



UNIVERSITATEA “POLITEHNICA“ din BUCUREȘTI

Facultatea de Automatică și Calculatoare

Departamentul de Calculatoare

REZUMAT TEZĂ DE DOCTORAT

Clasificarea în Imagini Digitale

Autor: ing. Octavian Georgescu

Conducător de doctorat: Prof. dr. ing. Florica Moldoveanu

BUCUREȘTI, 2012

## Prezentarea capitolelor lucrării

În cadrul analizei imaginilor, ne-am pus problema Clasificării în Imagini (CI) digitale. Cu toate că ne referim numai la aceste imagini, cele discutate de noi se pot generaliza în cadrul unei analize a semnalelor, imaginea putând fi văzută ca un semnal bidimensional. La modul general, clasificarea în (interiorul lui)  $X$  se poate face după anumite forme sau după anumite caracteristici ale primitivelor lui  $X$ . În cazul CI, aceste primitive sunt pixelii imaginii digitale. Așadar, CI se poate face prin metode de clasificare bazate pe identificarea de regiuni cu anumite geometrii (RB = ‘region based methods’) și / sau prin metode de clasificare bazate pe pixeli (PB = ‘pixel based methods’) cu anumite proprietăți. Putem vedea metodele RB ca metode de clasificare de nivel înalt comparativ cu metodele PB. Lucrarea noastră tratează în principal metodele PB. Vom menționa metodele RB, care mai poartă și denumirea generică de metode de Recunoaștere a Formelor, în special pentru a face trimitere la metodele PB.

Din punct de vedere al teoriei mulțimilor, CI înseamnă partiționarea imaginii (digitale) în mulțimi (clase de echivalență) de pixeli cu trăsături relativ comune. Partiționarea, în cele mai multe cazuri reale, nu se poate face exact, ci aproximativ, într-un anumit grad. În plus, partiționarea depinde de rezoluția / precizia / granularitatea conform căreia se face clasificarea.

Ca redactare, lucrarea e organizată conform ierarhiei: capitol – subcapitol – secțiune – subsecțiune. Din punct de vedere al contribuțiilor, lucrarea e orientată pragmatic și prezintă într-un tot unitar – poate fi și o carte folositoare pentru cei care, de la începători la avansați, se ocupă de analiza și, mai ales, de clasificarea în imagini – câteva contribuții (sinteze / observații) teoretice și contribuții practice. Întinderea lucrării a rezultat din scopul ei de a fi conținută în sine (totuși, s-a trecut repede peste teoria elementară). Prezentăm pe scurt conținutul lucrării, pe capitole, făcând precizări sumare în dreptul capitolelor de contribuție teoretică.

### **I§** Încadrarea problemei CI și descrierea ei în Domeniul Spațial și în Domeniul Probabilistic

- precizăm problemele practice ale CI și facem o scurtă incursiune teoretică a abordării lor în cel mai ușor de înțeles domeniu de reprezentare a imaginilor: cel spațial;

- plecăm de la principiul de bază de filtrare în scopul găsirii pixelilor de muchie (frontieră), a conturilor, pentru ca în subcapitolul “Reprezentarea formelor” să menționăm câteva dintre metodele RB tipice;

- subcapitolele “Tehnici de segmentare”, “Recunoașterea imaginilor prin potrivire” și “Tehnici de clasificare” duc discuția în aria efectivă a CI (metodele PB);

- subcapitolul “Domeniul Probabilistic” e prezent în acest capitol din două motive principale: 1) vedem Domeniul Probabilistic ca “istorie” a Domeniului Spațial și 2) acest domeniu e aproape la fel de ușor de abordat precum cel spațial.

### **II§** “Separabilitatea claselor”

- în primul subcapitol, “Descompunere ortogonală cu Unde și Undine”, tratăm acest subiect din perspectiva analizei semnalelor în Domeniul Frecvență și în Domeniul Frecvență-Timp (folosim denumirea “Frecvență-Timp” pentru că așa se regăsește în literatura de specialitate, deși, în cazul imaginilor, ar fi mai corect să spunem “Frecvență-Spațiu”, frecvențele de care vorbim fiind spațiale). Avem în vedere decelarea anumitor proprietăți în imagini după ce le trecem în domeniul frecvențelor prin aplicarea transformatei Fourier. Abordăm aici chestiunea dependenței rezoluției din domeniul Timp de cea din domeniul Frecvență, atât la nivelul teoremei eșantionării cât și la nivelul principiului incertitudinii (al lui Heisenberg). Ultima secțiune, II.1.5. “Relații fundamentale în jurul conceptului de corelare, între principalele domenii de reprezentare a semnalelor”, propune o sinteză originală în jurul problemei clasificării, problemă care poate fi văzută ca una de corelare sau de descompunere.

