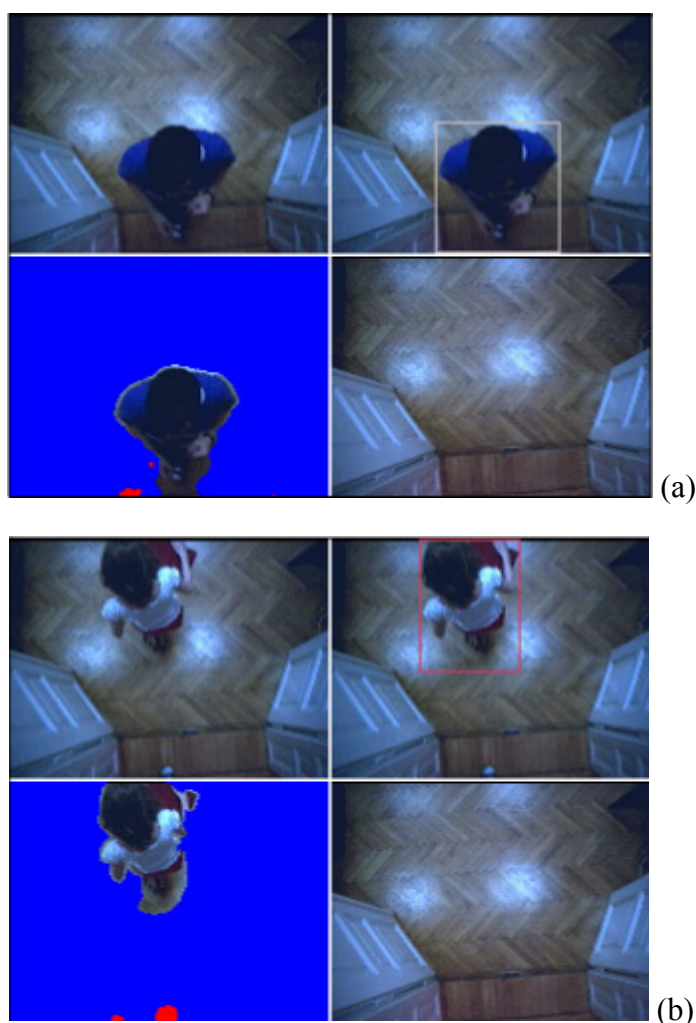


Fig. 3.3. Grupele de imagini a) și b) redau două cadre ale aceleiași secvențe obținute cu factori de scară în raport 1 la 3.

În figura 3.3.(a) factorul de scară s-a estimat cu $\alpha = 1$ în ecuația (3.5), în timp ce în figura 3.3.(b) s-a folosit $\alpha = 3$ în ecuația respectivă.

Fundalul este redat în imaginile din poziția dreapta-jos. Se poate observa că estimatele obținute la doi factori de scară ce diferă semnificativ sunt practic de nedeosebit, ceea ce demonstrează robustețea soluției bazate pe estimator nonparametric al FDP. Imaginile din poziția stânga-jos redau rezultatul substracției fundalului și al segmentării fundal/prim-plan. Fundalul este marcat cu albastru. Cu roșu s-au reprezentat regiunile de prim-plan invalidate prin dimensiune, obiectivele urmărite (persoane) fiind presupuse mai mari. Cadrul de imagine curent este redat în figura 3.3 în poziția stânga-sus. Aceeași imagine, cu dreptunghiul de încadrare ce marchează detecția, este prezentată în poziția dreapta-sus. În stânga-sus este redat cadrul curent iar în stânga-jos se prezintă rezultatul segmentării fundal / prim-plan. Imaginile din dreapta-jos redau fundalul estimat iar imaginile din dreapta-sus redau obiectele mobile valide extrase, marcate cu un dreptunghi.

3.5.2 Evaluarea cantitativă a estimatorului nonparametric recursiv

Pentru atestarea cantitativă a calității fundalului fundalului estimat prin metoda pe care propus-o, am generat un fundal constant de 200 de cadre, afectat de zgomot distribuit uniform între $-0,5$ și $0,5$ și am evaluat deviația standard a erorii de estimare a fundalului la metoda propusă de doctorand și la metoda de referință propusă de Elgamal [59], pentru cinci valori ale factorului de scară. Fereastra de timp de estimare a fost stabilită la o lungime de 40 de eșantioane.

Rezultatele testelor comparative sunt redate în figura 3.4, în care este reprezentată deviația standard a erorii de estimare a fundalului pentru cele 5 valori ale factorului de scară (0.2, 0.4, 0.6, 0.8, 1.0). În timp ce pentru estimatorul nonparametric de referință Elgamal deviația standard a erorii de estimare are valori cuprinse între 0,07 și 0,256, deviația standard a erorii de estimare obținută cu estimatorul nonparametric recursiv este sub 0.03. Raportul între deviația standard minimă obținută prin aplicarea estimatorului de referință Elgamal și deviația standard maximă obținută prin aplicarea estimatorului recursiv este de peste 2:1, adică de peste două ori mai mare.

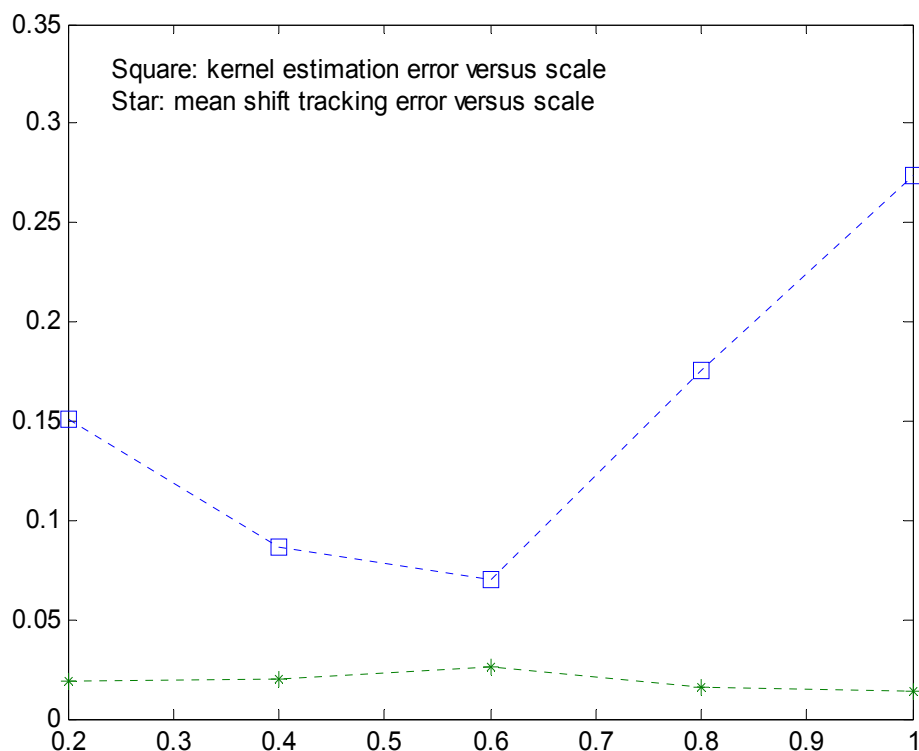


Fig. 3.4. Rezultate ale estimatorului nonparametric cu nucleu și ale estimatorului nonparametric cu urmărire propus de doctorand: deviația standard a erorii de estimare a fundalului calculată pentru cinci factori de scară diferiți.

Comparând medianele erorilor de estimare rezultate din experiment, obținem un raport de aproximativ 6 în favoarea soluției propuse.

De precizat că estimatorul cu urmărire a fost inițializat din poziție corectă, ceea ce ar putea fi considerat un start avantajat. Dincolo de acest aspect, dispersia mai mică a estimatorului pe care l-am propus confirmă:

- calitate lui și
- **netezimea superioară a fundalului estimat** prin metoda propusă.

Pentru a obține o imagine suplimentară referitoare la funcționarea estimatorilor de fundal, în figura 3.5 se redă comparativ secvența erorilor de estimare obținute prin aplicarea celor două metode. Rezultatele au fost consemnate pentru un factor de scală de 0.6, care corespunde celei mai mici deviații standard la estimatorul cu nucleu și celei mai mari la metoda cu urmărire propusă de doctorand (conform figurii 3.4).

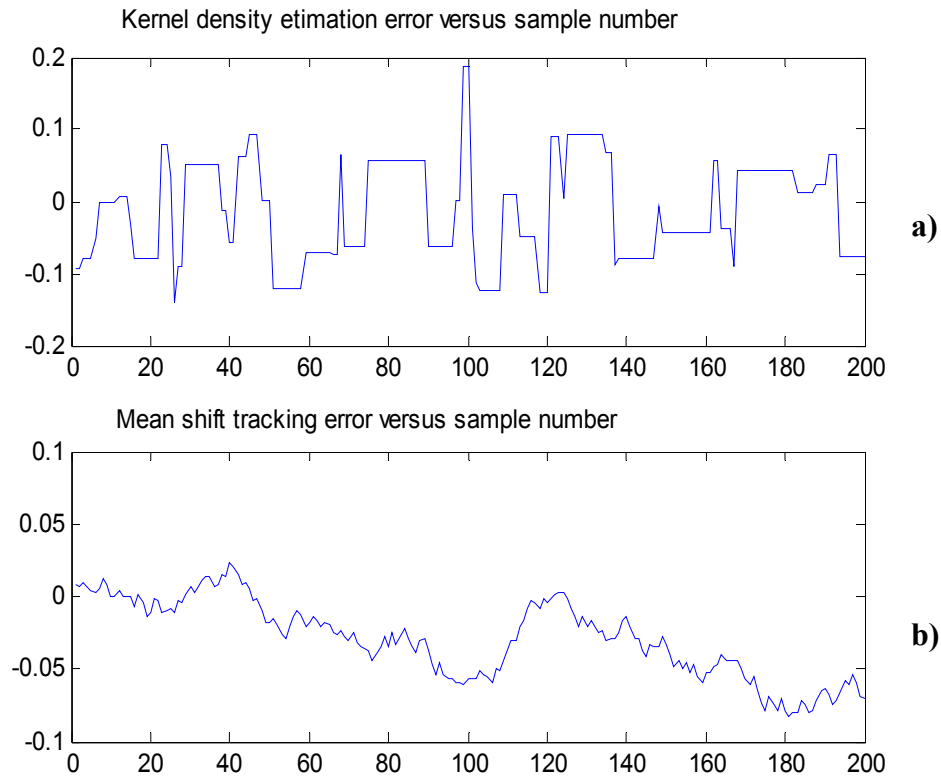


Fig. 3.5. Secvența erorilor de estimare instantanee pentru 200 de eșantioane la:
a) estimatorul nonparametric cu funcție nucleu și
b) estimatorul cu urmărire a fundalului propus de doctorand.

Se remarcă stabilitatea superioară a estimatorului cu urmărire pe care l-am propus.

3.5.3 Evaluarea vitezei estimatorului nonparametric recursiv

În situația cea mai defavorabilă, complexitatea de calcul a metodei propuse este $O(2N)$, la fel ca la estimarea lui Elgamal, bazată de Transformare Gauss Rapidă [59]. Argumentul forte în favoarea estimatorului recursiv pe care l-am propus este acela că asemenea situații defavorabile sunt extrem de infrecvente, așa cum o confirmă experimentele efectuate pentru numeroase secvențe de imagini de supraveghere și testul sintetic descris în continuare.

Testul simulează o schimbare radicală a fundalului, prin inserția unui pătrat de 32×32 pixeli într-o poziție centrală în imagine, pe durata primelor $(N/2 + 16) = 80$ de

cadre. Fundalul sintetic a fost inițial detectat ca fundal real la locațiile celor 1024 de pixeli aferenți. Treptat, fundalul sintetic a fost înlocuit de către modulul de estimare și urmărire a fundalului cu fundalul real, ce diferă radical de culoarea neagră a fundalului sintetic, după un număr de cadre.

În figura 3.6 se prezintă evoluția în timp a numărului de apeluri ale buclei lungi de estimare, cu complexitate de calcul $O(2N)$, ca un procentaj din numărul total de pixeli de fundal modificați.

