



UNIVERSITAT DE BARCELONA



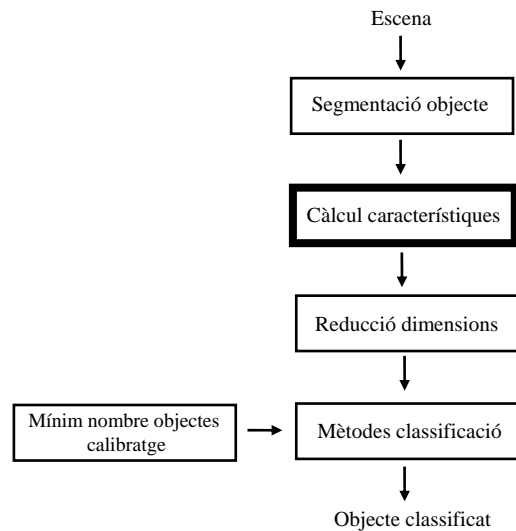
Departament de Física Aplicada i Òptica
Programa de Micro i Optoelectrònica Física
Bienni 1994-96

DISSENY D'UN PROTOCOL NUMÈRIC PER A LA
CLASSIFICACIÓ INVARIANT D'IMATGES APLICANT
TÈCNIQUES MULTIVARIANTS

Memòria presentada per optar al títol de doctor en Ciències Físiques

Directors:
Dr. Arturo Carnicer González
Dr. Ignacio Juvells Prades

Jordi-Roger Riba Ruíz
Barcelona, maig de 2000



2. Característiques discriminants

Per realitzar un procés de classificació d'imatges digitalitzades cal extreure la major quantitat d'informació possible de cada imatge. En el problema que es tracta, ens proporciona aquesta informació el nivell de gris de cada píxel de la imatge. Hi ha dues maneres de treballar: la primera consisteix a treballar directament amb la informació proporcionada per tots els píxels de la imatge, mentre que la segona consisteix a concentrar la major part d'aquesta informació en un nombre m reduït de característiques discriminants (generalment $m < 100$). És més eficient calcular les característiques que treballar directament amb els píxels de la imatge (en una imatge de 256×256 píxels hi ha 65.536 píxels). El càlcul d'un nombre reduït de característiques comporta una gran simplificació del problema i fa que les etapes posteriors del procés de reconeixement siguin més econòmiques en termes de temps de càlcul i memòria utilitzada.

En molts problemes de classificació, i en particular en els que es tracten en aquest treball, és necessari que aquestes característiques siguin invariants a les transformacions geomètriques que no deformen l'objecte com són les translacions, els girs i els canvis d'escala. D'aquesta manera ens assegurarem que el valor d'aquestes característiques es manté invariable quan l'objecte canvia de posició dintre de la imatge, quan l'objecte gira respecte al sistema d'adquisició de les imatges o quan l'objecte s'apropa o s'allunya, fet que seria molt més complicat treballant directament amb els píxels de la imatge.

S'acaba de veure que les característiques discriminants són un conjunt de variables que permeten definir o caracteritzar un objecte. Aquestes, si estan ben definides, han de proporcionar valors similars per a objectes que s'assemblin i valors força diferents per a objectes distints. Per tant, són els fonaments per dur a terme un procés de classificació.

És molt important disposar d'un bon conjunt de característiques adequades, que siguin capaces d'extreure la màxima informació possible dels objectes. Si aquestes no són adequades, difícilment es podrà dur a terme un procés eficient de reconeixement de patrons, ja que els algorismes matemàtics que s'apliquen a les dades proporcionades per les característiques faran servir informació parcial o distorsionada.

També és molt important que aquestes característiques siguin computacionalment eficients: han de ser ràpides de calcular i els seus algorismes han d'evitar en la mesura del possible realitzar suposicions i interpolacions que puguin distorsionar la informació.

A causa, en part, a la pròpia variabilitat dels objectes d'una mateixa classe i també, en part, al propi procés de digitalització, els objectes que pertanyen a una mateixa classe proporcionaran valors diferents per a la mateixa característica. Per tant, hi haurà una dispersió dels valors de la mateixa característica dintre de la mateixa classe d'objectes. És important que aquesta dispersió sigui al més petita possible i que el conjunt de valors d'una determinada característica es trobi al més separat possible entre classes diferents.

La figura següent mostra la variabilitat d'objectes hipotètics de dues classes diferents:

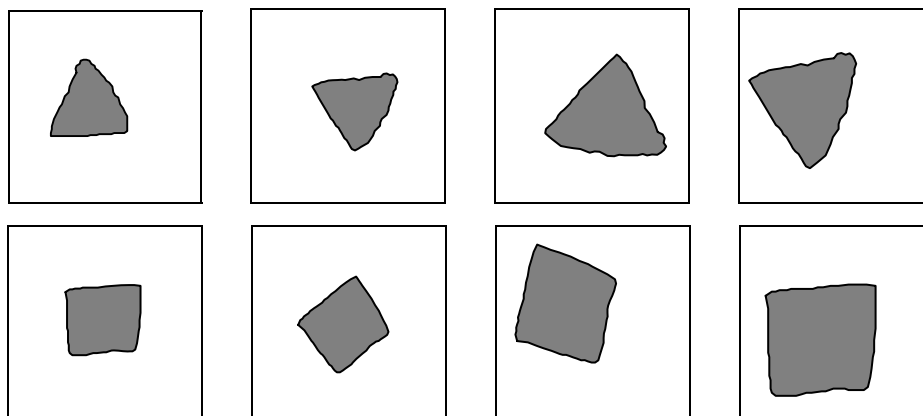


Figura 2.1.

Si es calculen dues característiques x_1 i x_2 per als objectes de la figura anterior, es poden obtenir gràfics de la dispersió del valor de les característiques de cada classe similars als següents:

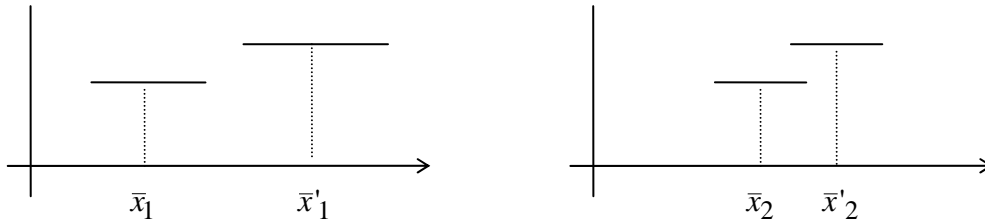


Figura 2.2

En la figura anterior, \bar{x}_i i \bar{x}'_i representen els valor mitjans de la característica i per a les dues classes d'objectes. En aquest cas, la característica x_i produeix més dispersió de les dades i , al mateix temps, més separació entre les dues classes.

2.1. Els moments geomètrics

Sota aquest nom es descriu un conjunt de característiques discriminants introduïdes per primera vegada per Hu [Hu61] l'any 1961 i ampliades per ell mateix l'any 1962 [Hu62]. Aquestes característiques tenen unes expressions matemàtiques molt similars a les dels moments estadístics. Molts autors han aplicat els moments definits per Hu per efectuar reconeixements invariants de formes. Vegeu les referències [Ree80a], [Tea80], [Abu84], [Dud87], [Teh88], [Jai89], [Bel91], [Sim91], [Flu93], [Gon93], [Mar93] i [She95]. També s'ha demostrat que aquests moments estan relacionats amb les intensitats obtingudes en fer la transformada de Fourier de l'objecte en qüestió ([Ree80b], [Vij86], [Cas87], [Duv87] i [Vij87]).

Moments geomètrics d'ordre $p + q$

Hu [Hu61], [Hu62] defineix els moments geomètrics d'ordre $p + q$ com:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p \cdot y^q \cdot f(x, y) \cdot dx \cdot dy \quad \forall p, q \in \mathbb{N}, \quad (2.1)$$

on $f(x, y)$ representa el nivell de gris de la imatge en el píxel (x, y) .

Si la imatge és discreta i de dimensions $(\text{dim}_x, \text{dim}_y)$, el càlcul es realitza de la manera següent:

$$m_{pq} = \sum_{x=0}^{\text{dim}_x-1} \sum_{y=0}^{\text{dim}_y-1} x^p \cdot y^q \cdot f(x, y) \quad \forall p, q \in \mathbb{N} \quad (2.2)$$

L'àrea de l'objecte, expressada en píxels, està determinada per m_{00} . Si es vol calcular l'àrea real de l'objecte, cal multiplicar aquest valor per l'àrea real d'un píxel. Es necessita, per tant, tenir calibrat el sistema d'adquisició de dades (s'ha de conèixer la relació entre píxels i unitats reals d'àrea).

El **centre de masses** de l'objecte (*CDM* o centre de gravetat) es calcula com:

$$x_{CDM} = \bar{x} = m_{10} / m_{00}, \quad y_{CDM} = \bar{y} = m_{01} / m_{00}. \quad (2.3)$$

Aquests moments no són invariants i, per tant, serà poc útil utilitzar-los de forma directa per a l'extracció de característiques invariants, però són la base per determinar la resta dels moments invariants a translacions, canvis d'escala i rotacions.

Depenent de la forma com agafem els valors de la funció $f(x,y)$, hi ha quatre formes possibles de calcular els moments:

1) $f(x,y) = \begin{cases} 1 & \text{contorn} \\ 0 & \text{resta} \end{cases}$	2) $f(x,y) = \begin{cases} f(x,y) & \text{contorn} \\ 0 & \text{resta} \end{cases}$
3) $f(x,y) = \begin{cases} 1 & \text{objecte} \\ 0 & \text{resta} \end{cases}$	4) $f(x,y) = \begin{cases} f(x,y) & \text{objecte} \\ 0 & \text{resta} \end{cases}$

Taula 2.1.

Cal distingir dos casos: el cas dels objectes formats bàsicament per línies (text, signatures, etc) i el cas dels objectes plens (objectes plans en general, per exemple segells). En el cas dels objectes formats per línies, la possibilitat més utilitzada serà la 1), ja que, en general, aquest tipus d'objectes són binaris. Si en aquest cas ens interessés conservar el nivell de gris, hauríem de considerar la possibilitat 2). Quan treballem amb objectes plens amb nivells de grisos, el cas més general és treballar amb la possibilitat 4), tot i que moltes vegades aquest tipus d'objectes es binaritzen i llavors es treballarà amb la possibilitat 3). Si es treballa amb objectes plens binaris, es passa a treballar directament amb la possibilitat 3).

