

# Projecte de Recerca

## DETECCIÓ AUTOMÀTICA DE PUNTS FACIALS CARACTERÍSTICS UTILITZANT RANDOM FERNS I EL CRITERI DE LA INFORMACIÓ MÚTUA SIMPLIFICADA

David Masip Rodo, Marc Argüelles Puigibet.  
*Universitat Oberta de Catalunya.*

**Resum** — En aquest article comparem el rendiment que presenten dos sistemes de reconeixement de punts característics en imatges: en el primer utilitzem la tècnica Random Ferns bàsica i en el segon (que anomenem Ferns amb Informació Mútua o FIM) apliquem una tècnica d'obtenció de Ferns utilitzant un criteri simplificat de la Informació Mútua. FIM ens permet eliminar l'aleatorietat durant la selecció dels test que formen part dels Ferns amb l'objectiu de seleccionar aquells que ens resultin més favorables a l'hora del reconeixement. Si bé la creació del classificador amb FIM és més lenta que amb Random Ferns, a l'article demostrarem que el procés de reconeixement i classificació d'imatges millora tant en eficàcia com en rendiment. Les proves experimentals s'han centrat en el reconeixement de punts facials característics, generant els classificadors a partir de múltiples imatges distorsionades de les zones que envolten aquests punts. El classificador és calculat aplicant múltiples test binaris simples, basats en el nivell d'intensitat dels píxels de les imatges. La tècnica dels Random Ferns selecciona píxels aleatoris on aplicar els càlculs d'intensitat, mentre que la tècnica FIM que presentem a l'article intenta descartar els tests poc útils i selecciona els que ens han d'afavorir a l'hora de generar el classificador. Això ens permet obtenir uns classificadors més petits amb uns índex de reconeixement iguals o superiors que en el cas dels Random Ferns.

**Paraules Clau** — Reconeixement de Punts Crítics, Ferns amb Informació Mútua, Random Ferns.

### 1 INTRODUCCIÓ

Moltes aplicacions de visió per computador depenen d'algorismes de reconeixement de punts característics en imatges. L'objectiu d'aquests algorismes és implementar sistemes que siguin capaços d'analitzar els punts característics dels objectes o les cares model per, posteriorment, identificar-los i classificar-los correctament en imatges noves.

Les tècniques que presentem en aquest article analitzen la zona de la imatge que envolta els punts clau a reconèixer. Les zones que embolcallen el punt clau són estudiades mitjançant comparacions d'intensitat de dos píxels. Executant moltes d'aquestes proves i tractant correctament les dades obtingudes generem un classificador robust que ens proporciona la capacitat de detectar els punts clau en noves imatges.

A l'article comparem el rendiment de dues tècniques diferents. La primera, Random Ferns, s'ha utilitzat àmpliament en molts articles d'investigació i és molt coneguda en el camp de la visió per computador. Aquesta tècnica selecciona els tests d'intensitat de manera aleatòria, és a dir, els píxels analitzats que envolten el punt clau es seleccionen totalment a l'atzar. Aquests tests s'agrupen posteriorment en grups anomenats Ferns. El fet

que els tests es seleccionin a l'atzar no implica que el mètode no sigui ràpid o robust. De fet, s'ha demostrat experimentalment que els Random Ferns són ràpids, simples, eficaços i robustos [1]. Al nostre article intentem sacrificar la simplicitat dels Random Ferns per obtenir una rapidesa i un rendiment en temps d'execució més elevats. Per aconseguir-ho, utilitzem un criteri simplificat de la informació mútua[6]. La idea de fons és executar proves sistemàtiques per analitzar quins píxels de la zona que envolta el punt crític ens interessin més per crear el classificador. Això permet descartar tests de píxels poc interessants amb l'objectiu d'obtenir uns resultats de detecció similars als de Random Ferns però utilitzant una quantitat de tests Ferns menor. La tècnica presentada és anomenada en aquest article *Ferns amb Informació Mútua (FIM)*.

Per crear els classificadors RF o FIM utilitzem múltiples imatges de model on seleccionem un punt característic a analitzar. A partir d'aquestes imatges model generem noves imatges distorsionades de la zona que envolta el punt característic. Aquestes imatges ens permetran entrenar el classificador amb l'objectiu que estigui el més preparat possible per tal de classificar el punt en una nova imatge (a la vida real, una imatge d'un

objecte concret sempre pot presentar modificacions d'angle, distància, lluminositat o soroll quan s'obté una nova imatge del mateix). Concretament, si l'objectiu del sistema final és analitzar imatges de model de baixa qualitat (com imatges de vídeo obtingudes a través d'una webcam) és necessari entrenar el classificador amb moltes més imatges de model distorsionades que si només s'utilitzen imatges d'alta qualitat que presentin pocs canvis.