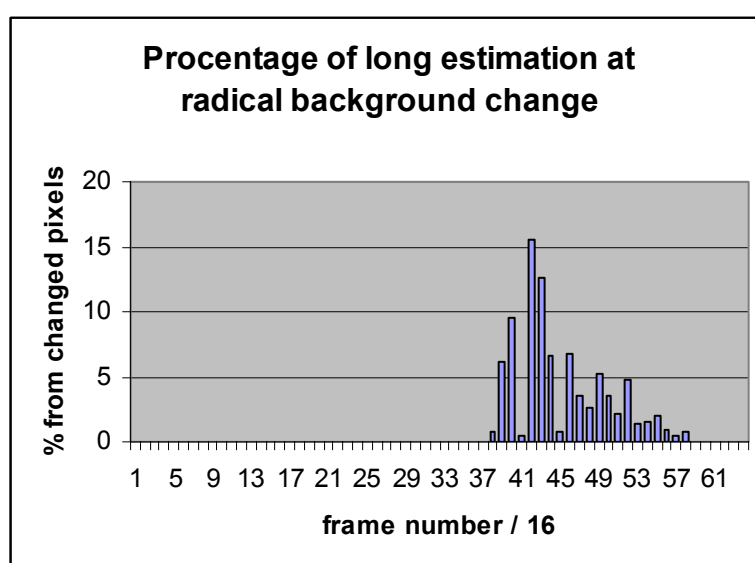


Fig. 3.6. Procentul de apeluri ale buclei lungi de calcul, $O(2N)$, pentru estimarea fundalului într-un experiment ce implică o schimbare radicală a fundalului, obținută prin scoaterea unui obiect fix din scenă. Fiecare bară corespunde unui grup de 16 cadre, necesare pentru actualizarea completă a fundalului.

Se poate observa că, exceptând o perioadă tranzitorie de circa 20 de cadre, bucla lungă a fost practic nefolosită, ceea ce confirmă și testele referitoare la timpul mediu de calcul per pixel la estimarea fundalului, care a fost de 1,5 microsecunde per pixel în cazul folosirii unui calculator cu procesor Pentium 3 cu tact de 750 MHz. Această valoare este independentă de lungimea bufferului de memorie cadre.

La o secvență de imagini cu rezoluția 352×240 , la o subeșantionare spațio-temporală de $1/16$ pentru urmărirea fundalului, rezultă un timp de procesare de circa 8 ms pe cadru, ceea ce asigură rulare în timp real și un timp de procesare confortabil disponibil pentru procesările de nivel înalt.

Comparația referitoare la viteza de calcul a fost completată prin rularea pe același calculator a versiunii celei mai rapide raportate curent în literatură, a Transformatei Gauss Rapide, denumită de autori *Îmbunătățită* și pusă la dispoziție cu generozitate de autorii ei, Yang et al. [60]. Timpul de calcul per pixel la un buffer de memorie de 128 de cadre a fost de 0,387 ms per pixel, ceea ce înseamnă un factor mai mare de 200 față de metoda propusă de doctorand, bazată pe urmărire.

Având în vedere frecvența extrem de scăzută a schimbărilor de fundal drastice, se poate concluziona că metoda de estimare și urmărire a fundalului propusă de doctorand are o complexitate ce nu depinde de lungimea bufferului de memorie cadre folosit la estimare, adică o complexitate de calcul $O(N^0)$. Chiar dacă se folosește Transformarea Gauss Rapidă (TGR), complexitatea de calcul la aplicarea directă a estimării nonparametrice a FDP pentru estimarea fundalului este fundamental mai mare, respectiv $O(2N)$.

Comparația referitoare la viteza de calcul a demonstrat obținerea unei performanțe de peste 200 de ori mai mare a estimatorului nonparametric recursiv de fundal față de metoda generală ce utilizează Transformarea Gauss Rapidă *Îmbunătățită*. Este de subliniat că această comparație nu diminuează cu nimic meritele generale ale Transformării Gauss Rapide *Îmbunătățite*. Comparația este menită doar a confirma faptul că soluția de estimare nonparametrică a fundalului propusă de doctorand nu poate fi implementată mai rapid folosind metoda mai generală a Transformării Gauss Rapide *Îmbunătățite*. Transformarea Gauss Rapidă devine cu adevărat utilă pentru un număr N de eșantioane sursă și țintă de estimare mult mai mare, situație prezentă în numeroase alte aplicații, inclusiv în unele de vedere artificială.

Viteza superioară a metodei de estimare și substracție a fundalului propusă de doctorand rezultă în esență din formularea recursivă a problemei, prezentă de altfel și la abordările parametrice.

3.6 Concluzii

Metoda de substracție a fundalului prezentată în acest capitol, bazată pe estimatorul nonparametric recursiv, conduce la fundaluri estimate cu o calitate vizuală similară cu cele obținute prin metoda de referință Elgamal. Totuși, la o analiză vizuală mai atentă, se poate observa un efect de netezire cu păstrarea contururilor prezent la fundalul obținut prin metoda propusă în teză, ceea ce reprezintă un plus calitativ.

Testele cantitative efectuate au demonstrat obținerea unor eroari de estimare semnificativ mai mici la metoda de substracție a fundalului bazată pe estimatorul nonparametric recursiv comparativ cu eroarile de estimare obținute în urma aplicării metodei de referință Elgamal. Mediana deviației standard a acestor erorilor de estimare, pentru experimentul prezentat în paragraful 3.5.2, este de peste 6 ori mai mică în cazul metodei propuse ce folosește estimatorul nonparametric recursiv.

Rezultatul poate fi explicat prin faptul că:

- estimatorul nonparametric recursiv include o operație de mediere condiționată (în conformitate cu ecuația 3.14);
- teoretic, estimatorul nonparametric recursiv estimează maximul unei densități continue, spre deosebire de estimatorul de referință Elgamal care estimează densitatea numai în punctele spațiului caracteristic pentru care există eşantioane de date.

În paragraful 3.5.3 a fost comparată viteza de calcul pentru estimatorul nonparametric recursiv cu cea corespunzătoare estimatorului de referință Elgamal.

Experimentele prezentate au confirmat faptul că estimatorul nonparametric recursiv implică o complexitate de calcul $O(N^0)$, adică independentă de dimensiunea bufferului de date folosit la estimare. Timpul de calcul pentru estimarea recursivă a fundalului obținut este de peste 200 de ori mai mare decât cel necesar în cazul utilizării estimatorului ce folosește metoda generală bazată pe Transformarea Gauss Rapidă Îmbunătățită (TGRI) în implementarea Yang.

În sinteză, estimatorul nonparametric recursiv elaborat în cadrul capitolului 3 asigură o funcționare precisă și rapidă, adecvată aplicațiilor de supraveghere în timp real.

Extensia spre aplicații ce utilizează vectori caracteristici combinați de culoare, textură și adâncime [2] nu implică o modificare de principiu a metodei, fiind un posibil argument suplimentar pentru adoptarea ei.

CAPITOLUL 4.

OPTIMIZĂRI ALE ESTIMATORULUI NONPARAMETRIC RECURSIV DE FUNDAL

Pornind de la necesitatea de a asigura o funcționare cât mai precisă a estimatorului nonparametric recursiv în condițiile schimbărilor fundalului scenei supravegheate, întâlnite într-o gamă largă de aplicații, în acest capitol sunt descrise două soluții noi care au condus la îmbunătățirea preciziei de estimare. Prima soluție vizează optimizarea estimatorului nonparametric recursiv la schimbări moderate de fundal, iar cea de-a doua soluție optimizează estimatorul în cazul schimbărilor drastice de fundal. Ambele soluții vizează comportamentul dinamic al estimatorului. Preliminar introducerii soluțiilor menționate se întreprinde o optimizare a formei nucleului folosit în dezvoltarea estimatorului recursiv.

Capitolul este structurat în șase paragrafe. În primul sunt trecute în revistă tehnicile adaptive de estimare a fundalului folosite curent. Paragraful 4.2. face o dezvoltare teoretică a ecuației de actualizare a fundalului, care este derivată din algoritmul de detecție a modului cu deplasare la medie pentru cazul unui nucleu de formă generală. Analiza efectului formei nucleului pentru estimarea funcției densitate de probabilitate asupra performanțelor statice și dinamice ale estimatorului nonparametric recursiv este subiectul paragrafului 4.3. În paragraful 4.4 este dezvoltată o variantă îmbunătățită a estimatorului recursiv pentru urmărirea mai rapidă a schimbărilor moderate de fundal. Acest nou estimator l-am denumit *autoadaptiv*. Performanțele noului estimator sunt evaluate comparativ cu cele ale estimatorului neadaptiv. Răspunsul estimatorului la schimbări drastice ale fundalului scenelor supravegheate este îmbunătățit semnificativ prin dezvoltarea unei versiuni noncauzale, descrise în paragraful 4.5. Concluziile capitolului sunt incluse în paragraful final.

4.1 Tehnici adaptive de estimare a fundalului

Adaptarea presupune evaluarea unor informații locale, regionale sau globale care pot fi asociate direct schimbărilor din secvența de imagini în raport cu care se urmărește adaptarea. Aceste informații pot fi clasificate în raport cu nivelul de prelucrare la care sunt extrase, în:

- informații extrase la nivel de prelucrare de pixel;
- informații extrase la nivel de prelucrare intermediar;
- informații extrase la nivel de prelucrare semantic.

Informațiile extrase la nivel de pixel sunt culoarea, intensitatea, adâncimea (distanța) și cele derivate din acestea: gradienti, muchii, texturi, câmp optic de mișcare. Nivelul de prelucrare intermediar corespunde etapei de segmentare și extragere de informații regionale, ce caracterizează segmentele: culoare, intensitate, viteză, traiectorie, diametru, arie, parametri de formă etc.

Informațiile extrase la nivel semantic provin de la module de prelucrare semantică. Aceste module exploatează informații extrase după ce etapele de detecție a fundalului, segmentare a obiectelor mobile și interpretare a mișcării au fost deja parcurse într-o primă etapă. Ele acționează de o manieră similară reacției negative globale, în sensul că mărimea de reacție este măsurată direct la ieșirea procesului. Indiciile folosite la nivel semantic prezintă specificitate pronunțată, în sensul că tind să fie mai puternic legate de particularitățile aplicației. Motivul este foarte simplu: nivelul semantic este necesar pentru obținerea acestor informații. Prin comparație, indiciile extrase la nivel intermediar sau la nivel de pixel tind să fie cu utilitate mai generală și în consecință mai puțin legate de o aplicație anume. Nivelul intermediar utilizează informații de natură statistică sau caracteristici extrase din imagini: muchii, gradienti, câmp de mișcare, traiectorii ale unor segmente.

În cele ce urmează se prezintă **soluții elaborate de doctorand în cadrul cercetării curente, referitoare la adaptarea estimatorului de fundal la schimbări de iluminare**. Sunt avute în vedere numai soluții ce lucrează la nivel de pixel și nu exploatează nici un fel de informații extrase la niveluri de prelucrare superioare. Sunt soluții intrinseci tehnicii de estimare și le consider potențial utile într-o categorie de

aplicații ce exced estimarea fundalului. Tehnicile adaptive bazate pe informații extrase la niveluri de prelucrare ierarhic superioare sunt complementare și pot fi combinate de o manieră directă cu soluțiile propuse în paragrafele următoare. Având în vedere faptul că adaptarea se bazează exclusiv pe informația brută existentă în fluxul de date la un pixel, vom denumi estimatorul autoadaptiv. Estimatorul cu urmărire poate fi adaptat în privința următorilor parametri funcționali:

- parametrii de scară ai funcției nucleu [73], [77], [100], [101];
- rata de învățare.

4.2 Ecuția de actualizare a fundalului

Ecuția de actualizare a fundalului, care stă la baza studiului experimental ce urmează, este derivată din algoritmul de detecție a modului cu deplasare la medie (mean shift) [91], [99]. Fie $\mathbf{b}_{old}$ vectorul de fundal curent estimat, cu funcția densitate de probabilitate estimată cu profilul gaussian $g()$ ca:

$$\hat{p}_G(\mathbf{b}_{old}) = \frac{1}{N} \sum_{i=1}^N g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) \quad (4.1)$$

Presupunem că s-a înregistrat un nou cadru și eșantionul nou este în regiunea activă a nucleului estimatorului centrat pe fundalul curent. Pornind de la $\mathbf{b}_{old}$, mi-am propus să găsesc noul maxim al funcției densitate de probabilitate a fundalului folosind algoritmul mean shift. Ecuția iterată de algoritm este:

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}, \quad j = 1, 2, \dots \quad (4.2)$$

Rezultatul primei iterații se obține înlocuind în ecuația (4.2) vectorul $\mathbf{y}_j$ cu fundalul precedent $\mathbf{b}_{old}$ și adăugând contribuția eșantionului nou:

$$\begin{aligned}
 \mathbf{b}_{new} &= \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + \mathbf{x}_{new} g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)} = \\
 &= \frac{\sum_{i=1}^n \mathbf{x}_i g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right)} \cdot \frac{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)} + \\
 &+ \frac{\mathbf{x}_{new} g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)} \\
 &\quad (4.3)
 \end{aligned}$$

Introducând notația:

$$\begin{aligned}
 \alpha &= \frac{g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)} \\
 &\quad (4.4)
 \end{aligned}$$

și observând că $\mathbf{b}_{old}$ este punctul de convergență al algoritmului mean shift la cadrul precedent, se obțin următoarele ecuații de actualizare a estimatei fundalului:

$$\begin{aligned}
 \mathbf{b}_{new} &= (1 - \alpha)\mathbf{b}_{old} + \alpha\mathbf{x}_{new}, \\
 &\quad (4.5)
 \end{aligned}$$

$$\begin{aligned}
 \alpha &= \alpha_0 g\left(\left\|\frac{\mathbf{x}_{new} - \mathbf{b}_{old}}{h}\right\|^2\right) \\
 &\quad (4.6)
 \end{aligned}$$

$$\begin{aligned}
 \alpha_0 &= \frac{1}{\sum_{i=1}^n g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_i}{h}\right\|^2\right) + g\left(\left\|\frac{\mathbf{b}_{old} - \mathbf{x}_{new}}{h}\right\|^2\right)} = \frac{1}{(n+1)\hat{p}_G(\mathbf{b})} \\
 &\quad (4.7)
 \end{aligned}$$

Factorul α poate fi interpretat ca o *rată de învățare* dependentă de distanța dintre eșantionul nou înregistrat și fundalul precedent estimat.

