

Contingut

1 OBJECTIU I ESPECIFICACIONS DEL PROJECTE	1
1.1 Objectiu	1
1.2 Definició del problema. Especificacions	1
2 MOTIVACIÓ DEL PROJECTE	1
3 ABAST DEL PROJECTE	2
4 L'ALGORISME DE RECONeixEMENT	2
4.1 Per què una tècnica activa?	2
4.2 L'algorisme de Grimson i Lozano-Pérez	4
4.3 L'arbre d'interpretacions	5
4.4 Utilització de restriccions geomètriques	6
4.5 Un exemple il·lustratiu	7
5 CARACTERÍSTIQUES DEL SOFTWARE DESENVOLUPAT	7
6 CONCLUSIONS	9
REFERÈNCIES	10

PERCEPCIÓ ACTIVA PER A LA IDENTIFICACIÓ I LOCALITZACIÓ DE PATRONS 2D

1 OBJECTIU I ESPECIFICACIONS DEL PROJECTE

1.1 Objectiu

L'objectiu d'aquest projecte és la realització d'un sistema per a aplicacions robòtiques d'identificació i localització d'objectes de geometria bàsicament bidimensional, en base a informació sensorial dispersa adquirida de forma activa.

1.2 Definició del problema. Especificacions.

Es vol identificar i localitzar de manera automàtica un objecte present en una determinada zona limitada de treball. L'objecte ha de pertànyer obligadament a un conjunt de formes prèviament apreses pel sistema. El sistema desenvolupat ha de tenir, per tant, la capacitat d'aprendre nous objectes.

Els objectes tenen geometria bidimensional amb 3 graus de llibertat (dos de translació, més un de rotacional) en el pla.

El procediment emprat per al reconeixement dels objectes ha de ser un mètode aplicable a diversos sistemes de captació sensorial (ja siguin basats en visió, tacte, proximitat, rastreig per làser, o d'altres). Aquest aspecte dóna flexibilitat al sistema i el fa idoni per a robots que vulguin ser dotats de capacitat multisensorial per interpretar el seu entorn.

La informació sensorial que s'utilitza no és exhaustiva: només es tenen dades de poques, disperses i petites zones de l'objecte a reconèixer. Exigint aquesta especificació pretenem que el sistema pugui funcionar sobre sensors de petita zona activa de lectura (p. ex. els de tacte). Amb aquest tipus de sensors és necessari efectuar diverses lectures sobre l'objecte per conèixer la seva forma.

2 MOTIVACIÓ DEL PROJECTE

La inclusió de sistemes robotitzats en les línies de fabricació i muntatge és actualment un fet cada vegada més habitual. D'aquesta manera, el que es pretén és assolir un augment de la productivitat amb el menor cost unitari possible. És també creixent la necessitat de flexibilitzar al màxim els processos productius i, per tant, es fa desitjable que els robots manipuladors siguin capaços de moure's i actuar dins d'ambients no estructurats prèviament.

En aquest context, el problema principal que sorgeix a la majoria de les aplicacions avançades de robòtica, és la presència d'incertesa en la identitat i posició dels objectes que ha de manipular un robot.

Per tant, es fa necessari que el robot estigui dotat de capacitat per sondejar l'entorn, adquirir-ne dades, interpretar-les i fer-se una idea de la seva estructura per tal d'actuar en conseqüència i sense errors. És precisament dins el marc d'aquest panorama que sorgeix la motivació d'aquest treball.

A més, un dels reptes actuals és poder arribar a equipar els robots amb capacitat multisensorial, és a dir, amb la possibilitat de poder combinar informació procedent de diversos sensors a la vegada per obtenir un model de l'entorn avalat pel major nombre de dades possible. La solució adoptada en aquest projecte és molt satisfactòria en aquest sentit, donat que el mètode proposat per a la identificació i localització de peces resta obert a un gran ventall de sensors i permet aquesta combinació.

3 ABAST DEL PROJECTE

Per tal d'assolir l'objectiu fixat s'han desenvolupat dos mòduls (o paquets) de software:

- el mòdul 3PO d'aprenentatge d'objectes.
- el mòdul R2 per al reconeixement d'objectes.

El mòdul 3PO

3PO efectua l'aprenentatge de nous objectes mitjançant l'obtenció dels models matemàtics d'aquests. 3PO actua a partir d'una imatge binaritzada que conté els contorns de l'objecte a aprendre (fig. 1 dreta). Aquesta imatge, en el nostre cas és facilitada pel sistema de visió de l'Institut de Cibernètica (IC), però en general pot ésser obtinguda per altres tipus de sistemes sensorials.

