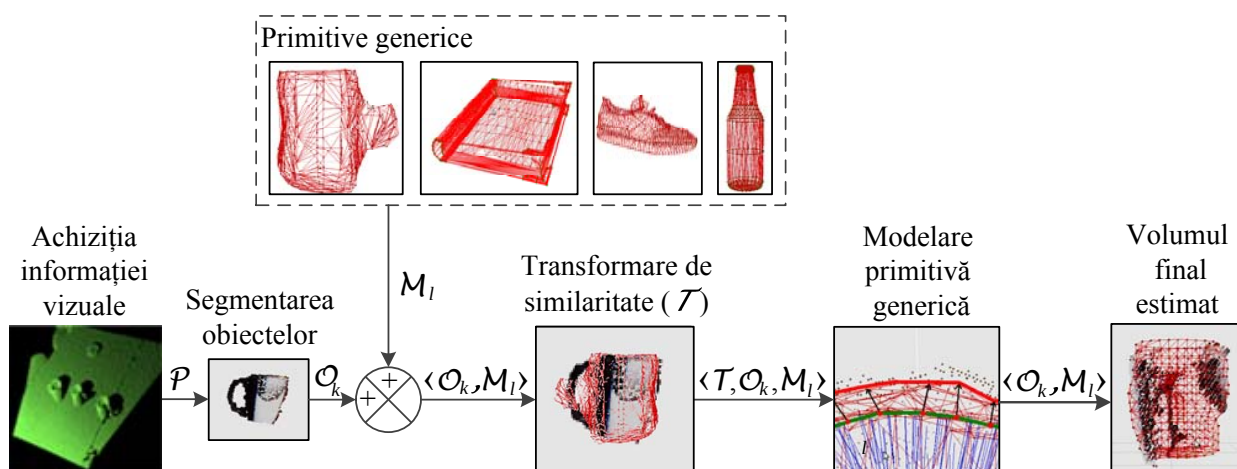


# Estimarea volumetrică 3D utilizată în structurile robotice pentru prehensarea obiectelor

<http://rovis.unitbv.ro>

**Tiberiu T. Cociaș, Florin Moldoveanu**

**Sorin M. Grigorescu**



# Prefață

---

Această carte tratează problematica simultană a segmentării și reconstrucției tridimensionale a unui obiect. Coceptul introdus în lucrare are la bază formele rigide, intitulate *primitive generice*. Scopul este acela de a permite structurilor robotice pentru prehensare, echipate cu sisteme de percepție precum camere stereo sau sensori capabili să furnizeze imagini de adancime, să prindă și să manipuleze obiectele astfel vizualizate.

Lucrarea de față se adresează în special studenților secțiilor de *Automatică și Informatică Tehnică* și *Robotică*, precum și acelor care doresc să se familiarizeze cu metodele și abordările de proiectare a unor asemenea sisteme. În prima parte a cărții sunt tratate noțiunile fundamentale de analiză semantică a spațiului tridimensional urmată de o prezentare detaliată a unora dintre cele mai relevante metode de estimare volumetrică a obiectelor din domeniul vederii mașinilor.

Ultima parte a cărții, este rezervată simulării unor scenarii de manipulare de obiecte utilizând prehensoare robotizate. Planificarea și controlul procesului de prehensare și manipulare al obiectelor sunt principalele elemente tratate în acest sens.

Autorii aduc pe această cale mulțumiri domnilor profesori Florin Moldoveanu și Claudiu Pozna, referenții științifici ai prezentei lucrări, care, prin citirea atentă a manuscrisului și prin observațiile și sugestiile făcute, au contribuit la aducerea îndrumarului la forma pe care autorii o prezintă cititorilor.

De asemenea, mulțumim membrilor și studenților grupului de cercetare *Robust Vision and Control Laboratory* (ROVIS) <http://rovis.unitbv.ro>, din care fac parte și autorii cărții, pentru sprijinul acordat de-a lungul ultimilor ani.

Brașov, Septembrie 2017

Tiberiu T. Cociaș  
Florin Moldoveanu  
Sorin M. Grigorescu

# Cuprins

---

|          |                                                                                                |           |
|----------|------------------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introducere</b>                                                                             | <b>1</b>  |
| <b>2</b> | <b>Analiza semantică a spațiului tridimensional</b>                                            | <b>7</b>  |
| 2.1      | Achiziția informației vizuale . . . . .                                                        | 8         |
| 2.2      | Reprezentarea informației vizuale . . . . .                                                    | 14        |
| 2.2.1    | Reprezentarea suprafețelor în spațiul tridimensional . . . . .                                 | 14        |
| 2.2.2    | Reprezentarea formelor în spațiul tridimensional . . . . .                                     | 18        |
| 2.3      | Manipularea informației vizuale în spațiul cartezian . . . . .                                 | 23        |
| 2.3.1    | Transformarea de similaritate . . . . .                                                        | 23        |
| 2.3.2    | Alinierea eficientă a formelor. Metoda iterativă a celui mai apropiat punct<br>vecin . . . . . | 25        |
| 2.4      | Caracteristici punctiforme 3D . . . . .                                                        | 31        |
| 2.4.1    | Identificarea celui mai apropiat punct vecin în spațiul 3D . . . . .                           | 32        |
| 2.4.2    | Estimarea normalei la suprafață . . . . .                                                      | 36        |
| 2.5      | Segmentarea prin partiționare a norilor de puncte 3D . . . . .                                 | 38        |
| 2.6      | Analiza statistică a formelor 3D . . . . .                                                     | 41        |
| 2.6.1    | Analiza Procrustes clasică . . . . .                                                           | 42        |
| 2.6.2    | Analiza Procrustes generalizată . . . . .                                                      | 44        |
| 2.7      | Concluzii . . . . .                                                                            | 44        |
| <b>3</b> | <b>Estimarea volumelor obiectelor vizualizate unidirecțional</b>                               | <b>47</b> |
| 3.1      | Estimarea volumetrică 3D utilizând forme implicite . . . . .                                   | 48        |
| 3.1.1    | Generarea unei forme implicite . . . . .                                                       | 49        |
| 3.1.2    | Estimarea formei optime a unei supercuadrice . . . . .                                         | 51        |

|          |                                                                               |            |
|----------|-------------------------------------------------------------------------------|------------|
| 3.1.3    | Modelarea supercuadricelor . . . . .                                          | 53         |
| 3.2      | Estimarea volumetrică 3D utilizând forme predefinite . . . . .                | 53         |
| 3.2.1    | Descriptori 3D . . . . .                                                      | 55         |
| 3.2.2    | Estimarea corespondențelor 3D . . . . .                                       | 56         |
| 3.2.3    | Filtrarea corespondențelor 3D . . . . .                                       | 59         |
| 3.3      | Concluzii . . . . .                                                           | 62         |
| <b>4</b> | <b>Reconstrucția volumetrică 3D a obiectelor utilizând primitive generice</b> | <b>65</b>  |
| 4.1      | Primitive generice . . . . .                                                  | 66         |
| 4.1.1    | Determinarea primitivei generice a unei clase de obiecte . . . . .            | 69         |
| 4.1.2    | Puncte de control . . . . .                                                   | 69         |
| 4.1.3    | Structura de date a unei primitivei generice . . . . .                        | 71         |
| 4.2      | Structura sistemului de reconstrucție volumetrică . . . . .                   | 72         |
| 4.3      | Recunoașterea obiectelor din scenă . . . . .                                  | 73         |
| 4.4      | Transformarea de similaritate a primitivei generice . . . . .                 | 74         |
| 4.5      | Concluzii . . . . .                                                           | 78         |
| <b>5</b> | <b>Modelarea activă a primitivelor generice</b>                               | <b>81</b>  |
| 5.1      | Contururi active . . . . .                                                    | 82         |
| 5.2      | Contururi active 2D . . . . .                                                 | 84         |
| 5.2.1    | Deformarea conturilor active prin minimizarea funcționalei de energie . .     | 84         |
| 5.2.2    | Modelarea dinamică a conturului utilizând forțe multiple . . . . .            | 90         |
| 5.2.3    | Potrivirea conturului pe granițele obiectelor din imagine . . . . .           | 91         |
| 5.2.4    | Reparametrizarea conturului activ . . . . .                                   | 92         |
| 5.2.5    | Considerații privind inițializarea conturului activ . . . . .                 | 93         |
| 5.3      | Forme active 3D . . . . .                                                     | 93         |
| 5.3.1    | Modelarea formelor active 3D prin minimizarea funcționalei de energie . .     | 96         |
| 5.3.2    | Modelarea dinamică a formei active utilizând forțe multiple . . . . .         | 102        |
| 5.3.3    | Potrivirea formei active pe granițele obiectului din scena de lucru . . . . . | 102        |
| 5.3.4    | Reparametrizarea formei active . . . . .                                      | 104        |
| 5.3.5    | Considerații privind forma activă inițială . . . . .                          | 107        |
| 5.4      | Concluzii . . . . .                                                           | 108        |
| <b>6</b> | <b>Prehensarea și manipularea obiectelor tridimensionale</b>                  | <b>111</b> |
| 6.1      | Considerații . . . . .                                                        | 112        |
| 6.2      | Sinteza operației de prehensare a unui obiect . . . . .                       | 119        |
| 6.2.1    | Planificarea procesului de prehensare a unui obiect . . . . .                 | 119        |
| 6.2.2    | Controlul operației de prehensare a unui obiect . . . . .                     | 128        |
| 6.3      | Concluzii . . . . .                                                           | 130        |
|          | <b>Bibliografie</b>                                                           | <b>133</b> |



# 1. Introducere

---

---

## Estimarea volumetrică 3D

---

---

Dezvoltarea sistemelor de vedere artificială a condus, în ultima vreme, la utilizarea, în diverse aplicații, a unui număr important de informații imagistice tridimensionale colectate de la diferiți senzori de percepție. Capacitatea unui sistem de calcul de a procesa și a înțelege astfel de informații a deschis noi orizonturi spre rezolvarea unor probleme din domenii diferite precum ar fi, de exemplu, vederea asistată de calculator sau grafica computerizată.

Percepția și înțelegerea semantică a unui obiect sau a mediului înconjurător este o sarcină ușoară pentru om. Totuși, o astfel de misiune nu este trivială pentru un sistem de calcul. În cazul unui astfel de sistem, percepția este o acțiune simplă ce presupune doar vizualizarea obiectului sau a unei scene, nu nereapărat și înțelegerea semantică a acestora. Inteligența și modul de gândire uman sunt dobândite printr-un lung și complex proces de învățare continuă. Limitate de puterea de calcul inferioară precum și de capacitatea redusă de înțelegere și învățare, sistemele de calcul pot fi utilizate doar pentru îndeplinirea unor sarcini simple precum recunoașterea de forme, navigația, diverse intervenții medicale, diagnosticare medicală, supravegherea video, prehensarea sau manipularea obiectelor, diverse aplicații militare, interacțiunea om-mașină etc. În majoritatea aplicațiilor enumerate anterior scena de lucru este vizualizată doar dintr-un singur punct. O astfel de vizualizare poartă denumirea de vizualizare unidirecțională. Contrar caracterului tridimensional (3D) al datelor percepute (vizualizate) acestea sunt denumite în literatura de specialitate date 2.5D, deoarece informațiile descriu parțial entitățile vizualizate din scenă. Tocmai acest neajuns a condus la apariția unui număr considerabil de metode și algoritmi care au ca principal scop reconstrucția și reprezentarea sub o formă solidă (volumetrică) a obiectelor sau scenelor vizualizate

[28, 29, 50, 78, 120, 161].

În ultimele decenii, motivată de numeroasele arii în care se poate aplica, reconstrucția tridimensională a obiectelor a devenit una dintre cele mai intens studiate ramuri ale vederii artificiale. Beneficiul unei reprezentări rigide a obiectului vizualizat este important oferind sistemului de procesare o serie de proprietăți unice ce permit mai departe înțelegerea complexă a datelor. Uzual, problema reconstrucției volumetrice se poate rezuma la identificarea (recunoașterea) obiectului din scenă și potrivirea unei forme rigide similare peste informația vizualizată a acestuia. Astfel, utilizarea unei forme rigide (volum) în locul reprezentării 2.5D a obiectului vizualizat este cea mai simplă și eficientă abordare pentru aproximarea informației lipsă [119].

O utilizare intensivă a tehnicilor și algoritmilor de reconstrucție volumetrică poate fi observată în domeniul roboticii unde un interes tot mai crescut este acordat roboților de serviciu care degreveză persoanele cu dezabilități de efectuarea unor sarcini care pentru ele ar fi aproape imposibile [132]. Una dintre cele mai comune activități umane de zi cu zi este prehensarea sau manipularea de obiecte<sup>1</sup>. În ultimele decenii, problema manipulării, într-o manieră autonomă, a obiectelor necunoscute a câștigat tot mai multă atenție din partea comunității științifice, în principal, datorită dezvoltării spectaculoase, așa cum s-a menționat, a roboților de serviciu [72, 92, 117]. Problema manipulării a fost studiată din punct de vedere fizic, biologic și ingineresc însă, o soluție general valabilă nu a fost identificată încă.

Prehensarea obiectelor reprezentate 2.5D este dificilă datorită inconsistenței suprafeței vizualizate ceea ce conduce la obținerea unor configurații de prehensare eronate și nesigure. Totuși, reconstrucția obiectelor 2.5D prin intermediul unor volume rigide permite obținerea unor configurații de prehensare corecte și sigure. În acest sens, interesul pentru reconstrucția tridimensională a obiectelor utilizând diferite forme volumetrice este perfect justificat.

### **Estimarea volumetrică 3D**

Reconstrucția volumetrică a unui obiect sau a unei scene nu este o sarcină ușoară mai ales atunci când aceasta este vizualizată dintr-o singură perspectivă. Având în vedere complexitatea geometrică a scenei și incertitudinea introdusă de vederea unidirecțională, capabilitatea unui sistem robotic dotat cu vedere artificială de a interacționa (intra în contact) cu obiectele vizualizate (percepute) este dificil de asigurat. Una dintre cele mai comune interacțiuni este manipularea obiectelor. Privite prin prisma acestei aplicații, principalele probleme pe care sistemul de vedere artificială trebuie să le rezolve sunt reconstrucția cu precizie a suprafeței obiectului vizualizat și calcularea celor mai sigure și precise configurații de prehensare a acestuia [129]. Utilizarea unor volume predefinite (universale sau particulare), care să înlocuiască reprezentarea incompletă a obiectului considerat, poate furniza informația volumetrică necesară determinării punctelor de contact dintre obiectul reconstruit și brațul mecanizat. Aspectul (înfățișarea) volumului predefinit este relevant și afectează direct numărul configurațiilor (modului) în care obiectul de interes poate fi prehensat (apucat).

În cartea de față, termenul volum este referit ca forma rigidă tridimensională a obiectului per-

---

<sup>1</sup>Eng. *Pick and Place Operations*

ceput. În acest sens, procesul de estimare a volumului unui obiect presupune reconstrucția tridimensională a formei obiectului vizualizat unidirecțional.

Indiferent de domeniul de aplicare (de ex., diagnoză [52], studierea structurilor anatomice [42], evitarea obstacolelor [87], înțelegerea complexă a scenei [155], manipularea [132] etc.), utilizarea unor *primitive* ca volume predefinite ce pot reconstrui și reprezenta suprafața obiectului considerat este o direcție de cercetare relativ nouă. Importante studii au fost întreprinse în acest sens de Marton în [68, 118, 119, 151]. Acesta propune estimarea volumelor obiectelor vizualizate utilizându-se o serie de *primitive geometrice* cu forme circulare, rectangulare sau toroidale. Principalul atu al acestei abordări este acela că nu este necesară niciun fel de informație a priori referitoare la obiectul a cărui volum urmează a fi estimat și, de asemenea, nu este nevoie de niciun proces de identificare și recunoaștere a acestuia. Volumul astfel estimat descrie o simplă dizlocare de spațiu din scenă. Motivația utilizării acestor primitive regulate este justificată de ipotezele conform cărora multiple obiecte de uz general prezintă forme cilindrice, rectangulare sau măcar conțin părți ce descriu astfel de forme. Fiind forme implicite, primitivele geometrice sunt definite utilizându-se un număr redus de parametri care, pe baza formei vizualizate a obiectului (2,5D), modelează rapid și eficient conturul primitivei. O abordare similară cu aceea a lui Marton din [118] este prezentată și în [134]. De această dată însă, obiectele vizualizate sunt identificate și descrise prin intermediul unor grafuri. Prin potrivirea grafurilor obiectelor percepute cu cele ale unor primitive geometrice predefinite se poate stabili o conexiune (similaritate) între acestea. Ulterior, primitiva a cărui graf este cel mai similar cu cel al obiectului vizualizat va fi identificată și utilizată pentru a reprezenta rigid volumul acestuia.

O abordare inovatoare, de asemenea bazată pe primitive geometrice, este prezentată și în [156]. Ceea ce este interesant în această abordare este faptul că autorii nu se limitează doar la reconstrucția volumelor obiectelor simple și regulate. Utilizându-se un proces de descompunere spațială complex, obiectul perceput este partiționat într-o serie de regiuni, de mici dimensiuni, ce pot fi descrise ulterior mult mai ușor și intuitiv prin intermediul unui număr mare de forme (primitive) parametrice (geometrice). Grupându-se toate aceste primitive, la nivel global se obține un volum compact, foarte bine definit, a cărui suprafață o aproximează fidel pe aceea vizualizată (a obiectului).

Pe lângă primitivele geometrice cu formă cilindrică, sferică sau rectangulară mai pot fi definite și numeroase alte tipuri de primitive. Utilizate cu scopul comun de a reconstrui obiecte, în [103] autorii utilizează ca primitive geometrice o serie de suprafețe generate folosind polinoame implicite. Pe de altă parte, în [169], se redefineste modul clasic de identificare al parametrilor de modelare ai primitivelor geometrice propunându-se o metodă eficientă care are ca principal scop obținerea unor noi primitive geometrice. Nu în ultimul rând, formele parametrice prezentate în lucrarea [156] reprezintă o alternativă în cadrul procesului de estimare volumetrică a obiectelor.

Principala problemă a primitivelor utilizate în lucrările citate anterior este că acestea pot aproxima doar obiecte a căror suprafață este regulată. Totuși, într-o scenă obiectele pot prezenta și forme neregulate. Spre exemplificare, o ceașcă, deși în esență poate fi reprezentată printr-un cilindru, în realitate aceasta poate prezenta numeroase curburi și deformări puternice. În contextul unei aplicații de manipulare, acest aspect afectează direct calitatea volumului obiectului estimat și implicit poate produce puncte de contact (dintre manipulator și obiectul reconstruit) eronate.

Totodată, utilizarea primitivelor geometrice prezintă o acuratețe ridicată (în cadrul unui proces de estimare volumetrică) doar dacă obiectul reconstruit prezintă ample suprafețe regulate (obiecte cu formă generică).

Cu scopul creșterii acurateței volumului reconstruit, obiectele cu formă complexă (suprafețe puternic neregulate) pot fi descompuse într-o serie de regiuni de mici dimensiuni care, la acest nivel, pot descrie o suprafață regulată. Totuși, nu în toate scenariile, această abordare este și cea mai eficientă. Zgomotele (punctele singulare), ocluziile și caracterul dispersat al reprezentării obiectului perceput pot conduce la obținerea unui volum inexact. Utilizarea obiectului ca un întreg, și nu descompus, produce un volum final cu un înțeles mai concludent decât unul format din multiple volume fuzionate. Totuși, reconstrucția volumetrică a unui obiect utilizând doar o simplă primitivă geometrică, după cum s-a spus și anterior, nu este întotdeauna cea mai bună soluție, datorită caracterului generic introdus de aceasta (suprafețe regulate introduse de forme precum cilindrii, paralelipede, sfere, etc.). Similar primitivelor geometrice, forma obiectului poate fi aproximată utilizându-se o serie de forme predefinite (modele rigide ale diferitelor obiecte). Numărul mare de înfățișări pe care le poate avea un obiect, spre exemplu diferite forme de câni, precum și numărul mare de obiecte diferite ce pot apărea într-o scenă implică definirea unei baze de date care să depoziteze toate instanțele formelor anterior menționate. Identificarea formei predefinite optime relativ la aceea vizualizată a obiectului se face prin compararea fiecărei reprezentări volumetrice (forme) din baza de date cu forma vizualizată a obiectului considerat [79]. Principala problemă într-o astfel de abordare nu este procesul de comparare a reprezentărilor similare dintre forme ci dimensiunea considerabilă a bazei de date în interiorul căreia se face căutarea.

O reducere însemnată a dimensiunii bazei de date poate fi obținută dacă în loc de multiple forme predefinite aceasta ar reține doar primitive ale unor clase de obiecte (cană, sticlă, farfurie, carte, etc.). Spre exemplificare, se consideră clasa de obiecte de tip *cană*. Pot exista în acest sens un număr relevant de câni cu diferite înfățișări (subclase). Tocmai acest număr important de înfățișări face ca dimensiunea bazei de date cu forme să aibă o dimensiune considerabilă și implicit identificarea unei anumite forme predefinite din baza de date să fie greoaie și consumatoare de timp. În schimb, lucrurile ar fi mult mai simple dacă, clasa *cană* ar fi definită doar printr-o singură primitivă cu o înfățișare generică similară cu majoritatea formelor primitivelor din aceeași clasă. Acest concept, prezentat în cartea de față, este unul nou și are ca principal scop definirea unor forme universale ce pot fi utilizate cu încredere pentru a estima volumele obiectelor reprezentate 2.5D. Noile forme introduse au fost denumite intuitiv *primitive generice*. Această abordare reprezintă, în fapt, nucleu prezentei cărți. În acest fel, numărul de primitive generice din baza de date s-ar reduce considerabil. Problema însă, în această abordare, este legată de alegerea sau generarea formei primitivei generice. Acesta nu poate fi ales la întâmplare deoarece ea poate fi foarte diferită de restul formelor predefinite din clasă. O soluție în acest sens o reprezintă determinarea formei medii identificate ca fiind suma tuturor formelor predefinite dintr-o anumită clasă. Forma medie a primitivei generice a unei clase de obiecte poate fi obținută utilizându-se o abordare ca aceea descrisă în [18]. Astfel, în final, se obține o bază de date redusă ca dimensiune și ale cărei primitive descriu volume ce pot fi utilizate cu succes în procesul de reconstrucție 3D a formelor vizualizate unidirecțional.

În continuare, pentru a crește similaritatea dintre obiectul perceput și primitiva generică co-

respondentă, se poate utiliza un proces complex de deformare. Acesta are ca principal scop particularizarea volumului primitivei generice relativ la informația percepută din scenă. Modelarea volumelor obiectelor este o practică des întâlnită mai ales în domeniul graficii asistate de calculator. Unul dintre pionierii acestei direcții de cercetare este Terzopoulos. Acesta este autorul multora dintre metodele actuale de modelare a formelor. Dintre acestea se pot menționa: modelarea hibridă a formelor implicite [171]; conturile active [99]; modelarea suprafețelor utilizând particule dinamice [167]; modelele deformabile neelastice [172] etc. Cea mai relevantă metodă a sa este aceea a conturilor active [99] care modelează un contur inițial (de ex. un cerc) utilizând o serie de forțe ce "*imping*" acest contur către informația percepută. În final, se obține un contur foarte apropiat de cel al obiectului din scenă. Pornind de la acest principiu, primitiva generică poate fi modelată astfel încât să preia cât mai multe similarități de la obiectul vizualizat.

În figura 1.1 este descris un exemplu simplu de estimare a volumului unei căni. Se poate observa cu ușurință faptul că volumul ideal al căni (v. fig. 1.1(b)) nu poate aproxima fidel silueta reală (v. fig. 1.1(a)), motiv pentru care acesta este modelat pentru a se realiza un volum final apropiat de cel real (v. fig. 1.1(d)).

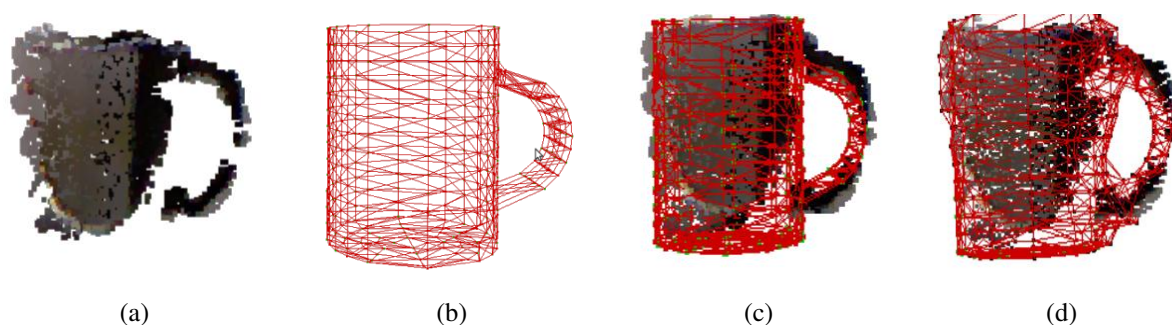


Fig. 1.1 Estimarea volumului unei căni: (a) forma percepută a căni; (b) volumul ideal al unei căni reprezentat sub forma unei rețele interconectate de triunghiuri de culoare roșie; (c) volumul căni aliniat peste forma percepută; (d) volumul modelat al căni relativ la obiectul perceput.



## 2. Analiza semantică a spațiului tridimensional

---

Achiziția informației vizuale  
Reprezentarea informației vizuale  
Manipularea informației vizuale în spațiul cartezian  
Caracteristici punctiforme 3D  
Segmentarea prin partiționare a norilor de puncte 3D  
Analiza statistică a formelor 3D  
Concluzii

---

A percepe, înțelege și reacționa în consecință, sunt o serie de acțiuni cognitive care transformă simpla mașină computerizată într-un sistem complex asemănător celui uman. Modul în care informațiile furnizate de aceste acțiuni sunt prezentate sistemului poate afecta direct reușita sarcinii impuse. Fiecare etapă a unei astfel de sarcini este importantă și oferă informații esențiale nivelurilor imediat următoare. Capitolul de față are rolul de a prezenta o serie de concepte de bază utilizate la percepția și analiza semantică a spațiului tridimensional. Sunt tratate astfel principalele tehnici de achiziție și reprezentare a informației vizuale, metode de manipulare și înțelegere a formelor tridimensionale și, nu în ultimul rând, o serie de tehnici de analiză semantică a acestora. Percepția, văzută din perspectiva unui sistem de calcul, este referită în cadrul prezentei cărți drept acțiunea prin care se vizualizează o scenă sau un obiect, și nu presupune oferirea unui sens acestei informații. Sensul a ceea ce a fost perceput este obținut în urma unei analize semantice complexe.

## 2.1 Achiziția informației vizuale

Informația vizuală, în cadrul sistemelor de vedere artificială, reprezintă variații ale unor semnale electrice provenite de la un senzor. În acest context, informațiile vizuale sunt materializate printr-o serie de imagini care reprezintă date provenite de la senzorul de percepție. Elementul de bază al unei imagini este *pixelul*. Acesta codifică intensitatea luminoasă a unui punct din spațiul sub forma unei *culori*. Matematic, o imagine poate fi reprezentată printr-o funcție  $f(x,y)$ , unde argumentele funcției reprezintă coordonatele punctelor imaginii pe cele două direcții ale spațiului  $\mathbb{R}^2$ , așa cum se poate vedea și în figura 2.1. O formă compactă de reprezentare a unei imagini digitale monocrome (gri) este dată sub forma unei matrice ale cărei elemente sunt *pixeli* [67]:

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}, \quad (2.1)$$

unde, valorile pe care le poate lua  $f(x,y)$  aparțin unui domeniu finit:

$$0 \leq f(x,y) \leq L-1, \quad (2.2)$$

iar,  $L$  reprezintă intensitatea luminoasă a unui anumit pixel.

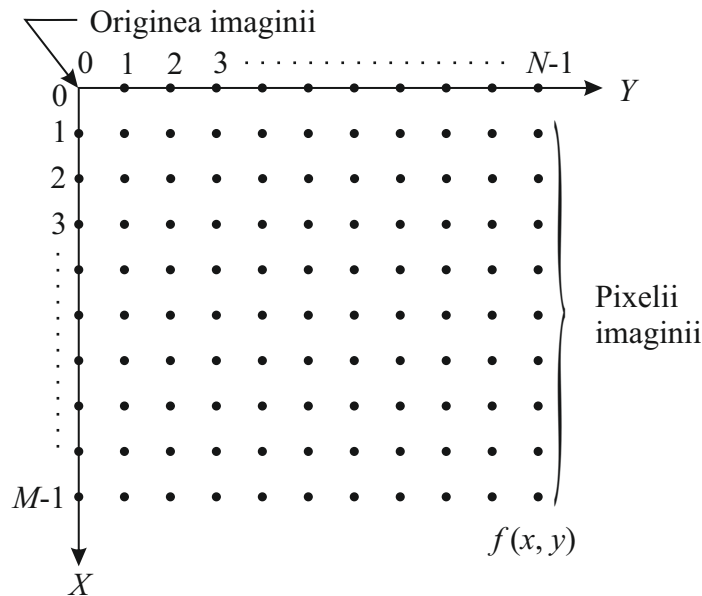


Fig. 2.1 Coordonatele pixelilor unei imagini digitale  $f(x,y)$ .

Valoarea lui  $L$  este, în general, 256 (pentru imagini gri și color). Pentru cazul imaginilor color,  $f(x,y)$  va conține cele trei culori fundamentale (roșu, albastru și verde)<sup>1</sup> și poate fi reprezentată matematic în două forme:

<sup>1</sup>Eng. *RGB* - Red Green Bblue

- prin intercalarea pixelilor, unde fiecare linie este o matrice bidimensională, cu fiecare element reprezentând o listă cu trei valori;
- prin intercalarea culorilor, în care pe fiecare linie datele sunt separate în matrice 2D, una pentru fiecare canal de culoare:

$$f_{RGB}(x,y) = \begin{bmatrix} f_R(x,y), & f_G(x,y), & f_B(x,y) \end{bmatrix}. \quad (2.3)$$

Relația exactă dintre imaginea digitală și spațiul tridimensional este determinată strict de procesul de achiziție care depinde exclusiv de tipul senzorului utilizat.

Având în vedere conceptul de imagine, formulat anterior, se pot defini următoarele tipuri de imagini, și anume:

- imagini de *intensitate*;
- imagini de *adâncime*;
- imagini *mixte*.

Imaginile de intensitate sunt imaginile fotografice în care intensitatea luminoasă percepută, prin intermediul unui senzor video, este transformată într-o valoare numerică ce poate fi identificată mai departe prin intermediul unei culori. Imaginea color, referită anterior ( $f_{RGB}(x,y)$ ), este un bun exemplu al unei astfel de imagini de intensitate. Un exemplu similar de imagine de intensitate este și imaginea monocromă.

Imaginea de adâncime este structural similară cu o imagine monocromă ( $f(x,y)$ ). Diferența este aceea că, intensitatea fiecărui punct din imagine în loc să codifice intensitate luminoasă, codifică distanța relativă dintre centrul senzorului și punctul omolog din spațiul tridimensional. Cu cât valoarea intensității unui pixel  $p$  al imaginii de adâncime este mai mare (nuanță deschisă la culoare), cu atât punctul omolog  $p$  din spațiul tridimensional se află mai aproape de centrul senzorului de percepție respectiv, o valoare mică a aceleiași intensități simbolizează o depărtare mai mare a punctului real  $p$ . Spre exemplificare, figura 2.2 prezintă grafic o imagine de intensitate, respectiv de adâncime.



(a)



(b)

Fig. 2.2 Tipuri de imagini digitale: (a) imagine de intensitate; (b) imagine de adâncime.

Imaginile mixte prezintă, pe lângă informația de intensitate (culoare), și informație de adâncime.

Matematic, acest lucru poate fi descris după cum urmează:

$$f_{RGB-D}(x,y) = \left[ f_R(x,y), f_G(x,y), f_B(x,y), f_D(x,y) \right], \quad (2.4)$$

unde,  $D$  reprezintă distanța dintre senzor și punctul real omolog pixelului de la coordonata  $(x,y)$ . Relația dintre  $D$  și coordonata  $(x,y)$  este strict legată de parametri și tipul senzorului de percepție.

În forma în care au fost achiziționate, imaginile de intensitate respectiv adâncime descriu doar puncte în spațiul  $\mathbb{R}^2$  (bidimensional). Utilizând valoarea  $D$ , pixelul din imaginea de adâncime poate fi *reproiectat* obținându-se în acest fel un punct tridimensional omolog [78]. Acesta constituie elementul de bază ce poate fi utilizat la reprezentarea informației în spațiul tridimensional. Prin reproiecția tuturor pixelilor din imaginea de adâncime, se obține o colecție de astfel de puncte tridimensionale care, văzute per ansamblu, oferă un sens spațiului înconjurător. Intuitiv, această colecție de puncte 3D este denumită *nor de puncte tridimensionale* și va fi notată în continuarea lucrării cu  $\mathcal{P}$ . Coordonatele  $\{x,y,z\}$  ale unui punct tridimensional  $\mathbf{p}_i$ ,  $i = 0 \dots N_p$ ,  $N_p$  fiind numărul total de puncte 3D din  $\mathcal{P}$ , sunt raportate la un sistem fix de coordonate care, în mod uzual, își are originea în centrul senzorului cu care s-au perceput (colectat) informațiile vizuale.

Având ca principal scop măsurarea distanțelor dintre senzor și scena vizualizată precum și transformarea acestora în puncte 3D, în literatura de specialitate se pot identifica o serie de metode ce satisfac aceste cerințe [133]. Principalele tehnici de percepție tridimensională definite în acest sens sunt:

- sistemele *ToF* (*Time of Flight*), care măsoară timpul scurs până când un semnal emis este reflectat de o anumită suprafață către un receptor;
- tehnicile de *triangulație*, care estimează adâncimea prin identificarea punctelor corespundente în imaginile percepute de către doi senzori diferiți la același moment de timp;
- sistemele cu *lumină structurată*, care estimează distanța dintre scenă și senzor prin proiectarea unui *șablon* luminos cu structură cunoscută. Relația liniară dintre grosimea dungilor proiectate și distanța reală permite estimarea cu precizie a adâncimii obiectului relativ la senzor. Datorită influențelor luminoase (în special în spectrul infraroșu) acest tip de percepție poate fi utilizat doar în scenarii cu spații închise. Această limitare a condus la o utilizare restrictivă a sistemelor cu lumină structurată în aplicațiile robotice generale. Totuși, datorită prețului redus, a ratei de achiziție bune (aproximativ 30 cadre per secundă) și preciziei bune tot mai mulți roboți mobili au început să utilizeze un astfel de senzor de percepție.

Sistemele *ToF* se mai numesc și sisteme de percepție *active* deoarece acestea folosesc, pentru estimarea distanței, o serie de semnale acustice sau luminoase. Totuși, aceste semnale nu sunt invazive și nu alterează sub nicio formă scena percepută. Cele mai întâlnite astfel de metode de percepție sunt:

- sistemele de măsurare cu laser *LMS* (*Laser Measuring Systems*). Acestea mai sunt cunoscute și sub numele de *LIDAR*;
- radarele;
- camerele *ToF*;
- sistemele de tip sonar.

Principiul metodelor enunțate este unul comun prin prisma faptului că toate se bazează pe emiterea unei serii de *unde*, luminoase sau sonore, care sunt reflectate mai apoi către un receptor. Dintre toate sistemele *ToF*, cel mai des utilizat este *LMS*. Nu doar timpul de achiziție foarte mic sau rezoluția foarte bună recomandă această metodă ci și precizia foarte mare. Cunosând viteza de propagare a unei luminoase (laser) în spațiu, și măsurând cu precizie timpul dintre emisia și recepția acesteia, distanța  $D$  poate fi estimată utilizându-se următoare relație:

$$D = \frac{c \cdot t}{2}, \quad (2.5)$$

unde,  $c$  reprezintă viteza de propagare a unei (pentru cazul unei laser aceasta reprezintă viteza luminii), iar  $t$  este timpul scurs de la emitere până la recepție. Norul de puncte astfel perceput se bucură atât de o rezoluție bună a reprezentării cât și de o acuratețe ridicată.

Una dintre cele mai utilizate tehnici de percepție este tehnica triangulației [78], iar cea mai comună tehnică de triangulație este vederea stereo. Utilizându-se această tehnică, mediul înconjurător este perceput prin intermediul a două imagini digitale  $I_s$  și  $I_d$ , ale căror centre optice  $O_s$ , respectiv  $O_d$  sunt decalate cu o distanță  $b$  numită și *distanță de bază* sau *linie de bază* [38, 78]. Preluate la același moment de timp, cele două imagini ( $I_s$  stângă, respectiv  $I_d$  dreaptă) descriu două perspective vizuale ușor diferite ale aceleiași scene. Tocmai această perspectivă de vizualizare diferită creează efectul de adâncime. Sistemul de vedere stereo, spre deosebire de cel *ToF* sau mixt, este unul pasiv deoarece procesul de achiziție al imaginilor nu implică emiterea de semnale în scena considerată. Totuși, conceptul nu este unul simplu, existând o serie de probleme care trebuie rezolvate. Prima, cunoscută și sub numele de *problema corespondenței* [176], constă în stabilirea legăturii dintre punctele din imaginea achiziționată de camera din partea stângă și punctele din imaginea achiziționată din partea dreaptă. A doua problemă, *problema reconstrucției* [176], reprezintă capacitatea de modelare a spațiului 3D având la bază informația video achiziționată de la sistemul stereo și informațiile legate de parametrii interni și externi ai camerelor (*parametrii intrinseci* și *parametrii extrinseci*) [114].

Descrierea geometrică a modelului camerei stereo poate fi urmărită în figura 2.3. Geometria unui sistem video stereo se numește *geometrie epipolară*. Elementele ce fac parte din acest ansamblu geometric sunt definite astfel:

- *epipoli*: reprezintă punctele de intersecție ale liniei ce unește centrele camerelor (*linia de bază*) cu *planurile imaginilor*, fiind simbolizați în figura 2.3 prin  $e_s$  și  $e_d$ ;
- *plan epipolar*: planul ce conține linia de bază;
- *linie epipolară*: intersecția dintre planul epipolar și planul imaginii.

Elementele de geometrie epipolară sunt conținute într-o matrice, numită *matrice fundamentală*. Această matrice are următoarea proprietate [78]: dat fiind un punct  $p$ , din spațiul 3D, care este proiectat pe cele două planuri ale imaginii, în punctele  $p_s$  respectiv  $p_d$ , este îndeplinită următoarea condiție:

$$p_d F p_s^T = 0, \quad (2.6)$$

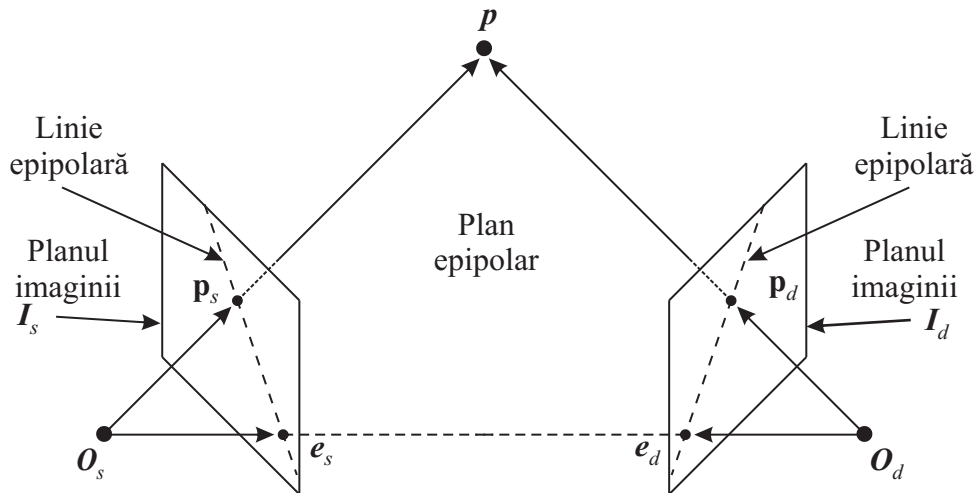


Fig. 2.3 Modelul simplificat al camerei stereo.

unde,  $F$  este matricea fundamentală, de dimensiune  $3 \times 3$  și rang 2. Literatura de specialitate propune mai multe metode pentru determinarea matricei fundamentale. Dintre acestea, cea mai cunoscută este *algoritmul celor opt puncte*<sup>2</sup> [111, 113], conform căreia, un minim de opt puncte corespondente între  $I_s$  și  $I_d$  sunt de ajuns pentru a se determina matricea fundamentală  $F$ . Avantajul major al geometriei epipolare este acela că restricționează căutarea corespondenței din imaginea din dreapta ( $I_d$ ) de-a lungul unei singure linii (cea epipolară). Spre exemplificare, se consideră dimensiunea imaginii digitale ca fiind de  $640 \times 480$  pixeli. Astfel, un pixel aparținând imaginii  $I_s$  comportă 307200 poziții candidate în imaginea  $I_d$ . Utilizând linia epipolară căutarea se reduce la doar 640 poziții candidate. Acest mod de căutare reduce semnificativ complexitatea problemei corespondențelor. Totuși, corespondențele dintre pixelii imaginilor  $I_s$  și  $I_d$  nu pot fi identificate comparând pixel cu pixel. Pentru determinarea acestora se pot folosi o serie de caracteristici cum ar fi: culoarea, nuanța, intensitatea sau informații generale obținute prin filtrarea imaginilor. În funcție de modul de determinare al corespondențelor, în literatura de specialitate sunt prezentate o serie de metode rapide și eficiente pentru stabilirea acestora. Conform autorilor lucrării [176], metodele de estimare a corespondențelor între imagini stereo pot fi împărțite în:

- *metode ce au la bază corelații*, în cazul cărora determinarea corespondențelor se aplică asupra tuturor punctelor din imaginile achiziționate. Elementele de potrivire între cele două imagini sunt ferestre de dimensiuni fixe, iar criteriul după care se stabilește similitudinea dintre imagini este o măsură a corelației dintre ferestre;
- *metode ce au la bază caracteristici*, în cazul acestor tipuri de metode se încearcă căutarea de similarități între seturi risipite de caracteristici din imagini (se utilizează proprietăți numerice sau simbolice ale acestora). Criteriul după care se realizează similitudinea este o măsură ce înglobează distanța dintre descriptorii caracteristicilor (detalii despre descriptorii caracteristici vor fi prezentate în secțiunile următoare).

Printre cele mai utilizate metode de estimare a corespondențelor se numără și metoda pe bază de corelații intitulată *potrivirea blocurilor BMA (Block-Matching Algorithm)* [75]. Metoda re-

<sup>2</sup>Eng. Eight Point Algorithm

alizează potrivirea informației din interiorul unei ferestre pătratică de dimensiune redusă ( $5 \times 5$  sau  $7 \times 7$  etc.) centrată în pixelul din imaginea stângă cu o fereastră glisantă cu dimensiune identică centrată secvențial în punctele de-a lungul unei linii epipolare din imaginea din dreapta. Măsura similarității dintre ferestre este evaluată utilizându-se *suma diferențelor absolute SAD* (*Sum of Absolute Difference*)[114]. Astfel, intensitatea fiecărui pixel din fereastra (blocul) din imaginea stângă este scăzută din intensitatea pixelului omolog din imaginea din dreapta. Suma acestor diferențe indică o măsură a similarității punctelor în care au fost centrate ferestrele. Corespondențele determinate cuprind doar punctele cu o textură puternică. O informație de culoare constantă în jurul unui anumit punct face imposibilă determinarea unei corespondențe discriminative în imaginea din dreapta. Modul de lucru al acestei abordări face ca doar suprafețele texturate să poată produce informație de adâncime.

Eliminarea corespondențelor greșite se poate realiza utilizând următoarea relație de filtrare [75]:

$$g_r = \frac{m - m_{\min}}{m_{\min}}, \quad (2.7)$$

unde,  $m$  este valoarea *SAD*, iar  $m_{\min}$  reprezintă valoarea de minim pentru *SAD*. Pentru a se putea considera existența unei corespondențe între două puncte,  $g_r$  trebuie să aibă o valoare mai mare decât o valoare de prag:

$$g_r > T_g, \quad (2.8)$$

unde,  $T_g$  este valoarea de prag.

Alegerea valorii de prag reprezintă o altă problemă importantă. O valoare mare a lui  $T_g$  poate mări considerabil zgomotul din scena reconstruită, în timp ce o valoare prea mică a lui  $T_g$  poate reduce semnificativ calitatea percepției. Totuși, există foarte multe situații în care o valoare constantă a lui  $T_g$  poate produce rezultate bune [21]. Există însă și metode adaptive pentru problema enunțată. În lucrarea [75] este prezentată o metodă prin care este posibilă adaptarea valorii de prag în funcție de un indicator  $i_m$ , care conține informații despre conectivitatea pixelilor din imaginea segmentată. Un astfel de indicator poate avea o expresie matematică de forma [75]:

$$i_m = -\log_2 p_8, \quad (2.9)$$

unde,  $p_8$  reprezintă raportul dintre numărul total de pixeli segmentați, ce au în jurul lor opt pixeli segmentați vecini, și numărul total de pixeli segmentați.

O corespondență corectă între pixelii  $p_s$  și  $p_d$  conduce la identificarea cuantumului deplasării lui  $p_d$  relativ la  $p_s$ . Diferența de poziție dintre doi pixeli în domeniul imaginii poate coda distanța relativă dintre punctul real  $\mathbf{p}$  (cel care a produs, prin proiecție, pixelii  $p_s$  și  $p_d$ ) și centrul camerei. În literatura de specialitate, această distanță poartă numele de *disparitate* (*disp*). Matematic, aceasta poate fi exprimată după cum urmează:

$$disp = p_s - p_d. \quad (2.10)$$

Cu toate acestea, căutarea similitudinilor din imaginea din partea stângă în imaginea din partea dreaptă, se poate realiza doar în interiorul unui interval definit după cum urmează [74]:

$$H = [d_{min}, d_{max}], \quad (2.11)$$

unde,  $H$  este denumit *horopter*, putând fi definit ca volumul 3D acoperit de regiunea de căutare a algoritmului *BMA* [75]. Capetele intervalului de existență pentru valoarea de horopter, sunt  $d_{min}$  și  $d_{max}$ . Acestea reprezintă valorile de minim și maxim pentru *disparitate*. În figura 2.4 este prezentată o imagine de disparitate împreună cu perechea de imagini stereo care a produs-o. Cu cât valoarea disparității este mai mare (intensitate mai luminoasă în imaginea de disparitate) cu atât punctul  $p$  se află mai aproape de centrul focal al camerei, respectiv o valoare mică a intensității simbolizează o depărtare mai mare a aceluiași punct.

Toate metodele de percepție prezentate anterior au ca rezultat un nor dens de puncte tridimensionale ce ilustrează mediul înconjurător. În cadrul prezentei cărți, norii de puncte înfățișând scene în spații deschise au fost percepute utilizându-se un sistem de vedere artificială stereo produs de compania PointGrey®, în timp ce pentru scene înfățișând spații închise a fost folosit un senzor cu lumină structurată produs de compania Microsoft®.

## 2.2 Reprezentarea informației vizuale

Informațiile vizuale, în forma în care au fost achiziționate utilizându-se una dintre tehnicile de percepție prezentate anterior, reprezintă o simplă distribuție de puncte 3D lipsită de sens. Prin operația de mapare, acestea sunt convertite pentru a se da o nouă reprezentare mediului perceput. Având ca principal subiect modalitățile de reprezentare a informației vizuale, prezenta secțiune va urmări două direcții de abordare. Prima, oferă o vedere de ansamblu asupra modurilor prin care informația vizuală poate reprezenta suprafețe în spațiul Euclidian, respectiv, cea de-a doua, abordează subiectul dintr-o perspectivă semantică în care informația vizuală utilizată ilustrează obiecte sau alte entități ce pot duce la îndeplinirea cu succes a unei anumite misiuni.

### 2.2.1 Reprezentarea suprafețelor în spațiul tridimensional

O suprafață tridimensională este considerată a fi o mulțime de puncte tridimensionale distribuite uniform sau neuniform în spațiu. Așa cum a fost prezentat în secțiunea anterioară, punctul este cel mai simplu element ce poate descrie o informație în spațiul  $n$ -dimensional. Complexitatea reprezentării unui punct  $p$  în spațiu este dată de proprietățile ce îl definesc. O primă proprietate fundamentală a acestuia este aceea *geometrică* (spațială). Considerând spațiul Euclidian  $\mathbb{R}^3$ , punctul  $p$  este definit utilizând cele trei coordonate carteziane  $(x, y, z)$ . În acest sens, punctul devine 3-dimensional. Cea de-a doua proprietate de bază o reprezintă *culoarea*. Asemenea pixelului utilizat la redare imaginilor, culoarea punctul  $p$  poate fi reprezentată prin cele trei culori fundamentale (roșu, verde și albastru). Luând în considerare și această proprietate, dimensiunea lui  $p$  devine 6. Sub această formă punctul 3D începe să capete un sens particular.

Pe lângă proprietățile fundamentale descrise, se pot defini și o serie de proprietăți adiacente. Spre exemplu, în procesul de navigare, este imperativ necesar să se cunoască distanța relativă

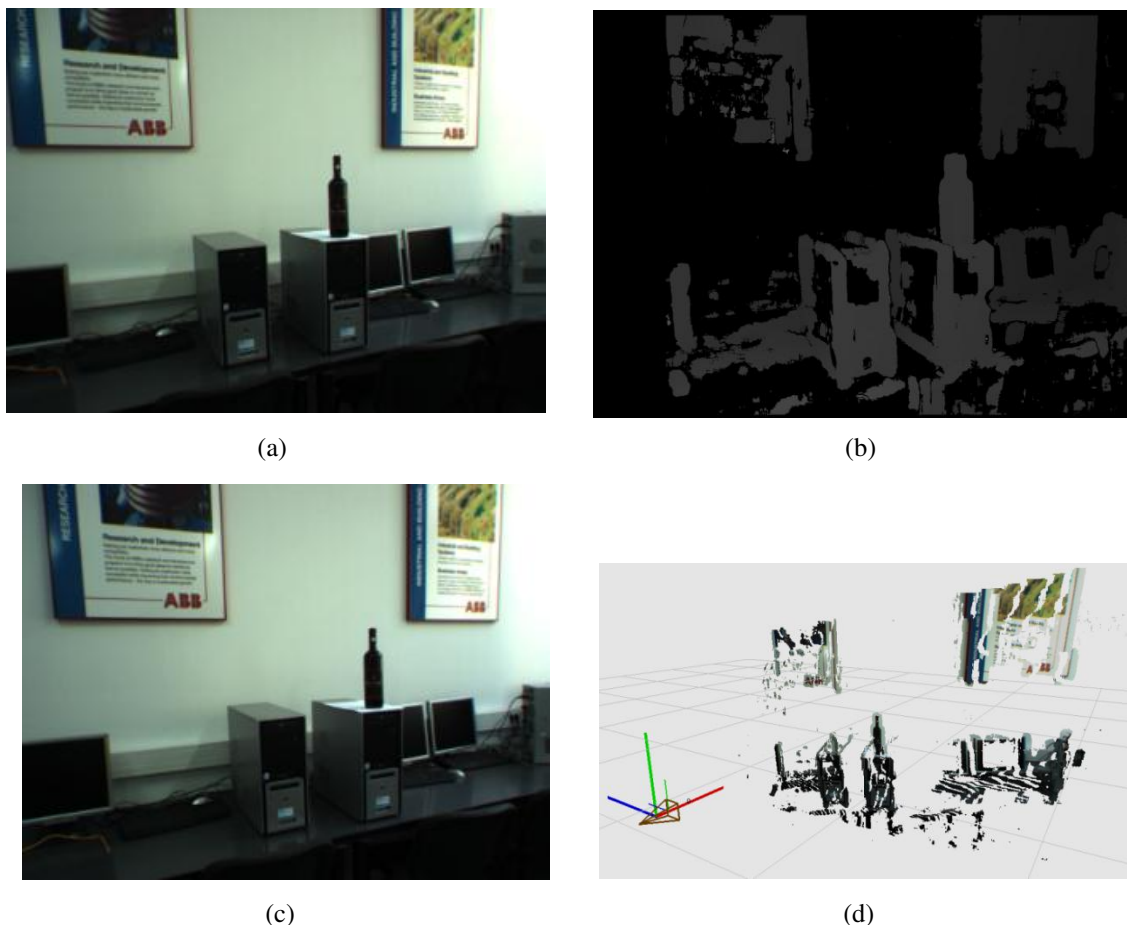


Fig. 2.4 Percepția mediului înconjurător utilizând un sistem de vedere artificială stereo: (a) imaginea vizualizată de camera din stânga; (b) imaginea de disparitate rezultată; (c) imaginea vizualizată de camera din dreapta; (d) norul de puncte 3D obținut prin repoiecția imaginii de disparitate.

dintre punctele 3D din scenă și punctul de vizualizare din care privește senzorul de percepție. Astfel, o proprietate adiacentă a lui  $p$  poate fi reprezentată de distanța  $d$  dintre acesta și senzorul de percepție. În continuare, o altă proprietate importantă este normala la suprafață ( $\vec{n}$ ). Utilizând informația geometrică vecină (punctele vecine), aceasta poate să ofere informații relevante despre orientarea punctului  $p$  în spațiu. În forma lui singulară, înțelesul unui punct poate fi limitat, însă prin alăturarea acestuia la un anumit grup de puncte vecine, capabilitatea de înțelegere a acestuia crește considerabil. În acest sens, punctul poate primi o etichetă  $e$  prin care se definește gradul de apartenență al acestuia la un anumit grup compact de puncte<sup>3</sup>. Această etichetă poate fi obținută în urma unui proces complex de partiționare a spațiului ce va fi detaliat într-o secțiune următoare. Concluzionând, complexitatea punctului  $p$  poate avea următoarea dimensiune: 3 dimensiuni definite de proprietatea geometrică la care se adaugă alte 3 dimensiuni introduse de cele trei culori fundamentale plus 3 dimensiuni definite de normala  $\vec{n}$  și alte 2 dimensiuni provenite de la eticheta  $e$  și distanța  $d$ . Se obține astfel un total de 11 dimensiuni caracteristice. Un nor de puncte 3D cu o astfel de complexitate poate fi utilizat eficient pentru analiza semantică a mediului perceput. Spre exemplificare, în figura 2.5 este prezentat un nor de puncte ce înfățișează o serie de obiecte situate pe o masă (un scenariu uzual). Scena a fost percepută utilizându-se, așa cum s-a mai prezentat,

<sup>3</sup>Eng. *Cluster*

un senzor cu lumină structurată produs de compania Microsoft®. În cadrul aceleiași figuri pot fi observate principalele proprietăți descrise anterior.

Din punctul de vedere matematic, un nor de puncte  $\mathcal{P}$  poate fi reprezentat printr-o matrice  $N_p \times l$ , în care numărul de linii este definit de numărul total de puncte  $N_p$  utilizat la reprezentarea lui  $\mathcal{P}$ , iar numărul de coloane este dat de numărul de proprietăți (caracteristici) pe care le poate avea un punct. Considerând aspectele menționate anterior, matricea norului de puncte poate fi descrisă după cum urmează:

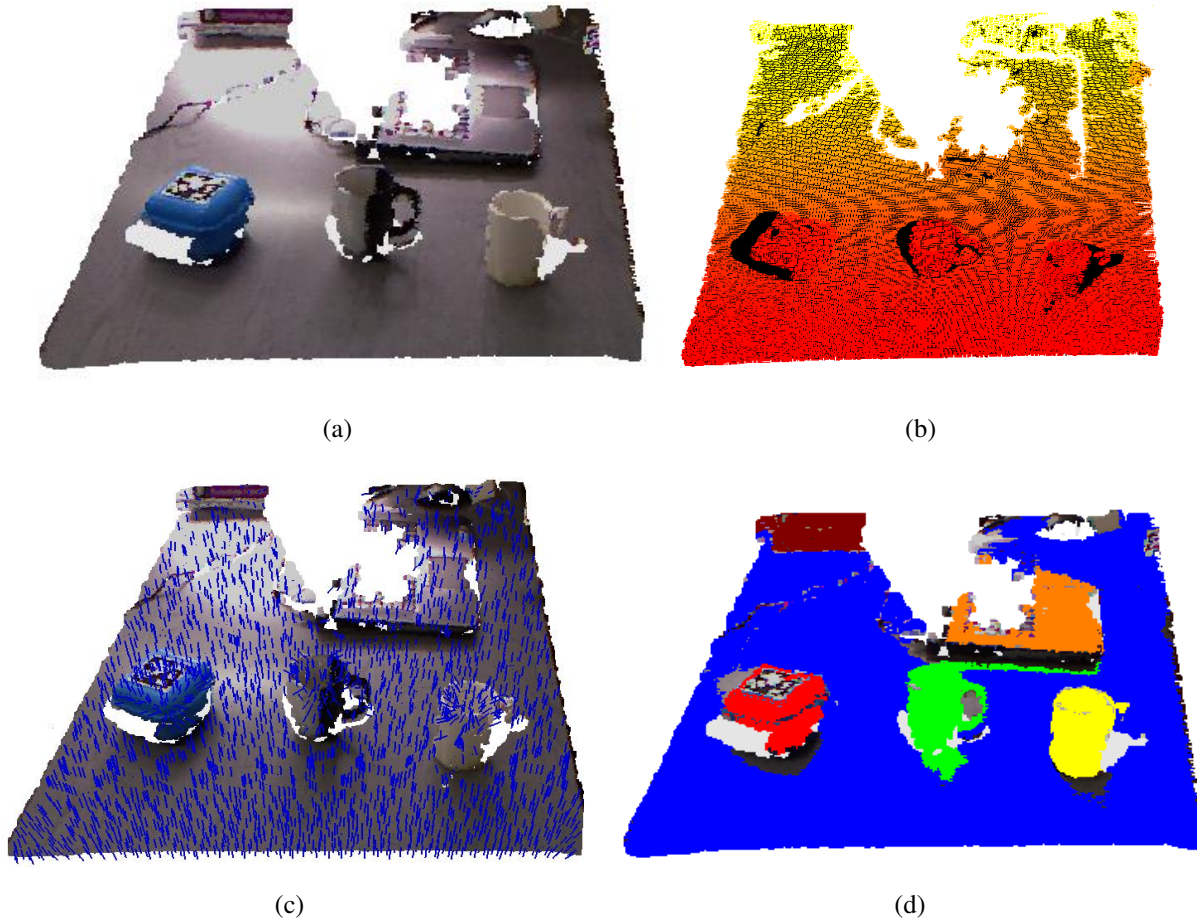


Fig. 2.5 Proprietățile punctelor  $n$ -dimensionale; (a) nor de puncte înfățișând proprietatea geometrică și de culoare; (b) nor de puncte a căror culoare ilustrează distanța relativă față de senzorul de percepție (cu culoare roșie sunt ilustrate punctele cele mai apropiate senzor); (c) normalele la suprafață în diferite puncte ale norului (cu linii de culoare albastră sunt ilustrate direcțiile normalelor); (d) gruparea punctelor norului în clase de apartenență.

$$\mathcal{P} = \begin{bmatrix} x_1 & y_1 & z_1 & r_1 & g_1 & b_1 & \mathbf{n}_{x_1} & \mathbf{n}_{y_1} & \mathbf{n}_{z_1} & e_1 & d_1 \\ x_2 & y_2 & z_2 & r_2 & g_2 & b_2 & \mathbf{n}_{x_2} & \mathbf{n}_{y_2} & \mathbf{n}_{z_2} & e_2 & d_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_{N_p} & y_{N_p} & z_{N_p} & r_{N_p} & g_{N_p} & b_{N_p} & \mathbf{n}_{x_{N_p}} & \mathbf{n}_{y_{N_p}} & \mathbf{n}_{z_{N_p}} & e_{N_p} & d_{N_p} \end{bmatrix}. \quad (2.12)$$

Coordonatele fiecărui punct sunt raportate la un sistem fix de coordonate care, în mod uzual, își are originea în centrul senzorului cu care s-au perceput datele.

După cum se poate observa în figura 2.5(a) și din relația 2.12, datele utilizate la reprezentarea norului de puncte prezintă un sens și ele pot fi utilizate, în continuare, pentru a analiza, cu scopul înțelegerii, scena. Reprezentarea suprafețelor prin intermediul unui nor de puncte nu este însă întotdeauna o soluție. Spre exemplificare, un punct oarecare  $p_i \notin \mathcal{P}$  poate aparține suprafeței descrise de norul de puncte  $\mathcal{P}$  doar dacă, coordonatele carteziane ale acestuia coincid cu cele ale unui punct din  $\mathcal{P}$ . Șansa ca acest lucru să se întâmple este foarte mică cu toate că, punctul oarecare considerat se află totuși pe suprafața descrisă de  $\mathcal{P}$ . O formă propice de reprezentare a norului de puncte, în acest sens, este prin utilizarea unei rețele interconectate de poligoane<sup>4</sup> [174]. Acest tip de reprezentare poate fi construită utilizându-se punctele norului perceput. Un poligon descrie o suprafață definită prin minim 3 puncte necoliniare. Acestea, în esență, definesc o formă geometrică triunghiulară. O suprafață astfel reprezentată este cunoscută în literatura de specialitate sub numele de rețea triunghiulară de puncte<sup>5</sup>. Nodurile poligoanelor sunt chiar punctele norului  $\mathcal{P}$ . Interconectarea acestora se realizează utilizând o serie de algoritmi asemănători celor prezentați în [54]. Un dezavantaj al acestui tip de reprezentare este timpul de calcul ridicat. Mai mult, dacă distribuția punctelor norului  $\mathcal{P}$  se modifică, întreaga rețea de triunghiuri trebuie generată din nou. Pe de altă parte, avantajul acestui tip de reprezentare este descrierea exhaustivă a unei suprafețe. Practic, nu mai există acele *goluri* dintre punctele norului  $\mathcal{P}$ . O astfel de rețea triunghiulară de puncte poate fi observată în figura 2.6(b).

Un alt mod inovativ de a reprezenta un nor de puncte poate fi obținut utilizându-se o serie de pixeli volumetrici cu formă cubică<sup>6</sup>. Astfel de entități pot fi obținute utilizând un proces de descompunere al spațiului care permite partiționarea norului  $\mathcal{P}$  într-o serie de volume cubice identice ca dimensiune. Astfel, punctele din interiorul unei anumite regiuni partiționate sunt echivalente printr-un simplu volum cubic care are rolul de a reprezenta brut regiunea respectivă. O astfel de reprezentare reduce considerabil numărul de elemente utilizate pentru a descrie o anumită suprafață. Avantajul major oferit de o astfel de abordare poate fi privit din prisma necesității identificării punctelor vecine. Dacă pentru un punct  $p_i \in \mathcal{P}$  identificarea vecinilor presupune compararea acestuia cu fiecare punct din  $\mathcal{P}$  (proces de o complexitate relativ mare), identificarea vecinilor prin metoda volumetrică implică doar identificarea celei mai apropiate forme cubice care va conține cu siguranță vecinii punctul considerat  $p_i$ . Principalele metode de descompunere a spațiului sunt *octree* [23] și *bd-tree* [41]. Dintre acestea, cea mai utilizată metodă de descompunere a spațiului este metoda *octree*. Conform acesteia, spațiul inițial definit de suprafața considerată este divizat în opt regiuni rectangulare egale. În continuare, fiecare regiune astfel obținută este divizată din nou pentru a se obține alte 8 regiuni rectangulare derivate din cea anterioară. Divizarea se oprește în momentul în care regiunea rectangulară descrie volumul minim stabilit de utilizator. Pornind de la structura definită anterior, poate fi identificată rapid, o anumită regiune prin parcurgerea arborelui de divizare. O astfel de partiționare a norului de puncte  $\mathcal{P}$  poate fi observată în figura 2.6(a). Pe lângă volumele rectangulare generate, figura prezintă și punctele 3D din interiorul acestora.

<sup>4</sup>Eng. *Polygon Mesh*

<sup>5</sup>Eng. *Triangle Mesh*

<sup>6</sup>Eng. *Voxels*

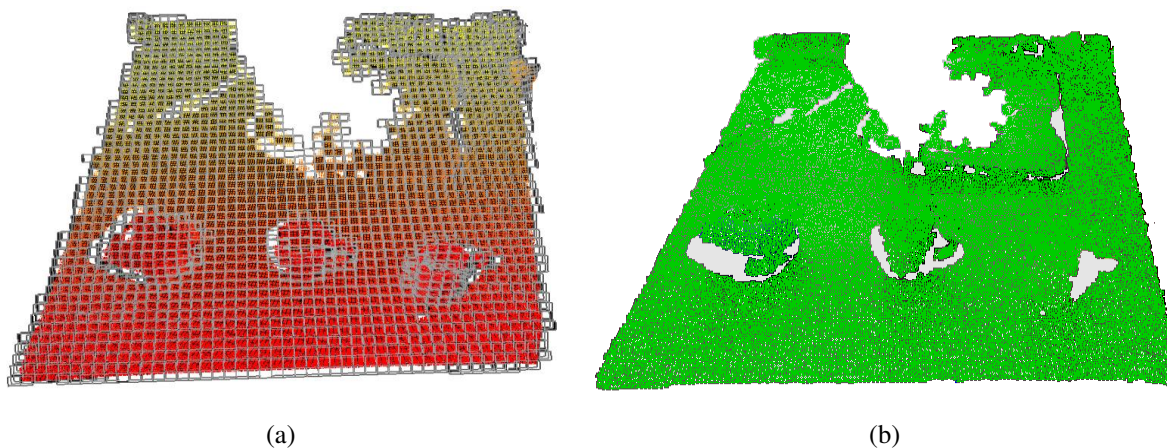


Fig. 2.6 Alte metode de reprezentare a suprafețelor; (a) utilizând volume rectangulare (*octree*); (b) utilizând o rețea triangulară de puncte (cu culoare verde sunt reprezentate laturile triunghiurilor definite de punctele din norul  $\mathcal{P}$ ).

### 2.2.2 Reprezentarea formelor în spațiul tridimensional

După prezentarea, în secțiunile anterioare, a modurilor de reprezentare a informației vizuale, în această secțiune atenția se va deplasa asupra descrierii modalităților de reprezentare al unui nor de puncte în scopul compunerii unei forme. În funcție de distribuția lor în spațiu, punctele unui nor pot căpăta un înțeles, definind astfel o formă a unui obiect. Modul de reprezentare al acestor puncte precum și tehnicile prin care ele pot fi conectate (linii, triunghiuri, suprafețe, curbe etc.), afectează direct capacitatea de înțelegere a formei considerate. Un prim criteriu de clasificare al formelor este după tipul structurii globale pe care o descriu. Aspectul global este de fapt ceea ce conferă unicitate unei forme și în același timp permite identificarea obiectului vizualizat. În acest sens, se pot contura două tipuri de forme:

- *statice*, a căror structură rămâne rigidă în timp;
- *dinamice*, a căror structură variază în timp (formele articulate sunt un bun exemplu de astfel de forme dinamice).