În implementarea estimatorului de fundal cu urmărire, am utilizat un factor α_0 constant. Motivația nu este legată de prelucarea mai rapidă ci de observația că creșterea ratei de învățare în regiunile cu densitate a fundalului estimată mai mică are drept consecință creșterea ratei de învățare în regiunile cu activitate (mișcare) intensă, în care fundalul este mai rar vizibil. O asemenea regulă ar conduce la creșterea riscului de eroare, introdus de obiecte mobile cu culoare similară fundalului. Chiar dacă un asemenea risc este relativ redus, el poate fi evitat ușor, cu avantajul suplimentar al simplificării calculelor.

4.3 Alegerea nucleului pentru estimarea funcției densitate de probabilitate

În estimarea nonparametrică a densității, se folosesc în mod curent nucleul uniform, nucleul Epanechnikov și nucleul Gaussian. Opinia larg răspândită este că forma funcției nucleu este puțin importantă, esențială fiind alegerea bine fundamentată a factorului de scară. Întrucât și experimentele pe care le-am efectuat în cadrul tezei confirmă faptul că alegerea factorului de scară este cea mai importantă în privința obținerii unor estimate de încredere în aplicații cu număr relativ redus de eșantioane, **mi-am propus un studiu comparativ al funcțiilor nucleu asupra rezultatelor estimatorului cu urmărire** [106]. Studiul este motivat de faptul că estimatorul cu urmărire are o comportare diferită atât față de estimatorii nonparametrici tradiționali cât și față de cei parametrici. La fel ca în cazul estimatorilor parametrici, forma funcției nucleu nu implică nici o concluzie referitoare la forma distribuției. În același timp, ecuațiile de actualizare a estimatei, prezintă unele similitudini cu estimarea parametrică, sensibilă la abaterile de la forma presupusă a distribuției. De subliniat însă că ecuațiile de actualizare a fundalului pe care le folosim sunt derivate din algoritmul de optimizare cu deplasare la medie (mean shift). În deducerea algoritmului mean shift, se face ipoteza că funcția nucleu este derivabilă, iar iterațiile se fac cu ajutorul profilului derivat. Din

acest motiv, am exclus din studiul prezent nucleul uniform, care este nederivabil la margini și are derivata nulă pe domeniul activ.

Pentru nucleul Epanechnikov, derivata profilului este constantă pe domeniul activ al profilului și ecuația se reduce la forma tradițională de actualizare a mediei la estimarea gaussiană parametrică, în care rata de învățare este constantă. În schimb, pentru profilul gaussian, profilul derivat este, de asemenea gaussian și în consecință rata de învățare este variabilă. Mai precis, rata de învățare este mai mare la eșantioanele din imediata vecinătate a estimatei și scade după o lege gaussiană de distanță la eșantioanele mai depărtate. Este o calitate prezentă la toți estimatorii robuști. De observat că și estimatorul pentru profilul Epanechnikov este robust, pentru că rata de învățare se anulează pentru eșantioanele ce ies din fereastra operatorului. Totuși, estimatorul cu profil gaussian are o proprietate mai puternică, fiind variabil în interiorul ferestrei active. Este de anticipat o sensibilitate mai redusă la zgomot, dar și la alegerea factorului de scară al estimatorului. Este de așteptat ca estimatorul cu profil Epanechnikov să răspundă mai rapid la schimbările fundalului.

4.3.1 Rezultate experimentale

Pentru o comparație cât mai obiectivă, am prevăzut două teste: unul static și unul dinamic.

În primul experiment, am generat un fundal constant, afectat de zgomot alb, distribuit uniform între -0,5 și 0,5. Secvența de test a constat din 200 de eșantioane afectate de zgomot. Parametrul de scară din ecuația (4.6) a fost ales $h = 0,33$ pentru estimatorul gaussian, trunchiat la implementare la lungimea $3h$. Pentru estimatorul bazat pe nucleul Epanechnikov am ales valoarea $h = 1$, pentru a obține același domeniu activ. Am evaluat deviația standard a erorii de estimare pentru ambele nuclee. Rezultatele a cinci experimente succesive sunt redate în figura 4.1. Sunt confirmate așteptările teoretice expuse în preambul.

În al doilea experiment, am folosit un fundal dinamic, conținând o treaptă cu amplitudinea 0,1 și același zgomot aditiv ca în primul experiment. Rezultatele sunt redate grafic în figura 4.2. Se observă o creștere ușoară a erorilor de estimare la ambii

estimatori, datorate procesului tranzitoriu, dar erorile rămân considerabil mai mici (aproximativ de cinci ori) la estimatorul Gaussian.

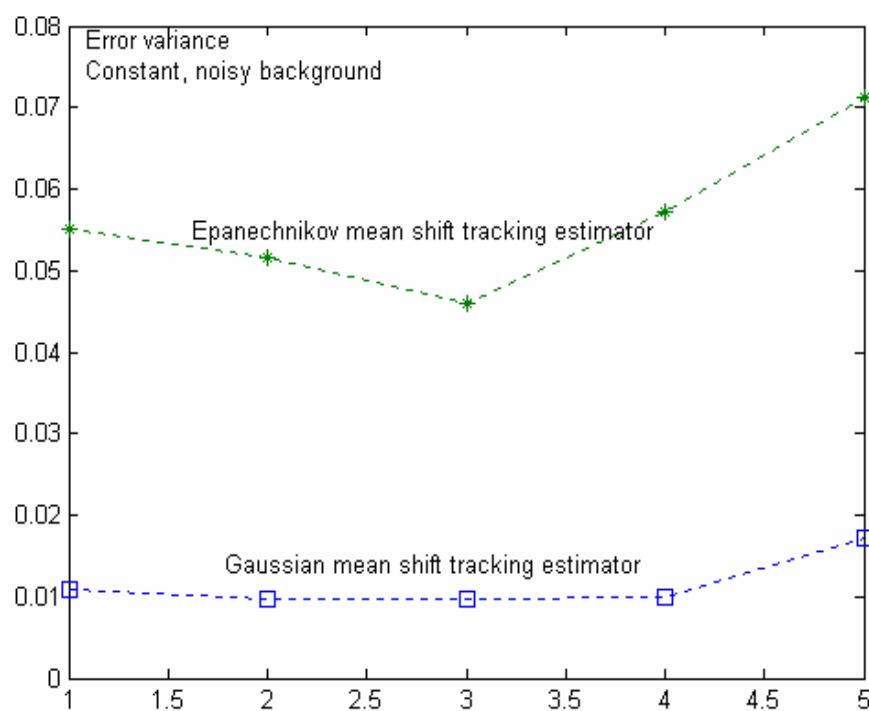


Fig. 4.1. Rezultate comparative ale testelor pentru estimatorii cu nucleele Epanechnikov și Gaussian la fundal constant, cu zgomot alb aditiv.

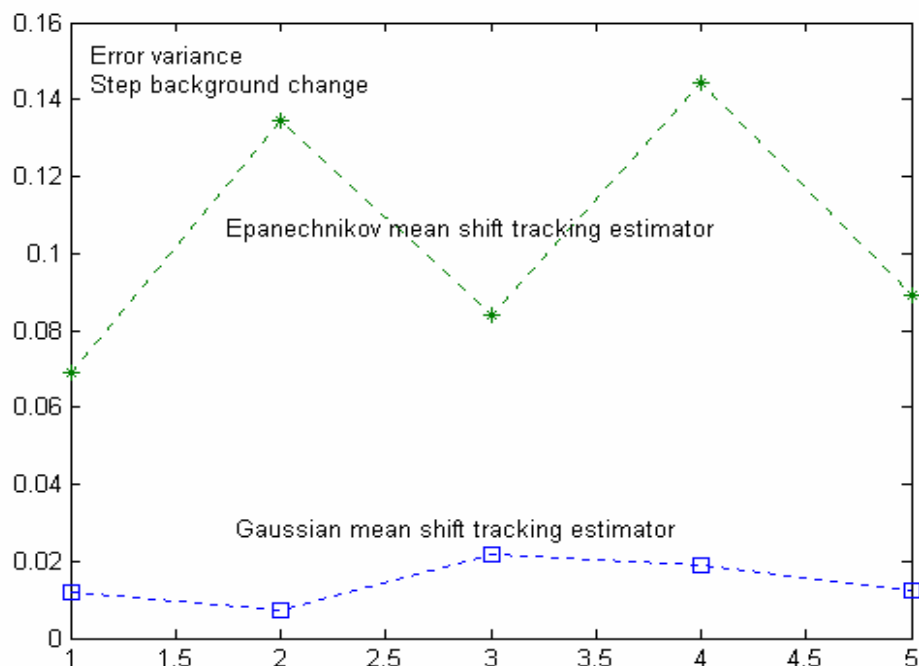


Fig. 4.2. Rezultate comparative ale testelor pentru estimatorii cu nucleele Epanechnikov și Gaussian la fundal dinamic, cu zgomot alb aditiv.

Având în vedere faptul că testele comparative au fost efectuate pentru ferestre de estimare de dimensiuni egale, diferența de performanță în favoarea estimatorului cu fereastră de tip Gaussian se poate explica teoretic prin profilul diferit al acestuia. Ponderarea mai slabă a eșantioanelor ce prezintă diferențe mai mari în raport cu estimata curentă aduce stabilitate, respectiv sensibilitate la zgomot mai mică pentru estimatorul cu urmărire de mod ce folosește fereastra de tip Gaussian.

Conform raționamentului de mai sus, este de așteptat ca stabilitatea superioară a estimatorului cu nucleu Gaussian să se obțină în detrimentul vitezei de răspuns la schimbările reale ale fundalului. Experimentele testelor cu fundal dinamic arată însă că efectul este neînsemnat cantitativ și estimatorul cu urmărire cu nucleu de profil Gaussian rămânând net superior și în condiții de schimbare a fundalului. Raportul dispersiei erorilor se reduce totuși de la aproximativ 6:1 la teste statice la aproximativ 4:1 pentru testele dinamice indicând o ușoară diminuată a superiorității nucleului Gaussian, ce nu modifică concluzia generală, și anume, că un nucleu cu profil monoton descrescător este preferabil unui profil rectangular

4.4 Estimator nonparametric recursiv autoadaptiv

Ideea de estimare adaptivă a fundalului în sine nu este nouă. De exemplu, în estimarea *parametrică* a fundalului descrisă în [80], activitatea detectată la o locație de un modul de estimare și interpretare a mișcării este asociată cu reducerea sau blocarea completă a procesului de estimare a fundalului pe durata cât acesta este obturat de obiecte active. Astfel, două persoane care s-au oprit temporar pentru a avea o conversație vor fi mai târziu încorporate în fundal. În același timp, un scaun introdus în zona supravegheată va fi încorporat mai rapid în fundal.

În [86], rata de învățare este calculată ca produsul a două funcții. Prima este o funcție de evaluare a “confidenței locale”, definită ca

$$f(d) = \exp(-d^2/2\sigma^2) \quad (4.8)$$

unde d este diferența dintre eșantionul curent și media locală estimată. Observăm că acest factor există în mod natural, în virtutea metodei și în varianta de bază a estimatorului nonparametric și în estimatorul cu urmărire dezvoltat în prezentul capitol. În [86], factorul suplimentar de adaptare apare ca o generalizare la algoritmul de estimare parametrică EM. Al doilea factor în calculul ratei de învățare introdus în [86] este o funcție de corelație globală folosit cu intenția de a detecta mișcări de rotație ale camerei. Asemenea mișcări sunt excluse din studiul prezent, care se referă la aplicațiile de supraveghere cu camere fixe.

