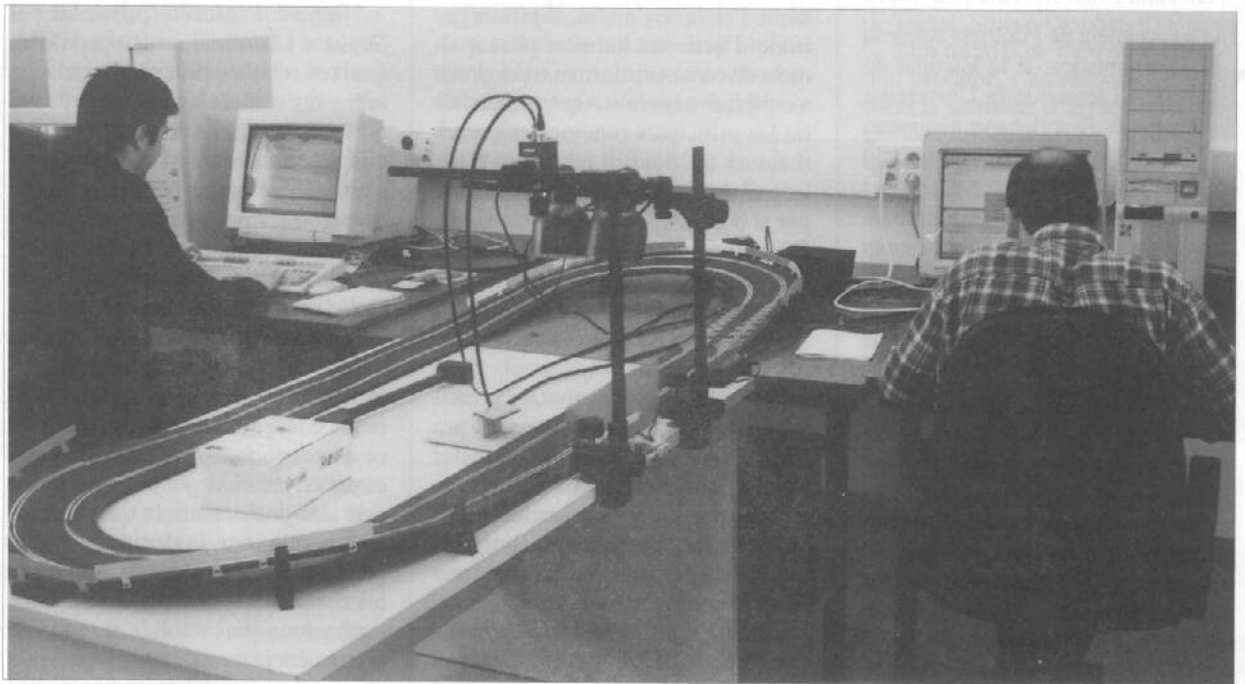


Models computacionals per a la visió artificial

La visió és el sentit més elaborat i com a tal ha estat objecte d'estudi durant segles per explicar el seu funcionament. L'arribada de la tecnologia digital ha permès el desenvolupament de teories computacionals que han intentat explicar com un agent dotat del sentit de la visió pot interactuar amb el món a partir de la percepció basada en imatges, proposant quina informació es pot extreure, com es processa i com s'utilitza. Fruit d'això ha estat l'aparició d'un nou objectiu científic: la síntesi —en contraposició a l'estudi precomputacional, de característiques fonamentalment analítiques— de sistemes de visió artificial i la seva aplicació a problemes del nostre entorn.

L'equip del professor Jordi Vitrià, del Departament d'Informàtica i del Centre de Visió per Computador de la UAB, està treballant en un projecte CICYT sobre els aspectes bàsics de la visió: la percepció de les distàncies entre l'observador i els objectes del seu entorn. Aquesta informació es pot obtenir per diferents mitjans (fusió d'imatges binoculars, anàlisi del moviment, etc.), però l'estudi s'ha centrat en un aspecte relativament poc estudiat, com són els gradients d'enfocament que es donen quan un sistema òptic amb una profunditat de camp limitada canvia la seva distància focal. Hi ha evidències que aquest sistema és present en alguns sistemes visuals naturals. El cas més paradigmàtic és el del camaleó, que basa la seva supervivència en la capacitat de realitzar mesures monoculars precises de la distància a què es troben les seves preses. La necessitat d'un sistema d'aquest tipus no es basa únicament en la possibilitat de percepció de la tercera dimensió del món representat a les imatges, sinó també en el mateix procés d'adquisició de la imatge, per assegurar que el sistema pot obtenir la millor imatge possible —en el sentit de ser tan informativa com es pugui—, sigui quina sigui la distància a què es troba l'objecte visualitzat. Això significa que l'observador ha d'estar dotat d'un criteri d'enfocament ràpid i precís, que permeti una adaptació contínua als canvis del món.



