

Implementacija algoritma Condensation v domeni zaprtega sveta

Matej Kristan, Janez Perš, Matej Perše, Stanislav Kovačič
Fakulteta za elektrotehniko
Univerza v Ljubljani
Tržaška 25, 1000 Ljubljana, Slovenija
{matej.kristan}, {janez.pers}, {matej.perse}, {stanislav.kovacic}@fe.uni-lj.si

Implementation of CONDENSATION algorithm in a closed world domain

People tracking in general is a challenging task and over the last two decades various computer vision algorithms dealing with this problem were proposed. Given a highly unpredictable nature of human motion stochastic based approaches such as CONDENSATION introduced by M. Issard and A. Blake in 1998, gained a lot of popularity among researchers in this field. In this paper we present an implementation of CONDENSATION algorithm for tracking people in sports. Since sport games usually take place in semi-controlled environments a closed world assumption, introduced by S.S. Intille and A. Bobick in 1995 has been adopted. We present an architecture of such condensation based tracking algorithm within a closed world domain and show some results.

1 Uvod

Računalniško podprta analiza gibanja postaja vedno bolj popularen pripomoček pri treningih v športnih igrah. Za uspešno analizo je potrebno zajeti podatke o položajih igralcev v vsakem trenutku, kar lahko storimo z ročnim označevanjem ali avtomatskim sledenjem.

V članku je predstavljen princip sledenja objektov v športnih igrah, ki združuje stohastični pristop k sledenju z uporabo algoritma CONDENSATION [2] in nekatere postavke domene zaprtega sveta [1].

2 Uporaba domene zaprtega sveta

Domena zaprtega sveta je pojem, ki sta ga v računalniško podprtem sledenju leta 1995 vpeljala S.S. Intille in A.F. Bobick [1] ter zajema tisto informacijo, ki je v nekem času in kraju zadostna za razlago opazovane scene. Pri poljubni športni igri so to pravila igre, število ter izgled opazovanih igralcev, oznake na igrišču, omejitve in barva igrišča, itd..

Domeno zaprtega sveta v športnih igrah definiramo s sledečimi postavkami:

1. Igrišče športne igre je omejeno, njegova slika je znana, osvetljava pa se s časom le malo spreminja, oziroma je konstantna.
2. Teksture igralcev so znane v naprej, in se z gibanjem preko igrišča spreminjajo.
3. Oblike igralcev se med igro močno spreminjajo.

3 CONDENSATION

CONDENSATION je stohastični pristop k sledenju, ki sta ga v računalniškem vidu leta 1998 predstavila M. Issard in A. Blake [2]. Pristop bazira na faktorskem vzorčenju, kjer v vsakem časovnem koraku opišemo zvezno distribucijo gostote verjetnosti stanj objekta z diskretnim naborom stanj in njihovimi utežmi. Princip je zelo privlačen tako zaradi svoje preprostosti, kakor zaradi dejstva, da je lahko v splošnem distribucija stanj multimodalna oziroma vsaj ne-Gaussova. Stanje objekta $s_t^{(n)}$ je definirano kot bistvene lastnosti objekta, utež $\pi_t^{(n)}$ pa je verjetje s katerim trdimo, da je objekt zavzel to stanje. Diskretni nabor N -tih stanj ob času t označimo z $S_t = \{(s_t^{(n)}, \pi_t^{(n)}); n = 1 \dots N\}$, kjer velja $\sum_n \pi_t^{(n)} = 1$. Nabor stanj S_{t+1} ob času $t+1$, dobimo s širjanjem nabora S_t preko dinamičnega modela sistema. Pričakovano stanje objekta v časovnem trenutku t , $s_t^{(A)}$, se določi kot povprečna vrednost uteženega nabora stanj S_t .

$$s_t^{(A)} = E(S_t) = \frac{1}{\sum_{n=1}^N \pi_t^{(n)}} \sum_{n=1}^N \pi_t^{(n)} s_t^{(n)} \quad (1)$$

4 Predstavitev objekta v zaprtem svetu

Stanje objekta določajo njegove lastnosti, ki vsebujejo informacije za opis objektovega zaprtega sveta, ključne lastnosti za opis igralca pa so izgled in oblika.