El nostre article s'ha centrat a analitzar les diferències entre RF i FIM en condicions bastant favorables, és a dir, fotografies de sets d'imatges amb poc soroll i bona qualitat d'imatge. S'ha de tenir en compte que en entorns més complicats normalment es requereix un entrenament més exhaustiu del classificador, amb més tests Ferns i més imatges de model, i que també existeixen tècniques específiques que permeten millorar els resultats quan ens enfrontem amb aquests reptes (suposem que algunes d'aquestes tècniques també podrien incorporar el concepte dels Ferns amb Informació Mútua al seu disseny per beneficiar-se dels seus avantatges).

## 2 PLANTEJAMENT

### 2.1 Aspectes Generals

Per crear i entrenar un classificador que ens permeti detectar els punts característics en una imatge, hem de disposar prèviament d'una imatge de model com a mínim. En aquest article ens hem centrat a detectar punts facials humans característics, especialment els situats al centre dels ulls. Alguns dels punts facials més habituals utilitzats en les proves de reconeixement facial són els que es poden observar a la Figura 1.

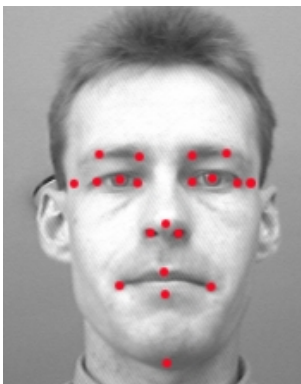


Fig. 1. Possibles punts crítics per realitzar una detecció facial.

Per obtenir les coordenades dels punts facials es poden utilitzar algoritmes específics que permetin automatitzar el procés. Una altra opció és proporcionar al classificador les coordenades dels punts facials manualment. Aquesta última opció té més sentit per realitzar proves experimentals, ja que es poden especificar les coordenades exactes de la zona a detectar i ens permet executar les proves exactament on desitgem.

El classificador és l'element principal en el procés de detecció de punts crítics. Aquest ens permet obtenir la probabilitat que un píxel d'una nova imatge formi part d'un punt crític analitzat anteriorment mitjançant les imatges de model. A la vida real la imatge de model i la imatge que s'està intentant classificar poden diferir en molts aspectes (angle, distància, lluminositat, soroll, distorsió, etc.). Per altra banda, alguns escenaris poden presentar reptes específics, com executar la detecció en temps real o l'ús d'imatges de baixa qualitat. Per entrenar bé el classificador s'utilitza una tècnica que consisteix a aplicar diverses modificacions i distorsions a les imatges model. Això permet millorar la robustesa del classificador, deixant-lo preparat per reconèixer una nova imatge que presenti canvis importants respecte la imatge model. A la nostra part experimental hem decidit aplicar modificacions d'angle, zoom i un filtre Gaussià de soroll.

A la figura 2 es mostren algunes de les imatges d'entrenament creades per entrenar el classificador en la detecció d'un ull esquerre. Es coneix com a classe el conjunt d'imatges que pertanyen a un punt crític analitzat concret.

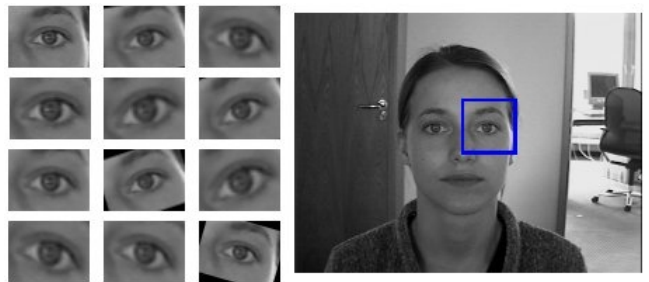
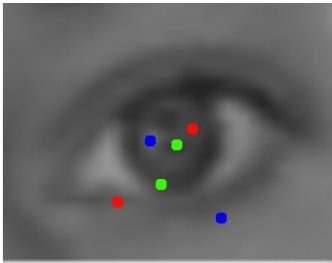


Fig. 2. Creació d'imatges d'entrenament per al punt crític del centre de l'ull esquerre. Es poden observar girs d'angle en dos sentits, zoom i l'aplicació d'un filtre de soroll Gaussià. El conjunt d'imatges formarien part de la classe "ull esquerre".

La quantitat d'imatges generades i les distorsions aplicades idealment haurien de ser seleccionades en funció de l'entorn en el qual ha de treballar l'aplicació. S'ha de destacar que existeixen mètodes específics que ens permeten crear classificadors robustos per tractar escenaris que presenten reptes concrets, com poden ser el reconeixement aplicat en imatges que presenten una alta densitat d'objectes [4][5]. Una altra possibilitat és implementar sistemes SLAM per rastrejar imatges en temps real o el reconeixement d'objectes en 3D[7][8].

