

## ЈЕДАН ПРИСТУП ДЕТЕКЦИЈИ И ПРАЋЕЊУ СПОРОПОКРЕТНИХ ОБЈЕКТА У ВИДЕО СЕКВЕНЦИ

Бобан Бонцулић, Слободан Симић, Бојан Зрнић, Војна академија, Ратка Ресановића 1, 11000 Београд

### Награда за најбољи рад младог истраживача на комисији

**Садржај** – У овом раду је предложен један метод за детекцију и праћење споропокретних циљева у видео секвенци. Метод се заснива на детекцији покрета и примени морфолошких оператора. Овај метод детекције и праћења може се применити за контролу саобраћаја, праћење објеката у ваздушном простору као и за видео надзор.

### 1. УВОД

Технике праћења споропокретних објеката можемо поделити у две групе: технике праћење на основу препознавања и технике праћења на основу детекције покрета (промена) [1,2]. Код техника за праћење на основу препознавања у сукцесивним фрејмовима врши се препознавање и издвајање позиције препознатих објеката (циљева). Главна предност ових техника праћења састоји се у томе што можемо естимирати translацију и ротацију објеката. Лоша страна ових техника је што се једино препознати објекти могу пратити и што су перформансе праћења зависне од комплексности метода за препознавање. Код друге групе техника, праћење се врши на основу естимације параметара покрета. Њихова предност је у томе што нам дају могућност да пратимо било који покретни објекат, независно од његовог облика или димензија.

У овом раду је предложен један метод за детекцију и праћење споропокретних циљева у видео секвенци. Метод се заснива на детекцији покрета и примени морфолошких оператора.

### 2. ОПИС МЕТОДА

Генерална шема процеса праћења примењена у овом раду дата је на сл.1. Стационарна CCD камера је коришћена за добијање секвенце сцене која се надгледа. У сваком фрејму  $I(x,y)$  се детектују промене и добијамо бинарну слику где сваки скуп пиксела представља могући објекат (ауто, камион, пешак,...). Промене се детектују на основу модула разлике  $D(x,y)$  између текуће слике (фрејма)  $I(x,y)$  и слике позадине  $BCK(x,y)$  сваких  $t_k$  временских размака

$$D(x,y) = \left| \begin{array}{c} I(x,y) - BCK(x,y) \\ \forall (x,y) \in [1,N] \times [1,M] \end{array} \right| \quad (1)$$

где је  $N \times M$  димензија слике у пикселима. Затим се одреди праг  $T$  ( $T \in [0,1]$ ) који ће се применити на слику  $D(x,y)$  да би се добила бинарна слика у којој ће пиксели бити класификовани у две групе (позадина и објекат) тј.

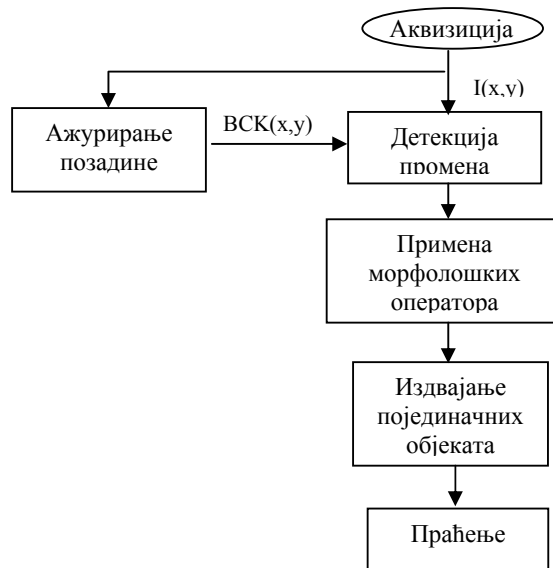
$$B(x,y) = \begin{cases} 0, & \text{за } D(x,y) < T \\ 1, & \text{за } D(x,y) > T \end{cases} \quad (2)$$

Овај праг смо за нашу секвенцу покушали да одредимо уз помоћ Отсуовог метода [3], међутим он није дао задовољавајуће резултате, па смо овај праг одредили експериментално. Отсуов метод није дао задовољавајуће

резултате јер су нивои осветљености појединих објеката и позадине слични. На крају, добијамо бинарну слику у којој пиксели позадине имају вредност нула, а пиксели могућих објеката имају вредност један.

