

(12)

## CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: **a 2012 00126**

(22) Data de depozit: **27.02.2012**

(41) Data publicării cererii:  
**30.08.2012** BOPI nr. **8/2012**

(71) Solicitant:  
• UNIVERSITATEA POLITEHNICĂ DIN  
TIMIȘOARA, STR. PIAȚA VICTORIEI NR.2,  
TIMIȘOARA, TM, RO

(72) Inventatori:  
• GUI VASILE, STR. CÂNTULUI NR. 20,  
AP. 12, TIMIȘOARA, TM, RO;  
• ALEXA FLORIN, STR. IOACHIM MILOIA  
NR. 14, TIMIȘOARA, TM, RO;

• CĂLEANU CĂTĂLIN-DANIEL,  
STR. CONSTANȚA NR. 2, TIMIȘOARA, TM,  
RO;  
• POPA GHEORGHE-DANIEL, ALEEA ICAR  
NR. 12, ET. 3, AP. 7, TIMIȘOARA, TM, RO;  
• DAVID CIPRIAN, STR. BRÂNDUȘA,  
BL. G7, AP. 3, SATU MARE, SM, RO;  
• SIMION GEORGIANA, STR. COTUL MIC  
NR. 27, LUGOJ, TM, RO

(54) **METODĂ DE URMĂRIRE A DEGETELOR MĂINII ȘI MIJLOC  
DE COMUNICARE OM- CALCULATOR**

(57) Rezumat:

Invenția se referă la o metodă de urmărire a degetelor mâinii și la un mijloc de comunicare om-calculator, bazat pe gesturi efectuate cu mâna. Metoda conform invenției constă într-o succesiune de operații de prelucrare numerică a semnalului generat de o cameră (1) video monoculară, conectată sau inclusă în componența unui calculator PC, ce permite urmărirea degetelor extinse ale unei mâini, care gesticulează pe un fundal complex, la o distanță variabilă în raport cu camera (1) video, inițializarea sistemului se realizează automat, la expunerea mâinii deschise în fața camerei (1) video; pentru fiecare deget, se generează un set de patru parametri, ce permit generarea unor traiectorii 4D, pe baza secvențelor de imagini 2D.

Revendicări: 1  
Figuri: 3

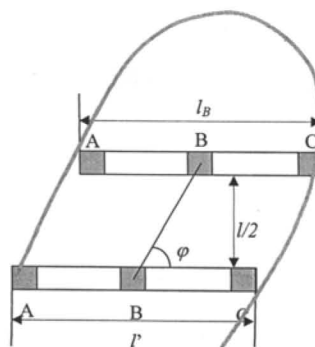


Fig. 1



36

## METODĂ DE URMĂRIRE A DEGETELOR MĂINII ȘI MIJLOC DE COMUNICARE OM-CALCULATOR

Invenția se referă la o metodă de urmărire a degetelor mâinii și un mijloc de comunicare om-calculator bazat pe gesturi efectuate cu mâna.

