



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Escola Tècnica
Superior d'Enginyeria
Informàtica

Escola Tècnica Superior d'Enginyeria Informàtica
Universitat Politècnica de València

Control de qualitat semiautomàtic de raïm de vinificació basat en visió artificial

Projecte Final de Carrera

Enginyeria Informàtica Superior

Autor: Josep Gaietà Camarasa i Baixauli

Director: Antonio José Sánchez Salmerón

Agraïments

Aquest projecte representa el final d'una etapa molt important de la meua vida. Caient de ple en el tòpic, he de dir que arribar fins ací no hauria sigut possible sense certes persones que m'han acompanyat al llarg del camí.

Primer de tot, agrair al meu director, Antonio, per donar-me l'oportunitat de treballar amb ell durant uns sis mesos que han constituït una de les experiències més educatives de la meua vida. El tracte i el recolzament rebut durant la meua participació en el projecte deixen el llistó desafortunadament alt per al que em queda de vida laboral.

Per extensió, vull agrair també als companys del departament AI2. Entre ells, a Eugenio per estar tres mesos aguantant les meues visites diàries carregades de preguntes. I a David, per haver-me ensenyat tantíssimes coses que no sabia ni que necessitava saber, per motivar-me a desenvolupar-me com a programador i per convertir-se en amic ara que ja no serem companys.

Per a arribar a ací, però, han calgut cinc anys. Cinc anys que de vegades han paregut interminables però que en retrospectiva se'm fan tan curts. Agrair també, per tant, a les persones que m'han acompanyat durant aquest temps.

Als meus companys de classe Juaneu, Santi, Vicent, Anna i molts altres que em deixen però que deuen de donar-se per al·ludits, i concretament a Jorge per haver estat al meu costat tot aquest temps i haver-me fet que fins i tot (o especialment?) les pitjors parts de la carrera hagen sigut increïblement divertides. O, com a mínim, passables, que ja te mèrit si et fixes.

Als meus amics, a tots ells sense distinció. Però sobretot a Marcos, David, Tony i Víctor (en cap ordre en particular, que vos veig vindre), per encarregar-se que no em falte mai ningú, per recolzar-me en tot el que he decidit emprendre i per fer-me entendre que no hi ha mal lloc ni moment si és entre bandada de llops.

Menció especial també per a la meua amiga Pau, per oferir-se (de pròpia voluntat, ho promet) a revisar aquest document, per ajudar-me a una cosa tan important com és expressar-me correctament i per ser l'única persona que riu el 100% de les meues bromes.

Finalment, a la meua família en tota la seua extensió, i sobretot a les meues germanes Alba i Elena, a mon pare Josep Joan i a ma mare Pilar per ser i estar, i per ensenyar-me allò que més aprecie en el món: a ser persona.

Resum

En aquest document presentem el desenvolupament d'una ferramenta semiautomàtica per a la reconstrucció de raïm amb l'objectiu de contribuir a l'automatització del control de qualitat sobre raïm de vinificació. Discutirem breument la importància del sector vinícola i el potencial impacte socioeconòmic que suposa control de qualitat automàtic. Seguidament explicarem les tècniques utilitzades tant per a la reconstrucció automàtica de part del raïm com de la reconstrucció manual del mateix. Detallarem també alguns aspectes d'implementació i potencials ampliacions. Finalment, comentarem els resultats obtinguts i n'extraurem conclusions a partir de la informació presentada en aquest document.

Continguts

1	Introducció al problema	7
1.1	Context.....	7
1.2	Objectius	8
1.3	Estat de l'art.....	10
1.4	Dificultats tècniques	11
1.4.1	Ocultació.....	11
1.4.2	Soroll visual.....	11
1.4.3	Geometria dels grans.....	12
1.5	Solució proposada.....	13
2	Tècniques utilitzades.....	14
2.1	Generació automàtica d'hipòtesis.....	14
2.1.1	Generació de múltiples hipòtesis mitjançant la transformada de Hough	16
2.1.2	Reconstrucció dels grans inferior i superior.....	24
2.2	Edició manual.....	26
2.2.1	Visualització i edició d'hipòtesis.....	26
2.2.2	Generació d'hipòtesis.....	27
2.2.3	Projecció d'hipòtesis.....	30
2.3	Refinament automàtic d'hipòtesis.....	31
2.3.1	Optimització d'una sola hipòtesi.....	31
2.3.2	Optimització d'un conjunt d'hipòtesis	33
3	Disseny i implementació	34
3.1	Tecnologia i ferramentes utilitzades	34
3.2	Estructura interna	35
3.2.1	<i>ApplicationData</i>	35
3.2.2	Supervisor i accions	37
3.3	Interfície gràfica	38
4	Resultats.....	42
4.1	Reconstrucció manual.....	42
4.2	Reconstrucció automàtica	47
5	Conclusions	51
Annex A.	Manual d'usuari	52
	Introducció.....	52
	Operacions generals	53
	<i>Viewport2D</i>	54
	<i>Viewport3D</i>	57

Annex B. Manual de programador 59

1 Introducció al problema

1.1 Context

Els sistemes d'inspecció visual automàtica 2D basats en visió per computador han demostrat en els últims anys ser una ferramenta fonamental per a la millora de la qualitat dels processos i productes. Aquests sistemes permeten la inspecció contínua, evitant fatigues i distraccions, i facilitant la quantificació de les variables de qualitat en pràcticament el 100% dels productes. Açò és tradueix, no sols en la millora de la qualitat final dels productes, sinó també en un estalvi en termes econòmics i mediambientals.

Hui en dia es disposa d'una gran quantitat d'exemples d'implantació de sistemes d'inspecció 2D en l'indústria agroalimentària. D'altra banda, els enormes avanços tecnològics en l'àmbit dels sensors 3D propicien nous reptes.

Aquest projecte s'emmarca dins del sector vitivinícola. Espanya compta amb casi 1,2 milions d'hectàrees de vinyes, cosa que la situa com al país amb la major extensió cultivada de vinyes del món. Les varietats natives de raïm a Espanya són abundants, amb més de 600 varietats natives plantades per tot el territori. Algunes de les zones vinícoles més conegudes internacionalment són Rioja, Ribera del Duero, famosa per la seua producció d'Ull de llebre (*Tempranillo*); el Marco de Jerez, pels seus vins fortificats; Rías Baixas, pels seus vins blancs d'Albariño; Penedès, per la producció de cava; i Priorat.

La qualitat del vi s'estima en part a partir de paràmetres subjectius (ja que depén del gust del consumidor) i també a partir de paràmetres objectius (ja que existeixen uns criteris establerts per a definir-la).

Entre els factors que defineixen la qualitat del vi destaca la importància de la qualitat del raïm com a matèria prima. Desgraciadament, el concepte de qualitat del raïm és difícil de valorar a causa de problemes, com ara, l'existència de paràmetres subjectius, el curt període de temps disponible per a realitzar els anàlisis en temps de verema, la manca d'instruments de mesura i el seu elevat cost, i la mescla de raïm de bona i mala qualitat a les grones de descàrrega.

Incrementar l'objectivitat i automatitzar les tasques de supervisió de la qualitat del raïm pot ser un gran avanç tecnològic respecte al sistema tradicional d'avaluació del raïm basat en els coneixements d'un expert, i suposaria un gran impacte en la indústria vinícola.

1.2 Objectius

El present projecte tracta, per tant, de contribuir a automatitzar el procés de valoració del raïm rebut en els centres d'acollida de les cooperatives mitjançant l'ús de tècniques de visió artificial tridimensional.

Per a aconseguir-ho es proposa el desenvolupament d'una ferramenta semiautomàtica de reconstrucció dels grans visibles d'un raïm. Aquest procés es realitzarà a partir d'imatges obtingudes mitjançant un sistema de visió estereoscòpica amb configuració de càmeres en paral·lel.

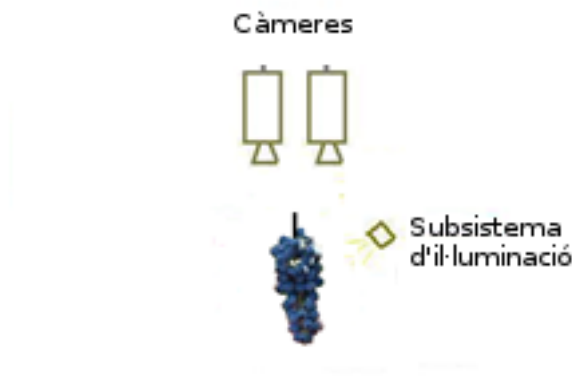


Figura 1.2-1 Esquema del sistema que captura les imatges del raïm.

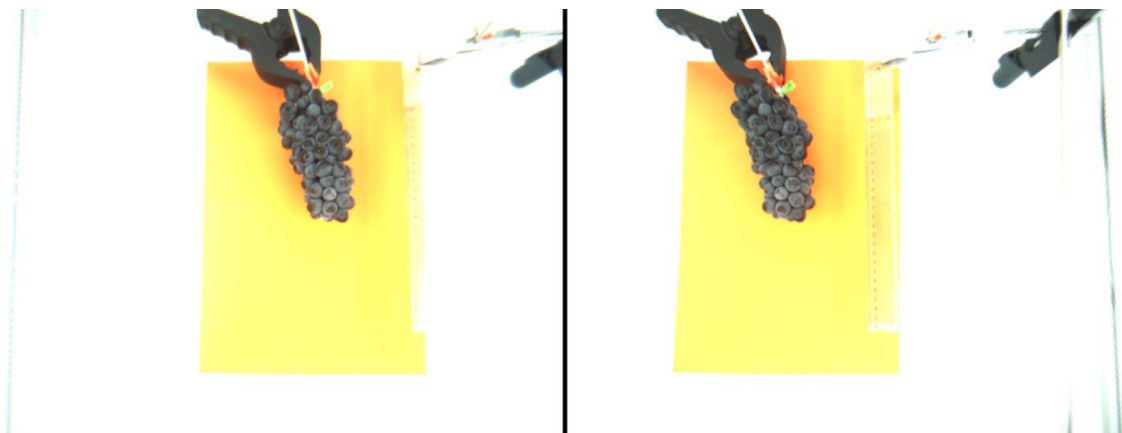


Figura 1.2-2 Dues imatges obtingudes mitjançant el sistema descrit.

Aquesta ferramenta rebrà com a únics paràmetres el calibratge de les càmeres i les dues imatges capturades i tindrà l'objectiu d'obtenir un model tridimensional dels grans del raïm. A més, aquest model es podrà emmagatzemar a disc per a tornar a ser carregat en qualsevol moment per aquesta o altres ferramentes.

Tot i que l'objectiu principal del projecte és la realització d'una ferramenta amb aquestes funcionalitats i característiques, es pretén també definir les bases per a la realització d'un procés (quasi) automàtic de reconstrucció del raïm. Per tant, a part de desenvolupar tècniques de reconstrucció basades en visió artificial tridimensional també s'intentarà dissenyar les bases d'un sistema de supervisió automàtic, en el qual

el programador pugua afegir intel·ligència i coneixement *a priori* al procés de reconstrucció (veure secció 3.2.2).

Podem dir, doncs, que la ferramenta i les tècniques presentades han sigut desenvolupades amb la seua posterior extensió en ment. És per això que, a més de realitzar un manual d'usuari per a la ferramenta, s'ha generat també un manual del programador amb documentació del codi font.

1.3 Estat de l'art

En l'actualitat la classificació del raïm es realitza a partir d'una inspecció per part d'un expert que el puntua en funció dels paràmetres que determinen la seua qualitat. També es realitzen diferents mostres en el laboratori per a estimar la quantitat de sucre, el pH, l'acidesa total i la qualitat fenòlica. Recentment, investigadors de Logronyo han definit un procediment objectiu que permet avaluar la qualitat del raïm en la vinya o en la cooperativa mitjançant l'estimació d'un conjunt de paràmetres definits en la denominada fitxa VITUR. Els paràmetres a estimar en aquesta fitxa són, entre altres, la grandària de la baia, la grandària del raïm i la compacitat del raïm.

Per una altra part, actualment s'estan desenvolupant projectes d'investigació on s'utilitzen tècniques de visió artificial per a determinar automàticament paràmetres de qualitat basats en el color, com la textura. No obstant això, no es coneix cap projecte que avalue automàticament el raïm mitjançant tècniques de visió 3D.

La introducció d'un sistema de visió tridimensional en la classificació del raïm permetrà estimar els paràmetres geomètrics dels raïms i, en conseqüència, avaluar la seua qualitat.

1.4 Dificultats tècniques

Donada la naturalesa del problema, existeixen diversos factors que dificulten la reconstrucció i que s'han tingut en compte a l'hora d'elegir tècniques utilitzades.

1.4.1 Ocultació

Segons la seua posició els grans es poden ocultar entre si. És a dir, alguns grans sols seran visibles parcialment des del punt de vista de la càmera. A més, donada la diferència de posició entre les dos càmeres, la ocultació entre els grans variarà d'una càmera a l'altra. En definitiva, alguns grans s'hauran de reconstruir a partir de informació limitada.

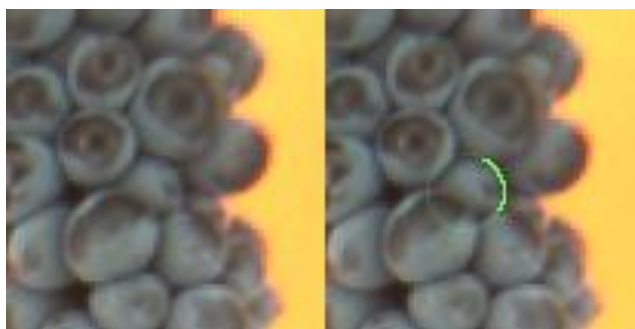


Figura 1.4-1 Exemple de gra del qual sols tenim informació parcial.

1.4.2 Soroll visual

A les fotografies originals es poden observar diversos tipus de soroll sobre el raïm. Això dificulta la detecció dels grans sobre la imatge i pot induir a assumpcions errònies.

Per una part, els grans no són homogenis quant a color sinó que tenen nombroses imperfeccions a la seua superfície, bé per brutícia, bé per variacions de colors inherents al gra.



Figura 1.4-2 Irregularitats a la superfície del gra.

Per una altra part, les fotografies originals estan fetes amb il·luminació frontal. Això provoca l'aparició de reflexos especulars sobre la superfície dels grans. Aquests reflexos varien entre càmera i càmera, la qual cosa suposa una dificultat afegida per a la detecció de característiques.



Figura 1.4-3 Reflexions circulars sobre alguns dels grans.

1.4.3 Geometria dels grans

Com s'explicarà en seccions posteriors, per simplicitat cada gra es modela com a una esfera perfecta. Els grans reals, en canvi, no s'ajusten perfectament a aquest model (semblant-se més a un ovoide), la qual cosa dificulta la seua reconstrucció.

1.5 Solució proposada

La solució que es proposa consisteix en una aplicació que ajude a reconstruir el model 3D d'un raïm donat dos fotografies 2D estereoscòpiques i el calibratge de les respectives càmeres. Entenem la reconstrucció d'un raïm com a la reconstrucció de cadascun dels seus grans visibles.

L'aplicació ha de reconstruir la major quantitat de grans possibles de manera automàtica. És a dir, l'aplicació ha de ser capaç de generar un percentatge dels grans totals sense que l'usuari intervinga, utilitzant com a única ajuda per a la reconstrucció les dos fotografies originals i el calibratge de les càmeres.

Per una altra part, també ha de ser possible la reconstrucció de grans de manera manual, és a dir, amb intervenció de l'usuari. Per a facilitar aquesta tasca se li proporcionarà a l'usuari eines que acceleren el procés de creació de nous grans i d'edició dels ja existents.

Finalment, tota l'aplicació serà presentada mitjançant una interfície gràfica que permeti a l'usuari visualitzar i interactuar amb el model de manera intuïtiva.

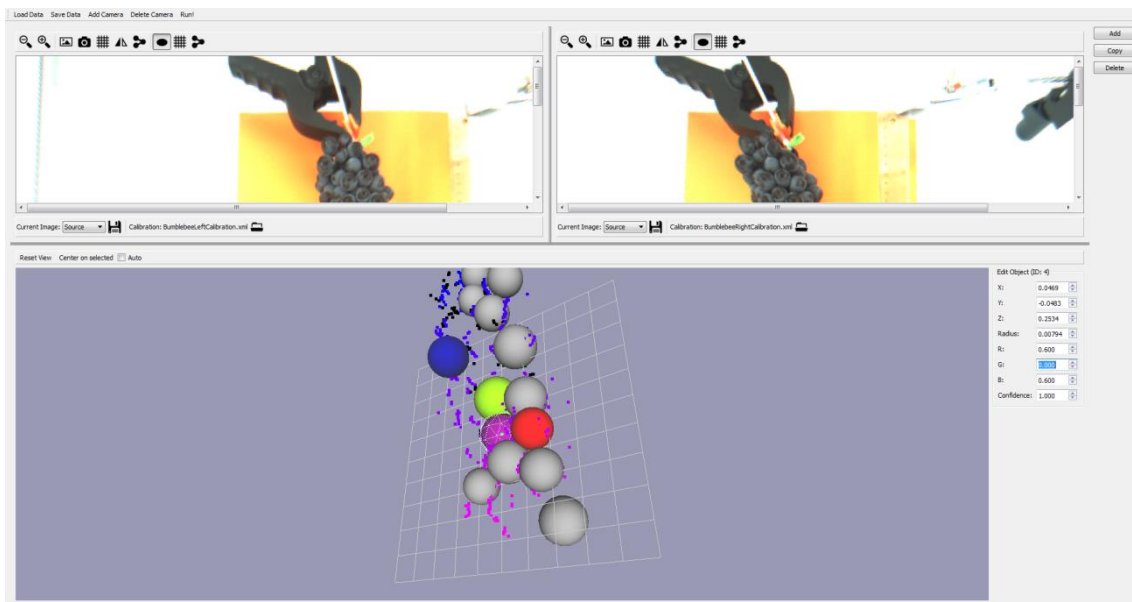


Figura 1.5-1 Captura de pantalla de l'aplicació final.

2 Tècniques utilitzades

2.1 Generació automàtica d'hipòtesis

En aquesta secció i durant tot el document el terme hipòtesi farà referència a una esfera 3D que representa un gra del raïm. Aquesta hipòtesi pot ser més o menys fidel al gra que representa. Entenem com a hipòtesi correcta, per tant, aquella que representa el seu gra amb un nivell mínim de fidelitat. Tal com ha sigut comentat anteriorment, aquesta etapa té com a objectiu generar tantes hipòtesis correctes de manera automàtica com siga possible.

Per a poder realitzar la generació d'hipòtesis és necessari la identificació de característiques especials a les fotografies originals. Una aproximació possible al problema seria la cerca d'el·lipses a les fotografies originals i realitzar la projecció inversa per a obtindre la esfera 3D.

