

Modelare și modele matematice în recunoașterea obiectelor și clasificarea automată a imaginilor

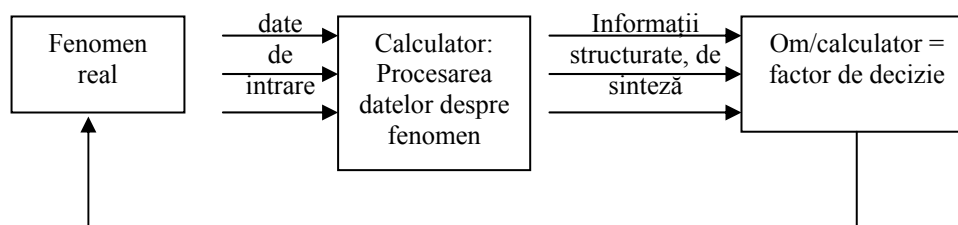
Ioan Ispas,

Catedra de Matematică-Informatică,
Universitatea Petru Maior, Târgu Mureș,
e-mail: john@upm.ro

Abstract: This paper summarize and classify the most important approaches, methods and techniques used in mathematical modeling and algorithms design for solving the *Image Recognition and Classification Problem*. The classification is done with respect to the fundamental mathematical modeling process of this difficult but challenging problem.

Keywords: Pattern recognition, Image classification, Image retrieval approaches, Mathematical modeling.

Se poate spune că schema următoare conține modul ideal în care oamenii își propun să folosească calculatorul. Când omul este pus în situația obositoare de a lua în mod repetat aceleași decizii "de rutină", ideal este ca el să fie înlocuit de un calculator, cu softul său adecvat.



Bucă inversă (de control a fenomenului real)

În domeniul prelucrării digitale a imaginilor (DIP), în care există multe operații "de uzură" ce nu implică un grad înalt de răspundere, această înlocuire este unul dintre cele mai importante obiective și se studiază mai ales atunci când se vorbește despre *Problema recunoașterii obiectelor și clasificării automate a imaginilor (pattern recognition and classification)*. Prin această sintagmă

generală se înțelege capacitatea calculatorului, prin softul său specializat, de a determina conținutul unei imagini imitând într-o cât mai bună măsură capacitatea umană de "citire" a imaginilor în vederea luării unor decizii ulterioare.

Prin *clasificarea automată a imaginilor* se înțelege de fapt o categorie de tehnici, metode și algoritmi de recunoaștere a obiectelor și formelor în imagini ce permit clasificarea imaginilor pe baza obiectelor recunoscute.

Câteva domenii care necesită aplicații DIP de clasificare automată a imaginilor pe baza recunoașterii obiectelor componente sunt:

Astronomie - pentru clasificarea automată a imaginilor preluate de la diferitele telescoape în benzi de frecvențe extinse;

Fizica pământului - pentru clasificarea automată (de multe ori în timp real) a imaginilor preluate de la sateliții geostaționari în vederea prognozei meteo, a stării atmosferei, a controlului creșterii vegetației (culturilor) și a comportamentului vulcanilor, etc.;

Fizica atomică sau cuantică - pentru clasificarea automată a imaginilor preluate în experimente cu particule subatomice;

Microelectronică - pentru clasificarea automată și detectarea calității circuitelor VLSI pe bază de imagini microscopice;

Automatică - pentru clasificarea automată și controlul automat al calității diferitelor produse, în diverse domenii economice, ce trec pe bandă rulantă prin fața unui controlor automat al calității (aici este inclus și domeniul alimentară);

Televiziune - pentru asigurarea controlului și securității pe bază de camere video de urmărire continuă prin detectarea și recunoașterea în timp real a situațiilor (eventual a persoanelor) frauduloase (aici este inclus controlul circulației prin camere video și radar, controlul bagajelor pe aeroporturi, urmărirea persoanelor în incinta băncilor, etc.);

Armată - pentru recunoașterea și detectarea țintelor și pentru dirijarea în mod direct (orientare) a rachetelor sau avioanelor de luptă;

Inspekția non-chirurgicală - mai ales pentru detectarea și marcarea automată (un fel de diagnoză asistată) a zonelor sau organelor suspecte, din imaginile preluate *in vivo* prin microscopie electronică, tomografie (raze gamma), fotografierea cu raze X și prin rezonanță magnetică, radiologie;