Urmărirea degetelor mâinii este o componenta esențială a sistemelor de comunicare om-calculator bazate pe gesturi efectuate cu mâna. Sunt cunoscute mai multe metode de urmărire a mâinii. Metodele („model based”) bazate pe modele articulate și analiza prin sinteză [B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla. Model-based hand tracking using a hierarchical bayesian filter. IEEE Trans. Pattern Analysis and Machine Intell., 28(9):1372–1384, 2006],[ M. de la Gorce, D. Fleet & N. Paragios. Model-based 3D Hand Pose Estimation from Monocular Video. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 33 (9), 2011, pp 1793-1805] permit caracterizarea cea mai nuanțată a posturii mâinii, în particular a poziției degetelor și au o robustețe bună la acțiuni perturbatoare, cum sunt camuflajul sau ocluziile parțiale, dar au un cost de calcul ridicat, ceea ce a împiedicat până în prezent răspândirea lor în aplicații în timp real. Metodele („view based”) bazate pe aspect sunt eficiente din punctul de vedere al calculelor, conducând la implementări de timp real. Intre acestea, unele detectează o poziție globală a mâinii [C. Shan, T. Tan, Y. Wei, Real-time hand tracking using a mean shift embedded particle filter, in Pattern Recognition No. 40, 2007, pp 1958-1970][ M. Kolsch and M. Turk, “Fast 2D hand tracking with flocks of features and multi-cue integration,” in IEEE Workshop on Real-Time Vision for Human-Computer Interaction, 2004, pp. 158–165], ceea ce conduce la o definiție globală, mai puțin exactă a posturii, limitând aria de aplicații posibile. Asemenea metodei ce face obiectul invenției, în [K. Oka, Y. Sato, and H. Koike. Real-time fingertip tracking and gesture recognition. IEEE Computer Graphics and Applications, 22(6):64–71, 2002] și [Song P., Winkler S., Gilani S. O. & Zhou Z. (2007). Vision-Based Projected Tabletop Interface for Finger Interactions. In Lecture Notes in Computer Science, 4796, 49-58], postura mâinii se definește pe baza poziției degetelor. Ambele se bazează pe detecția vârfurilor degetelor. Niciuna din ele nu este invariantă la scară, ceea ce le face impracticabile în aplicațiile la care distanța între persoana care gesticulează și camera se modifică apreciabil pe durata comunicației. Metoda [K. Oka, Y. Sato, and H. Koike. Real-time fingertip tracking and gesture recognition. IEEE Computer Graphics and Applications,

22(6):64–71, 2002] are dezavantajul ca presupune cunoașterea fundalului, care este un panou informatic. Metoda [Song P., Winkler S., Gilani S. O. & Zhou Z. (2007). Vision-Based Projected Tabletop Interface for Finger Interactions. In Lecture Notes in Computer Science, 4796, 49-58] se bazează pe detecția vârfurilor și folosește segmente de dreapta pentru a detecta prezența degetului în vecinătatea vârfului, în scopul eliminării detecțiilor fals pozitive, dar este are dezavantajul ca este vulnerabilă la ocluzii parțiale pe porțiunea superioară a degetului deoarece metoda de detecție a segmentelor este secvențială.

**Scopul** invenției este de a obține o metodă de urmărire a degetelor mâinii, ce poate fi utilizată pentru o interfață pentru comunicare om-calculator bazată pe gesturi efectuate cu mâna, care să înlăture dezavantajele identificate, permițând urmărirea în timp real a degetelor deschise ale mâinii, efectuate pe un fundal variabil, în condiții de ocluzie parțială, de la o distanță variabilă față de cameră, generând vectori de date 4D pentru fiecare deget detectat.

Metoda propusă are la bază detecția maximelor funcției densitate de probabilitate ale unor caracteristici special concepute pentru caracterizarea locațiilor din imagine care conțin degete și care sunt prezente în număr mare în regiunea degetelor. Denumim aceste caracteristici „fingerlets”. Robustețea este conferită de faptul că pozițiile maximelor de densitate a caracteristicilor nu sunt afectate de pierderea unor eșantioane sau de prezenta unor eșantioane deviate, ce nu aparțin țintei reale. Definim în continuare caracteristicile „fingerlet” pentru scanare orizontală. Pentru scanare verticala, se procedează similar. Presupunem ca la momentul detecției este disponibilă o estimatie a grosimii degetului,  $l_e$ .

Conform invenției, o caracteristica „fingerlet” este prezenta la coordonata  $B = (x_B, y_B)$  a unei imagini, daca exista șase puncte,  $A, B, C, A', B', C'$ , ilustrate în Fig. 1, pentru care sunt întrunite simultan următoarele condiții:

1. Punctele  $A, B, C$  aparțin aceleiași linii de imagine.
2. Distanța între  $A$  și  $C$ ,  $l$ , nu se abate de la grosimea estimata a degetului,  $l_e$ , cu mai mult decât un procent  $p\%$  (valorile uzuale sunt  $\pm 50\%$  din valoarea estimată a grosimii degetului  $l_e$ ).
3. Punctele  $A$  și  $C$  sunt puncte de tranziție complementare, caracterizate prin gradienti orizontali de semn contrar.
4. Punctul  $B$ , aflat la jumătatea segmentului  $AC$ , nu este un punct de tranziție.