Moments centrals d'ordre $p + q$

Hu [Hu62] defineix els moments centrals d'ordre $p + q$ d'un objecte, referits al seu centre de masses, com:

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p \cdot (y - \bar{y})^q \cdot f(x, y) \cdot dx \cdot dy \quad \forall p, q \in N \quad (2.4)$$

Si la imatge és discreta, el càlcul es realitza de la manera següent:

$$\mu_{pq} = \sum_{x=0}^{\dim x-1} \sum_{y=0}^{\dim y-1} (x - \bar{x})^p \cdot (y - \bar{y})^q \cdot f(x, y) \quad (2.5)$$

$$\mu_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \cdot \binom{q}{s} \cdot (-\bar{x})^r \cdot (-\bar{y})^s \cdot \left[\sum_{x=0}^{\dim x-1} \sum_{y=0}^{\dim y-1} x^{p-r} \cdot y^{q-s} \cdot f(x, y) \right]$$

$$\mu_{pq} = \sum_{x=0}^{\dim x-1} \sum_{y=0}^{\dim y-1} \left[\sum_{r=0}^p \binom{p}{r} \cdot x^{p-r} \cdot (-\bar{x})^r \right] \cdot \left[\sum_{s=0}^q \binom{q}{s} \cdot y^{q-s} \cdot (-\bar{y})^s \right] \cdot f(x, y)$$

Resulta:

$$\mu_{pq} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \cdot \binom{q}{s} \cdot (-\bar{x})^r \cdot (-\bar{y})^s \cdot m_{p-r, q-s} \cdot f(x, y) \quad \forall p, q \in N \quad (2.6)$$

De la relació anterior, en podem calcular fàcilment els moments centrals a partir dels moments geomètrics:

$$\begin{aligned} \mu_{00} &= m_{00}, \\ \mu_{01} &= \mu_{10} = 0, \\ \mu_{11} &= m_{11} - m_{10} \cdot m_{01} / m_{00}, \\ \mu_{20} &= m_{20} - m_{10}^2 / m_{00}, \\ \mu_{02} &= m_{02} - m_{01}^2 / m_{00}, \\ \mu_{21} &= m_{21} - 2 \cdot \bar{x} \cdot m_{11} - \bar{y} \cdot m_{20} + 2 \cdot \bar{x}^2 \cdot m_{01}, \\ \mu_{12} &= m_{12} - 2 \cdot \bar{y} \cdot m_{11} - \bar{x} \cdot m_{02} + 2 \cdot \bar{y}^2 \cdot m_{10}, \\ \mu_{30} &= m_{30} - 3 \cdot \bar{x} \cdot m_{20} + 2 \cdot \bar{x}^2 \cdot m_{10}, \\ \mu_{03} &= m_{03} - 3 \cdot \bar{y} \cdot m_{02} + 2 \cdot \bar{y}^2 \cdot m_{01} \dots \end{aligned} \quad (2.7)$$

Aquests moments, en estar referits al centre de masses de l'objecte, són invariants respecte a translacions d'aquest dintre de la imatge, però no ho són ni a canvis d'escala ni a rotacions.

Moments invariants a canvis d'escala

Hu [Hu62] els defineix de la manera següent:

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma}, \quad \text{on } \gamma = \frac{p+q+2}{2}. \quad (2.8)$$

El fet de dividir els moments centrals per $(\mu_{00})^\gamma$ respon a la necessitat de normalitzar-los en relació amb la seva àrea.

Aquests moments estan definits a partir dels moments invariants a translacions i, per tant, a part de ser invariants a canvis d'escala, també ho són a translacions.

També es poden definir a partir dels moments geomètrics:

$$\eta_{pq} = m_{pq}/(m_{00})^\gamma \quad (2.9)$$

En aquest últim cas, però, les magnituds que resulten només són invariants a canvis d'escala i no ho són a translacions de l'objecte.

Moments invariants a girs

Hu [Hu62] els defineix a partir dels moments invariants a canvis d'escala (que al mateix temps són invariants a translacions) de la manera següent:

$$\begin{aligned} \varphi_1 &= \eta_{20} + \eta_{02}, \\ \varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4.\eta_{11}^2, \\ \varphi_3 &= (\eta_{30} - 3.\eta_{12})^2 + (3.\eta_{21} - \eta_{03})^2, \\ \varphi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2, \\ \varphi_5 &= (\eta_{30} - 3.\eta_{12}).(\eta_{30} + \eta_{12}).[(\eta_{30} + \eta_{12})^2 - 3.(\eta_{21} + \eta_{03})^2] + \\ &\quad + (3.\eta_{21} - \eta_{03}).(\eta_{21} + \eta_{03}).[3.(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\ \varphi_6 &= (\eta_{20} - \eta_{02}).[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4.\eta_{11}.(\eta_{30} + \eta_{12}).(\eta_{21} + \eta_{03}), \\ \varphi_7 &= (3.\eta_{21} - \eta_{03}).(\eta_{30} + \eta_{12}).[(\eta_{30} + \eta_{12})^2 - 3.(\eta_{21} + \eta_{03})^2] + \\ &\quad + (3.\eta_{12} - \eta_{30}).(\eta_{21} + \eta_{03}).[3.(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned} \quad (2.10)$$

Aquests moments són invariants a translacions, a canvis d'escala, i a canvis d'orientació o girs de l'objecte. Per tant, en ser invariants respecte a totes les transformacions

geomètriques que no deformen l'objecte, seran els moments més interessants que caldrà tenir en compte amb vista a extreure característiques d'aquelles, per aplicar-les posteriorment a un reconeixement de formes.

El procés matemàtic desenvolupat per Hu [Hu62] que porta a obtenir les expressions anteriors és molt llarg i bastant complex, però es pot resumir de la manera següent:

Per cada objecte, tal com mostra la figura 2.3., es pot definir una direcció única de referència anomenada *eix de mínima inèrcia* (eix x' de la figura) i també anomenada *eix principal* de l'objecte, que es defineix com la recta que, passant pel centre de masses de l'objecte, compleix que la suma dels quadrats de les distàncies d'aquesta a tots els píxels de l'objecte és mínima. L'eix y' de la figura és la recta que passa pel centre de masses de l'objecte i és perpendicular a l'eix de mínima inèrcia.

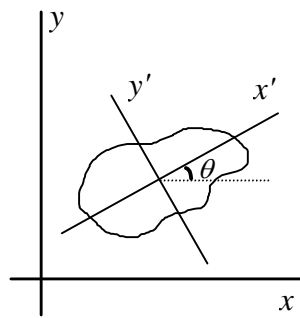


Figura 2.3.

Si es transformen els moments η_{pq} i es calculen respecte als nous eixos x' i y' , es pot arribar a les expressions dels moments invariants a girs.

2.2. Moments obtinguts a partir de la transformada de Fourier-Mellin

Yajun Li [Li92] va definir una integral generadora de moments invariants. Va prendre com a base la transformada de Mellin-Fourier (transformada radial de Mellin i circular de Fourier):

$$F_{s,l} = \int_0^{+\infty} \int_0^{2\pi} r^{s-1} \cdot f(r, \theta) \cdot \exp(-jl\theta) \cdot dr \cdot d\theta \quad \text{amb } s = 1, 2, 3, \dots \text{ i } l = 0, 1, 2, \dots \quad (2.11)$$

En l'expressió anterior, $f(r, \theta)$ és la funció que dona el nivell de gris de cada píxel, expressada en coordenades polars ($x = r \cdot \cos \theta$, $y = r \cdot \sin \theta$, $r = (x^2 + y^2)^{1/2}$).

La integral $F_{s,l}$ és invariant respecte a translacions, ja que tots els càlculs s'efectuen agafant com a origen de les coordenades el CDM de l'objecte.

Per tal de dotar la transformada de Fourier-Mellin d'invariància als canvis d'escala, es normalitza pel factor $F_{2,0}^{s/2} = m_{0,0}^{s/2}$, resultant la funció invariant a translacions i a canvis d'escala següent:

$$G_{s,l} = \frac{F_{s,l}}{m_{0,0}^{s/2}} \quad (2.12)$$

Si girem l'objecte un angle α , la versió girada de la transformada integral $G_{s,l}$ anterior és:

$$I_{s,l} = \exp(jl\alpha) \cdot G_{s,l} \quad (2.13)$$

Per tant, es pot preveure que una combinació apropiada de $I_{s,l}$ amb el seu complex conjugat pot eliminar l'exponencial complexa procedent d'haver efectuat el gir. Aquest fet dota la funció resultant d'invariància respecte a translacions, canvis d'escala i girs.

Per exemple, es pot fer:

$$I_{s,l} \cdot I_{s,l}^* = |G_{s,l}|^2 \quad (2.14)$$

La transformada a partir de la qual es podran calcular els invariants absoluts ve donada per l'expressió:

$$I_{s,l} = \frac{\exp(jl\alpha)}{m_{0,0}^{s/2}} \cdot \int_0^{+\infty} \int_0^{2\pi} r^{s-1} \cdot f(r, \theta) \cdot \exp(-jl\theta) \cdot dr \cdot d\theta \quad \text{amb } s = 1, 2, 3, \dots \text{ i } l = 0, 1, 2, \dots \quad (2.15)$$

Yajun Li demostra que:

$$I_{s,l} = \exp(jl\alpha) \cdot (a_{s,l} - j \cdot b_{s,l}), \quad (2.16)$$

on:

$$a_{s,l} = \sum_{i=0,2,4,\dots}^{s-2} \sum_{j=0,2,4,\dots}^l (-1)^j \binom{l}{j} \binom{n}{(i-j)/2} \eta_{p,q}$$

$$b_{s,l} = \sum_{i=1,3,5,\dots}^{s-2} \sum_{j=1,3,5,\dots}^l (-1)^{(j-1)/2} \binom{l}{j} \binom{n}{(i-j)/2} \eta_{p,q}$$

amb $l > j$, $p + q = s-2$ i $q = i$.

A partir de les expressions de Yajun Li s'obtenen, entre altres, els moments de Hu:

$$\begin{aligned} \varphi_1 &= I_{4,0} \\ \varphi_2 &= I_{4,2} \cdot I_{4,2}^* \\ \varphi_3 &= I_{5,1}^3 \cdot I_{5,3}^* + (I_{5,1}^3 \cdot I_{5,3}^*)^* \\ \varphi_4 &= I_{5,3} \cdot I_{5,3}^* \\ \varphi_5 &= [I_{5,1}^3 \cdot I_{5,3}^* - (I_{5,1}^3 \cdot I_{5,3}^*)^*] / j \\ \varphi_6 &= I_{5,1} \cdot I_{5,1}^* \\ \varphi_7 &= I_{5,1}^2 \cdot I_{4,2}^* + (I_{5,1}^2 \cdot I_{4,2}^*)^* \end{aligned} \quad (2.17)$$

Combinant adequadament les expressions $I_{s,l}$ es poden calcular moltes més característiques invariants. En l'article de Yajun Li n'hi ha més de cinquanta, però les més adequades són les de menor ordre (de φ_1 a φ_7), ja que no es veuen tan afectades per variacions aleatòries i sorolls, tal com es demostra experimentalment en [Teh88].

