

COMPUTADORS I PERCEPCIÓ

per

ALBERT SANFELIU

Institut de Cibernètica

Amb l'aparició dels computadors en el decenni dels anys seixanta es produeix un canvi profund en el disseny de la gestió i automatització dels processos industrials. De les diverses àrees que foren tractades mitjançant el computador, els sistemes de percepció induïren a un intens estudi pel potencial d'aplicació que tenen en camps tan diversos com és ara la robòtica, la medicina o l'astronomia. La raó era evident, pel fet que en la majoria de les àrees científiques i industrials foren utilitzats els sistemes de percepció per a obtenir informació de l'entorn i poder interaccionar-s'hi. Dins aquest ampli camp, la visió per computador és la que ha adquirit un major impuls gràcies a l'ampli espectre d'aplicacions que té. A la següent explicació ens centrarem en la investigació que hom duu a terme a l'Institut de Cibernètica en el camp de la visió per computador.

L'Institut de Cibernètica, des de fa uns quants anys, s'ha dedicat a la investigació de la percepció visual per a aplicacions en la robòtica (Basañez, L., i altres, 1981). El seu principal interès se centra a reconèixer objectes en 2 i 3 dimensions (objectes reals) tant aïllats com parcialment ocults o cavalcats, amb l'objecte de situar-los en l'àrea de treball, d'obtenir els punts d'agafada requerits a les tasques d'acoblament, de determinar certes característiques necessàries de les tasques de supervisió i d'adquirir la informació precisa per a incorporar nous objectes a la base de coneixement del sistema de percepció.

Els últims treballs desenvolupats són orientats cap a l'aprenentatge i el reconeixement d'objectes en 2 dimensions (Sanfeliu, 1984) (a partir d'una sola projecció plana) i 3 dimensions (Sanfeliu, Llorens, 1985; Sanfeliu, 1984), per tal de poder tractar amb les peces habituals en les tasques d'acoblament. Alhora hom estudia noves tècniques de cooperació entre els siste-

mes de percepció, principalment el de visió i un braç robot (Sanfeliu i Torras, 1984; Sanfeliu, i altres, 1985) de manera que puguin ésser resoltes les incerteses ocasionades per un entorn ambigu i el temps enorme de càlcul de computador necessari.

A continuació resumirem molt breument el plantejament general de la visió-planificació per a objectes 3D i els punts més interessants dels treballs mencionats:

PLANTEIG GENERAL DE LA VISIÓ-PLANIFICACIÓ PER A LA INTERPRETACIÓ D'ESCENES

L'enfocament general considerat en els treballs d'investigació és representat a la Fig. 1. Hom hi pot veure dues vies que depenen de la informació disponible en el moment de fer la interpretació.

Quan no són coneguts els objectes que poden presentar-se en una escena o bé hom vol fer una anàlisi completa d'una escena, la via que seguim és la marcada en línia contínua. Els mòduls s'interrelacionen d'una forma global de manera que qualsevol d'ells pot demanar més informació a la resta.

En el cas contrari, és a dir, quan hom disposa d'informació sobre l'escena o interessa d'identificar un objecte específic, llavors es requereix la planificació global i el nivell de cada un dels mòduls. Això permet de reduir el temps de càlcul i preveure els possibles errors que poden sorgir-hi. La planificació pot ésser duta a terme a diferents nivells i fins i tot pel que fa a la base de coneixements orientant-la de forma que hom pugui obtenir fàcilment les dades i les estratègies apropiades per a cada domini de l'escena.

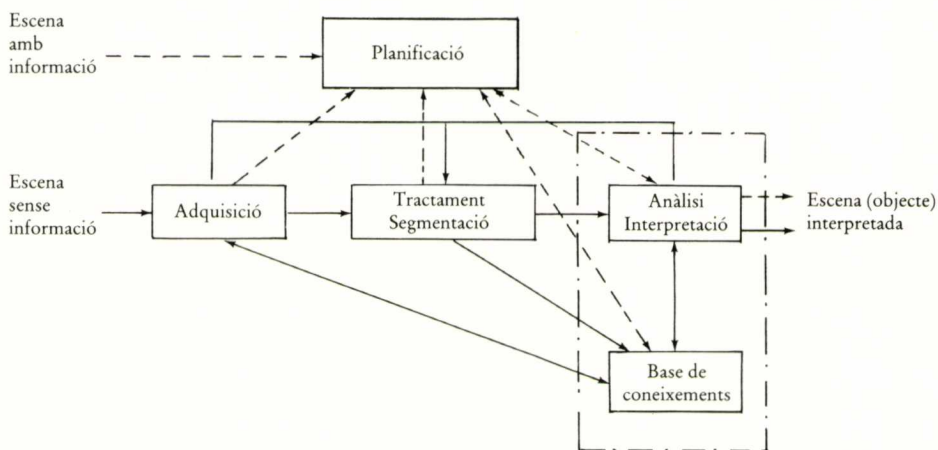


Fig. 1. Enfocament general de la visió-planificació per a la interpretació d'imatges.

