



Interfície head-tracking per a dispositius mòbils

*Grau d'Enginyeria Informàtica
Intel·ligència artificial*

Estudiant: Miquel Àngel Alcalde Velado

Tutor: Dr. Joan M. Nuñez Do Rio

Juny 2019

INDEX

1. Introducció.
2. Disseny modular.
3. Algorisme de visió.
4. Metodologia de verificació.
5. Conclusions.

INTRODUCCIÓ

Els avanços tecnològics en les darreres dècades han permès el desenvolupament de nous paradigmes d'**interacció usuari-màquina**.

La disciplina HCI estudia de diferents aspectes de la interacció usuari-màquina com la usabilitat, l'ergonomia i l'**accessibilitat**.

Les **interfícies d'interacció natural** són aquelles en les quals l'usuari interactua amb la màquina utilitzant les seves capacitats naturals, sense necessitat d'elements auxiliars (*Kinect, Alexa, Siri*).

PROBLEMA A RESOLDRE

Segons dades de l'OMS s'estima que el **15.6% de la població mundial**, el 8.6% a Espanya, pateix algun tipus de discapacitat.

L'accés a la tecnologia **no és igualitari** per a la totes persones, les persones amb discapacitat té dificultats per a usar la tecnologia.

Aquestes barreres d'accés creen una **bretxa digital** que té com a conseqüència l'aparició de discriminació i marginalitat.

Avui dia els **telèfons mòbils** s'han convertit en un element essencial de les nostres vides, cal dotar d'eines a les persones discapacitades que els facilitin el seu ús.

OBJECTIUS DEL PROJECTE

L'objectiu d'aquest projecte és el disseny i implementació d'una **interfície d'interacció natural per a dispositius mòbils** que permeti als desenvolupadors d'aplicacions la integració d'aquest tipus d'interfície en les seves aplicacions.

Aquest tipus d'interfície no invasives és especialment útil per a permetre l'ús de les aplicacions a persones discapacitades amb **problemes de motricitat als braços**.

La interfície ha de permetre el moviment d'un cursor virtual per la superfície de la pantalla utilitzant per a la **detecció dels moviments del cap** mitjançant la implementació d'un algorisme de visió per a tal finalitat.

DISSENY MODULAR

“L'objectiu principal del projecte és el desenvolupament d'una interfície d'usuari per head-tracking per a dispositius mòbils”.

Aquest criteri de portabilitat determina una de les principals decisions de disseny aplicada: el **disseny ha de permetre que el sistema sigui portable a noves plataformes.**

El disseny de components de la solució s'estructura de forma modular en les següents capes:

- Capa d'aplicació: defineix la interfície pública amb les aplicacions.
- Capa de plataforma: implementa els detalls de la interfície particulars de l'entorn.
- Capa de portabilitat: implementa la lògica comuna de l'algorisme de visió implementat.

CAPA DE PLATAFORMA

En aquesta capa s'implementa la lògica de la interfície natural implementada en el projecte que és **dependent de cada plataforma mòbil**.

És responsabilitat d'aquesta capa:

- La **comunicació amb la capa de portabilitat** encarregada de processament de les imatges obtingudes des de la càmera i l'aplicació de la lògica de visió.
- Gestió de la resposta del processament de la imatge per la capa de portabilitat i publicació dels resultats a la capa d'aplicació mitjançant un **esquema basat en esdeveniments**.
- Implementació de la lògica associada a la transformació de l'esdeveniment de selecció en un **esdeveniment de tipus touch** compatible amb la plataforma.

CAPA DE PORTABILITAT

En aquesta capa s'implementa la lògica de la interfície natural implementada en el projecte que és **comú a totes les plataformes**.

És responsabilitat d'aquesta capa:

- La implementació de la lògica de visió encarregada de processar les imatges de la càmera per determinar el moviment del cursor virtual per detecció del **moviment del nas** de l'usuari.
- La implementació de la lògica de visió encarregada de processar les imatges de la càmera per determinar el gest de selecció per detecció del **gest de picada d'ullet** per part de l'usuari.
- La implementació de la lògica de visió encarregada d'analitzar les **condicions d'il·luminació** de l'entorn i el **posicionament de l'usuari**.

