

O tehnică de urmărire a bilelor pentru aplicații de realitate îmbogățită

Dragoș Cătălin Barbu

ICI București

Bd. Mareșal Averescu,
Nr.8-10, București

dbarbu@ici.ro

Silvia-Ramona Bunea

Universitatea Politehnică București

Splaiul Independenței, Nr.313,
București

ramona.s.bunea@gmail.com

Ștefan Măgureanu

Universitatea Politehnică București

Splaiul Independenței Nr.313,
București

smagureanu@gmail.com

REZUMAT

În acest articol se prezintă o tehnică de urmărire robustă și cu un necesar redus de resurse de calcul bazată pe bile colorate. Tehnica se bazează pe un algoritm de transformare Hough aleatoare, care este optimizat pentru a fi folosit în aplicații în timp real, cum sunt sistemele de realitate îmbogățită. Necesitând doar o cameră color convențională atașată la o unitate PC obișnuită, această abordare oferă posibilitatea de a determina poziția 3D a unor bile colorate, pe baza procesării de cadre interactive. Algoritmii sunt suficient de rapid pentru a fi combinat ușor cu alte motoare de urmărire în timp real. Spre deosebire de tehnicile de urmărire bazate pe recunoașterea planară a markerilor, această tehnică oferă robustețe față de ocluzia parțială, ceea ce facilitează interacțiunea cu obiectele reale.

Cuvinte cheie

Segmentare, analiza scenei, urmărire, AR.

INTRODUCERE

O problemă cheie în proiectarea și implementarea interacțiunii cu obiecte reale în sistemele de realitate îmbogățită – AR (*Augmented Reality*) este urmărirea (*tracking*). În scopul furnizării unui feedback în timp real, este necesară actualizarea imaginii generate de calculator atât în funcție de poziția utilizatorului cât și a obiectelor reale.

De exemplu, în cazul unui sistem educațional bazat pe AR pentru învățarea chimiei, se utilizează două tipuri de obiecte reale: un tabel periodic al elementelor și un set de bile reprezentând atomi. Tehnica de interacțiune trebuie să permită atât urmărirea tabelului periodic cât și a bilelor ce vor fi asociate cu atomi.

Urmărirea unei bile într-un mediu dinamic și imprevizibil folosind tehnicile de vizualizare 3D devine dificilă atunci când bila este ascunsă sau mascată de alte obiecte. În [10] se prezintă o metodă de urmărire a unei bile într-un mediu de interacțiune care reproduce un teren de foosball semi-automat, folosind cunoștințe despre mediul staționar. Terenul de foosball este semi-automat permițând folosirea sa de către un singur jucător.

Tehnicile low-cost din domeniul vederii automate (*computer vision*) se referă la tehnicile ieftine din punct de vedere financiar și cu cerințe reduse față de complexitatea de calcul fiind mai interesante din motive de productivitate și de acceptare pe scară largă.

O tehnică *low-cost* larg utilizată în urmărirea obiectelor în sistemele AR este bazată pe recunoașterea optică a markerilor fiduciar planari (de ex. ARToolKit [5], ARTag[8], reacTIVision [1], ARToolKitPlus [21]). Această tehnică permite estimarea pozițiilor relative și a orientărilor a mai multor markeri vizibili în timp real. Principala limitare a acestei tehnici este necesitatea unei zone planare neflexibile fixată în apropiere sau pe un obiect real. O astfel de modificare nefirească scade, de obicei, ușurința de manipulare și afectează feedback-ul haptic.

Un alt dezavantaj al markerilor fiduciar este sensibilitatea lor mare la ocluzii. Chiar dacă o mică parte din marker este obstrucționată, de exemplu, ca urmare a manipulării utilizatorului, urmărirea obiectului nu se mai poate realiza. În plus, markerii sunt obiecte artificiale și un număr mare de markeri într-o zonă de interacțiune cu obiecte reale provoacă o perturbare vizuală care afectează experiența utilizatorului.