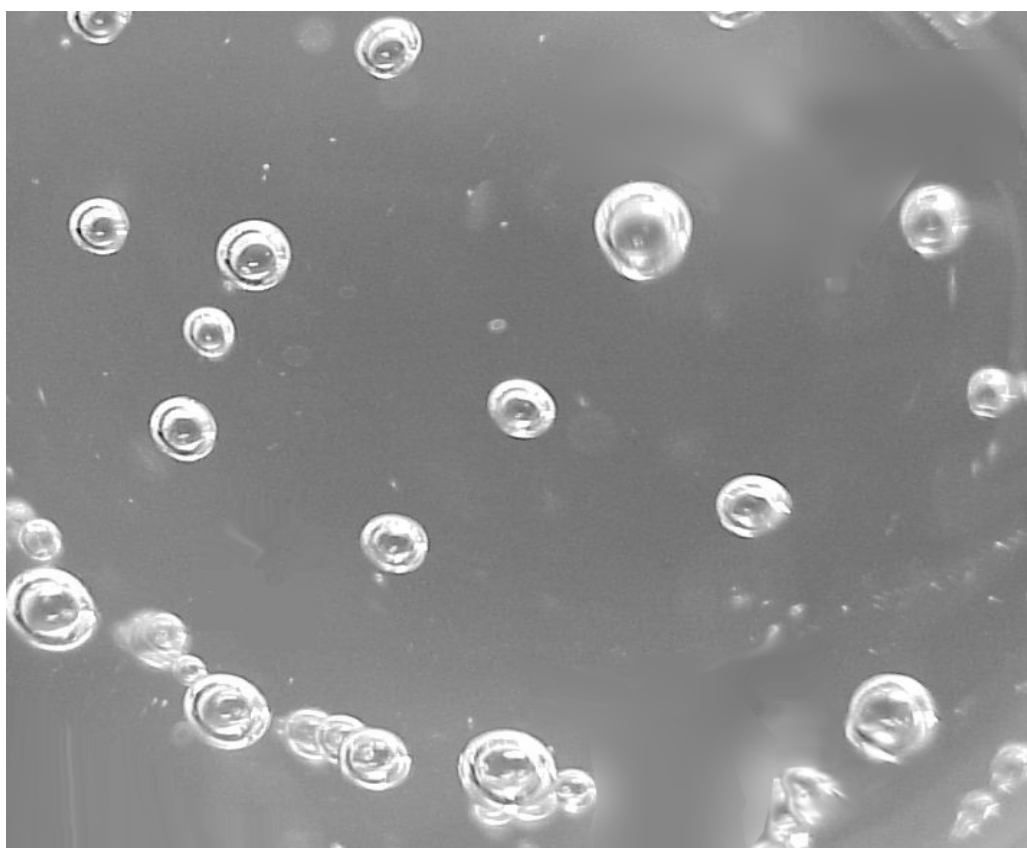
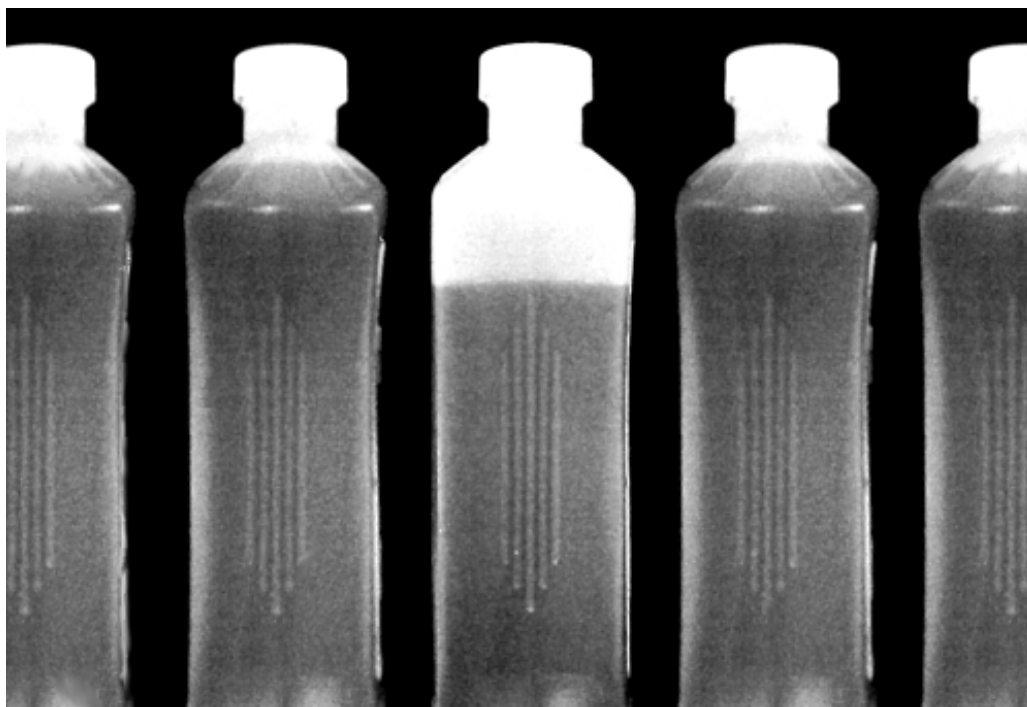
Analiza compoziției chimice - prin clasificarea și prelucrarea automată a informației achiziționate pe criterii cromatice, spectrale, etc.;

Microscopie - pentru recunoașterea și clasificarea automată a structurilor moleculare și atomice ale substanțelor studiate;

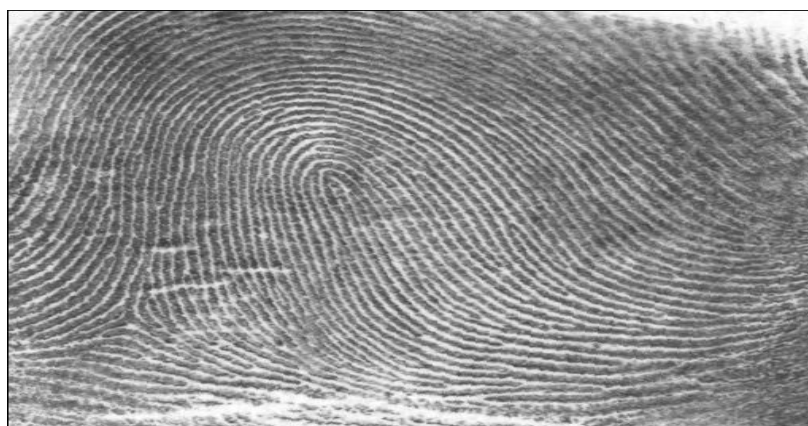
Data mining – clasificare și regăsirea informațiilor de diverse tipuri (multimedia) în volume mari de date structurate, mai mult sau mai puțin, ca baze de date, depozite de date sau bănci de date;

Document image analysis – presupune clasificare, căutarea, regăsirea cu sau fără conversia totală în text a unor documente păstrate sub formă de imagini (copii fax, scanner, etc.);

Multimedia database querying and content retrieval – se referă mai ales la regăsirea unor secvențe video după criterii de căutare specifice inclusiv, să zicem, după chipul unui personaj;



Iată două exemple ce ilustrează folosirea metodelor automate de recunoaștere, detecție și control a calității produselor farmaceutice: în primul caz, camera video detectează trecerea prin fața ei (pe bandă rulantă) a unor recipiente necorespunzătoare, iar în al doilea caz, prin inspecție microscopică se detectează prezența unui număr prea mare de bule de aer în compoziția unei substanțe (prin recunoaștere și numărare automată).



