

# DOLOČANJE POLOŽAJA TELES V PROSTORU IZ ENE SAME 2-D SLIKE

INFORMATICA 1/91

Barbara Lakner  
Institut Jožef Stefan, Ljubljana

Keywords: attitude determination, methods, polyhedral objects.

Podan je pregled metod za določanje položaja poliedrskih teles v prostoru iz ene same 2-D slike. Opisane so nekatere metode, ki pridejo v poštev pri prepoznavanju poliedrskih teles: Robertsova, analitični Horaudova in Dhomejeva ter iteracijska Lowejeva.

**ABSTRACT:** A survey of methods for determination of the attitude of 3-D polyhedral objects from a single perspective view is presented. These methods are used in systems for recognizing polyhedral objects from gray-level pictures. Methods described: Roberts' method, analytical Horaud's method, analytical Dhome's method and Lowe's iterative method.

## 1. UVOD

Ogledali si bomo nekaj metod za določanje položaja poliedrskega telesa iz ene same 2-D slike. Opisane metode so nastale pri prepoznavanju poliedrskih teles, kjer moramo ugotoviti, ali so na sliki telesa, opisana z vnaprej danimi modeli in kje v prostoru ležijo. Pri tem je treba določiti 6 prostostnih stopenj (tri kote zavrtitve okrog koordinatnih osi  $x, y, z$  in translacijski vektor). Iskana rotacija in translacija postavita model v prostor tako, da se projekcija transformiranega modela ujema z dano sliko.

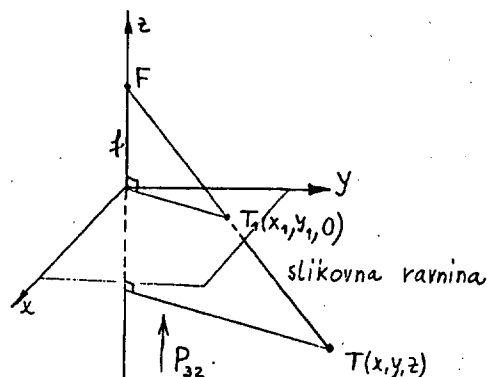
Položaj telesa v prostoru lahko določimo na različne načine: z aktivnimi metodami zaznavanja (dobivanje globinske slike s pomočjo laserskih žarkov, fotometrične metode), s stereovidom, z informacijami, dobljenimi s pomočjo gibanja (obravnavava več zaporednih slik) in z merjenjem globine s pomočjo teksture in senc. Vendar aktivne metode ne pridejo vedno v poštev (če bi imeli npr. dva taka sistema za prepoznavanje drugega ob drugem, bi se laserski žarki lahko motili med sabo); stereovid ne pride v poštev za določanje položaja teles, ki so zelo oddaljena; informacija, dobljena s pomočjo gibanja je uporabna, kadar imamo dovolj veliko relativno gibanje med opazovalcem in objektom, kar je praktično izvedljivo za blizu stoječe objekte; merjenje globine s pomočjo teksture in senc pa pride v poštev za posebne primere in ni vedno dovolj natančno.

Pionirsko delo na področju prepoznavanja poliedrskih teles je opravil Roberts [1]; v njem se ukvarja tudi z določanjem položaja telesa. Kanade [2] je ta problem rešil analitično za ortografsko projekcijo, vendar imamo opravka s perspektivno projekcijo, zato si bomo ogledovali le take metode. Barnard [3] interpretira tro-

jico črt na sliki kot oglišče, kjer se stikajo trije robovi. Problem je rešil pri predpostavki, da so koti med robovi enaki  $\pi/2$ . Shakunagar in Kaneko [4] sta predpostavila, da je le en rob izmed trojice pravokoten na ravnino ostalih dveh. Horaudova metoda [5] nima omejitve za kote med robovi; problem prevede na reševanje sistema transcendentnih enačb, ki ga reši s tabeliranjem. Dhome in soavtorji [6] rešujejo problem analitično z iskanjem ničel polinoma 8 stopnje. Rešitev Huttenlocherja in Ullmana [7] je vsebovana v prej navedeni metodi. Lowe [8], [9] rešuje problem iterativno z Newtonovo metodo. Pri danem začetnem približku dobimo ob obravnavi treh točk eno samo rešitev. Zen Chen s soavtorji [10] določa položaj kocke s pomočjo točk izginjanja (vanishing points). Metoda pride v poštev pri objektih, ki so posneti tako od blizu, da se nosilke vzporednih stranic na sliki sekajo.

## 2. ROBERTSOVA METODA

Oglejmo si najprej relacijo med sliko in sceno: slikovna ravnina naj bo kar ravnina  $(x, y)$ , vse točke na sceni naj imajo z koordinato negativno in naj bo gorišče v točki  $F(0, 0, f)$ .