4.4.1 Adaptarea la schimbări ale fundalului de amplitudine moderată

Tehnicile de adaptare a estimatorilor, menționate mai sus, pot fi încorporate de o manieră directă și în implementarea descrisă în continuare, fiind complementare. Ceea ce urmărim aici este să creștem viteza de urmărire a estimatorului de fundal în situațiile în care apare o schimbare reală și rapidă de amplitudine moderată a fundalului, pe care estimatorul nonparametric recursiv nu o poate urmări suficient de prompt cu rata de învățare normală. O situație de acest gen apare frecvent la scenele în aer liber ca rezultat al mișcării norilor, sau în scene de interior când se mai aprinde sau stinge un bec.

Scopul fiind diferit față de adaptarea folosită de metodele discutate anterior, **rezultă o metodă de adaptare diferită, propusă de doctorand în [107]**. Pentru estimarea fundalului în situații în care se produc schimbări de amplitudine mare ale iluminării sau înlocuiri ale fundalului prin inserția sau îndepărtarea unor obiecte, se propune, de asemenea, o soluție nouă. Soluția este bazată pe o fereastră de estimare noncauzală și este descrisă în paragraful 4.5.

Pentru a se obține un estimator nonparametric asimptotic perfect este nevoie de o fereastră îngustă și un număr de eșantioane foarte mare. În cazul estimatorului cu urmărire, un număr de eșantioane foarte mare conduce la o rată de învățare foarte redusă. O rată de învățare foarte redusă reduce substanțial efectul zgomotului asupra valorii estimate pentru fundal. În același timp însă reducerea ratei de învățare are drept consecință și o întârziere mai mare în schimbarea modelului fundalului atunci când apar schimbări reale ale acestuia, cum sunt cele produse de schimbări ale iluminării. Aparent, numărul de eșantioane din memoria de cadre de la un estimator nonparametric convențional, respectiv rata de învățare la estimatorul cu urmărire, trebuie selectate prin optimizarea unui compromis între doi factori: stabilitatea estimatei (caracterizată printr-o varianță a erorii redusă) în condiții statice, pe de o parte, și eroarea de neurmărire redusă în condiții dinamice, cum sunt cele generate de schimbările iluminării fundalului.

Soluția pe care o propunem permite ieșirea din dilemă pornind de la observația că rata de învățare poate fi schimbată în funcție de situație, adică selectată adaptiv.

O metodă de adaptare bună trebuie să fie capabilă să discrimineze între schimbările fundalului provocate de zgomot și cele reale. Deosebirea fundamentală între cele două este că schimbările produse de zgomot tind să fie aleatoare, în timp ce cele produse de evenimente reale, de exemplu de schimbarea iluminării, sunt persistente. Acest fapt poate fi exploatat în modul descris în continuare.

Pentru discriminarea între cele două tipuri de schimbare (zgomot/schimbări reale) se introduce ca indiciu semnificativ *eroarea de neurmărire cumulată*, $\mathbf{d}_{cum}$ dintre estimată și datele de intrare. Notând cu $\mathbf{b}(t)$ estimata fundalului la timpul discret t , prin definiție, eroarea de neurmărire cumulată la momentul t este:

$$\mathbf{d}_{cum}(t) = \sum_{i=0}^t [\mathbf{x}(i) - \mathbf{b}(i)]$$

(4.9)

Ideea de bază este că zgomotul produce modificări aleatoare ale erorii de neurmărire cumulate. Media acestor erori este nulă și în consecință aceste erori nu afectează semnificativ eroarea de neurmărire cumulativă, o mărime obținută în urma unui calcul integrativ. Prin contrast, o schimbare reală a fundalului care produce apariția unei erori de neurmărire sistematice, determină la fiecare cadru nou modificarea erorii de neurmărire cumulative în același sens.

Prin compararea normei vectorului erorii de neurmărire cumulative cu un prag, $\mathbf{d}_{th}$, ales judicios, putem detecta cu ușurință situațiile în care estimatorul rămâne în urmă prea mult și pentru prea multe cadre. Când o asemenea situație este detectată, se propune înlocuirea estimatei curente a fundalului cu eșantionul curent. Concomitent, eroarea de neurmărire cumulativă este reinițializată la zero, pentru a se șterge efectul produs de o succesiune de evenimente care au fost deja tratate și a se începe monitorizarea unui nou ciclu de schimbări. Operația de schimbare a estimatei cu eșantionul curent corespunde învățării cu rata $\alpha = 1$.

O descriere prin pseudo-cod a funcționării estimatorului cu urmărire de mod adaptiv este dată în figura 4.3.

```

if(  $K_h(\mathbf{x}_{new} - \mathbf{b}) \neq 0$  )
    update(  $\mathbf{b}$  and  $\hat{p}(\mathbf{b})$  );
     $\mathbf{d}_{cum} = \mathbf{d}_{cum} + \mathbf{x}_{new} - \mathbf{b}$ ;
    if(  $\|\mathbf{d}_{cum}\| > \mathbf{d}_{th}$  )
         $\mathbf{b} =: \mathbf{x}_{new}$ ;
         $\mathbf{d}_{cum} = \mathbf{0}$ ;
    end if
else if(  $\text{Hist}(\mathbf{x}_{new}) > \text{threshold}$  )
    and (  $\hat{p}(\mathbf{x}_{new}) > \hat{p}(\mathbf{b})$  )
         $\mathbf{b} =: \mathbf{x}_{new}$ ;
    end else if
end if

```

Fig. 4.3. Descriere prin pseudo-cod a funcționării estimatorului cu urmărire de mod adaptive.

Schimbarea estimatei fundalului cu eșantionul curent pare o operație riscantă. Un detaliu important care elimină riscul unor erori extreme este acela că un eșantion curent contribuie la calculul erorii de neurmărire numai dacă se găsește în fereastra activă a estimatorului, centrată pe estimata curentă, ceea ce limitează domeniul de influență al eșantioanelor noi. Eroarea maximă la schimbarea fundalului cu rata de învățare $\alpha = 1$ nu poate depăși jumătate din lățimea ferestrei.

În practică asemenea erori sunt posibile, dar puțin probabile, așa cum o atestă și rezultatele experimentelor efectuate pentru evaluarea cantitativă a efectelor adaptării ratei de învățare a estimatorului, descrise la paragraful următor.

4.4.2 Rezultate experimentale pentru evaluarea performanțelor estimatorului nonparametric recursiv autoadaptiv

Performanțele estimatorului nonparametric cu urmărire de mod neadaptiv au fost deja evaluate și descrise în [105]. S-a constatat că estimatorul cu urmărire are eroarea medie patritică semnificativ redusă în comparație cu estimatorul nonparametric cu nucleu convențional, prin experimente realizate pe o mare varietate de scări. Din acest motiv, în analiza performanțelor estimatorului nonparametric recursiv (cu urmărire de mod) autodaptiv la schimbări de fundal de amplitudine moderată am folosit ca referință exclusiv performanțele estimatorului nonparametric cu urmărire de mod de bază, neadaptiv.

Testele au fost efectuate atât în condiții de fundal static cât și în condiții de fundal dinamic [107]. În conformitate cu așteptările teoretice, testele statice au condus la rezultate asimptotic identice pentru cei doi estimatori, diferențele fiind statistic ne semnificative. Din acest motiv, prezentăm în continuare numai rezultatele experimentelor cu fundal dinamic.

În primul experiment descris, am generat un fundal dinamic unidimensional de tip treaptă. Semnalul, compus din 400 de eșantioane, a fost afectat de zgomot alb aditiv cu distribuție gaussiană având deviația standard egală cu 10 % din amplitudinea treptei. Pentru simplitate, s-a folosit o treaptă unitară. Fundalul a fost estimat pe baza acestui semnal folosind atât estimatorul cu urmărire de mod neadaptiv cât și cel adaptiv. Ambii estimatori au folosit nuclee gaussiene trunchiate, cu parametrul de scară $h = 2\sigma^2 = 1$ și rata de învățare normală cu parametrul $\alpha_0 = 0,02$ în ecuația (4.6). Pragul de discriminare pentru eroarea de neurmărire cumulativă a estimatorului adaptiv s-a stabilit la nivelul $\mathbf{d}_{th} = 3h = 3$. Rezultatele unui asemenea experiment sunt redată în figura 4.4.

Se observă că estimatorul cu urmărire de mod adaptiv a avut nevoie de numai câteva eșantioane pentru a sesiza și a se adapta la prezența treptei. Teoretic, în absența zgomotului, ar fi nevoie de numai trei eșantioane.

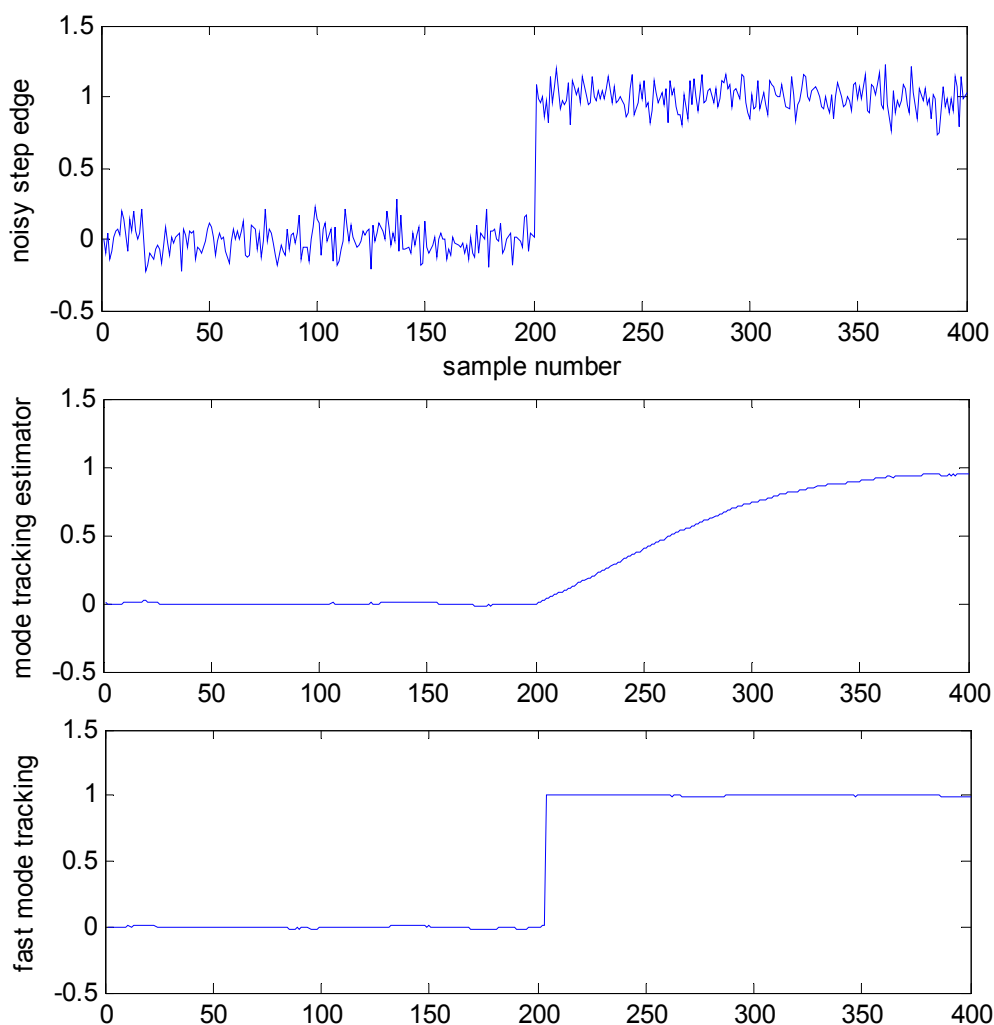


Fig. 4.4. Răspunsul estimatorilor nonparametrici recursivi la modificări moderate, de tip treaptă, ale fundalului.

- (sus): semnalul de fundal de tip treaptă, cu zgomot aditiv;
 (mijloc): răspunsul estimatorului nonparametric recursiv;
 (jos): răspunsul estimatorului nonparametric recursiv autoadaptiv.

În al doilea experiment, am evaluat deviația standard a erorii de estimare în funcție de amplitudinea treptei pentru cei doi estimatori nonparametric recursivi (cu urmărire de mod), adaptiv și neadaptiv, la două niveluri ale deviației standard a zgomotului aditiv, alb, gaussian suprapus semnalului treaptă.