Iată alte trei exemple de utilizare practică a metodelor de recunoaștere și clasificare a imaginilor: sus, o imagine preluată de o cameră radar a poliției care este prelucrată și se extrage în mod automat numărul de identificare a mașinii; la mijloc, o imagine folosită într-o aplicație de recunoaștere a poliției cu ajutorul unei baze de amprente digitale; jos, o imagine preluată cu raze X de o cameră de control a calității produselor alimentare și oferită unei aplicații care detectează prezența ascunsă a oaselor în pachetele de "carne fără oase"

Recunoașterea formelor. Noțiuni fundamentale.

Literatura despre *prelucrarea digitală a imaginilor (DIP)* în limba engleză folosește diferite exprimări, ce se acoperă parțial reciproc, pentru a desemna acest domeniu de interes. Din perspective diferite, exprimările întâlnite curent sunt: *object recognition*, *pattern recognition*, *pattern / template matching*, *image recognition*, *pattern / image classification* [GOW], [JDM], [BKPP]. În limba română acestor sintagme le corespund următoarele traduceri sau echivalări: *recunoașterea formelor*, *recunoașterea obiectelor*, *recunoașterea și clasificare imaginilor*, *analiză de imagini* [VLA], unde noțiunea de *pattern* (*șablon*, *model*, *formă primară*) este tradusă mai ales prin cuvântul cu înțeles general - *formă*.

Din punct de vedere al creșterii complexității metodelor și operațiilor implicate există următoarea incluziune logică evidentă: *forme* \subset *obiecte* \subset *imagini*. Dacă vrem să facem distincție între termenii de mai sus, ordinea firească de abordare teoretică a lor ar trebui să fie (pe nivele de complexitate):

1. *recunoașterea formelor*,
2. *recunoașterea obiectelor*,
3. *recunoașterea / clasificarea imaginilor*.

În modelarea matematică a problemei recunoașterii formelor există și abordări teoretice care eludează această ordine logică. De exemplu, abordarea cu *metode matematice statistice a recunoașterii*, face abstracție de incluziunea structurală a formelor în obiecte (cu forme complexe) și a obiectelor în imagini (conținând mai multe obiecte). Într-o astfel de modelare întreg procesul se reduce la un unic nivel al recunoașterii, ce conține la modul sintetic cele trei nivele de mai sus. Un exemplu semnificativ este *metoda de recunoașterea statistică a imaginilor /obiectelor pe baza conținutului lor cromatic*.

Înainte de a trece la prezentarea diferitelor modele teoretice și practice de soluționare a problemei recunoașterii vom prezenta mai întâi o scurtă analiză a acestei probleme din perspectiva sa fundamentală: *psihofiziologia umană*.

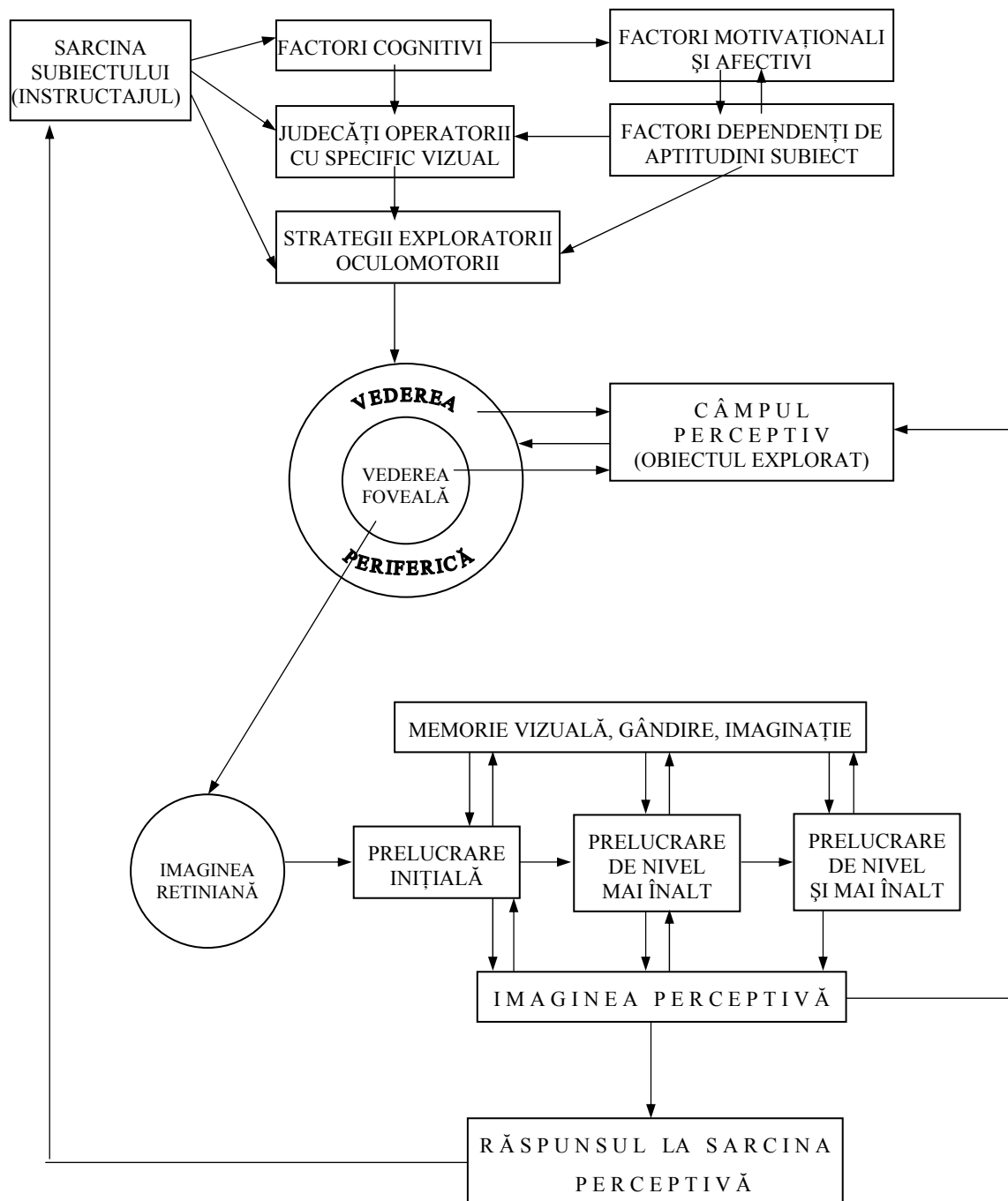
Din această perspectivă, problema recunoașterii formelor și imaginilor nu este o problemă de sine stătătoare ci este subordonată unei alte probleme mai generale: *problema luării unor decizii pe baza procesului psiho-mental de*