5. Pe linia decalata vertical cu jumătate din lungimea segmentului AC, exista trei puncte A',B',C' care întrunesc condițiile 1,2,3 si 4 și decalajul orizontal între cei doi tripleți nu depășește grosimea estimată a degetului.

Caracteristica detectată este descrisă prin vectorul  $(x,y,l,\varphi)$ , cu  $x=x_B, y=y_B, l=l_B, \varphi=\varphi_B$ , în care  $\varphi_B$  reprezintă unghiul dintre dreapta determinata de punctele BB' si o linie de scanare.

Pentru inițializare, utilizatorul stabilește grosimile minime si maxime ale degetului, exprimate in pixeli,  $l_{min}, l_{max}$ , în funcție de rezoluția camerei, parametrii sistemului optic și gama distanțelor la care se presupune a se afla utilizatorul sistemului. Inițializarea sistemului de urmărire se realizează automat, conform invenției, prin următoarele operații de prelucrare a imaginii cu utilizatorul prezentând palma în postura cu toate degetele extinse, în plan vertical:

1. Segmentare prim plan/fundal.
2. Extragere de caracteristici „fingerlets” din imaginea binara rezultata, cu parametrul de grosime a degetului,  $l$ , cuprins in intervalul  $[l_{min}, l_{max}]$ .
3. Determinarea histogramelor parametrilor,  $x, y$  si  $l$  ale caracteristicilor „fingerlets”.
4. Extragerea modului grosimii  $l_e$ , folosind estimarea nonparametrică a densității de probabilitate in histograma parametrului  $l$ , cu fereastra rectangulara având lățimea egală cu  $q\%$  din intervalul  $|l_{max}-l_{min}|$  (valorile uzuale sunt în intervalul  $\pm 50\%$ ).
5. Validarea modului si a prezentei mâinii, dacă numărul eșantioanelor asociate modului raportat la grosimea estimata a degetului depășește un prag de detecție.
6. Definirea regiunii de interes, ROI, prin centrul  $(x_e, y_e)$  reprezentat de modurile parametrilor  $x$  si  $y$  ale caracteristicilor „fingerlet” asociate modului grosimii prin algoritmul mean shift și prin dimensiuni (înălțime și lățime) egale cu un multiplu al grosimii degetului,  $n \times l_e$ .

Validarea prezenței mâinii determină trecerea sistemului în regimul de urmărire a mâinii.

Urmărirea mâinii, conform invenției, presupune următorii pași de prelucrare pentru fiecare cadru de imagine:

1. Segmentarea prim plan / fundal in regiunea de interes.
2. Extragerea de caracteristici „fingerlets” in imaginea rezultata.
3. Detecția si segmentarea degetelor folosind caracteristicile „fingerlets” .

4. Actualizarea modului grosimii  $l_e$ , folosind estimarea nonparametrică a densității de probabilitate în histograma parametrului  $l$ , cu fereastra rectangulară având lățimea egală cu  $q\%$  din intervalul  $|l_{\max}-l_{\min}|$  (valorile uzuale sunt în intervalul  $\pm 50\%$ ).

5. Actualizarea regiunii de interes, ROI, prin centrul  $(x_e, y_e)$  reprezentat de modurile parametrilor  $x$  și  $y$  ale caracteristicilor „fingerlet” asociate modului grosimii și prin lățimea egală cu un multiplu al grosimii degetului,  $n \times l_e$ .

În ambele situații, de inițializare și urmărire, alternativ, la pasul 1 se poate utiliza o segmentare bazată pe culoarea pielii sau detecție de muchii.

Extragerea eficientă a caracteristicilor „fingerlets”, conform invenției, constă în operațiile următoare:

1. Se calculează  $l_{\max} = l_e (1+\epsilon)$  și  $l_{\min} = l_e (1-\epsilon)$ , unde  $\epsilon$  este un număr subunitar, reprezentând toleranța abaterilor de grosime, ce definesc fereastra de detecție de caracteristici „fingerlets”.
2. Se scanează regiunea de interes, în ordine lexicografică, până la identificarea primei tranziții fundal-prim plan. Fie  $(x_A, y_A)$  coloana și linia primei tranziții, ce definesc un punct de tip A.
3. Dacă s-a găsit un punct de tip A, se continuă scanarea liniei curente, până la identificarea unei tranziții prim plan – fundal, la locația C  $(x_C, y_A)$ . Dacă distanța între C și A,  $l_B = |x_C - x_A|$  se încadrează în intervalul  $[l_{\min}, l_{\max}]$ , se determină poziția punctului de mijloc, B  $(x_B, y_B)$ , cu  $y_B = y_A$ ,  $x_B = (x_B + x_C)/2$  și se salvează, pe o listă asociată liniei  $y_A$ , segmentul candidat corespunzător punctului B, definit prin parametri  $(x_B, y_B, l)$ .
4. Se continuă scanarea regiunii de interes de la ultimul punct de tip C, pentru detectarea de noi segmente candidat.
5. Se parcurg în ordinea generată listele conținând parametri de segmente candidat. Pentru fiecare segment candidat,  $(x_B, y_B, l_B)$ , se caută segmente candidat pe lista segmentelor asociate liniei  $y_B + l_B/2$ . Dacă linia respectivă conține un segment candidat, corespunzător punctului, B', de coordonate  $(x_{B'}, y_{B'}, l_{B'})$ , cu decalajul  $|x_{B'} - x_B| < l_B$ , celor două segmente li se asociază o caracteristică „fingerlet” validă, cu vectorul de parametri  $(x_B, y_B, l_B, \varphi_B)$ .

Detecția și segmentarea degetelor folosind caracteristicile „fingerlets”, conform invenției, se realizează prin succesiunea următorilor pași de prelucrare:

1. Se parcurge lista caracteristicilor „fingerlet”, în ordinea generată.

2. Prima caracteristica identificata „fingerlet”, având parametrii  $(x_B, y_B, l, \varphi_B)$ , definește un segment de deget candidat și se marchează. Se reține poziția curentă în lista. Se continuă parcurgerea listei, începând de la linia următoare și  $i$  se asociază degetului candidat toate caracteristicile fingerlet nemarcate pentru care direcția locală,  $\varphi$ , satisface condiția de similaritate  $|\varphi - \varphi_B| < \varphi_T$ , unde  $\varphi_T$  reprezintă o valoare de prag pentru similaritatea orientării. Se continuă de la poziția curentă în listă, pentru extragerea tuturor degetelor candidat, la fel ca mai sus.
3. Degetele candidat cu un număr de caracteristici „fingerlets” mai mici decât, un prag definit ca un multiplu al grosimii estimate a degetului,  $k \cdot l_e$ , se consideră nevalide.
4. Caracteristicile fiecărui segment de deget valid se grupează folosind propagare morfologica pe lista coordonatelor asociate segmentului. Dacă se identifică mai multe sub-segmente, pentru fiecare se calculează parametrii  $(\rho, \varphi)$  corespunzători reprezentării în coordonate polare a dreptei definite de segment. Dacă diferența razelor,  $\rho$ , este mai mare decât grosimea unui deget, sub-segmentele valide definesc degete distincte.
5. Fiecare deget detectat, este caracterizat printr-un vector din spațiul 4D de forma  $(x, y, \varphi, l)$ , unde  $(x, y)$  sunt coordonatele spațiale ale primei caracteristici „fingerlet” asociate degetului,  $\varphi$  reprezintă valoarea medie a orientărilor locale ale caracteristicilor „fingerlet” asociate degetului și  $l$  grosimea medie a degetului.

Invenția prezintă următoarele **avantaje**:

Prin faptul că majoritatea operațiilor implicate în detecția și urmărirea mâinii se realizează pe lista de caracteristici „fingerlet” și nu la nivel de pixeli ai imaginii, invenția permite obținerea unui timp de execuție extrem de redus în comparație cu timpii raportați în literatură. În consecință, este posibilă rularea cu camere video de viteză mare și evitarea distorsiunilor de mânăjire a imaginii la mișcare rapidă („motion blur”) ce pot determina pierderea țintei. În același timp, rămâne timp de prelucrare mai mare pentru sistemul de interpretare a informației furnizate de sistemul de urmărire. Alternativ, permite rularea în timp real pe procesoare lente, specifice aplicațiilor mobile.