Per crear el classificador a partir de l'anàlisi d'aquestes imatges model disposem de múltiples tècniques. Algunes, com els histogrames de colors o els de gradients (HOG), presenten característiques que els fan ideals per certs entorns, però són complicats de calcular i requereixen un consum més alt de recursos [9]. En el nostre cas hem seleccionat tests d'intensitat en escala de grisos; aquesta tècnica consisteix a determinar la intensitat de dos píxels seleccionats a la imatge i normalitzar-la en una escala del 0 al 255. El test binari es calcula determinant si el primer píxel  $P_1$  presenta una

intensitat més elevada que el segon píxel  $P_2$ . Aquestes proves d'intensitat són simples i ràpides, però no serien útils per la caracterització de classes si els tests no estiguessin agrupats correctament.



**Fig. 3.** Tres tests binaris d'intensitats diferents, diferenciat per colors. La prova consisteix a obtenir la intensitat de cada parella de punts i validar posteriorment si el primer píxel presenta una intensitat major a la del segon.

A la Figura 3 es mostra una imatge d'una classe i 3 tests d'intensitat. Cada punt representa un píxel on s'avalua la intensitat. A cada parella de píxels s'esbrinarà si el primer píxel té una intensitat major a la del segon. Òbviament, un sol test binari per si sol tindria una capacitat de discriminació irrisòria, així que es realitzen múltiples tests que s'agrupen en conjunts especials anomenats Ferns. Desenvolupem el concepte de Ferns a l'apartat següent.

## 2.2 Classificadors Bayesianos

Per generar un classificador disposem de varies opcions. Els classificadors Bayesianos es basen en el teorema de Bayes:

$$P(H_i|E) = \frac{P(H_i)P(E|H_i)}{\sum_{k=1}^m P(H_k)P(E|H_k)}$$

**Eq. 1.** El teorema de Bayes en la seva forma extensa.

Aquest teorema permet calcular la probabilitat d'un determinat succés a partir de les probabilitats d'altres variables. Com que no s'assumeix independència entre cap de les seves variables el teorema és poc escalable: l'increment de variables al sistema impliquen un augment exponencial als càlculs probabilístics que s'han d'executar. El problema de l'escalabilitat fa que normalment no s'utilitzin classificadors Bayesianos purs en el camp de la visió per computador.

Una opció per solucionar aquest problema és suposar que existeix algun grau d'independència entre les variables. En un classificador Bayesianu Ingenu assumiríem la independència *total* entre les variables. D'aquesta manera provoquem que la presència o absència d'una característica de la classe no afecti la presència o absència d'una altra classe diferent. Es pot dir que aquesta seria una suposició falsa en la majoria de casos; per exemple, si

coneguéssim que un píxel concret d'una imatge presenta poca intensitat, podríem argumentar que la probabilitat que els píxels que l'envolten siguin foscos s'incrementa. Així i tot, un classificador Bayesianu Ingenu considerarà que les propietats de les variables contribueixen independentment a la probabilitat que es vol calcular. A la pràctica, els resultats que obtenen els classificadors Bayesianos Ingenuos són bons i ens permeten calcular probabilitats molt més ràpidament que els models purs, però és un model millorable.

A l'equació 2 es mostra com funcionaria un classificador Bayesianu Ingenu pur, on  $c$  representa les classes i  $f$  representa els tests (per exemple, els  $N$  tests d'intensitat binaris entre dos píxels).

$$P(f_1, f_2, \dots, f_N | C = c_i) = \prod_{j=1}^N P(f_j | C = c_i)$$

**Eq. 2.** assumim independència entre les variables (model Bayesianu Ingenu) per no aplicar el teorema de Bayes pur.

En aquest treball s'utilitzarà únicament un model intermedi entre el Bayesianu pur i l'Ingenu: *un model Bayesianu Semi-ingenu*. Aquest assumeix un compromís per reduir el grau d'independència entre els tests i així poder millorar l'eficàcia obtinguda amb el model Ingenu pur. A l'equació 3 i 4 es mostra el model matemàtic.

$$P(f_1, f_2, \dots, f_N | C = c_i) = \prod_{k=1}^M P(F_k | C = c_i)$$

**Eq. 3.** Els Ferns utilitzen el model Bayesianu semi-ingenu.

$F_k$  representen les agrupacions de Ferns per cada agrupació  $S$ . Com s'ha explicat anteriorment, un Fern és una agrupació de tests. La idea és mantenir la dependència entre els tests de cada Fern, però obtenir una independència total entre els mateixos Ferns:

$$F_k = \{f_{\sigma(k,1)}, f_{\sigma(k,2)}, \dots, f_{\sigma(k,S)}\}, k = 1, \dots, M$$

