

КОМБИНОВАНИ МЕТОД ДЕТЕКЦИЈЕ И ПРАЋЕЊЕ ПОКРЕТНИХ ОБЈЕКТА У ВИДЕО СЕКВЕНЦИ ПРИМЕНОМ КАЛМАНОВОГ ФИЛТРА

Димитрије Бујаковић, Бобан Бонцулић, Војна академија – Београд

Садржај – У раду је дат један приступ за детекцију и праћење покретних објеката. За детекцију покретних објеката се користи комбиновани метод детекције заснован на обради слике разлике три сукцесивна фрејма и слике разлике добијене одузимањем слике позадине од текућег фрејма. Комбиновани метод детекције је послужио за иницијализацију Калмановог филтра, који се користи за праћење детектованих објеката.

1. УВОД

Детекција и праћење покретних објеката у видео секвенци је битан задатак у великом броју апликација (видео надзор, контрола саобраћаја, обезбеђење државне границе, аеродрома и сл.). Динамика на слици може да постоји због кретања објеката на сцени али и због кретања саме камере или дејства оптике (зум). У раду се разматра динамика која постоји само због кретања објеката на сцени.

У првом делу рада дат је комбиновани метод детекције покрета, заснован на обради слике разлике три сукцесивна фрејма и слике разлике добијене одузимањем слике позадине од текућег фрејма. За добијање слике позадине користи се метод који даје слику позадине у неколико итерација.

Други део рада односи се на праћење покретних објеката помоћу Калмановог филтра.

2. ДЕТЕКЦИЈА ПОКРЕТНИХ ОБЈЕКТА НА СЛИЦИ СЦЕНЕ

За детекцију покрета на слици сцене може се користити обрада слике разлике која се добија одузимањем слике позадине од текућег фрејма или одузимањем сукцесивних фрејмова.

Означимо са $I^n(x, y)$ вредност интензитета пиксела фрејма на позицији (x, y) у тренутку $t = n$. Мера сличности фрејма у тренутку $t = n$ и у тренутку $t = n - 1$ може се узети као [1]:

$$D^n(x, y) = \left| I^n(x, y) - I^{n-1}(x, y) \right|, \quad \forall (x, y) \in [1, N] \times [1, M] \quad (1)$$

где је $N \times M$ димензија фрејма у пикселима и где интензитет пиксела фрејма може узети било коју вредност из скупа $[0, 1, 2, \dots, 255]$.

Означимо са $B^n(x, y)$ интензитет пиксела слике позадине на позицији (x, y) у тренутку $t = n$. За сада претпоставимо да нам је слика позадине позната. Детекција покрета се може извршити на основу разлике текућег фрејма и слике позадине [2]:

$$D^n(x, y) = \left| I^n(x, y) - B^n(x, y) \right|, \quad \forall (x, y) \in [1, N] \times [1, M] \quad (2)$$

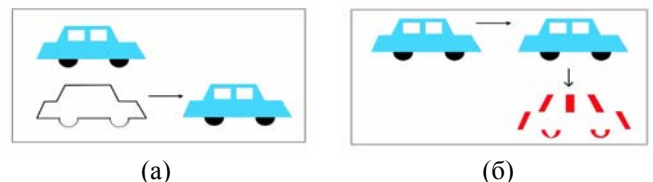
Код оба приступа за детекцију покрета на слици сцене јавља се утицај шума. Због утицаја шума поред детекције промене нивоа сивог, која је последица присуства покрета на сцени, детектује се и велики број пиксела код

којих је дошло до промене нивоа сивог због утицаја шума. Да би се елиминисао утицај шума и донела одлука о томе код којих је пиксела дошло до промене нивоа сивог због покрета на сцени, потребно је на слику разлике $D^n(x, y)$ применити праг T ($T \in [0, 255]$). На тај начин се добија бинарна слика $I_b^n(x, y)$ у којој су пиксели класификовани у две класе (позадина и покретни пиксел), тј.:

$$I_b^n(x, y) = \begin{cases} 0, & D^n(x, y) \leq T \\ 255, & D^n(x, y) > T \end{cases} \quad (3)$$

То значи да је пиксел покретан ако се његов интензитет битно променио у текућем и претходном фрејму или у текућем фрејму и слици позадине.

