

MULTISENZORSKO IZDVAJANJE I AŽURIRANJE SLIKE POZADINE U DETEKCIJI POKRETNIH OBJEKATA

Bondžulić P. *Boban*, Vojna akademija, Katedra vojnoelektronskih sistema, Beograd,
Petrović S. *Vladimir*, Imaging Science, University of Manchester

UDC: 621.397.3/4

Sažetak:

U radu je predložen postupak multisenzorskog izdvajanja i ažuriranja slike pozadine koji se koristi za video nadzor i detekciju pokretnih objekata. Izdvajanje slike pozadine zasnovano je na detekciji promena na slici i iterativnom postupku ažuriranja. Pri detekciji promena koriste se informacije raspoloživih senzora, čime se povećava pouzdanost detekcije i slike pozadine. Predloženi postupak testiran je na bazama registrovanih televizijskih i termovizijskih slika.

Ključne reči: multisenzorska detekcija, izdvajanje slike pozadine, inicijalizacija i ažuriranje slike pozadine.

Uvod

Istraživanja u oblasti obrade video signala često su posvećena video nadzoru. Da bi se dobile odgovarajuće performanse video nadzora potrebna su robustna rešenja detekcije i praćenja pokretnih objekata. Zbog različitih ambijentalnih uslova rešenja koja koriste jedan senzor nisu pouzdana u svim situacijama. Da bi se prilagodili različitim ambijentalnim uslovima potrebno je koristiti senzore različitih tipova [1, 2, 3, 4, 5, 6, 7, 8].

Za detekciju pokreta (promena) na slikama scene može se koristiti obrada slike razlike, koja se dobija oduzimanjem slike pozadine od tekućeg frejma ili oduzimanjem sukcesivnih frejmova. Detekcija je praćena uvođenjem praga kako bi se dobila binarna slika u kojoj su pikseli klasifikovani u dve klase: pozadina i pokretni objekti [9].

U mnogim sistemima video nadzora bitno je da se inicijalizuje (izdvoji) slika pozadine. Kako se za inicijalizaciju koristi trening video sekvencija koja sadrži pokretne objekte, potrebni su robustni metodi koji mogu izdvojiti sliku pozadine uz postojanje pokretnih objekata i šuma.

Dobar metod izdvajanja slike pozadine mora biti robustan u odnosu na promene u osvetljenosti scene, mora anulirati pojave na sceni, kao što su treperenje lišća, kiša, sneg i senke objekata i mora se prilagoditi promenama na sceni, kao što su polazak i zaustavljanje objekata.

Detekcija promena oduzimanjem sukcesivnih frejmova

Označimo sa $I^n(x,y)$ vrednost nivoa sivog piksela frejma (slike) na poziciji (x, y) u trenutku $t = t_n$. Mera sličnosti ovog frejma i frejma dobijenog u trenutku $t = t_{n-1}$ može se izraziti kao [10]:

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

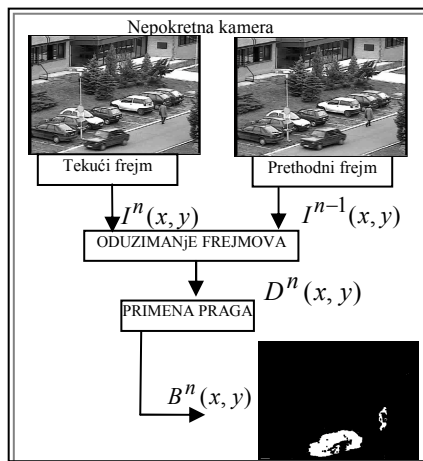
gde je $N \times M$ dimenzija frejma u pikselima. Na ovaj način dobija se slika razlike D^n .

Primenom praga T na sliku razlike $D^n(x,y)$ dobija se binarna slika u kojoj su pikseli klasifikovani u dve klase: pozadina i pokretni objekti, tj.:

$$B^n(x, y) = \begin{cases} 1 \equiv \text{pokretni objekat}, & D^n(x, y) > T \\ 0 \equiv \text{pozadina}, & D^n(x, y) \leq T \end{cases} \quad (2)$$

To znači da je piksel pokretan ako se njegov intenzitet nivoa sivog bitno promenio između dva frejma.

Generalna šema detekcije promena na slici scene oduzimanjem sukcesivnih frejmova dobijenih sa nepokretne kamere prikazana je na slici 1.



Slika 1 – Generalna šema detekcije promena oduzimanjem sukcesivnih frejmova

Sa slike 1 se vidi da na ovakav način nisu detektovani pikseli unutar objekata.

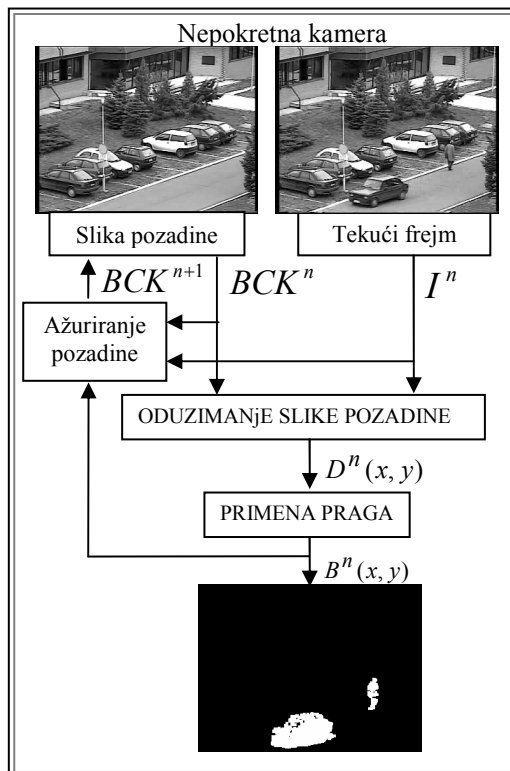
Detekcija promena oduzimanjem slike pozadine od tekućeg frejma

Označimo sa $BCK^n(x,y)$ intenzitet piksela slike pozadine na poziciji (x,y) u trenutku $t=t_n$. Detekcija pokreta može se izvršiti na osnovu razlike tekućeg frejma i slike pozadine [11]:

$$D^n(x,y) = |I^n(x,y) - BCK^n(x,y)| \quad (3)$$

Primenom praga T na sliku razlike $D^n(x,y)$ dobija se binarna slika u kojoj su pikseli klasifikovani u dve klase.

Generalna šema detekcije promena na slici scene oduzimanjem slike pozadine od tekućeg frejma prikazana je na slici 2.



Slika 2 – Generalna šema detekcije promena oduzimanjem slike pozadine od tekućeg frejma

Sa slike 2 vidi se da su oduzimanjem slike pozadine detektovana dva pokretna objekta (vozilo i pešak) i da je detekcija jasnija od detekcije sa sl. 1.