El mòdul R2

A partir de l'univers d'objectes coneguts que crea 3PO, R2 és capaç de donar la identitat, posició i orientació de l'objecte dipositat a l'espai de treball (fig. 1 esquerra).

R2 treballa a partir d'informació sensorial dispersa de l'objecte a reconèixer. Tanmateix no es restringeix R2 a l'ús d'un sensor en particular, restant en canvi obert a nombrosos sistemes, ja siguin basat en tacte, proximitat, rastreig per làser o visió artificial. Això és possible gràcies a les poques restriccions que fem sobre la forma i tipus de les dades sensorials necessàries, de manera que qualsevol d'aquests sistemes les podrien facilitar.

El sistema de visió

Cal disposar d'algun procediment per comprovar el correcte funcionament de R2 i 3PO. Amb aquesta finalitat, s'ha simulat la captació sensorial en base a imatges gravades pel sistema de visió de l'IC (constituït per una càmera CCD de nivells de gris, un PC 80486 i unes targetes MATROX-IMAGE per al tractament d'imatges). És damunt d'aquestes imatges que es mou un hipotètic sensor que va donant informació de les diferents zones de l'objecte amb les que 'entra en contacte'.

El posicionament d'aquest sensor fictici sobre l'escena de treball és, de moment, aleatori. Tanmateix en el terreny teòric i dins del mateix projecte, s'han estudiat i proposat certes tècniques que permetran efectuar la captació sensorial d'una manera més eficient, cercant trajectòries amb màxim poder discriminatori d'interpretacions falses.

Experimentalment s'ha comprovat el correcte funcionament d'R2 i 3PO aplicant-los a l'aprenentatge i reconeixement de juntes planes de formes diverses: de culata, carburador, tub d'escapament, etc. (fig. 2, (a) i (b)).

4 L'ALGORISME DE RECONeixEMENT

Justifiquem a continuació la necessitat d'emprar una tècnica activa de reconeixement. Així mateix es fa un resum introductori de l'algorisme utilitzat.

4.1 Per què una tècnica activa?

Els mètodes clàssics de reconeixement **passiu** basats en característiques globals dels objectes