Če je  $T(x, y, z)$  točka na sceni in  $T_1(x_1, y_1, 0)$  njena projekcija, je projekcija iz scene na slikovno ravnino

$$P_{32}(x, y, z) = \left( \frac{x}{f-z} f, \frac{y}{f-z} f \right).$$

Uvedemo še globino projicirane točke, katere  $z$  koordinato definiramo kot  $z_1 := \frac{x}{f-z} f$ . Zdaj imamo preslikavo  $P_{33}: R^3 \rightarrow R^3$ ;

$P_{33}(x, y, z) = \left( \frac{x}{f-z} f, \frac{y}{f-z} f, \frac{x}{f-z} f \right)$ . Predpis ni linearen, lahko pa ga lineariziramo z uvedbo homogenih koordinat:  $P_{33}$  predstavimo kot preslikavo  $P: R^4 \rightarrow R^4$ .

Vektor  $a = (x, y, z)$  v kartezičnih koordinatah izrazimo v homogenih koordinatah kot  $a_h = (wx, xy, zy, w)$ , kjer je  $w$  poljubna konstanta. Vektorja  $(wx, wy, wz, w)$  in  $(x, y, z, 1)$  sta ekvivalentna za  $w \neq 0$ , oziroma predstavljata isti vektor v  $R^4$ . V tem primeru zapišemo:  $(wx, wy, wz, w) \cong (x, y, z, 1)$ . Projekcija  $P$  torej je

$$\begin{pmatrix} x \\ y \\ z \\ w_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 1 \end{pmatrix} \begin{pmatrix} wx \\ wy \\ wz \\ w \end{pmatrix}.$$

Če vzamemo za  $w = 1$ , dobimo vektor  $(x, y, z, (f-z)/f)$ , kar se ujema s projektorjem  $P_{32}$ .

Naj bodo  $R_{33}^x, R_{33}^y, R_{33}^z$  rotacije za kot  $\varphi$  okrog osi  $x$ , za kot  $\psi$  okrog osi  $y$  oziroma za kot  $\vartheta$  okrog osi  $z$ . Tako kot to že nakazuje spodnji indeks, gledamo nanje kot na preslikave v 3-D prostoru. Zapišimo z matriko  $R_{33}^z$ :

$$R_{33}^z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}.$$

Podobno zapišemo matriki za  $R_{33}^y$  in  $R_{33}^x$  (glej [11]). Rotacije so v homogenih koordinatah predstavljene takole:

$$R_x = \begin{pmatrix} R_{33}^x & 0 \\ 0 & 1 \end{pmatrix}, \quad R_y = \begin{pmatrix} R_{33}^y & 0 \\ 0 & 1 \end{pmatrix}, \quad R_z = \begin{pmatrix} R_{33}^z & 0 \\ 0 & 1 \end{pmatrix}$$

in celotna rotacija je  $R = R_x \cdot R_y \cdot R_z$ .

Translacijo za vektor  $(x_i, y_i, z_i)$  predstavimo v homogenih koordinatah kot

$$T = \begin{pmatrix} 1 & & & x_i \\ & 1 & & y_i \\ & & 1 & z_i \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Naj bo  $Q := TR$  operator, s katerim je predstavljen poljuben premik telesa v prostoru. Preslikava  $Q$  postavi model na sceno tako, da se le-ta projicira s projektorjem  $P$  na dano sliko.

Če zapišemo komponente produkta  $H = PTR$ , vidimo, da lahko tretjo vrstico matrike  $H$  izračunamo iz njene četrte vrstice.

Naj bo  $\tilde{H}$  matrika  $H$  brez tretje vrstice. Če je  $H(x, y, z, w) = (x', y', z', w')$ , je  $\tilde{H}(x, y, z, w) = (x', y', w')$ .

Zdaj lahko definiramo problem: imejmo  $n$  točk  $T_i(x_i, y_i, z_i, 1)$

modela in prav toliko točk  $T'_i(x'_i, y'_i, 1)$  na sliki. Iščemo tako rotacijo in translacijo, ki bosta postavili model na sceno tako, da se bo sprojeiral v dano sliko. Iščemo torej tako preslikavo  $\tilde{H}$ , da bo

$$H(T_i) \cong T'_i, \quad i = 1, \dots, n \quad (2.1)$$

Iz  $\tilde{H}$  izračunamo  $H$  in od tod  $Q = P^{-1}H$ , kar je rešitev našega problema.

Sistem (2.1) rešimo takole:

$$\begin{aligned} \tilde{H}(T_i) &= w_i(T'_i), \quad i = 1, \dots, n \\ \tilde{H}[T_1, \dots, T_n] &= [T'_1, \dots, T'_n] \text{diag}(w_1, \dots, w_n), \end{aligned}$$