Да би се избегле грешке код одређивања могућих објеката потребно је вршити ажурирање позадине (различито доба дана, различити временски услови,...). У самом раду није вршено ажурирање позадине већ је за слику позадине узимана увек иста слика  $BCK(x,y)$ .

На сл.2(а) имамо једну реалну слику са више покретних објеката (аута, пешак, група пешака,...). На сл. 2(б) и 2(ц) можемо видети слику позадине и бинарну слику коју добијамо после примене прага за раздвајање могућих објеката од позадине (сегментација), респективно.

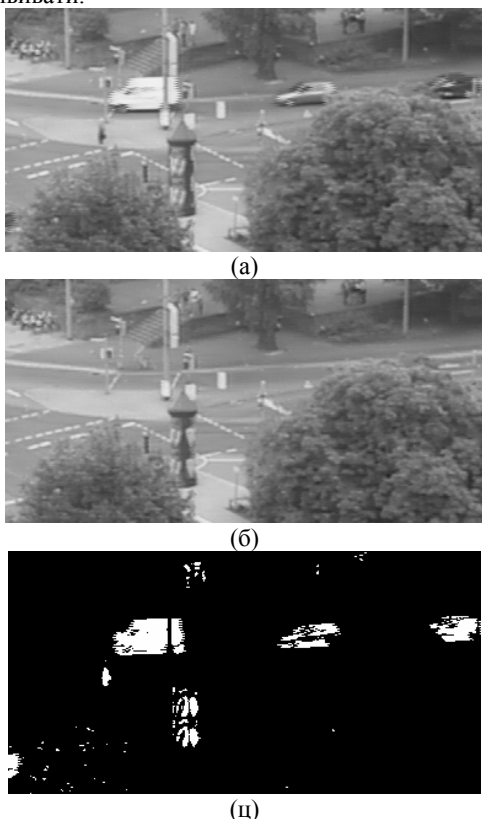


Сл.1. Генерална шема предложеног метода

На сл.2(ц) можемо видети да имамо велики број усамљених пиксела и велики број малих објеката, а којих нема у реалној (текућој) слици. Због тога се на добијену бинарну слику примењују морфолошке операције ерозије и дилатације.

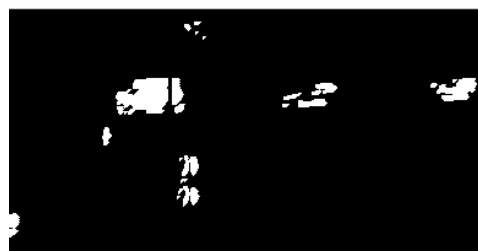
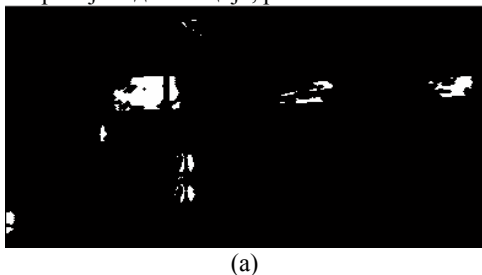
Операције ерозије и дилатације се састоје у интеракцији бинарне слике и структурног елемента којом се мења облик објекта. Структурни елемент чини група пиксела која формира одређени облик (квадрат, крст, круг, итд.). Структурни елементи које смо користили у самом алгоритму праћења за ерозију и дилатацију су имали димензије  $3 \times 3$  пиксела. Ерозијом бинарне слике  $B(x,y)$  структурним елементом  $S_e$  добијамо бинарну слику  $B_e(x,y)$  која представља скуп тачака  $(x,y)$  таквих да је после translације структурног елемента у тачку  $(x,y)$ ,