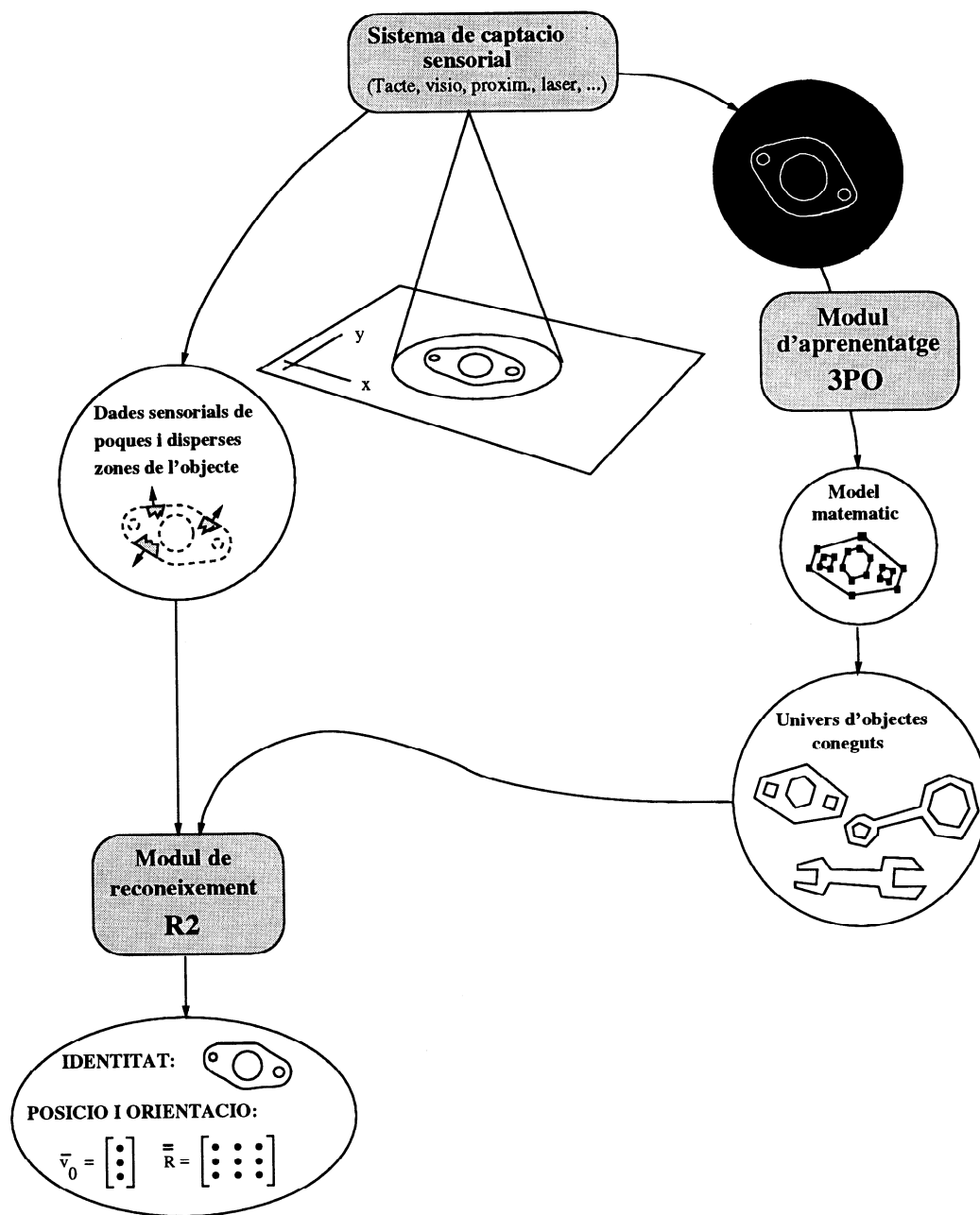


Fig. 1 Diagrama de fluxe de dades entre el sistema sensorial i els dos moduls: 3PO (d'aprenentatge) i R2 (de reconeixement).

(perímetre, àrea, moments d'ordre 2 i superior, etc.) [Hu 62] són realment efectius quan es pot disposar d'informació exhaustiva de l'objecte. Tanmateix hi ha situacions on això és impossible i es disposa únicament de dades locals de certes parts molt concretes de l'objecte. Situacions d'aquesta mena es troben, per exemple, en haver d'identificar objectes massa grans (de manera que no caben en una sola imatge) o, especialment, quan s'utilitzen sensors muntats sobre la pinça d'un robot (com per exemple els de tacte o proximitat) en els quals l'àrea activa és normalment petita si es compara amb

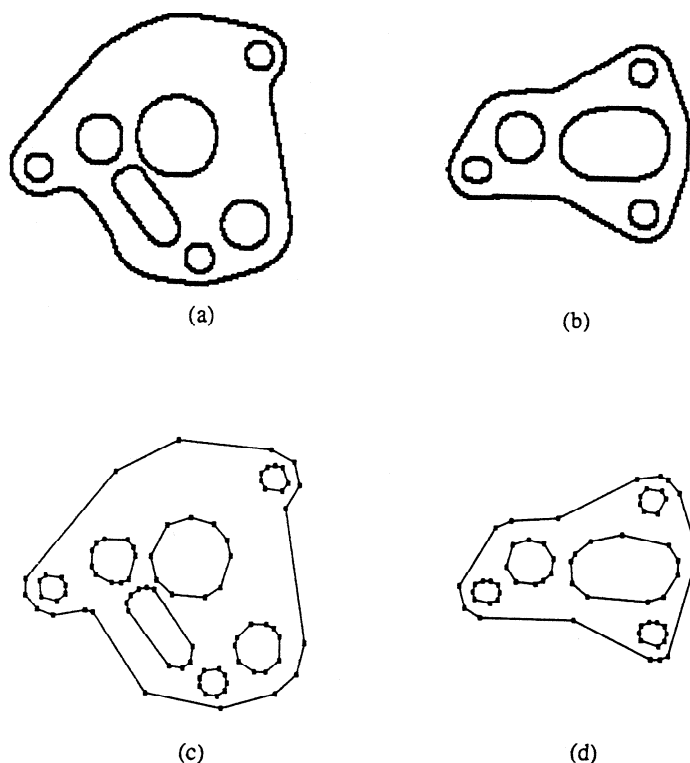


Fig. 2 (a) i (b) Contorns de les juntes 1 i 5 respect. (c) i (d) Aproximacions poligonals corresponents.

la de l'objecte a identificar. Per superar aquestes dificultats s'ha utilitzat una tècnica **activa** de reconeixement. Anomenem 'tècniques de reconeixement actiu' a aquelles que intenten identificar un objecte partint d'informació escassa d'aquest, de manera que només se'n va a buscar més si amb la que es disposa hi ha encara ambigüitat en la identitat i posició de l'objecte en qüestió. Però no només això, en la mesura que sigui possible, la cerca de noves dades ha de ser guiada per la informació prèvia que es disposa de l'objecte, intentant explorar aquelles zones d'aquest que descartin un major nombre d'interpretacions incorrectes.

