

L'ÚS DE LA MATRIU DE MARKOV EN EL RECONeixEMENT DE FORMES

per JORDI FLOS i BASSOLS

Departament d'Ecologia. Facultat de Biologia.
Universitat de Barcelona.

Sota el títol de «reconeixement de formes» (o «*patterns*») hom admet gran quantitat de coses diverses. Per un costat tenim els que tracten del reconeixement de formes sensorials, principalment formes visuals i auditives. Els psicòlegs i psicofisiòlegs analitzen quines són les formes discernibles i el perquè unes ho són i d'altres no, així com el procés mental involucrat. Per un altre costat hi ha els que tracten, des d'un punt de vista matemàtic, les distribucions naturals d'espècies en l'espai. En aquest camp es poden fer models matemàtics que donguin lloc a funcions de distribució estadística que s'amotllin, en molts casos força bé a la matriu original. La majoria de vegades, però, no expliquen res, puix que fins i tot models contradictoris poden portar a distribucions semblants o estadísticament idèntiques.

Al costat de tot això hi ha el problema del discerniment de formes visuals i auditives fora de tota psicologia. El problema que s'intenta resoldre és el de trobar la manera de caracteritzar al màxim una forma per tal que una màquina la pugui reconèixer. La idea de parlar als ordinadors i de fer que parlin és ja vella.

En el nostre cas concret, el nostre material de treball són fotografies en color de bentos sobre fons rocós, i el reconeixement de formes té dos caires: un és el de reconèixer les fotografies «pictòricament» semblants, l'altre és el de reconèixer les comunitats i ecosistemes semblants.

En el cas de la fotografia aèria o de satèl·lit el problema acostuma a reduir-se a la caracterització de les diferents taques a fi de treure'n un mapa temàtic (quan no es tracta de reconèixer carreteres, trens, soldats,...). Un mètode directe consisteix en fer una *cluster* anàlisi, de la qual surten una sèrie de grups (tonalitats de grisos per exemple), seguida d'una

anàlisi discriminat, adjudicant cada punt a un dels grups. En el cas de fer una classificació supervisada, es pren com a representants ideals dels grups que es volen fer una sèrie de zones limitades i de textura aparentment constant dins de cada una d'elles. Paral·lelament, es prenen mostres sobre el terreny i es qualifiquen aquestes zones (blat, ordi, patates, pins,...). S'obté llavors un mapa temàtic.

En el nostre cas podríem dir que la fotografia en color és un «mapa temàtic» instantani que s'ha d'interpretar. La interpretació es pot portar a dos nivells: en el primer es tracta de reconèixer grups d'espècies o d'individus diferents. En l'altre nivell es tracta de reconèixer, per la forma, la comunitat o fàcies.

El primer nivell es resol, després de digitalitzada la imatge amb un densímetre adequat, fent una classificació (supervisada o no) dels diferents tons que es troben. Aquest nivell queda pobre si després d'això el resultat és simplement un mapa temàtic que novament s'ha d'interpretar en el segon nivell. Per l'altre costat és possible que ens puguem saltar, en part, el primer nivell, i passar directament al segon. El sistema consisteix a digitalitzar la fotografia, i sobre la matriu de dades, buscar una sèrie de variables discriminants mesurades sobre la imatge que ens permeti, en un primer pas, de classificar les diferents fotografies en diferents grups. La hipòtesi de treball és, doncs, una que ve a dir que «*sobre fotografies en color de bentos rocós, la distribució, forma, grandària, orientació i diversitat de les taques, així com la presència o absència d'algunes característiques, permet discriminar les diferents comunitats bentòniques*». Darrera d'aquesta hipòtesi hi ha una altra d'implícita, i és que «*el funcionament des del punt de vista ecològic d'un sistema, està íntimament relacionat amb l'estructura espacial d'ell mateix*». Per entendre'ns: està relacionat amb la grandària, freqüència i distribució espacial dels seus individus o subsistemes.