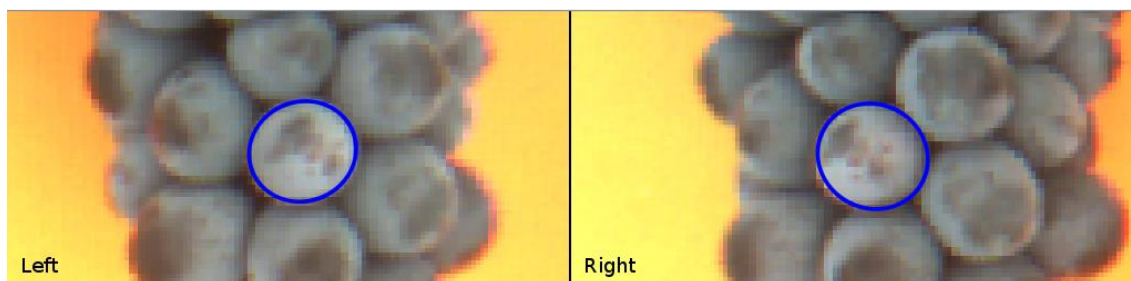


Figura 2.1-1 A partir de dos el·lipses corresponents a un mateix gra és possible obtindre l'esfera 3D corresponent.

Els problemes detallats en la secció 1.4, però, dificulten severament la identificació correcta de les vores 2D dels grans per a poder obtenir les el·lipses corresponents als grans. A més, ens obliga a treballar sobre l'espai 2D durant la major part del procés, el qual limita enormement la informació que podem utilitzar, ja que els coneixements dels que disposem *a priori* són del raïm en 3D, i no d'una projecció 2D. Cal tindre en compte que, tot i això, aquest principi s'utilitzarà per a la reconstrucció manual, explicada més avant en el document.

L'aproximació utilitzada a la solució proposada és realitzar la transformació de 2D a 3D en una etapa anterior. És a dir, processar les característiques 2D i transformar-les immediatament a 3D, el qual permet treballar sobre l'espai 3D pràcticament des del primer moment. D'aquesta manera treballarem de manera més propera al model que volem reconstruir i disposem d'una dimensió més (i, per tant, amb més informació). A més, ens permet aprofitar la informació que tenim sobre la geometria del raïm i obtenir-ne de nova per a ser utilitzada posteriorment.

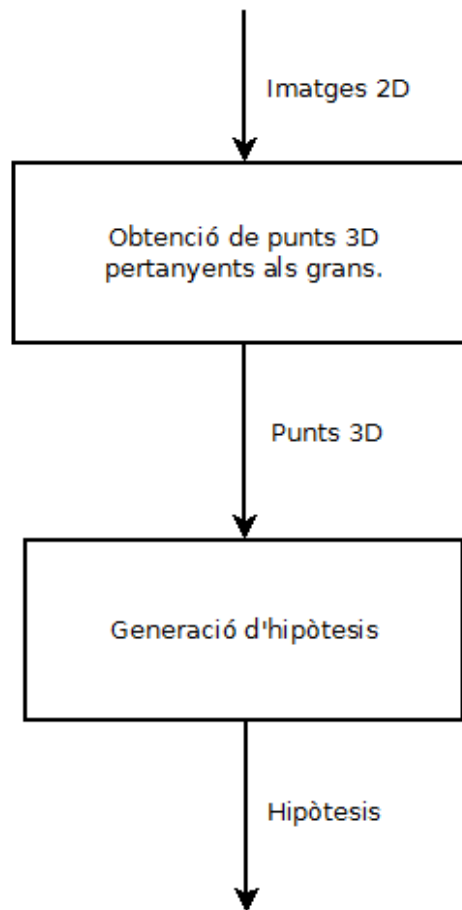


Figura 2.1-2 Procés genèric per a la reconstrucció d'hipòtesis.

Seguint aquest principi el problema de la generació d'hipòtesis a partir de les fotografies es pot dividir en dos subproblemes: l'extracció de les característiques i la generació de les esferes a partir d'aquestes. S'han utilitzat diverses combinacions de tècniques per a solucionar aquest problema, les quals explicarem a continuació.

2.1.1 Generació de múltiples hipòtesis mitjançant la transformada de Hough

La transformada de Hough és un algorisme que permet trobar certs cossos geomètrics en un espai. Tradicionalment, la transformada de Hough s'ha emprat per a la cerca de línies en una imatge 2D. Tot i això, aquest algorisme pot ser emprat per a trobar cossos en un espai n-dimensional.

En el cas que ens ocupa, s'ha utilitzat per a trobar esferes en un espai 3D. Cada esfera identificada mitjançant aquest algorisme serà considerada una hipòtesi que representa un gra del raïm real.

La transformada de Hough, en aquest cas, rep com a paràmetres dos elements: un domini i un conjunt de punts.

El domini representa tot l'espai possible de solucions. Com hem comentat anteriorment una solució en aquest cas és una esfera, la qual es modelitza segons quatre paràmetres: centre (x, y, z) i radi. Per tant, donat que una solució consta dels paràmetres $[x, y, z, r]$, l'espai de solucions ha de constar del mateix nombre de dimensions.

Aquest domini es divideix en elements discrets, anomenats cel·les acumuladores. L'algorisme repartirà vots entre les cel·les, les quals aniran acumulant els vots. A la fi del algorisme les cel·les que tinguen un nombre de vots superior a un cert llindar seran considerades solucions.

En resum, el domini és un espai de 4 dimensions discretitzat en cel·les, on cada cel·la correspon a una solució possible. Com a conseqüència, com major siga el nivell de discretització, major serà la precisió obtinguda en els resultats. Al mateix temps augmentarà també el cost computacional de l'algorisme, el qual depèn directament del nombre de cel·les del domini.

El segon paràmetre és un conjunt de punts que, idealment, pertanyen a la superfície dels grans. Aquests punts són els obtinguts a partir de la transformació de les característiques 2D a punts 3D. Com s'explicarà a continuació, en la pràctica és molt probable que alguns dels punts obtinguts no s'ajusten bé a la superfície o que directament siguin soroll. Això no suposa un problema atés que la transformada de Hough és tolerant al soroll gràcies al seu sistema de votacions.

2.1.1.1 Extracció dels punts 3D a partir de característiques 2D

L'objectiu d'aquest mètode és obtenir punts 3D que corresponguin a la superfície dels diferents grans del raïm per a poder aplicar la transformada de Hough i generar hipòtesis.

Disposant del calibratge de les càmeres és possible obtenir un punt concret de l'espai 3D a partir d'identificar aquest punt en les dos imatges 2D.

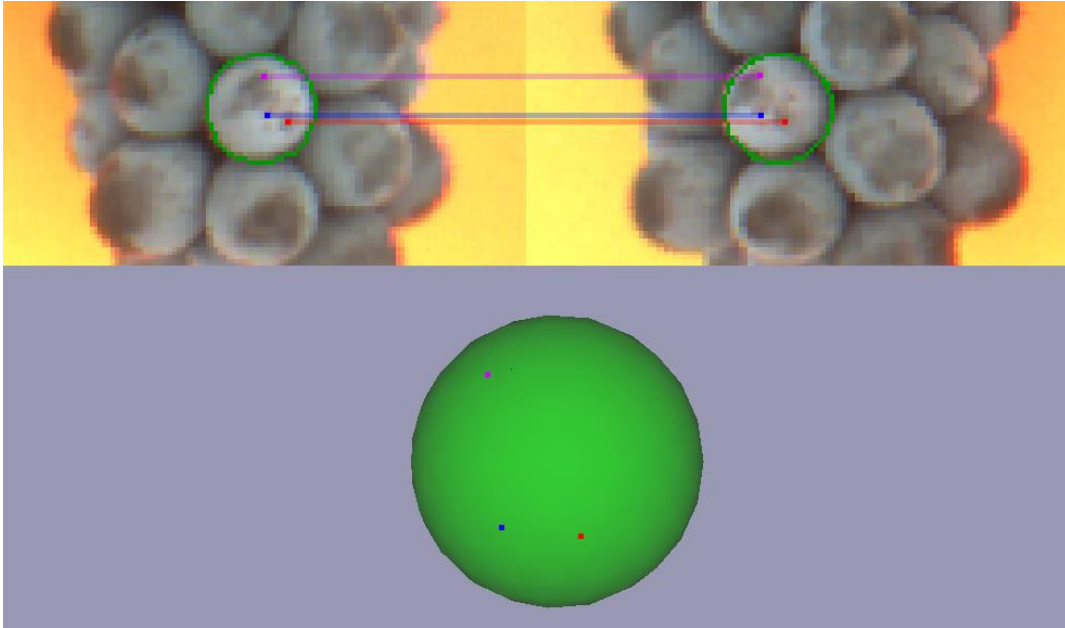


Figura 2.1-3 Exemple de tres punts identificats a les dos imatges i transformats a l'espai 3D.

A partir d'aquest principi el problema d'obtindre els punts 3D necessaris per a aplicar Hough es redueix a buscar correspondències entre punts de la imatge esquerra i punts de la imatge dreta¹.

Inicialment, es va intentar una aproximació directa a l'emparellament de característiques mitjançant les eines proporcionades per la llibreria *OpenCV*. Desafortunadament, prompte va quedar patent que una aproximació genèrica al problema de les correspondències era insuficient. Les característiques trobades entre les dos imatges no eren consistents i a sovint eren emparellades erròniament la qual cosa donava com a resultat punts a l'espai 3D que no corresponien a cap gra.

Com a conseqüència es va optar per aprofitar la informació que tenim sobre les fotografies originals per a introduir restriccions en l'emparellament.

2.1.1.1.1 Restriccions utilitzades

Per una part la disposició de les càmeres. Les càmeres estan situades en un arranjament horitzontal. És a dir, hi ha una diferència de posició a l'eix X mentre que ambdues tenen la mateixa posició a l'eix Y. Això implica que la correspondència **d'un punt de la imatge esquerra es trobarà a la mateixa fila en la imatge dreta** i viceversa (veure Figura 2.1-3 Exemple de tres punts identificats a les dos imatges i transformats a l'espai 3D.).

Per un altra part, la línia base(separació entre les càmeres). La separació entre les dues càmeres és d'aproximadament 11cm i la distància de les càmeres a l'objecte d'interès és d'aproximadament 25cm sobre l'eix Z. Això ens permet assumir com a certa una premissa que en general és errònia: **el contorn d'un gra vist des d'una càmera correspon al del mateix gra vist des de l'altra càmera**.

¹ Aquesta tècnica és coneguda com a *feature matching*.

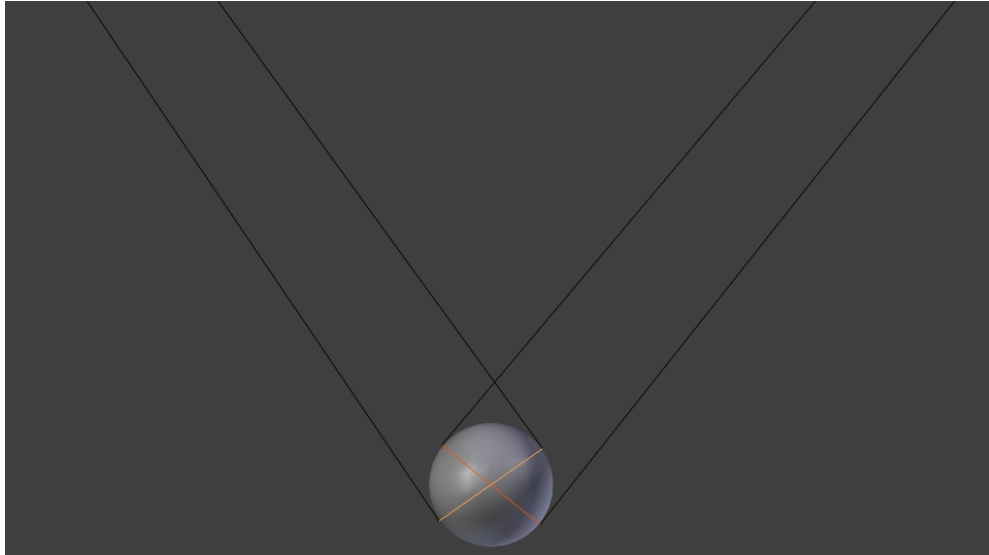


Figura 2.1-4 Per a un cas general, el contorn d'un gra (línies de color taronja) vist des de diferents càmeres difereixen d'una manera notòria.

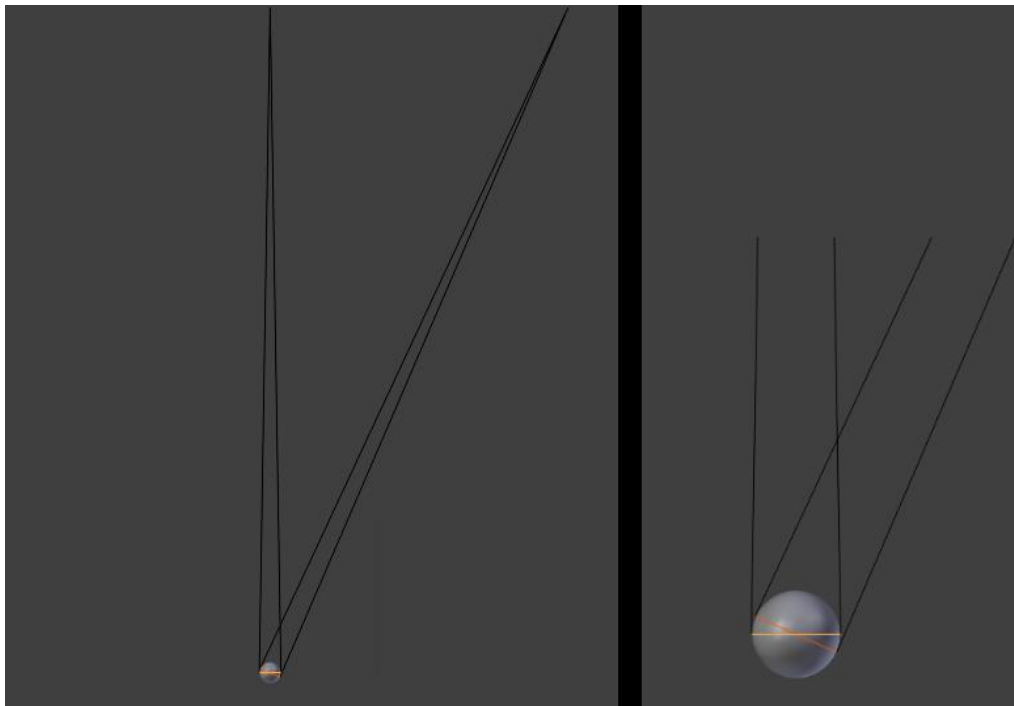


Figura 2.1-5 En el nostre cas, però, la diferència entre els contorns és negligible.

Aquesta assumció introdueix un error en la reconstrucció dels punts 3D. A continuació fem una anàlisi matemàtica d'aquest error per a determinar la seua magnitud.

Aquest error és la distància del punt d'intersecció entre les tangents de les dues càmeres al gra i el gra en si. El error ve determinat, per tant, per la línia base (distància entre les càmeres, b), distància del gra a les càmeres (z) i radi del gra (r).

Realitzant el càlcul per al cas que ens ocupa amb un radi estàndard ($b = 0.1197\text{m}$, $z = 0.24\text{m}$, $r = 0.0075\text{m}$), l'error és d'aproximadament 0.00026m , és a dir, 0.26mm .

Com podem comprovar, és un error negligible i que ens permet identificar un major nombre d'hipòtesis, encara que no siguin completament fidels als grans que representen. Tot i això, a l'etapa d'optimització d'hipòtesis aquest error serà redreçat mitjançant el refinament de les hipòtesis.

Adjuntem, a continuació, un estudi de l'error per a diferents configuracions de càmeres i radis de gra.

r = 0.005m	b = 0.1m	b = 0.2m	b = 0.4m
z = 0.1m	0.49mm	0.94mm	1.40mm
z = 0.2m	0.17mm	0.44mm	0.91mm
z = 0.3m	0.07mm	0.24mm	0.61mm

r = 0.0075m	b = 0.1m	b = 0.2m	b = 0.4m
z = 0.1m	0.81mm	1.5mm	2.14mm
z = 0.2m	0.27mm	0.7mm	1.4mm
z = 0.3m	0.13mm	0.38mm	0.94mm

r = 0.01m	b = 0.1m	b = 0.2m	b = 0.4m
z = 0.1m	1.1mm	2.07mm	2.90mm
z = 0.2m	0.39mm	0.98mm	1.9mm
z = 0.3m	0.19mm	0.54mm	1.29mm

Com podem observar a les taules, aquest error augmenta amb la línia base i amb el radi, i disminueix amb la z . Podem veure, també, que a partir d'un radi d'1cm, una línia base de 40cm i una profunditat de 10cm l'error deixa de ser negligible, i arriba quasi als 3mm.

Aquestes restriccions limiten en gran mesura les opcions d'emparellament entre els punts característics. Per una part, en compte d'emparellar punts característics de tota la imatge aquest procés es pot fer fila a fila. Per una altra part, ens podem centrar la cerca de característiques en el contorn del gra, on generalment la diferència de color és major i, per tant, és més probable trobar punts característics.

Si tenim això en compte aconseguim reduir el problema a l'emparellament de punts característics d'una fila de la imatge esquerra amb punts de la mateixa fila a la imatge dreta. Primerament intentem identificar les característiques mitjançant un algorisme de detecció de vores sobre les imatges originals i tot seguit les emparellem utilitzant programació dinàmica.

2.1.1.1.2 Extracció d'informació de vores mitjançant *Sobel*

Per a poder realitzar l'emparellament de punts hem de triar factors que identifiquen un mateix punt característic a les dues imatges. Com que els punts característics que busquem estaran als contorns dels grans o als punts d'alt contrast, un algorisme de detecció de vores sobre la imatge ens proporciona la informació per a identificar-los.

Amb aquest fi s'ha utilitzat el algorisme de *Sobel* pel gradient horitzontal de les imatges. Això ens permet detectar les vores verticals de la imatge i extraure la magnitud i l'angle del gradient com a factors. La detecció s'ha limitat a les vores verticals atés que l'exploració es fa fila a fila i, per tant, les vores horitzontals no aporten informació.



Figura 2.1-6 Visualització d'una dels raïms originals després d'haver aplicat *Sobel*.

Com a resultat obtenim, per a cada píxel de les imatges originals, els valors de la magnitud, el sinus i el cosinus de la imatge de gradients. Aquests seran utilitzats en la següent etapa per a emparellar punts característics que tinguen valors similars.

2.1.1.1.3 Emparellament mitjançant programació dinàmica

Per a realitzar l'emparellament ens basem en la similitud d'un punt determinat d'una imatge amb un punt de l'altra imatge. Per a mesurar aquesta similitud s'utilitzen els factors extrets a l'apartat anterior (magnitud, sinus i cosinus).

El problema de l'emparellament s'ha modelitzat com un problema d'optimització en el que hem de minimitzar el cost d'emparellar els punts de la fila de la imatge esquerra amb els de la mateixa fila de la imatge dreta.

Per a resoldre'l hem utilitzat un algorisme de programació dinàmica on s'explora un arbre de solucions. Un node d'aquest arbre consisteix en índexs per a les dues imatges (esquerra i dreta), el cost fins al moment i emparellaments realitzats fins al moment. Els índexs indiquen quins punts de les imatges estem intentant emparellar.