4.2 L'algorisme de Grimson i Lozano-Pérez

L'algorisme utilitzat [RoT 91] segueix molt de prop l'ideat per E. Grimson i T. Lozano-Pérez [GrL 84]. Per a poder-lo emprar és necessari que, abans que res, els contorns dels objectes suposats coneguts, hagin estat aproximats per polígons de costats rectes (fig. 2 (c) i (d)). A més, caldrà haver calculat certes matrius de restriccions geomètriques deduïdes a partir d'aquests polígons (ja veurem quines). Aquestes dues tasques les realitza el mòdul 3PO.

En resum, el procediment de reconeixement consisteix en els següents passos:

(1) *Lectura inicial de poques dades sensorials.*

Cada dada consisteix en:

- les coordenades d'un punt del contorn de l'objecte.
- el vector normal unitari cap a fora en cada un dels anteriors punts.

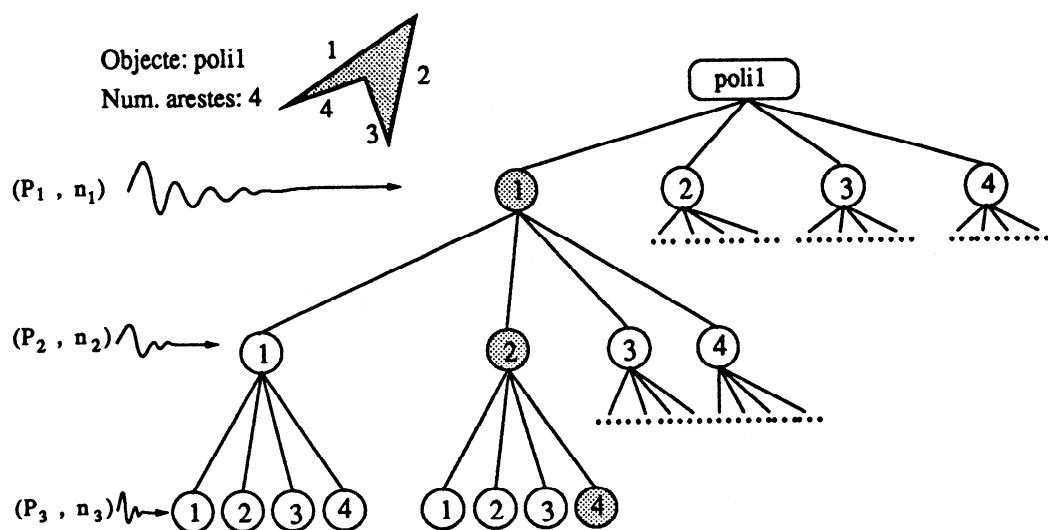


Fig. 3 *Arbre Complet per l'objecte 'poli1' i 3 punts llegits damunt del seu contorn.*

(2) *Generació d'interpretacions consistents.*

Anomenem 'interpretació' a cada conjunt d'emparellaments del tipus

dada sensorial ---> aresta del model suposat

per tots i cadascun dels punts llegits pel sensor. Amb la paraula 'consistent' volem dir 'que satisfaci certes restriccions geomètriques locals derivades dels models' (ja veurem quines).

(3) *Aplicació del 'test global'.* Es verifica la consistència global de les interpretacions generades en el pas anterior mirant si existeix alguna transformació (translació i rotació) del model suposat que faci que aquest encaixi amb les dades sensorials llegides.

(4) *Captació adicional d'informació sensorial* i repetició de (2) i (3) mentre sigui possible més d'una interpretació.

4.3 L'arbre d'interpretacions

La cerca d'interpretacions consistents es porta a terme en base a la generació i exploració d'una estructura de dades anomenada *arbre d'interpretacions*. En efecte, el conjunt de tots els emparellaments entre dades sensorials i arestes del model es pot representar en forma d'arbre (fig. 3). Cada nivell de l'arbre representa l'emparellament d'una dada sensorial amb totes les possibles arestes del model a que pot ésser assignada. Per tant l'arbre té tants nivells com dades sensorials disponibles. Una branca des de l'arrel de l'arbre fins al nivell inferior constitueix una possible interpretació de les dades. A més, hi ha un arbre d'aquest tipus per cada un dels models coneguts.

