

Asupra dificultății Problemei recunoașterii și clasificării imaginilor (IRC)

Ioan Ispas, UPM Tîrgu Mureș
(ispas@science.upm.ro)

Abstract: This paper it's an overview over the most important issues concerning the difficulty of the IRC Problem. It is focusing on the key step of the recognition process, the segmentation/discrimination/feature extraction step. Also, it emphasize on the less known importance of the second part of the human vision – the inverse optic or the tri-dimensional reconstruction of the scene.

Perspectiva algoritmică asupra dificultății problemelor

Definiția 1. Dificultatea unei probleme este dată de complexitatea algoritmului optim de soluționare a ei. ([ISIFL2008])

Matematic vorbind, pe baza Tezei Turing-Church, noțiunea de algoritm este echivalată printr-o Mașină Turing sau printr-o funcție recursiv-calculabilă a lui Church. Complexitatea unui algoritm, în acest context, este dată printr-o funcție O ce exprimă relația dintre numărul operațiilor de bază (pașilor) ale algoritmului și dimensiunea vectorului datelor de intrare.

De exemplu, fie $A(I_n)$ un algoritm având ca date de intrare vectorul I_n , atunci complexitatea algoritmului este:

$$O(A) = f(n),$$

unde funcția f este funcția de complexitate a algoritmului. Ordinul de mărime sau viteza de creștere a acestei funcții exprimă complexitatea algoritmului. Ea determină mai exact modalitatea de creștere a consumului de resurse (timp sau spațiu) calculator odată cu creșterea dimensiunii vectorului de intrare.

Dificultatea problemei va fi exprimată atunci în același mod în care este exprimată complexitatea algoritmului optim: problemă liniară, problemă polinomială, problemă exponențială, etc.

Propoziția 1. Se numește problemă închisă o problemă a cărei dificultate este cunoscută (demonstrată).

Evident, pentru închiderea unei probleme este necesară determinarea (și eventual proiectarea) algoritmului optim de soluționare a acesteia, măcar printr-o strategie de rezolvare. Din păcate, proiectarea și demonstrarea optimalității unui algoritm este fapt extrem de dificil. Se cunoaște că numărul algoritmilor optimali cunoscuți (de utilitate largă) nu depășește valoarea zece sau cinsprezece. De aceea, numărul problemelor închise în informatică este extrem de redus!

Definiția 2. Printr-un algoritm optim sau optimal se înțelege cel mai bun (cel mai eficient) algoritm dintr-o clasă bine determinată de algoritmi ce soluționează problema.

Cel mai bun (cel mai eficient) algoritm este acel algoritm care consumă cea mai redusă cantitate de resurse de timp sau spațiu. Optimalitatea se demonstrează matematic în clasa de algoritmi bine delimitată pe baza operației de bază.

Definiția 3. Printr-o clasă de algoritmi se înțelege mulțimea algoritmilor de soluționare a problemei ce au în comun aceeași strategie de rezolvare, ce se bazează pe aceeași operație de bază inevitabilă.

Exemple de operații de bază: adunarea (sau orice operație algebrică) a două valori numerice, comparația (numerică, logică sau complexă) a două elemente, etc. Operația de bază inerentă strategiei de rezolvare, fiind inevitabilă, permite includerea în clasa de algoritmi (delimitată prin ea) a tuturor algoritmilor, atât a celor proiectați în trecut cât și a celor ce vor fi proiectați în viitor pe baza aceleiași strategii. Acest fapt permite demonstrarea matematică a optimalității unui algoritm, fapt ce exclude posibilitatea descoperirii ulterioare a unui algoritm de soluționare mai bun, pe baza aceleiași strategii de soluționare.

De exemplu, căutarea într-o listă ordonată crescător după o cheie de căutare alfa-numerică este o problemă logaritmică. Acest fapt este dat de existența unui algoritm optimal de căutare prin comparații de complexitate logaritmică: algoritmul de căutare binară. Problema căutării în listă ordonată este închisă din perspectiva rezolvării pe bază de comparații. Problema sortării unei liste (șir) de valori numerice este deasemenea închisă din perspectiva soluționării pe baza comparațiilor, dar algoritmul optimal de complexitate liniar-polinomială nu a fost încă descoperit și proiectat.

