

Razpoznavanje objektov na osnovi superelipsoidnih modelov

Jaka Krivic

MAGISTRSKA NALOGA
IZ
RAČUNALNIŠTVA IN INFORMATIKE

predložena
Fakulteti za računalništvo in informatiko
Univerze v Ljubljani
kot delna izpolnitev pogoja za pridobitev naslova
magister računalništva in informatike

September 2000

Mentor:
prof. dr. Franc Solina

Magistrska naloga je bila izdelana pod mentorstvom prof. Franca Soline in je last Fakultete za računalništvo in informatiko v Ljubljani. Za objavljanje in uporabo rezultatov magistrskega dela je potrebno soglasje prej omenjene ustanove.

Razpoznavanje objektov na osnovi superelipsoidnih modelov

Jaka Krivic

Mentor:
prof. dr. Franc Solina

POVZETEK

Magistrska naloga se ukvarja z razpoznavanjem objektov na osnovi superelipsoidnih modelov. Superelipsoidi so tridimenzionalni modeli, določeni s parametri oblike, velikosti, orientacije in pozicije. Iz globinskih slik scene rekonstruiramo modele z uporabo segmentacijske paradigme gradi-in-izberi. Rekonstruirani modeli tako opisujejo dele, ki so vidni na prizoru. Z uporabo interpretacijskega drevesa lahko preverimo prisotnost vnaprej poznanih objektov. Če je objekt prisoten, ugotovimo tudi njegovo lokacijo v prostoru.

Object Recognition Using Superellipsoid Models

Jaka Krivic

Supervisor:
prof. dr. Franc Solina

ABSTRACT

This thesis deals with object recognition using superellipsoid built models. Superellipsoids are three dimensional models, defined with parameters of shape, size, orientation, and position. Superellipsoids are constructed from range images using recover-and-select paradigm. These superellipsoids describe the objects' constituent parts. The presence of an object from the model database can be determined by using an interpretation tree. If an object is present, its location in space is also determined.

Kazalo

Povzetek	iii
Abstract	iii
1 Uvod	1
1.1 Postavitev problema	1
1.2 Pregled literature	2
1.2.1 Razpoznavanje objektov	2
1.2.2 Modeliranje objektov	3
1.2.3 Rekonstrukcija delov	4
2 Superelipsoidi	6
2.1 Definicije	6
2.1.1 Površina superelipsoida	6
2.1.2 Normale	7
2.1.3 Implicitna enačba superelipsoida in funkcija znotraj-zunaj	8
2.2 Superelipsoid v splošni legi	8
2.3 Prostornina superelipsoida	9
2.4 Vztrajnostni momenti superelipsoida	10
2.5 Rekonstrukcija superelipsoidov z metodo gradi-in-izberi	11
2.5.1 Gradnja modelov	12
2.5.2 Izbira modelov	13
2.5.3 Izmenjevanje gradnje in izbiranja modelov	14
2.5.4 Prikaz poteka in uporaba opisov	14
3 Razpoznavanje objektov	17
3.1 Interpretacijska drevesa	17
3.1.1 Navidezna ujemanja in lažne značilke	20
3.2 Model objekta	22
3.3 Posebnosti pri uporabi superelipsoidov	24
3.3.1 Preverjanje konsistentnosti ujemanja	25
3.3.2 Verifikacija interpretacije	29

4	Rezultati	32
4.1	Razpoznavanje prizorov z enim samim objektom	32
4.2	Razpoznavanje pri kompleksnih prizorih	37
4.3	Napake pri razpoznavanju	43
4.4	Možne izboljšave	46
5	Zaključek	49
5.1	Nadaljnje delo	49
	Literatura	53
	Zahvala	54
	Izjava	55

Slike

2.1	Odvisnost oblike superelipsoidov od vrednosti parametrov ε_1 in ε_2	7
2.2	Primer rekonstrukcije superelipsoidov z metodo gradi-in-izberi.	15
2.3	Primer rekonstrukcije superelipsoidov z metodo gradi-in-izberi, kjer se predmeta ne da popolnoma opisati s superelipsoidi.	16
3.1	Interpretacijsko drevo, ki prikazuje celoten prostor iskanja.	18
3.2	Primer objekta in njegov model.	19
3.3	Del interpretacijskega drevesa za problem s slike 3.2	20
3.4	Človeška figurica kot objekt za razpoznavanje: intenzitetna slika.	22
3.5	Model človeške figurice.	23
3.6	Primer rekonstrukcije globinske slike človeške figurice s superelipsoidi.	24
3.7	Prikaz količin z enačbe 3.11.	31
4.1	Razpoznavanje prizora z enim objektom. (1).	34
4.2	Razpoznavanje prizora z enim objektom. (2).	35
4.3	Razpoznavanje prizora z enim objektom. (3).	36
4.4	Razpoznavanje prizora z več objekti. (1).	38
4.5	Razpoznavanje prizora z več objekti. (2).	39
4.6	Razpoznavanje prizora z več objekti. (3).	40
4.7	Razpoznavanje prizora z več objekti. (4).	41
4.8	Razpoznavanje prizora z več objekti. (5).	42
4.9	Primer napak pri razpoznavanju. (1).	43
4.10	Primer napak pri razpoznavanju. (2).	44
4.11	Primer napak pri razpoznavanju. (3).	45
4.12	Prilagajanje delov modela regijam globinske slike.	46
4.13	Primer dobrega položaja objekta.	48
4.14	Primer slabega položaja objekta.	48

Tabele

2.1	Vztrajnostni momenti superelipsoida	11
3.1	Algoritem za razpoznavanje z interpretacijskim drevesom.	21
3.2	Parametri, ki določajo model objekta s slike 3.4.	23
3.3	Uporabljen algoritem za razpoznavanje z interpretacijskim drevesom.	26
3.4	Deli modela objekta, ki ustrezajo ujemanjem na različnih globinah drevesa.	27
3.5	Intervali volumnov za dele objekta.	27
3.6	Omejitve za parametre superelipsoida, ki opisuje glavo.	27
4.1	Rezultati razpoznavanja na 56 prizorih z enim samim objektom. . .	33

1

Uvod

Računalniški vid naj bi strojem omogočil interakcijo z realnim zunanjim svetom. Omogočal naj bi zavedanje okolice, kar bi stroj lahko uporabil npr. za planiranje poti in manipuliranje z objekti. Tako bi bili stroji v svojem delovanju samostojni in uspešni tudi v spreminjajočem se okolju. V bližnji prihodnosti pa ni pričakovati splošnega sistema za računalniški vid, že zato ne, ker še ne razumemo povsem sistema človekovega vida.

Zavedanje okolice lahko dosežemo z razpoznavanjem objektov. Razvitih je bilo že mnogo sistemov za razpoznavanje objektov s takimi in drugačnimi lastnostmi. Pri tem je vsekakor pomembno, kako predstavljajo kompleksne 3D objekte. Predstavitev kompleksnih 3D objektov z njihovimi deli je vsekakor odlična za mnoge računske in kognitivne naloge. Ima dolgo tradicijo v računalniško podprtem načrtovanju, simulacijah, kognitivni psihologiji. Take predstavitve pa so se uveljavile tudi v računalniškem vidu. Predstavitev objekta z njegovimi deli omogoča, da kompleksne objekte predstavimo s kompaktnim opisom, kar ne zmanjša le količine podatkov, temveč lahko pričakujemo tudi, da vsaj približno opisuje strukturo objektov s slike. Različni objekti na neki sceni naj bi imeli tudi različne opise v smislu njihovih delov. Na ta način postane možno razločevanje med objekti in odločanje na višjem nivoju.

1.1 Postavitev problema

V tem delu se bomo ukvarjali z razpoznavanjem objektov iz globinskih slik, ki so segmentirane z metodo gradi-in-izberi in modelirane s superelipsoidi. Cilj te naloge je za razpoznavanje uporabiti rekonstrukcije superelipsoidov iz globinskih slik, ki jih dobimo z obstoječim sistemom za segmentacijo in rekonstrukcijo z metodo gradi-in-izberi. Rekonstrukcije v bistvu pomenijo opis objektov s slike na ravni njihovih sestavnih delov in njihove povezave oz. strukturo objektov. Naloga je tako na podlagi takih opisov preveriti prisotnost poznanega objekta na sliki in določiti njegovo lokacijo v 3D prostoru.

Kot objekt nas je specifično zanimala človeška figura, ki je zanimiva predvsem s

stališča fleksibilnosti. Ker so posamezni deli gibljivi, lahko figura nastopa v različnih pozah, kar močno poveča kompleksnost razpoznavanja objekta.

1.2 Pregled literature

Magistrska naloga se nanaša na področje razpoznavanja in modeliranja objektov ter rekonstrukcijo delov s slik, zato bomo na kratko predstavili dela s teh področij.

1.2.1 Razpoznavanje objektov

Sistem za razpoznavanje objektov išče ujemanja med opisi znanih objektov in opisom prizora, ki naj ga zazna. Na neki stopnji tega procesa sistem poveže ujemajoče značilke prizora s tistimi na modelu. Možnih ujemanj je mnogo, kar nadalje pomeni, da je za pravilno celotno ujemanje potrebno preiskusiti zelo veliko število kombinacij parov. Pri razpoznavanju objektov je torej potrebno učinkovito iskanje po prostoru možnih rešitev. Metode za razpoznavanje objektov, ki so bile razvite, rešujejo različne aplikacijske probleme. Tako je vsaka metoda tudi primerna le za majhno število opravil, ki lahko uporabljajo sistem razpoznavanja.

Nevatia in Binford [20] sta bila med prvimi, ki sta pri razpoznavanju uporabila 3D podatke. Model objekta in dele prizora sta opisala s posplošenimi cilindri in jih razvrstila v relacijski graf. Razpoznavanje objekta je tako ujemanje dveh grafov.

Morda najboljše dodelana je Biedermanova [3, 9] teorija o razpoznavanju po delih (*recognition by components, RBC*). Kompleksne objekte opišemo z majhno množico ($N \leq 36$) preprostih 3D oblik - geonov, skupaj z relacijami med njimi.

Flynn in Jain [11] sta razvila tehniko za razpoznavanje 3D objektov iz globinskih slik, ki temelji na indeksiranju invariantnih značilnosti in ujemanju 3D(podatki) - 3D(model). Na sliki zaznane značilnosti se uporabijo za izbiro hipotez za identiteto trojčkov površin s prizora. Uporabijo se kot večdimenzionalni indeks za naslavljanje tabele *proto-hipotez*, ki povezujejo površine modela s površino s slike. Te hipoteze se rangirajo s pomočjo ujemanja.

Johnson in Hebert [16] sta uporabila vrtilne slike (*spin images*) za ujemanje površin. Razpoznavanje temelji na iskanju ujemanja med površinami s pomočjo vrtilnih slik. Najprej se zgradi knjižnica vrtilnih slik znanih objektov. Za neko orientirano točko s prizora se pri razpoznavanju zgradi vrtilna slika, najde ujemanje z vrtilno sliko z modela, izračuna transformacijo med obema točkama in preveri, če je transformacija pravilna. Tehniko sta uspešno uporabila v prizorih z večimi objekti in zakrivanjem.

3DPO (Three-Dimensional Part Orientation) [6] je metoda, ki sta jo razvila Bolles in Horaud. Cilj je bil razpoznavanje in lociranje velikega števila kompleksnih industrijskih delov, da bi jih lahko robotska roka dosegala. Pri iskanju ujemanja sta uporabila le majhno število značilk oz. skupin značilk, ki pa so zelo značilne za objekt, in tako zmanjšala kompleksnost iskanja.

Grimson [12] je za razpoznavanje uporabil interpretacijska drevesa. Interpretacija se oblikuje s potjo v prostoru vseh možnih parov slika-model, ki so predstavljena z drevesom - interpretacijskim drevesom. Ker pa število možnih preslikav v realnem pomeni mnogo preveliko kompleksnost, je Grimson za zmanjšanje kompleksnosti uporabil zadovoljevanje geometrijskim omejitvam. Pot po interpretacijskem drevesu tako nadaljuje le, če se značilki lokalno ujemata.

Bischof in Caelli [4, 5] sta razvila metodo pogojnega generiranja pravil (*Conditional Rule Generation, CRG* [4]), ki se ukvarja z učenjem pravil o objektih, s spoznavanjem objektov. Ideja CRG je generirati klasifikacijska pravila, ki vsebujejo tudi strukturno informacijo o vzorcih. CRG išče pojavitve unarnih in binarnih značilnosti med deli vzorca, ki so povezani, in zgradi drevo hierarhično organiziranih pravil za klasifikacijo novih vzorcev. Pri uporabi teh pravil na novih pogledih, torej pri interpretaciji prizorov, avtorja predlagata uporabo ovrednotenja pravil (*Scene Understanding using Rule Evaluation, SURE*) [5]. SURE uporablja pravila, ki jih zgradi CRG, in upošteva splošne strukturne značilnosti prizorov, kar izboljša zanesljivost rezultatov.

1.2.2 Modeliranje objektov

Načrtovanje modelov je pomembna veja raziskovanja v računalniškem vidu, saj je uporaba pravih modelov ključnega pomena. Modeli so lahko geometrijske značilnosti objekta, lahko so značilnosti slik objekta (npr. robovi, barve, konture...) ali celo slike same (npr. predstavitev objekta z lastnimi slikami, *eigenimages* [19]). Če se omejimo le na geometrijske značilnosti [2], bi lahko bila definicija modela 3D objekta za potrebe računalniškega vida taka: to je računski predstavitev zadostnega števila geometrijskih lastnosti objekta, s katerimi lahko izvedemo določeno nalogo. To ne pomeni, da naj bo model izčrpen, potrebujemo le zadosti informacij, da lahko izvedemo nalogo. Če je na primer naša naloga razločiti med A in B, potrebujemo le najmanjšo količino informacije, ki zadostuje, da ju lahko razločimo.

Modele, s katerimi predstavljamo 3D točke, lahko delimo na lokalne in globalne. Lokalni modeli predstavljajo objekte kot množico nestrukturiranih podatkov, kot so robovi ali površine. Globalni modeli pa poskušajo objekt obravnavati kot celoto v lastnem koordinatnem sistemu. Od slednjih so najpogosteje uporabljeni posplošeni cilindri [20] in superkvadrki, podmnožica katerih so superelipsoidi. Seveda obstajajo tudi drugi modeli, kot npr. sferične harmonične površine [22] in hiperkvadrki [13].

Modeli, ki temeljijo na *videzu*, predstavljajo objekt z eno ali več slikami. Primer je metoda, ki uporablja lastne slike (*eigenimages* [19]). Ta dobro deluje za 2D objekte ali 3D objekte z omejenim številom položajev. Za ustrezno predstavitev kompleksnega 3D objekta potrebujemo veliko število slik. Razpoznavanje je iskanje slike v zbirki modelov, ki je najbolj podobna vhodni sliki. Podobnost določimo z merjenjem razdalje med slikami, npr. s korelacijo. Slike so torej direktno primerljive, kar je glavna prednost teh modelov. Dobra lastnost je tudi to, da lahko modeli-

ramo tudi objekte brez evidentnih značilnk. Po drugi strani pa pogoji zajemanja slik (osvetlitev, položaj, lokacija) za modeliranje vplivajo na slike in s tem na razpoznavanje.

Na drugi strani pa modeli, ki temeljijo na *značilnostih*, predstavljajo objekte z značilnkami, tipom značilnk in prostorskimi relacijami med njimi. To so lahko geometrijske značilke (npr. ploskve, robovi...) ali negeometrijske (npr. barva, odbojne lastnosti...). Identifikacija s takimi modeli pomeni najti množico značilnk, ki jasno razločujejo objekt. Lokacijo dobimo tako, da najdemo ujemanje za značilke s slike in modela, jih projeciramo na sliko in izračunamo 3D transformacijo za pozicijo in orientacijo. Taki modeli imajo zelo zgoščene opise objektov, nudijo nekaj robustnosti pri zakrivanju (značilke so lokalnega značaja), osvetlitvi in spreminjanju položaja. Dejstvo, da niso direktno primerljivi s sliko in je torej potrebno iz slike pridobiti značilke, je njihova glavna slabost. Nadalje jih lahko razdelimo glede na tip značilnk, npr. predstavitve z mejami (*B-reps*, *boundary representations*) predstavljajo objekt z njegovimi mejami (črtami, površinami); *volumetrične predstavitve* opisujejo pozicijo, velikost in obliko glavnih delov objektov ali pa samo obliko prostora, ki ga zavzema objekt. Primer volumetričnih modelov so superkvadrniki.

Odločitev o izbiri modela temelji predvsem na vrsti objektov, ki jih modeliramo. Če naj, na primer, predstavimo množico trianguliranih likov in imamo na voljo detektor linij, je predstavitev v obliki linij in njihovih medsebojnih pozicij prava: dobro predstavlja obliko objektov in značilke (linije) lahko detektiramo. Modeli na podlagi značilnk niso smiselni, če objekti nimajo ugotovljivih značilnk: če naj razpoznavamo obraze in imamo na voljo detektor linij, se je bolje odločiti za modele na podlagi videza.

1.2.3 Rekonstrukcija delov

Da bi iz slike lahko dobili njen opis z deli objektov, ki so na sliki, moramo rešiti dve nalogi. Sliko je potrebno razdeliti na območja, ki ustrezajo enemu samemu delu, čemur pravimo segmentacija. Nato moramo iz tega območja rekonstruirati model dela, to je dobiti parametre modela, ki dobro opisuje območje. Nalogi lahko opravimo posebej, najprej segmentacijo in nato še rekonstrukcijo, vendar pa tak način ne upošteva oblik, ki jih sploh lahko opišemo z določenimi modeli. Tako nekega segmenta slike z izbranimi modeli ne moremo opisati. Če združimo segmentacijo in rekonstrukcijo, lahko slike segmentiramo le na dele, ki jih dobro opisujejo dani modeli.

Na področju rekonstrukcije superelipsoidov iz globinskih slik je bilo razvitih že kar nekaj metod. Prvi, ki je uporabil superelipsoide v računalniškem vidu, je bil Pentland [21]. Predlagal je rekonstrukcijo z uporabo eksplisitne enačbe. Metoda se razen za nekaj preprostih sintetičnih slik ni obnesla.

Solina in Bajcsy [25] sta rešila problem rekonstrukcije posameznega superelipsoida iz že segmentiranih slik. Rekonstrukcijo sta definirala kot minimizacijo

funkcije prileganja z metodo najmanjših kvadratov. Za ta problem nelinearne optimizacije sta uporabila gradientno metodo. Funkcijo prileganja sta definirala na osnovi funkcije znotraj-zunaj.

Nato je bilo razvitih še kar nekaj metod, zato naj omenimo le nekatere. Yokoya s sodelavci [27] je predlagal novo funkcijo prileganja, katera poleg razlike med točkami na modelu in superelipsoidom uporablja še razliko v normalah. Za minimizacijo je uporabil algoritem ohlajanja.

Callari in Maniscalco [7] sta podala rekonstrukcijo s pomočjo senčenja oz. osvetlitve (*shading flow field*).

Leonardis [17] je podal segmentacijsko paradigmo 'gradi-in-izberi', v kateri se izmenjuje segmentacija z rekonstrukcijo. Jaklič [14] jo je uspešno uporabil s superelipsoidi kot modeli. Ta metoda je uporabljena tudi v tej nalogi.