Un arbre d'interpretacions complet seria aquell en el que cada dada pot ésser emparellada amb totes les arestes del model suposat (fig. 3). Ja es veu que una possible solució del problema fóra explorar

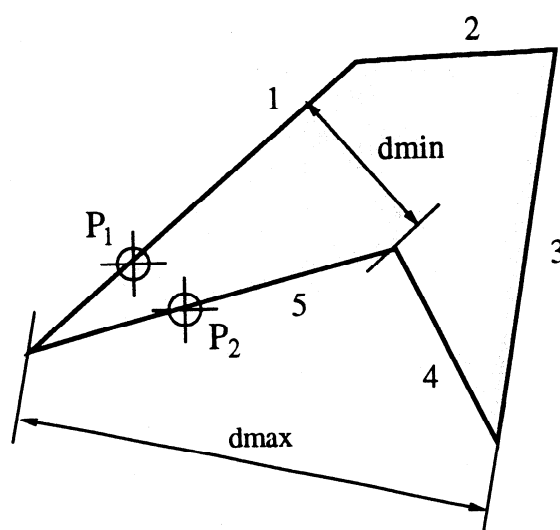


Fig. 4 *Restricció geomètrica de distància.*

totes les branques d'aquest arbre, aplicant-les-hi el pas (3) anterior. Computacionalment el cost d'això és exagerat. Es fa necessari algun mètode que permeti eliminar interpretacions falses de manera ràpida, per tal que al final només calgui aplicar el *test global* a poques d'elles.

4.4 Utilització de restriccions geomètriques

Per a això emprarem restriccions geomètriques deduïdes a partir dels models. S'usen tres restriccions de demostrada efectivitat [GrL 84] i [Gri 86a]:

- * restricció de distància
- * restricció d'angle
- * restricció de direcció

A tall d'exemple explicarem la primera. La distància entre cada parella de punts llegits pel sensor ha d'estar compresa dins l'interval de distàncies possibles entre parelles de punts de les arestes a les que han estat assignats aquests punts. Així, per exemple, per un objecte com el de la fig. 4, si s'han contactat P_1 i P_2 , i P_1 s'ha assignat a l'aresta 1, aleshores P_2 no es pot emparellar amb la 4 donat que:

$$dist(P_1, P_2) \notin [d_{\min}(1,4), d_{\max}(1,4)]$$

Els valors d_{\min} i d_{\max} poden calcular-se prèviament i ser emmagatzemats en una taula o matriu de restriccions de distància. La resta de les restriccions són semblants. Per cada objecte conegut 3PO calcularà doncs tres taules: la de restriccions de distància, la d'angles i la de direccions. Aquestes taules seran emprades durant el pas (2) de l'algorisme per "podar" aquelles branques de l'arbre que constitueixin interpretacions geomètricament inconsistentes.