Probleme elementare, aparent simple, cum ar fi înmulțirea matricilor numerice, sunt încă deschise, neexistând un algoritm optim de soluționare a lor. Majoritatea problemelor cu grafuri sunt deschise și au dificultate ridicată, ele intrând în categoria problemelor NP-complete. Soluțiile cunoscute pentru acestea sunt fie algoritmi exponențiali, fie mașini Turing nondeterministice cu complexitate polinomială (NP).

Problema Recunoașterii și Clasificării imaginilor (Image Recognition and Classification Problem - IRC)

Procesul Recunoașterii și Clasificării imaginilor este descris în următoarea diagramă de proces ([ISI2007], [ISIFO2008]).

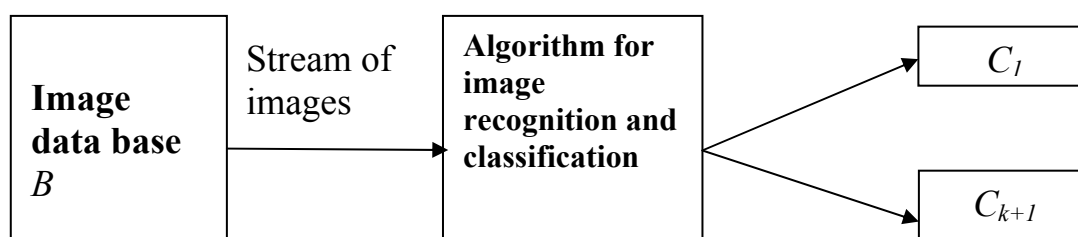


Diagrama 1. Procesul Recunoașterii și Clasificării imaginilor

Acest proces este des întâlnit în practică și este de importanță strategică. Imaginile provin dintr-un flux de date (bază de date) și trebuie să fie clasificate într-un timp cât mai scurt (optimal dacă se poate) pe baza unui obiect (subiect) ce este recunoscut în imagine. La clasele de imagini recunoscute și clasificate C_1, C_2, \dots, C_k se adaugă clasa C_{k+1} a imaginilor nerecunoscute.

Literatura de specialitate descrie acest proces în patru etape sau pași, ce pot fi văzuți în diagrama următoare ([GOW2002], [JAIN2000], [ISI2005]):

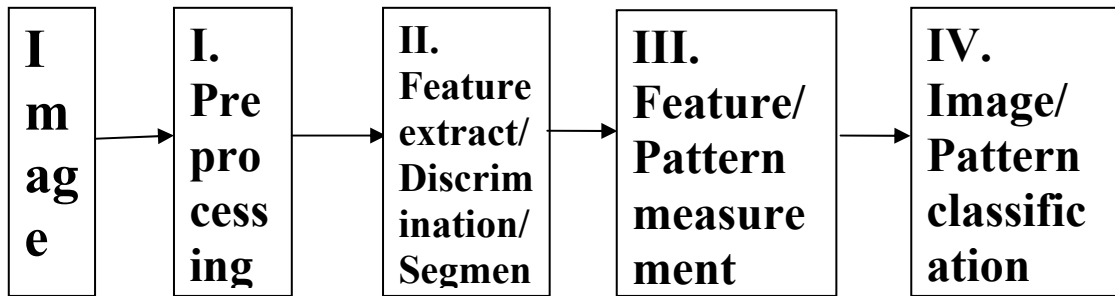
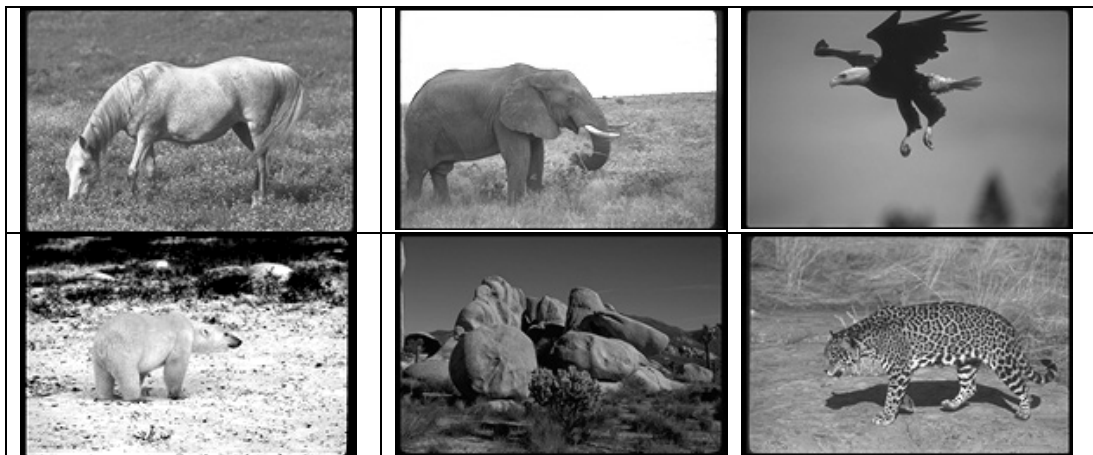