Ako se za klasifikaciju piksela slike koristi jedan prag, tada se on naziva globalni prag. Ovaj način klasifikacije najčešće se primenjuje za odvajanje objekata od pozadine. Takvi su slučajevi, na primer, pri izdvajanju pisanog ili štampanog teksta, analize nekih biomedicinskih slika, prepoznavanja tipa aviona koji leti, itd. Osim primene globalnog praga, slika se može podeliti na blokove i za svaki od blokova posebno određivati prag. Pragovi koji se koriste za klasifikaciju piksela blokova nazivaju se lokalni pragovi. Vrednosti pragova mogu biti određene i za svaki piksel ponaosob.

Većina tehnika u određivanju praga koristi histogram slike. U idealnom slučaju histogram slike je bimodalan, sa dva maksimuma između kojih postoji izražen minimum. Nivo sivog koji odgovara minimumu može se odabrati za prag. Međutim, histogram slike razlike često je unimodalan, čime se usložava postupak određivanja praga klasifikacije piksela slike razlike.

Kombinovani metod detekcije promena

Kombinovani metod detekcije predstavlja spoj tehnika oduzimanja sukcesivnih frejmova i oduzimanja slike pozadine [12]. Osnovni nedostatak tehnike oduzimanja slike pozadine od tekućeg frejma jeste što je ova tehnika osetljiva na situacije kada se stacionarni objekti scene počinju kretati. Iako će ovi objekti biti detektovani, oni iza sebe ostavljaju „rupe“ gde se novootkrivena pozadina razlikuje od poznate slike pozadine. Dok se model pozadine prilagodi ovim „rupama“, algoritam detekcije određeno vreme generiše lažne objekte. Razlika sukcesivnih frejmova nije osetljiva na pojavu „rupa“, ali ovaj algoritam menja oblik pokretnog objekta.

Da bi se izbegli navedeni nedostaci koristi se kombinacija ova dva metoda detekcije. Grupisanjem pokretnih piksela detektovanih oduzimanjem sukcesivnih frejmova dobijaju se regioni pokretnih piksela. Delovi tekućeg frejma u kojima se nalaze pokretni objekti određuju se na osnovu krajnjih tačaka svakog regiona pokretnih piksela, tj. tako što se svaki region pokretnih piksela obuhvati pravougaonikom minimalnog obima (*minimum bounding rectangle*).

Označimo sa R_i^n deo tekućeg frejma u kojem se nalazi pokretni objekat. Pokretni objekat b_i^n biće detektovan ako uzmemo sve piksele iz R_i^n , čiji intenzitet se bitno razlikuje od modela pozadine, tj.:

$$b_i^n = \left\{ (x, y) : \left| I^n(x, y) - BCK^n(x, y) \right| > T_i^n, (x, y) \in R_i^n \right\} \quad (4)$$

gde je T_i^n prag detekcije promena u delu tekućeg frejma u kojem se nalazi pokretni objekat.

Osim razlike dva sukcesivna frejma, za određivanje regiona pokretnih piksela može se koristiti razlika tri ili više sukcesivnih frejmova.

Zbog prednosti u odnosu na tehnike oduzimanja sukcesivnih frejmova i oduzimanja slike pozadine od tekućeg frejma u radu se za detekciju promena koristi kombinovani metod detekcije [12].

Inicijalizacija slike pozadine

U nekim primenama detekcije pokreta na sceni nije moguće da se iz same sekvence izdvoji jedan frejm i proglasi pozadinom. Ovaj problem se javlja ukoliko u sekvenci uvek postoje pokretni objekti (autoput, raskrsnice...). Zbog toga je potrebno formirati sliku pozadine $BCK^n(x,y)$ iz sekvence frejmova $I^t(x,y)$ koja može sadržati pokretne objekte. Kako je za detekciju pokretnih objekata potrebna slika pozadine, prvi korak u detekciji treba da bude izdvajanje pozadine.

Za izdvajanje pozadine može da se koristi srednja vrednost ili medijana određenog broja frejmova [13]:

$$BCK^n(x,y) = \text{mean}_t I^t(x,y),$$

$$BCK^n(x,y) = \text{med}_t I^t(x,y) \quad (5)$$

Sa dovoljno dugom sekvencom slika, usrednjavanje vrednosti nivoa sivog za svaki piksel daće prihvatljivu sliku pozadine. Međutim, broj frejmova koji je potreban ne može se odrediti unapred, a za dati broj frejmova rezultat zavisi od brzine pokretnih objekata. Medijana je robustnija od srednje vrednosti, ali može da toleriše do 50% uticaja šuma i pokretnih objekata.

U [14] se pošlo od pretpostavke da su vrednosti nivoa sivog piksela pozadine stabilne i relativno konstantne određeno vreme. Za određivanje vrednosti nivoa sivog piksela pozadine koristi se srednja vrednost nivoa sivog najstabilnijeg intervala nepreklapajućih subintervala $\{I_k\}$. Najstabilniji interval određuje se na osnovu:

$$\hat{I}_k = \arg \max_k \left(\frac{n_{I_k}}{S_{I_k}} \right) \quad (6)$$

gde su n_{I_k} i S_{I_k} dužina i varijansa nivoa sivog k -te subsekvence I_k . Međutim, za sekvence u kojima su pokretni objekti stacionarni određeno vreme mnogi pikseli se pogrešno klasifikuju. Takođe, vrednosti nivoa sivog objekata koji su veliki, spori i imaju slabu teksturu mogu biti stabilne i relativno konstantne određeno vreme.

Jedan od metoda koji se koristi daje sliku pozadine u nekoliko iteracija [11, 12]. Metod polazi od početne slike pozadine koja je jednaka prvom frejmu $BCK^0(x,y)=I^0(x,y)$. Da bi se dobila slika pozadine koja odgovara tekućem frejmu moraju se detektovati pokretni pikseli, a zatim izdvojiti pozadina na osnovu:

$$BCK^{n+1}(x,y) = \begin{cases} \alpha BCK^n(x,y) + (1-\alpha)I^n(x,y), (x,y) \text{ nepokreta } \bar{u} \\ BCK^n(x,y), (x,y) \text{ pokreta } \bar{u} \end{cases} \quad (7)$$