цео структурни елемент  $Se_{x,y}$  укључен у  $B(x,y)$ . На овај начин ћемо елиминисати усамљене пикселе и објекте којих у стварности немамо. Поступак ерозије може се рекурзивно примењивати. Да би се попуниле рупе у објектима настале применом ерозије примењује се дилатација. Дилатацијом бинарне слике  $Be(x,y)$  структурним елементом  $Sd$  добијамо бинарну слику  $Bed(x,y)$  која представља скуп тачака  $(x,y)$  таквих да, после translације структурног елемента у тачку  $(x,y)$ , пресек скупова  $Be(x,y)$  и  $Sd_{x,y}$  не буде празан скуп. Поступак дилатације се такође може рекурзивно примењивати.



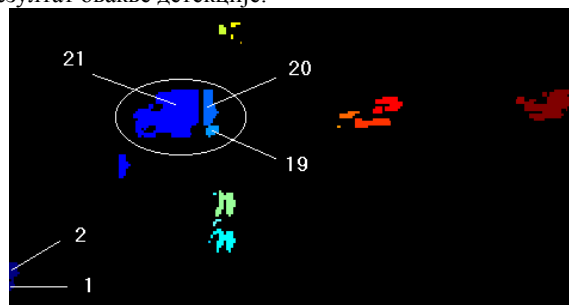
Сл.2. (а) Реална слика, (б) слика позадине и (ц) бинарна слика  $B(x,y)$  која се добија после примене прага  $T=0.125$  на слику разлике  $D(x,y)$

На сл.3 можемо видети резултате који се добијају након примене ерозије и дилатације, респективно.



Сл.3. (а) Реална слика након ерозије и (б) након дилатације

Да би се могли пратити потенцијални објекти потребно је извршити детекцију објеката након примене морфолошких операција ерозије и дилатације. Детекцију објеката смо вршили уз помоћ *Matlab*-ове функције *bwlabel*. Применом ове функције на бинарну слику димензија  $N \times M$  добијамо слику истих димензија као улазна бинарна слика ( $N \times M$ ), а у којој пиксели позадине имају вредност нула, пиксели једног сегмента вредност један, пиксели другог сегмента вредност два итд. За ову функцију детекције можемо сами бирати да ли желимо да користимо повезаност пиксела са 4 или са 8 суседних пиксела. У самом раду је изабрано да се користи повезаност са 4 суседна пиксела. На сл.4 можемо видети резултат овакве детекције.



Сл.4. Резултат раздвајања циљева функцијом *bwlabel*

Са сл.4 видимо да је један објекат (ауто, камион,...) подељен на више сегмената. У нашем случају имамо 21 сегмент (ради прегледности на слици су означени само сегменти 1,2,19,20,21), а свега 7 објеката (циљева). Ово може бити последица морфолошких операција ерозије и дилатације, наилазак објекта на неки природни или вештачки објекат који ће поделити један објекат на више сегмената (дрво, семафор, реклама,...), различите осветљености једног објекта итд. Због тога је потребно извршити спајање сегмената који припадају једном истом објекту. Спајање сегмената који припадају истом објекту је вршено применом следећих корака:

- израчунавање центроида за сваки сегмент,
- подела на парове сегмената који можда потичу од истог објекта (груба подела),
- фина подела на сегменте који потичу од истог објекта,
- спајање сегмената у објекте.