percepție vizuală. Aici trebuie deasemenea incluse fenomene umane complexe ca *orientarea spațială, explorarea vizuală, etc.*, în fiecare din aceste situații pornindu-se de la percepții vizuale primare (imagini). Această problemă generală și foarte complexă conține într-o formă particulară cel mai important mecanism ce stă la baza funcționării tuturor sistemelor vii: *mecanismul de feed-back*. Domeniul *Inteligenței artificiale (AI)* este cel care modelează și studiază aceste mecanisme umane de explorare, orientare și luare de decizii. El se bazează într-o mare măsură pe o altă importantă trăsătură a ființelor înzestrate cu capacitate de orientare: *învățarea*. În paginile ce urmează nu trebuie să se facă confuzie între abordarea problemei recunoașterii formelor și imaginilor din perspectiva DIP, cu abordarea aceleiași probleme din perspectiva AI, abordare care nu este obiectul acestui studiu.

Analiza și modelarea problemei recunoașterii.

Percepția vizuală - noțiuni fundamentale.

Din punct de vedere al percepției umane, problema recunoașterii este separată de problema captării senzațiilor primare (input-ul). Imitarea capacităților umane de percepție vizuală (oculară), perfecționarea și extinderea acestora este obiectul de studiu colaborat al medicinei, al ingineriei electronice și al fizicii optice. Pentru noi este importantă doar abordarea psihologică-mentală a problemei recunoașterii, în special pentru descoperirea și modelarea mecanismelor cerebrale de percepție, explorare și orientare vizuală. Din această perspectivă studiile de specialitate [VPRE, pag.40-56] constată că problema explorării vizuale la om are grad foarte mare de complexitate. Studiul acestei probleme se află încă în etapa de punere a fundamentelor și este plină de necunoscute. Numărul factorilor (parametrilor) care au fost descoperiți ca fiind implicați este atât de mare încât nu se poate vorbi încă de o modelare precisă a acestui fenomen, așa cum putem sesiza din schema de sinteză următoare [VPRE, pag. 53].



O altă concluzie care este extrasă din aceste studii este faptul că percepția și explorarea vizuală umană este o problemă avînd o complexitate atît "pe verticală" cît și "pe orizontală", fapt ce complică mult lucrurile din perspectiva modelării ei cu ajutorul calculatorului. Astfel, se vorbește de mecanisme de percepție vizuală structurate și îmbunătățite în timp, pe grupuri de vîrste: de la 0 la 2 ani, între 2 și 7 ani, între 7 și 12 ani și perioada peste 12 ani, aceasta constituind o extindere "pe verticală" a problemei și cuprinzînd numeroase mecanisme (de exemplu *învățarea*) care sînt modelate mai ales în domeniul inteligenței artificiale. Dar se constată la aceleași grupe de vîrstă, deci o extindere "pe orizontală", existența unor *strategii tipologice* de explorare vizuală în funcție de trăsăturile psihologice ale diferiților subiecți umani (și nu în funcție de starea de sănătate psiho-fiziologică a lor). Prin urmărirea mișcărilor și traseelor de explorare oculară a imaginilor-sarcină s-a putut constata chiar mai mult decît atît: există *strategii conjuncturale (orientate pe sarcină)* diferite de explorare vizuală la același subiect uman, ce depind atît de gradul de complexitate /importanța sarcinii de explorare vizuală asumate cît și de aptitudinile sau starea generală a subiectului.

În concluzie, din perspectivă psihologică-cognitivă, problema recunoașterii nu poate fi separată de cele două probleme cu care ea se întrepătrunde strîns:

1. *problema extragerii semnificațiilor* dintr-o formă sau imagine;
2. *problema interpretării conținutului* imaginilor pe baza tipologiei și formației cognitiv-psihologică a subiecților umani (tipologii ce diferă la vîrste diferite dar și la aceeași vîrstă).

Prin studiul sistematic al capacităților de explorare vizuală a copiilor foarte mici s-a putut constata existența următoarele trăsături fundamentale care sînt încă un deziderat major pentru sistemele automate de recunoaștere, urmărire și orientare spațială:

- a) *Fixarea și urmărirea*. În decursul primelor luni de viață copilul ajunge la o deplină capacitate de fixare și urmărire în cîmpul vizual (mai exact, e vorba de zona foveală) a subiectului (țintei) recunoscut și dorit.
- b) *Discriminarea figură – fond*. La numai 15 zile de la naștere copilul căruia i se arată o imagine conținînd un triunghi negru pe un fond alb sau o imagine bicoloră își îndreaptă cu precizie privirea spre contorul triunghiului (mai

ales spre colțuri) respectiv spre zona de graniță dintre cele două zone bicolore ceea ce permite să se tragă concluzia că procesul de discriminare dintre fond și figură (simplă) se însușește foarte repede.