În raport cu sistemele de urmărire bazate pe aspect („view based”), care furnizează doar coordonate ale vârfurilor degetelor, metoda de urmărire a degetelor

mâinii furnizează informație 4D pentru fiecare deget extins. În consecință, gama gesturilor ce pot fi definite și diferențiate fără erori se extinde semnificativ.

Datorită faptului că detecția degetelor se bazează pe o tehnică de grupare robustă a caracteristicilor „fingerlet”, ce nu presupune conectivitatea spațială a acestora, metoda este rezistentă la erori de segmentare produse de ocluzie parțială sau camuflaj.

Se dă în continuare un **exemplu de aplicare a invenției**.

Un exemplu mijloc de comunicare om-calculator bazat pe gesturi care folosește metoda de urmărire a degetelor mâinii, ce face obiectul propunerii de brevet este ilustrat în Fig. 2. Blocul 1 este o camera video, conectată sau inclusă în componența unui calculator PC, pe care se implementează componentele software reprezentate prin blocurile de prelucrare succesive 2, 3 și 4. Blocul 2 realizează preprocesarea și segmentarea imaginii de prim plan. Poate fi realizat prin substrație de fundal, detecție de piele bazată pe culoare, detecție de muchii sau orice combinație a acestor operații. Blocul 3 este sistemul de urmărire a mâinii. Acesta extrage caracteristici „fingerlet” din punctele de discontinuitate generate de Blocul 2, detectează degete și generează traiectorii în spațiul 4D. Blocul 4 exploatează informația furnizată de blocul 3 și determină acțiuni. Este în esență un clasificator de forma.

O aplicație posibilă a invenției constă în comanda prin gesturi efectuate cu mâna a unui panou informatic de mari dimensiuni, pentru selecția de informație. Panoul poate fi amplasat într-un muzeu, o stație de tren sau metrou, aerogara, magazin etc. Coordonatele indexului pot servi pentru selecția informației, înclinația lui sau numărul de degete extinse pot servi pentru confirmare sau infirmare de acțiuni, revenire la acțiunea anterioară, inițierea unei noi comenzi etc. Pentru fiecare deget extins, conform invenției, se generează o traiectorie în spațiul 4D. În Fig. 3 se exemplifică traiectoriile în subspațiul 3D, conținând informațiile de poziție și orientare, fără componenta de grosime, ce furnizează informații referitoare la modificarea distanței mâinii față de cameră. În figură, coordonatele  $x$  și  $y$  semnifică pozițiile ale vârfului degetului, măsurate în pixeli, în timp ce coordonata  $z$  semnifică unghiul de înclinație al degetului în plan vertical, măsurat în grade. Primul gest dinamic pentru care s-a extras traiectoria, redată în Fig. 3a, corespunde unei mișcări circulare. Următoarele două gesturi, compuse din mișcări cu traiectorii liniare pe porțiuni, au

trajectoriile reprezentate în Fig. 3b și Fig. 3c. Spre deosebire de primele trei gesturi, efectuate cu o orientare a degetului aproximativ constantă, la ultimul gest cu traiectoria prezentată în Fig. 3d, utilizatorul execută, cu degetul activ, o mișcare asemănătoare celei efectuate de un ștergător de parbriz, degetul pivotând în jurul vârfului. Se remarcă variația în limite largi a unghiului de inclinație și mișcarea redusă a vârfului. Trajectoriile exemplificate au fost extrase folosind un calculator de uz personal de tip laptop, cu tactul de 2,26 GHz, dotat cu camera web ce furnizează imagini cu rezoluția de 640×480 pixeli. Timpul consumat de sistemul de urmărire ce face obiectul invenției, a fost de 1,4 ms pe cadru. Substracția de fundal a necesitat un timp mediu de 4.5 ms pe cadru, ceea ce a condus la un timp mediu de prelucrare de 5.9 ms pe cadru. Folosind o camera video rapidă, sistemul poate rula la frecvențe mai mari de 150 de cadre pe secundă. Alternativ, se poate obține funcționarea în timp real cu procesoare mai puțin performante.



