

Sledilnik s korelacijskim filtrom

3. naloga pri predmetu Napredne metode računalniškega vida

Ivan Antešić¹

Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Večna pot 113,
Slovenija

`ia6382@student.uni-lj.si`

Keywords: Računalniški vid · Korelacija · Sledilnik · Mosse · Filter.

1 Uvod

V okviru tretje naloge smo razvili agloritem za sledenje na podlagi korelacijskega filtra [1].

Klasifikator je implementiran z korelacijskim filtrom H , ki ima z regijo slike velik korelacijski odziv na območju tarče in majhen odziv na območju ozadja. Med dolevanjem algoritma poskušamo naučiti filter, da bo čimbolj deskriminanten - želimo, da je velik odziv prisoten samo na majhnem območju, drugje pa je enak 0. Zato smo za odziv izbrali, kar gaussovo funkcijo. Ker se klasifikator med delovanjem prilagaja in uči pravimo, da je "online" algoritem.

Operacijo korelacije smo implemntirali z fourierjevo transformacijo, saj je v frekvenčnem prostoru hadamardov produkt dveh matrik (množenje po elementih) enak korelaciji in zato hitrejši ($O(n^2 \log n)$ v primerjavi z običajno korelacijo $O(n^4)$). Ker je korelacija cirkularna v diskretni fourierjevi transformaciji, je dobro da regijo slike utežimo z Hanningovim oknom, ker zmanjša vpliv robnih pikslov.

V vsakem okvirju (ang. frame) se premaknemo v lokacijo maksimalnega korelacijskega odziva.

V nadaljevanju analiziramo rezultate algoritem in njegove nadgradnje na testni sekvenci slik vot2014 (<http://box.vicos.si/vot/vot2014.zip>) z uporabo ogrodja Tracking Evaluation Toolkit (<https://github.com/alanlukezic/tracking-toolkit-lite>).

2 Parametri

Na delovanje algoritma vpliva več parametrov.

2.1 Alfa

V primerjavi z npr. barvnim vizuelnim modelom, je naš klasifikator občutljiv na transformacije. Ker se tarča skozi sekvenco lahko obrača in spreminja svoj

izgled je potrebno postopoma posodabljati filter H , da lahko z korelacijskim odziv uspešno sledimo tarči. V naslednji formuli je razvidno, da je hitrost posodabljanja odvisna od parametra alfa:

$$H_{novi} = (1 - alfa) * H_{stari} + alfa * H_{trenutni}$$

Pri velikem alfa (npr. 0.3), sledilnik bolj gladko spremlja tarčo in je bolj uspešen pri dinamičnih sekvencah. Odpove pa, ko se tarča premika počasi, saj se bo filter počasi prilagodil na ozadje (vidno na sliki 1) ali pa, ko nek objekt počasi preide čez našo tarčo in se filter prilagodi sledenju novega objekta (slika 2).

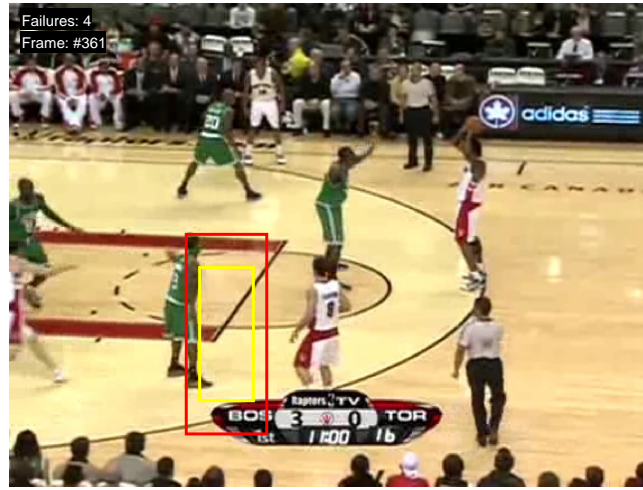


Fig. 1. Slaba lokalizacija tarče na sekvenci "basketball" - sledilnik se osredotoča na ozadje. Alfa = 0,3.