2.3. Moments de Zernike

Els moments de Zernike van ser desenvolupats per M. Reed [Ree80a] a partir dels polinomis de Zernike. Són moments invariants a translacions, girs i canvis d'escala. Realment, són molt similars als moments invariants φ_i estudiats anteriorment; de fet, només en difereixen en un terme constant i es poden expressar com:

$$\begin{aligned} S_1 &= 3 \cdot [2 \cdot (\eta_{20} + \eta_{02}) - 1] / \pi, \\ S_2 &= 9 \cdot [(\eta_{20} - \eta_{02})^2 + 4 \cdot \eta_{11}^2] / \pi^2, \\ S_3 &= 16 \cdot [(\eta_{03} - 3 \cdot \eta_{21})^2 + (\eta_{30} - 3 \cdot \eta_{12})^2] / \pi^2, \\ S_4 &= 144 \cdot [(\eta_{03} + \eta_{21})^2 + (\eta_{30} + \eta_{12})^2] / \pi^2, \\ S_5 &= 13824 \cdot [(\eta_{03} - 3 \cdot \eta_{21}) \cdot (\eta_{03} + \eta_{21}) \cdot [(\eta_{03} + \eta_{21})^2 - 3 \cdot (\eta_{30} + \eta_{12})^2]] / \pi^4 - \end{aligned}$$

$$\begin{aligned}
 & -13824 \cdot [(\eta_{30} - 3 \cdot \eta_{12}) \cdot (\eta_{30} + \eta_{12}) \cdot [(\eta_{30} + \eta_{12})^2 - 3 \cdot (\eta_{03} + \eta_{21})^2]] / \pi^4 \\
 S_6 & = 864 \cdot [(\eta_{02} - \eta_{20}) \cdot [(\eta_{03} + \eta_{21})^2 - (\eta_{30} + \eta_{12})^2] + 4 \cdot \eta_{11} \cdot (\eta_{03} + \eta_{21}) \cdot (\eta_{30} + \eta_{12})] / \pi^3 \\
 S_7 & = 25 \cdot [(\eta_{40} - 6 \cdot \eta_{22} + \eta_{04})^2 + 16 \cdot (\eta_{31} - \eta_{13})^2] / \pi^2, \\
 S_8 & = 25 \cdot [4 \cdot (\eta_{04} - \eta_{40}) + 3 \cdot (\eta_{20} - \eta_{02})]^2 + 4 \cdot [4 \cdot (\eta_{31} + \eta_{13}) - 3 \cdot \eta_{11}]^2 / \pi^2, \\
 S_9 & = 5 \cdot [6 \cdot (\eta_{40} + 2 \cdot \eta_{22} + \eta_{04}) - 6 \cdot (\eta_{20} + \eta_{02}) + 1] / \pi \dots \\
 (2.18)
 \end{aligned}$$

2.4. Moments de Reiss

Reiss [Rei91] va definir uns moments invariants a totes les transformacions geomètriques no deformatives. Aquests es calculen a partir dels moments invariants a translacions definits per Hu. Els moments invariants de Reiss es calculen a partir de les expressions següents:

$$\begin{aligned}
 I_1 & = \mu_{20} \cdot \mu_{02} - \mu_{11}^2 \\
 I_2 & = (\mu_{30} \cdot \mu_{03} - \mu_{21} \cdot \mu_{12})^2 - 4 \cdot (\mu_{30} \cdot \mu_{12} - \mu_{21}^2) \cdot (\mu_{21} \cdot \mu_{03} - \mu_{12}^2) \\
 I_3 & = \mu_{20} \cdot (\mu_{21} \cdot \mu_{03} - \mu_{12}^2) - \mu_{11} \cdot (\mu_{30} \cdot \mu_{03} - \mu_{21} \cdot \mu_{12}) + \mu_{02} \cdot (\mu_{30} \cdot \mu_{12} - \mu_{21}^2) \\
 I_4 & = \mu_{03}^2 \cdot \mu_{20}^3 - 6 \cdot \mu_{11} \cdot \mu_{12} \cdot \mu_{03} \cdot \mu_{20}^2 - 6 \cdot \mu_{02} \cdot \mu_{21} \cdot \mu_{03} \cdot \mu_{20}^2 + 9 \cdot (\mu_{20} \cdot \mu_{12})^2 \cdot \mu_{02} + 12 \cdot \mu_{20} \cdot \mu_{21} \cdot \mu_{03} \cdot \mu_{11}^2 \\
 & + 6 \cdot \mu_{20} \cdot \mu_{11} \cdot \mu_{02} \cdot \mu_{30} \cdot \mu_{03} - 18 \cdot \mu_{20} \cdot \mu_{11} \cdot \mu_{02} \cdot \mu_{21} \cdot \mu_{12} - 8 \cdot \mu_{30} \cdot \mu_{03} \cdot \mu_{11}^3 - 6 \cdot \mu_{20} \cdot \mu_{30} \cdot \mu_{12} \cdot \mu_{02}^2 \\
 & + 9 \cdot (\mu_{02} \cdot \mu_{21})^2 \cdot \mu_{20} + 12 \cdot \mu_{02} \cdot \mu_{30} \cdot \mu_{12} \cdot \mu_{11}^2 - 6 \cdot \mu_{11} \cdot \mu_{30} \cdot \mu_{21} \cdot \mu_{02}^2 + \mu_{30}^2 \cdot \mu_{02}^3 \quad (2.19)
 \end{aligned}$$

Finalment, resulten els moments de Reiss, donats per:

$$\begin{aligned}
 \text{Reiss}_1 & = I_1 / \mu_{00}^4, \quad \text{Reiss}_2 = I_2 / \mu_{00}^{10}, \\
 \text{Reiss}_3 & = I_3 / \mu_{00}^7, \quad \text{Reiss}_4 = I_4 / \mu_{00}^{10} \quad (2.20)
 \end{aligned}$$

Reiss també va definir tres moments més invariants a translacions, canvis d'escala, rotacions i canvis de contrast de l'objecte, donats per les expressions següents:

$$\Gamma_1 = \frac{I_4}{\mu_{00} \cdot I_2}, \quad \Gamma_2 = \frac{I_1^2}{\mu_{00} \cdot I_3} \quad \text{i} \quad \Gamma_3 = \frac{I_1 \cdot I_3}{I_4} \quad (2.21)$$

2.5. Característiques extrems d'un disc centrat en el CDM de l'objecte

En aquest treball es proposen unes noves característiques invariants que s'extrauen d'un disc. Aquest disc s'ha de trobar centrat en el CDM de l'objecte i el seu radi ha de ser proporcional a l'arrel quadrada de l'àrea de l'objecte.

El radi del disc es calcula com:

$$r = \text{factor} \cdot \sqrt{\text{àrea de l'objecte}}, \quad \text{on } \text{factor} \in \mathbb{R}^+ \quad (2.22)$$

Si les coses es fan d'aquesta manera, es poden obtenir característiques invariants a totes les transformacions geomètriques que no deformin l'objecte.

La invariància respecte a translacions de l'objecte dins de la imatge ve donada pel centrat del disc en el CDM de l'objecte.

La invariància respecte a canvis d'escala ve donada pel fet d'agafar un disc el radi del qual sigui proporcional a l'arrel quadrada de l'àrea de l'objecte. Així, en variar la mida de l'objecte, l'àrea del disc variarà proporcionalment amb la de l'objecte.

La invariància respecte a girs ve donada per la forma geomètrica del disc, el qual és simètric en totes les direccions.

La figura 2.4. mostra com s'agafa el disc centrat en el CDM de l'objecte.

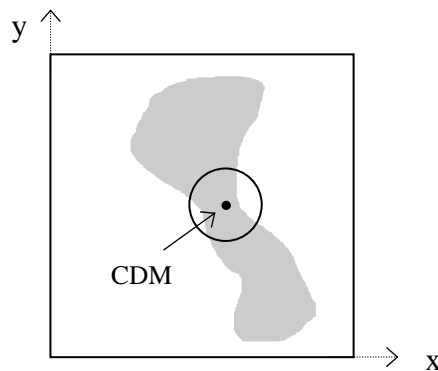


Figura 2.4.

Totes les característiques extrems del disc (es troben exposades en els apartats 2.5.1. i 2.5.2.) presenten un tret comú: per calcular-les cal fer una transformació de les coordenades cartesianes de l'objecte a les coordenades polars. En aquesta transformació s'han de fer interpolacions matemàtiques i aquest fet provoca una certa dispersió en els

valors d'aquestes característiques. Com menys nombre de píxels contingui l'objecte, més dispersió s'obindrà en els valors de les característiques.

Les característiques extretes del disc es poden calcular per cercles de radis diferents, simplement variant el valor del paràmetre factor ($r = factor \cdot \sqrt{\text{àrea de l'objecte}}$), i així s'obtinran diferents conjunts de característiques. Els valors òptims del paràmetre *factor* depenen del tipus d'objectes tractats i s'han d'estudiar en cada cas.

2.5.1. Imatges binàries

En el cas de considerar imatges binàries, les característiques que es poden extreure del disc són:

1. Tant per cent de píxels plens (pertanyents a l'objecte) en el disc respecte al nombre total de píxels del disc. L'anomenarem la característica **%AD** (àrea del disc).
2. Tant per cent de píxels plens en el disc respecte al nombre total de píxels de l'objecte. L'anomenarem la característica **%AO** (àrea de l'objecte).

2.5.2. Imatge amb nivells de grisos

Quan es treballa amb imatges amb nivells de grisos es pot definir un nombre més gran de característiques extretes del disc. Això es deu al fet que tenim més informació. En aquest cas, les característiques que es poden calcular són:

1. Tant per cent de píxels plens (pertanyents a l'objecte) en el disc respecte al nombre total de píxels del disc. L'anomenarem la característica **%AD**.
2. Tant per cent de píxels plens en el disc respecte al nombre total de píxels de l'objecte. L'anomenarem la característica **%AO**.
3. Nivell de gris mitjà dels píxels continguts en el disc. A aquesta característica ,l'anomenarem **GMD** (gris mitjà del disc).
4. Quocient entre el nivell de gris mitjà dels píxels del disc i el nivell mitjà total dels píxels de l'objecte. L'anomenarem la característica **QGM** (quocient de grisos mitjans)
5. Quocient entre l'energia tancada en el disc (l'energia d'un píxel és proporcional al quadrat del valor del nivell de gris d'aquest) i l'energia total de l'objecte. L'anomenarem característica **QE** (quocient d'energies).
6. Quocient entre el nombre de píxels de dintre del disc que tenen un nivell de gris superior a un valor donat (nivell de gris de tall) i el nombre total de

píxels del cercle. Aquesta característica pot funcionar millor o pitjor depenent del nivell de gris de tall que agafem. Per tant, cal temptejar per a quin valor ens resulta un poder discriminant superior. A aquesta característica, l'anomenarem *QGT* (quocient de gris de tall).