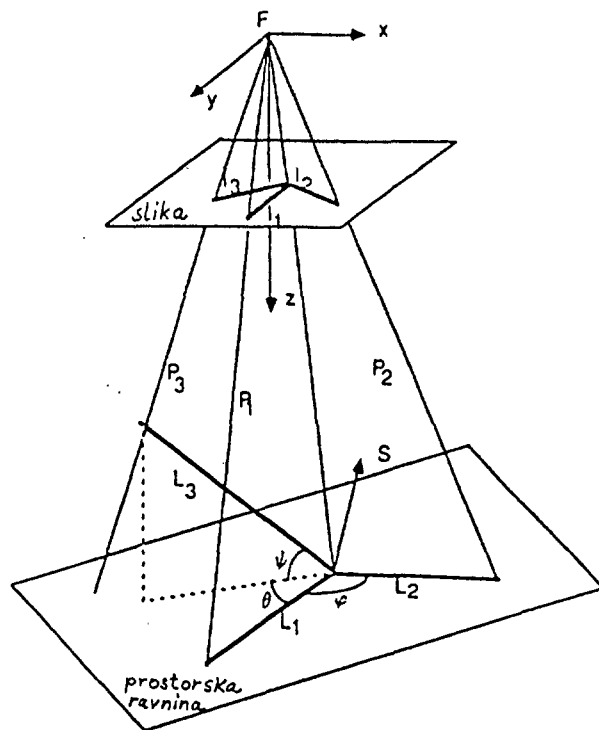
kjer so elementi matrike  $\tilde{H}$  in  $w_1, \dots, w_n$  neznanke.

Predpostavimo, da so elementi matrike  $\tilde{H}$  med sabo neodvisni in poiščemo rešitev z metodo najmanjših kvadratov [1]. Tako dobljena rešitev ne ustreza nujno pogojem, kakršnim morajo ustrezati elementi matrike  $H$ .

Koliko točk  $T_i$  potrebujemo za rešitev sistema? Imamo  $12 + n$  neznanek in  $3n$  enačb, torej potrebujemo vsaj 6 različnih točk za enolično rešitev.

### 3. ANALITIČNA METODA (HORAUD)

Gorišče postavimo v izhodišče koordinatnega sistema:



Zapišimo projekcijo  $P(x, y, z) = (xf/z, yf/z, f)$ .

Predstavimo vektorje s sferičnimi koordinatami:

$$x = \cos \alpha \cos \beta, \quad y = \sin \alpha \cos \beta, \quad z = \sin \beta,$$

kjer za parametra  $\alpha$  in  $\beta$  velja  $0 \leq \alpha \leq 2\pi$ ,  $-\pi/2 \leq \beta \leq \pi/2$ .

Imejmo normirane vektorje  $L_1, L_2, L_3$  s skupno začetno točko na modelu in naj bodo koti med njimi  $\varphi, \vartheta, \psi$ , tako kot je označeno na sliki.

Naj bo  $S$  normala na ravnino vektorjev  $L_1, L_2$ .

Naj bodo  $l_1, l_2, l_3$  slike vektorjev  $L_1, L_2, L_3$ . Položimo skozi gorišče in vektorje  $l_1, l_2, l_3$  ravnine (interpretacijske ravnine) z normalami  $P_1, P_2, P_3$ . Vektorji  $L_1, L_2, L_3$  leže na teh interpretacijskih ravninah z normalami  $P_1, P_2$  oziroma  $P_3$ . Velja:

$$L_1 = S \times P_1, \quad L_2 = S \times P_2. \quad (3.1)$$

Po krajši izpeljavi [5] dobimo

$$(S \times P_1)(S \times P_2) = \|S \times P_1\| \|S \times P_2\| \cos \varphi \quad (3.2)$$

$$L_3 = S \sin \psi + L_1 \cos \vartheta \cos \psi + (S \times L_1) \sin \vartheta \cos \psi \quad (3.3)$$

in

$$(S \cdot P_3) \|S \times P_1\| \sin \psi + S \cdot (P_1 \times P_3) \cos \vartheta \cos \psi + (S \cdot P_1)(S \cdot P_3) \sin \vartheta \sin \psi = (P_1 \cdot P_3) \sin \vartheta \sin \psi \quad (3.4)$$

Predpostavimo, da poznamo kote  $\varphi, \psi, \vartheta$ . Z rešitvijo enačb (3.2) in (3.4) dobimo normalo  $S$  in od tod s pomočjo (3.1) in (3.3) vektorje  $L_1, L_2$  in  $L_3$ . Še vedno pa ne vemo, v kateri točki prostora se ti vektorji stikajo.

Sistem enačb (3.2), (3.4) je nelinearen in transcendenten, zato ga avtor ne rešuje analitično, ampak rešitev konstruira. Enačbi (3.2) in (3.3) zapiše kot:

$$\cos \varphi = f(\alpha, \beta) \quad (3.5)$$

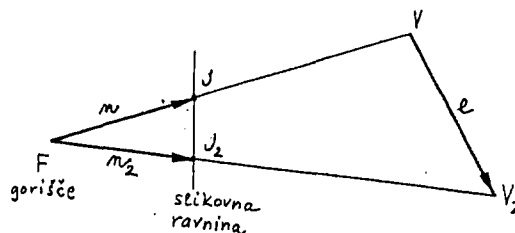
$$\sin \vartheta \cos \psi = g_1(\alpha, \beta) \sin \psi + g_2(\alpha, \beta) \cos \vartheta \cos \psi + g_3(\alpha, \beta) \sin \vartheta \cos \psi. \quad (3.6)$$

