

Reconeixement de les persones per la ditada

Francesc Serratosa

PID_00195432



Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-NoComercial-SenseObraDerivada (BY-NC-ND) v.3.0 Espanya de Creative Commons. Podeu copiar-los, distribuir-los i transmetre'ls públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), no en feu un ús comercial i no en feu obra derivada. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-nc-nd/3.0/es/legalcode.ca>

Índex

Introducció	5
Objectius	7
1. Dispositius de captura de la ditada	9
1.1. Sensors òptics	10
1.2. Sensors d'estat sòlid	11
1.3. Sensors d'ultrasò	12
2. Anàlisi i representació de la ditada	13
2.1. Nivell 1. Patró global de les serralades i valls	13
2.2. Nivell 2. <i>Minutiae</i> de les ditades	15
2.3. Nivell 3. Formes, porus i contorns de les ditades	17
3. Processament de la imatge de la ditada	18
3.1. Imatges de l'orientació i de la consistència local	18
3.1.1. Mètode basat en el gradient	19
3.2. Matriu de la coherència	24
3.3. Imatge segmentada	26
3.3.1. Algorisme de Bresenham	28
4. Detecció de singularitats i del nucli	30
4.1. Detecció de singularitats	30
4.2. Detecció del nucli	33
5. Millora de la imatge de la ditada	35
6. Detecció de les <i>minutiae</i>	38
6.1. Procés de binarització	38
6.2. Procés d'esqueletonització	40
6.3. Procés d'extracció de les <i>minutiae</i>	42
7. Comparació de ditades	46
7.1. Tècniques basades en la correlació d'imatges	49
7.2. Tècniques basades en la comparació de <i>minutiae</i>	50
8. Classificació de les ditades	56
8.1. Descripció formal de les classes de ditades	57
Resum	59

Activitats.....	61
Abreviatures.....	64
Bibliografia.....	65
Annex.....	66

Introducció

En aquest mòdul estudiarem la tècnica de la identificació de les persones a partir de la ditada. Tal com s'ha enfocat el desglossament de l'assignatura en mòduls, és imprescindible haver llegit i comprès abans el primer mòdul, "La biometria per a la identificació de les persones", i també el segon mòdul dedicat a l'estudi dels errors, "Avaluació dels sistemes biomètrics en aplicacions reals". En aquests primers mòduls expliquem els conceptes generals dels sistemes biomètrics que es necessiten en algunes parts d'aquest mòdul. A més, també hi ha alguns coneixements específics de la ditada, com, per exemple, els orígens, que s'ha decidit comentar-los en el primer mòdul en lloc de fer-ho en aquest perquè quedés una documentació més compacta. Així, és important remarcar que els orígens de la biometria aplicada a la ditada és un tema que s'explica en el primer mòdul.

L'ordre en què es comenten les tècniques i els processos en aquest mòdul té certa relació amb el flux de la informació en el procés general de la comparació de dues ditades o la cerca d'una ditada en una base de dades. Comentarem aquests mètodes de la manera següent:

- Primer de tot, ens trobem que hem d'escanejar el dit i generar-ne una imatge; per aquest motiu, en l'apartat "Dispositius de captura de la ditada" expliquem els diversos dispositius d'escàner de dit.
- Un cop obtinguda la imatge de la ditada, aquesta es transforma en un registre en què hi ha les característiques principals extretes; en l'apartat "Anàlisi i representació de la ditada" expliquem com es representa una ditada en tècniques que no són estrictament una imatge.
- Però per a convertir la imatge en un registre ens fa falta tot un procés de visió per computador en què calculem un seguit d'imatges intermèdies que ens serviran per a anar obtenint la informació de la representació que volem; en l'apartat "Processament de la imatge de la ditada" expliquem aquestes tècniques de visió per computador aplicades a la biometria.
- Amb aquestes imatges noves ja som capaços d'obtenir el nucli d'una ditada. El nucli representa el punt central de la ditada i és fonamental per a poder saber com s'ha ubicat el dit en el sensor i també com s'ha de deduir amb quin tipus de ditada ens trobem; en l'apartat "Detecció de singularitats i del nucli" s'explica l'obtenció del nucli i també d'altres parts de la ditada anomenades *singularitats*.
- D'altra banda, les *minutiae* són les parts de la ditada que més les caracteritzen; amb l'objectiu d'estudiar com s'extreuen, primer tenim un apartat

que torna a tractar de temes de visió per computador i està dedicat als mètodes de millora de la qualitat de la imatge de la ditada; s'explica en l'apartat "Millora de la imatge de la ditada".

- Després, ens centrem en les tècniques usades per a extreure les *minutiae*; aquestes tècniques s'expliquen en l'apartat "Detecció de les *minutiae*".
- Finalment, tenim dos apartats més dedicats estrictament a la biometria de la ditada: l'apartat "Comparació de ditades" explica com s'han de comparar dues ditades quan se n'han extret les característiques principals i l'apartat "Classificació de les ditades", com cal classificar-les.

Objectius

Els objectius bàsics d'aquest mòdul són els següents:

- 1.** Conèixer les etapes que hi ha en el procés de comparació de dues ditades o de cerca d'una ditada en una base de dades.
- 2.** Conèixer els tipus de dispositius per a la captura de la ditada.
- 3.** Conèixer les tècniques de la visió per computador aplicades a l'extracció de les característiques principals de la ditada.
- 4.** Representar una ditada en un registre d'una base de dades.
- 5.** Comparar dues ditades.
- 6.** Classificar les ditades en un nombre petit de classes.

1. Dispositius de captura de la ditada

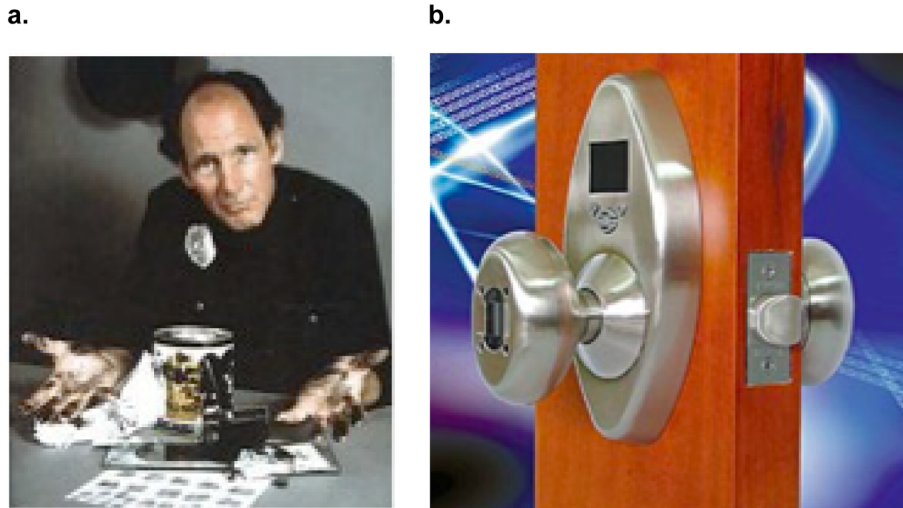
Històricament, en les aplicacions de les agències de l'ordre públic, l'adquisició de les imatges de les ditades es portava a terme amb el **mètode de la tinta**. El dit de la persona que s'havia de matricular es tacava amb tinta negra i es premia o es feia rodar en una cartolina. La cartolina s'escanejava més tard amb un escàner de propòsit general, cosa que produïa la imatge de la ditada. Aquest mètode per a adquirir les ditades s'anomena **captura fora de línia (off-line)**. Avui en dia, la majoria de les bases de dades de ditades accepten imatges *live-scan*, és a dir, que la imatge s'ha adquirit directament de la ditada amb un sensor de ditada. La tinta ja s'ha eliminat i només es requereix que l'usuari posi el dit sobre la superfície del sensor. Encara que les bases de dades s'han beneficiat d'aquest sistema ja que l'adquisició és més ràpida i les imatges de més bona qualitat, el gran beneficiari de l'adquisició *live-scan* han estat les aplicacions civils i comercials. És clar que ningú no acceptaria haver-se d'embrutar el dit cada cop que entra al gimnàs. No obstant això, sempre hi haurà sistemes fora de línia en aplicacions forenses, en què es recullen ditades d'escenes de crims.

En la figura 1 es mostra el sistema d'adquisició de la ditada amb els problemes que comporta i un dispositiu actual per a obrir una porta. El dispositiu llegeix els patrons de serralades i valls que hi ha a la superfície del dit i converteix la lectura analògica en digital per mitjà d'un convertidor analògic-digital. El mòdul de comunicació és el responsable d'enviar la imatge i també algunes ordres, com, per exemple, la finalitat de l'escaneig i de rebre l'ordre d'escanejar des d'un ordinador o dispositiu que incorpora un processador. El que aquí anomenem *sensor de ditada* és només el mòdul que llegeix la superfície del dit i genera un senyal analògic.

Vegeu també

L'estructura general d'un dispositiu de lectura de ditada es mostra en la figura 2, subapartat 1.1 d'aquest mòdul.

Figura 1. Evolució dels sistemes d'adquisició d'una ditada



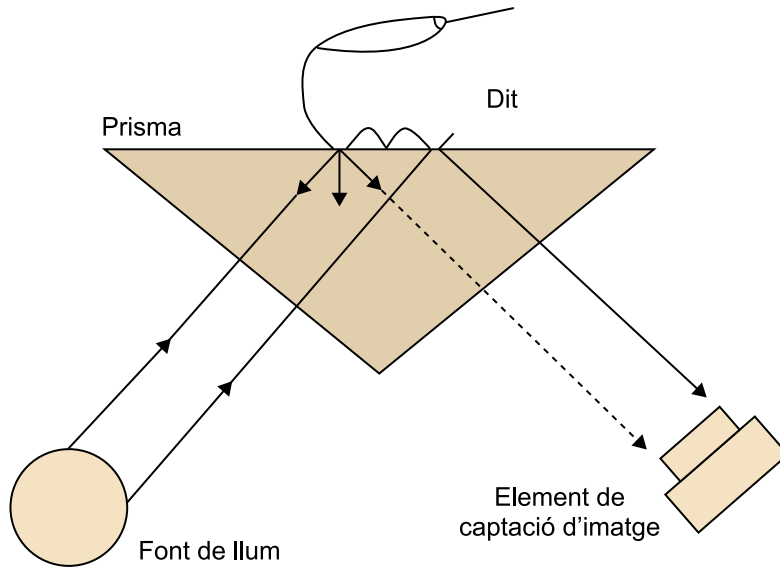
a. Sistema antic d'adquisició de les ditades amb tinta. b. Dispositiu actual per a llegir la ditada incorporat en un pany d'una porta

S'han dissenyat moltes tecnologies diferents per als sensors de les ditades que es poden agrupar en famílies. Algunes s'han arribat a comercialitzar i d'altres no han passat de prototip. Aquí comentarem les tres famílies més comunes (òptica, d'estat sòlid i ultrasò) i dins de les famílies, comentarem només la tecnologia més estesa.

1.1. Sensors òptics

Els **sensors òptics** utilitzen la llum per a adquirir la imatge. La tecnologia més comuna i la més antiga és l'anomenada *frustrated total internal reflection* (FTIR) (figura 2). Quan el dit toca la superfície de vidre o de plàstic, les serralades estan en contacte amb la superfície del prisma però les valls no. La part esquerra del prisma s'il·lumina amb una llum difusa. La llum que entra pel prisma es reflecteix a les serralades i és absorbida per les serralades. La manca de reflex permet a les serralades (que apareixen de color negre en la imatge) ser discriminades de les valls (que apareixen il·luminades). Els raigs de llum surten per la banda dreta del prisma i són captades per una lent cap al sensor d'imatge CCD. Com que els FTIR detecten la superfície tridimensional, és difícil que siguin enganyats presentant una fotografia d'una ditada. Tenen l'inconvenient que no funcionen gaire bé amb el dit moll.

Figura 2. Sensor de ditada tipus FTIR



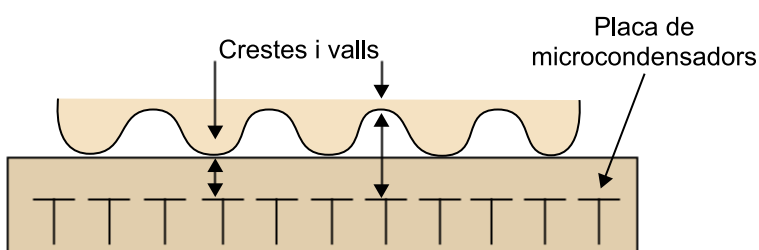
1.2. Sensors d'estat sòlid

Els **sensors d'estat sòlid** són aquests sensors, també anomenats *sensors de silici*, que ja es van proposar en diverses patents des del 1980, no obstant això, fins a mitjan 1990 no es van comercialitzar. L'objectiu d'aquests sensors és solucionar els problemes de mida i cost dels sensor òptics. Es considera que el gran desplegament de dispositius per a la captura de ditades no s'hauria portat a terme sense els sensors d'estat sòlid, ja que el preu és bastant més reduït i també la mida, llavors es poden incorporar a portàtils i dispositius de baix cost i mida. Tots els sensors d'estat sòlid consten d'una matriu de píxels en què cada píxel és un sensor petitet per ell mateix. L'usuari directament toca la superfície de silici i no hi ha components òptics ni sensors de captura de la imatge com CCD. Amb el dit moll continuen funcionant bastant bé.

Les dues tecnologies principals per a convertir el patró de la ditada en senyals elèctrics són capacitives i tèrmiques:

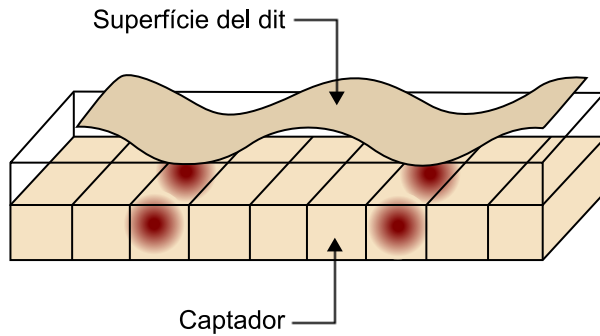
1) **Capacitiu**: aquests sensors (figura 3) detecten la diferència de càrrega elèctrica entre zones que toquen el sensor (serralades) i zones que no el toquen (valls). Són els sensors més comuns dins dels dispositius d'estat sòlid.

Figura 3. Sensor capacitiu



2) **Tèrmics:** aquests sensors (figura 4) detecten la diferència de temperatura entre serralada i vall. Estan formats per material piroelèctric. Les serralades, que estan en contacte amb el sensor, transmeten més temperatura que les valls. Com que en pocs mil·lisegons tota la superfície s'escalfa, de fet, no capten els punts calents sinó l'augment de temperatura en cada punt i en uns instants de mil·lisegons. Els afecta la temperatura ambient.

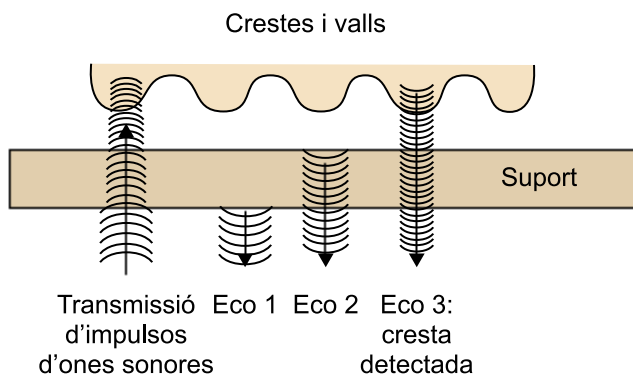
Figura 4. Sensor tèrmic



1.3. Sensors d'ultrasò

Els **sensors d'ultrasò** (figura 5) es poden veure com un tipus d'ecografia. La tecnologia es basa a enviar senyals acústics cap a la superfície del dit i capten l'eco del senyal. L'eco del senyal és el que ens aporta la informació de la profunditat en què es troba la pell, les serralades molt a prop i les valls lluny. Per tant, cada píxel del sensor està compost per dos components: un emissor de so i un micròfon. Són els dispositius que més toleren el fet que els dits estiguin bruts o molls però tenen el problema que són els més cars.

Figura 5. Sensor d'ultrasò



2. Anàlisi i representació de la ditada

La **ditada** és la reproducció de l'aspecte extern de l'epidermis dels dits. La característica més evident d'aquesta representació és un patró en el qual destaquen unes serralades i unes valls.

En una imatge d'una ditada, les serralades són fosques mentre que les valls són clares (aquesta representació és herència de la captura de la ditada basada en el mètode de la tinta). L'amplària de les serralades varia de 0,1 mil·límetres a 0,3 mil·límetres i, generalment, el període entre serralades i valls és aproximadament de mig mil·límetre. La majoria de ferides als dits, com són les cremades superficials, talls, desgast, no afecten l'estructura que hi ha sota de l'epidermis i, a més, el patró original es renova amb la mateixa estructura quan creix la pell nova.

Els detalls dels patrons de les ditades es descriuen a tres nivells diferents:

- **Nivell 1:** el flux global del patró de les serralades i valls que generen les **singularitats**.
- **Nivell 2:** les terminacions de les serralades, anomenades *minutiae*.
- **Nivell 3:** els **porus de suor**, formes petites locals o els contorns de les serralades.

Cadascun d'aquests nivells genera després uns algorismes i mètodes diferents per a la comparació i classificació de les ditades.