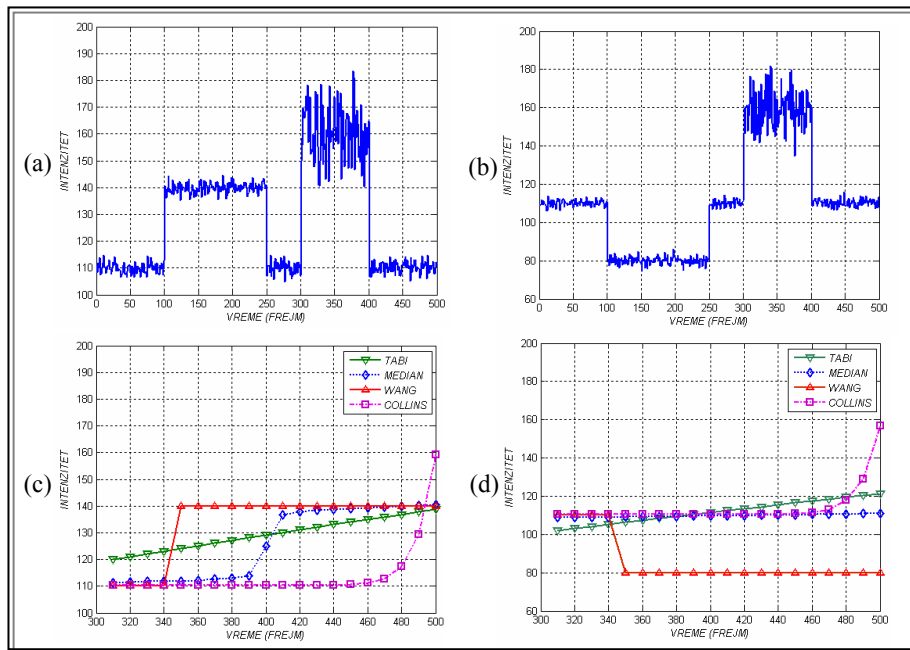
Iz (7) vidi se da do promene vrednosti piksela slike pozadine dolazi ako je piksel klasifikovan kao nepokretan, tj. ako je deo stacionarne pozadine. U suprotnom bi pokretni pikseli „poremetili“ sliku pozadine. Od koeficijenta α zavisi brzina promene. U radu se koristi koeficijent $\alpha = 0,9$ koji je predložen u literaturi [11,12].

Da bi se prilagodili promenama scene (dan-noć, promene u osvetljenju, promene uslovljene senkom...) potrebno je da se vrši ažuriranje pozadine, koje se vrši na osnovu (7), nakon svakog frejma ili u određenim intervalima.

Kako bi se izvršilo poređenje četiri metoda izdvajanja pozadine (usrednjavanjem – *TABI*, određivanjem medijane – *MEDIAN*, određivanjem stabilnih intervala – *WANG* i na osnovu detekcije promena – *COLLINS*) sproveden je eksperiment sa veštački generisanim nizom nivoa sivog piksela. Vrednosti nivoa sivog generisane su za 500 sukcesivnih frejmova nakon čega je određena vrednost nivoa sivog pozadine. U toku prvih 100 frejmova dostupne su vrednosti nivoa sivog pozadine – srednja vrednost nivoa sivog je 110, a varijansa 4. U narednih 150 frejmova otkriva se prvi objekat. Nakon prolaska prvog objekta otkriva se pozadina u 50 frejmova, a u poslednjih 200 frejmova otkrivaju se drugi objekat i pozadina. Srednja vrednost nivoa sivog drugog objekta je 160, a varijansa 100. Vreme otkrivanja drugog objekta menjano je od 1 do 200 frejmova. Za svaki od intervala otkrivanja drugog objekta određivana je vrednost nivoa sivog pozadine.

Na sl. 3a i 3b prikazane su realizacije veštački generisanih nivoa sivog piksela. U prikazanim realizacijama drugi objekat otkriva se u 100 frejmova. Srednja vrednost nivoa sivog prvog objekta na sl. 3a je 140, a na sl. 3b– 80. Varijansa nivoa sivog prvog objekta u oba slučaja je 4.

Na sl. 3c i sl. 3d prikazane su dobijene vrednosti nivoa sivog pozadine po četiri analizirana metoda i to za različite intervale otkrivanja drugog objekta.



Slika 3 – (a) (b) Realizacije generisanih nivoa sivog; (c) (d) rezultati izdvajanja slike pozadine

Sa sl. 3c i 3d vidi se da srednja vrednost nivoa sivog slike pozadine linearno raste sa vremenom otkrivanja drugog objekta. Korišćenjem stabilnih intervala pravilno se izdvaja slika pozadine sve dok je otkrivanje drugog objekta manje od 50 frejmova, nakon čega se za sliku pozadine proglašava srednja vrednost nivoa sivog prvog objekta.

Korišćenjem medijane na sl. 3c ispravno se izdvajaju vrednosti nivoa sivog pozadine sve dok je otkrivanje drugog objekta manje od 100 frejmova, nakon čega se za sliku pozadine proglašava srednja vrednost nivoa sivog prvog objekta. Korišćenjem medijane na sl. 3d ispravno se izdvajaju vrednosti nivoa sivog pozadine za sve intervale otkrivanja drugog objekta.

Iz sprovedenog eksperimenta vidi se da se najbolji rezultati izdvajanja slike pozadine dobijaju korišćenjem iterativnog postupka predloženog u [11,12], što je i bio razlog za njegovo prihvatanje u radu.

Multisenzorsko izdvajanje i ažuriranje slike pozadine

U jednom multisenzorskom sistemu dostupne su nam slike iste scene sa više senzora istog ili različitog tipa. Sprezanjem više senzora u jedinstven sistem i sjedinjavanjem informacija dobijenih od njih mogu se

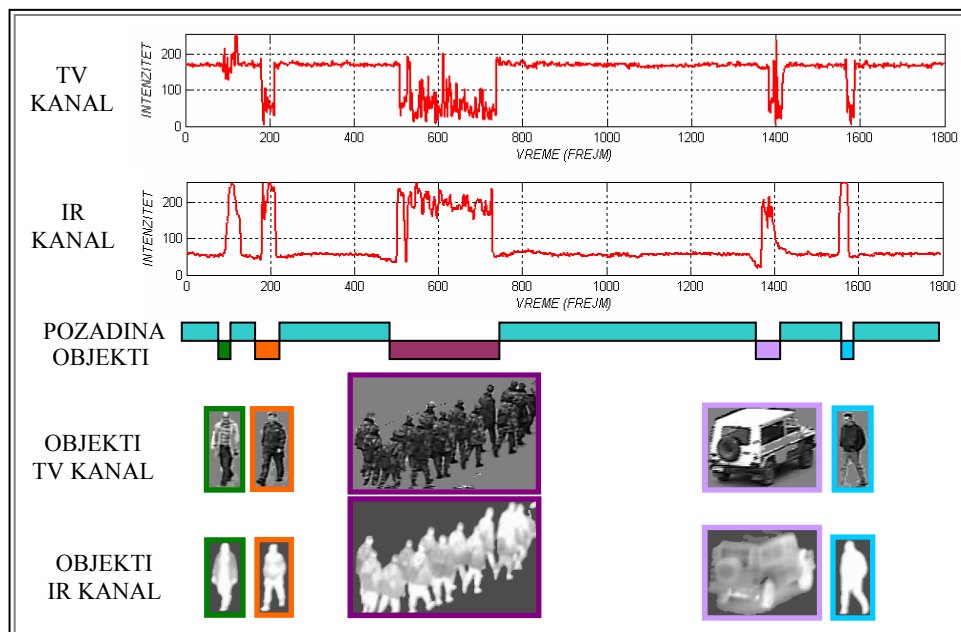
eliminirati nedostaci pojedinačnih senzora i time omogućiti detekcija u svim ambijentalnim uslovima.