Центроид (тежиште) сегмента је рачунат на основу једначина:

$$x_c = \frac{\sum_{(i,j) \in \text{сегменту}} i \cdot c(i,j)}{\sum_{(i,j) \in \text{сегменту}} c(i,j)}, \quad y_c = \frac{\sum_{(i,j) \in \text{сегменту}} j \cdot c(i,j)}{\sum_{(i,j) \in \text{сегменту}} c(i,j)} \quad (3)$$

где су:

- $x_c, y_c$  координате центроида дуж  $x$  и  $y$  осе,
- $i, j$  координате пиксела у  $x, y$  координатном систему,
- $c(i, j)$  функција класификације пиксела (0 ако је пиксел класификован као позадина, 1 ако је пиксел класификован као сегмент) [4].

Груба подела на парове сегмената који можда потичу од истог објекта је вршена на основу одређивања дистанце међу центронидима. Уколико је ова дистанца мања од задате (40 пиксела – одређено експериментално) пар сегмената је проглашаван кандидатом за спајање.

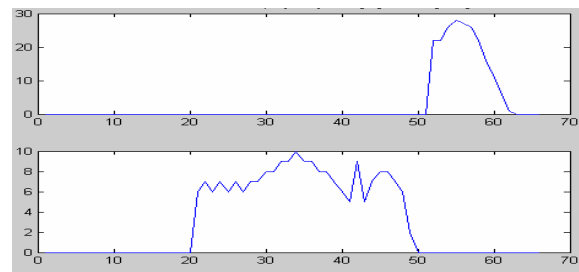
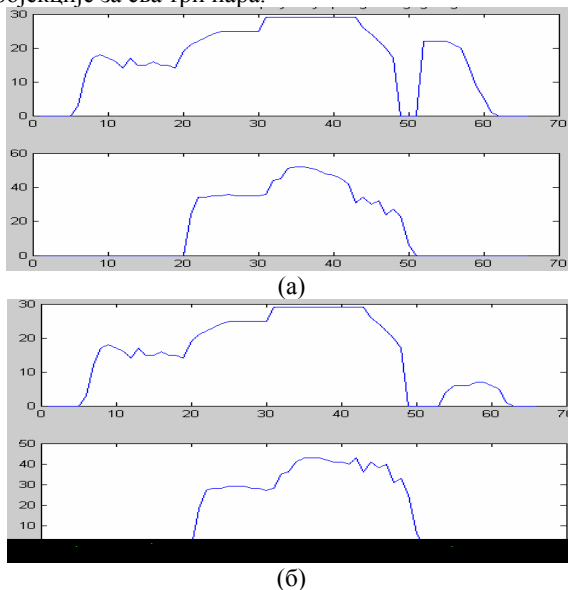
Како смо за грубу поделу користили дистанцу, а наши сегменти могу имати различите пречнике у било ком смеру, потребно је даље извршити фину поделу на сегменте који припадају или не припадају истом објекту. Фина подела је вршена на основу хоризонталне и вертикалне пројекције бинарних слика које се добијају спајањем два сегмента који су проглашени кандидатима за спајање. Пројекције представљају број пиксела који имају вредност један по хоризонтали и вертикали. На крају имамо спајање сегмената који потичу од једног објекта.

За пример узећемо камион са реалне слике након раздвајања сегмената функцијом `bwlabel` (сл. 4). Видимо да је овај камион подељен на три сегмента (семафор је један од разлога због којих је подељен на три сегмента – сл.2(a)). Израчунавањем центроида и рачунањем дистанци између њих добили смо да су ови сегменти кандидати за спајање (сваки са сваким). Да би проверили да ли они стварно припадају истом објекту извршили смо вертикалну и хоризонталну пројекцију за сва три пара у нашем случају (један-два, један-три и два-три). На сл.5 можемо видети ова три сегмента.



Сл.5. Сегменти камиона (први, други и трећи)

На сл.6 можемо видети вертикалне и хоризонталне пројекције за сва три пара.



