

OBOCV: Integració de la visió per computador en un entorn tridimensional

Gaspar Alomar Trias

Quint curs d'Enginyeria Informàtica
gaspar.alomar@gmail.com

Resum— L'objectiu del present projecte informàtic és integrar tècniques de visió per computador en un motor de renderització de gràfics tridimensionals. El projecte desenvolupa permet realitzar una interacció entre la persona i els objectes virtuals que conformen l'escena per tal de profunditzar en el desenvolupament d'aplicatius per a la rehabilitació de persones amb discapacitats psicomotrius.

I. INTRODUCCIÓ

La Unitat de Gràfics i Visió per computador i Intel·ligència Artificial (UGIVIA) del Departament de Ciències Matemàtiques i Informàtiques (DMI) treballa en el desenvolupament d'aplicacions que potenciïn l'accessibilitat a persones amb discapacitats físiques mitjançant la interacció persona-ordenador fent servir la visió per computador.

Primerament es varen desenvolupar aplicacions que detecten, a través d'una càmera web, certes parts del cos de l'usuari i en fan un seguiment de la seva posició per convertir-ho en el moviment del cursor del sistema operatiu. El primer d'aquests aplicatius fou el *Sistema d'Interacció Natural Avançat (SINA)*, que detecta el nas de l'usuari i fa un seguiment del moviment d'aquest [1]. Posteriorment es va desenvolupar el *SINA Eyes*, en el que els responsables del moviment del cursor passen a ser els ulls del pacient [2].

Mentre es provaven aquestes aplicacions amb persones discapacitades, els especialistes encarregats de les seves cures advertiren que es produïen millores en les seves capacitats motores. Això va motivar nous desenvolupaments com jocs per a la rehabilitació de persones amb problemes psicomotrius [3]. Aquests jocs, que detecten la presència de l'usuari a través d'algorismes de detecció de pell, permeten la interacció del pacient amb objectes sintètics que es mostren sobre la imatge capturada per la càmera web (veure Figura 1).



Figura 1: Videojoc per a rehabilitació mitjançant visió per computador.

II. MOTIVACIONS

La principal motivació del present projecte és la integració de la visió per computador en un entorn tridimensional que permeti interactuar l'usuari amb objectes virtuals. Amb aquest projecte es pretén obrir un ventall de noves aplicacions per a la rehabilitació de persones amb mancances psicomotrius ja que incorporen elements que es mouen dins un espai tridimensional, amb les seves corresponents dinàmiques.

A més, aquest projecte esdevé la primera passa per arribar a una futura interacció en l'espai tridimensional, que permetria al desenvolupador considerar la profunditat dels objectes de l'escena, tot a través de càmeres web capacitades per obtenir la informació de distància respecte de la càmera.

III. DESENVOLUPAMENT

OBOCV, acrònim obtingut dels noms de les llibreries *OGRE*, *Bullet* i *OpenCV* utilitzades per desenvolupar el present projecte, és una llibreria encastada dins el motor de renderització de gràfics tridimensionals *OGRE*. Permet la senzilla implementació d'aplicacions que permeten interactuar a l'usuari amb objectes virtuals a través de la visió per computador.

A continuació s'expliquen les etapes que s'han seguit pel desenvolupament d'*OBOCV*.

A. Composició de l'Entorn Tridimensional

L'entorn tridimensional està conformat per la llibreria *OGRE*, un motor de renderització multiplataforma escrit en C++ i que compta amb una llicència lliure. És un dels motors més utilitzats en l'actualitat, pel que per aquests motius ha estat l'elegit per desenvolupar el projecte.

A més, se li ha encastat el llibreria de físiques i dinàmiques *Bullet* per dotar de moviments als objectes, alhora de poder detectar col·lisions entre aquests.

B. Integració de la Visió per Computador

La llibreria encarregada de la captura, tractament i visualització de la imatge és *OpenCV*, una llibreria multiplataforma escrita en C++ i que també compta amb una llicència lliure. Per això -i per ser la llibreria utilitzada a l'UGIVIA pel desenvolupament d'aplicacions que fan ús de la visió per computador- s'ha triat *OpenCV* en el present projecte.

Per mostrar la imatge capturada amb la càmera web a l'entorn tridimensional d'*OGRE*, es dibuixa un plànol bidimensional a l'escena quan s'inicia l'aplicació, al qual se li associa una textura tridimensional. Aquesta textura tridimensional s'actualitza cada vegada que *OGRE* ha de renderitzar un nou *frame*, capturant la imatge amb la càmera web i projectant-la a la citada textura. A la Figura 2 se'n pot veure un esquema del resultat d'aquest procés.

