

Capítol 1

Introducció

Des de temps remots l'home ha tingut el somni de dotar les màquines de la capacitat de presa de decisions intel·ligents. Com és sabut la ciència està lluny d'assolir aquest objectiu tot i l'aparició de noves tecnologies sobretot, en el camp dels ordinadors a la dècada dels 60 i la notable evolució que han tingut les disciplines implicades com ara: el control, la intel·ligència artificial, la visió per ordinador, el processament d'imatges, la psicologia...

Una tendència actual dels científics molt important és la de desenvolupar sistemes d'assistència complexos, per tasques específiques, dotats d'interfícies persona-màquina molt sofisticades i ergonòmiques, intentant guanyar un grau d'autonomia cada cop més gran en funció de les capacitats tecnològiques del moment. Aquest plantejament permet alliberar les persones de tasques d'alta precisió o d'alt risc fent possible augmentar capacitats com precisió, repetitivitat, força, o accessibilitat, reservant les decisions més intel·ligents a les persones i evitant així les enormes dificultats que implica emular la intel·ligència humana. Exemples on aquesta filosofia dona resultats són : la robòtica mèdica en les seves vessants de cirurgia i assistencial, la robòtica submarina, en sistemes teleoperats en general, en processos productius automatitzats, anàlisis clínics assistits per ordinador, identificació d'espècies assistida.

La visió per computador juga un paper de primer ordre en aquestes línies de recerca. Es considera com el procés d'extracció d'informació del món físic a partir d'imatges i amb un suport informàtic adient. L'objectiu final és la correcta descripció i interpretació de les escenes representades per les imatges anteriors. Evidentment aquest objectiu es també molt ambiciós i implica a diferents disciplines.

En línies generals es pot dir que en l'extracció de la informació té gran importància el desenvolupament algorísmic clàssic, mentre que la descripció i la interpretació tindria més importància el tractament simbòlic que utilitza la intel·ligència artificial. No obstant, no és fàcil trobar la línia divisòria entre els dos tractaments, des dels que consideren la IA com l'etapa final de la visió per computador als que consideren aquesta última disciplina com una interfície d'entrada de la IA.



Figura 1.1: La simetria en construccions artificials

Per altra banda, quan la visió per computador és utilitzada com a informació sensorial de robots apareix l'exigència suplementària d'obtenir la informació en temps especificat i de tractar amb seqüències de imatges que canvien constantment. El factor temps de processament en aquest cas pren molta importància. Per tant els investigadors en el camp de

la visió per computador estudien algorismes i estratègies per reduir aquest temps, però també dissenyen arquitectures especialitzades quant el temps resposta ho requereix. En qualsevol cas, cal considerar l'extracció de característiques com una de les etapes bàsiques per a la consecució dels objectius plantejats. Aquest procés, per si sol, ja té un grau de complexitat elevada ja que tasques de detecció que poden semblar trivials, impliquen el tractaments d'enormes quantitat de dades i complexes manipulacions matemàtiques, amb la dificultat afegida del temps de processament.

1.1. El paper de la simetria en visió per ordinador

Des de fa molt temps es coneix en psicologia la importància de la simetria com una característica precognitiva que intervé en el reconeixement de formes i objectes [Attneave 55], [Bornstein 84] i com va mostrar Tyler [Tyler 96], el éssers humans presenten una alta sensibilitat a aquesta característica. La importància i utilitat d'aquesta informació també ha estat abastament estudiada per la comunitat científica implicada en visió per computador [Blum 73], [Blum74]

La simetria és present en molts entorns. La trobem de forma parcial o global en molts objectes quotidians, en edificacions artificials (Fig 1.1), textures naturals i artificials, en determinats entorns industrials, en imatges per la diagnosi mèdica. És evident doncs que la detecció d'aquesta característica pot ser de gran ajuda en visió per ordinador. La tasca de la detecció de la simetria és molt extensa i tal com es pot veure a la Fig 1.2 [Amat 95], es poden considerar diferents nivells de processament a partir de la captació sensorial i la digitalització, les quals normalment no comprenen cap tipus de tractament de la imatge.

En un primer nivell trobem el tractament anomenat preprocessament d'imatge, encaminat a millorar la relació entre les dades totals i les dades útils. D'aquest procés s'obté normalment una imatge simplificada de contorns o simplement una imatge binaritzada.

En un segon nivell que anomenat processament d'imatge, s'inclou un conjunt de tractaments destinats fonamentalment a l'extracció de diferents característiques de la imatge, de les quals es pretén obtenir una primera segmentació dels elements de la imatge. Cal precisar no obstant que no sempre s'aconsegueix aquest objectiu en aquesta etapa i es precisa d'un grau de coneixement més alt.