4.1 Oblika

Postavka 3 v razdelku 2 pravi, da se igralci močno spreminjajo v obliki. Obliko objekta i zapišemo z elipso $O^{(i)}$, ki je določena s centrom $(x^{(i)}, y^{(i)})$ in velikostjo $(a^{(i)}, b^{(i)})$. Elipso smo izbrali zaradi dejstva, da enaka elipsa pojasni položaj igralca pri različnih oblikah, kar pripomore k stabilosti sistema.

$$O^{(i)} = \{x^{(i)}, y^{(i)}, a^{(i)}, b^{(i)}\} \quad (2)$$

4.2 Tekstura

M.J. Swain in D.H. Ballard v [3] predlagata zapis teksture z barvnim histogramom. Ta princip se je v sledenju [4,5] že izkazal za primerega. Teksturo objekta znotraj elipse $O^{(i)}$ opišemo z uteženo vzorčenim barvnim histogramom $H^{(i)}$. Uteženo vzorčenje pomeni, da tisti slikovni elementi, ki so bližje centru elipse, vplivajo na histogram z večjo utežjo, kot tisti dlje od centra. Utežen pristop je smiseln, saj slikovni elementi na robu elipse objektu pripadajo z manjšo verjetnostjo, kot pa tisti na sredini (npr. pri zakritju z drugim objektom).

4.3 Zapis stanja igralca

Stanje igralca opišemo z njegovo obliko, teksturo ter položajem na igrišču. Za zapis teksture smo uporabili RGB histogram, ki vsebuje po 8 celic na barvni kanal. Vsebino histograma po vzorčenju normiramo tako, da je vsota njegove vsebine enaka ena. Obliko ter položaj igralca smo zapisali z elipso (En. 2). Ob času t formalno zapišemo n -to stanje objekta kot par elipse ter histograma.

$$s_t^{(n)} = \{O_t^{(n)}, H_t^{(n)}\} \quad (3)$$

4.4 Izločanje ozadja

Upoštevajoč prvo postavko v razdelku 2, lahko zgradimo masko, s katero izločimo slikovne elemente, ki bolj verjetno pripadajo ozadju, kot pa sledenemu objektu.

Masko dobimo z upragovljanjem evklidskih razlik med trenutno sliko ter sliko ozadja, slednja pa se izračuna z uporabo mediane vzdolž časovne osi vsakega slikovnega elementa.

5 Uteževanje stanj

Stanju $s_t^{(i)}$ dodelimo neko utež $\pi_t^{(i)}$, ki je proporcionalna verjetju, da je sledeni objekt zavzel izbrano stanje. To storimo s primerjanjem histogramov izbranega in pričakovanega stanja. Histogram pričakovanega stanja $H^{(M)}$ je histogram sledenega objekta, ki ga vzorčimo pred začetkom sledenja, izbranemu stanju pa histogram $H_t^{(i)}$ uteženo vzorčimo iz maskirane slike in ignoriramo ozadju pripadajoče slikovne elemente.

$$b = \sum_{r,g,b} (H_t^{(M)}(r,g,b) \cdot H_t^{(i)}(r,g,b))^{\frac{1}{2}} \quad (4)$$

$$d = \sqrt{1 - b^2} \quad (5)$$

$$occ = \frac{e^{vis} - 1}{e^1 - 1} \quad (6)$$

$$N(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (7)$$

$$\pi_t^{(i)} = \pi^t(s_t^{(i)}) = N(d, \sigma_{sim}) \cdot occ \quad (8)$$

V En. (6) z vis označimo delež slikovnih elementov znotraj elipse, ki ne pripadajo ozadju, σ_{sim} je eksperimentalno določen parameter, oceno verjetja $\pi_t^{(i)}$ pa izračunamo po En. (8).