Bralec najde podrobnejšo predstavitev metod v [24, 15].

Pregled naloge

Preostanek dela je sestavljen takole:

- V drugem poglavju z naslovom *Superelipsoidi* je podana definicija superelipsoidov kot matematičnih teles in nekatere njihove lastnosti. Opisana je tudi metoda gradi-in-izberi, ki smo jo uporabili za rekonstrukcijo superelipsoidov iz globinskih slik.
- Tretje poglavje z naslovom *Razpoznavanje objektov* opiše metodo razpoznavanja ter poda značilnosti modelov, ki so bile uporabljene v nalogi.
- Poglavje *Rezultati* predstavi eksperimentalne rezultate in na primerih pokaže dobre in slabe strani metode.
- V poglavju *Zaključek* je podan povzetek rezultatov dela. Navedene so tudi možnosti nadaljnjega dela.
- Sledijo še spisek uporabljene literature, *Zahvala* in *Izjava*.

2

Superelipsoidi

V tem poglavju bomo predstavili osnovne pojme superelipsoidov, ki jih bomo uporabljali skozi vso magistrsko nalogo. Superelipsoidi in njihove geometrijske lastnosti so podrobno opisani v [15].

2.1 Definicije

Kot smo že v uvodu povedali, so superelipsoidi podmnožica družine parametričnih modelov superkvadrikov. Definirali bomo vektor površine superelipsoida in funkcijo znotraj-zunaj, ki za dano točko v prostoru pove, ali leži znotraj, zunaj ali na površini superelipsoida.

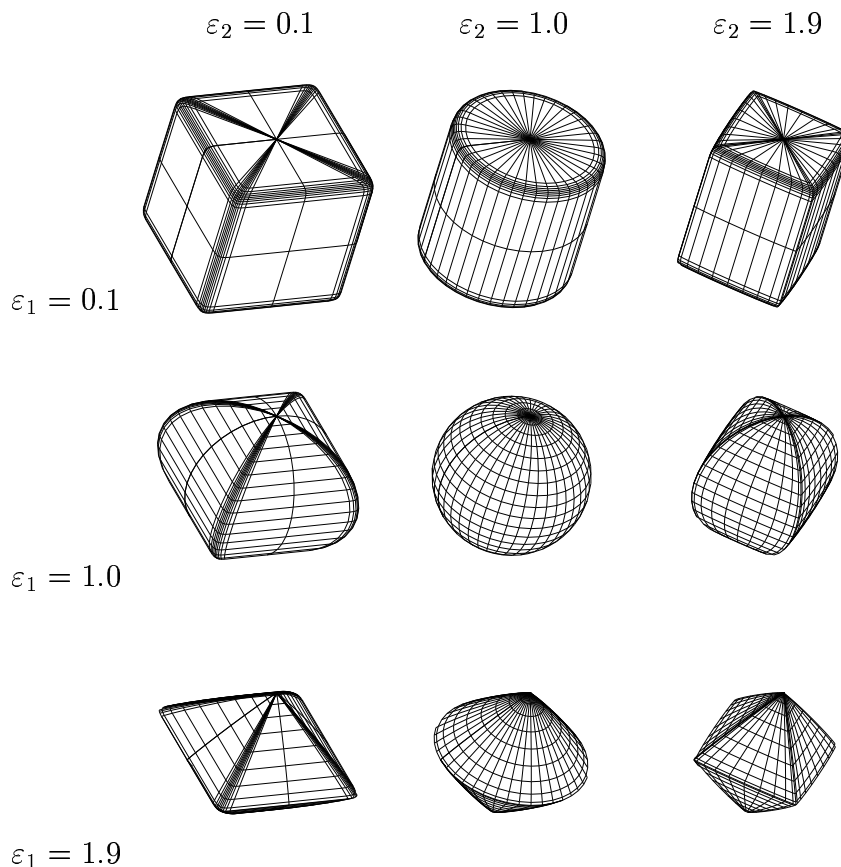
2.1.1 Površina superelipsoida

Superelipsoid je definiran z naslednjim 3D vektorjem:

$$\vec{s}(\eta, \omega) = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix} = \begin{bmatrix} a_1 \cos^{\varepsilon_1} \eta \cos^{\varepsilon_2} \omega \\ a_2 \cos^{\varepsilon_1} \eta \sin^{\varepsilon_2} \omega \\ a_3 \sin^{\varepsilon_1} \eta \end{bmatrix}, \quad \begin{array}{ccc} -\frac{\pi}{2} & \leq & \eta & \leq & \frac{\pi}{2} \\ -\pi & \leq & \omega & < & \pi \end{array} \quad (2.1)$$

Vektor površine \vec{s} izhaja iz koordinatnega izhodišča in določa površino superelipsoida. Parameter ω je kot med x osjo in projekcijo vektorja \vec{s} na x - y ravnino, medtem ko je parameter η kot med \vec{s} in njegovo projekcijo na x - y ravnino. V sferičnih koordinatah ω ustreza dolžini (latitude angle), η pa polarni razdalji (longitude angle). Parametri a_1 , a_2 , a_3 določajo velikost superelipsoida v smereh x , y in z osi. ε_1 in ε_2 sta parametra, ki določata obliko superelipsoida. ε_1 jo določa v smislu z osi, ε_2 pa jo določa v smislu x - y ravnine.

Zgoraj navedeni zapis vektorja površine je nekoliko poenostavljen (osnovna notacija [1]). Potenciranje negativne osnove na realen eksponent namreč daje kompleksne rezultate. Temu se izognemo tako, da potenciranje x^y preoblikujemo v $\text{sgn}(x) |x|^y$. Pri tem je $\text{sgn}(x)$ funkcija predznaka.



Slika 2.1: Odvisnost oblike superelipsoidov od vrednosti parametrov ε_1 in ε_2 (parametri, ki določajo velikost, so pri vseh superelipsoidih enaki). Povzeto iz [15].

V magistrski nalogi bomo uporabljali za preoblikovano potenciranje kar notacijo x^y .

2.1.2 Normale

Vektorji normal igrajo pomembno vlogo pri 3D geometrijskih telesih (npr. določanje vidne površine telesa). Poglejmo, kako so normale definirane pri superelipsoidih.

Vektor normale izračunamo kot vsak vektor normale ploskve – izračunamo vektorski produkt dveh vektorjev, ki ležita na tangantah na površino: $\vec{n} = \frac{\partial \vec{s}}{\partial \eta} \times \frac{\partial \vec{s}}{\partial \omega}$. Pri tem \times označuje vektorski produkt. *Vektor normale* superelipsoida torej je:

$$\vec{n}(\omega, \eta) = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} \frac{1}{a_1} \cos^{2-\varepsilon_1} \eta \cos^{2-\varepsilon_2} \omega \\ \frac{1}{a_2} \cos^{2-\varepsilon_1} \eta \sin^{2-\varepsilon_2} \omega \\ \frac{1}{a_3} \sin^{2-\varepsilon_1} \eta \end{bmatrix}. \quad (2.2)$$

2.1.3 Implicitna enačba superelipsoida in funkcija znotraj-zunaj

Površino superelipsoida lahko definiramo tudi kot množico rešitev implicitne enačbe superelipsoida, ki jo dobimo tako, da iz enačb 2.1 izločimo parametra ω in η

$$\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} = \cos^{\frac{2\epsilon_1}{\epsilon_2}} \eta \cos^2 \omega, \quad (2.3)$$

$$\left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} = \cos^{\frac{2\epsilon_1}{\epsilon_2}} \eta \sin^2 \omega, \quad (2.4)$$

$$\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} = \cos^{\frac{2\epsilon_1}{\epsilon_2}} \eta, \quad (2.5)$$

$$\left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} = \sin^2 \eta, \quad (2.6)$$

upoštevamo še $\sin^2 \eta + \cos^2 \eta = 1$ in dobimo implicitno enačbo superelipsoida

$$F(x, y, z) = \left(\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} = 1. \quad (2.7)$$

Če na $F(x, y, z)$ gledamo kot na funkcijo, razdeli prostor na tri področja. Točka (x, y, z) leži:

1. znotraj telesa $F(x, y, z) < 1$,
2. na površini telesa $F(x, y, z) = 1$ in
3. zunaj telesa $F(x, y, z) > 1$.

Tej funkciji pravimo *funkcija znotraj-zunaj* (inside-outside function).

2.2 Superelipsoid v splošni legi

Dosedanje definicije (vektor površine in funkcija znotraj-zunaj) so bile zapisane s koordinatami lokalnega koordinatnega sistema. Če želimo določiti lego superelipsoida v prostoru, moramo poznati transformacijo, ki lokalni koordinatni sistem pretvori v svetovni koordinatni sistem. Ta transformacija je sestavljena iz rotacije v lokalnem koordinatnem sistemu in premika izhodišča lokalnega koordinatnega sistema v izhodišče svetovnega.

Rotacijo izvedemo z ZYZ-Eulerjevo rotacijo [15]. Trije Eulerjevi koti φ , ϑ in ψ popolnoma določajo orientacijo transformiranega koordinatnega sistema. Kot φ pomeni rotacijo okoli z osi, kot ϑ okoli nove y osi in kot ψ okoli nove z osi. Izračunajmo *rotacijsko matriko* R :

$$R = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \vartheta & 0 & \sin \vartheta \\ 0 & 1 & 0 \\ -\sin \vartheta & 0 & \cos \vartheta \end{bmatrix} \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} = \quad (2.8)$$

$$\begin{bmatrix} \cos \varphi \cos \vartheta \cos \psi - \sin \varphi \sin \psi, & -\sin \varphi \cos \vartheta \cos \psi - \cos \varphi \sin \psi, & \sin \vartheta \cos \psi \\ \cos \varphi \cos \vartheta \sin \psi + \sin \varphi \cos \psi, & -\sin \varphi \cos \vartheta \sin \psi + \cos \varphi \cos \psi, & \sin \vartheta \sin \psi \\ -\cos \varphi \sin \vartheta, & \sin \varphi \sin \vartheta, & \cos \vartheta \end{bmatrix}. \quad (2.9)$$

Vektor površine \vec{s}' superelipsoida, zapisan z globalnimi koordinatami, dobimo tako, da vektor površine, zapisan z lokalnimi koordinatami, z leve pomnožimo z rotacijsko matriko R in prištejemo vektor premika $\vec{p} = [p_x, p_y, p_z]^T$:

$$\vec{s}' = R\vec{s} + \vec{p}. \quad (2.10)$$

Če želimo zapisati funkcijo znotraj-zunaj f' v svetovnem koordinatnem sistemu, moramo svetovni koordinatni sistem pretvoriti v lokalni koordinatni sistem. Funkcija znotraj-zunaj zapisana s svetovnimi koordinatami, torej je:

$$f'(\vec{s}') = f(R^{-1}(\vec{s}' - \vec{p})). \quad (2.11)$$

Za vsako rotacijsko matriko A velja $A^{-1} = \det(A)A^T$. Ker velja $\det(R) = 1$, je $R^{-1} = R^T$. Končna oblika funkcije znotraj-zunaj je:

$$f'(\vec{s}') = f(R^T(\vec{s}' - \vec{p})). \quad (2.12)$$

Če povzamemo, superelipsoid v splošni legi je določen z obliko, orientacijo in premikom:

$$\Lambda = \langle a_1, a_2, a_3, \varepsilon_1, \varepsilon_2, \varphi, \vartheta, \psi, p_x, p_y, p_z \rangle. \quad (2.13)$$

2.3 Prostornina superelipsoida

Če superelipsoid prerežemo vzdolž z osi z ravnino paralelno z ravnino $x - y$, dobimo superelipso [15]. z koordinata je odvisna le od η , tako da je površina te superelipse, ki jo izračunamo po Greenovi formuli

$$A(\eta) = 2a(\eta)b(\eta)\varepsilon_2 B\left(\frac{\varepsilon_2}{2}, \frac{\varepsilon_2 + 2}{2}\right), \quad (2.14)$$

kjer je $B(x, y)$ beta funkcija, ki je povezana z gama funkcijo in je definirana kot

$$B(x, y) = 2 \int_0^{\pi/2} \sin^{2x-1} \phi \cos^{2y-1} \phi d\phi = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \quad (2.15)$$

$$\begin{aligned}
V &= 2 \int_0^{a_3} A(z) dz = \\
&= 2 \int_0^{\pi/2} A(\eta) \dot{z}(\eta) d\eta = \\
&= 4a_1 a_2 a_3 \epsilon_1 \epsilon_2 B\left(\frac{\epsilon_2}{2}, \frac{\epsilon_2 + 2}{2}\right) \int_0^{\pi/2} \sin^{\epsilon_1 - 1} \eta \cos^{2\epsilon_1 + 1} \eta d\eta = \\
&= 2a_1 a_2 a_3 \epsilon_1 \epsilon_2 B\left(\frac{\epsilon_1}{2}, \epsilon_1 + 1\right) B\left(\frac{\epsilon_2}{2}, \frac{\epsilon_2 + 2}{2}\right). \tag{2.16}
\end{aligned}$$

Iz gornje enačbe lahko z nekaj algebre dobimo drugo obliko

$$V = 2a_1 a_2 a_3 \epsilon_1 \epsilon_2 B\left(\frac{\epsilon_1}{2} + 1, \epsilon_1\right) B\left(\frac{\epsilon_2}{2}, \frac{\epsilon_2}{2}\right). \tag{2.17}$$

2.4 Vztrajnostni momenti superelipsoida

Če superelipsoid razrežemo vzdolž z osi na rezine neskončno majhne debeline, paralelne z ravnino $x - y$, lahko z uporabo Steinerjeve formule določimo njegove vztrajnostne momente [15].

$$\begin{aligned}
I_{xx} &= \int \int \int_V (y^2 + z^2) dx dy dz = \tag{2.18} \\
&= \int_{-a_3}^{+a_3} \left(\int \int_{S(z)} y^2 dx dy + \int \int_{S(z)} z^2 dx dy \right) dz = \\
&= \frac{1}{2} a_1 a_2 a_3 \epsilon_1 \epsilon_2 \left(a_2^2 B\left(\frac{3}{2}\epsilon_2, \frac{1}{2}\epsilon_2\right) B\left(\frac{1}{2}\epsilon_1, 2\epsilon_1 + 1\right) + \right. \\
&\quad \left. + 4a_3^2 B\left(\frac{1}{2}\epsilon_2, \frac{1}{2}\epsilon_2 + 1\right) B\left(\frac{3}{2}\epsilon_1, \epsilon_1 + 1\right) \right),
\end{aligned}$$

$$\begin{aligned}
I_{yy} &= \int \int \int_V (x^2 + z^2) dx dy dz = \tag{2.19} \\
&= \int_{-a_3}^{+a_3} \left(\int \int_{S(z)} x^2 dx dy + \int \int_{S(z)} z^2 dx dy \right) dz = \\
&= \frac{1}{2} a_1 a_2 a_3 \epsilon_1 \epsilon_2 \left(a_1^2 B\left(\frac{3}{2}\epsilon_2, \frac{1}{2}\epsilon_2\right) B\left(\frac{1}{2}\epsilon_1, 2\epsilon_1 + 1\right) + \right. \\
&\quad \left. + 4a_3^2 B\left(\frac{1}{2}\epsilon_2, \frac{1}{2}\epsilon_2 + 1\right) B\left(\frac{3}{2}\epsilon_1, \epsilon_1 + 1\right) \right),
\end{aligned}$$

$$\begin{aligned}
I_{zz} &= \int \int \int_V (x^2 + y^2) dx dy dz = \tag{2.20} \\
&= \int_{-a_3}^{+a_3} \left(\int \int_{S(z)} (x^2 + y^2) dx dy \right) dz = \\
&= \frac{1}{2} a_1 a_2 a_3 \epsilon_1 \epsilon_2 (a_1^2 + a_2^2) B\left(\frac{3}{2}\epsilon_2, \frac{1}{2}\epsilon_2\right) B\left(\frac{1}{2}\epsilon_1, 2\epsilon_1 + 1\right).
\end{aligned}$$

Nekaj primerov vztrajnostnih momentov, izračunanih po zgornjih enačbah za kroglo, elipsoid in kocko, je prikazanih na tabeli 2.1.

Tabela 2.1: Vztrajnostni momenti posebnih primerov superelipsoidov. Povzeto po [15]

	<i>Krogla</i>	<i>Elipsoid</i>	<i>Kocka</i>
	$\epsilon_1 = 1, \epsilon_2 = 1$	$\epsilon_1 = 1, \epsilon_2 = 1$	$\epsilon_1 = 0, \epsilon_2 = 0$
I_{xx}	$\frac{8\pi}{15}r^5$	$\frac{4\pi}{15}abc(b^2 + c^2)$	$\frac{1}{12}abc(b^2 + c^2)$
I_{yy}	$\frac{8\pi}{15}r^5$	$\frac{4\pi}{15}abc(a^2 + c^2)$	$\frac{1}{12}abc(a^2 + c^2)$
I_{zz}	$\frac{8\pi}{15}r^5$	$\frac{4\pi}{15}abc(a^2 + b^2)$	$\frac{1}{12}abc(a^2 + b^2)$

2.5 Rekonstrukcija superelipsoidov z metodo gradi-in-izberi

V tem delu bomo grobo opisali metodo gradi-in-izberi, kot jo je podal Leonardis [17, 18] in je podrobno opisana v [15]. Kot že ime pove, sta bistvena gradnika metode gradnja modelov in izbira modelov, ki najboljše opisujejo dano množico točk.

Gradnja parametričnih modelov je težavno, ker mora rešiti dva problema:

1. najti množico elementov slike, ki pripadajo enemu modelu in
2. določiti vrednosti parametrov tega modela.

Za že segmentirane elemente slike lahko določimo vrednosti parametrov modela

$$\Lambda = (a_1, a_2, a_3, \epsilon_1, \epsilon_2, \phi, \theta, \psi, p_x, p_y, p_z) \quad (2.21)$$

z minimiziranjem funkcije napake

$$G(\Lambda) = a_1 a_2 a_3 \sum_{i=1}^N (F^{\epsilon_1}(x_i, y_i, z_i) - 1)^2, \quad (2.22)$$

če so $(x_i, y_i, z_i), i = 1 \dots N$ točke nekega segmenta slike, postavljene v kanonični koordinatni sistem superelipsoida. Metoda, ki sta jo razvila Solina in Bajcsy [23, 25], je praktično že standardna rešitev za problem rekostrukcije posameznih superelipsoidov iz globinskih slik.

Velja tudi obratno, če poznamo parametre modela, lahko s tehnikami klasifikacije vzorcev določimo kompatibilne elemente slike. Metoda gradi-in-izberi rešuje ta dva problema hkrati, z iterativnim kombiniranjem klasificiranja elementov slike in rekonstrukcije modelov.

2.5.1 Gradnja modelov

Eno glavnih vprašanj, ki bistveno vpliva na uspešnost celotne procedure, je, kako dobiti začetne opise (semena). Ena od možnosti je, da se točke, ki bi lahko pripadale enemu modelu, iščejo le na majhnem delu slike. Sliko tako mrežasto razdelimo na okna in potrebo po klasificiranju vseh točk nekega modela zmanjšamo na problem iskanja podmnožice, ki pripada istemu modelu. Seveda pa nimamo nikakršnih zagotovil, da bo vsako seme zraslo do dobrega opisa, ker bo nekaj semen lahko zgrajenih na območjih, ki pripadajo različnim modelom. Problem rešimo tako, da gradimo modele neodvisno iz vseh statistično konsistentnih semen in jih nato uporabimo kot hipoteze, ki lahko tvorijo končno rešitev.

