

METODE DE RECUNOAȘTERE BAZATĂ PE MODEL A OBIECTELOR TRIDIMENSIONALE

mat. Gabriela Dăncescu

Institutul de Cercetări în Informatică

Rezumat: În articolul de față, ne propunem să abordăm diferite metode de recunoaștere bazată pe model a obiectelor tridimensionale, problemele care apar în proiectarea unui sistem de vedere bazat pe model, reprezentări, scheme de modelare a obiectului etc.

Cuvinte cheie: CAD/CAM, recunoaștere obiect, sistem de vizualizare bazat pe model, scheme de modelare.

Scopul unui sistem de vedere este acela de a interpreta datele vizuale și de a utiliza această interpretare pentru a rezolva o anumită sarcină.

Utilizările specifice ale unui sistem de vedere sunt: navigația vehiculelor autonome (pe pământ, în aer sau sub apă), amplasarea sau inspecția anumitor piese manufacturate, analiza imaginilor microscopice sau a radiografiilor medicale.

În unele aplicații, scopul unui sistem de vedere artificială este de a identifica și de a localiza un obiect specificat într-o scenă. În aceste cazuri, un sistem de vizualizare trebuie să cunoască în întregime forma obiectului dorit. O astfel de cunoaștere dinainte a obiectului este furnizată prin intermediul unui model al obiectului. În cele mai multe cazuri el conține informații ale geometriei obiectului. Unele modele pot să conțină informații suplimentare cum ar fi: proprietățile termice și proprietățile referitoare la tensiunile ce apar în obiect la deformare.

Un sistem de vizualizare care utilizează un astfel de model al unui obiect se numește **sistem de vizualizare bazat pe model**, iar problema generală a identificării obiectului dorit face referire la **recunoașterea obiectului**.

Un sistem ideal de vedere bazat pe model ar trebui să fie capabil să localizeze obiectele dintr-o scenă presupunând că:

- obiectele pot avea forme arbitrare și complicate
- obiectele pot fi văzute din orice direcție, și

- obiectele pot fi parțial eclipsate de către alte obiecte.

Pentru a proiecta un astfel de sistem trebuie stabilite următoarele:

- a) tipul senzorului pentru culegerea datelor
- b) metodele de construcție a modelului obiectului
- c) mijloacele de descriere a datelor culese și a modelului
- d) metodele de împerechere a descrierilor obținute din datele de intrare, cu cele ale modelului.

Senzorul stabilește rezoluția și precizia. Mai mult, el stabilește dacă datele provin de la o informație 2D sau 3D a scenei. Modelele furnizează cunoștințe sistemului de vedere.

1. Probleme care apar în proiectarea unui sistem de vedere bazat pe model

a) Culegerea datelor

Prima problemă întâlnită într-un sistem de vedere este culegerea datelor. Această problemă, în cele mai multe sisteme, este rezolvată cu ajutorul unei camere de luat vederi care măsoară lumina vizibilă. Ieșirea (rezultatul furnizat de cameră) camerei este digitizată pentru a furniza un tablou bidimensional de numere; fiecare număr corespunde intensității medii, resimțite într-un pătrat foarte mic.

Există și alte modele de senzori, cum ar fi: camerele termice, care măsoară radiația termică emisă de un obiect, scanere cu laser și senzori sonari, care sunt utilizate pentru a calcula distanța la obiect.

b) Reprezentarea datelor

Tabloul bidimensional de numere furnizat de un senzor nu este prea util în forma lui brută. Pentru a descrie datele și modelul este necesară o schemă de reprezentare care trebuie să fie neambiguă (să nu existe două obiecte care să aibă aceeași reprezentare), unică (există o singură descriere pentru fiecare obiect), nesenzitivă (în ceea ce privește absența unor anumite puncte, cum

ar fi în cazul acoperirilor parțiale), convenabil de utilizat (în etapa de potrivire și de păstrare).

c) Construirea modelului

Pentru construirea modelului există două abordări principale.