Če je alfa premajhen ali pa sploh ne posodabljam pri vseh sekvencah, kjer tarča spreminja izgled odpove, saj se ne prilagodi (dovolj hitro) in preneha slediti tarči.

Potrebno je določiti kompromisno vrednost, kjer se filter dovolj počasi prilagaja, da ne zamenjamo tarčo in lahko sledimo majhnim spremembam. Na slikah 3, 4 in 5 lahko vidimo, kako zelo diskriminanten je filter H , ko ima tarča dobre značilne regije (v tem primeru črni nos na belem ozadju) in se počasi posodablja.

Tabela 1 in slika 6 primerjata delovanje sledilnika pri različnih vrednostih parametra alfa.

2.2 Sigma

Kot je omenjeno v uvodu smo za korelacijski odziv izbrali Gaussovo funkcijo. Širina vrha funkcije je odvisna od parametra sigma, kar pomeni, da sigma vpliva na občutljivost filtra H .



Fig. 2. Slaba lokalizacija tarče na sekvenci "basketball" - sledilnik je začel slediti belemu igralcu namesto zelenemu. Alfa = 0,3.

Manjši sigma (npr. $\sigma=1$ na sliki 7) je bolj občutljiv na visoke frekvence in je posledično bolj natančno deskriminira značilke v regiji. Filter na sliki 3 izgleda kot, da ima večjo ločljivost kot filter na sliki 8.

Prevelik sigma (npr. $\sigma=5$ na sliki 9) povzroči nenatančen filter in slabše sledenje. Še vedno pa mora biti dovolj velik, da odziv po nekaj okvirjih (angl. frame), ne vsebuje preveč šuma in je zadostno zglajen. Tako lažje najdemo maksimalen korelacijski odziv.

Tabela 2 in slika 10 primerjata delovanje sledilnika pri različnih vrednostih parametra sigma.

2.3 Velikost regije

Pri računanju filtra H in korelacijskega odziva lahko uporabljamo večjo regijo slike, kot je velika tarča. To je koristno pri sekvencah, kjer se tarča med dvema zaporednima okvirjema premakne veliko in lahko zaradi tega izgubimo značilke, ki smo jim sledili, saj so se premaknile izven mej regije tarče. Če zavzamemo večjo regijo lažje sledimo takšnim premikom (slika 11). Delovanje sledilnika se torej izboljša pri malo večji regiji (npr. dvakratna velikost tarče). Če je regija prevelika vsebuje večinoma le robne piksele, ker pade ven iz slike.

Tabela 3 in slika 12 primerjata delovanje sledilnika pri različnih vrednostih parametra sigma. Čeprav se je velikost trikratna izkazala za najboljšo je bilo delovanje algoritma prepočasno za tako majhno izboljšavo. Najboljše se je izkazala dvakratna velikost.

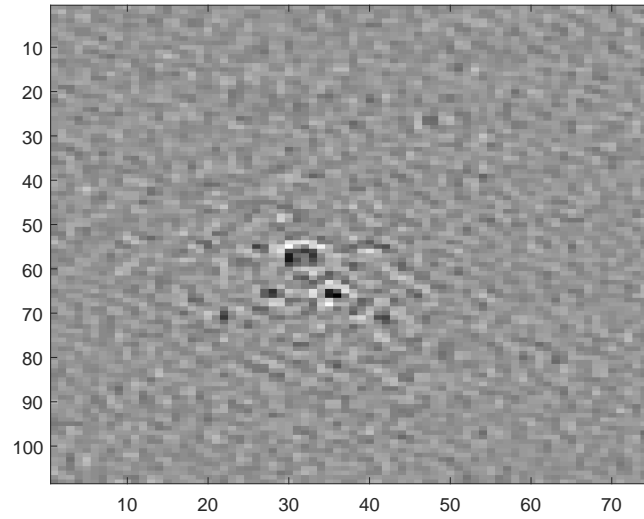


Fig. 3. Filter H za sliko 5, pri parametru $\sigma = 1$, v času prvega okvirja.