Za začetna območja torej uporabimo majhna okna, ki jih mrežasto postavimo na sliko. Vsakemu območju prilagodimo model – superelipsoid. Če vse točke območja res pripadajo izračunanemu superelipsoidu, je seme dobro in ga vključimo v začetni opis. Ta odločitev temelji na primerjavi povprečne napake ujemanja

$$\bar{\xi}_i = \frac{1}{|R_i|} \sum_{\mathbf{x} \in R_i} d(\mathbf{x}, M_i) = \frac{1}{|R_i|} \xi_i \quad (2.23)$$

s pragom in ni kritična zaradi redundančne narave metode. Tako le odstranimo modele, ki ležijo na mejah med deli slike in zmanjšamo število začetnih semen.

Ko imamo začetno množico točk (seme), lahko določimo parametre modela. Če dosežemo dovolj veliko ujemanje med modelom in podatki (odvisno od zastavljene naloge), začnemo z iskanjem novih kompatibilnih točk, sicer model zavrnamo. Nove točke iščemo v bližini trenutnih mejnih točk območja točk R_i , ki pripada modelu M_i . Točke morajo biti sosedne (na sliki) mejnim točkam območja in dovolj blizu danega modela. Nove točke vključimo v množico območja in ponovno določimo parametre modela. Izračunamo novo povprečno napako ujemanja $\bar{\xi}_i$, pri čemer za razdaljo od točke do modela uporabimo radialno evklidsko razdaljo, in jo primerjamo s pragom. Na podlagi tega se odločimo za nadaljno rast območja oz. zaključimo proceduro.

Da bi preprečili nenadne spremembe v orientaciji superelipsoida zaradi izbire smeri z osi, uporabimo pri rekonstrukciji modela iz novega območja R'_i parametre modela M_i kot začetni približek za model M'_i . Ker pa lahko taka rekonstrukcija sili model v lokalni minimum, rekonstruiramo tudi model M''_i po običajnem postopku [23] iz R'_i . Uporabimo tisti model, ki se bolje prilega novemu območju R'_i .

Gradnja modelov poteka neodvisno (lahko paralelno) za posamezne modele. Procedura gradnje za nek model M_i nam na koncu poda tri rezultate:

1. množica elementov slike (točk) R_i , ki pripada modelu M_i ,
2. tip in parametre P_i modela M_i in
3. stopnjo ujemanja ξ_i med modelom M_i in območjem R_i .

Te nadalje uporabimo pri izbiri modelov.

2.5.2 Izbira modelov

Predstavitev, ki jo tvori procedura gradnje modelov, vsebuje odvečne modele, kar je neposredna posledica tega, da iščemo parametrične modele povsod po sliki. Več modelov je tako delno ali v celoti prekritih. Kombiniranje različnih modelov omejimo na proceduro izbiranja, ki pretehta mnoge konkurenčne rešitve in izbere tiste modele, ki tvorijo najpreprostejšo rešitev, to je tisto, ki upošteva največ točk in hkrati ohranja čim manjše razlike med točkami in modeli. To zmanjšanje kompleksnosti se ujema z intuitivnim pojmom preprostosti, ki ima dolgo tradicijo v psihologiji (Gestalt pricip). Formalizacija tega principa je v informacijski teoriji privedla do metode minimalne dolžine opisov (*Minimum Description Length*, *MDL*), ki se uporablja tudi pri računalniškem vidu.

Objektna funkcija, ki jo moramo maksimizirati, da bi dobili „najboljši“ opis z modeli, ima obliko

$$F(\mathbf{m}) = \mathbf{m}^T \mathbf{Q} \mathbf{m} = \begin{bmatrix} c_{11} & \cdots & c_{1N} \\ \vdots & & \vdots \\ c_{N1} & \cdots & c_{NN} \end{bmatrix}, \quad (2.24)$$

kjer je $\mathbf{m}^T = [m_1, m_2, \dots, m_N]$ vektor prisotnosti. Če je $m_i = 1$, je model vključen v končni opis, če je pa $m_i = 0$, pa modela ni v končnem opisu. Diagonalni elementi matrike \mathbf{Q} izražajo vrednost cena - dobiček za dani model M_i

$$c_{ii} = K_1 |R_i| - K_2 \xi_i - K_3 |P_i|. \quad (2.25)$$

K_1, K_2, K_3 so uteži, ki jih lahko določimo zgolj na osnovi informacijske teorije (v smislu bitov), ali pa upoštevamo tudi razmerje med signalom in šumom. Nediagonalni elementi upoštevajo prekrivanje med modeli

$$c_{ij} = \frac{-K_1 |R_i \cap R_j| + K_2 \xi_{ij}}{2}. \quad (2.26)$$

$|R_i \cap R_j|$ je število točk, ki jih opisujeta oba modela. ξ_{ij} popravlja vrednosti napake na diagonali v slučaju, če sta oba modela izbrana

$$\xi_{ij} = \max\left(\sum_{\mathbf{x} \in R_i \cap R_j} d^2(\mathbf{x}, M_i), \sum_{\mathbf{x} \in R_i \cap R_j} d^2(\mathbf{x}, M_j)\right). \quad (2.27)$$

Vrednosti napak ujemanja $d(\mathbf{x}, M_i)$ in $d(\mathbf{x}, M_j)$ izračunamo na točkah iz preseka območij obeh modelov in ustrezata odstopanjem do i -tega in j -tega modela.

Maksimizacija objektna funkcije $F(\mathbf{M})$ pripada razredu problemov kombinatorične optimizacije (*Quadratic Boolean problem*). Ker število možnih rešitev raste eksponentno z velikostjo problema, je jasno, da vseh možnosti ne moremo pregledati in zato poiščemo neko podoptimalno rešitev, ki jo lahko izračunamo v nekem razumnem času.

Zaradi same narave problema se izkaže, da lahko tako rešitev poiščemo že z direktno uporabo *požrešnega algoritma*, ki zaporedoma izbira možnosti, ki so lokalno optimalne. Modele izbiramo v vrstnem redu, ki ustreza njihovem doprinosu k objektni funkciji. Model, ki na neki stopnji maksimizira objektno funkcijo, uvrstimo med izbrane, če je objektna funkcija z njim kaj pridobila. Algoritem se ustavi, ko najboljši model ne izboljša objektno funkcije oz. ko so vsi modeli izbrani. Hitra implementacija požrešnega algoritma ima najslabšo časovno kompleksnost $O(N^2)$ [14].

2.5.3 Izmenjevanje gradnje in izbiranja modelov

Da bi dosegli računsko učinkovit postopek iterativno izmenjujemo gradnjo in izbiranje modelov. Gradnjo trenutno aktivnih modelov prekinemo z izbiranjem modelov, ki trenutno najboljše opisujejo sliko. Tako nam ni potrebno izgubljati časa z grajenjem modelov, ki jih zelo verjetno ne bi bilo v končnem opisu slike. Postopek ponavljamo, dokler ne rekonstruiramo preostalih modelov. Dobre in slabe lastnosti izmenjevanja obeh postopkov so opisane v [17].

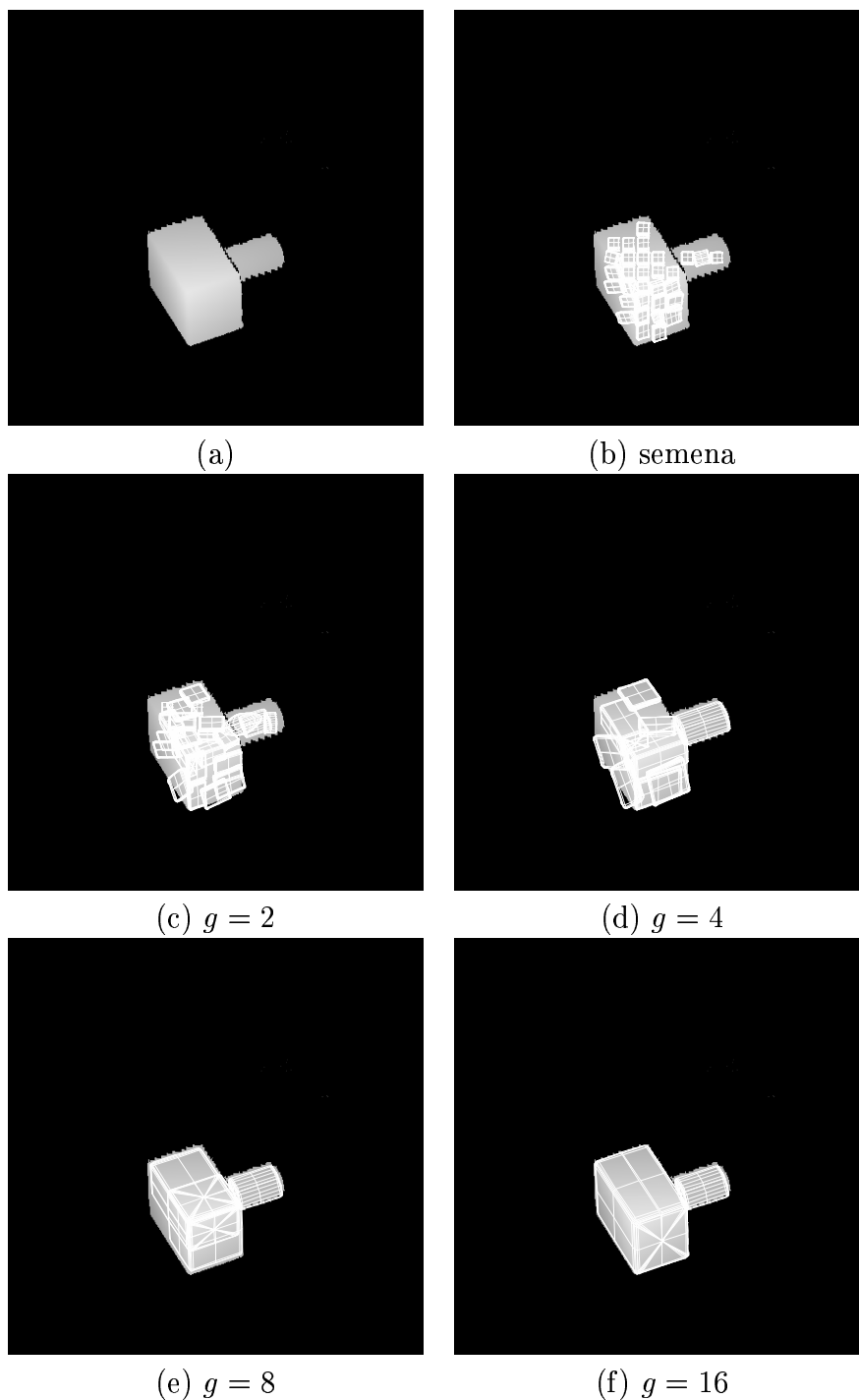
2.5.4 Prikaz poteka in uporaba opisov

Slika 2.2 prikazuje stopnje rekonstrukcije superelipsoidov z globinske slike z metodo gradi-in-izberi. Najprej se na globinsko sliko (slika 2.2(a)) postavijo semena opisov (slika 2.2(b)), sledi nekaj ponovitev gradnje modelov z vmesno izbiro najboljših opisov (slike 2.2(c) – 2.2(f)). Ko modeli ne morejo več rasti, dobimo končen opis prizora (slika 2.2(f)). Kot vidimo na vmesnih slikah 2.2(c) – 2.2(e), se opisi lahko prekrivajo. Predvsem pri objektih, ki se jih ne da popolnoma opisati s superelipsoidi, prekrivanje lahko ostane tudi v končnem opisu.

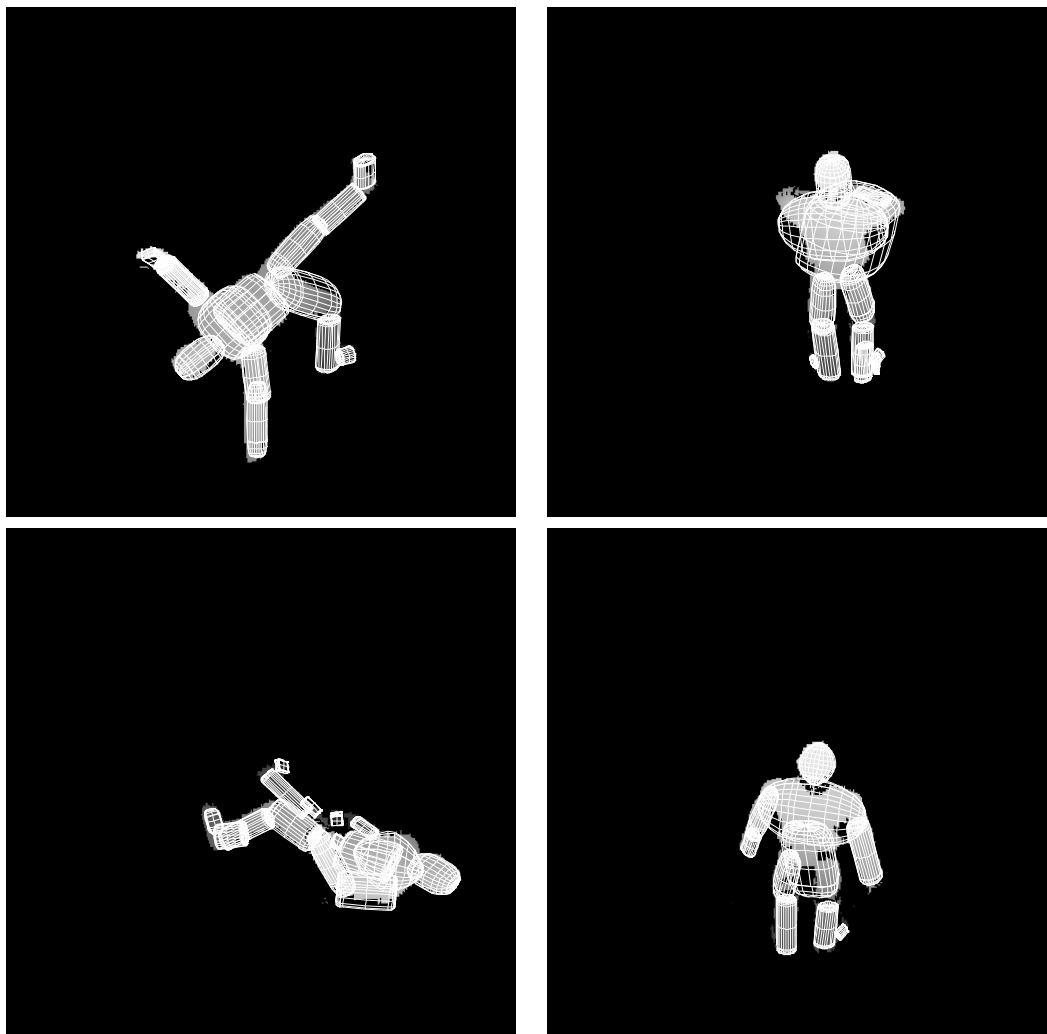
Rezultat rekonstrukcije globinske slike so opisi, ki sestojijo iz modela (superelipsoida) in ustrezne regije globinske slike, katero model opisuje. Cilj te magistrske naloge je bil razpoznavanje objektov na osnovi takih opisov.

Kot predmet razpoznavanja smo si izbrali človeško figurico, ki je podrobneje opisana v poglavju 3.2. Figurica je dokaj realistična, njenih sestavnih delov se ne da popolnoma opisati s superelipsoidi, zato rekonstrukcije vedno ne ustrezajo. Nekaj primerov rekonstrukcije takega objekta z metodo gradi-in-izberi prikazuje slika 2.3. Ponekod je nek del (npr. trup) opisan z večimi superelipsoidi, spet drugje pa en superelipsoid opisuje več sestavnih delov figure (npr. trup in nadlakti, trup in stegno, nogo ...).

Dlani in stopala so premajhna, da bi jih lahko z globinskim skenerjem dobro zajeli in nato dobili kolikor toliko dobre rekonstrukcije z metodo gradi-in-izberi. Zato jih v model objekta nismo vključili.



Slika 2.2: Primer rekonstrukcije superelipsoidov z metodo gradi-in-izberi, kjer se da predmet popolnoma opisati s superelipsoidi. (a) globinska slika, (b) postavljena semena, (c)-(f) po g korakih gradnje in izbiri. Po skupno 30 korakih gradnje z vmesnimi izbirami modeli ne morejo več rasti in dobimo končni opis (f).



Slika 2.3: Primer rekonstrukcije superelipsoidov z metodo gradi-in-izberi, kjer se predmeta ne da popolnoma opisati s superelipsoidi.

3

Razpoznavanje objektov

Faza razpoznavanja je verjetno najpomembnejša faza pri vsakem sistemu računalniškega vida. Sistem naj bi tako poiskal ujemanja med podatki, ki jih je pridobil z opazovanega prizora, in shranjenimi podatki, t.j. modeli, o objektih, ki naj bi jih razpoznal. Če so modeli shranjeni v obliki 3D geometrijskih modelov, tako kot v našem primeru, pa večina sistemov daje tudi informacijo o poziciji in orientaciji objekta. Tako lahko govorimo o identifikaciji in lokalizaciji objekta. Identifikacija govori o naravi oz. razredni pripadnosti objekta, lokalizacija pa objekt postavi v prostor. Meja med obema je pri večini sistemov dokaj zabrisana.

Razpoznavanje na podlagi rekonstrukcij globinskih slik s superelipsoidi je vsekakor problem iskanja ujemanja med deli s prizora (rekonstrukcije) in deli modela (*model-based matching*). Na osnovi analize rekonstrukcij globinskih slik objekta - človeške figure - smo se odločili v ta namen uporabiti interpretacijska drevesa. Metoda deluje v treh korakih:

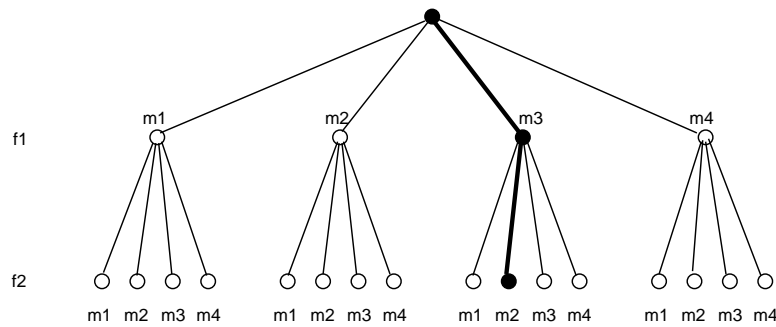
1. segmentacija globinskih slik in rekonstrukcija superelipsoidov z metodo gradi-in-izberi,
2. iskanje lokalno konsistentnih interpretacij z interpretacijskimi drevesi in
3. verifikacija interpretacije s pomočjo modela.

Drugi in tretji korak se prepletata, saj je potrebno z verifikacijo zavračati slabe interpretacije in tako priti do dobrih. V nadaljevanju bomo najprej predstavili interpretacijska drevesa, nato pa prilagoditev metode danemu problemu.