A ambdues vies hom pot utilitzar els quatre tipus de control d'interpretació d'escenes: l'ascendent, o dirigit per la informació de la imatge; el descendent, o dirigit per la meta a assolir; l'hetàrquic, que combina aspectes dels dos anteriors, i el model d'agenda, on alguns especialistes es comuniquen entre ells mitjançant una agenda.

RECONeixEMENT D'OBJECTES 2D AMB OCULTACIONS I CAVALCaments

En aquest treball ha estat estudiat i implementat un mètode per a reconèixer peces semiocultes i cavalcades en 2 dimensions, és a dir a partir d'una projecció plana de l'escena (Sanfeliu, 1984; Sanfeliu, Torras, Ruiz, 1985). El sistema consta de tres parts: adquisició, segmentació i anàlisi, i es basa en les tècniques de reconeixement de formes sintàctiques (Sanfeliu, Fu, 1983). El procés és el següent:

L'escena és digitalitzada en una matriu de 512 x 512 bytes amb 256 tonalitats de gris. La segmentació parteix de la binarització de la imatge mitjançant un "llindar" i continua amb l'extracció del contorn, el seu aïllament i l'obtenció de les característiques i primitives. A partir d'ací comença l'anàlisi, que consta de dues etapes. La primera es basa en la comparació de la cadena de primitives del contorn de la imatge amb cada un dels models de referència. Tots els objectes de l'escena que no han estat reconeguts passen a una segona etapa. En aquesta hom cerca l'orientació i la posició de translació de l'objecte quan comparem l'obtingut de la imatge amb cada model per mitjà de la transformada de Hough. Més tard és analitzat l'objecte i hom aplica una mesura de similitud entre els objectes de la imatge i cada un dels models de referència, i el model amb la mínima diferència és el reconegut.

Aquest mètode ha estat aplicat a diferents objectes en 2 dimensions i hom hi ha obtingut molt bons resultats.

VISIÓ-PLANIFICACIÓ PER A LA CAPTACIÓ D'OBJECTES CAVALCATS I PARCIALMENT OCULTS

La necessitat de captar objectes d'una àrea de treball on puguin estar cavalcats i parcialment ocults és una tasca bàsica d'acoblament. Hom s'ha plantejat la necessitat de dur a terme diferents estudis per a solucionar el problema mitjançant l'augment d'informació sensorial, refinar els procediments de reconeixement, captar els objectes directament a partir de l'obtenció dels punts d'agafada o bé identificar els objectes i ensorrar els cúmuls cercant les peces no reconegudes.

La nostra investigació s'ha basat en aquest darrer mètode amb la condició que partim d'un sistema sensorial molt escàs (Sanfeliu i Torras, 1984;