În această lucrare se prezintă o tehnică de urmărire destinată proiectării unui scenariu de învățare a chimiei pe o platformă AR. Tehnica se bazează pe un algoritm de urmărire a unor bile colorate cu diametrul cunoscut [19]. Această tehnică are două avantaje, în comparație cu tehnicile bazate pe markeri planari: bilele sunt în mod natural percepute ca obiecte reale, furnizând un feedback haptic plăcut și pot fi detectate chiar și în ocluzie parțială.

Restul acestui articol este organizat după cum urmează. În capitolul următor se prezintă abordările din tehnicile de urmărire. Apoi se va prezenta tehnica de urmărire implementată cu principalii săi pași: calibrarea, urmărirea în timp real, segmentarea și tehnica de ajustare. Lucrarea se încheie cu concluzii și direcții de continuare a cercetărilor.

ABORDĂRI ÎN TEHNICILE DE URMĂRIRE

Bilele au fost folosite ca dispozitive de interacțiune 3D în puține lucrări. Bruijnen et al. [3] în mediul Robocup folosesc roboți autonomi care trebuie să urmărească o minge și să o captureze. Un alt exemplu poate fi găsit în sistemul de computer vision Hawkeye [15] unde camerele sunt poziționate de-a lungul unui teren de tenis pentru urmărirea bilei. Bradley și Roth [2] folosesc o minge albastră acoperită cu puncte roșii și verzi pentru estimarea poziției și orientării. Algoritmii lor de detectarea bilei în timp real este bazat pe un prag de culoare simplă și estimarea cercului de legătură (*bounding circle estimation*) care, din păcate, este sensibilă la zgomotul de

fundal și ocluzii. Greenspan și Fraser [9] au introdus un algoritm de urmărire a bilei mai robust pentru estimarea poziției sferei dipol. În acest caz, dipol înseamnă două bile conectate cu un băț obținând cinci grade de libertate. Algoritmul lor are două faze. Primul pas este blocarea (1 sau 2 secunde) care permite găsirea bilelor în pozițiile comune și a doua fază este urmărirea în timp-real care ia în considerare poziția anterioară a bilei. Principalul dezavantaj al acestei abordări este faza de blocare, care introduce întârzieri notabile în interacțiune.

Detectarea bilei și urmărirea sa poate fi înțeleasă ca un caz special a unei probleme mai generale studiată în tehnicile *low-level vision*: detectarea curbilor parametrice în imagini. Deoarece proiecția perspectivă a unei sfere este întotdeauna un cerc, în cazul nostru considerăm doar detectarea cercului.

Cercetarea în acest domeniu a început cu seminarul *Circle Hough transform* (CHT) [7] care a extins principiile de bază ale transformării Hough la detectarea cercurilor. Ideea de bază este similară cu cea a detecției liniilor, adică de a extrage muchiile înainte iar apoi, pentru fiecare pixel de pe muchie să acumulăm voturi într-o submulțime adecvată a spațiului parametric (aici fiind reprezentat ca o histogramă 3D a centrelor cercurilor și a razelor) folosind toate cercurile posibile care trec prin el. Vârfurile semnificative dintr-o astfel de histogramă determină centrele și razele de cercuri importante din imagine. Deși transformarea Hough circulară originală (CHT) este robustă la zgomot și ocluzie, principalul său dezavantaj este complexitatea de calcul. Pentru a depăși această problemă și pentru a păstra robustețea CHT-ului original s-au dezvoltat diverse tehnici derivate.

Una din primele îmbunătățiri este bazată pe o paradigmă ierarhică descoperită de Li et al. [14] care începe cu grila grosieră (*coarse grid*) și efectuează subdiviziuni doar atunci când numărul de voturi depășește un anumit prag predefinit. O reducere similară a dimensionalității spațiului parametric este folosită pentru a reduce numărul de voturi [11]. În această metodă, o submulțime selectată de parametri este estimată la început (de ex. centrul unui cerc) și apoi această soluție inițială este folosită pentru a reduce numărul de voturi necesare pentru parametri rămași (de ex. raza). Deși aceste abordări pot reduce considerabil numărul de voturi, complexitatea lor de calcul este încă departe de a fi considerată utilizabilă în aplicații în timp real, deoarece numărul de pixeli de pe muchie din imagine poate fi foarte mare.