#### 2.2.2.1 Reprezentarea formelor statice

Formele statice compun majoritatea obiectelor din mediul înconjurător. Prin definiție, distanțele dintre punctele ce reprezintă o formă statică trebuie să fie constante în orice moment și indiferent de tipul de forțe ce acționează asupra lor. În contextul reprezentării formelor statice, se pot defini două noi tipuri de reprezentări și anume:

- reprezentarea bazată pe puncte tridimensionale;
- reprezentarea bazată pe diferite perspective de vizualizare.

#### *Reprezentarea bazată pe puncte tridimensionale*

Asemănător modurilor de reprezentare a suprafețelor descrise în secțiunea anterioară, formele, la rândul lor, pot fi descrise prin: *puncte* tridimensionale, *suprafețe* sau *volume*, cu mențiunea că, în această situație, datele (puncte 3D, pixeli volumetrici, etc.) ce descriu forma considerată prezintă un sens comun. Un nor de puncte 3D ce definește o structură distinctă poate, în esență, să

înfățișeze modelul tridimensional al unei forme. Având un înțeles particular, altul decât cel de a reprezenta simple suprafețe, norul de puncte ce descrie o formă va fi denumit în continuare *nor de puncte obiect* și va fi notat cu  $\mathcal{O}$ . Punctele norului obiect prezintă aceleași proprietăți (geometrice, de culoare, de distanță sau normale la suprafață) ca acelea ale unui nor de puncte uzual, singura diferență fiind aceea că punctele norului  $\mathcal{O}$  pot avea doar o singură etichetă, cea a obiectului căruia îi aparțin. O altă caracteristică importantă este aceea că  $\mathcal{O}$  nu poate fi partiționat pentru a se defini alte forme. În figura 2.7(a) poate fi observat un nor de puncte obiect ce înfățișează o cană.

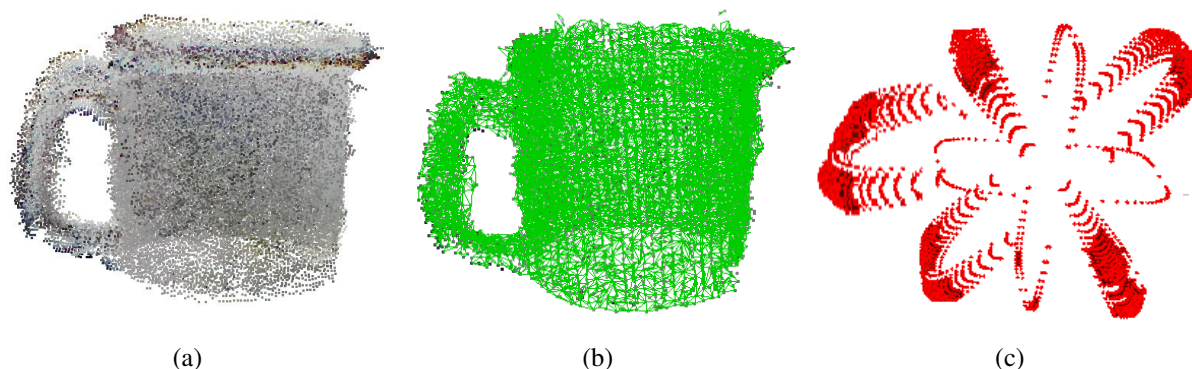


Fig. 2.7 Diferite moduri de reprezentare a formei unei căni: (a) utilizând un nor de puncte 3D; (b) utilizând o rețea triangulară (reprezentată prin linii de culoare verde); (c) utilizând forme parametrice.

Totuși, reprezentarea unei forme printr-un nor de puncte nu este suficientă, în special, în aplicații precum prehensarea sau manipularea formelor unde punctele de prehensare trebuiesc determinate cu precizie pe învelișul exterior al acestora. O alternativă la această problemă este reprezentarea formelor utilizând suprafețe. În funcție de modul în care aceste suprafețe sunt approximate, se pot defini trei metode de reprezentare:

- bazată pe rețele interconectate de triunghiuri;
- bazată pe modele parametrice;
- bazată pe forme implicite.

Reprezentarea formelor utilizând o rețea interconectată de poligoane este o tehnică de reprezentare des întâlnită în modelarea asistată de calculator a obiectelor [54]. Punctele ce descriu forma sunt unite în grupuri de minim trei pentru a defini o serie de mici suprafețe. Văzută per ansamblu, forma astfel reprezentată descrie un înveliș exterior neted, lipsit de *goluri* de informație și mult simplificat în raport cu reprezentarea prin puncte 3D. Reprezentarea unei forme utilizând o astfel de rețea triangulară poate fi observată în figura 2.7(b)

Pe de altă parte, în scopul reprezentării suprafețelor netede și mult simplificate, formele parametrice generează învelișul exterior al formei considerate utilizând o formulare matematică de tipul  $\mathbf{p} = F(\vec{p})$ , unde  $\mathbf{p}$  este un punct tridimensional generat, iar  $\vec{p}$  este vectorul parametrilor de modelare a suprafeței formei. Spre deosebire de metoda prezentată mai sus, în care punctele percepute erau cele care defineau forma considerată, de această dată forma reprezentată este definită, utilizând forme parametrice, prin intermediul unui nou set de puncte 3D. În acest caz, punctele percepute sunt folosite doar pentru a se determina parametrii optimi de modelare (generare) a noii forme parametrice. Prin definirea unui număr adecvat de parametri se poate obține o formă com-

plexă similară cu cea percepută. Un exemplu simplu de astfel de forme parametrice este reprezentat de modelele supercuadrice [9]. Acestea descriu o gamă largă de forme geometrice cum ar fi: cuburi, sfere, elipse, toroizi etc. Utilizând acest mod de reprezentare parametric, în figura 2.7(c) se poate observa reprezentarea parametrică a formei unei căni. Modul lor flexibil de definire face ca aceste modele să fie utilizate cu succes și la aproximarea volumelor formelor vizualizate. Nu se va intra mai mult în detalii în acest sens deoarece acest tip de modele va constitui subiectul unuia dintre capitolele următoare. Similar cu formele parametrice, modelele implicite pot fi generate utilizând o formulare matematică de tipul  $F(\mathbf{p}) = 0$  [66]. Din nou, punctele utilizate la reprezentarea formei sunt generate automat utilizându-se informația percepută a formei. Datorită flexibilității lor, formele supercuadrice pot fi formulate și într-o manieră implicită. Un dezavantaj al acestei abordări este faptul că noile puncte tridimensionale nu pot fi generate într-o ordine consecutivă. Pe de altă parte, avantajul acestui tip de reprezentare este că oferă o metodă simplă și rapidă pentru a se determina dacă un punct 3D se află sau nu pe suprafața nou generată.

Pe lângă metodele de reprezentare prin puncte tridimensionale, suprafețe poligonale sau implicite, o formă mai poate fi descrisă și prin intermediul unei reuniuni de volume. În funcție de tipul volumului descris, se pot defini următoarele tipuri de reprezentări volumetrice: reprezentarea bazată pe pixeli volumetrici, reprezentarea de tip *octree* [26]; reprezentarea prin partiționarea binară a spațiului; reprezentarea prin reuniunea unei sume de forme geometrice predefinite (sfere, cilindrii etc.).

Un pixel volumetric este cea mai simplă și cea mai mică formă de subdivizare a spațiului tridimensional. Ei sunt utilizați, cu precădere, în domeniul medical la reprezentarea volumelor organelor umane și în domeniul roboticii la estimarea volumelor obiectelor vizualizate unidirecțional (v. fig. 2.8(a)) [30, 187]. Poziția în spațiu a unui pixel volumetric nu este explicit definită. Ea se calculează relativ la pozițiile pixelilor volumetrici vecini.

Similar cu reprezentarea prin pixeli volumetrici, volumul unui obiect mai poate fi aproximat și printr-o structură de tip *octree*. Această structură ierarhică de date se formează prin divizarea recursivă a spațiului cubic în mai multe volume cu dimensiune redusă (v. fig. 2.8(b)) [23, 48]. Fiecare nod al structurii de date conține exact opt copii. Numărul de subdiviziuni ale spațiului determină acuratețea finală a reprezentării. Datorită structurii ei ierarhice, identificarea unui anumit punct sau a unei anumite regiuni a formei considerate se face foarte rapid și eficient. Reprezentarea prin *octree* este utilizată des în grafica 3D computerizată și la motoarele jocurilor tridimensionale. Pe de altă parte, partiționarea binară a spațiului tridimensional *BSP* (Binary Space Partitioning) oferă o serie de avantaje în raport cu reprezentarea de tip *octree* (v. fig. 2.8(c)). Pe lângă structura de căutare eficientă, *BSP* permite ca o singură reprezentare să înglobeze mai mult decât un singur obiect [89, 105]. Menținând același mod de construcție ierarhic, metoda geometriei constructive a rigidelor utilizează primitive geometrice multiple care, prin însumare creează o reprezentare globală foarte apropiată de aceea a formei considerate (v. fig. 2.8(d)). Prin conectarea unor sfere, prisme sau paralelipiede dreptunghice se poate obține un model compact ce poate fi folosit pentru a se reprezenta forma unui anumit obiect. Operația de concatenare a primitivelor geometrice este de tip boolean.

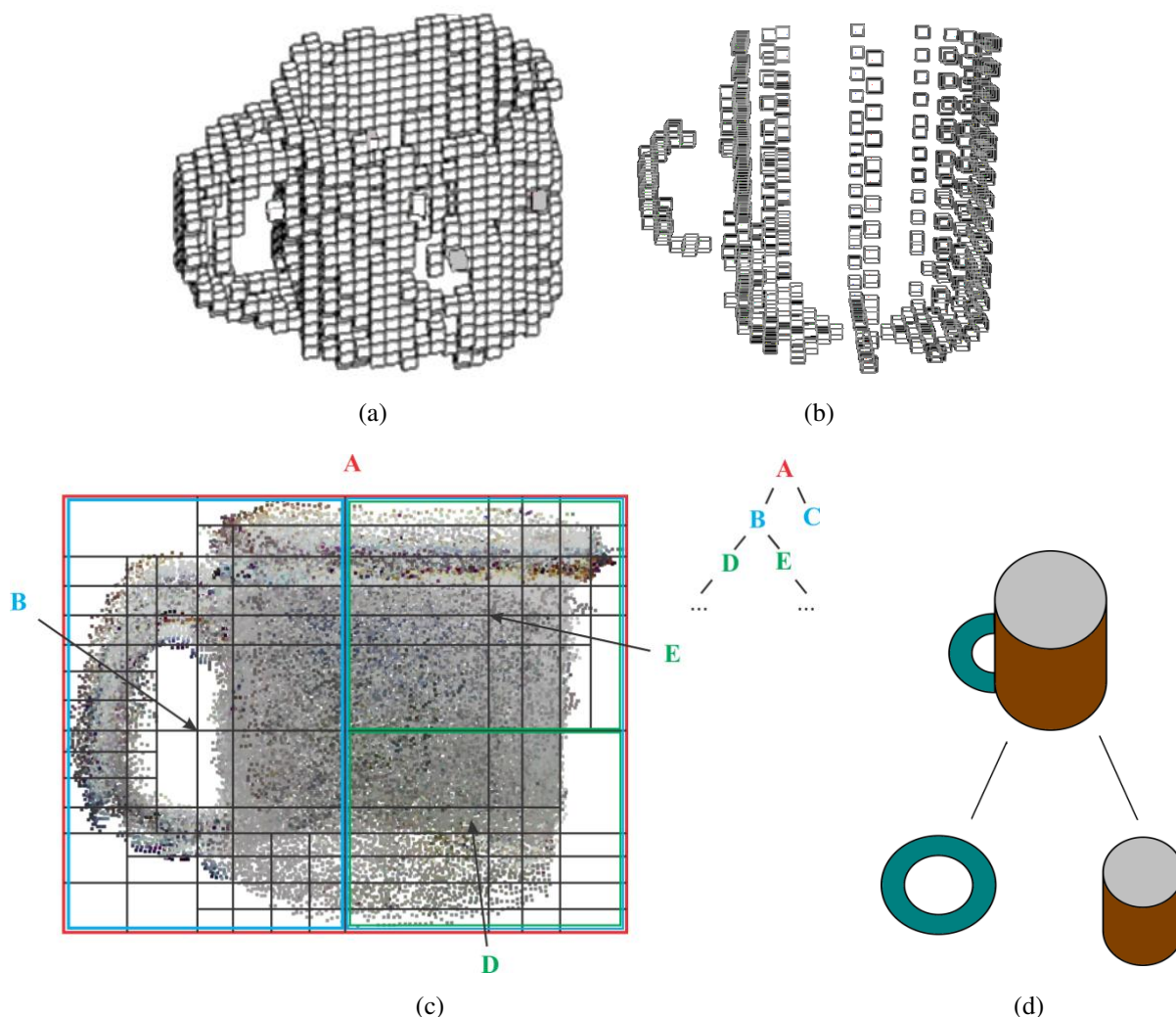


Fig. 2.8 Diferite tipuri de reprezentări volumetrice: (a) voxeli; (b) octree; (c) partiționarea binară a spațiului; (d) modele geometrice constructive.

### Reprezentarea bazată pe diferite perspective de vizualizare

Majoritatea senzorilor de percepție au o vedere unidirecțională. Astfel, o formă vizualizată poate fi reprezentată doar parțial, structura globală a acesteia fiind incompletă. Totuși, perspective multiple ale obiectului vizualizat pot fi aliniate și contopite pentru a obține o reprezentare rigidă a formei considerate [60]. Astfel, *silueta* și *aspectul* unei forme sunt singurele proprietăți ce pot fi utilizate pentru a o reprezenta.

*Silueta* descrie granițele formei vizualizate atunci când aceasta este privită dintr-un anumit punct. Reprezentarea tridimensională a unei forme se poate obține prin "*lipirea*" (concatenarea) diferitelor perspective ale sale când acestea sunt privite din diferite direcții. Abordarea este caracterizată de eficiență atât din punct de vedere temporal cât și al economiei de spațiu. Totodată, silueta unei forme poate permite identificarea obiectului pe care îl reprezintă prin compararea fiecărei perspective vizualizate cu o serie de perspective ale unor obiecte diferite depozitate într-o bază de date [28, 95]. Spre exemplificare, în figura 2.9(a) pot fi observate diferite siluete utilizate la reprezentarea unei căni. Acestea sunt proiecții în plan ale formei unei căni. Prin determinarea corespondențelor dintre diferitele siluete și "*lipirea*" acestora utilizând o serie de transformări de

similaritate, se poate obține o reprezentare rigidă a formei astfel vizualizate.

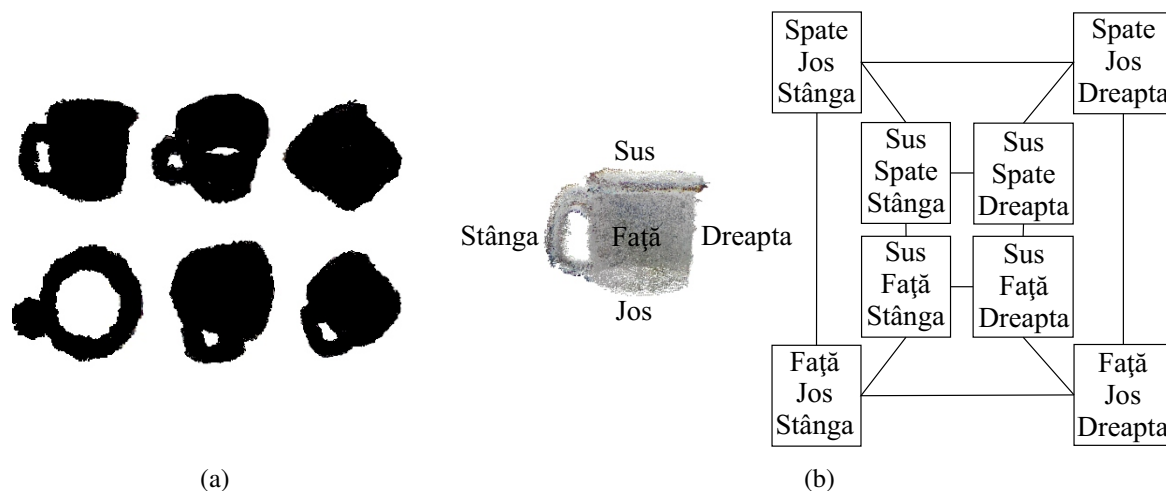


Fig. 2.9 Perspective vizuale ale unor modele: (a) perspective bazate pe siluete; (b) graf de aspecte.

Pe de altă parte, *aspectul* unei forme, identifică siluetele diferite pe care le poate lua o formă atunci când este văzută din perspective diferite. Mai multe aspecte pot fi grupate pentru a forma un graf de aspecte. Acestea pot fi utilizate, mai departe, pentru a se reprezenta în spațiu forma considerată [40]. Un exemplu de reprezentare bazată pe graful de aspecte poate fi observat în figura 2.9(b)

#### 2.2.2.2 Reprezentarea formelor dinamice

Spre deosebire de formele *statice*, formele cu aspect *dinamic* își pot schimba înfățișarea constrânse de o serie de forțe. Formele dinamice sunt utilizate cu precădere atât în robotică la estimarea volumelor obiectelor ce urmează a fi prehensate cât și în medicină la reconstrucția organelor umane. În categoria formelor dinamice se pot include contururile active [99], modele volumetrice deformabile [22, 159] și modelele de tip balon [34]. Un contur activ este o curbă care, plasată în interiorul unei imagini, se poate deplasa singură pentru a lua o formă particulară. Această curbă poate avea o reprezentare liniară (curbă deschisă) sau, poate fi descrisă prin intermediul unei suprafețe (curbă închisă). Legea prin care conturul activ se deformează are la bază un set de energii (detalii în acest sens vor fi oferite în capitolele următoare) a căror quantum este determinat utilizându-se o serie de factorii externi (imagine sau nor de puncte) și interni (lungime, consistența, etc.). Pe de altă parte, volumele deformabile capturează mișcarea suprafeței dinamice utilizând elemente volumetrice tetraedrice [159]. Comparativ cu metoda contururilor active, complexitatea atinsă de suprafața volumelor deformabile este net superioară. Avantajul în acest sens este dat de elasticitatea sporită a suprafeței deformabile. Dezavantajul este acela că pentru a se atinge această performanță este nevoie de un număr considerabil de deformări repetate. Numărul mare al acestor deformări determină un timp de procesare ridicat.

Combinând cele două moduri de reprezentare descrise se poate obține un nou tip de reprezentare numită și reprezentare balon. Denumirea provine de la faptul că orice formă poate fi aproximată inițial cu un balon. Considerând principiile de elasticitate ale acestuia, forma finală a

obiectului considerat se obține prin umflarea respectiv dezumflarea balonului. Cu scop ilustrativ, în figura 2.10 pot fi observate o serie de reprezentări construite pe baza metodelor prezentate.

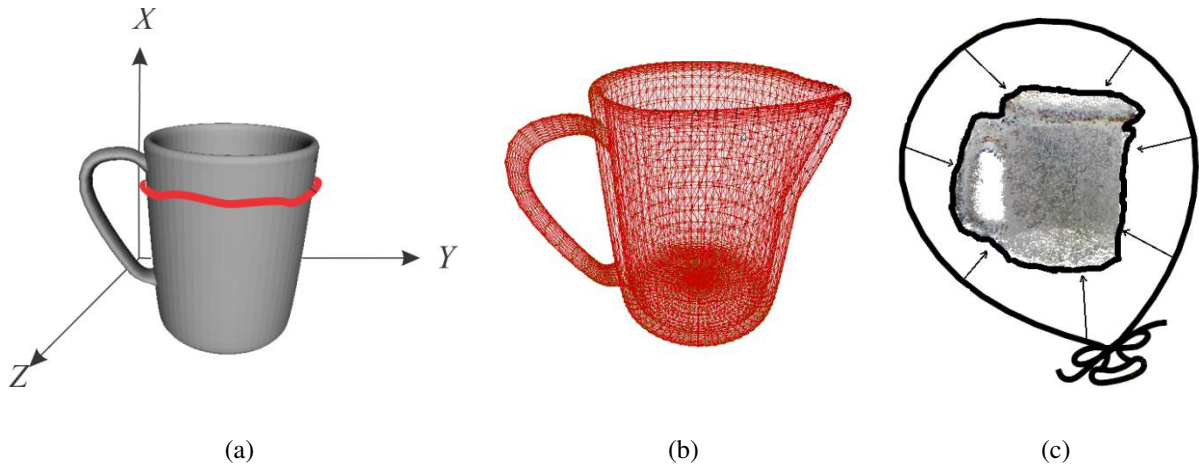


Fig. 2.10 Diferite reprezentări ale modelelor dinamice: (a) contur activ; (b) volum deformabil; (c) reprezentare de tip balon.

## 2.3 Manipularea informației vizuale în spațiul cartezian

În marea lor majoritate, formele tridimensionale sunt reprezentate în poziții și orientări arbitrare. Pentru a putea fi procesate și analizate corespunzător, acestea trebuie aduse într-un sistem de coordonate comun. Parametri ce permit manipularea unei forme în spațiul tridimensional pot fi considerați *gradele de libertate*.

### 2.3.1 Transformarea de similaritate

Se numește transformare de similaritate  $\mathcal{T}_{t_x, t_y, t_z, s, \mathbf{R}}(\mathcal{Q})$ , operația prin care o formă  $\mathcal{Q}$  (reprezentată prin intermediul unui nor de puncte) este manipulată în spațiul cartezian pentru a fi aliniată la o formă de referință  $\mathcal{P}$  (de asemenea reprezentată prin intermediul unui nor de puncte). Procesul de manipulare al formei  $\mathcal{Q}$  implică, în acest sens, utilizarea unei transformări de similaritate cu 7 grade de libertate după cum urmează:

- 1 grad de libertate introdus de factorul de scalare;
- 3 grade de libertate introduse de operația de translație;
- 3 grade de libertate introduse de operația de rotație.

Matematic, operația de transformare de similaritate a formei  $\mathcal{Q}$  relativ la forma  $\mathcal{P}$  poate fi scrisă astfel:

$$\mathcal{Q}_{transformat} \mapsto s \cdot \mathbf{R}(\mathcal{Q} - \mathbf{t}), \quad (2.13)$$

unde,  $\mathcal{Q}_{transformat}$  este forma ce are aplicată asupra ei transformarea de similaritate,  $\mathbf{t}$  reprezintă translația în spațiul  $\mathcal{R}^3$ ,  $\mathbf{R}$  este matricea de rotație având dimensiunea de  $3 \times 3$  și reprezintă rotația lui  $\mathcal{Q}$  relativ la  $\mathcal{P}$ , iar  $s$  este factorul de scalare care aduce forma  $\mathcal{Q}$  la aceeași mărime spațială cu aceea a formei  $\mathcal{P}$ . Relația 2.13 a fost formulată astfel încât transformarea de similaritate să poată fi aplicată direct pe reprezentarea prin puncte tridimensionale a formei  $\mathcal{Q}$ . Scopul acestei secțiuni

este acela de a prezenta modul în care o formă poate fi scalată, traslatată și rotită în spațiu.

### 2.3.1.1 Scalarea unei suprafețe

Factorul de scalare  $s$  are rolul de a aduce dimensiunea formei considerate  $\mathcal{Q}$  la o dimensiune similară cu aceea dată de forma de referință  $\mathcal{P}$ . Principial, dimensiunea unei forme poate fi determinată prin circumscrierea acesteia într-o sferă a cărei rază poate fi utilizată drept măsură a dimensiunii formei. Forma se consideră circumscrisă optim în momentul în care toate punctele acesteia se află în interiorul sferei. Cea mai mică rază care satisface condiția menționată este considerată a descrie dimensiunea formei studiate. Utilizând aceeași abordare, se determină și dimensiunea formei de referință ( $\mathcal{P}$ ). Raportul dintre raza formei de referință  $\mathcal{P}$  și aceea a formei  $\mathcal{Q}$  determină valoarea factorului de scalare [31]. Matematic, acesta poate fi calculat după cum urmează:

$$s = \frac{r_{\mathcal{P}}}{r_{\mathcal{Q}}}. \quad (2.14)$$

Astfel, redimensionarea lui  $\mathcal{Q}$  relativ la  $\mathcal{P}$  se realizează astfel:

$$\mathcal{Q}_{scalat} = s \cdot \mathcal{Q}. \quad (2.15)$$

### 2.3.1.2 Traslatarea unei suprafețe

Cel mai comun mod de a identifica poziția unei forme în spațiu este utilizând centrul de greutate  $\mu$ , al acesteia. Matematic, centrul de greutate al unei forme poate fi exprimat astfel:

$$\mu_{\mathcal{Q}} = \frac{1}{N_{\mathcal{Q}}} \cdot \sum_{i=1}^{N_{\mathcal{Q}}} \mathbf{q}_i, \quad (2.16)$$

unde,  $N_{\mathcal{Q}}$  reprezintă numărul total de puncte ce descrie forma  $\mathcal{Q}$ , iar  $\mathbf{q}_i$  este un punct al formei considerate. Deoarece forma este definită în spațiul  $\mathbb{R}^3$ , translația acesteia se realizează secvențial de-a lungul celor trei axe carteziane ( $X, Y$ , respectiv  $Z$ ). Matematic, translația  $\mathbf{t}$  poate fi descrisă prin intermediul unei matrice  $3 \times 1$ , după cum urmează:

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} 0 - \mu_{\mathcal{Q}}(x) \\ 0 - \mu_{\mathcal{Q}}(y) \\ 0 - \mu_{\mathcal{Q}}(z) \end{bmatrix}, \quad (2.17)$$

unde,  $t_x$ ,  $t_y$  respectiv  $t_z$  reprezintă cuantumul translației de-a lungul celor trei axe carteziane. Conform relației 2.17, forma  $\mathcal{Q}$  va fi traslatată în origine. Prin înlocuirea valorilor nule cu coordonatele unui anumit punct din spațiu, forma  $\mathcal{Q}$  va fi traslatată astfel încât noul ei centru de greutate să corespundă cu punctul considerat din spațiu.

### 2.3.1.3 Rotirea unei suprafețe

Cel mai simplu mod de a reprezenta o rotație este prin intermediul celor trei unghiuri Euler, după cum urmează:  $\phi$  pentru rotirea formei în jurul axei  $X$ ,  $\theta$  pentru rotirea în jurul axei  $Y$ , iar  $\psi$

pentru rotirea acesteia în jurul axei Z. Rotația globală  $\mathbf{R}$  se obține prin cumulara celor trei rotații independente:

$$\mathbf{R} = \mathbf{R}_x(\phi) \cdot \mathbf{R}_y(\theta) \cdot \mathbf{R}_z(\psi), \quad (2.18)$$

unde,  $\mathbf{R}_x(\phi)$ ,  $\mathbf{R}_y(\theta)$  respectiv  $\mathbf{R}_z(\psi)$  sunt rotațiile independente în jurul axelor considerate. Înmulțind cele trei rotații prezentate anterior, matricea finală de rotație poate fi descrisă astfel:

$$\mathbf{R} = \begin{pmatrix} \cos(\theta)\cos(\psi) & -\cos(\phi)\sin(\psi) + \sin(\phi)\sin(\theta)\cos(\psi) & \sin(\phi)\sin(\psi) + \cos(\phi)\sin(\theta)\cos(\psi) \\ \cos(\theta)\sin(\psi) & \cos(\phi)\cos(\psi) + \sin(\phi)\sin(\theta)\sin(\psi) & -\sin(\phi)\cos(\psi) + \cos(\phi)\sin(\theta)\sin(\psi) \\ -\sin(\theta) & \sin(\phi)\cos(\theta) & \cos(\phi)\cos(\theta) \end{pmatrix}. \quad (2.19)$$

### 2.3.2 Alinierea eficientă a formelor. Metoda iterativă a celui mai apropiat punct vecin

Date fiind două forme oarecare  $\mathcal{Q}$  și  $\mathcal{P}$ , fiecare reprezentate în sisteme de coordonate proprii, se pune problema identificării transformării de coordonate optime (rotație și translație) care să permită alinierea lui  $\mathcal{Q}$  relativ la  $\mathcal{P}$ . Metoda iterativă a celui mai apropiat vecin, cunoscută în literatura de specialitate sub titulatura *ICP* (*Iterative Closest Point*) [12, 191], prezintă o soluție eficientă și rapidă pentru o astfel de problemă. Metoda constă într-un proces simplu de minimizare a distanței Euclidiene, dintre punctele celor două forme, pentru a le aduce pe acestea într-un sistem de coordonate comun. Abordarea nu este dependentă de modul de reprezentare al formelor, permițând alinierea unui număr mare de tipuri de date geometrice [12]. Dintre acestea se pot menționa:

- seturi de puncte;
- segmente de dreaptă;
- curbe implicite:  $\vec{g}(x, y, z) = 0$ ;
- curbe parametrizate:  $(x(u), y(u), z(u))$ ;
- seturi de triunghiuri (forme reprezentate prin intermediul unui set de triunghiuri (*mesh*));
- forme implicite:  $g(x, y, z) = 0$ ;
- forme parametrizate:  $g(x(u, v), y(u, v), z(u, v))$ .

Din punctul de vedere al aplicabilității, algoritmul poate fi utilizat cu succes la fuzionarea diferitelor perspective ale unui obiect cu scopul obținerii unui volum rigid, la compararea a două sau mai multor scene (determinarea similarităților sau a diferențelor dintre scene), la stabilirea congruenței dintre două sau mai multe forme (identificarea formelor echivalente) și la estimarea mișcării în spațiu a senzorului de percepție. Considerând un caz concret, metoda *ICP* poate furniza o măsură (*confidență*) a apartenenței dintre o formă percepută (reprezentând o parte a corpului uman, de exemplu) și o formă de referință (întreg corpul uman). Acest proces poartă numele de *potrivire*. Exemplul prezentat recurge la o potrivire *locală* deoarece forma percepută corespunde doar unei anumite regiuni din suprafața de referință. Similar, poate exista și o potrivire *globală*, în care alinierea se produce între două forme cu structură constructivă similară ca dimensiune (de

exemplu două siluete ale aceleiași căni).

Principial, cele 6 grade de libertate, 3 pentru rotație și 3 pentru translație, sunt determinate utilizându-se o serie de corespondențe de tipul *cmapv* (*cel mai apropiat punct vecin*). Algoritmul *ICP* converge monoton către cel mai apropiat minim local, obținut utilizând ca măsură distanța medie pătratică dintre punctele ce urmează a fi aliniate (distanța Euclidiană) [191]. O astfel de abordare elimină, din start, necesitatea utilizării caracteristicilor formelor (normala la suprafață, densitatea de puncte, planeitatea, etc.) și totodată elimină și necesitatea stabilirii unor corespondențe între caracteristicile formelor implicate în proces. Singura operație de preprocesare necesară este aceea de eliminare a punctelor izolate (singulare).

Identificarea celui mai apropiat vecin se realizează diferit pentru fiecare tip de date geometrice în parte. Astfel, fie  $\mathcal{P}$  o formă descrisă prin  $N_p$  puncte 3D și  $\mathcal{Q}$  o a doua formă descrisă prin  $N_q$  puncte 3D. Generic, distanța Euclidiană dintre două puncte oarecare aparținând formelor definite anterior se poate determina, după cum urmează:

$$d(\mathbf{p}_i, \mathbf{q}_j) = \|\mathbf{p}_i - \mathbf{q}_j\|^2 = \sqrt{(\mathbf{p}_i(x) - \mathbf{q}_j(x))^2 + (\mathbf{p}_i(y) - \mathbf{q}_j(y))^2 + (\mathbf{p}_i(z) - \mathbf{q}_j(z))^2}, \quad (2.20)$$

unde,  $\mathbf{p}_i(x, y, z)$  și  $\mathbf{q}_j(x, y, z)$  sunt două puncte oarecare aparținând formelor  $\mathcal{Q}$ , respectiv  $\mathcal{P}$ , iar  $d$  este distanța Euclidiană dintre ele. În continuare, distanța dintre un punct  $\mathbf{q} \in \mathcal{Q}$  și cel mai apropiat punct din forma  $\mathcal{P}$  se poate determina utilizând relația următoare:

$$d(\mathbf{q}, \mathcal{P}) = \min_{i \in \{1, \dots, N_p\}} d(\mathbf{q}, \mathbf{p}_i). \quad (2.21)$$

Cel mai apropiat punct  $\mathbf{p}_i \in \mathcal{P}$  satisface egalitatea  $d(\mathbf{q}, \mathbf{p}_i) = d(\mathbf{q}, \mathcal{P})$ .

Fie  $l$  un segment ce unește două puncte  $\mathbf{p}_1$  și  $\mathbf{p}_2$ . Distanța dintre punctul  $\mathbf{q}$  și segmentul  $l$  se calculează după cum urmează:

$$d(\mathbf{q}, l) = \min_{u+v=1} \|u\mathbf{p}_1 + v\mathbf{p}_2 - \mathbf{q}\|, \quad (2.22)$$

unde,  $u \in [0, 1]$  și  $v \in [0, 1]$ . Considerând o suprafață  $L$  reprezentată prin  $N_l$  linii notate  $l_i$ , distanța dintre punctul  $\mathbf{q}$  și suprafața  $L$  devine:

$$d(\mathbf{q}, L) = \min_{i \in \{1, \dots, N_l\}} d(\mathbf{q}, l_i). \quad (2.23)$$

Cel mai apropiat punct  $\mathbf{p}_j \in \mathcal{P}$  situat pe unul dintre segmentele setului de linii  $L$  satisface egalitatea  $d(\mathbf{q}, \mathbf{p}_j) = d(\mathbf{q}, L)$ .

Fie  $t$  un triunghi definit de trei puncte oarecare  $\mathbf{p}_1(x, y, z)$ ,  $\mathbf{p}_2(x, y, z)$  și  $\mathbf{p}_3(x, y, z)$ . Distanța dintre punctul  $\mathbf{q}$  și triunghiul  $t$  este:

$$d(\mathbf{q}, t) = \min_{u+v+w=1} \|u\mathbf{p}_1 + v\mathbf{p}_2 + w\mathbf{p}_3 - \mathbf{q}\|, \quad (2.24)$$

unde,  $u \in [0, 1]$ ,  $v \in [0, 1]$ , și  $w \in [0, 1]$ . Considerând o suprafață  $T$  reprezentată prin  $N_t$  triunghiuri notate  $t_i$ , distanța dintre punctul  $\mathbf{q}$  și suprafața  $T$  este:

$$d(\mathbf{q}, T) = \min_{i \in \{1, \dots, N_t\}} d(\mathbf{q}, t_i). \quad (2.25)$$

Cel mai apropiat punct  $\mathbf{p}_j$  aparținând setului de triunghiuri  $T$  satisface egalitatea  $d(\mathbf{q}, \mathbf{p}_j) = d(\mathbf{q}, T)$ .

Fie  $F$  un set de  $N_e$  entități parametrice  $E_i$  (curbe sau suprafețe parametrice) și  $\mathbf{q} \notin E_i$  un punct oarecare din spațiu. Distanța Euclidiană dintre  $\mathbf{q}$  și setul  $F$  este:

$$d(\mathbf{q}, F) = \min_{i \in \{1, \dots, N_e\}} d(\mathbf{q}, E_i). \quad (2.26)$$

Cel mai apropiat punct  $\mathbf{p}_j$  ce aparține entității parametrice  $F$  satisface egalitatea  $d(\mathbf{q}, \mathbf{p}_j) = d(\mathbf{q}, F)$ .

Fie  $J$  un set de  $N_I$  entități parametrice notate  $I_k$  și  $J = \{I_k\}$  o suprafață implicită având  $k = 1, \dots, N_I$ . Distanța dintre un punct  $\mathbf{q}$  și suprafața implicită  $J$  este:

$$d(\mathbf{q}, J) = \min_{k \in \{1, \dots, N_I\}} d(\mathbf{q}, I_k). \quad (2.27)$$

Cel mai apropiat punct  $\mathbf{p}_j$  ce aparține entității implicite  $I_j$  satisface egalitatea  $d(\mathbf{q}, \mathbf{p}_j) = d(\mathbf{q}, J)$ . Problema calculării distanței dintre un punct și suprafața implicită este diferită de cazurile expuse anterior. Pentru a se determina cel mai apropiat punct  $\mathbf{q}$  relativ la o suprafață implicită  $g(\mathbf{p}) = 0$ , trebuie să se rezolve problema optimizării forțate ce minimizează funcția obiectiv care tratează constrângerea de nonlinearitate:

$$\text{minimizează } f(\mathbf{p}) = \|\mathbf{p} - \mathbf{q}\|^2, \quad g(\mathbf{p}) = 0. \quad (2.28)$$

O posibilă rezolvare a problemei de minimizare este obținută prin formarea unui sistem de ecuații Lagrange asemănător celui din [112]:

$$\begin{aligned} \nabla f(\mathbf{p}) + \lambda^T \nabla g(\mathbf{p}) &= 0, \\ g(\mathbf{p}) &= 0, \end{aligned} \quad (2.29)$$

unde,  $\nabla = |\partial/\partial \mathbf{p}|^T$ . Sistemul de ecuații neliniare poate fi rezolvat utilizându-se metode numerice.

Pe lângă procedura prezentată mai sus pentru estimarea celui mai apropiat vecin, metoda *ICP* mai implică utilizarea unui proces de transformare de similaritate inițială a datelor geometrice. Considerând spațiul tridimensional, identificarea rotației și translației se poate realiza eficient utilizându-se vectorii *quaternion*. Aceștia au fost dezvoltati pentru a evita singularitățile ce apar în metoda unghiurilor lui Euler. Un alt avantaj al acestora este și timpul de calcul redus [148]. Utilizându-se unghiurile Euler, rotația unui corp poate fi obținută prin definirea a trei rotații, câte una pentru fiecare unghi Euler. Prin folosirea quaternionului, rotația se obține într-o singură etapă care înglobează funcțiile trigonometrice sin și cos. Totodată, conversia din quaternion într-o matriță de rotație se obține cu ușurință. Soluția simplă pentru determinarea rotației și translației descrisă de Horn în [85], va fi expusă în continuare.

Un quaternion este un vector  $\vec{q}_R = [q_0, q_1, q_2, q_3]^T$  compus din 4 elemente  $q_i$  având proprietățile  $q_0 \geq 0$  și  $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ . Matricea de rotație  $3 \times 3$  generată de cuaternionul aferent este:

$$\mathbf{R} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{bmatrix}. \quad (2.30)$$

Fie  $\vec{q}_t = [q_4, q_5, q_6]^T$  vectorul de translație, și  $\vec{q}_R$  vectorul de rotație, atunci, vectorul procesului de aliniere se poate nota  $\vec{q} = [\vec{q}_R | \vec{q}_t]^T$ . Se consideră  $\mathcal{Q} = \{\mathbf{q}_i\}$  și  $\mathcal{P} = \{\mathbf{p}_i\}$ , doi nori de puncte 3D ce trebuie aliniați ( $\mathcal{Q}$  în raport cu  $\mathcal{P}$ ). Se definesc următoarele condiții:  $N_q = N_p$  și fiecare punct  $\mathbf{q}_i$  corespunde cu punctul omolog  $\mathbf{p}_i$  (punctele având același index). Funcția obiectiv de tip medie pătratică ce se dorește a fi minimizată este:

$$f(\vec{q}) = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\mathbf{p}_i - \mathbf{R}(\vec{q}_R)\mathbf{q}_i - \vec{q}_t\|^2. \quad (2.31)$$

Centrele de greutate,  $\mu_{\mathcal{Q}}$ , respectiv  $\mu_{\mathcal{P}}$ , corespunzătoare celor două densități de puncte, sunt date de următoarele ecuații:

$$\mu_{\mathcal{Q}} = \frac{1}{N_q} \sum_{i=1}^{N_q} \mathbf{q}_i \text{ și } \mu_{\mathcal{P}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbf{p}_i. \quad (2.32)$$

Matricea de covarianța  $\Sigma_{qp}$  corespunzătoare celor două entități geometrice  $\mathcal{Q}$ ,  $\mathcal{P}$  este:

$$\Sigma_{qp} = \frac{1}{N_q} \sum_{i=1}^{N_q} [(\mathbf{q}_i - \mu_{\mathcal{Q}})(\mathbf{p}_i - \mu_{\mathcal{P}})^T] = \frac{1}{N_q} \sum_{i=1}^{N_q} [\mathbf{q}_i \mathbf{p}_i^T] - \mu_{\mathcal{Q}} \mu_{\mathcal{P}}^T. \quad (2.33)$$

În continuare, componentele ciclice ale matricei asimetrice  $A_{ij} = (\Sigma_{qp} - \Sigma_{qp}^T)_{ij}$  sunt folosite pentru a forma vectorul coloană  $\nabla = [A_{23} \ A_{31} \ A_{12}]^T$ . Acest vector este utilizat la generarea matricei simetrice  $4 \times 4$   $\mathcal{Q}(\Sigma_{qp})$ :

$$\mathcal{Q}(\Sigma_{qp}) = \begin{bmatrix} tr(\Sigma_{qp}) & \delta^T \\ \delta & \Sigma_{qp} + \Sigma_{qp}^T - tr(\Sigma_{qp})\mathbf{I}_3 \end{bmatrix}, \quad (2.34)$$

unde,  $\mathbf{I}_3$  este matricea identitate de dimensiune  $3 \times 3$ . Vectorul propriu  $\vec{q}_R = [q_0, q_1, q_2, q_3]^T$  corespunzător celei mai mari valori proprii a matricei  $\mathcal{Q}(\Sigma_{qp})$  este vectorul ce conține rotația optimă. Translația optimă este descrisă de următoarea ecuație:

$$\vec{q}_T = \mu_{\mathcal{P}} - \mathbf{R}(\vec{q}_R)\mu_{\mathcal{Q}}. \quad (2.35)$$

Cuaternionul celor mai mici pătrate, calculat în  $O(N_q)$  timp, este:

$$(\vec{q}, d_{ms}) = \mathcal{Q}(\mathcal{Q}, \mathcal{P}), \quad (2.36)$$

unde,  $d_{ms}$  este eroarea mediei pătratice a procesului de potrivire a punctelor. Notăția  $\vec{q}(\mathcal{Q})$  este utilizată pentru a descrie setul de puncte  $\mathcal{Q}$  după transformarea vectorului  $\vec{q}$ .

Considerându-se cunoscute premisele descrise anterior, respectiv identificarea celui mai apropiat vecin și suprapunerea inițială a suprafețelor, algoritmul iterativ de aliniere, în forma în care este descris în [12], urmărește deplasarea (*alinieră*) formei  $\mathcal{Q}$  (denumită *formă translatată*) pentru a o alinia relativ la forma de referință  $\mathcal{P}$ .  $N_q$  și  $N_p$  reprezintă numărul de puncte ce descriu suprafețele considerate. Distanța metrică  $d$ , măsurată între punctele individuale  $\mathbf{q}_i$  și suprafața  $\mathcal{P}$ , se poate exprima astfel:

$$d(\mathbf{q}_i, \mathcal{P}) = \min_{\mathbf{p} \in \mathcal{P}} \|\mathbf{p} - \mathbf{q}_i\|. \quad (2.37)$$

Punctul din  $\mathcal{P}$  care satisface condiția de cel mai apropiat vecin se notează cu  $\mathbf{y}$ . Acesta trebuie să respecte afirmația  $d(\mathbf{q}_i, \mathbf{y}) = d(\mathbf{q}_i, \mathcal{P})$ , unde  $\mathbf{y} \in \mathcal{P}$ . Din punct de vedere temporal, identificarea lui  $\mathbf{y}$  se realizează în maxim  $O(N_p)$ , iar timpul total de calcul (a tuturor punctelor din  $\mathcal{Q}$ ) este  $O(N_q \cdot N_p)$ . Totuși, timpul de calcul estimat este  $O(\log(N_p) \cdot \log(N_p))$ . Fie  $\mathbf{Y}$  un vector ce conține cele mai apropiate puncte și  $\mathcal{C}$  operatorul de *mapv*, atunci se poate face afirmația:

$$\mathbf{Y} = \mathcal{C}(\mathcal{Q}, \mathcal{P}). \quad (2.38)$$

Cunoscând vectorul celor mai apropiate puncte  $\mathbf{Y}$ , alinierea celor două suprafețe, utilizând metoda celor mai mici pătrate, se calculează după cum urmează:

$$(\vec{q}, d) = \mathcal{Q}(\mathcal{Q}, \mathbf{Y}). \quad (2.39)$$

În continuare, setul de puncte  $\mathcal{Q}$  este actualizat utilizând relația  $\mathcal{Q} = \vec{q}(\mathcal{Q})$ .

Considerând cunoscute aspectele descrise anterior, algoritmul de aliniere poate fi descris pe scurt după cum urmează:

- se definesc datele de intrare sub forma a doi nori de puncte. Se notează cu  $\mathcal{Q}$  norul translatat și cu  $\mathcal{P}$  norul de referință. Dimensiunile celor doi nori sunt  $N_q$  și  $N_p$ ;
- iterațiile sunt inițializate prin setarea  $\mathcal{Q}_0 = \mathcal{Q}$ ,  $\vec{q}_0 = [1.0 \ 0.0 \ 0.0 \ 0.0]^T$  și iterația  $k = 0$ .  $\mathcal{Q}_k$  este unul și același cu norul de puncte  $\mathcal{Q}$ , cu singura diferență că  $\mathcal{Q}_k$  are aplicat asupra lui transformarea de similaritate corespunzătoare iterației  $k$ . Vectorii de aliniere sunt definiți relativ la setul translatat  $\mathcal{Q}_0$ , astfel încât alinierea finală să reprezinte transformarea completă. Următoarele etapele ( $a$ ,  $b$ ,  $c$  și  $d$ ) sunt aplicate până când algoritmul converge, către o valoare minimă a erorii, cu o toleranță  $\tau$ . Timpul de calcul al fiecărei etape este prezentat între paranteze;
  - a. Calculează cel mai apropiat punct:  $\mathbf{Y}_k = \mathcal{C}(\mathcal{Q}_k, \mathcal{P})$  (timp maxim:  $O(N_q \cdot N_p)$ , timp mediu:  $O(N_q \log N_p)$ );
  - b. Calculează alinierea  $(\vec{q}, d_k) = \mathcal{Q}(\mathcal{Q}_0, \mathbf{Y}_k)$  (timp:  $O(N_q)$ );
  - c. Aplică alinierea:  $\mathcal{P}_{k+1} = \vec{q}_k(\mathcal{P}_0)$  (timp:  $O(N_q)$ );
  - d. Încheie procesul de iterare dacă eroarea medie pătratică scade sub o valoare de prag  $\tau > 0$ . Acesta indică precizia de aliniere:  $d_k - d_{k+1} < \tau$ .

Acuratețea și rapiditatea procesului de aliniere sunt dependente de structura formelor implicate. Contextual, pot exista două situații:

- norul translatat descrie o formă globală similară cu aceea a norului de referință;
- norul translatat descrie o anumită regiune din norul de referință.

Aplicarea algoritmului în primul caz presupune, pe de o parte, translatarea centrului de greutate  $\mu_{\mathcal{Q}}$  al forme  $\mathcal{Q}$  peste centrul de greutate  $\mu_{\mathcal{P}}$  al forme  $\mathcal{P}$  iar, pe de altă parte, identificarea setului de rotații optim care produce o eroare de potrivire (convergență) minimă. Totuși, suprapunerea centrelor de greutate nu este obligatorie. Se poate recurge la o pretranslație utilizându-se un set de rotații inițiale corespunzătoare. Un exemplu de aliniere a doi nori de puncte ce înfățișează silueta a două câni diferite este prezentat în figura 2.11. Cânilor sunt vizualizate din aceeași perspectivă. Eroarea de potrivire (distanța Euclidiană dintre punctele formelor) pentru cazul considerat este  $3.1866 \cdot e^{-5}$ .

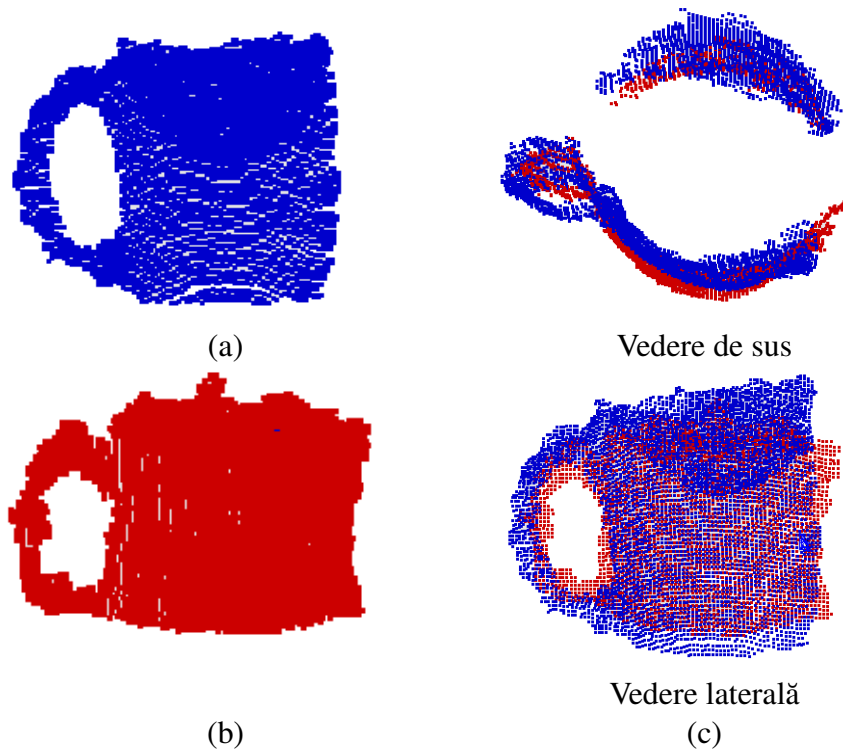


Fig. 2.11 Alinierea a două forme cu siluete similare: (a) forma translatată ( $\mathcal{Q}$ ); (b) forma de referință ( $\mathcal{P}$ ); (c) rezultatul procesului de aliniere utilizând metoda *ICP*.

Așa cum sunt prezentate în figura 2.11, cele două forme aliniate descriu o structură globală similară. Totuși, există numeroase situații în care forma ce urmează a fi aliniată la referință descrie doar o mică regiune din aceasta. Dacă forma aliniată descrie o regiune bine definită (cu iregularități geometrice proeminente) algoritmul *ICP* poate converge optim către un minim local al erorii de potrivire. Ce trebuie remarcat este faptul că numărul de translații necesar este considerabil mai mare deoarece cele două centre de greutate nu mai prezintă similaritate. Practic, forma ce urmează a fi aliniată este potrivită pe rând peste diferite regiuni din forma de referință până când este obținut acel minim local. Un exemplu de aliniere dintre o formă compactă și o formă ce descrie doar o anumită regiune a forme compacte este prezentat în figura 2.12. Eroarea de potrivire pentru cazul

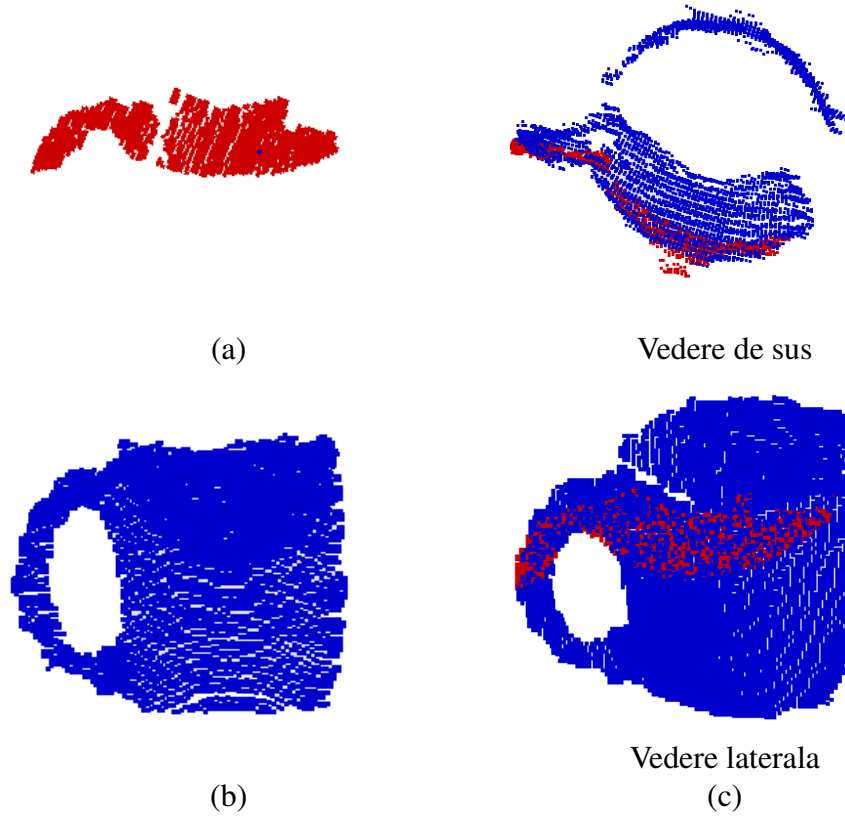


Fig. 2.12 Alinierea unei regiuni relativ la o formă compactă: (a) forma translatată ( $\mathcal{Q}$ ); (b) formă de referință ( $\mathcal{P}$ ); (c) rezultatul procesului de aliniere utilizând metoda *ICP*.

considerat este de  $7.489 \cdot e^{-8}$ .

## 2.4 Caracteristici punctiforme 3D

În forma în care a fost prezentat în secțiunile anterioare, punctul 3D poate fi văzut ca o entitate de sine stătătoare ce mapează o informație în spațiul  $n$ -dimensional. Cea mai importantă caracteristică a acestuia este cea geometrică. Totuși, privit dintr-o perspectivă singulară, geometria unui punct  $\mathbf{p}_i$  poate fi ambiguă. Astfel, la un anumit moment de timp  $t_k$  poate exista un alt punct  $\mathbf{p}_j$  care să prezinte aceleași coordonate carteziane ca cele ale lui  $\mathbf{p}_i$ . Sigur, se pot lua în considerare și celelalte caracteristici ale unui punct (culoare, apartenență la un anumit grup, distanță Euclidiană față de originea sistemului etc.) însă unicitatea punctului nu poate fi asigurată în totalitate. O posibilă îmbunătățire a proprietății de unicitate a unui punct poate fi obținută prin utilizarea informației geometrice de proximitate. Utilizând un *descriptor local*<sup>7</sup> geometria particulară din jurul fiecărui punct poate fi încapsulată într-un nou tip de caracteristică ce conferă punctului considerat unicitate. Literatura de specialitate prezintă numeroși descriptori locali ce pot fi utilizați eficient la reprezentarea unică a formelor în spațiu [79, 168].

Relativ la un anumit punct  $\mathbf{p}_i \in \mathcal{P}$ , vecinătate poate fi considerată o colecție de puncte  $\mathcal{P}^k = \{\mathbf{p}_1^k, \mathbf{p}_2^k, \dots\} \in \mathcal{P}$  a căror distanță față de un punct considerat  $\mathbf{p}_i$  nu depășește o valoare maximă impusă. Astfel, punctele din  $\mathcal{P}^k$  sunt considerate a fi puncte vecine lui  $\mathbf{p}_i$ . Matematic, punctele

<sup>7</sup>Eng. *Local Descriptor*

vecine lui  $p_i$  trebuie să respecte următoarea condiție:

$$\|p^k - p_i\|^2 \leq d_{max}, \quad (2.40)$$

unde,  $p^k$  este un punct aflat în vecinătatea lui  $p_i$ , iar  $d_{max}$  este distanța maximă dintre  $p_i$  și punctul  $p^k$  pentru ca acesta să poată fi considerat vecin. Cea mai uzuală metodă pentru determinarea distanței  $d$  este aceea care utilizează norma Minkowski ( $L_x$ )[173].

#### 2.4.1 Identificarea celui mai apropiat punct vecin în spațiul 3D

Modul în care punctele 3D sunt distribuite în spațiu reprezintă o proprietate tot mai importantă în contextul analizei complexe a norilor de puncte. Cunoașterea nu numai a punctelor vecine ci și a celui mai apropiat punct vecin ( $cmapv$ ), relativ la o anumită coordonată din spațiu, reprezintă, în acest context, un avantaj. Fie  $\mathcal{P}$  un nor de puncte 3D, iar  $p_v \in \mathcal{P}$  și  $p_c \in \mathcal{P}$  două puncte ce aparțin acestuia. Astfel,  $p_v$  are proprietatea de punct vecin dacă și numai dacă se află situat în interiorul unei sfere centrate în  $p_c$ . Sfera, de rază  $r$ , identifică vecinătatea din jurul punctului  $p_c$  (punct ai cărui vecini se doresc a fi identificați). Raza  $r$  este aleasă de către utilizator. După cum este descris și în relația 2.40, punctele vecine, ce compun norul de puncte vecine  $\mathcal{P}^k \in \mathcal{P}$  trebuie să se afle la o distanță, față de  $p_c$ , nu mai mare decât raza  $r$ . Mai departe, punctul vecin din  $\mathcal{P}^k$  a cărui distanță față de  $p_c$  este cea mai mică primește proprietatea de  $cmapv$ . Acesta va fi notat, în continuare, cu  $p_{cmapv}$ .

În conformitate cu notațiile menționate, în figura 2.13 poate fi observat grafic conceptul de  $cmapv$ . O metodă intuitivă de identificare a punctelor vecine lui  $p_c$  este și aceea prin care se realizează o căutare rudimentară(simplistă)<sup>8</sup> ce presupune calcularea distanței dintre  $p_c$  și toate punctele din norul  $\mathcal{P}$ . În contextul aplicațiilor în timp real, acest procedeu este complex și foarte lent. O rezolvare eficientă în acest sens o constituie descompunerea ortogonală a spațiului. În literatura de specialitate, există numeroase metode ce tratează cu succes problema descompunerii spațiului [6, 130, 182]. Toate aceste metode au ca obiectiv comun identificarea atât a setului de puncte vecine ( $\mathcal{P}^k$ ) cât și obținerea  $cmapv$  ( $p_{cmapv}$ ). Una dintre cele mai eficiente metode în acest sens utilizează o structură de date  $k$ -dimensională ( $kd-tree$ ) ce permite identificarea celui mai apropiat vecin într-un timp foarte scurt. Metoda  $kd-tree$  reprezintă de fapt o metodă robustă de organizare a punctelor într-un spațiu multidimensional și are ca principal scop reducerea timpului de căutare ce implică identificarea unei anumite chei multidimensionale (punct 3D) [10, 182]. Lucrul acesta este posibil datorită partiționării spațiului de definire a structurii.

Structura arborescentă  $k$ -dimensională este o structură binară în care fiecare punct  $p_i$  separă hiperspațiul norului  $\mathcal{P}$  în două părți. Astfel, prin separație, un punct se poate afla în hiperspațiul din stânga sau în cel din dreapta. Datorită acestui tip de partiționare binarizată, timpul de procesare se reduce considerabil. Principalele etape implicate în identificarea celui mai apropiat vecin al unui punct  $p_c$  relativ la un nor de puncte  $\mathcal{P}$ , sunt generarea arborelui  $k$ -dimensional și căutarea celui mai apropiat vecin.

<sup>8</sup>Eng. *Brute-Force Search*



în diferite planuri, se va realiza secvențial și ciclic de-a lungul axelor  $X$ ,  $Y$  respectiv,  $Z$ , odată cu fiecare adăugare a unui nou nivel (punct) (v. fig. 2.14(a)). Fiecare nod (punct) introdus este comparat cu nodul părinte de-a lungul unui anumit plan de separație. Pentru prima inserare, separația se va realiza de-a lungul axei  $X$ , pentru cel de-al doilea, de-a lungul axei  $Y$  iar pentru cel de-al treilea va fi  $Z$ , după care, ciclul se reia odată cu ce-a de-a patra inserare care va fi de-a lungul axei  $X$ . Pentru fiecare punct inserat, deciziile se iau plecând de la punctul rădăcină  $p_5$ . Astfel, deoarece  $p_4 < p_5$  de-a lungul axei  $X$ , punctul  $p_4$  va fi adăugat în stânga lui  $p_5$ . Deoarece, sub  $p_5$  nu mai există niciun alt punct,  $p_4$  va fi punct terminal temporar pentru ramura din stânga a lui  $p_5$  (v. fig. 2.14(a)). În continuare se trece la inserarea celui de-al treilea punct, și anume, punctul  $p_3$ . Acesta satisface următoarele condiții:  $p_3 < p_5$  de-a lungul axei  $X$ , deci arborele va fi parcurs pe partea stângă, în continuare  $p_3 \geq p_4$  de-a lungul axei  $Y$  și, deoarece nu mai există niciun punct cu care să îl comparăm, acesta va rămâne până în momentul curent ca punct terminal în dreapta lui  $p_4$  (deoarece  $p_3 > p_4$ ). În mod similar, inserarea celorlalte puncte comportă același procedeu. Un punct aparte este  $p_1$  care, satisfăcând condiția  $p_1 > p_5$  de-a lungul axei  $x$ , va fi inserat în dreapta lui  $p_5$  și astfel se începe o nouă ramură în partea dreaptă a rădăcinii. Arborele va fi complet odată cu inserarea ultimului element ( $p_9$ ). Folosind metoda descrisă mai sus, în figura 2.14(a) este prezentată structura arborescentă 3-dimensională generată utilizând norul de puncte  $\mathcal{P}$ . În continuare, pe baza arborelui  $k$ -dimensional generat se poate defini un vector ce conține indecși punctelor rearanjate din  $\mathcal{P}$  (v. fig 2.14(c)). Modul de completare al acestui vector este următorul: se identifică numărul total de puncte inserate la stânga și la dreapta punctului rădăcină utilizându-se arborele nou generat. În cazul de față vor exista 5 puncte la stânga lui  $p_5$  și 4 puncte la dreapta acestuia. Punctul rădăcină se introduce primul în vector. Acesta se află la poziția  $5 + 1 = 6$ , respectiv indexul 5. În continuare, ordinea de inserare în vector în partea stângă, respectiv dreaptă, se realizează prin alinierea proiecțiilor punctelor relativ la punctul rădăcină. Comparativ, în figurile 2.14(b,c) se poate observa componența acestui vector înainte și după reindexare.

#### 2.4.1.2 Identificarea celui mai apropiat vecin

Utilizând structura arborescentă 3-dimensională generată anterior, procesul de căutare a  $cmapv$  din  $\mathcal{P}$  relativ la punctul căutat  $p_c$  este rapid și eficient. Pentru exemplificare, se consideră punctul  $p_c$  având coordonatele  $(0.001, 0.001, 0.001)$ . Punctul  $p_c$  poate să aparțină norului  $\mathcal{P}$  sau poate să fie un punct oarecare din spațiu. Astfel, procesul de căutare a  $cmapv$  din  $\mathcal{P}$  se realizează după cum este descris în continuare. Plecând de la nodul rădăcină, algoritmul parcurge, în mod recursiv, arborele încercând inserarea lui  $p_c$  (utilizând algoritmul prezentat în secțiunea anterioară). Punctul terminal la care se ajunge va fi considerat  $cmapv$  al lui  $p_c$  până în prezent. Astfel, parcurgând arborele descris în figura 2.14(a),  $cmapv$  a lui  $p_c$  din norul  $\mathcal{P}$  este  $p_9$ , care este considerat, în momentul curent, noul  $p_{cmapv}$ . Distanța pătratică dintre coordonatele celor două puncte este:

$$|p_c - p_9| = (0.001 - 0.000)^2 + (0.001 - 0.005)^2 + (0.001 - 0.005)^2 = 0.000033. \quad (2.42)$$

Nu se poate spune cu certitudine în momentul de față dacă punctul  $p_9$  este cu adevărat  $cmapv$ . Pentru a confirma acest lucru, se apelează la o serie de comparații recursive față de vecinii acestuia. "Părinții" lui  $p_9$  sunt candidați viabili la titlul de  $cmapv$ . Pentru fiecare "părinte", algoritmul de

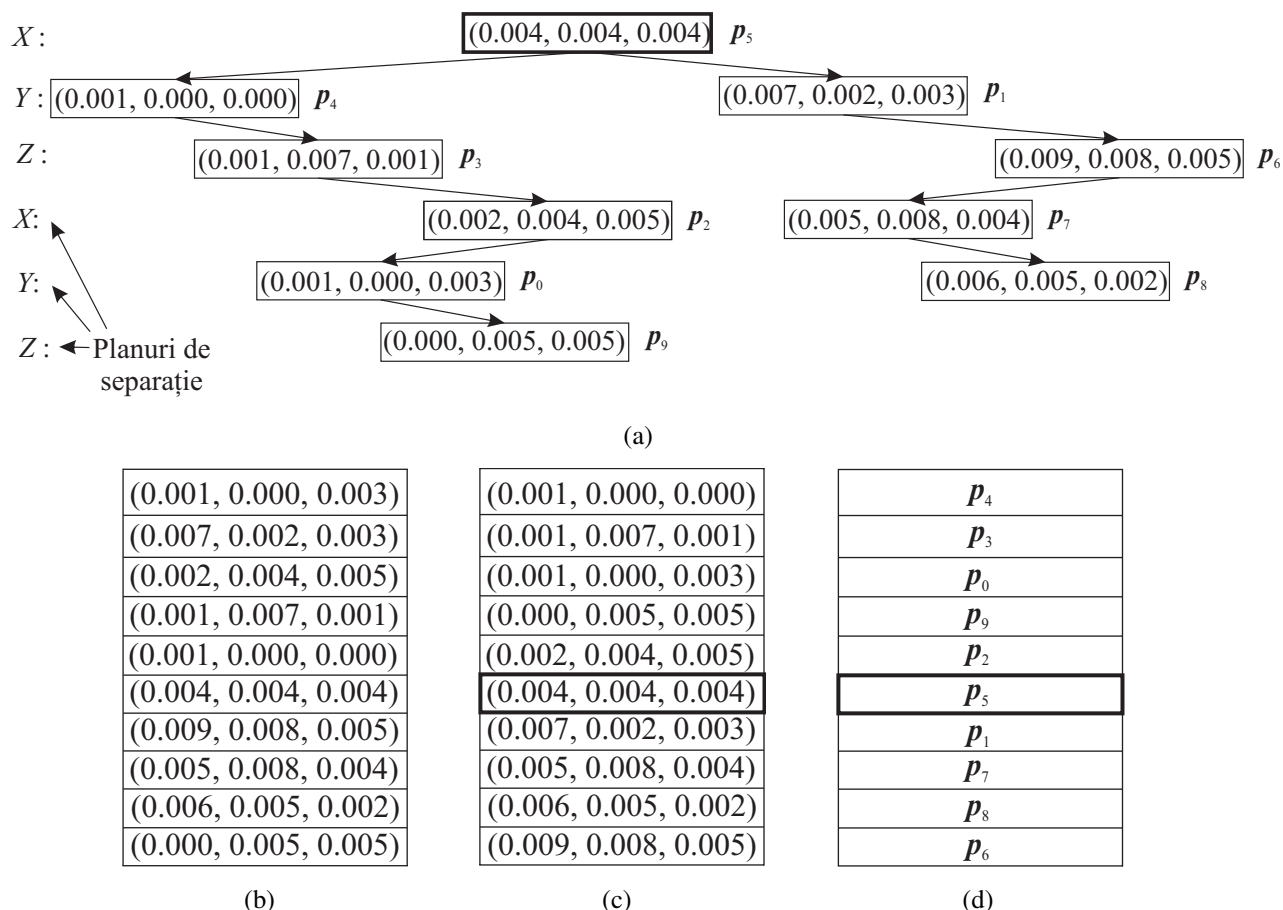


Fig. 2.14 Ilustrare grafică a unui arbore 3-dimensional: (a) structura arborescentă obținută; (b) vectorul inițial de puncte utilizat pentru a genera arborele 3-dimensional; (c) vectorul de puncte rearanjat în conformitate cu noua structură a arborelui; (d) vectorul indecșilor punctelor rearanjate.

căutare testează următoarele condiții:

1. dacă punctul curent este mai apropiat decât  $cmapv$  până la momentul considerat, acesta va deveni noul  $cmapv$ ;
2. dacă punctul se află aproape de planul de separație, pot exista vecini mai apropiați pe partea cealaltă a hiperspațiului (ramura paralelă). Acest lucru poate fi identificat prin intersectarea hiperplanului de separație cu o hipersferă centrată în punctul căutat. Sfera are raza egală cu distanța dintre punctul căutat și cel mai apropiat vecin până la momentul considerat. Dacă distanța dintre coordonata (de-a lungul axei considerate) punctului căutat și coordonata punctului curent este mai mică decât distanța dintre coordonata punctului căutat și coordonata punctului cel mai apropiat până la momentul considerat, atunci punctul curent va deveni  $cmapv$  relativ la punctul căutat;
  - a. dacă hipersfera intersectează planul de separație, pot exista puncte de cealaltă parte a planului, puncte ce pot fi mai apropiate decât cel mai apropiat până la momentul considerat. În această situație, algoritmul se deplasează în jos către o altă ramură a structurii plecând din punctul curent;
  - b. dacă hipersfera nu intersectează planul de separație, algoritmul este parcurs în continuare, cealaltă ramură a nodului fiind eliminată din căutare.