- c) *Percepția formelor, modelelor și obiectelor.* După numai 9 luni de viață copilul este capabil să recunoască și să deosebească obiectele noi de cele vechi, sau imagini conținând modele noi de imagini cu modele (texturi) vechi. El este capabil să recunoască și să diferențieze cu precizie obiectele după cele trei atribute de bază: *culoare, mărime și formă.*
- d) *Diferențierile perceptivă și preferințele.* Încă de la vârsta de două săptămâni copilul începe să aibă preferințe pentru unele forme și culori, preferințe care evoluează și se schimbă cu vârsta. Astfel, pînă la vârsta de două luni toți copiii preferă să examineze o imagine în dungii față de una cu cercuri, situație care se inversează după trecerea celor două luni. Copiii de 7 săptămâni nu pot încă reține nici una din atributele unui obiect dar îi țin minte prezența; după 11 săptămâni copiii rețin un atribut al obiectului – forma, iar după 14 săptămâni ei rețin deja două atribute – forma și culoarea.
- e) *Organizarea perceptivă.* Prin urmărirea și înregistrarea mișcărilor oculare (mai exact, a *sacadelor* oculare) s-a putut trage concluzia că, încă din primele zile după naștere, în timpul inspecțiilor oculare sînt prezente la noi născuți mecanisme de organizare și de structurare a procesului de percepție vizuală. Acest fapt are o importanță majoră pentru obținerea pînă la vârsta de 4-5 luni a coordonării ochi-mîna și a capacității de apucare și manipulare a obiectelor, în funcție de forma și de mărimea lor.
- f) *Constanțele perceptivă.* Este evident că, în timp ce un obiect își păstrează proprietățile tactile-kinestezice oriunde s-ar deplasa în câmpul vizual al copilului, totuși imaginea acestuia (deși obiectul este neschimbat) își modifică atît forma cît și mărimea în funcție de poziția pe care o ocupă în spațiu. S-a constatat că, pînă la împlinirea vârstei de un an, copilul își dezvoltă complet capacitatea de a recunoaște și identifica obiectele indiferent de proiecția imaginii lor pe retină, el reușind să stabilească o legătură cauzală între obiect și diferitele forme ale sale, așa cum sînt ele percepute din perspective vizuale diferite – din unghiuri și de la distanțe diferite.

Din scurtele extrase de mai sus [VPRE, pag.40-45] se poate conchide că, pentru cei ce se ocupă cu modelarea problemei recunoașterii formelor și imaginilor, este încă un țel deosebit de înalt acela de a imita capacitățile de explorare vizuală măcar ale unui copil de doi ani.

Totuși, din constatările enunțate putem să extragem niște idei prețioase pentru proiectarea algoritmilor de recunoaștere. Există șase trăsături fundamentale ale procesului de explorare vizuală, așa cum apare și se dezvoltă el la ființele umane:

1. *Fixarea și urmărirea figurii*
2. *Discriminarea figură – fond*
 1. *Percepția formelor, modelelor și obiectelor*
 2. *Diferențierile perceptivă / preferințele*
 3. *Organizarea perceptivă*
 4. *Constanțele perceptivă*

Prin trecerea în revistă a numeroaselor articole de specialitate se poate afirma că doar primele trei trăsături sunt luate în calcul la modelarea și proiectarea sistemele de recunoaștere automată a formelor și imaginilor. Este aceasta o dovadă în plus a complexității și dificultății în cazul modelării explorării vizuale.

Se poate constata că *fixarea și urmărirea* este imitată și realizată deja în cazul sistemelor de supraveghere automată, în cazul sistemelor de dirijare a rachetelor către ținte, etc. ce utilizează algoritmi foarte performanți (eficienți) de analiză în timp real a cadrelor video ce se succed. În această situație, zona centrală a imaginilor captate, ce este analizată foarte rapid, corespunde cu zona foveală de pe retina ochiului observatorului uman iar efortul algoritmului constă în a menține "ținta" aflată în mișcare în centrul imaginii, prin comenzi adecvate asupra orientării obiectivului camerei video. Un amănunt mai puțin cunoscut și oarecum trecut cu vederea, dar care constituie unul din eforturile majore ale algoritmilor de fixare și urmărire, este faptul că "ținta" trebuie mai întâi localizată și adusă în mijlocul imaginii pentru a putea fi apoi fixată și urmărită continuu, exact ca în cazul fixării privirii și implicit a imaginii obiectului în zona foveală. Ori acest lucru nu este deloc ușor în cazul unei ținte aflată în continuă mișcare într-un spațiu tridimensional! De aceea, de multe ori este necesar efortul conjugat al mai multor camere de luat imagini (sau sisteme

radar) pentru a se putea realiza rapid localizarea spațială 3-D a țintei, folosind de exemplu, o metodă topografică de *triangulație*.

Discriminarea figură – fond se realizează prin algoritmi de detecție automată a muchiilor și conturilor, algoritmi ce se bazează pe determinarea zonelor de discontinuitate din imagini.

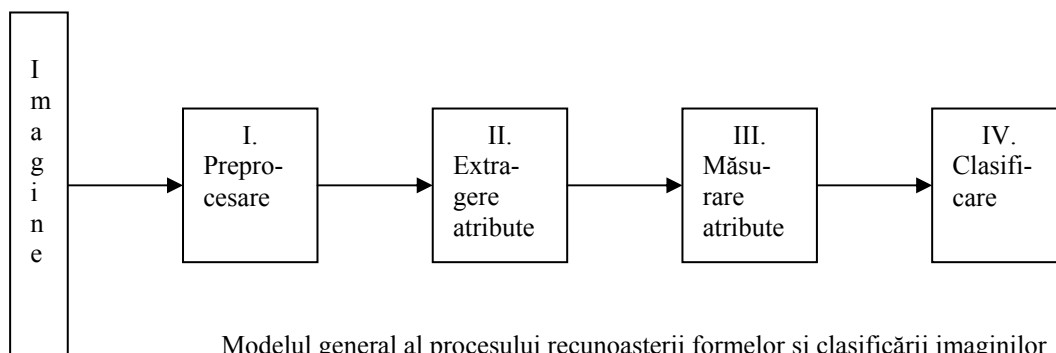
Percepția formelor, modelelor și obiectelor este trăsătura care ridică cele mai serioase probleme proiectanților de algoritmi de recunoaștere și constituie efortul principal al procesului de modelare, după cum se va putea vedea în continuare.