(ц)

Сл.6. Вертикалне и хоризонталне пројекције за сва три пара сегмената. (а) пар 1-2, (б) пар 1-3 и (ц) пар 2-3

Код вертикалних пројекција на сл.6(а) и 6(б) видимо да постоје одређене паузе између пројекција (вредност 0 на пројекцијама). Ове паузе нам користе за одређивање да ли два сегмента припадају истом објекту. Пауза за први случај износи два, а за други четири пиксела. Како је ова пауза мања од заданог прага, сегменти су спојени у један објекат.

Праг за спајање је одређен експериментално (7 пиксела). Он не сме да буде мали да не би дошло до раздвајања сегмената који потичу од истог објекта, али не сме да буде ни пуно велики да не би дошло до спајања сегмената који потичу од различитих објеката (нпр. спајање сегмената камиона са пешаком који долази на семафор – сл.2(a)).

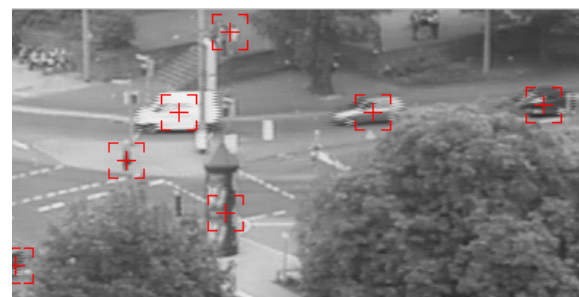
Такође, код хоризонталних пројекција на сл.6(а) и 6(б) и вертикалне и хоризонталне пројекције на сл.6(ц) видимо да не постоје паузе пројекција. Ово може бити последица простирања блиских сегмената по дијагонали. Самим тим може доћи до спајања два сегмента који потичу од различитог објекта а уколико се узимају само хоризонталне и вертикалне пројекције. Да би се ово избегло могу се пројекције одређивати у разним смеровима (за различите углове) а не само хоризонтална и вертикална пројекција. Ово отвара могућност примене Радонове трансформације.

Наш објекат после примене ових корака можемо видети на сл.7.



Сл.7. Објекат (камион)

После оваквог спајања сегмената у објекте, поново за све објекте одређујемо центроиде и на текућој слици обележавамо објекте (циљеве). Обележене објекте можемо видети на сл.8.



### Сл.8. Обележени објекти (циљеви)

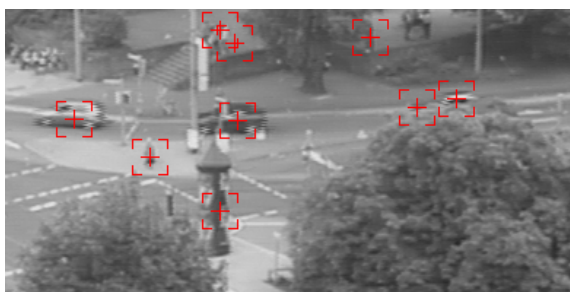
На сл.8 можемо видети да је овакав начин детекције успешно детектовао три аута, један камион, пешака, групу пешака и рекламу. Разлог због кога је детектовао рекламу је тај што ова реклама ротира, па је на основу разлике текуће слике и слике позадине и реклама детектована као један од покретних објеката. Да би се превазишао овај недостатак могуће је имплементирати логику која ће елиминисати циљ ако му се позиција центроида мало мења у сукцесивним фрејмовима. Исти је случај са објектима који не постоје али се детектују (нпр. сенка аутомобила која се креће заједно са њим).

Само праћење објеката вршено је на основу детекције циљева у сукцесивним фрејмовима, без било какве предикције.

### 3. РЕЗУЛТАТИ

У раду је коришћено 35 фрејмова секвенце Karl-Wilhelm-Strasse. Временски размак између фрејмова износио је 166 ms. Димензије оригиналних слика ове секвенце су 740x560 пиксела. Да би се убрзао алгоритам детекције и праћења користили смо исечке ових слика који су били димензија 201x391 пиксела. Други разлог због кога смо користили исечке је тај што овај алгоритам не даје задовољавајуће резултате ако нам циљеви долазе или одлазе фронтално у односу на камеру. У оваквим случајевима долазило је до спајања сегмената који припадају различитим објектима, па и до спајања самих објеката.