Rezultatele testelor sunt redată grafic în figura 4.5. Pe abscisă este reprezentată amplitudinea treptei, în timp ce pe verticală este reprezentată deviația standard a erorii pe un set de 400 de eșantioane.

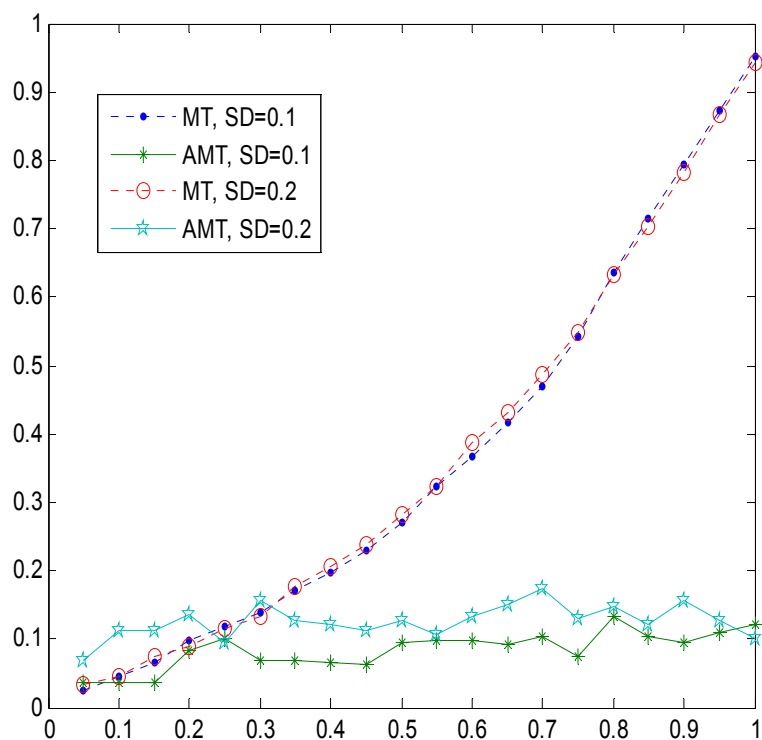


Fig. 4.5. Deviația standard a erorii de estimare a estimatorului cu urmărire de mod și a estimatorului cu urmărire de mod adaptiv în funcție de amplitudinea treptei calculate pentru două deviații standard ale zgomotului $SD = 0,1$ și $SD = 0,2$.

Conform așteptărilor teoretice, la amplitudini ale treptei foarte mici, cei doi estimatori au performanțe egale. Situația este corespunzătoare practic unui fundal static. Pe măsură ce fundalul devine mai dinamic, respectiv amplitudinea schimbărilor crește, fără a depăși domeniul de urmărire al celor doi estimatori, avantajele adaptării ratei de învățare devin mai mari. În timp ce deviația standard a erorii la estimatorul cu urmărire de mod neadaptiv crește cu amplitudinea treptei, ca rezultat al întârzierii mai mari în urmărirea treptei, eroarea este aproape complet independentă la estimatorul cu rată de învățare adaptivă.

4.5 Estimator nonparametric recursiv autoadaptiv noncauzal

Adaptarea estimatorului nonparametric recursiv prezentată în paragraful 4.4 nu rezolvă și problemele ce apar la estimarea fundalului în situații în care se produc schimbări de amplitudine mare ale iluminării sau înlocuiri ale fundalului prin inserția sau îndepărtarea unor obiecte. O soluție la problema menționată pe care am dezvoltat-o în cadrul tezei este descrisă în continuare.

4.5.1 Adaptarea la schimbări drastice ale fundalului

În paragraful curent, se analizează problema estimării prompte a schimbărilor radicale ale fundalului, cum sunt cele determinate de schimbări ale conținutului fundalului prin introducerea sau eliminarea unor obiecte, sau cele produse de schimbări drastice ale iluminării. Asemenea schimbări pun probleme atât metodelor parametrice de estimare a fundalului cât și celor nonparametrice.

Oricât de elaborate ar fi tehnicile de adaptare a parametrilor estimatorului, o problemă de fond rămâne relevantă: pentru obținerea unei informații de încredere referitoare la schimbarea fundalului, este nevoie de acumularea unui număr de cadre relativ mare. În caz contrar, obiecte mobile ce se suprapun temporar peste fundal pot schimba în mod eronat estimația acestuia.

Soluția pe care am propus-o [108] pornește de la observația că răspunsul la întrebarea dacă *o schimbare de iluminare este “permanentă” sau dacă un obiect nou introdus va rămâne sau nu pe termen lung în fundal* ar putea fi găsit mult mai ușor dacă ar exista posibilitatea de a se privi în viitor, respectiv de a se verifica dacă valoarea nouă înregistrată la un anumit pixel urmează a se recepționa pe o durată mai mare și/sau cu o frecvență mai mare în viitor.

La o analiză mai atentă, posibilitatea evocată este realizabilă relativ ușor în cazul estimatorilor nonparametrici de fundal, ce păstrează în memoria estimatorului un număr de N de cadre. Tot ce trebuie făcut este să se decaleze indicele cadrului la care se referă estimata, de exemplu cu $N/2$. Ceea ce propunem de fapt este să folosim pentru

estimarea fundalului la momentul t o fereastră de timp simetrică pentru a include cadrele cu indici între $t-t_w$ și $t+t_w$, pentru o fereastră de lungime $N = 2t_w + 1$.

Din informațiile pe care le deținem, toate metodele de estimare nonparametrică a fundalului se bazează exclusiv pe cadrul curent și un număr de cadre precedente. Există un motiv bine întemeiat pentru această opțiune: estimata pentru cadrul curent este disponibilă instantaneu. Dacă se aplică soluția pe care o propunem, estimata se decalează în urmă cu un număr de $N/2$ cadre.

Utilizarea ferestrelor de prelucrare simetrice este o practică folosită frecvent în prelucrarea imaginilor, unde filtrele noncauzale sunt mai degrabă regula decât excepția. Prelucrarea noncauzală nu prejudiciază în nici un fel funcționarea în timp real, în sensul că viteza de generare a datelor de ieșire poate fi aceeași cu viteza fluxului de date de la intrare. Singura implicație este întârzierea ieșirii în raport cu intrarea. Pentru o gamă largă de aplicații, cum sunt monitorizarea traficului într-o zonă de parcare sau analiza statistică a timpilor de așteptare la deservirea clienților în restaurante de tip fast food, o întârziere de ordinul minutelor în generarea rezultatelor este nederanjantă.

Un avantaj potențial al estimatorului noncauzal este acela, că pentru o valoare dată a intervalului de timp ales pentru definirea setului de cadre folosit în estimarea fundalului, decalajul în timp între cadrul pentru care se estimează fundalul și oricare cadru folosit la estimare este redus la jumătate în raport cu estimatorul cauzal. Mai mult, media decalajului în timp între cadrul curent pentru care se face estimarea și cadrele folosite la estimare este redusă, de asemenea, la jumătate în raport cu estimatorul cauzal. Cu alte cuvinte, datele sunt grupate mai strâns în jurul cadrului estimat.

Este ușor de anticipat faptul că răspunsul estimatorului noncauzal la o schimbare treaptă de amplitudine mare a fundalului este instantaneu, o caracteristică deosebit de puternică și aparent neposedată de niciuna din metodele descrise în literatură (desigur, generat ca la orice prelucrare numerică noncauzală, cu decalajul de timp corespunzător jumătății de fereastră). Decalarea răspunsului estimatorului rezolvă simplu și cu maximă siguranță o problemă vitală: evită erori de estimare grosolane altfel greu de evitat, dacă nu imposibil. În același timp, decalarea răspunsului estimatorului restrânge într-o oarecare măsură aria aplicațiilor care pot beneficia de avantajele

metodei. Este vorba de o subclasă a aplicațiilor de timp real care cer un timp răspuns inferior intervalului de timp utilizat pentru umplerea memoriei cadru.

O complicație minoră în aplicarea estimatorului noncauzal apare în situațiile în care se practică subeșantionarea cadrelor la estimarea fundalului. În implementările noastre curente folosim o subeșantionare cu factorul 1:16. Practic, în fiecare cadru nou, se extrage pentru estimarea fundalului, prin rotație, un singur pixel. Deoarece estimarea fundalului se face independent pentru fiecare locație spațială, acest procedeu nu afectează în nici o formă estimarea fundalului. Are în schimb o serie de avantaje:

- Eventualele erori în estimarea fundalului apar dispersate spațial, ceea ce facilitează enorm eliminarea lor prin tehnicile de filtrare spațială uzuale (morfologice).
- Pentru o dimensiune dată a memoriei cadru, este posibilă extinderea timpului de observare a fundalului, necesară pentru reducerea efectului obiectelor mobile asupra estimatei, în special în condiții de trafic intens și/sau opriri temporare.
- Complementar, pentru o durată de observare dorită la estimarea fundalului, se reduce dimensiunea bufferului de memorie necesar. Reducerea menționată nu poate fi exploatată în cazul estimatorului noncauzal, care are nevoie de toate cadrele din intervalul de timp dintre cadrul curent și cadrul pentru care se face curent estimarea fundalului.
- Subeșantionarea decalată permite repartizarea mai uniformă în timp a operațiilor de calcul ce intervin în estimarea fundalului.

4.5.2 Evaluarea performanțelor estimatorului nonparametric noncauzal

Pentru estimarea performanțelor estimatorului de fundal nonparametric, noncauzal, cu urmărire de mod, am conceput un test cu fundal dinamic 1D, în care se suprapun schimbări ale fundalului de amplitudine moderată cu o schimbare radicală. Pentru mai mult realism, am adăugat și zgomot. În acest fel se pot detecta eventuale consecințe ale estimării noncauzale asupra urmăririi adaptive a schimbărilor de fundal

moderare. Acestea sunt definite prin amplitudinea lor, ce nu depășește domeniul activ al ferestrei de estimare.

În cazul estimatorului gaussian cu fereastra trunchiată la domeniul $\pm 3h$, schimbările moderate de fundal nu pot avea o amplitudine mai mare de $3h$, unde h este parametrul de scară, respectiv dublul dispersiei. Pentru a avea rezultate comparative, fundalul a fost estimat atât cu estimatorul nonparametric, noncauzal, cu urmărire de mod cât și cu estimatorul nonparametric de densitate cu nucleu folosit de Elgamal [23], reprezentând metoda de referință curentă în estimarea nonparametrică a fundalului. Amplitudinea treptei moderate a fost stabilită la nivelul 1, în timp ce amplitudinea treptei de schimbare radicală a fundalului a fost stabilită la nivelul 5. Deviația standard a zgomotului gaussian s-a stabilit la o cincime din amplitudinea treptei moderate, respectiv 0,2. Pentru ambii estimatori, s-a ales același nucleu gaussian, cu parametrul de scară $h = 1$. Lungimea ferestrei de estimare s-a stabilit la valoarea $N = 100$ eşantioane (cadre), în timp ce semnalul generat conține 550 eşantioane. Cifra este aleasă cu obiectivul de a se obține o secvență de 400 de estimate ale fundalului. Primele 100 de eşantioane au fost rezervate inițializării estimatorului cu urmărire. Ultimele 50 de eşantioane sunt incluse exclusiv pentru a fi utilizate de estimatorul noncauzal, în calculul estimatelor cu index 450÷500.

Rezultatele ilustrate în figura 4.6 sunt reindexate între 0 și 400, pentru simplitate.

Întârzierea cu o jumătate de lungime de fereastră ce apare la estimatorul de referință în detecția treptei de amplitudine mare este clar vizibilă în figura 4.6, în contrast cu rezultatele estimatorului noncauzal. De asemenea, estimatorul nonparametric de referință răspunde cu întârziere la treptele de amplitudine moderată, în contrast cu comportarea estimatorului nonparametric, noncauzal, cu urmărire de mod adaptiv.

Se mai poate constata că performanțele menționate sunt obținute fără a se deteriora în vreun anumit fel capabilitățile de rejecție a zgomotului. Acest aspect a fost evaluat, de altfel și de o manieră cantitativă în teste care au precedat experimentul descris.