3 Prilagajanje velikosti

Občutljivost sledilnika na obračanja tarče smo zmanjšali z posodobitvijo filtra, še vedno pa je algoritem občutljiv na skaliranje tarče. Zato smo implementirali preprosto nadgradnjo, kjer vsak okvir izberemo regije treh različnih velikosti (+10%, -10% in isto velikost). normaliziramo regije na originalno velikost in izberemo tisto, pri kateri pride do največjega odziva.

Da ne bi bili preobčutljivi na skaliranje tarče nato končno velikost utežimo z prejšno po naslednji formuli, kjer je γ faktor hitrosti skaliranja (v našem primeru 0,3):

$$velikost_{nova} = (1 - \gamma) * velikost_{stara} + \gamma * velikost_{trenutna}$$

Če uporabljamo to nadgradnjo je pomembno, da je velikost regije enaka velikosti tarče, da lahko sledilnik ustrezno prilagaja velikost.

Pomembno je tudi znižati histrost posodabljanja filtra α , ker se pri statičnih tarčah regija postopoma povečuje, ker se filter začne prilagajati ozadju, dokler se regija ne poveča čez večino celotne slike (vidno na sliki 13). Nadgradnja deluje boljše kot sonovni algoritem pri skevencah kot je "car", kjer se tarča ne premika prehitro, se ne obrača in se postopoma povečuje. Slabše deluje pri bolj dinamičnih sekvencah.

Tabela 4 in slika 14 primerjata delovanje osnovnega sledilnika in sledilnika z nadgradnjo. Za parametre osnovnega sledilnika smo izbrali tiste, ki so se v prejšnjih testih izkazali za najboljše.

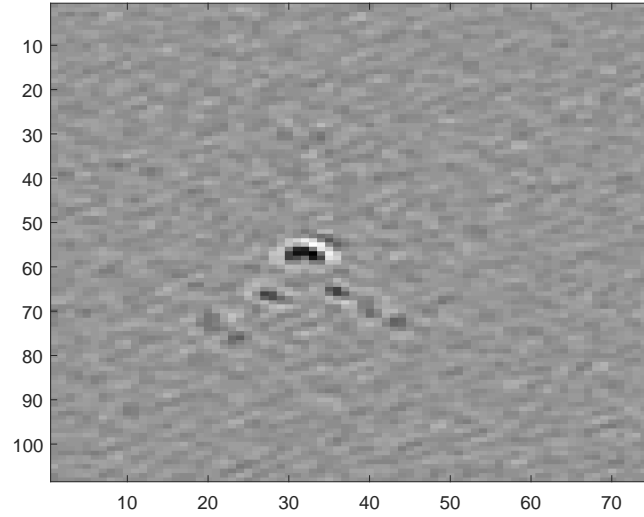


Fig. 4. Filter H za sliko 5, pri parametru $\sigma = 1$, v času petega okvirja. V primerjavi z sliko 3 je opazna zgladitev in večji fokus na značilke, ki predstavljajo nos in ušesa polarnega medveda.

4 Hitrost sledenja

Na hitrost sledenja vpliva predvsem velikost regije in sekvenca na kateri testiramo algoritem. Pri trikaratni velikosti tarče, na sekvenci "drunk" je osnovni sledilnik deloval z hitrostjo povprečno 25 fps (sličic na sekundo), medtem ko je pri dvakratni velikosti imel hitrost povprečno 68 fps.

Izbira sekvenca tudi vpliva na velikost regije tarče in ima morda zaradi tega velik vpliv na hitrost. V sekvenci "drunk" sledimo večjemu objektu - avtomobilu, kot v sekvenci "car" kjer začnemo slediti avtomobilu, ko je le ta daleč od kamere in zato manjši. Posledično je hitrost za "car" 324 fps.

Nadgradnja prilagajanja velikosti tudi upočasni algoritem, saj se večkrat spreminja velikosti slik in računa korelacijski odziv. Za sekvenco "car" ima hitrost 93 fps.

V tabeli 5 so prikazane hitrosti delovanja osnovnega algoritma na različnih sekvencah.

5 Zaključek

Po testiranju preproste implementacije sledilnika s korelacijskim filtrom smo ugotovili, da z ustrezno nastavitvijo glavnih parametrov lahko sledi zadovoljivo v realnem času (vsaj okoli 100 fps). Še vedno pa deluje slabše kot algoritmi, ki uporabljajo druge vizuelne modele (npr. mean shift), saj sledi le na podlagi frekvenc v sliki.