O altă idee este luarea în considerare și a orientării gradientului a pixelilor de pe muchie. Kimme et al. [13] folosesc această abordare pentru a estima segmente de cercuri în loc de cercurile întregi. Rad et al. [16] folosește perechi de vectori gradient opuse și gruparea pe distanțe Euclidiene pentru estimarea celor mai probabili parametri ai cercului.

Totuși, aceste metode au două mari dezavantaje. Primul este complexitatea de calcul ascunsă în faza de preprocesare unde un detector de muchii robust (de ex. [4]) este folosit pentru estimarea precisă a locației și orientării muchiilor în imaginile cu zgomot. Al doilea dezavantaj este cauzat de existența magnitudinilor neglijabile ale

gradientilor de pe frontierele bilei, datorită umbririi. Prezența lor împiedică estimarea corectă a orientării și forțează pragurile magnitudinilor joase ale gradientilor care fac să crească numărul total de pixeli de pe muchie care trebuie să fie procesați.

Abordarea prezentată în acest articol este inspirată din metodele bazate pe accesul aleator la datele de intrare. Această tehnică a fost prima oară introdusă în detectarea cercurilor de Xu et al. [22] care ia iterativ câțiva tripleți de pixeli de pe muchie aleși aleator, construiesc un cerc și folosesc cei trei parametri ai săi pentru votare în histograma 3D reprezentată de structura dinamică a datelor. Această procedură este repetată de mai multe ori și cercul cu cea mai mică eroare este ales. Deși abordările aleatoare câștigă semnificativ în viteză, ele necesită efectuarea a multor voturi aleatoare atunci când numărul de pixeli de pe muchie este mare, adică atunci când pragurile magnitudinii gradientilor sunt scăzute (cum este și în cazul nostru). Cu toate acestea, detectarea muchiilor bazate pe culoare permite reducerea numărului de pixeli de pe muchie, astfel că votul aleator devine trasabil computațional și mai eficient în comparație cu alte tehnici.

Deoarece abordările bazate pe CHT permit doar estimarea exactă a centrului cercului și a razei se folosesc tehnici de ajustare a erorilor pătratice minime pentru a obține o acuratețe sub-pixelară. Cu toate acestea, aceste metode sunt foarte sensibile la valori aberante și trebuie folosite cu atenție. În practică, este necesar să se preselecționeze o submulțime semnificativă a pixelilor de pe muchie folosind tehnici robuste (de ex. bazate pe CHT). Aceeași strategie vom folosi, dar în plus includem ponderarea după magnitudinea gradientului introdusă de Sykora et al. [19], pentru a îmbunătăți precizia.



Figura 1. (a) Imaginea color de intrare, (b)rezultatul segmentării imaginii color prin mean-shift și potrivirea cercului

TEHNICA DE URMĂRIRE IMPLEMENTATĂ

Ideea principală care a permis robustețe și viteză mare de procesare este utilizarea detecției muchiilor color bazată pe o segmentare color rapidă, care produce un număr mult mai mic de pixeli de muchie, spre deosebire de celelalte abordări standard bazate pe luminozitate. Această reducere scade dramatic numărul de voturi necesare pentru detectarea robustă a parametrilor cercului, astfel încât se poate estima poziția mai multor bile colorate în timp real.

Întregul cadru constă în două etape principale. În prima etapă de calibrare off-line parametri intrinseci ai camerei și denaturarea radială sunt estimați și un clasificator color este învățat pe un exemplu de imagine cu bile colorate.

Apoi, urmează etapa de urmărire on-line în timp real în care constă în trei pași principali: segmentarea imaginii color în care imaginea de intrare este convertită în mai multe regiuni; estimarea robustă a parametrilor cercului și rafinarea parametrilor cercului.