A primer cop d'ull, un biòleg de bentos acostumat a veure el fons submarí és capaç de classificar força bé les fotografies segons el tipus d'ecosistema al qual pertanyen. O sigui que el cervell humà entrenat «veu» variables discriminants evidents. L'interès que té llavors l'intent de fer-ho mecànicament resideix en el fet d'anar provant variables que ens permetin reconèixer formes o *patterns*, no només pictòrics, sinó d'altres que facin referència o estiguin relacionats directament amb les propietats d'ordre ecològic del sistema representant en la fotografia.

Aquí es presenta, encara que superficialment, les possibilitats que ofereix l'ús de la matriu de transició d'un procés de Markov per veure l'orientació i pes específic de les diferents taques d'una fotografia. El sistema consisteix a fer una sèrie d'escombrats sobre la matriu de dades (es treballa amb la imatge digitalitzada i amagatzemada a l'ordenador) en

angles diferents. Dins de cada escombrat es calcula, independentment per a cada transecte, la matriu de probabilitats de transició. Els colors possibles (quantitat establerta per l'analista) adjudicats a cada un dels elements de imatge, són els possibles «estats». Cada transecte és un «procés». Sobre la matriu de probabilitats es calcula una variable aleatòria que és funció de les probabilitats de transició, i que no és més que la probabilitat d'una sèrie de trajectòries. Simbòlicament:

a és l'angle d'escombrat, que pren valors de 0 a 180 graus.

Els possibles estats són:

$$T_{ij} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, n$$

Per cada transecte, ${}_k U$, es té una matriu de probabilitats:

$$({}_k U_{ij}) \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, n \quad a \text{ on}$$

$$\sum_{j=1}^n {}_k U_{ij} = 1 \quad \text{sent això la suma de les probabilitats de transició d'un estat } i \text{ a qualsevol altre } j.$$

Per cada angle a es té k_a transectes, i per cada probabilitat $P(T_{ij}) = U_{ij}$ es té k_a valors.

La forma de les variables aleatòries és del tipus:

$$X = P(T' \text{ ó } T'' \text{ ó } T''')$$

Això s'ha provat sobre matrius de 60×40 elements, fetes amb uns i dosos (dos estats). Les imatges analitzades han estat creades amb programes que utilitzen números a l'atzar i unes certes estratègies per fer que surtin les taques amb orientacions preestablertes (veure les figures).

Per la construcció de les variables aleatòries s'ha de tenir en compte què es vol posar en evidència: si es tracta, per exemple, de veure l'orientació de taques d'«uns» sobre «dosos» ens interessarà sobretot la distribució de la variable aleatòria que faci referència a la transició T_{11} . La probabilitat U_{11} serà la de passar a un 1 en el següent pas, estant ja en un 1. Per a figures compactes pot servir aquesta variable, però si les taques són una mica difuses s'haurà de fer servir variables d'un ordre més elevat, o sigui que representin probabilitats de trajectòries més llargues com per exemple:

T_{111} , T_{121} , T_{211} , T_{112} en les quals, de tres posicions seguides n'hi ha dues que són 1. La probabilitat de la unió de les quatre trajectòries seria una variable com les utilitzades en els exemples dels gràfics:

$$X_1 = U_{22}^2 + U_{21} U_{12} \circ T_{122} \circ T_{212} + U_{21} U_{12} = P (T_{222})$$

$$X_2 = U_{22}^3 + U_{12} | U_{22}^2 + 2 U_{21} U_{12} U_{12} = P (T_{222} \circ T_{1222} \circ T_{2122} \circ T_{2212})$$

Les imatges utilitzades en aquest treball estan molt lluny de les imatges (fotografies de bentos) que l'han provocat. Aquestes diferències es poden resumir en un parell de punts:

1. — Les imatges tractades aquí amb aquest mètode consten de dos estats (indicats per uns i dosos), mentre que les fotografies de bentos en presenten tants com permeten el mètode de digitilització (precisió) i l'analista.
2. — Les formes de les taques de les fotografies de bentos són rarament tant senzilles com les tractades aquí.