6 Adaptacija modela teksture objekta

Ker je igrišče v glavnem neenakomerno osvetljeno in se igralci različno obračajo proti kameri, se njihove teksture s časom spreminjajo (razdelek 2, postavka 2). Zaradi slednjega je potrebno teksture tekom sledenja adaptirati. Histogram modela $H_t^{(M)}$ se adaptira z nekim histogramom $H_t^{(i)}$ z utežjo p po En. (10). Utež adaptacije p se izračunava po En. (9), kjer vis predstavlja delež vidnih slikovnih elementov znotraj elipse ki pripada histogramu $H_t^{(i)}$, Ω_{max} je največja možna vrednost adaptacije, vsota na desni strani enačbe pa predstavlja vrednost preseka med histogramoma $H_t^{(M)}$ in $H_t^{(i)}$.

$$p_t^{(i)} = vis \cdot \Omega_{max} \cdot \sum_{j,k,l} \min(H_{t-1}^{(M)}(j,k,l), H_t^{(i)}(j,k,l)) \quad (9)$$

$$H_t^{(M)} = H_{t-1}^{(M)} \cdot p_t^{(i)} + H_t^{(i)} \cdot (1 - p_t^{(i)}) \quad (10)$$

7 Določanje maske

Masko potrebujemo takrat, ko se sledeni objekt nahaja na njemu podobnem delu ozadja, zato jo generiramo le, ko podobnost med objektom in ozadjem preseže nek prag podobnosti Φ_r . Podobnost je definirana kot koeficient Bhattacharyya (En. 9) med histogramoma ozadja na mestu kjer se nahaja objekt, ter histogramom modela.

Če vrednost podobnosti preseže neko predpisano mejno vrednost Φ_{tr} , se maska generira tako, da želi zadostiti pogoju, da se znotraj elipse sledenega objekta nahaja med 40 do 80 odstotkov slikovnih elementov, ki ne pripadajo ozadju.

7 Dinamični model sistema

Ob času t se stanje $s_t^{(i)}$ izračuna s premikom njegovega predhodnika $s_{t-1}^{(i)}$ preko dinamičnega modela.

$$cl(x, x_{min}, x_{max}) = \begin{cases} x_{min} & ; \quad x < x_{min} \\ x_{max} & ; \quad x > x_{max} \\ x & ; \quad sicer \end{cases} \quad (11)$$

$$\Theta = \{a_{\min}, a_{\max}, b_{\min}, b_{\max}\}$$

$$Clip(O_{t-1}^{(i)}, \Theta) = \{x_{t-1}^{(i)}, y_{t-1}^{(i)}, cl(a_{t-1}^{(i)}, a_{\min}, a_{\max}), cl(b_{t-1}^{(i)}, b_{\min}, b_{\max})\}$$

$$shift(O_{t-1}^{(i)}) = O_{t-1}^{(i)} + v_t^{(i)} + u_t^{(i)} \quad (12)$$

$$dyn(s_{t-1}^{(i)}, \Theta) = Clip(shift(O_{t-1}^{(i)}), \Theta)$$

Splošni model prikazujejo En. (12), kjer v_t predstavlja deterministični premik, u_t pa stohastično difuzijo. Za boljše stabilnost sistema, se velikost elipse razširjenega stanja omeji po En. (11) z Θ .