ALGORISME DE VISIÓ

És la lògica encarregada d'acceptar les imatges sense tractar des de la càmera del dispositiu per a determinar la **posició d'un element cursor** mitjançant l'aplicació de lògica de visió per computador adient per a realitzar el **seguiment d'un punt característic** situat en la **punta del nas**.

L'algorisme de visió es compon de diverses fases que s'executen de forma seqüencial:

- Fase de detecció.
- Fase de seguiment.
- Fase d'estimació.
- Fase de transferència.

FASE DE DETECCIÓ

L'objectiu d'aquesta fase és aplicar la lògica necessària per a determinar un **conjunt de punts característics** situats en la punta del nas sobre el qual aplicar l'algorisme de seguiment.

Aquesta fase es componen de següents tasques:

- Detecció de la **presència del rostre** de l'usuari mitjançant l'aplicació d'un classificador en cascada.
- Determinació de l'àrea del nas mitjançant aplicació de **regles antropomòrfiques del rostre**.
- Ajust de la zona del nas per aplicació de la **posició dels ulls** detectada mitjançant un classificador en cascada.
- Obtenció del conjunt de punts característics en la **zona del nas fitat**.

FASE DE SEGUIMENT

L'objectiu d'aquesta fase és l'aplicació de l'algorisme de **flux òptic** per a determinar la **direcció i magnitud del moviment** del conjunt de punts característics obtinguts en la fase de detecció.

Dues accions es realitzen per a eliminar del conjunt de punts característics aquells de menor rellevància:

- Filtrat per distància del punt central de la zona del nas a fi d'eliminar els punts no centrals.
- Aparellament dels punts de les regions dividides per l'eix vertical a fi d'eliminar la influència de la llum lateral.

La determinació del punt de seguiment associat al cursor virtual es calcula mitjançant **mitjana aritmètica dels punts característics** obtinguts.

FASE D'ESTIMACIÓ

El procés de determinació del punt mitjà es troba sotmès a **fontes d'error que introdueixen inconsistències** en la posició obtinguda.

Canvis en la il·luminació, la influència de les ombres, l'oclusió parcial del rostre o les imperfeccions dels sensors són algunes de les causes d'aquests errors.

L'aplicació del filtre de Kalman permet **estabilitzar la trajectòria de punt de seguiment** obtingut per a determinar la posició de manera més robusta.

FASE DE TRANSFERENCIA

Per a poder utilitzar les coordenades del punter de seguiment en la capa d'aplicació cal convertir aquesta posició a l'espai de **coordenades de pantalla**.

Factors com la posició de l'usuari en la imatge, la distància a la càmera i capacitats motrius de l'usuari són factors que influeixen en el **rang articular de l'usuari**. Cal adaptar la transformació a les característiques de moviment de l'usuari.

El procés realitza un **ajust dinàmic dels valors d'escalat** del moviment per a cadascun d'eixos i direccions sobre la base del rang articular detectat de l'usuari.

MECANISME DE TOLERÀNCIA A FALLADES

Per afegir **robustesa a l'algorisme** i evitar que les fallades esporàdiques en el processament de les imatges deteriori el rendiment s'ha incorporat un mecanisme de tolerància a fallades.

L'esquema de **finestra de validesa** defineix l'interval en nombre de *frames* durant el qual la darrera estimació de la posició continua sent vàlida.

Les fallades de detecció ocorregudes dins de l'interval de la finestra de validesa són ignorats permetent al sistema continuar el procés de seguiment.

MÀQUINA D'ESTAT DE L'ALGORISME

Durant l'execució de l'algorisme de visió, aquest pot trobar-se en **diferents estats**:

- Estat inicial.
- Estat de calibratge.
- Estat de seguiment.

Per a cadascun dels estats s'estableixen els llindars mesurats en número de *frames* processats amb èxit o fracàs per a determinar la **transició cap a altres estats**.

La divisió en estats permet **optimitzar el rendiment** de l'algorisme de visió pel fet que en cada estat s'executen els passos necessaris segons els requeriments de l'estat.