Fig. 5. Lokalizacija tarče na sekvenci "polarbear". $\text{Sigma} = 1$.

Nadgradnja prilagajanja velikosti regije, deluje boljše za nekatere specifične primere, povprečno pa zaradi svojih pomanjkljivosti deluje slabše kot osnovna implementacija.

Sledilnik bi lahko izboljšali, če bi velikost izbrane regije pomanjšali in tako pohitrili delovanje algoritma.

Literatura

1. Bolme, D.S., Beveridge, J.R., Draper, B.A., Lui, Y.M.: Visual object tracking using adaptive correlation filters. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 2544–2550 (June 2010). <https://doi.org/10.1109/CVPR.2010.5539960>

Sequence	ALFA3		ALFA0		ALFA0.05	
	Overlap	Failures	Overlap	Failures	Overlap	Failures
ball	0.42	8	0.47	8	0.33	3
basketball	0.46	8	0.62	16	0.60	6
bicycle	0.41	0	0.51	5	0.44	1
bolt	0.48	16	0.50	13	0.49	10
car	0.51	1	0.44	1	0.48	1
david	0.72	0	0.64	7	0.68	1
diving	0.38	4	0.39	4	0.38	4
drunk	0.53	2	0.39	3	0.43	3
fernando	0.22	2	0.33	6	0.50	5
fish1	0.35	12	0.47	18	0.45	18
fish2	0.32	10	0.40	12	0.39	11
gymnastics	0.62	5	0.58	3	0.53	2
hand1	0.55	6	0.41	10	0.48	10
hand2	0.46	14	0.48	16	0.46	14
jogging	0.76	2	0.51	3	0.74	2
motocross	0.50	2	0.48	2	0.41	2
polarbear	0.49	0	0.46	0	0.46	0
skating	0.58	2	0.60	8	0.48	5
sphere	0.53	1	0.65	3	0.67	3
sunshade	0.74	1	0.67	5	0.63	0
surfing	0.59	1	0.72	0	0.70	0
torus	0.51	5	0.61	9	0.43	6
trellis	0.49	3	0.64	13	0.71	3
tunnel	0.28	0	0.62	6	0.51	2
woman	0.50	3	0.67	10	0.67	3
Average	0.50	4.32	0.53	7.24	0.52	4.60

Table 1. Rezultati testiranja sledilnikov z različnimi vrednostmi parametra alfa (3, 0 in 0.05) na testnih sekvencah vot2014.

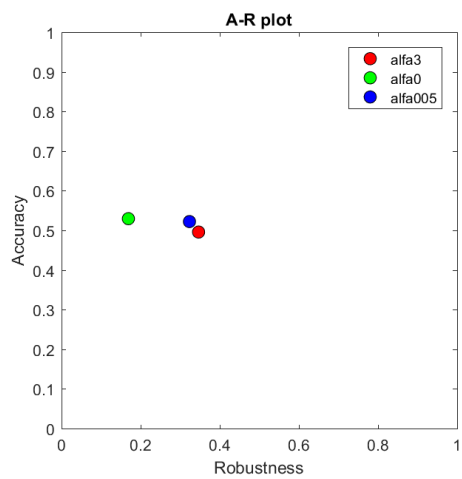


Fig. 6. A-R graf sledilnikov z različnimi vrednostmi parametra α .

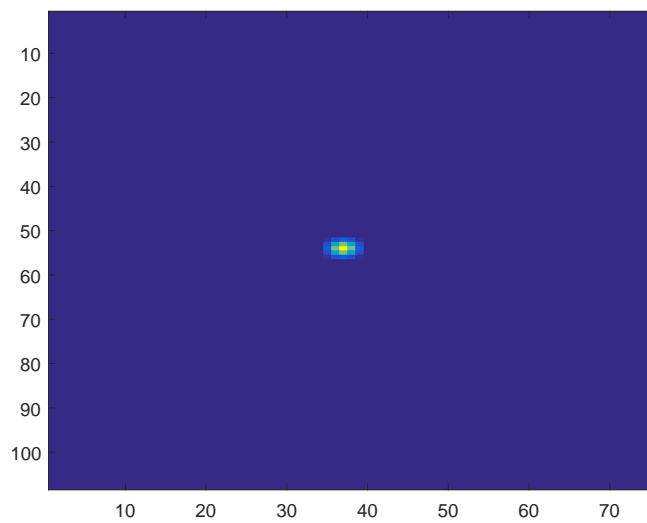


Fig. 7. Gaussova funkcija pri vrednosti parametra $\sigma = 1$.