Rezultati izdvajanja i ažuriranja slike pozadine korišćenjem relacije (7) zavise od: praga detekcije koji se koristi za određivanje pokretnih piksela. Primenom praga na sliku razlike dobija se matrica pokretnih piksela.

Na sl. 4 prikazane su promene nivoa sivog piksela televizijske (TV) i termovizijske (IR) sekvence u toku 1800 frejmova. Prikazane su promene nivoa sivog piksela na istoj poziciji prostorno i vremenski usaglašenih televizijskih i termovizijskih slika. Kako su slike usaglašene pikseli slika predstavljaju istu tačku scene.

Sa sl. 4 uočava se da postoji korelisanost pojava i procesa na televizijskim i termovizijskim slikama. Intervali u kojima se otkriva pozadina su stabilni i dugi (što nije uvek slučaj). Kontrast objekta u odnosu na pozadinu nekada je bolji u TV, a nekada u IR sekvenci slika.

Od kontrasta objekta u odnosu na pozadinu zavisi detekcija promena na slici. Zbog toga se u radu pošlo od ideje da se za sve senzore (kanale) u multisenzorskom sistemu formira jedna matrica pokretnih piksela za izdvajanje i ažuriranje slike pozadine pojedinačnih senzora. Matrica pokretnih piksela formira se sjedinjavanjem informacija o pokretu koje se dobijaju sa svih senzora.



Slika 4 – Promene nivoa sivog piksela televizijske i termovizijske sekvence u toku 1800 frejmova

Pri izdvajanju i ažuriranju slike pozadine potrebno je detektovati pokretne piksele. Za detekciju pokretnih piksela korišćen je kombinovani metod detekcije promena opisan jednačinama (1), (2) i (4). U jednačinama (2) i (4) potrebno je odrediti vrednosti pragova detekcije promena (pragova segmentacije). Ako slika sadrži šum ili ako su nivoi sivog objekta i pozadine slični, određivanje najbolje vrednosti praga postaje složeno. Problemi se mogu pojaviti i ako pozadina nije uniformno osvetljena.

Pri određivanju praga detekcije na slici razlike sukcesivnih frejmova (2) korišćen je metod zasnovan na Ojlerovom (*Euler*) broju [13]. Navedeni metod je korišćen zbog toga što su vrednosti praga koje se dobijaju nešto manje od očekivane vrednosti [15]. Time se povećava verovatnoća detekcije, ali i verovatnoća lažnog alarma. Lažni alarmi mogu se eliminisati primenom morfoloških operacija, zanemarivanjem detektovanih regiona čije su dimenzije manje od unapred zadatih ili se eliminišu u narednom frejmu (slučajan šum).

Umesto binarnog odlučivanja datog sa (4) pri oduzimanju slike pozadine korišćena su dva praga, T_1 i T_2 , ($T_1 < T_2$), i za svaki od detektovanih objekata određene su matrice pouzdanosti. Ukoliko su vrednosti nivoa sivog slike razlike veće od praga T_2 pouzdanost je jednaka 1. Ukoliko su vrednosti nivoa sivog manje od praga T_1 pouzdanost je jednaka 0. Za vrednosti nivoa sivog koje su između vrednosti pragova izvršeno je linearno skaliranje [16]:

$$C_{ch}^n(x, y) = \frac{D_{ch}^n(x, y) - T_1^{ch}}{T_2^{ch} - T_1^{ch}}, \quad T_1^{ch} < D_{ch}^n(x, y) \leq T_2^{ch} \quad (8)$$

U opštem slučaju $ch \in \{1, 2, \dots, k\}$, gde je k broj korišćenih senzora. Kako se u radu koriste dva senzora (televizijska – TV i termovizijska kamera – IR) u našem slučaju $ch \in \{TV, IR\}$.

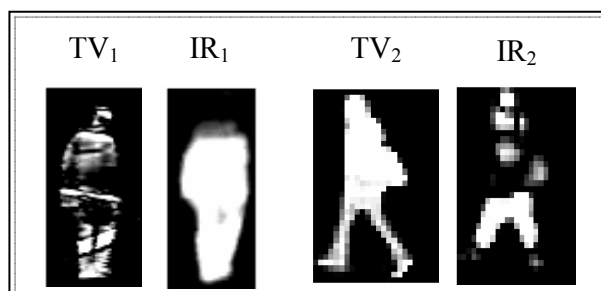
Za određivanje vrednosti pragova T_1 i T_2 mogu se koristiti neke od tehnika podele slike na dve ili više klasa predloženih u [17,18,19,20], koje pri određivanju optimalnih vrednosti pragova koriste histogram slike. U ovom radu vrednosti pragova određuju se korišćenjem tehnike predložene u [18], u kojoj se koristi Šenonova definicija entropije. Posmatraju se dve gustine verovatnoće, jedna za pozadinu, a druga za objekat, a optimalna vrednost praga dobija se maksimiziranjem zbira entropija pozadine i objekta. Ova tehnika je izabrana, jer je pokazano da ima bolje performanse u odnosu na ostale tehnike određivanja optimalne vrednosti praga [21].

Kako su promene nivoa sivog bilo kog senzora indikator regiona pokretnih piksela, za određivanje sjedinjene matrice pouzdanosti uzima se maksimum matrica pouzdanosti svih senzora:

$$C_{fus}^n(x, y) = \max \{C_{ch}^n(x, y)\}, \quad ch \in \{1, 2, \dots, k\} \quad (9)$$

Matrice pouzdanosti zavise od kvaliteta senzora koji se koriste, dimenzija pokretnih objekata, daljine do pokretnih objekata, kontrasta objekata u odnosu na pozadinu, vremenskih uslova u kojima se vrše snimanja, doba dana i sl. Zbog toga je bolje ići na formiranje sjedinjene matrice pouzdanosti nego za matricu pouzdanosti sistema uvek uzimati matricu pouzdanosti jednog senzora.