Prima abordare utilizează obiecte existente pentru a genera un model; adică punctele obținute, privind obiectul din mai multe poziții, sunt integrate astfel încât să se furnizeze o informație din toate unghiurile posibile.

A doua abordare utilizează un sistem CAD și un set de primitive predefinite, care permite utilizatorului construirea interactivă a modelului unui obiect.

d) Strategia de împerechere

O dată ce există descrieri corespunzătoare ale datelor culese și ale modelului sistemului de vizualizare se poate trece la împerecherea celor două descrieri. Aceasta se realizează în doi pași:

Pasul 1 se stabilește o corespondență între cele două mulțimi de descriere. Deoarece, în multe cazuri, datele sunt culese dintr-o singură vedere și pot apărea acoperiri parțiale, atunci strategia de împerechere trebuie să stabilească o corespondență între descrierea parțială a obiectului și descrierea completă a modelului.

Pasul 2 utilizând corespondențele stabilite se folosește o transformare geometrică (în general, o matrice de rotație compusă cu o matrice de translație) astfel încât modelul să poată fi transformat pentru a avea orientarea obiectului din scenă.

2. Percepția tridimensională

Sarcina achiziționării datelor 3D poate fi realizată în două moduri: pasiv și activ. În primul caz, informația 3D este preluată din scenă utilizând energia existentă în mediu, cum ar fi lumina reflectată. În al doilea caz, informația este obținută utilizând energia exterioară, cum ar fi undele sonore sau lumina laser. Datelor preluate în primul caz le lipsește precizia și rezoluția necesare în recunoașterea obiectelor tridimensionale.

3. Prelucrarea datelor achiziționate

Datele trebuie reprezentate utilizând descrieri simbolice pentru a permite sistemului realizarea proceselor de nivel înalt. Primul pas în obținerea reprezentării simbolice este împărțirea datelor de intrare, bazându-ne pe descrierea dorită. Metoda prin care sunt descrise în sistem formele are o mare importanță. Într-un număr mare de cazuri, se utilizează proprietățile locale ale suprafețelor, cum ar fi curbura și normala la suprafață. Curbura arată cât de mult deviază suprafața de la planul ei tangent; este invarianta la direcțiile din care se privește și nu se schimbă datorită acoperirilor. Are și avantajul că, în cazul suprafețelor netede, poate fi aproximată utilizând numai o mică vecinătate.

Înainte de calculul acestor proprietăți, datele trebuie să fie netezite din două motive: în primul rând, datorită faptului că măsurătorile făcute cu ajutorul senzorilor sunt deseori imprecise datorită zgomotului prezent în detector sau a erorilor de calibrare; în al doilea rând, calculul curburilor necesită derivate parțiale de ordinul 2, care conduc la majorarea efectului oricărui zgomot prezent.

4. Reprezentări

Pentru descrierea formelor se utilizează reprezentările. Spre deosebire de mărimile cuantificabile, cum ar fi mișcarea sau intensitatea, descrierea formelor 3D complexe necesită sute de parametri și este foarte complicată.

5. Reprezentări bazate pe suprafețe

În această clasă de reprezentări sunt approximate proprietățile suprafețelor (normala și curbura Gaussiană) și sunt utilizate în descrierea punctelor achiziționate.

a) Sfera gaussiană

În această reprezentare, normala la suprafață, în fiecare punct al ei, este mapată într-un punct corespunzător al sferei unitate. Reprezentarea este invariantă la rotație (dacă obiectul este rotit, atunci imaginile normalelor pe sfera gaussiană se rotesc corespunzător); totuși, reprezentarea nu este invariantă la translație și la mărirea dimensiunilor obiectului.