3.1 Interpretacijska drevesa

Interpretacijska drevesa so razred identifikacijskih algoritmov, katerih osnove ležijo v umetni inteligenci, posebno preiskovanju grafov in zadovoljevanju omejitev [12]. Interpretacija se oblikuje s potjo v prostoru vseh možnih parov slika-model, ki so predstavljena z drevesom. Primer interpretacijskega drevesa, ki prikazuje celoten

prostor iskanja za problem z dvema značilkama slike in štirimi značilkami modela, prikazuje slika 3.1.



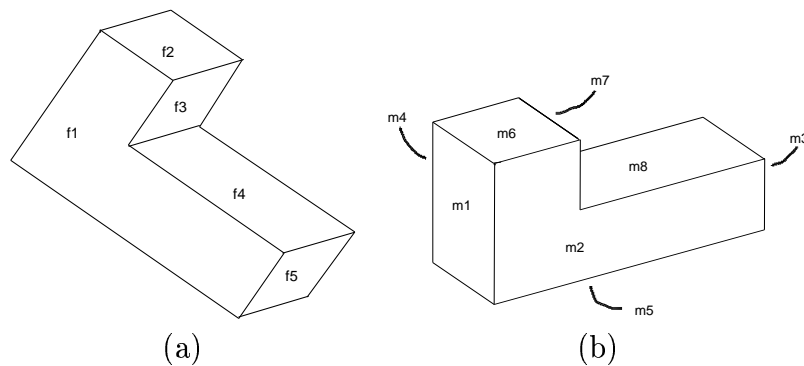
Slika 3.1: Interpretacijsko drevo, ki prikazuje celoten prostor iskanja za problem z dvema značilkama slike (f_1 , f_2) in štirimi značilkami modela (m_1 , m_2 , m_3 , m_4). Vsak nivo predstavlja vsa možna ujemanja med značilkami slike f_i in značilkami modela m_j . Temnejša pot tako prikazuje ujemanje (f_1, m_3) in (f_2, m_2).

Najprej predpostavimo, da so bile na sliki ugotovljene vse značilke, ki pripadajo določenemu razredu (npr. ploskve). Tako dobimo seznam simboličnih opisov značilnosti, ki so v bistvu numerične vrednosti lastnosti značilk (npr. površina, središče ploskve, itd.). Seveda lahko značilka modela ustreza značilki slike le, če so lastnosti obeh enaki (oz. v praksi dovolj podobni).

Pred razpoznavanjem potrebujemo tudi seznam značilk za znani objekt, torej model objekta. Seveda morajo biti lastnosti značilk modela in slik direktno primerljive in geometrijske. Potrebujemo tudi seznam geometrijskih omejitev, ki naj jih značilke modelov zadovoljijo. Značilke so lahko *unarne*, kot npr. ploščina, središče ali povprečna dolžina konture, *binarne*, ki nastopajo med parom značilk, kot npr. razdalja med središčema ali razlika med normalama ploskev, ali pa *n-arne*, ki povezujejo n delov. Omejitve označujejo značilke in njihove relacije ter nam omogočajo, da zavržemo ujemanja, ki so lokalno nekonsistentna.

Za razpoznavanje moramo najti preslikavo med značilkami modela in značilkami slike tako, da omejitve, ki veljajo za značilke modela, veljajo tudi za ustrezne značilke na sliki. Taki preslikavi pravimo *interpretacija*. Ta pokaže množico značilk slike, ki ustrezajo danemu modelu, kar je v bistvu pojavitev modela na sliki. Če ne moremo najti nobene interpretacije, lahko zaključimo, da se model ne pojavlja na sliki.

Problem je najti konsistentno interpretacijo, ne da preiščemo vse možne poti za ujemanje slike in značilk modela, kar bi bilo računsko prezahtevno. Za zmanjšanje kompleksnosti iskanja interpretacijska drevesa uporabljajo *zadovoljevanje omejitvam*. Interpretacija se oblikuje kot pot v prostoru vseh možnih ujemanj slike z modelom, ki so predstavljena z *interpretacijskim drevesom*. Najti interpretacijo pomeni najti pot po drevesu. Vsako vozlišče drevesa predstavlja hipotezo ujemanja, kot to prikazuje primer s slike 3.1.



Slika 3.2: (a) Idealna slika preprostega objekta z označenimi značilkami (ploskvami) slike (f_i). (b) 3D model istega objekta z označenimi značilkami modela (m_j). Ploskve m_1, m_5, m_8 so pravokotne, m_3, m_6, m_7 so kvadratne, m_2, m_4 pa sta v obliki črke L. Povzeto po [26]

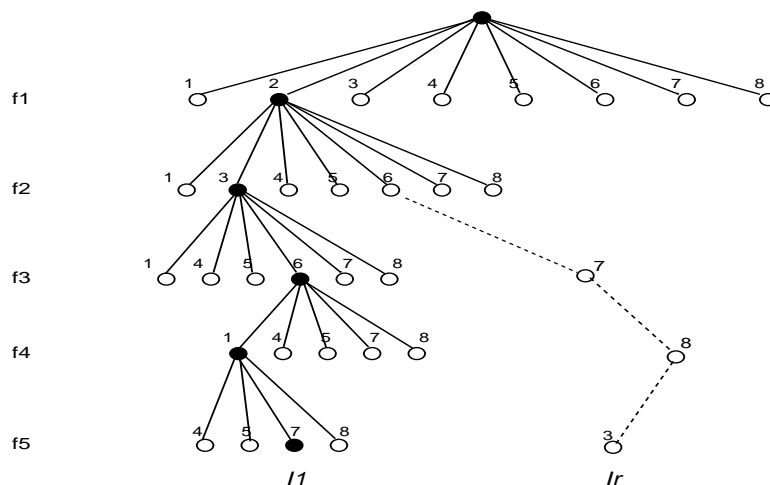
Poglejmo preprost primer, povzet po [26]. Slika 3.2(a) prikazuje idealno sliko predmeta in ploskve, ki smo jih identificirali, slika 3.2(b) pa 3D model objekta, ki temelji na ploskvah. Na sliki 3.2(a) je vidnih pet ploskev f_1, \dots, f_5 . Recimo, da je edina omejitev simboličnih opisov tip ploskve (kvadrat, pravokotnik, oblika črke L). Predpostavimo za zdaj tudi, da je na sliki en sam objekt. Del ustreznega interpretacijskega drevesa je na sliki 3.3. Na splošno velja, da vozlišča na nekem nivoju d predstavljajo vsa možna ujemanja za značilko f_d . Omejitve glede oblike pa izločijo nekaj nekonsistentnih ujemanj. Na primer značilka f_1 se lahko ujema le z m_2 ali m_4 . Drevo preiskujemo v globino, dokler ne dobimo ujemanja za vse značilke. Naslednji nivo v drevesu ne vsebuje značilk, ki se že ujemajo, saj lahko vsaka ploskev modela ustreza le eni ploskvi s slike. Prva konsistentna interpretacija, ki jo najdemo, je

$$I_1 = \{(f_1, m_2), (f_2, m_3), (f_3, m_6), (f_4, m_1), (f_5, m_7)\}.$$

Toda ta interpretacija je napačna. Zakaj? Ker so uporabljene omejitve povezane z lokalnimi značilkami in tako lokalne po naravi. Vidimo lahko, da imata npr. ploskvi f_2 in m_3 (ali f_3 in m_6 ali f_4 in m_1 ali f_5 in m_7) sicer enake lastnosti (tip ploskve), vendar pa ploskev f_2 ne ustreza ploskvi m_3 na modelu. Če imamo na voljo 3D informacijo, lahko globalno konsistentnost interpretacije preverimo npr. tako, da izračunamo transformacijo, ki preslika ujemajočo ploskev modela v ploskev objekta s slike, projeciramo model na sliko in preverimo, če so vse ploskve modela na pričakovanih mestih. Ko interpretacijo I_1 preverimo, ugotovimo, da je nekonsistentna, in s sledenjem nazaj skušamo dobiti dobro interpretacijo. Po še nekaj nekonsistentnih interpretacijah dobimo pravo:

$$I_r = \{(f_1, m_2), (f_2, m_6), (f_3, m_7), (f_4, m_8), (f_5, m_3)\}.$$

Pri preverjanju globalne konsistentnosti interpretacije, čemur bomo v nadaljevanju zaradi jasnosti rekli verifikacija, pa moramo biti pozorni. V primerih, kot je npr. zgornji, lahko interpretacijo verificiramo z N -arnimi omejitvami, ki omejujejo N delov z neko relacijo. Vendar pa nam v splošnem nič ne zagotavlja, da bodo vse interpretacije globalno smiselne.



Slika 3.3: Del interpretacijskega drevesa za problem s slike 3.2. Temnejša vozlišča predstavljajo prvo najdeno interpretacijo, I_1 , črtkana linija pa pravilno interpretacijo.

3.1.1 Navidezna ujemanja in lažne značilke

Primer je nerealen v vsaj dveh točkah. Kot prvo, realne slike pogosto vsebujejo več pojavitev objektov; seveda bi želeli razpoznati vse. Interpretacijsko drevo bi moralo odgovoriti na vprašanje: "je na sliki kakšen X?". Kot drugo pa ujemanje značilke ni nujno enolična preslikava. Nekatere značilke objekta so lahko na sliki skrite (kot m_1, m_4, m_5 na sliki 3.2(a)), druge značilke pa so lahko delno zakrite z drugimi objekti. Detektor značilk jih lahko nekaj ne najde, čeprav so prisotne na sliki, ali pa jih najde tam, kjer jih na sliki v resnici ni. Želeli bi razpoznati objekt, tudi če nekaj značilk manjka in tudi če je nekaj značilk lažnih.

Za lažne značilke lahko poskrbimo z navidezno značilko, ki se lahko ujema z vsako značilko s slike, za katero ne najdemo pravega ujemanja. V praksi le dodamo vozlišče z navidezno značilko v seznam značilk modela. Brez navidezne značilke bi se iskanje nadaljevalo z dviganjem (sledenjem nazaj), brž ko bi naleteli na lažno značilko in tako bi preskočili morebitne konsistentne interpretacije. To se z navidezno značilko ne bi zgodilo, vendar pa uvedba navidezne značilke močno poveča kompleksnost problema. Če ima npr. slika P pravih in L lažnih značilk, iskanje z navidezno značilko najde pravo interpretacijo, v kateri se vseh P pravih

značilke ujema, vendar pa najde tudi interpretacije, pri katerih se ujema manj kot P pravih značilke, saj se tudi prave značilke ujemajo z navidezno značilko.

Ena od metod za zmanjšanje nepotrebnega iskanja je metoda *veji-in-omeji*. Če iščemo vse pojavitve objekta na sliki in najdemo prvo interpretacijo, ki jo preverimo in je prava, zabeležimo p , število pravih značilke, ki jih vsebuje. S sledenjem nazaj želimo nato dobiti še druge interpretacije, vendar pa pot zavržemo, ko je jasno, da ne bomo mogli doseči več kot p realnih ujemanj. Če je npr. $p = 6$, imamo v trenutni interpretaciji dve pravi ujemanji in moremo dobiti še največ tri ujemanja, lahko to interpretacijo zavržemo in nadaljujemo s sledenjem nazaj, ker ne moremo doseči več kot 6 pravih značilke. Seveda pa moramo osvežiti vrednost p , ko dobimo interpretacijo, ki vsebuje $p' > p$ pravih ujemanj.

Algoritem INT_DREVO

```

WHILE (obstaja še možnih ujemanj)
  IF (naslednje ujemanje konsistentno AND
      največja možna velikost interpretacije na tej poti  $\geq MaxVel$ 
      dodaj ujemanje v Interp);
  IF (dosežen je list)
    IF (Interp je OK)
      shrani Interp;
       $MaxVel =$  velikost (Interp)
    ENDIF
  ELSE /* ni list, je konsistentno */
    razširi vozlišče z vsemi možnimi ujemanji na naslednjem nivoju
  ENDIF
ENDWHILE

```

Tabela 3.1: Algoritem za razpoznavanje z interpretacijskim drevesom.

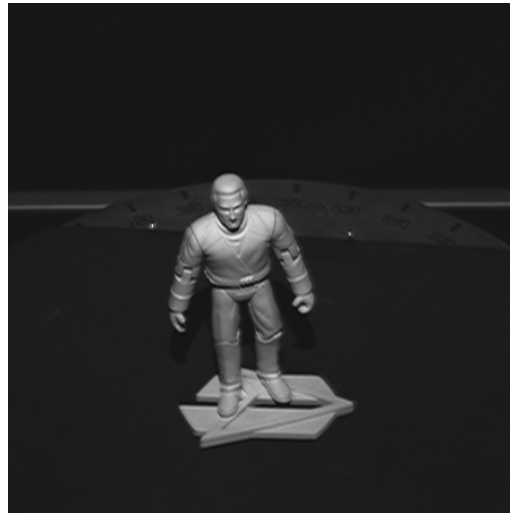
Čeprav je dokaj preprost, deluje algoritem INT_DREVO [26] s tabele 3.1 na realnih prizorih. Upošteva manjkajoče značilke, lažne oz. šumne značilke in večkratne pojavitve objekta. Navidezna ujemanja povečajo kompleksnost, ki so eksponentna glede na število značilke modela in slike. INT_DREVO uporablja metodo *veji-in-omeji* za zmanjšanje tega problema. INT_DREVO hkrati locira objekt in ga identificira: izbere značilke, ki najverjetneje pripadajo objektu in preveri identiteto. Neizogibna cena je velika kompleksnost algoritma. *Poravnalne* ali *hibridne metode* so drug način za zmanjšanje iskanja. Te metode najdejo ujemanje le za toliko značilke slike, kolikor jih potrebujemo za preverjanje identitete. Algoritem INT_DREVO preiskuje drevo najprej v globino, seveda pa lahko preiskujemo tudi najprej v širino ali pa najprej najboljšega. Pri slednjem potrebujemo pri vsakem vozlišču oceno kvalitete ujemanja, pot po drevesu pa nadaljujemo z vozliščem, ki

ima najkvalitetnejše ujemanje. Ocena kvalitete ujemanj lahko temelji npr. na podobnosti med ujemačimi deli.

Interpretacijska drevesa imajo težave s prizori z več objekti. Če ne moremo preveriti globalne konsistentnosti, kar drži pri številnih realnih aplikacijah, so rezultati nezanesljivi. Z uvedbo navideznih ujemanj, ki jih v realnih aplikacijah potrebujemo, pa povečamo kompleksnost algoritma.

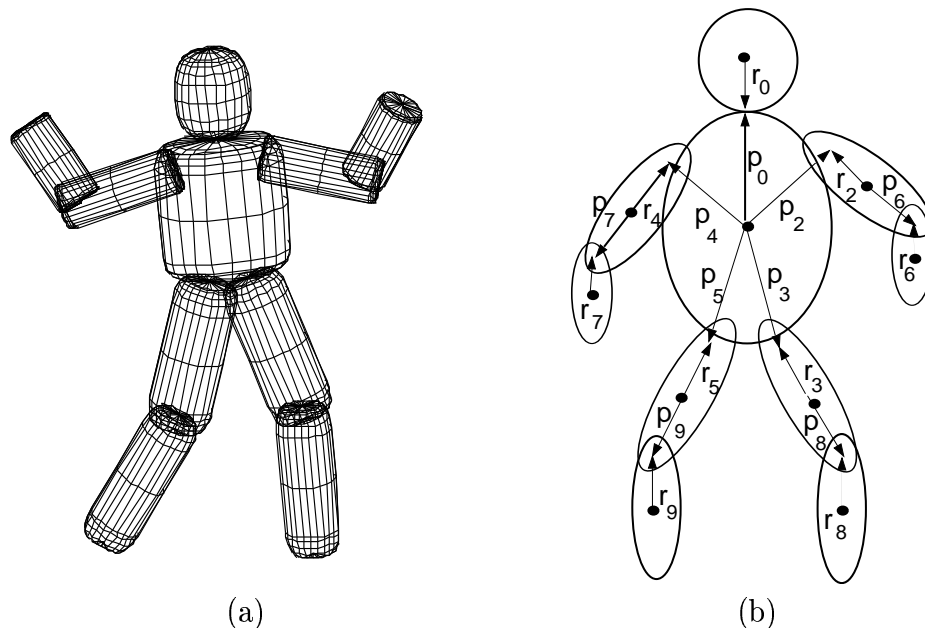
3.2 Model objekta

Kot objekt, ki naj ga naš sistem razpozna, smo si izbrali človeško figurico s slike 3.4. Njeni najvažnejši deli so glava, roke, noge in trup. Roke in noge so gibljive in tako lahko dobimo različne poze.



Slika 3.4: Intenzitetna slika človeška figurice, ki smo jo uporabili kot objekt za razpoznavanje.

Figurico smo modelirali s superelipsoidi, kot kaže slika 3.5(a). Figurico smo intuitivno razdelili na posamezne dele: glavo in trup ter po dve nadlakti, podlakti, stegni in goleni. Dlani in stopala smo pri modelu zanemarili. Vsak sestavni del je opisan s superelipsoidom, torej s parametri za velikost in obliko. Osrednji del je trup, kateremu določimo tudi pozicijo in orientacijo v prostoru. Na trup so pritrjeni glava in zgornji deli okončin, nanje pa še spodnji. Tako imamo za te dele podane še podatke o sklepu in sicer pozicijo sklepa relativno glede na središče samega dela (\vec{r}_i) in njegov položaj relativno glede na center njemu „nadrejenega“ dela (\vec{p}_i), kot to prikazuje slika 3.5(b). Vsem gibljivim delom (okončinam) lahko določimo parametre rotacije okoli sklepa. Vrednosti posameznih parametrov za dele objekta so prikazani na tabeli 3.2. Določili smo jih z meritvami ustreznih količin direktno na objektu.



Slika 3.5: Model človeške figurice: (a) model objekta, sestavljen iz superelipsoidov, in (b) struktura modela, ki prikazuje relacije med posameznimi deli modela.

Za izbiro takega objekta imamo več razlogov. Ker je dokaj realističen, se njegovih sestavnih delov ne da popolnoma opisati s superelipsoidi, zato rekonstrukcije vedno ne ustrezajo. Ponekod je nek del (npr. trup) opisan z večimi superelipsoidi, ki se deloma prekrivajo, spet drugje pa en superelipsoid opisuje več sestavnih delov objekta (npr. celotno nogo). Možnost transformacije figurice v različne poze še poveča te težave.