Diagrama 2. Recunoașterea și Clasificarea imaginilor în patru pași (etape)

Fiecare etapă a acestui proces este inevitabilă, chiar dacă în realitate cele patru etape nu sunt distincte ci întrepătrunse ([ISI2005], [ISI2007]). Este unanim cunoscut de cei care lucrează la implementarea acestui proces că etapa a II-a este etapa cheie și pasul critic al întregului proces. La această etapă este necesară segmentarea imaginii în subansamble în vederea extragerii unor atribute esențiale ce permit discriminarea obiectului de căutat în imagine.

Problema IRC. Să se clasifice următorul flux de imagini în $k+1$ clase, fiecare clasă fiind definită de unul din cele k “personaje” (subiecți de recunoscut): cai, elefanți, urși, vulturi, leoparzi, etc. așa cum se poate vedea în tabelul 1.



Tabelul 1

O strategie cu caracter general ce surprinde esențialul acestui proces este prezentată în următorul algoritm ([SIFO2008], [ISI2008]):

```

Algorithm IRC(image I);
For i=1,k do
  If Discrimination(Ok,I) return(k);
Return(k+1);
  
```

Algoritmul IRC de clasificare directă are ca și input imaginea I (matrice rectangulară de pixeli numerici) și apelează sub-algoritmul **Discrimination**(O_k,I) ce

inspectează imaginea I pentru recunoașterea obiectului O_k . Evident, complexitatea algoritmului IRC este, în cel mai rău caz, de k ori complexitatea sub-algoritmului **Discrimination**, care poate fi privit în același timp ca un algoritm de segmentare și de extragere de atribute.

Etapa Segmentării/ Discriminării/ Extragerii atributelor – etapă cheie în recunoașterea imaginilor

La modul esențial, mulțimea atributelor unui obiect (subiect) într-o imagine se reduce la vectorul pereche *Formă-Conținut* (F,C). Aceste două atribute independente permit atât delimitarea (*Forma*) cât și identificarea (*Conținutul*) oricărui obiect/subiect (fig.1, 2 și 3).

Definiția 4. Prin forma (conturul) unui obiect se înțelege contorul, muchiile sau granița delimitatoare ale obiectului.

Definiția 5. Prin conținutul (coloristica) unui obiect se înțelege mulțimea informațiilor despre culoarea, întinderea (suprafața) și ansamblare ale unui obiect.

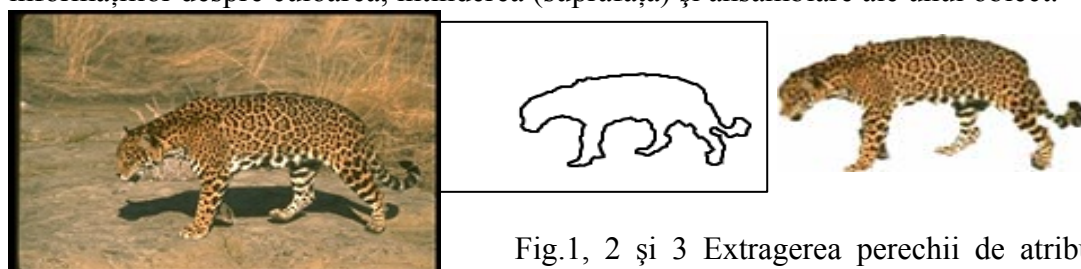


Fig.1, 2 și 3 Extragerea perechii de atribute (F,C) de *formă* (contur) și *conținut* (coloristică)

Acest demers de obținere a vectorului de atribute (F,C) este nu doar obligatoriu ci și un aspect dificil al întregii probleme. În figura 4 ([MALL2000]) este ilustrat faptul că este mai corect să afirmăm că input-ul procesului de recunoaștere este, nu o imagine color sau alb-negru, ci o matrice rectangulară de intensități luminoase. Privind imaginea din dreapta, poate este de înțeles de ce recunoașterea subiectului nu este o sarcină ușoară.