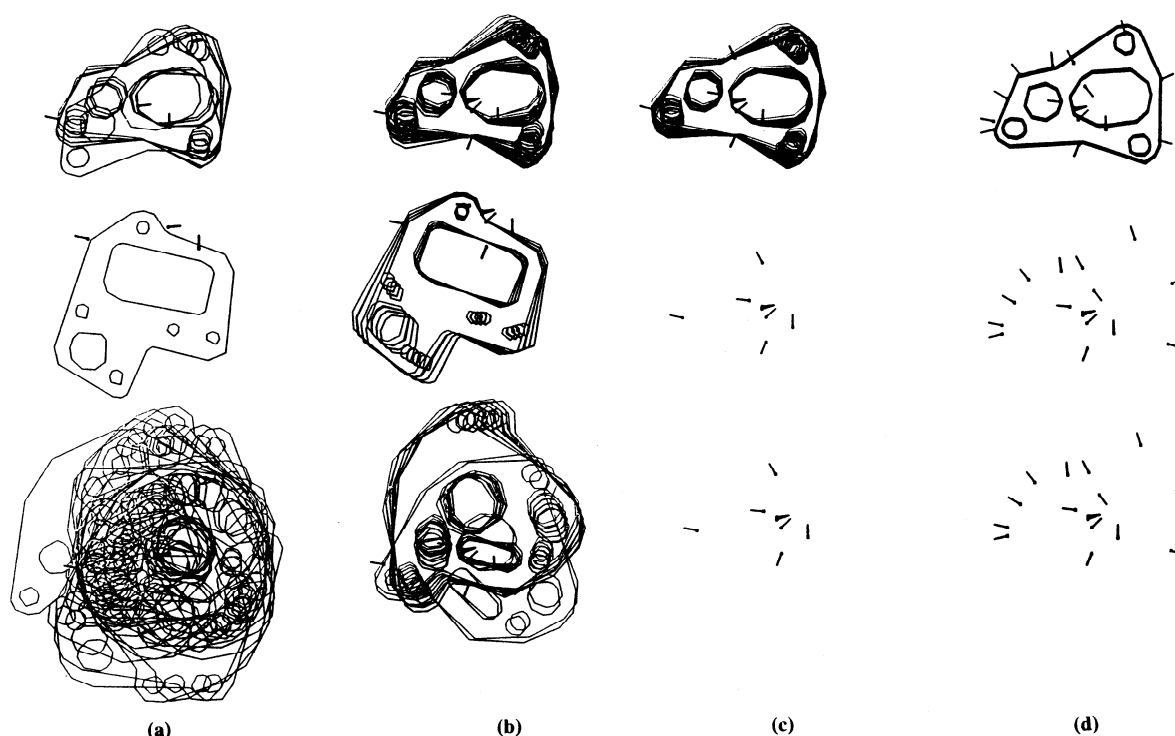


Fig. 5 Posicions dels models que encaixen amb 3, 7, 8 i 16 punts del contorn llegits pel sensor.

Després d'aplicar aquestes restriccions, el nombre de branques a explorar durant l'aplicació del pas (3) es redueix dràsticament (tal com es demostra a [Gri 86a] i [Gri 86b]) fent que l'algorisme arribi a la solució en pocs segons.

4.5 Un exemple il·lustratiu

Per il·lustrar gràficament l'aplicació d'aquest algorisme observem la fig. 5. En aquest exemple d'execució s'han suposat coneguts 3 models: junta1, junta4 i junta5. Se n'ha escollit un, el junta5, s'ha dipositat a la plataforma de treball i s'ha començat a explorar amb el sensor. Contactant l'objecte en tres punts, veiem a la fig. 5 (a) que hi ha nombroses posicions dels tres models que encaixen amb aquests. Observem que en anar captant més informació sensorial, es va tendint a una única interpretació de les dades (fig. 5 (d)), l'única que és vàlida. Amb 8 punts del contorn ja només és un el model que hi encaixa i amb 16 arribem a l'única posició vàlida d'aquest.

5 CARACTERISTIQUES DEL SOFTWARE DESENVOLUPAT.

R2 i 3PO han estat implementats en llenguatge C sobre una estació de treball SUN. Gràcies a això són programes fàcilment portables a altres tipus de màquines (p. ex. els PC's). L'aspecte extern ha estat especialment cuidat per a facilitar al màxim la seva utilització. En aquest sentit tant R2 com 3PO incorporen una agradable interfície d'usuari (figures 6 i 7) basada en l'ús del ratolí i amb finestres

superposables, construïda a partir de les llibreries SunView [Sun 90], però dissenyada de manera que sigui fàcil la seva traducció a altres entorns (OPENWINDOWS, MOTIF, WINDOWS per PC, etc.).

En el cas del mòdul R2, l'usuari visualitza fàcilment els resultats del programa en dues finestres gràfiques. A la finestra reticulada de la dreta (fig. 6) es mostra l'objecte dipositat a la taula de treball, així com les zones on el sensor l'ha contactat i les dades que n'ha extret. A la finestra fosca de l'esquerra el programa va dibuixant totes aquelles posicions dels models que encaixen amb les dades sensorials llegides. Automàticament el sensor és mogut damunt la simulada plataforma de treball per captar més informació, fins que no hi ha ambigüïtat en la solució i tant sols és un el model que encaixa en una sola posició sobre les dades preses (fig. 5 (d)). En general el temps d'execució depèn del nombre de models presents a l'univers d'objectes coneguts i de la seva complexitat. L'execució de la figura 5 ha tardat 14 segons en arribar a obtenir la solució (tolerant un error acceptable).

La pantalla d'usuari de 3PO es mostra a la fig. 7. En el seu pannel superior esquerre s'hi mostren disponibles en forma de botons 'clicables' totes les funcions d'aquest mòdul. A la pantalla gràfica inferior esquerra es mostra la plataforma de treball, amb l'objecte a aprendre dipositat damunt d'ella. A la pantalla fosca de la dreta s'hi mostren els resultats d'execució del programa: la figura que apareix a la part inferior és el model poligonal obtingut. A la part superior es sobreposa el model sobre el contorn original de l'objecte per tal d'apreciar-ne la qualitat de l'ajust (es poden obtenir models més o menys precisos variant els paràmetres SALT i Uc). Les matrius de restriccions s'obtenen automàticament i es guarden en fitxers. 3PO permet també la lectura/escriptura d'imatges en diversos formats.