El tercer nivell, que podem anomenar nivell de descripció, consisteix en un conjunt de funcions encaminades al reconeixement i descripció dels elements de l'escena. Si bé en el primer nivell i en part en el segon té especial rellevància l'arquitectura dels processadors utilitzat, en aquest tercer nivell el desenvolupament pren una gran importància.

Finalment, el quart nivell està orientat a la interpretació de l'escena. En aquesta etapa considerada la més intel·ligent, les tècniques anomenades d'intel·ligència artificial prenen el protagonisme.

A la Fig 1.2 es pot veure més detalladament els diferents processos que implica cada nivell i la relació entre ells, tot i que cal advertir que no es pot fer una generalització absoluta, donat que altres autors canvien l'ordre o la relació entre diferents funcions i fins i tot el nivell de tractament d'una determinada funció.

Segons la classificació anterior sembla lògic considerar la simetria com una característica més, obtinguda en el procés d'extracció de característiques, dins el nivell de processament de la imatge que ajudaria a la segmentació i en conseqüència el reconeixement. Tot i que creiem que aquest punt de vista és el més correcte, donada l'enorme dificultat que presenta normalment la segmentació, cap la possibilitat d'obtenir la simetria a partir de formes prèviament segmentades que ens ajudaria directament al reconeixement i a la descripció.

Troblem per tant un ampli ventall d'aportacions on s'utilitza la simetria amb diferents finalitats. Així [Blum 78] i [Rom 93] utilitzen els eixos de simetria obtinguts de les formes de la imatge segmentades prèviament, per tal de fer la seva descripció. En el treball de [Oh 88] s'utilitza la simetria per fer la correspondència de formes mentre que [Reisfeld 92] l'utilitza en el reconeixement formes i d'objectes.