**Eq. 4.** descripció de  $F_k$ .

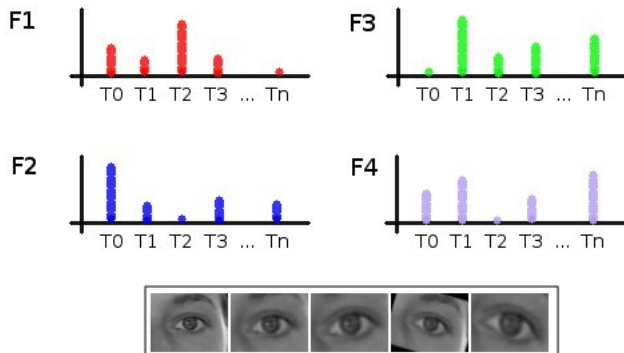
A l'hora de classificar un punt d'una imatge, es considerarà que el píxel analitzat pertany a la classe que hagi obtingut una probabilitat més elevada de totes les possibles ( $P$  més elevat a l'equació 3).

El model semi-ingenu utilitzat en aquest treball defineix  $M$  agrupacions (Ferns) de  $N$  tests d'una mida tal que  $S=N/M$ . Si els tests  $N$  son agrupats a l'atzar estarem aplicant el model Random Ferns (RF). Si en comptes de seleccionar els tests a l'atzar busquem els més útils utilitzant un criteri simplificat de la Informació Mútua estarem aplicant un model de Ferns amb Informació Mútua (FIM). En el nostre cas, quan parlem d'un test diferent estem fent referència a quines coordenades  $X$  i  $Y$

de la imatge es seleccionaran per aplicar les proves d'intensitat dels píxels.

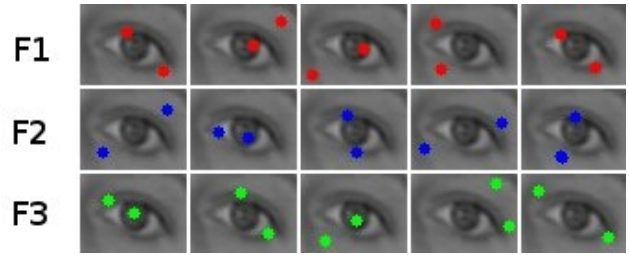
Seguint l'exemple mostrat a la Figura 3, si generem un Fern utilitzant els 3 tests de la imatge podem obtenir  $2^3$  possibles resultats binaris: 000, 001, 010, ... 111 (de 0 a 7 en decimal). Un test donarà 1 quan la intensitat del primer píxel sigui major a la del segon píxel, i un 0 en cas contrari. La idea és aplicar aquest Fern a totes les imatges generades per cada classe (com es veu a la Figura 2) i recopilar els resultats. Aplicarem tots els Ferns a totes les imatges d'entrenament generades; cada Fern resultarà en un test diferent per cada classe. Recopilant els resultats podrem obtenir la probabilitat de cada test per la zona analitzada.

A la Figura 4 es mostra aquest concepte d'una manera gràfica. A la imatge podem observar la distribució de probabilitats de quatre Ferns en aplicar-los a totes les imatges generades per la classe. Es pot observar com cada un dels possibles  $2^5$  resultats obté més o menys probabilitats d'ocórrer. A la fotografia s'identifica el test T0 com el resultat de tots els tests binaris d'intensitat a 0 (per exemple, 0000) i Tn com el resultat de tots els tests d'intensitat binaris a 1 (per exemple, 1111).



**Fig. 4.** Exemple gràfic de les probabilitats conjuntes per cada test de cada Fern al aplicar-los a totes les imatges d'un punt crític. Tn correspondrà al test  $2^5$ .

Així, aplicariem aquests Ferns a totes les imatges d'entrenament per conèixer el seu comportament (distribució de probabilitats) als tests d'intensitat. A la figura 5 es mostra un exemple de 3 Ferns de 5 tests aplicats a una de les imatges generades.



**Fig. 5:** Tres Ferns d'exemple, compostos cada un de 5 tests diferents, aplicats a una de les imatges d'entrenament. En aquest cas, tindriem  $2^5$  possibles resultats per cada Fern (00000, 00001, ... , 11111).

Un punt important a l'hora d'aplicar aquesta tècnica és la selecció dels tests que han de formar part dels Ferns. En aquest treball utilitzem dues tècniques diferents: la del Random Ferns i la dels Ferns amb Informació Mútua.

### 2.3 Classificador Random Ferns

El model Random Ferns ha demostrat presentar un compromís entre rendiment i eficàcia molt adequat en les aplicacions de visió per computador. Ha presentat un avenç comprovat en comparació amb altres estructures, com els Trees [3]. També ens proporciona una gran flexibilitat a l'hora de seleccionar les variables N, M o S, el que permet adaptar el model als recursos o necessitats de la part experimental.