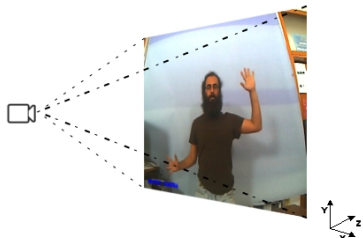


Figura 2: Ubicació de la càmera i del plànol amb la imatge real en l'entorn tridimensional

Posteriorment, s'ubica la càmera de renderització de l'escena d'*OGRE* de tal manera que, quan es visualitza l'escena tridimensional, el plànol on es mostra la imatge real quedi enquadrat dins la finestra de visualització, com es pot veure a la Figura 3.

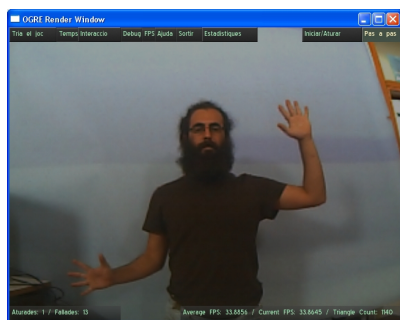


Figura 3: Resultat a la finestra de visualització

C. Interacció entre l'Usuari i els Objectes Virtuals

Per a determinar la interacció entre l'usuari i els objectes virtuals, primerament s'ha de detectar quan es produeix una col·lisió entre l'objecte i el plànol que conté la imatge. S'ha desenvolupat un algoritme de detecció de col·lisions que, una vegada que *OGRE* ha renderitzat un nou *frame* de l'escena, es comprova si qualche objecte dels que conforma l'escena té un punt de contacte amb el plànol. A més, es comprova que l'objecte no hagi col·lisionat anteriorment amb la imatge: és possible que es detecti el punt de contacte en dos *frames* consecutius, pel que cal evitar dobles positius.

Quan es detecta una col·lisió entre un objecte virtual i el plànol (és a dir, un punt de contacte entre ambdós), es procedeix a determinar si s'ha produït una interacció entre l'usuari i l'objecte. A continuació s'expliquen les diferents

fases d'aquest procés: detecció de l'usuari, detecció del punt de contacte i creuament d'informació.

Inicialment es detecta l'usuari a l'imatge a través de la heurística de Kovač, que determina quins píxels tenen una informació de color pròpia de la pell humana [4]. Per això és fa una comprovació píxel a píxel per veure quins compleixen les condicions que determina dita heurística (veure Taula I).

TAULA I

NORMES HEURÍSTIQUES DE KOVAČ PER A LA DETECCIÓ DEL COLOR DE PELL

(R, G, B) és pell si:	$(R > 95) \text{ and } (G > 40) \text{ and } (B > 20)$ $\text{and } (\max(R, G, B) - \min(R, G, B) > 15)$ $\text{and } (R - G > 15) \text{ and } (R > G) \text{ and } (R > B)$
----------------------------------	--

Aquesta informació es tradueix en una màscara binària: una imatge en la que es pinten de negre els píxels on no s'ha detectat pell i de blanc els píxels en els que sí s'ha detectat, com es pot veure a la Figura 4.



Figura 4: Detecció de pell i generació de la màscara de l'usuari

Seguidament es genera una segona màscara binària amb informació de la posició del punt d'impacte. Aquesta és tota negra excepte el punt al plànol on ha col·lisionat l'objecte que es pinta blanc, com es mostra a la Figura 5.



Figura 5: Generació de la màscara de la zona d'impacte

Finalment, es creua la informació d'ambdues màscares per determinar si la col·lisió s'ha produït en una zona on es troba l'usuari. En tal cas, s'hauria produït una interacció entre la persona i l'objecte virtual pel que al resultat de multiplicar ambdues màscares hi apareixerien píxels blancs.

D. Gestió de la Interacció

Tot el que s'ha vist fins aquest punt ho gestiona el propi *OBOCV* de tal manera que el desenvolupador d'aplicacions no s'ha de preocupar; cal recordar que *OBOCV* és una llibreria encastada dins *OGRE*. El desenvolupador que vulgui implementar una aplicació amb *OBOCV*, decideix quin comportament tendran els objectes col·lisionats en funció de que hagin interactuat amb l'usuari o no.

E. Seguiment de la Interacció

S'ha desenvolupat una finestra de seguiment de la interacció de l'usuari. D'aquesta manera, l'especialista és capaç de conèixer en directe la resposta de l'usuari als estímuls (objectes virtuals) i si aconsegueix o no interactuar amb aquests. A més, dita finestra permet que l'especialista realitzi una execució de l'aplicació personalitzada segons les necessitats de la persona tractada i en funció de com va respondre. La Figura 6 mostra una captura de dita finestra.