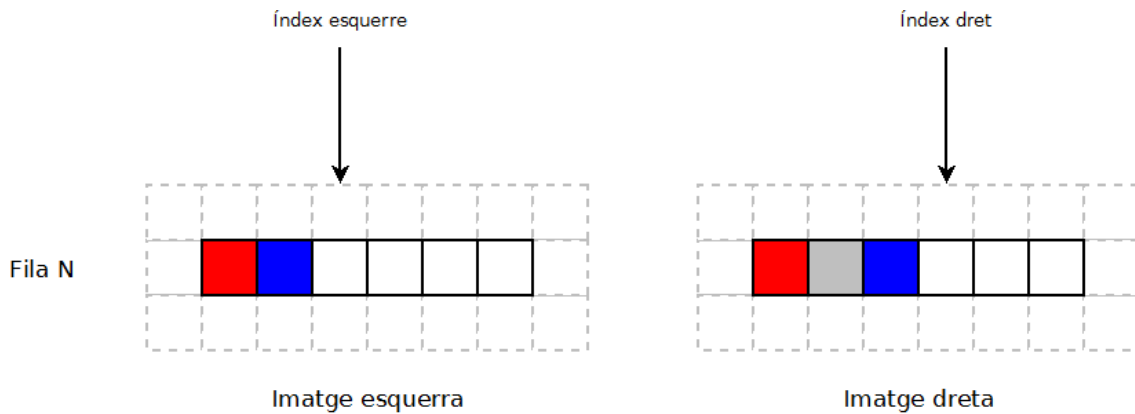


Figura 2.1-7 Exemple d'un node de l'arbre. Els punts del mateix color representen els emparellats. El punt gris representa que ha sigut considerat com a ocult a l'altra imatge.

De cada node pengen potencialment tres nodes més. El primer és el cas en el qual assumim que el punt de la imatge esquerra està ocult en la imatge de la dreta i per tant no es pot emparellar. En tal cas passem al punt següent a través d'incrementar l'índex esquerre i sumem una constant al cost total².

El segon cas és en el que decidim emparellar els dos punts. En tal cas afegim el nou emparellament i sumem el cost al cost total. Aquest cost està basat en la diferència entre la magnitud, sinus i cosinus (calculats prèviament) entre el punt de la imatge esquerra i el de la dreta.

$$\begin{aligned}
 C_{magnitud} &= \text{abs}(\text{normalize}(Mag_{left}) - \text{normalize}(Mag_{right})) \\
 C_{sinus} &= \text{abs}(\text{Sin}_{left} - \text{Sin}_{right})/2 \\
 C_{cosinus} &= \text{abs}(\text{Cos}_{left} - \text{Cos}_{right})/2 \\
 C_{total} &= C_{magnitud} + C_{sinus} + C_{cosinus}
 \end{aligned}$$

Figura 2.1-8 Càlcul del cost per a un emparellament.

Cadascun dels costos individuals estarà sempre normalitzat entre 0 i 1, per tant el cost total sempre estarà comprès entre 0 i 3. També s'incrementaran els dos índexs, així és procedeix a intentar emparellar els dos punts següents.

El tercer cas és equivalent al primer, però assumim que el punt de la imatge dreta és el que està ocult i, per tant, incrementem el índex dret i sumem la constant al cost.

L'algorisme expandeix l'arbre i realitza la cerca fins trobar la solució òptima, és a dir, el conjunt d'emparellaments de cost mínim.

Una volta realitzats els emparellaments obtenim els punts 3D a partir d'una simple projecció inversa. Aquests punts són els que utilitzarà la transformada de Hough per a generar hipòtesis.

² Sense aquesta constant l'algorisme sempre elegiria una solució on tots els punts estarien considerats ocults ja que tindria un cost total 0.

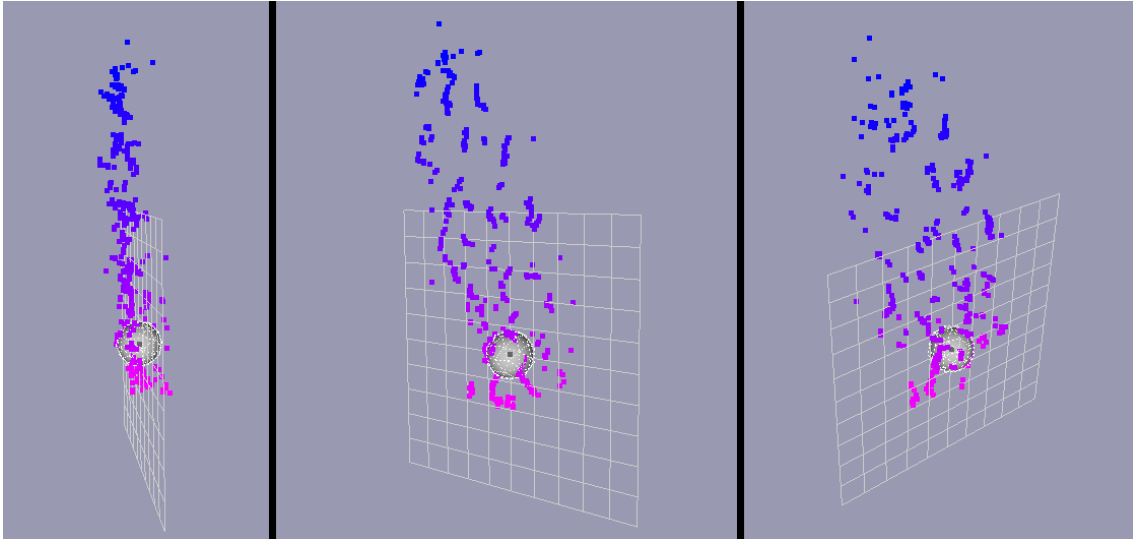


Figura 2.1-9 Resultat de l'extracció de punts característics junt a un gra de referència, vist des de diferents punts de vista.

2.1.1.1.4 Extracció de punts entre grans

Com veurem en la secció 2.1.1.2, és interessant disposar també dels punts entre grans com a informació addicional.

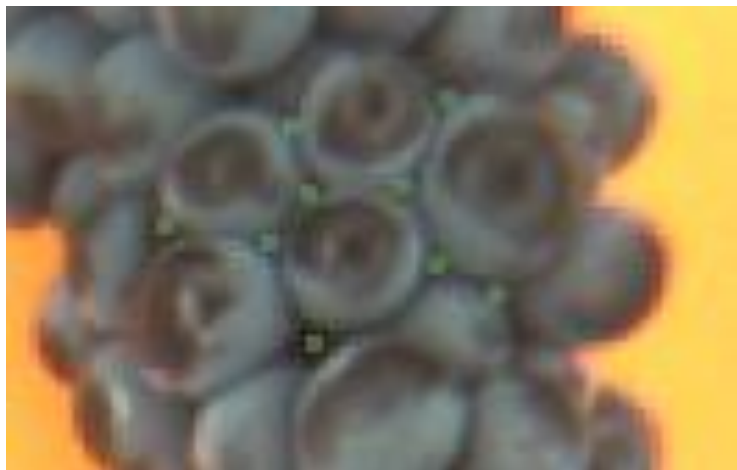


Figura 2.1-10 Punts entre grans.

El procés per a reconstruir aquests punts és molt similar al que hem comentat abans i per tant sols mostrarem el resultat d'exemple.

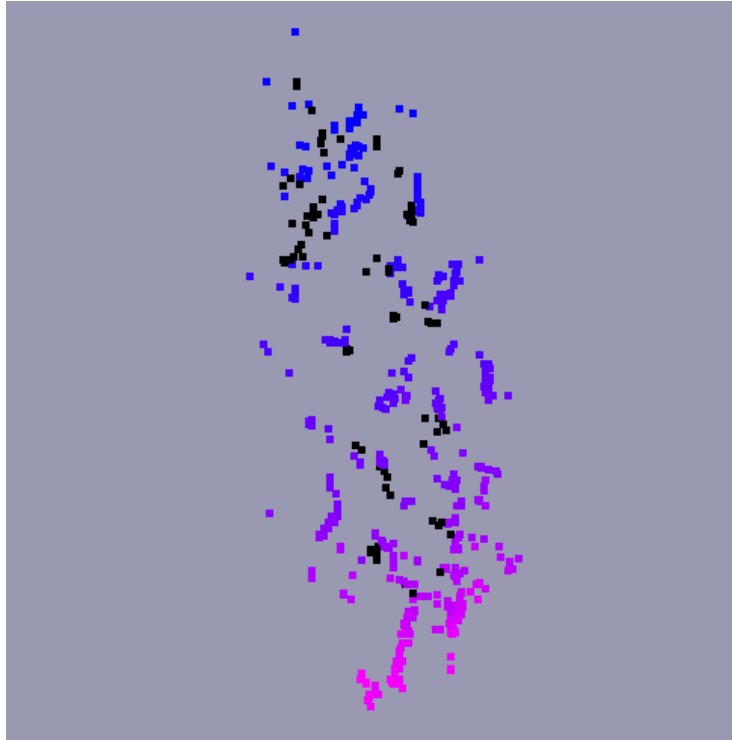


Figura 2.1-11 Resultat combinat de l'extracció de punts característics (color) i de punts entre grans (negres).

2.1.1.2 Aplicació de la transformada de Hough i filtrat del resultat.

Amb el domini definit i els punts característics reconstruïts sols resta aplicar la transformada de Hough per a obtenir les hipòtesis.

No descriurem el funcionament de l'algorisme genèric de Hough, doncs el lector pot trobar fàcilment informació al respecte a la xarxa³.

El resultat d'aplicar l'algorisme és una sèrie d'hipòtesis amb un nombre determinat de vots. Desafortunadament, part d'aquest conjunt d'hipòtesis pot ser erroni i cal ser processat abans de poder ser considerat com un resultat vàlid.

Generalment, com més vots reba una hipòtesi més fidel serà a un gra. Per tant el primer pas serà descartar aquelles hipòtesis que no arriben a un mínim de vots.

El segon pas serà utilitzar els punts entre grans calculats prèviament per a descartar hipòtesis. Sabem que idealment cap hipòtesi hauria de tindre cap d'aquests punts al seu interior per tant realitzarem la comprovació per a filtrar aquelles hipòtesis no desitjades⁴.

Finalment, és possible que la transformada de Hough retorne dues hipòtesis que en realitat corresponen al mateix gra. Per la qual cosa fusionarem aquelles hipòtesis que siguin molt similars quant a característiques.

³ http://ca.wikipedia.org/wiki/Transformada_de_Hough

⁴ En realitat es deixa un marge d'error: si el punt està prop de la superfície de la hipòtesi no serà necessari descartar-la.

Després d'aplicar aquestes mesures podem considerar el resultat final com a vàlid, encara que manca fer un últim refinament que comentarem més avant.

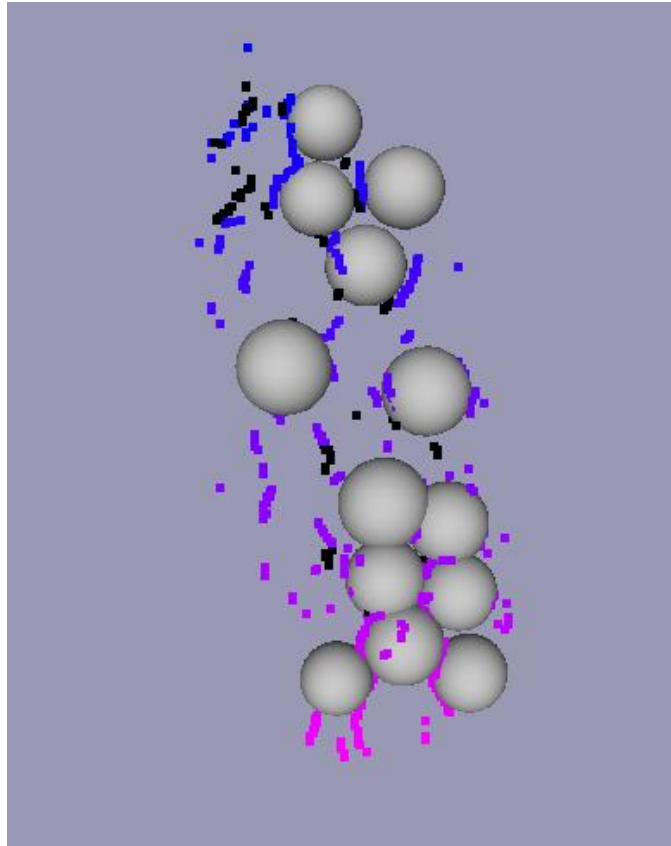


Figura 2.1-12 Hipòtesis obtingudes mitjançant la transformada de Hough després del postprocés.

2.1.2 Reconstrucció dels grans inferior i superior

La relació de posició entre les dues càmeres dota als grans situats als extrems inferior i superior del raïm de propietats especials que permeten la seua reconstrucció de manera més precisa mitjançant un mètode alternatiu al descrit en l'apartat anterior.

Aquest mètode es basa a aprofitar el fet que el punt més baix del gra més baix del raïm (gra inferior) coincideix a les dues imatges.

A partir d'aquest punt és possible reconstruir els punts adjacents que pertanyen al mateix gra, tant per l'esquerra com per la dreta. D'aquesta manera podem obtenir amb més precisió part del contorn del gra.

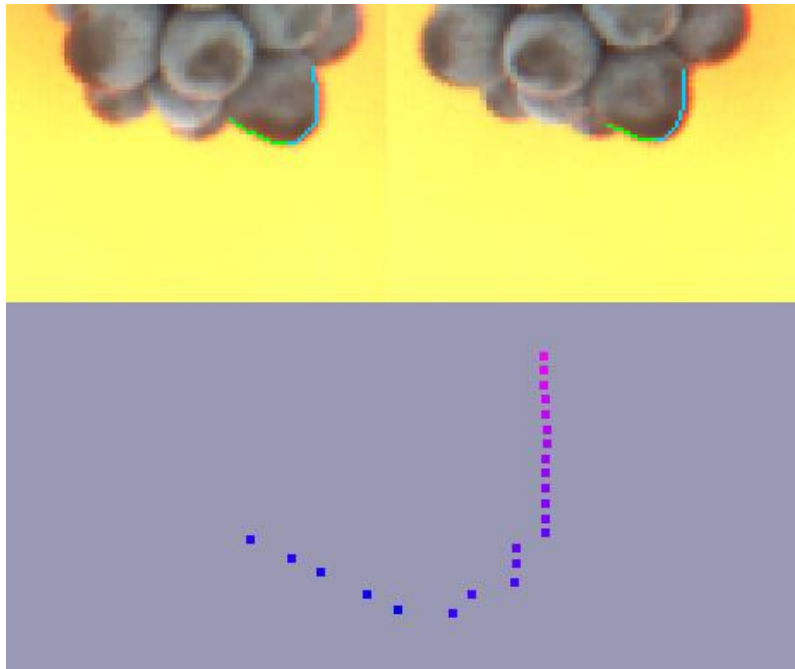


Figura 2.1-13 Exemple de reconstrucció del contorn parcial d'un gra inferior.

Sols tenim una porció del contorn, però és informació suficient per a poder reconstruir la hipòtesi del gra. Per a portar-ho a terme hem utilitzat un mètode d'optimització⁵ dels paràmetres de l'esfera que intentarà trobar aquella que s'ajuste millor als punts de contorn.

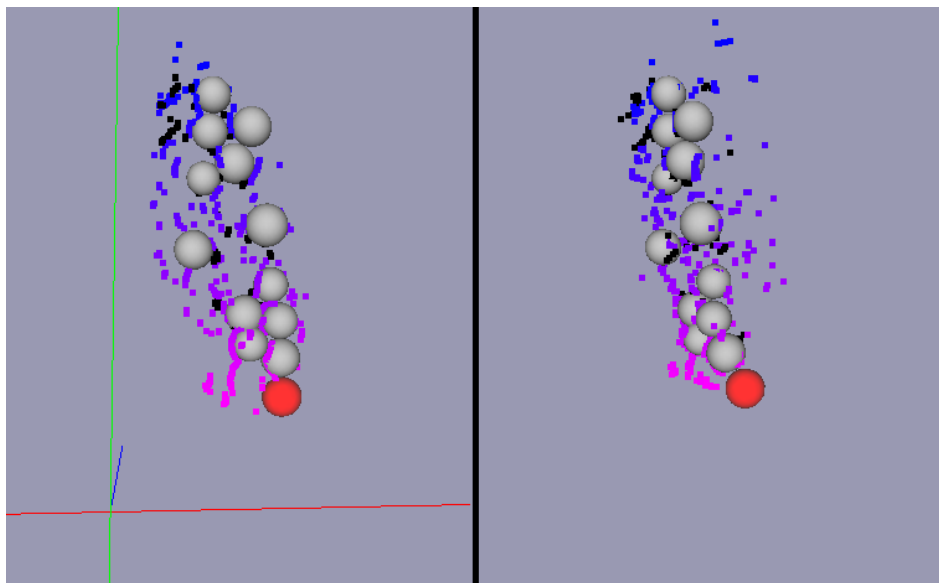


Figura 2.1-14 Gra inferior reconstruït (en roig).

Aquest mètode es pot fer servir anàlogament per al gra superior del raïm.

⁵ No confondre amb l'optimització respecte a vores de la secció 2.3, en la qual optimitzem hipòtesis ja existents en compte de crear-ne de noves.

2.2 Edició manual

Com hem pogut observar amb els exemples i com s'observarà a la secció 4, les tècniques utilitzades no permeten la reconstrucció de tots els grans visibles. Per això s'ha implementat una part d'edició manual del model que permet a l'usuari generar noves hipòtesis i editar les ja existents de forma ràpida i intuïtiva.

En aquesta secció ens centrarem en les diferents funcionalitats que la ferramenta ofereix per a generar i editar hipòtesis. El manual d'usuari que s'adjunta com a annex explica l'ús de la ferramenta en general.

2.2.1 Visualització i edició d'hipòtesis

La ferramenta té un visor de l'espai 3D on es veu representat el model, és a dir, les hipòtesis generades fins al moment.