Главна мана код технике одузимања позадине је та што је ова техника осетљива на ситуације када се стационарни објекти сцене почињу кретати. Иако ће ови објекти бити детектовани, они иза себе остављају "рупе" где се новооткривена позадина разликује од познате слике позадине (Сл.1(а)). Док се модел позадине прилагоди овим "рупама", сам алгоритам детекције одређено време генерише лажне објекте. Разлика сукцесивних фрејмова није осетљива на појаву "рупа" али овај алгоритам мења облик покретног објекта (Сл. 1(б)).



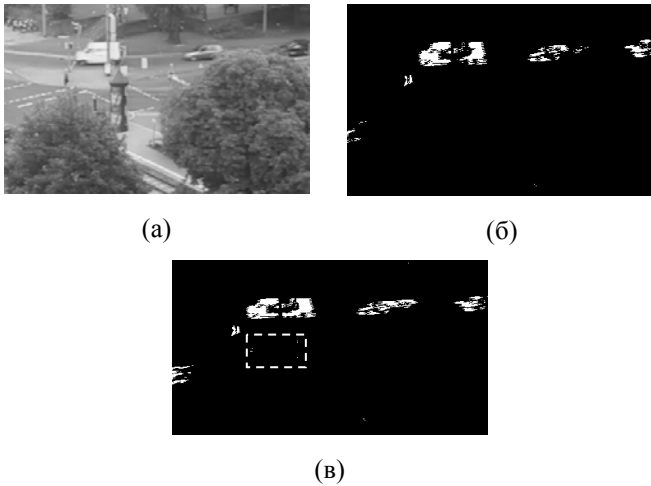
Сл. 1. (а) Одустражање слике позадине оставља "рупе" када стационарни објекат почиње да се креће; (б) Одустражање сукцесивних фрејмова мења облик покретног објекта

Да би се избегли наведени недостаци користи се комбинација ова два метода [1]. Одустражање сукцесивних фрејмова се користи за одређивање региона стварних покрета, а одустражање слике позадине за издвајање покретних објеката.

Осим разлике два сукцесивна фрејма може се користити разлика три или више сукцесивних фрејмова [1]. Правило разлике три сукцесивна фрејма каже да је пиксел на позицији (x, y) покретан ако је:

$$\left| I^n(x, y) - I^{n-1}(x, y) \right| > T \quad \text{и} \quad \left| I^n(x, y) - I^{n-2}(x, y) \right| > T \quad (4)$$

где је T одговарајући праг. На овај начин добија се бинарна слика $I_b^n(x, y)$ (Сл. 2(б)). Вредност прага је у раду одређивана на основу метода предложеног у [3]. Метод је заснован на ентропији густине вероватноће нивоа сивог слике разлике. Уместо једне густине вероватноће нивоа сивог слике разлике, *Kapur* је у [3] посматрао две густине вероватноће, једну за позадину и другу за објекат.



Сл. 2. (а) Текући фрејм; (б) Покретни пиксели у текућем фрејму добијени на основу разлике три сукцесивна фрејма; (в) Покретни пиксели у текућем фрејму добијени на основу разлике два сукцесивна фрејма

Применом разлике три сукцесивна фрејма елиминисани су неки покретни пиксели који су детектовани на основу разлике два сукцесивна фрејма а чија детекција није последица покрета објеката. Ови пиксели су обухваћени правоугаоником на Сл. 2(в).

Оно што се још може видети са Сл. 2(б) је да су групе пиксела које припадају једном објекту просторно одвојене. Да би се ове групе спојиле могу се применити морфолошке операције. Од морфолошких операција у детекцији покрета најчешће се користе дилатација и ерозија. Операције дилатације и ерозије се састоје у интеракцији бинарне слике и структурног елемента [4,5]. У раду се користио структурни елемент S димензија 3×3 пиксела.

На Сл. 3 можемо видети шта се добија након примене операција дилатације и ерозије на бинарну слику 2(б).

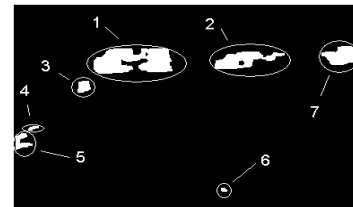


Сл. 3. (а) Примена дилатације на слику 2(б); (б) Примена ерозије на слику 3(а)

Након примене морфолошких операција пиксели који се налазе у унутрашњости објеката неће бити укључени у скуп покретних пиксела (Сл. 3(б)).