És fàcil d'entendre que totes les característiques extretes del disc, tant en el cas d'imatges binàries com en el de nivells de grisos, són invariants respecte a translacions, rotacions i canvis d'escala que pugui experimentar l'objecte dintre de la imatge. Això es deu al fet que totes les característiques són valors relatius, és a dir, magnituds adimensionals.

2.6. Contorn i perímetre de l'objecte

Es pot definir el **contorn** d'un objecte com el conjunt de punts que pertanyent a l'objecte i fan frontera entre l'objecte i el fons (el fons és la part de la imatge que no és l'objecte). És, per tant, el conjunt de punts que pertanyen a l'objecte i que tenen com a mínim un punt veí del fons.

El **perímetre** de l'objecte es defineix com la longitud del seu contorn, és a dir, la longitud total de la corba que descriuen els píxels que pertanyen al contorn. Per calcular el perímetre d'un objecte que ha estat prèviament segmentat, primer cal marcar el contorn de l'objecte i després calcular la longitud d'aquesta corba.

El perímetre ens serveix per caracteritzar un objecte. A més, és invariant a translacions i rotacions de l'objecte, però no és invariant a canvis d'escala. Relacionant adequadament l'àrea de l'objecte i el perímetre es pot arribar a una característica que també sigui invariant a canvis d'escala.

El procés de càlcul del perímetre d'un objecte és bastant més complex que el càlcul de la seva àrea. Per entendre el procediment de còmput del perímetre, primer cal introduir el concepte de *connectivitat* [Par93].

2.6.1. Connectivitat

Donat un píxel P qualsevol de l'objecte, sempre té vuit veïns immediats, tal com mostra la figura 2.5.:

8 $(i-1,j-1)$	1 $(i-1,j)$	2 $(i-1,j+1)$
7 $(i,j-1)$	<i>P</i> (i,j)	3 $(i,j+1)$
6 $(i+1,j-1)$	5 $(i+1,j)$	4 $(i+1,j+1)$

Figura 2.5.

Els veïns verticals i horitzontals (píxels 1, 3, 5 i 7) del píxel *P* es troben a una distància d'unitat d'aquest, mentre que els veïns diagonals (píxels 2, 4, 6 i 8) es troben a una distància $\sqrt{2}$ d'aquest.

Dos píxels plens són 4-connectius si són veïns verticals o horitzontals. La figura següent mostra les diferents possibilitats:

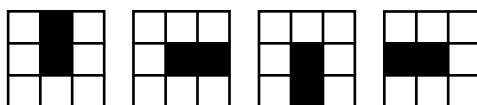


Figura 2.6.

Dos píxels plens són 8-connectius si són adjacents. La figura següent mostra les diferents possibilitats:

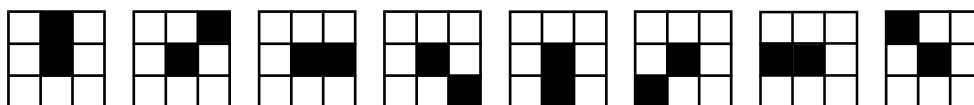


Figura 2.7.

2.6.2 Càlcul del perímetre

Perquè un píxel formi part del contorn de l'objecte, cal que sigui 4-connectiu o 8-connectiu amb el fons. Normalment, però, s'agafa connectivitat 4, ja que experimentalment s'ha comprovat que en general proporciona millors resultats.

L'algorisme de càlcul del perímetre d'un objecte es pot trobar en [Par93] i es detalla a continuació:

1. Per tal de determinar el contorn d'una regió, s'ha d'explorar aquesta píxel a píxel i s'ha d'associar un valor determinat, per exemple 255, als píxels que no pertanyen a l'objecte (píxels del fons).
2. Després cal identificar i marcar els píxels del contorn. És fàcil de trobar-los, ja que compleixen la condició que són 4-connectius (o 8-connectius) amb els píxels del fons. Per marcar-los, se'ls associa un valor determinat, per exemple 0.
3. Als píxels de l'objecte que no formen part del contorn, se'ls associa el mateix valor del fons: se'ls marca amb el valor 255.
4. Ara tenim una nova imatge, els píxels de la qual o bé tenen valor 0 (píxels del contorn) o bé tenen valor 255 (els píxels restants). Per calcular el valor del perímetre, s'ha d'explorar aquesta imatge nova píxel a píxel i anar acumulant la distància entre píxels veïns del contorn. La forma de calcular aquesta distància està detallada en el proper apartat.

La figura següent mostra com queda marcat el contorn d'un objecte:

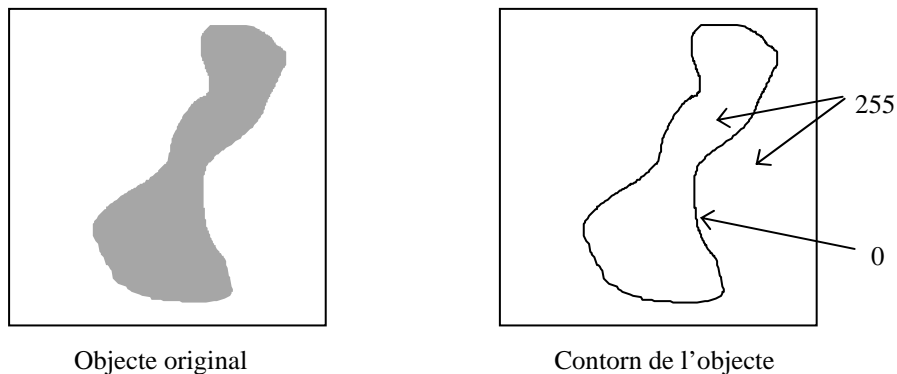


Figura 2.8.

Distància entre píxels veïns del contorn

Per calcular el valor numèric del perímetre, s'ha de determinar la suma de les distàncies entre els píxels veïns del contorn, o, dit d'una altra manera, s'ha de calcular la longitud total de la corba que formen els píxels del contorn.

Per tal de mesurar la distància entre píxels veïns del contorn, es fa servir una matriu de 3 x 3 píxels que se centrarà en cada un dels píxels del contorn. A cada casella d'aquesta matriu s'associa un valor determinat (aquests valors han de ser tals que la suma de qualsevol combinació de dos d'aquests no pugui coincidir amb la de qualsevol altra combinació), tal com mostra la figura següent:

1	2	4
8		16
32	64	128

Figura 2.9.

Segons el valor numèric que resulti, sabrem el valor que haurem de donar a la distància entre els píxels veïns, com s'explica a continuació.

Les figures següents mostren com es calcula la distància entre els píxels veïns del contorn. El píxel gris és el píxel del contorn on es troba centrada la matriu, mentre que els negres representen els píxels veïns que pertanyen al contorn de l'objecte.

- Grup amb $d = (1+1)/2 = 1$:

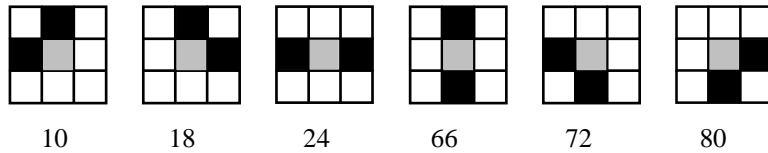


Figura 2.10.

- Grup amb $d = (\sqrt{2} + \sqrt{2})/2 = 1,4142$:

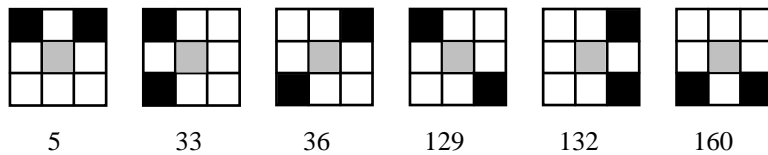
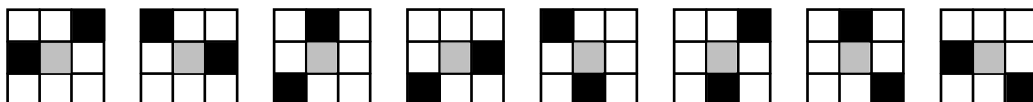


Figura 2.11.

- Grup amb $d = (1 + \sqrt{2})/2 = 1,2071$:



12 17 34 48 65 68 130 136

Figura 2.12.

Les distàncies anteriors s'han dividit per dos per tal de calcular el valor mitjà de les distàncies dels píxels del contorn al píxel central de la matriu.

Com que els casos anteriors no representen la totalitat dels casos possibles, associarem el valor $d = 1,2071$ a qualsevol altre cas. Aquest és el que resulta de fer la mitjana aritmètica dels tres casos considerats en les figures anteriors (valor 1, valor 1,4142 i valor 1,2071).

El càlcul del perímetre presenta una certa problemàtica a causa de la naturalesa discreta de les imatges digitalitzades. Per explicar aquesta problemàtica ens centrarem en l'objecte quadrat de 4×4 píxels que mostra la figura 2.13. El perímetre teòric d'aquest objecte és 16 ($4 + 4 + 4 + 4$), però a causa de la discretització, el perímetre que resulta és 12, tal com es pot veure a la figura. En el cas d'un quadrat de 5×5 píxels, el perímetre calculat resultaria de 16, mentre que el teòric és de 20.

1	2	3	4
12			5
11			6
10	9	8	7

Figura 2.13.

2.7. Factor de circularitat

Es pot definir el factor C de circularitat [Par93] de l'objecte com:

$$C = \frac{P^2}{4\pi \cdot A}, \quad (2.23)$$

on P és el perímetre de l'objecte i A n'és l'àrea.

Tot i que ni el perímetre ni l'àrea són invariants a canvis d'escala, el factor de circularitat depèn de la forma de l'objecte i és invariant a translacions, rotacions i canvis d'escala d'aquest, i, per tant, és una característica invariant. Això es deu al fet que és una magnitud adimensional.

El factor de circularitat pot adoptar valors diferents depenent de la forma geomètrica de l'objecte:

- Per a objectes circulars resulta $C = 1$.
- Per a objectes quadrats resulta $C = 1,2732$.
- Per a formes no circulars es té sempre $C > 1$.

2.8. Aprimament o esqueletització de l'objecte

Aquesta tècnica és molt emprada en objectes formats per línies, com ara imatges de text. La tècnica es basa a anar traient iterativament les capes del contorn fins a deixar un objecte format per línies d'un píxel d'amplada, aconseguint així un nou objecte molt més simple que l'original.