În lucrarea (Roach, 1987), se descrie o nouă reprezentare, bazată pe o extensie a sferei gaussiene, numită reprezentarea sferică duală. În această reprezentare, se utilizează fațete planare pentru a aproxima, atât obiecte poliedrale, cât și netede. Proprietățile sferei gaussiene sunt

combinat cu cele ale spațiului dual pentru a acoperi o gamă mai mare de obiecte ce pot fi reprezentate. Spațiul dual se bazează pe dualitatea care există între plane și puncte, adică pentru orice teoremă geometrică dată, există o teoremă corespunzătoare duală care se obține prin înlocuirea punctelor cu planele și invers. În reprezentarea duală planele sunt reprezentate prin puncte, iar orientarea lor are semnul determinat de poziția punctului din spațiul dual. Muchiile sunt definite ca linii ce leagă două puncte din spațiul dual. Spre deosebire de sfera gaussiană, obiectele neconvexe pot fi ușor reprezentate utilizând schema de reprezentare sferică duală. În plus, prin utilizarea reprezentării sferice duale, operațiile cu mulțimi sunt în mod natural regularizate. Regularizarea implică faptul că în urma operațiilor booleene, nu pot apărea elemente uni sau bidimensionale agățate.

Problema comună tuturor acestor reprezentări discutate mai sus este că, în cazul scenelor complexe, în care sunt prezentate mai multe obiecte, trebuie realizată în prealabil o partiționare a obiectelor pentru a evita suprapunerea reprezentărilor fiecărui obiect pe sferă.

b) Reprezentări bazate pe momente

În (Taubin, 1989), se definește forma ca o funcție de distribuție a masei, iar proprietățile formei sunt definite ca momente ale acestei funcții. Utilizând această definiție, se introduce un polinom al formei care se utilizează pentru a măsura asemănările dintre forme. Pentru împerecherea a două forme, se definește în mod corespunzător o ecuație de diferență a formelor.

6. Reprezentări bazate pe discontinuitate

Aceste reprezentări păstrează informații despre punctele unde caracteristicile unei zone de pe o suprafață se schimbă, ceea ce duce la curbele tridimensionale, cu cerințe mai mici de memorare, cu creșterea, în același timp, a eficienței algoritmilor de nivel superior. Curbele detectate sunt reprezentate prin mijloace matematice cum ar fi: curbe polinomiale parametrice și B-splinc-uri raționale.

În multe cazuri, reprezentarea nu este completă, deoarece informația privind concavitatea/convexitatea suprafeței se pierde.

a) Curbe spațiale

Curbele spațiale sunt utilizate pentru descrierea conturului frontierelor și a curbei generatoare a unei reprezentări "sweep."

b) Schița de bază a suprafeței

În această reprezentare, sunt detectate și modelate mai multe tipuri de discontinuități:

- pași- unde funcția de profunzime este discontinuă;
- acoperiri-unde normala la suprafață este discontinuă;
- îmbinări netede-unde curbura principală este discontinuă;
- ridicături - o combinație de două acoperiri.

c) Graful vederilor

Grafurile vederilor sunt grafuri în care fiecare nod reprezintă o vedere 2D dintr-un punct al unui obiect 3D, iar arcele reprezintă transformările sau "evenimentele vizuale".

Graful este o partiție a direcțiilor de vedere în regiuni stabile în cadrul cărora mici schimbări în direcția de vedere nu schimbă aspectul obiectului.

Trecerea între două regiuni stabile de direcții de vedere are asociat un eveniment vizual numit singularitate. Direcțiile din zona de tranziție între două regiuni stabile se numesc direcții de vedere accidentale. Au fost considerate cazuri relativ simple de obiecte poliedrale rigide pentru studiul grafurilor de vederi ale obiectelor poliedrale cu componente multiple în mișcare, dar și unele clase de obiecte curbate.

d) Etichetarea discontinuităților

Similar cu reprezentarea convențională prin etichetarea liniilor, a scenelor poliedrale în 2D, în graful de legare a muchiilor, nodurile reprezintă legăturile, iar arcele reprezintă muchiile care sunt legate. Fiecare arc poate fi convex, concav sau acoperitor.