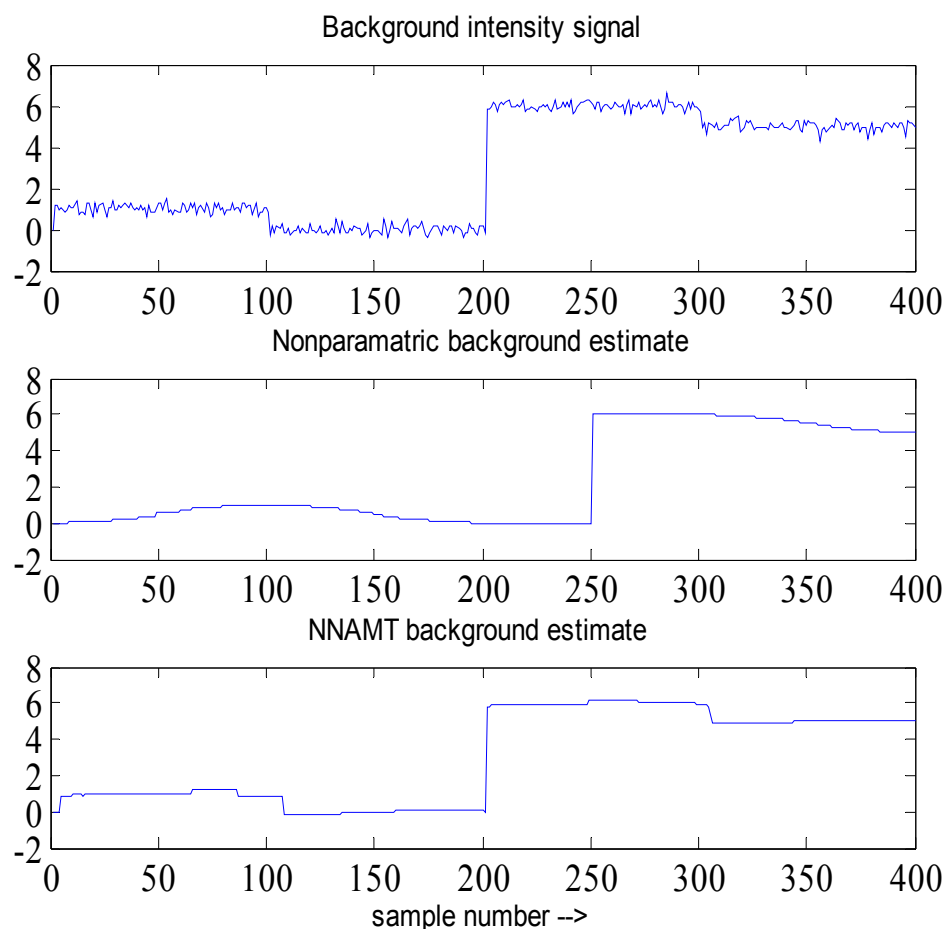


Fig. 4.6. Rezultate comparative pentru estimatorul de fundal nonparametric standard și estimatorul noncauzal, nonparametric, cu urmărire de mod adaptiv.

4.6 Concluzii

În capitolul prezent s-au descris dezvoltări ale estimatorului nonparametric recursiv propus în teză. Acestea au avut în vedere alegerea optimă a formei funcției nucleu pe baza căreia este proiectat estimatorul recursiv, creșterea vitezei de urmărire prin adaptarea la schimbări ale iluminării fundalului de amplitudine moderată, precum și reducerea erorilor de estimare la schimbările drastice ale fundalului.

Fiecare din soluțiile propuse au condus la reduceri semnificative ale erorilor de estimare, ultimele două vizând exclusiv comportarea dinamică a estimatorului nonparametric recursiv.

CAPITOLUL 5.

CONTRIBUȚII ȘI CONCLUZII

Estimarea fundalului în videosecvențe pentru aplicații în supraveghere a fost abordată inițial prin tehnici parametrice. Tehnicile nonparametrice, potențial mai robuste au fost introduse mai recent, o dată cu apariția unor soluții ce contracarează principalul lor dezavantaj: efortul de calcul superior. În acest context, doctorandul a propus estimatorul nonparametric recursiv (cu urmărire de mod), bazată pe un mod de calcul recursiv ce își are rădăcinile teoretice în algoritmul cu deplasare la medie (*mean shift*) pentru detecția maximelor funcțiilor densitate de probabilitate.

În teză s-au descris câteva optimizări ale estimatorului nonparametric recursiv (cu urmărire de mod), care-i conferă o adaptivitate mai bună. S-au avut în vedere două categorii de schimbări ale fundalului: schimbări moderate și schimbări radicale, specifice modificării reale a fundalului. Prima categorie de schimbări de fundal este rezolvată prin creșterea ratei de învățare în modul de urmărire. Pentru cea de a doua, s-a introdus modul de operare noncauzal.

Performanțele estimatorului cu urmărire de mod adaptiv au fost evaluate prin teste statice și dinamice, efectuate comparativ cu metodele de referință și versiunile precedente ale estimatorului cu urmărire de mod. Rezultatele experimentale au confirmat așteptările teoretice, de o manieră calitativă și cantitativă. S-a demonstrat posibilitatea creșterii vitezei de urmărire a schimbărilor fundalului concomitent cu păstrarea stabilității în condiții de funcționare statice.

Soluțiile propuse de doctorand completează tehnicile adaptive cunoscute, putând fi aplicate concomitent cu acestea. În același timp, este de menționat faptul că, deși studiul prezent s-a concentrat exclusiv pe cazul estimatorilor nonparametrici, soluțiile propuse sunt aplicabile potențial și estimatorilor parametrici, cu care estimatorul cu urmărire prezintă, de altfel, numeroase similitudini.

În lucrarea de față se pot evidenția un număr de 9 contribuții principale ale doctorandului, cu caracter de noutate în domeniul detecției și extragerii fundalului în supravegherea video prin tehnici nonparametrice, prezentate sintetic în continuare, fiind însoțite de referințe bibliografice, cu referire la articolele publicate de doctorand.

5.1 Contribuții teoretice

- (1) Analiza modului în care tehnicile de estimare nonparametrică a densității de probabilitate cu funcții nucleu pot fi utilizate în estimarea și substracția fundalului pentru aplicații de supraveghere video.**

Se tratează fundamentele teoretice ale metodelor de estimare nonparametrice. Se include o discuție referitoare la algoritmul cu translație la medie pentru detecția rapidă a modurilor funcției densitate de probabilitate [102], [103], [105].

- (2) Studiu comparativ asupra metodelor de detecție și estimare a fundalului în secvențe de imagini.**

Acest studiu comparativ mi-a permis să definească direcțiile în care se pot aduce îmbunătățiri pentru creșterea robusteții și adaptivității unui estimator de fundal pentru aplicații de supraveghere video [102], [103], [104], [105], [106], [107], [108].

- (3) Metodă de estimare nonparametrică recursivă a fundalului.**

Am elaborat o nouă soluție pentru estimarea nonparametrică a fundalului, cunoscută sub denumirea de metodă de estimare nonparametrică recursivă a fundalului [105]. Metoda propusă reprezintă în esență o metodă rapidă pentru urmărirea schimbărilor de fundal, ce îmbină avantajele tehnicilor de estimare nonparametrică a funcției densitate de probabilitate bazată pe funcții nucleu, cu viteza tehnicilor bazate pe histogramme. **Avantajele** metodei sunt:

- Estimatorul are performanțe neinfluențate de forma reală a distribuției, în contrast cu estimatorii parametrici.
- Conduce la implementări cu funcționare în timp real.

- Soluția propusă se bazează teoretic pe o estimare continuă a funcției densitate de probabilitate și a maximelor locale ale acesteia, rezultând o reducere a erorii de estimare cu un ordin de mărime.
- Se constată experimental stabilitatea rezultatelor estimatorului recursiv la variațiile factorului de scară al estimatorului, într-o plajă largă de valori.

(4) Dezvoltarea teoriei estimatorului nonparametric recursiv.

Pornind de la ecuațiile de funcționare ale estimatorilor nonparametrici cu nucleu și de la algoritmul cu deplasare la medie pentru detecția modurilor funcției densitate de probabilitate se deduc ecuațiile de funcționare ale estimatorului nonparametric recursiv cu urmărire de mod. Ecuațiile (4.5), (4.6), (4.7) au un caracter de noutate.

(5) Demonstrarea complexității de calcul de ordin $O(N^0)$ a estimatorului nonparametric recursiv cu urmărire de mod.

În teză se demonstrează că, independent de aspectele de implementare, complexitatea teoretică de calcul a estimatorului nonparametric recursiv cu urmărire de mod este $O(N^0)$ față de $O(2N)$ a metodei de referință nonparametrică, cea mai rapidă cunoscută până în prezent în literatură [60].

5.2 Contribuții aplicative

(6) Studiu pentru selecția optimizată a funcției nucleu a estimatorului nonparametric recursiv de fundal.

Studiul este motivat de faptul că estimatorul recursiv de fundal (cu urmărire) are o comportare diferită atât față de estimatorii nonparametrici tradiționali cât și față de cei parametrici [106].

(7) Estimator nonparametric recursiv cu urmărire de mod autoadaptiv.

Pornind de la estimatorul nonparametric recursiv de fundal, am propus și dezvoltat estimatorul nonparametric recursiv (cu urmărire de mod) autoadaptiv, cu performanțe superioare la schimbările de amplitudine moderată ale fundalului, ce survin

frecvent datorită unor schimbări ale condițiilor de iluminare, reflexii, penumbre etc. [107].

- Ideea de bază constă în faptul că adaptarea se realizează prin modificarea ratei de învățare în funcție de eroarea de urmărire cumulativă.
- Estimator nonparametric recursiv cu urmărire de mod autoadaptiv reduce substanțial deviația standard a erorii de estimare, progresiv cu amplitudinea modificărilor în fundal.

(8) Estimator nonparametric recursiv cu urmărire de mod autoadaptiv noncauzal.

Pentru eliminarea erorilor de estimare la schimbări drastice ale fundalului, se propune estimatorul nonparametric recursiv cu urmărire de mod autoadaptiv noncauzal [108].

- Răspunsul la schimbările drastice ale fundalului, ce apar la introducerea sau eliminarea din fundal a unor obiecte sau la schimbări drastice ale iluminării, devine cvasi-sincron cu evenimentul.
- Estimator nonparametric recursiv cu urmărire de mod autoadaptiv noncauzal încorporează avantajele de adaptivitate, viteză și acuratețe ale versiunilor cauzale.

(9) Se realizează în cadrul tezei implementări eficiente ale metodelor propuse, cu operare în timp real, pentru o aplicație de analiză a deplasărilor și timpilor de așteptare într-un magazin de tip fast-food.

- Se implementează un estimator de fundal care obține experimental timp de calcul redus de circa 200 ori în raport cu implementarea cea mai rapidă cunoscută în prezent în literatură, versiunea bazată pe Transformata Gauss Rapidă, denumită de autori *Îmbunătățită* [60].
- Se implementează un program VC++ cu următoarele capabilități: substracție de fundal și detecție de prim-plan; filtrare spațială; analiză și urmărire de regiuni; interpretare și numărare de evenimente de tip intrare-ieșire în zona supravegheată; calcule statistice referitoare la activitatea în zona supravegheată.

5.3 Considerații finale

Estimarea fundalului este o operație de bază în aplicațiile de supraveghere video. Ideal, fundalul la fiecare pixel este constant, în timp ce schimbările ce apar sunt datorate exclusiv obiectelor mobile de interes.

În realitate, intensitatea și culoarea fundalului se modifică pe parcursul estimării. Obiecte de prim plan se interpun temporar în câmpul captat. Estimatorul de fundal trebuie să fie capabil să discearnă între schimbările legitime ale fundalului și cele datorate mișcării în imagine. Robustețea și adaptivitatea sunt calități esențiale ale oricărui estimator de fundal, parametric sau nonparametric, conceput pentru a face față varietății condițiilor ce survin în aplicațiile reale.

Metodele nonparametrice, în cadrul cărora se situează preocupările doctorandului, descrise în lucrarea prezentă, au avantajul de principiu că nu presupun o formă anumită a distribuției parametrilor fundalului și permit incorporarea mai firească a adaptivității. Obstacolul principal în proliferarea acestor metode a constituit-o o bună vreme efortul de calcul mai mare necesar implementării. Dată fiind miza existentă, numeroase lucrări au fost consacrate elaborării unor tehnici de calcul rapid a densității de probabilitate prin metode nonparametrice. Soluțiile propuse în literatură sunt generale și au la bază tehnici de condensare și Transformarea Gauss Rapidă (TGR) cu perfecționările propuse în [60].

Teza aduce soluții particulare ce permit creșterea vitezei de prelucrare cu două ordine de mărime. Soluțiile adaptive propuse în cadrul lucrării de doctorat completează tehnicile adaptive cunoscute, putând fi aplicate concomitent cu acestea. Estimarea fundalului rămâne fără îndoială o problemă deschisă.

Teza rezolvă doar o mică parte din numeroasele dificultăți ce persistă în folosirea acestei tehnici în varietatea condițiilor ce apar la aplicațiile ce își așteaptă rezolvarea deplin satisfăcătoare.