L'aleatorietat és aplicada a l'hora de seleccionar la posició dels píxels  $P_1$  i  $P_2$  de cada test. D'aquesta manera els dos píxels a comparar es seleccionen totalment a l'atzar.

El principal avantatge dels Random Ferns és la gran velocitat i simplicitat a l'hora de crear els tests pel classificador [2]. La idea és que encara que alguns tests dels Fern seleccionats aleatòriament no siguin eficaços, això no importa perquè ho podem compensant creant Ferns més grans (amb més tests) o utilitzant una quantitat major de Ferns per generar el classificador[1].

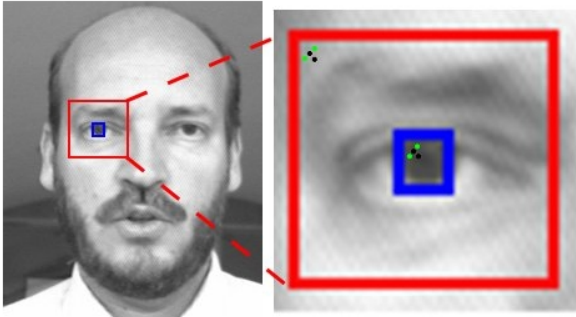
### 2.4 Classificador Ferns amb Informació Mútua

L'algorisme FIM implementat en aquest projecte elimina l'aleatorietat a l'hora de seleccionar els tests dels Ferns. Els tests són seleccionats entre tots els possibles aplicant un criteri simplificat de la informació mútua [6]. Aquest mètode necessita una fase d'aprenentatge extensa i sistemàtica. Això implica que el procés de generació del classificador és molt més lent que amb els Random Ferns. El resultat, però, és un classificador més precís, que pot identificar amb una probabilitat més alta les classes amb una quantitat menor de Ferns que el creat aplicant Random Ferns.

La idea és que alguns tests aplicats a certes classes poden ser més o menys útils per determinar si una

imatge forma part d'una classe o no. Per exemple, intentar detectar múltiples cops si un píxel presenta variacions d'intensitat en zones d'una imatge que no en presenten no ens serveix per identificar la classe. Per tant, l'objectiu és determinar quins test són millors per caracteritzar la classe que estem analitzant.

El nostre mètode consisteix a seleccionar aquells tests que resulten positius (1 al test binari) molts cops a la zona que envolta el punt crític de la classe i pocs cops fora de la zona que ens interessa identificar (Figura 6).



**Fig. 6.** Dos tests diferents d'exemple (punts verds i negres) aplicats a un píxel que pertany als punts crítics (requadre blau) i a un píxel que pertany a una zona que no ens interessa detectar (requadre vermell). Aquests dos Ferns seran aplicats a tots els píxels possibles dels dos requadres i s'emmagatzemaran els cops que els tests han donat positiu en cada cas.

Per fer-ho, delimitem dues finestres diferents. La primera i més petita envolta el punt crític que estem intentant detectar. La segona embolcalla la primera, i representa tots els píxels que no ens interessa detectar.

A la figura 6 mostrem un exemple del procés que es segueix. Per l'exemple de la imatge només es mostren 2 tests (punts negres i verds); en realitat aquest procés és aplicat a tots els tests possibles. El requadre blau identifica els píxels crítics que ens interessa detectar, i el requadre vermell identifica els píxels que no volem que siguin detectats per aquesta classe. Es generaran tots els tests possibles i aquests seran aplicats sistemàticament a tots els píxels de cada requadre. La idea és aplicar el test i emmagatzemar els cops que el test d'intensitat ha donat positiu en cada un d'ells. Quan s'hagi realitzat tots els tests possibles per totes les imatges d'entrenament, podrem determinar quin dels dos Ferns és més rellevant per la detecció dels píxels que embolcallen el punt crític central de l'ull (que formen part del requadre blau). Aplicarem una simple divisió tal que:

$$V_{fi} = \frac{N_{fi}}{M_{fi}}$$

**Eq. 5.** Obtindrem els valors  $V$  per cada test  $f_i$  a partir de la divisió entre la quantitat de cops que el Fern ha donat positiu pels punts que volem identificar ( $N$ ) i el nombre de cops que ha donat positiu en punts que no ens interessin ( $M$ ).

Aplicant la divisió que es mostra a l'equació 5 per a cada un dels test possibles obtindrem els índex  $V$  de cada test, que podríem veure com una puntuació de qualitat. Simplement seleccionarem els tests que presentin un índex  $V$  més alt i descartarem la resta.