Ако сада извршимо груписање покретних пиксела који припадају истом објекту, пиксели из унутрашњости се могу укључити одузимањем слике позадине. Груписање покретних пиксела се врши на основу повезаности пиксела. Повезаност пиксела може се испитивати поређењем са четири суседна пиксела или испитивањем са осам суседних пиксела. Груписањем пиксела добијају се региони покретних пиксела (Сл. 4). За добијање региона покретних пиксела се користила осмоповезаност.

Региони покретних пиксела служе да би се одредили делови текућег фрејма у којима се налазе покретни објекти.



Сл. 4. Региони покретних пиксела са слике 3(б) (региони су обележени бројевима од 1 до 7)

Делови текућег фрејма у којима се налазе покретни објекти су одређени на основу крајњих тачака сваког региона покретних пиксела, тј., тако што се сваки регион покретних пиксела обухвати са правоугаоником минималног обима. На Сл. 5 приказани су делови текућег фрејма (Сл 2(а)) у којима се налазе покретни објекти.

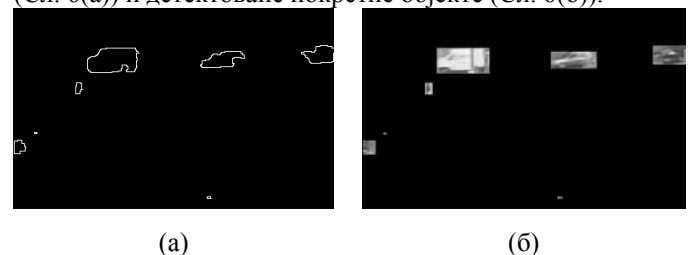


Сл. 5. Делови текућег фрејма (Сл. 2(а)) у којима се налазе покретни објекти

Означимо са R_i^n део текућег фрејма у коме се налази покретни објекат. Покретни објекат b_i^n ћемо коначно детектовати ако узмемо све пикселе из R_i^n чији интензитет се битно разликује од модела позадине, тј.:

$$b_i^n = \left\{ (x, y) : \left| I^n(x, y) - B^n(x, y) \right| > T, (x, y) \in R_i^n \right\} \quad (5)$$

На Сл. 6 можемо видети ивице детектованих објеката (Сл. 6(а)) и детектоване покретне објекте (Сл. 6(б)).



Сл. 6.(а) Ивице детектованих објеката; (б) Детектовани објекти

Детекција ивица је вршена на основу бинарне слике која се добија помоћу израза (5) (детекција прелаза 0-255 и обратно). Из наведеног примера видимо да је овакав начин детекције успешно детектовао пет покретних објеката (четири аутомобила и једног пешака) али је и детектовао два покретна објекта којих у стварности нема. Овај недостатак се може отклонити ако се узму у разматрање само детектовани покретни објекти димензија већих од унапред задатих. У раду су одбацивани детектовани објекти димензија мањих од 5×5 пиксела.

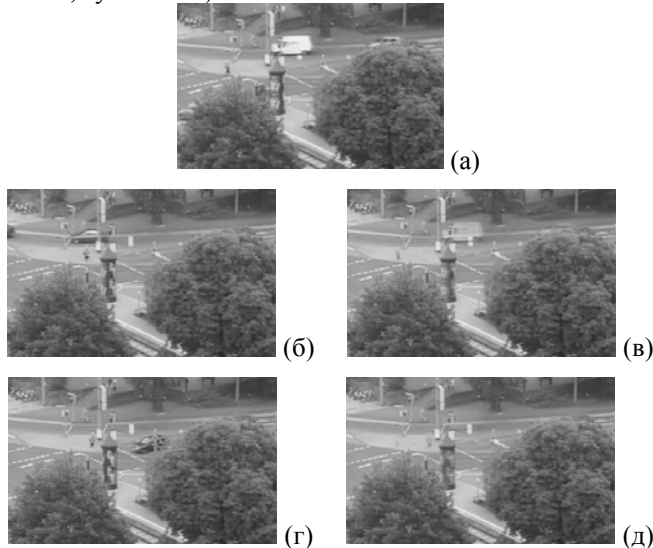
У неким применама детекције покрета на сцени, нећемо бити у могућности да из саме секвенце издвојимо један фрејм и прогласимо га позадином. Због тога је потребно формирати слику позадине из секвенце фрејмова $I^t(x, y)$, а која може садржати покретне објекте. Како је за детекцију покретних објеката потребна слика позадине, први корак у детекцији треба да буде издвајање позадине. Метод који се често користи даје слику позадине у неколико итерација [1,2]. Метод