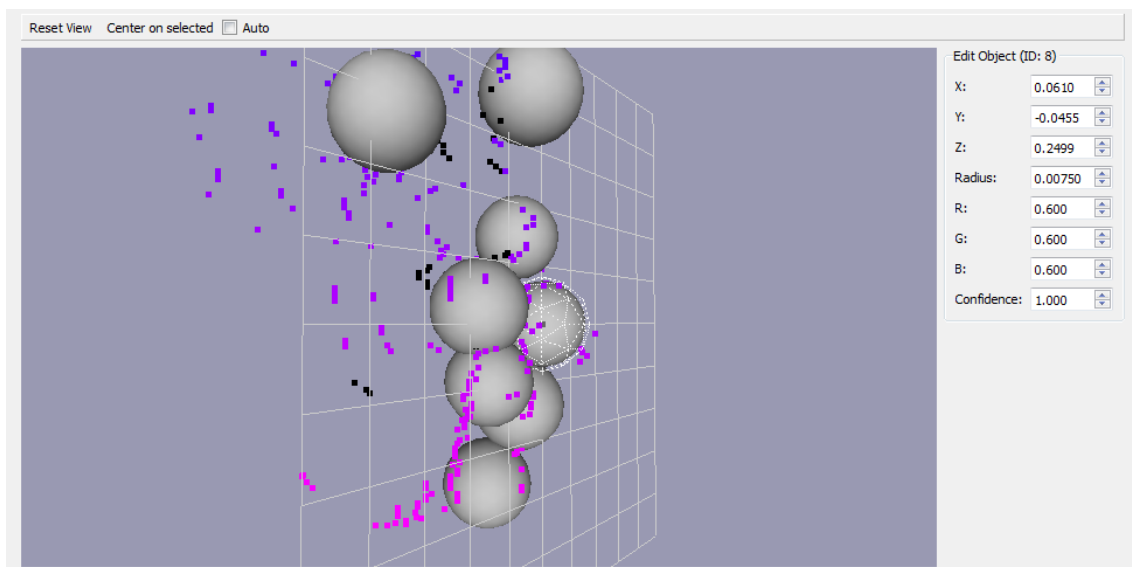


Figura 2.2-1 Visor 3D amb edició de paràmetres.

El visor ens permet, a més de visualitzar el model, navegar per l'espai 3D i seleccionar qualsevol de les hipòtesis existents.

La part interessant del procés, però, és l'edició dels paràmetres. Com podem observar a la part dreta de la Figura 3.3-1, l'usuari disposa d'uns controls amb els quals pot modificar els paràmetres de la hipòtesi seleccionada. Aquests paràmetres són la posició i el radi de l'esfera, el color amb el qual es mostra la hipòtesi i el paràmetre de confiança (que s'explicarà en seccions posteriors).

2.2.2 Generació d'hipòtesis

És possible afegir manualment noves hipòtesis de dues maneres diferents. Per una part, l'usuari té la possibilitat d'afegir una nova hipòtesi a partir dels seus paràmetres geomètrics (posició i radi). Aquesta opció requereix, però, que l'usuari conega aquests paràmetres *a priori*.

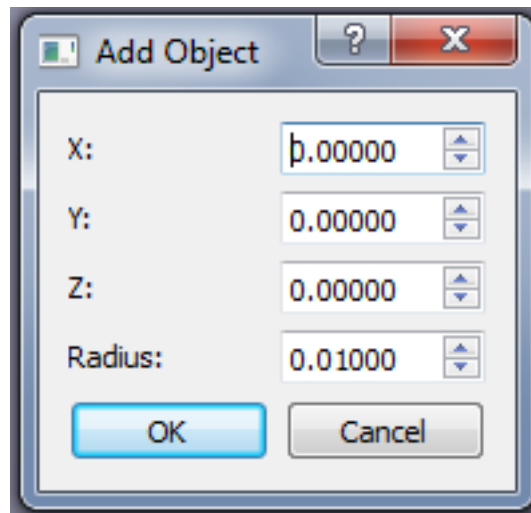


Figura 2.2-2 Diàleg de creació d'una nova hipòtesi a partir dels seus paràmetres.

Per això s'ha desenvolupat una funcionalitat que permet crear una hipòtesi corresponent a un gra de manera intuïtiva.

Aquesta funcionalitat fa ús dels visors 2D, els quals corresponen a les dues càmeres. Amb les fotografies originals com a referència, l'usuari pot definir visualment les hipòtesis en 2D sobre les imatges i la ferramenta s'encarrega de fer la reconstrucció de la hipòtesi 3D.

Per a poder definir una hipòtesi en 2D, l'usuari sols ha d'introduir 5 punts corresponents al gra en una imatge i el centre del gra a l'altra imatge. Per tant, amb sols 6 punts introduïts visualment per l'usuari es pot definir la hipòtesi. A més, la ferramenta proporciona ajudes i restriccions visuals per a facilitar la tasca. Expliquem el procés a continuació:

1. Definim 4 punts del contorn del gra en una imatge.

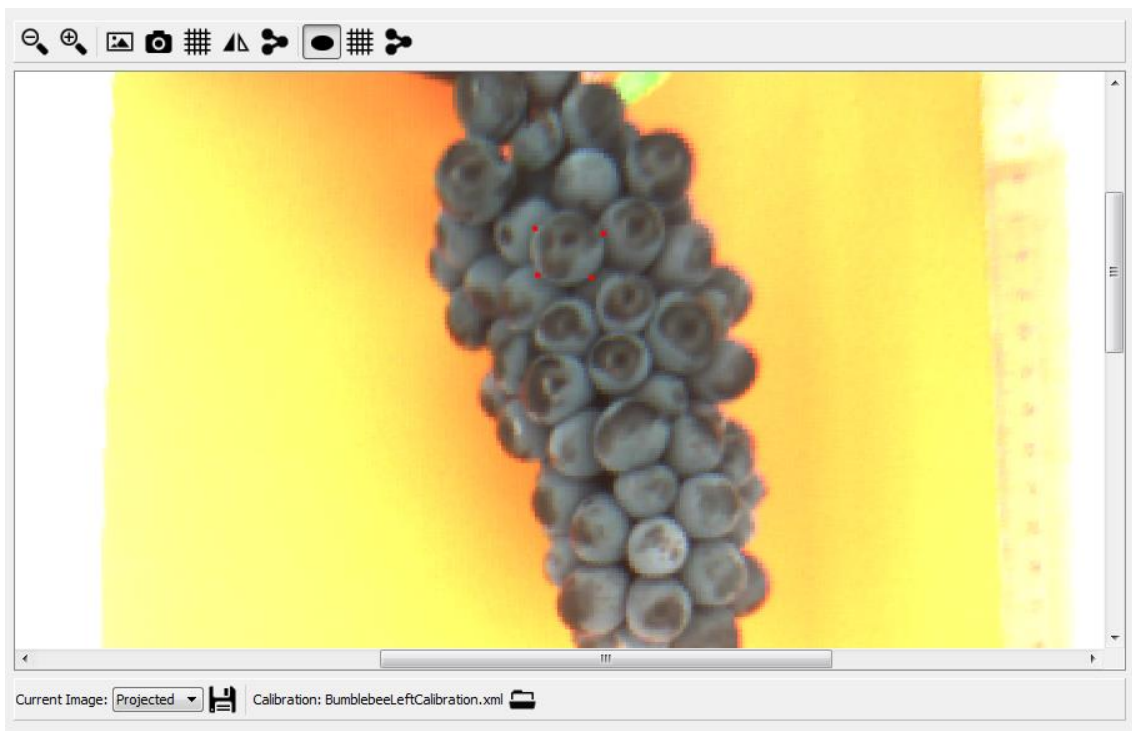


Figura 2.2-3 Quatre dels cinc punts necessaris.

2. Per al cinqué punt disposem d'una ajuda visual que ens indica quin serà el resultat final. D'aquesta manera podem ajustar l'el·lipse al gra de manera precisa.

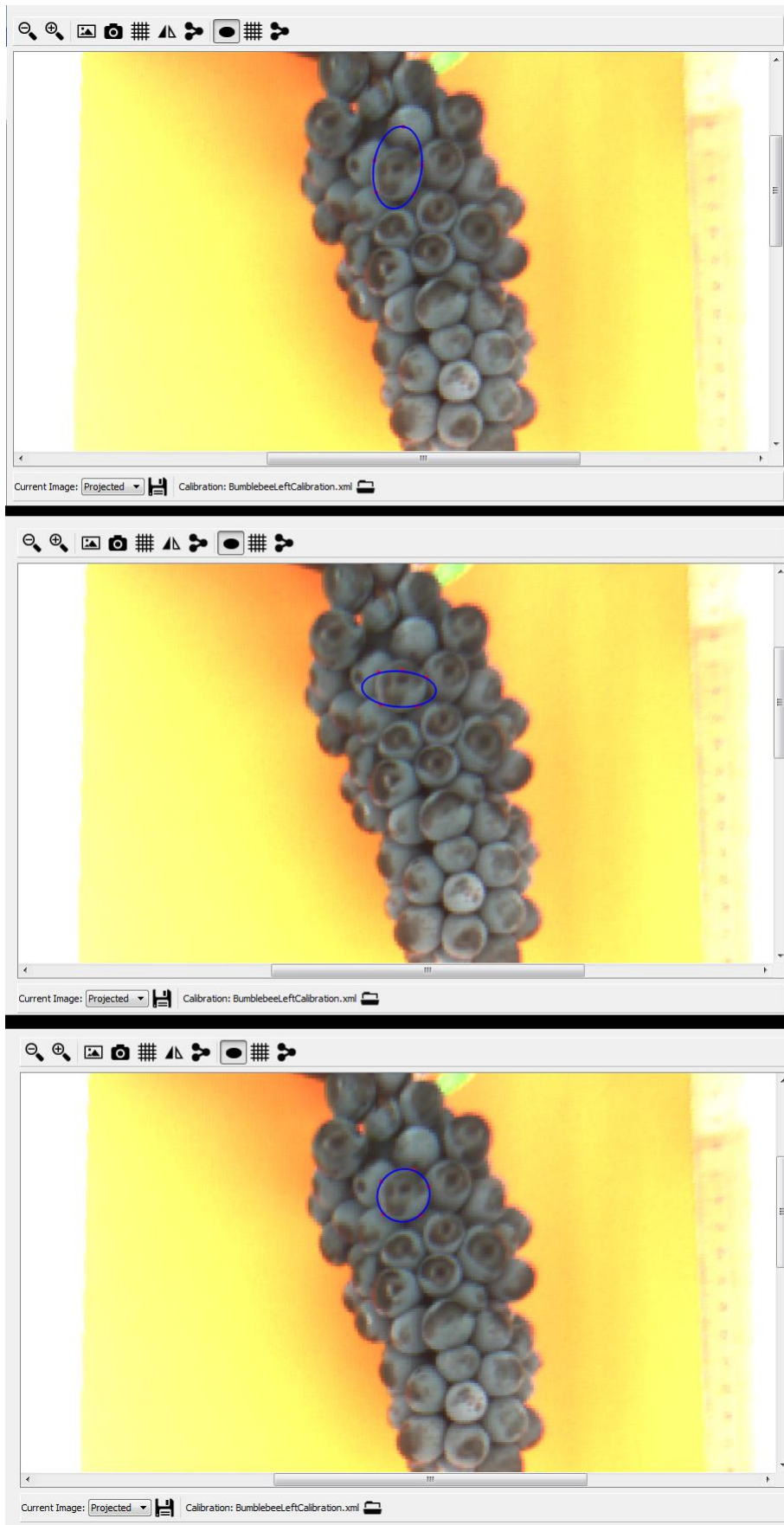


Figura 2.2-4 Per a ajustar l'últim punt es mostra una previsualització del resultat final.

- Finalment especifiquem el centre del gra en l'altra imatge. Per a fer-ho, sols cal especificar el punt sobre una línia. A més es mostra una previsualització del resultat final.

D'aquesta manera l'usuari defineix una hipòtesi ràpidament i de manera que

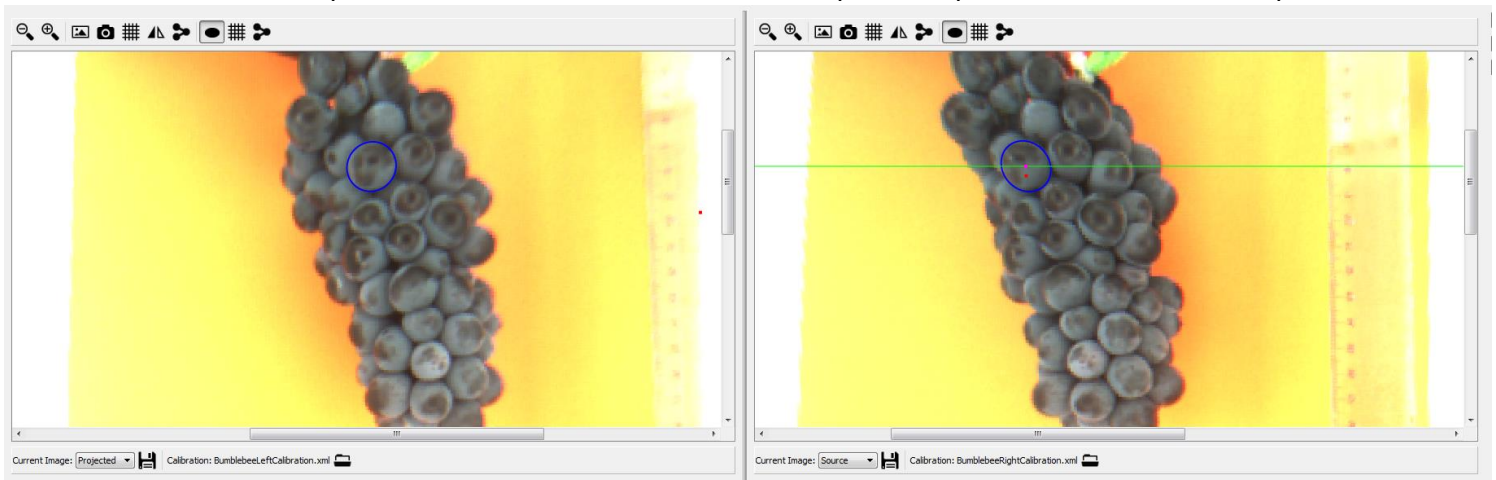


Figura 2.2-5 Ajust del centre del gra a l'altra imatge.

s'ajusta al gra original. En acabar aquest procés la ferramenta crea automàticament la hipòtesi a l'espai 3D.

2.2.3 Projecció d'hipòtesis

La ferramenta té la funcionalitat de mostrar la projecció de les hipòtesis sobre les imatges. Açò permet a l'usuari comprovar si la hipòtesi s'ajusta correctament al gra.

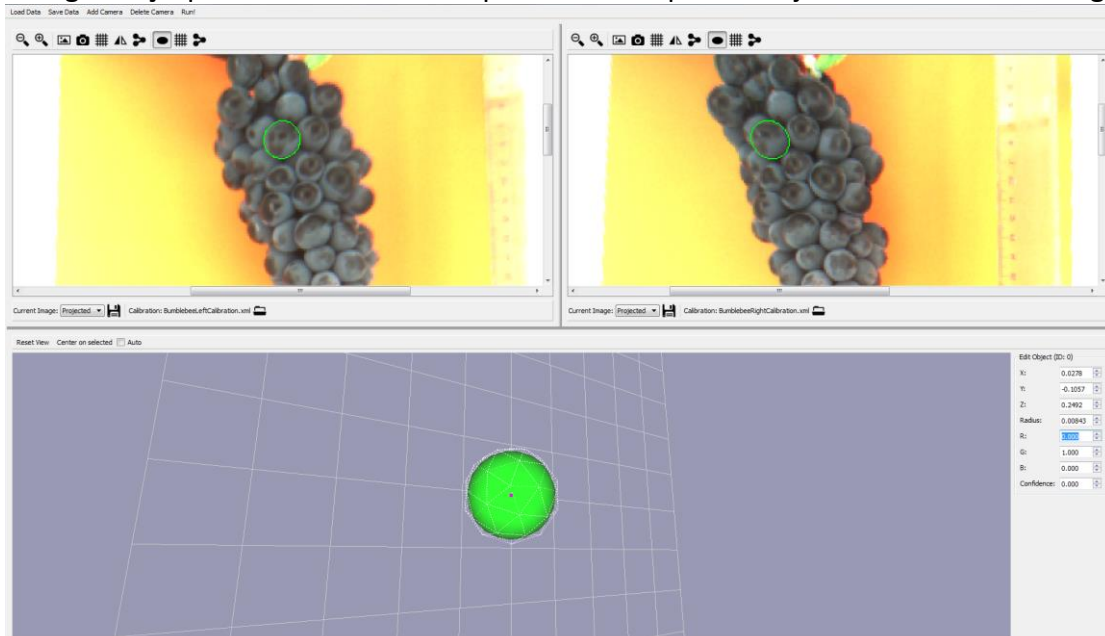


Figura 2.2-6 Hipòtesi 3D resultant de la generació manual i la seua projecció a les dos imatges.

2.3 Refinament automàtic d'hipòtesis

2.3.1 Optimització d'una sola hipòtesi

Cap la possibilitat que les hipòtesis generades tant pel mètode automàtic com per el mètode manual necessiten ser refinades perquè no s'ajusten al gra. Per a poder redreçar aquest problema s'ha desenvolupat un mètode automàtic d'optimització d'hipòtesis que intentarà ajustar correctament aquelles hipòtesis a les vores dels grans originals.

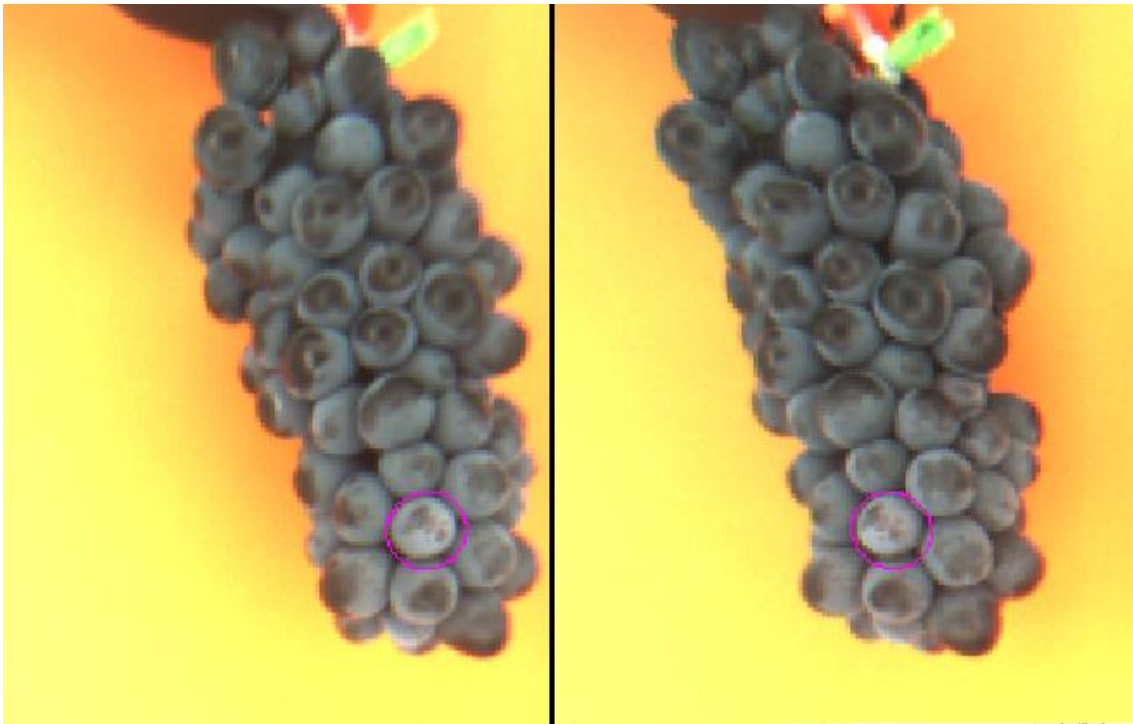


Figura 2.3-1 Exemple d'hipòtesi que no s'ajusta correctament al gra.