Tabeliramo funkcije  $f, g_1, g_2, g_3$  za  $\alpha, \beta$  iz intervalov  $0 \leq \alpha \leq 2\pi, 0 \leq \beta \leq \pi/2$  z določenim korakom in točko za točko konstruiramo krivulji, ki ju predstavljata enačbi (3.5) in (3.6). Presečišča krivulj so možne rešitve za smer vektorja  $S$ . Pri tem je treba pripomniti, da tabeliramo funkcije  $f, g_1, g_2, g_3$  le enkrat in rezultate uporabimo pri različnih predpostavkah za kote  $\varphi, \psi, \vartheta$ .

Določimo sedaj točen položaj trojice  $L_1, L_2, L_3$  s skupno točko  $V$  v prostoru! Projekcija točke  $V$  na slikovno ravnino je točka  $P(V) = J$  s koordinatami  $J(J_x, J_y, f, 0)$ . Vektor od gorišča  $F$  do točke  $J$  je  $n = (J_x, J_y, f, 0)$ . Vzemimo, da poznamo razdaljo  $d$  od gorišča  $F$  do točke  $V$ . Tedaj je translacijski vektor

$$t = dn / \|n\|. \quad (3.7)$$

Tako smo določili vektor translacije do skalarja natančno. Če poznamo še eno trojico vektorjev s skupno točko  $V_2$ , potem določimo  $d$  takole:

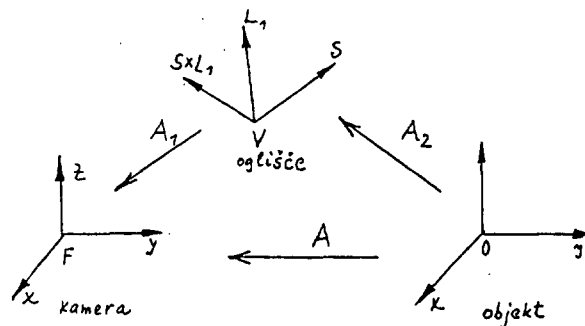


Naj bo  $e$  vektor od  $V$  do  $V_2$ . (Vektor  $e$  poznamo, saj primerjamo model telesa s sliko in tako  $V, V_2$  ustrežata ogliščem znanega modela.) Naj bo  $J_2$  projekcija  $V_2$  na slikovno ravnino in  $n_2$  vektor od  $F$  do  $J_2$ . Zaradi (3.7) dobimo

$$d = \frac{\|n\| \cdot \|n_2 \times e\|}{\|n \times n_2\|}$$

Tako smo za določitev vseh šestih prostostnih stopenj morali poznati dve trojici vektorjev oziroma 8 točk.

Zapišimo zdaj še premik  $A$  (translacija, rotacija) modela na sceno, tako da se izbrano oglišče modela s pripadajočimi vektorji projicira na izbrano trojico vektorjev!



$A = A_1 A_2$ , kjer je  $A_2$  prehodna matrika med različnima bazama vektorskih prostorov (bazo, v kateri je opisan model in bazo, sestavljeno iz  $S, L_1, S \times L_1$ ) in

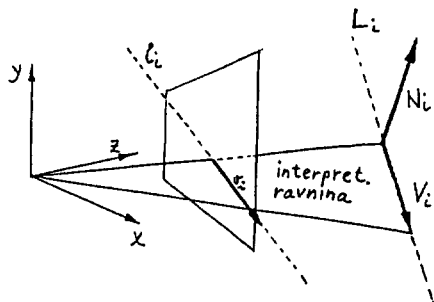
$$A_1 = \begin{pmatrix} p_x & q_x & r_x & t_x \\ p_y & q_y & r_y & t_y \\ p_z & q_z & r_z & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

kjer je  $S = (p_x, p_y, p_z), L_1 = (q_x, q_y, q_z), S \times L_1 = (r_x, r_y, r_z), t = (t_x, t_y, t_z)$ ; vse izraženo v koordinatnem sistemu kamere.

#### 4. ŠE ENA ANALITIČNA REŠITEV

(Dhome in soavtorji)

Koordinatni sistem kamere, projektor  $P$  in pojem interpretacijske ravnine poznamo že iz Horaudove metode.