полази од почетне слике позадине која је једнака првом фрејму $B^0(x, y) = I^0(x, y)$. Да би се добила слика позадине која одговара текућем фрејму, у сукцесивним фрејмовима морају се детектовати покретни пиксели, а затим издвојити позадина на основу:

$$B^{n+1}(x, y) = \begin{cases} \alpha B^n(x, y) + (1 - \alpha) I^n(x, y), & (x, y) - \text{непокретан} \\ B^n(x, y), & (x, y) - \text{покретан} \end{cases} \quad (6)$$

Од тежинског коефицијента α зависи брзина промене. Из (6) се види да до промене вредности пиксела слике позадине долази ако је пиксел класификован као непокретан, тј. ако је део стационарне позадине. У супротном би покретни пиксели "пореметили" слику позадине. У раду се користио коефицијент $\alpha = 0,9$ који је предложен у [1,2].

Пример издвајања позадине на овакав начин приказан је на Сл. 7. Сл. 7(а) представља почетну позадину на којој постоји више покретних објеката. Сл. 7(б) и 7(г) су слике 19-ог и 32-ог фрејма, а слике 7(в) и 7(д) су слике на овакав начин издвајане позадине након 19-ог и 32-ог фрејма. Видимо са слика 7(в) и 7(д) да се постепено покретни објекти из почетне позадине губе (бели камион, пешак, аутомобил).



Сл. 7. Пример издвајања позадине

Да би се прилагодили променама сцене потребно је да се врши ажурирање позадине. Ажурирање позадине може да се врши на основу израза (6) (након сваког фрејма или у одређеним временским интервалима).

3. ПРИМЕНА КАЛМАНОВОГ ФИЛТРА ЗА ПРАЋЕЊЕ ОБЈЕКТА У ВИДЕО СЕКВЕНЦИ

За праћење покретних објеката у видео секвенци делимично је примењена аналогија са праћењем покретних објеката помоћу радарског сензора. Наиме, применом Калмановог филтра може се естимирати позиција центроида објекта у слици, те описати око естимиране позиције прозор праћења (*gating*). Израчунавање центроида се врши према методу датом у [5]. У прозору праћења врши се детекција објеката одузимањем слике позадине. На овај начин врши се обрада само "интересантног" дела текућег фрејма.

За опис кретања користи се модел код кога је убрзање константно између тренутака одабирања ("piecewise

constant white acceleration"), где је вектор стања система описан:

$$\mathbf{x} = [\zeta_1 \ v_{\zeta 1} \ \eta_1 \ v_{\eta 1} \ \zeta_2 \ v_{\zeta 2} \ \eta_2 \ v_{\eta 2}]^T \quad (7)$$

при чему су:

ζ_i, η_i - координате положаја i -тог објекта дуж једне, односно друге координате, ($i=1,2$)

$v_{\zeta i}, v_{\eta i}$ - компоненте вектора брзине i -тог објекта дуж једне, односно друге координате, ($i=1,2$)

Једначина стања за овако дефинисан систем је:

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{w}(k) = \text{diag}(\mathbf{F}_{\text{pcwa}}, \mathbf{F}_{\text{pcwa}})\mathbf{x}(k) + \text{diag}(\mathbf{G}_{\text{pcwa}}, \mathbf{G}_{\text{pcwa}})\mathbf{w}(k) \quad (8)$$

при чему је \mathbf{F}_{cv} одређена у [6] као:

$$\mathbf{F}_{\text{pcwa}} = \text{diag}(\mathbf{F}_2, \mathbf{F}_2), \quad \mathbf{G}_{\text{pcwa}} = \text{diag}(\mathbf{G}_2, \mathbf{G}_2) \\ \mathbf{F}_2 = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad \mathbf{G}_2 = \begin{bmatrix} T^2 / 2 \\ T \end{bmatrix} \quad (9)$$

где је $\mathbf{w}(k) = [w_{\zeta 1}(k) \ w_{\eta 1}(k) \ w_{\zeta 2}(k) \ w_{\eta 2}(k)]^T$ секвенца дискретног белог шума и T периода одабирања.