INDEX FIGURI

Fig. 1.1. Schema bloc generală a unui sistem de video supraveghere	4
Fig. 1.2. Detecția schimbării prin diferența între cadre succesive:	9
Fig. 1.3. Deplasarea unei regiuni pe un fundal cunoscut și diferit poate fi detectată corect.	10
Fig. 1.4. Structura de date folosită la estimarea fundalului: buffer de N cadre, analiză la nivel de pixel (coordonate spațiale constante).....	11
Fig. 1.5. Exemplu de estimare a funcției densitate de probabilitate pe baza unui set finit de eșantioane observate.	11
Fig. 1.6. Fazele extragerii fundalului. (a) imaginea originală; (b) fundalul extras; (c) mișcarea din imaginea originală.	12
Fig. 1.7. Exemplu de segmentare a obiectului mobil (prim plan) într-o secvență de imagine, prin extracția fundalului estimat.	13
Fig. 1.8. Detecția defectuoasă a mișcării cu o tehnică standard de extragere a fundalului, într-o secvență limită (tufiș cu frunze mișcate de vânt).....	14
Fig. 3.1. Descriere concisă a algoritmului rapid de urmărire a fundalului.	47
Fig. 3.2. Performanțele privind estimarea și urmărirea fundalului:	49
Fig. 3.3. Grupele de imagini a) și b) redau două cadre ale aceleiași secvențe obținute cu factori de scară în raport 1 la 3.	50
Fig. 3.4. Rezultate ale estimatorului nonparametric cu nucleu și ale estimatorului nonparametric cu urmărire propus de doctorand: deviația standard a erorii de estimare a fundalului calculată pentru cinci factori de scară diferiți.	52
Fig. 3.5. Secvența erorilor de estimare instantanee pentru 200 de eșantioane la:.....	53
Fig. 3.6. Procentul de apeluri ale buclei lungi de calcul, $O(2N)$, pentru estimarea fundalului într-un experiment ce implică o schimbare radicală a fundalului,	

obținută prin scoaterea unui obiect fix din scenă. Fiecare bară corespunde unui grup de 16 cadre, necesare pentru actualizarea completă a fundalului.....	54
Fig. 4.1. Rezultate comparative ale testelor pentru estimatorii cu nucleele Epanechnikov și Gaussian la fundal constant, cu zgomot alb aditiv.....	64
Fig. 4.2. Rezultate comparative ale testelor pentru estimatorii cu nucleele Epanechnikov și Gaussian la fundal dinamic, cu zgomot alb aditiv.	65
Fig. 4.3. Descriere prin pseudo-cod a funcționării estimatorului cu urmărire de mod adaptive.....	69
Fig. 4.4. Răspunsul estimatorilor nonparametrici recursivi la modificări moderate, de tip treaptă, ale fundalului.	71
Fig. 4.5. Deviația standard a erorii de estimare a estimatorului cu urmărire de mod și a estimatorului cu urmărire de mod adaptiv în funcție de amplitudinea treptei calculate pentru două deviații standard ale zgomotului $SD = 0,1$ și $SD = 0,2$. 72	
Fig. 4.6. Rezultate comparative pentru estimatorul de fundal nonparametric standard și estimatorul noncauzal, nonparametric, cu urmărire de mod adaptiv.....	77

BIBLIOGRAFIE

- [1] A.J. Lipton, T. Kanade, guest editors, „Special Section on Video Surveillance”, *IEEE Trans. on PAMI*, 22(8), August 2000.
- [2] A. Harampur, L. Brown, J. Connell, A. Ekin, N. Haas, M. Lu, H. Merkl, s. Pankanti, A. Senior, C.-F. Shu, Y.L. Tian, „Smart Video surveillance: exploring the concept of multiscale spatiotemporal tracking”, *IEEE signal Processing Magazine*, March 2005, pp 38-51.
- [3] J. Manigel and W. Leonard, „Vehicle control by computer vision”, *IEEE Trans. On Industrial Electronics*, 39 (3), June 1992, pp. 181-188.
- [4] R. Cucchiara, D. Lovell, A. Prati and M.M. Trivedi, guest editors, „Special Section on In-Vehicle Computer Vision Systems”, *IEEE Trans. On Vehicular Technology*, 53(6), Nov. 2004.
- [5] C. Démonceaux, A. Potelle, D. Kachi-Akkouche, „Obstacle detection in a road scene based on motion analysis”, *IEEE Trans. On Vehicular technology* 53(6), Nov. 2004, pp. 1649-1656.
- [6] M. Bertozzi, A. Broggi, A. Fascioli, T. Graf, M.M. Meinecke, „Pedestrian detection for driver assistance using multiresolution infrared vision”, *IEEE Trans. On Vehicular Technology*, 53(6), Nov. 2004, pp. 1666-1678.
- [7] I. Haritaoglu, D. Harwood and L.S. Davis. W4: „Real-time surveillance of people and their actions”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):809–830, August 2000.
- [8] V. Kettner and R. Zabih, „Counting people from multiple cameras” , *Proc. Int. Conf. Multimedia Computing and Systems*, Florence, Italy, June 7–11, 1999, pp. 267–271.
- [9] J. M. Nash, J. N. Carter, and M. S. Nixon, „Extraction of moving articulated-objects by evidence gathering”, *Proc. British Machine Vision Conf.*, Southampton, U.K., Sept. 14–17, 1998, pp. 609–618.

-
- [10] C. Bregler, „Learning and recognizing human dynamics in video sequences”, *Proc. Conf. Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, June 17–19, 1997, pp. 568–574.
- [11] C. Eveland, K. Konolige, R. Bolles, „Background Modeling for Segmentation of Video-rate Stereo Sequences”, *CVPR’98*, pp. 266-271, June 1998.
- [12] N. Friedman, S. Russell. „Image Segmentation in Video Sequences: a Probabilistic Approach”, *13th Conf. on Uncertainty in Artificial Intelligence*, August 1997.
- [13] Y. Ivanov, A. Bobick, J. Liu, „Fast Lighting Independent Background Subtraction”, *Int. J. Comp. Vis.*, 37(2), pp. 199-207, June 2000.
- [14] G. Gordon, T. Darrell, M. Harville, J. Woodfill. „Background Estimation and Removal Based on Range and Color”, *CVPR’99*, Vol.2, pp. 459-464, June 1999.
- [15] D. Comaniciu, F. Berton and V. Ramesh, „Adaptive Resolution System for Distributed Surveillance”, *Real-Time Imaging*, pp. 427-437, Vol. 8, 2002.
- [16] D. Comaniciu, „Nonparametric Information Fusion for Motion Estimation,” *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. I, pp. 59-66, 2003.
- [17] S. Kahn, O. Javed and M. Shah, „Tracking in Uncalibrated Cameras with Overlapping Field of View”, *PETS*, 2001.
- [18] J. Krumm, S. Harris, B. Meyers, B. Brumitt, M. Hale and S. Shafer, „Multi-Camera Multi-Person Tracking for Easy Living”, *International Workshop on Visual Surveillance*, pp. 3-10, 2000.
- [19] M.M. Trivedi, I. Mikic and S.K. Bhonsle, „Active Camera Networks and Semantic Event Databases for Intelligent Environments”, *Proceedings of the IEEE Workshop on Human Modelling, Analysis and Synthesis*, 2000.
- [20] C.E. Shannon, „A mathematical theory of communication”, *Bell systems Technical Journal* 27, pp. 379-423, 1948.
- [21] M. D. Levine, *Vision in Man and Machine*, New York: McGraw-Hill, 1985.
- [22] C.I Toma, F. Alexa, R.A. Vasiu, *Principiile televiziunii analogice și digitale*, Editura Politehnica, Timișoara, 2006.

-
- [23] A. Elgamal, R. Duraiswami, D. Harwood, L. Davis, „Background and foreground modeling using nonparametric kernel density estimation for visual surveillance”, invited paper in *Proceedings of the IEEE*, Vol. 90, No.7, 2002, pp. 1151-1162.
- [24] C. Wren, A. Azarbayejani, T. Darrel and A. Pentland, „PFinder, Real Time Tracking of the Human Body”, *IEEE PAMI*, vol 19, no. 7, July 1997, pp. 780-785.
- [25] T. Horprasert, D. Harwood and L. Davis , „A Statistical Approach for Real Time Robust Background Subtraction and Shadow Detection”, *IEEE Frame Rate Workshop*, 1999.
- [26] P. Rosin and T. Ellis, „Image Different Threshold Strategies and Shadow Detection”, *6th British Machine Vision Conf.*, Birmingham, pp. 347-356 1995.
- [27] V. Gui, D. Lacrămă, D. Pescaru, *Prelucrarea imaginilor*, Editura Politehnica, 1999.
- [28] Rafael C. Gonzalez, Richard E. Woods, *Digital Image processing, Second Edition*, Prentice Hall International, 2002
- [29] A. Vlaicu. *Prelucrarea digitală a imaginilor*. Grupul Microinformatica, Cluj Napoca, 1997.
- [30] M. Sonka, V. Hlavac, R. Boyle, *Image processing, analysis and machine vision*, Chapman & Hall Computing, Glasgow, 1993.
- [31] J.F. Canny, „A computational approach to edge detection”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, 1986, pp 679-698.
- [32] J. Shi and C. Tomasi, „Good features to track”, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 593–600, Seattle, WA, 1994.
- [33] D. Comaniciu, V. Ramesh, and P. Meer, „Real-time tracking of non-rigid objects using mean shift”, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Hilton Head, SC, volume II, June 2000, pp. 142-149.
- [34] K. Nummiar, E. Koller-Meier, T. Svoboda, D. Roth and L. Van Gool, „Color-Based Object Tracking in Multi-Camera Environments”, in *Proceedings of the DAGM'03*, Springer LNCS 2781, pp. 591-599, September 2003.

-
- [35] P. Fieguth and D. Terzopoulos, „Color based tracking of heads and other mobile objects at video frame rates” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 21–27, Puerto Rico, 1997.
- [36] M. Isard and A. Blake, „Contour tracking by stochastic propagation of conditional density”, in *Proc. European Conf. Computer Vision*, pp. 343–356, Cambridge, UK, 1996.
- [37] X.S. Zhou, D. Comaniciu, A. Gupta, „An information fusion framework for robust shape tracking”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, NO. 1, Jan 2005, pp. 1-15.
- [38] B. Bascle and R. Deriche, „Region tracking through image sequences” in *Proc. Int’l Conf. Computer Vision*, pp. 302–307, 1995.
- [39] G. Hager and P. Belhumeur, „Efficient region tracking with parametric models of geometry and illumination”, *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1025–1039, 1998.
- [40] M. Irani and S. Peleg, „Motion analysis for image enhancement: Resolution, occlusion, and transparency”, *JVCIP*, 4, pp. 324–335, December 1993.
- [41] J. Yang and A. Waibel, „A real-time face tracker”, in *Proceedings of WACV*, pp. 142–147, Sarasota, FL, 1996.
- [42] G. Bradski, „Computer vision face tracking for use in a perceptual user interface”, *Intel Technology Journal*, (Q2), 1998.
- [43] G. Cheung, S. Baker, and T. Kanade, „Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture”, In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pp. 77–84, Madison, WI, 2003.
- [44] C. Sminchisescu and B. Triggs, „Kinematic jump processes for monocular 3D human tracking”, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pp. 69–76, Madison, WI, 2003.
- [45] S. Avidan, „Support vector tracking”, in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pp. 184–191, Kauai, HI, 2001.
- [46] O. Williams, A. Blake, and R. Cipolla, „A sparse probabilistic learning algorithm for real-time tracking”, in *Proc. Int’l. Conf. Computer Vision*, pp. 353–360, Nice, France, 2003.