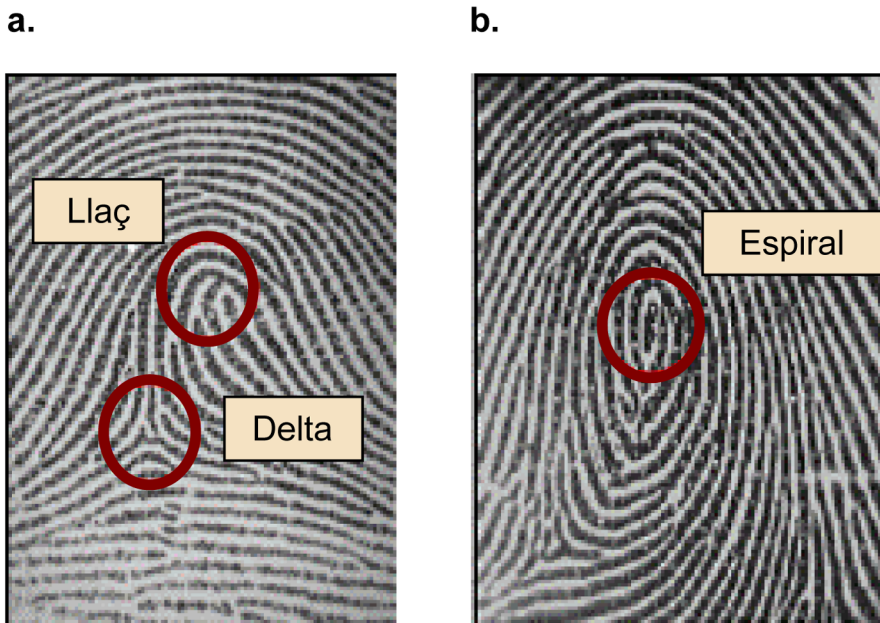
2.1. Nivell 1. Patró global de les serralades i valls

Normalment les serralades estan ubicades en una forma quasi paral·lela però hi ha unes petites regions on descriuen unes formes distintives, ja que generen una alta curvatura o terminacions freqüents de les serralades. Aquestes regions s'anomenen **singularitats** o **regions singulars** i es poden classificar, en línies generals, en tres classes: **llaç** (*loop*), **delta** (*delta*) i **espiral** (*whorl*). En general, quan es representen sobre d'una imatge d'una ditada es caracteritzen pels símbols U, Δ i O, respectivament.

En la singularitat tipus llaç, les serralades del voltant creen una forma que recorda la lletra U. D'altra banda, en les singularitats tipus delta, les serralades creen una forma semblant a un pic d'una muntanya o un triangle. Finalment, les serralades que creen una espiral tenen la forma d'un cercle. Algunes vegades les singularitats tipus espiral no es consideren, ja que se suposa que són dues singularitats tipus llaç encarades a molt poca distància. Aquesta representació

de les singularitats sense espirals té l'avantatge que no fa falta discernir quan ens trobem amb una espiral o dos llaços prou junts. En la figura 6 es mostren dues ditades amb les tres regions singulars, cadascuna de les quals amb el seu nucli.

Figura 6. Dues ditades en què apareixen singularitats



a. Llaç i delta. b. Espiral

Els algorismes per a comparar ditades (que es descriuen més endavant) poden superposar i alinear les ditades segons una marca o un punt central anomenat *nucli* (*core*).

El **nucli** serveix per a poder deduir el desplaçament en els dos eixos (vertical i horitzontal) que s'ha produït en ubicar el dit en el sensor, per a així poder alinear correctament les ditades a l'hora de fer-ne la comparació.

Henry va definir el nucli d'una ditada l'any 1900 com "el llaç superior de la imatge de la serralada més interna". A la pràctica, el nucli es considera el centre de la singularitat més superior (sigui quina sigui de les tres), ja que algunes vegades no hi ha cap llaç o espiral, com veurem més endavant. A més, i a causa de la gran variabilitat de les ditades, algunes vegades és realment difícil localitzar de manera metòdica el nucli de la ditada.

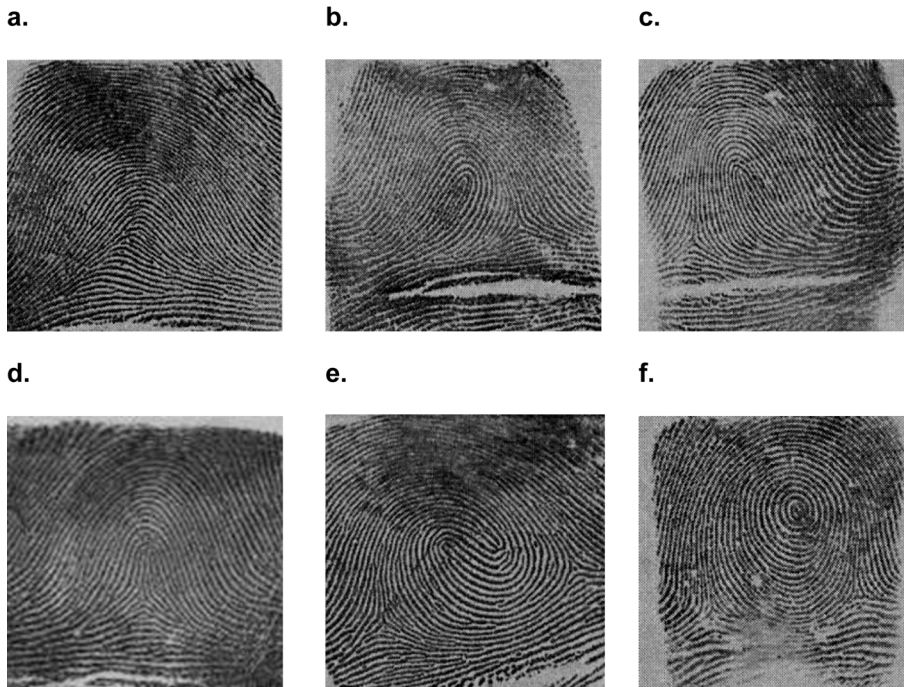
Les regions singulars, a més de ser útils per a l'alineació de les ditades, s'usen per a classificar les ditades, és a dir, donada una ditada, assignar una possible classe d'un conjunt molt reduït (normalment quatre o cinc classes). La classificació serveix per a simplificar la cerca d'una ditada en els sistemes d'identificació i accelerar l'obtenció de resultats. Això es deu al fet que, donada

Vegeu també

En l'apartat 8 d'aquest mòdul farem una descripció metòdica de les cinc classes.

una ditada, només es compara amb les ditades de la mateixa classe. La figura 7 mostra cinc imatges amb ditades de les cinc diferents classes en què es pot veure que segueixen patrons diferents.

Figura 7. Ditades de diferents classes



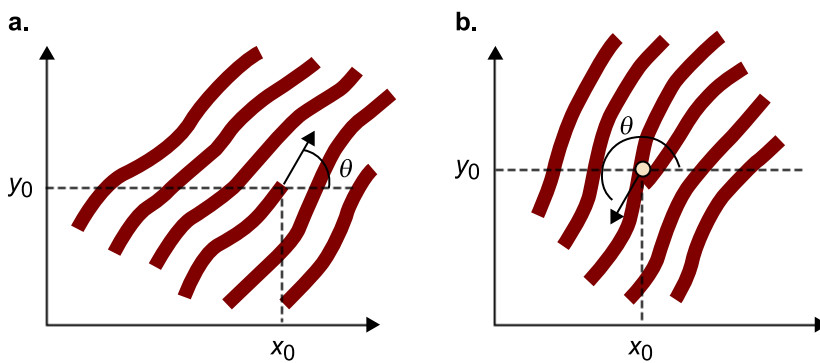
a. Arc. b. Llaç esquerre. c. Llaç dret. d. Arc tipus tenda. e. Doble llaç. f. Espiral

2.2. Nivell 2. Minutiae de les ditades

Les *minutiae* són unes altres característiques molt importants que es troben en els patrons de les ditades. En el context de les ditades, es refereixen a les diferents maneres com les serralades poden crear discontinuïtats. Per exemple, una serralada es pot acabar de sobte o es pot dividir en dues serralades. La figura 8 mostra els dos tipus de *minutiae* més comuns. També es mostra la informació de l'angle de la *minutia* i la seva posició en coordenades (x, y) dins la imatge.

Etimologia
<i>Minutia</i> vol dir 'detall petit', <i>minutiae</i> és el plural de <i>minutia</i> .

Figura 8. Dos tipus diferents de *minutiae*



a. Terminal. b. Bifurcació

La *minutia* és la característica més usada avui en dia en el reconeixement automàtic de ditades. Sir Francis Galton (1822-1911) va ser la primera persona a categoritzar les *minutiae* i va observar, l'any 1892, que es mantien incanviables al llarg de la vida d'una persona. Per aquest motiu, hi ha bibliografia que de les *minutiae* en diuen *detalls de Galton* en honor seu.

En una ditada completa, el nombre de *minutiae* pot ser superior a cent. No obstant això, s'assumeix que amb un nombre de dotze a quinze *minutiae* ja n'hi ha prou per a fer una bona comparació, si es demostra que les impressions són de la mateixa zona o l'angle d'atac del dit a la zona impresa és el mateix. Les estadístiques mostren que la densitat de les *minutiae* a les zones singulars és de $0,49 \text{ minutiae/mm}^2$ i la densitat a les zones fora de la zona singular és de $0,18 \text{ minutiae/mm}^2$.

Encara que es poden considerar molts tipus de *minutiae* (com les mostrades en la figura 8), usualment només se'n considera un grup reduït per representar la ditada, a causa de la dificultat pràctica de discernir entre classes de *minutiae* de manera automàtica. L'**American National Standard Institute** (ANSI) proposa una taxonomia de les *minutiae* basada en només quatre classes: **final de serralada, bifurcació, composta (trifurcacions o creuaments) i tipus indeterminat**. Però aquest estàndard no és el més usat i el model de *minutiae* establert per l'**FBI** considera només **final i bifurcació** (figura 8). Els altres tipus no es detecten a causa de l'alta probabilitat que siguin classificades en altres classes. Nosaltres usarem al llarg del mòdul el sistema FBI amb només dues classes de *minutiae*. En el que sí estan d'acord és en la manera de registrar les *minutiae*, que és la següent:

- **Classe:** *minutia* final o *minutia* bifurcació (en el sistema FBI). Fixeu-vos que és important que s'hagi estandarditzat que la serralada sempre genera valors foscos en la imatge (valors petits en els píxels) i que les valls generen valors clars en la imatge (valors alts en els píxels). Si no fos així, una *minutia* tipus final es podria confondre amb una *minutia* tipus bifurcació. Aquest problema s'anomena *la dualitat de les minutiae final i bifurcació*. Si generem la imatge negativa d'una ditada, llavors les *minutiae* finals es converteixen en bifurcació i viceversa.
- **Posició:** coordenada bidimensional (x, y) absoluta dins de l'àrea del sensor de la posició de la *minutia*. La posició $(0, 0)$ és el píxel inferior esquerre de la imatge generada. Aquesta dada depèn de com s'ha imprès el dit al sensor (o en una ditada latent, com s'ha fotografiat la ditada). Si el dit en la verificació s'ha imprès una mica més a l'esquerra que el dit en la matriculació, llavors el valor de l'eix horitzontal (normalment es considera l'eix x) tindrà un valor més alt en la verificació. Els algorismes de comparació "arreglen" aquestes diferències que no depenen del dit sinó de com s'ha imprès.

Referència bibliogràfica

La taxonomia de les *minutiae* proposada per l'American National Standard Institute es recull en el document ANSI/NIST-ITL 1 (2007).

Una curiositat

És curiós que l'**FBI** (institució pública dels Estats Units d'Amèrica) no faci cas de l'estandardització imposada per l'ANSI (que també és una institució pública dels Estats Units d'Amèrica).

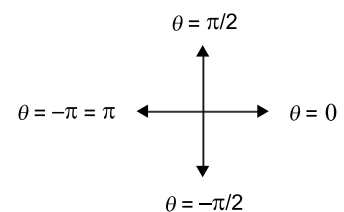


Figura 9. Orientació dels angles
Assumim que l'eix horitzontal genera l'angle 0 .

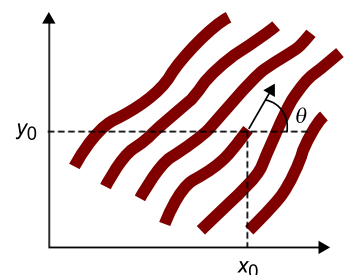


Figura 10. Exemple de *minutia* terminal
Es mostra l'angle de l'orientació de la serralada.

- **Angle:** angle generat entre la línia de la serralada i l'angle horitzontal del sistema de coordenades bidimensional. Aquest angle és de 360 graus $]-\pi, \pi]$, ja que es considera que el vector director generat per la serralada acaba a la *minutia*. De manera similar a la posició, si l'usuari orienta el dit al sensor d'una manera diferent en la matriculació que en la verificació, llavors apareixeran angles diferents per a la mateixa *minutia*. Els algorismes de comparació també "arreglen" aquestes diferències. En la figura 9 es mostren les orientacions dels angles. La figura 10 mostra una *minutia* terminal. Fixeu-vos que l'orientació de l'angle depèn del vector director i no del tipus de *minutia*.

Amb l'objectiu d'augmentar l'operativitat mitjançant els algorismes de detecció i comparació de *minutiae*, es consideren altres convencionalismes sobre l'orientació de l'angle i la posició de la *minutia*.

Referències bibliogràfiques

Altres convencionalismes sobre l'orientació de l'angle i la posició de la *minutia* es recullen en els documents ANSI/INCITS 378 (2004), ISO/IEC 19794-2 (2005) i CDEFFS (2008).

2.3. Nivell 3. Formes, porus i contorns de les ditades

Les formes de les serralades, els porus i els contorns són els detalls més fins que es poden extreure de les ditades. Inclouen tots els atributs dimensionals de les serralades com l'amplària, la forma, la llargària o el contorn. També trobem les serralades incipients o petites discontinuïtats d'aquestes serralades (que no es detecten en el nivell 2, a l'hora de deduir-ne les *minutiae*). A més, també hi ha els porus de suor, tenen una mida de 60 a 250 micres; s'ha observat que el nombre de porus en un centímetre d'una serralada varia de 9 a 18. Els especialistes reivindiquen que de 20 a 40 porus de suor pot ser suficient per a determinar la identitat d'una persona.

Encara que el nivell 3 sembla que sigui altament discriminador i molt important per als examinadors de ditades latents, actualment molt pocs mètodes per a comparar ditades l'usen, ja que són necessaris escàners de molta alta resolució (com a mínim 1.000 ppp) i imatges de molt bona qualitat. L'estàndard mencionat anteriorment, CDEFFS (2008), és l'últim esforç per a tirar endavant tecnologies amb aquest nivell. La figura 11 mostra una ditada en què són evidents els porus.



Figura 11. Imatge d'un detall d'una ditada on es poden veure els porus de suor. La imatge ha estat presa a 1.000 ppp.

3. Processament de la imatge de la ditada

En la resta d'aquest mòdul, anomenarem la imatge original que conté la ditada **I**. $I(x, y)$ representa el valor del píxel ubicat en la posició (x, y) de la imatge de la ditada **I**. Encara que per la metodologia i els algorismes comentats en aquest apartat és irrellevant, normalment, la posició $(0, 0)$ de la imatge es troba ubicada en el píxel inferior esquerre de la imatge.

3.1. Imatges de l'orientació i de la consistència local

L'orientació local de la serralada al píxel (x, y) de la imatge de la ditada **I** és l'angle θ_{xy} que forma la serralada que hi ha en un entorn local a aquesta posició amb l'eix horitzontal. Una posició concreta de la imatge de l'orientació es descriu amb el símbol $D(x, y)$:

$$D(x, y) = \theta_{xy}$$

El primer investigador a introduir aquesta imatge va ser **Grasselli** l'any 1969. Normalment, s'usa el terme **orientació** per a denominar un angle dins del rang $]0, \pi]$, ja que la serralada no té una direcció específica (usem només dos quadrants en lloc de quatre quadrants).

En cas que en la posició (x, y) no hi hagi serralada, $D(x, y)$ també pren un valor, encara que no tingui gaire significat real. En el paràgraf següent comentarem com es pot solucionar aquest problema. A més, fixeu-vos que en la caracterització d'una *minutia*, tenim la informació de la direcció que és de 360 graus, $]-\pi, \pi]$ (aquí usem els quatre quadrants). És a dir, que en la *minutia* hem de concretar la direcció.

Associada a l'orientació de la serralada θ_{xy} tenim la consistència r_{xy} .

La **consistència** representa fins a quin punt ens refiem de la informació de l'angle o com de vàlid és el valor de l'angle.

El valor de l'angle sempre hi és (és calculable), però a zones en què no hi ha serralades aquest valor no té sentit ja que no es pot parlar de l'orientació de la serralada si aquesta no existeix. La consistència pren valors no negatius. També ens fa falta calcular la matriu de la consistència \mathbf{r} , que té exactament les mateixes dimensions que la matriu d'orientació, i a cada píxel (x, y) hi trobem el valor $r(x, y)$ que representa la consistència de l'angle $D(x, y)$. En regions de la imatge **I** en què hi ha serralada però són molt sorolloses o seriosament corrompudes, \mathbf{r} pren valors positius baixos. D'altra banda, en les regions de bona

Vegeu també

Per a més informació vegeu el subapartat 7.2 d'aquest mòdul.

qualitat amb serralades, els valors de la consistència són alts. La consistència també pot servir per a saber on s'ha escanejat el dit dins de l'àrea del sensor. On no hi ha dit no hi ha serralades i llavors els valors de la consistència són baixos.

Cada valor de $D(x, y)$ i de $r(x, y)$ es calcula com la mitjana de l'orientació i la consistència de la imatge I en un entorn local quadrat de $n \times n$ píxels centrat en el píxel (x, y) . En els píxels propers als extrems de la imatge es redueix l'entorn en què es fa la mitjana. Hi ha diverses tècniques per a calcular aquestes dues matrius. Nosaltres ens centrarem en la més usada, que està basada en el **gradient** de la imatge.

Altres mètodes de càlcul

Altres mètodes per al càlcul de les dues matrius poden ser el mètode basat en les projeccions o el mètode basat en el domini freqüencial, que no es descriuen en aquest material.

3.1.1. Mètode basat en el gradient

El mètode més simple i natural d'extreure la informació de l'orientació i la consistència d'una ditada es basa a calcular primer el gradient de la imatge i després obtenir les imatges d'orientació i consistència per mitjà del gradient. El gradient $\Delta(x, y)$ en el punt (x, y) de I és un vector bidimensional tal que:

$$\Delta(x, y) = [\Delta_x(x, y), \Delta_y(x, y)]$$

en què els components Δ_x i Δ_y són les derivades horitzontal i vertical de I en el punt (x, y) . Primer de tot, hem d'explicar la tècnica usada en visió per computador per a extreure les imatges del gradient horitzontal $\Delta_x(x, y)$ i el gradient vertical $\Delta_y(x, y)$. Després, descriurem quina relació hi ha entre el gradient de la imatge i l'orientació de la serralada. Finalment, explicarem la tècnica per a obtenir la imatge d'orientació i la imatge de consistència.