Pentru cazul considerat, părintele lui  $p_9$  este  $p_0$ . Distanța dintre acesta și punctul căutat  $p_v$ , este

$0.000^2 + 0.001^2 + 0.002^2 = 0.000005$ . Aceasta, fiind mai mică decât aceea a punctului  $p_9$ , va desemna punctul  $p_0$  ca fiind noul *cmapv* până la momentul considerat în locul lui  $p_9$  care este eliminat din lista candidaților. Deoarece din  $p_0$  nu pleacă și cea de-a doua ramură (cea stângă), punctul  $p_9$  nu prezintă niciun punct candidat pe partea cealaltă a hiperspațiului (v. condiția 2) în consecință, algoritmul se poate deplasa în sus către "părintele" lui  $p_0$ . Dacă,  $p_0$  ar fi avut un "copil" (noduri în partea inferioară a acestuia) pe ramura din stânga, decizia de a parcurge și această ramură ar fi fost validată doar dacă distanța pătratică de-a lungul axei  $Y$  (deoarece este a doua iterație investigată) dintre  $p_0$  și punctul "copil" din stânga ar fi fost mai mică decât cea mai mică distanță descoperită până în momentul considerat (cea dintre  $p_0$  și  $p_c$ ). În continuare, se evaluează distanța dintre punctul căutat  $p_c$  și  $p_2$ , care este "părintele" lui  $p_0$ . Aceasta este  $0.001^2 + 0.003^2 + 0.004^2 = 0.000026$  ceea ce înseamnă că nici  $p_2$  nu poate fi *cmapv*. Neavând ramură în partea dreaptă, se continuă parcurgerea arborelui. Distanța dintre  $p_3$  și  $p_c$  este egală cu  $0.000^2 + 0.006^2 + 0.000^2 = 0.000036$  ceea ce plasează și acest punct în stiva celor rejectate. La fel, neavând ramură laterală, se continuă iterarea arborelui. Următorul punct comparat este  $p_4$  și datorită distanței mici ( $0.000^2 + 0.001^2 + 0.001^2 = 0.000002$ ), acesta devine noul *cmapv*. Deoarece  $p_4$  prezintă punct "părinte", căutarea continuă. Următorul punct verificat este chiar rădăcina, a cărei diferență este 0.000027. Pentru că, pe partea cealaltă a hiperspațiului (ramura din dreapta a rădăcinii) diferența rămâne mai mare decât cea descoperită până în prezent, parcurgerea arborelui se oprește. Numărul total de comparații efectuate este de 12. Cel mai apropiat vecin astfel identificat este  $p_4(0.001, 0.000, 0.000)$ .

#### 2.4.1.3 Timpii de procesare

Construcția unui arbore  $k$ -dimensional pentru un nor de puncte  $\mathcal{P}$  format din  $N_p$  puncte, durează  $O(N_p \cdot \log N_p)$  timp. Numărul de dimensiuni considerat este  $k = 3$ . Inserarea unui punct în arbore durează  $O(\log N_p)$  timp. Același interval de timp este necesar și pentru eliminarea unui punct din arbore.

Din punctul de vedere al procesului de căutare, în mod clasic, prin utilizarea algoritmului de căutare forțată, găsirea celui mai apropiat punct vecin ar dura  $k \cdot N_p$ . Factorul  $k$  este dat de numărul de caracteristici utilizat pentru a se reprezenta un punct. Pentru cazul considerat,  $k$  este 3. Totuși, utilizând arborele  $k$ -dimensional, complexitatea procesului de căutare scade la  $O(2 \cdot N_p \cdot \log N_p)$  timp. Acesta din urmă ia în calcul timpii tuturor operațiilor necesare căutării și anume: generarea arborelui ( $O(N_p \cdot \log N_p)$  timp) și căutarea în sine a *cmapv* ( $O(\log N_p)$  timp).

#### 2.4.2 Estimarea normalei la suprafață

În forma în care a fost definită anterior, vecinătatea  $\mathcal{P}^k$  unui punct  $p_c \in \mathcal{P}$  poate fi utilizată pentru a descrie în mod unic o serie de noi caracteristici ale lui  $p_c$ . Astfel, *normala* punctului  $p_c$  relativ la suprafață vecină descrisă de  $\mathcal{P}^k$ , este o caracteristică importantă. Normala la suprafață este utilizată cu succes în domenii precum grafica asistată de calculator (pentru a defini surse de lumină și a produce umbre) sau la modelarea formelor (pentru a defini direcțiile de deplasare ale punctelor în spațiu).

În literatura de specialitate se pot identifica numeroase metode de estimare a direcției normalei

[11, 20, 104]. Una dintre cele mai simple abordări este aceea prezentată în [11]. Soluția problemei estimării normalei unui punct  $\mathbf{p}_c$  poate fi descrisă astfel. Se generează un plan (artificial) cu o dimensiune impusă de utilizator. Pe acest plan, în centru, trebuie să se afle punctu  $\mathbf{p}_c$  a cărui normală urmează a fi identificată. Implicit, planul nou creat prezintă o normală relativ la suprafața acestuia. În continuare, planul creat este rotit (în jurul punctului  $\mathbf{p}_c$ ) încercând a-l face tangent cu suprafața vecină (percepută) descrisă de  $\mathcal{P}^k$ . Odată aliniată cele două suprafețe, cea descrisă de planul artificial (nou) și cea descrisă de punctele vecine lui  $\mathbf{p}_c$ , normala căutată (aceea a punctului  $\mathbf{p}_c$ ) este echivalentă cu aceea a panului nou creat. Procesul de aliniere implică minimizarea distanței dintre cele două suprafețe. Distanța minimă, poate fi calculată utilizându-se metoda celor mai mici pătrate<sup>9</sup> prezentată în [157]. Distanța dintre un punct vecin  $\mathbf{p}_i^k \in \mathcal{P}^k$  și planul tangent  $\mathfrak{x}$  poate fi definită ca fiind  $d_i = \mathbf{p}_i^k - \mathfrak{x}$ . În continuare, se notează  $\vec{n}$  ca fiind normala planului tangent. Aceasta, în esență va fi una și aceeași cu aceea a punctului  $\mathbf{p}_c$ . O soluție pentru identificarea normalei  $\vec{n}$  poate fi găsită prin analizarea valorilor și vectorilor proprii ai matricei de covarianță  $\mathcal{C} \in \mathbb{R}^{3 \times 3}$ . Matricea poate fi calculată utilizându-se punctele norului  $\mathcal{P}^k$  (v. relația 2.43).

$$\mathcal{C} = \frac{1}{k} \sum_{i=1}^k \zeta \cdot (\mathbf{p}_i^k - \bar{\mathbf{p}}) \cdot (\mathbf{p}_i^k - \bar{\mathbf{p}})^T, \quad \mathcal{C} \cdot \vec{v}_j = \lambda_j \cdot \vec{v}_j, \quad i = 0 \dots N_{\mathcal{P}^k}, \quad j \in \{0, 1, 2\}, \quad (2.43)$$

unde,  $\zeta_i$  este factorul de pondere al punctului  $\mathbf{p}_i$  (uzual acesta este egal cu 1),  $\mathcal{C}$  este matricea de covarianță,  $\lambda_j \in \mathbb{R}$  sunt valorile proprii,  $N_{\mathcal{P}^k}$  este numărul total de vecin ai punctului  $\mathbf{p}_i$ , iar,  $\vec{v}_j$  este vectorul propriu corespunzător componentelor principale ale lui  $\mathcal{P}^k$ . Dacă  $0 \leq \lambda_0 \leq \lambda_1 \leq \lambda_2$ , vectorul propriu  $\vec{v}_0$ , corespunzător celei mai mici valori proprii  $\lambda_0$ , identifică astfel direcția normalei  $\vec{n} = \{n_x, n_y, n_z\}$ . Similar reprezentării anterioare, normala  $\vec{n}$  poate fi exprimată folosind și perechea de unghiuri  $(\phi, \theta)$  după cum urmează[81]:

$$\phi = \arctan\left(\frac{n_z}{n_y}\right), \quad \theta = \arctan\left(\frac{\sqrt{(n_y^2 + n_z^2)}}{n_x}\right). \quad (2.44)$$

O proprietate importantă a normalei  $\vec{n}$  o reprezintă sensul acesteia. Prin definiție, sensul normalei poate fi pozitiv ( $+\vec{n}$ ) sau negativ ( $-\vec{n}$ ). Relativ la o anumită suprafață determinarea sensului normalei nu este trivial. Există două tehnici ce permit estimarea sensului normalei. Prima, utilizează un proces complex de analiză a componentelor principale ale suprafeței  $\mathcal{P}^k$ . Totuși, sensurile normalelor astfel obținute sunt ambigui. Spre exemplificare, considerând o suprafață cu geometrie regulată (de ex. suprafață plană), metoda mai sus amintită nu produce aceeași direcție și același sens pentru toate punctele suprafeței considerate. Normalele astfel determinate sunt inconsistente. Eliminarea acestei inconsistențe poate fi realizată prin utilizarea punctului din care este vizualizat norul  $\mathcal{P}$ . Acesta va fi notat, în continuare, cu  $\mathbf{p}_{viz} \notin \mathcal{P}$ . Prin definirea acestui punct, stabilirea sensului normalelor devine trivial. Pentru ca sensul normalei să fie orientat către  $\mathbf{p}_{viz}$ ,  $\vec{n}$  trebuie să respecte următoarea relație:

$$\vec{n} \cdot (\mathbf{p}_{viz} - \mathbf{p}_i) > 0. \quad (2.45)$$

<sup>9</sup>Eng. *Least Square*

Pe de altă parte, cea de-a doua metodă de estimare a sensului normalei  $\vec{n}$  încearcă să evite utilizarea lui  $p_{viz}$  deoarece, în nenumărate situații, acesta este inexistent. În schimb, această metodă asigură consistență normalelor prin utilizarea unui graf al orientărilor acestora. Astfel, două puncte apropiate  $p_i$  și  $p_j$  aparținând aceleiași suprafețe netede trebuie să aibă normalele orientate similar. Acest lucru poate fi descris prin intermediul relației următoare:

$$\vec{n}_i \cdot \vec{n}_j \approx 1. \quad (2.46)$$

Mai multe detalii despre această abordare pot fi găsite în lucrarea [84]. Utilizând această abordare, în figura 2.15, cu linii albastre sunt prezentate normalele punctelor unui nor ce descrie o scenă uzuală (o masă cu 3 obiecte pe ea) ce conține, la rândul ei, atât suprafețe plane cât și suprafețe cu geometrie complexă.

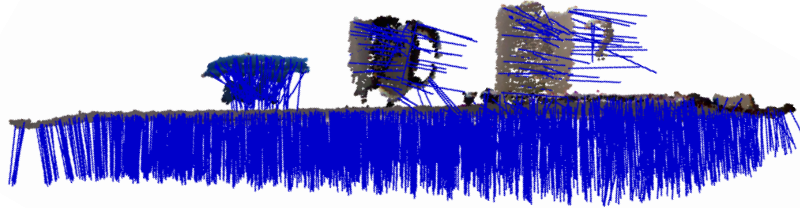


Fig. 2.15 Normale la suprafață orientate consistent, ilustrate prin intermediul liniilor de culoare albastră (vedre în secțiune a scenei percepute).

## 2.5 Segmentarea prin partiționare a norilor de puncte 3D

Scenele complexe conțin adesea un număr mare de obiecte sau elemente necunoscute. Reprezentarea unei astfel de scene implică utilizarea unui nor de puncte de dimensiune considerabil de mare. În consecință, operații precum manipularea sau analiza norului de puncte devin costisitoare atât din punctul de vedere al timpului de procesare cât și din punctul de vedere al puterii de calcul. Mai mult, privită ca un tot unitar, scena este lipsită de sens. O rezolvare posibilă a acestor probleme constă în partiționarea norului de puncte  $\mathcal{P}$ , ce descrie scena inițială, într-o serie de grupuri compacte de puncte tridimensionale. Fiecare dintre aceste grupuri va avea un înțeles și un sens propriu. Avantajele unei astfel de abordări sunt: interpretarea rapidă a datelor și gruparea punctelor cu proprietăți similare (culoare, curbura etc.)

Procesul de grupare compactă a punctelor tridimensionale poate fi considerat o segmentare trivială a spațiului. Rezultatul unei astfel de segmentări este un set  $\mathcal{O}$  de grupuri compacte, fiecare adunând puncte cu proprietăți similare. Ceea ce dă sens fiecărui grup de puncte este caracteristica pe baza căruia se realizează partiționarea spațiului (textură comună, densitate, geometrie etc.). Mai mult, în funcție de împrejurări, procesul de partiționare al spațiului poate folosi o serie de precondiții. Spre exemplificare, în aplicațiile robotice de manipulare, dimensiunea obiectului (reprezentat prin puncte 3D) ce urmează a fi prehensat și manipulat nu trebuie să depășească dimensiunea maximă pe care mecanic robotul o poate prehensa. Orice obiect mai mare decât unul maxim impus va fi etichetat ca nefezabil pentru o operație de prehensare. Totodată, natura acestor grupuri compacte de puncte simplifică din start procesul de segmentare. Fizic, niciodată o cană nu

o să poată sta pe tavan sau pe pereți. Așadar, procesul de partiționare a scenei în grupuri compacte de puncte utilizează o serie de cunoștințe contextuale pentru a da sens, la nivel local, unor densități de puncte care din punct de vedere global nu prezentau niciun înțeles.

Prima problemă ce necesită rezolvare devine astfel identificarea suprafețelor plane din scenă. În afară de utilizarea lor în navigare sau evitarea obstacolelor, acestea sunt redundante și pot fi eliminate scăzând considerabil numărul de puncte din scena de lucru. În figura 2.16 este prezentat un exemplu de identificare a suprafețelor plane din scena considerată. O metodă robustă de identificare a acestor suprafețe, utilizându-se imagini  $RGB - D$ , este expusă în [83, 186]. Algoritmul prezentat de autori este capabil să detecteze obstacole, să segmenteze obiecte ce pot fi manipulate sau să estimeze geometria globală a scenei. Principala caracteristică exploatată de autori în acest sens este normala la suprafață. Astfel, este considerată o suprafață plană, acea densitate de puncte compactă (distanță Euclidiană mică între puncte) ale cărei normale la suprafață sunt similare (consistente). De exemplu, trei puncte ne-coliniare ce prezintă aceeași direcție normală pot fi considerate un plan tridimensional. Mai exact, pentru a se detecta eficient o suprafață plană trebuie efectuați următorii pași:

- se selectează aleatoriu trei puncte necolineare  $\{\mathbf{p}_i, \mathbf{p}_j, \mathbf{p}_k\}$  din norul de puncte  $\mathcal{P}$  al scenei;
- se calculează coeficienții planului model utilizând cele trei puncte definite anterior ( $ax + by + cz + d = 0$ );
- se calculează distanța Euclidiană dintre toate punctele  $\mathbf{p} \in \mathcal{P}$  și planul model  $(a, b, c, d)$ ;
- se calculează numărul de puncte  $\mathbf{p}^* \in \mathcal{P}$  a căror distanță  $d$  până la plan respectă condiția  $0 \leq |d| \leq |d_{th}|$ , unde  $d_{th}$  reprezintă un prag de distanță definit de utilizator.

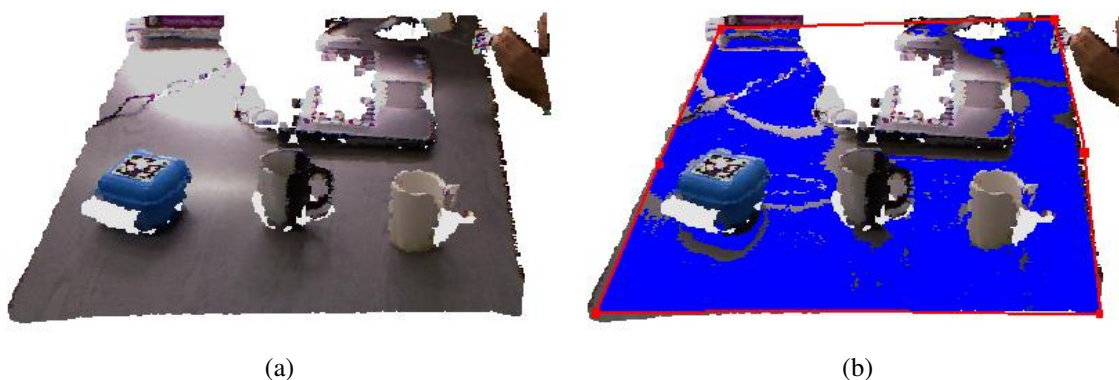


Fig. 2.16 Identificarea suprafețelor plane din scena de lucru: (a) nor de puncte inițial; (b) suprafața plană estimată a cărei puncte 3D sunt colorate cu culoare albastră.

Fiind un proces iterativ, pașii prezentați anterior sunt efectuați de  $k$  ori. În final, se obține o serie de grupuri de puncte  $\mathcal{P}^*$  candidate pentru a fi identificate ca plan. Setul cu cele mai mici distanțe față de planul  $(a, b, c, d)$  (puncte situate pe suprafața planului<sup>10</sup>) este selectat ca fiind cel mai bun suport pentru un plan model. Mai departe, pe baza punctelor  $\mathbf{p}^* \in \mathcal{P}$  sunt determinați coeficienții modelului  $a, b, c$  și  $d$ . Aceștia definesc, în fapt, regiunea rectangulară bidimensională ca fiind cea mai bună aproximare a planului considerat. Rezultatul unui astfel de proces de segmentare a suprafețelor plane este prezentat în figura 2.16. Cu linii de culoare roșie

<sup>10</sup>Eng. Inliers

sunt marcate laturile paralelogramului ce definește suprafața plană estimată, în timp ce, cu albastru sunt colorate punctele 3D ( $\mathbf{p}^*$ ) a căror distanță față de planul considerat respectă condițiile impuse.

Segmentarea suprafețelor plane prezentată mai sus este doar un pas premergător obținerii grupurilor compacte de puncte. Obiectivul real este acela de a partiționa un nor de puncte  $\mathcal{P}$  într-o serie de porțiuni mai mici și cu înțeles propriu. Problema poate fi abordată sub forma unei descompuneri spațiale a datelor. Astfel, prin identificarea subdiviziunilor și marginilor din norul  $\mathcal{P}$  se pot contura, într-o fază incipientă, o serie de grupuri compacte de puncte. Subdiviziunile și marginile anterior menționate sunt identificate pe baza unei măsuri a vecinătății dintre puncte.

Similar, o metodă simplă și eficientă de formare a grupurilor compacte de puncte este prezentată în [149]. Autorii, folosesc pentru gruparea seturilor de puncte o metodă de tip *octree* [23]. Astfel, folosind o subdivizare de tip rețea<sup>11</sup>, spațiul inițial poate fi partiționat rapid. Totuși, metoda poate fi folosită doar pentru aplicațiile ce necesită subdiviziuni spațiale egale. Pentru grupuri compacte cu dimensiuni variabile, algoritmul devine mult prea complex. Având în vedere obiectivul de partiționare a spațiului cu scopul identificării formelor ce pot fi manipulate, sistemul trebuie să înțeleagă, în primul rând, ce este acela *nor de puncte obiect* ( $\mathcal{O}$ ), și cu ce se diferențiază acesta de restul densităților de puncte compacte. Matematic, un grup de puncte compact poate fi definit astfel: fie  $\mathcal{O}_i$  un grup de puncte distincte, descris prin intermediul norului de puncte  $\mathcal{P}_i \in \mathcal{P}$ , și  $\mathcal{O}_j$  un alt grup de puncte distincte, descris prin intermediul norului de puncte  $\mathcal{P}_j \in \mathcal{P}$ , se poate face afirmația  $\mathcal{O}_i \neq \mathcal{O}_j$  dacă și numai dacă:

$$\min \|\mathcal{P}_i - \mathcal{P}_j\|^2 \geq d_{th}, \quad (2.47)$$

unde,  $d_{th}$  este distanța maximă dintre două puncte pentru a se afirma faptul că acestea aparțin aceluiași grup compact. Conform relației 2.47, dacă distanța minimă dintre un set de puncte  $\mathcal{P}_i$  aparținând lui  $\mathcal{P}$  și un alt set de puncte  $\mathcal{P}_j$  aparținând tot lui  $\mathcal{P}$  este mai mare decât o distanța impusă  $d_{th}$ , atunci punctele din  $\mathcal{P}_i$  aparțin unui grup compact de puncte  $\mathcal{O}_i$ , iar punctele din  $\mathcal{P}_j$  aparțin unui grup compact distinct de puncte notat  $\mathcal{O}_j$ . Aproximarea dintre două sau mai multe puncte tridimensionale poate fi determinată utilizându-se o structură de tipul *kd-tree* (v. sec. 2.4.1). Având în vedere toate premisele enunțate anterior, algoritmul de extragere a grupurilor compacte de puncte poate fi schițat astfel:

1. se creează o structură arborescentă de căutare de tipul *kd-tree* utilizând norul de puncte  $\mathcal{P}$  al scenei;
2. se creează o listă goală  $C$  menită să stocheze grupurile compacte ce vor fi identificate, și totodată, se creează o stivă  $Q$  cu punctele ce urmează a fi verificate;
3. pentru fiecare punct  $\mathbf{p}_i$  din setul  $\mathcal{P}_i$  se efectuează următorii subpași:
  - se adaugă  $\mathbf{p}_i$  la stiva  $Q$ ;
  - pentru fiecare  $\mathbf{p}_i \in Q$ :
    - se caută un set de vecini  $\mathcal{P}^k$  într-un perimetru sferic cu raza  $r < d_{th}$ ;
    - pentru fiecare punct vecin  $\mathbf{p}_j^k \in \mathcal{P}^k$  se verifică dacă punctul  $\mathbf{p}_i$  a fost procesat iar dacă nu a fost procesat se adaugă în lista  $Q$ ;

<sup>11</sup>Eng. 3D grid

- dacă toate punctele din  $Q$  au fost procesate, se adaugă  $Q$  la lista grupurilor compacte  $C$ . Se șterge  $Q$  pentru a se începe o nouă iterație;
4. algoritmul se termină atunci când toate punctele din  $\mathcal{P}$  au fost procesate și sunt parte componentă a listei de grupuri compacte  $C$ .

Utilizând algoritmul de extragere a grupurilor compacte de puncte prezentat, în figura 2.17 pot fi observate o serie de nori de puncte obiect înfățișând diferite obiecte de uz general. Conform condițiilor exprimate anterior, obiectele au fost situate pe o suprafață plană (masă).



Fig. 2.17 Segmentarea prin partiționare a norilor de puncte: (a) nor de puncte inițial; (b) grupuri de puncte segmentate.

## 2.6 Analiza statistică a formelor 3D

Interpretate prin prisma operațiilor descrise în secțiunile anterioare, datele au trecut de la etapa în care au fost considerate simple distribuții ale unor puncte în spațiul tridimensional la aceea în care pot descrie forme cu posibil înțeles. Semantica fiecărei forme poate fi extrasă utilizându-se un proces de analiză statistică. Spre exemplificare, se consideră o formă a unui obiect încă neidentificat. Forma provine dintr-o scenă percepută și a fost extrasă utilizându-se metoda de partiționare a spațiului prezentată mai sus. În paralel se definește un set de forme a căror clasă de apartenență este cunoscută. O analiză statistică a formei neidentificate poate stabili dacă aceasta este sau nu similară cu una dintre forme a priori definite.

În statistică, una dintre cele mai uzuale metode de analiză a formelor este *analiza Procrustes* [18, 44, 91]. Aceasta, este utilizată cu precădere la studierea distribuției seturilor de forme (modele) în scopul stabilirii similarității dintre acestea. Deoarece formele sunt definite în sisteme de coordonate proprii, orice proces de analiză este zadarnic fără o *alinieare* prealabilă a acestora. Stabilirea unui sistem de coordonate comun presupune identificarea transformării optime (rotație  $\mathbf{R}$ , translație  $\mathbf{t}$  și factor de scalare  $s$ ) pentru fiecare formă relativ la o formă de referință. Transformarea optimă se obține prin minimizarea distanței dintre punctele formelor. În literatura de specialitate aceasta distanță este cunoscută sub numele de *distanță Procrustes*. Dacă două forme sunt perfect similare, distanța Procrustes rezultată trebuie să tindă către 0.

Discriminativitatea este o proprietate importantă în cadrul analizei comparative a două forme. De exemplu, deși au siluete foarte apropiate, un pantof este diferit de o mașină de jucărie. În acest context, metoda de analiză a formelor trebuie să aibă capacitatea de a analiza diferit cele două forme. Un caz aparte al acestui aspect îl reprezintă reflexiile. Acestea sunt obiecte cu aceeași

*etichetă* (fac parte din aceeași clasă de obiecte) dar care prezintă forme văzute în oglindă. Spre exemplu, pantoful stâng trebuie să aibă o distanță Procrustes similară cu cel drept, piciorul stâng să fie similar cu cel drept, etc. Analiza Procrustes tratează fără dificultăți problemele enumerate [70].

În funcție de obiectivul urmărit, analiza Procrustes poate fi clasică sau generalizată. Primul tip de analiză are ca principal scop identificarea transformării liniare ortogonale ce suprapune optim două sau mai multe forme [18]. Pe de altă parte, analiza Procrustes generalizată își propune să realizeze o formă *medie* identificată ca fiind suma formelor supuse analizei [69].

### 2.6.1 Analiza Procrustes clasică

Se consideră  $\mathcal{P}$  o formă definită printr-un număr finit  $N_p$  de puncte tridimensionale. Fie  $S = S_1, \dots, S_{N_s}$  un set de  $N_s$  forme tridimensionale a priori definite. Toate formele menționate anterior sunt reprezentate prin intermediul unor nori denși de puncte 3D. Identificarea similarității dintre forma  $\mathcal{P}$  și formele descrise în  $S$  implică parcurgerea următorilor pași:

- se stabilește  $\mathcal{P}$  ca formă de referință;
- se aliniază formele din  $S$  relativ la  $\mathcal{P}$ ;
- se calculează distanța Procrustes dintre  $\mathcal{P}$  și forma considerată  $S_k \in Q$  unde,  $k = 1, \dots, N_s$ .

Dacă aceasta este mai mică decât o valoare prag, se stabilește  $S_k$  ca fiind cea mai apropiată formă din  $S$  relativ la  $\mathcal{P}$ , după care se trece la pasul 2.

Forma de referință  $\mathcal{P}$  poate fi o formă indicată explicit (de ex. extrasă utilizându-se metoda de partiționare a spațiului) sau poate fi prima formă din setul  $S$ . Aplicarea analizei Procrustes clasice presupune, după cum s-a mai spus, alinierea prealabilă a tuturor formelor implicate în proces. În acest sens trebuie identificate translația, rotația și factorului de scalare a tuturor formelor din  $S$  relativ la forma referință. Translația se obține prin simpla suprapunere a centrelor de greutate ale formelor din  $S$  peste centrul de greutate al referinței. Pentru simplitate, centrul de greutate al referinței se stabilește la coordonata  $(0, 0, 0)$ . Acest lucru presupune ca toate formele din  $S$  să fie translatare în origine. Măsura distanței cu care este traslatată fiecare formă este dat de media coordonatelor punctelor formelor de-a lungul fiecărei axe. Media acestor puncte este  $(\bar{x}, \bar{y}, \bar{z})$ , unde:

$$\begin{aligned}\bar{x} &= \frac{x_1 + x_2 + \dots + x_N}{N}, \\ \bar{y} &= \frac{y_1 + y_2 + \dots + y_N}{N}, \\ \bar{z} &= \frac{z_1 + z_2 + \dots + z_N}{N},\end{aligned}\tag{2.48}$$

unde,  $N$  este numărul total de puncte al formei. Deplasarea centrului de greutate în origine se realizează după cum urmează:

$$(x, y, z) \rightarrow (x - \bar{x}, y - \bar{y}, z - \bar{z}).\tag{2.49}$$

Alinierea continuă cu determinarea factorului de scalare. Rolul acestuia este acela de a aduce la o dimensiune comună toate formele implicate. Pentru simplitate, dimensiunea lui  $\mathcal{P}$  va fi considerată etalon. Celelalte forme vor fi scalate relativ la dimensiunea lui  $\mathcal{P}$  utilizându-se distanța

medie pătratică <sup>12</sup> dintre punctele fiecărei forme. O formă din  $S$  este considerată scalată atunci când distanța medie pătratică dintre punctele acesteia și originea în care a fost traslatată este 1. Măsura statistică a scalei formei se calculează cu următoarea relație:

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (y_1 - \bar{y})^2 + (z_1 - \bar{z})^2 + \dots}{N_p}}. \quad (2.50)$$

Factorul de scalare devine 1 în momentul în care punctele formei sunt împărțite la factorul de scalare inițial al acesteia  $((x_1 - \bar{x})/s, (y_1 - \bar{y})/s, (z_1 - \bar{z})/s)$ .

Alinierea se încheie odată cu estimarea rotației formelor din  $S$ . Acest lucru nu este un lucru trivial, deoarece o rotație standard de referință nu este tot timpul disponibilă. Considerând cazul cel mai simplu, cel al spațiului bidimensional, estimarea rotației unei forme relativ la o formă de referință se realizează în felul următor: se menține referința într-o poziție fixă și se rotește forma  $S_k$  în jurul referinței până în momentul în care suma distanțelor pătratice<sup>13</sup> devine minimă. Unghiul corespunzător poziției de minim este  $\theta$  și poate fi matematic determinat astfel:

$$\theta = \tan^{-1} \frac{\sum_{i=1}^{N_q} (g_i y_i - l_i x_i)}{\sum_{i=1}^{N_q} (g_i x_i - l_i y_i)}, \quad (2.51)$$

unde,  $(x, y)$  este un punct aparținând lui  $\mathcal{P}$ , iar  $(g, l)$  este un punct din  $S_{N_s}$ . Pentru formele reprezentate în trei dimensiuni, problema rotației devine complexă [70]. În cazul acesta, rotația optimă a lui  $S_k$  este o matrice pătratică de  $3 \times 3$ . Aceasta poate fi determinată utilizând algoritmul Kabsch [97] care are la bază principiul descompunerii în valori singulare<sup>14</sup> (SVD). Prin utilizarea unui astfel de algoritm pot fi soluționate probleme legate de: calculul rangului unei matrice, calculul proiecțiilor ortogonale pentru subspații liniare, metoda celor mai mici pătrate, etc [46].

### 2.6.1.1 Măsurarea similarității dintre forme

Similaritatea dintre două sau mai multe forme poate fi stabilită doar după încheierea procesului de aliniere, adică după efectuarea translației, rotației și scalării formelor analizate relativ la o formă de referință. Statistic, măsurarea similarității dintre două forme poate fi efectuată utilizându-se suma diferențelor pătratice, utilizată anterior la estimarea rotației formei. Matematic, aceasta poate fi descrisă după cum urmează:

$$d = \sum_{i=1}^{N_p} \sqrt{(u_{nn} - x_i)^2 + (v_{nn} - y_i)^2 + (w_{nn} - z_i)^2}, \quad (2.52)$$

unde,  $(u_{nn}, v_{nn}, w_{nn})$  sunt coordonatele celui mai apropiat punct din forma analizată (ce aparține setului de forme  $S$ ) relativ la coordonatele  $(x_i, y_i, z_i)$  ale unui punct din forma de referință. Distanța prezentată în ecuația 2.52 oferă cea mai simplă formă de evaluare a similarității dintre obiecte. O valoare mică a acestei distanțe indică o similaritate ridicată, în timp ce o valoare mare a acesteia

<sup>12</sup>Eng. Root Mean Square Distance

<sup>13</sup>Eng. Sum of Squared Distances

<sup>14</sup>Eng. Singular Value Decomposition

sugerează o potrivire nefavorabilă dintre forme.

### 2.6.2 Analiza Procrustes generalizată

Diferența dintre metoda Procrustes clasică și metoda generalizată este aceea că, în cel din urmă caz, alinierea formelor are loc utilizându-se ca referință toate formele setului și nu doar o singură formă. Astfel, referința se schimbă la fiecare aliniere. Dacă, în formularea clasică orientarea unui obiect era determinată, în formularea generalizată aceasta este selectată aleatoriu. Operațiile de scalare și translație rămân neschimbate față de formularea originală. Pentru compararea a doar două obiecte, analiza Procrustes clasică și cea generalizată sunt identice.

Aplicarea analizei Procrustes generalizate presupune efectuarea următorilor pași:

- 1 se selectează aleatoriu o formă de referință (dintre cele disponibile din  $S$ );
- 2 se aliniază formele disponibile la cea selectată la pasul anterior;
- 3 se calculează o formă medie utilizând setul de forme  $S$  aliniate la pasul anterior;
- 4 se calculează distanța Procrustes dintre punctele formelor. Dacă aceasta este mai mică decât o valoare de prag, se stabilește forma medie ca și formă de referință după care se trece la pasul 2.

Utilizând algoritmul descris, în figura 2.18 este prezentată o formă medie obținută ca sumă a formelor unui set  $S$ .

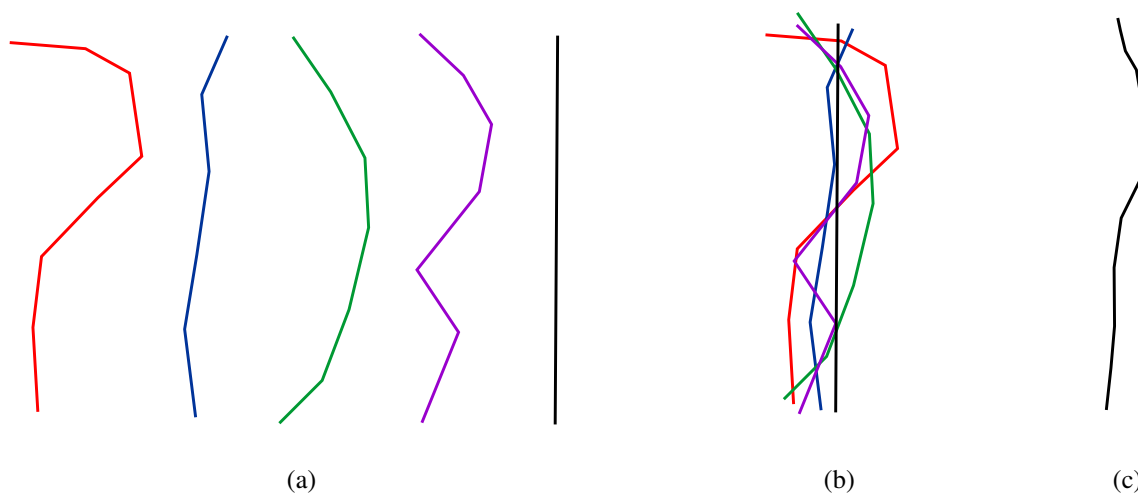


Fig. 2.18 Generarea formei medii a unui set de forme: (a) setul inițial de forme; (b) forme aliniate utilizând metoda de analiză Procrustes generalizată; (c) forma medie obținută.

## 2.7 Concluzii

În cadrul acestui capitol au fost prezentate câteva dintre cele mai importante aspecte legate de analiza și înțelegerea spațiului tridimensional. Aspectele sunt tratate în ordinea cronologică în care apar în cadrul unui proces de reconstrucție volumetrică a scenei/obiectelor vizualizate.

Capitolul a debutat printr-o prezentare succintă a celor mai utilizate și eficiente metode de percepție tridimensională. În acest sens, au fost evidențiate principalele atribute ale unor astfel de sisteme, și anume: acuratețe, timp achiziție, cadre/secundă, preț, etc. Fiind caracterizată de majoritatea atributelor enumerate anterior, metoda de percepție bazată pe lumină structurată a fost

aleasă pentru testarea metodei de reconstrucție volumetrică descrisă în prezenta carte. Rezultat operației de percepție este un nor dens de puncte tridimensionale ce ilustrează perspectiva vizualizată de senzor. Totuși, aceste date, pentru a putea fi interpretate de către sistemul de reconstrucție tridimensională, trebuie să fie reprezentate sub o formă adecvată. În acest sens, secțiunea 2.2 a prezentului capitol a fost realizată o descriere concisă a principalelor moduri de reprezentare atât a suprafețelor, în forma brută în care au fost percepute, cât și a formelor tridimensionale, considerate ca date ce prezintă un interes comun. Avându-se în vedere principalele metode de achiziție și reprezentare a suprafețelor, capitolul a continuat prin prezentarea unor modalități de manipulare a suprafețelor, respectiv formelor tridimensionale. Principalele operații de manipulare descrise în cadrul acestei secțiuni au fost: translația, rotația și scalarea. Rolul acestor operații este de a poziționa și orienta, în spațiul 3D, forma/suprafața considerată. Totodată, în secțiunea 2.3 a fost prezentată și o metodă robustă de aliniere a formelor/suprafețelor bazată pe metoda celui mai apropiat punct vecin (*ICP*).

În secțiunea 2.4 accentul a fost pus pe înțelegerea informației percepute, mai exact, pe determinarea caracteristicilor punctiforme 3D, caracteristici ce vor permite analizarea semantică formei considerate. Una dintre cele mai importante caracteristici este aceea geometrică. Modul în care sunt distribuite punctele 3D dintr-o anumită regiune poate oferi informații esențiale despre acea regiune și despre punctele ce o compun. Astfel, cunoașterea celui mai apropiat punct vecin (*cmav*) devine un atu foarte important în procesul de analiză a spațiului tridimensional. Identificarea celui mai apropiat punct vecin, dat fiind un nor de puncte tridimensionale, este o sarcină relativ ușoară și rapidă. Aceste proprietăți sunt atinse prin reprezentarea norului de puncte prin intermediul unei structuri arborescente  $k$ -dimensionale. Descrierea detaliată a principiului de identificare a celui mai apropiat punct vecin a fost prezentată, pe larg, în carul secțiunii 2.4.

În forma în care a fost perceput mediul înconjurător, acesta reprezintă o aglomerare de obiecte și suprafețe. Selectarea doar a norului de puncte ce descrie un anumit obiect din scenă implică o partiționare a spațiului inițial într-o serie de subregiuni. Generarea acestor subregiuni a fost realizată prin analiza distanțelor eculidiene dintre puncte. Astfel, punctele, a căror distanța Euclidiană respectă o valoare impusă, pot fi considerate ca făcând parte dintr-un grup compact care, cel mai probabil reprezintă un obiect sau o anumită entitate din scena percepută. Pe larg, metoda de partiționare a unui nor de puncte a fost prezentată în cadrul secțiunii 2.5.

Capitolul s-a încheiat prin prezentarea unei metode de analiză statistică a formelor. Scopul unei astfel de analize este acela de a identifica similaritățile dintre formele supuse analizei pentru a genera o formă medie ce astfel poate aproxima o gamă largă de obiecte.



### 3. Estimarea volumelor obiectelor vizualizate unidirecțional

---

---

Estimarea volumetrică 3D utilizând forme implicite  
Estimarea volumetrică 3D utilizând forme predefinite  
Concluzii

---

---

Unele dintre cele mai frecvente acțiuni din cadrul unui scenariu robotic uzual sunt prehensarea și manipularea obiectelor. Există totuși o discrepanță majoră între modalitățile în care un subiect uman, respectiv un sistem computerizat (sistem de calcul) se comportă în astfel de situații. Dacă pentru o persoană au fost necesari numeroși ani să își perfecționeze modul de percepție și înțelegere a mediului înconjurător, sistemul computerizat trebuie să fie capabil să se adapteze și să reacționeze pe loc la fiecare situație nou apărută. Cercetările în domeniu nu au reușit să definească încă pentru sistem o bază de *experiențe* similară cu aceea umană.

Considerându-se un obiect perceput din scenă, un subiect uman îi poate estima forma, volumul și modalitatea prin care îl poate *apuca* pe baza unor experiențe trăite anterior. Neavând la dispoziție aceste informații, un sistem de calcul este obligat să estimeze volumul obiectului vizualizat și mai departe, pe baza acestei percepții, să acționeze corespunzător.

Prehensarea, respectiv manipularea unui anumit obiect reprezintă una dintre cele mai comune acțiuni întreprinse în activitatea de zi cu zi. Pentru ca un sistem de calcul să poată îndeplini o astfel de sarcină acesta trebuie să cunoască structura (coloana vertebrală ce definește forma) obiectului considerat. Totodată, structura unui obiect descrie volumul brut pe care acesta îl ocupă în scenă. Obligatorietatea cunoașterii acestui volum este necesară deoarece diversele moduri de

apucare depind în exclusivitate de forma acestuia. La rândul ei, suprafața obiectului considerat trebuie să înglobeze cât mai multe particularități tocmai pentru a crește precizia punctelor de contact determinate între prehensor și suprafața obiectului.

Relativ la considerațiile prezentate, se poate defini o primă problemă și anume: *cum poate ca ceea ce a perceput sistemul de calcul să reprezinte un corp rigid, un volum?* Ceea ce vede sistemul, dintr-o singură perspectivă, este doar o mică parte din forma reală a întregului obiect. Ideal, pentru a avea un grad de acoperire vizuală cât mai mare, obiectul considerat trebuie "privit" din mai multe perspective. Spre exemplificare, în figura 3.1 sunt prezentate diferite perspective de vizualizare ale unei căni.

În forma în care sunt prezentate în figura 3.1, perspectivele vizualizate nu descriu un obiect rigid ci doar diferite perspective individuale ale acestuia. Pentru a putea reprezenta un volum rigid (o suprafață închisă) toate aceste perspective trebuiesc aliniate pentru a forma un singur obiect. Procesul de aliniere nu este dificil, în contextul unei vizualizări adecvate a fiecărei perspective și presupune identificarea corespondențelor utilizând diferite proprietăți comune (precum cele geometrice sau de textură) din perspectivele succesive [60, 150, 165]. La final, se obține un volum rigid fuzionat din mai multe perspective ce poate fi folosit pentru a interacționa cu obiectul considerat.

Situațiile în care sistemul de calcul poate capta diferite perspective de vizualizare sunt însă foarte rare. În majoritatea cazurilor, configurația scenei nu permite manevrarea sistemului în jurul obiectului pentru a lua perspective multiple. Ținând cont de această limitare, este evident că sistemul de calcul poate percepe util doar o singură perspectivă a scenei. Acesta, de altfel, este cel mai frecvent caz întâlnit în cadrul scenariilor de zi cu zi. Din moment ce doar o mică parte din obiect poate fi vizualizată, problema se poate rezuma la estimarea unui volum echivalent, creat implicit sau definit a priori, care să înlocuiască forma parțială vizualizată a acestui [30, 168]. Volumul echivalent va fi adăugat în scenă și va putea fi utilizat, mai departe, pentru a se reprezenta integral obiectul considerat. Ceea ce deosebește diferitele metode de estimare volumetrică a obiectelor este modul prin care este format volumul echivalent al acestora. Luând în considerare acest aspect, în continuare se vor prezenta două dintre cele mai uzuale metode de estimare volumetrică tratate în literatura de specialitate.



Fig. 3.1 Diferite perspective de vizualizare ale unei căni, reprezentate sub forma unor nori denși de puncte 3D.

### 3.1 Estimarea volumetrică 3D utilizând forme implicite

Una dintre cele mai rapide și simple metode de a genera un volum este prin intermediul forme-

lor implicite. Generarea norului de puncte 3D ce definește un astfel de volum se face automat utilizându-se o relație matematică. Ceea ce este specific formelor implicite este faptul că oferă o soluție simplă și eficientă pentru a identifica dacă un punct oarecare aparține sau nu volumului formei considerate. Unele dintre cele mai populare forme implicite sunt formele *supercuadrice*. Acestea reprezintă o clasă aparte de modele geometrice care, printr-un set redus de parametrii pot descrie o gamă largă de forme cum ar fi: elipse, cuburi, cilindrii, etc. Datorită avantajelor oferite, formele supercuadrice sunt considerate drept unelte comune de modelare geometrică. Supercuadricele au fost introduse pentru prima dată de Barr în [9]. Utilizarea lor a cunoscut o ascensiune impresionantă, acest subiect oferind nenumărate subiecte și teme de cercetare [30, 93, 108]. Fiind forme polimorfe, supercuadricele pot fi folosite cu succes în reprezentarea sau reconstrucția volumetrică a obiectelor. Printre avantajele utilizării acestor forme se menționează:

- parametri ce controlează forma supercuadrice sunt intuitivi și ușor de modelat;
- având o putere de deformare importantă, supercuadricele pot caracteriza atât forme naturale, descrise prin curburi proeminente, cât și forme geometrice solide, descrise prin muchii foarte ascuțite;
- datorită suprafeței închise, formele supercuadrice pot fi utilizate cu succes la reprezentarea volumelor obiectelor;
- prezintă metode robuste și rapide pentru potrivirea formelor supercuadrice peste datele percepute din scena de lucru.

### 3.1.1 Generarea unei forme implicite

Cea mai comună formă supercuadrice este *superelipsa* sau *superelipsoidul*. Prin forma ei, aceasta poate aproxima optim o mare varietate de obiecte. În literatura de specialitate, superelipsoidul este referit sub numele de *supercuadrice* [9], motiv pentru care, în continuare, acesta va fi denumit ca atare. Cel mai ușor mod de a genera o formă supercuadrice este prin rotirea în plane diferite ( $xOz$ , respectiv  $yOz$ ) a două elipse a căror axe principale sunt ortogonale. Matematic, cele două elipse pot fi descrise astfel:  $s_{a_3 a_2 \varepsilon_1}(\eta)$ , respectiv  $s_{a_1 a_2 \varepsilon_2}(\omega)$ . Parametrii  $a_1$ ,  $a_2$  și  $a_3$  descriu lungimile semiaxelor celor două elipse, iar  $\varepsilon_1$ , respectiv  $\varepsilon_2$  reprezintă curbura acestora. Prin pivotarea elipselor în jurul axelor ortogonale ale sistemului de coordonate se obține forma supercuadrice  $\mathcal{S}_{a_1 a_2 a_3 \varepsilon_1 \varepsilon_2}(\eta, \omega)$  (v. fig. 3.2). Explicit, aceasta poate fi descrisă prin relația următoare:

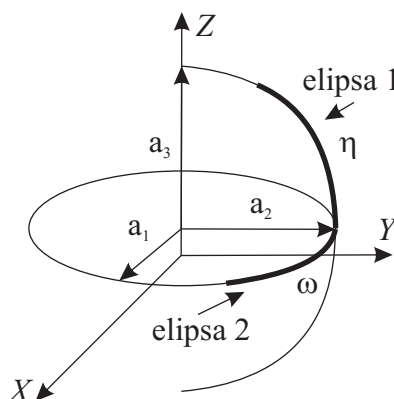


Fig. 3.2 Produsul sferic a două elipse.

$$\mathcal{S}_{\vec{p}}(\eta, \omega) = \begin{bmatrix} x(\vec{p}; \eta, \omega) \\ y(\vec{p}; \eta, \omega) \\ z(\vec{p}; \eta, \omega) \end{bmatrix} = \begin{bmatrix} a_1 \cos \eta^{\varepsilon_1} \cos \omega^{\varepsilon_2} \\ a_2 \cos \eta^{\varepsilon_1} \sin \omega^{\varepsilon_2} \\ a_3 \sin \eta^{\varepsilon_1} \end{bmatrix}, \quad (3.1)$$

unde,  $\eta$  și  $\omega$  sunt unghiurile cu care se rotesc elipsele  $s_{a_3 a_2 \varepsilon_1}(\eta)$ , respectiv  $s_{a_1 a_2 \varepsilon_2}(\omega)$  în jurul axelor  $Z$ , respectiv  $Y$ , iar  $\vec{p} = (a_1 a_2 a_3 \varepsilon_1 \varepsilon_2)$  este vectorul parametrilor ce controlează forma supercuadrice. Unghiurile  $\eta$  și  $\omega$  sunt definite în domeniul de valori  $[-\pi/2, \pi/2]$ , respectiv  $[-\pi, \pi]$ . Considerând vectorul  $\vec{p}$ , parametrii  $\varepsilon_1$  și  $\varepsilon_2$  controlează curbura formei supercuadrice, în timp ce,  $a_1, a_2, a_3$  ajustează dimensiunea acestora de-a lungul celor trei axe carteziene. Figura 3.3 prezintă o serie de forme supercuadrice obținute pentru diferiți parametri ai vectorului  $\vec{p}$ . Pentru cazurile considerate în figură, dimensiunea fiecărei forme a fost menținută constantă la valorile  $a_1 = 5\text{cm}$ ,  $a_2 = 3\text{cm}$ ,  $a_3 = 3\text{cm}$ , în timp ce parametrii  $\varepsilon_1$  și  $\varepsilon_2$  au fost modificați în plaja  $[0.1, 2]$ . Valorile extreme ale ultimilor doi parametri produc forme cu suprafețe ascuțite, în timp ce valorile mediane ale acestora tind să genereze forme cu aspect rotunjit.

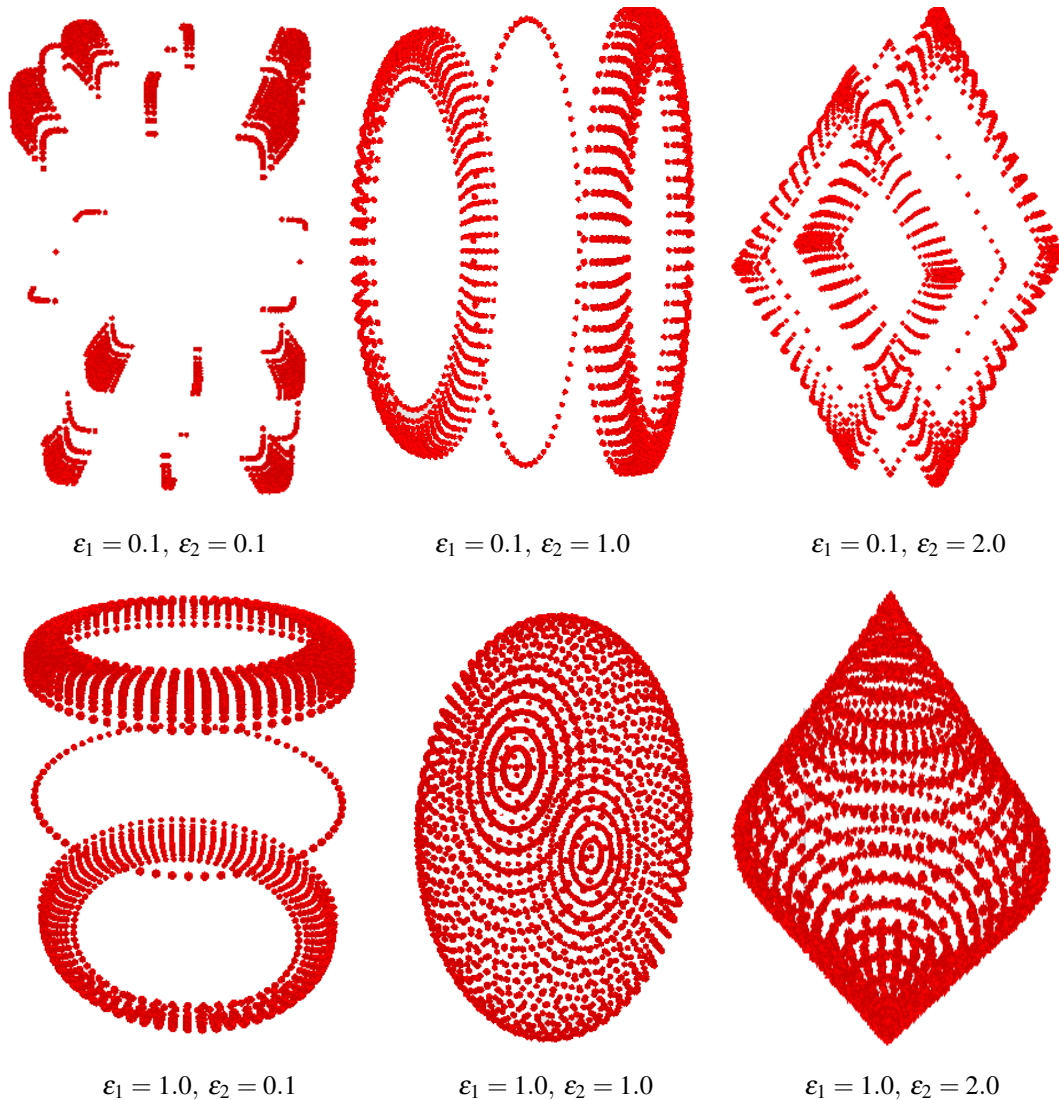


Fig. 3.3 Diferite forme supercuadrice.

Generarea norului de puncte ce descrie o formă supercuadrică în spațiul cartezian poate fi realizată utilizându-se următoarea ecuație implicită:

$$\left[ \left( \frac{x}{a_1} \right)^{\frac{2}{\varepsilon_2}} + \left( \frac{y}{a_2} \right)^{\frac{2}{\varepsilon_2}} \right]^{\frac{\varepsilon_2}{\varepsilon_1}} + \left( \frac{z}{a_3} \right)^{\frac{2}{\varepsilon_1}} = 1, \quad (3.2)$$

unde,  $(x, y, z)$  sunt coordonatele unui punct  $\mathbf{p}$  din spațiu. Dacă acesta satisface egalitatea din relația 3.2, atunci, punctul este situat pe suprafața supercuadrică. Membrul din partea stângă a egalității din ecuația 3.2 poartă denumirea de *funcție interior-exterior* și poate fi definită astfel:

$$F(\vec{p}; x, y, z) = \left[ \left( \frac{x}{a_1} \right)^{\frac{2}{\varepsilon_2}} + \left( \frac{y}{a_2} \right)^{\frac{2}{\varepsilon_2}} \right]^{\frac{\varepsilon_2}{\varepsilon_1}} + \left( \frac{z}{a_3} \right)^{\frac{2}{\varepsilon_1}}. \quad (3.3)$$

Dacă, funcția  $F(\vec{p}; x, y, z) < 1$ , punctul a cărui coordonate sunt  $(x, y, z)$  se găsește în interiorul volumului definit de forma supercuadrică, iar dacă funcția  $F(\vec{p}; x, y, z) > 1$  atunci, acesta se află în afara volumului descris de forma supercuadrică. Având în vedere acest comportament,  $F$  poate fi folosită pentru a se determina poziția unui punct oarecare relativ la forma implicită considerată.

Observând cu atenție formele descrise în figura 3.3 se poate remarca faptul că regiunile curbate prezintă o densitate mai mare de puncte 3D în comparație cu regiunile ce descriu suprafețe plane. Acest lucru se datorează ratei de eșantionare neuniforme introduse de ecuația implicită. Totuși, dacă se dorește folosirea unei forme supecuadrică pentru a aproxima și reconstrui volumul unui obiect vizualizat, norul de puncte ce descrie forma supercuadrică trebuie să fie uniform distribuit. Relativ la această problemă, în literatura de specialitate există numeroase soluții [8, 139]. Dintre toate metodele, cea expusă în [139] prezintă cea mai bună performanță.

### 3.1.2 Estimarea formei optime a unei supercuadrică

În forma în care a fost prezentat anterior, un supercuadric poate fi definit prin intermediul a doar 5 parametri ( $a_1, a_2, a_3, \varepsilon_1, \varepsilon_2$ ). Norul de puncte astfel generat se află la o poziție și o orientare oarecare în spațiu. Deoarece se dorește ca forma supercuadrică să aproximeze rigid volumul obiectului din scenă, aceasta trebuie generată în locația unde se afla norul de puncte al obiectului perceput. Practic, se urmărește "potrivirea" formei supercuadrică peste informația percepută a obiectului de interes (cel ce se dorește a fi reconstruit). Acest proces implică introducerea a încă 6 noi parametri de definire a formei supercuadrică. Trei dintre aceștia,  $t_x, t_y, t_z$ , realizează translația supercuadrică în poziția în care se află obiectul perceput, iar restul de trei parametri descriu unghiurile Euler cu care trebuie rotit supercuadricul pentru a se alinia perfect cu obiectul perceput. Cele trei unghiuri sunt  $\phi, \theta$ , respectiv  $\psi$ . Ele reprezintă rotația în jurul axelor  $X, Y$ , respectiv  $Z$ . Luând în considerare și acești 6 parametri, definirea unei forme supecuadrică la o poziție și orientare bine definite în spațiu poate fi realizată utilizând un vector  $\vec{p}$  compus din 11 parametri ( $a_1, a_2, a_3, \varepsilon_1, \varepsilon_2, x, y, z, \phi, \theta, \psi$ ).

Simpla aliniere și orientare a formei supercuadrică nu sunt de ajuns pentru a aproxima optim obiectul perceput. Un al doilea obiectiv în acest sens este identificarea parametrilor de modelare

cei mai adecvați care aduc forma supercuadrică foarte aproape de forma percepută a obiectului de interes din scenă. Parametri care modelează structura unui supercuadric sunt, așa cum s-a mai prezentat,  $a_1, a_2, a_3, \varepsilon_1, \varepsilon_2$ . O metodă eficientă pentru determinarea acestor parametri este prezentată în [93]. Conform acesteia, potrivirea formei supercuadrice peste un nor de puncte tri-dimensionale  $\mathcal{P}$ , se poate realiza prin minimizarea sumei pătratelor distanțelor dintre cele două suprafețe. Această sumă poate fi exprimată astfel:

$$L(\vec{p}) = \sum_{i=1}^{N_p} (F(\vec{p}; \mathcal{P}) - 1)^2, \quad (3.4)$$

unde,  $N_p$  este numărul total de puncte ce descrie obiectul perceput  $\mathcal{P}$ . Este imperativ necesar ca înainte de a aplica ecuația 3.4, cele două forme (supercuadrice și  $\mathcal{P}$ ) să fie aliniate [93]. Relația 3.4 are totuși o serie de neajunsuri. Minimizarea lui  $L(\vec{p})$  favorizează supercuadricele ce descriu volume mari. Această problemă apare, în special, în momentul în care obiectul perceput este văzut dintr-o perspectivă frontală care produce masive regiuni ocluzionate în părțile laterale și superioare. Practic, una din lungimi ( $a_1, a_2$  sau  $a_3$ ) nu poate fi aproximată din lipsă de informație 3D. Pe lângă această problemă, măsura potrivirii  $L(\vec{p})$  favorizează cu precădere formele reprezentate prin curburi puternice. O soluție la aceste probleme o reprezintă introducerea termenului  $\sqrt{a_1 a_2 a_3}$  în relația 3.4 [93, 162]. Acest termen reduce semnificativ efectele dimensiunii supercuadrice asupra procesului de minimizare a lui  $L(\vec{p})$ . Pe de altă parte, ridicarea lui  $F$  la puterea  $\varepsilon_1$ , elimină complet influența regiunilor curbate asupra lui  $L(\vec{p})$ . Ținând cont de mențiunile precedente, relația 3.4 poate fi rescrisă după cum urmează:

$$L(\vec{p}) = \sum_{i=1}^n (\sqrt{a_1 a_2 a_3} (F(\vec{p}; \mathcal{P}) - 1))^2. \quad (3.5)$$

Având în vedere toate aspectele expuse, în figura 3.4 este prezentată forma supercuadrice ce estimează optim forma sticlei percepute. Minimizarea lui  $L(\vec{p})$  s-a realizat utilizându-se metoda descrisă în [5].

Pentru obiecte cu formă regulată, estimarea volumelor utilizând forme supercuadrice este rapidă și eficientă. Nu același lucru se poate spune și în cazul obiectelor articulate sau care prezintă curburi proeminente. Pentru acestea, volumul estimat va reprezenta doar o formă *brută* a obiectului perceput. O posibilă soluție pentru cazul obiectelor complexe este partiționarea formei inițiale într-o serie de subregiuni, fiecare conținând o suprafață regulată, parte a celei inițiale [13, 47, 100]. O astfel de abordare este prezentată în [30] unde, autorii potrivesc în interiorul fiecărei subregiuni astfel determinate câte o formă supercuadrică de sine stătătoare. În final, prin însumarea formelor supercuadrice din toate subregiunile determinate se poate obține un volum global ce reprezintă fidel obiectul complex perceput.

Păstrând scenariul descris în figura 3.4, în figura 3.5 poate fi observat volumul estimat al sticlei utilizându-se multiple forme supercuadrice. Folosind două forme supercuadrice, una pentru gâtul sticlei și una pentru corpul principal, similaritatea dintre forme crește considerabil. Reprezentat de o formă supercuadrică independentă, corpul sticlei aproximează precis forma reală.

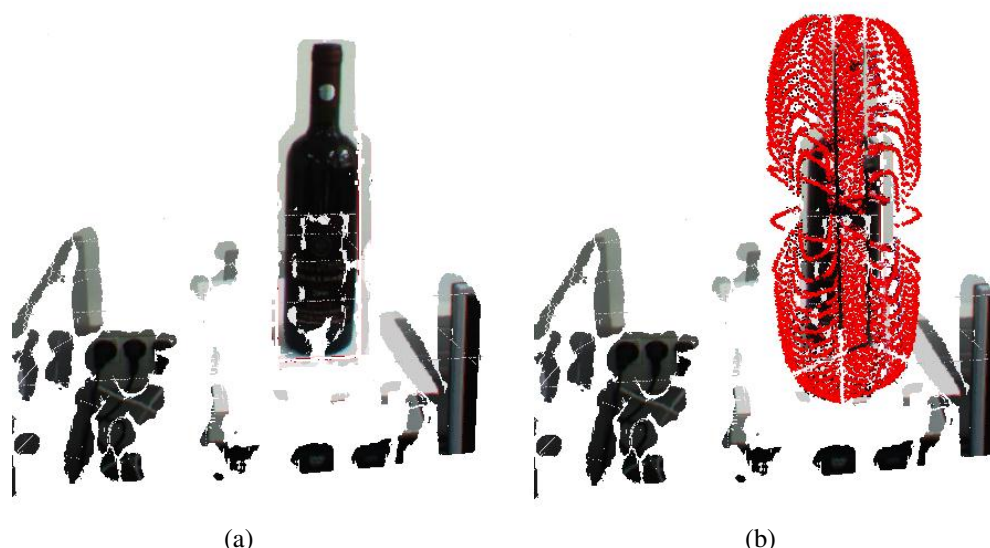


Fig. 3.4 Estimarea volumului unei sticle utilizând forme supercuadrice: (a) scena percepută; (b) forma supercuadrică optimă ce aproximează obiectul considerat din scenă.

### 3.1.3 Modelarea supercuadriceilor

Există situații în care forma generică a unei supercuadrice nu poate lua înfățișarea obiectului considerat din scenă. Totuși, prin deformare, capacitatea formelor supercuadrice de a aproxima obiecte este considerabil extinsă. Principalele acțiuni de deformare pe care un supercuadric le poate suferi sunt *îndoirea*, respectiv *alungirea* [93, 162]. Deformarea globală a formelor este o practică des întâlnită în reconstrucția tridimensională a obiectelor [101]. Scopul final al acestor deformări este acela de a reduce considerabil eroarea de potrivire dintre forma supercuadrică și cea percepută din scenă.

Una dintre cele mai populare metode de deformare este cea parabolică [162]. Metoda constă în trei etape. În prima etapă, se definește planul de deformare ce formează un unghi  $\alpha$ , de exemplu cu axa  $X$ , și care va reprezenta punctul de inflexiune al formei. Pe acest plan sunt proiectate, în continuare, toate punctele formei supercuadrice. În cea de-a doua etapă, punctele astfel proiectate sunt modelate utilizându-se o lege parabolică de deformare. În final, punctele sunt reprojectate înapoi în spațiu pentru a da naștere unei noi forme supercuadrice.