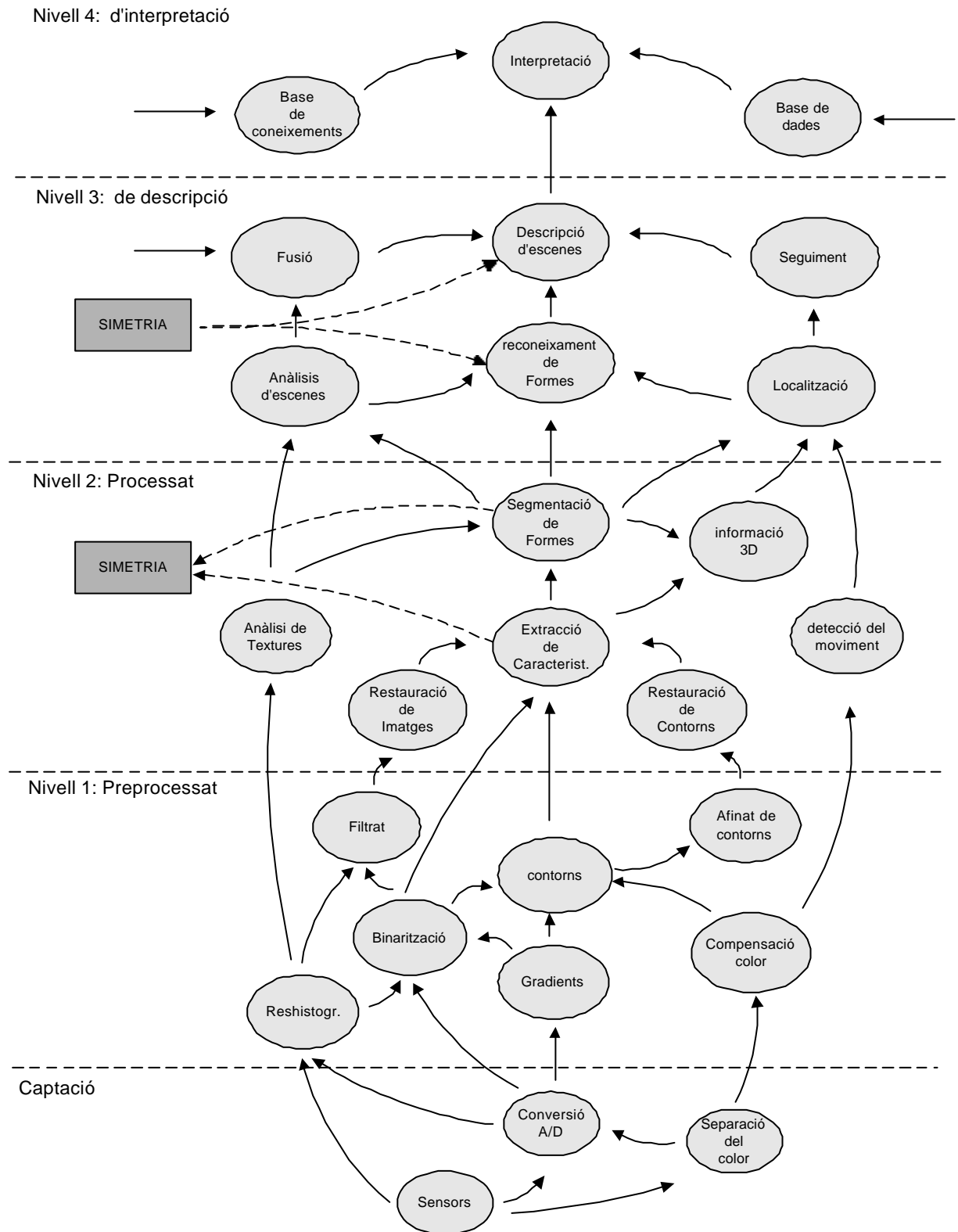


Fig 1.2 Nivells de tractament de la imatge en visió per ordinador

Per altra banda, en un treball força rellevant T.Kanade [Kanade 81] demostra que la simetria pot jugar un paper important en la inferència de la forma 3D, a partir d'un tipus de simetria anomenat simetria obliqua (veure ap. 2.1.3). Més endavant Ulupinar i Nevatia [Ulupinar 90] postulen que la simetria que anomenen simetria paral·lela (veure ap.2.1.5) també ajuda en aquesta finalitat. Aquests treballs han generat un gran interès en la detecció d'aquestes simetria; així [Saint-Marc 90] mostra que la utilització de B_splines facilita la detecció d'aquests tipus de simetria mentre que J. Ponce presenta un algorisme de detecció i infereix el 3D a partir de la simetria obliqua [Ponce 90].

Es poden citar altres aplicacions on s'utilitza la simetria com ara: segmentació [Geiger 98], discriminació de textures [Bonneh 93], compressió d'imatges [Kirby 90], reconstrucció de formes i objectes [Mitsumoto 92],[Zabrodsky 95]

En el cas de la robòtica, esmentat anteriorment, permet la localització i orientació d'objectes i formes [Marola 89]. Aquesta informació és molt important en la tasca d'aprensió dels braços dels robots [Blake 95] i en el seguiment i la navegació de robots mòbils [Kelly 95]

Com veurem més endavant, tot i que existeixen actualment molts mètodes de detecció de simetria, es tracta d'un tema que es manté obert per la seva extensió, complexitat, per les diferents maneres de tractar la simetria i per els diferents tipus de simetria que es poden considerar en visió per ordinador.

1.2. Objectius de la tesi

L'objectiu d'aquesta tesi doctoral és la de proposar una metodologia de detecció de simetria de reflexió. Evidentment treballarem amb imatges amb absència de deformacions de perspectiva de forma que la simetria de l'escena es conservi sobre el pla de projecció. Això comporta obtenir les imatges a partir d'escenes suficientment allunyades de forma que la projecció es pugui considerar ortogràfica i els objectes un cop projectats no presenten



Figura 1.3: Imatges amb projecció de perspectiva

deformació en cap de les direccions dels dos eixos de coordenades x , y tal com es pot veure a la Fig 1.1.

La conseqüència més evident d'utilitzar aquesta forma de captació de la imatge és que les línies paral·leles a l'escena es conserven un cop projectades a la imatge. En termes de simetria de reflexió això significa que es compleixen tres propietats:

- *Paral·lelisme de les línies de simetria*: les línies que uneixen dos punts simètrics a cada costat de la forma simètrica són paral·leles.
- *Ortogonalitat de les línies de simetria*: les línies que uneixen dos punts simètrics a cada costat de la forma simètrica són ortogonals a l'eix de simetria
- *Colinialitat dels punts mitjos de les línies de simetria* : Els punts mitjos de les línies que uneixen les parelles de punt simètrics formen un eix de simetria rectilini.

Així ens referirem al “ punt de vista ortogràfic “ sempre que es pugui assumir la projecció ortogràfica i en conseqüència es compleixen les condicions que hem de citar. A la Fig 1.3 es poden veure imatges que no compleixen aquestes condicions i presenten la projecció amb deformació de perspectiva. La proximitat i l'angle en que s'han captat les escenes fa que la distància entre els punt inicials i finals de les línies paral·leles sigui significativa respecte la distància a que s'observa l'escena, de tal manera que els punt finals de les línies es veuen més pròxims que els punt inicials més llunyans.