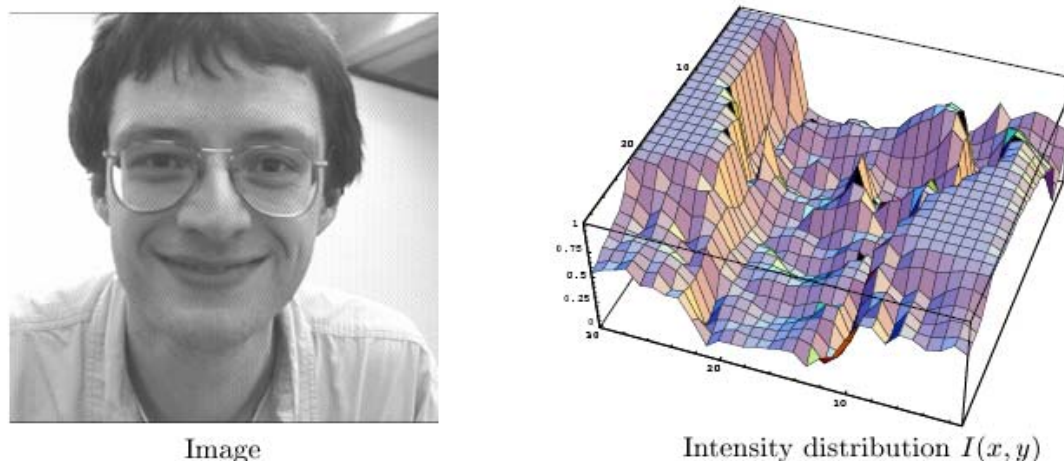
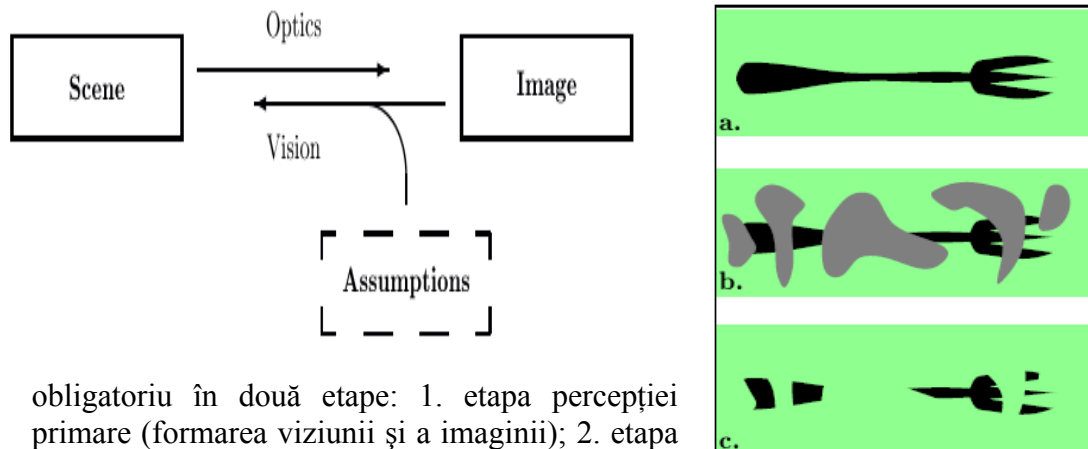


Fig.4 Recunoașterea pornește de la informațiile de intensitate (dreapta)

Studii ([MALL2000]) recente asupra întregului fenomen al viziunii umane confirmă faptul că recunoașterea subiectelor în cazul omului este un proces



obligatoriu în două etape: 1. etapa percepției primare (formarea viziunii și a imaginii); 2. etapa reconstrucției optice tridimensionale a scenei (“inverse vision”). Figurile 5 și 6 conțin diagrama sintetică actualizată a întregului fenomen, precum și un exemplu semnificativ ([MALL2000]):

Fig.5 și 6 Recunoașterea (viziunea) umană este un proces de reconstrucție optică



Fig. 7 Exemple de CAPTCHA ce nu pot fi recunoscute de softurile malware

Această idee poate fi regăsită și în cazul mecanismului CAPTCHA ("Completely Automated Public Turing test to tell Computers and Humans Apart") ce este folosit pe scară largă pentru discriminarea între utilizatorii umani și roboții malware (figura 7). Ceea ce părea, la apariția acestei modalități, să confirme doar dificultatea extremă a etapei de segmentare/discriminare, s-a dovedit a fi mai mult decât atât. Este chiar o confirmare a existenței etapei de reconstrucție optică (viziune inversă) în cazul procesului de recunoaștere la om.