- [47] C. Stauffer and W. Grimson, „Learning patterns of activity using real-time tracking”, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, no. 8, pp. 747-757, 2000.
- [48] B. Stenger, P. R. S. Mendonca and R. Cipolla, „Model-based hand tracking using an unscented kalman filter”, in *Proc. British Machine Vision Conference*, volume I, Manchester, UK, September 2001, pp. 63-72.
- [49] M. Isard and A. Blake, „Condensation – Conditional density propagation for visual tracking”, *Intl. J. of Computer Vision*, vol. 29, no. 1, 1998.
- [50] M. Isard and A. Blake, „Icondensation: Unifying low level and high-level tracking in a stochastic framework”, *Lecture Notes in Computer Science*, vol. 1406, pp. 893-908, 1998.
- [51] K. Nummiaro, E. Koller-Meier and L. V. Gool, „An adaptive color-based particle filter”, *Image and Vision Computing*, vol. 21, no. 1, pp. 99-110, 2003.
- [52] P. Perez, C. Hue, J. Vermaak and M. Gangnet, „Color-based probabilistic tracking”, in *Proc. European Conf. on Computer Vision*, Copenhagen, Denmark, volume I, 2002, pp. 661-675.
- [53] Y. Rui and Y. Chen, „Better proposal distributions: Object tracking using unscented particle filter”, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawaii, volume II, 2001, pp. 786-793.
- [54] L.W. Campbell, D. A. Becker, A. Azarbayejani, A. F. Bobick and A. Pentland, „Invariant features for 3-D gesture recognition”, in *Proc. Int. Conf. Automatic Face and Gesture Recognition*, Killington, VT, October 1996, pp. 157–162.
- [55] A. D. Wilson and A. F. Bobick, „Parametric hidden Markov models for gesture recognition”, *IEEE Trans. Pattern Anal. Mach Intell.*, vol. 21, pp. 884–890, Sept. 1999.
- [56] A. Bobick and J. Davis, „The Recognition of Human Movements Using temporal Templates”, *IEEE PAMI*, Vol 23, No. 3, March 2001.
- [57] A. Senior, A. Hampapur, Y.-L. Tian, L. Brown, S. Pankanti and R. Bolle, „Appearance models for occlusion handling”, in *Second International workshop on Performance Evaluation of Tracking and Surveillance systems*, 2001.

-
- [58] P. Meer, „Robust techniques for computer vision”, *Emerging Topics in Computer Vision*, G. Medioni and S. B. Kang (Eds.), Prentice Hall, 2004, pp. 107-190
- [59] A. Elgamal, R. Duraiswami, L.S. Davis, „Efficient kernel density estimation using the Fast Gauss Transform with applications to color modeling and tracking”, *IEEE Trans. Pattern Anal. Machine Intell.* Vol. 25, No. 11, 2003, pp. 1499-1504.
- [60] J. Yang, R. Duraiswami, N. Gumerov, L. Davis, „Improved Fast Gauss Transform for efficient kernel density estimation”, *IEEE Intl. Conference on Computer Vision, ICCV*, 2003, pp. 464-471.
- [61] Y-J. Zhang, Ed., *Advances in Image and Video Segmentation*, IRM Press, Hershey, PA, 2006.
- [62] Y.H. Gu and V. Gui, „Joint space-time-range mean shift-based image and video segmentation”, invited paper in Y-J. Zhang, Ed. *Advances in Image and Video Segmentation*, IRM Press, Hershey, PA, 2006.
- [63] Z.S. Bojkovic, C. I. Toma, V. Gui, R. Vasiu, *Advanced Topics in Digital Image Compresion*, Editura Politehnica, Timisoara, 1997.
- [64] R. Vasiu, *Compresie audio-video. Tehnici și Aplicații*, Editura Orizonturi Universitare, Timisoara, 2002.
- [65] F. Alexa, „Contributii la estimarea miscarii bazata pe regiuni pentru compresia secventelor de imagini”, *Teza de doctorat*, Universitatea “Politehnica” Timisoara, 1999.
- [66] D. De Menthon, R. Megret, „Spatio-Temporal Segmentation of video by hierarchical mean shift analysis”, *UIMACS-TR*, 2002.
- [67] O. Javed, K. Shafique and M. Shah, „A hierarchical approach to robust background Subtraction using color and gradient information”, *MVC*, pp 22-27, Florida, December 2002.
- [68] D. Koller, J. Weber, J. Malik, „Robust multiplecar tracking with ecclusion reasoning”, *ECCV*, pp.189-196, Stockholm, 1994.
- [69] M.P. Wand, M.C. Jones, *Kernel Smoothing*, Chapman and Hall, 1995.
- [70] K. Toyama, J. Krumm, B. Rumitt and B. Beyers, Wallflower, „Principles and practice of backround maintainance”, *ICCV*, pp 255-261, , Greece, Sept.1999.

-
- [71] J. Zhong and S. Schlaroff, „Segmenting foreground objects from a dynamic, textured background via robust kalman filter”, *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003)*, 2-Volume Set, 2003.
- [72] H. Chen and P. Meer, „Robust Computer vision through kernel density estimation”, *ECCV*, pp 236-250, Copenhagen, Denmark, May 2002.
- [73] I. Abramson, „On bandwidth variation in kernel estimates - a square root law”, *The Annals of Statistics*, pp. 1217-1223, 1982.
- [74] X. Gao, T.E. Boult, F. Coetzee and V. Ramesh, „Error analysis of background adaptation”, in *CVIR*, pp 503-510, Hilton Head Island, S.C., June 2000.
- [75] A. Monnet, A. Mittal, N. Paragios and V. Ramesh, „Background modeling and substraction of dynamic scenes”, in *ICCV*, pp. 1305-1312, Nice, France, October 2003.
- [76] Y. Cheng, „Mean Shift, mode seeking and clustering”, *IEEE Trans.on PAMI*, vol.17, pp.790-799, 1995.
- [77] D. Comaniciu, U. Ramesh and P.Meer, „The variable band-width mean shift and data – driver scale selection”, *Proc.8th Intl. Conf. on Comp. Vision*, Vancouver, Canada, vol.I, July 2001, pp. 438-445.
- [78] K. Toyama, J. Krumm, B. Brumitt, B. Meyers, „Wallflower: principles and practice of background maintenance”, *IEEE Conference on Computer Vision*, Kerkyra, Greece, 1999, pp. 255-261.
- [79] R. Cucchiara, M. Piccardi, A. Prati, „Detecting moving objects, ghosts and shadows in video streams”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 25, No. 10, October 2003, pp. 1337-1342.
- [80] M. Harville, G. Gordon, J. Woodfill, „Adaptive video background modeling using color and depth”, *International Conference on Image Processing ICIP 2001*, Tessaloniki, Greece, Vol.3, Oct. 7-10, 2001, pp. 90-93.
- [81] D. Farin, P. H. N. de With, W. Effelsberg, „Robust background estimation for complex video sequences“, *IEEE International Conference on Image Processing, ICIP 2003*, Barcelona, September 2003, pp. 145-148.

- [82] P. J. Rouseeuw, A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley, 1987.
- [83] W. E. L. Grimson, C. Stauffer, R. Romano, L. Lee, „Using adaptive tracking to classify and monitor activities in a site”, *IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA., 1998, pp. 22-29.
- [84] P. Kaew, T. K. Pong, R. Bowden, „An improved adaptive background mixture model for real-time tracking with shadow detection”, *Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems*, Sept. 2001.
- [85] P. Withagen, K. Schutte, F. Groen, „Object detection and tracking using a likelihood based approach”, *Proc. ASCI 2002 Conference*, Lochem, The Netherlands, June 2002, pp. 248-253.
- [86] M. Pic, L. Berthouze, T. Kurita, „Adaptive background estimation: Computing a pixel-wise learning rate from local confidence and global correlation values”, *IEICE Trans. Inf & Syst.*, Vol. E87-D, No.1, January 2004, pp. 1-7.
- [87] A. Dempster, N. Laird, D. Rubin, „Maximum likelihood from incomplete data via the EM algorithm”, *Journal of the Royal Statistical Society, Series B*, 39(1), 1977, pp. 1-38.
- [88] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, Wiley, Second Edition, 2000.
- [89] D. W. Scott, *Multivariate Density Estimation*, Wiley, 1992.
- [90] D. Comaniciu, P. Meer, „Mean shift analysis and applications”, *International Conference on Computer Vision*, Kerkyra, Greece, 1999, pp. 1197-1203.
- [91] D. Comaniciu, P. Meer, „Mean shift: A robust approach toward feature space analysis”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 24, No. 5, 2002, pp.603-619.
- [92] M. Girolami, C. He, „Probability density estimation from optimally condensed data sets”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 25, No. 10, 2003, pp. 1253-1264.

-
- [93] J. P. R. Renno, J. Orwell, G. A. Jones, „Evaluation of shadow classification techniques for object detection and tracking”, *International Conference on Image Processing, ICIP*, Vol.1, 2004, pp. 143-146.
- [94] M. Petrou, S. Barsky, „Shadows and highlights detection in 4-source color photometric stereo”, *Proc. International Conference on Image Processing, ICIP*, Vol. 3, 2001, pp. 967-970.
- [95] A. Mahalanobis, J. Cannon, S. Robert Stanfill, R. Muise, M. Shah, „Network video image processing for security”, *Surveillance and Situational Awareness Keynote at SPIE Conference of Digital Wireless Communication VI*, Orlando, April 2004.
- [96] A. Prati, R. Cucchiara, I. Mikic, M. Trivedi, „Analysis and detection of shadows in video streams: a comparative evaluation”, *International Conference on Computer Vision and Pattern Recognition*, 2001.
- [97] M. Harville, „Stereo Person Tracking with Adaptive Plan-view Templates of Height Occupancy Statistics”, *Journal of Image and Vision Computing*, Vol. 22, No.2, Feb. 2004, pp. 127-142.
- [98] M. Harville, „Stereo Person Tracking with Adaptive Plan-view Appearance Models of Shape and Color”, *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance*, Como, Italy, Sept. 2005, pp. 522-527.
- [99] M. Fashing, C. Tomasi, „Mean shift is a bound optimization”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, No. 3, 2005, pp. 471-474.
- [100] B. Han, D. Comaniciu, Y. Zhu, L. Davis, „Incremental Density approximation and Kernel-Based Bayesian Filtering for Object Tracking”, *IEEE Conf. Computer Vision and Pattern Recognition (CVPR'04)*, Washington DC, 2004.
- [101] A. Mittal, N. Paragios, „Motion-based background subtraction using adaptive kernel density estimation”, *IEEE Conf. Computer Vision and Pattern Recognition (CVPR'04)*, Washington DC, 2004.

- [102] **C. N. Ianăși**, „Stadiul actual și de perspectivă în detecția și extragerea fundalului în supravegherea video”, *Referatul de doctorat nr.1*, Universitatea “Politehnica”, Timișoara, 2005.
- [103] **C. N. Ianăși**, „Metode adaptive de estimare a fundalului în supravegherea video”, *Referatul de doctorat nr.2*, Universitatea “Politehnica”, Timișoara, 2006.
- [104] **C. N. Ianăși**, „Estimator de fundal nonparametric, noncauzal cu urmărire de mod”, *Referatul de doctorat nr.3*, Universitatea “Politehnica”, Timișoara, 2006.
- [105] **C. N. Ianăși**, V. Gui, C. I. Toma, D. Pescaru, „A Fast Algorithm for Background Tracking in Video Surveillance using Nonparametric Kernel Density Estimation“, *Facta Universitatis Niš*, Serbia and Montenegro, Series Electronics and Energetics, Vol. 18, No.1, April 2005, pp. 127-144.
- [106] **C. N. Ianăși**, C. I. Toma, V. Gui, D. Pescaru, „Kernel Selection for Mean Shift Background Tracking in Video Surveillance“, *Proceedings 4th Int. Conference on Microelectronics and Computer Science (ICMCS-05)*, Chișinău, Moldova, Vol. II, September 15-17, 2005, pp. 389-392.
- [107] **C. N. Ianăși**, V. Gui, F. Alexa, C. I. Toma, „Fast and Accurate Background Subtraction for Video Surveillance, Using an Adaptive Mode-tracking Algorithm”, *WSEAS Int. Conf. on Dynamical Systems and Control (WSEAS 2005)*, Venice, Italy, November 2-4, 2005, pp. 391-397.
- [108] **C. N. Ianăși**, V. Gui, F. Alexa, C. I. Toma, „Noncausal, Adaptive Mode-Tracking Estimation for Background Subtraction in Video Surveillance”, *WSEAS Transactions on Signal Processing*, Issue 1, Volume 2, January 2006, pp. 52-59, ISSN 1790-5022.
- [109] D.Al. Pescaru, D. Fuiorea, V. Gui, C.I. Toma, G.M. Muntean, A. Doboli, „Image-based Node Localiyation Algorithm for Wireless Video Sensor Networks”, *The Sixth Information Technologz and Telecommunication Conference IT&T'06*, Carlow, Ireland, October, 2006 – acceptată la conferință.

- [110] D. Fuiorea, D.Al. Pescaru, V. Gui, C.I. Toma, „Feature Based 2D Image Registration using Mean Shift Parameter Estimation”, *Proceedings of Etc'2006, 7th International Symposium on Electronics and Telecommunications*, “Politehnica” University, Timișoara, September 2006.
- [111] R. Collins, „A system for video surveillance and monitoring”, VSAM Final Report, Carnegie Mellon Univ., Pittsburgh, PA, May 2000.
- [112] Human ID at a Distance, DARPA Project, Electronic Frontier Foundation, [Online], Available: <http://www.eff.org/Privacy/TIA/hid.php>.