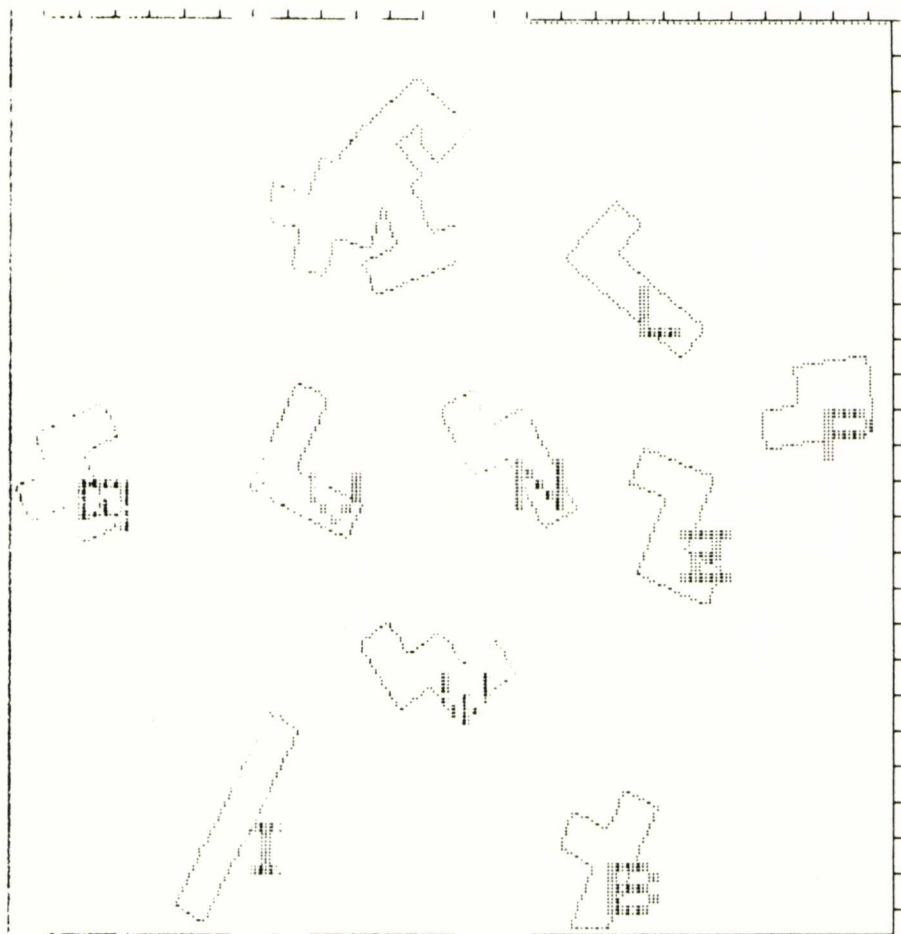


Fig. 2. Reconeixement de peces soltes i cúmul.

Sanfeliu i d'altres, 1985). La idea ha estat de desenvolupar un sistema basat en especialistes (operadors) que estudien la situació, i basant-se en llur punt de vista suggereixen a un interpretador la manera com cada un d'ells resol-dria el problema (Erman *et al.*, 1980). L'interpretador utilitza aquesta infor-mació i, a partir de l'estratègia imposada, determina què ha d'aplicar a con-tinuació l'operador. Els operadors considerats fins ara són set: anàlisi d'es-cenes, planificació en l'ensorrament de cúmul, planificació del desplaça-ment d'objectes reconeguts, planificació de la manera d'aconseguir punts d'agafada inaccessibles, planificació de l'escombratge de cúmul de peces reconegudes, desplaçament de peces i trasllat de peces.

L'operador d'anàlisi d'escena és una aplicació de l'operador esmentat a la secció anterior, i inclou una etapa de més que serveix per a reconèixer les peces dels cúmul (Fig. 2). En aquesta etapa hom identifica les peces seguint

el procediment esmentat però incloent-hi una sèrie de modificacions. La segmentació de les superfícies és duta a terme mitjançant un operador de gradient, i a continuació hom cerca el millor ajust de cada model amb el de la imatge, tant en la seva orientació com en la posició de translació. Un cop trobat, hom aplica una mesura de similitud i estudia les possibles configuracions que es poden donar resolent les possibles ambigüitats.

ANÀLISI D'OBJECTES COMPLEXOS EN 3D

La visió per computador aplicada a la robòtica, pretén, entre altres coses, de reconèixer els objectes del món real per tal de poder situar-los a l'espai, i adquirir certes mesures i inspeccionar determinades zones. Fins al pre-