Матрица коваријансе шума процеса помноженог са појачањем је дата у [6] као:

$$\mathbf{Q} = \text{cov}(\mathbf{G}\mathbf{w}(k)) = \text{diag}(\mathbf{Q}_1, \mathbf{Q}_2) \\ \mathbf{Q}_1 = \text{diag}[\text{var}(\mathbf{w}_{\zeta 1})\mathbf{Q}_3, \text{var}(\mathbf{w}_{\eta 1})\mathbf{Q}_3], \\ \mathbf{Q}_2 = \text{diag}[\text{var}(\mathbf{w}_{\zeta 2})\mathbf{Q}_3, \text{var}(\mathbf{w}_{\eta 2})\mathbf{Q}_3], \\ \mathbf{Q}_3 = \begin{bmatrix} T^4 / 4 & T^3 / 2 \\ T^3 / 2 & T^2 \end{bmatrix} \quad (10)$$

У овом раду посматра се да су варијансе шума процеса дуж координатних оса за оба објекта исте и да је њихова вредност 200 (добијено емпиријски).

Једначина мерења је:

$$\mathbf{z}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{v}(k) \quad (11)$$

при чему је:

$\mathbf{z}(k)$ - вектор опсервација у тренутку k , а

$\mathbf{v}(k) = [v_{\zeta 1}(k) \ v_{\eta 1}(k) \ v_{\zeta 2}(k) \ v_{\eta 2}(k)]^T$ - секвенца дискретног мерног шума, где је

\mathbf{H} - матрица мерења која се дефинише као

$$\mathbf{H} = \text{diag}(\mathbf{H}_1, \mathbf{H}_1), \\ \mathbf{H}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (12)$$

Матрица коваријансе шума мерења је дата у [6] као:

$$\mathbf{R} = \text{diag}(\mathbf{R}_1, \mathbf{R}_2) \\ \mathbf{R}_1 = \text{diag}[\text{var}(\mathbf{v}_{\zeta 1}), \text{var}(\mathbf{v}_{\eta 1})], \\ \mathbf{R}_2 = \text{diag}[\text{var}(\mathbf{v}_{\zeta 2}), \text{var}(\mathbf{v}_{\eta 2})] \quad (13)$$

У овом раду сматрано је да су варијансе шума мерења дуж обе координате исте и да њихова вредност износи 5 (добијено емпиријски).

За овако дефинисан математички модел кретања објекта, користећи познате релације које се користе при рачунању параметара Калмановог филтра [7], израчунава се предикција позиције објекта. Око предикције позиције објекта врши се издвајање дела слике која припада прозору праћења у наредном фрејму, на тај начин што се узима део текућег фрејма облика квадрата у чијем је пресеку дијагонала естимирана позиција објекта. Дужина стране квадрата је 30 пиксела. На Сл. 8 приказани су прозори праћења.



Сл.8. Прозори праћења: (а) група пешака; (б) аутомобил

Одузимањем слике позадине која одговара естимираном прозору праћења, добијају се покретни објекти. За детектоване објекте одређују се центроиди и поступак се понавља. На Сл. 9 приказано је обележавање објеката на слици сцене.



Сл. 9. Обележавање објеката на слици сцене

У иницијалном вектору стања система претпостављено је да су позиције центроида одређене на основу почетног фрејма секвенце (према методу описаном у другом поглављу овог рада) и да су компоненте брзина дуж координатних оса једнаке нули, а матрица коваријансе грешке естимације има велику вредност на главној дијагонали. Ово се може претпоставити на основу чињенице да се у секвенци врши праћење споропокретних објеката.

4. РЕЗУЛТАТИ

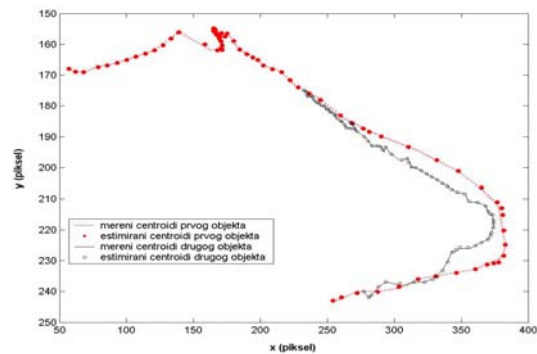
Тестирање предложеног алгоритма детекције и праћења извршено је на 80 фрејмова видео секвенце добијене фиксном телевизијском камером. Димензије фрејмова су 358x422 пиксела, због проширења слике са сваке стране са по 35 пиксела чији је ниво сивог 0. У секвенци су постојала два објекта чије су се трајекторије укрштале. Анализа је вршена у програмском пакету *Matlab* на *PC Pentium-IV* на 1,66 GHz.