Če je  $L_i$  premica v prostoru in  $l_i$  njena projekcija,  $V_i$  in  $v_i$  pa enotska vektorja na  $L_i$  oziroma  $l_i$ ,  $p_i$  pa točka na  $l_i$ , je normala  $N_i$  na interpretacijsko ravnino določena z

$$N_i = l_i \times p_i.$$

Za vektor  $V_i$  mora veljati

$$N_i \cdot V_i = 0. \quad (4.1)$$

Formulirajmo problem: imejmo tri premice  $L_{01}, L_{02}, L_{03}$ , ki leže na robovih modela (model je predstavljen v svojem lastnem koordinatnem sistemu  $(S_{0m})$ ), imejmo premice  $l_{01}, l_{02}, l_{03}$ , definirane v slikovnem koordinatnem sistemu  $(S_{0v})$ . Želimo najti rotacijo  $R_{\alpha\beta\gamma}$  in translacijo  $T_{uvw}$ , ki ju uporabimo na modelu, tako da se

$$T_{uvw} R_{\alpha\beta\gamma}(L_{0i}) = L_{s1} \quad \text{projicirajo na } l_{0i}, \quad i = 1, 2, 3$$

Problem rešimo v dveh korakih: najprej poiščemo rotacijo in nato še translacijo telesa. Rotacija je sestavljena iz zavrtitev za kote  $\alpha$  okrog osi  $x$ ,  $\beta$  okrog osi  $y$  in  $\gamma$  okrog osi  $z$ ; predstavljena je na običajen način (glej [11]). Če označimo z  $V_{01}, V_{s1}$  smerne vektorje premic  $L_{01}, L_{s1}$ , mora veljati:

$$V_{s1} = R_{\alpha\beta\gamma} V_{01}, \quad i = 1, 2, 3.$$

Zaradi (4.1) rešujemo sistem enačb

$$N_{s1} \cdot R_{\alpha\beta\gamma} V_{01} = 0, \quad i = 1, 2, 3.$$

za neznanke  $\alpha, \beta, \gamma$ . Sistem poenostavimo z uvedbo dveh novih koordinatnih sistemov  $(S_{1m})$  in  $(S_{1v})$ . Tako bomo imeli 4 koordinatne sisteme

$$S_{0m} \xrightarrow{R_m} S_{1m} \xrightarrow{R_{\alpha\beta\gamma}} S_{1v} \xrightarrow{R_v} S_{0v} \quad \text{in} \quad R_{\alpha\beta\gamma} = R_v \circ R_{\alpha\beta\gamma} \circ R_m$$

V koordinatnem sistemu  $(S_{0m})$  imamo normirane vektorje  $v_{01}, v_{02}, v_{03}$  (smerni vektorji na premicah  $l_{01}, l_{02}, l_{03}$ ). Novi koordinatni sistem  $(S_{1m})$  ima tako bazo, da je prvi bazni vektor  $e_{11} := v_{01}$ , drugi bazni vektor leži v ravnini  $v_{01}, v_{02}$ , torej  $e_{12} := e_{11} \times e_{13}$  in tretji bazni vektor je pravokoten na to ravnino

$e_{13} := v_{01} \times v_{02}$ . Prehodna matrika je

$$R_m = (v_{01}, v_{01} \times (v_{01} \times v_{02}), v_{01} \times v_{02})^{-1}.$$

Uvedemo še nov koordinatni sistem  $(S_{1v})$ , tak da  $(x, y)$  ravnina ustreza interpretacijski ravnini in da je os  $x$  kolinearna z  $e_{01}$ . Če ima smerni vektor  $v_{01}$  premice  $l_{01}$  koordinate  $v_{01} = (a_{01}, b_{01}, 0)$  in točka  $p_{01}$  na  $l_{01}$  koordinate  $p_{01} = (x_{01}, y_{01}, f)$ , lahko transformacijsko matriko  $R_v$  med  $(S_{1v})$  in  $(S_{0v})$  izrazimo

$$R_v = \frac{1}{\sqrt{f^2 + d_{01}^2}} \begin{pmatrix} \sqrt{f^2 + d_{01}^2} & 0 & 0 \\ 0 & f & d_{01} \\ 0 & -d_{01} & f \end{pmatrix} \begin{pmatrix} a_{01} & -b_{01} & 0 \\ b_{01} & a_{01} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

kjer je  $d_{01} = a_{01}y_{01} - b_{01}x_{01}$ .

Z uvajanjem novih koordinatnih sistemov smo dosegli, da iščemo rotacijo kot preslikavo med  $(S_{1m})$  in  $(S_{1v})$  v obliki  $R_{\alpha\beta}$  (zavrtitev za kot  $\alpha$  okrog osi  $x$  in zavrtitev za kot  $\beta$  okrog osi  $y$ ).

Veljati mora  $V_{s1} = R_{\alpha\beta} V_{11}$ . Zaradi (4.1) rešujemo sistem

$$N_{s1} V_{s1} = N_{s1} R_{\alpha\beta} V_{11}, \quad i = 1, 2, 3, \quad \text{za neznanke } \alpha, \beta.$$

Od teh treh enačb sta uporabni le dve, ki ju v [6] prevedejo v obliko

$$\sum_{i=0}^s \delta_i t^i = 0, \quad (4.2)$$

kjer je  $t = \tan \alpha/2$ ,  $\delta_i$  pa so znani koeficienti.

Ko rešimo enačbo (4.2), poiščemo še  $\beta$ , ki je izražen v obliki  $\cos \beta = f(\alpha)$ . Celotna rotacija je  $R_{\alpha\beta\gamma} = R_v R_{\alpha\beta} R_m$ .