La matriu de les probabilitats de transició permet fàcilment l'extensió del cas de dos estats a més de dos. Es poden estudiar taques formades per un o més estats, sobre un fons de diversos estats indiferenciats, ja que la matriu permet calcular sense dificultat les probabilitats de les trajectòries interessants que constituïran la variable aleatòria.

La variable escollida es distribuirà de manera diferent segons l'angle d'escombrat. Si agafem una figura semblant a la de la fig. 1 o 1 bis, veu-

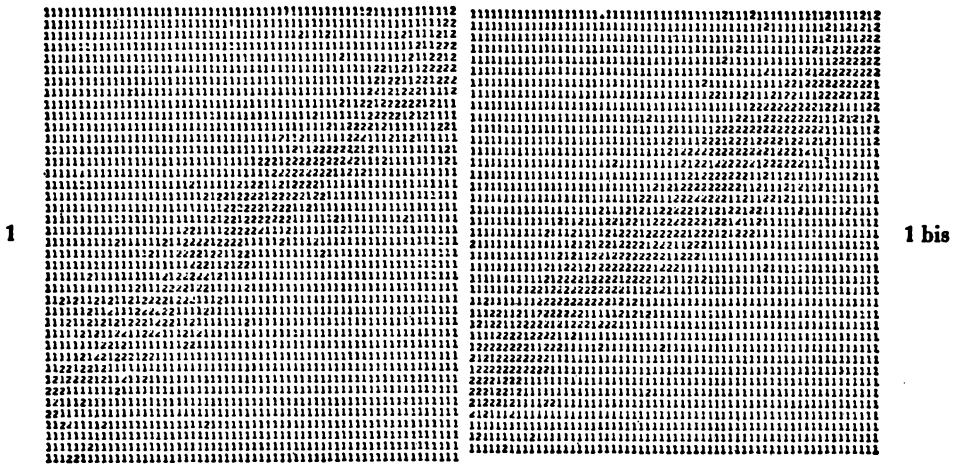


FIG. 1. — Els dosos estan col·locats amb probabilitat decreixent perpendicularment a una recta obliqua, segons una exponencial negativa.

FIG. 1 bis. — El mateix que la fig 1, però les probabilitats són més altes. El resultat és una taca més densa.

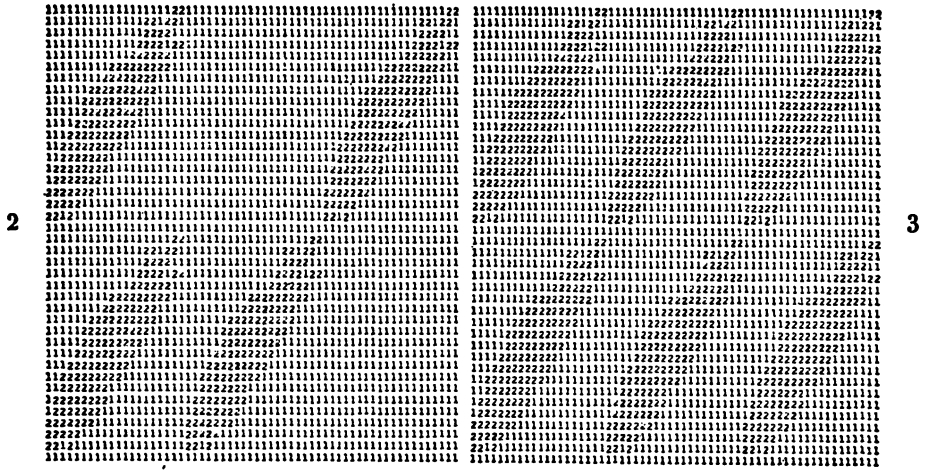


FIG. 2. — Taca repetida quatre vegades. En aquest cas la taca s'inicialitza en un punt central d'una matriu de 20 x 20 i els altres dosos es posen per iteració d'una subrutina que fa que la probabilitat que s'hi afegeix un dos en un punt és o si no n'hi ha cap al costat, que sigui màxima en la direcció obliqua i menor en la perpendicular a aquesta.