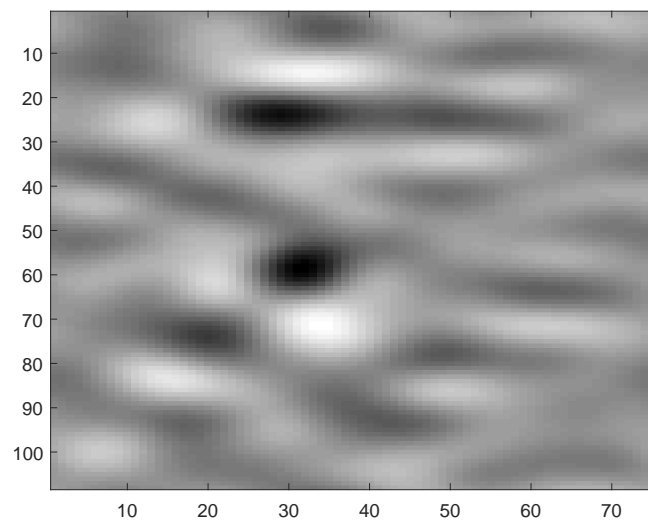


Fig. 8. Filter H za sliko 5, pri parametru $\sigma = 1$, v času petega okvirja.

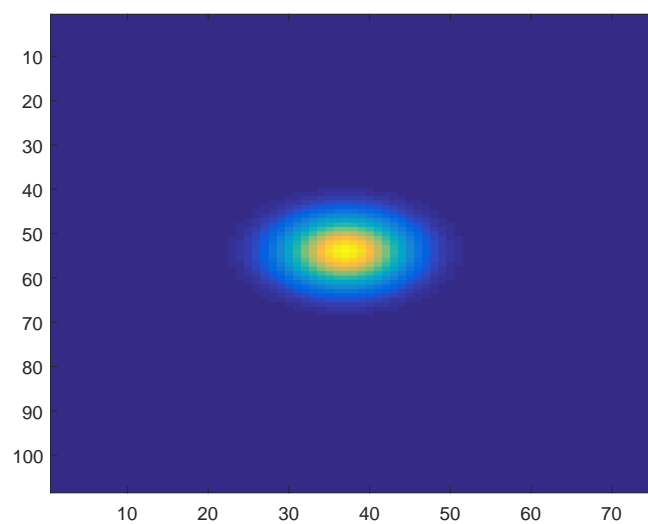


Fig. 9. Gaussova funkcija pri vrednosti parametra $\sigma = 5$.

Sequence	SIGMA2		SIGMA1		SIGMA0.1	
	Overlap	Failures	Overlap	Failures	Overlap	Failures
ball	0.48	3	0.33	3	0.48	13
basketball	0.54	8	0.60	6	0.38	28
bicycle	0.46	1	0.44	1	0.52	4
bolt	0.39	5	0.49	10	0.47	15
car	0.55	1	0.48	1	0.52	1
david	0.62	2	0.68	1	0.65	18
diving	0.39	4	0.38	4	0.38	4
drunk	0.31	0	0.43	3	0.39	4
fernando	0.45	5	0.50	5	0.48	8
fish1	0.45	14	0.45	18	0.43	26
fish2	0.34	8	0.39	11	0.34	11
gymnastics	0.58	3	0.53	2	0.62	4
hand1	0.50	6	0.48	10	0.34	13
hand2	0.51	15	0.46	14	0.36	14
jogging	0.74	2	0.74	2	0.54	10
motocross	0.40	2	0.41	2	0.50	2
polarbear	0.46	0	0.46	0	0.46	0
skating	0.55	4	0.48	5	0.54	9
sphere	0.46	1	0.67	3	0.71	5
sunshade	0.75	3	0.63	0	0.69	7
surfing	0.71	0	0.70	0	0.67	0
torus	0.44	9	0.43	6	0.35	9
trellis	0.72	11	0.71	3	0.64	20
tunnel	0.33	0	0.51	2	0.65	6
woman	0.74	5	0.67	3	0.70	12
Average	0.51	4.48	0.52	4.60	0.51	9.72