La figura següent mostra un exemple d'aprimament:

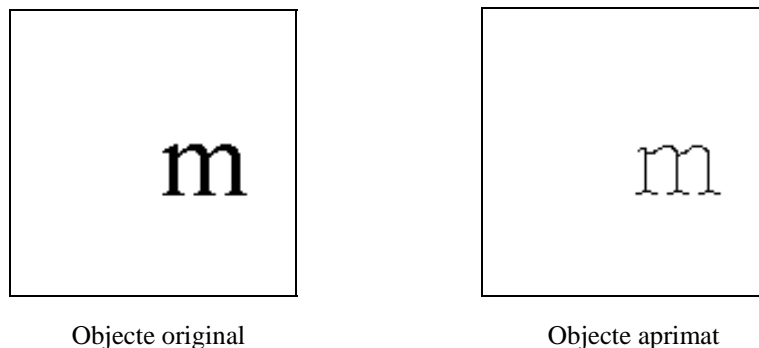


Figura 2.14.

Un dels algorismes d'aprimament més emprats és el de Zhang-Suen [Zha84]. És un algorisme d'aprimament per esborrat successiu i s'explica a continuació:

1. Es fa una escombrada de la imatge i es marquen els píxels del contorn que s'han d'eliminar.
2. En una segona escombrada de la imatge s'eliminen els píxels marcats. Les regles per decidir quan s'eliminen s'expliquen en l'apartat 2.8.2.

3. Es repeteixen els passos anteriors fins que no queden més píxels per eliminar. Els píxels supervivents són els que conformen l'esquelet de l'objecte.

2.8.1. Propietats de l'esquelet

L'esquelet té les propietats següents:

- 1.- Ha d'estar format per corbes primes, d'un píxel d'amplada.
- 2.- Els píxels de l'esquelet han d'estar propers a la secció transversal de l'objecte original.
- 3.- Els píxels de l'esquelet han d'estar connectats entre si i han de formar el mateix nombre de regions que hi havia en l'objecte original.

2.8.2. Regles per decidir si un píxel s'esborra

Per decidir si un píxel es pot esborrar o no, cal mirar els seus vuit veïns immediats i aplicar les regles següents:

- 1.- Un píxel pot ser esborrat si té un mínim de dos veïns i un màxim de sis. Aquesta regla assegura que els punts finals de l'esquelet no s'esborraran i que els píxels s'esborren des de fora cap a dintre del contorn, mai des de dintre de l'objecte cap a fora.
- 2.- Un píxel pot ser esborrat quan només es trobi connectat a una regió. Si es trobés connectat a dues regions i s'esborrés, les dues regions quedarien separades.

Les figures 2.15. i 2.16. mostren les dues condicions límit explicades. En la primera d'aquestes, un píxel es troba envoltat de sis veïns. Com que és un píxel frontera, es pot esborrar. En la segona, el píxel encerclat només té un veí i, per tant, no es pot esborrar.

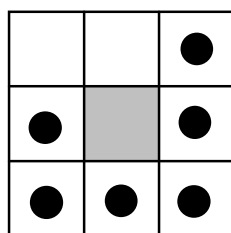


Figura 2.15.

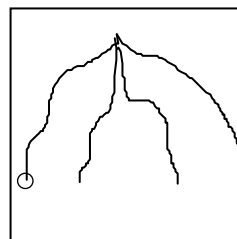


Figura 2.16.

Per poder aprimar una regió (eliminar els píxels redundants) cal aplicar les dues regles anteriors a tots els píxels de la regió i esborrar tots els píxels que satisfacin les dues condicions. El procés es repeteix iterativament fins que no queden més píxels per esborrar, escombrant la imatge de dalt a baix i d'esquerra a dreta. El problema que apareix és que el mateix objecte girat pot portar a esquelets lleugerament diferents, tal com mostra la figura 2.17.:

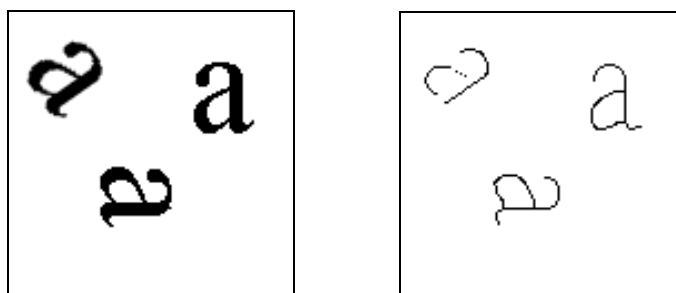


Figura 2.17.

2.9. Moments calculats a partir del contorn o de l'esquelet

Els moments del contorn van ser definits per Gupta i Srinath [Gup87]. Constitueixen un conjunt de característiques invariants, calculades, originalment, a partir del contorn de l'objecte. En aquest treball se n'ha ampliat l'àmbit d'aplicació a l'esquelet de l'objecte, en el cas d'imatges de text o imatges formades per línies.

El contorn d'un objecte es pot descriure com una seqüència ordenada de parelles de coordenades (x_i, y_i) amb $i = 1, 2, \dots, n$. D'aquí es pot calcular la distància (normalment s'agafa l'euclidiana) r_i (amb $i = 1, 2, \dots, n$) del vector que uneix el CDM de l'objecte amb el conjunt ordenat de píxels del contorn. Per tant, s'obté una representació unidimensional del contorn de l'objecte. La figura següent mostra la seqüència dels r_i :

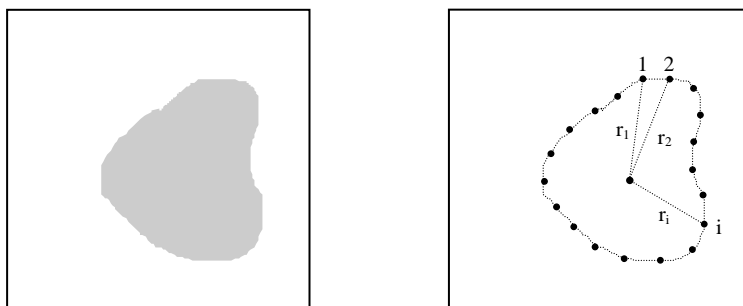


Figura 2.18.

Si tenim en compte contorns tancats, la seqüència obtinguda de valors r_i és cíclica i es compleix:

$$r_{i+n} = r_i \quad \text{amb } i = 1, 2, \dots, n \quad (2.24)$$

La forma de la seqüència r_i obtinguda és diferent per a objectes diferents, i per a objectes similars (de la mateixa classe) les seqüències haurien de ser semblants. La seqüència dels r_i es pot obtenir amb una algorisme que escombri el contorn, per exemple, en sentit horari.

2.9.1. Característiques invariants extrems de la seqüència del contorn

Interessa obtenir característiques invariants a translacions, girs i canvis d'escala a partir de la seqüència del contorn de l'objecte. Suposem que r_i és la seqüència del contorn de l'objecte original i r_i' és la seqüència del contorn després d'haver efectuat una transformació geomètrica a l'objecte.

Invariància a translacions

Encara que l'objecte canviï de posició dintre de l'escena, el seu contorn es manté invariant. Per tant, la nova seqüència r_i' tampoc varia, de manera que resulta:

$$r_i' = r_i \quad \text{amb } i = 1, 2, \dots, n \quad (2.25)$$

Invariància a girs

En girar un objecte respecte a un punt determinat, la forma del seu contorn es manté inalterada. Això ens indica que només canviarà l'ordre de la seqüència de píxels que formen el contorn. Si indiquem la seqüència del contorn girat com r_i' , s'haurà de complir la relació següent:

$$r_i' = r_{i+s} \quad \text{amb } i = 1, 2, \dots, n \quad \text{i } s = -n, -n+1, \dots, 1, 2, \dots, n \quad (2.26)$$

Invariància a canvis d'escala

Quan un objecte canvia d'escala, la forma del seu contorn resulta inalterada. Si el seu perímetre augmenta o disminueix en un factor λ , la seva àrea augmentarà o disminuirà en un factor λ^2 . Per tant, la longitud del contorn i els mòduls dels vectors que connecten el CDM amb els seus punts, augmentaran o disminuiran en un factor λ . Anomenant r_i' la seqüència escalada, s'haurà de complir:

$$r_i' = \lambda \cdot r_i, \text{ amb } i = 1, 2, \dots, n \quad (2.27)$$

Resumint, la seqüència r_i és invariant respecte a translacions de l'objecte. La rotació d'aquesta causa un desplaçament de la seqüència dels r_i i un canvi d'escala de l'objecte provoca un escalat de les amplituds dels vectors r_i .

Es defineixen els moments següents:

Moments del contorn d'ordre q :

$$m_{c,q} = \frac{1}{n} \cdot \sum_{i=1}^n r_i^q \quad (2.28)$$

Moments centrals d'ordre q :

$$M_{c,q} = \frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^q \quad (2.29)$$

Per si mateixos, tots els moments anteriors són invariants a translacions i girs. La invariància a translacions de l'objecte és trivial. La invariància a girs es deu al fet que en realitzar el sumatori, dues seqüències r_i iguals però desplaçades l'una respecte de l'altra condueixen al mateix resultat.

Si volem obtenir unes característiques també invariants a canvis d'escala de l'objecte, cal agafar les característiques anteriors i normalitzar-les de manera que s'obtinguin magnituds adimensionals. Fent això es poden calcular els següents moments invariants a translacions, girs i canvis d'escala de l'objecte:

2.9.2. Moments d'ordre q normalitzats del contorn

$$\bar{m}_{c,q} = \frac{m_{c,q}}{[M_{c,2}]^{q/2}} = \frac{\frac{1}{n} \cdot \sum_{i=1}^n r_i^q}{\left[\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^2 \right]^{q/2}} \quad (2.30)$$

2.9.3. Moments centrals normalitzats d'ordre q

$$\bar{M}_{c,q} = \frac{M_{c,q}}{[M_{c,2}]^{q/2}} = \frac{\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^q}{\left[\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^2 \right]^{q/2}} \quad (2.31)$$

Dels moments anteriors es pot derivar un nombre pràcticament il·limitat de característiques invariants. El problema que tenen és que, en augmentar l'ordre dels moments, al mateix temps se n'augmenta el rang dinàmic i els resultats que produeixen es tornen inestables. Per evitar aquest fet, Gupta i Srinath també van derivar uns moments invariants de baix ordre que són força estables. Aquests moments es troben exposats a continuació:

2.9.4. Variació normalitzada d'amplitud

$$F_{c,1} = \frac{[M_{c,2}]^{1/2}}{m_{c,1}} = \frac{\left[\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^2 \right]^{1/2}}{\frac{1}{n} \cdot \sum_{i=1}^n r_i} \quad (2.32)$$

$F_{c,1}$ és un estadístic de dispersió que es pot interpretar com una mesura de les variacions d'amplitud en la seqüència del contorn. En un contorn perfectament circular el seu valor seria 0.