Concluzie asupra dificultății Problemei recunoașterii și clasificării imaginilor (IRC)

Atât timp cât se va încerca modelarea matematică doar a primei faze a procesului recunoașterii ([ISI2008]), eficiența și corectitudinea soluțiilor problemei IRC va rămâne la actualul nivel. Cu excepția unor situații singulare și particularizate ([JAIN2000]), nu se cunoaște nici o soluție generală care să permită recunoașterea, identificarea și clasificarea imaginilor după conținut. Înafara mecanismului CAPTCHA care “profită” de slăbiciunea actuală a segmentării și discriminării, mai

există și alți algoritmi care s-au strecurat prin “breșa” creată și se constituie în adevărate “strategii adverse” ce pot construi automat imagini de test aproape imposibil de clasificat ([ISI2009]).

Amploarea pe care a luat-o studiul domeniului Human Vision, versus dezvoltare uriașă a domeniului Computer Vision, ne permit să concluzionăm că soluționarea problemei deschise IRC a intrat într-o a doua mare și importantă etapă. În această etapă se realizează înțelegerea, modelarea și transferarea abilităților umane ce țin de procesul de reconstrucție optică (inverse vision). Această etapă este în fază de derulare (acumulare) atât la nivel teoretic cât și la nivelul implementărilor și validărilor practice.

Bibliografie

[GOW2002] Gonzalez R., Woods R., *Digital Image Processing*, (2nd Edition), Prentice Hall, 2002

[ISI2005] Ispas Ioan, *Sinteză asupra metodelor de recunoaștere și clasificare automată a imaginilor*, Proc. Of the 1st International Conf. on European Integration – Between Tradition and Modernity, Târgu Mureș, 22-24 sept., pag. 732-751, 2005

[ISI2007] Ispas Ioan, *The image recognition and classification, a four-stepped modeling*, Proc. Of the 2nd International Conf. on European Integration – Between Tradition and Modernity: pag. 124-132, Petru Maior University, Tirgu Mures, Sept 20-21, 2007

[ISIFL2008] Ispas Ioan, Franti Eduard, Lazo Florin, *The Complexity of the Algorithms for the Image Recognition and Classification (IRC)*, Mathematics and Computers in Science and Engineering, WSEAS Press, Computational methods and applied computing Conference, Istanbul, Turkey, May 27-30, pag. 160-164, 2008

[ISIFO2008] Ispas Ioan, Franti Eduard, Osiceanu Sanda, Stoian Marius, *Basic elements in the modeling of the problem of the image recognition and classification (IRC)*, Proceedings of the 6th WSEAS International Conference on APPLIED ELECTROMAGNETICS, WIRELESS AND OPTICAL COMMUNICATIONS (ELECTROSCIENCE '08), Trondheim, Norway, July 2-4, 2008, pag. 135-139

[ISI2008] Ispas Ioan, *Modeling Of The Image Recognition And Classification Problem (IRC)*, STUDIA UNIV. BABES-BOLYAI, INFORMATICA, Volume LIII, Number 1, pag. 113-120, 2008

[ISI2009] Ispas Ioan, *Hybridization Of The Attributes And The Difficulty Of The Problem Of The Image Recognition And Classification (IRC)*, Bio-Inspired Computational Methods Used for Difficult Problems Solving: Development of Intelligent and Complex Systems, Romania, 5-7 November 2008, American Institute of Physics Conference Proceedings: AIP Volume 1117, 2009

[JAIN2000] Jain A., Duin R., Mao J., *Statistical Pattern Recognition: A Review*, IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 22, No.1, pag. 720-736, January 2000

[MALL2000] Mallot H.A., Allen J.S. *Computational Vision, Information Processing in Perception and Visual Behavior* (2ed), MIT, 2000