Table 2. Rezultati testiranja sledilnikov z različnimi vrednostmi parametra sigma (2, 1 in 0.1) na testnih sekvencah vot2014.

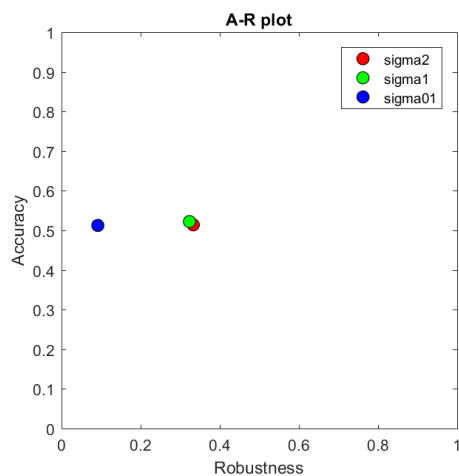


Fig. 10. A-R graf sledilnikov z različnimi vrednostmi parametra sigma.

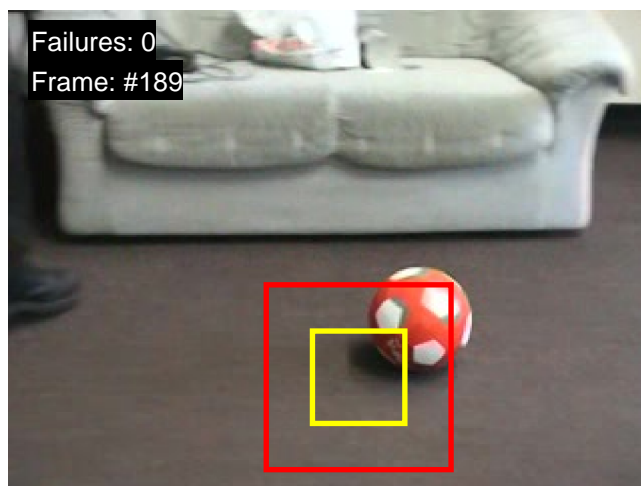


Fig. 11. Lokalizacija tarče na sekvenci "ball" - sledilnik se osredotoča na ozadje. Velikost regije (rdeč kvadrat) je dvakrat večja kot tarča (rumen kvadrat). Zaradi večje velikosti uspešno sledi tarči tudi pri večjem premiku.

Sequence	VELIKOST3		VELIKOST2		VELIKOST1	
	Overlap	Failures	Overlap	Failures	Overlap	Failures
ball	0.35	0	0.35	3	0.43	10
basketball	0.43	16	0.57	11	0.37	9
bicycle	0.44	0	0.48	2	0.58	3
bolt	0.55	9	0.54	10	0.47	11
car	0.44	0	0.44	0	0.45	1
david	0.60	1	0.54	0	0.55	4
diving	0.38	4	0.38	4	0.38	4
drunk	0.45	2	0.40	3	0.38	3
fernando	0.48	5	0.44	4	0.50	6
fish1	0.31	14	0.36	13	0.43	16
fish2	0.31	7	0.34	10	0.33	8
gymnastics	0.62	7	0.51	3	0.59	4
hand1	0.28	9	0.36	11	0.46	9
hand2	0.41	15	0.43	15	0.36	12
jogging	0.74	1	0.70	1	0.60	5
motocross	0.41	2	0.42	2	0.45	2
polarbear	0.47	0	0.47	0	0.45	0
skating	0.51	5	0.53	7	0.44	4
sphere	0.68	0	0.57	0	0.20	5
sunshade	0.75	0	0.78	0	0.74	2
surfing	0.72	0	0.72	0	0.71	1
torus	0.35	9	0.37	6	0.49	7
trellis	0.67	6	0.73	11	0.75	4
tunnel	0.66	1	0.29	0	0.27	0
woman	0.70	1	0.66	2	0.59	11
Average	0.51	4.56	0.49	4.72	0.48	5.64

Table 3. Rezultati testiranja sledilnikov z različnimi velikostmi regije na testnih sekvencah vot2014.