## REVENDICĂRI

Metodă de urmărire a degetelor mâinii, caracterizată prin utilizarea caracteristicii „fingerlet” orizontală respectiv verticală pentru detecția și urmărirea degetelor mâinii în timp real într-un sistem de comunicare om-calculator. Caracteristica „fingerlet” poate fi extrasă din imaginea binară obținută prin: segmentarea prim plan/fundal, segmentarea piele/non-piele, detecția de muchii sau orice combinație a acestor metode. Caracteristica este definită în spațiul  $(x, y, l, \varphi)$  și este considerată prezentă la coordonata  $B = (x_B, y_B)$  a unei imagini, dacă există șase puncte, A, B, C, A', B', C', pentru care sunt întrunite simultan următoarele condiții: punctele A, B, C aparțin aceleiași linii respectiv coloane de imagine; distanța între A și C,  $l_B$ , nu se abate de la grosimea estimată a degetului,  $l_e$ , cu mai mult decât un procent  $p\%$ ; punctele A și C sunt puncte de tranziție complementare, caracterizate prin gradienti orizontali respectiv verticali de semn contrar; punctul B, aflat la jumătatea segmentului AC, nu este un punct de tranziție; pe linia respectiv coloana decalată cu jumătate din lungimea segmentului AC, există alte trei puncte A', B', C' care întrunesc condițiile enunțate mai sus și decalajul orizontal respectiv vertical între cei doi tripleți de puncte nu depășește grosimea estimată a degetului. Parametrul  $\varphi_B$  al caracteristicii reprezintă unghiul dintre dreapta determinată de punctele BB' și o linie respectiv o coloană de scanare. Extragerea caracteristicilor „fingerlet” în timp real se realizează prin operațiile următoare: se calculează  $l_{max} = l_e (1+\varepsilon)$  și  $l_{min} = l_e (1-\varepsilon)$ , unde  $\varepsilon$  este un număr subunitar, reprezentând toleranța abaterilor de grosime, ce definesc fereastra de detecție de caracteristici „fingerlets”; se scanează regiunea de interes, în ordine lexicografică, până la identificarea primei tranziții fundal-prim plan, ce definește un punct de tip A, de coordonate  $(x_A, y_A)$ ; dacă s-a găsit un punct de tip A, se continuă scanarea liniei curente, până la identificarea unei tranziții prim plan – fundal, la locația C  $(x_C, y_A)$ ; dacă distanța între C și A,  $l_B = |x_C - x_A|$  se încadrează în intervalul  $[l_{min}, l_{max}]$ , se determină poziția punctului de mijloc, B  $(x_B, y_B)$ , cu  $y_B = y_A$ ,  $x_B = (x_A + x_C)/2$  și se salvează, pe o lista asociată liniei  $y_A$ , segmentul candidat corespunzător punctului B, definit prin parametrii  $(x_B, y_B, l_B)$ ; se continuă scanarea regiunii de interes de la ultimul punct de tip C, pentru detectarea de noi segmente candidat; se parcurg în ordinea generată listele conținând parametrii de segmente candidat; pentru fiecare segment candidat,  $(x_B, y_B, l_B)$ , se caută segmente candidat pe lista segmentelor asociate liniei  $y_B + l_B/2$ ; dacă linia respectivă conține un

segment candidat,  $B'$ , de coordonate  $(x_{B'}, y_{B'}, l_{B'})$ , cu decalajul  $|x_{B'} - x_B| < l_B$ , punctului  $B$  i se asociază o caracteristică „fingerlet” validă, cu vectorul de parametri  $(x_B, y_B, l_B, \varphi_B)$ .

Mijloc de urmărire a mâinii pentru mijloc de comunicare om-calculator bazat pe gesturi, caracterizat prin aceea că extrage caracteristici „fingerlet” și generează pentru fiecare deget extins traiectorii în spațiul 4D definit de  $(x, y, l, \varphi)$ .