Calibrarea

Faza de calibrare off-line constă în două etape. Prima

etapă estimează parametri intrinseci și parametri radiali de denaturare ai camerei din mai multe fotografii ale unui pattern de șah luate în conformitate cu diferite orientări și poziții folosind ArToolKit[5].

În etapa a doua se estimează distribuția de culori a fiecărei bile care va trebui recunoscută în timpul fazei de urmărire în timp real. În acest scop se aranjează bilele colorate pentru a fi vizibile de către cameră și se achiziționează o singura imagine colorată (Fig 1.a). Apoi se va realiza segmentarea imaginii folosind o versiune modificată a algoritmului mean-shift [6]. În acest caz este importantă doar constanta culorii regiunii, prin urmare se va efectua segmentarea mean-shift indiferent de luminanță (*luminance*) folosind doar componentele de culoare (din spațiul de culoare LUV [20]). Prin această modificare se oferă o robustețe la variațiile de luminozitate pe suprafața bilei (Fig. 1b).

Pentru etichetarea regiunilor, la segmentarea imaginii se efectuează analiza componentelor conectate [17]. Pentru fiecare regiune suficient de mare se aplică un algoritm de potrivire a cercului [19].

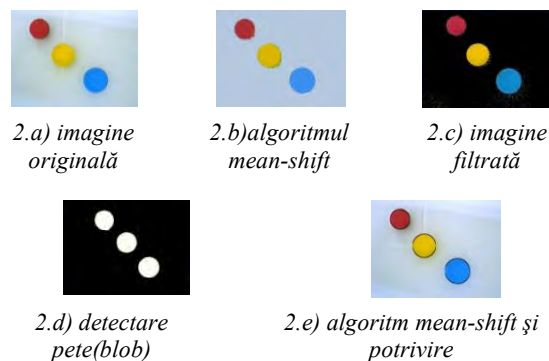
După ce toate probele de culoare sunt colectate, se aplică o operație de închidere [18] pe matricea de distribuție a culorilor în spațiul LUV, pentru a acoperi întreaga paletă de culori asociată unei bile. În final, se folosește distribuția ne-parametrică a culorilor pentru a pre-calcula un clasificator RGB simplu ce permite convertirea imaginii de intrare într-un singur indice. Clasificatorul este implementat ca o tabelă de aspect 3-D care convertește triplete RGB în valori întregi.

Segmentarea

Pentru început se achiziționează o imagine de la cameră (Fig. 2a) și apoi se aplică clasificatorul RGB pentru a obține un indice de culoare unic pentru fiecare pixel (Fig.2b). Apoi, ca și în faza de calibrare, se folosește analiza componentelor conectate (Fig. 2d) pentru a obține o listă a regiunilor și a relațiilor lor de vecinătate. Pentru reducerea zgomotului și pentru a îndepărta golurile micile regiuni sunt conectate la vecinătățile lor suficient de mari. După aceasta vom obține o mulțime de regiuni unde fiecare poate reprezenta una sau mai multe bile (Fig. 2c). Pentru procesările următoare, fiecare regiune este reprezentată de o mulțime de pixeli de frontieră nedistorsionați.

Acum problema este mult mai simplă în comparație cu detectarea cercului din muchiile de luminanță deoarece va trebui să căutăm doar într-o submulțime limitată a spațiului parametric 3D. Totuși, este necesară o estimare robustă pentru a depăși problemele ivite din cauza mai multor bile într-o regiune, de ocluzie și de zgomotul de fundal.

Folosim o abordare a transformării Hough în care în prima fază un cerc inițial este estimat robust folosind votarea aleatoare oarbă cu reducerea dimensionalității spațiului parametric și apoi parametrii cercului sunt rafinați pentru a se potrivi mai bine cu frontierele bilei reale folosind o tehnică a pătratului minim. Această combinație permite atingerea unui echilibru optim între robustețe, precizie și viteză de procesare.