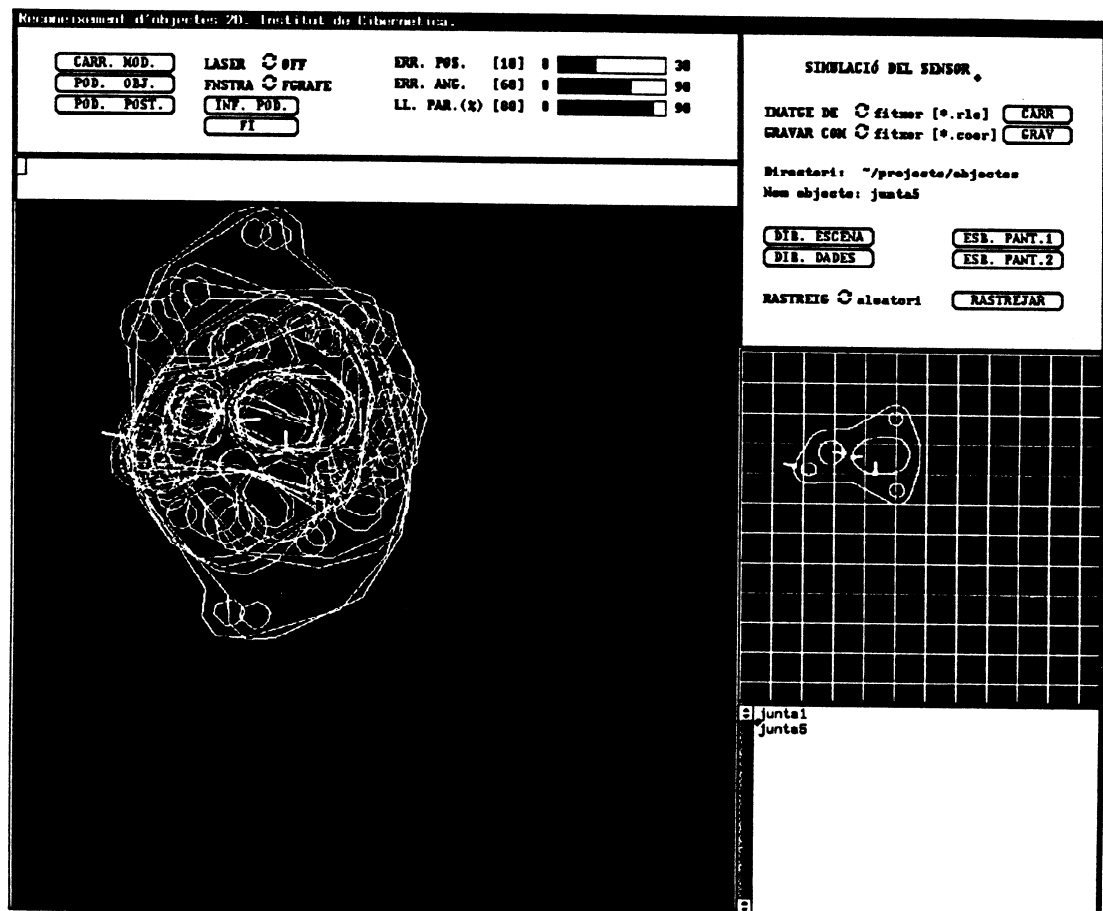


Fig. 6 Interfície d'usuari del mòdul de reconeixement R2.