Aquest mètode consisteix a tractar els paràmetres de la hipòtesi (centre i radi) com a variables a optimitzar en un problema d'optimització on la funció objectiu a minimitzar és la distància de les projeccions de la hipòtesi a les vores detectades en la imatge.

Primerament, cal obtenir la imatge de vores mitjançant un algoritme de detecció de vores. En aquest cas utilitzem *Canny*.

Després, per a poder calcular fàcilment aquesta distància generem una imatge de distàncies. Aquesta imatge és una matriu on cada píxel te codificada la distància en píxels a la vora més propera.

D'aquesta manera el càlcul de la funció objectiu es redueix a:

1. Projectar el gra en una càmera, per la qual cosa obtenim una el·lipse com a resultat.
2. Obtindre N punts del contorn de l'el·lipse.
3. Buscar per a cada punt el píxel associat en la imatge de distàncies corresponent i obtenir el seu valor (equivalent a la distància a la vora més pròxima).

4. Fer la mitja de tots els punts.
5. Repetir el procés per a l'altra càmera.
6. Quedar-nos amb el valor més alt de les dos càmeres.

Una vegada resolt aquest problema⁶ de minimització de costos obtenim l'esfera, les projeccions de la qual més s'ajusten a les vores detectades. Per tant, si apliquem aquest mètode a una hipòtesi que no acaba de ser acurada, la seua posició serà refinada de manera que s'ajuste millor al gra original.

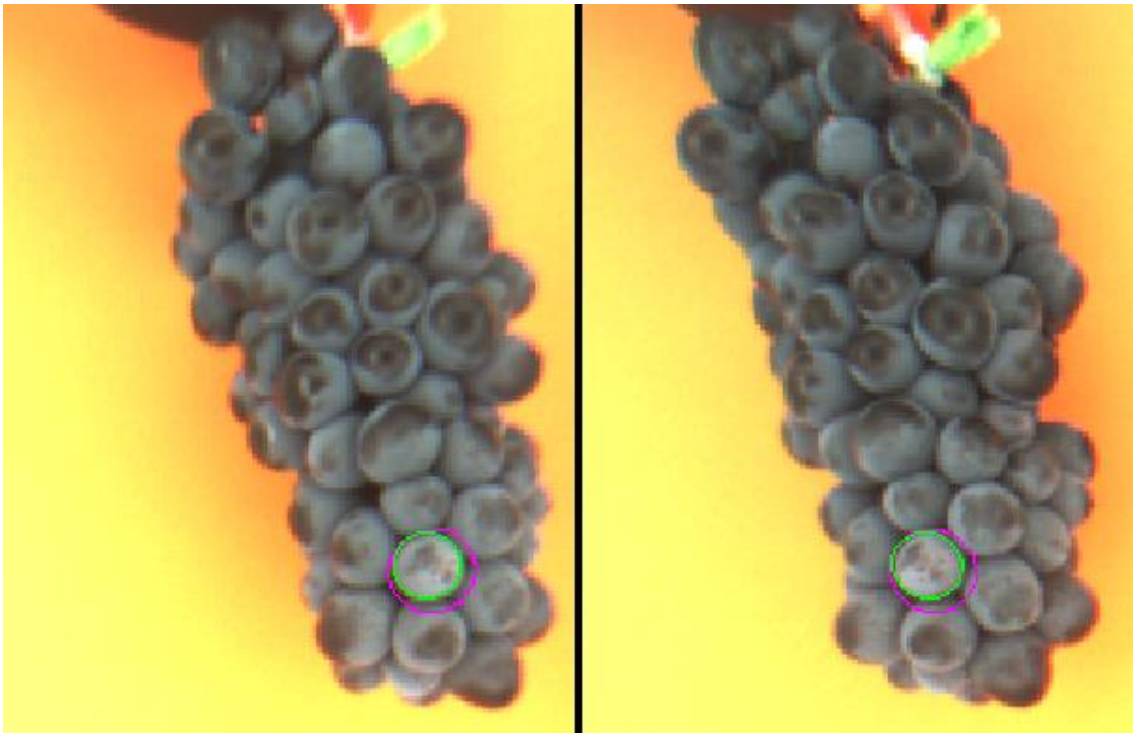


Figura 2.3-2 Comparació entre la hipòtesi sense optimitzar (color rosa) i la hipòtesi optimitzada (color verd). Podem observar que s'ajusta millor al gra.

⁶ Utilitzant la llibreria IPOPT. Veure la secció 3 per a més detalls.

2.3.2 Optimització d'un conjunt d'hipòtesis

Com hem comentat a la secció 1.4.1, és molt probable que alguns grans estiguen parcialment ocults per altres grans. És impossible optimitzar les hipòtesis corresponents a aquests grans de manera aïllada ja que part de les vores d'aquests grans estaran ocultes per altres grans més propers a la càmera. Això suposa un problema perquè estarem intentant ajustar projeccions les vores de les quals no disposem.

La solució consisteix a optimitzar les hipòtesis per ordre, de més properes a la càmera a més llunyanes. Això ens permet projectar sols les parts del gra que no estiguen ocultes per altres hipòtesis.

Malauradament, no coneixem l'ordre real de les hipòtesis abans d'optimitzar-les. Per tant, cal trobar l'ordre correcte a través d'un arbre de cerca en el qual anem optimitzant les hipòtesis una rere l'altra. Si en optimitzar obtenim un resultat no satisfactori⁷ significarà que l'ordre no és correcte i procedim a provar un ordre diferent.

Ens quedarem amb la solució òptima, és a dir, amb aquell ordre d'hipòtesis que sumen una menor distància a les vores. D'aquesta manera les hipòtesis s'hauran ordenat correctament i a més hauran sigut optimitzades de manera adient.

És possible que en algun punt de la reconstrucció automàtica s'intente optimitzar alguna hipòtesi que ja ha sigut optimitzada, però aquesta volta en condicions diferents⁸. Això pot portar que el resultat de l'optimització siga en realitat pitjor que l'anterior. En aquest cas cal que aquest fet quede patent i s'actue en conseqüència.

Per això calculem un paràmetre de confiança en acabar l'optimització, que indica el grau de fidelitat de la hipòtesi. Aquest paràmetre es calcula a partir de distàncies a vores i quantitat de vores visibles. Si el paràmetre de confiança és menor que el que tenia anteriorment, descartarem l'optimització. D'aquesta manera ens assegurem que la qualitat de la hipòtesi no empitjore.

⁷ Un resultat no satisfactori serà aquell en el que la projecció de la hipòtesi optimitzada segueix estant lluny de les vores detectades al gra.

⁸ Per exemple, pot ser que una hipòtesi nova haja sigut afegida, en eixe cas el resultat de l'optimització pot canviar.

3 Disseny i implementació

3.1 Tecnologia i ferramentes utilitzades

Durant el desenvolupament del projecte s'han utilitzat diverses llibreries i ferramentes, tant per a ajudar en la investigació com per a realitzar la implementació final.

- C++ com a llenguatge de programació ja que en el context de la solució del problema el temps, tot i no ser crític, és important. C++ ofereix la flexibilitat de l'orientació a objectes i la potència d'un llenguatge compilat de relatiu baix nivell.
- Visual Studio com a entorn de desenvolupament (IDE) i compilador de C++.
- OpenCV com a llibreria matemàtica i de tractament d'imatge. Aquesta llibreria ha sigut utilitzada extensivament en la ferramenta, sobretot per a operacions matricials i per a l'aplicació d'algorismes de processament d'imatge.
- IPOPT per a la resolució d'alguns dels problemes d'optimització.
- QT per al desenvolupament de la interfície gràfica d'usuari de la ferramenta.
- OpenGL per a la visualització 3D.
- Repositori SVN per al control de versions.

3.2 Estructura interna

La part interna de la ferramenta és la que realitza totes les operacions i emmagatzema la informació necessària (en contraposició a la part externa o interfície de la qual parlem en la secció 3.3). Aquesta part està organitzada de manera que la informació siga fàcilment accessible des dels diferents punts d'execució del programa. A continuació mostrem un diagrama de classes que mostra com es relacionen els diferents elements.

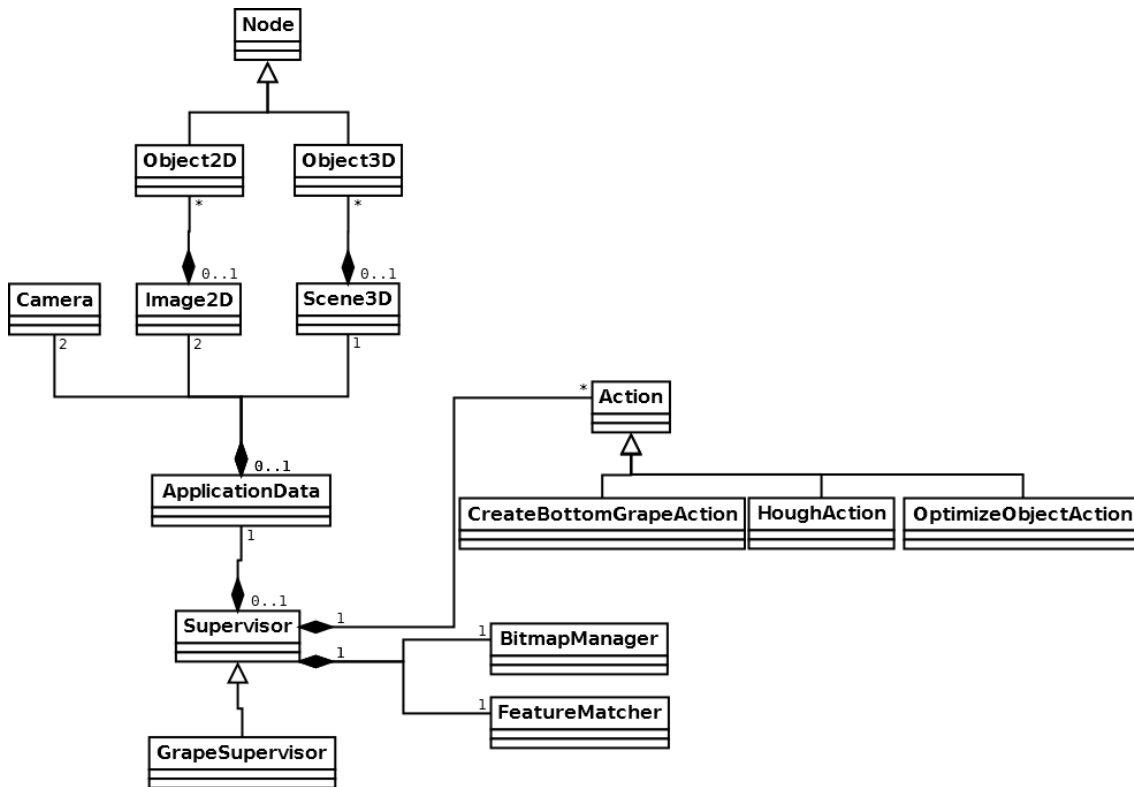


Figura 3.2-1 Diagrama de classes corresponent a la part interna de la ferramenta.

3.2.1 ApplicationData

Una de les classes més importants és *ApplicationData*. Aquesta classe encapsula les dades necessàries per a representar el problema de la reconstrucció d'hipòtesis i és, generalment, la unitat d'informació que es transmet entre diferents parts de la ferramenta. Podem dir que *ApplicationData* conté el model amb el qual treballa la ferramenta i el supervisor.

Com es pot observar al diagrama, aquesta classe conté dues instàncies d'*Image2D*⁹ i una *Scene3D*. Ambdues funcionen com a arbres formats per nodes. En el cas d'una *Image2D* els nodes són *Object2D* i en el cas d'una *Scene3D* els nodes són *Object3D*.

En el problema que ens ocupa¹⁰, un *Object2D* representa una el·lipse i *Object3D* una esfera. Un *Object2D* és, per tant, la projecció sobre una imatge d'un *Object3D*.

⁹ No confondre amb una imatge en el sentit tradicional de la paraula.

Conceptualment hem d'entendre un *Object3D* com el que hem anomenat hipòtesi al llarg del document. Una *Scene3D* és llavors el conjunt de totes les hipòtesis 3D, i les dues *Image2D* són el conjunt de projeccions d'aquestes hipòtesis en la càmera esquerra i la càmera dreta.

Per tant, quan al document diem de generar o optimitzar una hipòtesi, a nivell d'implementació estem parlant de generar o optimitzar un *Object3D*. De la mateixa manera, quan parlem de projectar una hipòtesi sobre una imatge, estem en realitat generant un *Object2D* a partir d'un *Object3D* i afegir-lo a una *Image2D*.

ApplicationData conté també dues instàncies de *Camera*, que representen els calibratges de la càmera esquerra i dreta. Cal destacar que s'estableix una associació implícita entre *Image2D* i *Camera* atés que els calibratges de la primera *Image2D* es troben a la primera *Camera* i així successivament.

¹⁰ Tal i com s'explica a la documentació adjunta, en realitat un *Object2D* és un conjunt d'el·lipses i un *Object3D* és un conjunt d'esferes. Això ocorre perquè s'ha volgut donar suport parcial a poliesferes com a objectes a optimitzar, en vistes a altres hipotètics projectes derivats del que ens ocupa. El lector pot assumir, però, que un *Object2D* i un *Object3D* fan referència directament a una el·lipse i a una esfera, respectivament.

3.2.2 Supervisor i accions

Tal com s'ha explicat en la secció 1.2 la ferramenta ha sigut dissenyada per a ser reutilitzada més enllà de l'abast d'aquest projecte.

S'han tingut en compte dues possibles necessitats. Per una part, el disseny respon a la possibilitat d'afegir noves tècniques de reconstrucció automàtica per a millorar els resultats en aquest apartat. Per una altra part, a la possibilitat de crear processos automàtics que resolguen problemes diferents però similars al quals ens ocupa (problemes de reconstrucció d'hipòtesis).

Per a satisfer aquestes necessitats s'ha creat la classe *Supervisor*, la qual serveix com a base a partir de la qual crear supervisors automàtics que guien el procés de reconstrucció d'hipòtesis. Podem veure representada aquesta classe i les relacionades a la Figura 3.2-1 Diagrama de classes corresponent a la part interna de la ferramenta.

Aquesta classe implementa un sistema que permet al programador indicar al supervisor quines accions cal executar i amb quina prioritat. Aquestes accions, a més, poden generar-ne de noves, les quals s'afegiran a la cua d'accions pendents ordenades per prioritat. En implementar el mètode abstracte *onStart()*, el programador pot establir afegir les accions necessàries amb la seua prioritat corresponent utilitzant el mètode *addAction()*.

Conseqüentment, també és proporciona la classe *Action* com a base a partir de la qual definir les accions a realitzar pel supervisor. El cos de l'acció s'implementarà al mètode *run()*.

La pròpia ferramenta que es presenta inclou un supervisor (anomenat *GrapeSupervisor*) que s'encarrega de portar a terme les tècniques explicades a la secció 2, implementades com a accions. Concretament les accions executades són les següents:

- *CreateBottomGrapeAction*: Creació de la hipòtesi corresponent al gra inferior.
 - *OptimizeObjectAction*: Optimització de la hipòtesi.
- *HoughAction*: Generació de les hipòtesis a partir de Hough.
 - *OptimizeObjectAction*: Optimització de les hipòtesis.

Com podem observar, aquest sistema adreça les dos necessitats comentades prèviament. D'una banda, el supervisor utilitzat al present projecte, *GrapeSupervisor*, pot ampliar-se fàcilment a través de definir noves accions i afegir-les al supervisor.

D'altra banda, és possible crear un nou supervisor derivat de la classe *Supervisor* amb el seu propi conjunt d'accions i prioritats per a resoldre problemes de reconstrucció automàtica.

A la documentació adjuntada com a annex en aquest document es poden trobar més detalls sobre la implementació d'aquest sistema i de com estendre'l.

3.3 Interfície gràfica

La interfície gràfica juga un paper fonamental en la utilització de la ferramenta a causa de diversos factors:

- *Feedback* de l'estat del model: Donada la naturalesa del problema és molt més senzill per a l'usuari interpretar la informació de manera visual.
- Interacció amb el model: De la mateixa manera, manipular el model mitjançant una interfície gràfica és més intuïtiu i ràpid que numèricament.
- Facilitat d'ús: La interfície permet a l'usuari carregar i guardar diferents recursos com imatges originals, calibratge, imatges comparatives i dades sobre el model en qualsevol moment mentre manipula el model, tot integrat en una sola ferramenta.

Donada la importància de la visualització del model s'ha optat per dissenyar la interfície de manera simple, donant molta importància a les diferents representacions del model, minimitzant l'espai que ocupen la resta d'elements de la interfície.

Com hem vist a la secció 3.2, el model consta del conjunt d'hipòtesis 3D (*Scene3D*) i les seues projeccions a les càmeres esquerra i dreta (*Image2D*). Per a visualitzar les hipòtesis 3D l'usuari disposa del *Viewport3D* situat a la part inferior de l'aplicació. Al *Viewport3D* l'usuari pot navegar per l'espai 3D, visualitzar les hipòtesis i operar sobre elles.

En canvi, per a treballar en el context 2D del model (projeccions), la ferramenta presenta a la part superior dos *Viewport2D*. Cada *Viewport2D* correspon a una càmera, i permet a l'usuari carregar el calibratge de la càmera i la fotografia original. A més, dona a l'usuari la opció de visualitzar les projeccions (*Image2D*) a través projectar les hipòtesis sobre la imatge (veure secció 2.2.3).

A més a més, en la part superior de la ferramenta l'usuari disposa d'una sèrie de controls, que permeten a l'usuari iniciar la reconstrucció automàtica (veure secció 2.1) i guardar el model, entre altres opcions.