Estimarea robustă a parametrilor cercului

Se pornește de la ideea că fiecare pată de culoare reprezintă o bilă, deci dacă alegem 3 pixeli de frontieră pentru o pată este o șansă foarte mare ca acestea să determine un cerc ce va încadra aproximativ bila corespunzătoare. Vom efectua un număr suficient de astfel de voturi și vom determina centrul cel mai frecvent determinat de acestea. Odată fixat un centru, vom stabili raza ca fiind distanța cea mai frecventă între acesta și fiecare pixel de frontieră și vom determina dacă această pată de culoare este într-adevăr un cerc folosind un coeficient:

$$Q = \frac{v_centru * v_raza}{N_voturi * N_pct}$$

unde v_centru este numărul de voturi obținute de centrul ales, v_raza este numărul de voturi obținute de raza aleasă, N_voturi este numărul de voturi totale pentru determinarea centrului iar N_pct este numărul pixelilor de frontieră al petei. Valoarea lui Q pentru pete ce nu reprezintă cercuri sau grupuri de cercuri este cu 3-4 ordine de mărime mai mare decât în cazul altor forme.

Rafinarea parametrilor cercului

După estimarea robustă a cercului trebuie să rafinăm centrul lui și raza pentru a se potrivi mai bine cu frontiera bilei reale. Pentru a simplifica această sarcină am presupus că pixeli care reprezintă frontiera bilei au o magnitudine mare a gradientului intensității și sunt simultan apropiați de cercul estimat. Aplicăm abordarea introdusă de Sykora et al.[19] care utilizează diferențele centrale simple pe intensitățile pixelilor într-un mic cerc în jurul cercului estimat și selectăm pixeli cu răspunsuri mari.

Rezultate

Acest algoritm a fost implementat și va fi testat pe o aplicație AR de învățare a chimiei. Vom detalia limitările care apar în cadrul acestui algoritm și în urma experimentelor efectuate.

Principala limitare a acestui algoritm este presupunerea făcută asupra spectrului de culori constant provenite de la lumină. Această limitare este strâns legată de numărul de culori distincte care vor fi recunoscute. Astfel, se face un compromis între numărul de culori active și dimensiunea cluster-elor de culori. Dacă clusterelor sunt mari oferă o mai mare robustețe pentru schimbarea de culori dar scad numărul de culori distincte și crește probabilitatea de coliziune cu obiectele de pe fundal.

Conform Sykora et al. [19] 4 clustere de culoare distincte oferă un compromis optim între numărul de culori și robustețea sistemului. Atunci când lumina zilei este folosită pentru iluminare și sunt necesare mai multe culori este necesar să se folosească faza de calibrare de mai multe ori în decursul unei zile.

O altă limitare importantă este precizia estimării adâncimii bilei. Când o bilă se află departe de cameră sau este puternic ocluzionată estimarea poziției cercului și a razei poate fi influențată și astfel adâncimea rezultată poate fi diferită în mod semnificativ față de cea reală. O situație similară are loc atunci când condițiile de iluminare sunt slabe și/sau se schimbă în spectrul luminii și cauzează o clasificare zgomotoasă a culorilor. Acest tip de deviație este de obicei tranzitorie și poate fi depășită prin filtre de trecere-jos sau filtrare Kalman [12] în detrimentul unei latențe în mișcare.

CONCLUZII ȘI DIRECȚII DE CONTINUARE

În această lucrare a fost prezentată o tehnică de urmărire a bilei pentru sisteme AR. Folosind o cameră color standard permite estimarea pozițiilor 3D a mai multor bile color în timp real. Comparativ cu markeri planari fiduciali această abordare este robustă față de ocluzia parțială dispunând de o manipulare mai confortabilă și de un feedback haptic adecvat.

Pentru a urmări instanțele bilelor în mod sistematic și pentru a evita eșecurile de detectare pe timp scurt cauzate de ocluzii sau schimbări bruște a condițiilor de iluminare (de ex., din cauza luminii fluorescente pâlpâitoare), dorim să dezvoltăm pe viitor un istoric temporal al pozițiilor bilei. Pentru aceasta se poate folosi o asignare a urmăririi celui mai apropiat vecin cu întârziere temporală. Fiecare bilă detectată va fi asignată celei mai apropiate urme. În practică, acest mecanism permite evitarea pâlpâirii (*flickering*) și brusca interschimbare a augmentării atunci când diferite obiecte virtuale sunt atribuite bilelor cu același indice de culoare.

