

## 1 Introducció

Aquest projecte s'emmarca dins l'àmbit de la recerca de la Visió per Computador del departament d'Arquitectura i Tecnologia de Computadors de la Universitat de Girona, i més concretament, dins la temàtica de reconstrucció d'informació tridimensional mitjançant sistemes de visió. En els últims anys, i degut a l'evolució de la informàtica i a l'ús d'informació tridimensional, la reconstrucció 3D i l'estudi de models deformables ha adquirit un gran protagonisme sobretot en camps com l'animació 3D, videojocs i en estudis mèdics sobre diferents parts del cos humà. L'objectiu principal d'aquest PFC és el d'estudiar i implementar diferents tècniques que permetin, donada una evolució temporal d'un objecte/estructura 3D, segmentar els punts 3D rígids dels que són deformables i varien la seva forma al llarg del temps, d'aquesta manera podrem separar els punts rígids dels punts no rígids mitjançant mètodes que analitzarem, dissenyarem i implementarem. El motiu d'aquest estudi de separar punts rígids de punts no rígids es bàsicament per poder diferenciar i poder fer un seguiment dels punts deformables i/o poder-ne fer posteriorment una reconstrucció 3D amb models deformables a partir de la identificació inicial de punts rígids.

Per avaluar aquests mètodes de segmentació de punts 3D, utilitzarem la informació tridimensional capturada a partir d'un sistema de visió estereoscòpica. Un sistema de visió estereoscòpic està format per dues càmeres les quals ens realitzaran captures reals des de dos punts de vistes diferents, on hi haurà solapament de les dos imatges, per poder-ne extreure informació tridimensional amb la qual treballarem amb els algorismes que haurem analitzat, dissenyat i implementat. En aquest PFC, això implicarà realitzar: 1) el desenvolupament i la calibració precisa d'un sistema de visió estèreo real; 2) la utilització del software necessari per realitzar la detecció i seguiment del punts d'interès en les imatges 2D capturades pel sistema, i per últim, 3) l'obtenció de les reconstruccions 3D a partir de la triangulació estèreo dels punts 2D. Les seqüències de reconstruccions 3D resultants, s'utilitzaran com a dades d'entrada per a la segmentació de punts rígids/deformables i la seva avaluació. On nosaltres ens centrarem específicament. Per a assolir aquests objectius es treballarà bàsicament amb la plataforma de desenvolupament Matlab. Un entorn molt potent, eficient i molt utilitzat dins el món de la recerca.

Resumint, aquest projecte el dividirem en dos grans parts que serien la part de dissenyar, desenvolupar i implementar els mètodes de segmentació que ens

serviran per separar els punts rígids dels punts no rígids/deformables. I l'altra part seria la d'obtenir reconstruccions 3D a partir d'un sistema estèreo, passant per a la calibració de les càmeres del sistema, la realització de captures d'experiments reals, la generació de reconstruccions 3D per finalment posar a prova els mètodes desenvolupats en la part anterior. A l'esquema de la figura 1.1 s'explica gràficament el que realitzarem al llarg d'aquest PFC.