Na sl. 5 prikazane su matrice pouzdanosti dobijene pomoću televizijskih i termovizijskih kamera dva multisenzorska sistema (sistemi 1 i 2).



Slika 5 – Matrice pouzdanosti dobijene sa TV i IR kamera dva multisenzorska sistema

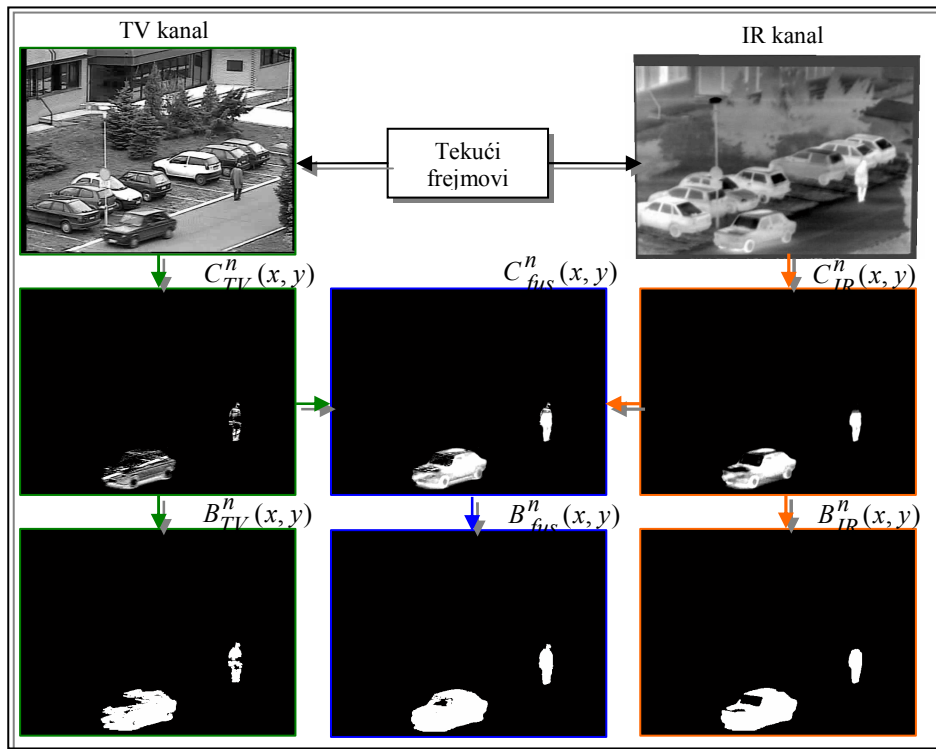
Sa sl. 5 vidi se da je u prvom slučaju kompletniji opis pokretnog objekta korišćenjem termovizijske, a u drugom slučaju korišćenjem televizijske kamere.

Nakon formiranja matrice pouzdanosti određuju se regioni povezanih pokretnih piksela i eliminišu regioni unutar kojih ne postoje segmenti čije su vrednosti pouzdanosti jednake 1. Preostali regioni pripadaju pokretnim objektima i na navedenim pozicijama vrednosti nivoa sivog pozadine ostaju iste kao u prethodnom frejmu, dok se na ostalim pozicijama vrši ažuriranje slike pozadine u skladu sa jednačinom (7).

Na sl. 6 prikazan je primer formiranja matrice pokretnih piksela sjedinjavanjem informacija od dva senzora. Prikazani su tekući frejmovi dve sekvence, matrice pouzdanosti i matrice pokretnih piksela oba senzora, sjedinjena matrica pouzdanosti i rezultati detekcije korišćenjem informacija iz oba kanala.

Sa sl. 6 vidi se da je matrica pokretnih piksela dobijena sjedinjavanjem informacija od dva senzora kompletnija od matrica pokretnih piksela pojedinačnih senzora. Na taj način, korišćenjem sjedinjene matrice pokretnih piksela u procesu izdvajanja i ažuriranja slike pozadine, neće biti “zagađivanja” slike pozadine sa pokretnim pikselima.

Jedno od alternativnih rešenja formiranja sjedinjene matrice pokretnih piksela jeste da se u toku vremena određuje pouzdanost svakog od senzora [22]. Poređenjem pouzdanosti senzora može se odabrati najpogodniji senzor i pokretni pikseli odrediti samo po najpouzdanijem senzoru.



Slika 6 – Rezultati detekcije promena korišćenjem pojedinačnih senzora i sjedinjavanjem informacija

Rezultati izdvajanja slike pozadine

Opisani pristup multisenzorskog izdvajanja i ažuriranja slike pozadine testiran je na bazama prostorno i vremenski usaglašenih televizijskih i termovizijskih slika. Korišćene su slike dostupne iz baze autora i iz OTCBVS Benchmark Dataset Collection baze [23].

Slike iz baze autora dobijene su pomoću televizijske kamere firme Samsung, tip SCL860 i termovizijske kamere firme ATIS (Advanced Thermal Imaging System). Formirane su baze od po jedanaest sekvenci televizijskih i termovizijskih slika dve scene, bliske scene na daljini od 70 m (parking u Vojnotehničkom institutu - Beograd) i daleke scene na daljini od 4 km (magistralni put Beograd–Obrenovac i njegova okolina). Baza slika je formirana upotrebom nepokretnih kamera postavljenih na platformu zgrade. U radu se koriste sekvence bliske scene.

OTCBVS Benchmark Dataset Collection je javno dostupna kolekcija namenjena za poređenje algoritama obrade video signala. Kolekcija se sastoji

od video sekvenci i slika iz vidljivog i infracrvenog dela elektromagnetnog spektra (EMS). Čini je sedam baza slika i sekvenci. Detaljniji podaci o kolekciji prikazani su u tabeli 1. U radu se koristi treća baza slika (Dataset 03) namenjena istraživanjima u oblasti multisenzorskog sjedinjavanja informacija.

Podaci o korišćenim sekvencama prikazani su u tabeli 2.