Obtenció de la imatge del gradient horitzontal Δ_x i vertical Δ_y

Per a calcular aquests gradients usarem el conegut **mètode de Sobel**. L'algorisme de Sobel, presentat l'any 1968, es basa a convolucionar unes màscares per tota la imatge. En aquest cas, una màscara W és una imatge quadrada de 3×3 píxels que conté els pesos W_0, \dots, W_8 ubicats com es mostra en la figura 12.

Figura 12. Màscara de convolució genèrica amb la ubicació dels pesos

W_0	W_1	W_2
W_3	W_4	W_5
W_6	W_7	W_8

La convolució és una funció c que per a cada píxel (x, y) de la imatge I obté el valor següent:

$$c(x, y, \mathbf{W}) = W_0 \cdot I[x - 1, y + 1] + W_1 \cdot I[x, y + 1] + W_2 \cdot I[x + 1, y + 1] + W_3 \cdot I[x - 1, y] + W_4 \cdot I[x, y] + W_5 \cdot I[x, y + 1] + W_6 \cdot I[x - 1, y - 1] + W_7 \cdot I[x, y - 1] + W_8 \cdot I[x, y - 1].$$

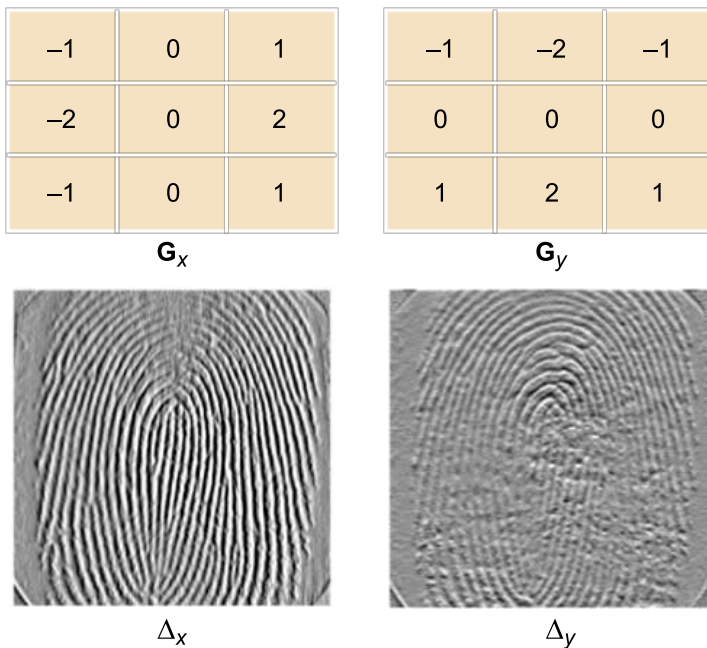
Diem que C és la imatge convolucionada de I i segons una màscara concreta \mathbf{W} obtenim:

$$c(x, y, \mathbf{W}) = C[x, y].$$

Fixeu-vos que els valors de les imatges resultants de la convolució C amb qualsevol tipus de màscara no han de ser dins del rang $[0, 255]$ i, per això, aquestes imatges no es poden visualitzar directament sense una normalització.

La figura 13 mostra les màscares de 3×3 $\mathbf{W} = \mathbf{G}_x$ i $\mathbf{W} = \mathbf{G}_y$ usades per a obtenir la imatge del gradient horitzontal Δ_x i la imatge del gradient vertical Δ_y , juntament amb aquestes imatges normalitzades. Fixeu-vos que en la imatge Δ_x es detecten les serres verticals, en canvi, en la imatge Δ_y es detecten les serres horitzontals.

Figura 13. Màscares per a obtenir el gradient horitzontal i vertical i imatges resultants normalitzades



Vegeu també

En l'annex d'aquest mòdul s'explica el procés de normalització d'imatges.

Relació entre el gradient, orientació i consistència

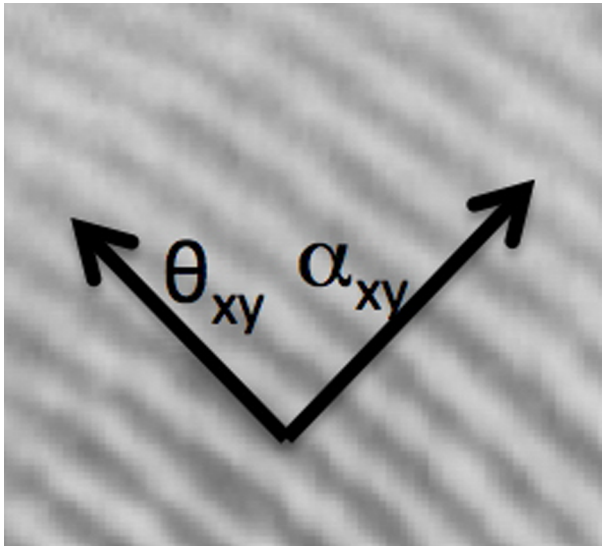
És ben sabut que l'angle del vector gradient d'una funció denota la direcció on hi ha el màxim canvi de valor de la funció i la magnitud del vector gradient denota la quantitat de canvi del valor de la funció. Per tant, l'angle en un punt (x, y) del vector gradient:

$$\Delta(x, y) = [\Delta_x(x, y), \Delta_y(x, y)]$$

de la imatge I mostra on hi ha el canvi d'intensitat màxim de la imatge I en aquest punt. Per tant, la magnitud del vector gradient mostra si hi ha molt canvi en els valors dels píxels en aquesta orientació.

On hi ha el canvi d'intensitat màxim en la imatge de la ditada? Donada una ditada, el canvi d'intensitat màxim apareix en els costats de les serralades. Els píxels de la serralada tenen un valor baix (gris que s'assembla a negre) i els píxels de les valls tenen un valor alt (gris molt clar, que s'assembla a blanc). Per tant, el canvi d'intensitat màxim apareix en els contorns o costats de les serralades. L'orientació θ_{xy} d'una serralada hipotètica que creua la regió centrada en el punt (x, y) és ortogonal a l'angle en què es genera la intensitat màxima, és a dir, és ortogonal a l'angle del gradient en el punt (x, y) . La consistència r_{xy} denota la magnitud (o també podríem dir quantitat) de la variació. La figura 14 mostra una serralada, l'angle del gradient α_{xy} en el punt (x, y) i l'orientació de la serralada θ_{xy} .

Figura 14. Angle del gradient α_{xy} i orientació de la serralada θ_{xy}



Amb les descripcions anteriors, la imatge d'orientació es defineix de la manera següent:

$$\alpha_{xy} = \arctan\left(\frac{\Delta_y(x, y)}{\Delta_x(x, y)}\right) \quad (1)$$

en què el rang de la funció és $]-\pi/2, \pi/2]$. En llenguatge de programació C o Java, aquesta operació es porta a terme amb la funció:

$$\alpha_{xy} = \text{atan2}(\Delta_y(x, y), \Delta_x(x, y)) \quad (2)$$

Així, es manté el signe del valor de les derivades vertical i horitzontal. Com que l'orientació de la serralada és sempre ortogonal a l'angle del gradient, assumim que:

$$\mathbf{D}(x, y) = \alpha_{xy} + \frac{\pi}{2} \quad (3)$$

Per tant, com s'ha comentat anteriorment, el rang dels valors de la matriu \mathbf{D} és $]0, \pi]$.

La imatge de consistència es defineix com la norma del vector gradient de la manera següent:

$$\mathbf{r}(x, y) = \sqrt{\Delta_x(x, y)^2 + \Delta_y(x, y)^2} \quad (4)$$

en què el rang de la funció és $[0, \infty[$.

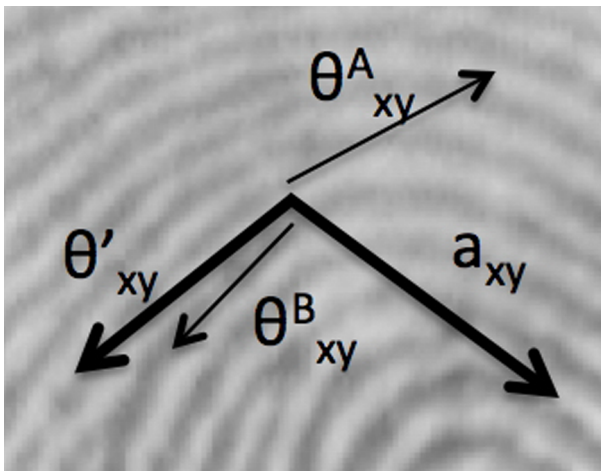
Obtenció de les imatges d'orientació i consistència

A l'hora de deduir l'orientació de les serralades mitjançant el gradient, no sols ens podem centrar en l'angle del gradient en un píxel concret ja que estimaríem l'orientació a una escala massa fina i seria molt sensible al soroll de la imatge de la ditada i a la mateixa resolució del sensor. Aquest és el motiu pel qual ja hem dit que calcularem el valor de l'orientació donada una finestra local (una zona entorn del píxel que cal analitzar). D'altra banda, no podem simplement calcular la mitjana dels angles dins d'aquesta finestra. L'orientació mitjana de dues serralades amb orientacions 5 i 175° no és 90° , com resultaria de portar a terme una mitjana aritmètica, sinó 0° . A més a més, el concepte de la *mitjana de l'orientació* no està sempre ben definit. Considerem dues serralades amb orientacions 0° i 90° . L'orientació mitjana pot ser de 45° o de 135° . En la figura 15 es mostren dues serralades, el valor de la mitjana d'orientació angular, θ'_{xy} , i el valor de la mitjana aritmètica, a_{xy} .

Reflexió

Com que el valor a_{xy} no és útil en aquests mètodes, no en tornarem a parlar.

Figura 15. Dues serralades amb els seus angles d'orientació θ^A_{xy} i θ^B_{xy}



També s'hi mostren la mitjana d'orientació angular, θ'_{xy} , i la mitjana d'orientació aritmètica, a_{xy}

A continuació, es descriuen dos mètodes per a calcular la mitjana d'orientació angular. Amb aquestes mitjanes s'obté la **imatge orientació filtrada** \hat{D} . Els algorismes de detecció de la ditada, comparació i classificació es basen en aquesta imatge nova i la imatge no filtrada ja no s'usa més, un cop s'ha obtingut la filtrada.

El primer mètode va ser presentat per **Kass i Witkin** l'any 1987. La idea bàsica és duplicar els angles, així l'estimador de l'orientació es defineix pel vector següent:

$$\hat{\Delta} = [\mathbf{r} \cdot \cos(2\theta), \mathbf{r} \cdot \sin(2\theta)]$$

Fixeu-vos que:

$$\Delta_x = \mathbf{r} \cdot \cos(\theta) \text{ i } \Delta_y = \mathbf{r} \cdot \sin(\theta)$$

Per tant, 2θ s'usa en lloc de θ per no haver de tenir en compte la circularitat dels angles. D'aquesta manera, es pot obtenir un vector mitjana tal que:

$$\hat{\Delta}(x, y) = [\hat{\Delta}_x(x, y), \hat{\Delta}_y(x, y)] \quad (5)$$

simplement amb la mitjana aritmètica en un entorn E de la manera següent:

$$\hat{\Delta}_x(x, y) = \frac{1}{n^2} \sum_{(x', y') \in E(x, y)} r \cdot \cos(2D(x', y')) \quad (6)$$

$$\hat{\Delta}_y(x, y) = \frac{1}{n^2} \sum_{(x', y') \in E(x, y)} r \cdot \sin(2D(x', y')) \quad (7)$$

L'entorn E representa tots els píxels de D dins d'una finestra local de $n \times n$ píxels. Normalment, $n = 11$ però aquest valor depèn de la densitat de píxels del sensor. Per obtenir finalment la matriu d'orientació filtrada \hat{D} , simplement aplicarem les fórmules exposades anteriorment però considerant que l'angle ha de ser la meitat:

$$\hat{D}(x, y) = \frac{1}{2} \arctan\left(\frac{\hat{\Delta}_y}{\hat{\Delta}_x}\right) \quad (8)$$

El segon mètode va ser presentat per **Ratha, Chen i Jain** l'any 1995. L'avaluació pràctica mostra que és molt més eficaç que el mètode presentat per Kass l'any 1987. A més, en aquest nou mètode eviten haver de calcular les matrius inicials d'orientació D i consistència \mathbf{r} ja que treballa directament amb les matrius dels gradients horitzontal i vertical:

$$\Delta(x, y) = [\Delta_x(x, y), \Delta_y(x, y)]$$

A la pràctica, els resultats són bastant semblants respecte del mètode de Kass i Witkin. Primer, s'han de calcular els valors que es podrien interpretar com la covariància i variància entre els gradients verticals i horitzontals en un entorn E de $n \times n$ píxels i centrat en el píxel que estem tractant (x, y) . La proposta inicial de n és 9.

$$G_{xy}(x, y) = \sum_{(x',y') \in E(x,y)} \Delta_x(x', y') \cdot \Delta_y(x', y') \quad (9)$$

$$G_{xx}(x, y) = \sum_{(x',y') \in E(x,y)} \Delta_x(x', y')^2 \quad (10)$$

$$G_{yy}(x, y) = \sum_{(x',y') \in E(x,y)} \Delta_y(x', y')^2 \quad (11)$$

Després, es calcula el valor final de la mitjana de l'angle:

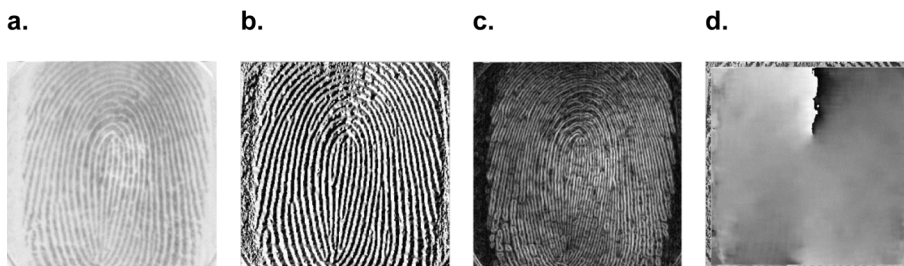
$$\hat{D}(x, y) = \frac{\pi}{2} + \frac{1}{2} \arctan\left(\frac{2G_{xy}}{G_{xx} - G_{yy}}\right) \quad (12)$$

La figura 16 mostra una imatge d'una ditada, la imatge que representa la matriu d'orientació original, la imatge que representa la matriu de consistència i la matriu d'orientació filtrada amb el mètode de Ratha. Totes aquestes imatges han estat normalitzades perquè els píxels prenguin un valor de 0 a 255 i es puguin visualitzar. El procés de normalització és diferent per a una imatge els píxels de la qual tenen valors numèrics positius i negatius, valors numèrics només positius o angles en radians.

Vegeu també

El procés de normalització d'imatges s'explica en l'annex d'aquest mòdul.

Figura 16. Imatges real, d'orientació, de consistència i filtrada



a. Imatge que representa una ditada. b. La matriu d'orientació D. c. La matriu de consistència r. d. La matriu d'orientació filtrada \hat{D}

3.2. Matriu de la coherència

La matriu de la coherència (C) d'una ditada ens informa per a cada píxel del grau en què les orientacions són paral·leles a elles mateixes en un entorn d'una finestra. Aquesta matriu ens pot servir per a veure si som en una zona on hi ha

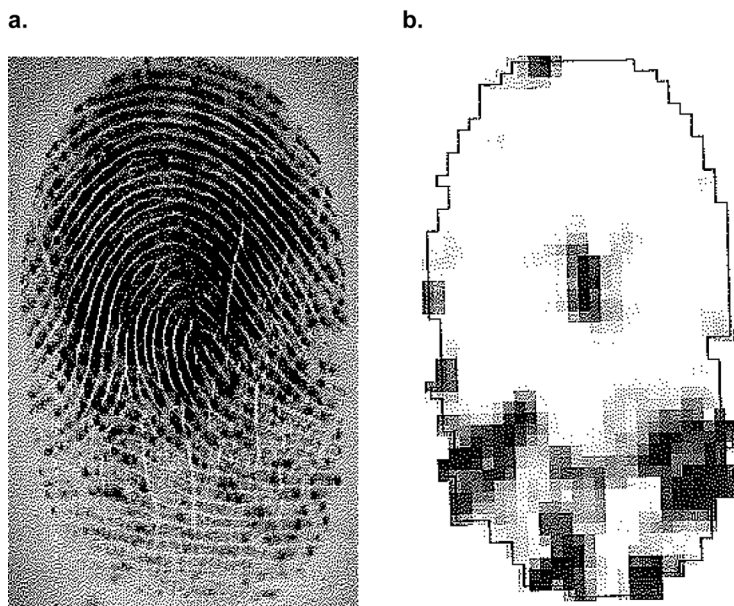
dit i també per a detectar singularitats, ja que canvis bruscos en les orientacions no passen en zones en què hi ha serralades i valls correctament marcades i paral·leles.

Kass i Witkin l'any 1987, van definir la *coherència* C com la norma de la suma dels vectors d'orientació \hat{D} en un entorn E i normalitzada de la manera següent:

$$C(x, y) = \frac{\left| \sum_{(x',y') \in E(x,y)} \hat{D}(x', y') \right|}{\sum_{(x',y') \in E(x,y)} |\hat{D}(x', y')|} \quad (13)$$

El rang de valors de $C(x, y)$ és $[0, 1]$. S'aconsegueix el valor màxim d'un quan totes les orientacions són paral·leles a elles mateixes i el valor 0 quan apunten en direccions diferents. La figura 17 mostra una ditada i la imatge de coherència corresponent. Fixeu-vos que la coherència disminueix (valors foscos) en les zones en què la ditada és més sorollosa.