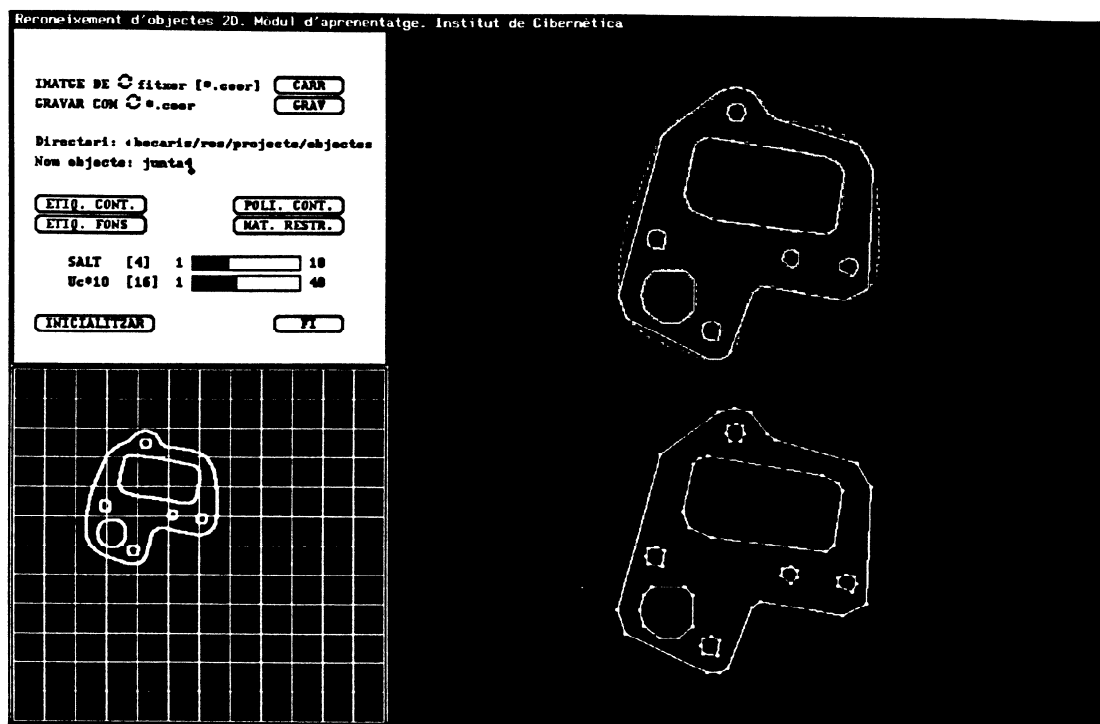


Fig. 7 Interfície d'usuari del mòdul d'aprenentatge 3PO.

6 CONCLUSIONS

La metodologia presentada i analitzada en aquest projecte ha permès resoldre satisfactoriament el problema de la identificació i localització d'un objecte a l'espai de treball, a partir d'informació dispersa.

Pel que fa referència als objectius fixats al començament, podem afirmar que s'han acomplert plenament. En efecte, tal i com es posa de manifest a la memòria, és possible identificar i localitzar un objecte dins l'espai de treball, repetint el procés cíclic de *generació d'hipòtesis-verificació* les vegades que convingui. Així mateix, mitjançant l'ús de restriccions geomètriques s'escurça dràsticament el temps necessari per aconseguir aquest reconeixement.

El sistema ha estat implementat en un prototipus software i validat experimentalment amb èxit, emprant objectes reals (concretament juntes planes de formes i tipus diversos: de culata, de carburador, de tub d'escapament, etc.), reconeguts a partir d'informació sensorial procedent del sistema de visió de l'IC.

És important destacar que per a l'aplicació de l'algorisme n'hi ha prou amb disposar d'informació dispersa de la peça a reconèixer i que, per tant, el 'software' dissenyat és especialment indicat per a sensors de tipus tacte, proximitat, sistemes combinats làser-càmera, etc., sensors (tots ells) en els que la zona activa de lectura és comparativament petita enfront de l'àrea total de l'objecte a identificar.

Convé dir també que tot i que el sistema dissenyat permet reconèixer objectes de geometria bidimensional (no necessàriament plans) pot també ésser emprat per a peces tridimensionals. En aquest cas el sensor hauria de prendre les dades sobre els contorns dels objectes situats a una certa alçada constant damunt del pla de treball.

El present projecte no ve a resoldre una necessitat industrial concreta i particular, ans al contrari, és un sistema obert a múltiples aplicacions on es requereixin robots dotats de capacitat sensorial sobre l'entorn. Tot i no ser un producte directament preparat per la seva comercialització, l'autor es mostra molt satisfet del seu grau d'acabat.

REFERÈNCIES:

- [Gri 86a] "The Combinatorics of Local Constraints in Model-Based Recognition and Localization from Sparse Data", J. ACM, vol. 33, no. 4, pp. 658-686, 1986.
- [Gri 86b] W.E.L. Grimson, "Disambiguating Sensory Interpretations Using Minimal Sets of Sensory Data", 1986 IEEE Conf. on Rob. and Autom., pp. 286-292, San Francisco, California, 1986.
- [GrL 84] W.E.L. Grimson i T. Lozano-Pérez, "Model-based and Localization from Sparse Range or Tactile Data", Inter. Journal of Robotics Research, vol. 3, no. 3, pp. 3-35, 1984.
- [Hu 62] M. Hu, "Visual Pattern Recognition by Moment Invariants". IRE Trans. Information Theory, pp. 179-187, febrer 1962.
- [RoT 91] L. Ros i F. Thomas. "Percepción activa para la Identificación y Localización de Juntas Parcialmente Solapadas". Actas del Segundo Congreso de la AER, pp. 249-255, noviembre 1991.
- [Sun 90] "SunView Programmer's Guide". Sun Microsystems, Incorporated. 1990.