FIG. 3. — La mateixa taca de la figura 2 repetida 6 vegades.

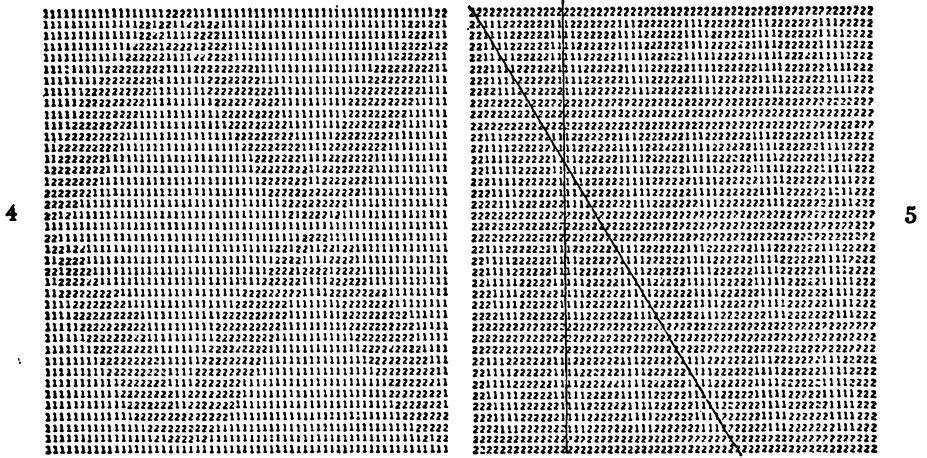


FIG. 4. — La mateixa taca: 3 del dret i 3 girades 90 graus. Apareixen dues direccions creuades.

FIG. 5. — 24 el·lipses col·locades regularment, girades 60 graus cadascuna (eix gran=8, eix petit=4).

rem que pels transectes que vagin en la direcció de l'eix gran de la taca la variable prendrà o bé valors molt alts, o bé valors molt baixos. En les altres direccions prendrà valors semblants. L'histograma de la figura 6 representa dues distribucions imaginàries: una bimodal que correspondria a un angle d'escombrat en direcció de l'eix principal i un altre unimodal per una direcció perpendicular a l'anterior.

Explicació de les figures

Totes les figures són matrius de 60×40 elements formades per uns i dosos. La sortida per impressora de l'ordenador fa que la distància entre dos elements sigui més petita horitzontalment que verticalment, i que quedin afectats els angles.

Explicació dels gràfics

Sent la imatge rectangular i discreta (60×40 elements), el número de transectes i la seva longitud (considerats com processos, en número d'estats) dependrà de l'angle d'escombrat. Per un angle de 0 graus tindrem 40 transectes de longitud 60, i per un angle de 90, 60 transectes de longitud 40. Pels angles intermedis la cosa es complica i la longitud dels transectes varia molt. En el cas de $\alpha = 45$ graus, tindrem 99 transectes, dels quals 2 seran de longitud 1, que no es poden considerar pròpiament processos. Els transectes més llargs seran de $60/\cos 45$ estats i d'aquests n'hi haurà 20 només.

Per cada transecte es calcula la matriu de probabilitats de transició i de retruc el valor de la variable associat. Per un procés de longitud 2, només hi haurà un transició i la seva probabilitat serà 1. Aquesta diversitat de longituds dels transectes, fa que la distribució de la variable aleatòria es vegi molt falsejada pels valors deduïts de transectes curts. Per tant, a l'hora de calcular un paràmetre que caracterizi la distribució, hem prescindit d'aquells valors deduïts de transectes de longitud més curta de 21, i hem perdut una informació indesitjable sobre les cantonades del rectangle.