Figura 17. Imatges real i de coherència



a. Imatge que representa una ditada. b. Matriu de coherència C

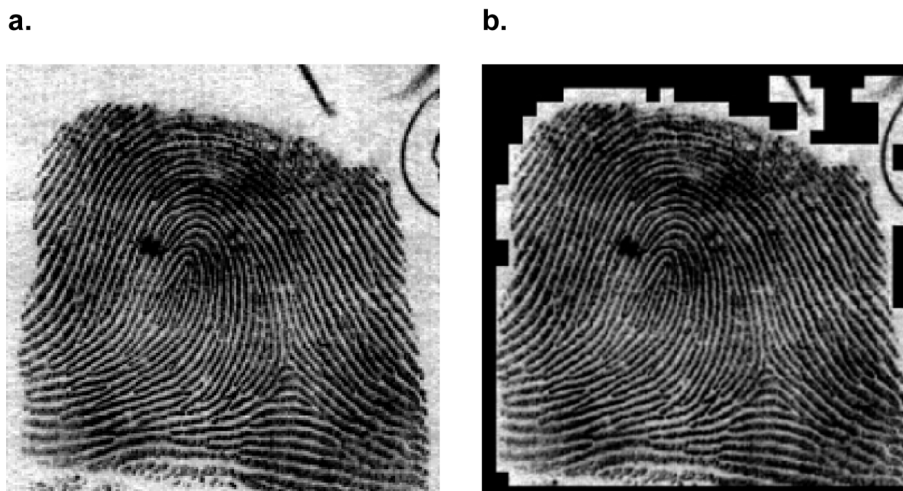
Si s'usa el mètode de Ratha per a obtenir la imatge d'orientació filtrada, llavors s'ha demostrat que l'equació anterior es pot simplificar per la següent:

$$C(x, y) = \frac{\sqrt{(G_{xx}(x, y) - G_{yy}(x, y))^2 + 4G_{xy}(x, y)^2}}{G_{xx}(x, y) + G_{yy}(x, y)} \quad (14)$$

3.3. Imatge segmentada

El terme **segmentació** és generalment usat per a denotar la separació de la zona de la ditada (també anomenat *primer pla*) respecte de la zona en què no hi ha ditada (també anomenat *fons*). Una imatge segmentada **S** és una imatge que només conté dos nivells de color. El negre representa el fons i el blanc el primer pla. La figura 18 mostra una ditada i la seva imatge segmentada superposada a l'original. Fixeu-vos que hi ha dues ratlles en el paper en què s'ha imprès la ditada. En aquestes zones, la imatge segmentada dedueix que hi ha dit. Per solucionar aquest problema, se n'ha de dur a terme un tractament posterior i allisar la forma que genera la imatge segmentada.

Figura 18. Imatges real i segmentada



a. Imatge que representa una ditada. b. La mateixa imatge del punt a superposada a la imatge segmentada. Les zones en què no hi ha ditada s'han marcat en negre.

La separació del fons és important per a evitar extreure característiques de la ditada com són les *minutiae* o singularitats que segur són errònies. A més, permet accelerar els algorismes d'extracció d'aquestes característiques, ja que només s'apliquen en els píxels continguts en el primer pla.

Com que les imatges de les ditades són patrons ratllats en què s'alternen nivells alts dels píxels amb nivells baixos, no és vàlid usar un algorisme de binarització amb un llindar global. Aquests algorismes de llindar global simplement posen en negre els píxels amb valor inferior al llindar i en blanc els píxels amb valor superior o igual al llindar. De fet, el que sí discrimina entre ditada i fons és la presència de patrons amb orientació definida. En la ditada, hi ha una orientació dominant en els píxels en una finestra local. D'altra banda, en el fons, el patró és isotròpic (no hi ha orientació dominant).

S'han publicat diversos mètodes per obtenir la imatge segmentada. En aquest mòdul descriurem un mètode senzill d'entendre i bastant eficaç.

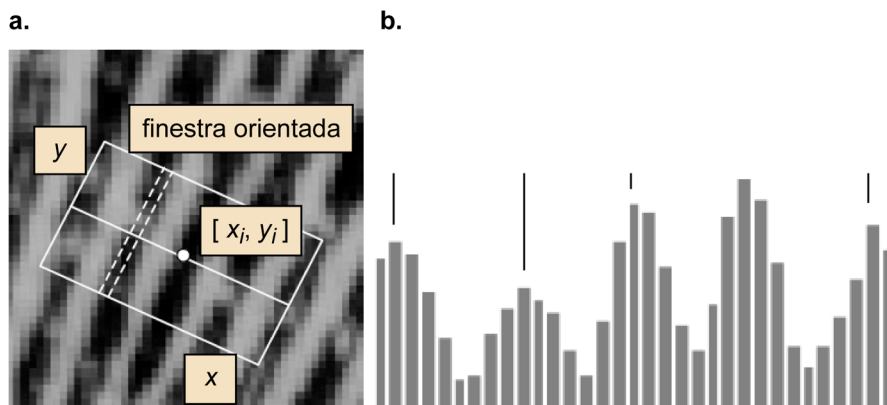
Ratha, Chen i Jain, l'any 1995, aïllen els píxels de la zona de la ditada segons la variància dels valors de la imatge original I en un segment local i unidimensional i orientat segons l'angle ortogonal a l'orientació de les serralades obtingut de la imatge de l'orientació \hat{D} .

Primer de tot, s'ha de definir què és una *signatura*.

Una **signatura** de la imatge original I s'identifica com a $f(x, y, n, \alpha)$ i és un segment recte unidimensional centrat en el píxel (x, y) i que té n elements i està orientat segons l'angle α .

El valor de l'element i -èsim de la signatura es representa per f_i . En la figura 19 es mostra a l'esquerra una imatge amb una ditada en què es destaca la signatura amb una línia blanca. A la dreta els valors de la signatura i la relació entre les posicions bidimensionals (x, y) de la imatge i les unidimensionals de la signatura.

Figura 19. Imatge real i signatura



a. Imatge que representa una ditada. b. Una signatura

Normalment, n pren un valor senar perquè (x, y) estigui centrat al mig de la signatura. Si no és senar, simplement queda desplaçat un element però aquest fet no és problemàtic per a l'obtenció de la segmentació. D'altra banda, si el píxel central (x, y) es troba ubicat a prop d'un marge de la imatge, llavors pot ser que n s'hagi de reduir perquè la signatura no quedi fora de la imatge. Finalment, a l'hora de decidir els n píxels que componen la signatura, s'han de fer aproximacions, ja que és usual que no s'obtinguin valors naturals en les posicions (x, y) quan s'aplica l'angle α . Un dels mètodes més usats per a decidir per quins píxels d'una imatge passa una recta amb angle α és el clàssic **mètode de Bresenham**.

Donada una imatge I , la seva imatge d'orientació filtrada corresponent \hat{D} , un valor de n (normalment pren valors prop de $n = 15$) i un llindar t , el mètode de Ratha calcula per a cada píxel (x, y) si $S(x, y)$ és fons (negre o valor 0) o és ditada (blanc o valor 255) com s'exposa a continuació.

Vegeu també

El mètode de Bresenham el descrivim en el subapartat 3.3.1.

Per a tot (x, y) calculeu:

$$f\left(x, y, \hat{D}(x, y) - \frac{\pi}{2}\right) \quad (15)$$

El terme $-\frac{\pi}{2}$ serveix perquè la signatura sigui ortogonal a la serralada. Volem que l'angle de la signatura sigui ortogonal a l'orientació de la serralada, per això:

$$\alpha = \hat{D}(x, y) - \frac{\pi}{2} \quad (16)$$

i llavors:

$$\alpha = \left] -\frac{\pi}{2}, \frac{\pi}{2} \right] \quad (17)$$

Un cop obtinguda la signatura f , en calcula la variància $v(f)$. Si $v(f) < t$, llavors hi ha poca variació en els valors dels píxels i per tant es considera que és fons. Altrament, es considera que és ditada ja que hi ha un patró amb molta variació.

La variància es calcula amb la fórmula següent:

$$v(f) = \frac{1}{n} \sum_{i=1}^n (f_i - \mu(f))^2 \quad (18)$$

en què $\mu(f)$ representa la mitjana de la signatura, és a dir:

$$\mu(f) = \frac{1}{n} \sum_{i=1}^n f_i \quad (19)$$

3.3.1. Algorisme de Bresenham

L'**algorisme de Bresenham**, de 1962, determina quins píxels se seleccionen en una imatge per dibuixar una recta entre dos píxels determinats (x_0, y_0) i (x_1, y_1) . Suposem que la recta que s'ha de dibuixar no és vertical, és a dir, que $x_1 \neq x_0$, ja que dibuixar una recta vertical és trivial. A més, en l'algorisme que mostrem suposem que ens trobem amb el cas que $x_0 \leq x_1$ i $y_0 \leq y_1$. Els altres casos són molt similars.

És un dels primers algorismes dissenyats en el camp de les gràfiques per computador. S'han definit nous algorismes per dibuixar rectes i també s'ha generalitzat per a dibuixar circumferències. A més, algunes plaques gràfiques el tenen implementat en maquinari.

```
algorisme Bresenham(x0, x1, y0, y1)
```

```
Delta :=  $\left| \frac{Y_1 - Y_0}{X_1 - X_0} \right|$ 
```

```
Error := 0
```

```
y := y0
```

```
per a x des de x0 a x1
```

```
selecciona el píxel (x,y)
```

```
Error := Error + Delta
```

```
si Error ≥ 0.5 llavors
```

```
    y := y + 1
```

```
    Error := Error - 1.0
```

```
fi per
```

```
fi algorisme
```

4. Detecció de singularitats i del nucli

La majoria dels mètodes presentats en la bibliografia per detectar singularitats es basen en la **matriu d'orientació normalitzada** o la usen \hat{D} . En aquesta assignatura farem servir el mètode més usat que, partint d'aquesta matriu d'orientació, calcula l'**índex de Pointcaré**. Altres mètodes exploren les característiques locals de la matriu d'orientació. Un cop localitzats els punts singulars, deduirem el nucli de la ditada.

4.1. Detecció de singularitats

Kawagoe i Tojo van presentar l'any 1984 un mètode basat en l'índex de Pointcaré.

L'**índex de Pointcaré** en un punt es defineix com la integral dels vectors d'un camp de vectors mitjançant una corba tancada que envolta aquest punt. Serveix per a saber el vector resultant en aquest punt.

Suposem que G és un camp discret de vectors associat a una imatge d'orientacions \hat{D} . Haurem de tenir en compte que \hat{D} no és realment un camp de vectors ja que conté les orientacions de les serralades dins del rang $]0, \pi]$ però no les direccions dins del rang $]-\pi, \pi]$. Per aquest motiu, haurem d'estimar quina és la direcció dels vectors tenint-ne en compte l'orientació. L'índex de Pointcaré en el punt (x, y) de la imatge \hat{D} es calcula com s'exposa a continuació.

La corba que conté els vectors està formada pels vuit píxels més propers al punt (x, y) que anomenarem v_0, v_1, \dots, v_7 . És indiferent quin píxel veí es considera que comença la corba, és a dir, quin veí és v_0 . Direm que θ_k és l'angle que correspon al valor del píxel veí v_k a \hat{D} (recordeu que el rang de valors és de $]0, \pi]$) i α_k és l'estimació que ara explicarem de la seva direcció (el rang de valors de la direcció estimada és $]-\pi, \pi]$). Llavors l'índex de Pointcaré es calcula de la manera següent:

$$P(x, y) = \sum_{k=0..7} \alpha_k \quad (20)$$

Ara toca estimar les direccions dels veïns. Considerem que:

$$\alpha_0 = \theta_0$$

Llavors, la resta de valors de la direcció $\alpha_1, \dots, \alpha_7$ dels veïns v_1, \dots, v_7 , segons la direcció de les agulles del rellotge, es calcula segons indica una de les opcions següents:

$$\alpha_k = \begin{cases} \theta_k - \theta_{k-1} - \pi & \text{si } (\theta_k - \theta_{k-1}) > \frac{\pi}{2} \\ \theta_k - \theta_{k-1} + \pi & \text{si } (\theta_k - \theta_{k-1}) < -\frac{\pi}{2} \\ \theta_k - \theta_{k-1} & \text{altrament} \end{cases} \quad (21)$$

No és difícil demostrar que l'índex de Pointcaré en un punt i calculat de la manera descrita només pot adoptar cinc valors diferents. Segons aquests valors, detectarem si ens trobem en una singularitat i quin tipus de singularitat:

$$P(x, y) = \begin{cases} 0 & \text{si } (x, y) \text{ no pertany a cap regió singular} \\ 2\pi \text{ o } -2\pi & \text{si } (x, y) \text{ pertany a una espiral} \\ \pi & \text{si } (x, y) \text{ pertany a un llaç} \\ -\pi & \text{si } (x, y) \text{ pertany a un delta} \end{cases} \quad (22)$$

En la figura 20 es mostren quatre exemples amb els quatre possibles valors del Pointcaré. Perquè sigui més fàcil d'interpretar, els valors dels angles es descriuen amb graus $[0, 180]$. Llavors, si hi apliquem la fórmula anterior, tenim:

- 0: no pertany a cap regió singular.
- 180: pertany a un llaç.
- -180: pertany a un delta.
- 360 o -360: pertany a una espiral.

Les taules en negreta representen els valors de la matriu d'orientació. A sota de les taules en negreta tenim el valor de la resta dels angles de cada veí. El valor en negreta del centre de les taules que no són en negreta és el valor del Pointcaré.

Figura 20. Valors diferents de Pointcaré segons diverses combinacions de les orientacions

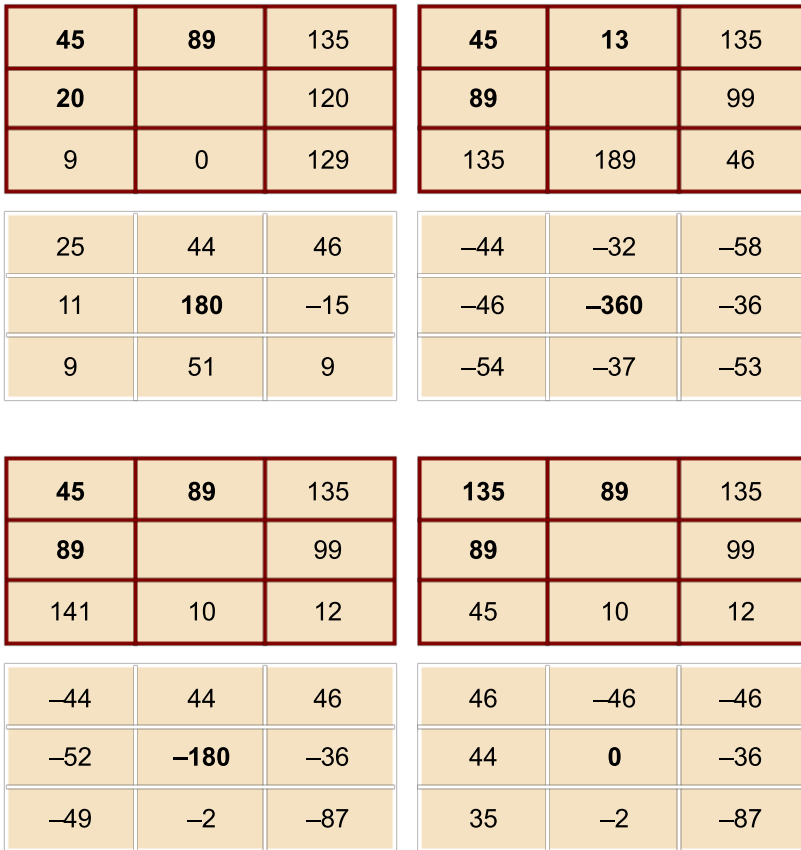


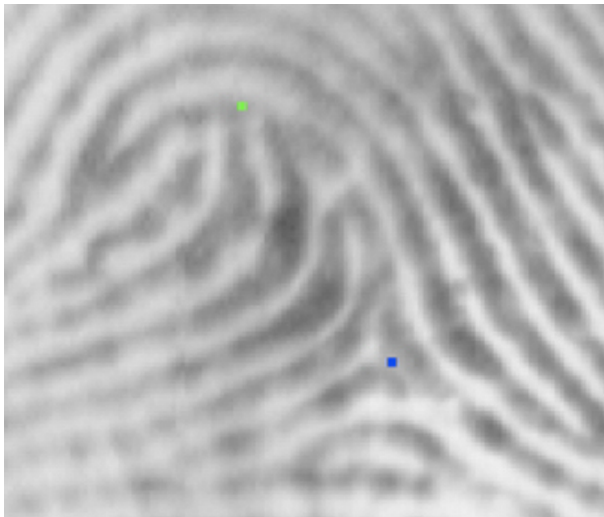
Figura 20

Per exemple, el valor 25 de la cel·la superior esquerra de la primera taula que no és en negreta resulta de fer la resta dels angles 45-20 de la taula en negreta que té a sobre.

En la primera fila tenim un llaç i una espiral. En la segona fila tenim un delta i una figura sense patró.

La figura 21 mostra un tros d'una ditada en què s'ha marcat un delta en blau i un llaç en verd.

Figura 21. Tros d'una ditada en què s'ha marcat un delta en blau (a baix a la dreta) i un llaç en verd (a dalt a l'esquerra).



En la majoria dels casos, quan s'aplica aquest mètode per detectar els punts singulars, apareix un nombre més gran de punts que el real. Per tal de solucionar aquest problema, els autors proposen filtrar consecutivament la imatge

d'orientació filtrada. En cada procés de filtratge van desapareixent els punts singulars falsos. De manera similar al mètode que hem utilitzat per a obtenir la matriu d'orientació filtrada, calculem per a cada píxel un vector mitjana:

$$\hat{\Delta}(x, y) = [\hat{\Delta}_x(x, y), \hat{\Delta}_y(x, y)] \quad (23)$$

simplement amb la mitjana aritmètica en un entorn E de la manera següent (ara la imatge inicial és \hat{D} en lloc de D i no fa falta considerar el gradient r):

$$\hat{\Delta}_x(x, y) = \frac{1}{n^2} \sum_{(x', y') \in E(x, y)} \cos(2\hat{D}(x', y')) \quad (24)$$

$$\hat{\Delta}_y(x, y) = \frac{1}{n^2} \sum_{(x', y') \in E(x, y)} \sin(2\hat{D}(x', y')) \quad (25)$$

Llavors, tornem a obtenir una nova matriu d'orientació filtrada:

$$\hat{D}(x, y) = \frac{1}{2} \arctan\left(\frac{\hat{\Delta}_y}{\hat{\Delta}_x}\right) \quad (26)$$

D'altra banda, també es poden aplicar tècniques com suposar que els punts singulars no poden aparèixer als extrems de la ditada (ajudat amb la imatge segmentada). Finalment, la matriu de la coherència també pot ser útil per a deduir les singularitats. Els punts singulars són aquells en què el vectors d'orientacions no són paral·lels, per tant són poc "coherents".