7.1 Deterministični premik

Deterministični premik je izračunan kot utežena vsota T-tih predhodnih stanj objekta. Ker je vsako stanje uteženo z utežjo, ki je proporcionalna verjetju v njegovo pravilnost, se napovedani premik v_t izračuna z upoštevanjem teh uteži ter časovega sosledja stanj.

$$\Delta O_i^{(A)} = \{x_i^{(A)} - x_{i-1}^{(A)}, y_i^{(A)} - y_{i-1}^{(A)}, a_i^{(A)} - a_{i-1}^{(A)}, b_i^{(A)} - b_{i-1}^{(A)}\}$$

$$k = \sum_{i=t-T+1}^{t-1} \pi_{i-1}^{(A)} \cdot \pi_i^{(A)} \cdot (i-t+T) \quad (13)$$

$$v_t = \frac{1}{k} \sum_{i=t-T+1}^{t-1} \Delta O_i^{(A)} \cdot \pi_{i-1}^{(A)} \cdot \pi_i^{(A)} \cdot (i-t+T)$$

7.2 Stohastična difuzija

Parameter $u_t^{(i)}$ v En. (12) označuje stohastično difuzijo ter predstavlja naključno spremembo stanja $s_t^{(i)}$ glede na predhodnika $s_{t-1}^{(i)}$, izračun pa prikazujeta naslednji enačbi.

$$u_t^{(i)} = \{\Delta x_t, \Delta y_t, \Delta a_t, \Delta b_t\} \quad (14)$$

$$[\Delta x_t, \Delta y_t, \Delta a_t, \Delta b_t]^T = \begin{bmatrix} \alpha & 0 & \gamma & 0 \\ 0 & \beta & 0 & \delta \end{bmatrix} \cdot \begin{bmatrix} a_{t-1} \\ b_{t-1} \end{bmatrix}$$

V En. (14) se parametra α ter β porazdeljujeta v bivariatni normalni porazdelitvi z enkima standardnima deviacijama $(\alpha, \beta) \sim N(0, \sigma_{din})$, enako pa velja tudi za γ ter δ ; $(\gamma, \delta) \sim N(0, \sigma_{siz})$. Za naključni generator smo uporabili Mersenne twister [6].

8 Zgradba sledilnika in rezultati

Sledilnik opisan v Sliki (1) (imenujemo ga *razširjeni sledilnik*) smo primerjali z izvedbo, ki ne upošteva domene zaprtega sveta (imenujemo jo *naivna izvedba*). V naivni izvedbi se tekstura ne adaptira, ozadje se ne izloča, nabor stanj se ne zgoščuje, deterministični premik pa je ničtega reda.

Z obema sledilnikoma smo posledili trajektorije osmih igralcev rokometu, ki so vsak posebej šestkrat pretekli na igrišču narisani krog premera osmih metrov (Slika 2). Poizkus je bil izveden na šest minutnem

posnetku [7], kjer je bila osvetljava igrišča vzdolž kroga neenakomerna, ter se je časovno spreminjala, za manjšo pristranskost poizkusa, pa so igralci nosili drese različnih barv. Posledene podatke smo kalibrirali po postopku [8]. Za vsakega igralca smo izračunali iz znanega obsega kroga teoretično pretečeno pot L_{teor} , preko kalibriranih rezultatov sledenja pa izmerjeno pretečeno pot L_{izm} . Razliko med izmerjeno in izračunano pretečeno potjo zapišemo po naslednji enačbi, kjer i označuje zaporedno številko igralca.

$$\mathcal{E}^{(i)} = |L_{teor}^{(i)} - L_{izm}^{(i)}| \quad (15)$$

Tabela (1) prikazuje pričakovano število odpovedi na minuto obeh sledilnikov, ter pričakovano napako μ_{err} in standardno deviacijo σ_{err} v odstotkih pretečene poti.