Una de les sales on es fan les proves de models computacionals per a la visió artificial.

Aplicacions

A banda de l'interès científic de l'estudi, les possibilitats d'aplicació directa d'aquestes tècniques són moltes i variades. El projecte n'ha abordat dues: les aplicacions en l'àmbit de la microscòpia i la utilització com a part del sistema perceptiu d'un robot. En les dues aplicacions s'ha estudiat el procés d'autoenfocament i la possibilitat d'extreure informació mètrica de l'escena visualitzada.

Un dels aspectes importants del projecte ha estat el treball dins d'un nou paradigma científic conegut per «visió activa». Aquest model computacional va néixer fa pocs anys com a resposta a un cert esgotament del paradigma «reconstruccionista», desenvolupat per David Marr a l'Artificial Intelligence Laboratory del MIT durant la dècada dels setanta, i que fins ara ha estat pràcticament únic dins del camp de la visió artificial. Mentre que el reconstruccionisme basa la seva teoria en un procés passiu dirigit a recuperar les característiques físiques del món (semblant al que podria ser una teoria de l'«òptica inversa»), la visió activa considera l'observador

com un element crític del procés perceptiu i l'introdueix com una variable dinàmica més. Aquest nou paradigma ha fet desplaçar l'interès cap a l'estudi de les necessitats perceptives de l'agent en funció de la tasca que realitza, i limitar la «reconstrucció» del món als aspectes estrictament necessaris.

Dins del paradigma reconstruccionista, el principal resultat del projecte és un mètode per recuperar simultàniament informació sobre la forma i la reflectància dels objectes presents en una seqüència d'imatges preses amb distàncies focals diferents. Aquest mètode, desenvolupat en col·laboració amb el professor J. Llacer, de la Universitat de Califòrnia, a Berkeley, s'ha basat en un model bayesià per a la resolució de problemes inversos. Aquest resultat té una aplicació directa en l'àmbit de la microscòpia, ja que permet la visualització d'imatges totalment enfocades més enllà de les possibilitats òptiques del microscopi, i al mateix temps proporciona mapes de profunditat de la mostra.

Visió activa

Ja dins del camp de la visió activa, els objectius s'han dirigit a determinar l'estímul visual més adequat per a un procés d'autoenfocament i un sistema òptic de control per manipular el procés d'adquisició de la imatge. El resultat ha estat un mètode basat en les variacions del contrast local de les imatges, que, comparat amb les possibilitats dels estímuls considerats fins ara en l'àmbit de la visió artificial, ha resultat ser més precís, més robust i amb una complexitat computacional més petita. Això ha permès de dissenyar un sistema d'autoenfocament en temps real per a microscòpia i un sistema d'acomodació òptica per a un sistema de visió antropomòrfic

dotat de lents motoritzades. El mètode, a més de ser capaç d'enfocar automàticament un sistema òptic en un rang de condicions molt ampli, també pot proporcionar informació sobre la distància a què es troben els objectes enfocats, fet d'importància vital en l'àmbit de la robòtica.

Actualment s'està treballant en dues vies de continuació. D'una banda, l'estudi de nous estímuls visuals útils per a l'enfocament i relacionats amb el color. Aquest estímul, molt poc estudiat fins ara, podria ser crític en circumstàncies on d'altres són inoperants, com per exemple objectes que presentin poca textura visual. De l'altra, s'està estudiant la integració dels processos oculars bàsics en la plataforma de visió antropomòrfica. Aquesta plataforma està formada per un sistema robotitzat amb cinc graus de llibertat que imiten els moviments d'un cap i per dues càmeres de CCD amb òptiques motoritzades que permeten el control de l'enfocament, l'obertura i el zoom. Les possibilitats mecàniques d'aquesta plataforma, que s'acosten molt a un model antropomòrfic, són una bona base per al desenvolupament de sistemes de visió actius, útils en problemes com la conducció automàtica de vehicles, el control del trànsit o la vigilància. Però, per ser operatius, cal tenir-ne resolt el control a nivell ocular de manera independent de l'aplicació. És a dir, cal que el sistema sigui capaç d'obtenir automàticament imatges enfocades, de mantenir la vergència de les dues càmeres, d'ajustar el seu iris a les condicions d'il·luminació i d'estabilitzar la seva mirada en els objectes. La integració de tots aquests processos en un únic esquema de control, guiat pels estímuls visuals, continua sent un tema apassionant de recerca, necessari per construir sistemes visuals artificials.



Instal·lacions del Centre de Visió per Computador.