2.9.5. Coeficient d'asimetria

$$F_{c,2} = \frac{M_{c,3}}{[M_{c,2}]^{3/2}} = \frac{\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^3}{\left[\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^2 \right]^{3/2}} \quad (2.33)$$

$F_{c,2}$ és un factor de forma que fa referència al grau de simetria de la funció r_i . El seu valor és positiu quan la majoria dels píxels del contorn tenen un r_i més petit que el r_i mitjà.

2.9.6. Coeficient de curtosi

$$F_{c,3} = \frac{M_{c,4}}{[M_{c,2}]^{4/2}} = \frac{\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^4}{\left[\frac{1}{n} \cdot \sum_{i=1}^n [r_i - m_{c,1}]^2 \right]^{4/2}} \quad (2.34)$$

$F_{c,3}$ és un factor de forma que fa referència al grau d'agudeses de la funció r_i . El seu valor és positiu quan la majoria dels píxels del contorn tenen un r_i concentrat al voltant del r_i mitjà.

Nota: tots els moments explicats en aquest apartat es poden calcular a partir de l'esquelet, i en aquest cas es canvia el subíndex c pel subíndex e . Per exemple: $m_{e,q}$, $M_{e,q}$, $F_{e,q}$, etc.

2.10. Eix principal d'inèrcia (EPI) i eix menor (EM)

En aquest treball es defineix un conjunt de característiques invariants calculades a partir de l'eix principal d'inèrcia i l'eix menor. Abans, però, s'ha d'explicar quins són aquests eixos.

Es defineix l'eix principal d'inèrcia de l'objecte o EPI [Hu62] [Tea80] com la recta que passant pel CDM de l'objecte, compleix la condició que la suma dels quadrats de les distàncies de tots els punts de l'objecte a aquesta recta és mínima. Donat un objecte determinat, en general l'EPI serà únic, fora que l'objecte en qüestió sigui circular sense forats.

Es defineix l'eix menor de l'objecte o EM com la recta que, passant pel CDM de l'objecte, és perpendicular a l'EPI.

Aquests dos eixos que hem definit són solidaris a l'objecte i sempre estan centrats en el seu CDM. Aquest fet ens indica que si l'objecte es desplaça o gira dintre de l'escena, l'EPI i l'EM també ho faran, fet molt important per poder definir característiques invariants.

2.10.1. Càlcul de les equacions de l'EPI i l'EM

En aquest treball es proposa el següent mètode de càlcul de l'EPI i l'EM.

L'equació general d'una recta es pot expressar com:

$$Ax + By + C = 0. \quad (2.35)$$

La distància al quadrat d'un punt qualsevol $P(x_i, y_i)$ a una recta es calcula com:

$$D = d^2 = \frac{(Ax_i + By_i + C)^2}{A^2 + B^2}. \quad (2.36)$$

$$\text{Equació de l'EPI: } y - y_{cdm} = m.(x - x_{cdm}) \Rightarrow m.(x - x_{cdm}) - (y - y_{cdm}) = 0. \quad (2.37)$$

El pendent de la recta ve donat per m . Per tant, per saber l'expressió de l'EPI només falta per determinar el valor de m . Resulta:

$$A = m, B = -1 \text{ i } C = -m.x_{cdm} + y_{cdm}.$$

Com que l'EM és perpendicular a l'EPI, el producte dels seus pendents ha de ser -1 , resultant que el pendent de l'EM és $-1/m$.

$$\text{Equació de l'EM: } y - y_{cdm} = (-1/m).(x - x_{cdm}) \Rightarrow (x - x_{cdm}) + m.(y - y_{cdm}) = 0 \quad (2.38)$$

La suma dels quadrats de les distàncies de tots els punts de l'objecte a l'EPI es calcula a partir de la fórmula:

$$\sum_{i=1}^n D = \sum_{i=1}^n \frac{[m(x_i - x_{cdm}) - (y_i - y_{cdm})]^2}{m^2 + 1}. \quad (2.39)$$

L'EPI és la recta que minimitza l'expressió anterior. Per determinar-ne el valor del pendent hem d'igualar la primera derivada a zero:

$$\frac{\partial \sum_{i=1}^n D}{\partial m} = 0 \quad (2.40)$$

Desenvolupant l'expressió anterior, resulta que:

$$0 = \sum_{i=1}^n \left[\frac{2[m(x_i - x_{cdm}) - (y_i - y_{cdm})](x_i - x_{cdm})(m^2 + 1) - [m(x_i - x_{cdm}) - (y_i - y_{cdm})]^2 2m}{(m^2 + 1)^2} \right]$$

Arreglant l'expressió anterior i simplificant alguns termes s'arriba a:

$$0 = \sum_{i=1}^n \left[m[(x_i - x_{cdm})^2 - (y_i - y_{cdm})^2] + (m^2 - 1)(y_i - y_{cdm})(x_i - x_{cdm}) \right]$$

D'aquí resulta que:

$$\frac{(m^2 - 1)}{m} = \frac{-\sum_{i=1}^n [(x_i - x_{cdm})^2 - (y_i - y_{cdm})^2]}{\sum_{i=1}^n [(y_i - y_{cdm})(x_i - x_{cdm})]} = c \Rightarrow m^2 - c.m - 1 = 0.$$

Finalment s'obtenen els resultats següents de l'equació anterior:

$$m_1 = \frac{c}{2} + \sqrt{\left(\frac{c}{2}\right)^2 + 1} \quad \text{i} \quad m_2 = \frac{c}{2} - \sqrt{\left(\frac{c}{2}\right)^2 + 1}. \quad (2.41)$$

En ser una equació de segon grau, s'obtenen dues solucions. Fàcilment es pot comprovar que $m_1.m_2 = -1$, fet que ens indica que resulten dos pendents perpendiculars. Per tant, un dels dos pendents correspon a l'EPI i l'altre correspon a l'EM. Per saber quin és quin, s'han de calcular els valors $\sum_{i=1}^n D$ amb cada un dels pendents. El pendent pel qual el valor de l'expressió anterior és més petit correspon a l'EPI, mentre que l'altre correspon a l'EM.

2.11. Característiques que es dedueixen de l'EPI i l'EM

La figura següent mostra un objecte amb els seus EPI i EM. La figura també indica l'angle entre l'EPI i l'eix horitzontal (així definim l'orientació de l'objecte), el mínim rectangle que conté l'objecte i les longituds a i b dels eixos major i menor d'aquest rectangle, respectivament.

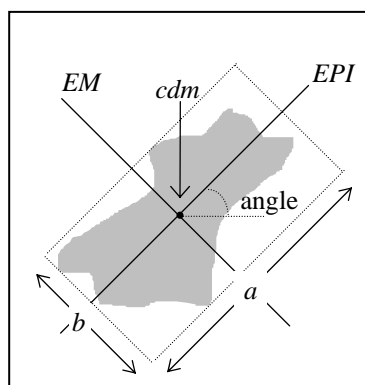


Figura 2.19.

Es poden definir les característiques següents:

2.11.1. Angle α entre l'EPI i l'eix horitzontal

L'orientació de l'objecte es defineix com l'angle α que forma l'EPI de l'objecte amb l'eix horitzontal. Trobar-lo és molt senzill:

$$\operatorname{tg}(\alpha) = m \Rightarrow \alpha = \operatorname{arctg}(m), \quad (2.42)$$

on m és el pendent de l'EPI.

2.11.2. Quocient entre l'àrea de l'objecte i l'àrea del rectangle que el conté

Si A_o és l'àrea de l'objecte i A_r és l'àrea del mínim rectangle que conté l'objecte, el quocient:

$$A_1 = A_o/A_r \quad (2.43)$$

és una característica invariant a translacions, canvis d'escala i girs de l'objecte.

2.11.3. Quocient entre l'àrea de l'objecte i l'àrea de l'el·lipse que el conté

Si A_o és l'àrea de l'objecte i A_e és l'àrea de la mínima el·lipse que conté l'objecte, el quocient:

$$A_2 = A_o/A_e \quad (2.44)$$

és una característica invariant a translacions, canvis d'escala i girs de l'objecte.

L'àrea de l'el·lipse es calcula com:

$$A_e = \pi \cdot (a/2) \cdot (b/2), \quad (2.45)$$

on a és la longitud del costat del rectangle paral·lel a l'EPI i b és la longitud del costat del rectangle paral·lel a l'EM.

2.11.4. Quocient entre la longitud de l'eix menor i la de l'eix major

El quocient

$$E = b/a \quad (2.46)$$

és una característica invariant a translacions, canvis d'escala i girs de l'objecte.

2.11.5. Excentricitat de l'el·lipse

L'excentricitat de la mínima el·lipse que conté l'objecte:

$$\varepsilon = c/a = \sqrt{a^2 - b^2} / a \quad (2.47)$$

és una característica invariant a translacions, canvis d'escala i girs de l'objecte.

2.11.6. Quocient D

El quocient entre la suma del quadrat de les distàncies de tots els píxels de l'objecte a l'EPI i la suma del quadrat de les distàncies de tots els píxels de l'objecte a l'EM:

$$D = \frac{\sum_{i=1}^{\text{núm. píxels}} d_{i,\text{EPI}}^2}{\sum_{i=1}^{\text{núm. píxels}} d_{i,\text{EM}}^2} \quad (2.48)$$

és una magnitud invariant a translacions, girs i canvis d'escala de l'objecte.

2.12. Moments projectats sobre l'EPI i l'EM.

En aquest apartat es calculen uns moments a partir de les projeccions de l'objecte. Els van definir R. Bajaj i S. Chaudhury [Baj97] a partir de les projeccions de l'objecte sobre els eixos naturals (horitzontal i vertical) de la imatge i, per tant, no són invariants. En aquest treball s'han fet invariants calculant les projeccions respecte a l'EPI i l'EM de l'objecte, obtenint-se uns moments invariants a totes les transformacions geomètriques no deformatives de l'objecte.

Es defineixen les projeccions horitzontal i vertical de l'objecte com:

$$P^H(x) = \sum_{y=1}^{\text{dimy}} f(x,y) \quad \text{i} \quad P^V(y) = \sum_{x=1}^{\text{dimx}} f(x,y), \quad (2.49)$$

on $f(x,y)$ és el nivell de gris del punt (x,y) , mentre que dimx i dimy són, respectivament, les dimensions horitzontal i vertical de la imatge.