Els càlculs s'han fet tots a l'ordenador IBM-1130 de l'Institut de Investigacions Pesqueres de Barcelona.

El paràmetre que s'ha fet servir per caracteritzar les distribucions és

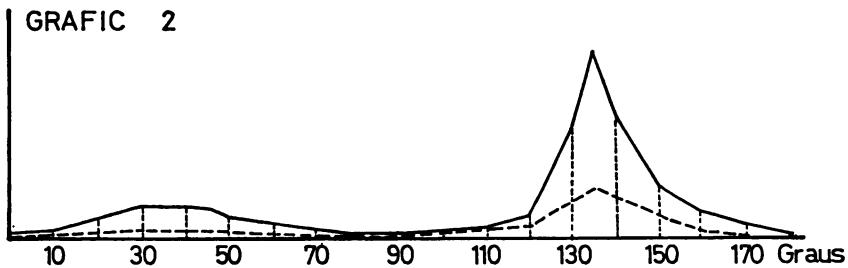
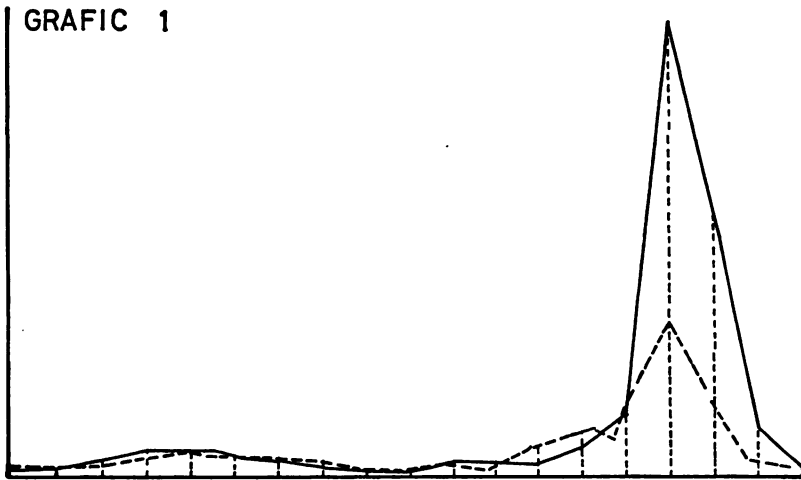
$$V = (\text{mitjana dels 5 valors màxims-mitjana total})^2$$

Aquest paràmetre ben senzill s'ha mostrat útil a l'hora de posar en evidència les direccions de les taques, presentant pics quan l'angle d'escom-

brat coincideix amb la direcció de l'eix gran de la taca. El que ens donarà més informació serà, però, el nombre de pics i la seva forma (amplada, forma arrodonida o punxeguda,...).

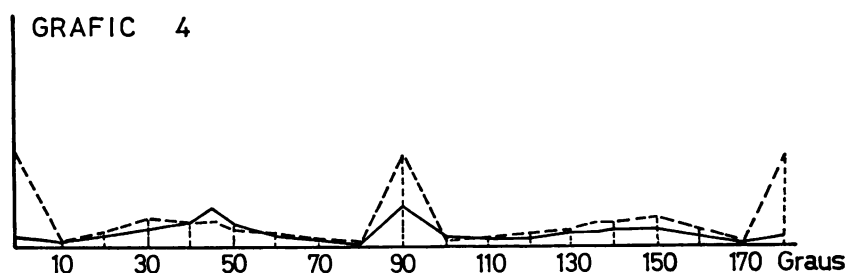
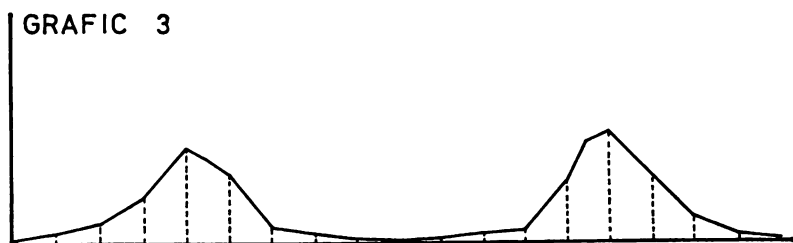
Conclusions

Amb variables successives de diferents ordres sembla ser que es pot posar en evidència la densitat de les taques.