Če imamo opravka s koplanarnimi premicami ali s premicami, ki imajo eno točko skupno, je polinom (4.2) le četrte stopnje.

Izračunajmo zdaj še translacijo  $T_{uvw}$ ! Naj bo  $t := (u, v, w)$  vektor translacije. Vzemimo, da za vsako premico  $L_{0i}$  poznamo tudi točko  $P_{0i}$  na njej. Točke  $P_{0i}$  se morajo po rotaciji in translaciji projicirati na premice  $l_{0i}$ , torej morajo zadoščati pogoju (4.1):

$$(R(P_{0i}) + t) \cdot N_{s1} = 0, \quad i = 1, 2, 3 \quad (4.3)$$

Tako imamo linearen sistem treh enačb za neznanke  $u, v, w$ , ki ga brez težav rešimo. V posebnem primeru, ko se premice  $L_{0i}$  stikajo v skupni točki  $P_{00}$ , imamo v zgornjem sistemu (4.3) le dve linearno neodvisni enačbi, zato si tretjo enačbo dobimo takole: imejmo točke  $T_{0i}$  (drugo oglišče roba, ki leži na  $L_{0i}$ ) in njihove slike  $t_{s1}$ . Veljati mora

$$P(R(T_{0i}) + t) = t_{s1}, \quad i = 1, 2, 3$$

Če je torej v celoti viden vsaj eden od robov, lahko določimo translacijo. Če je v celoti vidnih več robov, izberemo tisto translacijo, ki ima najmanjšo vrednost vzdolž osi  $z$ . V praksi je težko določiti, kdaj je rob objekta v celoti viden, zato se v takih primerih

na izračunano translacijo ne smemo preveč zanašati.

Po tej metodi dobimo za translacijo in rotacijo več rešitev. Vse te rešitve moramo preveriti, ali so prave, še prej pa jih lahko veliko večino izločimo s temi preprostimi pravili:

- translacija vzdolž osi  $z$  ne sme biti negativna
- izbrani robovi morajo biti na sliki vidni [9]
- rešitev pri kateri slika transformiranega roba ne leži na sliki roba, ne pride v poštev.

## 5. LOWEJEVA METODA

Pri tem pristopu so linearizirane enačbe projekcije in uporabljena je Newtonova metoda za potrebno število iteracij. Še prej pa so enačbe reparametrizirane, tako da je poenostavljeno računanje potrebnih odvodov.

Zapišimo preslikave

$$(x, y, z) = R(p - t), \quad (u, v) = \left( \frac{fx}{z}, \frac{fy}{z} \right) = P(x, y, z), \quad (5.1)$$

kjer je  $t$  translacijski vektor v  $R^3$ ,  $R$  rotacija, ki zavrti model iz osnovnega položaja v položaj, ki ga ima telo na sceni,  $f$  goriščna razdalja in  $(u, v)$  koordinati točke na sliki, kamor se projicira točka  $p$  iz modela.

Pri danih točkah  $p_i$ ,  $i = 1, \dots, n$  modela in njihovih slikah  $(u_i, v_i)$  na projekcijski ravnini iščemo tak premik  $t$  in rotacijo  $R$ , da bo

$$PR(p_i - t) = (u_i, v_i), \quad i = 1, \dots, n \quad (5.2)$$

Da bi lažje računali parcialne odvode, reparametriziramo enačbi (5.1) takole:

$$(x, y, z) = R \cdot p$$

$$(u, v) = \tilde{P}(x, y, z) = \left( \frac{fx}{z + D_x} + D_x, \frac{fy}{z + D_x} + D_y \right) \quad (5.3)$$

Spremenljivki  $R, f$  sta isti kot v enačbah (5.1), vektor  $t$  pa smo nadomestili z vektorjem  $(D_x, D_y, D_z)$ . Transformaciji (5.1), (5.3) sta ekvivalentni, če je

$$t = R^{-1} \left[ -\frac{D_x(z + D_x)}{f}, -\frac{D_y(z + D_x)}{f}, -D_z \right]^T.$$

$D_x$  in  $D_y$  določata položaj telesa na sliki,  $D_z$  pa oddaljenost objekta od kamere.

Kako predstaviti rotacijo  $R$  s parametri  $\varphi_x, \varphi_y, \varphi_z$  (zavrtitev o krog  $x, y, z$  osi)? Izhajamo iz začetnega približka  $R_0$ , ob spremembi kotov  $\varphi_x, \varphi_y, \varphi_z$  za  $\Delta\varphi_x, \Delta\varphi_y, \Delta\varphi_z$  pa dobimo novo rotacijo  $R_1$  takole:

$$R_1 = \begin{bmatrix} \cos \Delta\varphi_x & -\sin \Delta\varphi_x & 0 \\ \sin \Delta\varphi_x & \cos \Delta\varphi_x & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos \Delta\varphi_y & 0 & \sin \Delta\varphi_y \\ 0 & 1 & 0 \\ -\sin \Delta\varphi_y & 0 & \cos \Delta\varphi_y \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \Delta\varphi_z & -\sin \Delta\varphi_z \\ 0 & \sin \Delta\varphi_z & \cos \Delta\varphi_z \end{bmatrix} R_0$$