És necessari remarcar que el procés de calcular tots els tests per crear els Ferns per cada un dels píxels de cada una de les imatges és llarg, i no té res a veure amb la rapidesa que s'obté en seleccionar-los utilitzant Random Ferns. D'aquesta manera estem sacrificant molta velocitat a l'hora de calcular el classificador amb l'objectiu d'obtenir uns Ferns més discriminants per realitzar la detecció. Aquests seran de molta més qualitat classificatòria, ja que es donen sovint quan estem propers al punt crític i es donen pocs cops quan hi estem lluny.

## 2.5 Dirchett

El classificador és calculat a la fase d'entrenament utilitzant les propietats de Dirchett. En aquesta fase l'objectiu és calcular les probabilitats de cada classe per cada Fern. La probabilitat del Fern  $k$  per la classe  $i$  podria ser la mostrada a l'equació següent:

$$p_{k,c_i} = \frac{N_{k,c_i}}{N_{c_i}}$$

**Eq. 6.** Probabilitat  $p_{k,c_i}$ .

Però l'equació anterior presenta un problema: com que calculem les probabilitats conjuntes per totes les classes multiplicant-les per  $p_{k,c_i}$ , (veure equació 3) si una mostra d'una classe no s'avalua a  $k$ , el resultat de  $p_{k,c_i}$  total és 0. Com que es multiplica  $p_{k,c_i}$  per tots els Ferns, si aquesta és 0 la imatge avaluada mai podria ser classificada a la classe  $c_i$  (al marge de les probabilitats dels altres Ferns). Aquest és un problema greu, ja que si es segueix aquest model un sol test que resulti en  $P=0$  faria impossible assignar aquesta classe al punt avaluat, al marge dels resultats de la resta de tests. El classificador no tindria marge d'error i seria poc flexible, i els resultats finals serien ineficients. Per solucionar aquest problema s'aplica la propietat de Dirchett:

$$p_{k,c_i} = \frac{N_{k,c_i} + N_r}{N_{c_i} + K \times N_r}$$

**Eq. 7.** Probabilitat  $p_{k,c_i}$  aplicant la propietat de Dirchett.

A l'equació 7,  $N_r$  és un valor arbitrari major a 0. A les proves experimentals s'ha assignat aquesta variable a 1; és un possible valor recomanat, ja que si  $N_r$  s'incrementa molt, els tests començarien a degradar resultats.  $K$  és igual a  $2^S$ , on  $S$  és la quantitat de tests per Fern [1]. Aplicant la fórmula de l'equació 7 aconseguim que les probabilitats de tots els tests sempre siguin més elevades



a 0, al marge de si el test s'ha donat o no. Evidentment, si un test no s'ha donat en un punt concret, la probabilitat resultant serà baixa (però mai nul·la). La resta de tests podrà així contribuir a augmentar o disminuir la probabilitat final del Ferns.

Per tant, el classificador calculat final és, en definitiva, una taula que mostra les probabilitats que presenta cada classe per cada possible  $2^5$  resultat de cada N Fern.

### 3 PROVES EXPERIMENTALS

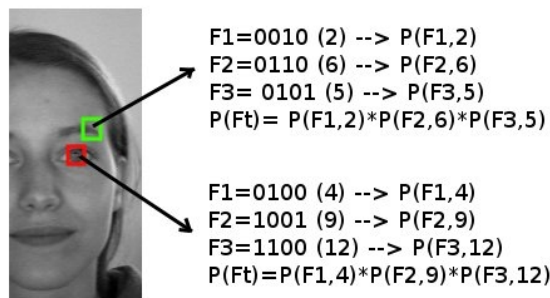
Un cop generat el classificador el podem utilitzar per executar proves en imatges. L'objectiu és detectar els punts crítics que hem entrenat durant les fases anteriors.

Les proves experimentals d'aquest article s'han realitzat utilitzant una finestra lliscant; s'ha definit una zona àmplia d'una fotografia i el classificador ha de ser capaç de determinar la probabilitat que cada un dels píxels de la imatge pertanyi a un punt crític analitzat per ell. Les proves les hem executat buscant els píxels que formen part del centre d'un ull. Si el test funciona correctament, els píxels més propers al centre de l'ull han d'obtenir una probabilitat més alta de formar part del nucli de l'ull que els que estan situats més lluny.

Per obtenir aquesta probabilitat el programa aplica cada un dels Ferns al píxel avaluat actualment i n'obté quin ha estat el resultat final per cada Fern (per exemple, Fern1: 1001, Fern2: 1110, etcètera.). Aquests resultats es contrasten al classificador: es multipliquen les probabilitats que ha obtingut el punt crític per cada test obtingut al píxel, i el resultat és un valor superior a zero. Com que s'han utilitzat propietats de Dirchett, encara que un test no s'hagi donat en cap cas durant la creació del classificador no existirà cap test amb una probabilitat 0 absoluta. S'executen els tests per tots els píxels i es normalitzen de 0 a 1. Els píxels amb probabilitats més altes haurien de ser els que formen part de la classe cercada.