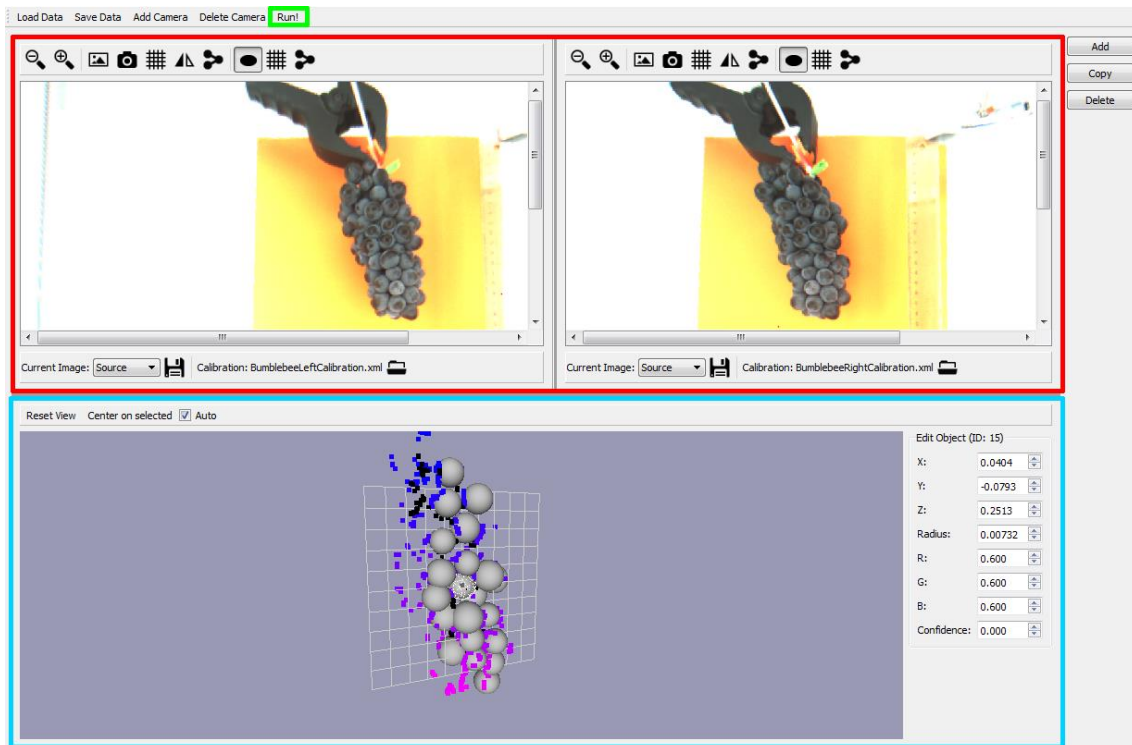


Figura 3.3-1 Dos *Viewport2D* (roig), *Viewport3D* (blau) i botó per a iniciar la reconstrucció automàtica (verd)

Com es pot observar, la interfície ha sigut dissenyada prioritzant la visualització del model i establint un *workflow* molt definit:

1. Càrrega de calibratges i imatges originals.
2. Execució de la reconstrucció automàtica.
3. Comprovació de resultats parcials mitjançant projeccions.
4. Reconstrucció manual.
5. Exportació del model a un fitxer *XML*.

La ferramenta ofereix, però, altres funcionalitats que són tangencials a aquest *workflow* establert:

- Carregar models generats amb la ferramenta amb la fi de visualitzar-los o modificar-los.
- Generar una imatge sintètica (*Render*) de les hipòtesis 3D.
- Visualitzar el resultat de la detecció de vores sobre la fotografia original.
- Comparar dues imatges (per exemple, la fotografia original i la imatge sintètica generada amb la fi de comprovar la fidelitat dels resultats).
- Generar punts en l'espai 3D a partir de punts de les imatges 2D.

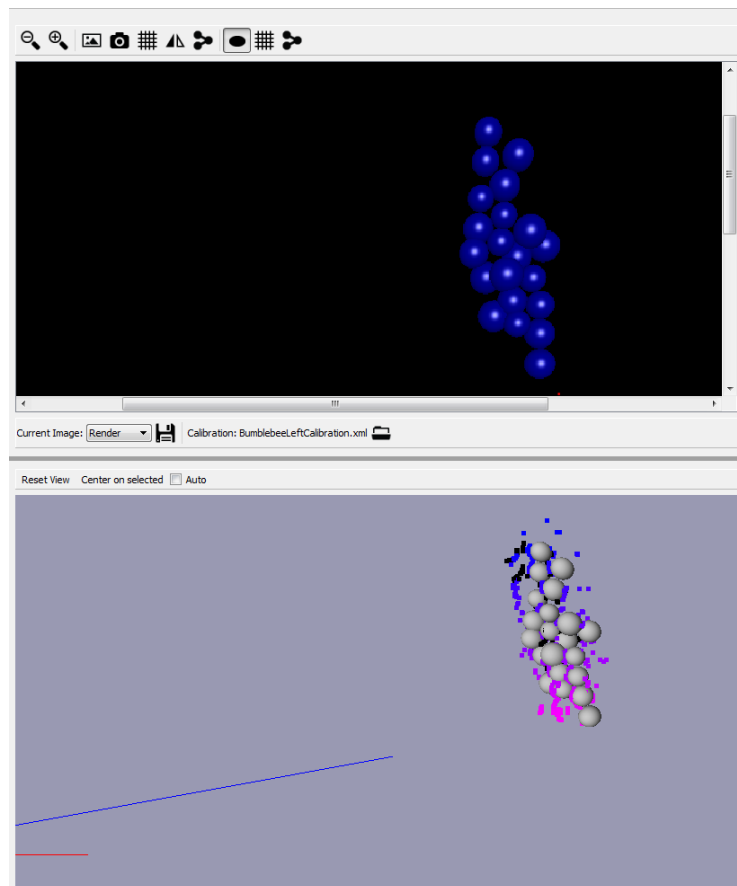


Figura 3.3-2 *Render* d'un conjunt d'hipòtesis des del punt de vista de la càmera esquerra.

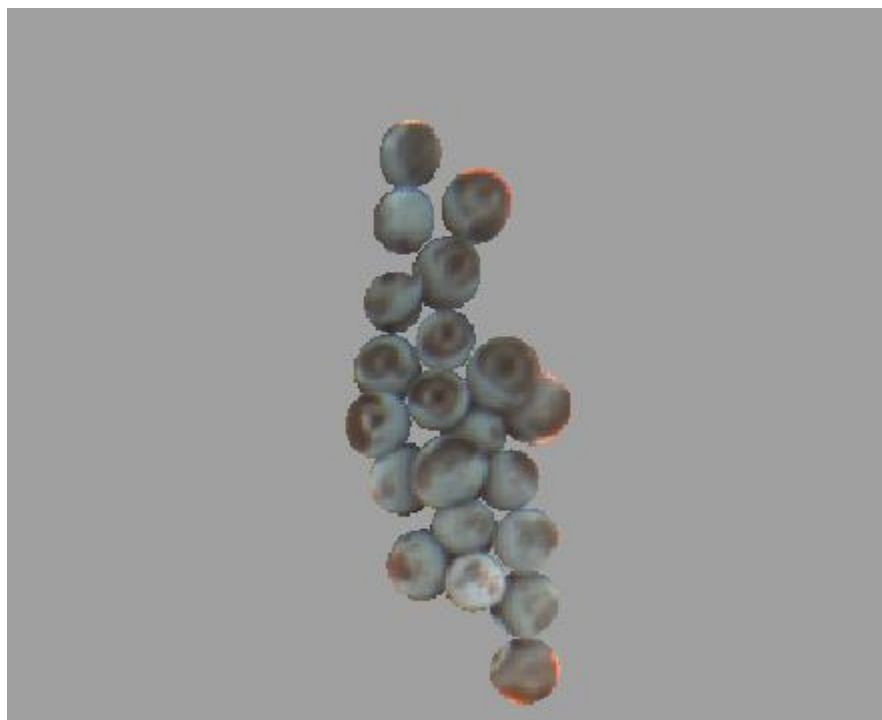


Figura 3.3-3 Comparació de la imatge sintètica (*Render*) i imatge original. Es pot fer servir com a mètode alternatiu per a comprovar com s'ajusten les hipòtesis als grans reals.

A la figura Figura 3.3-4 Diagrama de classes de la interfície. podem observar el diagrama de classes corresponent a la interfície. En ell veiem representada la jerarquia dels diferents elements de la interfície. Cal notar com la interfície manté la seua pròpia còpia del model (*ApplicationData*). Per una altra part, tal com es mostra a la Figura 3.2-1, el supervisor també té el seu propi model.

Aquesta duplicitat és intencionada i s'ha fet amb l'objectiu que el supervisor realitze totes les seues operacions sobre una còpia privada del model i aquest es retorne a l'aplicació sols quan haja acabat el procés. Quan això es dona, la interfície actualitza el seu model a partir del que rep del supervisor.

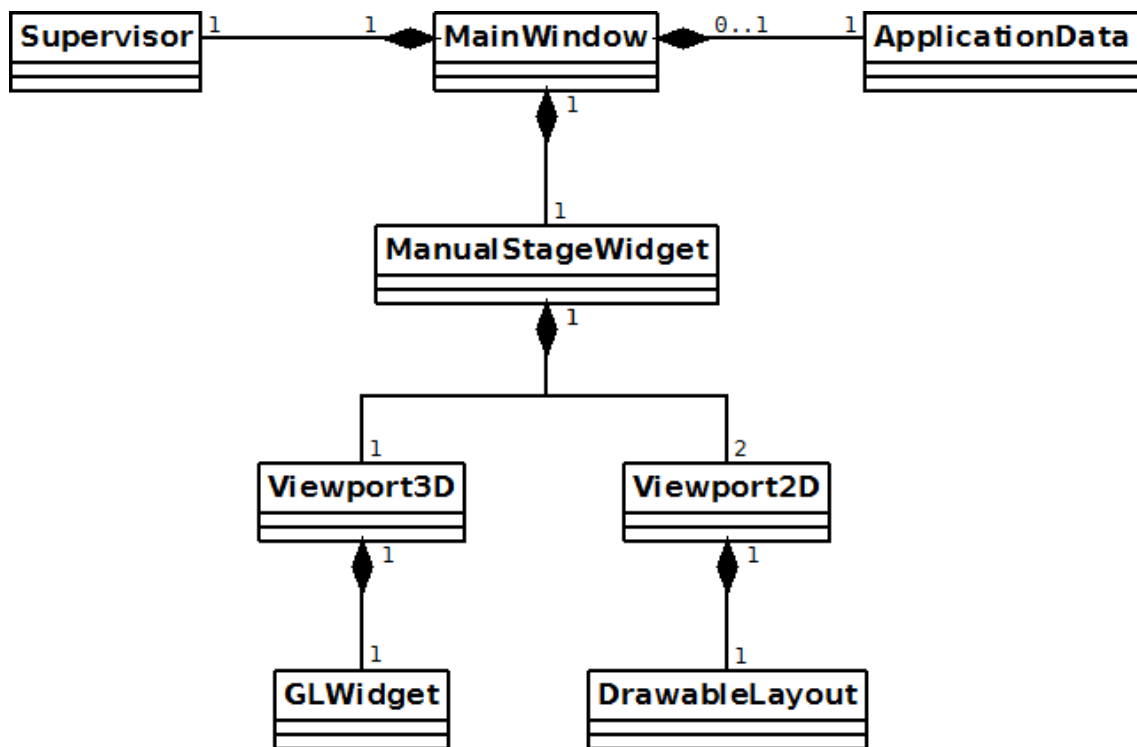


Figura 3.3-4 Diagrama de classes de la interfície.

4 Resultats

4.1 Reconstrucció manual

A continuació, mostrem diversos tipus de raïm reconstruïts manualment mitjançant la ferramenta comparats amb les imatges originals.

Cinsault

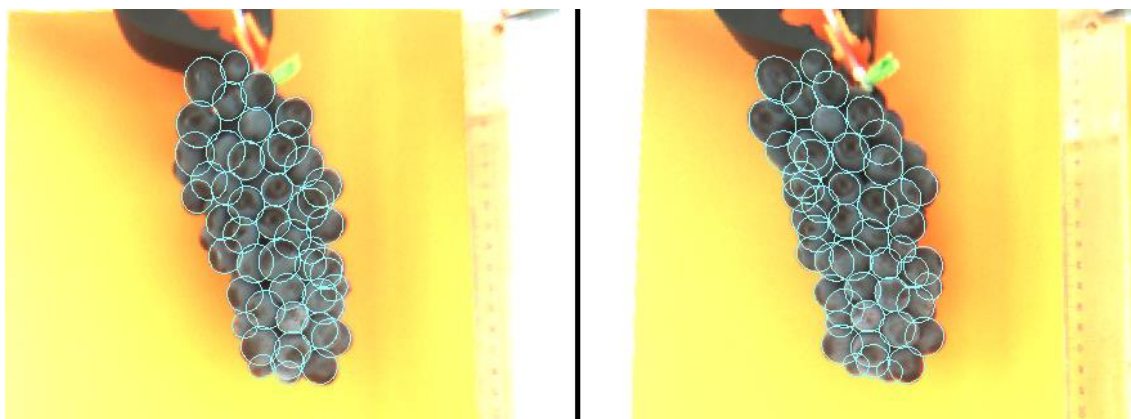


Figura 4.1-1 Projecció de les hipòtesis sobre un raïm de la varietat *Cinsault*.

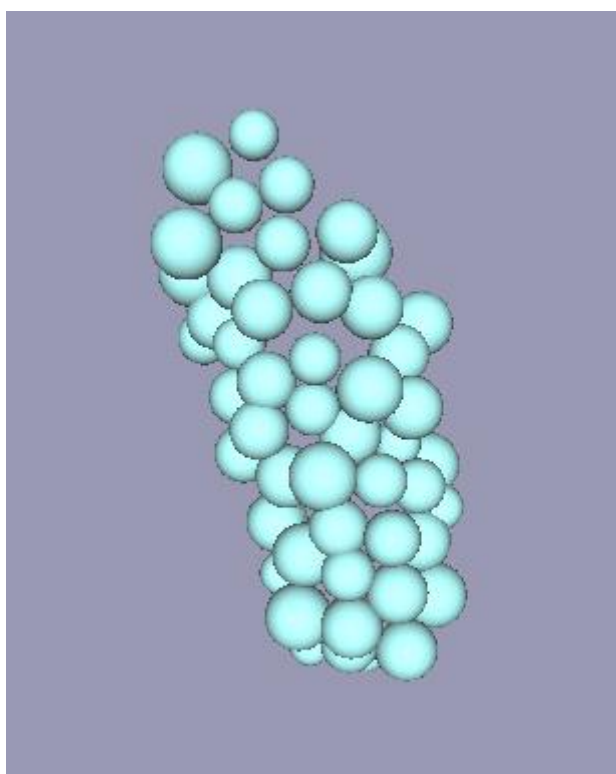


Figura 4.1-2 Reconstrucció d'un raïm de tipus *Cinsault*.

Bobal

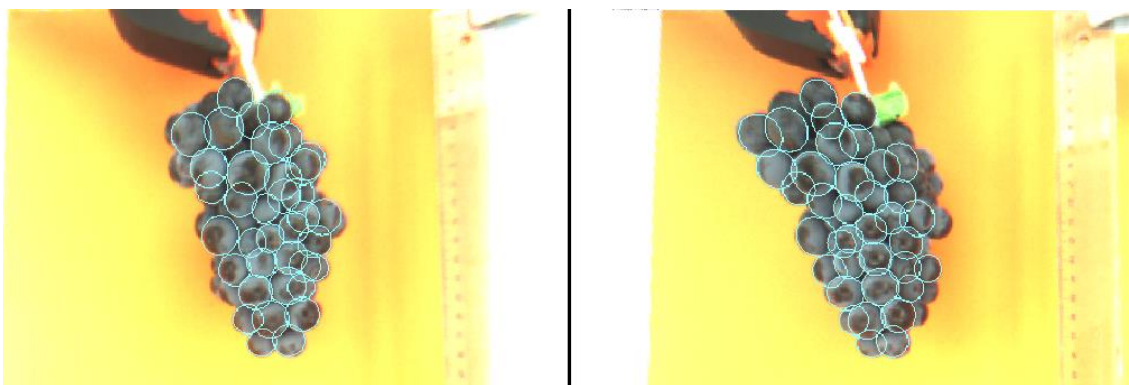


Figura 4.1-3 Projecció de les hipòtesis sobre un raïm de la varietat *Bobal*.

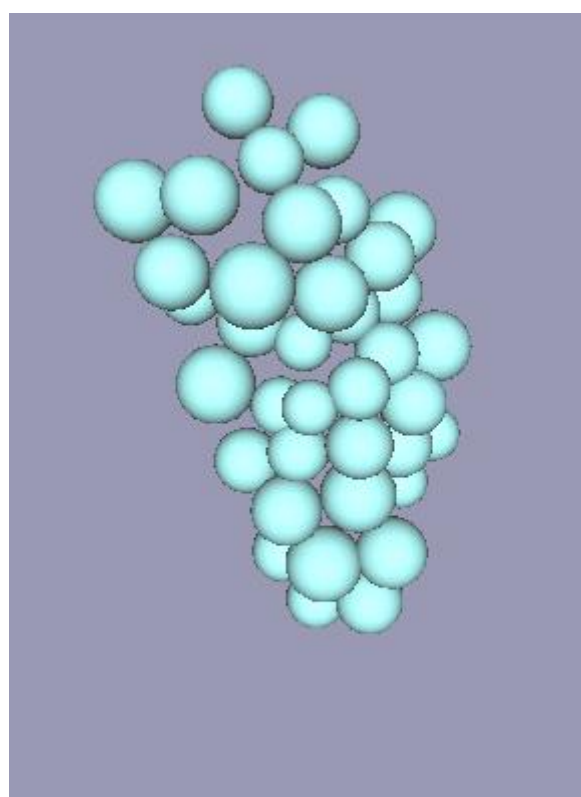


Figura 4.1-4 Reconstrucció d'un raïm de la varietat *Bobal*.

Monastrell

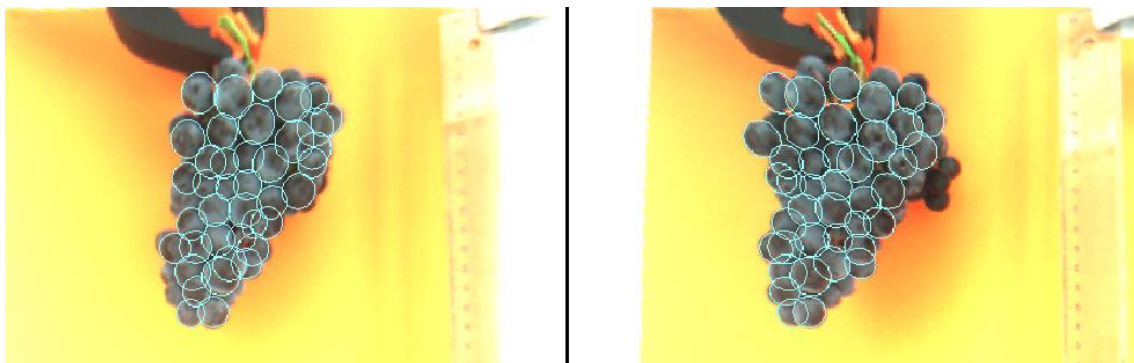


Figura 4.1-5 Projectió de les hipòtesis sobre un raïm de la varietat *Monastrell*.

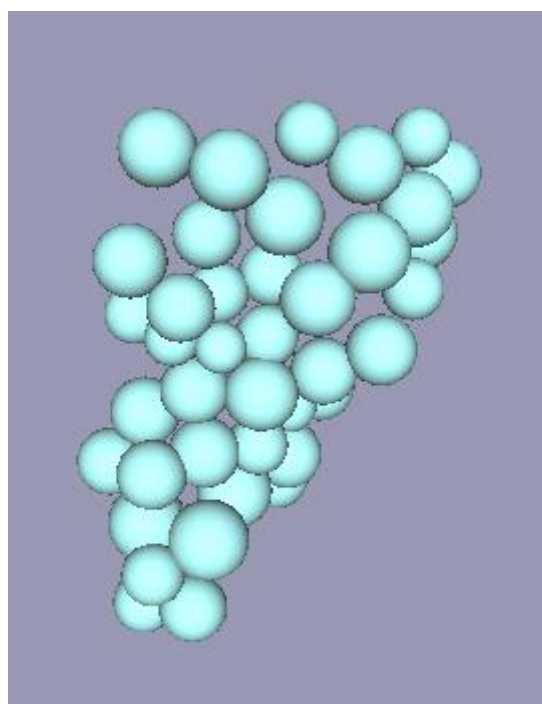


Figura 4.1-6 Reconstrucció d'un raïm de la varietat *Monastrell*.