Existeixen molts mètodes de detecció de simetria que proporcionen els eixos de simetria de formes planes. El problema és que la majoria treballen amb objectes o formes segmentades. Com és sabut la segmentació dels objectes en una escena és una de les tasques més difícils de realitzar. Per altra banda, la simetria, com hem comentat en l'apartat anterior, pot ajudar també a la segmentació i a la reconstrucció de contorns. És per tot això que sembla més útil detectar la simetria sense segmentació de formes prèvia. Volem proposar doncs un mètode capaç de detectar les simetries, obtenint els eixos de simetria més importants de la imatge, en escenes reals on poden haver-hi diferents objectes, simètrics o no.

Un altre objectiu és la detecció de la llargada dels eixos a part d' obtenir la posició i orientació. La majoria de mètodes de detecció proporcionen els eixos de simetria mitjançant línies infinites de manera que no mostren l'abast de la simetria. Aquesta informació pot ser vital en casos com el de la reconstrucció, localització etc.

En aquesta tesi pretenem també obtenir un mètode que presenti un bon nivell de robustesa en front al soroll i a les imperfeccions i degradacions més corrents en les imatges.

Un problema molt freqüent que dificulta la detecció de simetria es la degradació de les formes a causa de la presència d'ombres, oclusions parcials o soroll. La majoria d'autors treballen a partir de contorns i es troben amb aquesta dificultat. Estudiarem doncs aquesta problemàtica i la possibilitat de proposar una alternativa.

Finalment, com a darrer objectiu, ens proposem obtenir un mètode que sigui útil en aplicacions de temps real com les descrites a la introducció. Buscarem doncs la forma de millorar el temps de processament desenvolupant si cal processadors específics.

El cost de computació elevat és una constant que apareix sempre a l'hora d'obtenir algorisme de detecció de simetria per la qual cosa la millora del temps de processament és una tasca ineludible quan es precisa temps real.

1.3. Organització de la tesi

Aquesta tesi ha estat organitzada en set capítols incloent-hi aquesta primera introducció. En el segon capítol presentem l'estat de l'art. Per tal de fer més comprensibles les diferents formes de tractar la simetria, així com les diferents metodologies en la detecció, la primera part d'aquest capítol inclou un estudi dels diferents tipus de simetria existents des del punt de vista de la visió per computador, amb les definicions, propietats i aplicacions particulars si s'escau.

En el tercer capítol presentem el mètode de detecció de simetria proposat per tal d'assolir el objectius plantejats anteriorment. A la primera part d'aquest, fem un estudi teòric de les condicions d'existència de simetria entre segments rectilinis i definim la contribució a la simetria d'aquests segments. A la segona part presentem el processament d'imatge previ a l'aplicació del mètode. Finalment presentem el mètode proposat, tot aplicant-lo sobre una imatge de prova per tal de fer l'exposició més comprensible.

En el quart capítol ens hem plantejat la millora del temps de processament. S'ha desenvolupat un processador específic per una de les etapes de més alt cost de computació del mètode que consisteix en la generació d'un mapa d'acumulació de segments rectilinis. Com veurem més endavant l'obtenció d'aquest mapa precisa d'obtenir els píxels que aproximem un segment rectilini especificat per els seus punt inicial i final. Aquest problema coincideix amb el problema gràfic conegut a la bibliografia com "traçat de línies". Per aquest motiu en la primera part d'aquest capítol s'estudia aquest problema i es presenta un algorisme paral·lel de traçat de línies que, a més a més, té la propietat de facilitar l'obtenció d'un processador paral·lel específic. La segona part del capítol està dedicada al desenvolupament del processador paral·lel específic dedicat al càlcul del mapa d'acumulació esmentat. Es discuteixen unes primeres simplificacions de forma que el disseny resultant permeti l'obtenció d'un circuit implementat amb tecnologia d'alta escala d'integració (VLSI).

Al capítol cinquè presenta l'avaluació de resultats sobre diferents imatges. Es presenten inicialment sobre imatges sintètiques per tal d'explicitar els errors inherents al propi mètode i posteriorment s'avalua les possibilitat de funcionament sobre imatges d'escenes reals. Es presenten també els resultats respecta el temps de processament així com les millores que aporta el processador específic.

El sisè capítol presenta l'adaptació del disseny del processador específic a un disseny que permeti la implementació del circuit amb tecnologia d'alta escala d'integració. D'aquesta forma s'obté també una primera aproximació de la capacitat de càlcul del processador i de l'àrea de silici necessària.

Finalment presentem les conclusions amb les aportacions en el capítol 7 i concloem amb les referències bibliogràfiques respectivament.