A la figura 5 es mostra un exemple sobre com s'executa el procés de detecció. A la imatge només es mostren les proves en dos punts de la imatge, però tal i com hem comentat anteriorment, els tests s'executen utilitzant una finestra lliscant per sobre de tota la fotografia. S'ha de tenir present que en deteccions reals la quantitat de Tests òptims per Fern normalment és superior a 4 i es creen molts més Ferns.

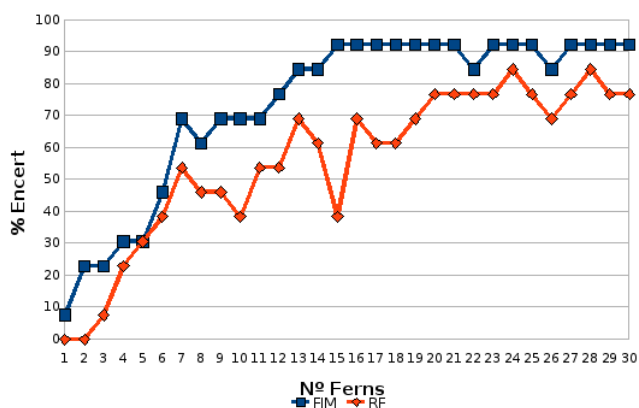
Executant aquest procés a tota la imatge obtenim les probabilitats que cada un dels píxels formi part del centre de l'ull. Els resultats es normalitzen de 0 a 1 per processar amb més facilitat les dades, i es comparen i s'observen quins punts són els que el classificador ha identificat com a pertanyents a l'ull.



**Fig. 5.** Exemple de detecció en una nova imatge. S'apliquen els Ferns a dos píxels diferents i s'obté el resultat dels tests. Posteriorment s'obtenen les probabilitats que el classificador ha guardat per cada Fern i es multipliquen. El resultat es normalitza en valors de 0 a 1. En el nostre cas, com que hem generat el classificador situat al centre de l'ull, la probabilitat  $P(Ft)$  del píxel situat al quadre vermell hauria de ser força superior a la del situat al quadre verd.

Per generar els classificadors RF i FIM hem utilitzat 3 imatges de model, creant en total 300 imatges distorsionades de l'ull dret. Per la detecció amb el classificador hem utilitzat 13 imatges del mateix model. A les nostres proves s'ha considerat encert aquells tests que han mostrat com a punt més probable del centre de l'ull ( $P=1$ ) un píxel situat a un màxim 3 píxels de distància del punt central de l'ull analitzat. Sempre que s'ha donat un empat de píxels amb probabilitat màxima (és a dir, si hi ha més d'un píxel detectat amb probabilitat màxima), s'ha considerat un encert si tots aquests punts es troben situats a un màxim de 3 píxels del centre de l'ull (aquest efecte només s'ha donat a les proves executades amb pocs Ferns, a mesura que anem afegint Ferns al sistema la capacitat de detecció del punt crític analitzat es va afinant). Per executar les proves experimentals s'han utilitzat imatges del Dataset BioID [10]. El programa implementat ha utilitzat les biblioteques OpenCV per l'anàlisi i tractament de les imatges [11].

Per explicar les conclusions sobre el nostre classificador utilitzarem els resultats obtinguts en un dels experiments aplicats a la detecció d'un ull pel model mostrat a la Figura 2. Els resultats es poden observar a la taula següent:



A partir de les dades que s'observen a la taula podem extreure algunes conclusions interessants. Primerament, els resultats obtinguts amb FIM son superiors als obtinguts amb RF. El percentatge d'encert utilitzant Ferns amb Informació Mútua es manté sempre per sobre dels obtinguts a l'atzar. Un altre detall és que FIM presenta una corba de creixement constant, on no s'aprecien baixades brusques en el percentatge d'encert. En contrast, RF es mostra més imprevisible, presentant freqüents pics i valls. Aquest és un efecte que es podia preveure, ja que amb RF els Ferns s'elegeixen a l'atzar cada cop que es genera un nou classificador. Això fa que sigui molt possible que un classificador 15 Ferns presenti un nombre d'encerts molt inferior a un classificador de 13 Ferns (tot depèn del grau de discriminació que tinguin els tests elegits a l'atzar). Aquest efecte es va moderant a mesura que anem afegint Ferns al sistema ja que anem disminuint el grau de variances al sistema.