4.2. Detecció del nucli

Un cop s'han aconseguit detectar els punts singulars, la posició del nucli es pot definir simplement com el punt singular superior. No obstant això, apareixen alguns problemes amb les ditades tipus arc que no tenen singularitats i per tant el nucli no queda definit. La localització correcta del nucli és crucial ja que serveix per a obtenir una imatge de la ditada invariant al desplaçament aplicat al dit respecte del sensor. Un error en aquesta etapa comporta errors en les etapes posteriors com la comparació.

Weigstein, l'any 1982, va proposar un mètode per trobar el nucli sense haver de localitzar primer les singularitats. El mètode, anomenat **R92**, escaneja fila per fila la matriu d'orientació filtrada per trobar **arcs ben formats**. Un arc ben format és aquell que crea un segment horitzontal de set posicions com el que indica la figura 22. Encara que aquest mètode és molt heurístic, obté força bons resultats. Mostra d'això és el fet que és un dels components fonamentals del sistema d'identificació de ditades usat per l'FBI i altres aplicacions civils.

Figura 22. Localització del nucli mitjançant la localització prèvia de set arcs ben formats



Figura 22

El nucli es troba en el punt en què convergeixen les rectes ortogonals al segment.

5. Millora de la imatge de la ditada

L'extracció de les singularitats i del nucli es basa en la matriu de l'orientació filtrada. No obstant això, per a extreure les *minutiae* també es necessita, a més de la matriu de l'orientació filtrada, la imatge original de la ditada. Aquest apartat, en què s'explica com s'ha de millorar la imatge de la ditada, es podria haver ubicat dins de l'apartat de processament de la imatge, ja que els processos que s'hi apliquen són típics dels mètodes de processament d'imatges. Tot i així, s'ha decidit ubicar-lo entre l'apartat d'extracció de singularitats i l'apartat d'extracció de les *minutiae* per dos motius. El primer és que aquesta millora de la imatge no s'aplica quan s'extreuen les *minutiae* i el segon és que algunes tècniques més complexes usen la matriu d'orientació filtrada per a millorar la imatge.

Les tècniques aplicades a l'extracció de les *minutiae* confien en gran manera en la qualitat de la ditada original. En una imatge de la ditada ideal, les serralades i les valls s'alternen i flueixen de manera local en una mateixa direcció. En aquesta situació, les serralades es poden detectar i ubicar fàcilment i, per tant, també se'n poden detectar els finals i separacions, que és on s'ubiquen les *minutiae*. No obstant això, a la pràctica i a causa de les condicions de la pell (per exemple, humida, bruta, amb talls o cops), el soroll propi del sensor, pressió incorrecta del dit o simplement, dits amb poca qualitat inherent, un percentatge important de ditades té poca qualitat. La figura 23 mostra tres imatges de ditades amb molta, mitjana i poca qualitat.

Figura 23. Imatges de ditades amb diferent qualitat



En molts dels casos, les imatges contenen regions amb bona, mitjana i baixa qualitat i és la tasca del sistema automàtic detectar aquestes regions. De les zones amb qualitat alta, se n'extreuen les *minutiae* sense problemes. Es consideren que són **regions ben definides**. A les zones amb qualitat mitjana, es poden aplicar tècniques de visió per computador per a millorar-ne la qualitat. A més es pot informar l'usuari (especialista en ditades) que aquelles *minutiae* són poc fiables. Es consideren que són **regions recuperables**. De les zones amb qualitat baixa no se n'intenten extreure les *minutiae*. Simplement es considera

que són **zones no recuperables**. A més, es considera que hi ha tres tipus diferents de degradació. Segons el tipus de degradació, es pot aplicar una tècnica de recuperació diferent o prendre decisions específiques. Els tres tipus són:

- Les serralades no són estrictament contínues, és a dir, que tenen petits forats.
- Les serralades paral·leles no estan separades correctament o totalment.
- Hi ha molts talls, arrugues o cops al dit.

Aquests tres tipus de degradacions generen problemes a l'hora d'extreure les *minutiae* i d'ubicar-les en el pla bidimensional de la manera següent:

- 1) S'extreuen *minutiae* falses,
- 2) desapareixen *minutiae* genuïnes i
- 3) apareixen desplaçaments en la ubicació de les *minutiae* en el pla bidimensional.

Un expert en ditades normalment té molta capacitat d'extreure o identificar les *minutiae* en zones d'alta i mitjana qualitat usant diversos trucs visuals com la imatge de l'orientació, la suposada continuïtat de la serralada o la tendència de les serralades en un entorn local. En teoria, si ho pot fer un humà també es pot desenvolupar un algorisme capaç d'actuar de la mateixa manera i de manera automàtica. No obstant això, és difícil que aquests algorismes aconseguixin el mateix nivell de qualitat de l'humà.

S'han desenvolupat diverses tècniques per millorar la imatge de la ditada. De més i de menys sofisticades. Algunes usen la matriu d'orientació filtrada o la matriu de coherència. També s'apliquen tècniques basades en l'estudi freqüencial de la ditada, com, per exemple, els filtres de Gabor. I altres tècniques que apliquen mètodes basats en diverses resolucions dels píxels en la imatge. El mètode explicat en aquesta assignatura és un dels més senzills ja que treballa només a escala de píxel. Per tant, no canvia les estructures serralada i vall que es considerin incorrectes com ho fan els algorismes més complexos. En particular, no és capaç d'omplir petits trencaments de les serralades, omplir forats dins de les serralades o separar serralades que s'han ajuntat a causa de la humitat o massa pressió.

Hong, Wan i Jain, l'any 1998, van definir una normalització de la imatge per augmentar-ne la qualitat basada estrictament en el nivell del píxel. Això vol dir que el valor del píxel depèn únicament del seu valor original i d'alguns paràmetres globals però no d'un entorn local al píxel. Aquesta tècnica no genera uns resultats gaire satisfactoris o definitius en imatges molt corrompudes; això

no obstant, s'ha usat com a pas previ a altres tècniques més sofisticades (que considerem que estan fora de l'abast d'aquesta assignatura). El nou valor de cada píxel s'obté de la manera següent:

$$I[x, y] = \begin{cases} m_0 + \sqrt{(I[x, y] - m)^2 \cdot \frac{v_0}{v}} & \text{si } I[x, y] > m \\ m_0 - \sqrt{(I[x, y] - m)^2 \cdot \frac{v_0}{v}} & \text{altrament} \end{cases} \quad (27)$$

en què m i v són la mitjana i la variància de la imatge original I que es calculen de la manera següent:

- **Mitjana:**

$$m = \frac{1}{n^2} \sum_{x,y=0}^{n-1} I[x, y] \quad (28)$$

- **Variància:**

$$v = \frac{1}{n^2} \sum_{x,y=0}^{n-1} (I[x, y] - m)^2 \quad (29)$$

En contrapartida, m_0 i v_0 són la nova mitjana i la variància desitjada després de la normalització.

Kim i Park, l'any 2002, van introduir una implementació d'aquest mètode que treballa per blocs. La imatge es divideix en uns quants blocs (es crea una quadrícula de blocs). Llavors, s'aplica la mateixa equació però en valors diferents per a cada bloc. La mitjana i la variància de cada bloc de la imatge es representen per m^i i v^i , respectivament. A més, es poden definir una mitjana i variància volgudes, m^i_0 i v^i_0 , diferents per a cada bloc dependent d'algunes característiques del bloc, com, per exemple, si es troba en una zona ben definida, recuperable o no recuperable.

6. Detecció de les *minutiae*

La majoria dels mètodes per a comparar ditades es basen en la comparació de les *minutiae*. Per això, la detecció de les *minutiae* és una tasca extremament important i s'hi han dedicat molts esforços de recerca. La majoria dels mètodes requereixen convertir prèviament la imatge original en binària. Normalment, l'entrada d'aquests mètodes és la imatge de la ditada millorada (en lloc de la imatge original). Un cop se n'ha obtingut la imatge binària, es requereix un procés d'afinament de les serralades amb l'objectiu que l'amplària de les serralades sigui només d'un píxel. Normalment, s'anomena *esqueletonització*, ja que l'objectiu és extreure "l'esquelet" de les serralades. I finalment, amb la imatge dels esquelets de les serralades s'obtenen les *minutiae*.

La figura 24 mostra el procés d'extracció de les ditades de manera esquemàtica amb una imatge d'exemple en cadascun dels processos. La ditada original, binaritzada, esqueletonitzada, i les *minutiae* que s'han detectat.

Figura 24. Esquema del procés de detecció de les *minutiae*



6.1. Procés de binarització

El problema general de binaritzar una imatge ha estat àmpliament estudiat en els camps de la visió per computador i reconeixement de patrons des dels orígens d'aquestes ciències en la dècada de 1970. Sovint és una tasca primordial, ja que serveix per a seleccionar quins píxels contenen l'objecte que cal considerar (en la imatge binària tenen el valor blanc) i quins píxels són fons (en la imatge binària tenen el valor negre). Un cop se n'ha obtingut la imatge binària, és trivial deduir la mida, alçària, amplària o posició de l'objecte que cal considerar.

En aquests apunts, considerem que un píxel d'una imatge binària només pot adoptar dos valors: $B[x, y] = \{0, 255\}$. El color és negre quan $B[x, y] = 0$ i el color és blanc quan $B[x, y] = 255$.

Els mètodes més senzills treballen amb un únic llindar per a tota la imatge. Si el valor del píxel és superior a aquest llindar, llavors situen el nou valor a 255. Altrament, el situen a 0. Algunes vegades, podem caracteritzar diferents zones de la imatge (una malla de blocs quadrats) amb diferents llindars, considerant els nivells d'il·luminació. Normalment, les zones són rectangulars i

Vegeu també

El procés de la imatge de la ditada millorada s'ha explicat en l'apartat 5 d'aquest mòdul.

Reflexió

En el que resta d'apartat, explicarem aquests processos, excepte el primer que ja ha estat comentat en l'apartat anterior.

se'n detecten unes setze o trenta-dues. En aquests casos, es pot aplicar un llindar local a cada zona i automàticament adaptar el llindar segons la intensitat mitjana de la zona de la imatge. En el cas de les ditades, i com que les imatges originals (o fins hi tot les que han estat millorades) tenen una qualitat molt baixa, és difícil usar aquests mètodes generals de binarització i se n'han deduït d'altres d'específics. S'han presentat força mètodes de binarització de ditades. En aquest document en presentarem un dels més usats.

Ratha, Chen i Jain, l'any 1995, van introduir un mètode de binarització basat en la matriu d'orientació filtrada. El procés de binarització es porta a terme per a cada píxel seguint els passos següents:

1) Es genera una finestra de 15×15 orientada segons la matriu d'orientació filtrada i centrada en el píxel que volem tractar. Atenció en no sortir de la matriu; en cas que alguna part de la finestra quedés fora de la matriu, es redueixen les dimensions de la finestra.

2) Es calcula la signatura mitjana amb les 15 signatures ortogonals a la matriu d'orientació.

3) Se suavitza la signatura mitjana. Cada píxel de la signatura mitjana suavitzada es calcula com la mitjana d'ell mateix i dels seus dos veïns en la signatura mitjana (atenció amb els píxels extrems).

4) Amb la signatura mitjana, es calcula un únic valor, que és la mitjana dels quinze valors obtinguts de la signatura mitjana. Aquest nou valor és el llindar local per a binaritzar el píxel central de la signatura.

Amb l'objectiu de millorar la qualitat de la imatge binaritzada, alguns sistemes detecten petites taques negres o blanques en la imatge binària que es consideren que han de ser soroll, ja que les serralades o valls mai no són tan petites. Per a eliminar aquestes taques, es pot usar el mètode anomenat *acoloriment de taques*¹.

Donada una imatge binària, el mètode d'acoloriment de taques per a la neteja de petites àrees és un procés format per tres etapes:

1) **Segmentació de la imatge binària en regions:** s'obté una imatge nova en què cada taca té un nivell de gris diferent. A l'hora d'assignar el nivell de gris per a cada regió no s'usa el 0 ni el 255, ja que ja estan inicialment assignats a la imatge binària. Si s'usa un format d'imatge estàndard, només tenim $256 - 2 = 254$ possibles taques diferents. De vegades, aquest valor és massa baix. En aquest cas, s'ha de treballar amb un format d'imatge que permeti tenir més

Reflexió

És important considerar alguns trucs per a optimitzar el procés; si no, aquest algorisme és molt lent. Per exemple, si l'angle en un píxel és molt semblant a l'angle del píxel veí, alguns càlculs es poden evitar.

Vegeu també

Què és una signatura i com s'obtenen els píxels que la formen s'explica en el subapartat 3.3 d'aquest mòdul.

⁽¹⁾En anglès, *blob colouring*.

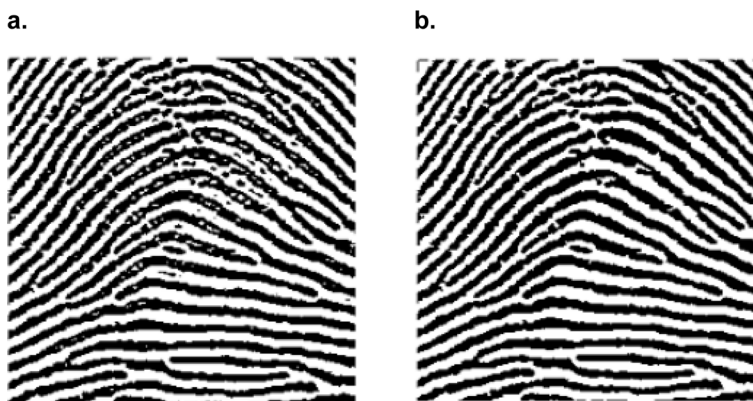
profunditat de píxel. És a dir, els píxels, en lloc d'estar definits per un byte, han d'estar definits per 2 bytes, llavors podem arribar a tenir $2^{16} - 2 = 65.534$ taques diferents.

2) **Detectar les àrees de les taques:** s'obté un vector en què cada posició representa una taca. La taca que s'ha acolorit amb el valor X en el procés anterior queda representada per la posició X del vector. L'àrea d'aquesta taca s'emmagatzema a Vector[X]. Després de la segmentació (etapa anterior) no hi pot haver taques blanques ni negres ja que totes s'han d'haver reassignat. Per això Vector[0] = 0 i Vector[255] = 0.

3) **Les taques amb àrea superior a un llindar (Vector[X] > Llindar) s'esborren:** per fer-ho, prendrem la imatge original binària i les tornarem a pintar. Això vol dir que si són taques negres (blanques) estan envoltades d'una sola taca blanca (negra) i, per tant, s'han de pintar de color blanc (negre).

En la figura 25 es mostra una imatge binària original i la imatge netejada. S'hi aprecia l'eliminació de les petites taques i dels forats. A més d'aquest procés de neteja de la imatge binària, també es pot portar a terme un procés més intel·ligent en què s'esborren els petits sortints de les serralades o s'uneixen serralades que es considera que han estat trencades per talls o brutícia. Aquestes tècniques poden usar la imatge d'orientació.

Figura 25. Imatge binària original i netejada



a. Detall d'una ditada que ha estat binaritzada. b. Imatge després de la binarització amb neteja de taques.

6.2. Procés d'esqueletonització

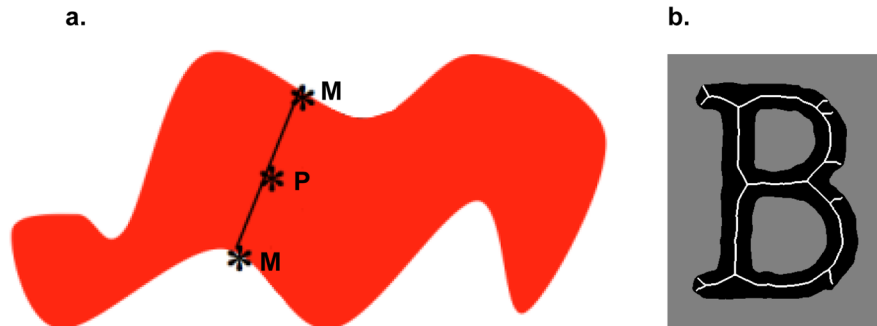
Un píxel d'una imatge esqueletonitzada només pot adoptar dos valors $E[x, y] = \{0, 255\}$. El color negre, $E[x, y] = 0$, representa un píxel de la serralada i el color blanc, $E[x, y] = 255$ representa un píxel d'una vall.

El procés de generar l'esquelet també s'anomena *transformació de l'eix mitjà*². Un píxel forma part de l'esquelet si passa el següent: se'n calcula la distància respecte a tots els píxels del contorn de la figura. En el nostre cas, es calcula la distància amb els píxels del costat de la serralada. Si hi ha més d'un píxel contorn amb la mateixa distància, llavors vol dir que és exactament al mig de

⁽²⁾En anglès, *medial axis transform*.

la serralada i és l'eix mitjà. La figura 26 mostra a l'esquerra una figura amb un píxel p que forma l'esquelet ja que té dos píxels contorn M a la mateixa distància. A la dreta es mostra l'esquelet de la lletra B.