Težave povzročata tudi razdelitev na dele, ker ne ustreza povsem rekonstrukcijam, ki jih dobimo z metodo gradi-in-izberi. Razlika je predvsem pri trupu, kjer dobimo

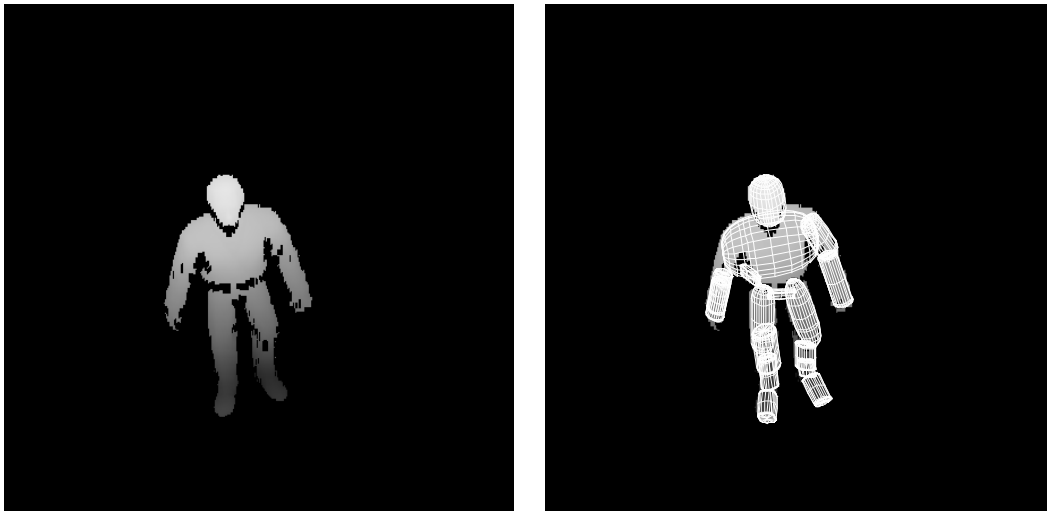
$\tilde{S}t.$	Del	a_1	a_2	a_3	ϵ_1	ϵ_2	\vec{p}_i	\vec{r}_i
0	glava	8	8	10	0.7	1.0	$[0, 0, 15]^T$	$[0, 0, -10]^T$
1	trup	14	10	15	0.3	0.9	-	-
2	nadlakt ₁	5	5	13	0.1	1.0	$[15, 0, 8]^T$	$[0, 0, 7]^T$
3	nadlakt ₂	5	5	13	0.1	1.0	$[-15, 0, 8]^T$	$[0, 0, 7]^T$
4	stegno ₁	7	7	17	0.3	1.0	$[5, 0, -22]^T$	$[0, 0, 7]^T$
5	stegno ₂	7	7	17	0.3	1.0	$[-5, 0, -22]^T$	$[0, 0, 7]^T$
6 in 7	podlakt _x	5	5	10	0.1	1.0	$[0, 0, -9]^T$	$[0, 0, 10]^T$
8 in 9	goleno _x	6	6	17	0.3	1.0	$[0, 0, -14]^T$	$[0, 0, 17]^T$

$x = \{1, 2\}$

Tabela 3.2: Parametri, ki določajo model objekta s slike 3.4.

superelipsoid v zgornjem delu, ki opisuje regijo pleč in vsaj delno tudi nadlakti, včasih pa tudi superelipsoid v spodnjem, trebušnem delu, ki lahko delno opisuje tudi regijo stegen. Primer vidimo na sliki 3.6(b): na trupu je superelipsoid, ki zajema tudi del leve nadlakti in večino desne nadlakti, tako da so superelipsoidi, ki so opisovali desno nadlakt, v fazi izbire izpadli. Glava je tipično dobro rekonstruirana. Na predelu okončin lahko ločimo več možnih rekonstrukcij:

1. superelipsoid s trupa sega na področje okončine in opisuje zgornji del okončine (npr. desna roka s slike 3.6(b)).
2. en sam superelipsoid opisuje celotno okončino (npr. če je iztegnjena in pravilno obrnjena).
3. dva superelipsoida pravilno opisujeta okončino (npr. leva roka s slike 3.6(b)).
4. več superelipsoidov, ki se deloma prekrivajo, opisuje okončino (npr. desna noga s slike 3.6(b)).



(a)

(b)

Slika 3.6: Primer rekonstrukcije globinske slike človeške figurice s superelipsoidi: (a) globinska slika objekta, (b) primer rekonstrukcije objekta s superelipsoidi, ki je prikazana na globinski sliki.

3.3 Posebnosti pri uporabi superelipsoidov

Za razpoznavanje našega objekta smo uporabili algoritem za preiskovanje interpretacijskih dreves, ki ga prikazuje tabela 3.3. Da bi pohitrili iskanje dobrih rešitev,

smo glede na prej opisan algoritem v tabeli 3.1 obrnili ujemanja. Na nekem nivoju drevesa tako iščemo ujemanja za značilko modela. Vozlišča na nivoju d tako predstavljajo vsa možna ujemanja za značilko modela m_d (in ne značilko slike f_d kot prej).

Rekonstrukcija glave je najbolj stabilna. Zato za večino rekonstrukcij velja (glej poglavje 3.3.1), da lahko ujemanje superelipsoida z modelom glave dokaj dobro preverimo. Tako je smiselno najprej preveriti interpretacije, pri katerih imamo realna ujemanja za glavo. Podobno velja tudi za trup in okončine.

Pri preiskovanju interpretacijskega drevesa je bil za hitro doseganje rešitev pomemben vrstni red delov objekta. Kot že omenjeno, smo lahko v našem primeru najbolje določili potencialne glave, v povezavi glave s trupom pa lahko objektu približno določimo pozicijo v prostoru. Vrstni red delov, ki hkrati ustreza globini drevesa, prikazuje tabela 3.4.

Zaradi pohitritve samega iskanja je bilo smiselno poiskati ujemanja samo za glavo, trup, nadlakti in stegna, saj bi drugače preveč povečali kompleksnost iskanja rešitev. Ujemanja za podlakti in goleni poiščemo pri preverjanju interpretacije, kot je opisano v poglavju 3.3.2. Upoštevali smo tudi omejitve, ki sledijo iz same strukture objekta in vrstnega reda iskanja ujemanj. Če interpretacija nima realnega ujemanja za prvo nadlakt, velja, da tudi za drugo roko ni realnega ujemanja, ker je objekt simetričen in so bila vsa možna realna ujemanja za prvo nadlakt že preizkušena. Podobno velja za drugo stegno: če v interpretaciji ni realnih ujemanj za prvo nadlakt in prvo stegno, so bila vsa možna ujemanja za stegno že preizkušena. V teh primerih vozlišč ne razširjamo z realnimi ujemanji, temveč le z navideznim.

Ko se algoritem 3.3 konča, dobimo na izhodu interpretacijo, ki je najdaljša, kar nekako pomeni, da najbolje določa pojavitev objekta v prizoru. Če je v prizoru več pojavitev objekta, dobimo z enkratnim izvajanjem algoritma le najboljšo. Če potem iz prizora odstranimo dele (superelipsoide), ki so v dobljeni interpretaciji, lahko s ponovnim izvajanjem dobimo naslednjo najboljšo interpretacijo. Postopek ponavljamo, dokler algoritem še daje dobre (preverjene) interpretacije.

3.3.1 Preverjanje konsistentnosti ujemanja

Rekonstrukcije objekta v različnih pozah in z različnih gledišč se bistveno razlikujejo med seboj, razen v primeru glave. Za primerjavo dveh superkvadrikov med seboj tudi sicer ne moremo direktno primerjati njihovih parametrov [15], zato smo pri preverjanju konsistentnosti ujemanja del modela - del s prizora kot primarno lastnost uporabili prostornino superelipsoida. Za odgovor na vprašanje ali se del s prizora in objekta ujemata, smo uporabili preprosto primerjavo volumnov posameznih delov na njih, kar kaže tabela 3.5. Vrednosti volumnov smo določili na osnovi tridesetih rekonstrukcij.

Uporabljen algoritem

Sklad - sklad vozlišč za razširitev

Interp - seznam konsistentnih ujemanj, ki tvorijo interpretacijo

MaxVel - največje število realnih ujemanj do zdaj

velikost(Interp) - število realnih ujemanj v *Interp*

koren - oznaka za koren drevesa

konsistentna(X) - če je *X* realno ujemanje, vrne DA, če je *X* konsistentno z omejitvami, ali $X = koren$, sicer NE. Navidezno ujemanje je po definiciji konsistentno z vsemi omejitvami.

Sklad = [*koren*], *Interp* = [], *MaxVel* = 0;

WHILE (*Sklad* ni prazen)

 vzemi s *Sklada* naslednje ujemanje $X = (f_j, m_k)$

 IF (konsistentno(*X*) AND

 največja možna velikost interpretacije na tej poti $\geq MaxVel$)

 dodaj *X* v *Interp*;

 IF (dosežen je list)

 preveri *Interp*;

 IF (*Interp* je v redu)

 shrani *Interp*;

MaxVel = velikost (*Interp*)

 ENDIF

ELSE /* ni list, je konsistentno */

 dodaj na sklad (*W*, m_{k+1})

 FOR $i = 1..N$

 IF (del f_i še ni v interpretaciji)

 dodaj na sklad (f_i , m_{k+1})

 ENDIF

 dodaj *L* na levo stran *Vozli*;

 ENDIF

ELSE /* nekonsistentno ujemanje */

 nadaljuj;

 ENDIF

ENDWHILE

Tabela 3.3: Uporabljen algoritem za razpoznavanje z interpretacijskim drevesom [26]

<i>Globina drevesa</i>	1	2	3	4	5	6
<i>Ustrezni del objekta</i>	glava	trup	nadlakt ₁	stegno ₁	nadlakt ₂	stegno ₂

Tabela 3.4: Deli modela objekta, ki ustrezajo ujemanjem na različnih globinah drevesa.

<i>Del</i>	<i>Min. volumen</i>	<i>Max. volumen</i>
glava	1000	6000
trup	5500	18000
trup (*)	3500	13000
nadlakti	300	6500
podlakti	300	3500
stegna	500	8000
goleni	300	3200

Tabela 3.5: Intervali volumnov za dele objekta.

Ujemanje glave. Z analizo rekonstrukcij objekta smo ugotovili, da je najbolj dosledno rekonstruiran del objekta glava, ki je hkrati glede na trup edini praktično negibljev del. Zato je bilo poleg volumna smiselno ugotoviti tudi omejitve za vrednosti drugih parametrov, s katerimi bi izločili čim več neprimernih delov. Te omejitve so prikazane v tabeli 3.6. Vrednosti smo določili na podlagi tridesetih rekonstrukcij.

ϵ_1	ϵ_2	$\min(a_1, a_2, a_3)$	$\text{middle}(a_1, a_2, a_3)$	$\max(a_1, a_2, a_3)$
0.5 – 1.4	0.4 – 1.5	3.0 – 9.0	5.0 – 11.0	6.5 – 15.0

Tabela 3.6: Omejitve za parametre superelipsoida, ki opisuje glavo.

Ujemanje trupa. Za superelipsoide, rekonstruirane na predelu trupa, velja, da se nekako najbolj razlikujejo od modela trupa. Rekonstrukcija lahko tam zgradi več superelipsoidov, ki se lahko raztezajo tudi čez prostor okončin. Za ugotavljanje prisotnosti trupa smo uporabili preprosto oceno z volumnom:

- če naj bo nek superelipsoid s prizora trup, mora biti vsota volumnov superelipsoidov, katerih središča so v neki oddaljenosti d od njegovega središča v dovoljenem intervalu (glej tabelo 3.5). Oddaljenost d smo določili kot polmer največje krogle, ki jo lahko včrtamo v model trupa, torej $d = 12.0$.
- če v taki oddaljenosti ni nobenega drugega superelipsoida, velja interval označen z * v tabeli 3.5.

Primeri obravnavamo ločeno (z različnimi vrednostmi), ker ne upoštevamo preseka dveh superelipsoidov, ki ga tako štejemo dvakrat. Kot približek lahko središče trupa \vec{c}_b' določimo kot povprečje središč zaobjetih superelipsoidov.

Če je poleg trupa prisotna tudi glava, lahko natančneje izračunamo parametre pozicije in dodatno še (nepopolne) parametre orientacije za celoten model, kar nam v nadaljevanju služi za omejevanje iskanja oz. pri določanju konsistentnosti ujemanja. Medtem ko so središča superelipsoidov, ki jih rekonstruiramo na območju trupa, dokaj blizu prave vertikalne osi trupa, pa je razlika do samega središča večja. To popravimo z upoštevanjem razdalje med središčema glave in trupa modela. Orientacijo dobimo tako, da izračunamo parametre rotacije, ki z -os preslika v smerni vektor $c_h - c_b$. Pri tem ne določimo parametra ψ_b , ki določa zasuk okoli objektive z osi.

Naj bo \vec{c}_h središče superelipsoida, ki opisuje glavo, in \vec{c}_b' zgoraj omenjeni približek središča trupa. Potem je boljši približek središča trupa \vec{c}_b :

$$\vec{c}_b = \vec{c}_h - |\vec{s}_h| \vec{s} \quad (3.1)$$

$$\phi_b = \arctan \frac{-s.x}{-s.y} \quad (3.2)$$

$$\theta_b = -\arctan \frac{\sqrt{s.x^2 + s.y^2}}{s.z}, \quad (3.3)$$

pri čemer je

$$\begin{aligned} \vec{s}_h &= \vec{r}_h - \vec{p}_h \\ \vec{s} &= \frac{\vec{c}_h - \vec{c}_b'}{|\vec{c}_h - \vec{c}_b'|} \end{aligned} \quad (3.4)$$

\vec{c}_b' , \vec{c}_b in \vec{c}_h so 3D vektorji oblike $[x, y, z]^T$, ϕ_b , θ_b in ψ_b pa so koti, ki določajo orientacijo trupa in s tem celotne figurice v prostoru.

Ujemanje okončin. Pri ugotavljanju konsistentnosti ujemanja delov za nadlakti oz. stegni smo razlikovali dva primera:

1. v interpretacijo sta vključeni realni ujemanji tako za glavo kot tudi za trup.
2. vsaj eno od zgornjih ujemanj je navidezno.

V prvem primeru imamo že izračunane parametre pozicije in (delno) orientacije za celoten model, kar lahko uporabimo pri preverjanju konsistentnosti posameznega ujemanja. Ko se nek superelipsoid s prizora po volumnu ujema z nadlaktjo oz.

stegnom, lahko izračunamo parameter ψ_b , ki obrne model objekta v prostoru tako, da se (približno) stakneta sklepa modela objekta in morebitne nadlakti oz. stegna:

$$\begin{aligned}\psi_b &= \arctan \frac{s.y}{s.x} \\ \vec{s} &= R^{-1}(\vec{c}_x - \vec{c}_b),\end{aligned}\quad (3.5)$$

pri čemer je \vec{c}_x središče superelipsoida, katerega ujemanje z nadlaktjo oz. stegnom preverjamo, R je rotacijska matrika s parametri $(\phi_b, \theta_b, 0)$. Če naj bo obravnavani del res okončina, mora biti njegovo središče (približno) toliko oddaljeno od sklepa, kot je to primer pri modelu objekta:

$$\begin{aligned}m_1|r_x| &< |\vec{s}| < m_2|r_x| \\ \vec{s} &= \vec{p}_x - R^{-1}(\vec{c}_x - \vec{c}_b).\end{aligned}\quad (3.6)$$

Skalarjema m_1 in m_2 smo določili vrednosti $m_1 = 0.25$ in $m_2 = 2.1$ na podlagi rekonstrukcij. Ustreznemu delu izračunamo še parametre rotacije z

$$\phi_x = \arctan \frac{-s.x}{-s.y} \quad (3.7)$$

$$\theta_x = -\arctan \frac{\sqrt{s.x^2 + s.y^2}}{s.z}, \quad (3.8)$$

$$\psi_x = 0. \quad (3.9)$$

V drugem primeru pa nimamo informacije o poziciji in orientaciji trupa (oz. samo o orientaciji, če je v interpretaciji realno ujemanje za trup). V tem primeru preverimo, ali je del, ki ga želimo vključiti v interpretacijo, približno toliko oddaljen od delov, ki so že vključeni v interpretacijo, kot ustrezni deli z modela. Ujemanje (x, y) (x je del s prizora, y del objekta) je nekonsistentno, če vsaj pri enem ujemanju (u, v) , ki je že vključeno v interpretacijo ne velja

$$|\vec{p}_y - \vec{p}_v| - m_3(|\vec{r}_y| + |\vec{r}_v|) < |\vec{c}_x - \vec{c}_u| < |\vec{p}_y - \vec{p}_v| + m_4(|\vec{r}_y| + |\vec{r}_v|), \quad (3.10)$$

pri čemer sta \vec{c}_x in \vec{c}_u središči delov s prizora x in u . Skalarjema m_3 in m_4 smo določili vrednosti $m_3 = 0.5$ in $m_4 = 1.5$ na podlagi rekonstrukcij.

Spodnjih delov okončin z interpretacijskim drevesom nismo preverjali, ker lahko že brez njih dobro verificiramo interpretacijo in bi takó iskanje spodnjih delov okončin le povečalo prostor iskanja.

3.3.2 Verifikacija interpretacije

Preverjanje globalne konsistentnosti pomeni, da sistem odgovori na vprašanje „ali dana množica delov res sestavlja objekt X ?“ Najprej lahko zavrnamo interpretacije, ki vsebujejo premalo realnih ujemanj. Interpretacija, ki ima le navidezna ujemanja, nam o pojavitvi objekta ne pove nič. Interpretacija z enim realnim ujemanjem

pomeni, da se le en del s modela ujema z enim delom s slike. V našem primeru smo se odločili, da interpretacija, ki ima manj kot štiri realna ujemanja, ni dobra. S tem morda zavrnilo tudi pravilne interpretacije, vendar pa zavrnilo več nepravilnih, saj je manjša verjetnost, da bo več delov „slučajno” razporejenih v podobni strukturi, kot je značilna za človeško figuro.

Pri številnih realnih aplikacijah globalne konsistentnosti ne moremo preveriti. V našem primeru, ko za verifikacijo interpretacije uporabljamo superelipsoide oz. njihove lastnosti, naletimo na težave. Vzrok je predvsem v tem, da uporabljamo superelipsoide za opis (dokaj) realnega sveta, torej takega, ki se ga ne da popolnoma opisati s superelipsoidi. Rekonstrukcije superelipsoidov so nestabilne, kar pomeni, da moramo za posamezne dele modela iskati ujemanja pri zelo različnih superelipsoidih ali za en del celo pri večih superelipsoidih. Za morebitne pojavitve objekta nam je torej bolj kot lastnosti posameznih superelipsoidov pomembnejša geometrijska struktura, ki jih „povezuje”. Zato smo za verifikacijo interpretacije uporabili tiste regije na globinski sliki, ki pripadajo superelipsoidom iz interpretacije. Na vsako tako regijo smo z metodo za rekonstruiranje posameznih superelipsoidov [23] poskušali prilagoditi ustrezen superelipsoid iz modela tako, da smo minimizirali le po parametrih pozicije in orientacije. Ostali parametre superelipsoida (za velikost in obliko) pa smo pustili enake kot pri superelipsoidu modela.

Tako kot pri preverjanju konsistentnosti ujemanja, lahko tudi pri verifikaciji interpretacije ločimo primera, ko imamo za glavo in trup realni ujemanji, in primer, ko za glavo ali trup ni realnega ujemanja.

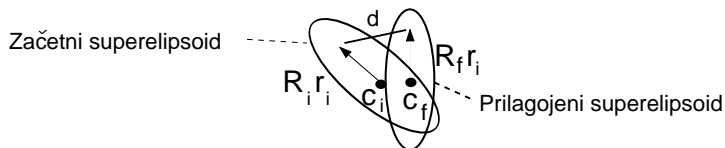
V prvem primeru lahko preprosto določimo parametre pozicije in orientacije celotnega objekta (glej enačbo 3.1). Izkaže se, da tako izračunani parametri dokaj natančno opisujejo položaj figurice. V drugem primeru bi lahko na podlagi pozicije sklepov, ki okončine povezujejo s trupom, izračunali parametre pozicije in orientacije za celoten objekt. Vendar moramo najprej izračunati pozicije sklepov, kar pa na osnovi rekonstrukcije ni ravno preprosto oz. natančno. Zato smo v nadaljevanju predpostavili, da je interpretacija dobra le, če sta ujemanji za glavo in trup realni. Problemi in morebitne rešitve, ko temu ni tako, pa so predstavljeni v poglavju 4.4.