La teoritzada reducció de Ferns necessaris quan utilitzem FIM queda palesa a l'exemple si acceptem un llinar d'encert del 70%. Aquesta eficiència a la detecció l'aconseguim de manera estable a partir del Fern número 9 en el cas de FIM; amb RF no podem assegurar aquest percentatge fins que no fem servir 19 Ferns per el classificador. I tot i això, degut a l'aleatorietat que presenta RF probablement hauríem de mantenir un llinar de seguretat més ampli. Com que FIM demostra un creixement d'eficiència estable, el llinar de seguretat que hauríem de deixar seria molt menor. Amb RF depenem de la sort a l'hora de seleccionar els tests; amb FIM eliminem el factor aleatori i obtenim un sistema més estable per la detecció. El que implica finalment un augment a la capacitat de detecció i generar classificadors amb menys Ferns (un consum de memòria menor i un procés de detecció més eficient o ràpid).

Per desgràcia, FIM no presenta millores en tots els aspectes quan el comparem amb RF. Un possible desavantatge és l'obligació de definir les dues finestres a l'hora de crear el classificador. Aquest és un pas totalment inexistent amb RF, el que implica un lleuger augment de complexitat al procés. També hem pogut constatar durant l'execució de múltiples proves experimentals que si no es defineixen unes mides adequades de les finestres, els Ferns seleccionats per generar el classificador perden força i obtenen un classificador menys potent. Per exemple, per detectar el centre d'un ull hem pogut observar que quan la finestra de punts crítics envoltava tot l'iris el classificador era més eficient que quan envoltava només la pupil·la. Una finestra amb una mida inadequada pot degradar la capacitat discriminatòria del classificador notablement.

No obstant, el desavantatge de més pes que hem de considerar és la velocitat en el procés de generació del classificador FIM. La selecció dels Ferns a l'hora de crear el classificador FIM fa que aquesta fase sigui molt més lenta que amb RF. Amb RF seleccionar els tests és trivial; amb FIM hem de provar tots els tests possibles a tots els

píxels de totes les imatges generades de model. Aquest és un procés lent, que en cap cas seria adequat per sistemes que hagin de generar classificadors en temps real.

## 4 CONCLUSIONS

En aquest article hem presentat FIM, una modificació del model Random Ferns amb l'objectiu de crear classificadors perfeccionats tant en eficiència com en eficàcia. Per fer-ho hem utilitzat un criteri simplificat de la Informació Mútua, que ens ha permès seleccionar els tests d'intensitat més discriminatòris i descartar els menys útils. S'ha demostrat a la part experimental que el sistema FIM funciona correctament, permetent generar classificadors més precisos amb una quantitat menor de Ferns. Utilitzar menys Ferns implica executar menys tests, per tant es pot incrementar la velocitat al procés de detecció.

Però el sistema presenta alguns inconvenients que hem de considerar abans de seleccionar-lo per davant d'altres models. Quan el comparem amb RF aquests desavantatges es presenten només en el procés de generació del classificador. Calcular el classificador consumeix molts més recursos de processament, i generar-lo és més complex ja que s'han de definir dues finestres: una que embolcalla els píxels que volem detectar i una altra que defineix els píxels que no volem detectar. La selecció adequada de les finestres és un aspecte clau si es vol garantir un percentatge d'encert millor que el de RF. En una frase, podem dir que FIM sacrifica senzillesa durant el procés d'entrenament per obtenir millores quantificables durant el procés de detecció.

Això implica que FIM no és adequat per aplicacions que necessitin generar o modificar el classificador en temps real d'aplicació. Però és una excel·lent elecció en escenaris que requereixin prioritzar el procés de detecció, tant en rendiment com en eficàcia.

## 5 REFERÈNCIES

- [1] Mustafa Özuysal, Michael Calonder, Vincent Lepetit i Pascal Fua. Fast Keypoint Recognition using Random Ferns.
- [2] Mustafa Özuysal, Pascal Fua, Vincent Lepetit. Fast Keypoint Recognition in Ten Lines of Code.
- [3] Anna Bosch, Andrew Zisserman, Xavier Muñoz. Image Classification using Random Forests and Ferns.
- [4] M. Villamizar, J. Andrade-Cetto, A. Sanfeliu i F. Moreno-Noguer. Bootstrapping Boosted Random Ferns for Discriminative and Efficient Object Classification.
- [5] Michael Villamizar, Francesc Moreno-Noguer, Juan

Andrade-Cetto, Alberto Sanfeliu. Detection Performance Evaluation of Boosted Random Ferns.

- [6] David J.C. MacKay. Information Theory, Inference, and Learning Algorithms.
- [7] David G. Lowe. Object recognition from local scale-invariant features.
- [8] Brian Williams, Georg Klein, Ian Reid. Real-Time SLAM Relocalisation.
- [9] Navneet Dalal and Bill Triggs. Histograms of Oriented Gradients for Human Detection.
- [10] BioID Face Database:  
<http://www.bioid.com/index.php?q=downloads/software/bioid-face-database.html>
- [11] OpenCV (Open Source Computer Vision):  
<http://opencv.org/>