7. Reprezentări volumetrice

Această clasă de reprezentări descrie volumele și nu suprafețele sau discontinuitățile suprafețelor. Reprezentările volumetrice au fost utilizate mai des pentru modelare CAD decât

pentru rezolvarea problemei de recunoaștere a obiectelor.

a) Supercuadrice

Supercuadricele au mai fost numite și superelipsoizi. Ele sunt o extensie a solidelor și a suprafețelor cuadrice. Pentru a determina îndoirea, ascuțirea la vârf, răsucirea, mărirea supercuadrice sunt utilizate mai multe tehnici de minimizare. Totuși, problema minimizării fiind neliniară, pentru stabilitate numerică sunt necesare estimări inițiale bune. De asemenea, dacă obiectul nu se presupune a fi simetric, sunt necesare date de intrare asociate mai multor unghiuri de vedere. Ca și în cazul altor scheme volumetrice, principalul dezavantaj este neunicitatea.

b) Octree-uri

"Octree-uri" constituie o altă clasă de reprezentări volumetrice. Prin descompunerea recursivă a unui volum cubic, până când fiecare subcub rezultat este omogen în raport cu un anumit criteriu, se construiește o structură arborescentă numită "octree". Există mai multe extinderi față de structura de bază a octree-urilor. De exemplu, prismele triunghiulare și paralelipipedele se mai utilizează ca elemente de "bază" ale octree-urilor.

c) Geometrie Solidă Constructivă (CSG)

În CSG, solidele sunt reprezentate prin combinații boolcne ale primitivelor solide. Fiecare obiect este reprezentat printr-un arbore unde frunzele sunt primitivele solide (cum ar fi prisme sfere, cilindri etc.), iar nodurile sunt operații boolcne folosite pentru construirea solidului; aceste operații pot fi regularizate pentru a asigura validitatea obiectelor construite. Reprezentarea CSG este mai ales utilă în construirea modelului unui obiect (ca într-un sistem CAD/CAM) și, într-o măsură mai mică, în recunoașterea obiectului.

Metoda prezintă mai multe dezavantaje din punct de vedere al recunoașterii obiectelor:

- obiectele pot fi sculptate, nu pot fi reprezentate foarte precis
- reprezentarea nu este unică
- cunoașterea primitivelor solide utilizate în timpul construirii obiectului nu este suficient de folositoare sistemului de

recunoaștere, deoarece numai suprafețele sunt vizibile.

d) Reprezentări de tip "sweep"

O metodă interactivă puternică de care pot beneficia modelatoarele geometrice este măturarea spațiului cu contururi bidimensionale pentru a defini solide - se descrie profilul 2D al unui solid și o curbă spațială care precizează traiectoria pe care se deplasează profilul când "mătură spațiul".

Dintre avantajele acestei reprezentări enumerăm:

- cel mai ridicat grad de interactivitate - datorită bidimensionalității sale, profilul poate fi introdus cu multă ușurință
- reprezentare internă compactă - se memorează numai profilul și traiectoria
- accesibilitate sporită.

Ca dezavantaj, menționăm necesitatea introducerii unui test de validitate, deoarece nu orice profil închis bidimensional generează prin măturare un obiect real.

8. Scheme de modelare a obiectelor

Modelele constituie cunoștințele geometrice și topologice a priori pe care le are sistemul de vizualizare despre o mulțime de obiecte. Aceste cunoștințe sunt comparate cu descrierile care provin din datele de intrare (obținute prin procesele de nivel jos). Există două categorii de modalități de construire ale modelului: unele bazate pe senzori, iar altele prin intermediul sistemelor CAD/CAM. Prima metodă are mai multe dezavantaje în raport cu a doua metodă:

- erori prezente în pasul de achiziție a datelor datorită erorilor de calibrare sau a zgomotului, pot afecta modelarea și complica foarte mult faza de potrivire (împerechere);
- trebuie dezvoltate tehnici pentru a înregistra date prețiate din toate direcțiile de vedere;

- în unele cazuri imaginile au puncte lipsă, datorită efectului de umbră;
- un model complet nu se poate construi, deoarece în el nu se încorporează vederile din toate unghiurile posibile; în plus și senzorul utilizat poate limita domeniul de preluare a datelor.