Diferențierile perceptive și preferințele, Organizarea perceptivă și Constanțele perceptive sînt trăsături ale mecanismelor de percepție vizuală care fac subiectul de studiu al *inteligenței artificiale* și presupun modelarea și simularea unor mecanisme extrem de complexe ca *extragerea semnificațiilor* și *recunoașterea prin învățare*.

Recunoașterea formelor și clasificarea imaginilor – modelare teoretică

Așa cum am spus deja, pentru a se face distincție clară între problema recunoașterii din perspectiva prelucrării digitale a imaginilor și aceeași problemă din perspectiva inteligenței artificiale și roboticii, terminologia de specialitate folosită în limba engleză este *pattern recognition* [GOW, pag. 693] și mai puțin *object recognition*. Chiar dacă vom folosi expresii ca *recunoașterea formelor*, *recunoașterea obiectelor* sau *clasificarea imaginilor*, în continuare ne vom referi doar la ceea ce literatura engleză de specialitate numește *pattern recognition*.

Recunoașterea formelor și/sau clasificarea imaginilor este un proces ce are la bază un model în patru pași sau etape. Chiar dacă în sistemele automate de recunoaștere și clasificare etapele nu apar întotdeauna ca fiind separate, ele vor fi în continuare prezentate distinct. Se urmărește astfel scoaterea în evidență a trăsăturile esențiale ale fiecărui pas /etapă, așa cum se poate observa în schema următoare:



Modelul general al procesului recunoașterii formelor și clasificării imaginilor

I. **Preprocesarea imaginii.** De obicei această etapă este privită împreună cu etapa a doua dar importanța capitală pe care noi dorim să o acordăm etapei de extragere a atributelor ne-a făcut să le tratăm separat. Prin această etapă de preprocesare se înțelege de fapt aplicarea, unor algoritmi DIP specializați de îmbunătățire a calității imaginii. Unii dintre cei mai uzitați algoritmi de preprocesare, prezentați în [ISP 1], sînt:

- algoritmi de amplificare a contrastului;
- algoritmi de eliminarea paraziților și a "zgomotului" (*noise reduction*);
- algoritmi de transformare Fourier a imaginii, folosiți mai ales pentru analizarea texturii imaginii;

Rezultatul final al acestei etape poate fi: aceeași imagine dar îmbunătățită, avînd unele detalii puse în evidență; aceeași imagine dar într-o altă descriere ce facilitează recunoașterea obiectelor componente; doar părți componente ale imaginii în diverse descrieri și formalizări.

II. **Extragerea atributelor sau descriptorilor de imagine (*feature extraction*).** Este etapa cheie, este cea care dă măsura performanțelor și a profesionalismului aplicației de recunoaștere. Ea constă în aplicarea, împreună sau separat (după necesități), a unor algoritmi DIP specializați care urmează să fie prezentați în această lucrare, cum ar fi:

- îngroșarea, extragerea și/sau detectarea liniilor, muchiilor sau conturilor;
- concatenarea segmentelor și conectarea conturilor;

- segmentarea imaginii în zone de interes;
- introducerea unor reprezentări sau descrieri ajutătoare a conturilor, zonelor sau obiectelor cu contururi delimitate prin metodele anterioare;

Alegerea unor atribute sau caracteristici cât mai potrivite este cheia succesului algoritmilor de recunoaștere. Deși, în general vorbind, obiectele sînt înzestrate cu doar trei atribute principale *mărimea*, *culoarea* și *forma*, totuși în practică determinarea cu exactitate și completă a celor trei caracteristici fundamentale este extrem de dificilă sau chiar imposibilă și presupune proiectarea și implementarea unor metode de aproximare a lor. Iar numărul posibilităților de aproximare a mărimii și formei combinate între ele și înmulțite cu numărul imens de nuanțe de culori (ce sînt conținute de exemplu în textura obiectelor) conduce la un număr foarte mare de alegeri posibile. Secretul reușitei în problema recunoașterii constă tocmai în reducerea acestui număr uriaș de posibilități de a alege dintre mărimile, formele și culorile "sugerate" de imagine. Rezultatul final al acestei etape este un vector de n atribute extrase (v_1, v_2, \dots, v_n) nu neapărat numerice.

III. **Măsurarea atributelor sau descriptorilor (*feature/pattern measurement*).** În practică această etapă nu este separată de cea anterioară întrucît, odată alegerea atributelor fiind făcută, se presupune ca pentru ele a fost aleasă deja o metodă de măsurare, evaluare și comparare (discriminare). Totuși, cărțile de specialitate scot în evidență această etapă și, deseori, doar ea este cea care e prezentată în capitolul *Object recognition and pattern clasification*. Un motiv pentru care această etapă este atît de bine tratată teoretic este faptul că există un aparat matematic bine pus la punct – *Teoria măsurii* – cu ajutorul căruia se pot introduce diferite metrici n -dimensionale sau metode discriminative eficiente. În practică se constată că folosirea unei singure măsuri (a unei singure metrici) nu este o metodă suficient de discriminativă (nu conduce la separarea precisă în clase distincte) și de aceea majoritatea aplicațiilor folosesc mai multe metrici simultan (așa cum, parafrazînd, și omul folosește doi ochi în loc de unul). Atragem

atenția că aceasta este etapa "vulnerabilă" a întregului proces și poate conduce în final la confuzii, rezultate eronate sau chiar mistificări. Exprimarea populară cu tîlc "*Rezultatul măsurătorii depinde de ce măsură ai și de cine măsoară*" se potrivește și în această situație. Rezultatul final al acestei etape este de obicei o valoare numerică uni- sau multi-dimensională (un vector) privită ca fiind "distanța" vectorului de attribute față de "bornele" de clasificare.