GRAFIC 1. — La variable és X_3 (tercer ordre) per la imatge de la fig. 1 (línia discontinua), i amb X_3 també per la imatge de la fig. 1 bis (línia contínua).

GRAFIC 2. — Utilitzant X_1 per les imatges de la fig. 2 (discontinua), i fig. 3 (contínua).



GRAFIC 3. — Utilitzant X_1 per la imatge de la fig. 4. Queden en evidència les dues direccions, amb els pics desplaçats 5 graus respecte als 45 i 135 originals que eren d'esperar.

GRAFIC 4. — Utilitzant X_1 per la imatge de la figura 5 (contínua) i per una figura similar, però formada per circumferències de radi 4 en lloc d'el·lipses (discontínua). En el cas de les el·lipses hi ha dos pics: a 90 i 45 graus (angles marcats sobre la figura), i no es posa en evidència l'orientació de 60 graus de cada el·lipse. En el cas de les circumferències els pics es presenten a 0 i 90 graus.

Sembla ser que l'estudi de la forma dels pics (que no hem fet aquí) pot donar informació sobre la regularitat dels contorns, la importància de la superfície de la taca respecte a la total, i l'elongació. Possiblement escollint algun altre paràmetre per a caracteritzar les distribucions (potser del tipus de quartils), es podrien fer càlculs més precisos.

Els resultats amb les figures de les el·lipses i les circumferències mostra clarament que el nostre sistema, tal com l'hem fet servir, posa en evidència una estructura d'un cert nivell. Les estructures locals (orientació de les el·lipses) podrien posar-se en evidència reduint la superfície escombrada. En aquest últim cas, es podria utilitzar una estimació del la mitjana del diàmetre petit de la taca (o taques) per tal de fixar la superfície per escombrar. Tenim un programa que fa una estimació d'aquest tipus.

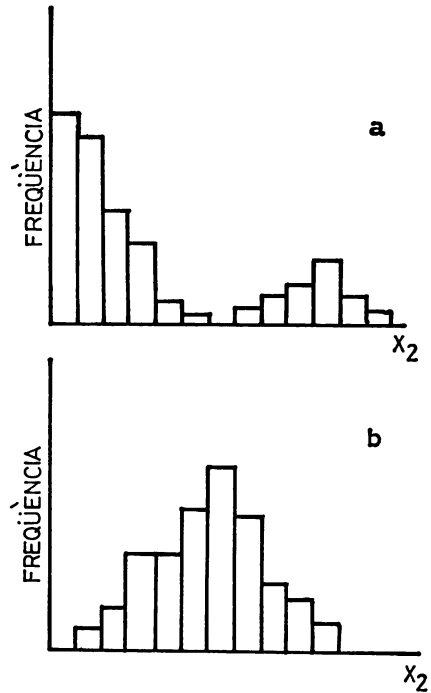


FIG. 6. — Exemples de distribució esperada de la variable X_2 en dos escombrats d'una figura formada per una taca allargada com la de les figures 1 i 1 bis. a) Direcció d'escombrat paral·lel a l'eix principal de la taca. b) Direcció d'escombrat perpendicular a l'eix principal de la taca.

A la pràctica el que s'haurà de fer és fixar las variables aleatòries, els paràmetres per calcular la forma de la distribució d'aquestes i les transformacions (fruit de l'estudi dels pics) que donin els valors de les variables discriminants. L'estimació d'una «mitjana del diàmetre petit de les taques» donaria lloc a una altra variable discriminant.