DESENE EXPLICATIVE

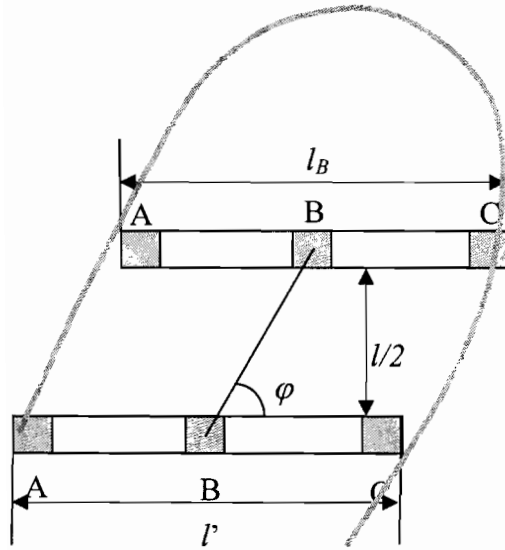


Fig. 1

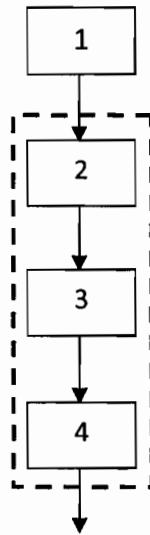
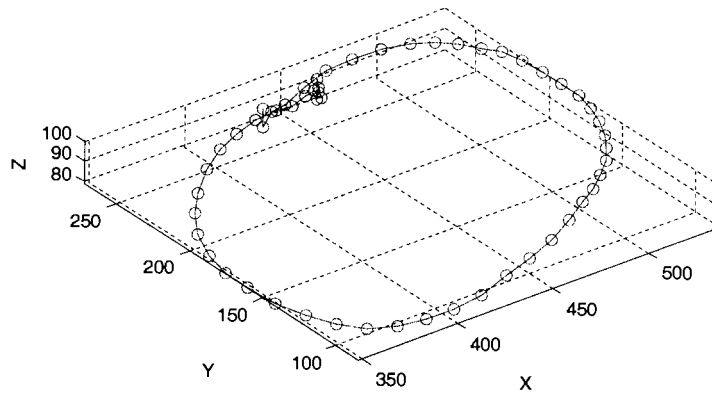
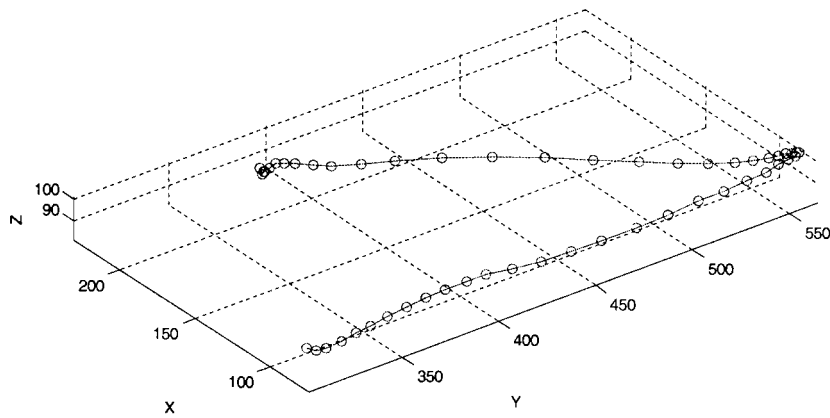


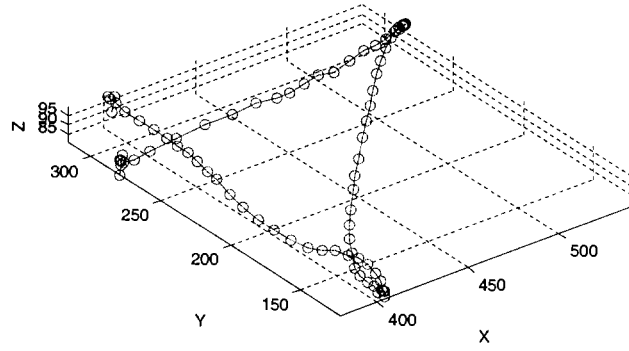
Fig. 2



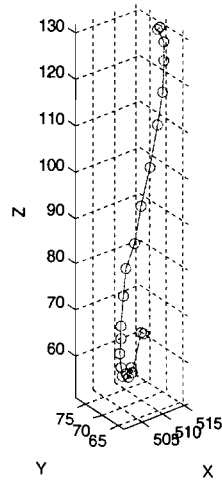
a)



b)



c)



d)

Fig. 3