Tabela 1

Podaci o OTCBVS Benchmark Dataset Collection

Baza	Opis	Namena
Dataset 01	Baza termalnih slika pešaka	Detekcija pešaka u termalnim sekvencama
Dataset 02	Baza kolor/termalnih slika ljudskog lica	Neregistrovane termalne/kolor slike ljudskog lica snimljene za različite osvetljenosti, ekspresije lica i poze
Dataset 03	Baza kolor/termalnih slika	Sjedinjavanje informacija kolor i termalnih slika i detekcija pokretnih objekata
Dataset 04	Baza IR slika ljudskog lica	Analiza ljudskog lica u termalnim slikama
Dataset 05	Baza pokretnih objekata u IR domenu	Detekcija i praćenje objekata u IR domenu
Dataset 06	Baza IR slika naoružanja	Detekcija naoružanja i detekcija opaljivanja u termalnim slikama
Dataset 07	Baza slika ljudskog lica u bliskom infracrvenom delu EMS (Near-infrared – NIR)	NIR detekcija ljudskog lica, NIR detekcija očiju, NIR prepoznavanje lica

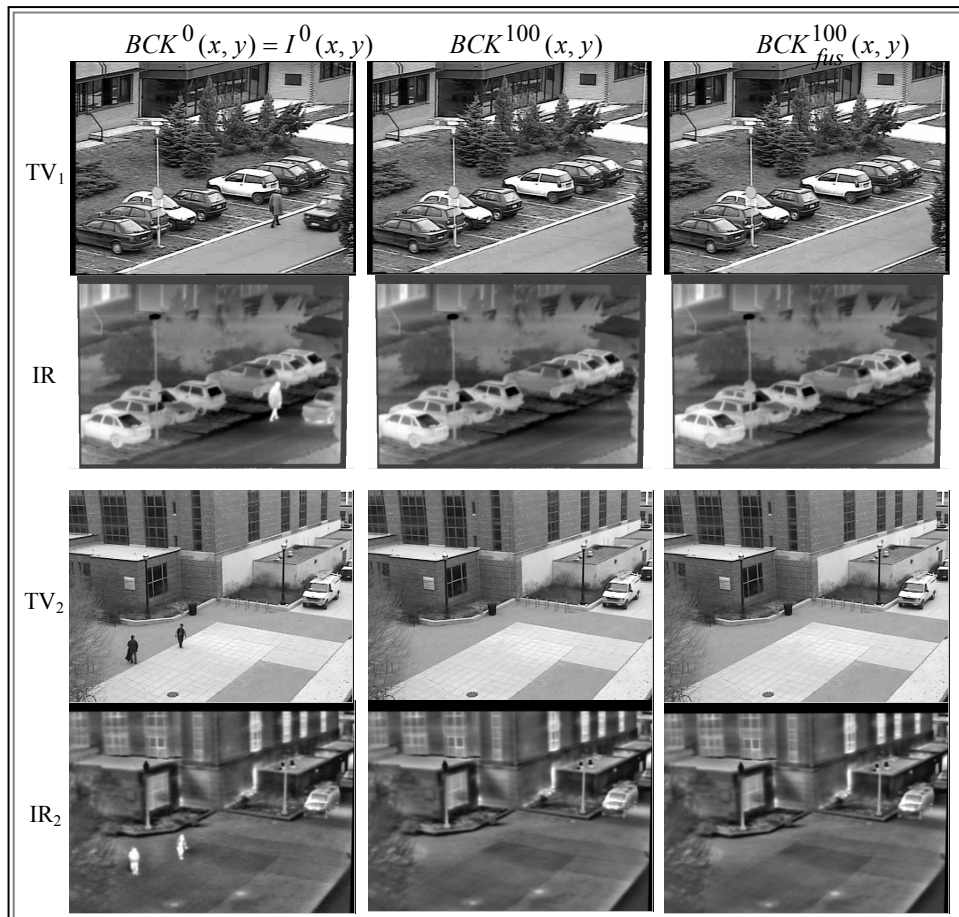
Tabela 2

Podaci o korišćenim sekvencama slika

Baza		Baza autora	OTCBVS Dataset 03
Podaci		Parking	Raskrsnica
Podaci o senzoru	Termalni	ATIS, (8-12) μm	Raytheon PalmIR 250D
	Kolor	Samsung SCL860	Sony TRV87 Handycam
Broj sekvenci		6	6
Slike	Termalne	8-bitne	8-bitne
	Kolor	24-bitne	24-bitne
Dimenzije [piksela]		480 x 640	240 x 320
Broj frejmova u s		30 Hz	30 Hz
<ul style="list-style-type: none"> ✓ Kamere postavljene jedna do druge ✓ Registracija slika izvršena je korišćenjem ručno odabranih tačaka 			

Na sl. 7 prikazani su rezultati izdvajanja (ažuriranja) slike pozadine dve usaglašene sekvence slika. Prikazane su inicijalne slike pozadina, slike pozadina dobijene korišćenjem informacija pojedinačnih senzora i slike pozadina dobijene sjedinjavanjem informacija raspoloživih senzora.

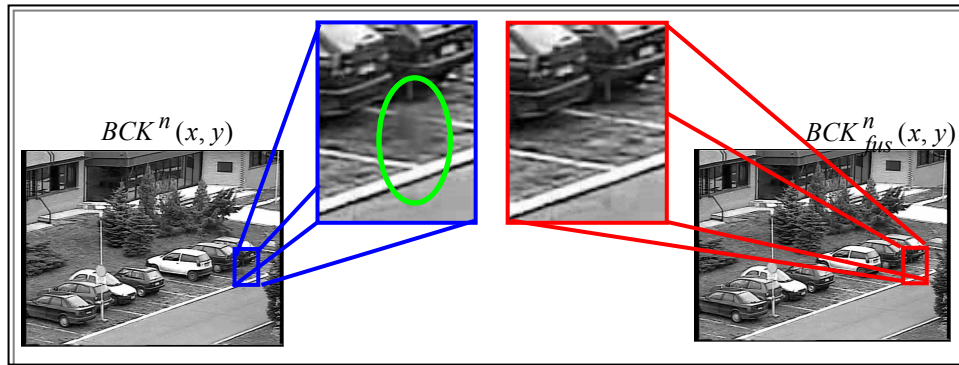
U oba slučaja prikazane su slike pozadina dobijene nakon 100. frejma sekvence, a za inicijalne frejmove odabrani su frejmovi u kojima postoje pokretni objekti.



Slika 7 – Rezultati izdvajanja (ažuriranja) slike pozadine

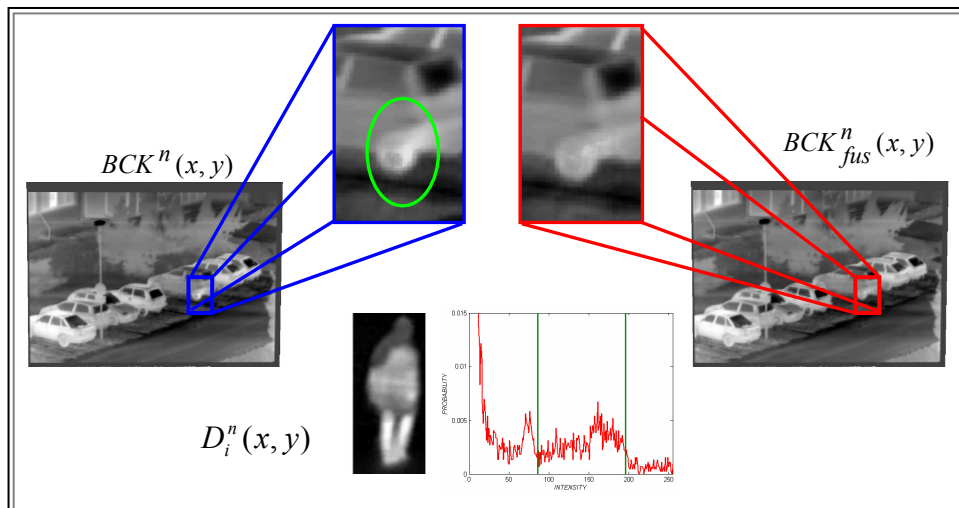
Na slici 7 se vidi da su izdvajanjem slike pozadine nestali objekti prisutni u inicijalnim slikama.