Резултате детекције и праћења можемо видети на сл.9. На њој су издвојени 11-ти, 24-ти и 35-ти фрејм.



(a)

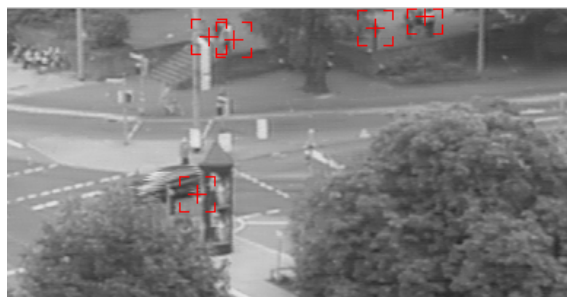


(б)

Детекцију и праћење радили смо на PC Pentium-IV на 1,5 GHz. Средње време обраде једног фрејма укључујући и приказивање обележених објеката износило је 1.57 s а средње време обраде без приказивања 1.35 s.

У доступној литератури дошли смо до податка да је за процес праћења споропокретних објеката довољно обрађивати 2 до 3 фрејма у секунди. На основу

временских перформанси нашег праћења, видимо да би систем радио у реалном времену потребно убрзати извршење алгоритма 2 до 3 пута. Извршење се може убрзати употребом бржих процесора, применом паралелног рада више процесора, поделом текуће слике на више сегмената и њиховом засебном обрадом итд.



(ц)

Сл.9. Резултати детекције различитих фрејмова. (а) 11-ти, (б) 24-ти и (ц) 35-ти

На сл.9(а) видимо да је због шуме дошло до раздвајања аута на два одвојена објекта. Такође, на сл.9(ц) видимо да је дошло до спајања три објекта (аута, рекламе и пешака). У зависности од примене алгоритма праћења имплементирала би се одређена логика која би ове појаве елиминисала. Један од начина да се елиминишу ове појаве је предикција положаја центроида објекта праћења уз помоћ Калмановог филтера.

### 4. ЗАКЉУЧАК

Овај алгоритам детекције и праћења може се применити за одређивање положаја и димензија објеката, контролу саобраћаја, видео надзор итд. Даље правце рада можемо усмерити ка убрзавању самог алгоритма детекције и праћења, употреби одговарајућих процедура за елиминисање наведених недостатака, препознавању детектованих објеката и предикцији положаја центроида.

### ЛИТЕРАТУРА

- [1] G.L.Foresti, "Object recognition and tracking for remote video surveillance", *IEEE Trans. Circuits Syst.*, vol. 9, pp. 1045-1062, Oct. 1999.
- [2] G.L.Foresti, "A real time system for video surveillance of unattended outdoor environments", *IEEE Trans. Circuits Syst.*, vol. 8, pp. 697-704, Oct. 1998.
- [3] М.Поповић, "Дигитална обрада слике", Рукопис за припремање испита, Београд, 1998.
- [4] D.R.Van Rheeden, R.A.Jones, "Noise effects on centroid tracker aim point estimation", *IEEE Trans. on Aerospace and Electronic Syst.*, vol. 24, March 1988.

**Abstract** – In this paper the method for detection and tracking of slow moving objects in video sequence is proposed. The method is based on a frame by frame motion detection and morphological operators. The proposed method for detection and tracking is able to use for ground traffic control, tracking of airspace object and video surveillance.

### ONE APPROACH TO DETECTION AND TRACKING OF SLOW MOVING OBJECTS IN VIDEO SEQUENCE

Boban Bondzulich, Slobodan Simic, Bojan Zrnich