## 3.2 Estimarea volumetrică 3D utilizând forme predefinite

În forma în care este el perceput, obiectul vizualizat nu poate descrie o formă rigidă (o suprafață închisă). Cea mai naturală și precisă metodă de reconstrucție a volumului obiectului considerat este prin înlocuirea acestuia cu o replică exactă. Această abordare presupune definirea a priori a unui număr relevant de forme rigide predefinite (replici) ce sunt posibil să apară în scena de lucru. Principala provocare în acest sens este identificarea sau căutarea formei predefinite care se potrivește optim cu aceea a obiectului vizualizat. O soluție la această problemă o reprezintă motoarele de căutare a formelor care, prin utilizarea unor corespondențe 3D, identifică similaritățile geometrice dintre forme [168]. În literatura de specialitate aceste motoare sunt cunoscute sub acronimul *3DOR* (*3D Object Retrieval*). În esență, sistemul *3DOR* încearcă potrivirea formei vizibile a obiectului peste o serie de forme predefinite depozitate în interiorul unei baze de date. Forma predefinită, care

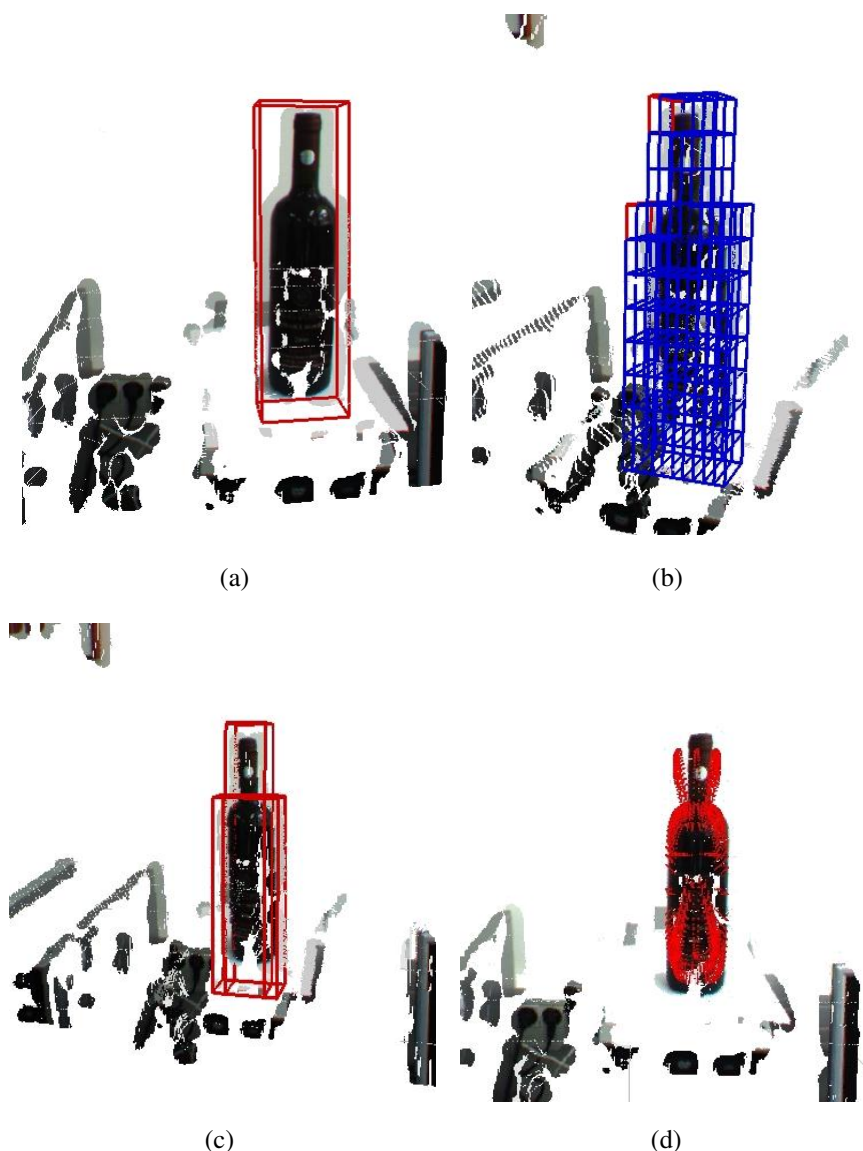


Fig. 3.5 Estimarea volumului unei căni utilizând multiple forme supercuadrice: (a) regiunea de interes inițială dislocată de obiect; (b) partiționarea spațiului definit de regiunea de interes inițială; (c) subregiunile de interes formate în jurul suprafețelor cu caracter regulat ale obiectului considerat; (d) volumul final estimat utilizând două forme supercuadrice reunite.

prezintă cel mai mare număr de similarități față de obiectul perceput, va fi adnotată în scenă și va fi utilizată, în continuare, pentru a reprezenta obiectul vizualizat în etapele de procesare următoare.

Un avantaj important al acestei abordări este precizia. Spre deosebire de celelalte metode de reconstrucție volumetrică, aceasta nu necesită nicio etapă suplimentară de modelare deoarece, forma predefinită aproximează fidel forma percepută. În acest sens trebuie definit un număr ridicat de forme predefinite tocmai pentru a mări șansele ca forma percepută să prezinte o formă rigidă corespondentă în baza de date. Pe de altă parte, motoarele de căutare *3DOR* sunt cunoscute ca mari consumatoare de timp și putere de calcul. Totuși, motivația utilizării unei astfel de abordări este întreținută de numeroasele baze de date cu forme predefinite ușor accesibile prezente pe internet [154, 158]. Pe lângă utilitatea sistemului *3DOR* de a reconstrui volume ale obiectelor, acesta poate fi folosit cu succes la reconstrucția sau identificarea clădirilor precum și a locațiilor acestora. În acest sens, Microsoft® și Google® au demarat un amplu proiect de scanare și reprezentare

tridimensională a orașelor ce are ca principal obiectiv crearea unui sistem inovativ de identificare și documentare a poziției globale.

Generic, un motor de căutare *3DOR* are la bază o schemă-bloc ca aceea prezentată în figura 3.6. Fiecare dintre elementele schemei-bloc are un rol important în obținerea unui rezultat rapid și corect. Pe baza obiectului vizualizat (unidirecțional) se formează o cerere de identificare. Este necesar ca forma obiectului vizualizat să conțină particularități (puternice proeminențe geometrice) pentru a putea fi identificat în mod unic în baza de date. O formă cu foarte multe suprafețe generice (plane sau foarte puțin curbate) nu poate fi deosebită în comparație cu celelalte forme predefinite. Pentru a crește șansa ca forma vizualizată să fie identificată cu una din baza de date, numărul de forme predefinite din baza de date trebuie să fie considerabil de mare.

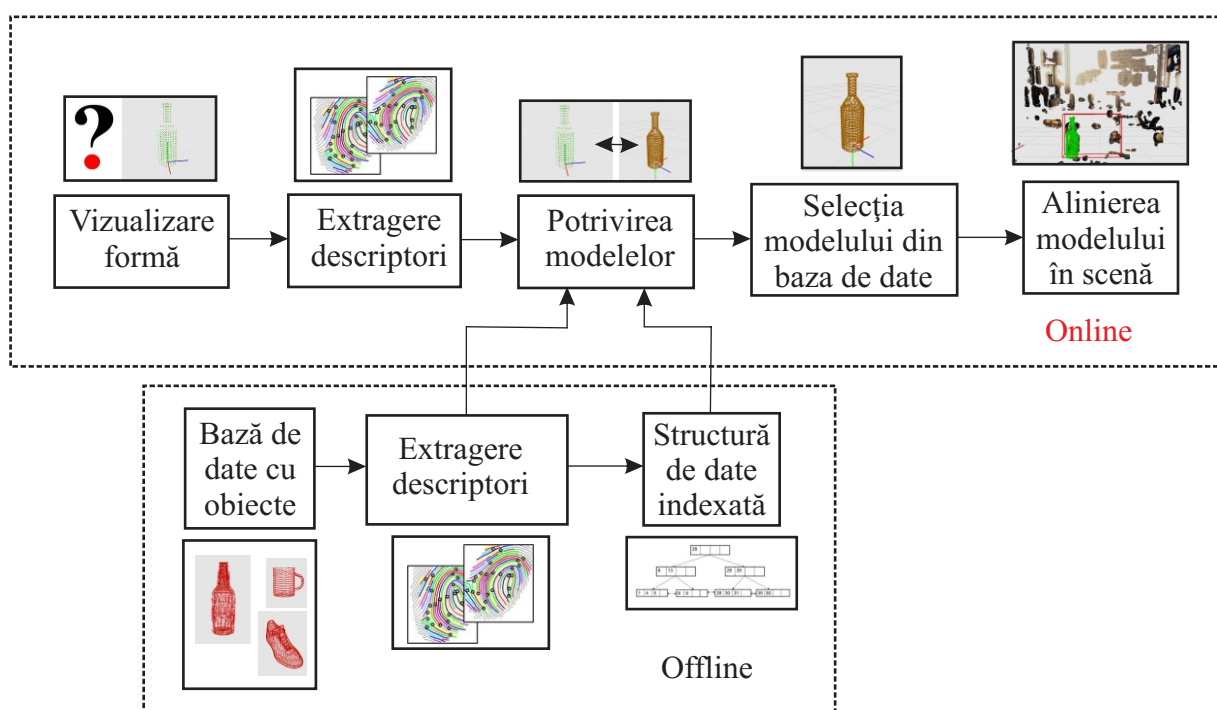


Fig. 3.6 Schema-bloc a unui motor de căutare a formelor.

În procesul de comparare a două forme, identificarea similarităților dintre două sau mai multe puncte nu se poate realiza utilizându-se doar coordonatele carteziene ale acestora deoarece poziția în spațiu a unui punct este mult prea generală. Ceea ce însă mai specific pentru punctul respectiv este relația acestuia cu punctele din jur. Este foarte puțin probabil ca două puncte oarecare să prezinte exact același model de distribuție spațială a punctelor vecine. O întrebare esențială în acest sens este cum pot punctele vecine să reprezinte distinct un anumit punct în spațiu. Răspunsul este dat de o metodă inovativă de reprezentare a suprafețelor bazată pe așa numiții *descriptori*.

### 3.2.1 Descriptori 3D

Rolul principal al unui descriptor este acela de a integra geometria suprafeței din jurul unui punct într-un mod unic de reprezentare [168]. Un descriptor este o histogramă ce mapează distribuția punctelor unei suprafețe într-un vector multidimensional.

În funcție de tipul proprietăților punctelor integrate, descriptorii pot fi împărțiți în trei mari categorii:

- descriptori având la bază *caracteristici*;
- descriptori având la bază *grafuri*;
- alte tipuri de descriptori.

Descriptori având la bază grafuri sunt utilizați cu precădere atunci când forma supusă analizei descrie o formă rigidă (completă) a obiectului considerat. Având în vedere faptul că majoritatea scenelor sunt percepute dintr-o singură perspectivă, descriptorii ce au la bază caracteristici sunt, de departe, cea mai bună alegere pentru acest tip de percepție. În funcție de tipul de caracteristici folosite, descriptorii pot fi împărțiți în următoarele subcategorii:

- descriptori ce utilizează caracteristici *globale* [39, 136];
- descriptori ce utilizează caracteristici *locale* [79];
- descriptori ce utilizează *distribuția* caracteristicilor [194];
- descriptori ce utilizează *hărți spațiale*.

Suprafețele percepute unidirecțional mai sunt cunoscute, așa cum a mai fost prezentat, și sub numele de suprafețe 2.5D. Din moment ce există foarte puțină informație globală despre aceste suprafețe, descriptorii globali nu pot fi utilizați eficient decât dacă obiectul perceput poate fi vizualizat în ansamblul său. În contextul suprafețelor percepute unidirecțional descriptorii ce utilizează caracteristici locale sunt însă cei mai eficienți [79]. Principalii descriptori locali definiți în literatura de specialitate sunt: *SHOT* (*Signature of Histograms of Orientations*) [175], imagini rotative [94], *FPFH* (*Fast Points Features Histograms*) [150], armonice sferice [153], etc.

Un descriptor poate integra informația înconjurătoare sub două forme: *semnătură* și *histogramă*. Prima formă de înglobare descrie suprafața tridimensională din jurul unui anumit punct sub forma unei amprente lăsate pe o sferă centrată în punctul considerat. Pe de altă parte, cea de-a doua abordare integrează geometria și topologia locală (număr puncte, rețea triangulară, etc) într-o histogramă relativ la un *cadru de referință* centrat în punctul considerat. Spre exemplu, descriptorul *SHOT* [175] folosește o combinație între semnătura și histograma unei suprafețe pentru a produce unul dintre cele mai robuste descriptoare ce utilizează caracteristici locale. Cadrul de referință introdus de autori are rolul de a crește și mai mult eficiența acestui descriptor prin asigurarea invarianței la rotație și translație a suprafețelor ce urmează a fi comparate. Figura 3.7 prezintă cadrul de referință și histograma rezultantă pentru o suprafață dată, reprezentată prin intermediul unui descriptor de tip *SHOT*.

### 3.2.2 Estimarea corespondențelor 3D

Stabilirea unui grad de similaritate între doua sau mai multe forme, considerând un motor de căutare al formelor *3DOR*, se realizează prin intermediul corespondențelor 3D. Acestea pot fi determinate după cum urmează. Fie  $\mathcal{P}$  norul de puncte al formei percepute (denumită, în continuare, *forma căutată*) și  $\mathcal{Q}$  norul de puncte al unei forme predefinite (denumită, în continuare, *forma țintă*). Numărul punctelor 3D utilizat la descrierea norilor este  $N_p$ , respectiv  $N_q$ , iar  $N_p \neq N_q$ . Fiecare formă predefinită din baza de date este considerată, la rândul ei, o posibilă formă țintă. Pentru ca procesul de identificare a corespondențelor să decurgă corect, forma țintă

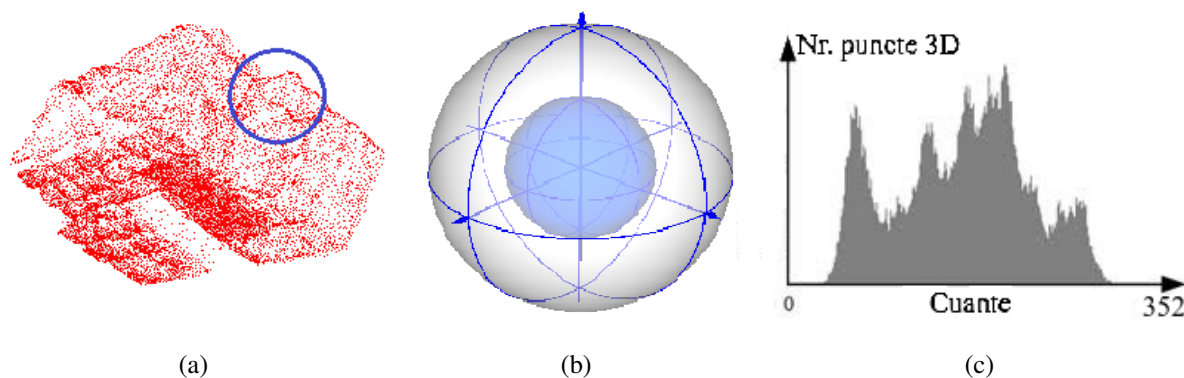


Fig. 3.7 Reprezentarea suprafețelor utilizându-se descriptori de tip *SHOT*: (a) nor de puncte inițial împreună cu regiunea vecină a unui punct 3D considerat, marcată printr-un cerc de culoare albastră; (b) cadrul de referință aferent centrat în punctul considerat; (c) histograma regiunii definite de cadrul de referință.

trebuie scalată la dimensiunea formei căutate. Pentru fiecare punct din  $\mathcal{P}$ , respectiv din  $\mathcal{Q}$  ( $i \neq j$ ,  $i = 1, \dots, N_p$ ;  $j = 1, \dots, N_q$ ) se calculează câte un descriptor local. Rezultatul fiecărui descriptor este o histogramă  $H_{\mathcal{P}_i}$ , respectiv,  $H_{\mathcal{Q}_j}$ . Modul în care aceste histogramme sunt create depinde de descriptorul utilizat (de modalitatea în care acesta înglobează informația). Histogrammele punctelor dintr-o formă au toate aceeași dimensiune (definită de numărul de cuante (elemente)<sup>1</sup>, notate în continuare cu  $k$ , generate de descriptor). Unii descriptori generează histogramme cu 153 de cuante [94], alții cu 352 de cuante [175]. Procesul de identificare a corespondențelor dintre cele două forme presupune compararea, pas cu pas, a fiecărei histogramme dintre cele două forme. Algoritmul de comparare se mai numește și *algoritm brut de potrivire rudimentară* [137], deoarece compară descriptorul fiecărui punct din forma căutată cu descriptorii tuturor punctele din forma țintă. Ceea ce se urmărește în acest fel este să se identifice regiunile din  $\mathcal{P}$  ce se regăsesc și în  $\mathcal{Q}$ . Astfel, numărul de comparații efectuat pentru a se putea stabili un grad de similaritate între cele două forme este  $N_p \times N_q$ .

Gradul de asemănare dintre două histogramme,  $H_{\mathcal{P}_i}$ , respectiv  $H_{\mathcal{Q}_j}$ , este dat de o măsură specială care, în esență descrie distanța dintre cele două histogramme. În literatura de specialitate se poate identifica un număr relativ ridicat de metode de stabilire a acestor măsuri. Dintre cele mai importante se menționează: coeficientul de corelare [94], pătratul factorului *chi*<sup>2</sup> [107], intersecția histogramelor, norma L1, norma L2 și, nu în ultimul rând, distanța *Battacharyya* [14]. Formulările matematice ale acestor măsuri de similaritate pot fi observate în tabelul 3.1.

Fie că descrie un punct din  $\mathcal{P}$  sau unul din  $\mathcal{Q}$ , dimensiunea oricărei histogramme este constantă tot timpul. Spre exemplificare, descriptorul *SHOT* creează o histogramă cu 352 cuante. Similaritatea dintre două histogramme se stabilește apoi prin compararea fiecărei cuante  $k = 1, \dots, 352$ , în parte. Pentru ca cele două histogramme să fie identice, fiecare cuantă din prima histogramă trebuie să fie egală cu cuanta de la aceeași poziție din cea de a doua histogramă. Satisfăcând o serie de cerințe, precum complexitate matematică redusă sau timp de calcul redus, norma L2 este una dintre cele mai robuste soluții pentru stabilirea gradului de similaritate dintre două histogramme. Totuși, dacă cele două cerințe nu trebuiesc imperativ satisfăcute, distanța *Battacharyya* este una dintre

<sup>1</sup>Eng. *Bins*

<sup>2</sup>Eng. *Chi Squared*

Tab. 3.1 Măsurile ale similarităților dintre două sau mai multe histogramme.

| Distanța<br>(măsura)<br>$d(H_{\mathcal{P}_i}, H_{\mathcal{Q}_j})$ | Formulare matematică                                                                                                                                                            | Rezultat<br>similaritate<br>perfectă | Rezultat<br>similaritate<br>pe jumătate | Rezultat<br>similaritate<br>nereușită |
|-------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|-----------------------------------------|---------------------------------------|
| Coeeficient de<br>corelare                                        | $d = \frac{\sum_k H_{\mathcal{P}_i}'(k) \cdot H_{\mathcal{Q}_j}'(k)}{\sqrt{H_{\mathcal{P}_i}'(k)^2 \cdot H_{\mathcal{Q}_j}'(k)^2}}$ $H_G' = H_G(k) - \frac{1}{N} \sum_l H_G(l)$ | 1.0                                  | 0.7                                     | -1.0                                  |
| ChiSqr                                                            | $d = \sum_k \frac{(H_{\mathcal{P}_i}(k) - H_{\mathcal{Q}_j}(k))^2}{H_{\mathcal{P}_i}(k) + H_{\mathcal{Q}_j}(k)}$                                                                | 0.0                                  | 0.67                                    | 2.0                                   |
| Intersecția<br>histogramelor                                      | $d = \sum_k \min(H_{\mathcal{P}_i}(k), H_{\mathcal{Q}_j}(k))$                                                                                                                   | 1                                    | 0.5                                     | 0.0                                   |
| Norma L1                                                          | $d = \sum_k (H_{\mathcal{P}_i}(k) - H_{\mathcal{Q}_j}(k))$                                                                                                                      | 0.0                                  | -                                       | 1.0                                   |
| Norma L2                                                          | $d = \sqrt{\sum_k (H_{\mathcal{P}_i}(k) - H_{\mathcal{Q}_j}(k))^2}$                                                                                                             | 0.0                                  | -                                       | 1.0                                   |
| Distanța<br>Battacharyya                                          | $d = \sqrt{1 - \sum_k \frac{\sqrt{H_{\mathcal{P}_i}(k) \cdot H_{\mathcal{Q}_j}(k)}}{\sqrt{\sum_k H_{\mathcal{P}_i}(k) \cdot \sum_k H_{\mathcal{Q}_j}(k)}}}$                     | 0.0                                  | 0.55                                    | 1.0                                   |

cele mai bune soluții pentru compararea aceluiași două histogramme [96]. Luând în considerare norma L2, o valoare mică a măsurii indică o similaritate ridicată între cele două regiuni descrise prin intermediul histogramelor, în timp ce, o valoare apropiată de 1 reprezintă o nepotrivire între acestea. La finalul comparațiilor se obține un vector  $C$  format din  $N_c = N_p$  corespondențe  $c$ . O corespondență  $c$  este definită astfel:

$$c(i) = \{i, j\}, \quad i = 1, \dots, N_p, \quad j = 1, \dots, N_q, \quad (3.6)$$

unde,  $i$ , respectiv  $j$  reprezintă indecși punctelor din  $\mathcal{P}$ , respectiv  $\mathcal{Q}$  indicate de histogrammele lor aferente. Practic, o corespondență  $c$  leagă un punct din  $\mathcal{P}$  cu un punct din  $\mathcal{Q}$ . Astfel, toate punctele din  $\mathcal{P}$  trebuie să prezinte corespondențe în  $\mathcal{Q}$ , nu și invers deoarece forma căutată este  $\mathcal{P}$ . În realitate, datorită generalității diferitelor regiuni (suprafețe plane sau constante), corectitudinea corespondențelor din vectorul  $C$  este îndoielnică. Spre exemplificare, formele din figura 3.8 sunt reprezentate prin aproximativ 214 de puncte 3D, dar prezintă doar 49 de puncte corespondente. Restul de 165 de corespondențe nu au fost afișate deoarece indicau toate același punct din  $\mathcal{Q}$ . Acestea sunt, de fapt, corespondențe invalide ce descriu regiuni cu o formă comună. Numărul echivalent al punctelor celor doi nori ( $N_p \simeq N_q$ ) este întâmplător. Descriptorul utilizat pentru reprezentarea celor două forme a fost *SHOT*.

Cadrul de referință utilizat pentru calcularea descriptorului unui anumit punct controlează și numărul de corespondențe valide stabilite între forme. Astfel, un cadru de referință redus ca dimensiune ( $\simeq 5\text{mm}$ ) descrie o suprafață relativ redusă în jurul punctului din  $\mathcal{P}$ . Fiind așa de mică, numeroase suprafețe din  $\mathcal{Q}$  vor prezenta similaritate (față de o serie de suprafețe din  $\mathcal{P}$ ) deoarece, deformarea geometrică din interiorul unui cadru suport de  $\simeq 5\text{mm}$  nu poate avea unicitate ridicată.

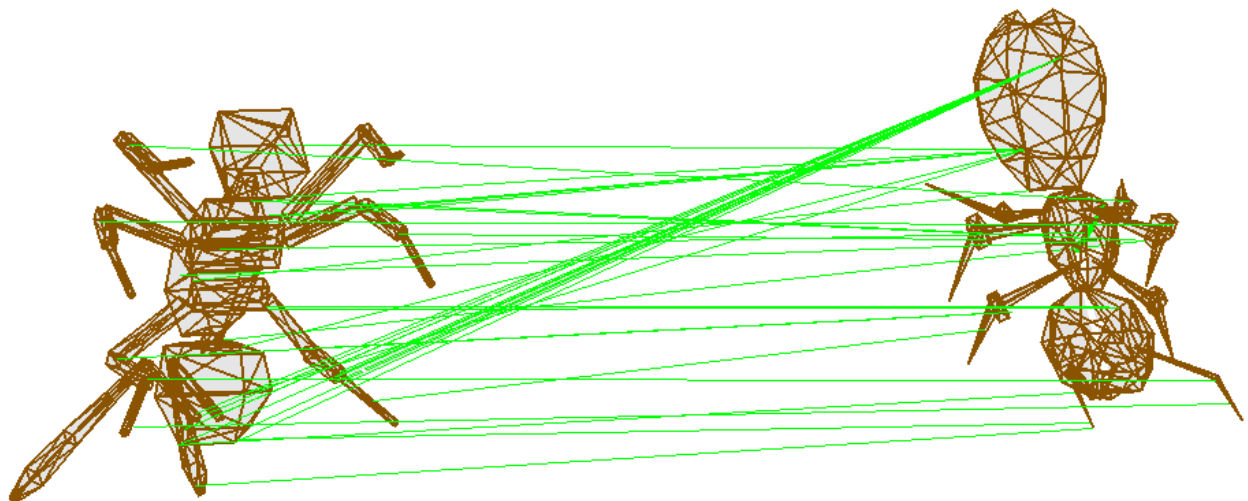


Fig. 3.8 Set de corespondențe (marcate prin linii de culoare verde) stabilite între două forme similare.

Pe de altă parte, un cadru de referință larg ( $\simeq 5\text{cm}$ ) descrie o suprafață considerabil mai mare. Astfel, descriptorul trebuie să reprezinte o regiune cu un potențial geometric proeminent ce contribuie la scăderea ambiguității corespondenței. În concluzie, cadrul de referință trebuie ajustat în funcție de dimensiunea formelor precum și de complexitatea geometrică descrisă de acestea.

După cum se poate observa, nu toate corespondențele valide din figura 3.8 sunt corecte (*valide*). Pentru a se putea stabili corect gradul de similaritate dintre formele considerate, corespondențele eronate trebuie filtrate.

### 3.2.3 Filtrarea corespondențelor 3D

Scopul principal al procesului de filtrare al corespondențelor este de a elimina acele corespondențele care conectează regiuni greșite din punct de vedere al poziționării lor în spațiu. Validarea corespondențelor se poate realiza prin utilizarea filtrelor de corespondențe. Unele dintre cele mai relevante metode de validare a corespondențelor sunt:

- validarea distanței celui mai apropiat vecin [94];
- filtrarea probabilistică a pantelor corespondențelor [73].

Scopul validării distanței celui mai apropiat vecin este acela de a se păstra în final doar corespondențele cu adevărat discriminatorii. O corespondență este considerată discriminatorie dacă măsura (distanța) dintre histogramele punctelor indicate ( $H_{\mathcal{P}_i}$ , respectiv  $H_{\mathcal{Q}_j}$ ) este cea mai mică (v. tab. 3.1). Prin compararea histogramei punctului  $\mathbf{p}_i \in \mathcal{P}$  cu histograma unui punct alt  $\mathbf{q}_l \in \mathcal{Q}$ ,  $l = 1, \dots, N_q$  ( $\mathbf{q}_j \neq \mathbf{q}_l$ ), este posibil să se obțină o măsură cu valoare ușor mai mare (grad de similaritate ușor mai scăzut) decât aceea a punctului  $\mathbf{q}_j$  (care a fost stabilit ca fiind punctul ce cea mai mare similaritate în raport cu punctul căutat  $\mathbf{p}_i$ ). Conform măsurilor lor, punctelor  $\mathbf{q}_j$ , respectiv  $\mathbf{q}_l$  descriu suprafețe din  $\mathcal{Q}$  foarte asemănătoare și sunt cele mai mici stabilite între punctul  $\mathbf{p}_i$  și toate punctele din  $\mathcal{Q}$ . Astfel,  $\mathbf{p}_i$  prezintă două corespondențe candidate în  $\mathcal{Q}$  ( $\mathbf{q}_j$ , respectiv  $\mathbf{q}_l$ ). Aceasta este, de fapt, o ambiguitate, motiv pentru care punctul  $\mathbf{p}_i$  nu trebuie să prezinte corespondențe în  $\mathcal{Q}$ , deoarece regiunea descrisă de aceste corespondențe este prea generală pentru forma considerată. O astfel de anomalie poate fi detectată făcându-se un raport între primele două cele mai mici

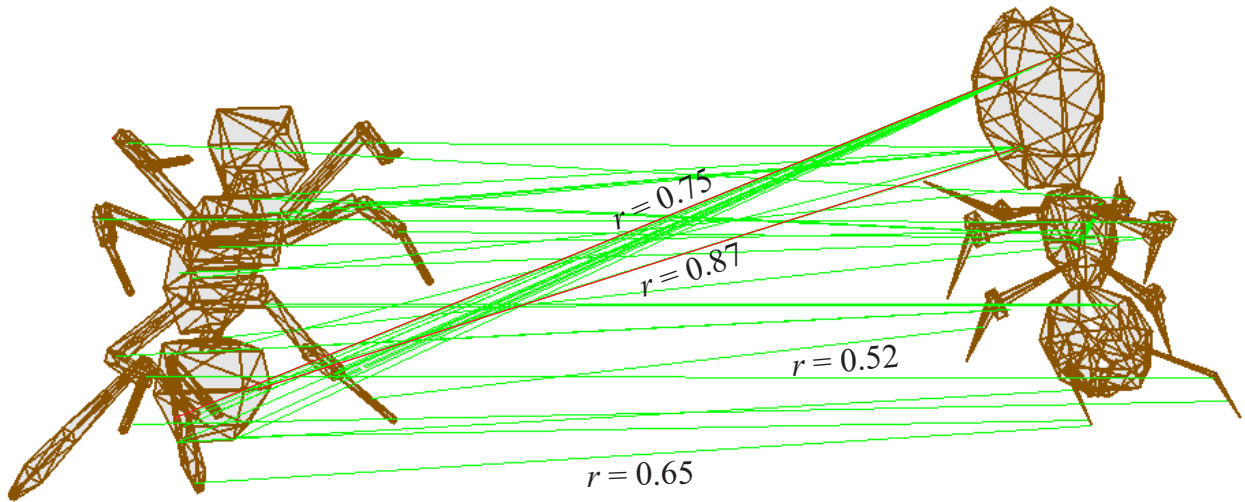


Fig. 3.9 Diferite valori ale raportului  $r$  pentru o serie de corespondențe stabilite între două forme similare. Cu culoare roșie sunt indicate câteva dintre corespondențele eliminate.

măsurile stabilite între  $p_i$  și toate punctele din  $\mathcal{Q}$ . Un astfel de raport este descris matematic prin relația:

$$r = \frac{dist_{NN}}{dist_{sNN}}, \quad (3.7)$$

unde,  $r$  este raportul menționat,  $dist_{NN}$  este cea mai mică măsură dintre histograma punctului  $p_i$  considerat și forma  $\mathcal{Q}$  (având în vedere exemplul anterior, acest punct este  $q_j$ ), iar  $dist_{sNN}$  este cea mai mică a doua măsură dintre histograma punctului  $p_i$  considerat și forma  $\mathcal{Q}$  (conform exemplului anterior, acest punct este  $q_l$ ). Notăția  $NN$  este folosită pentru a nota cel mai apropiat vecin<sup>3</sup>.

Dacă cele mai mici două măsuri din  $\mathcal{Q}$  au valori apropiate, atunci rezultatul raportului se apropie de valoarea 1. Acest lucru arată faptul că două regiuni din  $\mathcal{Q}$  sunt similare cu o regiune din  $\mathcal{P}$  motiv pentru care corespondența punctului din  $\mathcal{P}$  nu trebuie luată în considerare. În caz contrar, dacă raportul  $r$  are o valoare ce se apropie de 0, cea de-a doua mai mică măsură este foarte îndepărtată de prima cea mai mică măsură și astfel unicitatea acesteia este solidă. Punctul din  $\mathcal{P}$  are astfel o singură corespondență în  $\mathcal{Q}$  care este și cea mai corectă (*de încredere*). Pragul de la care raportul unei corespondențe prezintă *încredere* este stabilit empiric la valoarea 0.7. Menținând acest raport sub valoarea indicată, la finalul operației de filtrare vor fi păstrate doar corespondențele corecte care indică cu adevărat suprafețe unice între cele două forme. Pentru a se înțelege mai bine procesul de validare, utilizând raportul definit prin relația 3.7, în figura 3.9 sunt prezentate grafic valorile raportului dintre un punct din  $\mathcal{P}$  și o serie de puncte corespondente (candidate) din  $\mathcal{Q}$ .

Cea de a doua metodă de filtrare își propune eliminarea corespondențelor încrucișate. Acest principiu de validare este simplu și se bazează pe calculul pantelor corespondențelor dintre punctele formelor considerate. Panta unei corespondențe se poate calcula pe baza celor două puncte ce de-

<sup>3</sup>Eng. Nearest Neighbour

finesc o corespondență, după cum urmează:

$$m_{xOy} = \frac{[\mathbf{p}_i(x) - \mathbf{q}_j(x)]}{[\mathbf{p}_i(y) - \mathbf{q}_j(y)]}, \quad m_{xOz} = \frac{[\mathbf{p}_i(x) - \mathbf{q}_j(x)]}{[\mathbf{p}_i(z) - \mathbf{q}_j(z)]}, \quad (3.8)$$

unde,  $m_{xOy}$ , respectiv  $m_{xOz}$  sunt pantele proiectate pe planele  $xOy$ , respectiv  $xOz$ , iar  $\mathbf{p}_i$  și  $\mathbf{q}_j$  aparțin norilor  $\mathcal{P}$  și  $\mathcal{Q}$ . Pentru eliminarea corespondențelor ce indică pante greșite este necesar un *model* comun al distribuției pantelor. Acesta este obținut utilizându-se o funcție probabilistică de densitate<sup>4</sup> asemănătoare celei descrise prin relația 3.9.

$$\mathcal{L} = (\mathcal{S} | \theta_\mu, \theta_\sigma), \quad (3.9)$$

unde,  $\mathcal{L}$  este funcția probabilităților de densitate,  $\mathcal{S}$  este vectorul pantelor, iar  $\theta_\mu$  și  $\theta_\sigma$  sunt media, respectiv varianța distribuției pantei de-a lungul unei anumite axe ( $X$ ,  $Y$  sau  $Z$ ).

Elementul central al unei astfel de funcții este un estimator de probabilitate asemănător celui din relația următoare [164]:

$$\theta = \arg \max_{\theta \in \Theta} \mathcal{L}(\theta | \mathcal{S}), \quad (3.10)$$

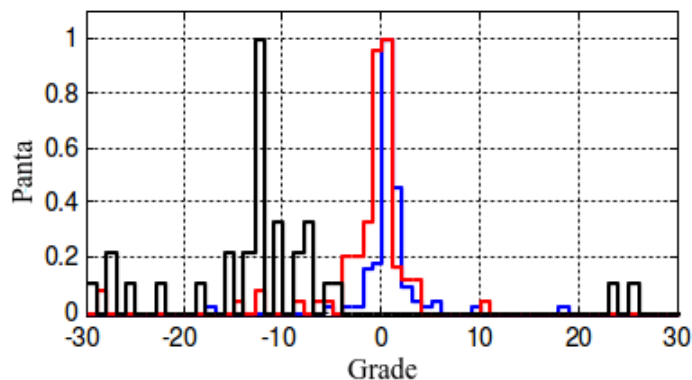
unde,  $\theta$  este estimatul de probabilitate maximă<sup>5</sup> al funcției  $\mathcal{L}$ . Corespondențele care depășesc deviația standard a acestui model al pantelor sunt eliminate. Modelului estimat al pantelor precum și distribuția lor de-a lungul celor 3 axe pot fi observate în figura 3.10

Utilizându-se în cascadă cele două filtre prezentate anterior, în ordinea în care au fost descrise, se obține un nou set  $C$  de corespondențe a căror *încredere* (confidență) este ridicată. Calitatea setului  $C$  este importantă deoarece la finalul procesului de comparare, formele țintă din baza de date cu similaritatea cea mai mare vor primi o *confidență de potrivire* relativ la forma căutată. Aceasta este calculată utilizându-se numărul de corespondențe valide. *Încrederea* cea mai mare o va primi forma țintă cu cel mai mare număr de corespondențe relativ la forma căutată. Spre exemplificare, în partea dreaptă a figurii 3.11 este prezentată forma țintă din baza de date cu cea mai mare *încredere* (confidență) relativ la modelul căutat (aflat în partea stângă). Cu culoare verde sunt marcate corespondențele corecte, în timp ce, cu culoare roșie sunt marcate corespondențele eliminate. Din totalul de 214 puncte utilizate la reprezentarea formei căutate, 49 dintre acestea prezintă corespondențe valide înainte de etapa de filtrare. După această etapă, dintre aceste 49 de corespondențe doar 28 au fost considerate valide. Raportând această valoare la numărul total de puncte, rezultă o confidență de 13.08%. Următoarea formă țintă din punctul de vedere al similarității prezintă o confidență de 5.13%. Formele țintă supuse procesului de căutare au fost scalate iar granulația dintre punctele acestora a fost menținută între 1mm și 3mm.

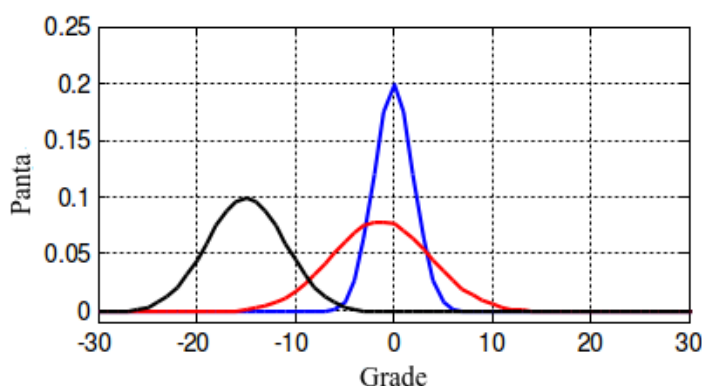
Motorul de căutare al formelor nu necesită exclusiv utilizarea unor forme globale. Algoritmul poate fi utilizat cu succes și la identificarea unor regiuni anumite dintr-o formă mai mare. Astfel, motorul de căutare a formelor poate identifica porțiuni din clădiri, regiuni din obiecte cu dimensi-

<sup>4</sup>Eng. *Probabilistic Density Function*

<sup>5</sup>Eng. *Maximum Likelihood Estimate*



(a)



(b)

Fig. 3.10 Filtrarea corespondențelor utilizând filtrul probabilistic de densitate: (a) distribuția pantelor de-a lungul celor trei axe; (b) modelele estimate ale pantelor.

uni mari sau chiar fizionomii umane. În figura 3.12 se pot observa o serie de exemple de căutare a unor astfel de regiuni particulare. Cu culoare albastră sunt desenate punctele ce aparțin formei țintă, iar cu culoare roșie sunt reprezentate punctele formei căutate. Modelul din partea stângă a fiecărei figuri reprezintă modelul căutat iar modelul din partea dreaptă reprezintă modelul din baza de date a cărui *încredere* (confidență) este cea mai mare.

### 3.3 Concluzii

În cadrul acestui capitol au fost prezentate două dintre cele mai uzuale metode de estimare a volumelor obiectelor vizualizate unidirecțional. Abordarea problemei reconstrucției obiectelor vizualizate, în aceste cazuri, presupune potrivirea (alinieră) unui model rigid peste informația obiectului percepută din scenă. În funcție de modul de generare al modelului, metodele tratează abordări diferite. Cele două metode descrise în cadrul prezentului capitol au fost: estimarea volumetrică 3D utilizând forme implicite, respectiv estimarea volumetrică 3D utilizând forme predefinite.

Prima metodă de estimare volumetrică, descrisă în detaliu în cadrul secțiunii 3.1, generează o formă rigidă cu suprafață regulată, cunoscută și sub numele de formă supercuadrică, direct în locația unde se află forma incompletă (ce trebuie reconstruită). Fiind o formă implicită, geometria acesteia este ușor controlabilă prin intermediul a 11 parametrii. Modul în care acești parametrii pot fi determinați a fost prezentat în cadrul secțiunii 3.1.2. Sunt obținute în acest sens forme 3D

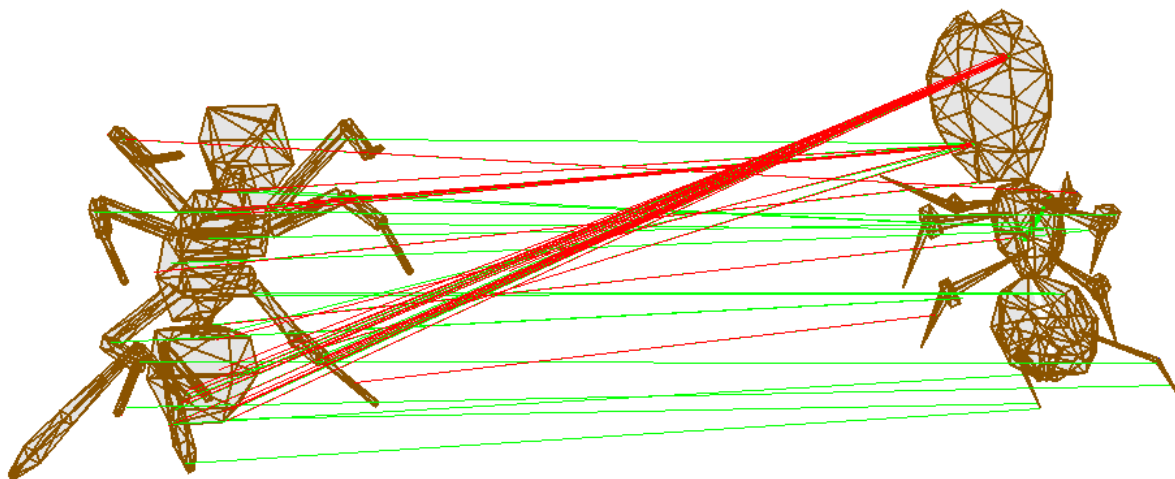


Fig. 3.11 Corespondențele stabilite între forma căutată și forma țintă cea mai asemănătoare din baza de date. Cu culoare roșie sunt reprezentate corespondențele eliminate în timp ce cu culoare verde sunt marcate corespondențele valide.

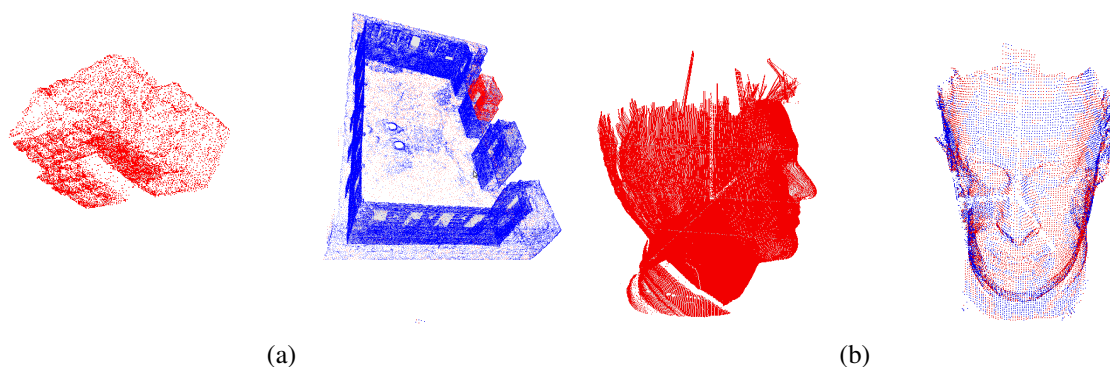


Fig. 3.12 Diferite aplicații în care pot fi utilizate motoarele de căutare a formelor: (a) identificarea unei regiuni particulare dintr-o clădire; (b) identificarea fizionomiei unei fețe umane.

înfățișând elipse, cuburi, cilindrii, toroizi, paralelipipezi etc. Datorită faptului că forma rigidă este generată "direct la fața locului" în funcție de informația percepută a obiectului de interes, sistemul de reconstrucție volumetrică nu necesită niciun fel de informație a priori despre scenă sau obiectul vizualizat, și implicit nu este nevoie de nicio bază de date cu forme. Timpul de reconstrucție mic, precum și simplitatea aparatului matematic, ce a fost descris în secțiunea 3.1.1, fac din această metodă una dintre cele mai bune abordări pentru reconstrucția volumetrică a obiectelor simple, a căror formă poate fi ușor aproximată prin intermediul unor forme geometrice de bază (sfere, cuburi, cilindri, etc.). Aplicarea acestei metode pe obiecte cu geometrie complexă (deformări puternice ale suprafeței) sau pe obiecte ce prezintă articulații, nu este recomandată datorită generalității, mult prea mare, a formelor geometrice utilizate. Totuși, pentru aplicarea metodei de reconstrucție și în aceste cazuri, se poate recurge la o partiționare obiectului vizualizat într-o serie de regiuni cu dimensiune redusă. Datorită dimensiunii reduse, aceste regiuni au un caracter regulat, și implicit, în fiecare astfel de regiune poate fi aliniată (inserată) o formă supercuadrică. Cumulând toate aceste forme, din toate regiunile în care a fost partiționat obiectul de interes, se poate obține o reprezentare rigidă cu o acuratețe net mai ridicată decât aceea obținută în cazul aproximării obiectului de interes

doar prin intermediul unei singure forme supercuadrice. Nu în ultimul rând, pentru a extinde capabilitatea formelor supercuadrice de a aproxima o gamă mai largă de forme, în secțiunea 3.1.3 a fost prezentată o soluție prin care formele supercuadrice pot fi modelate.

Secțiunea 3.2 a prezentului capitol a fost rezervată descrierii metodei de reconstrucție volumetrică bazată pe forme predefinite. Aceasta, prezintă o abordare diferită de aceea a formelor supercuadrice. În contrast cu metoda anterioară, formele utilizate pentru a reconstrui obiectul vizualizat nu sunt generate implicit ci, sunt depozitate într-o bază de date cu forme. Dintre toate formele din baza de date, aceea care prezintă cele mai multe similarități (de natură geometrică) cu forma incompletă vizualizată (din scenă) va fi folosită pentru a reprezenta rigid obiectul vizualizat. Un atu important al acestei abordări este precizia. Dacă obiectul vizualizat (perceput) prezintă un omolog cu formă rigidă în baza de date cu forme, atunci acuratețea procesului de reconstrucție volumetrică este foarte ridicată. Pentru a putea fi aplicat acest principiu de reconstrucție, suprafața obiectelor trebuie să fie reprezentată prin intermediul unor descriptori 3D. În secțiunea 3.2.1 au fost prezentați uni dintre cei mai uzuali descriptori precum și contextul în care aceștia oferă eficiență maximă. Totuși, datorită procesului foarte laborios de identificare a similarităților dintre descriptorii formelor, timpul de reconstrucție prezintă un dezavantaj major. Reprezentarea tuturor formelor implicate prin intermediul descriptorilor precum și algoritmul brut de "potrivire" al formelor în vederea identificării corespondențelor dintre acestea fac total nefezabilă reconstrucția volumetrică în timp-real. O paralelizare a procesului de identificare a corespondențelor ar putea reduce semnificativ timpul final de reconstrucție. Totuși, reducere a acestui timp, până la pragul de timp-real, este foarte dificil de obținut.

## 4. Reconstrucția volumetrică 3D a obiectelor utilizând primitive generice

---

Primitive generice  
Structura sistemului de reconstrucție volumetrică  
Recunoașterea obiectelor din scenă  
Transformarea de similaritate a primitivei generice  
Concluzii

---

Reconstrucția și segmentarea, de către un sistem de calcul, a scenelor complexe nu sunt triviale, în special atunci când scenele sunt privite doar dintr-o singură perspectivă. Avându-se în vedere complexitatea geometrică a scenei și incertitudinea introdusă de vizualizarea unidirecțională, îndeplinirea cu succes a unei sarcini poate fi greu de atins. De exemplu, considerându-se ca sarcină primară prinderea și manipularea unui anumit obiect din scenă, metoda de reconstrucție a obiectelor, ce asigură această capabilitate, trebuie să producă o reprezentare rigidă precisă a acestuia ce permite calcularea celor mai sigure și echilibrate configurații de prehensare [129].

În capitolul anterior au fost prezentate două dintre cele mai utilizate metode de estimare a volumului unui obiect de interes, fie el vizualizat chiar și dintr-o singură perspectivă. Performanțele acestor metode ating în mare parte cerințele unei acțiuni de prindere și manipulare din cadrul unui scenariu robotizat (acuratețe, timp de calcul redus, stabilitate, echilibru, etc.). Motivată de atingerea, cu o acuratețe mai bună, a acestor caracteristici, în capitolul curent este prezentat un nou concept de forme rigide 3D denumite simbolic *primitive generice* [31–33]. Capacitatea primitivelor generice de a aproxima un număr ridicat de obiecte precum și simplitatea utilizării lor fac din această metodă o alternativă viabilă la metodele tradiționale de reconstrucție și segmentare a

scenelor tridimensionale [33].

Principiul de aproximare al volumului obiectului de interes (al obiectului perceput incomplet) este asemănător cu cel al metodelor prezentate în capitolul anterior. Similar, obiectul de interes din scenă este identificat și segmentat urmând ca peste informația percepută a acestuia să fie aliniată o formă rigidă 3D (o primitivă generică) pentru a produce un estimat volumetric al obiectului de interes. Ulterior, pentru a crește asemănarea dintre volumul estimat și informația percepută din scenă se poate utiliza un proces de modelare. În cadrul prezentei cărți a fost adresată o tehnică de modelare derivată din conceptul *contururilor active* [1, 36]. Metoda a fost modificată pentru a putea fi aplicată și suprafețelor tridimensionale din spațiul cartezian [33]. Conform acestui principiu, forma primitivei generice este împinsă către informația vizuală a obiectului de interes căpătând astfel particularitățile acestuia. Atât conceptul de primitivă generică cât și abordarea deformării acesteia prin intermediul metodei propuse (ce va fi prezentată în detaliu în capitolul următor) sunt aspecte noi în literatura de specialitate și ele reprezintă nucleul prezentei cărți [33]. În continuare va fi descris în detaliu conceptul de primitivă generică precum și procesul prin care aceasta poate fi utilizată pentru a reprezenta volumetric obiectele din scena de lucru.

Abordarea prezentată în acest capitol este una similară cu modul uman de percepție al obiectelor. În acțiunile întreprinse, o persoană utilizează "concepte" (prototipuri) ale lucrurilor cu care intră în contact și nu forme particulare ale acestora. De exemplu, când aceasta dorește să utilizeze o "cană", caută să identifice în scenă o formă generică a acelei căni și nu una particulară (cana verde cu formă ovală, puțin ciobită și cu mânerul bombat), urmând ca, după identificare, să poată face mai multe observații asupra acesteia și să o poată particulariza astfel încât să o utilizeze în siguranță. În metoda propusă, prototipul obiectului de interes este definit de primitiva generică a clasei de obiecte considerată, iar particularizarea acesteia, pe baza observațiilor din scenă, este obținută prin utilizarea unei metode de modelare.

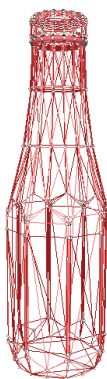
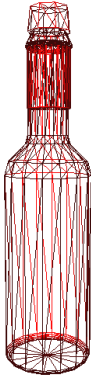
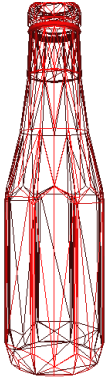
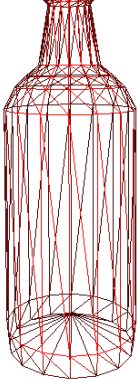
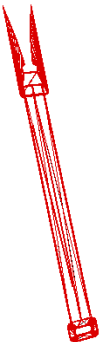
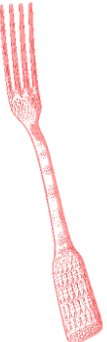
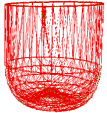
#### 4.1 Primitive generice

După cum a fost descris în secțiunea introductivă a prezentului capitol, primitivele generice sunt un nou tip de forme rigide a căror înfățișare se află la granița dintre generalitate și particularitate. Cu alte cuvinte, o primitivă generică nu este atât de generală încât să poată aproxima orice obiect ce poate apărea în scenă (așa cum este cazul formelor supercuadrice), și, în același timp, nu este nici îndeajuns de particulară încât să poată aproxima, cu precizie ridicată, un obiect particular din scenă. Pentru a putea explica mai clar conceptul de primitivă generică trebuie menționat termenul de *clasă de obiecte*. O clasă de obiecte reprezintă o colecție de forme particulare a căror înfățișare este similară. Toate aceste forme particulare nu sunt altceva decât diferite aspecte ale unui obiect de bază. Pentru a susține această afirmație, în cadrul tabelului 4.1 pot fi observate 4 astfel de clase de obiecte. Spre exemplificare, obiectul "stică" prezintă 6 aspecte diferite dar care în esență descriu același obiect de bază, și anume, "sticlă". Numărul de înfățișări din fiecare clasă de obiecte poate varia, în funcție de câte aspecte poate avea obiectul considerat. Pentru a evita ambiguitatea legată de ce înfățișare optimă din clasa de obiecte considerată se potrivește cu aceea percepută din scenă, pentru fiecare clasă de obiecte în parte se definește doar o singură primitivă generică a cărei formă este o formă medie a înfățișărilor din clasa considerată. Astfel, o primitivă

este îndeajuns de generică pentru a aproxima orice obiect dintr-o clasă de obiecte și, în același timp, este îndeajuns de particulară pentru a aproxima, cu precizie, doar o anumită clasă de obiecte (de ex. sticlă, furculiță, cană, pantof, etc.). Primitivele generice ale claselor obiectelor ilustrate în tabelul 4.1 în pot fi vizualizate în tabelul 4.2(c, f, i).

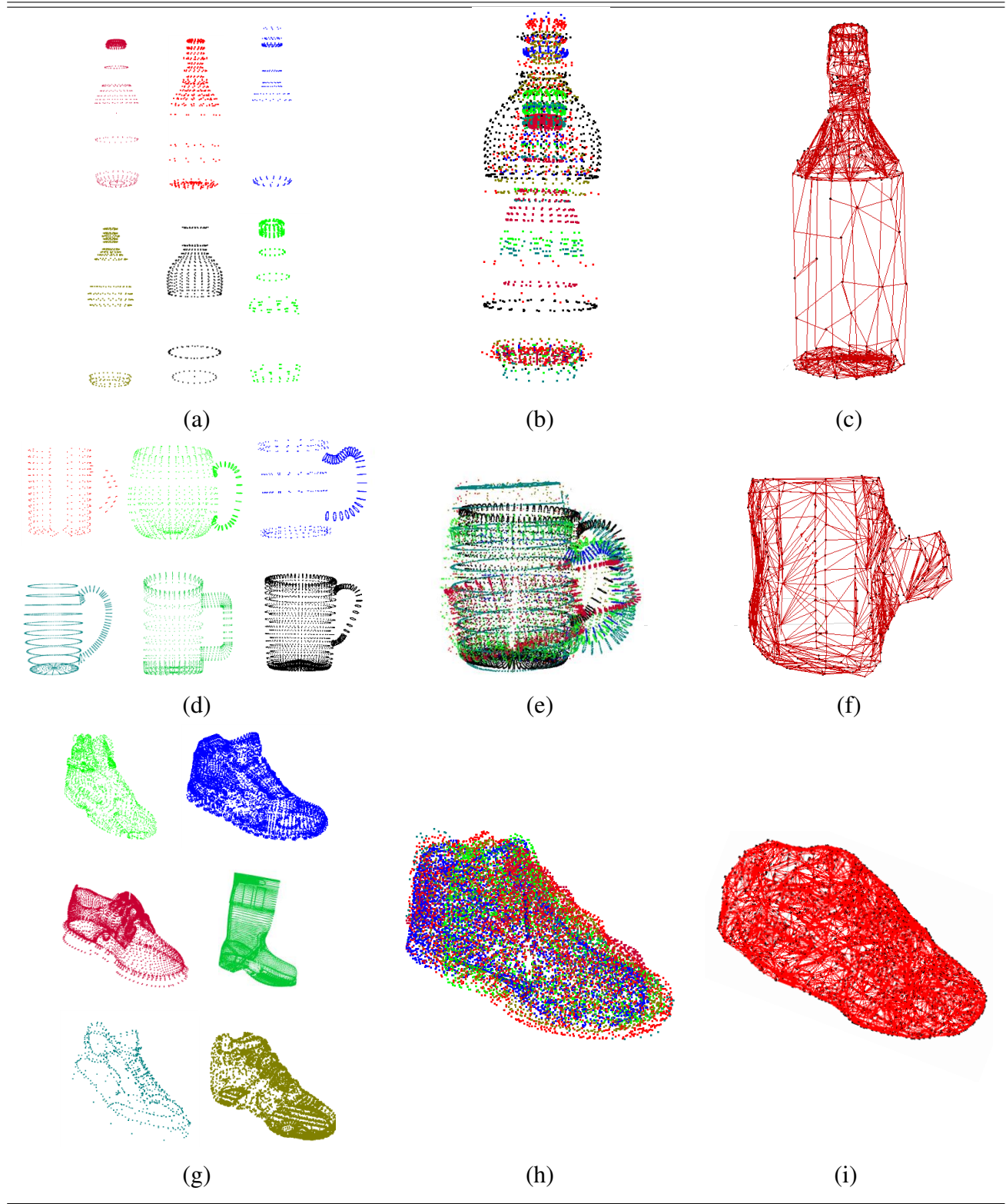
Utilizând conceptul nou introdus, pentru un obiect din scenă, obiect alocat unei anumite clase de obiecte, va exista doar o singură formă rigidă candidată ce va fi utilizată în procesul de estimare volumetrică și anume aceea a primitivei generice.

Tab. 4.1 Diferite volume ale unor obiecte de uz general.

|                  |           |                                                                                     |                                                                                     |                                                                                     |                                                                                      |                                                                                       |                                                                                       |
|------------------|-----------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------|
| Clasă de obiecte | Sticlă    |   |   |   |   |   |   |
|                  | Furculiță |  |  |  |  |  |  |
|                  | Cană      |  |  |  |  |  |  |
|                  | Pantof    |  |  |  |  |  |  |

Totuși, datorită formei comune, o primitivă generică nu poate fi similară întotdeauna cu orice formă din clasa de obiecte. Poate exista oricând un obiect cu o formă complexă ieșită din comun care să fie disimilară relativ la primitiva generică aferentă. O rezolvare robustă în acest sens poate fi obținută prin aplicarea unui proces de modelare a primitivei generice ce are rolul de a aduce forma comună a primitivei cât mai aproape de forma obiectului considerat din scenă.

Tab. 4.2 Determinarea formelor medii pentru diferite clase de obiecte: (a)(d)(g) Nori de puncte ce înfățișează diferite obiecte din clasele sticlă, cană, respectiv pantof; (b)(e)(h) Nori de puncte suprapuși ce însumează toate înfățișările claselor considerate; (c)(f)(i) forme medii (primitive generice) ale claselor de obiecte sticlă, cană și pantof (reprezentate prin intermediul unor rețele triangulare de culoare roșie);



### 4.1.1 Determinarea primitivei generice a unei clase de obiecte

După cum se poate observa în tabelul 4.1 și tabelul 4.2(a, d, g) obiectele pot avea înfățișări foarte diferite. Totuși, o formă comună (similară) relativ la toate formele dintr-o anumită clasă poate fi obținută utilizând un proces complex de analiză intitulat *Analiză Procrustes* (v. sec. 2.6)[18, 69]. Scopul acestuia este de studia distribuția formelor unei clase de obiecte pentru a se identifica posibilele similarități dintre acestea. Mai departe, utilizând analiza Procrustes generalizată, aceste similarități pot fi utilizate pentru a se determina o formă medie identificată ca fiind suma tuturor formelor supuse analizei. Ulterior, această formă medie este considerată a fi primitiva generică a clasei de obiecte studiate.

Fie  $\mathcal{U}$  un set de forme particulare  $\mathcal{U}_1, \mathcal{U}_2, \dots, \mathcal{U}_n$  ce descriu diverse forme ce aparțin aceleiași clase de obiecte. Fiecare astfel de formă este descrisă prin intermediul unui nor dens de puncte 3D. Prin interconectarea acestor puncte se poate obține o rețea triunghiulară de puncte care, în esență, poate descrie o suprafață. Algoritmul care stă la baza determinării formei medii a unei primitive generice poate fi enunțat astfel:

1. se alege arbitrar o formă  $\mathcal{U}_i$ , unde  $i = 1, \dots, N_{\mathcal{U}}$ , care va fi considerată formă de referință curentă;
2. se suprapun celelalte forme peste forma de referință curentă (v. tab. 4.2(b, e, h));
3. se calculează forma medie a setului curent de forme aliniate la pasul anterior (v. tab. 4.2(c, f, i));
4. dacă distanța Procrustes (v. sec. 2.6.1.1) dintre forma medie determinată și forma de referință curentă este mai mare decât un anumit prag definit de utilizator, se consideră ca formă de referință forma medie obținută și se trece la pasul 2, iar în caz contrar forma medie obținută va fi utilizată pentru a reprezenta clasa de obiecte considerată.

Procesul de suprapunere (alinieri) a formelor din  $\mathcal{U}$  produce un nor dens de puncte tridimensionale similar celor prezentate în tabelul 4.2(b, e, h). În centrul acestui nor se află și forma de referință candidată la titlul de formă medie. Rând pe rând, punctele acesteia sunt deplasate (de-a lungul direcției normalei la suprafață) pentru a fi poziționate la mijlocul densității de puncte (descrise de formele predefinite suprapuse din clasă) de-a lungul direcției considerate. În final, după ce toate punctele formei de referință au fost repositionate, aceasta poate fi considerată ca fiind forma medie curentă. O descriere ilustrată a deplasării unui anumit punct poate fi observată în figura 4.1.

Formele medii (primitivele generice) prezentate în tabelul 4.2(c, f, i) sunt descrise, pentru o vizualizare mai elocventă, sub forma unei rețele triunghiulare de puncte. După cum se poate observa, primitivele generice rezultate au o formă comună relativ la clasa de obiecte considerată. Având în vedere baza de date cu forme descrisă în [158], din totalul de 1814 de forme particulare ce au fost supuse analizei Procrustes au fost obținute 142 de primitive generice, fiecare definind o clasă de obiecte individuală.

### 4.1.2 Puncte de control

Ca orice altă formă tridimensională, forma unei primitive generice este descrisă prin intermediul unui nor de puncte 3D. Densitatea (granularitatea) punctelor norului poate varia în funcție

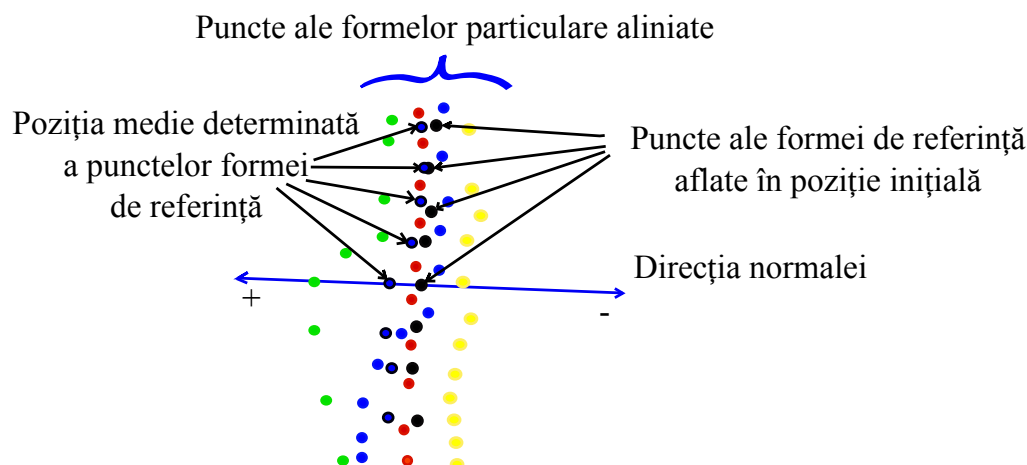


Fig. 4.1 Determinarea poziției medii a punctelor forme de referință în raport cu punctele vecine ce aparțin formelor particulare suprapuse (vedere în secțiune prin norul de puncte al formelor suprapuse).

de dimensiunea obiectului considerat precum și în funcție de complexitatea suprafeței acestuia. Totodată, densitatea norului de puncte poate influența timpii de procesare și analiză ai primitivei generice. Una dintre cele mai costisitoare etape în acest sens este modelarea primitivei. Considerând norul dens de puncte 3D, utilizat la reprezentarea unei primitive generice, modelarea fiecărui punct 3D ar putea conduce la un timp de procesare dincolo de performanța de timp-real. Dacă, generarea unei primitive generice se realizează offline, particularizarea acesteia, pe baza observațiilor din scenă, este obligatorie a se realiza online.

În scopul eficientizării procesului de modelare, punctele 3D ce descriu primitiva generică pot fi împărțite în două clase, după cum urmează:

- *punctele de control* ( $p_c$ ), ce definesc scheletul (coloana vertebrală) primitivei generice considerate;
- *punctele normale sau puncte ordinare* ( $p_o$ ), ce compun majoritatea punctelor din primitivă și au rolul doar de a crea o suprafață netedă prin interpolarea punctelor de control.

Acest mod de clasificare a punctelor este intens utilizat în domeniul medical, în special la modelarea în timp real a organelor umane [192].

Având în vedere principiul de clasificare prezentat anterior, norul de puncte redefinit al primitivei generice ce descrie un pantof poate fi observat în figura 4.2. Atribuirea proprietății de punct de control, respectiv ordinar se poate realiza manual sau automat. În procedura de atribuire manuală se utilizează un program software specializat, prin care, un operator specifică explicit proprietatea fiecărui punct în parte. Această abordare, deși este consumatoare de timp, este considerată a fi foarte precisă. Totuși, datorită implicării factorului uman, este caracterizată de subiectivism.

Pe de altă parte, procedura de atribuire automată este rapidă, obiectivă și nu implică sub nicio formă factorul uman. Proprietățile punctelor sunt acordate pe baza unui set de reguli derivate din metoda formelor active prezentată în [36]. Aceste reguli determină dacă un punct poate fi catalogat ca fiind ordinar sau de control. Astfel, un punct al primitivei generice este considerat a fi de control dacă respectă una din următoarele reguli:

1. marchează centrul unei anumite regiuni;

2. este localizat pe o muchie ascuțită sau pe o graniță;
3. este localizat pe extremitatea unei curbe;
4. este punct de inflexiune al unei curbe;
5. este cel mai sudic punct al primitivei (pentru acest caz, primitiva este rotită constant de-a lungul celor trei axe carteziane pentru a se obține punctele sudice din toate perspectivele de vizualizare posibile);
6. este situat la jumătatea distanței de-a lungul unei granițe definite de două puncte ce satisfac regula 1.

Punctele ce nu satisfac condițiile enunțate anterior sunt implicit considerate puncte ordinare.