Figura 26

a. Exemple d'un píxel p que forma part de l'esquelet. b. Esquelet d'una figura

González i Woods, l'any 2001, van presentar un algorisme per a generar l'esquelet d'un objecte donada una imatge binaritzada. Definim el veïnat del píxel $p = (x, y)$ de la manera següent:

$p_0 = (x, y + 1)$, $p_1 = (x + 1, y + 1)$, $p_2 = (x + 1, y)$, $p_3 = (x + 1, y - 1)$, $p_4 = (x, y - 1)$, $p_5 = (x - 1, y - 1)$, $p_6 = (x - 1, y)$, $p_7 = (x - 1, y + 1)$

p_3	p_2	p_1
p_4	p_3	p_0
p_5	p_6	p_7

Els comentaris i explicacions d'aquest algorisme se centren en el cas específic que volem trobar l'esquelet de les serralades. Primer de tot, hem d'establir quins píxels es consideren frontera de la serralada:

- p pertany a la frontera esquerra si p_4 és vall i p és serralada.
- p pertany a la frontera dreta si p_0 és vall i p és serralada.
- p pertany a la frontera superior si p_2 és vall i p és serralada.
- p pertany a la frontera inferior si p_6 és vall i p és serralada.

En aquest algorisme usarem expressions lògiques. Per això, hem de definir el valor lògic $n_i = \text{CERT si, i només si, } p_i \text{ és serralada, per a } i = 0 \dots 7$. Amb aquestes expressions lògiques senzilles podem definir les següents expressions més complexes que s'usen en l'algorisme:

$$\begin{aligned}
 B_0 &= n_4 \cdot (n_2 + n_3 + n_5 + n_6) \cdot (n_6 + \bar{n}_7) \cdot (\bar{n}_1 + n_2) \\
 B_2 &= n_6 \cdot (n_0 + n_4 + n_5 + n_7) \cdot (n_0 + \bar{n}_1) \cdot (\bar{n}_3 + n_4) \\
 B_4 &= n_0 \cdot (n_1 + n_2 + n_6 + n_7) \cdot (n_2 + \bar{n}_3) \cdot (\bar{n}_5 + n_6) \\
 B_6 &= n_2 \cdot (n_0 + n_1 + n_3 + n_4) \cdot (n_4 + \bar{n}_5) \cdot (n_0 + \bar{n}_7)
 \end{aligned} \tag{30}$$

No explicarem en detall com es dedueixen aquestes expressions i definirem directament l'algorisme. Aquestes expressions són combinacions de situacions amb què ens podem trobar i que són píxels potencials que passaran de ser serralada a ser vall per així anar aprimant successivament les serralades i quedar-nos només amb l'esquelet. Durant l'algorisme es van seleccionant píxels de la imatge binària per després tractar-los. Inicialment, no hi ha cap píxel seleccionat. La manera més senzilla d'implementar el mètode de selecció de píxels és a partir d'una imatge binària auxiliar.

algorisme d'esqueletonització de González i Woods

repetir

```
per a tot píxel p a la imatge binària E[x, y]
  si p pertany a la frontera esquerra i B4
    llavors selecciona p
fi per a tot
esborra els píxels que han estat seleccionats.
```

```
per a tot píxel p a la imatge
  si p pertany a la frontera dreta i B0
    llavors selecciona p
fi per a tot
esborra els píxels que han estat seleccionats.
```

```
per a tot píxel p a la imatge
  si p pertany a la frontera superior i B2
    llavors selecciona p
fi per a tot
esborra els píxels que han estat seleccionats.
```

```
per a tot píxel p a la imatge
  si p pertany a la frontera inferior i B6
    llavors selecciona p
fi per a tot
esborra els píxels que han estat seleccionats.
```

fins que no hi hagi canvis

fi algorisme

6.3. Procés d'extracció de les *minutiae*

Un cop s'ha extret l'esquelet de la imatge binària, hem de detectar la posició de les ditades, el tipus de ditada (terminal o bifurcació) i també la seva orientació.

Un dels mètodes més senzills s'anomena *creuament de nombres*³.

Arcelli i Baja, l'any 1984 van proposar el mètode de creuament de nombres.

⁽³⁾En anglès, *crossing number*.

Donat un píxel de la imatge esqueletonitzada $\mathbf{p} = (x, y)$ i el valor del qual és $E[x, y] = E[\mathbf{p}]$, el **creuament de nombres** $cn(\mathbf{p})$ es defineix com

la meitat de la suma de les diferències entre parells de píxels adjacents en els vuit veïns d'aquest píxel:

$$cn(p) = \frac{1}{2} \sum_{i=1..8} |E(p_{i \bmod 8}) - E(p_{i-1})| \quad (31)$$

En què $\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_7$ són els píxels veïns a \mathbf{p} que pertanyen a una seqüència específica i que defineixen els vuit veïns de \mathbf{p} . Podria ser la seqüència de l'apartat anterior, per exemple:

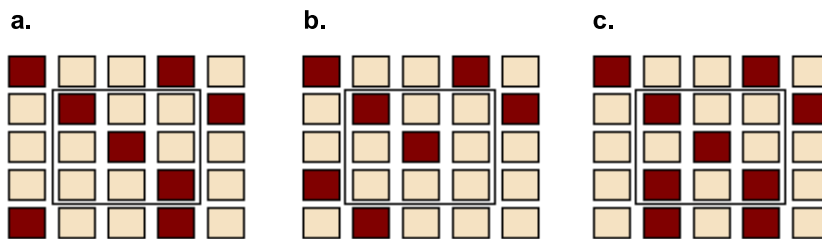
$$\mathbf{p}_0 = (x - 1, y - 1), \mathbf{p}_1 = (x - 1, y), \mathbf{p}_2 = (x - 1, y + 1), \mathbf{p}_3 = (x, y + 1), \mathbf{p}_4 = (x + 1, y + 1), \mathbf{p}_5 = (x + 1, y), \mathbf{p}_6 = (x + 1, y - 1), \mathbf{p}_7 = (x, y - 1)$$

Com que en aquests apunts sempre considerem que els píxels de la imatge binària i esqueletonitzada només poden adoptar els dos valors $\{0, 255\}$, llavors és senzill que ens adonem que un píxel (x, y) que forma part d'una cresta (color negre), és a dir, que té valor $E[x, y] = 0$ és:

- Un punt aïllat si $cn(\mathbf{p}) = \frac{1}{2} \cdot 0 \cdot 255 = 0$.
- Una *minutia* tipus terminal si $cn(\mathbf{p}) = \frac{1}{2} \cdot 2 \cdot 255 = 255$.
- Un punt intermedi d'una cresta si $cn(\mathbf{p}) = \frac{1}{2} \cdot 4 \cdot 255 = 510$.
- Una *minutia* tipus bifurcació si $cn(\mathbf{p}) = \frac{1}{2} \cdot 6 \cdot 255 = 765$.
- Un creuament de serralades si $cn(\mathbf{p}) = \frac{1}{2} \cdot 8 \cdot 255 = 1.020$.

És matemàticament impossible que el nombre en negreta dels casos exposats anteriorment prengui un valor senar. La figura 27 mostra un punt intermedi: $cn(\mathbf{p}) = 510$, una *minutia* terminal: $cn(\mathbf{p}) = 255$ i una *minutia* bifurcació: $cn(\mathbf{p}) = 765$.

Figura 27. Tres exemples de píxels que componen una serralada



a. Un punt intermedi. b. Una *minutia* terminal. c. Una *minutia* bifurcació

Ara hem de detectar la direcció de la *minutia* ubicada en la posició (x, y) que descriurem per α_{xy} . Com s'ha comentat en apartats anteriors, aquesta direcció és de 360 graus (els quatre quadrants) i es calcula mitjançant la imatge d'orientació. Com que la imatge d'orientació només té la informació de dos quadrants, haurem de fer la conversió de dos quadrants $]-\pi/2, \pi/2]$ a quatre quadrants $]-\pi, \pi]$.

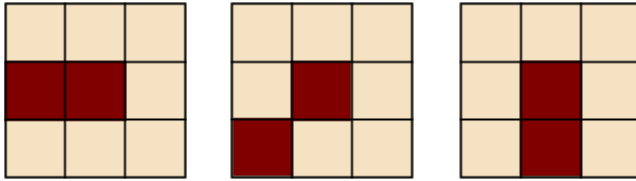
Vegeu també

La figura 9 del subapartat 2.2 mostra l'orientació dels angles.

1) En els casos en què tinguem una *minutia* tipus terminal, $cn(\mathbf{p}) = 255$:

a) Si $\theta_{xy} \in [0, \pi/2]$ i ens trobem en una de les situacions que presenta la figura 28:

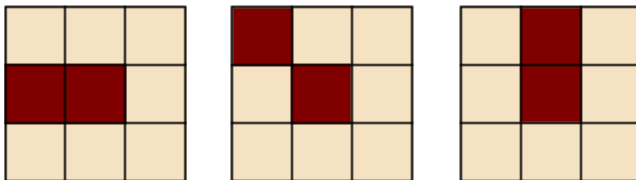
Figura 28



llavors: $\alpha_{xy} = \theta_{xy} - \pi$

b) Si $\theta_{xy} \in]0, -\pi/2]$ i ens trobem en una de les situacions que presenta la figura 29:

Figura 29



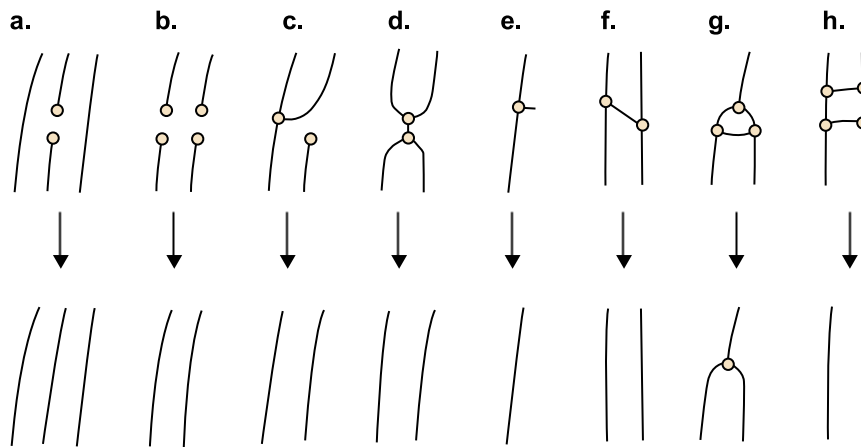
llavors: $\alpha_{xy} = \theta_{xy} + \pi$

c) En la resta de casos: $\alpha_{xy} = \theta_{xy} - \pi$

2) En els casos en què tinguem una *minutia* tipus bifurcació, $cn(\mathbf{p}) = 765$, també es podria trobar l'angle de la *minutia* fent una casuística semblant. No obstant això, el nombre de combinacions és molt superior. Per aquest motiu, diversos investigadors opten per una solució diferent. Partint de la imatge binaritzada, llavors s'inverteix la imatge. D'aquesta manera, les crestes queden representades en color blanc i les valls en color negre. Amb aquesta imatge invertida, s'aplica tot el mateix procés esmentat anteriorment fins a arribar a la definició de les *minutiae* terminals i el seu angle. Les *minutiae* terminals de la imatge invertida són les *minutiae* bifurcació de la imatge binària original.

Finalment, és interessant esmentar que, amb l'algorisme que hem presentat, s'acostumen a detectar força *minutiae* falses. Un dels mètodes més usats per a eliminar falsos positius (és a dir, esborrar *minutiae* falses) és l'algorisme de Xiao i Raafat presentat l'any 1991. L'algorisme, usant la informació de la imatge i el conjunt de *minutiae* trobat, dedueix les combinacions o estructures que han de ser eliminades si es troben a una distància inferior a un llindar. La figura 30 mostra aquestes combinacions.

Figura 30



a i b. Connecta *minutiae* terminals encarrades segons els seus angles. c i d. Esborra bifurcacions encarrades segons els seus angles. e. Esborra petites serralades amb distància inferior a un llindar. f. Esborra ponts entre serralades i considera l'angle de la imatge d'orientació. g. Esborra triangles i deixa només una *minutia* bifurcació depenent d'una distància mínima. h. Esborra estructures tipus escala (generalització del cas f)

7. Comparació de ditades

En les biblioteques informàtiques per a desenvolupar aplicacions amb dispositius biomètrics hi ha les rutines de comparació de ditades. Aquestes, donades un parell d'imatges de ditades, poden retornar un resultat continu que normalment està limitat de 0 a 1 o poden retornar directament una decisió binària: semblants/diferents. El resultat continu entre 0 i 1 pot representar dues coses. De vegades, representa una probabilitat, llavors, 1 significa que les dues ditades són completament iguals i 0 completament diferents. Però, d'altres, representa una distància, llavors, el valor 0 indica que són completament iguals i 1 completament diferents. Tots els algorismes bàsics de comparació de ditades retornen uns valors continus, per tant, en els casos en què la sortida de la rutina és binària és perquè hi ha un llindar modificable pel supervisor del sistema. Si la sortida contínua de l'algorisme és superior (inferior) al llindar de la probabilitat (distància), llavors es decideix que les ditades són iguals. Altrament, es decideix que són diferents.

La majoria dels algorismes de comparació de ditades operen amb la informació del conjunt de *minutiae* extretes de les ditades que cal comparar, però també s'han presentat alguns algorismes que treballen directament amb la imatge. La qualitat de la comparació d'aquests últims sempre ha estat més baixa, però tenen l'avantatge que no fan falta tots els processos d'obtenció de la imatge binària i extracció de les *minutiae* comentats anteriorment.

A partir d'aquest punt, anomenarem la representació de la ditada que s'ha usat en el procés de matriculació de l'usuari T (prové de l'anglès *template*). I anomenem I (prové de l'anglès *input*) la representació de la nova ditada que volem comparar. En el cas dels algorismes que comparen directament les ditades, les representacions de les ditades són exactament les imatges de les ditades I i T. En els altres algorismes, normalment és un conjunt de *minutiae*. Al llarg d'aquest apartat, anomenem tant la ditada com la seva representació T i I.

Els processos d'extracció de les característiques de les ditades i de comparació són normalment molt similars tant si estem duent a terme una identificació com una verificació. Això es deu al fet que la identificació es pot portar a terme fent un conjunt de comparacions una per una entre parelles de representacions de ditades. Els processos de classificació de ditades normalment s'usen per a reduir el temps de cerca d'una ditada en una base de dades gran en un procés d'identificació.

El procés de comparació de dues ditades és un problema molt difícil que fa molts anys que s'està investigant i desenvolupant. Això es deu a la diferència enorme que hi pot haver entre dues ditades del mateix dit i posades en el

Vegeu també

Els sistemes d'identificació i verificació han estat explicats en el mòdul "La biometria per a la identificació de les persones" i el processos de classificació s'explicaran en l'apartat següent, "Classificació de les ditades" d'aquesta assignatura.

sensor en un interval molt curt de temps. Es diu que les ditades tenen una gran **variabilitat** dins de la classe. És a dir, que la impressió de la mateixa ditada pot variar molt. Els principals factors són:

- **Desplaçament:** el mateix dit es pot posar en diferents posicions en l'àrea de lectura del sensor en diverses adquisicions. Aquest fet genera un desplaçament de la regió on hi ha el dit en la imatge. Per exemple, només que desplacem el dit 2 mm (imperceptible per l'usuari) genera un desplaçament de 40 píxels en un escàner de dit usual amb resolució de 500 ppp.
- **Rotació:** de manera similar al desplaçament, el mateix dit es pot posar en diferents angles respecte de l'eix vertical del sensor. Normalment els algorismes toleren una rotació de fins a 20 graus cap a la dreta o cap a l'esquerra. Si la rotació és superior, llavors els algorismes no són capaços d'adonar-se que és el mateix dit.
- **Solapament parcial:** el desplaçament i la rotació del dit causen que part de la ditada caigui fora de l'àrea del sensor. Aquest fet provoca que la part de la ditada que apareix en les dues imatges sigui més petita del normal.
- **Distorsió no lineal:** a causa de la plasticitat del dit, que la superfície del dit és tridimensional i que el sensor és bidimensional, la ditada és una imatge bidimensional que mostra una superfície tridimensional però amb distorsions difícils de controlar. Aquestes distorsions poden ser del tipus **compressió de la imatge** (normalment en les parts centrals de la ditada) o **extensió** (normalment en les parts externes de la ditada) i depenen de la pressió que s'ha aplicat amb el dit cap al sensor.
- **Condicions de la pell:** les serralades i valls són captades pel sensor si estan tocant la superfície del sensor de manera uniforme. No obstant això, la pressió, la brutícia de la pell, les malalties de la pell, la suor o els talls fan que la imatge canviï en successives adquisicions de la ditada.
- **Soroll:** el mateix sensor per a adquirir la imatge genera soroll i distorsiona la imatge. Un exemple habitual és el greix que queda del dit a l'àrea de vidre del sensor. Depenent de la netedat d'aquest vidre, la imatge serà canviant. Una altra font de distorsió pot ser la comunicació del sensor amb l'ordinador o sistema usat per a comparar les imatges. Depenent del protocol de comunicació o de la llargària del cable o de la xarxa, es poden generar errors de comunicació.
- **Errors en l'extracció de característiques:** els mateixos algorismes d'extracció de característiques introdueixen errors. Aquests errors poden aparèixer en qualsevol dels processos interns de l'algorisme. Per exemple, hem de fer aproximacions que indueixen a errors quan estimem l'orientació, o quan estimem la posició de la ditada.

La figura 31 mostra unes imatges del mateix dit que no van poder ser reconegudes per la majoria dels algorismes presentats en el concurs Fingerprint Verification Competition de l'any 2002.

Figura 31. Dues ditades d'un mateix dit però amb característiques molt diferents



D'altra banda, en bases de dades grans, ens podem trobar dos dits diferents que tenen una estructura molt semblant (posició de la singularitat i imatge d'orientació). No obstant això, s'ha vist que la probabilitat que les *minutiae* caiguin en posicions similars és molt baixa. En la figura 32 es mostren dues ditades de dits diferents i amb una estructura molt similar. Però si ens hi fixem atentament, veurem que les *minutiae* estan ubicades en posicions diferents.

Figura 32. Dues ditades de dos dits diferents però amb característiques semblants



S'han presentat molts algorismes de comparació de ditades i amb metodologies molt diverses durant els darrers vint anys. La majoria d'aquests algorismes no tenen problemes a l'hora de comparar ditades de bona qualitat. No obstant això, la comparació de ditades de baixa qualitat continua sent un repte avui

en dia. En els casos dels AFIS supervisats per un humà, s'hi acostuma a aplicar un algorisme de detecció de la qualitat de la ditada. Les ditades que no superen el nivell de qualitat no són introduïdes al sistema en el procés de matrícula.