La figura següent mostra les projeccions horitzontal i vertical de l'objecte:

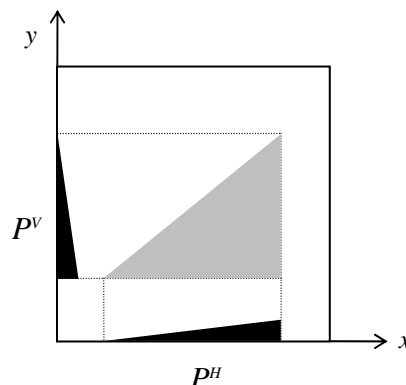


Figura 2.18.

A partir de les projeccions de l'objecte sobre els eixos originals x i y de la imatge, [Baj97] defineix els moments projectats següents:

$$\begin{aligned}\mu_r^H &= \sum_{x=1}^{\dim x} (x - x_{cdm})^r \cdot P^H(x) \\ \mu_r^V &= \sum_{y=1}^{\dim y} (y - y_{cdm})^r \cdot P^V(y)\end{aligned}\tag{2.50}$$

El problema dels moments anteriors és que no són invariants a girs ni a canvis d'escala.

2.12.1. Moments invariants calculats a partir de les projeccions

En aquest treball s'han dotat aquests moments d'invariància respecte a totes les transformacions geomètriques. La invariància respecte a translacions i rotacions s'obté calculant les projeccions de l'objecte respecte a l'EPI i l'EM en comptes de fer-ho sobre els eixos originals x i y de la imatge. Com que l'EPI i l'EM estan centrats en el CDM de l'objecte, la invariància a translacions queda garantida. La invariància a girs es deu al fet que tant l'EPI com l'EM de l'objecte giren en fer-ho aquest. Perquè aquests moments també siguin invariants a canvis d'escala, cal que els moments resultants siguin magnituds adimensionals, i això s'obté fent la normalització següent:

$$\begin{aligned}\eta_r^H &= \frac{\mu_r^H}{(\text{\grave{a}rea de l'objecte})^{r/2+1}} \\ &\text{i} \\ \eta_r^V &= \frac{\mu_r^V}{(\text{\grave{à}rea de l'objecte})^{r/2+1}}\end{aligned}\tag{2.51}$$

L'àrea de l'objecte es pot calcular com:

$$\text{\grave{à}rea de l'objecte} = \sum_{x=1}^{\dim x} P^H(x) = \sum_{y=1}^{\dim y} P^V(y)\tag{2.52}$$

2.12.2. Altres característiques invariants

[Baj97] també defineix les magnituds següents, calculades a partir dels moments de les projeccions:

Coefficient de curtosi de les projeccions horitzontal i vertical:

$$K^H = \frac{\mu_4^H}{(\mu_2^H)^2} \quad \text{i} \quad K^V = \frac{\mu_4^V}{(\mu_2^V)^2} \quad (2.53)$$

Coefficient d'asimetria de les projeccions horitzontal i vertical:

$$S^H = \frac{\mu_3^H}{(\mu_2^H)^{1.5}} \quad \text{i} \quad S^V = \frac{\mu_3^V}{(\mu_2^V)^{1.5}} \quad (2.54)$$

Quocient entre els moments de les projeccions vertical i horitzontal:

$$VH_1 = \frac{\mu_2^V}{\mu_2^H} \quad \text{i} \quad VH_2 = \frac{\mu_4^V}{\mu_4^H} \quad (2.55)$$

Les magnituds acabades de definir, si es calculen respecte a les projeccions sobre l'EPI i l'EM, són invariants respecte a translacions, girs i canvis d'escala de l'objecte.

2.13. Moments calculats a partir de l'histograma

Donada una imatge qualsevol, codificada en L nivells de grisos x , on $x \in [0,1,\dots,L-1]$, definim la freqüència $f(x)$ d'aparició d'un nivell de gris x com:

$$f(x) = \frac{\text{núm. de píxels amb nivell de gris } x}{\text{núm. total de píxels}}, \quad \text{amb } x = 0,1,\dots,L-1. \quad (2.56)$$

L'histograma d'una imatge és una correspondència en la qual s'associa a cada nivell de gris la seva freqüència d'aparició en la imatge. Com que els moments que es descriuen a continuació tenen en compte la freqüència d'aparició de cada nivell de gris, es diu que estan calculats a partir de l'histograma. Es troben definits en [Jain89] com:

Moments generals

$$m_i = \sum_{x=0}^{L-1} x^i \cdot f(x), \quad \text{amb } i = 1,2,\dots \quad (2.57)$$

Moments centrals

$$\mu_i = \sum_{x=0}^{L-1} (x - m_1)^i \cdot f(x), \quad \text{amb } i = 1, 2, \dots \quad (2.58)$$

Entropia de la imatge

$$H = - \sum_{x=0}^{L-1} f(x) \cdot \log_2[f(x)] \quad (2.59)$$

Tots aquests moments calculats a partir de l’histograma tenen en compte el nivell de gris de cada píxel, però no tenen en compte la posició d’aquest dintre de la imatge. Per tant, són invariants a translacions, rotacions i canvis d’escala i poden ser molt útils amb vista a extreure característiques d’imatges amb nivells de grisos. No serviran, però, per caracteritzar imatges binàries.

Tenen el problema que no tenen en compte la posició de cada píxel dintre de l’objecte. Per tant, dos objectes amb histogrames idèntics generaran idèntics valors dels moments calculats a partir de l’histograma.

Els dos objectes següents, com que tenen el mateix histograma, presenten els mateixos moments calculats a partir de l’histograma.

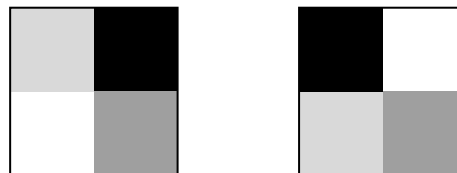


Figura 2.21.

2.14. Descriptors de Fourier

Els descriptors de Fourier constitueixen un conjunt de característiques que poden ser invariants a translacions, canvis d’escala i rotacions de l’objecte. Aquestes característiques fan servir les propietats de les transformades de Fourier i d’aquí en resulta el nom. Han estat àmpliament emprats; consulteu [Gra72], [Zah72], [Per77], [Cri82], [Che84], [Lin87], [Mah94], [Wan94], [Rot96] i [Tri96]. En la literatura sobre aquest tema hi ha una gran varietat de descriptors, calculats a partir de definicions diferents. El problema que presenten molts d’aquests és que no són invariants a canvis d’escala i a girs. Per aquest fet,

s'utilitzaran els descriptors de Fourier definits per Granlund en [Gra72], els quals són invariants a totes les transformacions lineals no deformatives.

Els descriptors definits per Granlund només tenen en consideració el contorn de l'objecte; per tant, primer s'ha de calcular el contorn i a partir d'aquí tots els càlculs es realitzen sobre els píxels del contorn.

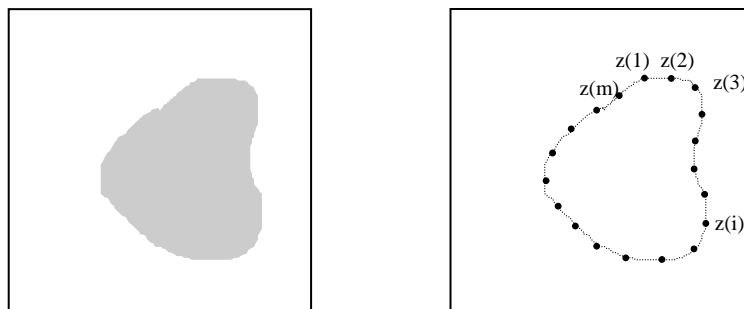


Figura 2.22.

Les coordenades dels píxels del contorn (vegeu la figura 2.22.) se suposen nombres complexos. A més, si el contorn és tancat, els píxels formen la successió periòdica següent:

$$z(i) = x(i) + j.y(i), \quad \text{amb } \Delta i = 1 \quad (2.60)$$

El període de la successió és T (nombre de píxels del contorn). La figura 2.22. mostra que es compleix:

$$T = m,$$

i també:

$$z(i + n.T) = z(i), \quad \text{on } n \in N. \quad (2.61)$$

Com que tenim una funció unidimensional $z(i)$ complexa i periòdica, es pot determinar la seva transformada de Fourier. Els coeficients de la transformada de Fourier vénen donats per l'expressió:

$$a_n = \frac{1}{T} \int_0^T z(i) \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot i}{T}\right) di, \quad \text{amb } n \in N. \quad (2.62)$$

El centre de gravetat del contorn ve donat per a_0 , on:

$$a_0 = \frac{1}{T} \int_0^T z(i) di = z_{cdm}. \quad (2.63)$$

També es compleix:

$$z(i) = \sum_{-\infty}^{\infty} a_n \cdot \exp\left(\frac{j \cdot 2\pi \cdot n \cdot i}{T}\right), \text{ amb } n \in \mathbb{N}. \quad (2.64)$$

Els coeficients a_n són invariants a translacions de l'objecte (per les propietats de la transformada de Fourier), però no són invariants a canvis d'escala i a girs.

A continuació es demostren les relacions entre els coeficients a_n de l'objecte original, el contorn del qual és $z(i)$, i els coeficients a'_n de l'objecte transformat, el contorn del qual és $z'(i)$.

2.14.1. Efecte d'una translació sobre els coeficients a_n

Quan l'objecte es mou de posició, tots els punts del seu contorn es mouen el mateix vector complex r , resultant que:

$$z'(i) = \sum_{-\infty}^{\infty} a_n \cdot \exp\left(\frac{j \cdot 2\pi \cdot n \cdot i}{T}\right) + r = z(i) + r. \quad (2.65)$$

Es pot fer:

$$a'_n = a_n, \text{ per } n \neq 0, \quad \text{i} \quad a'_0 = a_0 + r. \quad (2.66)$$

Resulta que tots els coeficients es mantenen inalterats excepte el corresponent a $n = 0$.