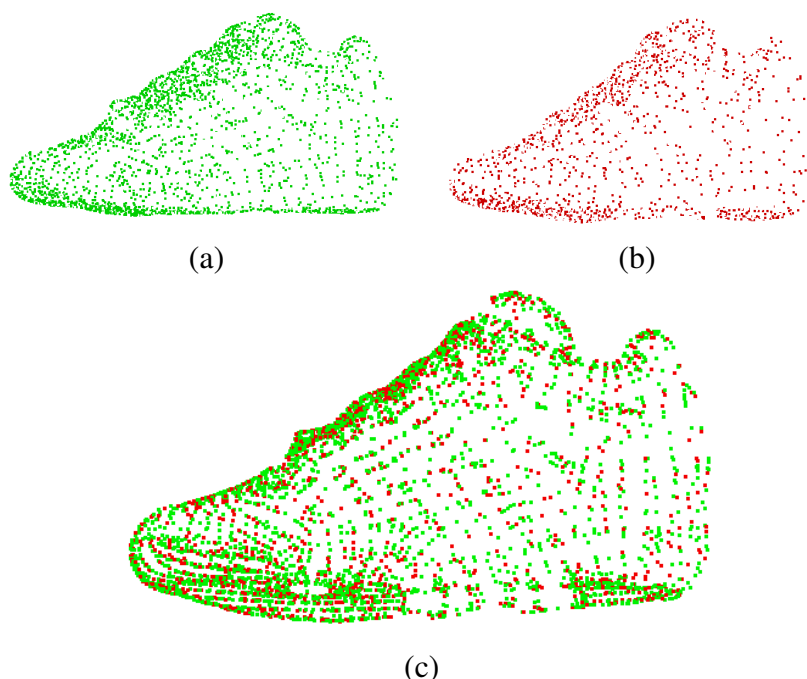


Fig. 4.2 Clasificarea punctelor primitivei generice a unui pantof: (a) nor de puncte înfățișând punctele ordinare (culoare verde, 3691 puncte); (b) nor de puncte înfățișând punctele de control (culoare roșie, 1284 puncte); (c) nor de puncte complet al primitivei unui pantof (4975 puncte).

#### 4.1.3 Structura de date a unei primitivei generice

Structural, o primitivă generică, notată în continuare cu  $\mathcal{M}_l$ , se compune dintr-o serie elemente după cum urmează: un nor de puncte  $\mathcal{P}^{\mathcal{M}}$  ce stochează coordonatele Euclidiene ale tuturor punctelor ce definesc primitiva generică considerată (v. fig. 4.2(c)); un vector  $\mathcal{V}$  ce conține proprietățile, ordinar, respectiv de control (v. sec. 4.1.2) ale punctelor norului  $\mathcal{P}^{\mathcal{M}}$ ; o matrice  $\mathcal{F}$ , de dimensiune  $3 \times N_{\mathcal{P}^{\mathcal{M}}}$  (unde  $N_{\mathcal{P}^{\mathcal{M}}}$  este numărul total de puncte folosit la reprezentarea norului  $\mathcal{P}^{\mathcal{M}}$ ), matrice  $\mathcal{F}$  ce stochează punctele 3D ce descriu fețele rețelei triangulare (v. tab. 4.2(c, f, i)), utilizată pentru reprezentarea volumetrică a lui  $\mathcal{P}^{\mathcal{M}}$ ; și o serie de informații cu privire la proprietățile globale ale primitivei, cum ar fi lungimea  $w$ , lățimea  $l$ , înălțimea  $h$  sau transformarea de similaritate ( $\mathcal{T}$ ) (factorul de scalare  $s$ , rotația  $\mathbf{R}$  sau translația  $\mathbf{t}$ ). Matematic, primitiva generică  $\mathcal{M}_l$  poate fi descrisă după cum urmează:

$$\mathcal{M}(\mathcal{P}^{\mathcal{M}}, \mathcal{Y}, \mathcal{F}, A) = \begin{cases} \mathcal{P}^{\mathcal{M}} = (\mathbf{p}_1^{\mathcal{M}}, \dots, \mathbf{p}_n^{\mathcal{M}}), \\ \mathcal{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_n), & \text{unde, } \mathbf{y}_i = \begin{cases} 0 & \text{punct ordinar;} \\ 1 & \text{punct de control;} \end{cases} \\ & i = 1, \dots, n; \\ \mathcal{F} = (\mathbf{f}_1, \dots, \mathbf{f}_l), & \text{unde, } \mathbf{f}_j = (\mathbf{p}_m^{\mathcal{M}}, \mathbf{p}_n^{\mathcal{M}}, \mathbf{p}_o^{\mathcal{M}}); \\ & j = 1, \dots, l, i \neq j; \\ & m, n, o = 1, \dots, n; \\ A = \{l, h, w, \mathcal{T}\}. \end{cases} \quad (4.1)$$

Se notează în continuare  $\mathcal{S}_l$  ca fiind primitiva generică  $\mathcal{M}_l$  aliniată relativ la norul de puncte obiect  $\mathcal{O}_k$  (perceput). Această notație face diferența dintre o primitivă generică aflată la o poziție și orientare aleatoare în spațiu și una aliniată optim, relativ la obiectul de interes din scenă. Procesul de aliniere implică aplicarea cu succes a unei operații de transformare de similaritate.

#### 4.2 Structura sistemului de reconstrucție volumetrică

Având ca obiectiv final reconstrucția volumetrică 3D a obiectelor vizualizate, în figura 4.3 se poate observa schema-bloc a sistemului de reconstrucție volumetrică propus.

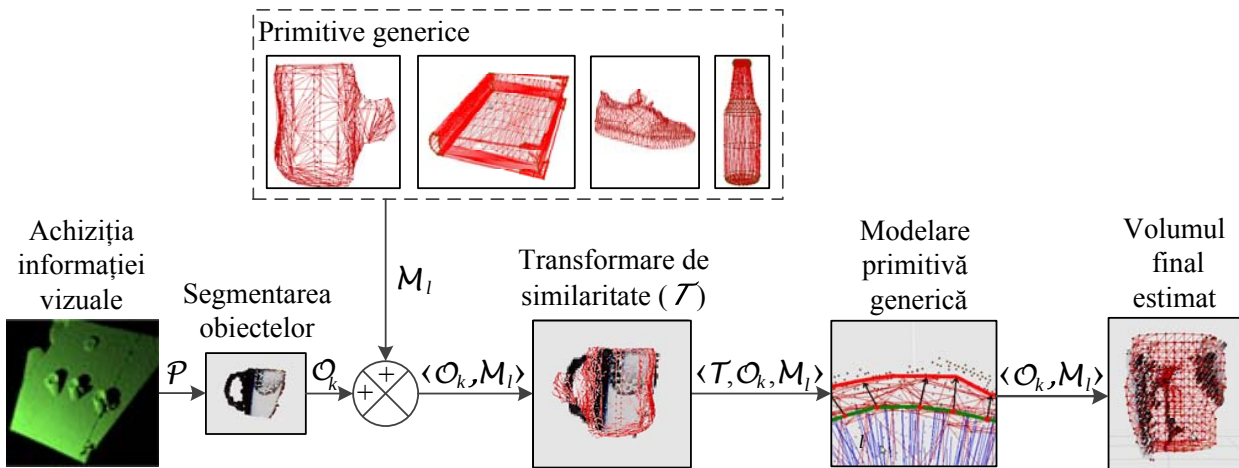


Fig. 4.3 Schema-bloc a sistemului de reconstrucție volumetrică propus.

Rolul primei etape este acela de a achiziționa informația vizuală ce descrie scena de lucru. Tot aici, informația vizuală este convertită într-un nor dens de puncte tridimensionale ( $\mathcal{P}$ ) ce descrie în adâncime scena considerată. În cea de-a doua etapă, norul de puncte 3D  $\mathcal{P}$ , este partiționat obținându-se  $\mathcal{O}$  grupuri compacte de puncte. Fiecare grup  $\mathcal{O}_k \in \mathcal{O}$ , iar  $\mathcal{O} \in \mathcal{P}$ ,  $k = 1, \dots, N_k$ , unde  $N_k$  reprezintă numărul total de grupuri compacte identificate în scenă. Norul de puncte utilizat pentru descrierea grupului compact de puncte poartă denumirea, așa cum a mai fost menționat, de nor de puncte obiect. O observație importantă este aceea că distanța dintre punctele ce compun

grupul de puncte  $\mathcal{O}_k$  nu poate fi mai mare decât un anumit prag impus de proces. În acest mod, fiecare grup  $\mathcal{O}_k$  descrie o entitate distinctă din scenă (v. sec. 2.5) [83]. Mai departe, identificarea și selectarea a grupului compact de interes  $\mathcal{O}_k \in \mathcal{P}$ , care satisface sarcina impusă, se realizează utilizându-se un proces de clasificare care va fi prezentat în secțiunea următoare [59]. La sfârșitul acestei etape, sistemul de reconstrucție dispune atât de norul de puncte 3D  $\mathcal{O}_k$  ce descrie doar obiectul de interes cât și de eticheta acestuia ce definește contextual clasa din care face parte (cană, sticlă, tacâm etc.).

În cea de-a treia etapă, este selectată, dintr-o serie  $\mathcal{M}$  de primitive generice (bază de date cu forme), primitiva generică  $\mathcal{M}_l$  ce descrie optim clasa obiectului considerat, unde  $l = 1, \dots, N_M$ , iar  $N_M$  reprezintă numărul total de primitive din baza de date. În continuare, norul de puncte  $\mathcal{P}^{\mathcal{M}}$  al primitivei  $\mathcal{M}_l$  este aliniat relativ la norul de puncte obiect  $\mathcal{O}_k$  utilizându-se o *transformare de similaritate* ( $\mathcal{T}$ ). În urma acestei alinieri, se obține un estimat al volumului obiectului de interes  $\mathcal{O}_k$ , a cărui formă este definită de primitiva generică aferentă. Particularizarea lui  $\mathcal{P}^{\mathcal{M}}$  se obține printr-un proces de modelare. Astfel, în decursul etapei numărul patru a lanțului de procesare, primitiva  $\mathcal{S}_l$  este deformată pentru a o aduce cât mai aproape de forma percepută a acestuia. În final, norul de puncte  $\mathcal{P}^{\mathcal{M}}$  al primitivei generice  $\mathcal{S}_l$ , ce aproximează rigid obiectul vizualizat, descrie un volum comprehensibil ce poate fi utilizat cu încredere pentru îndeplinirea diferitelor sarcini precum, analiza complexă a scenei, prehensarea sau manipularea obiectelor etc.

### 4.3 Recunoașterea obiectelor din scenă

Mecanismul de identificare (recunoaștere) a obiectelor din scenă are la bază un clasificator [122, 180]. Această tehnică folosește principiul conform căruia obiectele dintr-o mulțime dată (scenă) sunt distribuite în submulțimi numite *clase* în funcție de asemănările și deosebirile dintre ele. O clasă identifică, în mod unic, un anumit obiect. Compararea obiectelor se face dintr-un anumit punct de vedere, numit *criteriu* [125]. În domeniul vederii artificiale, aceste criterii sunt identificate ca fiind caracteristici extrase din spațiul imaginii sau din spațiul tridimensional [4]. Un exemplu de caracteristici extrase din spațiul imaginii sunt *caracteristicile Haar* [122]. Astfel, pentru fiecare obiect din scenă detectat clasificatorul formează un set de regiuni de interes (rectangulare) numite și *caracteristici*. Acestea sunt definite în interiorul unei ferestre numite *fereastră de detectare* care mărginește obiectul supus analizei de clasificare. Mai departe, pentru fiecare astfel de regiune se calculează suma tuturor pixelilor din interior. Diferența dintre sumele diferitelor ferestre formează o caracteristică de tip Haar. Acest tip de caracteristici se mai numesc și *caracteristici rectangulare de ordin 2*. Derivate din acestea, se mai pot defini caracteristici rectangulare de ordinele 3 și 4 [180]. Valoarea unei astfel de caracteristici indică o proprietate importantă a unei arii din interiorul imaginii de intrare. Astfel, fiecare caracteristică semnalează existența, respectiv absența unei muchii, a modificării de textură sau a unor colțuri. Datorită naturii caracteristicilor Haar, clasificatorul poate fi aplicat doar pe imagini gri.

Caracteristicile Haar astfel detectate sunt comparate, în continuare, cu o serie de caracteristici "învățate" anterior (în urma unui proces de *antrenare*) care au rolul de a separa entitățile obiect de cele non-obiect. Pentru o acuratețe ridicată, numărul determinat de astfel de caracteristici trebuie să fie ridicat. Spre exemplificare, pentru cazul câinii din figura 4.4 (centru) au fost determinate 192

de caracteristici Haar normalizate în intervalul  $[-1, 1]$ . În [180], pentru obținerea unui clasificator mai puternic, caracteristicile sunt organizate în cascadă. Un atu important al acestui mod de recunoaștere este viteza de calcul care este ridicată.

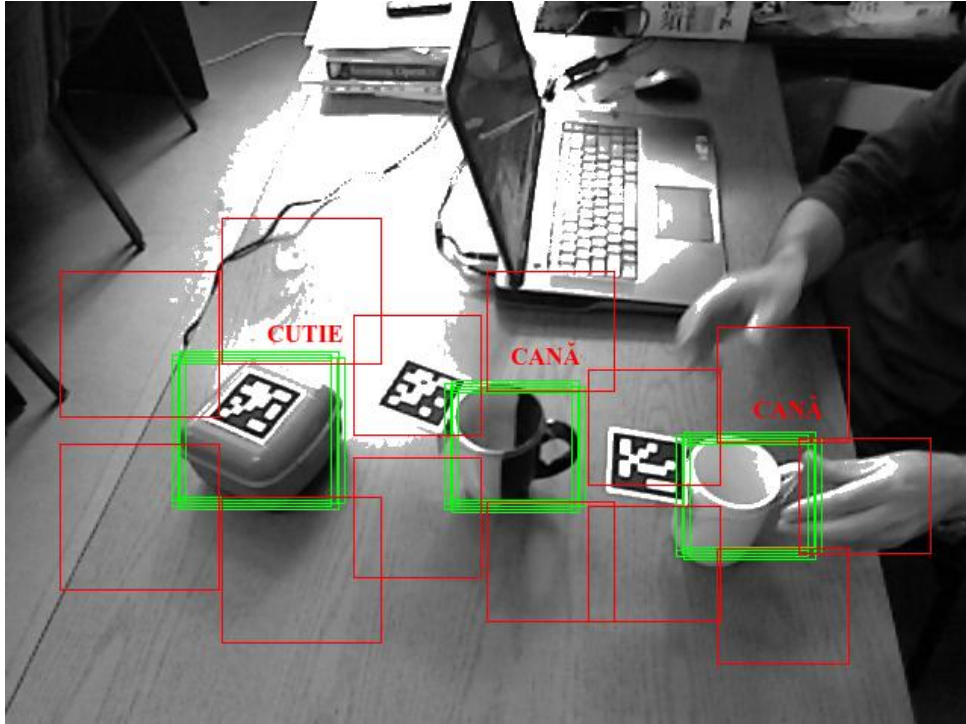


Fig. 4.4 Clasificarea obiectelor din scenă. Cu forme rectangulare de culoare roșie sunt identificate caracteristicile non-obiect (fundal), în timp ce, cu forme rectangulare de culoare verde sunt ilustrate caracteristicile Haar aparținând obiectelor considerate.

La finalul procesului de clasificare se obține o măsură a apartenenței obiectului vizualizat la o anumită clasă (cană, sticlă, carte, pantof, etc.). Astfel, o valoare mică a acestei confidence indică faptul că obiectul nu a putut fi alocat cu siguranță ridicată unei anumite clase, în timp ce, o valoare mare a acesteia denotă o etichetare corectă și de încredere. Rolul identificării clasei obiectului este strict necesară deoarece, pe baza acesteia se va face selectarea primitivei generice din baza de date.

#### 4.4 Transformarea de similaritate a primitivei generice

Deoarece atât norul de puncte obiect  $\mathcal{O}_k$ , segmentat din scenă, cât cel al primitivei generice aferentă acestuia,  $\mathcal{P}^M$ , sunt definite în sisteme de coordonate diferite, scopul operației de transformare de similaritate este acela de a aduce cele două obiecte într-un sistem de coordonate comun. Întrucât norului de puncte obiect înglobează date *reale* (dimensiune, poziție, sau structură), sistemul de coordonate al acestuia poate fi considerat drept referință (sistem de coordonate comun) relativ la care poate fi transformată primitiva generică.

Conform definiției prezentate în secțiunea 2.3.1, problema transformării de similaritate, notată în continuare cu  $\mathcal{T}_{t_x, t_y, t_z, s, \mathbf{R}}(\mathcal{P}^M)$ , se rezumă la determinarea celor 7 grade de libertate după cum urmează:

- un grad de libertate introdus de factorul de scalare ( $s$ );

- trei grade de libertate introduse de operația de translație ( $\mathbf{t} = t_x, t_y, t_z$ );
- trei grade de libertate introduse de operația de rotire ( $\mathbf{R}$ ).

După cum a fost prezentat și în cadrul capitolului 2, ordinea în care cele 7 grade de libertate sunt determinate este scalare, translație și rotire (a primitivei generice în raport cu obiectul de interes). Matematic, transformare de similaritate este definită de magnitudinea normei Euclidiene a diferenței dintre cele două forme aliniate:

$$\mathcal{T} \doteq ||\mathcal{O}_k - s \cdot \mathbf{R} \cdot (\mathcal{P}^{\mathcal{M}} - \mathbf{t})||^2, \quad (4.2)$$

unde,  $\mathcal{O}_k$ , respectiv  $\mathcal{P}^{\mathcal{M}}$  sunt nori de puncte al obiectelor implicate în proces, iar  $s$ ,  $\mathbf{R}$ , respectiv  $\mathbf{t}$  sunt operațiile procesului de transformare de similaritate. Având în vedere relația 4.2, factorul de scalare  $s$  are rolul de a aduce dimensiunea norului de puncte  $\mathcal{P}^{\mathcal{M}}$  la dimensiunea de referință dată de norul obiect  $\mathcal{O}_k$ . Quantumul factorului de scalare este determinat urmându-se etapele: se circumscriu, pe rând, ambii nori de puncte, aflați încă în sistemele de coordonate proprii, în câte o sferă separată (v. fig. 4.5); se ajustează continuu magnitudinea razelor celor două sfere până în momentul în care punctele fiecărui nor, în parte, se află în interiorul sferei aferente; cea mai mică rază care satisface condiția prezentată anterior este considerată a descrie dimensiunea norului analizat. În continuare, raportul dintre raza sferei care circumscrie norul de puncte al primitivei generice și aceea a sferei care circumscrie norul obiect, determină valoarea factorului de scalare [31]:

$$s = \frac{r_{\mathcal{P}^{\mathcal{M}}}}{r_{\mathcal{O}_k}}, \quad (4.3)$$

unde,  $r_{\mathcal{P}^{\mathcal{M}}}$  este raza sferei care circumscrie optim norul de puncte al primitivei, iar  $r_{\mathcal{O}_k}$  este raza sferei care circumscrie optim norul de puncte al obiectului perceput. Redimensionarea lui  $\mathcal{P}^{\mathcal{M}}$  relativ la  $\mathcal{O}_k$  se realizează astfel:

$$\mathcal{P}_{scalat}^{\mathcal{M}} = s \cdot \mathcal{P}^{\mathcal{M}}, \quad (4.4)$$

unde,  $\mathcal{P}_{scalat}^{\mathcal{M}}$  reprezintă norul de puncte  $\mathcal{P}^{\mathcal{M}}$  scalat.

La nivelul fiecărui punct din  $\mathcal{P}^{\mathcal{M}}$ , factorul de scalare se poate aplica după cum urmează:

$$\begin{aligned} \mathbf{p}_{scalat}^{\mathcal{M}}(x) &= s \cdot \mathbf{p}^{\mathcal{M}}(x), \\ \mathbf{p}_{scalat}^{\mathcal{M}}(y) &= s \cdot \mathbf{p}^{\mathcal{M}}(y), \\ \mathbf{p}_{scalat}^{\mathcal{M}}(z) &= s \cdot \mathbf{p}^{\mathcal{M}}(z), \end{aligned} \quad (4.5)$$

unde,  $\mathbf{p}^{\mathcal{M}}$  este un punct al norului  $\mathcal{P}^{\mathcal{M}}$ , iar  $\mathbf{p}_{scalat}^{\mathcal{M}}$  este același punct doar că de această dată are operația de scalare aplicată asupra sa.

Cea de-a doua operație a procesului de transformare de similaritate, după scalare, este translația. Aceasta are rolul de a suprapune cei doi nori de puncte,  $\mathcal{P}^{\mathcal{M}}$  peste  $\mathcal{O}_k$  (v. fig. 4.6). Aceștia se consideră suprapuși optim atunci când centrele lor de greutate sunt identice. Pentru aceasta, trebuie calculată o matrice de translație  $\mathbf{t}$ , de dimensiune  $3 \times 1$ . Utilizând centrele de greutate ( $\mu$ ) ale

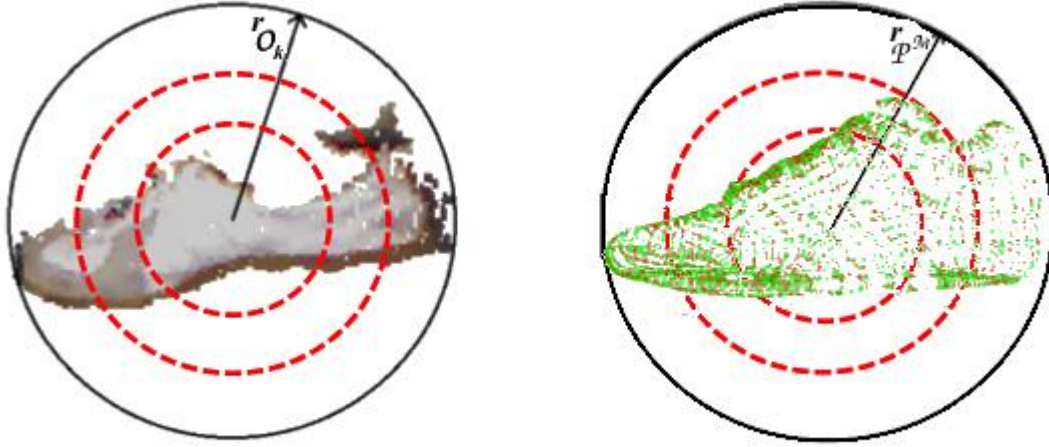


Fig. 4.5 Identificarea sferelor care circumscriu optim formele considerate. Cu culoare roșie sunt marcate sferele necorespunzătoare (vedere în secțiune).

celor două obiecte, matricea  $\mathbf{t}$  poate fi obținută în felul următor:

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \begin{bmatrix} \mu_{\mathcal{O}_k}(x) - \mu_{\mathcal{P}^M}(x) \\ \mu_{\mathcal{O}_k}(y) - \mu_{\mathcal{P}^M}(y) \\ \mu_{\mathcal{O}_k}(z) - \mu_{\mathcal{P}^M}(z) \end{bmatrix}, \quad (4.6)$$

unde:

$$\mu_{\mathcal{O}_k} = \frac{1}{N_{\mathcal{O}_k}} \cdot \sum_{i=1}^{N_{\mathcal{O}_k}} \mathbf{o}_i, \quad \mu_{\mathcal{P}^M} = \frac{1}{N_{\mathcal{P}^M}} \cdot \sum_{j=1}^{N_{\mathcal{P}^M}} \mathbf{p}_j^M, \quad (4.7)$$

în care,  $N_{\mathcal{O}_k}$ , respectiv  $N_{\mathcal{P}^M}$  reprezintă numărul total de puncte ce descrie obiectul  $\mathcal{O}_k$ , respectiv  $\mathcal{P}^M$ , iar  $\mathbf{o}_i$ , respectiv  $\mathbf{p}_j^M$  sunt puncte ale acestora.



Fig. 4.6 Suprapunerea centrelor de greutate ale norilor de puncte ce înfățișează o primitivă generică ( $\mathcal{P}^M$ ), respectiv un nor de puncte obiect ( $\mathcal{O}_k$ ). Pentru o vizualizare mai elocventă,  $\mathcal{P}^M$  este reprezentat sub forma unei rețele triangulare de puncte (linii de culoare roșie).

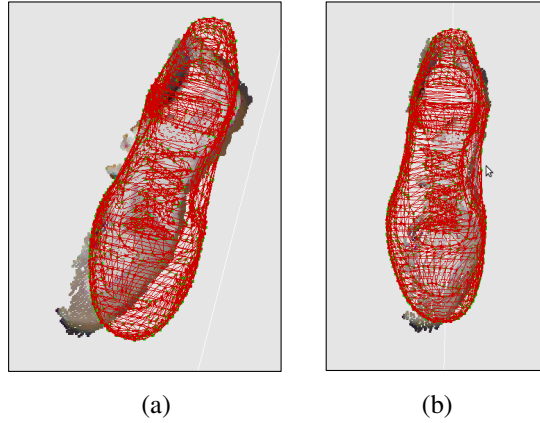


Fig. 4.7 Alinierea unei primitive generice ce înfățișează un pantof (reprezentată prin intermediul unei rețele triangulare de culoare roșie) la norul de puncte obiect aferent: (a) primitiva generică scalată, translatată și rotită brut relativ la norul de puncte  $\mathcal{O}_k$ ; (b) alinierea finală obținută după aplicarea metodei ICP.

La nivelul fiecărui punct, matricea de translație se poate aplica conform relației:

$$\mathbf{p}_{translatat}^{\mathcal{M}} = \mathbf{t} * \mathbf{p}^{\mathcal{M}} = \begin{bmatrix} t_x & 0 & 0 \\ 0 & t_y & 0 \\ 0 & 0 & t_z \end{bmatrix} \cdot \begin{bmatrix} \mathbf{p}^{\mathcal{M}}(x) \\ \mathbf{p}^{\mathcal{M}}(y) \\ \mathbf{p}^{\mathcal{M}}(z) \end{bmatrix} = \begin{bmatrix} t_x \cdot \mathbf{p}^{\mathcal{M}}(x) \\ t_y \cdot \mathbf{p}^{\mathcal{M}}(y) \\ t_z \cdot \mathbf{p}^{\mathcal{M}}(z) \end{bmatrix}, \quad (4.8)$$

unde,  $\mathbf{p}^{\mathcal{M}}$  este un punct oarecare al norului  $\mathcal{P}^{\mathcal{M}}$ , iar  $\mathbf{p}_{translatat}^{\mathcal{M}}$  reprezintă același punct doar că de această dată are operația de transformare aplicată asupra sa.

Dat fiind norul de puncte al primitivei generice  $\mathcal{P}^{\mathcal{M}}$  scalat și translatat, relativ la un nor de puncte obiect  $\mathcal{O}_k$ , următorul obiectiv al procesului de transformare de similaritate este determinarea matricei de rotație  $\mathbf{R}$ , de dimensiune  $3 \times 3$ . Acesta este determinată utilizând un proces multirezoluție care, într-o manieră repetitivă, încearcă să rotească norul  $\mathcal{P}^{\mathcal{M}}$  pentru a-l alinia cu norul  $\mathcal{O}_k$ . Evaluarea gradului de "potrivire" dintre cei doi nori de puncte se realizează utilizând suma distanțelor Euclidiene dintre punctele acestora. Cu cât cei doi nori sunt aliniați mai bine, cu atât distanța dintre punctele acestora este mai mică. Distanța  $d$  se poate determina utilizându-se următoarea relație matematică:

$$d = \arg \min_d \sum_{i=0}^{N_{\mathcal{O}_k}} \sqrt{(\mathbf{o}_i(x) - \mathbf{p}_{nn}^{\mathcal{M}}(x))^2 + (\mathbf{o}_i(y) - \mathbf{p}_{nn}^{\mathcal{M}}(y))^2 + (\mathbf{o}_i(z) - \mathbf{p}_{nn}^{\mathcal{M}}(z))^2}, \quad (4.9)$$

unde,  $N_{\mathcal{O}_k}$  este numărul de puncte ce descrie norul obiect, iar  $\mathbf{p}_{nn}^{\mathcal{M}}(x)$ ,  $\mathbf{p}_{nn}^{\mathcal{M}}(y)$  și  $\mathbf{p}_{nn}^{\mathcal{M}}(z)$  sunt coordonatele celui mai apropiat punct (din  $\mathcal{P}^{\mathcal{M}}$ ) relativ la punctul din norul obiect  $\mathbf{o}$  cu indexul  $i$ . A se remarca faptul că distanța  $d$  este o sumă a distanțelor Euclidiene dintre fiecare punct al lui  $\mathcal{O}_k$  și cel mai apropiat punct din  $\mathcal{P}^{\mathcal{M}}$ . Pentru eficiența calculului,  $\mathcal{P}^{\mathcal{M}}$  este rotit cu un increment de  $10^\circ$ . Acesta este un compromis ce asigură calcularea metrici de rotație într-un interval de timp rezonabil. Rotația optimă astfel obținută este, de fapt, o rotație brută ( $\mathbf{R}_{brut}$ ) deoarece precizia acesteia este de  $\pm 10^\circ$  (v. fig. 4.7).

Luând în considerare rotația brută determinată anterior, o aliniere fină a norului de puncte  $\mathcal{P}^{\mathcal{M}}$ , relativ la norul de puncte  $\mathcal{O}_k$ , se poate obține eficient utilizându-se metoda ICP (Iterative

*Closest Point*) [12]. O descriere detaliată a acestei metode a fost prezentată în secțiunea 2.3.2. Asemănător metodei de estimare a rotației brute, aceasta are la bază un proces de minimizare a distanței Euclidiene dintre punctele celor doi nori implicați. Diferență dintre cele două abordări, brută și fină (*ICP*), este aceea că metoda brută de estimare a rotației converge global către un minim al distanței Euclidiene, în timp ce, metoda *ICP* converge către un minim local. Astfel, distanța Euclidiană dintre punctele formelor poate fi considerată ca o măsură a convergenței procesului de transformare de similaritate. Așadar, procesul de aliniere brută este strict necesar pentru a se asigura o convergență locală corectă a metodei *ICP*.

Faptul că algoritmul *ICP* poate converge către minime locale multiple se datorează unui prag de validare a distanței dintre puncte. Astfel, dacă distanța (corespondența) dintre două puncte (unul din  $\mathcal{P}^M$  și celălalt din  $\mathcal{O}_k$ ) este mai mare decât un prag stabilit (de ex. 1 cm), aceasta nu va fi luată în considerare la calculul convergenței (distanței dintre punctele celor două forme) cumulate. Acest tip de comportament face ca, la aplicarea de mai multe ori, pe aceiași nori de puncte, a metodei *ICP*, să se obțină rate de convergență diferite (foarte probabil din ce în ce mai mici), după cum se poate observa și în figura 4.8 unde, sub fiecare aplicare succesivă a metodei *ICP* sunt afișate valorile de convergență. Așadar, matricea de rotație fină  $\mathbf{R}_{fin}$  este obținută după aplicarea de mai multe ori a algoritmului *ICP*. În urma aplicărilor succesive, valoarea convergenței are tendința de a se stabili datorită obținerii deja a minimului local al rotației. De remarcat este faptul că tendința convergenței este de scădere. Totuși, numărul de aplicări nu poate să fie fix din moment ce convergența poate fi, pentru un caz, constantă după 3 aplicări, iar pentru al caz, constantă după 13 aplicări. Rezolvarea acestei probleme poate fi obținută prin monitorizarea ratei de scădere a convergenței. Astfel, dacă rata de scădere a convergenței este mai mică decât un prag stabilit (de ex. 5%, relativ la convergența anterioară), procesul repetitiv de aliniere prin metoda *ICP* poate lua sfârșit deoarece convergența tinde să fie constantă. În acest fel, se poate aplica un număr optim de alinieri fine utilizând metoda *ICP*. Cumulat, rotația  $\mathbf{R}$  finală, ce are rolul de a alinia norul de puncte  $\mathcal{P}^M$  relativ la norul  $\mathcal{O}_k$ , poate fi descrisă astfel:

$$\mathbf{R} = \mathbf{R}_{brut} \cdot \mathbf{R}_{fin}. \quad (4.10)$$

În forma în care este prezentată în figura 4.8(d), secvența numărul 4, primitiva generică  $\mathcal{S}_l$  este scalată, translatată și rotită optim astfel încât să existe o aliniere ideală între aceasta și norul de puncte obiect  $\mathcal{O}_k$ . Deși, odată cu finalizarea transformării de similaritate se obține o formă volumetrică estimată a obiectului de interes, primitiva generică aliniată este caracterizată totuși de un anumit grad de generalitate. În continuare, pentru a particulariza volumul astfel estimat, se poate recurge la un proces complex de modelare ce are ca principal scop deformarea primitivei generice pentru a crește și mai mult similaritatea dintre forma acesteia și forma obiectului perceput.

## 4.5 Concluzii

În cadrul acestui capitol a fost tratată reconstrucția volumetrică 3D utilizând un nou tip de forme 3D, acela al primitivelor generice. Acestea sunt un concept nou de forme rigide a căror înfățișare este menită să aproximeze o gamă largă de obiecte. Prin utilizarea unei singure primitive

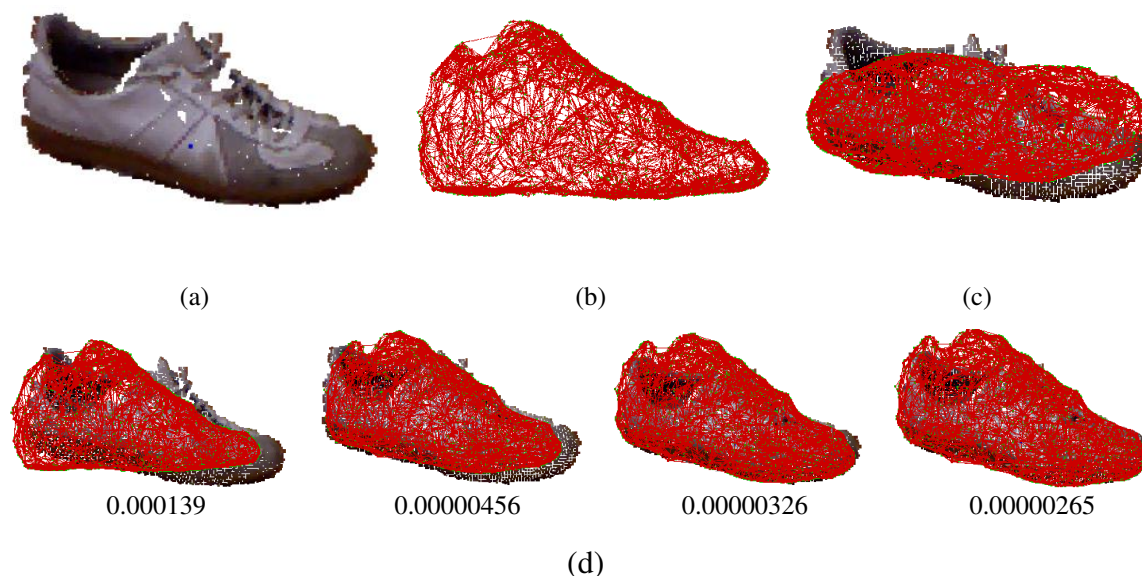


Fig. 4.8 Determinarea rotației optime dintre un nor de puncte obiect și primitivă generică aferentă: (a) nor de puncte obiect ( $\mathcal{O}_k$ ) înfățișând un pantof; (b) primitivă generică aferentă obiectului pantof (reprezentată printr-o rețea triunghiulară de puncte de culoare roșie); (c) rotația brută obținută între cele două forme; (d) rotația fină obținută între cele două forme împreună cu diferite valori ale convergenței (afișate sub fiecare imagine) pentru una, două, trei, respectiv patru aplicări succesive ale algoritmului ICP.

generice, ca reprezentant pentru fiecare clasă de obiecte în parte, se obține o bază de date cu forme redusă ca și dimensiune. Conceptul, a cărui structură a fost prezentată în secțiunea 4.2, este similar cu cel al metodei de reconstrucție volumetrică bazată pe forme predefinite, prezentat în cadrul capitolului anterior, cu mențiunea că, procesul laborios de identificare al similarităților dintre obiectul vizualizat și formele din baza de date cu forme este înlocuit cu unul de recunoaștere al obiectelor din scenă care, direcționează procesul de selecție a formei din baza de date direct către forma optimă. Pentru recunoașterea obiectelor din scenă a fost folosit un clasificator care a fost "antrenat" să caute, în scenă, obiectul ce satisface sarcina impusă. Rezultatul unei astfel de clasificări, este atât o regiune de interes, care localizează în scenă obiectul căutat, cât și o etichetă care poate fi utilizată mai departe pentru a extrage primitivă generică aferentă din baza de date cu primitive. O descriere detaliată a procesului de recunoaștere a obiectelor a fost oferită în cadrul secțiunii 4.3.

Secțiunea 4.1 a fost rezervată oferirii unor detalii amănunțite referitoare la structura de date, proprietățile sau modul de obținere al unei primitive generice. În fapt, o primitivă generică este un model distribuit de puncte 3D ce înfățișează forma comună a obiectelor dintr-o anumită clasă de obiecte. Nu toate punctele 3D ale unei primitive generice caracterizează efectiv forma acesteia. O mare parte dintre punctele 3D sunt utilizate doar pentru a crea o suprafață netedă. Utilizând un concept intens utilizat în modelarea tridimensională a formelor organelor umane, în secțiunea 4.1.2, punctele primitivei au fost împărțite în două categorii: de control (ce descriu și controlează efectiv forma primitivei generice), respectiv ordinare/normale (ce interpolatează punctele de control pentru a crea o suprafață lină). Totodată acest concept permite accelerarea tuturor operațiilor (manipulare, modelare, analiză) aplicate asupra primitivei generice.

Tot în cadrul acestei secțiuni a fost descris și modul prin care este generată forma medie a

primitivei generice. La baza acestui proces se află analiza Procrustes generalizată. O descriere detaliată a acestui tip de analiză fost realizată în cadrul capitolului cu numărul doi. În medie, pentru generarea unei primitive generice, au fost folosite un număr de șapte forme predefinite de obiecte, fiecare făcând parte din aceeași clasă de obiecte. Folosirea unui număr redus de forme predefinite poate conduce la obținerea unei primitive generice mult prea particulare, motiv pentru care, pentru generarea fiecărei primitive generice se recomandă utilizarea unui număr de minim șase forme predefinite de obiecte (ce aparțin aceleiași clase). Numărul maxim de forme predefinite poate fi însă nelimitat. Fiind un proces independent de acela al reconstrucției propriu-zise, generarea primitivei generice se poate realiza exclusiv "offline".

Ultima secțiune a capitolului curent prezintă procedeele prin care primitiva generică este scalată, translatată și rotită pentru a fi aliniată cu obiectul vizualizat din scenă. Dacă primele două operații pot fi realizate relativ cu ușurință, rotirea primitivei generice este o operație mult mai dificilă. Problema a fost abordată prin obținerea, în primă fază, a unei rotații brute iar, în cea de a doua fază, a unei rotații fine. După cum a fost descris în cadrul ultimei secțiuni, rotația brută a fost obținută prin rotirea succesivă a primitivei generice, cu un pas de  $10^\circ$ , în jurul celor trei axe carteziane, și măsurarea distanței Euclidiene dintre punctele acesteia și cele ale obiectului vizualizat. Rotația care produce cea mai mică distanță Euclidiană însumată a tuturor punctelor formelor este considerată rotația brută optimă. În continuare, rotația fină s-a obținut prin simpla aplicare a algoritmului iterativ de identificare a celui mai apropiat punct vecin (*ICP*). Odată aplicată asupra sa operația de transformare de similaritate, primitiva generică poate fi folosită de către sistem pentru a reprezenta volumetric obiectul de interes.

## 5. Modelarea activă a primitivelor generice

---

Contururi active  
Contururi active 2D  
Forme active 3D  
Concluzii

---

Acuratețea volumului estimat a obiectului de interes depinde, în mare măsură, de forma rigidă utilizată. După cum a fost prezentat în capitolele anterioare, pentru a îndeplini această sarcină pot fi utilizate, în principal, forme predefinite [102], forme implicite [30] sau primitive generice [31]. Aliniată peste datele 3D percepute, acestea au rolul de a reprezenta volumetric obiectul considerat. În funcție de forma utilizată, reprezentarea volumetrică finală poate fi, sau nu, similară cu obiectul real. De exemplu, utilizarea modelelor supercuadrice pentru aproximarea volumelor cu suprafață complexă nu este indicată dat fiind faptul că această abordare folosește doar forme geometrice regulate pentru a reconstrui volumul obiectului considerat. Pe de altă parte, utilizarea formelor predefinite (a unei replici ideale a obiectului vizualizat) oferă un estimat net mai precis dar la un cost de timp foarte ridicat. Deși se folosește o replică rigidă a obiectului perceput, rezultatul nu este în totdeauna cel scontat datorită posibilelor erori de achiziție, de aliniere și, de ce nu, a micilor neconcordanțe dintre regiunile formei predefinite și regiunile reale percepute. Chiar și primitivele generice introduse în prezenta lucrare oferă doar un aproximat al volumului considerat. O soluție la problema acurateței poate fi oferită de un proces de modelare al volumului estimat. Acesta are rolul de a "împinge" suprafața volumului reconstruit, către aceea a obiectului de interes din scenă.

În literatura de specialitate există o serie de abordări ce tratează cu succes acest tip de problemă. În funcție de domeniul căruia i se adresează pot exista algoritmi de modelare 2D (pentru domeniul imaginii), respectiv 3D (pentru spațiul Euclidian). Luând în considerare primul caz,

se pot menționa trei metode de modelare a formei unui anumit obiect, și anume metoda *contururilor active* [34, 177], algoritmul *Active Shape Model (ASM)* [36, 37], respectiv algoritmul *Active Appearance Model (ASM)* [121]. Plasat în domeniul unei imagini, conturul activ 2D se va deplasa, constrâns fiind de o serie de forțe, către caracteristicile proeminente din imagine (colțuri, muchi, etc.). Această abordare beneficiază de o acuratețe ridicată, modelul final al conturului activ urmărind îndeaproape forma obiectului din imagine. Utilizând un proces similar de modelare, formele active (*ASM*) prezentate în [37] sunt deformate într-o serie de moduri definite de un set de forme de antrenare. Acest set definește variabilitatea modelului formei active. Dezavantajul acestei abordări este acela că modelul poate fi deformat doar în maniera descrisă de setul de antrenare, fapt ce limitează capacitatea acestuia de a se mula pe orice tip de suprafață. Cea de-a treia abordare folosește o tratare similară cu *ASM*, noutatea fiind aceea că de această dată, pe lângă modelul de variație al formei considerate, generat folosind setul de forme de antrenare, este utilizat și un model al variației aspectului acestei forme.

Considerând spațiul Euclidian, principal pot fi modelate forme reprezentate prin puncte 3D conectate între ele printr-o serie de segmente ce formează o rețea poligonală (modelare poligonală<sup>1</sup>), forme reprezentate prin curbe<sup>2</sup> și forme reprezentate mixt (puncte 3D, voxelii și rețea poligonală). Având în vedere acest mod de reprezentare, principalele metode de modelare prezente în literatura de specialitate sunt: modelarea utilizând multiple forme geometrice solide care prin concatenare pot lua diferite forme [7], modelarea formelor utilizând suprafețe implicite (v. sec. 3.1) [30], modelarea formelor utilizând principiul contururilor active și, nu în ultimul rând, modelarea prin partiționare a suprafețelor ce are rolul de a nivela ("sculpta") o suprafață inițială pentru a o face similară cu obiectul de interes [145].

Luând în considerare avantajele și dezavantajele algoritmilor de modelare enumerați anterior, atât cei utilizați în domeniul 2D, cât și cei utilizați în spațiul Euclidian, precum și modelul dat al primitivei generice (drept volum estimat inițial), ce face subiectul prezentei lucrări, cea mai oportună abordare pentru modelarea primitivei generice este aceea a *contururilor active*. Principalele atuuri sunt flexibilitatea sporită oferită suprafeței primitivei generice precum și acuratețea ridicată a formei finale modelate a acesteia.

## 5.1 Contururi active

Un contur activ, denumit în literatura de specialitate și *șarpe*<sup>3</sup>, reprezintă un model de contur deformabil [99]. Acesta poate avea o reprezentare liniară (curbă deschisă) sau, poate fi descris prin intermediul unei suprafețe (curbă închisă). Pentru cel de al doilea caz, trebuie avut în vedere faptul că punctele ce compun sfârșitul curbei nu trebuie să se deplaseze în aceeași direcție cu celelalte puncte. În acest fel se va evita aglomerația punctelor și implicit dispariția conturului. Deplasarea acestei curbe în spațiu se realizează ordonat, constrânsă fiind de o serie de forțe. Capacitatea conturului activ de a efectua lucru mecanic când trece dintr-o stare în alta (când își modifică forma) este măsurată prin intermediul așa numitei energii. Schimbarea poziției punctelor conturului activ

<sup>1</sup>Eng. *Polygonal Modelling*

<sup>2</sup>Eng. *Curve Modelling*

<sup>3</sup>Eng. *Snake*

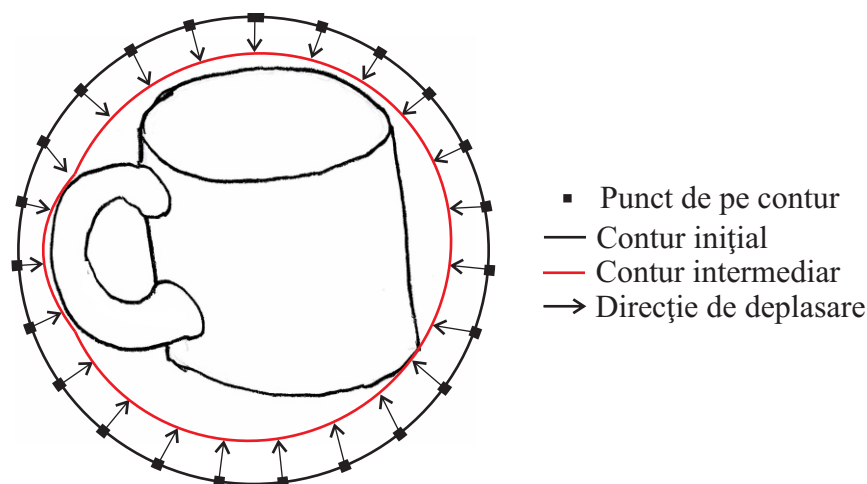


Fig. 5.1 Principiul de deformare al unui contur activ bidimensional.

determină câștigarea, respectiv pierderea unui quantum de energie. Să presupunem că un punct al conturului activ este împins, de către o forță, din poziția *A* în poziția *B*. În această situație, punctul aflat poziția *B* va câștiga o energie egală cu forță ce a produs deplasarea, multiplicată cu distanța pe care s-a realizat deplasarea. Modul explicit de calcul al acestor energii va fi prezentat în secțiunile următoare. Poziția corectă a fiecărui punct al conturului activ este identificată prin minimizarea funcționalei de energie. Inițial, forma conturului activ descrie o siluetă regulată (cerc pentru spațiul imaginii, respectiv o sferă sau o formă particulară pentru spațiul Euclidian). Plecând de la această formă incipientă, plasată în interiorul obiectului a cărui contur se dorește a fi identificat, conturul activ este "împins" către granițele obiectului considerat consumând două tipuri de energii: *interne*, respectiv *externe*. Primul tip de energie se numește *internă* deoarece determinarea cuantumului ei se realizează utilizând informații legate de conturul în sine (lungime, elasticitate, continuitate etc.). Uzual, energia internă este susceptibilă de variație continuă. Cel de-al doilea tip de energie se numește *externă* deoarece aceasta este strict determinată utilizându-se informația din jurul conturului (informația din imagine, respectiv din spațiul Euclidian). Împreună, cele două energii conduc conturul către zonele proeminente ce definesc silueta obiectului considerat. Un minim al acestor energii denotă atingerea granițelor obiectului.

Intuitiv, modul de comportare al conturilor active poate fi asemănat cu cel al unui balon (v. fig. 5.1). Prin umflare, un balon crește liber în dimensiune. Dacă în decursul acestui proces, porțiuni din balon intră în contact cu diferite părți externe acestuia, granițele respectivelor porțiuni din balon vor lua forma entității exterioare colizionate. Astfel, dacă se introduce un balon în interiorul unui obiect, prin umflare, membrana balonului se va apropia din ce în ce mai mult de granițele obiectului considerat fără însă a le depăși. În final, forma balonului va fi identică cu aceea a obiectului în cauză. Procesul poate fi aplicat și viceversa, respectiv se poate plasa un obiect în interiorul unui balon umflat, urmând ca, prin dezumflare, acesta să capete forma obiectului introdus [34]. Modul de comportare al balonului are și o serie de neajunsuri. Unul dintre acestea îl reprezintă semnul forței care asigură proprietatea de umflare, respectiv de dezumflare a balonului. O rezolvare a acestei probleme este prezentată în [98].

Poziționarea corectă a conturului pe granițele obiectului este asigurată în întregime de forțe

interne și externe, forțe ce consumă în acest sens diferite energii. Poziția finală a punctelor conturului activ se obține în momentul în care se atinge un echilibru între energiile interne și cele externe [179], echilibru ce asigură o energie minimă consumată de contur pentru a atinge o formă finală corectă. Fiecare dintre aceste forțe identifică o proprietate importantă a membranei balonului. Forțele interne, intuitiv asigură proprietățile de elasticitate și continuitate ale conturului activ. Proprietatea de continuitate caracterizează capacitatea conturului de a păstra o distanță egală între puncte. Această proprietate asigură totodată și coerența marginilor conturului. Pe de altă parte, proprietatea de elasticitate asigură maleabilitatea conturului activ. Astfel, un contur activ maleabil se poate undula cu ușurință peste granițele obiectului considerat, în timp ce, un contur activ mai rigid va conduce la obținerea unei forme brute a aceluiași obiect considerat.

Ghidarea corectă a conturului către granițele obiectului considerat este controlată de forțele externe. Fie că vorbim de spațiul 2D (pixeli din imagini) sau de spațiul 3D (densități de puncte 3D), trebuie să evidențiem granițele obiectului a cărui formă se dorește a fi extrasă. Acest proces se realizează prin definirea unor funcții de potențial ce prezintă minime locale în dreptul granițelor obiectului considerat. Spre exemplificare, în cazul 2D, aceste minime locale sunt identificate ca fiind muchii de intensitate din cadrul imaginii [177]. Similar, în cazul 3D, aceste minime sunt asigurare de variații puternice ale densităților de puncte din spațiu. Nu în ultimul rând, în cadrul forțelor externe se pot introduce o serie de constrângeri definite de utilizator.

În funcție de tipul spațiului în care se deplasează, contururile active se împart în:

- bidimensionale (ce se deplasează liber în domeniul imaginii) [34];
- tridimensionale (ce se deplasează liber în spațiul Euclidian) [1].

## 5.2 Contururi active 2D

Un contur activ 2D este definit inițial prin intermediul unui model parametric generic (un cerc). Relativ la acest tip de modele parametrice, procesul de deformare poate decurge în două sensuri [146, 177], după cum urmează:

- prin minimizarea energiei;
- prin forțe.

Primul tip de formulare prezintă avantajul satisfacerii principiului variațional, în timp ce, formularea dinamică prin forțe oferă o flexibilitate importantă ce permite utilizarea unor forțe generice.

### 5.2.1 Deformarea contururilor active prin minimizarea funcționalei de energie

Un contur activ bidimensional poate fi descris ca o funcție  $v: s \rightarrow \mathbb{R}^2$ , unde  $s \in [0, 1]$ , iar  $v(s) = [x(s), y(s)]$ ,  $x(s)$ , respectiv  $y(s)$  fiind coordonatele punctelor de-a lungul conturului. În domeniul discret,  $v(s)$  este o înșiruire de puncte bidimensionale  $\{\mathbf{p}_1, \dots, \mathbf{p}_{N_v}\}$  unde,  $\mathbf{p}_1 = (x_1(s), y_1(s))$ , iar  $N_v$  este numărul total de puncte ce descrie conturul. Lăsat liber în domeniul unei imagini  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , acesta se va deplasa spre o poziție optimă ghidat fiind de minimizarea unei funcționale de energie [25]. Figura 5.2 ilustrează un astfel de contur bidimensional. Procesul este unul iterativ, conturul înaintând în domeniul imaginii în trepte. Pentru fiecare remodelare, determinată de o serie de forțe, conturul consumă o anumită cantitate de energie care, în esență, este un scalar, o funcție a

formeii curente. Matematic, modelarea conturului implică minimizarea unei funcționale de energie [146]. Aceasta, în fapt, este o sumă a energiilor consumate de către procesul de modelare al conturului. Funcționala de energie poate fi descrisă astfel:

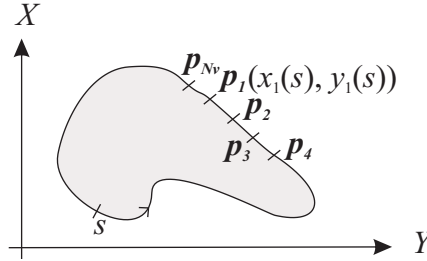


Fig. 5.2 Contur activ (bidimensional).

$$\begin{aligned}
 \varepsilon(v, f) &= \varepsilon_{int}(v) + \varepsilon_{ext}(v, f) = \\
 &= \int_0^1 [E_{contur}(v(s), f)] ds = \\
 &= \int_0^1 [E_{int}(v(s)) + E_{ext}(v(s), f)] ds = \\
 &= \int_0^1 [\alpha(s)E_{cont}(v(s)) + \beta(s)E_{curb}(v(s)) + \gamma(s)E_{imagine}(v(s), f)] ds,
 \end{aligned} \tag{5.1}$$

unde,  $E_{contur}$  este energia totală a conturului,  $E_{int}$ , respectiv  $E_{ext}$  reprezintă energia internă, respectiv externă a conturului,  $\alpha$ ,  $\beta$  și  $\gamma$  sunt factorii de pondere ce controlează influența energiilor de continuitate ( $E_{cont}$ ), de curbura ( $E_{curb}$ ), respectiv externă ( $E_{imagine}$ ) asupra conturului,  $f$  este imaginea scenei considerate, iar  $v(s)$  este curba ce reprezintă conturul activ.

În cazul discret, funcționala de energie poate fi descrisă după cum urmează:

$$\varepsilon(v, f) = \sum_{i=1}^{N_v} (\alpha(i)E_{cont}(i) + \beta(i)E_{curb}(i) + \gamma(i)E_{imagine}(i)), \tag{5.2}$$

unde,  $N_v$  este numărul total de puncte (noduri) ce compun conturul în cauză. În continuare, se va face o descriere a energiilor ce realizează modelarea conturului activ.

### 5.2.1.1 Energia internă

Rolul energiei interne este acela de a asigura continuitatea și elasticitatea conturului activ. Energia internă poate fi scrisă matematic sub următoarea formă:

$$\begin{aligned}
 \varepsilon_{int}(v) &= \int_0^1 [E_{int}(v(s))] ds = \\
 &= \int_0^1 [\alpha(s)E_{cont}(v(s)) + \beta(s)E_{curv}(v(s))] ds = \\
 &= \int_0^1 \frac{1}{2} [\alpha(s)|v'(s)|^2 + \beta(s)|v''(s)|^2] ds,
 \end{aligned} \tag{5.3}$$

unde,  $\alpha(s)$  și  $\beta(s)$  sunt factorii ce ponderează influența derivatelor de ordinul întâi ( $v'(s)$ ), respectiv doi ( $v''(s)$ ), în raport cu  $s$ . Se poate observa că factorii de ponderare se aplică independent pentru

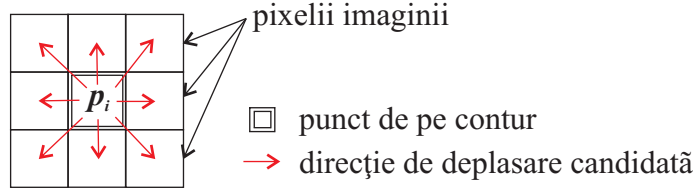


Fig. 5.3 Direcții de deplasare candidate ale unui punct 2D al conturul activ (considerându-se ca spațiu de deplasare domeniul imaginii).

fiecare punct. În cazul discret, energia internă pentru un anumit punct al conturului poate fi scrisă astfel:

$$\begin{aligned} E_{int} &= \frac{1}{2} \left( \alpha(s) |v'(s)|^2 + \beta(s) |v''(s)|^2 \right) = \\ &= \frac{1}{2} \left( \alpha(s) \left| \frac{dv}{ds} \right|^2 + \beta(s) \left| \frac{d^2v}{ds^2} \right|^2 \right). \end{aligned} \quad (5.4)$$

Această ecuație este cunoscută în literatura de specialitate sub numele de *ecuația membranei* și este utilizată cu precădere în domeniul mecanicii [135]. După cum se poate observa din ecuația 5.4, energia internă este alcătuită din două componente. Acestea sunt numite *energie de continuitate* respectiv *energie de curbura*. Matematic, energiile sunt asociate cu derivatele de ordinul unu, respectiv doi ale conturului  $v$ . Rolul fiecăreia precum și efectul lor asupra contului activ va fi detaliat în continuare.

Fie un contur  $v(s)$  format din  $N_v$  puncte  $\mathbf{p}_1, \dots, \mathbf{p}_{N_v}$ . În cazul discret, derivata de ordinul unu a conturului activ, asociată energiei  $E_{cont}$ , este descrisă de următoarea ecuație cu diferențe finite:

$$\begin{aligned} E_{cont}(i) &= \left| \frac{dv}{ds} \right|^2 = \\ &= \|\mathbf{p}_i - \mathbf{p}_{i-1}\|^2 = \\ &= (x_i - x_{i-1})^2 + (y_i - y_{i-1})^2, \end{aligned} \quad (5.5)$$

unde,  $\mathbf{p}_i$  este un punctul curent ce se dorește a fi deplasat.

În decursul procesului de deformare, punctele conturului activ sunt libere să se deplaseze în orice direcție, bineînțeles ghidate de diferitele forțe ce acționează asupra acestora. Considerând imaginea  $f$ , în care conturul se deplasează, un punct al conturului poate fi deplasat în oricare dintre cele 8 direcții prezentate în figura 5.3. Această dimensiune este de fapt rezoluția imaginii  $f$ . Având în vedere toate aceste direcții, energia  $E_{cont}$  are valoarea cea mai mică în direcția naturală (aceea de înaintare către granițele obiectului din imagine) ce poate fi observată și în figura 5.1. În timp, conturul are intenția de a se micșora datorită tendinței de minimizare a distanței dintre puncte, minimizare asociată cu dezumflarea balonului. Pentru a se evita acest aspect, se poate introduce un element nou ce conține distanța medie dintre puncte. Acest element va păstra o distanță constantă între punctele conturului. Astfel, ecuația 5.5 poate fi rescrisă astfel:

$$\begin{aligned} E_{cont}(i) &= (\bar{d} - \|\mathbf{p}_i - \mathbf{p}_{i-1}\|)^2 = \\ &= (\bar{d} - \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2})^2, \end{aligned} \quad (5.6)$$

unde,  $\bar{d}$  reprezintă distanța medie dintre toate puncte conturului [176]. O valoare mare a derivatei

de ordinul unu indică o distanță mare între poziția curentă și poziția candidată, poziție ce coincide cu o deformare puternică (neobișnuită) a conturului activ. Un minim al energiei  $E_{cont}$  va asigura o echidistanță între punctele conturului ceea ce va evita o aglomerare a acestora în anumite regiuni. Influența energiei asupra conturului poate fi ponderată prin intermediul factorului  $\alpha$ . Pentru valori mari ale energiei  $E_{cont}$ , factorul  $\alpha$  poate fi redus pentru a se evita efectul acesteia. Dacă  $\alpha$  ar fi setat la valoarea 0, șansa apariției unor goluri (porțiuni cu puncte relativ îndepărtate) în contur ar fi foarte mare.

Cel de al doilea tip de energie internă este energia de curbura, sau  $E_{curb}$ . Matematic, ea poate fi descrisă utilizându-se derivata de ordinul doi (v. ec. 5.4). Rolul ei este de a elimina oscilațiile conturului prin penalizarea curburilor pronunțate.  $E_{curb}$  definește, de fapt, elasticitatea conturului. Așa cum este prezentat în ecuația 5.7, aceasta este calculată utilizându-se poziția punctului anterior și următor relativ la un punct considerat. Cu cât acesta este îndepărtat mai mult de pe trasa ideală (definită de punctul anterior și cel următor), cu atât valoarea energiei de curbura crește. Similar, o valoare mică a derivatei de ordinul doi indică o curbura redusă a conturului și implicit un efort minim la înaintare. Curbura unui contur poate fi aproximată utilizând următoarea ecuație cu diferențe finite:

$$\begin{aligned} E_{curb}(i) &= \left| \frac{d^2 v}{ds^2} \right|^2 = \\ &= \| \mathbf{p}_{i-1} - 2\mathbf{p}_i + \mathbf{p}_{i+1} \|^2 = \\ &= (x_{i-1} - 2x_i + x_{i+1})^2 + (y_{i-1} - 2y_i + y_{i+1})^2. \end{aligned} \quad (5.7)$$

Influența acestei energii asupra conturului poate fi ponderată prin intermediul factorului  $\beta$ . Pentru valori mari ale energiei  $E_{curb}$ , factorul  $\beta$  poate fi redus pentru a se evita efectul acesteia. Dacă  $\beta$  ar fi setat la valoarea 0, conturul va fi caracterizat de o serie de curbură discontinue asociate colțurilor.

### 5.2.1.2 Energia externă

Dacă energiile definite anterior exprimau diferite stări interne ale conturului, energia externă nu folosește niciuna dintre proprietățile fizice ale acestuia. Rolul ei este acela de a "ghida" conturul către diferitele caracteristici întâlnite în imagine. Aceste caracteristici sunt, de fapt, frontierele obiectului a cărui formă trebuie să o ia conturul activ. Matematic, energia externă poate fi exprimată după cum urmează:

$$\varepsilon_{ext}(v(s), f) = \int_0^1 [\gamma(s) E_{ext}(v(s), f)] ds. \quad (5.8)$$

În cazul discret, ecuația 5.8 poate fi scrisă în felul următor:

$$E_{ext}(v(s), f) = E_{imagine}(v(s), f). \quad (5.9)$$

$E_{imagine}(v(s), f)$  se comportă ca o forță ce atrage conturul spre o formă particulară (caracteristică proeminentă din imagine). În forma în care aceasta va fi definită, energia  $E_{imagine}(v(s), f)$  va prezenta valori minime doar în dreptul caracteristicilor asociate conturilor din imagine. Datorită

fenomenului de atragere,  $E_{\text{image}}(v(s), f)$  poartă denumirea de *image atractor* [1]. Un exemplu de image atractor este *imagea gradient*. Aceasta descrie o hartă a tranzițiilor intensității pixelilor din image de-a lungul axelor  $x$ ,  $y$  și pe diagonală. Tranzițiile intensității pixelilor evidențiază, în fapt, granițe ale diferitelor entități din imagea  $f$ . Matematic, energia externă provenită din această image gradient este descrisă prin relația:

$$E_{\text{image}}(v, f) = -\gamma(s) \|\nabla I\|, \quad (5.10)$$

unde,  $\nabla$  este operatorul gradient,  $\gamma$  este factorul ce ponderează energia externă, iar  $I$  este imagea gri.  $E_{\text{image}}(v(s), f)$  devine foarte mică atunci când punctul conturului se apropie de o cantă, respectiv muchie. Un exemplu de image atractor, ce are la bază gradientul intensității unei imagini gri, poate fi observat în figura 5.4

Totuși, imagea gradient nu este singurul model de forță externă disponibilă. În literatura de specialitate se mai pot întâlni numeroase alte modele precum *GVF* (*Gradient Vector Flow*) [188], *GGVF* (*Generalized Gradient Vector Flow*) [189], *BVF* (*Boundary Vector Field*) [166], *FVF* (*Fluid Vector Flow*) [184], etc. Dintre acestea, unul dintre cele mai performante este modelul de forță externă bazat pe varianță [177]. După cum o subliniază și numele, caracteristicile importante din image (asociate granițelor obiectelor) sunt evidențiate utilizând varianța intensităților pixelilor din image. Considerând o image de intrare gri  $I(x, y)$ , varianța se calculează local pentru fiecare pixel utilizându-se o fereastră glisantă pătratică cu dimensiune impară (de ex.  $75 \times 75$  sau  $3 \times 3$ , etc.). O valoare mare a acestei ferestre va produce o serie de granițe (asociate variației intensității pixelilor) cu difuzie mare, în timp ce, o valoare mică a acesteia va produce o serie de granițe subțiri. Informația astfel calculată poate fi folosită ulterior ca energie externă pentru a ghida conturul către muchiile proeminente din imagea  $I(x, y)$ . Spre exemplificare, în figura 5.5(c) se poate observa imagea atractor bazată pe varianță (calculată utilizându-se o fereastră cu dimensiunea  $3 \times 3$ ). Formula de calcul a varianței prezintă următoarea formă:

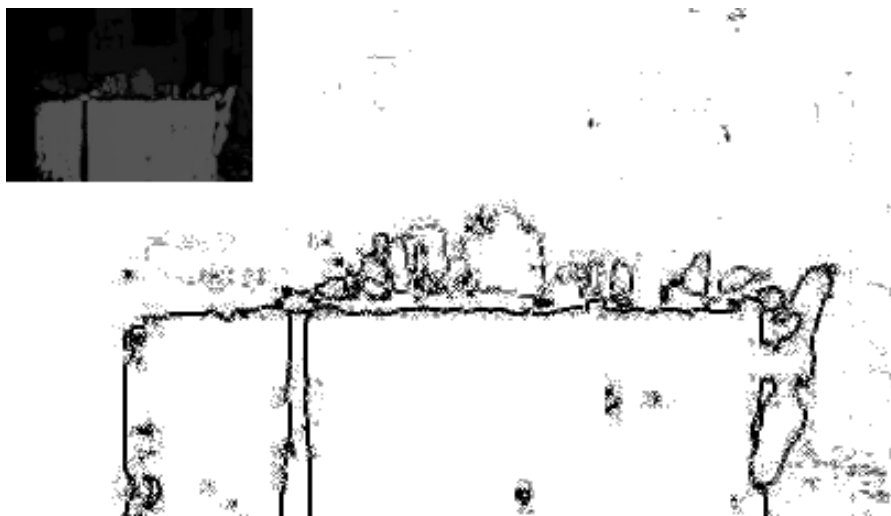


Fig. 5.4 Image atractor definită utilizând gradientul unei imagini. În colțul din stânga sus este ilustrată imagea de adâncime ce a condus la obținerea imaginii atractor.

$$\sigma^2 = \frac{1}{N_v} \sum_{i=1}^{N_v} (X_i - \bar{X})^2, \quad (5.11)$$

unde,  $\bar{X}$  reprezintă valoarea medie a elementelor din fereastra considerată,  $N_v$  este numărul total de puncte 2D ce compun conturul activ, iar  $\sigma$  este varianța. O dimensiune mare a ferestrei pe baza căreia se calculează varianta va atrage conturul de la o distanță mare, în timp ce, o dimensiune mică a acestei ferestre (de ex.  $3 \times 3$ ) va permite atragerea conturului doar de la o distanță foarte mică. Principiul de modelare a conturului activ utilizând energia externă bazată pe varianță este următorul (v. fig. 5.5). Se deformează conturul utilizând energia externă calculată cu fereastra cea mai mare ( $75 \times 75$ ). Se utilizează această energie pentru fiecare punct al conturului atâta timp cât diferența dintre energiile din forma conturului curent și următor (calculată cu fereastra  $65 \times 65$ ), este monoton crescătoare. Acest aspect este datorat tendinței de umflare a conturului activ inițial, umflare ce permite adăugarea de noi puncte și deci o creștere a valorii funcționalei de energie. Punctul de inflexiune în care cele două funcționale de energie încep să descrească indică faptul că este suficientă energie în pasul următor (calcularea energiei externe cu o fereastră cu dimensiunea  $55 \times 55$ ) pentru a atrage conturul. Procesul este iterativ și se realizează până în momentul când se atinge dimensiunea minimă a ferestrei ( $3 \times 3$ ). Totodată acest moment indică un contur subțire și precis [177]. În final, se obține un contur ce urmărește perfect silueta obiectului considerat.