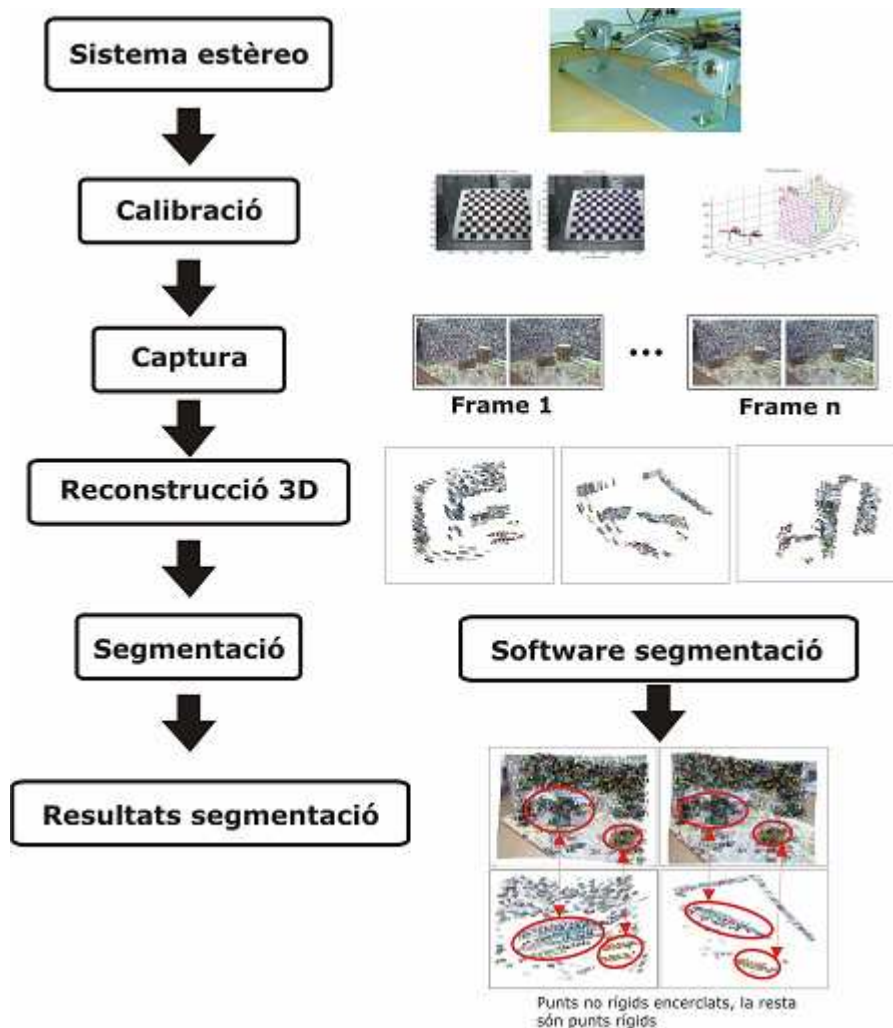


Figura 1.1 Esquema general del projecte.

## 2 Mètodes de segmentació

Primer de tot hem analitzat, dissenyat i implementat tres mètodes per a la segmentació. Aquests tres mètodes són: Mètode Otsu, Mètode K-Means i el mètode LDA (Linear Discriminant Analysis). Són mètodes que s'utilitzaven en altres camps com en la segmentació d'imatges, però que en aquest PFC els hem adaptat per a poder-los tractar amb les dades que nosaltres obtindrem a partir de

seqüències de vídeo, on obtindrem informació sobre punts 3D que hauran aparegut i hàgim trobat amb el software corresponent al llarg de tota la seqüència de vídeo. Per segmentar assumim que tenim la reconstrucció 3D d'un objecte durant un període de temps (seqüència de vídeo). Llavors registrem totes les reconstruccions en el frame inicial que és la primera reconstrucció de la seqüència, assumint també, que hi ha punts rígids a l'objecte obtenim la transformació tant en rotació com en translació a través del mètode Ransac. Al registrar totes les reconstruccions a la inicial observem que els punts rígids no tenen o tenen molt poc error, en el cas dels punts no rígids l'error de posició serà més elevat respecte els no rígids, tal com es pot veure a la figura 2. Llavors al calcular l'error acumulat per a cada punt 3D i sobre aquests errors del registre serà on nosaltres aplicarem els mètodes de segmentació per realment diferenciar els punts rígids dels no rígids. Si tenim més deformació, lògicament, tindrem més error en el registre. Si tenim punts rígids i no hi ha soroll, l'error d'aquests serà 0.

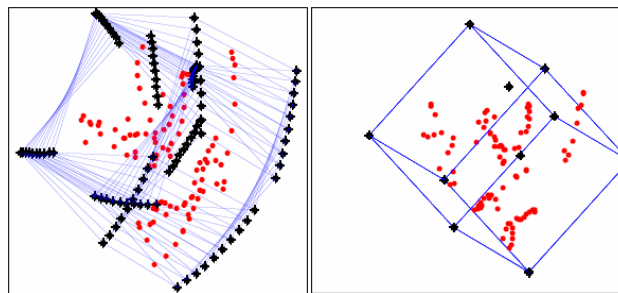


Figura 2. A l'esquerra punts 3D sintètics al llarg de la seqüència. A la dreta registre de tots els punts respecte el frame inicial.

S'han fet proves exhaustives amb dades sintètiques, utilitzant un cub i punts aleatoris escampats per l'escena tant rígids com no rígids, per avaluar els tres mètodes en diferents condicions, com ara: canvis de número punts rígids/no rígids (10/10, 10/30, 10/50, 10/100, 30/10, 30/30, 30/50, 30/100, 50/10, 50/30, 50/50, 50/100), nivell de deformació, soroll gaussià en les imatges 2D (0, 0.5, 1, 1.5, 2). I finalment després d'haver avaluat els mètodes en dades sintètiques ho hem fet amb dades reals provinents d'un sistema estèreo.