DETECCIÓ DEL GESTO DE CLICK

És la lògica encarregada de determinar si l'usuari està realitzant el **gesto de picada d'ullet** mitjançant l'anàlisi de la imatge obtinguda des de la càmera.

La lògica està basada en tècniques de visió per **aproximació de la forma general de l'ull** mitjançant un polígon convex.

El sistema proveeix d'una implementació alternativa basada en el **temps de permanència** del cursor virtual en una mateixa posició.

CONTROL DE LES CONDICIONS DE L'ENTORN

La influència de les condicions de l'entorn i de les condicions dependents de l'usuari són **factors rellevants** en el rendiment de l'algorisme de visió.

- En **condicions d'il·luminació** adequades el rendiment dels classificadors i algorismes de visió usats és òptim.
- Un correcte **posicionament de l'usuari** afavoreix el rendiment de la fase de transferència.

El sistema **comprova de forma periòdica l'estat** d'aquestes dues variables per a informar del seu estat a la capa d'aplicació.

METODOLOGIA DE VERIFICACIÓ (I)

En aquesta fase es verifica que la interfície head-tracking **compleix amb els objectius** definits en el projecte. Es va comprovar els següents requisits funcionals:

- L'usuari sigui capaç de moure el cursor virtual per la pantalla de l'aplicació.
- Que el moviment de cursor sigui fluid.
- Que l'algorisme sigui tolerant a fallades i pèrdues de seguiment.

Per a això s'ha **implementat una aplicació** que actua com a prova de concepte, aquesta aplicació va ser instal·lada en diversos terminals i posada a la disposició de diversos usuaris.

METODOLOGIA DE VERIFICACIÓ (II)

L'aplicació desenvolupada s'implementa com un petit joc, l'objecte del joc és **tocar els cercles de color** que apareixen en la pantalla utilitzant el cursor virtual controlat per moviment del nas.

Les posicions possibles en les quals els cercles es defineixen per la divisió de la pantalla en una **gralla de 11x11** (121 posicions).

El cada cicle de joc es **validen totes les posicions** possibles.

L'aplicació té dues modalitats de joc:

- Modalitat de cobertura: els cercles són de major dimensió i el temps de selecció és menor.
- Modalitat de precisió: els cercles són de menor dimensió i el temps de selecció és més major.

METODOLOGIA DE VERIFICACIÓ (III)

Durant el cicle de joc, l'aplicació recull **indicadors i mètriques** de rendiment de l'algorisme i, del comportament de l'usuari.

Es va realitzar una anàlisi del conjunt de les dades recollides en totes les **sessions d'usuari**.

Es van extreure les següents **hipòtesis**:

- En relació al **rendiment del sistema es va establir que el valor més determinant és el desplaçament màxim**, s'atribueix aquesta relació al fet que a major intenció de moviment, l'angle d'orientació frontal d'usuari es perd de forma més freqüent, disminuint el rendiment del classificador.
- En relació al **rendiment de l'usuari es va establir que el valor més determinant és el valor d'exposició de l'escena**, s'atribueix aquesta relació al fet que a millor condició d'il·luminació, l'algorisme és més estable i es produeixen menors perdudes de seguiment.

CONCLUSIONS

L'algorisme de visió basat en la detecció de moviments del cap implementat en aquest projecte és una **interfície d'usuari NUI viable** per a ser usada en entorns mòbils.

El disseny modular del sistema permet **minimitzar l'esforç d'integració** i aïllar als desenvolupadors dels detalls de l'algorisme de visió.

La implementació d'interfícies *head-tracking* en dispositius mòbils no aquesta exempta de dificultats, per la qual cosa és necessari que s'implementin **funcionalitats per a millorar la robustesa i tolerància a fallades de la solució**.

Es proposen com línies de treball futur:

- Realitzar la implementació de la capa de plataforma necessària per a portar la solució a **dispositius IOS**.
- Establiment d'un acord de col·laboració amb associacions de discapacitats, entitats sanitàries i empreses de desenvolupament interessades, a fi de l'elaboració d'un estudi formal que permeti extreure conclusions sobre el **benefici d'ús de la solució en persones discapacitades**.