Un sistem CAD de modelare a obiectelor are multe avantaje. Dintre acestea enumerăm:

- informația de densitate și de culoare a obiectului, proprietățile de reflexie, termice și texturi pot fi incluse în bazele de date CAD. Aceste informații sunt utile dacă se utilizează camere termale și color ca senzori;
- multe obiecte manufacturate au fost proiectate utilizând un sistem CAD, deci există deja un model pentru a recunoaște ulterior obiectul.

9. Stabilirea corespondenței

Acest pas presupune găsirea unui model corespunzător pentru fiecare trăsătură găsită în scenă. Datele care provin de la senzor le notăm: f_i , $i=1,2,\dots,N$. Se presupune că prima trăsătură de tip dată corespunde fiecărei trăsături extrase din model, notate cu m_j , $j=1,2,\dots,M$. Aceasta se poate reprezenta prin primul nivel al arborelui (f_1, m_1), (f_1, m_2) ... (f_1, m_j).

Trăsătura de tip dată, f_2 , se presupune că ar corespunde cu fiecare trăsătura din model, m_j ; aceasta se reprezintă prin al doilea nivel în arbore. Continuând în felul acesta, se adaugă câte un nivel în arbore pentru fiecare trăsătură de tip dată, f_i .

Fiecare drum în acest arbore, de la rădăcină la o frunză reprezintă o soluție posibilă a problemei găsirii corespondenței. Ideea este de a căuta un drum astfel încât să se găsească o transformare rigidă de la datele de intrare la trăsăturile modelului. Căutarea întregului arbore este irealizabilă în cele mai multe cazuri.

Reducerea arborelui este realizată folosind două tipuri principale de constrângeri: unare și binare.

Constrângerile unare consideră numai o trăsătură la un moment dat și folosesc informații cum ar fi lungimea, raza, direcția trăsăturilor.

Constrângerile binare utilizează perechi de trăsături și consideră relațiile care pot să apară între două trăsături. Astfel, intervalul în care se găsesc trăsăturile detectate în scenă (unghiuri între normale, distanța între două muchii, distanța între centrele a două arce etc.) trebuie să fie înăuntrul intervalului determinat de trăsăturile corespunzătoare ale modelului.

Bibliografie

1. ASTHEIMER, F.D.P.: Dynamic Objects in Virtual Worlds Integrating Simulations in a Virtual Reality Toolkit. In: *Proceedings of the 1993 European Simulation Symposium, ESS'93*, pp. 299-304.
2. BLAKE, E.H., WISSKIRCHEN, P. (eds): *Advances in Object Oriented Graphics I*, Springer Verlag, Berlin, 1993.
3. FOLEY, J.D., van DAM, A., FEINER, S.K., HUGHES, J.F.: *Computer Graphics. Principles and Practice*, Addison Wesley, Massachusetts, SUA, 1992.
4. KOPARKAR, P.A., MUDUR, S.P.: Subdivision Techniques for Processing Geometric Objects. In: *Proceedings of the NATO Advanced Study Institute on Fundamental Algorithms for Computer Graphics* (R.A. Earnshaw, ed.), Springer-Verlag, Berlin, pp.751-801.
5. KRIEGMAN, D.J., PANCE, J.: On Recognizing and Positioning Curved 3D Objects from Image Contours. In: *IEEE Transactions on*

Pattern Analysis and Machine Intelligence,
Vol.12, No.12, pp.1127-1137.

6. KUIHN, V., MULER, W.: Advanced Object Oriented Methods and Concepts for Simulations of Multi-body Systems. In: *The Journal of Visualization and Computer Animation*, Vol.4, No.2, pp.95-111.

7. ROACH, F.W., PARIPATI, P.K., WRIGHT, J.S.: A CAD System Based on Spherical Dual Representations, 1987, pp. 37-44.

8. TAUBIN, G., BOLLE, R.M., COOPER, D.B.: Representing and Comparing Shapes using Shape Polynomial. In: *Proceedings of Computer Vision and Pattern Recognition*, 1989, pp.510-516.