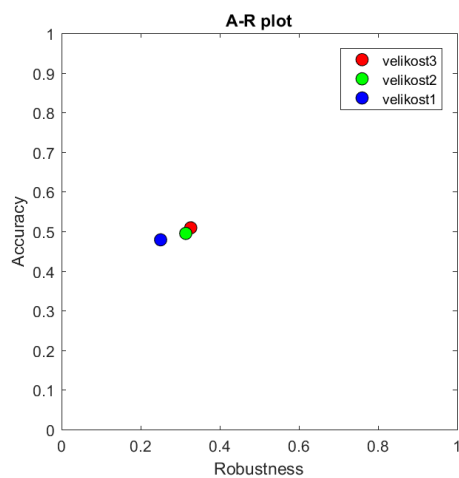


Fig. 12. A-R graf sledilnikov z različnimi velikostmi regije.



Fig. 13. Lokalizacija tarče na skevenci "drunk". Uporabljeno je prilagajanje velikosti, $\alpha = 0.05$. Regija se je prekomerno razširila.

Sequence	BASIC		SCALE ADAPTATION	
	Overlap	Failures	Overlap	Failures
ball	0.35	3	0.42	1
basketball	0.57	11	0.43	22
bicycle	0.48	2	0.60	6
bolt	0.54	10	0.40	11
car	0.44	0	0.56	0
david	0.54	0	0.38	9
diving	0.38	4	0.40	5
drunk	0.40	3	0.26	4
fernando	0.44	4	0.42	7
fish1	0.36	13	0.41	21
fish2	0.34	10	0.46	15
gymnastics	0.51	3	0.42	6
hand1	0.36	11	0.42	10
hand2	0.43	15	0.40	14
jogging	0.70	1	0.37	7
motocross	0.42	2	0.51	4
polarbear	0.47	0	0.32	3
skating	0.53	7	0.50	6
sphere	0.57	0	0.37	5
sunshade	0.78	0	0.64	6
surfing	0.72	0	0.33	4
torus	0.37	6	0.41	9
trellis	0.73	11	0.60	10
tunnel	0.29	0	0.50	19
woman	0.66	2	0.49	13
Average	0.49	4.72	0.44	8.68

Table 4. Rezultati testiranja osnovnega sledilnika (mosse) in sledilnika (mosseScale) z nadgradnjo, na testnih sekvencah vot2014.

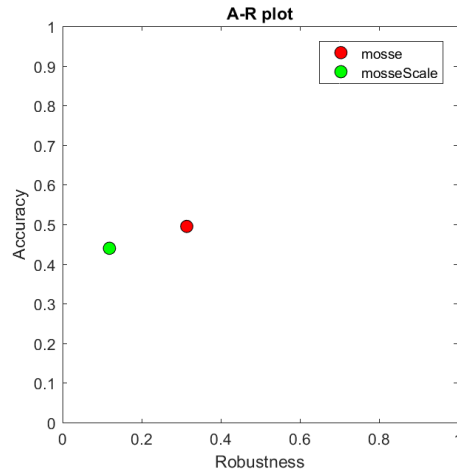


Fig. 14. A-R graf osnovnega sledilnika (mosse) in sledilnika z nadgradnjo (mosseScale).

Sequence	FPS
ball	210
basketball	153
bicycle	459
bolt	192
car	324
david	114
diving	140
drunk	68
fernando	60
fish1	300
fish2	173
gymnastics	202
hand1	279
hand2	330
jogging	242
motocross	93
polarbear	221
skating	162
sphere	151
sunshade	297
surfing	510
torus	334
trellis	294
tunnel	206
woman	275

Table 5. Hitrost delovanja (v sličicah na sekundo) sledilnika na sekvencah vot2014.