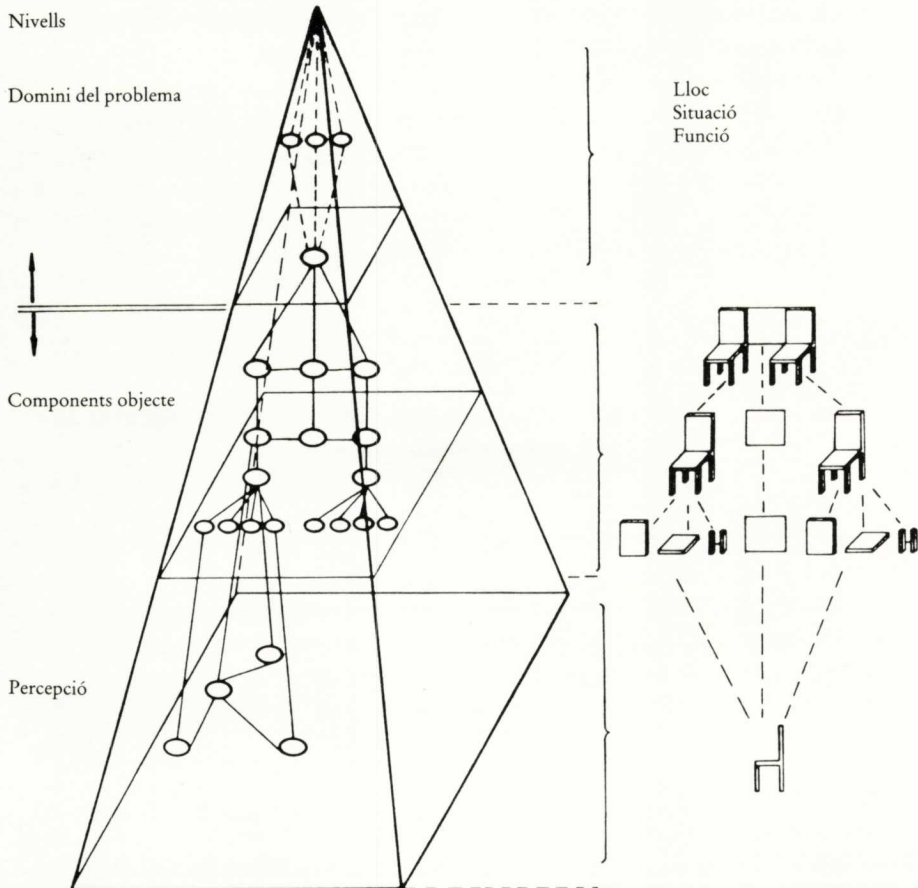


Fig. 3. Representació del model global 3D.

sent, els sistemes comercialitzats de visió es basen en escenes 2D, tenint en compte la complexitat dels de 3D i que en molts casos industrials una única projecció és suficient. La tercera dimensió en visió (Basañez i altres, 1981) és essencial quan el robot ha de poder resoldre els seus propis errors o superar l'imprevist que pot esdevenir, i sobretot quan ha de moure's en un entorn quasi-industrial o no industrial.

La investigació que hom fa a l'Institut de Cibernètica ha seguit dos camins, un de teòric i un altre de pràctic. En el primer (Sanfeliu, 1984) ha estat presentat un model de representació d'objectes 3D (Fig. 3) que permet d'interpretar una escena sota dos punts de vista: el de l'objecte físic, i el de domini del problema. Sota el prisma de l'objecte físic i dels sistemes de percepció presents, hom transforma el model comparant-lo amb l'objecte de la imatge que cal reconèixer. Per a la seva anàlisi hom utilitza una mesura de similitud jeràrquica entre grafs, la diferència dels quals és utilitzada per a classificar l'objecte de la imatge entre els models de referència. Un cop classificat, quan el comparem amb els models dels objectes físics, hom estudia l'objecte segons on resideix el domini del problema.

En el camí pràctic (Sanfeliu i Llorens, 1985) han estat desenvolupats una sèrie d'algoritmes per a obtenir una segmentació idònia de la imatge de manera que poden separar-se fàcilment les diferents superfícies que componen un objecte. A partir d'ací han estat implementats algoritmes d'anàlisi basats en el còmput de transformacions error entre grafs.

CONCLUSIONS

Les investigacions que hom duu a terme a l'Institut de Cibernètica en l'àrea de la percepció visual se centren principalment en la seva aplicació a la robòtica. L'interès està a reconèixer objectes a partir d'una única projecció (2D) o bé amb informació 3D, per a situar-los en l'espai, adquirir llurs punts d'agafada i inspeccionar-los.

Han estat duts a terme diferents treballs en aquesta àrea sota dues orientacions, l'una teòrica i l'altra pràctica. En la segona han estat desenvolupats algorismes eficaços per a analitzar escenes.

BIBLIOGRAFIA

- AMAT, J., BASAÑEZ, L., TORRAS, C. (1983): "Visión tridimensional en los robots industriales". *Mundo Electrónico* 130.
- BASAÑEZ, L., HUBER, R., JUAN, J., TORRAS, C., (1981): "Percepción Visual en Robótica: Técnica y Algoritmos". *Revista Regulación y Mando Automático*, 144, 61-72. Des. 1981.

- SANFELIU, A., (1984): "A Distance Measure Based on Tree-Graph Grammars: A Way of Recognizin Hidden and Deformed 3D Complex Objects" 7th Int. Conf. on Pattern Recognition. Montreal (Canadà).
- SANFELIU, A., i FU, K. S. (1983): "A Distance Measure between Atributed Relational Graphs for Pattern Recognition" IEEE Trans. on System, Man and Cybernetics.
- SANFELIU, A., LLORENS, A. (1985): "Reconocimiento de objetos 3D para la visión en robótica", VI Congreso de Informática y Automática AEIA, Madrid, Oct. 1985.
- SANFELIU, A. TORRAS, C. (1984): "Cooperación visión-robot en tareas de ensamblaje" II Simposium Nacional sobre Automática en la Industria, Saragossa, 14-16 Nov. 1984.
- SANFELIU, A., TORRAS, C., FONT, J., RUIZ, J., (1985): "Active-Recognition system for the Acquisition of Overlapping and partially Hidden Workpieces", SYROCO (Symposium on Robot Control), Barcelona 6-8 Nov. 1985.
- TORRAS, C., i BASAÑEZ, L. (1982): "The Sweep Mapping: A way to perform Digital Image Processing for Robot Vision" 12 th Int. Symposium on Industrial Robots, Juny, París.