Prednosti združenog izdvajanja slike pozadine ilustrirani su na slikama 8 i 9, na kojima su dati rezultati izdvajanja slike pozadine korišćenjem informacija pojedinačnih senzora i sjedinjavanjem informacija raspoloživih senzora. Prednosti združenog izdvajanja uočavaju se u regionima u kojima se trenutno nalaze pokretni objekti (pešak). Regionima u kojima se trenutno nalaze pokretni objekti izdvojeni su i uvećani.



Slika 8 – Prednosti združenog izdvajanja slike pozadine

Na slici 8 uočava se da je na slici pozadine dobijenoj korišćenjem informacija od TV kamere, u delu slike u kojem se nalazi pokretni objekat, došlo do ugrađivanja nivoa sivog pokretnog objekta u sliku pozadine. Do ugrađivanja je došlo zbog sličnih nivoa sivog pozadine i pokretnog objekta (kaput pešaka), pa su pikseli koji pripadaju pokretnom objektu proglašeni nepokretnim. Korišćenjem informacija sa oba senzora u delu slike u kojem se nalazi pokretni objekat pozadina se u potpunosti otkriva.



Slika 9 – Prednosti združenog izdvajanja slike pozadine

Slična situacija uočava se i na sl. 9. U ovom slučaju objekat (pešak) topliji je u odnosu na pozadinu, ali je ipak došlo do ugradnje nivoa sivog pokretnog objekta u sliku pozadine. Histogram slike razlike u delu slike u kojem se nalazi

pokretni objekat je multimodalan. Primenom pragova T_1 i T_2 na sliku razlike detektuju se dve klase nivoa sivog koje su tople, a ne detektuje se treća klasa koja, takođe, potiče od pokretnih piksela. Na sl. 9 prikazane su slika razlike i njen histogram, na kojem su vertikalnim linijama obeležene vrednosti pragova detekcije. Sa sl. 9 vidi se da je grupa pokretnih piksela (pikseli sa uočljivim pikom ispod praga T_1) proglašena nepokretnim, pa je došlo do ažuriranja slike pozadine sa pokretnim pikselima. Korišćenjem informacija sa oba senzora u delu slike u kojem se nalazi pokretni objekat pozadina se u potpunosti otkriva.

Rezultati izdvajanja slike pozadine zavise od detekcije promena. Kako je detekcija promena zasnovana na razlici frejmova i primeni praga, opisani metod izdvajanja slike pozadine ima određene mane koje ograničavaju primenu u svim sistemima video nadzora: 1) loše osobine u slučaju slika sa jakim šumom ili slabim kontrastom; 2) osetljiv na pomeraj platforme sa koje se snima, odnosno osetljiv na pomeranje pozadine i 3) mala verovatnoća detekcije objekata malih dimenzija. Neki od navedenih nedostataka mogu se eliminisati korišćenjem vremensko-frekvencijske analize video signala [24, 25].

Zaključak

U radu je predložen robustan postupak multisenzorskog izdvajanja i ažuriranja slike pozadine koji se koristi u detekciji pokreta na slici scene. Opisani postupak koristi poznate koncepte u detekciji promena kako bi se pikseli klasifikovali u pokretne piksele i piksele pozadine, nakon čega se koristi iterativni postupak izdvajanja – ažuriranja slike pozadine. Sjedinjavanjem informacija raspoloživih senzora o trenutnim promenama na sceni vrši se ažuriranje slike pozadine u regionima u kojima ne postoji pokret.

Predloženi postupak testiran je na sekvencama registrovanih televizijskih i termovizijskih slika. Pokazano je da se sjedinjavanjem informacija raspoloživih senzora dobija kompletniji opis pokretnih objekata i pouzdanija slika pozadine.

Radi se o veoma jednostavnom postupku, što ga čini idealnim za primenu u sistemima multisenzorske detekcije sa radom u realnom vremenu.

U daljem radu planira se predlaganje funkcionalnog sistema detekcije i praćenja pokretnih objekata. Trenutna istraživanja usmerena su na robustnije metode kombinovanja informacija o promenama. Kako bi se rešio problem regularnih pokreta pozadine, biće razmotreni i mnogo kompleksniji modeli pozadine, kao što su lokalne multimodalne distribucije.

Literatura

[1] Jones, G. D.; Hodgetts, M. A.; Allsop, R. E.; Sumpter, N.; Vicencio-Silva, M.-A.: A novel approach for surveillance using visual and thermal images, DERA/IEE Workshop on Intelligent Sensor Processing, 2001.