За детекцију и праћење покретних објеката без примене Калмановог алгоритма, средње време обраде фрејма је 1,4 s. Применом Калмановог алгоритма средње време обраде је 0,5 s. Осим смањења средњег времена обраде, применом Калмановог алгоритма детектовани су и објекти који се налазе далеко од камере и који се удаљавају. Ови објекти применом само алгоритма датог у другом делу овог рада нису детектовани, јер је први корак детекције одузимање сукцесивних фрејмова, а природа кретања у овом случају је таква да у сукцесивним фрејмовима нема видљиве промене њихове позиције. Применом Калмановог филтра смањен је шум који је последица дигитализације аналогног видео сигнала (осим уколико није “напао” прозор праћења).

Још један допринос примене Калмановог алгоритма у праћењу покретних објеката је решавање ситуације када се два или више покретних објеката сусрећу. Уколико би се користила метода дата у другом делу овог рада, детектовао би се само један објекат, док се у случају примене Калмановог алгоритма детектује тачан број објеката.

На Сл. 10 приказани су естимиране и измерене позиције центроида објеката, на основу примене предложеног алгоритма.

Како постоји добро поклапање естимираних позиција центроида са стварно измереним, може се сматрати да је изабрани модел процеса добар.



Сл. 10. Измерени и естимирани центроиди објеката

5. ЗАКЉУЧАК

У раду је дат један приступ за детекцију и праћење покретних објеката. За детекцију објеката се користио комбиновани метод детекције заснован на обради слике разлике три сукцесивна фрејма и слике разлике добијене одузимањем слике позадине од текућег фрејма. За праћење објеката се користио Калманов филтар. Применом Калмановог филтра средње време обраде једног фрејма се смањило у односу на случај када се за праћење користи само комбиновани метод. Позитивни ефекти примене овог приступа су и: детекција и праћење објеката који се налазе далеко од камере и удаљавају од ње, смањење утицаја шума текућег фрејма и решавање ситуација када се два објекта сусрећу.

У даљем раду било би потребно применити поступке елиминације сенке која се јавља у појединим деловима секвенце, као и развити механизам променљиве димензије прозора праћења.

ЛИТЕРАТУРА

- [1] R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, O. Hasegawa, “A system for video surveillance and monitoring: VSAM final report”, *Robotics Inst., CMU-RI-TR-00-12, 2000*.
- [2] S. Gupte, O. Masoud, R. Martin, N. Papanikolopoulos, “Detection and Classification of Vehicles”, *IEEE Trans. on ITS, Vol.3, No.1, March 2002*.
- [3] J.N. Kapur, P.K. Sahoo, A.K.C. Wong, “A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram”, *Computer Vision, Graphics, and Image Processing 29, pp. 273-285, 1985*.
- [4] М. Поповић, “Дигитална обрада слике”, *Ручник за припремање истама, Београд, 1998*.
- [5] Б. Бонцулић, С. Симић, Б. Зрнић, “Један приступ детекцији и праћењу споропокретних објеката у видео секвенци”, *Proc. XLVIII ETRAN Conference, Čačak*
- [6] Y. Bar-Shalom, X.R. Li, “Estimation and Tracking: Principles, techniques and software”, *Artech House 1993*.
- [7] B. Kovačević, Ž. Đurović, “Fundamentals of Stochastic Signals, Systems and Estimation Theory with Worked Examples”, *Akademska misao, Beograd, 1999*.

Abstract – In this paper the approach for detection and tracking of moving objects is proposed. Detection moving objects is based on three successive frames differencing and background subtraction. Tracking is based on Kalman filtering.

HYBRID METHOD FOR DETECTION AND TRACKING MOVING OBJECTS IN VIDEO SEQUENCE USING KALMAN FILTERING

Dimitrije Bujakovic, Boban Bondzulic