### 5.2.1.3 Minimizarea funcționalei de energie

După cum a fost descris anterior, forma optimă a conturului activ este atinsă atunci când se constată un minim al funcționalei de energie. Minimizarea acesteia este echivalentă cu rezolvarea următoarei ecuații Euler-Lagrange:

$$\frac{1}{2} \left( \alpha(s) \left| \frac{dv}{ds} \right|^2 + \beta(s) \left| \frac{d^2v}{ds^2} \right|^2 \right) = -\gamma(s) \|\nabla I\|, \quad (5.12)$$

care implică, ca energiile interne ale conturului (prezentate în partea stângă a egalității) să echilibreze energiile externe (aflate în dreapta egalității). Fie un contur activ  $v(s)$  descris prin intermediul a  $N_v$  puncte bidimensionale, atunci, ecuația 5.12 poate fi rescrisă după cum urmează:

$$A v(s) + \nabla I = 0, \quad (5.13)$$

unde,  $A$  este matricea derivatelor de ordinele unu și doi prezentate în relația 5.12 și completate de factorii de pondere  $\alpha$  și  $\beta$ . Conform metodei lui Euler descrisă în [176], un contur poate fi calculat iterativ în felul următor:

$$(I + A) v_{k+1}(s) = v_k(s) + \nabla I, \quad (5.14)$$

unde,  $v_0(s)$  este vectorul inițial (nedeformat) descris prin intermediul a  $N_{v_0}$  puncte, iar  $I$  este matricea unitate. Dimensiunile conturului (numărul de puncte) la iterația  $v_{k+1}(s)$  și  $v_k(s)$  nu sunt necesare a fi egale. Pe măsură ce conturul crește/scade, dimensiunea acestuia va varia în timp.

Conform ecuației 5.14, punctele conturului  $v(s)$  vor fi deplasate în direcția indicată de *imaginea atractor*, direcție care balansează cuantumul energiilor interne. Acest proces este iterativ și are loc pentru fiecare punct în parte, până în momentul când se atinge un minim al funcționalei de energie.

### 5.2.2 Modelarea dinamică a conturului utilizând forțe multiple

Acest tip de modelare este o alternativă la modelarea clasică prin minimizarea funcționalei de

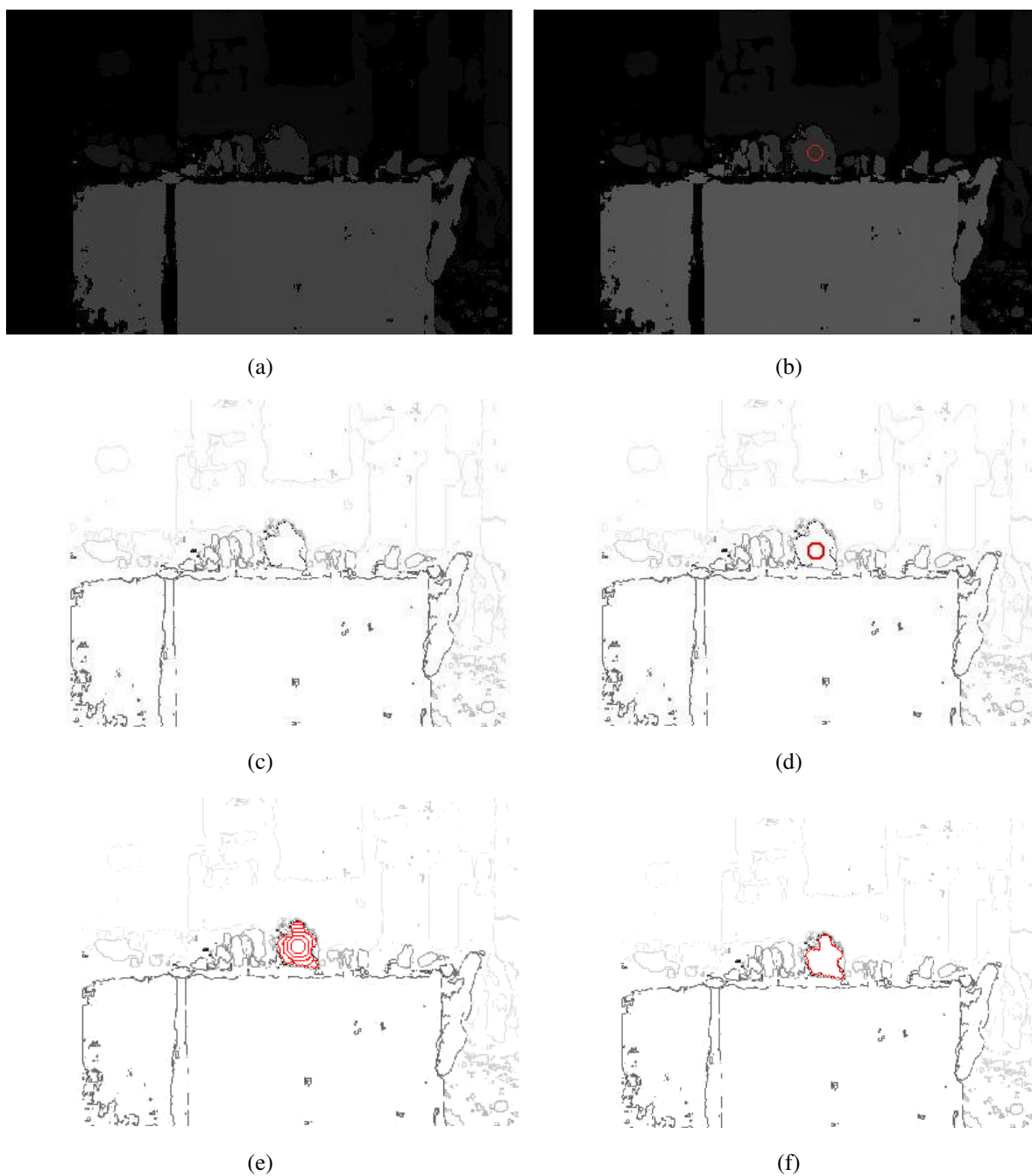


Fig. 5.5 Exemplu de segmentare, prin metoda conturilor active, a unui obiect din scena vizualizată: (a) imaginea gri a scenei considerate; (b) inițializarea unui contur activ în domeniul imaginii (considerându-se un obiect de interes dat); (c) imaginea atractor; (d) inițializarea conturului activ în interiorul imaginii atractor (considerându-se un obiect de interes dat); (e) forma conturului activ pentru diferite iterații; (f) forma finală a conturului activ.

energie prezentată anterior. Spre deosebire de aceasta, utilizarea a mai multor forțe externe pentru a controla forma conturului poate fi, în unele situații, mai avantajoasă [146]. Practic, numărul mare de forțe crește acuratețea procesului de atracție a conturului. Cel mai ușor și convenabil se pot defini forțe noi în rândul celor externe care controlează, de fapt, forma conturului.

În afară de forța externă bazată pe gradientul imaginii gri, se mai pot defini noi tipuri de forțe externe mai generale. Acestea pot fi exprimate ca superpoziția mai multor tipuri de forțe, după cum urmează:

$$F_{ext}(v(s)) = F_1(v(s)) + F_2(v(s)) + \dots + F_n(v(s)), \quad (5.15)$$

unde,  $n$  este numărul total de forțe externe distincte ce acționează asupra conturului.

În literatura de specialitate sunt prezentate numeroase tipuri (modele) de forțe externe dintre care, în ordine cronologică, se menționează următoarele:

- modelul curgerii vectorilor gradient *GVF* (*Gradient Vector Flow*) [188]. Avantajele acestui model de forță externă sunt posibilitatea atragerii conturului de la distanțe mari precum și calitatea de a pătrunde în concavitățile proeminente din imagine;
- modelul generalizat al curgerii vectorilor gradient *GGVF* (*Generalized Gradient Vector Flow*) [189];
- modelul câmpului electric virtual *VEF* (*Virtual Electric Field*) [138]. Relativ la modelul *GVF*, *VEF* este considerabil mai ușor și mai rapid de calculat;
- modelul convoluțional al câmpului electric virtual *CONVEF* (*CONvolutional Virtual Electric Field*) [185]. Spre deosebire de *VEF*, acesta prezintă o rază de atracție mai mare în contrast cu un timp de procesare mult mai mic. Conform autorilor, acest model poate fi calculat în timp real;
- modelul curgerii vectorilor de curbura *CVF* (*Curvature Vector Flow*) [63]. Evaluând geometria conturilor proeminente din imagine, modelul *CVF* este capabil să ghideze conturul către orice caracteristică din imagine, indiferent de geometria acesteia;
- modelul câmpului vectorilor limită *BVF* (*Boundary Vector Field*) [166]. Acest model reduce considerabil puterea de calcul necesară și totodată îmbunătățește raza de atracție a conturului;
- modelul curgerii vectorilor fluizi *FVF* (*Fluid Vector Flow*) [184]. Datorită ariei mari de atracție precum și a capacității de a pătrunde în concavități proeminente, acest model este superior modelelor *GVF* și *BVF*;
- modelul varianței [177]. În comparație cu modelele de energie *GVF*, *GGVF*, modelul bazat pe varianță este mult mai simplu și necesită mai puține iterații pentru difuzie.

### 5.2.3 Potrivirea conturului pe granițele obiectelor din imagine

Potrivirea conturului  $v(s)$  pe granițele obiectului de interes din imagine presupune parcurgerea a trei etape. Spre exemplificare, în figura 5.5 sunt prezentați grafic pașii necesari modelării conturului activ. Se consideră scena ilustrată în figura 5.5(a) mediul în interiorul căruia se va deplasa conturul activ.

Prima etapă a procesului de potrivire presupune poziționarea conturului inițial în interiorul obiectului de interes (a cărei formă se dorește a fi determinată). Din această poziție, conturul activ va fi atras, mai departe, de către diferitele caracteristici din imagine luând, în final, o formă particulară foarte apropiată de aceea a obiectului de interes. Pentru exemplul analizat, forma inițială a conturul activ înfățișează un cerc. Poziția acestuia în interiorul imaginii poate fi stabilită fie *manual* de către un operator uman, fie *automat* utilizându-se un proces de segmentare [142]. Figurele 5.5(b) și 5.5(d) înfățișează forma inițială a conturului activ poziționat în interiorul obiectului de interes din imaginea de intrare inițială (imagine gri) respectiv, din imaginea atractor utilizând metoda *manuală*.

Cea de a doua etapă a procesului de potrivire presupune determinarea *imaginii atractor*. Pentru exemplul utilizat, imaginea atractor a fost calculată utilizându-se metoda descrisă în [177]. Figura 5.5(c) prezintă imaginea atractor astfel calculată. Pentru aceasta, varianța a fost calculată utilizându-se o fereastră pătratică de dimensiune  $3 \times 3$ . Acesta este și motivul pentru care granițele obiectelor din imagine sunt subțiri. Pe măsură ce se mărește dimensiunea ferestrei pătratice, se mărește și difuzia acestor granițe.

Cea de a treia și ultima etapă este modelarea continuă a conturului. Printr-un proces iterativ, ecuația 5.14 împinge conturul activ către granițele obiectului din imaginea atractor. Procesul se încheie în momentul în care conturul devine stabil (respectiv, există un echilibru între energiile conturului). Figura 5.5(e) descrie diferitele forme luate de către contur de-a lungul procesului iterativ. Silueta finală a conturului activ este obținută după doar 38 de iterații (deplasări succesive ale punctelor conturului). Aceasta poate fi observată în figura 5.5(f).

#### 5.2.4 Reparametrizarea conturului activ

În cursa lor spre muchiile proeminente din imaginea atractor, punctele conturului activ se pot îndepărta sau apropia unele de altele. Spre exemplificare, în forma sa incipientă, conturul activ (descriș sub forma unui cerc) prezintă 48 puncte, în timp ce, în forma sa finală (ce urmărește perfect silueta obiectului considerat) acesta prezintă 256 puncte. Menținerea numărului inițial de puncte la 48 nu este recomandată deoarece, prin umflare/dezumflare distanța dintre acestea s-ar mări/micșora considerabil de mult și astfel muchii proeminente din contur ar putea fi pierdute. Pentru a se evita acest aspect, la fiecare iterație (umflare/dezumflare) conturul activ trebuie reparametrizat. Acest proces presupune adăugarea de noi puncte în contur pentru cazul procesului de umflare, respectiv eliminarea acestora pentru cazul procesului de dezumflare. Adăugarea punctelor în contur se realizează după o anumită regulă. Astfel, dacă distanța dintre punctul curent ( $\mathbf{p}_i$ ) și punctul următor ( $\mathbf{p}_{i+1}$ ) din contur este mai mare decât o distanță prag  $d_{max}$ , atunci se va recurge la adăugarea unui nou punct situat la jumătatea distanței dintre acestea. Similar, se va proceda și dacă distanța dintre aceste puncte este prea mică. În acest caz, dacă distanța dintre aceste două puncte este mai mică decât un anumit prag  $d_{min}$  atunci punctul  $\mathbf{p}_{i+1}$  va fi eliminat, punctul  $\mathbf{p}_{i+2}$  devenind astfel noul punct  $\mathbf{p}_{i+1}$ . Figura 5.6 prezintă intuitiv procesul de reparametrizare a conturului activ.

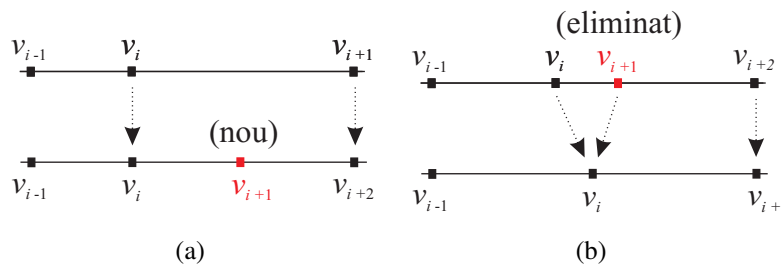


Fig. 5.6 Reparametrizarea conturului activ: (a) adăugarea unui nou punct la contur; (b) eliminarea unui punct din contur.

### 5.2.5 Considerații privind inițializarea conturului activ

Forma inițială a conturului activ este foarte importantă pentru acesta. O inițializare greșită poate conduce la un contur final parțial sau total eronat, în timp ce, o inițializare corespunzătoare a acestuia poate produce un contur precis și de încredere. O primă problemă, în acest sens, o reprezintă poziția de inițializare. Astfel, conturul inițial trebuie poziționat corespunzător în interiorul obiectului de interes. Identificarea acestei poziții se poate realiza fie automat, utilizând tehnici specifice de segmentare [67], recunoaștere [35] sau clasificare [125], fie manual de către un operator uman. Totuși, implicarea factorului uman presupune existența unei interfețe care să permită acest tip de interacțiune.

Un al doilea element relevant, considerând inițializarea conturului activ, este forma acestuia. Așa acum a fost descris în secțiunile anterioare, conturul activ inițial poate avea înfățișarea unei forme regulate precum un cerc, pătrat sau dreptunghi. Atâta timp cât granițele conturului inițial nu le depășesc pe cele ale muchiilor obiectului real, oricare dintre formele menționate anterior poate fi folosită în acest scop. Cu cât forma inițială aproximează mai fidel conturul obiectului real, cu atât numărul de iterații necesar pentru a se atinge forma finală este mai mic. Acest aspect a motivat, de fapt, și apariția unei metode de definire manuală a conturului inițial. Astfel, punctele conturului inițial pot fi definite manual de către un operator uman cât mai apropiate de muchiile reale ale obiectului. În acest fel, numărul de iterații se reduce drastic și astfel timpul necesar obținerii unui contur final este foarte mic.

Pentru a se eficientiza și mai mult timpul de calcul, dacă se dorește obținerea conturului într-o serie de imagini succesive, conturul activ final al imaginii anterioare poate fi considerat contur inițial în imaginea curentă. În acest fel, numărul iterațiilor necesar convergenței conturului se reduce simțitor.

## 5.3 Forme active 3D

Un contur tridimensional poate fi descris prin intermediul unei funcții  $v : s \times r \rightarrow \mathbb{R}^3$  unde,  $s \in [0, 1]$ , respectiv  $r \in [0, 1]$ . În domeniul discret, conturul  $v(s, r)$  reprezintă o înșiruire de puncte 3D ( $v(s, r) = \{p_i\}$ ,  $i = 1 \dots N_v$  unde,  $N_v$  este numărul total de puncte ce descrie conturul  $v(s, r)$ ). Asemănător figurii 5.2, figura 5.7 descrie un astfel de contur tridimensional. Plasat în spațiul  $\mathbb{R}^3$ , conturul activ tridimensional se va deplasa (deforma) într-o manieră similară ca aceea utilizată în spațiul bidimensional. Intuitiv, dacă *balonul* bidimensional prin umflare/dezumflare căpăta forma obiectului vizualizat, în spațiul 3D, acesta va replica același fenomen. Studiul conturilor ac-

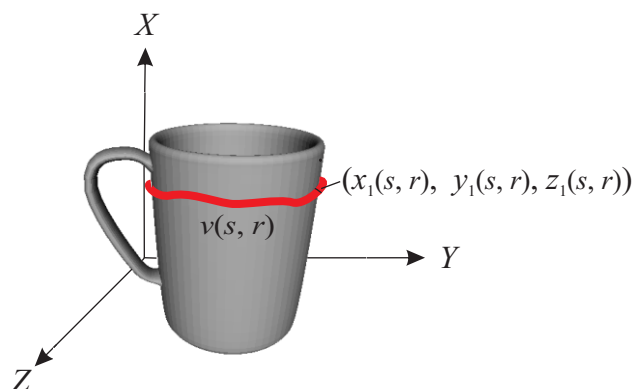


Fig. 5.7 Contur activ tridimensional.

tive tridimensionale nu este nou, multe lucrări din literatura de specialitate făcând referire la acest subiect [1, 106]. Totuși, un contur tridimensional nu este la fel de expresiv ca și o suprafață tridimensională. În spațiul Euclidian, obiectele sunt descrise prin forme și nu prin contururi asemănător spațiului bidimensional. În consecință, obiectivul principal, în continuare, va fi modelarea suprafețelor (formelor) utilizând o abordare similară cu aceea a conturilor active. Pentru a se evita orice confuzie, termenul de *contur activ tridimensional*, impropriu în această situație, va fi înlocuit cu termenul *formă activă tridimensională*. Această expresie este perfect justificată considerând scopul comun al unui contur (ca și în cazul 2D), de a prelua (atrage) forma unui obiect de interes.

Relativ la literatura de specialitate, termenul de formă activă este des utilizat în domeniul vederii artificiale [36, 183]. Totuși, în majoritatea studiilor aceasta este asociată spațiului bidimensional. Principiul de funcționare al formelor active bidimensionale este foarte similar cu cel al conturilor active. O formă predefinită (fie ea generică sau particulară) este potrivită (aliniată, inițializată) peste informația din scena de lucru (domeniul imaginii pentru cazul 2D, respectiv spațiul Euclidian pentru cazul 3D). În continuare, forma activă inițială este deformată atrăsă fiind de caracteristicile obiectului a cărui formă urmează a fi estimată. O abordare similară este folosită de autori și în [36], diferența fiind aceea că, în locul forțelor de atracție, autorii folosesc un *Model de Puncte Distribuite (MPD<sup>4</sup>)*. Acesta reprezintă, de fapt, un nor de puncte ce descrie un set de forme aliniate (suprapuse). Mai departe, norul va defini intervalul (variabilitatea) în care punctele formei active se pot deplasa pentru a se păstra forma inițială consistentă. Dezavantajul major al acestei abordări este acela că forma se poate deforma (modela) doar în maniera definită de setul de forme cu care se face antrenarea *MPD*-ului. Dacă forma din imagine este disimilă relativ la *MPD*, aceasta nu va putea fi modelată eficient.

În consecință, utilizarea energiilor pentru a controla modul în care forma activă este "sculptată" oferă atât flexibilitate crescută cât și eficiență sporită. Un exemplu de o astfel de formă activă modelată pentru a capta silueta unei căni poate fi observat în figura 5.8.

Pentru a putea utiliza principiul de modelare bazat pe energii (principiul conturilor active) forma activă 3D trebuie intuitiv reprezentată prin intermediul unui contur activ tridimensional ce

<sup>4</sup>Eng. Point Distribution Model (PDM)

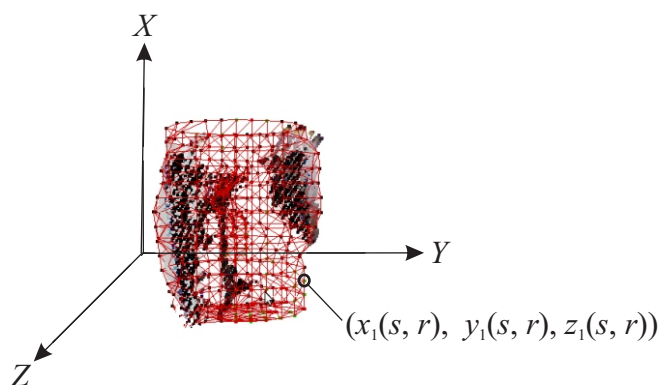


Fig. 5.8 Formă activă 3D, reprezentată prin intermediul unei rețele interconectate de triunghiuri de culoare roșie, potrivită optim peste un nor de puncte obiect, ce înfățișează o cană.

"șerpuiește" (țese) în spațiul Euclidian pentru a descrie o suprafață închisă (aceea a formei active 3D). Deplasarea punctelor formei active se realizează controlat sub influența forțelor interne, respectiv externe ce minimizează funcționala de energie  $\mathcal{E}(v(s, r), \mathcal{O}_k)$  unde,  $\mathcal{O}_k$  este norul de puncte obiect (a cărei siluetă trebuie să o capete forma activă 3D), iar  $v(s, r)$  este conturul tridimensional țesut pentru a descrie suprafața formei active considerate.  $v(s, r)$  poate fi considerat ca o aglomerație topologică de puncte. Cele mai comune astfel de suprafețe (forme) topologice sunt granițele obiectelor solide tridimensionale, de exemplu forma (suprafața) unei mingi. Punctele conturului activ  $v(s, r)$  sunt utilizate și pentru a descrie forma activă, motiv pentru care notația va fi comună. Ambele entități reprezintă densități de puncte tridimensionale.

Silueta inițială a formei active poate fi diferită de la caz la caz. Unele dintre cele mai eficiente forme inițiale sunt și *primitivele generice* (v. cap. 4). O descriere detaliată a procesului de inițializare a formelor active va fi prezentată în secțiunea 5.3.5.

Forțele interne și externe vor fi cele care vor asigura consistența și integritatea formei active 3D. În acest sens, forțele interne se definesc pentru a menține intactă coloana vertebrală (structura) a formei obiectului considerat, coloană ce oferă acestuia unicitate (particularitate). O primă energie internă definită în acest scop este  $E_{cont}$ . Aceasta asigură echidistanța dintre punctele formei active 3D. În lipsa ei, șansa apariției unor goluri în suprafața formei ar fi considerabil mai ridicată. O altă energie internă relevantă este cea de elasticitate,  $E_{curb}$ . Aceasta asigură maleabilitatea suprafeței formei active, respectiv capacitatea sa de a se mula pe diferitele protuberanțe (colțuri sau diferite deformări geometrice) ale norului de puncte obiect. O rigiditate mare a formei active 3D va produce, în final, un model relativ neted (fără prea multe caracteristici geometrice evidente).

Pe de altă parte, energiile externe,  $E_{ext}$ , sunt cele care ghidează suprafața formei active către granițele obiectului de interes. Datorită potențialului de atracție, dintre punctele formei active și cele ale obiectului de interes, energia externă poate fi considerată o funcție de potențial. Aceasta ia valori minime doar în vecinătatea norului de puncte obiect, mai precis pe granița acestuia.

Similar conturilor active bidimensionale, modelarea formelor active tridimensionale poate fi realizată utilizându-se următoarele abordări:

- prin minimizarea funcționalei de energie;
- dinamic utilizând forțe multiple .

### 5.3.1 Modelarea formelor active 3D prin minimizarea funcționalei de energie

Se consideră o formă activă  $v(s, r) = \mathbf{p}_1, \dots, \mathbf{p}_{N_v}$ , unde  $N_v$  este numărul total de puncte  $\mathbf{p}(x(s, r), y(s, r), z(s, r))$  ce descrie suprafața forme active tridimensionale. Având în vedere această reprezentare a forme active, funcționala de energie poate fi descrisă după cum urmează:

$$\begin{aligned} \varepsilon(v(s, r), \mathcal{O}_k) &= \varepsilon_{int}(v(s, r)) + \varepsilon_{ext}(v(s, r), \mathcal{O}_k) = \\ &= \int_0^1 [E_{int}(v(s, r)) + E_{ext}(v(s, r), \mathcal{O}_k)] ds dr = \\ &= \int_0^1 [\alpha(s, r)E_{cont}(v(s, r)) + \beta(s, r)E_{curb}(v(s, r)) + \\ &\quad + \gamma(s, r)E_{attractor}(v(s, r), \mathcal{O}_k)] ds dr, \end{aligned} \quad (5.16)$$

unde,  $\alpha$ ,  $\beta$  și  $\gamma$  sunt o serie de factori ce ponderează influența energiilor interne  $E_{int}$  ( $E_{cont}$ , respectiv  $E_{curb}$ ) și externe  $E_{ext}$  asupra forme active.

În domeniul discret, ecuația 5.16 poate fi scrisă astfel:

$$\varepsilon(v(s, r), \mathcal{O}_k) = \sum_{i=1}^{N_v} (\alpha(i)E_{cont}(i) + \beta(i)E_{curb}(i) + \gamma(i)E_{attractor}(i)), \quad (5.17)$$

unde,  $N_v$  este numărul total de puncte 3D (noduri) ce reprezintă forma activă considerată. În continuare, se va face o descriere a energiilor ce realizează modelarea forme active.

#### 5.3.1.1 Energia internă

Rolul unei astfel de energii este acela de a asigura o suprafață netedă descrisă prin puncte uniform distribuite. În esență, energia internă este compusă, la rândul ei, din alte două energii; energia de continuitate ( $E_{cont}$ ), respectiv energia de curbura ( $E_{curb}$ ). Matematic, energia internă poate fi exprimată prin următoarea expresie:

$$\begin{aligned} \varepsilon_{int} &= \int_0^1 [E_{int}(v(s, r))] ds dr = \\ &= \int_0^1 [\alpha(s, r)E_{cont}(v(s, r)) + \beta(s, r)E_{curb}(v(s, r))] ds dr = \\ &= \int_0^1 [\alpha(s)|v'(s)|^2 + \alpha(r)|v'(r)|^2 + \beta(s)|v''(s, s)|^2 + \beta(r)|v''(r, r)|^2 + \\ &\quad + \beta(s, r)|v''(s, r)|^2] ds dr, \end{aligned} \quad (5.18)$$

unde,  $\alpha(s)$  și  $\alpha(r)$  sunt factorii de ponderare ai elasticității de-a lungul axelor considerate  $s$ , respectiv  $r$ ,  $\beta(s)$  și  $\beta(r)$  sunt factorii de ponderare ai rigidității, iar  $\beta(s, r)$  determină rezistența conturului la răsucire.

Elementele derivativ de ordinul întâi din ecuația 5.18,  $(|v'(s)|^2 + |v'(r)|^2)$ , descriu energia internă de continuitate  $E_{cont}$ . Efectul acestei energii este de a simula proprietatea de membrană (întindere/contractare) a forme active. Se reamintește faptul că, un contur activ se comportă asemănător unui balon. Energia de continuitate asigură o poziționare consistentă și echidistantă a punctelor forme active. În cazul discret,  $E_{cont}$  poate fi descrisă astfel:

$$\begin{aligned}
E_{cont}(i) &= \left| \frac{dv}{ds} \right|^2 + \left| \frac{dv}{dr} \right|^2 = \\
&= \| \mathbf{p}_i(s) - \mathbf{p}_{i-1}(s) \|^2 + \| \mathbf{p}_i(r) - \mathbf{p}_{i-1}(r) \|^2 = \\
&= (x_i(s) - x_{i-1}(s))^2 + (y_i(s) - y_{i-1}(s))^2 + (y_i(r) - y_{i-1}(r))^2 + \\
&\quad + (z_i(r) - z_{i-1}(r))^2,
\end{aligned} \tag{5.19}$$

unde,  $\mathbf{p}_i$  este un punct aparținând formei  $v(s, r)$ .

În forma în care a fost descrisă, ecuația 5.19 implică utilizarea coordonatelor punctelor 3D curente ( $\mathbf{p}_i$ ) și anterioare ( $\mathbf{p}_{i-1}$ ) ale suprafeței active. Dacă, pentru un contur activ bidimensional identificarea punctului (nodului) anterior este trivială (pur și simplu nodul cu indexul curent minus o unitate), în cazul unei suprafețe tridimensionale problema este considerabil mai complexă. În cursul operației de țesere cu scopul creerii unei suprafețe, conturul activ tridimensional se intersectează cu el însuși. În acest fel vor apărea puncte duplicate (din punct de vedere al coordonatelor Euclidiene). Pentru a evita acest aspect, punctele duplicate sunt eliminate. Eliminarea acestora face imposibilă indexarea secvențială a punctelor în cadrul conturului. O abordare eficientă pentru rezolvarea acestei probleme a fost prezentată în [1]. Conform autorilor, pentru o identificare mai ușoară a punctelor următoare/anterioare relativ la un punct curent, se poate utiliza o reprezentare de tip rețea triunghiulară de puncte (v. sec. 2.2.1). Conform acestei reprezentări, cu fiecare cele mai apropiate 3 puncte se poate crea un triunghi unic. Utilizând astfel caracteristica de cel mai apropiat vecin al punctelor fiecărui triunghi se poate stabili proprietatea de punct anterior/următor. Totuși, această reprezentare nu este potrivită întotdeauna deoarece distribuția punctelor pe suprafața obiectului considerat poate fi neregulată și implicit, ecuația 5.18 este greu de verificat. Acest aspect este valabil pentru formele active inițiale de formă sferică [34]. Pentru a se evita acest neajuns și totuși a se păstra un aspect consistent al formei inițiale, poate fi utilizată o reprezentare de tip poliedric. De exemplu, pentru o formă inițială descrisă prin intermediul a 8 puncte 3D, forma poliedrică aferentă înfățișează un cub a cărui puncte pot fi tangente sferei imagine circumscrise acestuia. Pe măsură ce numărul punctelor crește, aspectul sferic al poliedrului este tot mai concludent. Un exemplu de poliedru descris prin intermediul a 24 puncte 3D poate fi observat în figura 5.9. Fiind un poliedru regulat, distanța dintre punctele adiacente este egală. Numărul de fețe ( $N_f$ ), respectiv muchi ( $N_e$ ) este dependent de numărul de puncte ( $N_p$ ) al poliedrului  $P$ . Un poliedru regulat poate fi creat în mod recursiv plecând de la un icosaedru (poliedru format din 20 de fețe) parcurgând următoarele etape:

- se poziționează un nou punct la mijlocul fiecărei muchii;
- se șterg toate muchiile anterioare și se crează unele noi între punctele având distanța mutuală egală cu jumătate din distanța vechilor muchii;
- se proiectează punctele nou introduce pe sfera care circumscrie poliedrul inițial.

Dacă pentru un icosaedru parametrii caracteristici sunt  $(F_0, E_0, \mathcal{P}_0) = (20, 30, 12)$  unde cu  $F$  este notată o față a poliedrului, iar cu  $E$  este notată o muchie a acestuia, pentru un poliedru regulat,

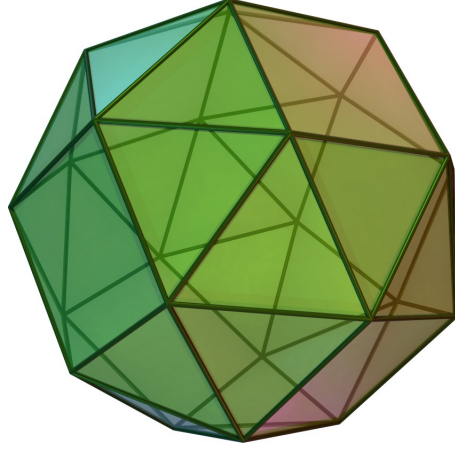


Fig. 5.9 Poliedru regulat (sursă: [www.wikipedia.org](http://www.wikipedia.org)).

obținut după  $j$  regresii aceștia sunt:

$$\begin{cases} F_j = 4F_{j-1}, \\ E_j = 2E_{j-1} + 3F_{j-1}, \\ \mathcal{P}_j = \mathcal{P}_{j-1} + E_{j-1}. \end{cases} \quad (5.20)$$

După efectuarea substituțiilor, ecuația 5.20 devine:

$$\begin{cases} F_j = 20 \cdot 4^j, \\ E_j = 30 \cdot 4^j, \\ \mathcal{P}_j = 2 + 10 \cdot 4^j. \end{cases} \quad (5.21)$$

Considerând reprezentarea poliedrică descrisă anterior, fiecare punct  $\mathbf{p}$  din  $v(s, r)$  va fi astfel conectat cu vecinii acestuia. Stabilirea punctului anterior ( $\mathbf{p}_{i-1}$ ) implică identificarea conexiunii (liniei de legătură) celei mai scurte dintre punctul curent și punctele vecine, puncte care, de altfel, compun poliedrele din jurul punctului curent. Stabilirea punctului următor ( $\mathbf{p}_{i+1}$ ) relativ la punctul curent presupune mascarea punctului cu indexul  $\mathbf{p}_{i-1}$  și identificarea din nou a liniei de legătură celei mai scurte. Acest mod de identificare este valabil doar în cazul suprafețelor reprezentate prin puncte uniform distribuite.

Pentru cazul în care forma activă inițială este o primitivă generică, reprezentarea poliedrică nu poate fi aplicată. În acest caz, se poate folosi, în locul acesteia, reprezentarea de tip rețea triangulară de puncte [54].

Revenind la formularea în domeniul discret a energiei  $E_{cont}$ , ecuația 5.19 poate fi îmbunătățită prin adăugarea termenului  $\bar{d}$ . Acesta reprezintă de fapt distanța medie dintre punctele forme active tridimensionale. Astfel, ecuația 5.19 poate fi rescrisă în felul următor:

$$E_{cont}(i) = (\bar{d} - (\|\mathbf{p}_i(s) - \mathbf{p}_{i-1}(s)\| + \|\mathbf{p}_i(r) - \mathbf{p}_{i-1}(r)\|))^2. \quad (5.22)$$

Prin adăugarea termenului  $\bar{d}$  se urmărește păstrarea unei distanțe egale nu doar față de punctul anterior  $\mathbf{p}_{i-1}$  ci față de toate punctele vecine acestuia. Efectul acestei energii poate fi controlat

prin intermediul factorilor  $\alpha(s)$ , respectiv  $\alpha(r)$ . Modul în care energia continuă influențează forma activă este descris intuitiv în figura 5.10(b).

Elementele derivative de ordinul 2 din ecuația 5.18,  $(|v''(s,s)|^2 + |v''(r,r)|^2 + |v''(s,r)|^2)$ , sunt cele care asigură elasticitatea (maleabilitatea) formei active. O valoare mare a acestei energii permite suprafeței active să pătrundă în proeminențele geometrice ale obiectului de interes din spațiul Euclidian. În cazul discret, aceasta poate fi descrisă prin relația:

$$E_{curb}(i) = \left| \frac{d^2v}{ds^2} \right|^2 + \left| \frac{d^2v}{dr^2} \right|^2 + \left| \frac{d^2v}{dsdr} \right|^2. \quad (5.23)$$

Factorii  $\beta(s)$ ,  $\beta(r)$  și  $\beta(s,r)$  sunt responsabili de netezirea formei active. O valoare mică a acestora determină energia internă să ia valori mici, mai ales, în zonele puternic deformate geometric, zone unde forma nu va fi capabilă să se muleze peste aceste caracteristici. Trebuie precizat faptul că factorii pondere sunt stabiliți independent pentru fiecare punct al formei active în parte. Modul în care energia de curbare influențează forma activă este descris intuitiv în figura 5.10(a).

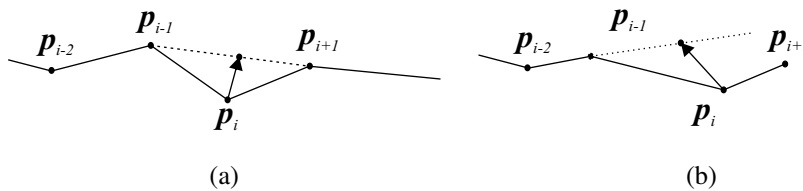


Fig. 5.10 Modul de acționare a energiei interne: (a) forța de elasticitate ce deplasează punctul  $p_i$  pentru a se asigura netezimea conturului; (b) forța de continuitate ce asigură liniaritatea punctului  $p_i$ , relativ la punctele precedente  $p_{i-1}$  și  $p_{i-2}$ .

### 5.3.1.2 Energia externă

Dacă energiile definite anterior exprimau diferite stări interne ale formei active, acest tip de energie nu folosește niciuna dintre proprietățile fizice ale acesteia. Rolul ei este acela de a împinge suprafața formei către diferitele caracteristici întâlnite în spațiu. Aceste caracteristici reprezintă frontierele obiectului a cărui înfățișare trebuie să o ia forma activă. Matematic, energia externă poate fi descrisă după cum urmează:

$$\varepsilon_{ext}(v(s,r), \mathcal{O}_k) = \int_0^1 [\gamma(s,r) E_{ext}(v(s,r), \mathcal{O}_k)] ds dr. \quad (5.24)$$

În cazul discret, ecuația 5.24 poate fi scrisă după cum urmează:

$$E_{ext}(v(s,r), \mathcal{O}_k) = E_{attractor}(v(s,r), \mathcal{O}_k). \quad (5.25)$$

$E_{attractor}(v(s,r), \mathcal{O}_k)$  se comportă ca o energie potențială de atracție. În forma în care este aceasta definită, energia  $E_{attractor}(v(s,r), \mathcal{O}_k)$  va prezenta valori minime doar în dreptul caracteristicilor asociate granițelor obiectului perceput din spațiu. Datorită fenomenului de atracție,  $E_{ext}$  poartă denumirea de *energie attractor* [1]. Prezența granițelor unui obiect în spațiul tridimensional este sesizabilă prin variația densității de puncte 3D din regiunea luată în considerare. Astfel, considerând un punct al formei active, capabil să se miște liber în spațiu, acesta se va deplasa în direcția

în care densitatea de puncte 3D (aparținând obiectului din scenă și nu a formei active) este cea mai mare. În acest sens, se consideră scalarul  $\mathcal{D}$  ca fiind o densitate de puncte obiect măsurată în jurul unui punct al formei active. Gradientul acestuia ( $\nabla \mathcal{D}$ ) de-a lungul unei anumite direcții va descrie variația densității de puncte obiect. Magnitudinea lui ( $\|\nabla \mathcal{D}\|$ ) va defini rata de creștere/descreștere a densității de puncte considerate. Ceea ce se urmărește prin utilizarea gradientului este *atingerea* graniței obiectului din spațiu. Astfel, în momentul în care magnitudinea densității de puncte va prezenta o fluctuație relevantă, energie de atracție va fi minimă fapt ce va determina funcționala de energie să oprească punctul considerat în poziția respectivă. Dacă în loc de gradient (derivata de ordinul unu) s-ar fi utilizat derivata de ordinul doi a densității de puncte 3D de-a lungul unei anumite direcții, atunci poziționarea optimă a punctului considerat nu ar fi fost pe graniță ci în mijlocul densității maxime de puncte. Acest aspect este eronat deoarece, în cazul formelor active 3D se urmărește identificarea formei exterioare a obiectului considerat. Matematic, energia de atracție poate fi exprimată prin relația:

$$E_{attractor}(v(s, r), \mathcal{O}_k) = -\gamma(s, r)\|\nabla \mathcal{D}\|. \quad (5.26)$$

Procesul de modelare al formei active fiind unul iterativ, deplasarea unui punct al formei active de-a lungul unei anumite direcții se realizează în trepte cu o anumită cantă<sup>5</sup> (de ex. 5 mm). Pentru fiecare această nouă poziție candidată, se măsoară densitatea de puncte  $\mathcal{D}$  din jur folosind doar informația (punctele) de proximitate din norul de puncte obiect  $\mathcal{O}_k$ . Pot exista numeroase situații (iterații) în care  $\mathcal{D}$  are valoarea 0. Acest lucru presupune că punctul formei active se află la o depărtare relativ mare față de norul de puncte obiect. Pentru acest caz, forma activă va prezenta un quantum al energiilor mare fapt ce va determina înaintarea (umflarea/dezumflarea) cu o cantă a formei active. Densitatea din jurul unui anumit punct se măsoară prin centrarea unei sfere în punctul considerat și contorizarea tuturor punctelor obiect din interiorul acesteia. Raza  $a$  a acestei sferei se stabilește dinamic în funcție de granularitatea norului de puncte obiect  $\mathcal{O}_k$ , precum și în funcție de dimensiunea globală a acestuia. Considerând o serie de obiecte de uz casnic cu dimensiune relativ redusă (aprox. 10 cm), o rază  $a$  candidată în acest sens poate fi de 0.005 m (5 mm). O ilustrare grafică a energiei externe pentru un punct considerat poate fi urmărită în figura 5.11

Energiile, atât cele interne cât și cele externe, sunt calculate pentru fiecare punct al formei active în parte. Dacă pentru cazul bidimensional existau 8 direcții de deplasare, considerând reprezentarea de tip rețea (v. fig. 5.3), în spațiul tridimensional, avându-se în vedere aceeași reprezentare de tip rețea, vor exista 26 de astfel de direcții candidate de deplasare. Acestea pot fi urmărite în figura 5.12.

Calcularea energiilor interne/externe în toate cele 26 de direcții posibile candidate și pentru fiecare punct al formei active în parte poate fi costisitoare. Pentru a se evita acest lucru se poate utiliza ca direcție de deplasare exclusiv direcția normalei la suprafață. Aceasta descrie, relativ la punctele formei active vecine, direcția naturală în care punctul considerat se poate deplasa (v. fig. 5.11(b)). În acest fel, în loc de 26 direcții candidate posibile vor exista doar 2, definite de sen-

<sup>5</sup>Eng. Step

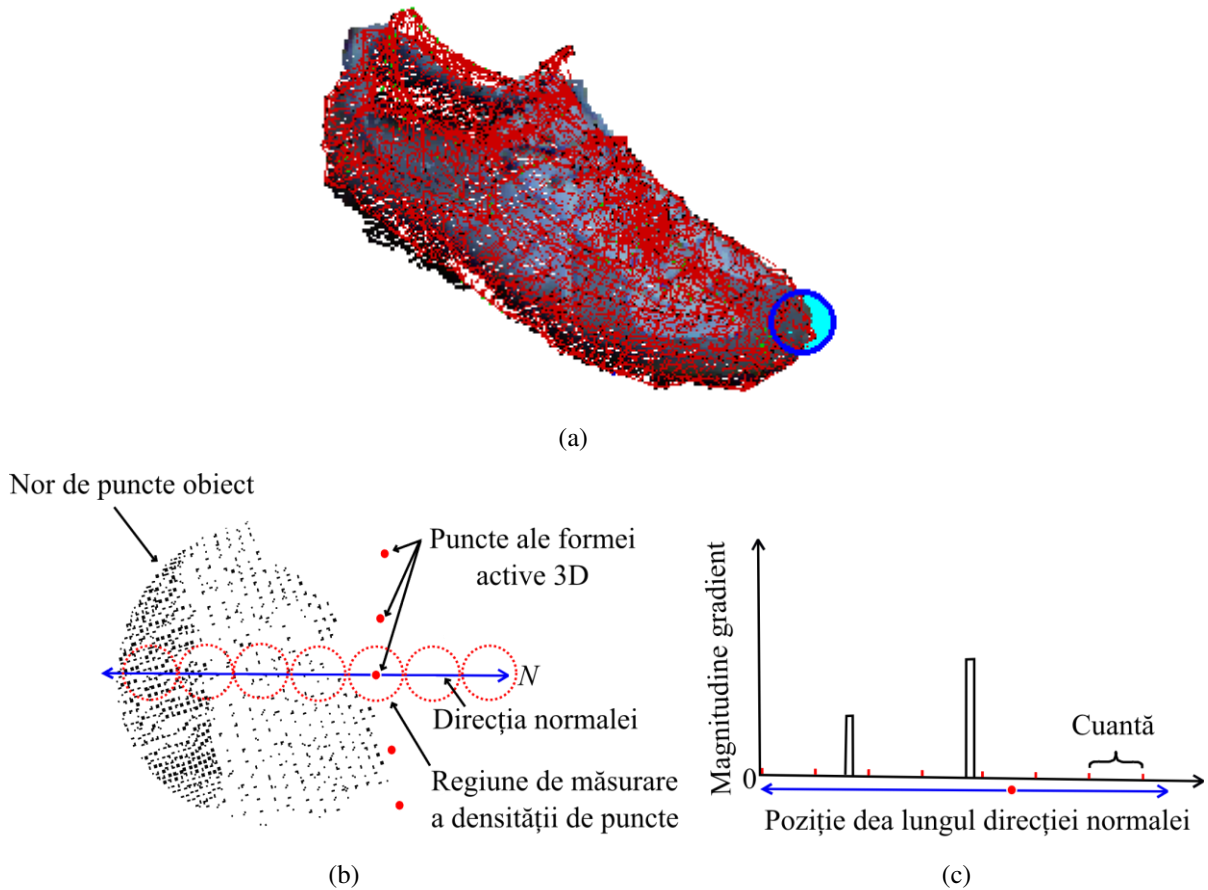


Fig. 5.11 Determinarea energiei externe pentru un punct de control aparținând formei active tridimensionale: (a) regiune de interes afectată, ilustrată cu privire la norul de puncte obiect și la forma activă inițială (înfățișată prin intermediul unei rețele triangulare de culoare roșie); (b) secțiune laterală (în planul  $xOy$ ) a regiunii afectate; (c) magnitudinea gradientului densității de puncte 3D de-a lungul direcției normalei a punctului de control considerat din regiunea afectată.

surile normalei (pozitiv/negativ) și asociate proceselor de umflare/dezumflare ale formei active. Un algoritm simplu și eficient pentru calcularea direcției normalei a fost prezentat în secțiunea 2.4.2. Normalele la suprafață, calculate pentru o formă activă tridimensională ce înfățișează un pantof pot fi observate în figura 5.14. Utilizând acest principiu, modul de deplasare al punctelor formei active poate fi urmărit în figura 5.13.

### 5.3.1.3 Minimizarea funcționalei de energie

Așa cum a fost menționat anterior, silueta optimă a formei active este atinsă în urma minimizării funcționalei de energie. Acest proces implică identificarea formei optime  $v(s, r)$  care satisface următoarea ecuație Euler-Lagrange [1]:

$$\gamma(s, r) \|\nabla \mathcal{D}\| + \alpha(s) \left| \frac{dv}{ds} \right|^2 + \alpha(r) \left| \frac{dv}{dr} \right|^2 - \beta(s) \left| \frac{d^2v}{ds^2} \right|^2 - \beta(r) \left| \frac{d^2v}{dr^2} \right|^2 - \beta(s, v) \left| \frac{d^2v}{dsdr} \right|^2 = 0. \quad (5.27)$$

La fel ca și în cazul 2D, ecuația 5.27 presupune ca energiile implicate în proces să fie în echilibru. Condițiile de îndeplinire a echilibrului sunt  $\alpha(s) = \alpha(r)$ ,  $\beta(s) = \beta(r)$  și  $\beta(s, r) = 0$ . Factorii de ponderare își mențin aceeași semnificație ca și în cazul conturului bidimensional, și anume, factorul  $\alpha$  controlează efectul energiei de continuitate, în timp ce,  $\beta$  ajustează quantumul energiei

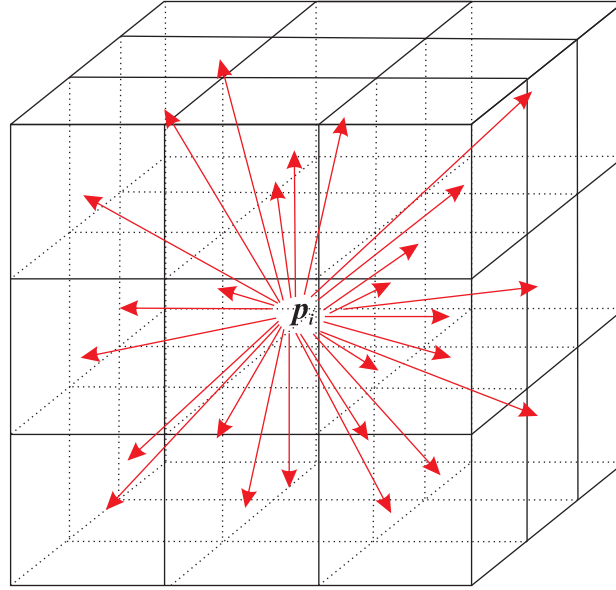


Fig. 5.12 Direcții de deplasare candidate ale unui punct de control al formei active (considerându-se o reprezentare a spațiului tridimensional de tip rețea).

de curbură. Ambii factori sunt aplicați pentru toate direcțiile spațiului cartezian.

Datorită influenței mult prea mari a energiilor interne asupra modului de deplasare a conturului activ apare o oarecare instabilitate. Limitarea energiilor interne conduce la o influență mai mare a densității punctelor (energiei externe) asupra conturului în sine. Totodată această abordare stabilizează conceptul de modelare de tip *balon* [34].

### 5.3.2 Modelarea dinamică a formei active utilizând forțe multiple

Acest tip de modelare este o alternativă la modelarea clasică prin minimizarea funcționalei de energie prezentată anterior. Spre deosebire de aceasta, utilizarea mai multor forțe externe pentru a controla forma activă poate fi mai avantajoasă în unele situații [146]. Astfel, energia poate fi exprimată ca superpoziția mai multor tipuri de forțe, după cum urmează:

$$F_{ext}(c) = F_1(c) + F_2(c) + \dots + F_n(c), \quad (5.28)$$

unde,  $n$  este numărul total de forțe externe distincte ce acționează asupra formei active. Dat fiind faptul că densitatea de puncte 3D din spațiu este singura caracteristică ce poate fi utilizată pentru a controla din exterior forma activă, principiul de modelare dinamică utilizând forțe multiple nu poate fi aplicat cu succes în acest caz.

### 5.3.3 Potrivirea formei active pe granițele obiectului din scena de lucru

Fie o formă activă descrisă prin intermediul unei funcții  $v(s, r)$  și un nor de puncte  $\mathcal{O}_k$  ce descrie o densitate de puncte înfățișând un obiect. Modelarea formei active, care în esență descrie tot un nor de puncte, relativ la norul de puncte obiect se realizează în trei etape. Schema-bloc a întregului proces poate fi observată în figura 5.15.

În prima etapă a procesului de modelare are loc inițializarea și poziționarea formei active în

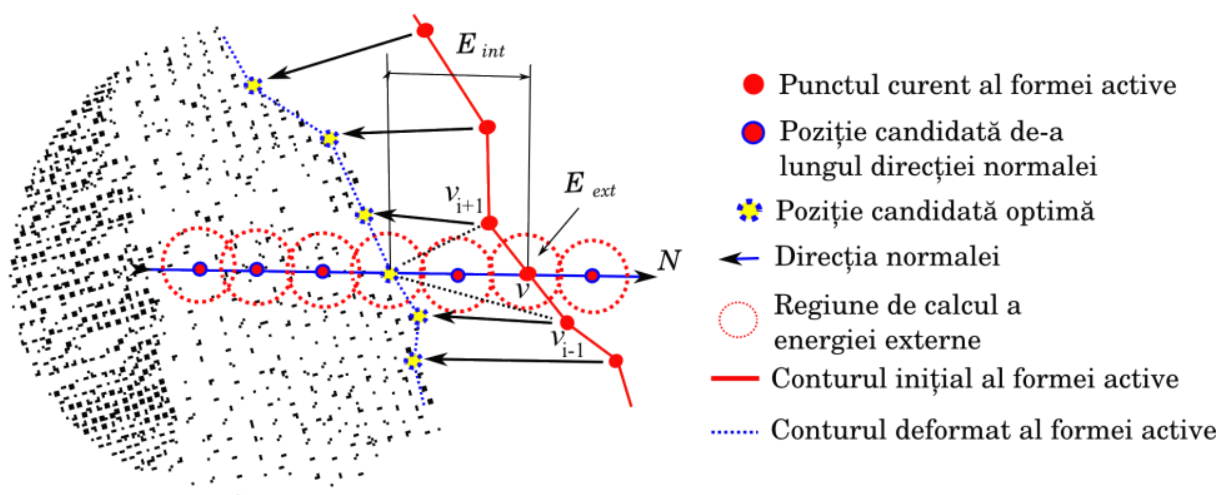


Fig. 5.13 Identificarea poziției candidate optime a unui punct de control ce aparține forme active.

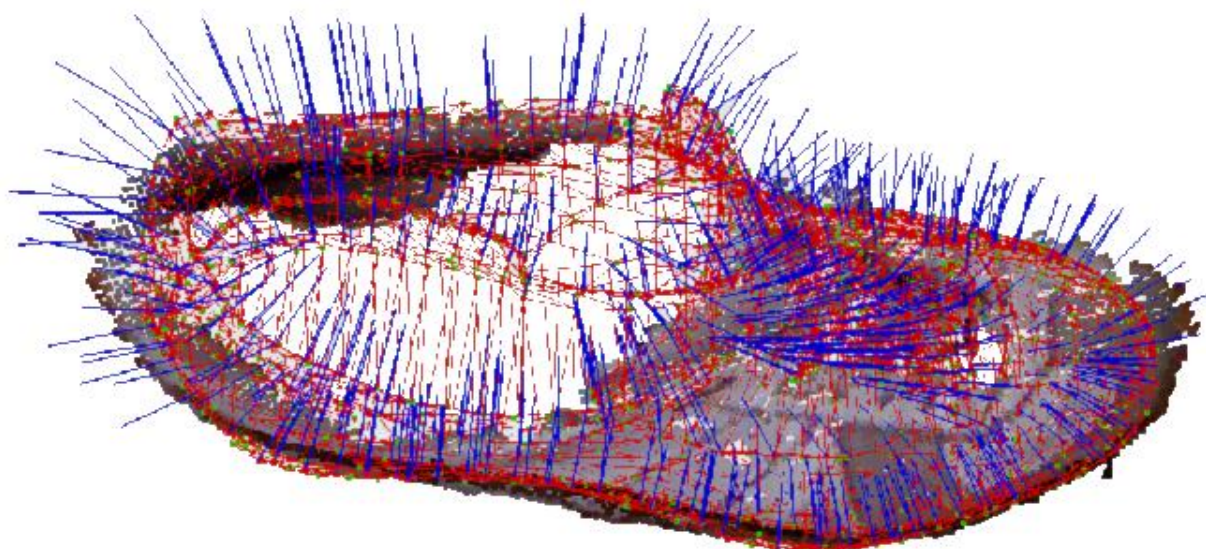


Fig. 5.14 Normalele la suprafață, descrise prin linii de culoare albastră, calculate în punctele de control ale forme active (reprezentată prin intermediul unei rețele triangulare de culoare roșie).

scenă. Pot exista, în acest sens, trei tipuri de forme active inițiale. Primul tip de formă activă este produs implicit și presupune utilizarea unei relații matematice pentru a genera un nor de puncte ce va reprezenta forma rigidă inițială. Cea mai comună astfel de formă este sfera. Intuitiv aceasta poate descrie cel mai eficient principiului balonului enunțat în secțiunile anterioare. Cel de al doilea tip de formă activă inițială este primitiva generică. Relativ la forma sferică menționată mai sus, o primitivă generică poate fi similară, încă de la început, cu obiectul de interes. Acest aspect influențează direct calitatea modelului final precum și timpul total de calcul. În final, ultimul tip de formă activă inițială este aceea creată manual de către un subiect uman utilizând un program special ce permite acestuia să interacționeze cu scena. Dintre toate formele active inițiale descrise, cea mai robustă este primitiva generică. Odată creată forma activă, poziționarea acesteia se realizează printr-un proces de transformare de similaritate (v. sec. 2.3.1) care are rolul de a alinia norul de puncte al primitivei generice cu norul de puncte obiect din scenă.

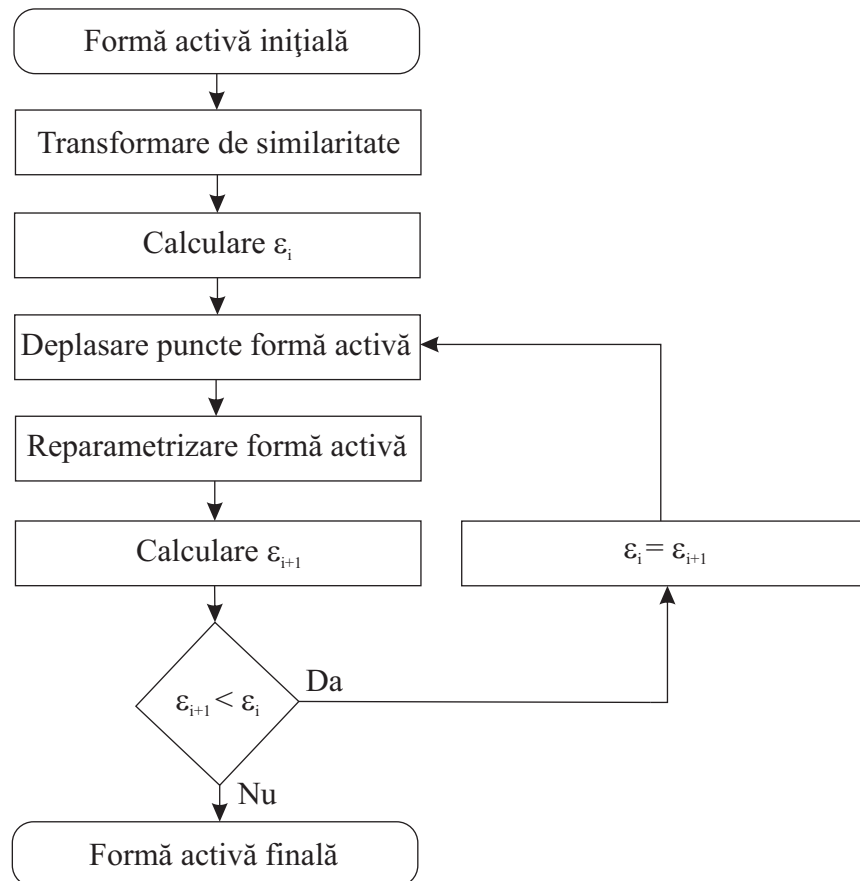


Fig. 5.15 Schema-bloc a procesului de potrivire a formei active pe granițele obiectului de interes.

Cea de a doua etapă a procesului de potrivire presupune calcularea normalelor la suprafață (v. sec. 2.4.2). Pentru eficiența calculului, normala la suprafață va fi calculată doar pentru punctele de control ce compun forma activă (v. fig. 5.14). De asemenea, aceste puncte de control vor fi singurele care vor fi modelate utilizându-se principiul contururilor active. Pentru fiecare astfel de punct, atât pentru sensul pozitiv cât și pentru cel negativ al direcției normalei, se va genera o hartă a gradientului densității de puncte asemănătoare celei din figura 5.11(c). Acesta din urmă va fi utilizat pentru calcularea energiei externe a punctului de control în diferitele poziții candidate.

A treia și ultima etapă presupune modelarea formei active propriu-zise relativ la norul de puncte obiect. Acesta este un proces iterativ și implică deplasarea fiecărui punct de control către caracteristici proeminente din spațiu. Un punct se poate deplasa doar o cantă în timpul unei iterații (de ex. 5 mm). După deplasarea, cu o cantă, a tuturor punctelor de control se determină quantumul funcționalei de energie  $\varepsilon_{i+1}(v(s, r), \mathcal{O}_k)$  pentru iterația considerată ( $i + 1$ ). Dacă aceasta este mai mică decât valoarea  $\varepsilon_i(v(s, r), \mathcal{O}_k)$  determinată anterior, atunci punctele de control mai pot înainta încă o cantă, în caz contrar, forma activă definită de  $\varepsilon_i(v(s, r), \mathcal{O}_k)$  este considerată a fi optimă, relativ la norul de puncte obiect (există un echilibru între energiile conturului). O formă activă finală a unei căni, obținută după doar 37 iterații, poate fi observată în figura 5.16.

#### 5.3.4 Reparametrizarea formei active

Odată cu evoluția formei active, punctele acesteia se vor deplasa controlat, conduse de diferitele

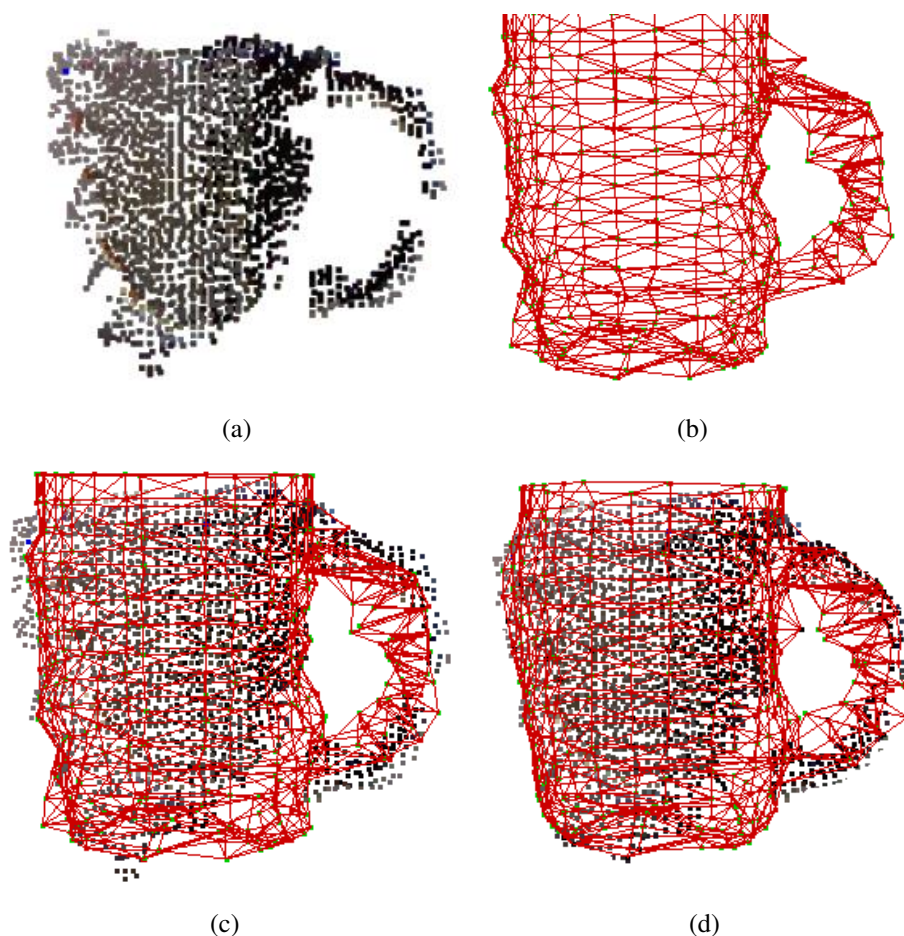


Fig. 5.16 Potrivirea formei active pe granițele obiectului de interes: (a) nor de puncte obiect; (b) formă activă inițială; (c) forme aliniate utilizându-se o transformare de similaritate; (d) formă activă modelată în raport cu norul de puncte obiect.

forțele interne și externe. Acest lucru atrage după sine ca distanța dintre puncte să scadă, respectiv să crească. Pentru a nu se pierde structura inițială al formei active se apelează la un proces de reparametrizare a suprafeței formei active în două etape.

În prima etapă se adresează un proces liniar de re poziționare a punctelor ordinare ce definesc forma generică (v. sec. 4.1.2). Se reamintește faptul că, potrivite (modelate) pe granițele obiectului din spațiu au fost doar punctele de control nu și cele ordinare, acestea din urmă rămânând în poziția în care primitiva generică a fost inițializată. Procesul liniar de re poziționare poate fi comparat intuitiv cu acțiunea de "tracțiune" a unei bucăți de material textil. Se consideră astfel o bucată plană de material textil. Aceasta se prinde de zona mediană și se încearcă tragerea lui într-o anumită direcție. Materialul din jurul punctului de prindere va urmări îndeaproape această regiune (v. fig. 5.17). Un comportament similar poate fi aplicat și la deplasarea fiecărui punct de control al formei active. Astfel, dacă un punct de control se mută de pe poziția inițială *A* pe poziția optimă candidată *B*, atunci toate punctele din jur vor fi trase după acesta păstrând, în acest sens, consistența suprafeței formei active. Quantumul punctelor afectate depinde de distanța dintre *A* și *B*. Dacă aceasta este foarte mică, regiunea afectată (ce conține punctele ce vor fi atrase) este și ea mică. Regiunea afectată prezintă o formă sferică. În centrul acestei sfere se află punctul de control considerat. Punctele aflate foarte aproape de centru vor fi afectate într-un quantum mai mare în

raport cu punctele aflate la marginea acestei sfere. Din centru sferei spre margine acest quantum afectează liniar noile coordonate ale punctelor. Relația matematică ce realizează repoziționarea liniară poate fi exprimată astfel:



Fig. 5.17 Comportamentul unui material textil atunci când este supus unei forțe de tracțiune.

$$\mathbf{p}_{nou} = \mathbf{p}_{vechi} \cdot \left(1 + \frac{d_c}{d_{max}}\right), \quad \mathbf{p} \in \{\mathbf{p}_c, \mathbf{p}_o\}, \quad (5.29)$$

unde,  $\mathbf{p}_{nou}$  și  $\mathbf{p}_{vechi}$  sunt coordonatele noi, respectiv vechi ale punctului aflat în interiorul regiunii afectate de rază  $d_{max}$ , iar  $d_c$  este distanța Euclidiană dintre punctul de control central și punctul  $\mathbf{p}$  vecin din interiorul regiunii afectate. Cu  $\mathbf{p}_c$  sunt notate punctele de control, iar cu  $\mathbf{p}_o$  sunt notate punctele ordinare.  $d_{max}$  reprezintă distanța Euclidiană dintre poziția  $A$  inițială a punctului de control central și poziția  $B$  finală a acestuia. Dacă un punct de control, altul decât cel central utilizat la reparametrizare, se află în interiorul regiunii afectate, acesta va fi considerat ca și un punct oarecare. Pentru a susține aceste afirmații, în figura 5.18 este prezentat un exemplu grafic în care un punct de control este tras într-o anumită direcție.