IV. **Clasificarea imaginii (*pattern classification*).** Este etapa finală în care se colaborează rezultatele măsurătorilor multiple anterioare (realizate cu mai multe metrice diferite) și se stabilește apartenența formei, obiectului sau imaginii descris prin vectorul de attribute la o clasă. Pentru creșterea vitezei de clasificare apartenența la o clasă se stabilește cu ajutorul unor metode ce sînt specifice strategiei de recunoaștere folosite. Ele poartă denumirea de *criterii matematice sau funcții de apartenență*, ce se bazează pe poziția spațială a vectorului de semnificați față de granițele (liniile, planele sau hiperplanele) de delimitare între clase, sau *reguli structurale (sintactice) de decizie (acceptare)* a apartenenței la o clasă. Aplicarea practică a acestor metode de clasificare nu este întotdeauna o problemă simplă de calcul ci, de cele mai multe ori, necesită un aparat teoretic sofisticat și niște algoritmi pe măsură. Rezultatul final al etapei de clasificare este numărul *C* al clasei de apartenență sau direct denumirea ei, de unde se obține implicit și numele obiectului ce se dorea a fi recunoscut.

Literatura de specialitate grupează modelele și metodele generale de recunoaștere și clasificare în patru mari categorii sau strategii, denumite *abordări (pattern recognition approaches)* [JDM, pag.6] :

- A. Recunoașterea prin potrivirea cea mai bună (*template matching approach*);
- B. Recunoașterea prin metode statistice (*statistical approach*);
- C. Recunoașterea cu ajutorul rețelelor neuronale (*neural networks approach*);
- D. Recunoașterea sintactică sau structurală (*syntactic or structural approach*);

Există și o altă clasificare în doar două mari categorii [GOW, pag.XIV]:

A. Recunoașterea bazată pe metode teoretice de decizie (*decision-theoretic methods*);

B. Recunoașterea sintactică sau structurală (*syntactic or structural methods*);

În cea de a doua clasificare, metodele A, B, C ale clasificării anterioare sînt cuprinse toate într-o singură categorie. Diferența între cele două clasificări constă în faptul că cele trei metode ce au fost grupate împreună, deși diferite între ele, au comună totuși la etapa clasificării aceeași abordare funcțională. Categoria D din prima clasificare este aceeași cu categoria B de metode din cea de-a doua clasificare și conține în etapa finală a procesului de recunoaștere (etapa clasificării) aceeași metodă de discriminare bazată pe algoritmi de apartenență, folosind *arbori sintactici de acceptare*. Această metodă de recunoaștere a apartenenței la o clasă este o metodă structurală ce este inspirată din *Teoria gramaticilor și limbajelor formale*.

Tabelul următor, adaptat după [JDM, pag.6], ne oferă o privire de ansamblu asupra strategiilor (abordărilor) de recunoaștere a formelor și clasificare a imaginilor scoțînd în evidență modul specific, particular, de abordare a fiecărei strategii. De asemenea este prezentat distinct aparatul matematic implicat în analiza și modelarea procesului de recunoaștere în cazul fiecăreia din cele patru abordări.

Strategia de abordare / Modelarea matematică	Modele de reprezentare a atributelor (descriptorilor)	Modele / funcții de discriminare (recunoaștere)	Modele / criterii (condiții) de discriminare
Potrivirea cea mai bună (<i>Template matching approach</i>)	mostre de pixeli, contururi, forme	funcția de corelație, diferite metrice	Minimizarea erorii/maximizare a potrivirii
Metode statistice (<i>Statistical approach</i>)	valori numerice, contururi, forme	funcția de discriminare	Minimizarea pierderii (riscului)
Rețele neuronale (<i>Neural networks approach</i>)	forme și mărimi numerice de antrenament	funcție de decizie neuronală	Minimizarea erorii la învățare
Sintactică / structurală (<i>Syntactic / structural approach</i>)	șiruri de coduri (simboluri)	reguli de derivare (sintactice)	arbore de derivare sintactică

Tabelul sintetic de descriere a celor patru strategii de modelare a problemei recunoașterii și clasificării automate a imaginilor

Tabelul sintetic al algoritmilor specifici folosiți la recunoașterea imaginilor grupați pe modele /strategii și repartizați pe cele patru etape ale procesului de recunoaștere

Etapa în procesul recunoașterii → Modelul /Strategia de abordare ↓	I. Preprocesarea imaginii	II. Extragerea atributelor (<i>feature extraction</i>)	III. Măsurarea atributelor (<i>feature measurement</i>)	IV. Clasificarea (<i>pattern classification</i>)
A. Potrivirea cea mai bună (<i>Template matching approach</i>)	restaurarea imaginii, îmbunătățirea calității, amplificarea contrastului; eliminarea paraziților și a "zgomotului" (<i>noise reduction</i>); transformarea Fourier a imaginii, folosită mai ales pentru analizarea texturii	algoritmi morfologici: <i>dilatate, eroziune, umplere, înfășurătoarea convexă, scheletizare</i> ; algoritmi de segmentare a imaginii: <i>detectarea discontinuităților - puncte, linii, muchii (edges), conectarea segmentelor (edge linking), determinarea conturilor (boundaries), filtre globale și adaptative (thresholdings), histograma</i> ; algoritmi de reprezentare și descriere a formelor: <i>descrierea conturilor, momente statistice invariante, descriptori Fourier, texturi</i> ;	algoritmi de determinare a potrivirii: <i>minimum distance classifier, matching by correlation</i>	algoritmi de determinare a graniței (<i>decision boundary</i>) între clase;
B. Modelare prin metode statistice (<i>Statistical approach</i>)			metode statistice de minimizare a pierderii (a riscului) (<i>conditional average risk statistical equation</i>)	algoritmi și funcții de discriminare statistică: <i>funcția de clasificare Bayes (optimum statistical Bayes classifier)</i> ;
C. Modelare prin rețele neuronale (<i>Neural networks approach</i>)			perceptroni, metode de antrenare, rețele neuronale multistrat, algoritmi de învățare	metode neuronale de antrenare și învățare, de exemplu <i>training by back-propagation</i> ;
D. Modelare sintactică / structurală (<i>Syntactic / structural approach</i>)			algoritmi de reprezentare și descriere a formelor: <i>înlănțuire de coduri, aproximări poligonale, semnături, descriptori topologici, shape numbers</i>	gramatici și reguli de derivare sintactică, arbori de derivare (analiză) sintactică, automate finite de recunoaștere