2.14.2. Efecte d'un canvi d'escala sobre els coeficients a_n

Quan l'objecte es veu afectat per un canvi d'escala, les coordenades de tots els seus punts es modifiquen en un factor real R (magnitud del canvi d'escala), resultant que:

$$z'(i) = z(i) \cdot R = \sum_{-\infty}^{\infty} a_n \cdot R \cdot \exp\left(\frac{j \cdot 2 \cdot \pi \cdot n \cdot i}{T}\right). \quad (2.67)$$

Resulta que:

$$a'_n = a_n \cdot R. \quad (2.68)$$

2.14.3. Efecte d'un gir sobre els coeficients a_n

Quan l'objecte gira, les coordenades de tots els seus punts es veuen afectades en un factor $\exp(j\phi)$, on ϕ és l'angle girat, resultant que:

$$z'(i) = z(i).exp(j\phi) = \sum_{-\infty}^{\infty} a_n.exp(j\phi).exp\left(\frac{j.2.\pi.n.i}{T}\right). \quad (2.69)$$

Per tant:

$$a'_n = a_n.exp(j\phi). \quad (2.70)$$

2.14.4. Descriptors de Fourier proposats per Granlund

Granlund [Gra72] va definir els següents descriptors de Fourier, invariants a translacions, canvis d'escala i girs de l'objecte:

$$b_n = \frac{a_{1+n}.a_{1-n}}{a_1^2}, \quad \text{amb } n \neq 1.$$

$$D_{m,n} = \frac{a_{1+m}^{n/k}.a_{1-n}^{m/k}}{a_1^{(m+n)/k}}, \quad \text{amb } n \neq 1 \text{ i } k = \text{m.c.d.}(m,n). \quad (2.71)$$

Quan $n = m$, resulta que $b_n = D_{n,n}$. Això indica que els descriptors b_n són un cas particular dels descriptors $D_{m,n}$.

Si m i n tenen un màxim comú divisor k , és convenient treure aquest factor k de l'expressió $D_{m,n}$ (tindríem l'expressió elevada innecessàriament a una potència k) per tal de no perdre informació per raó dels càlculs.

El coeficient de normalització que es posa en el denominador és a_1 , perquè normalment és el coeficient amb valor més gran, i s'ha comprovat que funciona bé en aquesta transformació d'invariància.

Davant d'una transformació geomètrica al més general possible, els coeficients a_n es veuen transformats de la manera següent:

$$a'_n = R.exp(j\phi).a_n, \quad \text{per } n \neq 0. \quad (2.72)$$

Anem a demostrar la invariància a translacions, canvis d'escala i rotacions dels descriptors de Fourier b_n i $D_{m,n}$ definits per Granlund:

Invariància dels coeficients b_n

Anem a trobar la relació entre els coeficients calculats a partir de l'objecte original (b_n) i els coeficients resultants d'efectuar a l'objecte anterior una transformació geomètrica al més general possible (b'_n):

$$b'_n = \frac{a'_{1+n} \cdot a'_{1-n}}{a_1'^2} = \frac{a_{1+n} \cdot R \cdot \exp(j\phi) \cdot a_{1-n} \cdot R \cdot \exp(j\phi)}{a_1^2 \cdot (R \cdot \exp(j\phi))^2} = \frac{a_{1+n} \cdot a_{1-n}}{a_1^2} = b_n. \quad (2.73)$$

Acabem de veure que el valor teòric del coeficient b_n no es veu afectat en sotmetre l'objecte a una transformació geomètrica no deformativa.

Invariància dels coeficients $D_{m,n}$

Anem a procedir de la mateixa manera que en el cas anterior:

$$D'_{m,n} = \frac{a_{1+m}^{n/k} \cdot a_{1-n}^{m/k}}{a_1^{(m+n)/k}} = \frac{a_{1+m}^{n/k} \cdot R^{n/k} \cdot \exp(j\phi \cdot n/k) \cdot a_{1-n}^{m/k} \cdot R^{m/k} \cdot \exp(j\phi \cdot m/k)}{a_1^{(m+n)/k} \cdot R^{(m+n)/k} \cdot \exp(j\phi(m+n)/k)} = \frac{a_{1+m}^{n/k} \cdot a_{1-n}^{m/k}}{a_1^{(m+n)/k}}.$$

Resulta que:

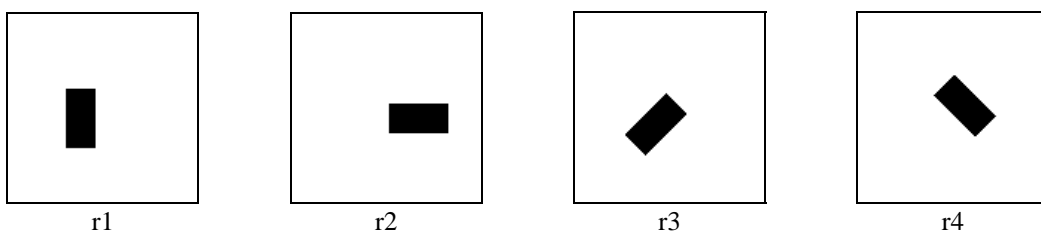
$$D'_{m,n} = D_{m,n}. \quad (2.74)$$

Acabem de veure que el valor teòric del coeficient $D_{m,n}$ no es veu afectat en sotmetre l'objecte a una transformació geomètrica no deformativa.

2.15. Exemple del càlcul de les característiques en imatges sintètiques

En aquest apartat es calculen la majoria de les característiques desenvolupades en aquest capítol i se n'analitza el comportament. Aquest apartat és important perquè, en el cas que alguna de les característiques explicades anteriorment no tingui un bon comportament, pugui ser rebutjada. Si hi ha alguna característica que o bé necessita un temps de computació massa elevat, o bé no discrimina els objectes de classes diferents, o bé té una variabilitat massa elevada (i aporta confusió al problema), serà descartada.

Es treballarà amb dues classes d'imatges, cada una de les quals conté vuit elements de dues mides diferents. Per tant, cada classe conté dues subclasses de quatre objectes de les mateixes dimensions, però amb orientacions i posicions diferents. Les imatges tenen 128 x 128 píxels i han estat generades artificialment (no són imatges digitalitzades).



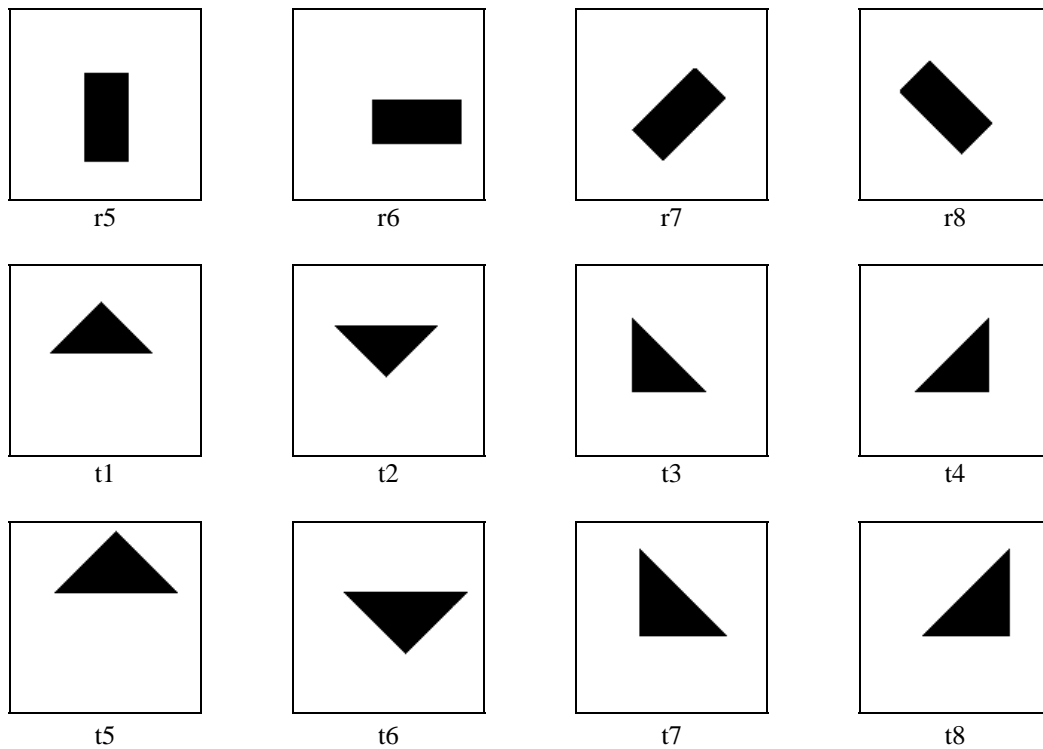


Figura 2.24

La figura anterior mostra tot el conjunt d'objectes amb els quals es calcularan les característiques invariants.

2.15.1. Anàlisi dels resultats

Per no allargar innecessàriament aquest apartat, no es donen els valors proporcionats per totes les característiques avaluades. A excepció dels descriptors de Fourier, totes les altres característiques han superat aquest test.

S'ha cregut convenient, però, mostrar el resultat proporcionat per alguns dels descriptors de Fourier aportats per Granlund. En la taula següent se'n pot comprovar el comportament deficient:

	$ b_2 $	$ b_3 $	$ b_4 $	$ b_5 $	$ b_6 $	$ b_7 $	$ b_8 $	$ b_9 $
r1	2,33e-2	0	7,13e-4	0	8,28e-4	0	6,81e-5	0
r2	2,33e-2	0	7,13e-4	0	8,28e-4	0	6,81e-5	0
r3	2,28e-2	0	7,59e-4	0	8,22e-4	0	6,21e-5	0
r4	2,28e-2	0	7,59e-4	0	8,22e-4	0	6,21e-5	0

r5	2,28e-2	0	7,66e-4	0	8,21e-4	0	6,10e-5	0
r6	2,28e-2	0	7,66e-4	0	8,21e-4	0	6,10e-5	0
r7	2,31e-2	0	7,14e-4	0	7,99e-4	0	6,27e-5	0
r8	2,31e-2	0	7,14e-4	0	7,99e-4	0	6,27e-5	0
t1	1,24e-3	0	4,47e-3	0	9,19e-5	0	2,58e-4	0
t2	1,24e-3	0	4,47e-3	0	9,19e-5	0	2,58e-4	0
t3	0	1,57e-2	0	0	8,25e-4	0	0	1,60e-4
t4	0	1,57e-2	0	0	8,25e-4	0	0	1,60e-4
t5	1,24e-3	0	4,46e-3	0	9,15e-5	0	2,56e-4	0
t6	1,24e-3	0	4,46e-3	0	9,15e-5	0	2,56e-4	0
t7	0	1,57e-2	0	0	8,23e-4	0	0	1,59e-4
t8	0	1,57e-2	0	0	8,23e-4	0	0	1,59e-4

Taula 2.2.

Els descriptors de Fourier presenten dos problemes. El primer és el seu elevat cost computacional: són molt lents de càlcul. El segon problema és que els resultats que proporcionen no són massa discriminants, ja que presenten una variabilitat elevada. Una possible explicació de l'elevada variabilitat dels resultats proporcionats pels descriptors de Fourier és que estan calculats respecte a la transformada de Fourier del contorn de l'objecte i que les freqüències del contorn varien força en fer una transformació geomètrica de l'objecte, pel caràcter digital de les imatges. Per tant, no s'utilitzaran en aquest treball.