Vhod:
nabor prejšnje iteracije: $S_{t-1} = \{s_{t-1}^{(n)}, \pi_{t-1}^{(n)}, n = 1 \dots N\}$
histogram modela: $H_t^{(M)}$
predikcija stanja za čas t: s_t^P
omejitev velikosti elipse: Θ

Izhod: *nov nabor* S_t , *novo stanje* $s_t^{(A)}$

1. Določi utež napovedanemu stanju (En.8):

$$\pi_t^{(P)} = \pi'(s_t^{(P)})$$

2. Propagiraj S_{t-1} v S_t :

za vsak $n = 1$ do N naredi

(a) izberi vzorec: $\{s_{t-1}^{(i)}, \pi_{t-1}^{(i)}\} \in S_{t-1}$

(b) propagiraj $s_{t-1}^{(i)}$ (En.12): $s_t^{(n)} = dyn(s_{t-1}^{(i)})$

(d) uteži vzorec $s_t^{(n)}$ po En.8: $\pi_t^{(n)} = \pi'(s_t^{(n)})$

konec za vsak n

3. Zgosti distribucijo položajev stanj v S_t

za vsak $n = 1$ do N naredi

$$s_t^{(n)}_x = \frac{s_t^{(n)}_x \cdot \pi_t^{(n)} + s_t^{(P)}_x \cdot \pi_t^{(P)}}{\pi_t^{(n)} + \pi_t^{(P)}}$$

$$s_t^{(n)}_y = \frac{s_t^{(n)}_y \cdot \pi_t^{(n)} + s_t^{(A)}_y \cdot \pi_t^{(A)}}{\pi_t^{(n)} + \pi_t^{(A)}}$$

konec za vsak n

4. Normaliziraj uteži S_t z njihovo vsoto, da velja:

$$\sum_n \pi_t^{(n)} = 1$$

5. Določi povprečno stanje in njegovo utež (En.1,8)

$$\{s_t^{(A)}, \pi_t^{(A)}\} = \{E(S_t), \pi(s_t^A) \frac{1}{N}\}$$

6. Adaptiraj histogram modela $H_t^{(M)}$ s histogramom povprečnega stanja $H_t^{(A)}$ z močjo $p^{(A)}$ (En.10)

7. Posodobi predikcijo z uteženim stanjem $\{s_t^{(A)}, \pi_t^{(A)}\}$

8. Če je potrebno, naredi masko (razdelek 7)

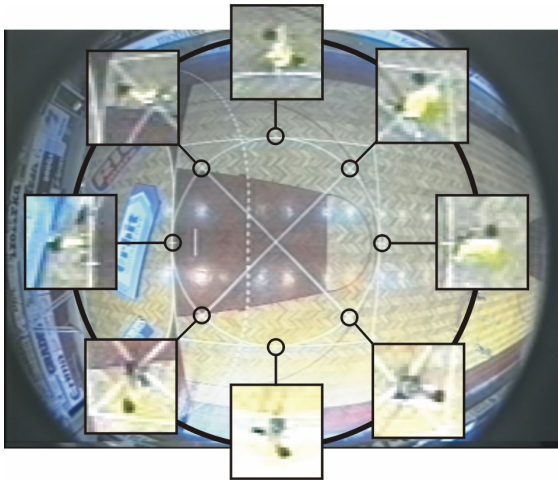
Slika 1: Iteracija sledenja.

Parametri naivnega in razširjenega sledilnika so prikazani v Tabeli (2). Tipični kalibrirani posledeni

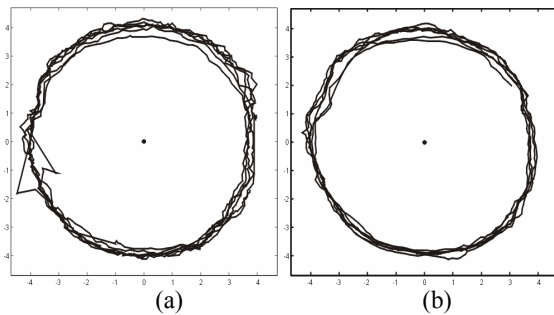
trajektoriji naključnega igralca prikazuje Slika 3(a,b), kjer je izboljšava gladkosti sledenja z uporabo razširjenega sledilnika očitna.