- [2] Toet, A.: Detection of dim point targets in cluttered maritime backgrounds through multisensor image fusion, *Proceedings of SPIE*, Vol. 4718, pp. 118–129, 2002.
- [3] Ye, G.; Wei, J.; Pickering, M. R.; Frater, M.R.; Arnold, J. F.: Simultaneous tracking and registration in a multisensor surveillance system, *International Conference on Image Processing – ICIP*, 2003.
- [4] Kang, J.; Gajera, K.; Cohen, I.; Medioni, G.: Detection and Tracking of Moving Objects from Overlapping EO and IR Sensors, *Joint IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum*, 2004.
- [5] O Conaire, C.; O'Connor, N. E.; Cooke, E.; Smeaton, A. F.: Multispectral Object Segmentation and Retrieval in Surveillance Video, *IEEE International Conference on Image Processing – ICIP*, 2006.
- [6] Dwyer, D.; Hickman, D.; Riley, T.; Heather, J.; Smith, M.: Real time implementation of image alignment and fusion on a police helicopter, *Proceedings of the Defense and Security Symposium 2006*, Orlando, FL, 2006.
- [7] Li, J.; Nikolov, S. G.; Benton, C. P.; Scott-Samuel, N. E.: Motion-Based Video Fusion Using Optical Flow Information, *9th International Conference on Information Fusion*, 2006.
- [8] Davis, J. W.; Sharma, V.: Background-subtraction using contour-based fusion of thermal and visible imagery, *Computer Vision and Image Understanding* 106 (2007) 162–182
- [9] Foresti, G.L.: A real time system for video surveillance of unattended outdoor environments, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, pp. 697–704, October 1998.
- [10] Lipton, A.; et al.: Moving target classification and tracking from real-time video, *IEEE Workshop on Application of Computer Vision*, pp. 19–21, 1998.
- [11] Gupte, S.; et al.: Detection and Classification of Vehicles, *IEEE Trans. on ITS*, Vol.3, No.1, 2002.
- [12] Collins, R.; et al.: A system for video surveillance and monitoring: VSAM final report, *Robotics Institute, CMU-RI-TR-00-12*, 2000.
- [13] Rosin, P. L.; Ellis, T.: Image difference threshold strategies and shadow removal, *British Machine Vision Conference*, pp. 347–356, 1995.
- [14] Wang, H.; Suter, D.: A Novel Robust Statistical Method for Background Initialization and Visual Surveillance, *7th Asian Conference on Computer Vision*, 2006.
- [15] Snidaro, L.; Foresti, G. L.: Real-time thresholding with Euler numbers, *Pattern Recognition Letters* 24 (2003) 1533–1544
- [16] Jabri, S.; et al.: Detection and Location of People in Video Images Using Adaptive Fusion of Colour and Edge Information, *Proc. of the International Conference on Pattern Recognition – ICPR*, 2000.
- [17] Otsu, N.: A Threshold Selection Method from Gray-Level Histograms, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. SMC-9, No. 1, January 1979.
- [18] Kapur, J. N.; et al.: 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.

[19] Tsai, W. H.: Moment-Preserving Thresholding: A New Approach, Computer Vision, Graphics, and Image Processing 29, pp. 377–393, 1985.

[20] Kittler, J.; Illingworth, J.: Minimum Error Thresholding, Pattern Recognition, Vol. 19, No. 1, pp. 41–47, 1986.

[21] Rosin, P.L.; Ioannidis, E.: Evaluation of global image thresholding for change detection, Pattern Recognition Letters 24, pp. 2345–2356, 2003.

[22] Snidaro, L.; et al.: Sensor Fusion for Video Surveillance, 7th International Conference on Information Fusion, 2004.

[23] <http://www.cse.ohio-state.edu/otcbvs-bench/>, 15. 09. 2007.

[24] Munno, C. J.; et al.: Automatic Video Image Moving Target Detection for Wide Area Surveillance, IEEE International Carnahan Conference on Security Technology, 1993.

[25] Simić, S.: Analiza kontinualnih frekvencijskih moduliranih radarskih signala primenom vremenski – frekvencijskih transformacija i korelacionih tehnika, Vojnotehnički glasnik br. 4/2007, str. 400–413, Beograd, ISSN 0042–8469.

MULTISENSOR BACKGROUND EXTRACTION AND UPDATING FOR MOVING TARGET DETECTION

Summary:

This paper proposes a new method for multisensor background extraction and updating aimed at surveillance and target detection applications. The background scene extraction is based on robust multi-sensor change detection of moving objects in the scene. An iterative mechanism updates the background estimate using this information thereby ignoring transient objects but allowing for slow changes in scene illumination. This new multisensor confidence framework combines change information from all available sensors to greatly improve the reliability of true scene change detection, i.e. the reliability of the background image estimate. The proposed method is demonstrated on several sets of registered multisensor images.

Introduction

Video surveillance is one of the most common applications of video processing research. To achieve reliable surveillance performance, a simultaneous and consistent solution for the detection and tracking of moving objects is required. In real world, due to various environmental conditions, a solution based on a single imaging modality is not reliable in all cases. The use of multiple sensors is required in order to handle different environmental conditions.

In such systems as in many other applications a critical issue in this process is the initialization of a background image. As the availa-

ble data for background image construction is rarely perfect, i.e. without moving objects and noise, robust methods of extraction that can deal with these problems are required.

Change detection in difference images and through background subtraction

In order to detect movement and changes in the observed scene one may well start by looking at the difference images obtained by subtracting some estimate of the scene background from the currently observed scene (frame). Detection is then easily obtained by thresholding the difference image into two classes: unchanged background and moving objects.

A hybrid change detection algorithm

A hybrid method for change detection is the combination of the two techniques presented above. The main weakness of the background subtraction method is exposed when a hitherto stationary object begins to move. Even if such objects are easily detected when they leave their place they leave "holes" where the newly discovered background differs from the current background estimate. Until the background model adjusts to these "holes" the detection algorithm generates false moving objects at that location. Frame differencing is not sensitive to this effect but has the already highlighted weakness of not being able to detect the entire object.

Background initialization and update

As our chosen detection method requires a good background estimate, the first step in the process is to initialise the background estimate, in a robust manner. In many surveillance and detection applications it is not possible to choose a single frame as the background estimate. This is most pertinent in scenes containing many moving objects such as roads, paths, etc. So our background needs to be estimate from frames containing moving objects.

Multisensor background extraction and update

In a multisensor surveillance system several images of the same scene are available, albeit seen through different sensor modalities. By combining the information from such multiple sources we can eliminate the weaknesses of each individual source and provide robust detection in all conditions.

The moving object contrast largely determines the success of change detection in the images. Consequently, the idea proposed in this paper is to combine the output of all sensors to obtain a more accurate and more robust estimate of moving and stationary pixels. A single multisensor movement (change) estimate is generated by fusing

change information obtained from individual sensors. Such robust change estimates could then be used to extract and update the background estimates for each sensor in a more robust manner making the subsequent detection more reliable.

Background Extraction Results

The proposed multisensor background extraction approach was tested on a number of registered sets of multisensor sequences, collected by the group and data widely available in the community, e.g. OTCBVS (Object Tracking and Classification in and Beyond the Visible Spectrum) Benchmark Dataset Collection.

Conclusion

This paper presents a robust multisensor background extraction and update scheme intended for use in target detection applications. The method uses a number of well known concepts in detection to provide a basic classification of the scene into moving and stationary objects and an iterative procedure to update the background model and deal with transient (slow) changes in illumination conditions. A new fusion-based confidence framework uses the information from multiple registered views of the scene obtained by different sensors in order to update the background model only in areas where there are no moving objects. It was shown on a number of representative multisensor sequences that the proposed method obtains a more complete and robust representation of the background than any estimate based on the individual sensors. Finally, a relatively simple computational architecture of the proposed system makes it ideal for demanding real time multisensor detection applications

Key words: multisensor detection, background extraction, background initialization and updating.

Datum prijema članka: 13. 10. 2009.

Datum dostavljanja ispravki rukopisa: 22. 09. 2009.

Datum konačnog prihvatanja članka za objavljivanje: 24. 09. 2009.