Plantamura

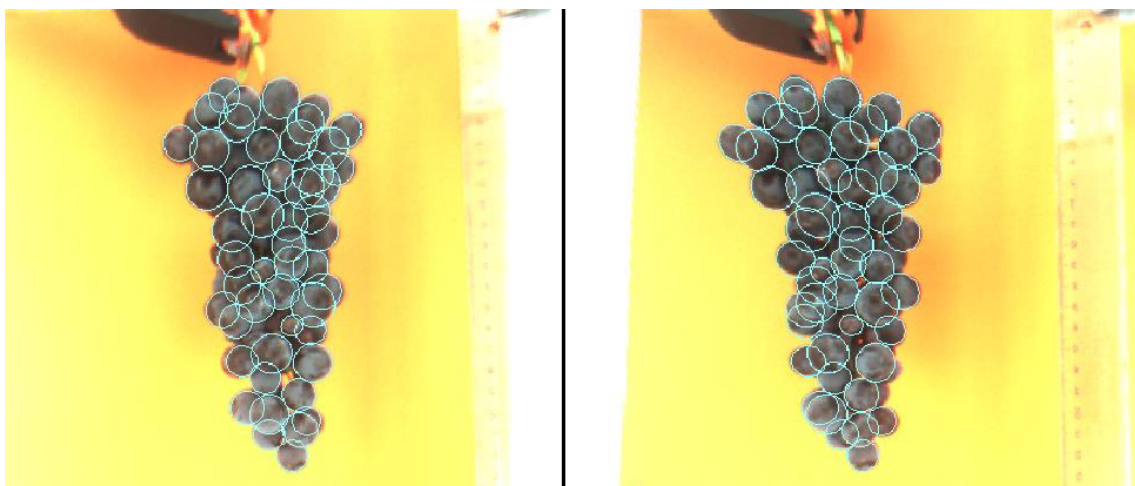


Figura 4.1-7 Projectió de les hipòtesis sobre un raïm de la varietat *Plantamura*.

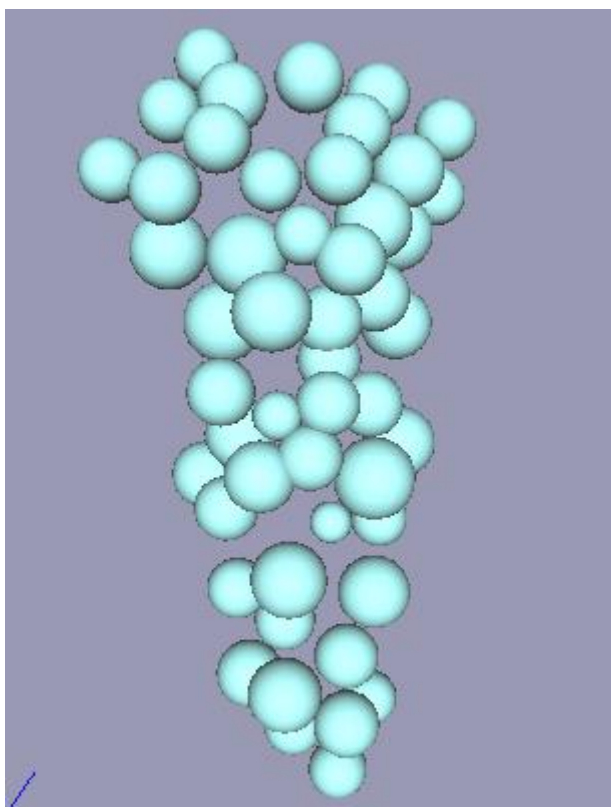


Figura 4.1-8 Reconstrucció d'un raïm de la varietat *Plantamura*.

Cal observar que sols és possible reconstruir aquelles hipòtesis que siguen visibles des de les dues càmeres. Aquest efecte es veu clarament a la Figura 4.1-5 Projecció de les hipòtesis sobre un raïm de la varietat *Monastrell.*, on des de la càmera de la dreta apareixen uns grans que no són visibles des de la càmera esquerra i, per tant, són impossibles de reconstruir per falta d'informació. Aquest problema era conegut des del principi com a una de les limitacions de l'aproximació estereoscòpica al problema i no afecta més que a un nombre limitat de grans en tan sols alguna de les imatges.

La reconstrucció manual, en definitiva, dóna com a resultat un ajustament quasi perfecte de les hipòtesis al raïm gràcies al sistema implementat que permet a l'usuari definir els grans a partir de les seues vores i centre (veure secció 2.2.2). Aquests mateixos models generats manualment seran considerats correctes i s'utilitzaran com a referència (*ground truth*) per a avaluar els resultats de la generació automàtica.

4.2 Reconstrucció automàtica

En aquesta secció avaluem els resultats de la reconstrucció automàtica.

A continuació mostrem les hipòtesis generades per a els diferents raïms i mostrem aquelles que són considerades com a correctes (color verd) i aquelles que són considerades incorrectes (roig), basades en la comparació amb el raïm generat manualment (veure secció anterior). El criteri per a decidir si una hipòtesi es correcta o no és la diferència quant a posició i radi respecte les hipòtesis del raïm generat manualment.

Cinsault

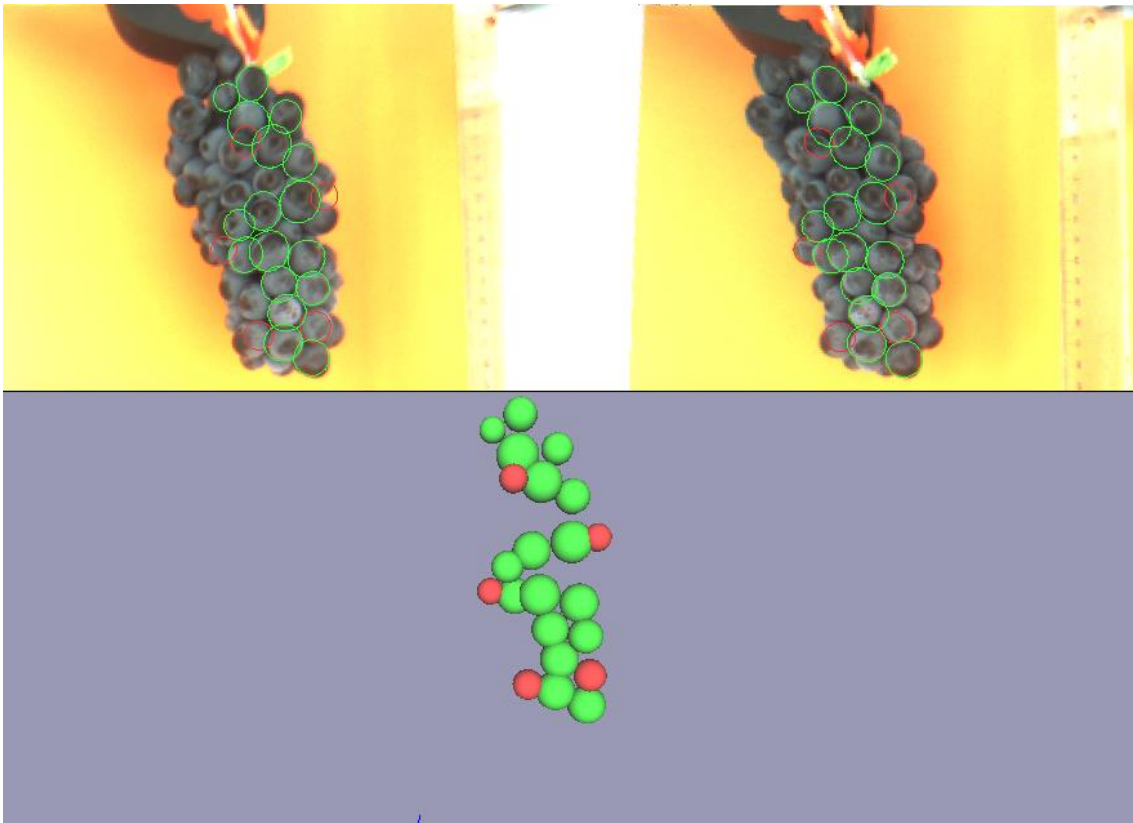


Figura 4.2-1

Bobal

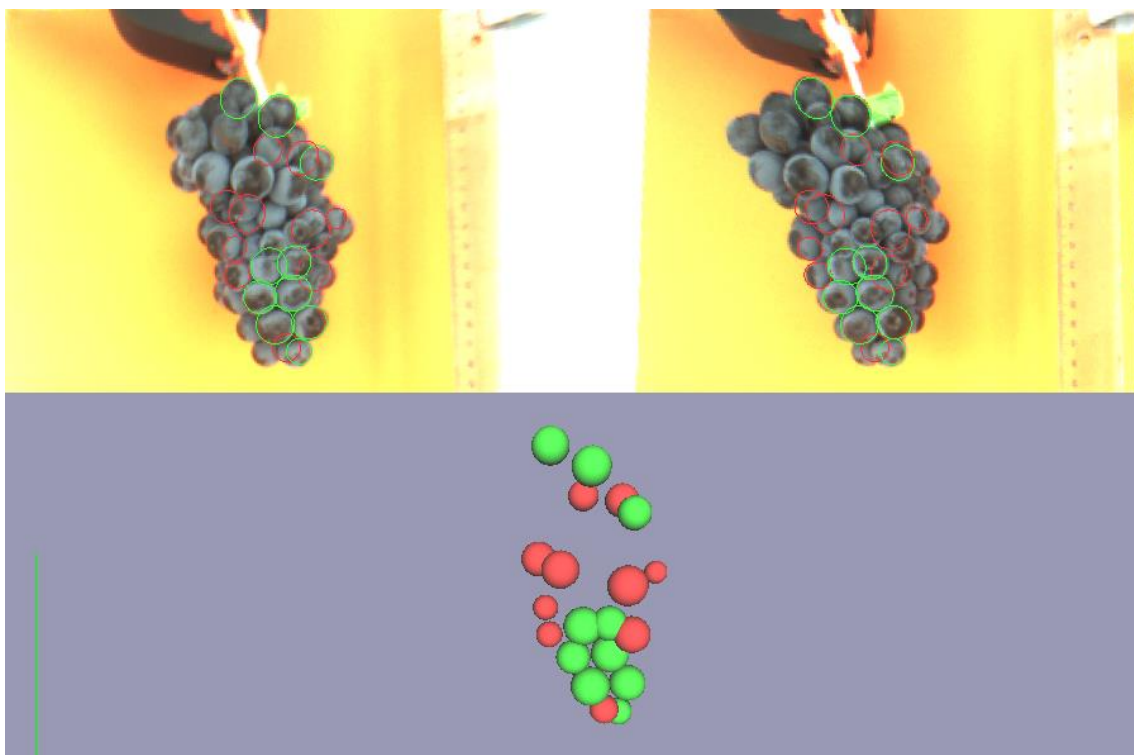


Figura 4.2-2

Monastrell

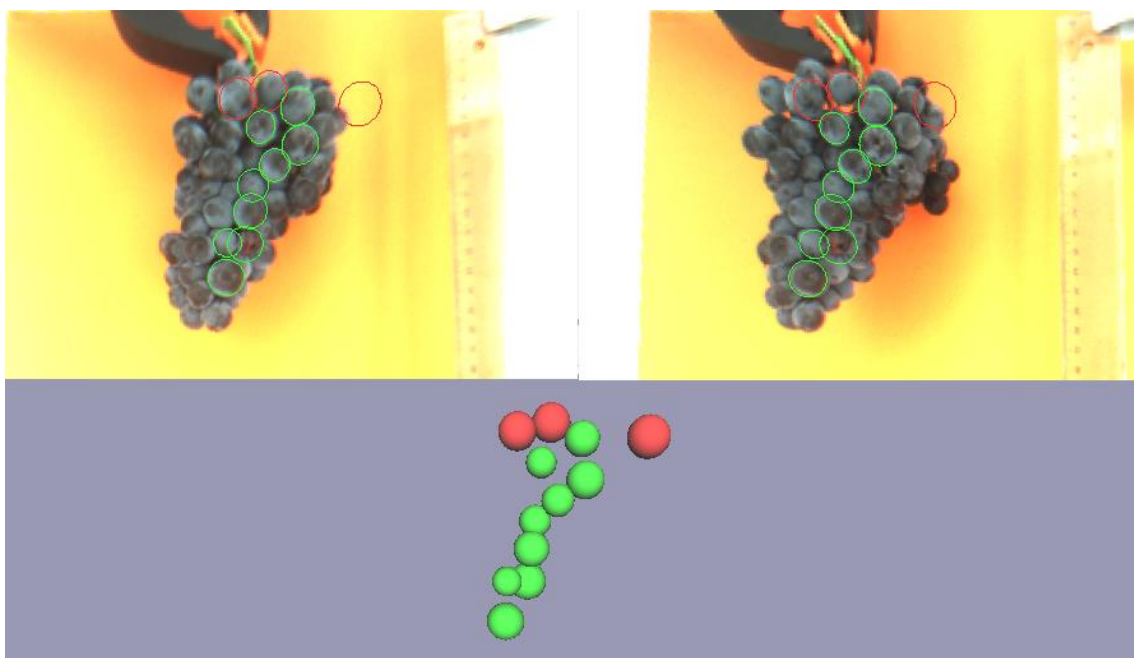
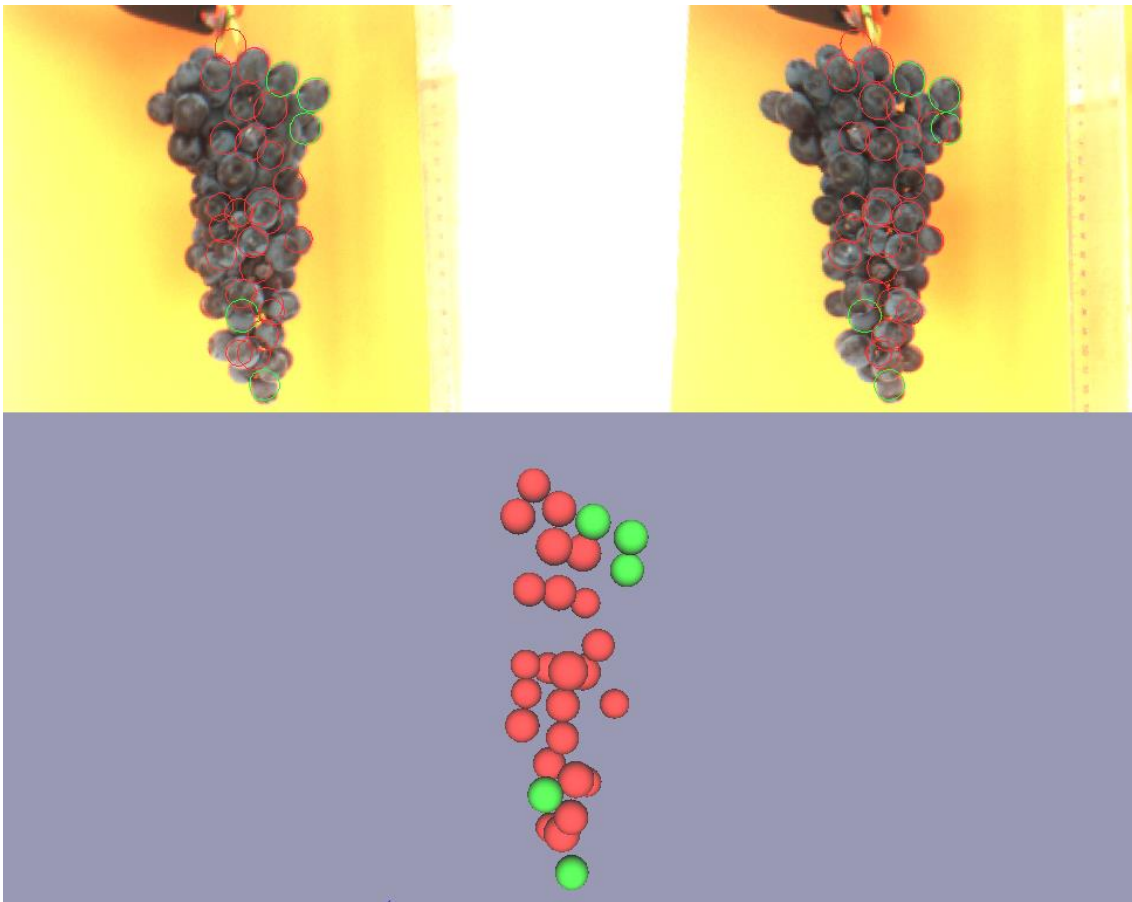


Figura 4.2-3

Plantamura



Com podem observar, els resultats varien en gran mesura depenent del raïm. Mentre que a la reconstrucció del raïm de la varietat *Cinsault* el nombre d'hipòtesis reconstruïdes correctament és alt, la reconstrucció de *Plantamura* dóna una gran quantitat d'hipòtesis incorrectes.

Per una altra part, si analitzem més detalladament aquestes hipòtesis incorrectes podem observar que, generalment, es tracta d'un problema d'error en la correspondència de posició entre les dues càmeres. És a dir, que mentre que en una imatge una hipòtesi s'ajusta correctament, a l'altra imatge ha patit un desplaçament. Això provoca que mentre que considerem aquestes hipòtesis incorrectes perquè la seua posició no és correcta, el seu radi sí que s'aproxima al dels grans reals.

És una cosa que cal destacar ja que, tot i ser hipòtesis incorrectes pel que respecta a la reconstrucció del raïm, l'error en la grandària dels grans no és tan accentuat, com podem observar en la següent taula.

Varietat	Nombre de grans	Grans correctes	Grans incorrectes	% de grans reconstruïts	Error de radis (mitja)
<i>Cinsault</i>	50	17	5	34.00%	0.7mm (8.64%)
<i>Bobal</i>	36	10	10	27.78%	0.31mm (4.05%)
<i>Monastrell</i>	40	9	3	22.50%	0.26mm (3.46%)
<i>Plantamura</i>	50	5	25	10%	0.16mm (2.04%)

L'error dels radis està calculat com a diferència entre la mitja dels radis en el model correcte i la mitja dels radis en el model reconstruït. Cal tindre en compte, finalment, que aquesta mesura és important ja que, a efectes d'avaluar la qualitat del raïm un dels paràmetres més importants és la grandària dels seus grans.

5 Conclusions

L'objectiu marcat a l'inici del projecte era aconseguir una ferramenta semiautomàtica per a la reconstrucció de raïm de vinificació que contribuïra al control de qualitat del raïm de vinificació.