### 3 Obtenció de dades a partir d'un sistema estèreo

Per a l'obtenció de dades reals hem desenvolupat un sistema estèreo amb dues càmeres Fire-i Unibrain, on necessitarem fer-hi una calibració del sistema utilitzant la toolbox "Camera Calibration Toolbox for Matlab". Després de fer una acurada calibració del sistema ja podrem realitzar la captura real de vídeo grabant amb les

dues càmeres alhora sobre una escena on hi hagi suficient sol·lapament, bàsicament el que farem serà simular deformacions a partir del moviment en diferents sentits d'objectes rígids, com la imatge real que es pot apreciar en la figura 1.1. Els fitxers de vídeo seran convertits en imatges (una per cada frame). Aleshores s'ha utilitzat un software cedit per en Jordi Ferrer Plana del grup VICOROB per a resoldre el problema de la correspondència i per a fer un seguiment (*tracking*) al llarg de tota la seqüència dels punts 2D trobats en les correspondències esquerra-dreta. Les dades amb les que treballarem i que seran retornades per aquest software són les coordenades 3D de cada un dels punts en cada frame de la seqüència. Llavors al tenir la informació de cada un d'aquests punts 3D per a cada frame de la seqüència ja podem realitzar el mateix procés que hem explicat en l'apartat anterior per tal de realitzar la segmentació de punts rígids i no rígids. Hem avaluat els tres mètodes de segmentació en múltiples experiments reals.

## **4 Conclusions**

A partir d'aquest projecte ens hem introduït al món de la reconstrucció 3D fent unes proves reals de reconstrucció a partir d'un sistema estèreo. Després de desenvolupar el sistema estèreo, hem après a realitzar un acurat procés de calibració de les càmeres, primer calibrant individualment les càmeres per a després realitzar el procés de calibració estèreo per a l'extracció dels paràmetres extrínsecs. Amb una bona calibració llavors hem pogut fer unes captures reals de vídeo per a generar-ne la reconstrucció 3D a partir d'una seqüència d'imatges on hem realitzat un seguiment de diferents punts que es trobaven en moviment per a després posar a prova els mètodes que hem desenvolupat per a realitzar una segmentació de punts rígids i no rígids de l'estructura 3D al llarg d'una seqüència. Els mètodes de segmentació que hem analitzat, dissenyat i implementat per a segmentar punts rígids i no rígids, s'han avaluat primer sobre un conjunt gran de dades sintètiques, amb les que teníem control total d'aquestes, on podíem jugar amb tots els paràmetres possibles d'aquestes dades, tals com serien: la deformació que patirien els punts 3D no rígids, afegint-hi soroll gaussià, augmentant o disminuint tots els punts, tant rígids com no rígids.

Els primers resultats obtinguts han sigut durant l'avaluació dels mètodes de segmentació sobre les dades sintètiques. Els mètodes que hem avaluat han sigut el mètode d'Otsu, el mètode K-Means (utilitzant dos classes) i el mètode LDA (Linear Discriminant Analysis). Amb aquestes dades sintètiques cap mètode ens ha donat

un 100% de segmentació correcta amb tot aquest conjunt de dades, en cada un dels mètodes hem obtingut errors de segmentació, tot i que s'ha de dir que algun mètode ha donat més errors de segmentació que els altres. Això a mesura que variàvem els paràmetres de les dades sintètiques (variant nombre de punts rígids o no rígids, soroll i deformació) aconseguíem diferents resultats de segmentació amb diferents errors de segmentació. El mètode que ens ha donat més bons resultats en aquest apartat és el mètode d'Otsu on hem obtingut molts pocs errors de segmentació en comparació amb els altres dos mètodes, llavors el segueix el mètode K-Means en què també ens ha donat forces bons resultats i finalment el mètode LDA que ens ha donat els pitjors resultats.

Els factors que han influït més alhora de generar-nos errors de segmentació són: el nivell de deformació, el soroll i la proporció entre punts rígids i no rígids. En el cas de tenir poc nivell de deformació les dades tenen un valor més semblant, això s'explica perquè nosaltres ens basem en l'error de posició que tenen punts, i al tenir menys deformació, lògicament, tindrem menys error de posició. A més a més, si també hi tenim més soroll l'error de posició dels punts rígids serà més proper a l'error dels punts no rígids. L'altre cas, al tenir diferents proporcions de punts rígids i no rígids doncs variarà per a cada mètode donant-nos així llinars diferents.

Els resultats que hem obtingut amb les dades reals han estat satisfactoris, ja que en tots els experiments que hem realitzat, excepte un on teníem una deformació progressiva punt a punt, doncs no hem obtingut cap error de segmentació. El motiu principal d'aquest fet és que en les proves reals, els objectes els quals hem donat moviment (simulant deformacions) els hi hem donat molta més deformació que en les dades sintètiques, on aquesta deformació era molt més petita. I en el cas excepcional que hem comentat, la deformació on era progressiva (uns punts tenien més deformació que els altres progressivament), el mètode que ens ha donat més bons resultats és el mètode LDA, on hem obtingut menys errors de segmentació que els altres dos mètodes, tot i ser el mètode que pitjors resultats ens havia donat quan hem avaluat els mètodes sobre dades sintètiques. Amb aquests resultats que hem obtingut al llarg del projecte, tant en dades sintètiques com en dades reals podem arribar a les conclusions que cada mètode pot ser utilitzat depenent dels factors que hem comentat obtenint així menys errors de segmentació utilitzant un mètode en comptes d'un altre. Tot i que en general, podem concloure que Otsu és una bona opció.