Vrnimo se k verifikaciji interpretacije s prilagajanjem modela regiji globinske slike. Za začetni približek smo uporabili parametre superelipsoida modela, kot jih lahko izračunamo (glej prejšnje poglavje, 3.3.1). To smo uporabili le na superelipsoidih, ki naj bi se ujemali z okončinami, saj so le te podolgovate in je zato prilagojen model dobro določil pravo lego sklepa. Ko dobimo parametre prilagojenega superelipsoida, lahko izračunamo razdalje med položajem sklepa na prilagojenem in začetnem superelipsoidu:

$$d = |(R_i \vec{r}_i + \vec{c}_i) - (R_f \vec{r}_i + \vec{c}_f)|, \quad (3.11)$$

pri čemer je R_i rotacijska matrika, ki del i zavrti v začetni položaj, r_i je položaj sklepa v delu, \vec{c}_i je središče začetnega superelipsoida, R_f rotacijska matrika, ki del i zavrti v prilagojeni položaj in \vec{c}_f središče prilagojenega superelipsoida. Situacijo

prikazuje slika 3.7. Dobljeno razdaljo primerjamo s pragom p in če je vsaj pri eni okončini ta večja od praga, interpretacijo zavrnilo kot neveljavno. Vrednost praga $p = 11.0$ smo določili na podlagi rekonstrukcij.



Slika 3.7: Prikaz količin z enačbe 3.11.

Opazimo lahko, da bi zgornje sklepanje lahko uvrstili v preverjanje konsistentnosti ujemanja med delom iz modela in prizora, saj preverjamo le posamezne dele. Tako bi lahko preverjanje uvrstili v preverjanje konsistentnosti ujemanja, vendar pa za preverjanje razdalje med morebitnimi sklepi potrebujemo prilagajanje superelipsoida na regijo globinske slike, kar je časovno zelo potratno. Tako v primeru, da se interpretacija izkaže za nekonsistentno zaradi premajhnega števila realnih ujemanj, ne bi bilo smiselno računati prilagajanja. Po drugi strani, pa lahko v primeru, da prilagajamo superelipsoide v fazi verifikacije, to uporabimo za zmanjšanje prostora iskanja. Nekonsistentnost pri nekem ujemanju (m_i, f_j) (pri čemer so ujemanja za m_0 do m_{i-1} konsistentna) pomeni, da lahko iskanje nadaljujemo z naslednjim ujemanjem na nivoju i drevesa.

Spodnji deli okončin. Ko tako dobimo neko verodostojno interpretacijo, moramo preveriti še prisotnost delov, ki jih z interpretacijskim drevesom nismo, torej podlakti in goleni. Prisotnost teh delov preverjamo le, če je prisoten ustrezen del, ki je z danim povezan (nadlakt, stegno), saj bi sicer zaradi nekonsistentnih opisov dobili rezultate, ki jih ne moremo dobro preveriti.

Naj bo j superelipsoid iz interpretacije, ki se ujema z neko okončino (nadlaktjo oz. stegnom). Potem za superelipsoid i , ki je ustrezno nadaljevanje okončine (podlakt oz. golen) in ga še ni v interpretaciji, velja:

1. volumen superelipsoida i je v danem intervalu (tabela 3.5) in
2. razdalja med sklepom, ki povezuje dela ustrezajočima j in i , in središčem superelipsoida j je v danem intervalu (enačba 3.6) in
3. razdalja med sklepi izračunanega in prilagojenega dela modela je v danem intervalu (enačba 3.11).

Če omejitve veljajo za več delov, izberemo tistega z najmanjšo razdaljo d iz enačbe 3.11.

4

Rezultati

V tem poglavju bomo prikazali rezultate, ki smo jih dosegli z opisano metodo. Najprej bomo predstavili rezultate razpoznavanja prizorov z enim samim objektom, nato pa kompleksnejše prizore.

Slike smo zajemali z globinskim skenerjem, ki deluje na podlagi strukturirane svetlobe. Na tako zajetih slikah smo rekonstruirali superelipsoide po metodi gradi-in-izberi s sistemom „SEGMENTOR” [14, 15]. Rekonstrukcija ene slike je na PC računalniku s procesorjem *Pentium Celeron 400MHz* porabila od uro in pol (preprosti prizori) do tri ure (kompleksnejši prizori) procesorskega časa. Dobljene opise smo uporabili kot vhodne podatke za v prejšnjem poglavju predstavljen algoritem za razpoznavanje, ki je bil implementiran v jeziku C++. Na prej omenjenem računalniku je razpoznavanje porabilo od 3 do 50 sekund procesorskega časa.

Vsak primer razpoznavanja je predstavljen s štirimi slikami:

- levo zgoraj je intenzitetna slika prizora,
- desno zgoraj je globinska slika z rekonstruiranimi superelipsoidi,
- levo spodaj so le tisti superelipsoidi z rekonstrukcije, ki so vključeni v interpretacije, in
- desno spodaj na intenzitetni sliki prikazan prilagojen model človeške figure.

4.1 Razpoznavanje prizorov z enim samim objektom

Metodo smo najprej preizkusili na prizorih, ki vsebujejo le poznani objekt. Objekt smo postavili v sedem različnih poz in za vsako pozo zajeli osem globinskih slik iz različnih gledišč, skupno torej 56 globinskih slik. Rekonstrukcije superelipsoidov s teh slik smo uporabili kot vhod za razpoznavanje.

Metoda je na 39 prizorih zaznala poznani objekt, od tega se je model objekta, izračunan iz najboljše interpretacije, v 24 primerih zelo dobro prilagal objektu. V

Skupaj prizorov 56			
Objekt razpoznan 39		Objekt ni razpoznan 17	
Prileganje modela:		Vzrok neuspeha:	
Dobro 24	Slabo 15	Glava ali trup zakrita 9	Premajhna interpretacija 8

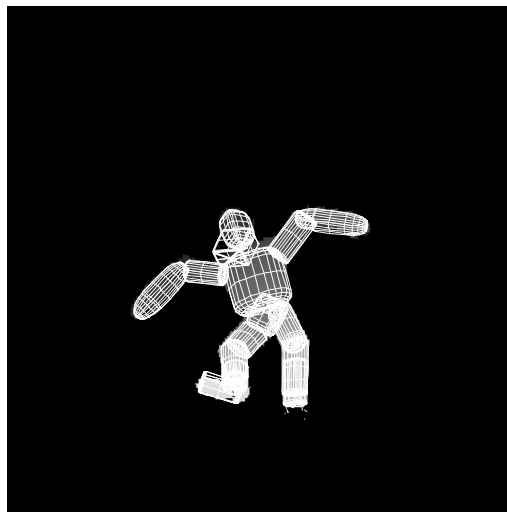
Tabela 4.1: Rezultati razpoznavanja na 56 prizorih z enim samim objektom.

interpretaciji sta bili povprečno 7.2 realni ujemanji. Na 17 prizorih objekt ni bil zaznan. Vzrok sta bila v 9 primerih zakrita glava ali trup, tako da se rekonstruirani superelipsoidi niso ujemali z glavo oz. trupom, v 8 primerih pa so bile rekonstrukcije take, da je najboljša interpretacija vsebovala manj kot štiri realna ujemanja in je bila kot taka označena kot slaba. Rezultate povzema tabela 4.1.

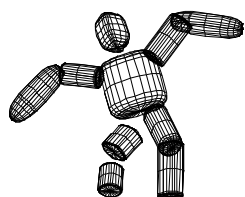
Na slikah 4.1 do 4.3 so prikazani primeri, ko se je model dobro prilegal objektu. Objekt je bil dobro razpoznan, brez manjkajočih delov. Model objekta zavzema isto pozo kot objekt in vsi njegovi deli modela se dobro prilegajo delom objekta. Napake pri izračunu poze modela so minimalne. K temu prav gotovo prispevajo dokaj dobri rezultati rekonstrukcije.



(a)



(b)

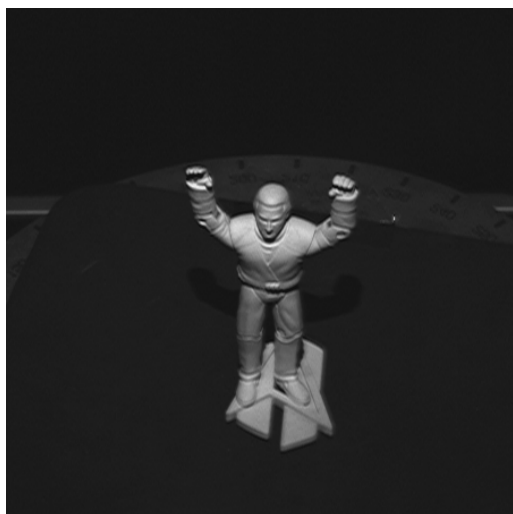


(c)

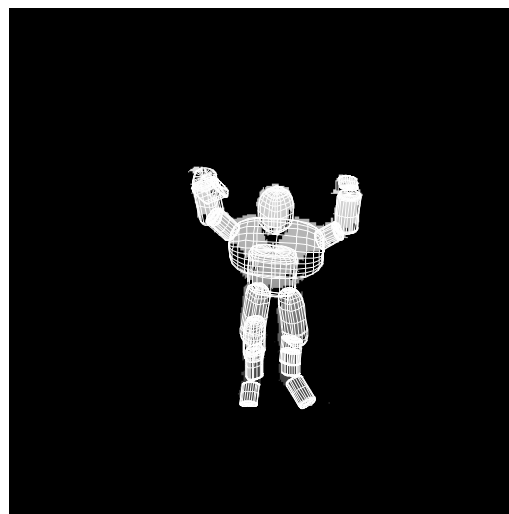


(d)

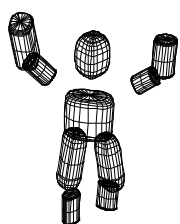
Slika 4.1: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



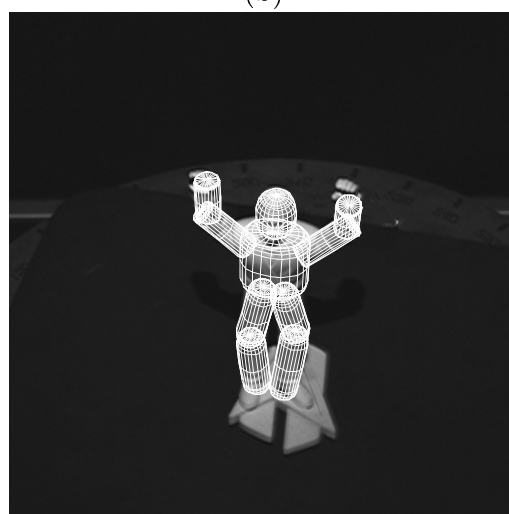
(a)



(b)



(c)

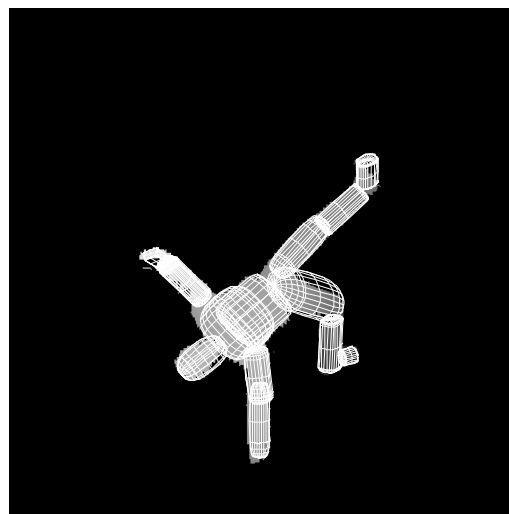


(d)

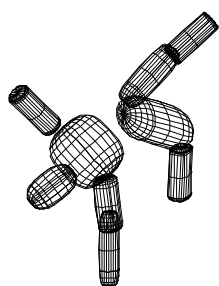
Slika 4.2: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



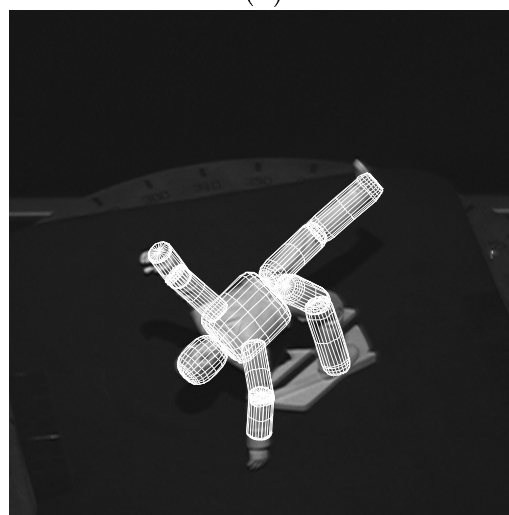
(a)



(b)



(c)



(d)

Slika 4.3: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.

4.2 Razpoznavanje pri kompleksnih prizorih

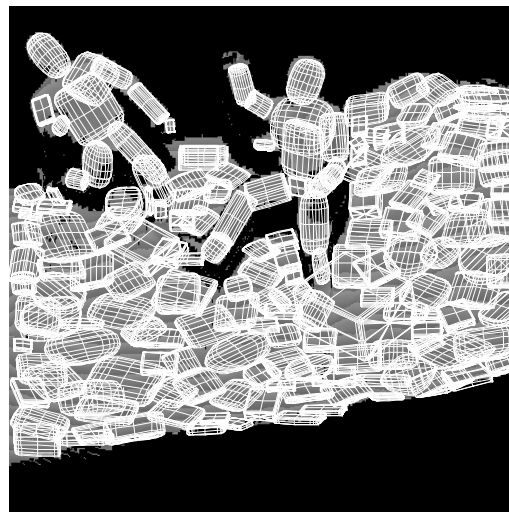
Metodo smo preizkusili tudi na bolj kompleksnih prizorih. Ti so tipično vsebovali več pojavitev človeške figure kakor tudi neznane objekte oziroma ozadje. Na primerih s slik 4.4 do 4.6 vidimo, da so rekonstrukcije „tal” dokaj neprimerne, saj se takih površin ne da dobro opisati s superelipsoidi. Rekonstrukcije vsebujejo mnogo superelipsoidov, ki so lokalno konsistentni z deli našega modela figure. V nobenem od takih primerov pa ni bil kot objekt izbrana množica delov, ki ne bi (razen izjem) res pripadala objektu. Interpretacije prizora s slike 4.4 imajo pomanjkljivosti, predvsem za goleni, kjer je eno ujemanje nepravilno (desna golen levega objekta), ostala pa so navidezna. Razen v primeru desne goleni desnega objekta je vzrok temu rekonstrukcija, ki je na teh predelih dala premajhne opise.

Na primeru s slike 4.5 vidimo učinkovitost preverjanja konsistentnosti ujemanja dela s pomočjo rekonstrukcije modela: ker objekt leži med „smetmi”, ki jih povzroča rekonstrukcija tal, je v okolici okončin precej delov, ki so potencialne okončine. S preverjanjem konsistentnosti dobimo najboljše, ki se v tem primeru ujemajo z okončinami objekta.

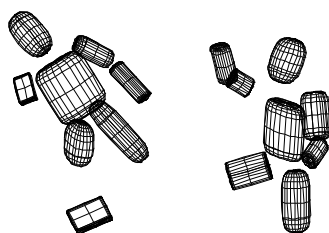
Na slikah 4.7 in 4.8 vidimo še dva primera razpoznavanja. Prizora vsebujeta poleg poznanih še več drugih objektov. Interpretacije prizorov so zelo dobre, le nekaj delov manjka oz. je kvantitetno narobe postavljenih.



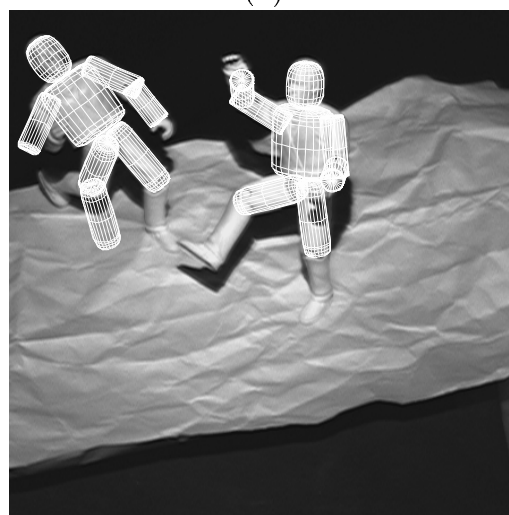
(a)



(b)

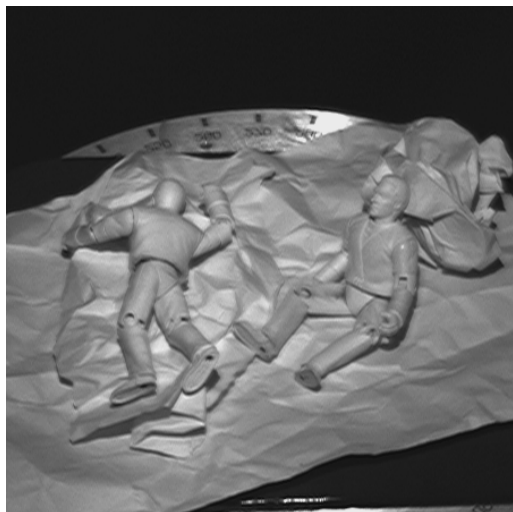


(c)



(d)

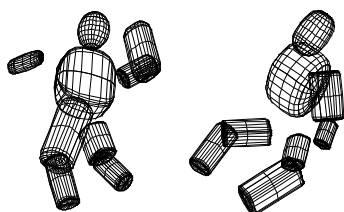
Slika 4.4: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



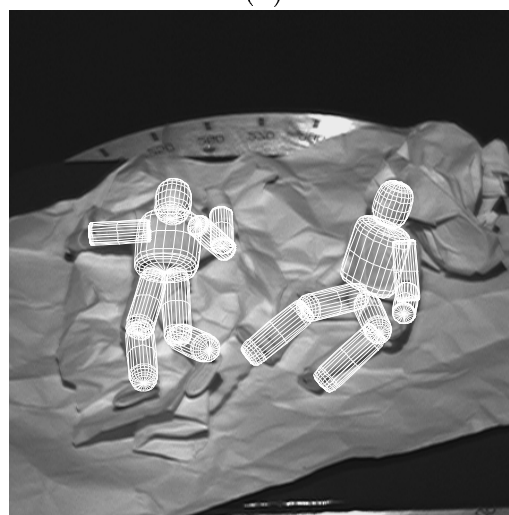
(a)



(b)