Če so spremembe kotov  $\Delta\varphi_x, \Delta\varphi_y, \Delta\varphi_z$  dovolj majhne, so približno neodvisne druga od druge, tako da lahko upoštevamo zveze:  $(x, y, z) = (r_x \cos \varphi_x, r_x \sin \varphi_x, z) = (x, r_x \cos \varphi_x, r_x \sin \varphi_x) = (r_y \sin \varphi_y, y, r_y \cos \varphi_y)$ , kjer so  $r_x, r_y, r_z$  razdalje točke od osi  $x, y, z$ . Od tod izračunamo parcialne odvode  $x, y, z$  glede na  $\varphi_x, \varphi_y, \varphi_z$ :

	$x$	$y$	$z$
$\varphi_x$	0	-z	y
$\varphi_y$	z	0	-x
$\varphi_z$	-y	x	0

in nato z zaporednim odvajanjem še parcialne odvode  $u$  in  $v$  glede na iskane parametre  $D_x, D_y, D_z, \varphi_x, \varphi_y, \varphi_z$ .

	$u$	$v$
$D_x$	1	0
$D_y$	0	1
$D_z$	$-fc^2x$	$-fc^2y$
$\varphi_x$	$-fc^2xy$	$-fc(z + cy^2)$
$\varphi_y$	$fc(z + cx^2)$	$fc^2xy$
$\varphi_z$	$-fcy$	$fcx$
$h$	$cx$	$cy$

kjer je  $c := (z + D_x)^{-1}$ . Naj bo

$$h := [\Delta D_x, \Delta D_y, \Delta D_z, \Delta\varphi_x, \Delta\varphi_y, \Delta\varphi_z].$$

Če je neznana tudi goriščna razdalja, dodamo vektorju  $h$  še  $\Delta f$ . Zdi se lahko direktno uporabimo Newtonovo metodo: imejmo začetne približke  $D_{x_0}, D_{y_0}, D_{z_0}, \varphi_{x_0}, \varphi_{y_0}, \varphi_{z_0}$  za količine  $D_x, D_y, D_z, \varphi_x, \varphi_y, \varphi_z$ . Od tod izračunamo začetni približek  $\tilde{P}_0, R_0$  za  $\tilde{P}$  oziroma  $R$ . Velja:

$$\tilde{P}_0 R_0(p_i) = (u_{i0}, v_{i0}) = (E_{u_0} + u_i, E_{v_0} + v_i), \quad i = 1, \dots, n$$

kjer sta  $E_{u_0}$  in  $E_{v_0}$  napaki, ki ju naredi začetna preslikava  $\tilde{P}_0 R_0$ . Za napako  $E_{u_0}$  vzamem

$$E_{u_0} = \frac{\partial u_i}{\partial D_x} \Delta D_x + \frac{\partial u_i}{\partial D_y} \Delta D_y + \frac{\partial u_i}{\partial D_z} \Delta D_z + \frac{\partial u_i}{\partial \varphi_x} \Delta\varphi_x + \frac{\partial u_i}{\partial \varphi_y} \Delta\varphi_y + \frac{\partial u_i}{\partial \varphi_z} \Delta\varphi_z, \quad i = 1, \dots, n \quad (5.4)$$

Podoben izraz velja za  $E_{v_0}$ . Če imamo korespondenco med temi točkami ( $n = 3$ ), imamo linearen sistem 6 enačb, ki jih rešimo za neznanke  $\Delta D_x, \Delta D_y, \Delta D_z, \Delta\varphi_x, \Delta\varphi_y, \Delta\varphi_z$ . Določimo novo preslikavo  $\tilde{P}_1 R_1$  in ponovimo iteracijo. Po [8] je potrebnih le nekaj iteracij, da pridemo do dovolj natančnega rezultata.

Če imamo imamo več povezav med točkami ( $n > 3$ ), uporabimo za reševanje sistema metodo najmanjših kvadratov. Zapišimo sistem (5.4) bolj kompaktno:

$$Jh = e,$$

ker je  $J$  Jacobijeva matrika, ki vsebuje parcialne odvode,  $h$  vektor, ki vsebuje neznane funkcije in  $e$  vektor napak, ki smo jih izmerili na sliki. Rešimo normalni sistem

$$J^T Jh = J^T e.$$

Tako dobimo dokaj natančno rešitev, če imamo dober začetni približek. Poleg tega ta metoda konvergira, če imamo znanih veliko ujemanj med robovi in njihovimi slikami, kar zahteva visoko kompleksnost pri iskanju pravega ujemanja med robovi modela in njihovimi slikami. Zato je dobro po prejšnji metodi (Dhome) dobiti vse možne transformacije z interpretacijo treh robov. Na ta način opustimo hipotezo (da se slike robov ujemajo z izbranimi črtami na sliki) ali pa dobimo zanjo dober začetni približek.