Els mètodes per a comparar un parell de ditades es poden classificar en tres grans famílies:

1) **Comparació basada en la correlació:** les imatges de les ditades s'ubiquen una sobre de l'altra i la correlació entre els píxels es calcula aplicant diferents alineaments, com, per exemple, diferents rotacions i translacions.

2) **Comparació basada en les *minutiae*:** és la tècnica més usada amb diferència. A més, és la base de la metodologia que usen els experts per a comparar ditades de manera manual. La ditada queda representada per un conjunt de *minutiae* i la comparació es basa a alinear el conjunt de *minutiae* generat per la ditada *template* amb el conjunt de *minutiae* generat per la ditada *input*.

3) **Comparació basada en altres característiques que no són les *minutiae*:** extreure les *minutiae* és una tasca difícil en ditades de molt baixa qualitat. Tanmateix, hi ha altres característiques que són més fàcils d'extreure, encara que la imatge sigui molt sorollosa. Per exemple, la informació de l'orientació local, la freqüència de les serralades en un punt, la forma de les serralades o la informació de la textura. La comparació també es basa a alinear aquestes característiques locals.

Tot seguit, ens endinsarem a explicar les dues primeres famílies de mètodes de comparació.

7.1. Tècniques basades en la correlació d'imatges

Suposem que T' i I' són dues imatges **vectoritzades** de ditades corresponents a la imatge registrada T i la imatge I nova que es vol identificar.

Una **imatge vectoritzada** és aquella la matriu de la qual s'ha convertit en un vector simplement concatenant les columnes de la imatge.

Una mesura intuïtiva de la diversitat d'aquestes imatges és la suma de les diferències al quadrat, SSD, entre les intensitats dels píxels:

$$SSD = \|T' - I'\|^2 = (T' - I')^t (T' - I') = \|T'\|^2 + \|I'\|^2 - 2T'tI' \quad (32)$$

Reflexió

La tercera família de mètodes de comparació no s'explicarà en aquest material.

en què el superíndex t denota la matriu transposada. Si els termes $\|T\|^2$ i $\|I\|^2$ són constants i independents de la correlació entre imatges, la diversitat entre les dues imatges vectoritzades es minimitza quan la relació creuada, $CC(T, I) = T^t I$, entre T i I es maximitza.

Com que el terme $T^t I$ apareix en forma negativa a la funció de la diversitat, llavors podem dir que $CC(T, I)$ és una funció que representa la similitud entre imatges. No obstant això, a causa del possible desplaçament i rotació que apareixen inevitablement en posicionar el dit al sensor, la similitud entre les imatges no es pot calcular directament superposant ambdues imatges i aplicant-hi la funció CC .

Suposem que $I(\Delta_x, \Delta_y, \theta)$ és la imatge I en què s'ha aplicat una rotació amb angle θ amb origen al centre de la imatge i un desplaçament Δ_x i Δ_y píxels en les direccions x i y , respectivament. Suposem que $I(\Delta_x, \Delta_y, \theta)$ és la imatge vectoritzada de $I(\Delta_x, \Delta_y, \theta)$. Llavors, la similitud entre les dues imatges es pot mesurar de la manera següent:

$$S(T, I) = \max_{\Delta_x, \Delta_y, \theta} CC(T, I(\Delta_x, \Delta_y, \theta)) \quad (33)$$

L'aplicació directa d'aquesta equació poques vegades aporta resultats acceptables pels problemes de variabilitat que hem comentat anteriorment. Aquesta equació només té en compte la rotació i el desplaçament de les ditades però no considera tota la resta dels factors de distorsió o variabilitat. A més, el còmput directe d'aquesta equació és molt costós. S'han definit alguns algorismes per trobar solucions aproximades. Una solució senzilla seria adaptar l'algorisme de la transformada generalitzada de Hough.

Reflexió

Els algorismes per a trobar solucions aproximades no s'expliquen en aquest material pel poc ús que es fa d'aquesta tècnica.

7.2. Tècniques basades en la comparació de *minutiae*

Com s'ha comentat anteriorment, la tècnica de comparació de ditades basada en la comparació de *minutiae* és la més estesa. El principal motiu d'això és l'analogia estricta amb la manera com els especialistes forenses comparen les ditades i l'acceptació a tribunals com a prova d'identitat de persones a la majoria dels països de tot el món. Si la metodologia del sistema automàtic és la mateixa que la del sistema manual, llavors, els especialistes forenses poden defensar les proves d'una manera més convincent.

Suposem que T i I són la representació de la ditada *template* i la ditada *input*. Assumirem que aquesta representació és un conjunt de *minutiae* del tipus:

$$T = \{m_1, m_2, \dots, m_m\}, m_i = \{x_p, y_p, \theta_p, t_i\}, i = 1 \dots m$$

$$I = \{m'_1, m'_2, \dots, m'_n\}, m'_j = \{x'_j, y'_j, \theta'_j, t'_j\}, j = 1 \dots n$$

en què m i n denoten el nombre de *minutiae* en els conjunts **T** i **I**, respectivament. La posició i l'orientació de la *minutia* en la imatge són x_j, y_j, θ_j i el tipus de *minutiae* és $t_j \in \{Terminal, Bifurcació\}$

Considerem que les *minutiae* m_i i m_j s'emparellen si són del mateix tipus, la distància espacial, sd , entre elles és inferior a un llinard de tolerància r_0 i la diferència de direccions, dd , entre elles és també inferior a un llinard de tolerància θ_0 . Aquests llinars són necessaris per a compensar els errors inevitables que sorgeixen en els algorismes d'extracció de característiques i les distorsions causades per la ditada en si (plasticitat, brutícia...). Formalment, dues *minutiae* s'emparellen si es compleixen les tres condicions següents:

$$t_i = t_j \quad (34)$$

$$sd(m'_j, m'_i) = \sqrt{(x'_j - x'_i)^2 + (y'_j - y'_i)^2} \leq r_0 \quad (35)$$

$$dd(m'_j, m'_i) = \min(|\theta_j - \theta_i|, 2\pi - |\theta_j - \theta_i|) \leq \theta_0 \quad (36)$$

La definició de la distància de direccions pren el mínim de $|\theta_j - \theta_i|$ i de $2\pi - |\theta_j - \theta_i|$ per considerar la circularitat dels angles. Per exemple, la diferència entre 5° i 350° no és 345° sinó $360^\circ - (350^\circ - 5^\circ) = 150$.

Alinear les dues ditades és un pas obligat per a maximitzar el nombre de *minutiae* emparellades. L'alineament correcte de dues ditades requereix que puguem recuperar la informació del desplaçament en les x i les y i també la rotació θ que hem portat a terme quan s'ha ubicat el dit al sensor. Suposem que $map()$ és una funció que transforma una *minutia* m'_j de **I** amb una *minutia* m''_j segons les transformacions de desplaçament $[\Delta_x, \Delta_y]$ i rotacions θ . Usualment, el centre de la rotació es considera la posició mitjana de les *minutiae*. Definim $map()$ formalment com a:

$$map_{\Delta_x, \Delta_y, \theta}(m'_j) = m''_j \quad (37)$$

en què:

$$m''_j = \{x''_j, y''_j, \theta''_j, t''_j\} \quad (38)$$

La nova posició i l'angle de rotació es calculen de la manera següent:

$$\theta_j^* = \theta_j + \theta \quad (39)$$

$$\begin{bmatrix} x_j^* \\ y_j^* \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_j \\ y_j \end{bmatrix} + \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} \quad (40)$$

I òbviament, el tipus de *minutiae* no es veu afectat per aquests tipus de transformacions, $t_j^* = t_j$.

D'altra banda, definim $mm()$ com una senzilla funció que retorna 1 si dues *minutiae* s'emparellen, és a dir, si són del mateix tipus i la seva diferència és inferior als llindars r_0 i θ_0 . Formalment:

$$mm(m_j^*, m_i) = \begin{cases} 1 & \text{si } t_j^* = t_i \text{ i } sd(m_j^*, m_i) \leq r_0 \text{ i } dd(m_j^*, m_i) \leq \theta_0 \\ 0 & \text{altrament} \end{cases} \quad (41)$$

Llavors, el problema d'etiquetar les *minutiae* de dues ditades es pot formular com un problema de trobar un punt màxim. Formalment:

$$\max_{\Delta_x, \Delta_y, \theta, P} \sum_{i=1}^m mm(\text{map}_{\Delta_x, \Delta_y, \theta}(m^{P(i)}), m_i) \quad (42)$$

en què $P(i) = j$ pot ser qualsevol funció que determina l'etiquetatge entre *minutiae* d'una de les ditades amb *minutiae* de l'altra ditada. Una *minutia* té exactament una parella a l'altra ditada o no en té cap.

- $P(i) = j$: indica que la parella de m_i de T és la *minutia* m'_j de I.
- $P(i) = \text{Nul}$: indica que la *minutia* m_i de T no té parella.
- Un *minutia* m'_j de I no té parella a T si $P(i) \neq j, \forall i = 1 \dots m$.
- Cada *minutia* de I s'associa com a molt a una *minutia* de T, és a dir, P és una funció injectiva dins del conjunt de les ditades. O s'associa a una sola *minutia* o s'associa al *Nul*. La seva expressió matemàtica és:

$$\forall i = 1 \dots m, k = 1 \dots m, i \neq j \Rightarrow P(i) \neq P(k) \text{ o } P(i) = P(k) = \text{Nul} \quad (43)$$

És important remarcar que quan s'associen dues *minutiae*, $P(i) = j$, no significa que les *minutiae* m'_j de I i m_i de T siguin exactament iguals després de la transformació deguda a $\Delta_x, \Delta_y, \theta$, sinó vol dir que és probable que siguin la mateixa *minutia*. També s'ha de remarcar que el problema d'etiquetar dos conjunts de *minutiae* s'ha convertit en un problema de trobar un màxim en una funció. Aquest màxim depèn de dos factors: el primer és la transformació global d'una ditada respecte de l'altra $\Delta_x, \Delta_y, \theta$ i el segon és l'associació de *minutiae* P . Però a l'hora de trobar la solució òptima, s'han de considerar aquests dos

factors alhora. En el camp de la visió per computador, aquest és el problema d'emparellar imatges (*image registration*) i és molt complex. S'han presentat diversos algorismes i encara se'n presenten de nous.

Nosaltres us mostrem una solució bastant senzilla al problema. La senzillesa de l'algorisme es basa a suposar que primer es pot calcular la transformació de la ditada (és a dir, els paràmetres Δ_x , Δ_y , θ) i després, tenint aquests paràmetres, es pot calcular l'associació de ditades P . El mètode aporta una solució subòptima, que vol dir que potser hi ha una solució més bona però no l'hem sabut trobar. Per tant, un dels paràmetres de transformació pot tenir un valor una mica diferent o una associació de dues *minutiae* ser diferent.

Ratha, Rover i Jain, l'any 1996, van proposar l'algorisme següent per trobar la transformació d'una ditada. Primer de tot, van discretitzar l'espai de transformació:

$$\Delta_x^+ \in \{\Delta_{x_1}^+, \Delta_{x_2}^+, \dots, \Delta_{x_a}^+\} \quad (44)$$

$$\Delta_y^+ \in \{\Delta_{y_1}^+, \Delta_{y_2}^+, \dots, \Delta_{y_a}^+\} \quad (45)$$

$$\theta^+ \in \{\theta_1^+, \theta_2^+, \dots, \theta_b^+\} \quad (46)$$

El nombre d'elements discretitzats en el desplaçament, a , i en l'angle, b , depèn de l'aplicació. Com més grans són, més lent és l'algorisme. A més, si es discretitza en molts valors, ens trobem que totes les mostres cauen gairebé en valors discrets diferents i llavors l'algorisme no funciona. Amb aquesta discretització, es defineix una matriu tridimensional A en què cada eix és un dels tres paràmetres de transformació i en què cada cel·la és una possible combinació d'aquests valors de manera discreta:

$$A[\Delta_x^+, \Delta_y^+, \theta^+] \quad (47)$$

A continuació, es mostra l'algorisme.

algorisme de Ratha, Rover i Jain

$A[\Delta_x^+, \Delta_y^+, \theta^+] = 0$ per a tot valor de Δ_x^+ , Δ_y^+ i de θ^+

```

per a tota minutia  $m_i$  de  $T$ 
  per a tota minutia  $m'_j$  de  $I$ 
    si  $t'_j = t_i$  llavors
      per a tot valor de  $\theta^+$ 
         $\theta_{aux} = \theta'_j + \theta^+$ 
        si  $dd(m_{aux}, m_i) \leq \theta_0$ 
          
$$\begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} - \begin{bmatrix} \cos \theta^+ & -\sin \theta^+ \\ \sin \theta^+ & \cos \theta^+ \end{bmatrix} \begin{bmatrix} x'_j \\ y'_j \end{bmatrix}$$

           $\Delta_x^+ =$  discretització de  $\Delta_x$ 
            en el valor discret més proper
           $\Delta_y^+ =$  discretització de  $\Delta_y$ 
            en el valor discret més proper
          si  $\Delta_x^+$  i  $\Delta_y^+$  són dins del domini
            llavors  $A[\Delta_x^+, \Delta_y^+, \theta^+] = A[\Delta_x^+, \Delta_y^+, \theta^+] + 1$ 
          fi si
        fi si
      fi si
    fi si
  fi si
fi algorisme

```

Fixeu-vos que aquest algorisme usa només el llindar θ_0 i no fa falta usar el llindar r_0 . Un cop s'ha acabat aquest procés d'acumular candidats a la matriu A , llavors es decideix que els valors resultants de la transformació entre ditades, Δ_x , Δ_y , θ són els que han aparegut més vegades. És a dir, la posició de la cel·la de A el contingut de la qual és més gran:

$$\{\Delta_x, \Delta_y, \theta\} = \max_{\{\Delta_x^+, \Delta_y^+, \theta^+\}} A[\Delta_x^+, \Delta_y^+, \theta^+] \quad (48)$$

Podrien aparèixer diverses cel·les amb el màxim valor; en aquests casos, es tria una de les cel·les a l'atzar.

Un cop tenim els valors de transformació, Δ_x , Δ_y , θ , proposem un algorisme per emparellar les *minutiae* molt senzill:

1) Primer de tot, s'han de convertir totes les *minutiae* de I amb els valors obtinguts:

$$m'_j = \text{map}_{\Delta_x, \Delta_y, \theta}(m_j) \quad (49)$$

2) Després, calcularem les distàncies de totes les *minutiae* m_j^* amb totes les *minutiae* m_i i obtindrem una matriu **D** en què cada cel·la representa la distància entre les *minutiae*. Aquesta distància es pot calcular com la distància ponderada de la distància espacial i la distància angular:

$$[j, i] = \alpha \cdot sd(m_j^*, m_i) + (1 - \alpha) \cdot dd(m_j^*, m_i)$$

en què $0 \leq \alpha \leq 1$.

3) Després, anirem recorrent la matriu **D** i n'ordenarem les cel·les de més petites a més grans.

4) Amb les cel·les de **D** ordenades, anirem seleccionant les cel·les de menys a més valor. Si el valor més petit és $D[j, i]$ llavors:

a) Portarem a terme l'assignació $P(i) = j$.

b) Anul·larem totes les cel·les de la matriu **D**, de manera que $D[j, k]$ i $D[k, i]$ per a tot k . Aquest pas es duu a terme per garantir que l'etiquetatge sigui bijectiu.

c) Aquest procés s'acaba quan ja no queden més cel·les per assignar o el seu valor supera un llindar. Una manera de decidir aquest llindar pot ser:

$$sd(m_j^*, m_i) \leq r_0 \quad i \quad dd(m_j^*, m_i) \leq \theta_0 \quad (50)$$

5) Si han quedat *minutiae* sense assignar, ja que el procés d'assignació ha acabat perquè s'ha superat el llindar, llavors les *minutiae* restants s'han d'assignar a *Nul*:

$$P(i) = Nul \quad o \quad P(i)^{-1} = Nul$$

8. Classificació de les ditades

Com s'ha explicat en el mòdul "La biometria per a la identificació de les persones", la identificació d'una persona requereix la comparació de la seva ditada amb totes les ditades de la base de dades. Aquesta base de dades podria ser molt gran en algunes aplicacions forenses o civils (més de cinquanta milions de ditades). En aquests casos, si no s'hi aplica cap tècnica de classificar les ditades, la resposta temporal del sistema no és admissible. Es pot reduir el temps emprat en el procés d'identificació si es redueix el nombre de comparacions que cal fer. Algunes vegades, es pot usar la informació del sexe, raça, edat i altres dades relacionades amb la persona amb intenció de reduir la porció de la base de dades per cercar. No obstant això, algunes vegades aquesta informació no és accessible (per exemple, en la identificació de criminals). En aquests casos, la ditada de la persona desconeguda és comparada només amb les ditades que pertanyen a la mateixa classe de ditada que ella.

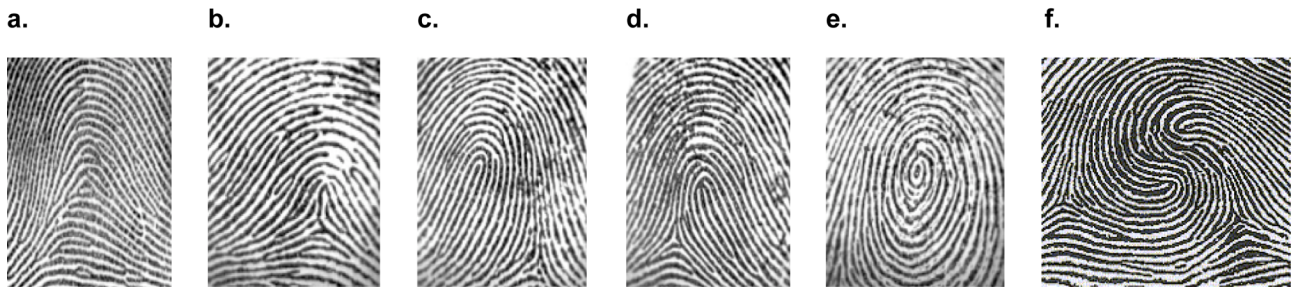
El problema de la classificació de les ditades es refereix al problema d'assignar les ditades en una classe específica d'una manera **consistent** (cada ditada va a parar només a una classe) i **segura** (sempre va a parar a la mateixa classe). El mètode més usat per a generar aquestes classes es basa en les característiques globals de la ditada, és a dir, l'estructura global de les serralades i valls i les singularitats.