Înainte de trecerea în revistă a numeroaselor strategii de modelare și a algoritmilor utilizați în sistemele de recunoaștere a formelor și de clasificare a imaginilor, tabelul de mai sus cuprinde clasificarea sintetică unitară - pe strategii și modele - a tuturor algoritmilor folosiți în recunoaștere, atât de numeroși și de diferiți. Scopul urmărit este de a aduce un spor de claritate, care poate să conducă mai apoi la un plus de utilitate practică. Parcurgerea tabelului pe linii și pe coloane permite realizarea de corelații între modelul matematic și strategia de recunoaștere folosite de un algoritm și etapa din procesul de recunoaștere la care acel algoritmul este adecvat, stabilindu-se astfel cu acuratețe locul și rolul lui practic.

Parcurgând literatura de specialitate – de exemplu colecția articolelor din *International Journal on Document Analysis and Recognition, Springer-Verlag, 2004* – se pot identifica o varietate impresionantă de metode, tehnici și algoritmi diferiți folosite în abordarea dificilei probleme a recunoașterii și clasificării automate. Pentru edificare și pentru a crea o imagine justă a multitudinii de perspective diferite existente, le vom trece în continuare în revistă. Pentru acuratețe și pentru a permite rafinarea ulterioară a termenilor ce apar, ei vor apărea în limba engleză fără traducere în limba română.

Să observăm mai întâi că în articolele de sinteză pe această temă se oferă explicit strategiile de modelare folosite, numite tehnici, metode sau abordări:

În [TCY] sunt oferite trei strategii generale de recunoaștere și regăsire a imaginilor: *the signature-based technique, the partition-based approach and the cluster-based approach*.

[MSTS&N] propune o clasificare asemănătoare, tot în trei mari strategii: *text-based retrieval, content-based retrieval, and semantic-based retrieval*. Tot aici sunt trecute în revistă metodele utilizate de sistemele autoamte de recunoaștere și regăsire a imaginilor existente pe piață. Există astfel șase criterii generale de căutare /recunoaștere /regăsire: *Color Content (CC), Shape Content (SC), Texture Content (TC), Color Structure (CS), Brightness Structure (BS) and Aspect Ratio (AR)*.

Pe lângă aceste strategii sau modele generale, pot fi găsite și alte abordări generale: *component classification using fuzzy approach, User and Task-Based Approach, contextual clues and automatic pseudofeedback, relevance feedback, etc.*

Fără a epuiza subiectul, lista următoare conține o alte metode /tehnici /algoritmi ce combină strategiile de mai sus sau introduc altele noi: *Fourier transform for*

segmentation, wavelets analysis methods, multi-level color histogram, Similarity Measure methods, Dominant color classification, joint histograms, Edge angle distribution, 3D neighborhood graph model, Hough transform based methods, data covariance matrix based methods, connected component analysis, Statistical image differences methods, degradation features based techniques, Clustering methods, Skew estimation methods, skew detection using morphology, classification and segmentation using support vector machines, Multilevel thresholding – Region growing – Complex background analysis, classification and segmentation using boundary characteristics, etc.

BIBLIOGRAFIE

[VPRE] – VASILE PREDA – *Explorarea vizuală. Cercetări fundamentale și aplicative*, Editura științifică și enciclopedică, București, 1988

[GOW] - GONZALEZ R., WOODS R. - *Digital Image Processing*, Prentice Hall, 2002, 2nd Edition

[JDM] – A. JAIN, R. DUIN, J. MAO – *Statistical Pattern Recognition: A Review*, IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 22, No. 1, January 2000

[BKKP] - BEZDEK J., KELLER J., KRISHNAPURAM R., PAL N.- *Fuzzy Models and Algorithms for Pattern Recogniton and Image Processing*, Kluwer Academic Publishers, 1999

[VLA] AUREL VLAICU – *Prelucrarea digitală a imaginilor*, MicroINFORMATICA, Cluj-Napoca, 1997

[ISP 1] IOAN ISPAS – *Algoritmi de prelucrare digitală a imaginilor*, Referat, Univ. Babeș-Bolyai, Facultatea de Matematică-Informatică, Cluj-Napoca, 2003

[ISP 2] IOAN ISPAS – *Baze de date de imagini*, Referat, Univ. Babeș-Bolyai, Facultatea de Matematică-Informatică, Cluj-Napoca, 2003

[ISP 3] IOAN ISPAS – *Algoritmi de recunoașterea formelor și clasificarea automată a imaginilor*, Referat, Univ. Babeș-Bolyai, Facultatea de Matematică-Informatică, Cluj-Napoca, 2003

[TCY] KIAN-LEE TAN, BENG CHIN OOI, CHIA YEOW YEE - *An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases*, *Multimedia Tools and Applications*, 14, 55–78, 2001, Kluwer Academic Publishers

[MSTS&N] MARJO MARKKULA, EERO SORMUNEN, MARIUS TICO, BEMMU SEPPONEN AND KATJA NIRKKONEN - *A Test Collection for the Evaluation of Content-Based Image Retrieval Algorithms - A User and Task-Based Approach*, *Information Retrieval*, 4, 275–293, 2001, Kluwer Academic Publishers

În lucrare s-au folosit spre ilustrare imagini oferite public: www.prenhall.com/gonzalezwoods/ - „the companion web site for students, instructors and practitioners” oferit de autorii bestseller-ului *Digital Image Processing*