št. stanj-N	μ_{err} [%]	σ_{err} [%]	št. odpovedi / min
<i>brez uporabe domneve zaprtega sveta</i>			
20	18.0	10.0	15
40	12.0	8.0	10
100	12.0	8.0	8
<i>z uporabo domneve zaprtega sveta</i>			
20	3.0	2.0	0.5
40	3.0	1.0	0.5
100	3.0	1.0	0.5

Tabela 1: Rezultati sledenja pri različnem številu stanj v diskretnem naboru, z ali brez uporabe domneve zaprtega sveta.



Slika 2: Igrišče z narisanim krogom, ter povečane slike igralca na različnih položajih med sledenjem.



Slika 3: Tipična trajektorija igralca, ki teče po krogu posledena z naivnim sledilnikom (a), ter z razširjeno izvedbo (b).

Ω_{max}	Φ_{tr}	σ_{sim}	σ_{din}	σ_{siz}	$(a,b)_{min}$	$(a,b)_{max}$	T
0.1	0.3	0.2	0.6	0.1	[slik. el.] (4,4)	[slik. el.] (5,5)	10

Tabela 2: Parametri sledilnikov. Naivni sledilnik je uporabljal le parametre (σ_{sim} , σ_{din} , σ_{siz} , a_{min} , a_{max} , b_{min} , b_{max}), razširjeni pa vse v tabeli.

10 Zaključek

V članku smo predstavili sledilnik, ki temelji na principu CONDENSATION algoritma, ter upošteva domnevo zaprtega sveta. Domnevo smo definirali s tremi postavkami o igrišču ter barvi in obliki igralcev. Rezultati na šest minutnem posnetku gibanja igralcev po krogu z znano dimenzijo so pokazali, da je tak sledilnik precej bolj točen v ocenjevanju pretečenih poti kot sledilnik, ki ne upošteva domnev. Kar zadeva števila odpovedi na minuto sledenja je nedvomno robustnejši od naivnega, prav tako pa ga odlikuje lastnost, da za dobro delovanje ne zahteva velikega števila vzorcev v naboru stanj. Izkaže se namreč, da sledilnik s štiridesetimi vzorci v naboru stanj enako uspešno sledi igralce kot sledilnik s stotimi.

V luči dosedanjega dela nameravamo v prihodnosti vključiti še več domnev zaprtega sveta za boljšo uspešnost pri trkih med podobnimi igralci, ter izboljšati določanje maske igralca z bolj naprednimi metodami.

11 Literatura

[1] S.S. Intille, A.F. Bobick: "Closed-World Tracking", *ICCV*, pp. 672-678, 1995

[2] M. Isard, A. Blake: "CONDENSATION-conditional density propagation for visual tracking", *Int. J. Journal of Computer Vision*, vol. 29 (1), pp. 5-28, 1998

[3] M.J. Swain, D.H. Ballard: "Color indexing", *Int. J. Journal of Computer Vision*, vol. 7(1), pp. 11-32, 1991

[4] K. Nummiaro, E. K. Meier, L. Van Gool: "A Color-Based Particle Filter", *ECCV02*, pp. 53-60, 2002

[5] Chris Needham and Roger Boyle: "Tracking multiple sports players through occlusion, congestion and scale", *BMVC*, vol. 1, pp. 93-102, 2001

[6] M. Matsumoto, T. Nishimura: "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator", *TOMACS*, vol. 8(1), pp. 3-30, 1998

[7] J. Perš, M. Bon, S. Kovačič, M. Šibila, B. Dežman: "Observation and analysis of large-scale human motion", *Human Movement Science*, vol. 21(2), pp. 295-311, 2002

[8] G. Klančar, M. Kristan, R. Karba: "Wide-angle camera distortions and non-uniform illumination in mobile robot tracking", *Robotics and Autonomous Systems*, vol. 46(2), pp. 125-133, 2004