Aquest objectiu ha sigut assolit mitjançant funcionalitat per a la reconstrucció automàtica de part del raïm així com mecanismes per a la reconstrucció manual d'aquest. La combinació d'aquestes dues tècniques junt a la visualització i interacció amb el model mitjançant una interfície gràfica d'usuari dona coma resultat la ferramenta presentada en aquesta memòria.

Com hem pogut observar a la secció de resultats, la ferramenta permet la reconstrucció de la part visible de qualsevol raïm a partir de dues fotografies i el calibratge de les càmeres. A partir del model reconstruït la extracció de característiques geomètriques és trivial i instantània.

La millora dels resultats de la part automàtica és definitivament possible, particularment en usar informació geomètrica per a anar realitzant la reconstrucció de manera iterativa. De fet, la ferramenta te parcialment implementada la funcionalitat que permet aquest mètode.

Destaquem així el potencial de la solució proposada per a ser ampliada, mitjançant la seua arquitectura modular i a certes decisions de disseny que han sigut preses amb la intenció de facilitar tant la millora de les tècniques presentades com la realització d'aplicacions similars, tot i estar fora de l'abast d'aquest projecte concret.

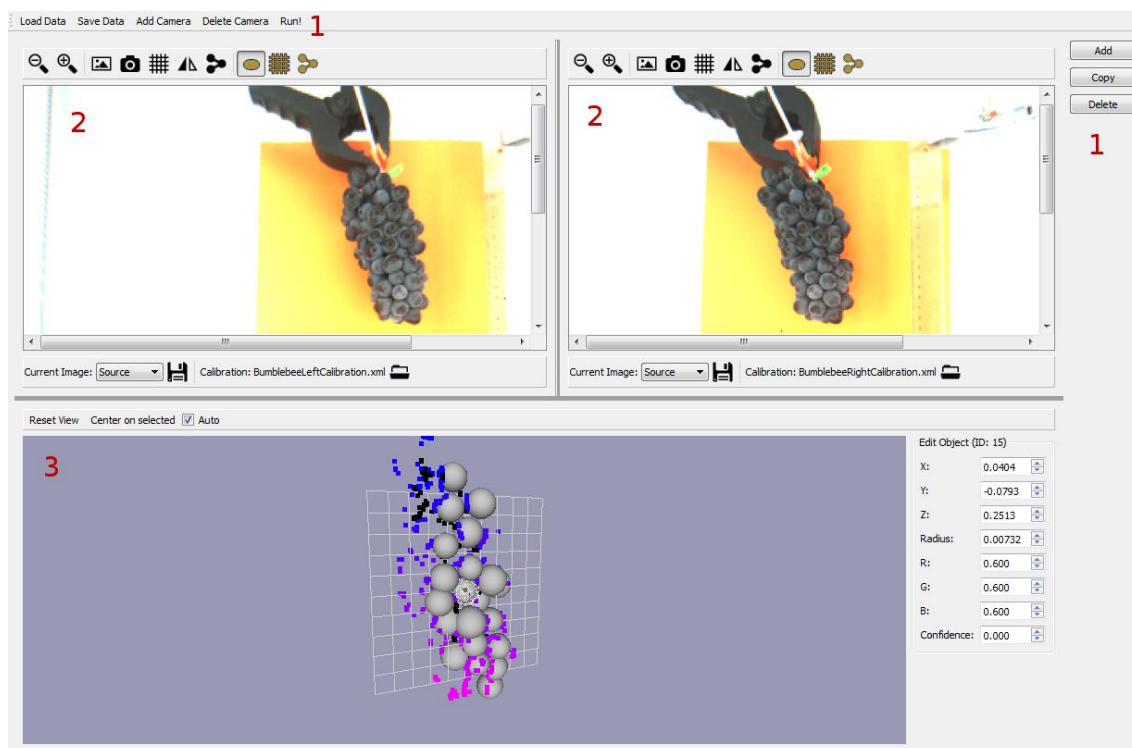
Podem establir, per tant, que aquesta ferramenta representa un avanç respecte a l'extracció de les característiques geomètriques de la manera tradicional, ja que requereix menys temps i esforç una volta s'ha instal·lat el sistema de captura de fotografies.

Amb tot, aquest projecte representa un pas lògic cap a l'automatització total del procés, el qual representaria un gran impacte socioeconòmic, si tenim en compte la importància del sector vinícola a Espanya, així com ha sigut discutit a la secció 1.1.

Finalment, concloem remarcant que els objectius marcats han sigut definitivament assolits i que el projecte estableix un suport a partir de la qual continuar la investigació en un camp tan important i amb tant de potencial com el control de qualitat vinícola a mitjançant tècniques de visió artificial tridimensional.

Annex A. Manual d'usuari

Introducció



La ferramenta permet a l'usuari visualitzar i interactuar amb el model així com iniciar la reconstrucció automàtica.

La interfície es divideix en tres parts: operacions generals (1), *Viewports2D* (2) i *Viewport3D* (3). A continuació detallarem les diferents funcionalitats que ofereixen cadascuna d'aquestes parts.

Operacions generals

Aquestes operacions són independents del tipus de visualització (2D o 3D) i és divideixen en dos conjunts.

Per una part, el primer conjunt d'operacions generals es troba a la part superior de la ferramenta:

- *Load Data*: Carrega les hipòtesis i els punts 3D des de d'un arxiu XML.
- *Save Data*: Guarda les hipòtesis i els punts 3D a un arxiu XML.
- *Add Camera*: Afegeix un nou *Viewport2D*.
- *Remove Camera*: Elimina l'últim *Viewport2D* afegit.
- *Run!*: Inicia la reconstrucció automàtica.

Sobre aquestes operacions cal destacar dues coses. Per una part, tot i que l'usuari pot afegir i eliminar *Viewports2D*, es recomana no fer-ho ja que el projecte està dissenyat per a treballar amb dues càmeres. Aquesta funcionalitat existeix per l'única raó de donar suport a possibles nous projectes que sí que utilitzen un nombre diferent de càmeres.

Per una altra part, cal tindre en compte que la reconstrucció automàtica és costosa temporalment i que la ferramenta no podrà ser utilitzada fins que no acabe la reconstrucció.

L'altre conjunt d'operacions generals, localitzades a la part dreta de la ferramenta, són les que permeten treballar sobre les hipòtesis:

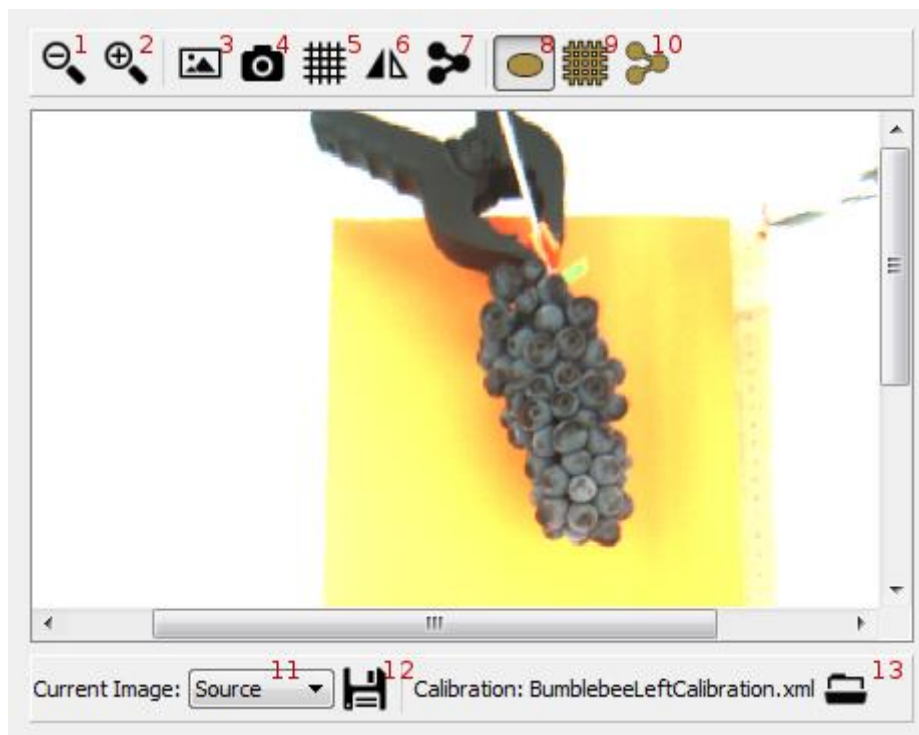
- *Add*: Afegeix una nova hipòtesi a partir dels paràmetres introduïts per l'usuari.
- *Copy*: Copia la hipòtesi seleccionada.
- *Delete*: Elimina la hipòtesi seleccionada.

Viewport2D

Un *Viewport2D* representa tota la informació corresponent a una de les càmeres des d'on es prenen les fotografies originals del raïm. Aquesta informació és:

- **Imatge original.** La imatge o fotografia original capturada des de la càmera.
- **Calibratge de la càmera.** Arxiu XML amb el calibratge de la càmera. Atés que totes les fotografies han sigut realitzades amb el mateix sistema de captura, els calibratges de la càmera esquerra i dreta són els mateixos per a totes les imatges originals.
- **Imatges derivades.** Imatges generades a partir d'alguna de les funcionalitats proporcionades pel *Viewport2D* (les quals expliquem a continuació).

Com s'ha comentat prèviament, la ferramenta permet afegir o eliminar *Viewports2D* però per a aquest projecte sols té sentit utilitzar-ne dos, ja que les fotografies originals s'han capturat mitjançant dues càmeres (esquerra i dreta).



A continuació llistem els controls disponibles al *Viewport2D*:

- Carregar calibratge (13).
- Carregar la imatge original (3).
- Zoom (1, 2).
- Generar imatge sintètica (*render*) (4).
- Generar imatge de vores (5).
- Comparar dues imatges (6).
- Projectar les hipòtesis sobre la imatge (7).
- Mode de generació d'hipòtesis (8).
- Mode de generació d'espina 3D(9).
- Mode de generació de punt 3D (10).

- Selecció d'imatge (11).
- Guardar imatge actual (12).

Calibratge i imatge original

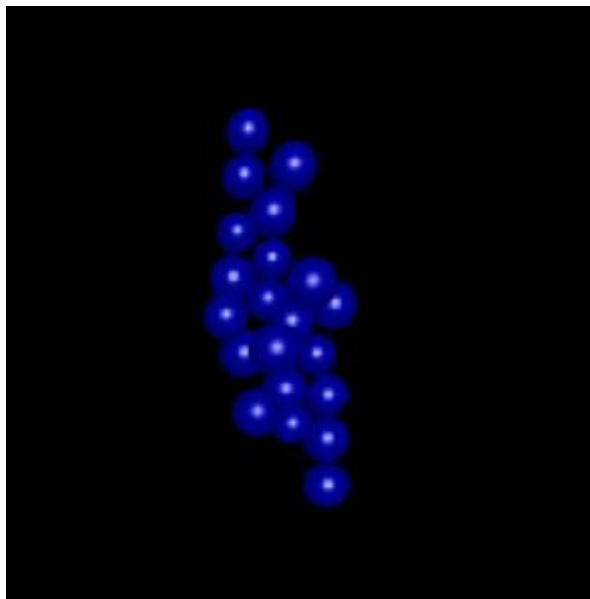
La càrrega del calibratge i de la imatge original són les operacions més essencials de la ferramenta i és el primer que cal de fer abans d'utilitzar qualsevol altra funcionalitat, ja que la reconstrucció no és possible sense aquestes dades.

Per conveniència, però, la ferramenta ja porta els calibratges carregats per defecte, així com les imatges originals corresponents a un dels tipus de raïm.

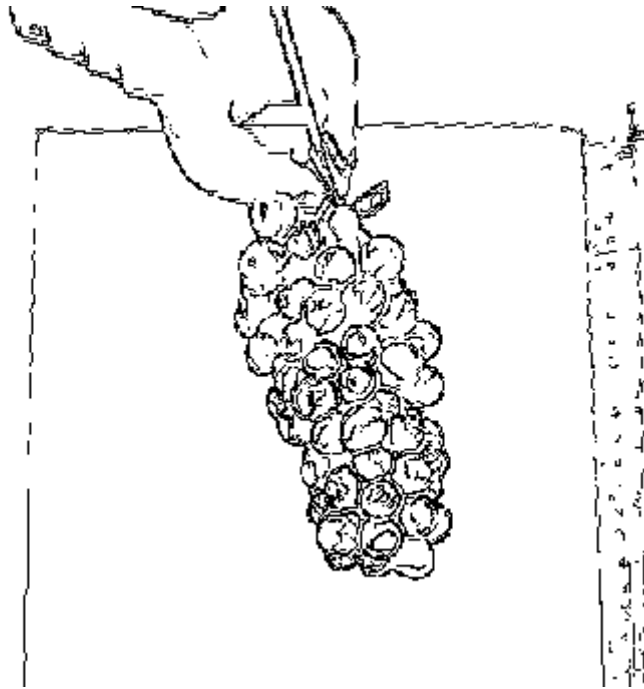
Imatges derivades

Com hem comentat, la ferramenta pot generar imatges derivades per a donar més informació a l'usuari, com la imatge sintètica, imatge de vores o imatge comparativa. L'usuari pot elegir quina visualitzar mitjançant el control de selecció d'imatge.

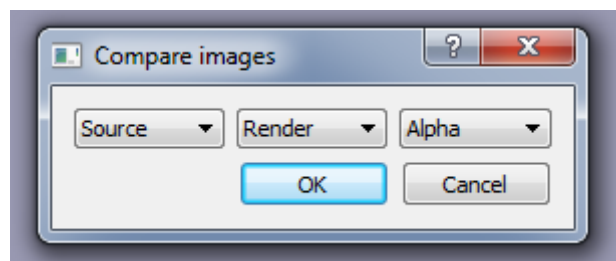
- La imatge sintètica consisteix en un *render* de les hipòtesis des del punt de vista de la càmera realitzat mitjançant un *raytracer*.



- La imatge de vores genera la imatge de vores detectades en la imatge original. Això permet, entre altres, predir a l'usuari quines hipòtesis seran més dificultoses per a reconstruir de manera automàtica.



- La imatge comparativa permet realitzar una comparació entre dues imatges ja existents (per exemple, imatge sintètica i imatge original) mitjançant un mètode de diferències o un mètode de superposició de la transparència (*alpha*). L'ús més habitual d'aquesta funcionalitat és el de comparar la imatge original i la imatge sintètica mitjançant el mètode *Alpha* per a obtindre una imatge comparativa en la qual podem observar visualment com s'ajusten les hipòtesis als grans.



El *Viewport2D*, a més de visualitzar informació, es pot fer servir perquè l'usuari en genere de nova utilitzant les imatges com a guia visual. L'usuari pot especificar quina funcionalitat vol utilitzar a través de seleccionar un dels tres modes: generació d'hipòtesis, generació d'espina i generació de punts.

La generació de punts i generació d'espina s'utilitzen principalment amb objectius de test i investigació i no tenen, ara per ara, repercussió en el procés de reconstrucció.

La generació d'hipòtesis és una de les parts més importants de la ferramenta i és el nucli de la reconstrucció manual. Aquest mode permet a l'usuari generar una nova hipòtesi a partir de les imatges 2D.

Per a generar una hipòtesi mitjançant aquest mètode l'usuari ha de seguir els següents passos:

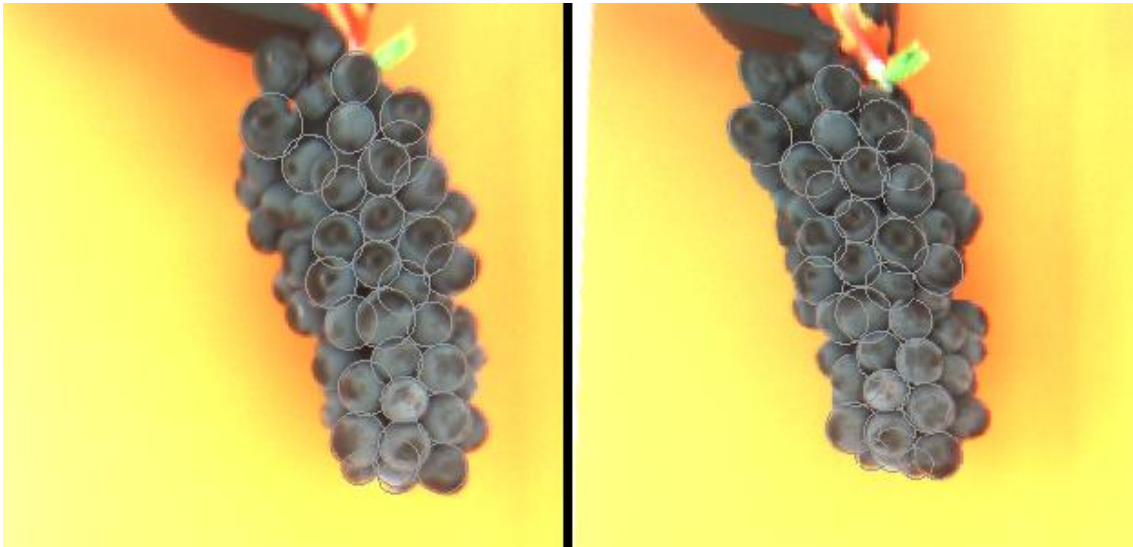
- Activar el mode de generació d'hipòtesis a un dels *Viewport2D*.
- Introduir 5 punts corresponents al gra mitjançant el ratolí.
- A l'altre *Viewport2D*, introduir el centre del gra.

Una volta realitzats aquests passos es crearà la nova hipòtesi i apareixerà al *Viewport3D*.

Projecció d'hipòtesis

Projectar les hipòtesis sobre les imatges és la manera de visualitzar les hipòtesis en l'espai 2D i, per tant, comprovar en quina mesura s'ajusten al gra real.

Mitjançant el botó de projecció d'hipòtesis es generarà una nova imatge resultat de representar les projeccions de les hipòtesis sobre la imatge que s'estava visualitzant.



Viewport3D

El *Viewport3D* és on es presenta a l'usuari les hipòtesis a l'espai 3D. L'usuari pot navegar per aquest espai de les següents maneres:

- Rotar càmera: Mantindre botó dret del ratolí.

- Desplaçar càmera: Mantindre boto central del ratolí.
- Controlar zoom: Roda del ratolí.
- Seleccionar hipòtesi: Clic esquerre.

A més es proporcionen diferents accions que faciliten la navegació, disponibles a la part superior del *Viewport3D*:

- *Reset View*: Reinicia la càmera al seu estat original.
- *Center on selected*: Centra la vista en la hipòtesi seleccionada.
- *Auto*: Si està activada, la càmera es centrarà automàticament al canviar d'hipòtesi seleccionada.

El *Viewport3D* també permet interactuar amb les hipòtesis mitjançant la modificació dels seus paràmetres:

- *X, Y, Z*: Posició.
- *Radius*: Radi.
- *R, G, B*: Color.
- *Confidence*: Paràmetre de confiança.

Annex B. Manual de programador

El manual de programador consisteix en el llistat i documentació de les classes utilitzades per a implementar la ferramenta. Per conveniència s'adjunta com a document a part.