## 6. ZAKLJUČEK

Ogledali smo si štiri metode za določanje položaja telesa iz ene same 2-D slike. Različni avtorji so vsak na svoj način poenostavljali reševanje nelinearnih transcendentnih enačb.

Roberts operira s samimi obrnljivimi operatorji zaradi uvedbe homogenih koordinat in rešuje linearen problem, ker predpostavi, da so elementi transformacijske matrike med sabo neodvisni. Ko dobi rešitev, mora preveriti medsebojno odvisnost matričnih elementov. Za rešitev potrebuje najmanj šest točk modela in njihovih slik.

Horaud rešuje sistem dveh transcendentnih enačb z dvema neznančkama s tabeliranjem funkcij dveh spremenljivk, konstruiranjem nivojnic teh funkcij in iskanjem njihovih presečišč. To je precej zamuden postopek, ki nam da rezultate z omejeno natančnostjo. Za določitev petih prostostnih stopenj potrebuje oglišče modela, iz katerega izhajajo trije robovi in slika tega oglišča in robov, za določitev šeste prostostne stopnje pri translaciji pa potrebuje še eno tako oglišče.

Dhome in soavtorji morajo poiskati ničle polinoma 8 stopnje, kar tudi ni trivialno. Do polinoma pridejo po uvedbi takih koordinatnih sistemov, da iščejo rotacijo kot funkcijo dveh in ne treh kotov. Na ta način dobimo vse možne rešitve, ki nam lahko služijo za začetni približek pri Lowejevi metodi, ki nam ob dobrem začetnem približku da natančnejši rezultat. Za določitev rotacije potrebujemo tri smerne vektorje iz robov modela ter smerne vektorje iz slik robov ter po eno točko na vsaki sliki roba. Če se obravnavani robovi modela ne stikajo v skupnem oglišču, moramo za določitev translacije za vsak izbrani rob poznati eno njegovo oglišče, če pa se obravnavani robovi modela stikajo v skupni točki, moramo vsaj za en rob poznati še drugo oglišče in njegovo sliko na projekcijski ravnini.

Lowe rešuje problem iterativno z Newtonovo metodo; računanje parcialnih odvodov močno poenostavi z uvedbo novih spremenljivk. Za rešitev potrebuje korespondenco med vsaj tremi točkami modela in slike ter dovolj dober približek za začetno rešitev. Natančnost dobljene transformacije lahko povečamo, če poznamo več točk modela in njihovih slik.

Robertsovo metodo smo si ogledali, ker je ena prvih; naslednji dve metoda nam dasta vse rešitve, vendar z omejeno natančnostjo, medtem ko z Lowejevo metodo dobljeno rešitev lahko izboljšamo.

Zanimivo bi bilo vedeti, kako se spremeni rešitev problema, če malo premaknemo slike točk. Tako bi vedeli, katera metoda se najboljše obnese pri realnih slikah, ki imajo tudi šume. Zanimive bi bile tudi ocene, kako natančne so dobljene transformacije in kako so opisane metode hitre.

## LITERATURA

- [1] L. G. Roberts: Machine perception of three-dimensional solids, *Optical and Electro Optical Information Processing*, J.T. Tippet; Ed. Cambridge, MA: MIT Press, 1965, pp. 159-197
- [2] T. Kanade: Recovery of the three dimensional shape of an object from a single view, *Artificial Intell.*, Special Volume on Computer Vision, vol. 17, no. 1-3, Aug. 1981
- [3] S. T. Barnard: Choosing a basis for perceptual space, *Comput. Vision Graph. Image Processing*, vol. 29, no. 1, pp. 87-99, 1985
- [4] T. Shakunagar, H. Kaneko: Perspective angle transform and its application to 3-D configuration recovery, *Proc. Int. Comput. Vision Pattern Recogn.*, Miami Beach, FL, June 1986, pp. 594-601
- [5] R. Horaud: New methods for matching 3-D objects with single perspective views, *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. PAMI-9, No. 3, May 1987
- [6] M. Dhome et al: Determination of the attitude of 3-D objects from a single perspective view, *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 11, No. 12, Dec. 1989
- [7] D. P. Huttenlocher, S. Ullman: Object recognition using alignment, in *Proc. First Int. Conf. Comput. Vision*, London, England, June 1987, pp. 102-111.
- [8] D. G. Lowe: Three-dimensional object recognition from single two-dimensional images, *Artificial Intell.*, 31(1987), pp. 355-395
- [9] D. G. Lowe: Perceptual organisation and visual recognition, Boston, MA: Kluwer, 1985, ch. 7
- [10] Zen Chen et al: A simple vision algorithm for 3-D position determination using a single calibration object, *Pattern Recognition*, vol. 22, no. 2, 1989
- [11] Yoshiaki Shirai: Three-dimensional computer vision, Springer Verlag 1987, ch. 5