Un alt aspect care merită să fie luat în considerare este precizia de estimare a razei bilei în cazul ocluziilor esențiale.

CONFIRMARE

Această lucrare a fost suportată din proiectul de cercetare PN0503/2009.

REFERINȚE

- Bencina R., Kaltenbrunner M., Jorda S.: Improved topological fiducial tracking in the reactivation system. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2005), vol. 3, pp. 99–106.
- Bradley D., Roth G.: Natural interaction with virtual objects using vision-based six DOF sphere tracking. Proceedings of the International Conference on Advances in Computer Entertainment Technology (2005), pp. 19–26.
- Bruijnen, D., Aangent, W., Helvoort, J., v.d. Molengraft, J.G., Steinbuch, M.: From Vision To Realtime Motion Control For The Robocup Domain. IEEE Conference on Control Applications, Singapore (2007).
- Canny J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 6 (1986), pp. 679–698.
- Claus D., Fitzgibbon A.: Reliable fiducial detection in natural scenes. Proceedings of European Conference on Computer Vision (2004), pp. 469–480.
- Comaniciu D., Meer P.: Robust analysis of feature spaces: Color image segmentation. Proceedings of Conference on Computer Vision and Pattern Recognition (1997), pp. 750–755.
- Duda R. O., Hart P. E.: Use of the hough transformation to detect lines and curves in pictures. Communications of the ACM 15, 1 (1972), pp. 11–15.
- Fiala M.: ARTag, a fiducial marker system using digital techniques. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2005), vol. 2, pp. 590–596.
- Greenspan M., Fraser I.: Tracking a sphere dipole. Proceedings of International Conference on Vision Interface (2003), pp. 154–161.
- Janssen, R., Jeroen de Best, Van de Molengraft, R.: Real-Time Ball Tracking in a Semi-Automated Foosball Table, Robocup 2009, J. Baltés et al. (Eds), pp. 128-139, (2010), Springer-Verlag
- Jiang X., Cheng D.-C.: Fitting of 3D circles and ellipses using a parameter decomposition approach. Proceedings of International Conference on 3-D Digital Imaging and Modeling (2005), pp. 103–109.
- Kalman R. E.: A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82 (1960), pp. 35–45.
- Kimme C., Balard D., Sklansky J.: Finding circles by an array of accumulators. Communications of the ACM 18, 2 (1975), pp. 120–122.
- Li H., Lavin M. A., Lemaster R. J.: Fast hough transform: A hierarchical approach. Computer Vision, Graphics and Image Processing 36, (1986), 139–161.
- Owens, N., Harris, C., Stennet, C.: Hawk-Eye Tennis System. International Conference on Visual Information Engineering, pp. 182–185 (2004) ISBN 0-85296-757-8.
- Rad A. A., Faez K., Qaragozlou N.: Fast circle detection using gradient pair vectors. Proceedings of International Conference on Digital Image Computing (2003), pp. 879–888.
- Rosenfeld A., Kak A. C.: Digital Picture Processing, vol. 1. Academic Press, Orlando, USA, 1982.
- Serra J.: Image Analysis and Mathematical Morphology. Academic Press, 4th Edition, 1993
- Sýkora, D., Sedlaček, D., Riege, K.: Real-time Color Ball Tracking for Augmented Reality. Proceedings of Eurographics Symposium (2008)
- Wyszecki G., Stiles W. S.: Color Science: Concepts and Methods, Quantitative Data and Formulae. Wiley, 2nd Edition, 1982.
- Wagner D., Schmalstieg D.: ARToolKitPlus for pose tracking on mobile devices. Proceedings of Computer Vision Winter Workshop (2007).
- Xu L., Oja E., Kultanen P.: A new curve detection method: Randomized hough transform (RHT). Pattern Recognition Letters 11, 5 (1990), 331–338.