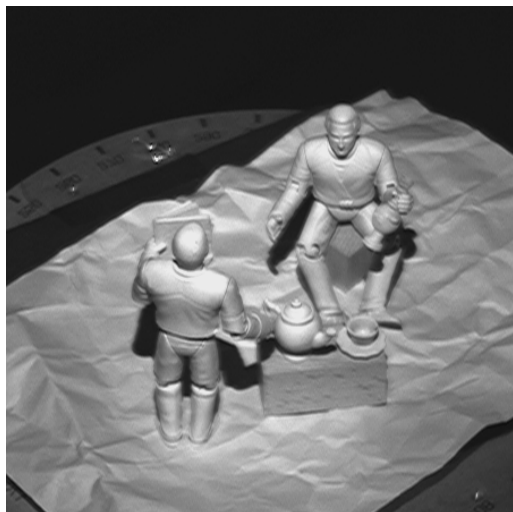


(c)

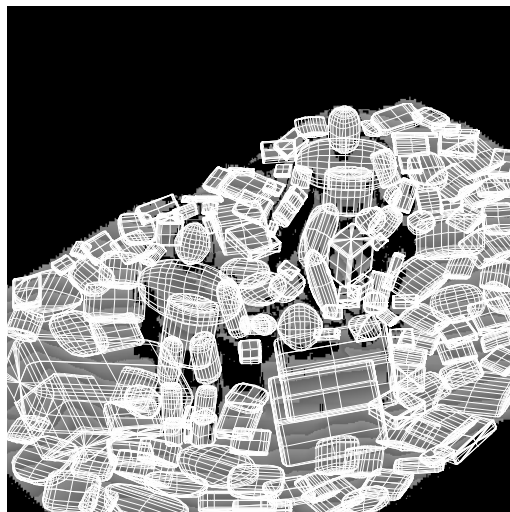


(d)

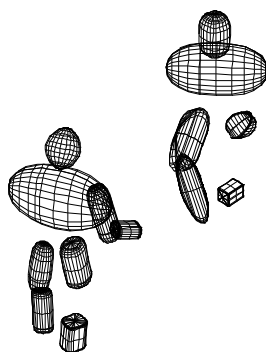
Slika 4.5: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



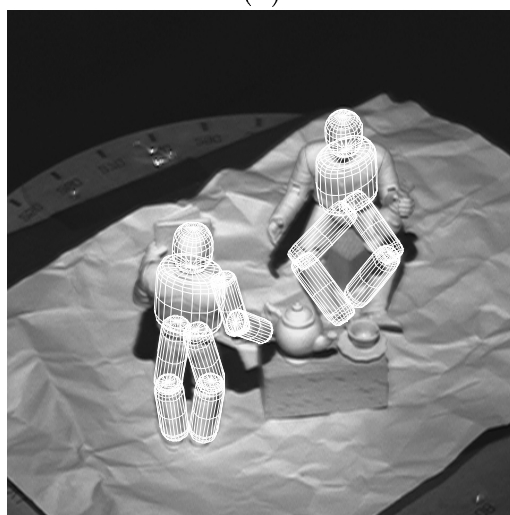
(a)



(b)

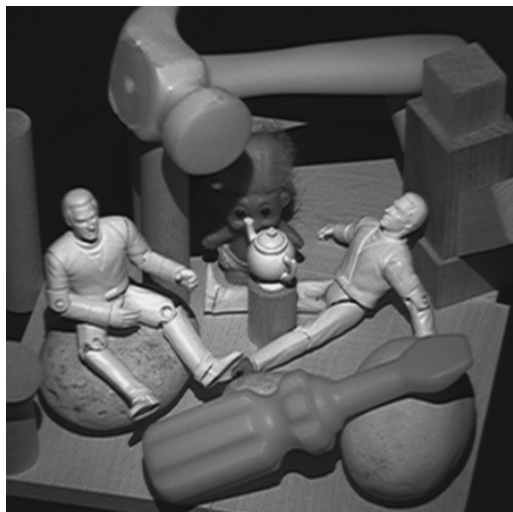


(c)

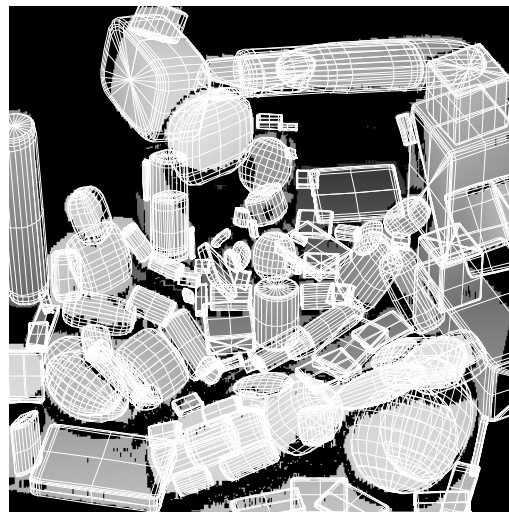


(d)

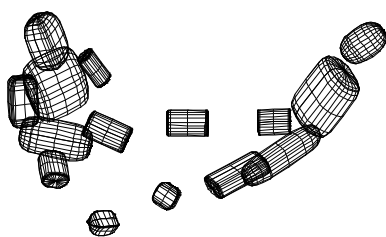
Slika 4.6: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



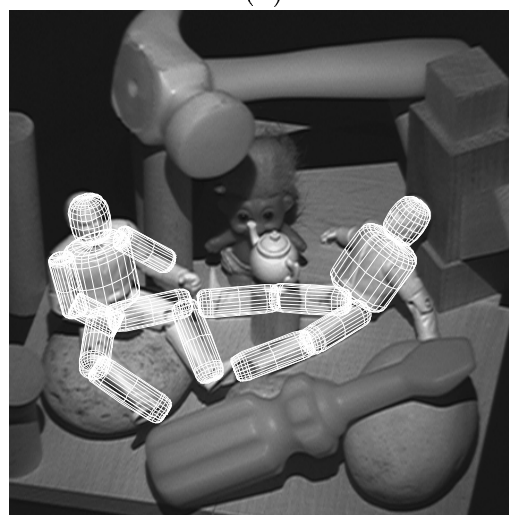
(a)



(b)

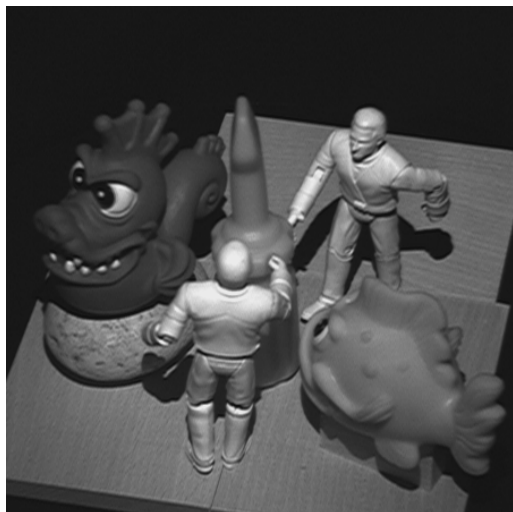


(c)

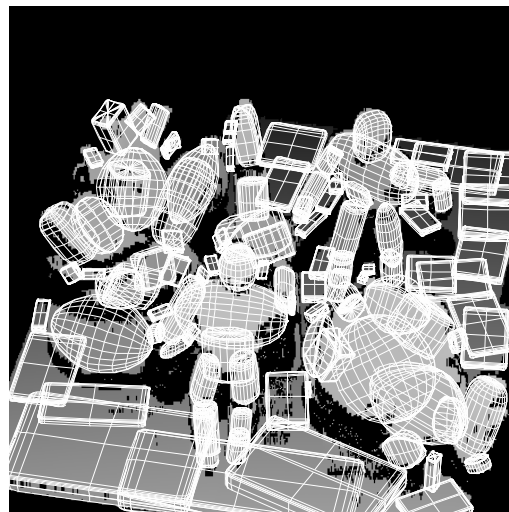


(d)

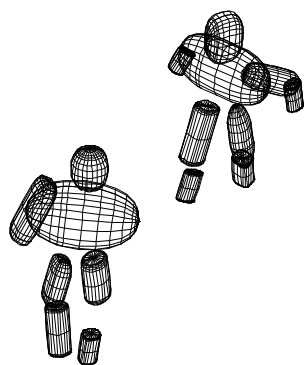
Slika 4.7: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.



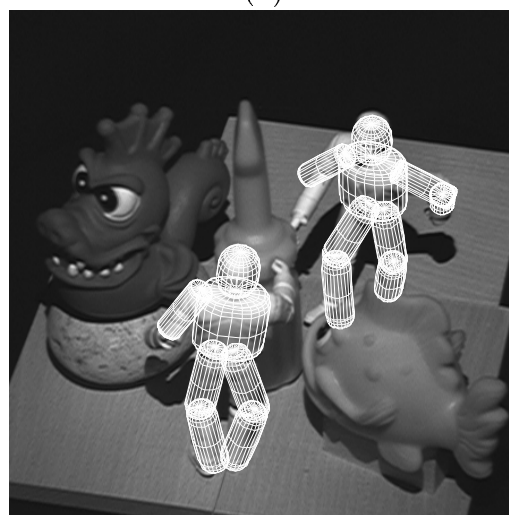
(a)



(b)



(c)



(d)

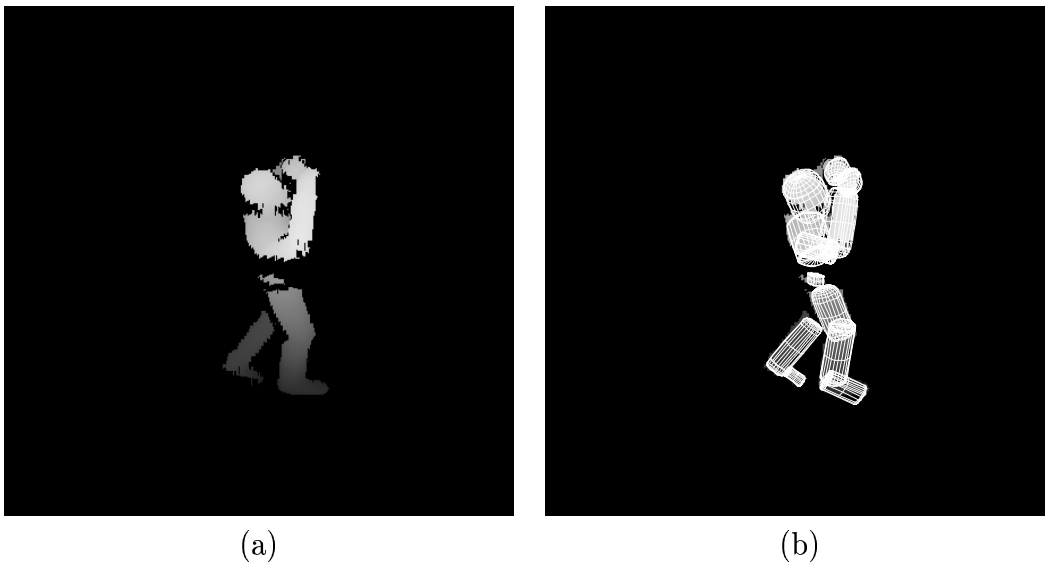
Slika 4.8: Primer razpoznavanja. (a) intenzitetna slika prizora, (b) globinska slika prizora z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.

4.3 Napake pri razpoznavanju

Metodo razpoznavanja smo preizkusili na 56 prizorih z enim objektom in na približno 20 kompleksnih prizorih. V nobenem primeru ni bila kot človeška figura interpretirana množica rekonstruiranih superelipsoidov, ki ne bi bili v večini rekonstruirani z območja figure. Tako smo zasledili le dve vrsti napak:

1. sistem prisotnega objekta ni razpoznal in
2. sistem je prisotni objekt razpoznal, prišlo pa je do napak pri ugotavljanju poze objekta.

V prvem primeru je bil najpogostejši vzrok nerazpoznanja objekta (tabela 4.1) zakrit trup oz. glava. Primer vidimo na sliki 4.9. Ker desna roka zakriva in meče senco na trup, tega praktično ni na globinski sliki in tako tudi ne ustreznega superelipsoida na rekonstrukciji.

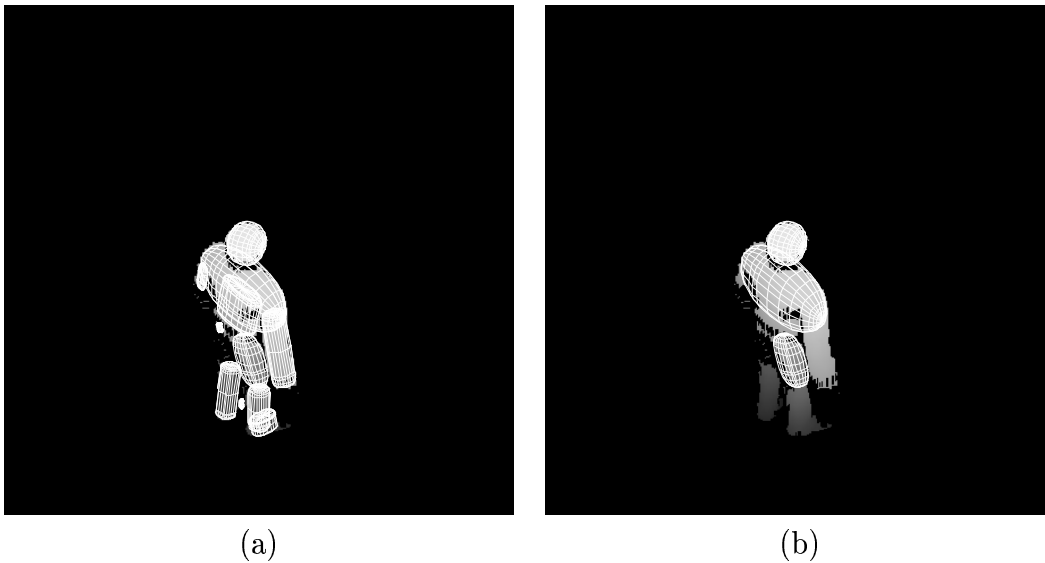


Slika 4.9: Primer, ko ni dobre interpretacije, zaradi manjkajočega superelipsoida s področja trupa: a) globinska slika, b) rekonstruirani superelipsoidi. Opazimo lahko, da na področju trupa praktično ni nobenega superelipsoida.

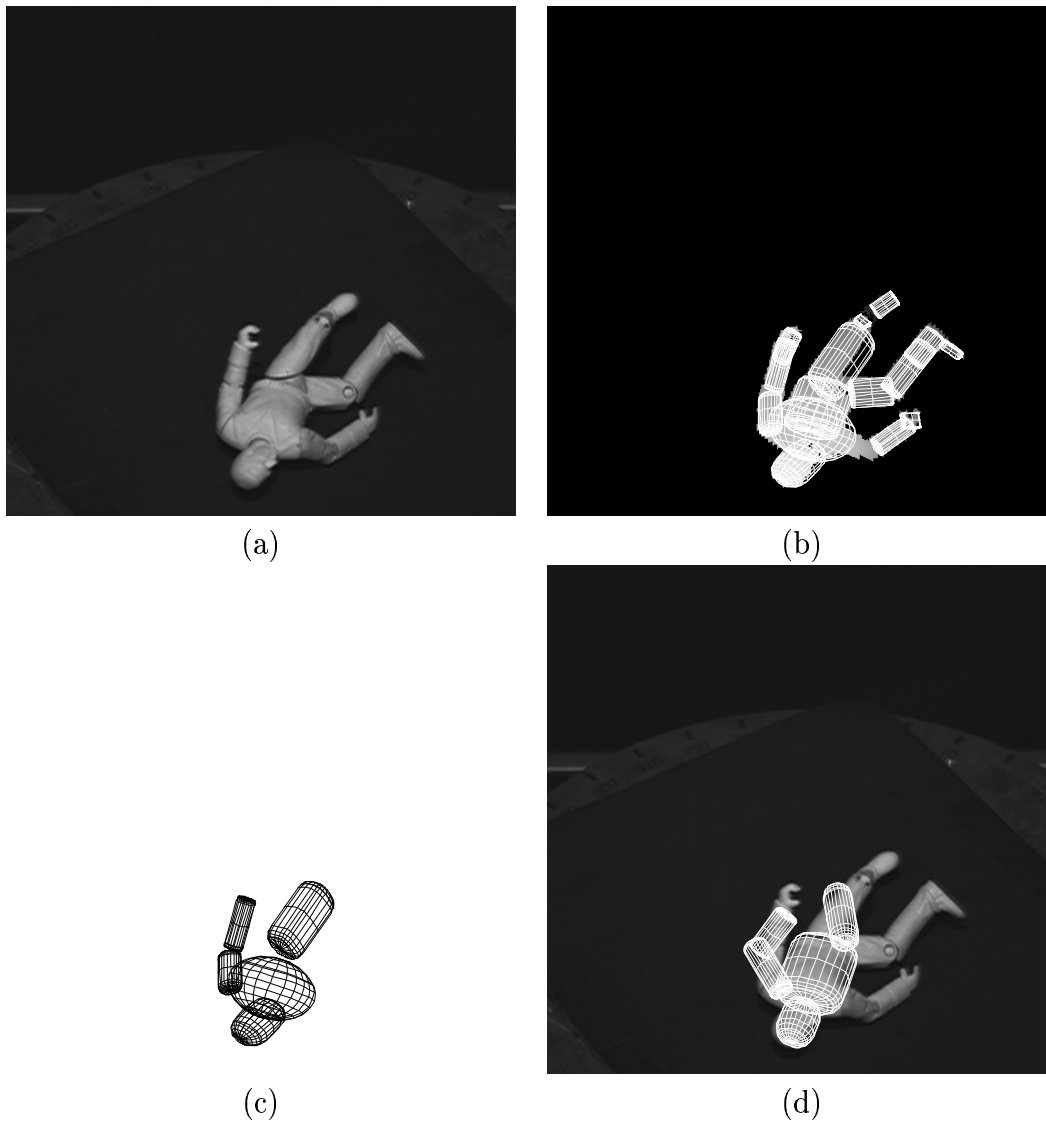
Vzrok nerazpoznanja je lahko tudi rekonstrukcija, ki ne zagotavlja zadostnega števila realnih ujemanj za interpretacijo. Primer kaže slika 4.10. Najboljšo interpretacijo tvorijo ujemanje za glavo, trup in desno stegno. Ker je superelipsoid, ki opisuje podlakti in del nadlakti predaleč od potencialnega sklepa, ga sistem ne vključi v interpretacijo, ki tako vsebuje manj kot štiri realna ujemanja in jo sistem zavrne kot nepravilno.

Napake pri ugotavljanju poze objekta izvirajo iz napak pri določanju pozicije in orientacije objekta in okončin. Primer kaže slika 4.11. Zaradi slabe rekonstrukcije

superelipsoidov s trupa se izračunani pozicija in orientacija figure ne ujemata s realnim stanjem. Zato je kvantitetno napačen izračun zadnjega parametra rotacije (enačba 3.5). Superelipsoid z levega stegna se tako bolje ujema z „desnim” stegnom modela in zato sistem zgreši ujemanja za levega. Te napake smo skušali odstraniti tako, da bolje določimo položaje sklepov, ki povezujejo okončine s trupom, in tako bolje določimo položaj trupa v prostoru. Na ta način bi lahko preverjali tudi interpretacije, ki nimajo realnega ujemanja za glavo ali trup.



Slika 4.10: Primer, ko ni dobre interpretacije. (a) globinska slika z rekonstruiranimi superelipsoidi, (b) globinska slika z deli iz najboljše interpretacije.