La primera classificació de ditades coneguda va ser generada per **John Evangelist Purkinje** l'any 1823. Purkinje era professor de la Universitat de Breslau (Polònia) i va classificar les ditades en nou categories encara que mai no en va mencionar la utilitat per a la identificació personal. El primer estudi seriós i profund va ser elaborat per **Sir Francis Galton**, l'any 1892. Galton era un antropòleg britànic i va dividir les ditades en tres classes principals (arc, llaç i espiral) i més tard hi va afegir unes subdivisions. Més o menys pel mateix temps, **Juan Vucetich**, que era el cap de la policia a Buenos Aires, a l'Argentina, va dissenyar un sistema de classificació semblant. Aquest sistema es va fer famós perquè va ser el primer sistema que es va usar per a identificar un assassí l'any 1892. Uns deu anys més tard, el britànic **Sir Edward Henry** va redefinir el mètode Galton l'any 1900 i en va incrementar el nombre de classes.

El **sistema Galton-Henry** de classificació de ditades es va adoptar en molts països per primer cop i van començar a crear les primeres base de dades.

De fet, els sistemes actuals es basen encara en aquella primera classificació. La figura 33 mostra les sis classes més comunes de l'esquema Galton-Henry: arc, arc tipus tenda, llaç esquerre, llaç dret, espiral i doble llaç.

Figura 33. Classes més comunes de l'esquema de Galton-Henry



Ditades. a. Tipus arc. b. Arc tenda. c. Llaç esquerre. d. Llaç dret. e. Espiral. f. Doble llaç

8.1. Descripció formal de les classes de ditades

La descripció formal de les classes de ditades és la següent:

- **Arc:** la ditada té serralades que entren per una banda, s'enfilen cap a un petit turó i surten per l'altra banda. No hi ha punts singulars de cap tipus. No hi ha ni llaços ni deltes.
- **Arc tipus tenda:** són ditades semblants a l'arc però el turó és molt més pronunciat, i això fa que hi hagi una curvatura molt més pronunciada, quasi vertical. Hi ha dos punts singulars, hi ha un llaç i a sota hi ha un delta.
- **Llaç esquerre:** unes quantes serralades entren per la banda esquerra, fan una corba i se'n tornen cap a l'esquerra. Hi ha dos punts singulars, un llaç i un delta. A diferència de l'arc tipus tenda, el delta està ubicat a sota del llaç però una mica a la dreta.
- **Llaç dret:** és semblant al llaç esquerre però simètric a l'eix vertical.
- **Espiral:** conté com a mínim una serralada que fa un tomb de 360 graus entorn del centre de la ditada. Hi ha quatre singularitats. Dos llaços encarats i pràcticament en la mateixa posició i dos deltes més a baix i un a cada costat.
- **Doble llaç:** aquest tipus de ditada és semblant a l'espiral però les singularitats tipus llaç estan separades. D'aquesta manera, ja no trobem una serralada que faci un tomb de 360 graus i hi ha una serralada que recorre la ditada d'un dels llaços a l'altre.

La classificació de ditades és una tasca molt difícil a causa de la gran variabilitat dins de les classes i la petita variabilitat entre una classe i l'altra. A més, les ditades tenen soroll, cosa que fa que la classificació encara sigui més complexa. S'han presentat diversos mètodes per deduir automàticament les classes de ditades. Els més coneguts es basen en:

- **Sistema de regles:** cada classe de ditada es representa per un conjunt de normes. Per exemple, perquè sigui un arc tendra hi ha d'haver un llaç a sobre d'un delta.
- **Sintàctic:** els patrons es descriuen amb una gramàtica. És a dir, es defineixen unes regles de producció i uns punts terminals. Una ditada pertany a una classe si és acceptada per l'analitzador sintàctic.
- **Estructural:** la ditada es defineix amb un graf. Els nodes del graf són regions de la ditada en què l'angle és similar. Els arcs indiquen que les regions estan a tocar.
- **Estadístic:** de cada ditada es genera un vector de característiques i llavors s'hi aplica un classificador de propòsit general estadístic. El més conegut és el classificador de Bayes.
- **Xarxa neuronal:** els elements de la imatge d'orientació s'usen com a característiques d'entrada d'una xarxa neuronal. Normalment la xarxa conté diverses capes de perceptrons.

La selecció del mètode de classificació depèn del nombre de classes i de la distribució natural de les ditades dins les classes. Desafortunadament, el nombre de classes és molt petit i, a més, les ditades no es distribueixen de manera uniforme dins les classes. Estudis pràctics han demostrat que la distribució és la següent: llaç esquerre 32%, llaç dret 32% i espiral 28%. Només el 8% restant són arcs i arcs tipus tenda de manera equitativa. Un problema afegit a la classificació de ditades és que hi ha moltes ditades la classe de les quals és molt ambigua i es dubta entre una classe i l'altra. Per exemple, el pas de llaç esquerre a arc o d'arc a llaç dret és continu. Per això, s'ha de posar un llindar i decidir a partir de quin punt creiem que és d'un tipus o de l'altre. Es pot usar un sistema que indiqui el nivell d'ambigüitat de la ditada, per exemple una probabilitat de pertànyer a una classe.

Reflexió

En aquest mòdul no explicarem en detall els mètodes per a deduir automàticament les classes de ditades.

Resum

En aquest mòdul hem explicat el procés de comparació de dues ditades i també el procés de classificació. S'han descrit els algorismes de visió per computador que són necessaris en aquests processos perquè no podem suposar que l'alumnat del màster tingui coneixements d'aquests algorismes. També s'ha descrit com es representa una ditada en un registre que no sigui una imatge.

S'han descrit els sensors bàsics usats per a capturar una ditada i obtenir-ne la imatge digital i també la diferència que hi ha des del punt de vista tant tecnològic com d'aplicació d'obtenir una imatge en línia i fora de línia.

També és important mencionar la descripció de les diferents classes de ditades. Aquest coneixement pot servir per a saber quins tipus de ditades tenim. També hem après a saber comparar dues ditades donades dues imatges i deduir (òbviament no ens hem fet uns experts forenses) si pertanyen al mateix dit o no.

Finalment, hem après de retruc alguns mètodes de visió per computador. Aquest mètodes són normalment usats en processos industrials més senzills. Però com que el camp de la visió per computador està entrant amb molta força a la indústria, són coneixements importants per a un enginyer o un informàtic.

Activitats

1. Empleneu la taula 3 consultant els apunts i altres fonts d'informació.

Taula 3

	Tecnologia	Avantatges	Inconvenients
Òptic			
Capacitància			
Tèrmic			
Ultrasò			

2. En un sensor de ditada de 200 píxels per polsada i en un de 600 píxels per polsada, quants píxels ocupa una serralada i una vall? Quina creieu que pot ser la resolució mínima d'un sensor?

3. Marqueu sobre les imatges de la figura 34 les tres possibles singularitats i classifiqueu-ne les ditades.

Figura 34



4. Marqueu sobre la imatge de la figura 35 dues *minutiae* que siguin una de cada tipus. A més, empleneu la taula 4. Suposem que la coordenada (0, 0) és la inferior esquerra. Suposem que 1 píxel ocupa 1 mil·límetre.

Figura 35

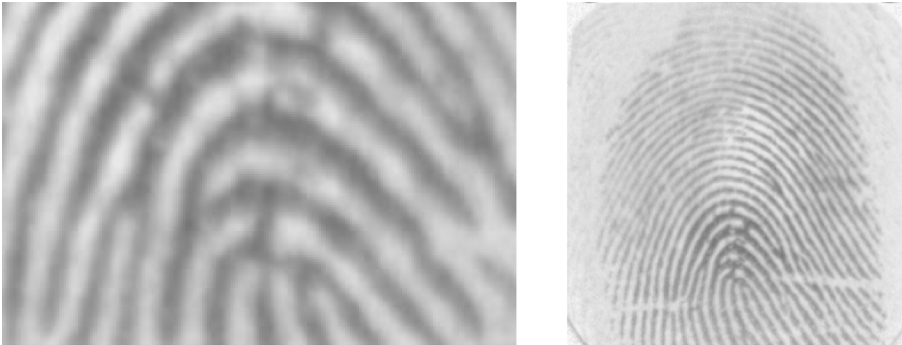


Taula 4

	Tipus	Coordenada x	Coordenada y	Angle (radians)
1	Terminal			
2	Bifurcació			

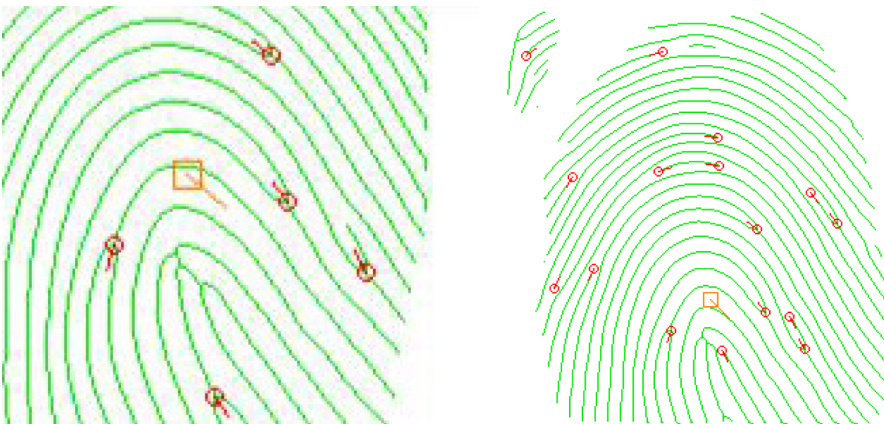
5. Comenteu quins són els tres nivells d'abstracció d'informació de les ditades. Descriviu també quins nivells són els més usats i des de quan.
6. Comenteu per què a la matriu d'orientació només hi ha angles compresos en dos quadrants.
7. Per què és necessari treballar sobre la imatge d'orientació filtrada en lloc de la imatge d'orientació original en els processos finals de la comparació de ditades?
8. Comenteu quin és l'objectiu de la matriu de consistència i la seva necessitat.
9. Volem verificar si una imatge d'una ditada parcial (figura 36a) forma part d'una ditada (figura 36b).

Figura 36



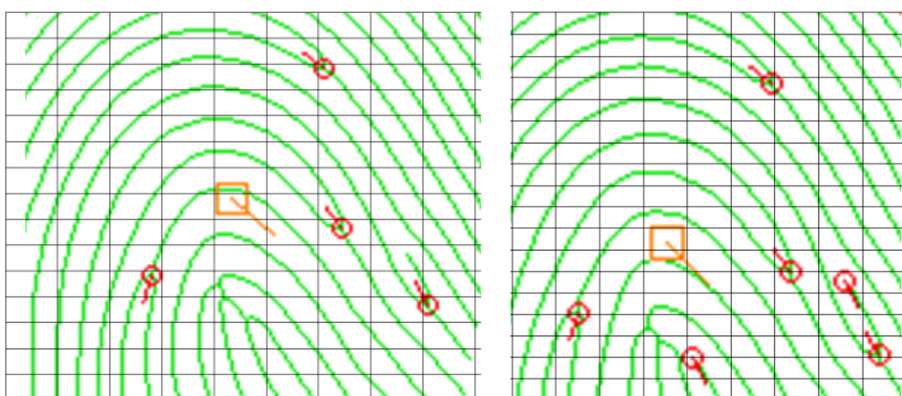
Per fer-ho, primer hem extret l'esquelet de les serralades i també les *minutiae* (cercles en roig) i el punt central de la ditada (quadrat roig).

Figura 37



Llavors, hem alineat les imatges com es mostra en la figura 38.

Figura 38



Decidiu un bon alineament entre *minutiae* i calculeu el cost d'aquest alineament. Les cel·les representen la resolució que cal considerar a l'hora de calcular la distància. La cel·la (0, 0) és la de l'esquerra a baix. Mostreu l'equació usada per a calcular el cost de l'alineament.

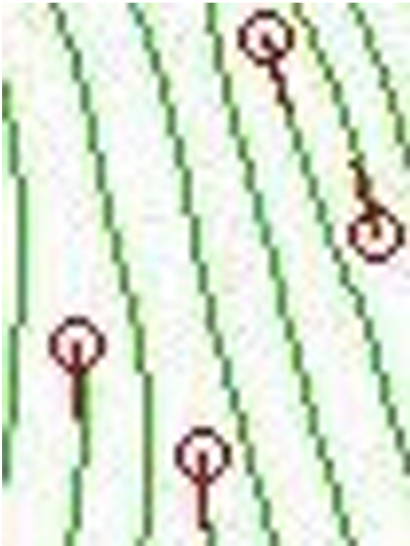
10. Donades les imatges de ditades parcials de les figures 39a i 39b i els registres obtinguts, calculeu l'alineació i el seu cost. Essent els valors de les *minutiae*:

a) {[T, (9, 44), 180°], [T, (25, 58), 181°], [T, (37, 8), 202°], [T, (36, 24), 10°]}

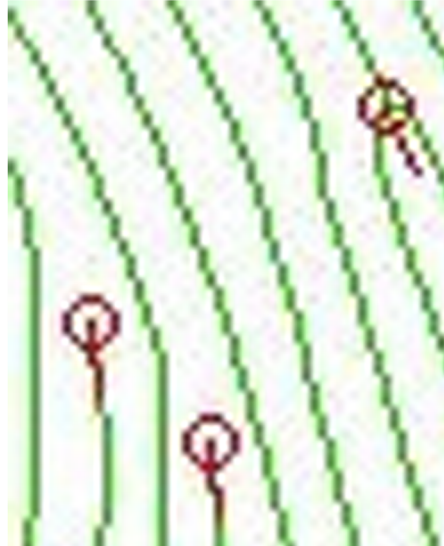
b) {[T, (10, 49), 181°], [T, (26, 58), 182°], [B, (39, 16), 206°]}

Figura 39

a)



b)



Abreviatures

- α_{xy} Angle del gradient màxim de la serralada en el punt (x, y) . És ortogonal a θ_{xy}
- Δ Singularitat tipus delta (delta)
- Δ_x Matriu del gradient horitzontal. $\hat{\Delta}_x$ matriu del gradient horitzontal filtrada
- Δ_y Matriu del gradient vertical. $\hat{\Delta}_y$ matriu del gradient vertical filtrada
- B** Imatge binària
- C** Matriu de coherència
- C** Matriu resultat de convolucionar una imatge **I** amb una màscara **W**
- D** Matriu direccional $D[x, y] \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Matriu dels angles de les serralades $D(x, y) = \theta_{xy}$.
- \hat{D} Matriu direccional filtrada
- E** Imatge esqueletonitzada
- I** Imatge original de la ditada
- I'** Imatge original de la ditada normalitzada
- I i T** Imatges *input* i *template* de les ditades
- O** Singularitat tipus espiral (*whorl*)
- r** Matriu de consistència $r[x, y] \in [0, \infty]$
- S** Matriu segmentada
- U** Singularitat tipus llaç (*loop*)
- W** Màscara de convolució. Normalment és una imatge de 3×3

Bibliografia

Jain, Anil; Flynn, Patrick; Ros, Arun (editors) (2008). *Handbook of biometrics*. Editorial Springer.

Maltoni, Davide; Maio, Dario; Jain, Anil; Prabhakar, Salil (2009). *Handbook of fingerprint recognition*. Editorial Springer.

Annex

Les matrius de píxels o imatges no normalitzades que anem obtenint en el procés de comparació de dues ditades, com, per exemple, la matriu del gradient horitzontal Δ_x i la matriu del gradient vertical Δ_y , la matriu de consistència (norma del gradient) r o la matriu de l'angle del gradient D o la matriu de l'angle filtrada \hat{D} no es poden visualitzar directament, ja que no es troben en el rang normalitzat dels píxels $[0, 255]$. Per aquest motiu, quan ens mostren aquestes imatges, en realitat no és directament el valor sinó una imatge normalitzada. El procés de normalització depèn del rang de valors original d'aquestes matrius. A continuació, es mostren tres processos de normalització que hem comprovat que funcionen bé.

1) Normalització de les matrius dels gradients

Mostrem els valors per a Δ_x però lògicament també serveix per a Δ_y . Les imatges normalitzades són Δ'_x i Δ'_y . Veiem que el rang de valors és:

$$\Delta_x[x, y] \in [-4 \cdot 255, 4 \cdot 255] \quad (51)$$

Com que poques vegades s'assoleixen valors extrems i la mitjana està centrada en el zero, llavors es decideix truncar aquests valors extrems i afegir-hi un desplaçament que és la meitat del rang màxim normalitzat:

$$\Delta'_x[x, y] = \text{Truncate}_{[0, 255]}(\Delta_x[x, y] + 127) \in [0, 255] \quad (52)$$

2) Normalització de la matriu de consistència

Com que la consistència sempre pren valors positius, simplement truncarem el valor inicial:

$$r[x, y] = \text{Truncate}_{[0, 255]}(r[x, y]) \in [0, 255] \quad (53)$$

3) Normalització de la matriu direccional

El rang de valors de la matriu direccional és:

$$D[x, y] \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad (54)$$

La transformació que es proposa està formada per tres passos. Primer, convertim el rang de valors a positiu sumant $\frac{\pi}{2}$. Segon, normalitzem perquè el rang sigui de 0 a 1 dividint per π . I tercer, multipliquem per 255 per tenir el rang complet. L'expressió final és:

$$D[x, y] = \frac{D[x, y] + \frac{\pi}{2}}{\pi} \cdot 255 \in [0, 255] \quad (55)$$

Fixeu-vos que el negre absolut, valor 0, equival a un angle original de $-\frac{\pi}{2}$. El valor gris, 127, equival a l'angle 0. El valor blanc, 255, equival a l'angle $\frac{\pi}{2}$. Òbviament, es podria haver usat un altre conveni, en què, per exemple, l'angle 0 fos el negre absolut.