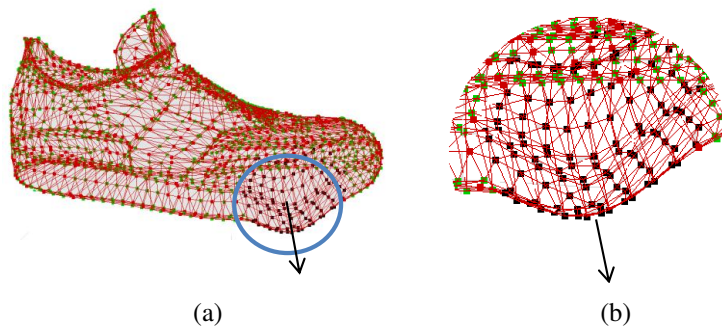


Fig. 5.18 Reparametrizarea forme active pentru cazul deplasări a unui punct de control într-o anumită direcție: (a) vedere globală; (b) vedere locală a regiunii afectate.

Cea de a doua etapă a procesului de reparametrizare implică adăugarea/eliminarea de puncte

ale forme active (de control sau ordinare). Procesul este identic, ca și în cazul reparametrizării conturilor active bidimensionale. Se definește astfel o distanță prag inferioară ( $d_{p-min}$ ) și una superioară ( $d_{p-max}$ ). Dacă distanța Euclidiană dintre cele mai apropiate două puncte vecine ( $p_i$ , respectiv  $p_{i+1}$ ) este mai mică decât  $d_{p-min}$  atunci punctul  $p_{i+1}$  este candidat pentru eliminare. Viceversa, dacă distanța Euclidiană dintre cele mai apropiate două puncte vecine ( $p_i$ , respectiv  $p_{i+1}$ ) este mai mare decât  $d_{p-max}$ , atunci se va adăuga un nou punct la jumătatea distanței dintre  $p_i$  și  $p_{i+1}$ , iar punctul inițial  $p_{i+1}$  va deveni  $p_{i+2}$ . În acest fel, odată cu înaintarea punctelor forme active se va asigura o granularitate constantă a reprezentării acesteia.

### 5.3.5 Considerații privind forma activă inițială

Inițializarea forme active este o etapă esențială a procesului de modelare a forme active. Poziția și forma acesteia sunt două aspecte foarte importante. Cu privire la poziția forme active inițiale, pot exista două situații: fie se poziționează manual de către un operator uman utilizând o interfață om-mașină specială, fie se utilizează o abordare automată. Deși este foarte precisă, metoda manuală prezintă dezavantajul utilizării factorului uman, care în unele situații poate fi incomod. În acest sens, metoda automată rămâne cea mai viabilă soluție. Astfel, în primă instanță aceasta presupune recunoașterea obiectului de interes (a cărui volum final se dorește a fi reconstruit utilizând primitive generice). Pentru aceasta se poate utiliza o abordare precum aceea prezentată în secțiunea 4.3 [180]. Odată identificat obiectul de interes, se poate calcula centrul de greutate al acestuia, centru ce va reprezenta poziția în care va fi potrivită (aliniată) forma activă inițială. Pentru aliniere se utilizează un proces de transformare de similaritate.

Pe de altă parte, aspectul forme active inițiale este foarte important. Relativ la acesta pot exista trei metode de a crea o formă activă inițială. Prima, este o metodă manuală prin care, un operator uman indică o serie de puncte tridimensionale cu ajutorul unei interfețe speciale. Aceste puncte descriu forma activă inițială ce va fi modelată și reparametrizată continuu pentru a se mula pe informația obiectului considerat. Forma astfel creată se bucură de o particularitate mare deoarece operatorul poate crea aceste puncte foarte aproape de granițele obiectului din scenă. Acest lucru conduce, mai departe, la reducerea semnificativă a numărului de iterații necesar pentru a converge către forma finală. Principalele neajunsuri sunt timpul mare de producere și, nu în ultimul rând, caracterul subiectiv al acesteia.

Cu scopul eliminării neajunsurilor mai sus menționate, cea de-a doua metodă de generare a forme active inițiale utilizează o formulare matematică ce produce diferite forme geometrice implicite [66]. Sferele, cuburile sau cilindrii sunt doar câteva dintre formele geometrice astfel generate. Timpul de generare redus precum și complexitatea scăzută le recomandă cu încredere pentru a fi folosite în acest scop. Dezavantajul acestora însă este generalitatea formelor create. Acest aspect influențează direct timpul de modelare (de convergență către o formă finală), fiind astfel necesare mai multe iterații pentru a împinge forma regulată către granițele obiectului din scena de lucru.

Ultima și cea mai relevantă metodă de generare a unei forme active este aceea care folosește o serie de primitive generice (v. cap. 4). Motivația unei astfel de abordări este simplă. Din moment ce obiectul de interes din scenă este identificat și segmentat (v. sec. 4.3) utilizarea unei forme

cunoscute a acestuia, pe post de formă activă inițială, este firească. Comparativ cu forma activă inițială ce înfățișează o sferă, primitivele generice oferă o acuratețe mai ridicată a formei active finale obținute, mai ales pentru obiecte cu structură geometrică relativ complexă. Un atu important al utilizării primitivelor generice ca forme active inițiale este și faptul că, în părțile ocluzionate, în care percepția informației obiectului de interes este imposibilă datorită vederii unidirecționale, forma aproximată este naturală și nu regulată, ca în cazul utilizării unei forme sferice. Figura 5.19 ilustrează grafic procesul de inițializare al unei forme active în cazul ultimelor două tipuri de forme inițiale mai sus prezentate, și anume forme regulate (sferă), respectiv primitive generice.

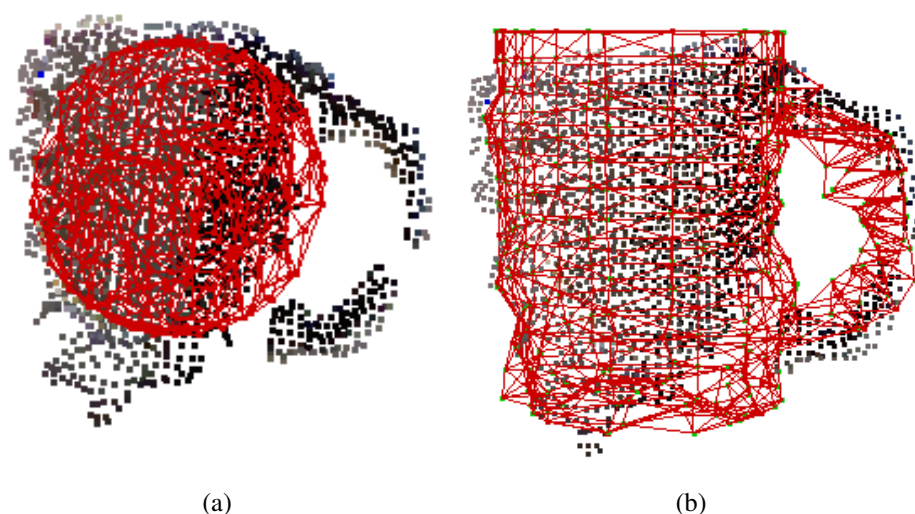


Fig. 5.19 Tipuri de forme active inițiale (descrise prin intermediul unor rețele triangulare de culoare roșie), aliniate la norii de puncte aferenți: (a) formă inițială generată implicit (sferă); (b) primitivă generică.

## 5.4 Concluzii

În formă în care a fost generată și aliniată, relativ la obiectul incomplet vizualizat din scenă, primitiva generică reprezintă doar o aproximare, un estimat ce poate prezenta totuși diferențe față de obiectul de interes perceput. Pentru unele obiecte, această formă estimată poate fi prea brută (imprecisă) pentru a asigura un proces de prehensare și manipulare stabil și sigur. În acest sens, prezentul capitol tratează problema creșterii asemănării dintre primitiva generică aliniată și obiectul perceput din scenă adresând un proces de modelare activă. Obiectivul este ca, odată aliniată forma primitivei generice în raport cu reprezentarea parțială a obiectului vizualizat, suprafața primitivei să fie condusă (împinsă) către aceea percepută a obiectului din scenă crescând astfel asemănarea dintre acestea. O soluție simplă și eficientă pentru manipularea punctelor acestor forme este oferită de metoda contururilor active. În forma în care acestea au fost introduse de către Kass în [99], ele se adresează doar contururilor bidimensionale. În acest sens, secțiunile 5.1, respectiv 5.2, au prezentat o descriere sumară a evoluției acestora, precum și a avantajelor oferite în special în cadrul proceselor de segmentare ale obiectelor. Totodată, în cadrul secțiunii 5.2, a fost prezentat și aparatul matematic care stă la baza modelării unui contur activ bidimensional.

Necesitatea pregnantă a modelării formelor tridimensionale a condus la extinderea conceptului de contur activ și pentru spațiul tridimensional. Astfel, în secțiunea 5.3 a fost prezentat aparatul matematic, într-o formă derivată din aceea a conturilor active bidimensionale, ce permite mode-

larea unei forme tridimensionale. Pentru a putea considera similare cele două abordări, bidimensională, respectiv tridimensională, suprafața primitivei generice a fost considerată drept un contur activ tridimensional ce șerpuiește în spațiu pentru a crea o suprafață închisă. Pentru a avea o raportare corectă între puncte, suprafața primitivei generice a fost reprezentată prin intermediul unei rețele triangulare de puncte 3D. În continuare, modelarea unei suprafețe tridimensionale poate decurge în două sensuri, și anume: prin minimizarea funcționalei de energie, respectiv utilizându-se forțe multiple. Primul tip de modelare, ce a fost folosit și în cadrul prezentei cărți, este și cel mai des utilizat. În momentul când un punct al primitivei generice se deplasează, atras fiind de o anumită caracteristică proeminentă din spațiu, acesta consumă o serie de energii. În funcție natura și proveniența lor, energiile pot fi interne, respectiv externe. Modalitatea prin care aceste energii pot fi determinate a fost descrisă în cadrul secțiunii 5.3.1. Rolul energiilor este acela de a menține forma primitivei generice compactă și, totodată, de a ghida corespunzător punctele acesteia către informația percepută a obiectului vizualizat. Pentru accelerarea întregului proces, modelarea a fost aplicată doar asupra punctelor de control. Punctele ordinare au fost modelate utilizându-se un proces de reparametrizare liniară, proces prin care coordonatele punctelor ordinare din jurul unui punct de control sunt ajustate (în conformitate cu noua poziție a punctului de control) utilizând o formulare matematică simplă ca aceea descrisă în secțiunea 5.3.4.

La finalul prezentului capitol au fost formulate o serie de considerații privind inițializarea procesului de modelare. În formularea originală a conturilor active, conturul inițial este descris printr-o formă regulată (sferă, cub, etc.). În funcție de tipul obiectului reconstruit/segmentat, această reprezentare inițială poate fi improprie. Utilizarea unei forme particulare (primitivă generică) drept formă activă inițială determină ca punctele ce compun această formă să străbată o distanță mai scurtă în spațiul 3D, și implicit procesul de modelare să convergă mult mai rapid. Un efect direct al acestui aspect este un timp de modelare foarte scăzut. Desigur, dacă forma inițială, descrisă prin intermediul primitivei generice, este foarte asemănătoare cu aceea a obiectului perceput din scenă, procesul de modelare va decurge foarte rapid deoarece majoritatea punctelor forme inițiale se află deja în poziție optimă. Forma finală obținută, în regiunile în care a existat informație percepută, va fi similară cu obiectul vizualizat, iar în regiunile ocluzionate (fără informație 3D prezentă) va prezenta un caracter generic introdus de forma primitivei generice utilizate.



## 6. Prehensarea și manipularea obiectelor tridimensionale

---

Considerații

Sinteza operației de prehensare a unui obiect

Concluzii

---

În ultimele decenii, roboții autonomi au fost implicați, din ce în ce mai mult, în activități ce vizează persoanele. Odată cu evoluția tehnologiei, aceștia au putut îndeplini multiple sarcini făcând astfel viața oamenilor mai ușoară. Numeroase sarcini, care pentru o persoană par simple, pentru un robot pot fi dificil de realizat. Odată cu creșterea complexității sarcinii crește, în mod evident, și complexitatea structurii hardware și software a robotului în special în cazul utilizării, de către robot, a unui sistem de percepție a mediului înconjurător. Vederea artificială a robotului, așa cum este numită în literatura de specialitate, lărgeste considerabil capabilitatea acestuia de a îndeplini diverse obiective. Acest domeniu de studiu este cunoscut în literatura de specialitate sub numele de *vederea mașinilor*<sup>1</sup> și reprezintă o combinație între diferitele tehnici ce definesc *vederea artificială asistată de calculator* și o serie de soluții hardware. Acest tip de "percepție" a permis rezolvarea, de către robot, a unor sarcini diferite din diverse domenii precum industria [58], construcțiile [71], domeniul social [45], agricultura [82], medicina [76], reabilitarea persoanelor cu handicap [116] etc.

Având în vedere ariile de aplicabilitate menționate anterior, roboții de serviciu au cunoscut cea mai mare atenție din partea comunității științifice. Roboții de serviciu reprezintă o clasă

---

<sup>1</sup>Eng. *Machine Vision*

aparte de roboți care, în mod autonom, sau semi-autonom, îndeplinesc diferite sarcini ce aduc beneficii directe persoanelor [181]. Un tip aparte al roboților de serviciu îl reprezintă roboții de asistență, al căror principal rol este acela de a ajuta persoanele cu dizabilități să realizeze o serie de operații pe care independent (fără ajutor) le-ar fi fost imposibil să le poată efectua. Una dintre cele mai importante capabilități a unui robot de asistență este aceea de a prehensa și manipula obiecte. Necesitatea obținerii unei forme rigide, utilă asigurării unei astfel de capabilități, a motivat studierea aplicabilității metodei de reconstrucție a formelor propusă în prezenta carte, cu precădere, în această ramură a roboților autonomi. Un astfel de robot, ce poate folosi metoda de reconstrucție propusă de autor, este, de exemplu, *FRIEND*<sup>2</sup> ce se află deja la a treia generație [116]. *FRIEND* se încadrează în gama roboților de asistență semi-autonomi destinați să ajute persoanele cu invaliditate atât la nivelul membrelor inferioare cât și la nivelul celor superioare. Capabilitatea sistemelor motrice superioare este asigurată prin utilizarea unui manipulator cu 7 grade de libertate. Autonomia robotului este obținută prin utilizarea unor senzori de percepție diferiți ce îi permit să observe și să înțeleagă mediul în care se deplasează. Astfel, unul dintre cei mai importanți senzori, în acest sens, este cel prin care se percepe scena în adâncime. O descriere a principalelor tipuri de percepție, între care se regăsește și cel utilizat de *FRIEND*, a fost prezentată în secțiunea 2.1. Pe baza informației primite de la acești senzori, robotul autonom, poate să stabilească o traiectorie de navigare, să evite obstacole, să localizeze obiecte, să le manipuleze, etc. Obiectele, în cadrul unui scenariu comun, reprezintă aproximativ 90% din totalul entităților cu care, un subiect uman intră în contact [88]. Datorită acestui aspect, o direcție importantă a cercetărilor din domeniu s-a orientat în sensul rezolvării problemei prehensării și manipulării obiectelor de către structurile robotice autonome.

## 6.1 Considerații

Prinderea/apucarea (prehensarea) obiectelor din scena de lucru de către o structură mecanică reprezintă o problemă dificilă, a cărei rezolvare implică utilizarea unor cunoștințe complexe de mecanică și electronică. Asemănător comportamentului uman, roboții au învățat să meargă, să sară, să alerge sau să interacționeze și să "socializeze" cu persoanele din jur. Astfel, modul uman de comportare poate furniza un răspuns și la problema prehensării și manipulării mecanizate a obiectelor.

În urma observațiilor efectuate pe subiecți umani, s-a constatat că aceștia folosesc, pentru prehensarea obiectelor, o serie de cunoștințe anterioare (instanțe vizualizate a priori ale obiectului sau ale unui similar). Practic, o persoană poate învăța cum să apuce un obiect necunoscut având în vedere o serie de experiențe anterioare legate de prinderea unor obiecte asemănătoare. Aceste modalități de prehensare pot fi obținute atât prin învățare (analitic sau empiric) cât și prin demonstrare (proces realizat cu ajutorul unui operator uman). Astfel, în cazul unui scenariu real, robotul, în primă fază identifică obiectul de interes, după care, utilizând modalitățile de prindere ale obiectului identificat (stocate într-o bază de date), execută, în deplină siguranță, prehensarea și manipularea acestuia.

---

<sup>2</sup>Eng. *Functional Robot with dexterous arm and user frIENdly interface for Disabled people*

Conform studiului efectuat de Helbig în [80], creierul uman are o metodă simplă și eficientă de a planifica o operație de prehensare și anume, la observarea unui obiect nou, în subconștient este evaluată accesibilitatea acestuia cu privire la operația de prehensare. Când este observat un obiect nou, diferit de cel anterior dar cu o accesibilitate similară, în mod automat din memorie este extrasă informația referitoare la obiectul anterior, informație care dă posibilitatea de prehensare a obiectului nou vizualizat. Acest comportament a fost studiat de numeroși cercetători confirmându-se astfel teoria prezentată anterior [178]. Cu alte cuvinte, memoria noastră vizuală poate fi "corelată" cu modalități prin care diferite obiecte pot fi prehensate. Conform acestor observații, se pare că nu este imperativ necesar ca obiectul să fie identificat cu acuratețe maximă. Mai degrabă, este suficient ca obiectul identificat să fie destul de asemănător cu unul din baza de date deoarece este foarte posibil ca cele două obiecte să poată fi prehensate într-o manieră similară. Studiind îndeaproape acest concept, numeroși cercetători au reușit să-l implementeze și în structurile robotice pentru prehensarea obiectelor [65, 128].

Un aspect esențial în abordarea prezentată mai sus este acela că se pot prehensa atât obiecte 3D cât și obiecte a căror structură prezintă un caracter 2.5D (obiecte vizualizate unidirecțional). Deoarece sistemul de recunoaștere a obiectelor se bazează, în exclusivitate, pe geometria obiectului și nu pe culoarea sau textura acestuia, se pot determina cu ușurință regiuni similare între obiectul vizualizat (ce urmează a fi prehensat) și obiectele din baza de date [64, 102, 128, 168]. Atașat fiecărui obiect din baza de date se pot corela și o serie de modalități (configurații) de prehensare a acestuia. Odată identificat, în baza de date, obiectul cu cea mai mare similaritate, în raport cu obiectul vizualizat, se pot utiliza, în continuare, configurațiile de prehensare ale acestuia și pentru a se prehensa obiectul nou vizualizat.

O abordare ușor diferită a proceselor de prehensare și manipulare a obiectelor 2.5D este și aceea prin care "prinderea" obiectelor se realizează cu ajutorul unor senzori tactili. Astfel, prin pipăire, se poate construi atât un model tridimensional al obiectului de interes cât și o configurație de prindere a acestuia. Numeroase studii au fost îndreptate în această direcție, studii care s-au concretizat prin dezvoltarea unor algoritmi eficienți și robuști [140, 163]. În acest sens, Bierbaum în [16] propune potrivirea, peste datele provenite de la senzorii tactili, a unor forme supercuadrice (forme implicite) cu scopul obținerii unui volum estimat al obiectului vizualizat.

Un studiu relevant, realizat în sensul reconstrucției obiectelor vizualizate unidirecțional în scopul prehensării și manipulării acestora, este și cel prezentat de Marton în [118]. Acesta propune o metodă de a se estima volumul unui obiect plecând de la premiza că acestea prezintă o axă verticală de simetrie. Informația lipsă a obiectului considerat este reconstituită pe baza considerentelor de simetrie ale suprafețelor. Totuși, în practică, abordarea menționată încearcă potrivirea unor primitive geometrice (cilindrii, sfere, etc.) peste informația percepută. Deși este un proces foarte rapid, această abordare nu este fezabilă pentru obiecte complexe.

Nu în ultimul rând, primitivele generice introduse în prezenta carte [33] oferă o soluție viabilă la problema prehensării obiectelor vizualizate unidirecțional. În acest caz, structura aparatului de reconstrucție volumetrică a obiectelor face ușoară atașarea, fiecărei primitive generice din baza de date, a unui set de configurații de prehensare, existând astfel posibilitatea replicării comportamentului uman, prezentat la începutul acestei secțiuni, vis-a-vis de procesul de prehensare și

manipulare a obiectelor. Odată identificată primitiva generică cea mai apropiată de obiectul vizualizat, în mod implicit sistemul va beneficia și de o serie de configurații de prehensare a acestuia. În plus, în funcție de constrângerile externe (obstacole, limitări mecanice sau software) se poate alege configurația de prindere cea mai oportună ce satisface sarcina de prehensare și manipulare a obiectului de interes.

Configurația optimă, extrasă din baza de date, trebuie să ofere informații legate de orientarea prehensorului, cinematica articulațiilor acestuia sau quantumul forțelor ce trebuiesc aplicate, pe fiecare punct de contact, pentru a se asigura o strângere corespunzătoare.

În literatura de specialitate, termenul de prehensor este uzual asociat cu cel de efector final. În fapt, efectorul final reprezintă o denumire generală pentru elementul de acționare ce se află atașat brațului robotic și care intră în contact direct cu obiectele manipulate. În funcție de complexitatea operației executate, efectoarele finale pot avea de la 2 grade de libertate până la 22 de grade de libertate. Astfel, manipularea poate fi realizată, fie cu simplul scop de a prinde și transporta un anumit obiect, fie pentru a efectua o operație de îndemânare<sup>3</sup> asupra obiectului considerat. Spre exemplificare, în figura 6.1 pot fi observate cele două sensuri menționate ale operației de manipulare a unui pantof.



Fig. 6.1 Tipuri de manipulare; (a) prehensare în scopul transportării; (b) prehensare în scopul efectuării unei operații de îndemânare.

Manipularea prin apucare are ca obiectiv fixarea unui anumit obiect prin prinderea lui. Astfel, capacitatea de prindere a obiectului este definită de configurația manipulatorului (poziția celor  $n$  grade de libertate). Prehensorul, în calitate de efector final este, așa cum s-a mai prezentat, cel care intră în contact direct cu obiectul în cauză în urma operației de prindere. Principial, pentru a se putea apuca eficient un obiect sunt necesare minim 3 puncte de contact între prehensor și obiect [129]. Cele mai simple și eficiente exemple, în acest sens, sunt prehensoarele cu 3 degete Barrett [147] și LMS [62]. Designul inovativ precum sistemul de acționare eficient fac din aceste prehensoare unele dintre cele mai utilizate elemente din robotică, pentru prinderea diferitelor obiecte.

Pe de altă parte, manipularea cu scopul efectuării unei operații de îndemânare presupune ca obiectul de interes să fie manevrat cu o îndemânare similară cu aceea a degetelor unei persoane. Acest lucru implică atingerea unei dexterități foarte ridicate a prehensorului. Pe lângă complexitatea mecanică a acestuia intervine și un aspect foarte relevant, și anume controlul degetelor. Un

<sup>3</sup>Eng. *Dexterous Manipulation*

astfel de prehensor complex încearcă să relice, cu acuratețe, modelul unei mâini umane. Un exemplu în acest sens este prehensorul antropomorf cu 5 degete FRH-4 [61]. Deși prezintă o complexitate extrem de ridicată, datorată mobilității celor 5 degete, acesta poate satisface cu succes și cerințele unei prinderi normale (ce poate fi realizată prin intermediul unui număr redus atât de degete cât și de grade de mobilitate). Totodată, fiind acționat hidraulic, acesta poate fi utilizat în siguranță și la interacționarea cu persoane.

Unele dintre cele mai importante proprietăți ale unui prehensor sunt dexteritatea, echilibrul (capacitatea de a imobiliza un obiect), stabilitatea și comportamentul dinamic. Controlul tuturor acestor aspecte necesită o serie de metode și algoritmi care, în literatura de specialitate sunt referite ca sinteza metodelor de prehensare robotizată. Un prim obiectiv al oricărei operații de prindere a unui obiect este acela de a asigura stabilitate. O configurație (modalitate) de prindere a unui obiect este considerată stabilă dacă orice perturbație, ce apare și forțează obiectul să își schimbe poziția, produce o forță de restabilire ce va aduce sistemul din nou la starea inițială [86]. În legătură directă cu această proprietate se află modalitatea prin care un obiect poate fi prins. Se pot defini, în acest sens, două tipuri de prinderi:

- prindere prin strângere<sup>4</sup>;
- prindere prin blocare<sup>5</sup>.

Prinderile prin strângere reprezintă o categorie a prinderilor echilibrate. Acest tip de prindere prezintă proprietatea importantă de a fi stabile. Totuși, nu toate prinderile stabile sunt realizate prin strângere. Conform studiului prezentat în [15], analiza stabilității acestui tip de prindere este echivalentă cu stabilitatea și rezolvarea unei ecuații diferențiale ordinare.

Pe de altă parte, prinderea prin blocare oferă o condiție mai puternică decât aceea prin strângere. Analiza unei prinderi prin blocare este intrinsec geometrică. O operație de prindere atinge proprietatea "de blocare" dacă și numai dacă poate realiza operația de prindere prin strângere însă fără a aplica forțe de presiune asupra punctelor de contact. Practic, se încearcă prinderea obiectului în cauză fără a îl strânge în prehensor. În această situație, prinderea prin strângere și aceea prin blocare sunt duale [193]. Evident, stabilitatea este necesară dar nu și suficientă pentru o strategie de prindere. Uzual, prehensarea unui anumit obiect se realizează pentru a se îndeplini o anumită sarcină. Evident, pentru a o finaliza cu succes, configurația de prindere trebuie să fie compatibilă cu cerințele sarcinii. În acest sens, datorită formelor și dimensiunilor variate ale obiectelor de interes, strategia de prindere trebuie să fie capabilă să prehenseze și obiectele noi pe care robotul le poate observa. Prin obiect nou se înțelege acel obiect care este văzut pentru prima oară de către sistemul robotic. Strategia de prindere trebuie să ofere, așa cum s-a mai menționat, stabilitate, să fie compatibilă cu sarcina impusă și, nu în ultimul rând, să poată fi aplicată și pentru obiecte noi. Sinteza unei astfel de strategii de prindere trebuie să răspundă la următoarea întrebare: *Cum să prind obiectul pentru a îndeplini sarcina impusă?* Două tipuri de strategii răspund, în mod diferit, la această întrebare, și anume:

- analitică;
- empirică.

---

<sup>4</sup>Eng. *Force-Closure*

<sup>5</sup>Eng. *Form-Closure*

Prehensarea obiectelor utilizându-se o strategie analitică ia în considerare aspecte legate de cinematica și dinamica elementelor implicate în proces. Fiind o metodă analitică, complexitatea calculelor este în strânsă legătură cu numărul condițiilor care trebuie satisfăcute pentru a duce la bun sfârșit operația de prindere. În esență, abordarea analitică a problemei prehensării implică determinarea pozițiilor de contact dintre obiect și prehensor și, foarte important, configurația prehensorului prin care se poate asigura îndeplinirea cu succes a sarcinii impuse [152]. Figura 6.2 ilustrează schema-bloc a unei strategii generale de prindere ce adoptă o abordare analitică.

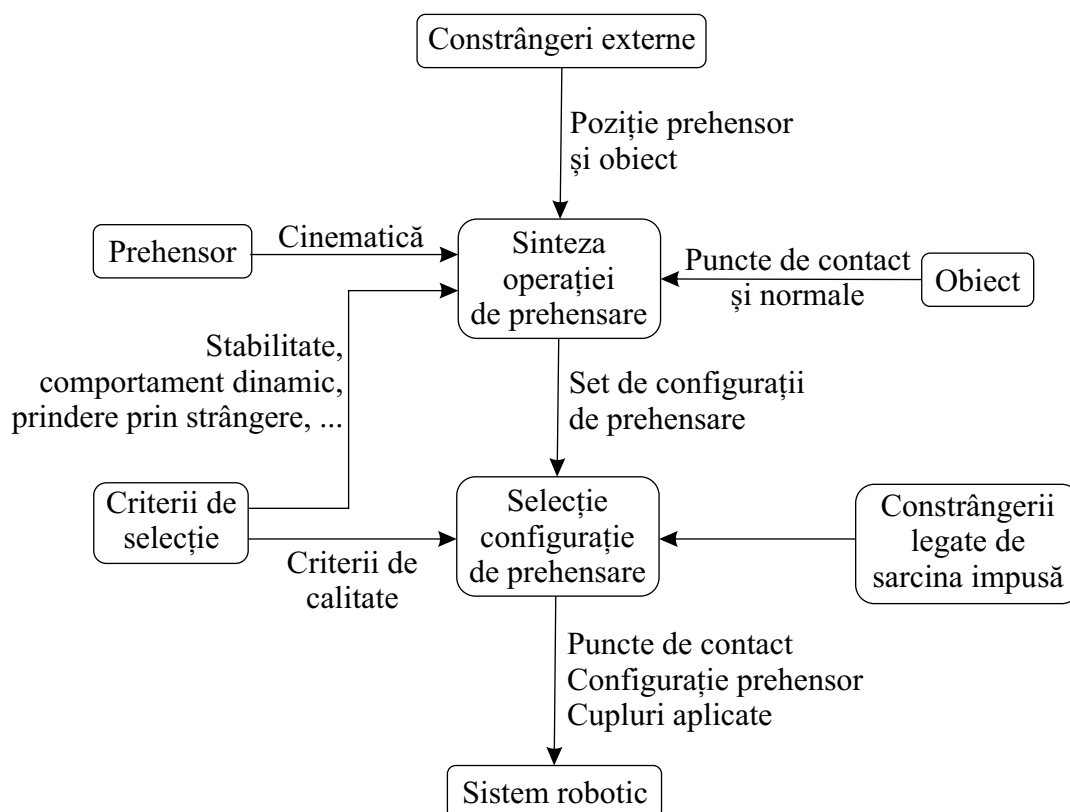


Fig. 6.2 Strategie de prehensare a obiectelor utilizând o abordare analitică [152].

Considerându-se strategia de prehensare prezentată în figura 6.2, nu toți algoritmi din literatură de specialitate conțin toate elementele prezente în schemă. Majoritatea acestora nu iau în considerare modelul prehensorului și/sau constrângerile legate de sarcină. Pe de altă parte, un element important din componența strategiei din figură este modelul obiectului. Evident, pentru a se asigura un proces de prindere și manipulare stabil, obiectul de interes trebuie să aibă o reprezentare rigidă. Primitivele generice introduse în prezenta carte reprezintă o soluție corespunzătoare pentru obținerea unor volume precise ce pot asigura un proces de manipulare stabil și eficient.

Majoritatea metodele analitice descrise în literatura de specialitate sunt centrate pe analiza unor configurații particulare de prindere ale obiectelor de interes sau pe analiza modalităților prin care acestea pot fi prinse [86, 109, 110]. Soluțiile, în acest sens, sunt cel mai dificil de obținut, mai ales când vine vorba de optimizarea procesului de prindere. Criteriile, introduse de strategia analitică, pe baza cărora are loc procesul de strângere și blocare a unui obiect, recomandă această abordare pentru operațiile de prehensare simple ce au ca principal scop transportul. Totuși, majoritatea

acțiunilor efectuate zi de zi implica mult mai mult decât simple operații de prehensare, motiv pentru care cercetătorii au demarat o serie de studii și în direcția dezvoltării unor soluții analitice pentru operații de prehensare și manipulare mai complexe (axate, de exemplu, pe sarcini) [131].

În scopul evitării complexității calculului introduse de abordarea analitică, cercetătorii și-au îndreptat atenția către o serie de soluții empirice, pentru rezolvarea problemei prehensării și manipulării obiectelor. Importanța și eficiența unei astfel de abordări este confirmată de numeroasele studii întreprinse, majoritatea foarte recente [3, 43, 53, 190].

Prehensarea obiectelor utilizându-se o strategie empirică presupune folosirea unor tehnici precum cele de clasificare și/sau învățare<sup>6</sup> care au o complexitate a calculului mai redusă. Considerându-se aceste tehnici de clasificare și învățare, strategiile de prehensare empirice pot fi împărțite în două categorii, după cum urmează:

- tehnici centrate pe observarea mișcărilor de prehensare pe care le efectuează un subiect uman [53];
- tehnici centrate pe observarea obiectului ce urmează a fi preluat. [3].

Prima tehnică mai poartă denumirea de "profesor" sau "expert". Această denumire provine de la faptul că mișcarea de prindere și manipulare a brațului robotic este învățată prin observarea unui subiect uman ce efectuează această mișcare. Practic, sistemul privește, în primă fază, mișcarea făcută de om, după care încearcă să imite aceeași mișcare ori de câte ori obiectul nou vizualizat este similar cu cel manevrat de către subiectul uman [190].

Pe de altă parte, cea de-a doua tehnică permite sistemului robotic să învețe asocierea dintre diferitele caracteristici ale obiectului de interes și o serie de tipuri de preluare pentru a calcula punctele de prehensare optime ce vor asigura un proces echilibrat și stabil [152]. Structura generică a strategiei de prehensare ce folosește metode empirice poate fi observată în figura 6.3

O serie de algoritmi empirici nu urmăresc îndeaproape strategia prezentată în figură motiv pentru care sunt limitați doar pentru un anumit domeniu de utilizare. Prin folosirea unei etape de evaluare a procesului de prehensare, se poate obține un răspuns (feedback) ce conferă "profesorului" un rol activ în cursul procesului de învățare. De fapt, sistemul de învățare poate fi dezvoltat pentru a permite celui care învață (sistemului robotic) să își îmbunătățească strategia, dincolo de ceea ce a învățat prin demonstrare. Numeroși cercetători utilizează, pentru achiziția semnalelor, o serie de mănuși biometrice ce mapează mișcarea mâinii unei persoane într-un spațiu artificial [51, 57]. O altă abordare, diferită față de aceea expusă anterior, este și aceea prin care observarea poate decurge prin intermediul tehnicilor stereoscopiei ce permite urmărirea mâinii persoanei umane (demonstratorului) și învățarea dinamicii acesteia [90].

Pentru situațiile în care este disponibil un model 3D al obiectului ce urmează a fi preluat, caz în care se poate utiliza cu foarte bune rezultate de metoda de reconstrucție volumetrică bazată pe primitive generice descrisă în prezenta carte, problema prehensării, privită din prisma metodelor empirice, poate fi o problemă de identificare a similarităților dintre forme [190]. Plecând de la ideea că obiectele similare pot fi preluate asemănător, este suficientă identificarea obiectului celui mai asemănător din baza de date în raport cu obiectul nou vizualizat, pentru a se folosi configurația

---

<sup>6</sup>Eng. *Learning*

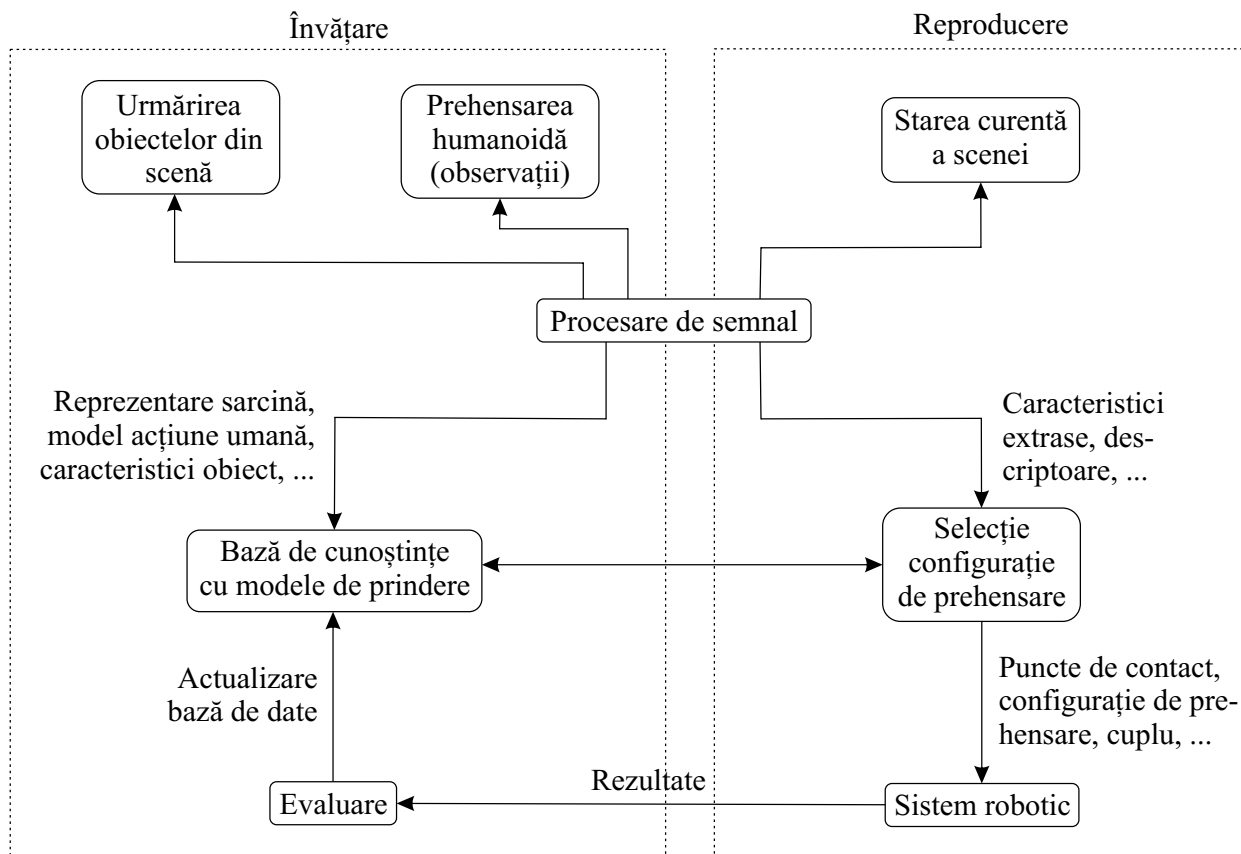


Fig. 6.3 Strategie de prehensare a obiectelor utilizând o abordare empirică [152].

acestui de prehensare.

O problemă relevantă a strategiei de prehensare bazate pe metode empirice este aceea că nu sunt autonome în momentul în care sunt aplicate asupra unor obiecte observate pentru prima dată. Utilizând abordarea expusă anterior strategia de prehensare poate fi astfel generalizată pentru a putea fi aplicată și asupra obiectelor noi vizualizate [190]. Totuși, problema este încă studiată de comunitatea științifică deoarece, și în această situație, sistemul nu poate selecta, în mod automat, configurația de prehensare optimă. În acest sens, este necesară o selecție manuală a configurației optime. Metode pentru asigurarea unui proces cu adevărat autonom în acest sens nu sunt, până în acest moment, prezente în literatura de specialitate motiv pentru care metodele empirice mai au nevoie de timp pentru a ajunge la maturitate.

Considerațiile prezentate s-au concentrat doar asupra sintezei controlului prehensorului (efectorului final), element ce intră practic în contact cu obiectul de interes. Plecând de la aceste aspecte, în continuare va fi prezentat un studiu detaliat al sintezei operației de prehensare cu aplicare asupra formelor rigide reconstruite utilizându-se metoda descrisă în capitolul 4. Poziționarea corespunzătoare a sistemului mecanic (atingerea unui punct din spațiul de lucru) și orientarea adecvată a efectorului terminal (în funcție de modalitatea de prindere a obiectului) reprezintă probleme de cinematică respectiv control al manipulatorului, probleme ce nu fac obiectul prezentului studiu motiv pentru care ele vor fi tratate ca atare. Soluțiile referitoare la aceste probleme pot fi studiate în lucrările [24, 77, 181].

## 6.2 Sinteza operației de prehensare a unui obiect

Sinteza operației de prehensare are ca principal obiectiv determinarea punctelor de contact dintre efectorul final și obiectul de interes, puncte ce asigură o manevrare echilibrată și sigură. După cum a fost prezentat și în cadrul considerațiilor expuse anterior, procesul de imobilizare a obiectului poate avea două obiective, și anume: blocarea în scopul atingerii unui grad de imobilitate (poziționarea degetelor efectorului final astfel încât prehensarea să aibă loc fără aplicarea vreunei forțe pe suprafața obiectului), respectiv strângerea obiectului de interes cu scopul efectuării unei operații de manipulare [15]. Fiecare tip de prehensare prezintă o serie de avantaje și dezavantaje. Spre exemplificare, prehensarea prin blocare se utilizează, în general, atunci când operația de prindere a obiectului nu se bazează pe forțe de frecare, în timp ce, prehensarea prin strângere se utilizează cu precădere pentru apucarea și manipularea obiectelor cu ajutorul prehensoarelor mecanice sau a mâinilor bionice.

Referitor la punctele de contact aferente procesului de prindere, în medii nestructurate, acestea sunt planificate, în mod uzual, utilizându-se o serie de măsurători vizuale care, de alt fel, sunt folosite și în procesul de reconstrucție 3D a scenei. Un stadiu actual al cercetărilor în acest sens este prezentat în lucrarea [141]. Acest studiu prezintă concis principalele metode prin care se poate controla simplu și eficient un proces de prehensare și manipulare. Totuși, în prezența unor obiecte reconstruite imperfect precum și în cazul unor erori de poziționare și orientare a brațului robotic, se poate ajunge la o serie de configurații (modalități) de prehensare greșite. Așadar, calitatea operației de reconstrucție a obiectului ce urmează a fi apucat se reflectă direct asupra echilibrului și a siguranței operației de prehensare. Cât privește partea de reconstrucție 3D, problemă ce face, de altfel, și subiectul prezentei cărți, un model volumetric precis și sigur poate fi obținut utilizându-se conceptul introdus de primitivele generice. Utilizând această abordare, pe lângă creșterea considerabilă a numărului de configurații de prehensare candidate, datorită preciziei ridicate a reconstrucției obiectului de interes, se obține și o poziționare și orientare precisă a brațului robotic. Identificarea corectă a punctelor de contact dintre prehensor și obiect poate fi formulată sub aspectul unei probleme de *panificare*. Pe de altă parte, prehensarea (prinderea) propriu-zisă, în contextul aplicării unor forțe asupra obiectului, poate fi adresată sub forma unei probleme de *control*. O serie de soluții la aceste probleme sunt prezentate în lucrarea [49]. Cele două probleme descrise anterior pot fi privite din perspectiva identificării punctelor de contact optime, respectiv din perspectiva detectării coliziunilor dintre prehensor și obiectul propriu-zis.

### 6.2.1 Planificarea procesului de prehensare a unui obiect

Datorită mediului înconjurător nestructurat în care procesele de prehensare și manipulare pot avea loc, utilizarea unui algoritm generic de planificare nu este indicată. De exemplu, nu toate obiectele din scenă pot fi manipulate, unele dintre acestea prezentând un caracter fix (rigid în raport cu mediul). Lucrurile se pot complica și mai mult în sensul că în timp, obiectele pot apărea și dispărea din scenă sau, acestea își pot schimba poziția. Considerând situațiile expuse anterior, spațiul format din totalitatea configurațiilor de prehensare posibile, a tipurilor de prehensare precum și a proprietăților obiectelor implicate în proces, este foarte redus. Găsirea unei soluții pentru

această problemă poate fi dificilă. Datorită caracterului nestructurat al scenei precum și a formei foarte variate a obiectelor ce pot fi prehensate, utilizarea în toate situațiile a unei configurații comune (predefinite), atât pentru manipulator cât și pentru prehensor, este imposibilă.

Totodată, îndeplinirea cu succes a unor operații de prehensare și manipulare depinde, într-o măsură foarte mare, de capacitatea robotului de a ajunge la obiectul de interes și de a-l observa și reprezenta corespunzător. Majoritatea metodelor de planificare utilizează, pentru aceasta, un model al obiectului. Corespondența (similaritatea) dintre acesta și obiectul real poate afecta direct rezultatul metodelor de planificare utilizate. Această condiție poate fi îndeplinită prin folosirea primitivelor generice introduse în prezenta carte. În acest fel, asemănarea dintre volumul reconstruit și cel real este foarte mare.

Date fiind toate combinațiile posibile de construcție ale brațului robotic, de formă ale obiectului, de control sau de poziție și orientare, poate exista un număr foarte restrâns de configurații ce pot conduce la o prehensare corectă a unui obiect considerat. Identificarea celei corecte poate fi un proces complicat. Această situație, privită din perspectiva unei probleme de planificare a procesului de prehensare, poate fi abordată în trei moduri diferite după cum urmează:

- căutarea în întreg spațiul a combinațiilor posibile;
- partiționarea spațiului de căutare a combinațiilor posibile;
- planificarea utilizându-se demonstrații.

Spațiul combinațiilor posibile reprezintă o colecție de modalități în care un obiect poate fi prehensat. Căutarea unei soluții în întreg spațiul combinațiilor poate fi dificilă de realizat. Pe lângă faptul că această abordare nu ia în calcul cinematica manipulatorului, metoda implică utilizarea unor ipoteze, cum ar fi, de exemplu, identificarea punctelor de contact doar pe conturul ce definește obiectul ce urmează a fi prehensat [17]. Rezultatele obținute, utilizându-se această abordare, sunt relevante și oferă în același timp o serie de puncte de contact optime relativ la obiectul considerat. Totuși, această metodă se pretează a fi utilizată în situațiile unei manipulari ce necesită îndemânare.

Pe de altă parte, în abordarea numărul doi, identificarea unei configurații de prehensare fezabile nu necesită căutarea soluției în întreg spațiul combinațiilor. Caracteristicile obiectului de interes precum și o serie de experiențe de prehensare anterioare pot fi de mare folos în procesul de identificare a configurației de prehensare optime [123]. Flexibilitatea ridicată a acestei abordări este asigurată și de faptul că astfel se poate lua în calcul atât cinematica brațului robotic cât și o serie de informații a priori despre fezabilitatea procesului de prehensare. De asemenea, sub forma unor informații a priori, pot fi definite o serie de puncte de contact pentru fiecare primitivă generică în parte.

Cea de a treia abordare a problemei de planificare implică învățarea robotului cum să apuce un obiect utilizându-se în acest scop programarea prin demonstrații. Practic, o persoană îi arată robotului cum să realizeze o anumită operație, pe care robotul trebuie să fie capabil ulterior să o repete în mod fidel. Datorită diferențelor morfologice, acest lucru nu poate fi atins cu așa mare ușurință. Procesul de învățare se bazează, în principal, pe informația vizuală percepută de robot. Informații detaliate referitoare la acest tip de planificare a procesului de prehensare pot fi găsite în lucrarea [160].

Planificarea operației de prindere a unui obiect se află în strânsă legătură cu numărul de degete

(în cazul unei mâini antropomorfe), respectiv ghiare (în cazul unui prehensor) necesar apucării acestuia. În continuare vor fi prezentate o serie de aspecte ce permit identificarea numărului optim de degete/ghiare necesar apucării diferitelor obiecte.

#### 6.2.1.1 Numărul de degete necesar prinderii unui obiect

Numărul de degete (ghiare) necesar prehensării unui obiect depinde de tipul de contact dintre obiectul propriu-zis și degete/ghiare prehensorului. Se pot defini trei tipuri de astfel de contacte, și anume: puncte de contact fără frecare<sup>7</sup> (*PCFF*), puncte de contact cu frecare<sup>8</sup> (*PCCF*, similare cu punctele de contact ale mâinii umane) și puncte de contact "slabe"<sup>9</sup> (*PCS*). Prinderea prin intermediul punctelor *PCFF* este cunoscută și ca prindere pozitivă. Pentru acest caz, poziția punctelor de contact se poate determina în două moduri după cum este descris în [127]. Algoritmii sunt liniari în timp și depind de numărul de fețe prin care este descris obiectul reconstruit (considerându-se o reprezentare poliedrică a acestuia). Un obiect ce prezintă simetrii nu poate fi prins prin strângere cu puncte de contact fără frecare. Numărul optim de puncte de contact necesar prinderii prin strângere a unui obiect poate fi urmărit în cadrul tabelului 6.1 [2].

Tab. 6.1 Numărul optim de puncte de contact necesar pentru a prinde prin strângere un obiect.

| Obiect                        | Tip obiect | <i>PCFF</i> | <i>PCCF</i> | <i>PCS</i> |
|-------------------------------|------------|-------------|-------------|------------|
| Bidimensional<br>( $d = 3$ )  | Simetric   | Imposibil   | 3           | 3          |
|                               | Nesimetric | 4           | 3           | 3          |
| Tridimensional<br>( $d = 6$ ) | Simetric   | Imposibil   | 4           | 4          |
|                               | Nesimetric | 7           | 4           | 4          |

Contactele cu sau fără frecare pot fi studiate de asemenea și din punct de vedere geometric [115]. Pentru prehensarea, cu puncte de contact de tipul *PCCF* sau *PCS*, a oricărui obiect bidimensional sau tridimensional sunt necesare 3 sau 4 puncte de prehensare. Totuși, în funcție de obiectul ce urmează a fi prehensat precum și în funcție de tipul prehensorului în sine (număr de degete/ghiare, forma, capacitate etc.) numărul de puncte de contact prezentat în tabelul 6.1 poate varia. În prezenta carte se urmărește prehensarea obiectelor utilizându-se puncte de contact de tipul *PCCF*. În continuare vor fi prezentate o serie de modalități prin care pot fi determinate aceste puncte de contact.

#### 6.2.1.2 Determinarea punctelor de contact

Se consideră zonă de contact, regiunea în care două entități sunt separate de o distanță mai mică decât un prag impus. Uzual, acest prag poate fi considerat în jurul valorii de 0.1 mm [123]. O a doua condiție importantă a zonei de contact este aceea că, cele două entități nu trebuie să se întrepătrundă deoarece, în acest caz, zona de contact va deveni zonă de coliziune.

Datorită faptului că suprafața elementelor ce intră în contact, fie că este vorba de prehensor sau de obiectul de interes, nu este deformabilă, zonele de contact pot fi descrise prin intermediul unor

<sup>7</sup>Eng. *Frictionless Point Contact*

<sup>8</sup>Eng. *Point Contact With Friction*

<sup>9</sup>Eng. *Soft Finger Contacts*

puncte de contact. Principial, această abordare este mai simplă prin prisma faptului că, astfel, nu va mai fi necesar calcularea unor arii de contact ci doar a unor simple puncte. Pot exista însă o serie de situații în acest sens. În prima, regiunea de contact dintre două corpuri poate fi similară (de ex. două suprafețe plane). În acest caz, pot fi identificate numeroase puncte de contact, dintr-o arie restrânsă, care satisfac condiția de prag (distanță mai mică de 0.1 mm). Pentru a se evita această situație, punctele de contact din această regiune sunt eliminate păstrându-le doar pe acelea care se află pe conturul regiunii de contact. Calitatea prehensării nu va fi afectată în acest fel și, în plus, timpul de calcul al punctelor de contact va fi considerabil mai redus.

O altă situație este și prehenarea obiectelor cu textură moale (slabă). Pentru a se evita o serie de calcule complexe, introduse de procesul de deformare a suprafeței de contact, se poate utiliza o abordare similară cu aceea prezentată în lucrarea [123]. Un efect relevant al acestei abordări este acela că, la punctele de contact dintr-o anumită regiune pot apărea forțe de torsiune. Figura 6.4 prezintă intuitiv punctele de contact, dintre un prehensur (model Barrett) și suprafața unui obiect (un pantof). Cu  $f$  este notată forța de contact totală, cu  $f_{\perp}$  este notată componenta normală a forței de contact  $f$ , iar cu  $\mu$  coeficientul de frecare.

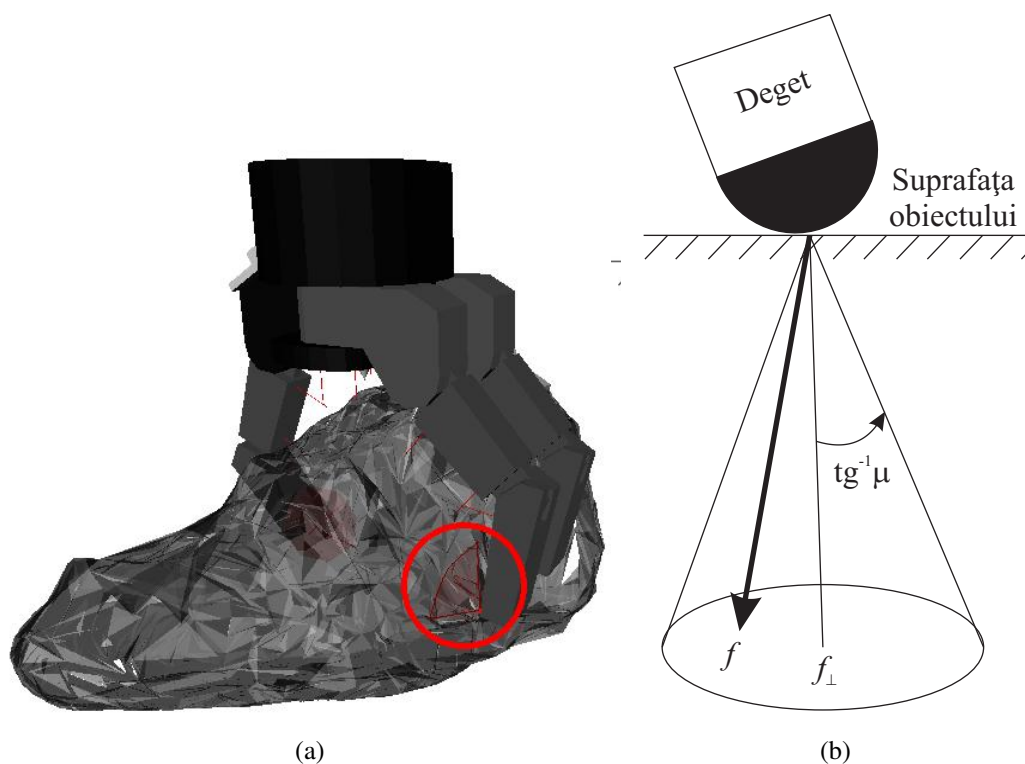


Fig. 6.4 Puncte de contact dintre prehensur și obiect: (a) vedere globală a punctelor de contact; (b) distribuția forțelor de contact pentru un anumit punct.

Conform aspectelor prezentate în figura 6.4, forțele de frecare pot avea o magnitudine maximă de  $\mu f_{\perp}$ . Pentru a se evita alunecarea ghierei prehensurului pe suprafața obiectului, forța  $f$  trebuie să se afle în interiorul conului ce descrie direcțiile fezabile ale acesteia. Acest con poartă denumirea de *con de frecare*<sup>10</sup>.

<sup>10</sup>Eng. *Friction Cone*

Problema determinării punctelor de contact dintre prehensor și obiect poate fi soluționată utilizând atât o strategie empirică, cât și una analitică.

#### *Determinarea empirică a punctelor de contact*

O rezolvare empirică a problemei identificării punctelor de contact este prezentată în [124]. În această abordare, procesul de planificare al punctelor de contact tinde să urmărească îndeaproape comportamentul uman care, înainte de a apuca un obiect, încearcă să simplifice problema prin selectarea doar a unei singure posturi (poziții și orientări) a prehensorului (mâini), în funcție de o serie de criterii legate de geometria obiectului precum și de sarcina impusă. Planificarea decurge în două etape. În prima, se generează un set de posturi inițiale ale prehensorului. Aceste posturi sunt produse având în considerare o reprezentare brută a obiectului ce urmează a fi prehensat de tipul primitivă geometrică (sferă, cub, cilindru, etc.). Fiecare astfel de primitivă geometrică este asociată cu o strategie de prindere diferită care generează puncte de contact cu probabilitate mare de a produce o apucare echilibrată. În cea de-a doua etapă, pentru fiecare postură din set, se utilizează simulatorul *GraspIt* [123] pentru a închide prehensorul și pentru a determina punctele de contact reale de pe obiect. Calitatea operației de apucare este evaluată și comparată cu aceea a celorlalte configurații obținute din setul de posturi inițial. În final, configurația cu ce prezintă calitatea cea mai bună este aleasă pentru a prehensa obiectul considerat. Utilizând abordarea descrisă anterior, în figura 6.5 este prezentată o strategie de planificare empirică a procesului de prehensare a primitivei generice a unei căni, în contextul în care aceasta a fost utilizată pentru a reconstrui obiectul perceput din scenă. În experimentul prezentat, forma brută a primitivei generice a fost echivalată cu aceea a unui paralelipiped dreptunghic. Similar, în figura 6.5(a) sunt ilustrate formele brute ale obiectului cană obținute utilizând un con, o prismă și un cilindru.

Plecând de la forma brută a obiectului cană, setul inițial de posturi este obținut prin poziționarea, de jur împrejurul obiectului, a câte unui prehensor din  $15^\circ$  în  $15^\circ$ . Pentru efectuarea testelor a fost utilizat un prehensor de tip Barrett® [170]. Sunt obținute astfel un număr de 288 de posturi inițiale. Datorită faptului că obiectul de studiu se află poziționat pe o suprafață plană, aproximativ jumătate din aceste posturi sunt nefezabile. Filtrând mai departe posturile rămase (eliminând posturilor aflate în oglindă sau cele care nu produc minim 3 puncte de contact) în final este obținut un număr de 31 posturi. Dintre acestea, 7 pot fi observate în figura 6.5(b). Utilizând simulatorul "GraspIt", posturile inițiale au fost închise dând naștere configurațiilor de prehensare prezentate în figura 6.5(c). Fiecare configurație prezentată este însoțită de o măsură a calității operației de apucare. O valoare mare a acestei măsuri indică o configurație mai stabilă. Aspectele legate de modalitatea prin care este obținută această măsură vor fi detaliate într-o secțiune viitoare. Conform acestui criteriu, configurația cu numărul 7 prezintă cea mai mare stabilitate.

Având un caracter detaliat, în figura 6.6 pot fi observate atât punctele de contact dintre prehensor și primitiva generică a căni cât și conurile de frecare aferente acestora. Totodată, cele din urmă componente oferă informații referitoare la direcția forței aplicate în fiecare punct de contact. Pentru fiecare configurație de prindere sunt definite 3 conuri de frecare. Conurile ce nu pot fi observate în figură sunt ocluzionate de corpul primitivei generice a căni.

Pentru a se forma o opinie referitoare la regiunile, din primitiva generică a căni, cu cele mai

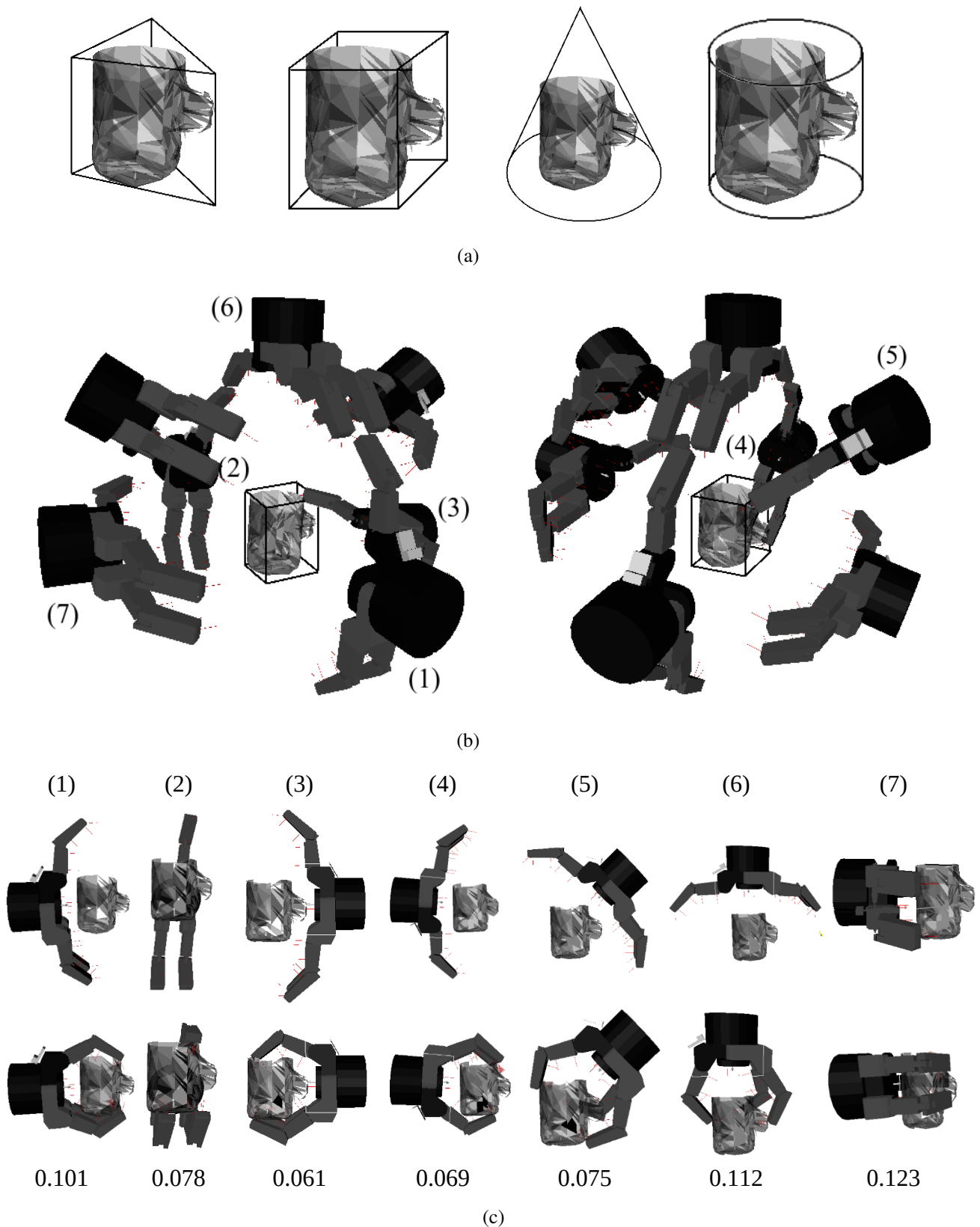


Fig. 6.5 Prehensarea primitivei generice ce ilustrează o cană utilizând o strategie de planificare empirică; (a) primitive geometrice care circumscriu optim obiectul de interes (în ordine pornind de la stânga la dreapta: prismă triunghiulară, paralelipiped dreptunghic, con și cilindru); (b) Set de posturi inițiale, văzut din două perspective diferite, pentru o primitivă geometrică de tipul paralelipiped dreptunghic ; (c) diferite configurații de prindere candidate (rândul din mijloc) a setului de posturi inițiale (rândul de sus) împreună cu măsura calității operației de prehensare (rândul de jos).

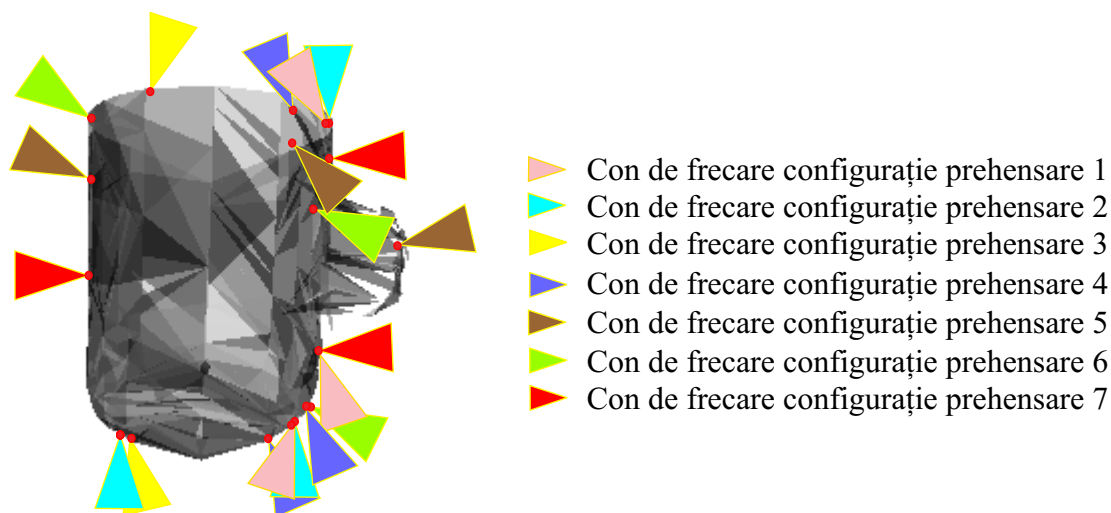


Fig. 6.6 Conuri de frecare aferente configurațiilor de prehensare obținute din posturile inițiale ale prehenzorului pentru primitiva generică a obiectului cană.

mari densități de puncte de contact, figura 6.7 prezintă toate punctele de contact stabilite în urma tuturor configurațiilor de prindere candidate (spațiul configurațiilor posibile).

Pentru a se evita generarea setului de posturi inițiale în momentul execuției scenariului, acesta poate fi generat și salvat în momentul definirii fiecărei primitive generice în parte. În timpul procesului de reconstrucție, atunci când este extrasă primitiva generică din baza de date, sistemul va beneficia implicit și de posturile inițiale ale prehenzorului. Astfel, operațiile executate "online" sunt eliminarea configurațiilor nefezabile, închiderea posturilor și evaluarea calității acestora. Procesul de evaluare al calității configurațiilor de prehensare (obținute în urma închiderii posturilor) este obligatoriu să se execute "online" datorită faptului că primitiva generică este deformată pentru fiecare scenariu în parte. Dacă acest proces nu ar fi avut loc și primitiva ar avea o formă rigidă constantă în toate scenariile, în baza de date s-ar fi putut reține și această măsură a calității operației de prehensare.