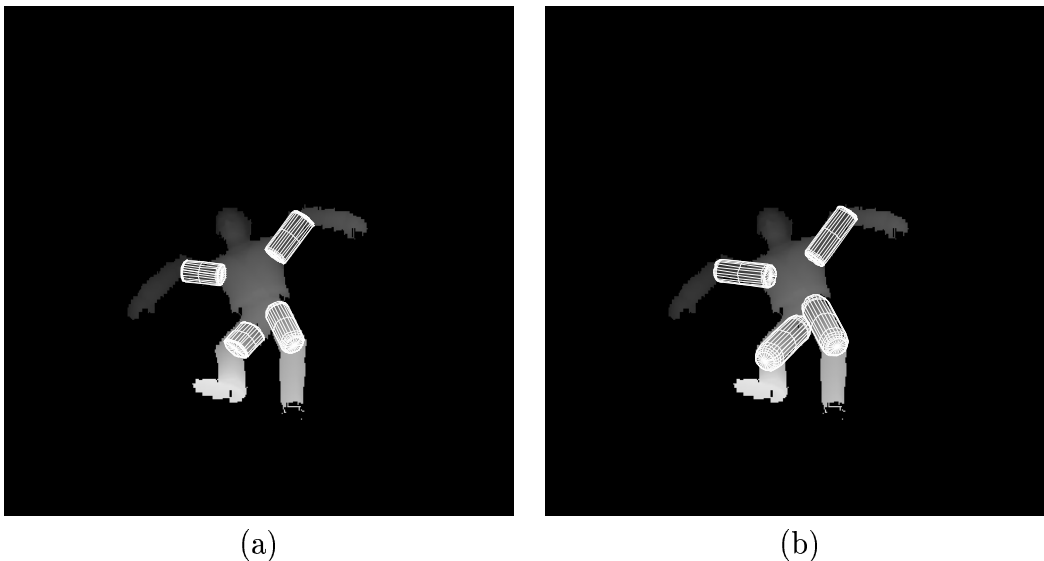


Slika 4.11: Primer napak pri razpoznavanju. (a) intenzitetna slika, (b) globinska slika z rekonstruiranimi superelipsoidi, (c) superelipsoidi iz interpretacije in (d) interpretacija.

4.4 Možne izboljšave

Napake pri izračunu pozicije objekta in položaja okončin bi se bistveno zmanjšale, če bi pri izračunu uporabljali položaj sklepov, ki povezujejo okončine s trupom. Tako bi lahko izračunali tudi pozicijo in orientacijo objekta, ko v interpretaciji ni realnih ujemanj za glavo ali trup.

Sklepe smo skušali najprej iskati na osi najmanjšega momenta. Ker so okončine podolgovate, je pričakovati, da bo sklep ležal na osi najmanjšega momenta. Vendar pa zaradi nestabilnih rekonstrukcij ne dobimo vedno pričakovanih superelipsoidov. Zato smo za namen boljšega določanja položaja sklepov lahko uporabili prilagajanje superelipsoidov modela na regije globinske slike, ki ustrezajo neki interpretaciji. Podobno kot pri verifikaciji interpretacije uporabimo za začetne parametre pri minimizaciji kar parametre superelipsoidov iz rekonstrukcije. Primer prilaganja delov modela kaže slika 4.12. Deli modela, prilagojeni na okončine, dobro nakazujejo smer, v katero kaže okončina. V takem primeru se zgoraj omenjeni sklepi nahajajo zelo blizu premice - z -osi superelipsoida. To smo hoteli uporabiti za točnejši izračun položaja objekta oz. delov - okončin, predvsem v primeru, ko interpretacija nima realnih ujemanj za glavo in trup.



Slika 4.12: Prilagajanje delov modela regijam globinske slike. (a) globinska slika s superelipsoidi, ki so bili uporabljeni kot začetni približki, (b) prilagojeni superelipsoidi modela.

Boljše položaje sklepov smo dobili s pomočjo relaksacije. Za začetne približke smo vzeli kar središča prilagojenih delov modela. Nek sklep smo nato premaknili po premici v smeri z -osi prilagojenega superelipsoida tako, da smo minimizirali napako

razdalj do ostalih sklepov:

$$\alpha_i = \min_{\alpha} |\vec{r}_i + \alpha \vec{e}_i - \vec{r}_j| - d_{ij} \quad (4.1)$$

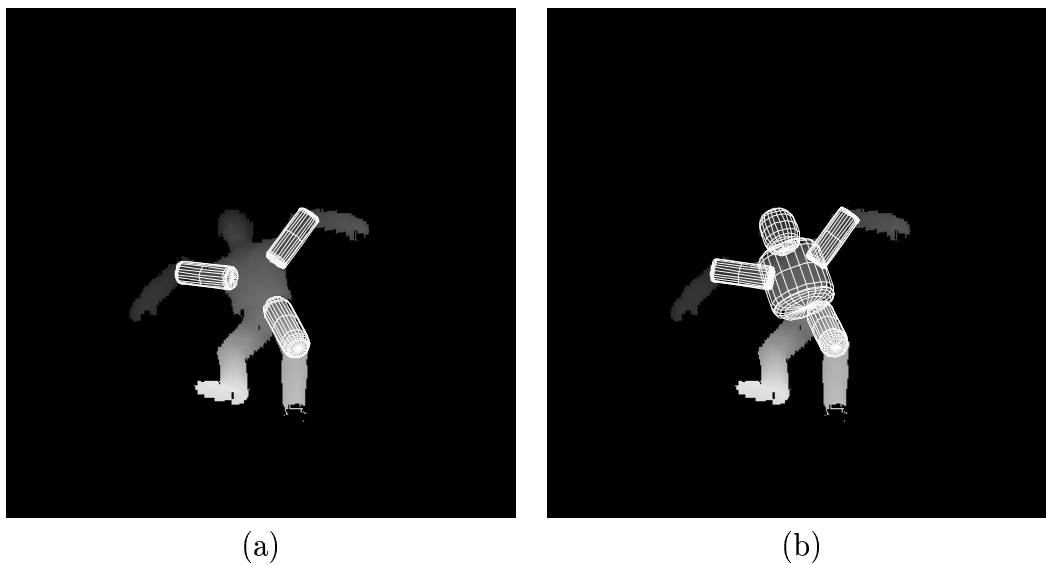
pri čemer je \vec{r}_i približek položaja i -tega sklepa, \vec{e}_i je smerni vektor, ki je vzporeden z z -osjo prilagojenega dela, in d_{ij} razdalja med sklepoma i in j na modelu. Sklep j smo nato premaknili za povprečen α_i :

$$\alpha = \frac{1}{N} \sum_{i=1}^N \alpha_i \quad (4.2)$$

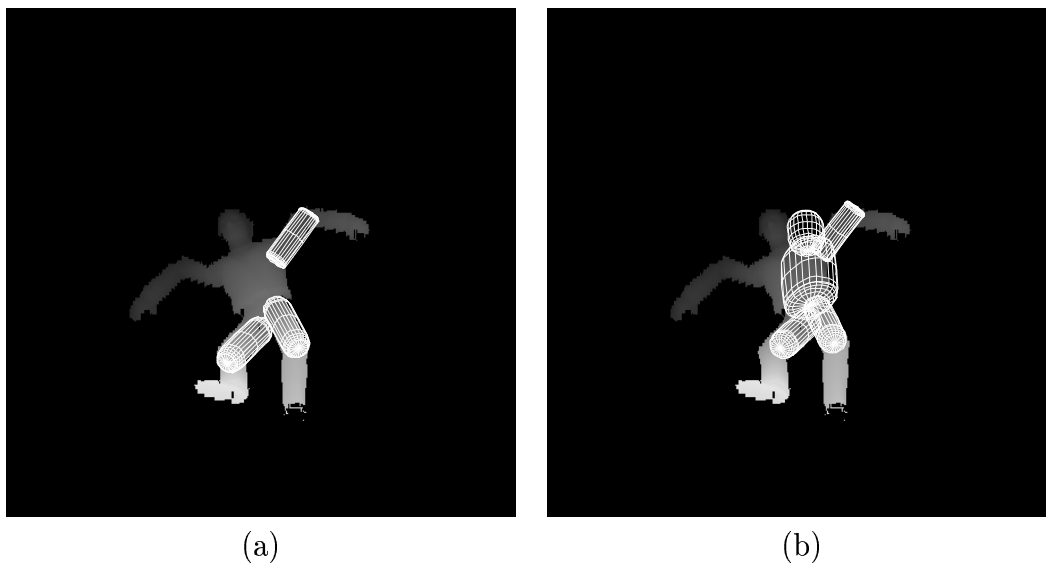
$$\vec{r}'_j = \vec{r}_j + \alpha \vec{e}_j \quad (4.3)$$

Ko popravimo položaje vseh sklepov, lahko postopek ponavljamo. Po stotih ponovitvah se sklepi niso več premikali in smo dobili boljše približne položaje sklepov. Te smo nato uporabili za izračun položaja figure. Če imamo podane položaje treh sklepov lahko izračunamo položaj figure: tri točke določajo ravnino v prostoru, na kateri naj bi ležalo središče, iz modela pa poznamo tudi kote med premicami, ki povezujejo pare sklepov in sklep s središčem. Tako lahko izračunamo položaj središča, nato pa lahko podobno kot pri enačbi 3.1 izračunamo še parametre orientacije.

Sliki 4.13 in 4.14 kažeta primer izračunanega položaja modela, ko so bili za izračun uporabljeni trije deli. Na sliki 4.13 je izračunani položaj dober, na sliki 4.14 pa ne. Upoštevati moramo, da so deli modela na teh primerih dobro prilagojeni (slika 4.12). Tipično prilagajanja teh delov niso tako dobra, zato so bili tudi rezultati slabši in vključitev takih izračunov ne bi bistveno doprinesla k izboljšanju rezultatov.



Slika 4.13: Primer, ko je položaj objekta dobro izračunan iz treh delov s pomočjo prileganja. (a) globinska slika z uporabljenimi deli, (b) prilagojen objekt, prikazani so le glava trup in tri okončine, na podlagi katerih je bil položaj izračunan.



Slika 4.14: Primer, ko je položaj objekta slabo izračunan iz treh delov s pomočjo prileganja. (a) globinska slika z uporabljenimi deli, (b) prilagojen objekt, prikazani so le glava trup in tri okončine, na podlagi katerih je bil položaj izračunan.

5

Zaključek

V magistrski nalogi je prikazana metoda za razpoznavanje objektov na podlagi opisov globinskih slik s superelipsoidi. Problem razpoznavanja je problem iskanja ujemanja med deli modela objekta in deli rekonstruiranimi s prizora. Metoda temelji na interpretacijskih drevesih, s katerimi iščemo tako podmnožico rekonstruiranih superelipsoidov, ki se ujema z modelom objekta. Kot objekt je bil uporabljena človeška figurica. Objekt je bil modeliran s superelipsoidi in zanimiv iz večih razlogov. Ne da se ga popolnoma modelirati s superelipsoidi, kar drži pri večini realnih objektov. Rekonstrukcije, ki jih dobimo z metodo gradi-in-izberi tako niso stabilne. Problem se še poveča zaradi fleksibilnosti objekta, ki ima gibljive dele in lahko zavzame različne poze.

Podana je metoda za razpoznavanje z interpretacijskimi drevesi. Splošen algoritem za iskanje po interpretacijskem drevesu smo prilagodili na dani problem. Definirali smo preverjanje konsistentnosti ujemanja delov na podlagi ujemanja volumnov in razdalj med ustreznimi deli. Prav tako smo predlagali način za verifikacijo interpretacije. Zaradi nenatančnih izračunov za pozicijo in orientacijo objekta v prostoru v primeru, ko v interpretaciji ni glave oz. trupa, predpostavimo, da je v tem primeru interpretacija nepravilna, sicer pa jo verificiramo z uporabo prilagajanja delov modela na globinsko sliko.

Rezultati razpoznavanja s podano metodo so dokaj dobri. Po interpretacijah prilagojeni modeli se dobro ujemajo z dejanskim stanjem na prizoru. Napake pri postavljanju modela in določanja pozicije okončin bi se z boljšimi rekonstrukcijami verjetno zmanjšale. Glavni problem, na katerega smo naleteli, je zagotovo netočnost izračuna pozicije in orientacije objekta v primeru, ko interpretacija nima realnih ujemanj za glavo ali trup.

5.1 Nadaljnje delo

V metodi smo sicer uporabili lastnosti strukture objekta in tipičnih rekonstrukcij, s katerimi najprej preverimo verjetneje boljše rešitve, vendar pa razširjamo drevo najprej v globino (*depth-first*). Lahko bi implementirali razširjanje najprej

najboljšega vozlišča (*best-first*). Pri vsakem razširjanju bi tako sortirali vozlišča po neki oceni kvalitete vozlišča. Ta bi lahko bila kombinacija ujemanja volumnov in bližine superelipsoidov.

Razviti bi bilo potrebno metodo za robusten izračun pozicije in orientacije objekta. V poštev bi prišla minimizacija razdalj med ustreznimi sklepi na modelu in možnimi pozicijami sklepov na sliki.

Preizkusiti bi bilo potrebno izboljšanje razpoznavanja z boljšimi rekonstrukcijami. Sam preizkus bi lahko naredili z objekti, ki so v bistvu sestavljeni iz superelipsoidov, ali pa z modifikacijami metode gradi-in-izberi oz. kakšno drugo metodo.

Ker se mnogo superelipsoidov na rekonstrukcijah delno prekriva, bi morali metodo razširiti tako, da bi omogočala ujemanja večih delov s prizora z enim delom modela. Za preverjanje konsistentnosti ujemanj bi lahko uporabili oceno, ki bi temeljila na preseku volumnov, ki jih zasedajo določeni superelipsoidi [8].

Boljše rezultate bi lahko dosegli tudi z vključevanjem znanja o objektih v samo rekonstrukcijo. Uporabili bi lahko kinematično obogaten model človeške figure (model, pri katerem upoštevamo močne položaje človeka) [10]. Fazo razpoznavanja objektov bi lahko vključili v fazo rekonstrukcije delov in tako izboljšali rezultate obeh.

Literatura

- [1] A.H. Barr, „Superquadrics and Angle-Preserving Transformations,” IEEE Computer Graphics and Applications, vol. 1, pp. 11–23, January 1981.
- [2] P.J. Besl: „Geometric Modelling and Computer Vision,” Proceedings of the IEEE, Vol. 76, str 936-958, 1998.
- [3] I. Biederman: „Human Image Understanding: Recent Research and a Theory,” Computer Vision, Graphics and Image Processing, Vol. 32, No. 1, 1985.
- [4] W.F. Bischof, T. Caelli, „Learning Structural Descriptions of Patterns: A New Technique for Conditional Clustering and Rule Generation,” Pattern Recognition, Vol. 27, No. 5, str. 689-697, 1994.
- [5] W.F. Bischof, T. Caelli, „Scene Understanding by Rule Evaluation,” IEEE Trans. PAMI, Vol. 19, No. 11, str. 1284-1288, 1997.
- [6] R.C. Bolles in P. Horaud, „3DPO: A Three-Dimensional Part Orientation System,” The International Journal of Robotic Research, Vol. 5, No. 3, str. 3-26, Jesen 1986.
- [7] F.G. Callari in U. Maniscalco, „A New Robust Approach to Image Shading Analysis and 3-D Shape Reconstruction,” Proc. Int. Conf. Pattern Recognition, str. 103-107, 1994.
- [8] L.H. Chen, Y.T. Liu, H.Y. Liao: „Similarity Measure for Superquadrics,” IEEE Proc. Vis. Image Signal Process., Vol. 144, No. 4, 1997.
- [9] S.J. Dickinson et al.: „Panel report: the potential of geons for generic 3-D object recognition,” Image and Vision Computing 15, str. 277-292, 1997.
- [10] V. Filova, F. Solina, J. Lenarčič: „Automatic reconstruction of 3D human arm motion from a monocular image sequence,” Machine Vision and Applications, 10, str. 223-231, 1998.
- [11] P.J. Flynn in A.K. Jain, „3D Object Recognition Using Invariant Feature Indexing of Interpretation Tables,” CVGIP: Image Understanding, Vol. 55, No. 2, str. 119-129, Marec 1992.

- [12] W.E.L. Grimson, „Object Recognition by Computer,” MIT Press, Cambridge (MA), 1990.
- [13] A.J. Hanson: „Hyperquadrics: smoothly deformable shapes with complex polyhedral bounds,” *Comp. Vision Graphics Image Proc.* 44, str. 191-210, 1988.
- [14] A. Jaklič, „Construction of CAD Models from Range Images,” doktorska disertacija, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 1997.
- [15] A. Jaklič, A. Leonardis, F. Solina: „Segmentation and Recovery of Superquadrics,” Kluwer Academic Publishers, Dordrecht, Oktober 2000.
- [16] A.E. Johnson in M. Hebert, „Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes,” *IEEE Trans. PAMI*, Vol. 21, No.5, str. 433-449, Maj 1999.
- [17] A. Leonardis, „Image analysis using parametric models: model-recovery and model-selection”, doktorska disertacija, Univerza v Ljubljani, Fakulteta za elektrotehniko in računalništvo, 1996.
- [18] A. Leonardis, A. Gupta, R. Bajcsy: „Segmentation of range images as the search for geometric parametric models,” *International Journal of Computer Vision*, 14, str 253-277, 1995.
- [19] H. Murase, S.K. Nayar: „Visual Learning and Recognition of 3-D Objects from Appearance,” *International Journal of Computer Vision*, Vol. 3, pp. 181-209.
- [20] R. Nevatia, T. Binford, „Description and Recognition of Curved Objects,” *Artificial Intelligence*, Vol. 38, str. 77-98, 1977.
- [21] A.P. Pentland, „Perceptual organization and the representation of natural form,” *Art. Intell.* 28, str. 293-331, 1986.
- [22] R.B. Schudy in D.H. Ballard: „Towards an anatomical model of heart motion as seen in 4-D cardiac ultrasound data,” *Proc. 6th Conference on Computer Applications in Radiology and Computer-Aided Analysis of Radilogical Images*, 1979.
- [23] F. Solina, „Shape Recovery and Segmentation with Deformable Part Models,” Ph.D. Thesis, University of Pennsylvania, 1987.
- [24] F. Solina, „Segmentation with Volumetric Part Models,” *Computing Suppl.* 11, str. 201-220, 1996.
- [25] F. Solina in R. Bajcsy, „Recovery of parametric models from range images: The case for superquadrics with global deformations,” *IEEE Trans. Pattern Anal. Machine Intell.* 12, str. 131-147, 1990.

- [26] E. Trucco, A. Verri, „Introductory techniques for 3-D computer vision,” Prentice Hall, NJ, 1998.
- [27] N. Yokoya et al., „Recovery of Superquadric Primitives from a Range Image Using Simulated Annealing,” Proc. Int. Conf. Pattern Recognition, str. 168-172, 1992.

Zahvala

Iskreno se zahvaljujem prof. dr. Francu Solini za mentorstvo.

Brez podpore staršev in družine seveda ne gre: hvala Vida, Boro, Diana.

Za pomoč pri raznih zapletih in stiskah gre zahvala Alešu, Bojanu, Danijelu, Igorju, Marjanu, Matjažu, Petru in vsem, ki sem jih pozabil v tej zahvali.

Izjava

Izjavljam, da sem magistrsko nalogo izdelal samostojno pod vodstvom mentorja prof. dr. Franca Soline. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.