Figura 6: Finestra per a realitzar un seguiment de les interaccions

F. Recull de Resultats

Finalment, una vegada s'acaba l'execució de l'aplicació, es desa automàticament un fitxer d'imatge amb els resultats de la interacció. Les col·lisions on s'ha produït una interacció es pinten de color verd, les que no es pinten de vermell (és semblant a la imatge del cantó inferior-dret de la Figura 6) i s'hi imprimeix un recompte dels resultats. Aquesta informació desada permet a l'especialista tenir al seu abast un historial de l'evolució del pacient, poden determinar quins aspectes està millorant de la seva capacitat motora.

IV. RESULTATS

Per avaluar els resultats de la llibreria OBOCV, s'ha desenvolupat un joc que consisteix en llençar esferes des de la posició de la càmera cap al plànol on es troba l'usuari, com es mostra a la Figura 7.

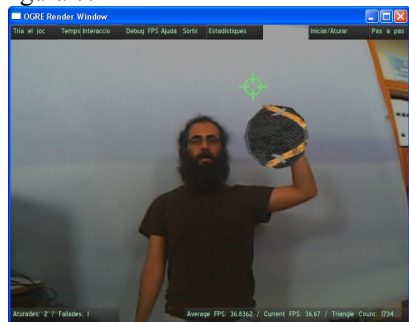


Figura 7: Captura del joc desenvolupat amb OBOCV

El joc té disponibles dos modes de funcionament:

- **manual:** les esferes les llença una segona persona (l'especialista) mitjançant el ratolí, decidint on es dirigeixen.
- **automàtic:** cada vegada que impacta una esfera en el plànol, es llença una altra amb un angle de tir aleatori.

La lògica desenvolupada pels esdeveniments d'entrada es mostra a la Taula II:

TAULA II

COMPORTAMENT DEL JOC EN FUNCIO DE LA INTERACCIÓ

Esdeveniment		Resposta
Teclat	Espai	Canvia mode joc manual/automàtic
	'+'	Augmenta velocitat llançament.
	'-'	Disminueix velocitat llançament.
Ratolí	Botó esquerre	Llença esfera a posició cursor.

A més, s'ha desenvolupat la lògica de la Taula III per a gestionar les col·lisions en funció de si s'ha produït interacció.

TAULA III

COMPORTAMENT DEL JOC EN FUNCIO DE LA INTERACCIÓ

Esdeveniment	Resposta
Interacció	S'aplica una força cap avall a l'esfera.
No Interacció	S'elimina l'esfera.

V. CONCLUSIONS

S'han assolit els principals objectius que s'havien marcat pel projecte: s'ha integrat la visió per computador en un entorn 3D, s'ha desenvolupat la interacció persona-objectes virtuals, s'ha desenvolupat un seguiment en directe i posterior avaluació de la interacció i s'ha utilitzat maquinari de baix cost i programari lliure.

Per tant, es pot concloure que és possible desenvolupar noves aplicacions d'interacció persona-ordinador d'una manera econòmica, lliure i, com s'ha vist a l'aplicació desenvolupada a l'apartat IV Resultats, de manera senzilla i intuïtiva.

També s'han desenvolupat les eines necessàries per a que l'especialista que utilitzi l'aplicació pugui fer un seguiment de la resposta del pacient, a més de poder governar l'aplicatiu en funció de les necessitats/capacitats de la persona a rehabilitar.

Finalment, cal apuntar que mitjançant la introducció de càmeres RGB-D, que permeten obtenir informació de la profunditat en la que es troben els elements dins l'escena, és possible avançar cap una interacció tridimensional absoluta.

REFERÈNCIES

- [1] UGIV et al. Sina (sistema natural d'interacció avançat). <http://www.caib.es/sacmicrofront/contenido.do?idsite=1625&cont=29559&lang=ca>, 22 novembre 2010. xiii
- [2] S.Garcés; J.Varona; F.J.Perales. Raó'n funcional no invasivo y estudio funcional mediante bci. *Simposio CEA de Bioingenier'a 2010. Redes REDINBIO y RETADIM*, 2010. xiii
- [3] B.Moyà; A.Jaume; J.Varona; P.Martinez-Bueso; A.Mesejo; Use of serious games for motivational balance rehabilitation of cerebral palsy patients. *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility Pages 297-298*, 2011. xv,13
- [4] Jure Kovač; Peter Peer; and Franc Solina. Human skin colour clustering for face detection. In *EUROCON 2003. Computer as a Tool. The IEEE Region 8, volume 2, pages 144-148*, 2003. 31