Plecând de la configurația optimă de prindere a obiectului determinată anterior, se pot genera o familie de noi configurații (identificate în vecinătatea acesteia) a căror măsură a calității prehensării este cel puțin egală cu aceea de bază (de la care s-a plecat). O astfel de abordare este prezentată în [143]. Metoda prezentată de autor permite astfel identificarea unei configurații de prehensare mai precise și echilibrate decât aceea de la care s-a plecat.

#### *Determinarea analitică a punctelor de contact*

Problema identificării punctelor de contact dintre efectorul final și obiectul de interes poate avea și o rezolvare analitică. În continuare, problema va fi tratată avându-se în vedere obiectul definit prin primitiva generică aferentă și prehenzorul cu 3 degete de tip Barrett<sup>®</sup>. Problema prehensării obiectelor tridimensionale folosind doar trei puncte de contact este tratată cu succes în [126] prin utilizarea unui criteriu de optimizare bazat pe o serie de torsoare de decuplare. Configurația optimă a celor trei puncte de contact este aleasă ca fiind aceea care rezistă cel mai bine la o serie de forțe externe de perturbare. Cele trei puncte de contact formează un plan, denumit și *plan de prindere*.

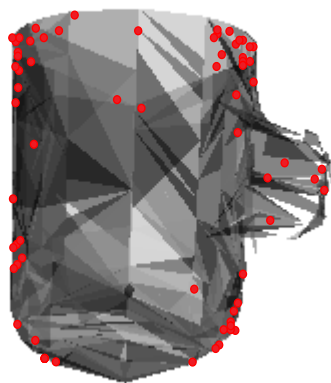


Fig. 6.7 Totalitatea punctelor de contact reale obținute prin închiderea tuturor posturilor inițiale candidate pentru primitiva generică a obiectului cană.

Totodată, dreptele care trec prin cele trei puncte și care străpung obiectul trebuie să fie concurente. Cuplurile perturbatoare sunt aplicate perpendicular pe acest plan de prindere. Utilizându-se această abordare, configurația optimă de prehensare pentru un obiect neted este determinată de prisma care circumscrie optim obiectul de interes. Considerându-se cazul prehensării unei căni, în figura 6.8 se poate observa prisma care circumscrie optim obiectul considerat.

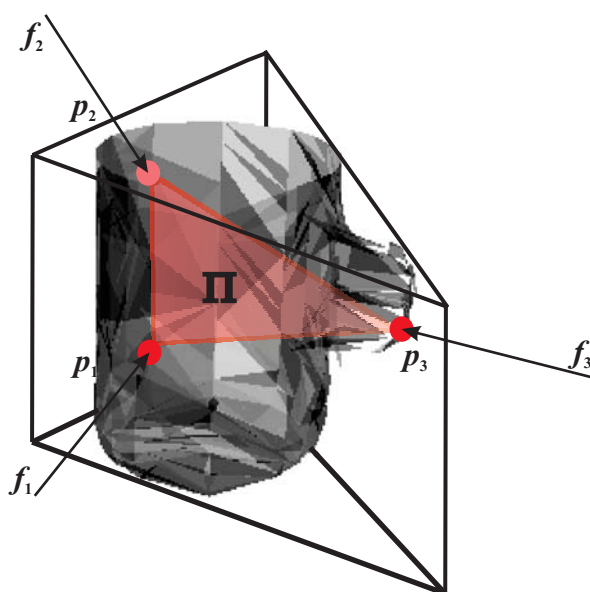


Fig. 6.8 Configurația optimă a celor trei puncte de prehensare dată de prisma care circumscrie obiectul cană.

În cazul particular al reprezentării poliedrice a obiectului de interes, punctele de contact coincid cu punctele utilizate la generarea fețelor ce descriu obiectul. Prehensarea prin strângere a unui obiect implică satisfacerea unor cerințe după cum urmează [144]:

- condiție necesară: există un punct aflat la intersecția planului format de cele trei puncte de contact cu conurile de frecare aferente acestor puncte de contact;
- condiție suficientă: există un punct aflat la intersecția celor trei conuri (în lungul direcției negative a forței) descrise de punctele de contact cu triunghiurile formate de aceste puncte. Conurile de frecare negative formate în lungul forței de presiune sunt acelea care sunt orientate

înspre interiorul obiectului de interes. Pe de altă parte, conurile pozitive sunt conurile de frecare ce sunt orientate înspre exteriorul obiectului de interes. Condiția suficientă descrisă anterior este destul de restrictivă deoarece pot exista numeroase configurații de prindere a căror puncte nu se află în interiorul conului de frecare. În acest sens, în [144] este propusă o nouă condiție suficientă, și anume:

- condiție suficientă: trei puncte,  $\mathbf{p}_1$ ,  $\mathbf{p}_2$  și  $\mathbf{p}_3$  formează o configurație de prehensare validă dacă unghiurile pereche dintre normalele interne ale acestor puncte sunt mai mici decât  $\pi - 2\theta$ . Există trei vectori ce realizează unghiuri mai mici de  $\theta/3$  cu normalele acestor puncte. Totodată, există un punct  $\mathbf{p}_0$  astfel încât pentru  $i = 0, 1, 2$  vectorul  $\mathbf{p}_0 - \mathbf{p}_i$  se află la intersecția dintre planului descris de cele trei puncte de contact (plan notat cu  $\Pi$ ) și conurile (cu două fețe, pozitivă, respectiv negativă) ale jumătății unghiului  $\theta/3$  din punctul  $\mathbf{p}_i$ .

Planul ce conține cele trei puncte de contact nu este fix. Fiecare punct de contact poate fi parametrizat prin intermediul a două variabile ( $u_i$ ,  $v_i$ ) ce definesc poziția punctului de contact în planul feței poliedrice corespunzătoare. Aceste variabile pot fi utilizate pentru a se specifica dacă  $\mathbf{p}_i$  se află sau nu pe o față convexă. Astfel, problema se poate reduce la identificarea a 9 elemente după cum urmează: 6 necunoscute ce definesc pozițiile punctelor de contact  $\mathbf{p}_i$  ( $i = 0, 1, 2$ ), respectiv 3 ce definesc poziția punctului de intersecție  $\mathbf{p}_0$ . Pentru a se simplifica problema,  $\mathbf{p}_0$  poate fi eliminat rămânând, în final, doar 6 elemente necunoscute, elemente ce plasează punctele de contact pe suprafața obiectului de interes. Aceste elemente pot fi determinate utilizând metode de programare liniară.

Sinteza operației de prehensare a obiectelor, în faza inițială de cercetare, a fost tratată doar din prisma obiectelor poliedrice ceea ce este eficient doar pentru situațiile în care numărul de fețe este redus. Totuși, pentru tratarea situațiilor în care obiectul de interes este reprezentat printr-un număr ridicat de fețe precum și pentru situațiile în care acesta este descris prin intermediul unui nor de puncte sau cel al unei rețele triangulare de puncte au fost dezvoltate o serie de metode noi [27, 56].

În [56] este prezentată o metodă euristică pentru prehensarea prin strângere cu trei puncte de contact a obiectelor reprezentate prin intermediul suprafețelor continue (nori de puncte sau rețea triangulară). Metoda propune încadrarea obiectului de interes în interiorul unei regiuni de interes rectangulare și generarea unui sistem de coordonate (cu origine arbitrară și orientare aleatoare). În continuare, sunt generate, în direcții predefinite dinspre origine către cadrul regiunii de interes, trei raze. Prima rază este emisă de-a lungul axei  $Y$ , cea de a doua, respectiv cea de a treia tot în sensul axei  $Y$ , dar rotite cu un unghi de  $120^\circ$ , respectiv  $240^\circ$ . Fiecare rază prezintă un punct de penetrare pe suprafața obiectului de interes. Aceste puncte de penetrare sunt puncte candidate pentru a fi puncte de contact dintre obiect și prehensor. Pentru obiecte non-convexe, se obțin mai mult de un punct de penetrare în lungul unei anumite direcții, motiv pentru care se alege, în mod aleator, doar un singur punct de penetrare. Procesul de căutare euristică nu ia în considerare coeficienții de frecare asociați unui punct de contact, dar configurația de prehensare respectă îndeaproape cerințele impuse prezentate anterior. Utilizând abordarea prezentată anterior, în figura 6.9 sunt ilustrate 4 astfel de configurații de prehensare. După cum se poate observa în figură, sistemul de coordonate a fost generat într-o poziție și orientare arbitrară.

În continuare, configurațiile de prehensare obținute sunt filtrate în scopul eliminării acelor

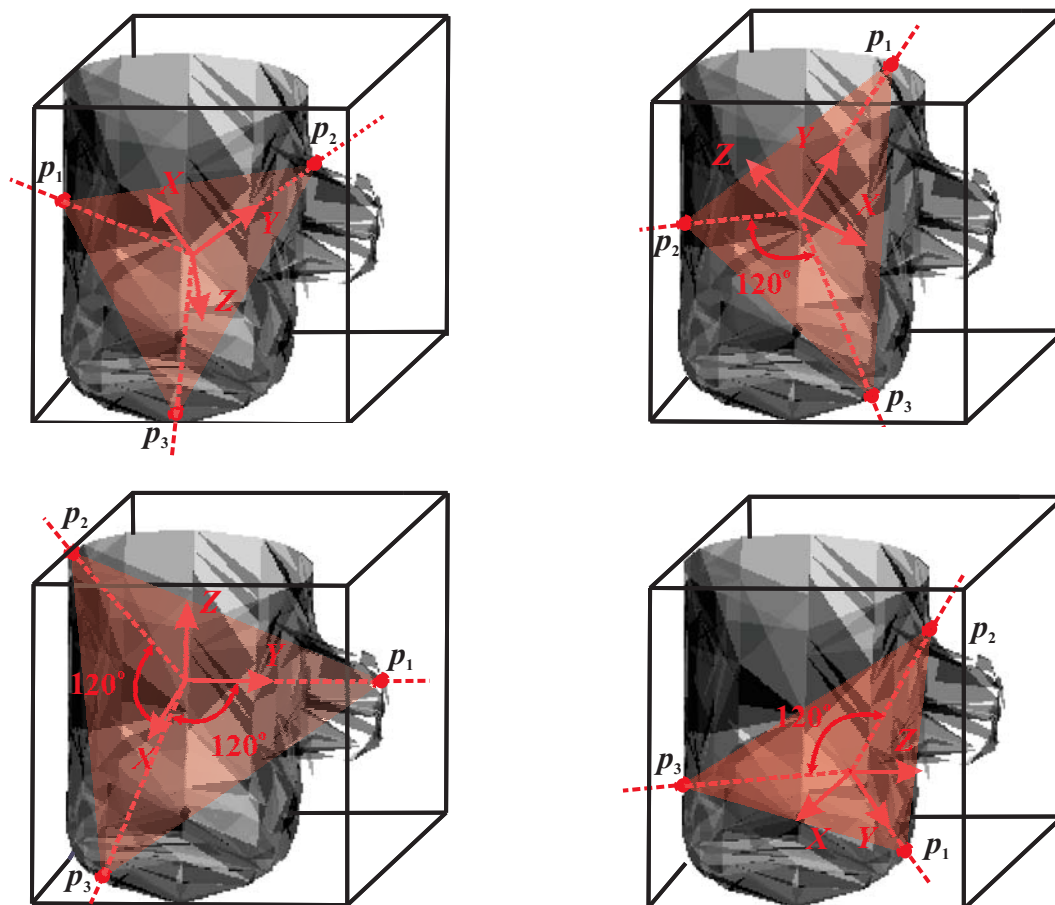


Fig. 6.9 Puncte de contact stabilite utilizându-se metoda analitică prezentată în [56].

nefezabile. Primele configurații eliminate sunt acelea ale căror puncte sunt prea apropiate de suprafața suport. Următoarele eliminate sunt acelea ale căror puncte sunt prea distanțate pentru a putea fi prehensate de către efectorul final (din punct de vedere mecanic imposibile). În final, sunt verificate configurațiile ce produc coliziuni cu efectorul final sau cu mediul. Configurațiile de prehensare, care trec de filtrele descrise, sunt ordonate în funcție de o măsură calitativă pentru ca în final, configurația cu cea mai bună măsură să fie aleasă ca și configurație optimă de prehensare a obiectului de interes. Un studiu similar, însă pentru configurații de prindere cu patru puncte de contact este prezentat în [19]. Aparatul matematic utilizat la identificarea configurației de prehensare optime este perfect similar cu acela pentru 3 puncte de contact.

### 6.2.2 Controlul operației de prehensare a unui obiect

Prehensarea unui anumit obiect presupune controlarea degetelor/ghiarelor efectorului final astfel încât acestea să aplice, pe suprafața obiectului de interes, forțe de strângere. Poziția în care acestea se aplică, magnitudinea și direcția lor vor fi discutate în detaliu în continuare.

#### 6.2.2.1 Considerații privind forțele de strângere

Se consideră un sistem de coordonate fix, localizat în centrul de greutate al obiectului reconstruit (primitiva generică). Relativ la acest sistem vor fi determinate pozițiile  $p$  ale punctelor de

contact precum și forțele aplicate în aceste puncte. O forță  $f_i$  aplicată asupra punctului  $\mathbf{p}_i$  de pe suprafața obiectului reconstruit, generează un cuplu  $\boldsymbol{\tau}_i = \mathbf{p}_i \times f_i$  (relativ la centrul de greutate al primitivei generice). Forța și cuplul pot fi grupate pentru a forma un "tossor"<sup>11</sup> (cunoscută și sub titulatura de vectorul forței generalizate). Acest tossor poate fi descris matematic prin următoarea relație:

$$\tilde{\omega}_i = \begin{pmatrix} f_i \\ \boldsymbol{\tau}_i \end{pmatrix} = \alpha_i \begin{pmatrix} \hat{n}_i \\ \mathbf{p}_i \times \hat{n}_i \end{pmatrix}, \quad (6.1)$$

unde,  $\tilde{\omega} \in \mathbb{R}^d$ ,  $\hat{n}$  este normala la suprafață din jurul punctului de contact, iar  $d$  reprezintă dimensiunea spațiului în care este definit tossorul ( $d = 3$  pentru spațiul bidimensional, respectiv  $d = 6$  pentru prehensarea obiectelor tridimensionale în spațiul fizic).

#### 6.2.2.2 Detectarea coliziunilor

O coliziune poate fi considerată ca fiind procesul prin care două entități intră în contact direct sau se pot întrepătrunde. În cazul concret al unui scenariu de prehensare, elementele ce intră în coliziune trebuie să se oprească din mișcarea lor pentru a se evita deformarea acestora. Ca efect secundar, coliziunea poate produce puncte de contact.

Dat fiind faptul că obiectul de interes este reprezentat prin intermediul unui nor dens de puncte  $\mathcal{P}$ , detectarea atât a regiunilor de coliziune cât și a punctelor de contact este aproape imposibilă, deoarece este foarte puțin probabil ca, coordonatele unui punct al obiectului să fie identice cu cele ale unui punct de pe prehensor. O rezolvare a acestei probleme constă în reprezentarea norului de puncte prin intermediul unei rețele interconectate de triunghiuri (v. sec. 2.2). Astfel reprezentat, obiectul de interes va avea o înfățișare consistentă permițând identificarea regiunilor de contact.

În general, majoritatea metodelor de detectare a coliziunilor consideră coliziunea dintre două elemente ca fiind o stare "ilegală" a acestora (configurație a articulațiilor nefezabilă). În consecință, sistemele de detectare și evitare a coliziunilor vor încerca identificarea configurațiilor manipulatorului care nu produc astfel de stări. Atunci când este detectată o posibilă coliziune, sistemul încercă interpolarea dintre configurația brațului robotic ce a produs coliziunea și ultima configurație fezabilă a acestuia, până când cele două elemente nu se vor mai ciocni.

#### 6.2.2.3 Evaluarea calității prehensării

Pentru a se putea evalua calitativ o configurație de prehensare (apucare) a unui obiect, trebuie să se stabilească proprietățile ce caracterizează o configurație corectă. Pentru o astfel de evaluare, măsura calității prehensării trebuie descrisă prin intermediul unei valori scalare. În acest sens, lucrarea [55] descrie o astfel de mărime definită de mărimea celei mai mari perturbații care poate determina punctul de contact să își piardă aderența. Cu cât această mărime este mai mare cu atât punctul de contact prezintă o aderență mai mare și implicit o prindere mai sigură. Astfel, în primă fază, procesul de evaluare a calității prehensării implică aproximarea conurilor de frecare aferente punctelor de contact (v. fig. 6.4(b)), cu suma numărului finit de forțe aflate pe mar-

<sup>11</sup>Eng. *Wrench*

ginea acestor conuri imaginare. În continuare, pentru fiecare forță de pe acest contur se determină punctul de presiune (de strângere) omolog de pe obiect. În final, se determină învelișul convex, de pe suprafața obiectului, în interiorul căruia cad aceste puncte. Presupunând că fiecare con de frecare prezintă o anumită înălțime (asociată cu magnitudinea forței de presiune necesare menținerii aderenței), învelișul exterior ce conține acest con corespunde spațiului punctelor de strângere  $L_1$  [55]. Acest spațiu reprezintă spațiul *torsoarelor de strângere*<sup>12</sup> ce poate fi aplicat pentru a prehensa obiectul considerat. Pentru a fi valabilă această afirmație, suma totală a forțelor normale de contact  $f_{\perp}$  este 1. Dacă originea forțelor normale de contact nu se află în interiorul spațiului punctelor de strângere (asociat conului de frecare), pot exista o serie de forțe de perturbație care pot destabiliza procesul de prehensare. În această din urmă situație, calitatea configurației de prehensare este 0. În caz contrar, calitatea configurației de prehensare este egală cu distanța dintre origine (punctul de presiune) și cea mai apropiată față a învelișului convex ce conține conurile de frecare. Acest comportament este similar cu cel al menținerii verticalității unui obiect. Dacă, prin înclinarea obiectului, centrul de greutate al acestuia se află în afara amprenteii, acesta se va răsturna, respectiv în cazul considerat va aluneca.

Magnitudinea forței de frecare, la care un punct de contact poate fi supus fără să își piardă aderența, influențează direct cuantumul valorii de măsurare a calității. La rândul ei, forța de frecare este dependentă, într-o mare măsură, de materialul din care este făcut învelișul exterior al obiectului (plastic, sticlă, metal etc.).

### 6.3 Concluzii

În cadrul acestui capitol a fost studiată problema prehensării și manipulării unui obiect de către o structură robotică pentru prehensare. Asigurarea unui proces de prehensare echilibrat și stabil necesită ca obiectul vizualizat (care duce la bun sfârșit sarcina impusă) să fie reprezentat sub o formă rigidă (volumetrică). În acest sens, primitivele generice introduse în capitolele precedente oferă o reprezentare rigidă precisă a obiectului vizualizat, reprezentare care poate conduce la îndeplinirea cu succes a sarcinii impuse.

În contextul unei operații de prindere, echilibrul, stabilitatea sau comportamentul dinamic al obiectului prehensat depind în exclusivitate de punctele de contact (strângere) pe care le exercită efectorul final (prehensorul) asupra obiectului manipulat. În cadrul secțiunii introductive a prezentului capitol au fost descrise două dintre cele mai uzuale strategii prin care efectorul final poate realiza o operație de prindere a unui obiect. Prima abordare, aceea analitică implică, pe de o parte, determinarea locațiilor de coliziune (contact) dintre obiect și prehensor și, pe de altă parte, configurația prehensorului prin care se poate asigura îndeplinirea cu succes a sarcinii impuse. Cea de-a doua abordare, aceea empirică, permite sistemului robotic să învețe asocierea dintre diferite caracteristici ale obiectului de interes și o serie de tipuri de prehensoare pentru a calcula punctele de prehensare optime ce vor asigura un proces echilibrat și stabil. Ambele strategii de prehensare sunt mari consumatoare de timp, motiv pentru care, aplicarea acestora în timp-real este aproape imposibilă.

<sup>12</sup>Eng. Wrench Space

O soluție posibilă la problema legată de timpul de calcul se poate obține dacă, atașat structurii de date a primitivei generice pot fi salvate și o serie de configurații de prindere (puncte de contact), astfel încât, în momentul când primitiva generică este extrasă din baza de date pentru a reconstrui un anumit obiect, robotul să beneficieze și de o serie de configurații de prindere candidate ce pot fi utilizate la prehensarea acesteia.

Plecând de la o poziționare și orientare adecvată a prehensorului, secțiunea 6.2 prezintă sinteza operației de prehensare a obiectului vizualizat. Dacă, identificarea poziției optime a punctelor de interes face subiectul unei operații de planificare, prehensarea (prinderea) propriu-zisă, în contextul aplicării unor forțe asupra obiectului, poate fi adresată sub forma unei probleme de control. Astfel, secțiunea 6.2.1 s-a concentrat asupra prezentării aspectelor legate de modalitatea de determinare a punctelor de contact, fiind prezentate aici două strategii diferite (empirică, respectiv analitică), în timp ce secțiunea 6.2.2 a avut ca principal subiect studiul forțelor ce intervin într-un astfel de proces de prehensare. Poziția, magnitudinea și direcția forțelor de strângere aplicate au fost discutate, pe larg, în cadrul aceleiași secțiuni. Aspecte legate de numărul de degete (ghiere) necesare prinderii unui obiect au fost prezentate în cadrul secțiunii 6.2.1.1. Pentru operații simple de prindere și transport, unul dintre cele mai comune efectoare finale este cel bazat pe trei degete, motiv pentru care, în carul sesiunii de implementare și testare a fost utilizat un prehensor cu trei degete denumit Barrett.

În cadrul procesului de manipulare a unui obiect, brațul robotic poate întâlni o serie de obstacole. În acest sens, un sistem pentru evitarea coliziunilor este strict necesar tocmai pentru a se evita hazardurile de acest tip. Secțiunea 6.2.2.2 a prezentat, succint, conceputul de coliziune dintre două corpuri precum și modul în care aceasta poate fi evitată.

Pentru a avea un proces controlabil precum și pentru a putea îmbunătății constant configurațiile de prindere ale diferitelor obiecte, în ultima parte a secțiunii 6.2 a fost prezentată o metodă de evaluare a calității prehensării. Pe baza acestei informații, ce poate fi folosită drept reacție, sistemul robotic poate ajusta poziția punctelor de contact, asigurând astfel o stabilitate mai ridicată a procesului de prehensare.



# Bibliografie

---

- [1] J. Ahlberg, “Active Contours in Three Dimensions,” Master’s Thesis, Linköping University, Linköping, Sweden, Sept. 1996.
- [2] M. Alejandro, R. Garzón and R. S. Feijoo, “Grasp Synthesis for 3D Objects,” Politechnical University of Catalunya, Catalunya, Spain, Tech. Rep., July 2006.
- [3] J. Aleotti and S. Caselli, “Robot Grasp Synthesis from Virtual Demonstration and Topology-Preserving Environment Reconstruction,” *Proc. of the Inter. Conf. on Intelligent Robots and Systems - IROS 2007*, San Diego, CA, USA, Nov. 2007, pp. 2692–2697.
- [4] D. Anguelov, B. Taskar, V. Chatalbashev, D. Koller, D. Gupta, G. Heitz and A. Ng, “Discriminative Learning of Markov Random Fields for Segmentation of 3D Scan Data,” *Proc. of the Conf. on Computer Vision and Pattern Recognition - CVPR 2005*, Vol. 2, San Diego, CA, USA, 2005, pp. 169–176.
- [5] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, 1st Ed. Berlin, Germany: Springer Publishing Company, Incorporated, 2010.
- [6] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman and A. Y. Wu, “An Optimal Algorithm for Approximate Nearest Neighbor Searching Fixed Dimensions,” *Journal of the Association for Computing Machinery*, Vol. 45, No. 6, Nov. 1998, pp. 891–923.
- [7] D. Baldwin, “Surface Reconstruction from Constructive Solid Geometry for Interactive Visualization,” *Proc. of the 3rd Inter. Conf. on Advances in Visual Computing - ISVC 2007*, Vol. 1, No. 1. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 321–330.
- [8] E. Bardinet, L. D. Cohen and N. Ayache, “A Parametric Deformable Model to Fit Unstructured 3D Data,” *Computer Vision and Image Understanding*, Vol. 71, No. 1, 1995, pp. 39–54.
- [9] A. Barr, “Superquadrics and Angle-Preserving Transformations,” *IEEE Computer Graphics and Applications*, Vol. 1, No. 1, 1981, pp. 11–23.

- [10] J. L. Bentley, "Multidimensional Binary Search Trees Used for Associative Searching," *Communications of the Association for Computing Machinery*, Vol. 18, No. 9, Sept. 1975, pp. 509–517.
- [11] J. Berkmann and T. Caelli, "Computation of Surface Geometry and Segmentation Using Covariance Techniques," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 11, 1994, pp. 1114–1116.
- [12] P. J. Besl and N. D. McKay, "A Method for Registration of 3D Shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, Feb. 1992, pp. 239–256.
- [13] T. Bhabhrawala, *Shape Recovery from Medical Image Data Using Extended Superquadrics*, 1st Ed. New York, USA: State University of New York at Buffalo, 2005.
- [14] A. Bhattacharyya, "On a Measure of Divergence Between Two Statistical Populations Defined by Their Probability Distributions," *Bulletin of the Calcutta Mathematical Society*, Vol. 35, 1943, pp. 99–109.
- [15] A. Bicchi and V. Kumar, "Robotic Grasping and Contact: A Review," *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 2000*, Vol. 1, San Francisco, CA, USA, April 2000, pp. 348–353.
- [16] A. Bierbaum, I. Gubarev and R. Dillmann, "Robust Shape Recovery for Sparse Contact Location and Normal Data from Haptic Exploration," *Proc. of the Inter. Conf. on Intelligent Robots and Systems – IROS 2008*, Nice, France, Sept. 2008, pp. 3200–3205.
- [17] G. M. Bone, A. Lambert and M. Edwards, "Automated Modeling and Robotic Grasping of Unknown Three-Dimensional Objects," *Proc. of the 2008 IEEE Inter. Conf. on Robotics and Automation*, Pasadena, CA, USA, 2008, pp. 292–298.
- [18] I. Borg, *Procrustes Analysis*. John Wiley and Sons, Ltd, 2005, Ch. 5, pp. 337–357.
- [19] C. Borst, M. Fischer and G. Hirzinger, "A Fast and Robust Grasp Planner for Arbitrary 3D Objects," *Proc. of the Inter. Conf. on Robotics and Automation – ICRA 1999*, Vol. 3, Atlanta, Georgia, USA, May 1999, pp. 1890–1896.
- [20] A. Boulch and R. Marlet, "Fast and Robust Normal Estimation for Point Clouds with Sharp Features," *Computer Graphics Forum*, Vol. 31, No. 5, 2012, pp. 1765–1774.
- [21] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. Sebastopol, Canada: O'Reilly Media, 2008.
- [22] M. Bronielsen and S. Cotin, "Real-time Volumetric Deformable Models for Surgery Simulation using Finite Elements and Condensation," *Computer Graphics Forum*, Vol. 15, No. 3, 1996, pp. 57–66.
- [23] P. Brunet and I. Navazo, "Solid Representation and Operation Using Extended Octrees," *ACM Trans. on Graphics*, Vol. 9, No. 2, April 1990, pp. 170–197.
- [24] S. R. Buss, "Introduction to Inverse Kinematics with Jacobian Transpose, Pseudoinverse and Damped Least Squares Methods," *IEEE Journal of Robotics and Automation*, Tech. Rep., 2004.
- [25] T. Charlton, *Energy Principles in Theory of Structures*. UK: Oxford University Press, 1973.
- [26] H. H. Chen and T. S. Huang, "A Survey of Construction and Manipulation of Octrees,"

- Computer Vision Graphics and Image Processing*, Vol. 43, No. 3, Sept. 1988, pp. 409–431.
- [27] I.-M. Chen and J. Burdick, “Finding Antipodal Point Grasps on Irregularly Shaped Objects,” *IEEE Trans. on Robotics and Automation*, Vol. 9, No. 4, 1993, pp. 507–512.
- [28] K.-M. G. Cheung, S. Baker and T. Kanade, “Shape-from-Silhouette Across Time Part I: Theory and Algorithms,” *Inter. Journal on Computer Vision*, Vol. 62, No. 3, May 2005, pp. 221–247.
- [29] T. T. Cocias, S. M. Grigorescu and F. Moldoveanu, “On Performance Evaluation of 3D Scene Reconstruction Systems,” *Proc. of the 20th Inter. Workshop on Robotics in Alpe-Adria-Danube Region – MM Science Journal (Special Edition) - RAAD 2011*, Brno, Czech Republic, 2011, pp. 55–60.
- [30] T. T. Cocias, S. M. Grigorescu and F. Moldoveanu, “Multiple-Superquadrics Based Object Surface Estimation for Grasping in Service Robotics,” *Proc. of the 13th Inter. Conf. on Optimization of Electrical and Electronic Equipment - OPTIM 2012*, Brasov, Romania, 2012, pp. 1471–1477.
- [31] T. T. Cocias, S. M. Grigorescu and F. Moldoveanu, “Object Volumetric Estimation Based on Generic Fitted Primitives for Service Robotics,” *Proc. of the Inter. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications – VISIGRAPP 2012*, Vol. 2, Rome, Italy, 2012, pp. 191–197.
- [32] T. T. Cocias, S. M. Grigorescu and F. Moldoveanu, “3D Structure Estimation from a Single View Using Generic Fitted Primitives (GFP),” *Computer Vision, Imaging and Computer Graphics. Theory and Application*, ser. Communications in Computer and Information Science. Springer Berlin, Heidelberg, 2013, Vol. 359, No. 2, pp. 369–382.
- [33] T. T. Cocias, S. M. Grigorescu and F. Moldoveanu, “Generic Fitted Shapes (GFS): Volumetric Object Segmentation in Service Robotics,” *Robotics and Autonomous Systems*, Vol. 61, No. 9, Sept. 2013, pp. 960–972.
- [34] L. D. Cohen, “On Active Contour Models and Balloons,” *Computer Vision Graphics and Image Processing - CVGIP 1991*, Vol. 53, No. 2, March 1991, pp. 211–218.
- [35] D. Cojocaru, *Achiziția, Prelucrarea și Recunoașterea Imaginilor*. Craiova, Romania: Ed. Universitaria, 2002.
- [36] T. F. Cootes, C. J. Taylor, D. Cooper and J. Graham, “Active Shape Models: Their Training and Application,” *Computer Vision and Image Understanding*, Vol. 61, No. 1, 1995, pp. 38–59.
- [37] T. F. Cootes, A. Hill, C. J. Taylor and J. Haslam, “The Use of Active Shape Models for Locating Structures in Medical Images,” *Proc. of the 13th Inter. Conf. on Information Processing in Medical Imaging - IPMI 1993*, London, UK, 1993, pp. 33–47.
- [38] P. Corke, *Visual Control of Robots : High-performance Visual Servoing*. Taunton, England: Research Studies Press Ltd, 1996.
- [39] J. Corney, H. Rea, D. Clark, J. Pritchard, M. Breaks and R. Macleod, “Coarse Filters for Shape Matching,” *IEEE Computer Graphics and Applications*, Vol. 22, No. 3, 2002, pp. 65–74.
- [40] C. M. Cyr and B. B. Kimia, “A Similarity-Based Aspect-Graph Approach to 3D Object

- Recognition,” *Inter. Journal of Computer Vision*, Vol. 57, No. 1, April 2004, pp. 5–22.
- [41] S. P. Dandamudi and P. G. Sorenson, “Algorithms for BD Trees,” *Software: Practice and Experience*, Vol. 16, No. 12, 1986, pp. 1077–1096.
  - [42] C. Davatzikos and J. L. Prince, “An Active Contour Model for Mapping the Cortex,” *IEEE Trans. on Medical Imaging*, Vol. 14, 1995, pp. 65–80.
  - [43] J. Demiris and G. M. Hayes, “Imitation in Animals and Artifacts.” Cambridge, MA, USA: MIT Press, 2002, Ch. Imitation as a Dual-Route Process Featuring Predictive and Learning Components: A Biologically Plausible Computational Model, pp. 327–361.
  - [44] I. Dryden and K. Mardia, *Statistical Shape Analysis*, ser. Probability and Statistics. Chichester, UK: Wiley, 1998.
  - [45] B. Duffy, “Fundamental Issues in Affective Intelligent Social Machines,” *The Open Artificial Intelligence Journal*, Vol. 2, No. 14, 2008, pp. 21–34.
  - [46] B. Dumitrescu, C. Popeea and B. Jora, *Metode de Calcul Numeric Matriceal. Algoritmi Fundamentali*. Bucuresti: All, 1998.
  - [47] R. Durikovic, K. Kaneda and H. Yamashita, “Reconstructing a 3D Structure with Multiple Deformable Solid Primitives,” *Computer Graphics Forum*, Vol. 21, No. 5, 1997, pp. 611–624.
  - [48] H. Eberhardt, V. Klumpp and U. Hanebeck, “Density Trees for Efficient Nonlinear State Estimation,” *Proc. of the 13th Conf. on Information Fusion*, Edinburgh, UK, 2010, pp. 1–8.
  - [49] A. Edsinger, “Robot Manipulation in Human Environments,” Ph.D. Thesis, Massachusetts Inst. of Technology, USA, 2007.
  - [50] P. Eisert, *Reconstruction of Volumetric 3D Models*. John Wiley and Sons, Ltd, 2006, Ch. 8, pp. 133–150.
  - [51] S. Ekvall and D. Kragic, “Interactive Grasp Learning Based on Human Demonstration,” *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 2004*, Vol. 4, New Orelans, LA, USA, May 2004, pp. 3519–3524.
  - [52] A. El-Baz, M. Nitzken, A. Elnakib, F. Khalifa, G. Gimel’farb, R. Falk and M. A. El-Ghar, “3D Shape Analysis for Early Diagnosis of Malignant Lung Nodules,” *Proc. of the 14th Inter. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, Vol. 3, No. 1. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 175–182.
  - [53] S. El-Khoury and A. Sahbani, “A New Strategy Combining Empirical and Analytical Approaches for Grasping Unknown 3D Objects,” *Robotics and Autonomous Systems*, Vol. 58, No. 5, May 2010, pp. 497–507.
  - [54] J. E. Fenwick, “A Survey of Meshing Algorithms and a Multi-Purpose Mesh Generator,” Ph.D. Thesis, University of California at San Diego, La Jolla, CA, USA, 2005.
  - [55] C. Ferrari and J. Canny, “Planning Optimal Grasps,” *Proc. of the 1992 IEEE Inter. Conf. on Robotics and Automation - ICRA 1992*, Nice, France, May 1992, pp. 2290–2295.
  - [56] M. Fischer and G. Hirzinger, “Fast Planning of Precision Grasps for 3D Objects,” *Proc. of the Inter. Conf. on Intelligent Robots and Systems – IROS 1997*, Vol. 1, Osaka, Japan, Nov. 1997, pp. 120–126.
  - [57] M. Fischer, P. van der Smagt and G. Hirzinger, “Learning Techniques in a Dataglove Based

- Telemanipulation System for the DLR Hand,” *Proc. of the Inter. Conf on Robotics and Automation - ICRA 1998*, Vol. 2, Leuven, Belgium, May 1998, pp. 1603–1608.
- [58] E. Freund and J. Rossmann, “Intelligent Autonomous Robots for Industrial and Space Applications,” *Proc. of the IEEE/RSJ/GI Inter. Conf. on Intelligent Robots and Systems - IROS 1994*, ser. Advanced Robotic Systems and the Real World, Vol. 2, Munich, Germany, Sept. 1994, pp. 1072–1079.
- [59] Y. Freund and R. E. Schapire, “Experiments with a New Boosting Algorithm,” *Proc. of the Thirteenth Inter. Conf. on Machine Learning*, San Francisco, USA, 1996, pp. 148–156.
- [60] H. Fukai and G. Xu, “Fast and Robust Registration of Multiple 3D Point Clouds,” *Proc. of the 20th IEEE Inter. Symp. on Robot and Human Interactive Communication*, Atlanta, Georgia, USA, 2011, pp. 331–336.
- [61] I. Gaiser, S. Schulz, A. Kargov, H. Klosek, A. Bierbaum, C. Pylatiuk, R. Oberle, T. Werner, T. Asfour, G. Bretthauer and R. Dillmann, “A New Anthropomorphic Robotic Hand,” *Proc. of the 8th Inter. Conf. on Humanoid Robots*, Daejeon, South Korea, Dec. 2008, pp. 418–422.
- [62] J.-P. Gazeau, S. Zehloul, M. Arsicault and J. P. Lallemand, “The LMS Hand: Force and Position Controls in the Aim of the Fine Manipulation of Objects,” *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 2001*, Vol. 3, No. 1, Shanghai, China, May 2001, pp. 2642–2648.
- [63] D. Gil and P. Radeva, “Curvature Vector Flow to Assure Convergent Deformable Models for Shape Modelling,” *Proc. of the 9th Inter. Conf. on Energy Minimization Methods in Computer Vision and Pattern Recognition - EMMCVPR 2003*, Lund, Sweden, 2003, pp. 357–372.
- [64] J. Glover, D. Rus and N. Roy, “Probabilistic Models of Object Geometry with Application to Grasping,” *Inter. Journal of Robotics Research*, Vol. 28, No. 8, Aug. 2009, pp. 999–1019.
- [65] C. Goldfeder, “Data-Driven Grasping,” Ph.D. Thesis, Columbia University, New York, NY, USA, 2010.
- [66] A. Gomes, I. Voiculescu, J. Jorge, B. Wyvill and C. Galbraith, *Implicit Curves and Surfaces: Mathematics, Data Structures and Algorithms*, 1st Ed. New York, USA: Springer Publishing Company, Incorporated, 2009.
- [67] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd Ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2006.
- [68] L. C. Goron, Z. C. Marton, G. Lazea and M. Beetz, “Segmenting Cylindrical and Box-like Objects in Cluttered 3D Scenes,” *Proc. of the 7th German Conf. on Robotics - Robotik*, Munich, Germany, May 2012, pp. 1–6.
- [69] J. Gower, “Generalized Procrustes Analysis,” *Psychometrika*, Vol. 40, No. 1, March 1975, pp. 33–51.
- [70] J. C. Gower and G. B. Dijksterhuis, *Procrustes Problems*, ser. Oxford Statistical Science. Oxford, UK: Oxford University Press, Jan. 2004, Vol. 70, No. 4.
- [71] F. Gramazio and M. Kohler, “Towards a Digital Materiality,” *Proc. of the Manufacturing Material Effects: Rethinking Design and Making Architecture*, New York, USA, 2008, pp. 103–118.

- [72] S. M. Grigorescu, "Robust Machine Vision for Service Robotics," Ph.D. Thesis, Faculty of Physics and Electrical Engineering, University of Bremen, Bremen, Germany, 2009.
- [73] S. M. Grigorescu, T. T. Cocias, G. Macesanu and F. Moldoveanu, "Stereo Vision-based 3D Camera Pose and Object Structure Estimation - An Application to Service Robotics," *Proc. of the Inter. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications - VISIGRAPP 2012*, Vol. 2, Rome, Italy, 2012, pp. 355–358.
- [74] S. M. Grigorescu, G. Macesanu, T. T. Cocias and F. Moldoveanu, "On the Real-time Modelling of a Robotic Scene Perception and Estimation System," *Proc. of the 15th Inter. Conf. on System Theory, Control, and Computing*, Sinaia, Romania, 2011, pp. 273–276.
- [75] S. M. Grigorescu, G. Macesanu, T. T. Cocias, D. Puiu and F. Moldoveanu, "Robust Camera Pose and Scene Structure Analysis for Service Robotics," *Robotics and Autonomous Systems*, Vol. 59, No. 11, 2011, pp. 899–909.
- [76] U. Hagn, M. Nickl, S. Jörg, A. Tobergte, B. Kübler, G. Passig, M. Gröger, F. Fröhlich, U. Seibold, R. Konietzke, L. Le-Tien, A. Albu-Schäffer, M. Grebenstein, T. Ortmaier and G. Hirzinger, "DLR MiroSurge - Towards Versatility in Surgical Robotics," *Proc. of the 7 Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie e.V.*, Vol. 1, Leipzig, Germany, Sep. 2008, pp. 143–146.
- [77] R. S. Hartenberg and J. Denavit, *Kinematic Synthesis of Linkages*, ser. Mechanical Engineering. New York: McGraw-Hill, 1965.
- [78] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. New York, NY, USA: Cambridge University Press, 2000.
- [79] P. Heider, A. Pierre-Pierre, R. Li and C. Grimm, "Local Shape Descriptors, A Survey and Evaluation," *Proc. of the 4th Eurographics Conf. on 3D Object Retrieval - EG 3DOR 2011*. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2011, pp. 49–56.
- [80] H. Helbig, M. Graf and M. Kiefer, "The Role of Action Representations in Visual Object Recognition," *Experimental Brain Research*, Vol. 174, No. 2, 2006, pp. 221–228.
- [81] G. Hetzel, B. Leibe, P. Levi and B. Schiele, "3D Object Recognition from Range Images Using Local Feature Histograms," *Proc. of the 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition - CVPR 2001*, Vol. 2, Kauai, HI, USA, 2001, pp. 394–399.
- [82] J. Hollingum, "Robots in Agriculture," *Industrial Robot: An Inter. Journal*, Vol. 26, No. 6, 1999, pp. 438–446.
- [83] D. Holz, S. Holzer and R. B. Rusu, "Real-Time Plane Segmentation Using RGB-D Cameras," *Proc. of the 15th RoboCup Symp.*, Istanbul, Turkey, 2011, pp. 306–317.
- [84] H. Hoppe, "Surface Reconstruction from Unorganized Points," Ph.D. Thesis, Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA, 1995.
- [85] B. K. P. Horn, "Closed-Form Solution of Absolute Orientation Using Unit Quaternions," *Journal of the Optical Society of America*, Vol. 4, No. 4, 1987, pp. 629–642.
- [86] W. Howard and V. Kumar, "On the Stability of Grasped Objects," *IEEE Trans. on Robotics and Automation*, Vol. 12, No. 6, 1996, pp. 904–917.
- [87] S. Hrabar, "Reactive Obstacle Avoidance for Rotorcraft UAVs," *Proc. of the IEEE/RSJ Inter.*

- Conf. on Intelligent Robots and Systems - IROS 2011*, San Francisco, USA, 2011, pp. 4967–4974.
- [88] K. Huebner, K. Welke, M. Przybylski, N. Vahrenkamp, T. Asfour, D. Kragic and R. Dillmann, “Grasping Known Objects with Humanoid Robots: A Box-Based Approach.”
- [89] J. Huerta, M. Chover, R. Quiros, R. Vivo and J. Ribelles, “Binary Space Partitioning Trees: A Multiresolution Approach,” *Proc. of the IEEE Conf. on Information Visualisation - IV 1997*. Washington, DC, USA: IEEE Computer Society, 1997, pp. 148–156.
- [90] M. Hueser, T. Baier and J. Zhang, “Learning of Demonstrated Grasping Skills by Stereoscopic Tracking of Human Head Configuration,” *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 2006*, Orlando, Florida, USA, May 2006, pp. 2795–2800.
- [91] A. Huitson, “Problems with Procrustes Analysis,” *Journal of Applied Statistics*, Vol. 16, No. 1, 1989, pp. 39–45.
- [92] O. Ivlev, C. Martens and A. Graeser, “Rehabilitation Robots FRIEND-I and FRIEND-II with the Dexterous Lightweight Manipulator,” *Technology and Disability*, Vol. 17, No. 2, 2005, pp. 111–123.
- [93] A. Jaklic, A. Leonardis and F. Solina, *Segmentation and Recovery of Superquadrics: Computational Imaging and Vision*. Norwell, MA, USA: Kluwer Academic Publishers, 2000.
- [94] A. E. Johnson and M. Hebert, “Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 21, No. 5, 1999, pp. 433–449.
- [95] T. Joshi, N. Ahuja and J. Ponce, “Structure and Motion Estimation from Dynamic Silhouettes under Perspective Projection,” *Inter. Journal of Computer Vision*, Vol. 31, No. 1, 1999, pp. 31–50.
- [96] A.-L. Jousselme and P. Maupin, “Distances in Evidence Theory: Comprehensive Survey and Generalizations,” *Inter. Journal of Approximate Reasoning*, Vol. 53, No. 2, 2012, pp. 118–145.
- [97] W. Kabsch, “A Solution for the Best Rotation to Relate Two Sets of Vectors,” *Acta Crystallographica Section*, Vol. 32, No. 5, 1976, pp. 922–923.
- [98] T. Kapur, “Segmentation of Brain Tissue from Magnetic Resonance Images,” Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA, Tech. Rep., 1995.
- [99] M. Kass, A. Witkin and D. Terzopoulos, “Snakes: Active Contour Models,” *Inter. Journal of Computer Vision*, Vol. 1, No. 4, 1988, pp. 321–331.
- [100] D. Katsoulas, C. C. Bastidas and D. Kosmopoulos, “Superquadric Segmentation in Range Images via Fusion of Region and Boundary Information,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 5, 2008, pp. 781–795.
- [101] D. Katsoulas and A. Jaklic, “Fast Recovery of Piled Deformable Objects Using Superquadrics,” *Proc. of the 24th DAGM Symposium on Pattern Recognition*, London, UK, 2002, pp. 174–181.
- [102] M. Kazhdan, “Shape Representations and Algorithms for 3D Model Retrieval,” Ph.D. Thesis, Princeton University, USA, June 2004.

- [103] D. Keren, D. Cooper and J. Subrahmonia, "Describing Complicated Objects by Implicit Polynomials," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 1, Jan. 1994, pp. 38–53.
- [104] K. Klasing, D. Althoff, D. Wollherr and M. Buss, "Comparison of Surface Normal Estimation Methods for Range Sensing Applications," *Proc. of the 2009 IEEE Inter. Conf. on Robotics and Automation - ICRA 2009*, Piscataway, USA, 2009, pp. 1977–1982.
- [105] R. Krishnaswamy, G. S. Alijani and S.-C. Su, "On Constructing Binary Space Partitioning Trees," *Proc. of the 1990 ACM Annual Conf. on Cooperation - CSC 1990*, New York, USA, 1990, pp. 230–235.
- [106] M. Krueger, P. Delmas and G. Gimel'Farb, "Active Contour Based Segmentation of 3D Surfaces," *Proc. of the 10th European Conf. on Computer Vision - ECCV 2008*, No. 2. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 350–363.
- [107] H. O. Lancaster, *The Chi-Squared Distribution*, ser. Probability and Mathematical Statistics. New York, USA: J. Wiley, 1969.
- [108] A. Leonardis, A. Jaklic and F. Solina, "Superquadrics for Segmenting and Modeling Range Data," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 11, Nov. 1997, pp. 1289–1295.
- [109] J.-W. Li, H. Liu and H.-G. Cai, "On Computing Three-Finger Force-Closure Grasps of 2-D and 3-D Objects," *IEEE Trans. on Robotics and Automation*, Vol. 19, No. 1, 2003, pp. 155–161.
- [110] Y.-H. Liu, M. ling Lam and D. Ding, "A Complete and Efficient Algorithm for Searching 3-D Form-closure Grasps in the Discrete Domain," *IEEE Trans. on Robotics*, Vol. 20, No. 5, 2004, pp. 805–816.
- [111] H. C. Longuet, "A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature*, Vol. 293, 1981, pp. 133–135.
- [112] D. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, 3rd Ed. New York, USA: Springer, 2008.
- [113] Q. Luong and O. Faugeras, "The Fundamental Matrix: Theory, Algorithms, and Stability Analysis," *Inter. Journal of Computer Vision*, Vol. 17, 1995, pp. 43–75.
- [114] G. Macesanu, T. T. Cocias, F. Moldoveanu and M. Cernat, "Active Vision System for 3D Scene Reconstruction," *Proc. of the Inter. Conf. on Interdisciplinarity in Education*, Athens, Greece, 2011, pp. 376–381.
- [115] X. Markenscoff, L. Ni and C. Papadimitriou, "The Geometry of Grasping," *Inter. Journal on Robotics Research*, Vol. 9, No. 1, Feb. 1990, pp. 61–74.
- [116] C. Martens, O. Prenzel and A. Graeser, "The Rehabilitation Robots FRIEND-I & II: Daily Life Independency Through Semi-Autonomous Task-Execution," *Rehabilitation Robotics*, Vol. 17, No. 1, Aug. 2007, pp. 137–162.
- [117] C. Martens, N. Ruchel, O. Lang, O. Ivlev and A. Graser, "A FRIEND for Assisting Handicapped People," *IEEE Robotics Automation Magazine*, Vol. 8, No. 1, 2001, pp. 57–65.
- [118] Z. C. Marton, L. Goron, R. B. Rusu and M. Beetz, "Reconstruction and Verification of 3D Object Models for Grasping," *Proc. of the 14th Inter. Symp. on Robotics Research - ISRR*

- 2009, Lucerne, Switzerland, Sept. 2009, pp. 315–328.
- [119] Z. C. Marton, D. Pangercic, N. Blodow, J. Kleinhellefort and M. Beetz, “General 3D Modelling of Novel Objects from a Single View,” *Proc. of the 2010 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems - IROS 2010*, Taipei, Taiwan, Oct. 2010, pp. 3700–3705.
- [120] T. Matsuyama, X. Wu, T. Takai and T. Wada, “Real-Time Dynamic 3D Object Shape Reconstruction and High-Fidelity Texture Mapping for 3D Video,” *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 14, No. 3, 2004, pp. 357–369.
- [121] I. Matthews and S. Baker, “Active Appearance Models Revisited,” *Inter. Journal on Comput. Vision*, Vol. 60, No. 2, Nov. 2004, pp. 135–164.
- [122] C. H. Messom and A. L. Barczak, “Stream Processing for Fast and Efficient Rotated Haar-like Features Using Rotated Integral Images,” *Inter. Journal of Intelligent Systems Technologies and Applications*, Vol. 7, No. 1, May 2009, pp. 40–57.
- [123] A. Miller, *GraspIt!: A Versatile Simulator for Robotic Grasping*. New York, USA: Columbia University, 2001.
- [124] A. T. Miller, S. Knoop, H. I. Christensen and P. K. Allen, “Automatic Grasp Planning Using Shape Primitives,” *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 2003*, Taipei, Taiwan, Sept. 2003, pp. 1824–1829.
- [125] P. Mills, “Efficient Statistical Classification of Satellite Measurements,” *Inter. Journal of Remote Sensing*, Vol. 32, No. 21, 2012, pp. 6109–6132.
- [126] B. Mirtich and J. Canny, “Easily Computable Optimum Grasps in 2D and 3D,” *Proc. of the Inter. Conf. on Robotics and Automation - ICRA 1994*, San Diego, CA, USA, May 1994, pp. 739–747.
- [127] B. Mishra, J. Schwartz and M. Sharir, “On the Existence and Synthesis of Multifinger Positive Grips,” *Algorithmica*, Vol. 2, No. 4, Nov. 1987, pp. 541–558.
- [128] A. Morales, T. Asfour, P. Azad, S. Knoop and R. Dillmann, “Integrated Grasp Planning and Visual Object Localization for a Humanoid Robot with Five-Fingered Hands,” *Proc. of the Inter. Conf. on Intelligent Robots and Systems*, Beijing, China, Oct. 2006, pp. 5663–5668.
- [129] A. Morales, P. J. Sanz, A. P. D. Pobil and A. H. Fagg, “Vision-Based Three-Finger Grasp Synthesis Constrained by Hand Geometry,” *Robotics and Autonomous Systems*, Vol. 54, No. 6, 2006, pp. 496–512.
- [130] M. Muja and D. G. Lowe, “Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration,” *Proc. of the Inter. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications - VISAPP 2009*, Lisboa, Portugal, 2009, pp. 331–340.
- [131] A. Nakashima, T. Uno, Y. Hayakawa, T. Kondo, S. Sawada and N. Nanba, “Synthesis of Stable Grasp by Four-Fingered Robot Hand for Pick-And-Place of Assembling Parts,” *Proc. of the Inter. 5th IFAC Symp. on Mechatronic Systems*, Munich, Germany, July 2010, pp. 669–676.
- [132] S. Natarajan, D. Ristic-Durrant, A. Leu and A. Graser, “Robust Stereo-Vision Based 3D Modelling of Real-World Objects for Assistive Robotic Applications,” *Proc. of the 2011 IEEE/RSJ Inter. Conf. on Intelligent Robots and Systems - IROS 2011*, San Francisco, California, USA, 2011, pp. 786–792.

- [133] A. Nchter, *3D Robotic Mapping: The Simultaneous Localization and Mapping Problem with Six Degrees of Freedom*. Heidelberg, Germany: Springer Publishing Company, Incorporated, 2010.
- [134] M. Nieuwenhuisen, J. Stückler, A. Berner, R. Klein and S. Behnke, "Shape-Primitive Based Object Recognition and Grasping," *Proc. of the 7th German Conf. on Robotics*, Munich, Germany, 2012, pp. 289–297.
- [135] K. Nordberg, "Signal Representation and Processing using Operator Groups," Thesis, Linköping University, Sweden, Linköping, Sweden, 1995.
- [136] R. Osada, T. Funkhouser, B. Chazelle and D. Dobkin, "Shape Distributions," *ACM Trans. Graphics*, Vol. 21, No. 4, 2002, pp. 807–832.
- [137] C. Paar and J. Pelzl, *Understanding Cryptography: A Textbook for Students and Practitioners*, 1st Ed. Heidelberg, Germany: Springer Publishing Company, Incorporated, 2009.
- [138] H. K. Park and M. J. Chung, "A New External Force for Active Contour Model: Virtual Electric Field," *Proc of the 2005 Inter. Conf. on Visualization, Imaging, and Image Processing*, Marbella, Spain, Sept. 2002, pp. 135–139.
- [139] M. Pilu and R. B. Fisher, "Equal-distance sampling of superellipse models," *Proc. of the British Machine Vision Conf. - BMVC 1995*, Birmingham, UK, 1995, pp. 1–10.
- [140] R. Platt, A. Fagg and R. Grupen, "Nullspace Composition of Control Laws for Grasping," *Proc. of the Inter. Conf. on Intelligent Robots and Systems - IROS 2002*, Vol. 2, Lausanne, Switzerland, Oct. 2002, pp. 1717–1723.
- [141] R. J. Platt, "Learning and Generalizing Control-based Grasping and Manipulation Skills," Ph.D. Thesis, University of Massachusetts, USA, 2006.
- [142] C. Pluempitiwiriwaj and S. Sotthivirat, "Active Contours with Automatic Initialization for Myocardial Perfusion Analysis," *IEEE Engineering in Medicine and Biology Magazine*, Vol. 3, 2005, pp. 3332–3347.
- [143] N. S. Pollard, "Closure and Quality Equivalence for Efficient Synthesis of Grasps from Examples," *Inter. Journal of Robotic Research*, Vol. 23, No. 6, 2004, pp. 595–613.
- [144] J. Ponce, S. Sullivan, J. Boissonat and J. Merlet, "On Characterizing and Computing Three- and Four-finger Force-closure Grasps for Polyhedral Objects," *Proc. of the Inter. Conf. on Robotics and Automation*, Atlanta, Georgia, USA, May 1993, pp. 821–827.
- [145] K. Qin, H. Wang, D. Li, R. Kikinis and M. Halle, "Physics-Based Subdivision Surface Modeling for Medical Imaging and Simulation," *Proc. of the Inter. Workshop on Medical Imaging and Augmented Reality - MIAR 2001*. Washington, DC, USA: IEEE Computer Society, 2001, pp. 117–124.
- [146] T. Radulescu, "Modele de Contur Deformabile pentru Segmentarea Obiectelor Puternic Non-convexe," Ph.D. Thesis, Universitatea Politehnica din Bucuresti, București, Romania, 2009.
- [147] B. Rooks, "The Harmonious Robot," *Industrial Robot*, Vol. 33, March 2006, pp. 125–130.
- [148] B. Rozenfeld, *The History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*, ser. Studies in the History of Mathematics and Physical Sciences Series. Heidelberg, Germany: Springer-Verlag, 1988.

- [149] R. B. Rusu, "Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments," Ph.D. Thesis, Computer Science Department, Technische Universitat Munchen, Munchen, Germany, Oct. 2009.
- [150] R. B. Rusu, N. Blodow and M. Beetz, "Fast Point Feature Histograms (FPFH) for 3D Registration," *Proc. of the 2009 IEEE Inter. Conf on Robotics and Automation - ICRA 2009*. Piscataway, NJ, USA: IEEE Press, 2009, pp. 1848–1853.
- [151] R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha and M. Beetz, "Towards 3D Point Cloud Based Object Maps for Household Environments," *Journal on Robotics and Autonomous Systems (Special Issue on Semantic Knowledge in Robotics)*, Vol. 56, No. 11, 2008, pp. 927–941.
- [152] A. Sahbani, S. El-Khoury and P. Bidaud, "An Overview of 3D Object Grasp Synthesis Algorithms," *Robotics and Autonomous Systems*, Vol. 60, No. 3, 2012, pp. 326 – 336.
- [153] A. Sajjanhar, G. Lu, D. Zhang, J. Hou, W. Zhou and Y.-P. P. Chen, "Spectral Shape Descriptor Using Spherical Harmonics," *Integrated Computer-Aided Engineering*, Vol. 17, No. 2, 2010, pp. 167–173.
- [154] E. Saltel, "INRIA Gamma Team Research Database," 2008. [Online]. Available: <http://www-roc.inria.fr/gamma/download/download.php>.
- [155] S. Satkin, J. Lin and M. Hebert, "Data-Driven Scene Understanding from 3D Models," *Proc. of the 23rd British Machine Vision Conf.*, Guildford, UK, 2012, pp. 412–422.
- [156] R. Schnabel, R. Wahl and R. Klein, "Efficient RANSAC for Point-Cloud Shape Detection," *Computer Graphics Forum*, Vol. 26, No. 2, 2007, pp. 214–226.
- [157] C. M. Shakarji, "Least-Squares Fitting Algorithms of the NIST Algorithm Testing System," *Journal of Research of the National Institute of Standards and Technology*, 1998, pp. 633–641.
- [158] P. Shilane, P. Min, M. Kazhdan and T. Funkhouser, "The Princeton Shape Benchmark," *Proc. of the 2004 Inter. Conf. on Shape Modeling and Applications*, Genova, Italy, June 2004, pp. 167–178.
- [159] A. Singh, D. Goldgof and D. Terzopoulos, *Deformable Models in Medical Image Analysis*. Alamitos, CA: IEEE Press, 1998.
- [160] A. Skoglund, B. Iliev and R. Palm, "A Hand State Approach to Imitation with a Next-State-Planner for Industrial Manipulators," *Proc. of the 4th Inter. Conf. on Cognitive Systems - CogSys 2008*, Zurich, Switzerland, Jan. 2008, pp. 130–137.
- [161] G. Slabaugh, B. Culbertson, T. Malzbender and R. Schafer, "A Survey of Methods for Volumetric Scene Reconstruction from Photographs," *Proc. of the 2001 Eurographics Conf. on Volume Graphics - VG 2001*, Aire-la-Ville, Switzerland, Switzerland, 2001, pp. 81–101.
- [162] F. Solina and R. Bajcsy, "Recovery of Parametric Models from Range Images: The Case for Superquadrics with Global Deformations," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 2, 1990, pp. 131–147.
- [163] J. Steffen, R. Haschke and H. Ritter, "Experience-Based and Tactile-Driven Dynamic Grasp Control," *Proc. of the Inter. Conf. on Intelligent Robots and Systems - IROS 2007*, San Diego, California, USA, Nov. 2007, pp. 2938–2943.

- [164] S. M. Stigler, *Statistics on the Table : the History of Statistical Concepts and Methods*. Cambridge (Mass.), London: Harvard University Press, 1999.
- [165] A. J. Stoddart and A. Hilton, "Registration of Multiple Point Sets," *Proc. of the 13th Inter. Conf. on Pattern Recognition*, Loire Valley, France, 1996, pp. 40–44.
- [166] K. W. Sum and P. Y. S. Cheung, "Boundary Vector Field for Parametric Active Contours," *Pattern Recognition*, Vol. 40, No. 6, June 2007, pp. 1635–1645.
- [167] R. Szeliski, D. Tonnesen and D. Terzopoulos, "Modeling Surfaces of Arbitrary Topology with Dynamic Particles," *Proc. of the 1993 IEEE Inter. Conf. on Computer Vision and Pattern Recognition*. New York, USA: IEEE Computer Society, 1993, pp. 82–87.
- [168] J. W. Tangelder and R. C. Veltkamp, "A Survey of Content Based 3D Shape Retrieval Methods," *Multimedia Tools Applications*, Vol. 39, No. 3, 2008, pp. 441–471.
- [169] G. R. Taylor and L. Kleeman, "Robust Range Data Segmentation Using Geometric Primitives for Robotic Applications," *Proc. of the 2003 Inter. Conf. on Signal and Image Processing - SIP 2003*, Honolulu, HI, USA, Aug. 2003, pp. 467–472.
- [170] B. Technology, "WAM Arm," 2010. [Online]. Available: <http://www.barrett.com/robot/products-arm-specifications.html>.
- [171] D. Terzopoulos and D. Metaxas, "Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, 1991, pp. 703–714.
- [172] D. Terzopoulos and K. Fleischer, "Modeling Inelastic Deformation: Viscoelasticity, Plasticity, Fracture," *Intern. Journal on Computer Graphics*, Vol. 22, No. 4, 1988, pp. 269–278.
- [173] A. C. Thompson, *Minkowski Geometry*. Cambridge University Press, 1996.
- [174] R. F. Tobler and S. Maierhofer, "A Mesh Data Structure for Rendering and Subdivision," *Proc. of the Inter. Conf. in Central Europe on Computer Graphics, Visualization and Computer Vision - WSCG 2006*, Plzen, Czech Republic, Feb. 2006, pp. 157–162.
- [175] F. Tombari, S. Salti and L. Di Stefano, "Unique Signatures of Histograms for Local Surface Description," *Proc. of the 11th European Conf. on Computer Vision - ECCV 2010*, Vol. 2, Berlin, Heidelberg, 2010, pp. 356–369.
- [176] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ, USA: Prentice Hall, 1998.
- [177] M. Turiac, M. Ivanovici, T. Radulescu and V. Buzuloiu, "Variance-Driven Active Contours," *Proc. of the 2010 Inter. Conf. on Image Processing, Computer Vision, and Pattern Recognition*, Las Vegas, Nevada, USA, 2010, pp. 83–86.
- [178] L. Vainio, E. Symes, R. Ellis, M. Tucker and G. Ottoboni, "On the Relations Between Action Planning, Object Identification, and Motor Representations of Observed Actions and Objects," *Cognition*, Vol. 108, No. 2, 2008, pp. 444–465.
- [179] O. Veksler, "Efficient Graph-based Energy Minimization Methods in Computer Vision," Ph.D. Thesis, Cornell University, Ithaca, NY, USA, 1999.
- [180] P. A. Viola and M. J. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," *Proc. of the 2011 Inter. Conf. on Computer Vision and Pattern Recognition -*

- CVPR 2001*, Vol. 1, Kauai, HI, USA, Dec. 2001, pp. 511–518.
- [181] I. Volosyak, O. Ivlev and A. Graser, “Rehabilitation Robot FRIEND II - The General Concept and Current Implementation,” *Proc. of the 9th Inter. Conf. on Rehabilitation Robotics - ICORR 2005*, Chicago, Illinois, Jun. 2005, pp. 540–544.
- [182] I. Wald and V. Havran, “On Building Fast kd-Trees for Ray Tracing, and On Doing That in  $O(N \log N)$ ,” *Proc. of the 2006 IEEE Symp. on Interactive Ray Tracking*, Salt Lake City, USA, 2006, pp. 61–70.
- [183] Q. Wang and K. L. Boyer, “The Active Geometric Shape Model: A New Robust Deformable Shape Model and Its Applications,” *Computer Vision and Image Understanding*, Vol. 116, No. 12, Dec. 2012, pp. 1178–1194.
- [184] T. Wang, I. Cheng and A. Basu, “Fluid Vector Flow and Applications in Infant Brain Tumor Segmentation,” *IEEE Trans. on Biomedical Engineering*, Vol. 56, No. 3, 2009, pp. 781–789.
- [185] Y. Wang and Y. Jia, “Convolutional Virtual Electric Field External Force for Active Contours,” *Proc. of the 9th Asian Conf. on Computer Vision - ACCV 2009*, Vol. 3. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 11–20.
- [186] Z. Wang, H. Liu, Y. Qian and T. Xu, “Real-Time Plane Segmentation and Obstacle Detection of 3D Point Clouds for Indoor Scenes,” *Proc. of the 12th European Conf. on Computer Vision - ECCV 2012*, Vol. 2, Berlin, Heidelberg, 2012, pp. 22–31.
- [187] H. Xie, G. Li, H. Ning, C. Menard, C. Coleman and R. Miller, “3D Voxel Fusion of Multi-Modality Medical Images in a Clinical Treatment Planning System,” *Proc. of the 17th IEEE Symp. on Computer-Based Medical Systems*, Bethesda, MD, USA, June 2004, pp. 48–53.
- [188] C. Xu and J. L. Prince, “Snakes, Shapes, and Gradient Vector Flow,” *IEEE Trans. on Image Processing*, Vol. 7, No. 3, 1998, pp. 359–369.
- [189] C. Xu and J. L. Prince, “Generalized Gradient Vector Flow External Forces for Active Contours,” *Inter. Journal on Signal Processing*, Vol. 71, No. 2, 1998, pp. 131–139.
- [190] L. Ying, J. Fu and N. Pollard, “Data-Driven Grasp Synthesis Using Shape Matching and Task-Based Pruning,” *IEEE Trans. on Visualization and Computer Graphics*, Vol. 13, No. 4, 2007, pp. 732–747.
- [191] Z. Zhang, “Iterative Point Matching for Registration of Free-Form Curves and Surfaces,” *Inter. Journal on Computer Vision*, Vol. 13, No. 2, 1994, pp. 119–152.
- [192] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering and D. Comaniciu, “Four-Chamber Heart Modeling and Automatic Segmentation for 3D Cardiac CT Volumes Using Marginal Space Learning and Steerable Features,” *IEEE Trans. on Medical Imaging*, Vol. 27, No. 11, 2008, pp. 1668–1681.
- [193] X. Zhu and J. Wang, “Synthesis of Force-Closure Grasps on 3D Objects Based on the Q Distance,” *IEEE Trans. on Robotics and Automation*, Vol. 19, No. 4, 2003, pp. 669–679.
- [194] Y. Zyrianov, “Distribution-Based Descriptors of the Molecular Shape,” *Journal of Chemical Information and Modeling*, Vol. 45, No. 3, 2005, pp. 657–672.