- în al doilea subcapitol, “Granularitate și gradație în Domeniul Fuzzy”, ne referim la clase separabile într-un anumit grad, clasa corespunzând unei “partiții fuzzy”. În literatura de specialitate “partițiile fuzzy” au diverse înțelesuri. Plecăm de la așa-numitele partiții / clase fuzzy, care se pot suprapune, putând fi proiectate ca seturi fuzzy ( [Shakhatreh & Hayajneh 2010] ), pentru a propune un termen care ni se pare mai semnificativ decât acela de partiție fuzzy, granulă ș.a.: PROTOTIPUL (vezi ultima secțiune: II.2.5. Clasa fuzzy = PROTOTIPUL). Generalizând conceptul de prototip, în domeniul frecvență, prototipul corespunde undei / undinei, iar în domeniul fuzzy, prototipul corespunde clasei. Cu precizarea că, în general (conform teoriei clasice), dacă în primul caz, separarea semnalului în unde / undine se face printr-o descompunere ortogonală, în cazul al doilea, separarea în clase admite o suprapunere.

Domeniul fuzzy permite generalizarea / integrarea conceptuală a mai multor abordări ale clasificării. Mă refer în special la acele abordări bazate pe noțiunea de probabilitate sau posibilitate (acesta e un concept fuzzy mai larg decât probabilitatea – vezi [Zadeh, 1978]). În secțiunea II.2.2., “Domeniul posibilului: o fuziune fuzzy a diverselor tipuri de cunoștințe”, scoatem în evidență această versatilitate a conceptului de fuzzy. Vom continua în această direcție în II.2.4., “Fuzzy logic = calculul cu cuvinte” [Zadeh, 1996], arătând că prin termenul lexical se poate îngloba în raționamentul ingineresc intuiția, flerul, calități care, desigur, sunt favorizate de o experiență îndelungată, experiență pe care o putem vedea și ca pe o statistică implicită, mai greu de formalizat (e.g., față de teoria probabilităților). Domeniul fuzzy îmi permite să vorbesc într-un cadru unitar despre două din cele trei contribuții practice (redate în capitolul IV)

### **III§ Domeniul multispectral al imaginilor color – pregătiri teoretice pentru capitolul IV.1.**

- în subcapitolul III.3., “Caracterul perceptual al culorii”, evidențiem calitatea domeniului fuzzy de a se preta procesării multi-canal – în cazul contribuțiilor noastre practice, clasificării în imagini color

### **IV§ Contribuții practice – vezi Cuprinsul de mai jos**

### **V§ Tabel sinoptic al contribuțiilor originale. Studiul viitor**

## CUPRINSUL SUBCAPITOLELOR

<b><i>IȘ Încadrarea problemei CI și descrierea ei în Domeniul Spațial și în Domeniul Probabilistic</i></b> .....	<b>6</b>
I.1 Introducere .....	6
Scopul sistemelor Computer Vision și al CI.....	6
Organizarea unui sistem Computer Vision (CV / CI).....	10
I.2. Detecția muchiilor (edge detection).....	11
I.3. Extragerea contururilor .....	11
I.4. Reprezentarea formelor.....	12
Simplificarea formelor.....	13
Funcții asociate formelor .....	14
Parametrii de formă scalari.....	14
Concluzii.....	15
I.5. Tehnici de segmentare.....	16
I.6. Clasificarea imaginilor prin potrivire ('Image Matching').....	16
Potrivirea pe baza corelării ('Template Matching and Area Correlation').....	16
I.7. Tehnici de clasificare .....	18
Clasificarea supervizată .....	19
Clasificarea nesupervizată ('clustering') .....	20
I.8. Domeniul probabilistic.....	21
<b><i>IIȘ Separabilitatea claselor</i></b> .....	<b>21</b>
II.1. Descompunere ortogonală cu Unde și Undine .....	21
O viziune abstractă asupra recunoașterii formelor, în jurul conceptului de corelare .....	22
II.2. Granularitate și gradație în Domeniul Fuzzy .....	25
Logica fuzzy și "oficializarea" unei alte paradigme de cunoaștere.....	25
Domeniul posibilului: o fuziune fuzzy a diverselor tipuri de cunoștințe .....	27
Conceptele de bază ale logicii fuzzy .....	28
"Fuzzy logic = calculul cu cuvinte" [Zadeh, 1996] .....	30
PROTOTIPUL = clasa fuzzy.....	30
<b><i>IIIȘ Domeniul multispectral al imaginilor color</i></b> .....	<b>31</b>
III.1. Introducere în colorimetrie .....	31
III.2. O trecere în revistă a spațiilor de culoare .....	31
III.3. Caracterul perceptual al culorii.....	32
Metamerismul sau granularitatea spectrului vizibil .....	32
Fuzitatea culorii și procesarea multi-canal .....	32

<b><i>IVS Contribuții practice</i></b> .....	<b>33</b>
IV.1 Transformări eficiente între spații de culoare.....	33
Implementarea conversiei între două spații de culoare pentru echipamente .....	33
Rezultatele conversiei.....	36
Concluzii.....	38
IV.2 Definirea formei spațiului unei probleme de recunoaștere nedeterminate și modelarea acestuia. Cazul 'skin detection' .....	39
Alegerea spațiului de culoare și colectarea pixelilor de piele.....	40
Reprezentarea Skin Locus-ului în vederea eficienței Skin Detection.....	41
Modelarea distribuției SL .....	42
Concluzii.....	44
IV.3 Folosirea subclasării și a marcajelor pentru reducerea fals pozitivelor. Cazul 'skin detection' .....	45
Abordare .....	45
Clasificarea skin / non skin .....	46
Discuție în urma experimentelor .....	52
Concluzii.....	53
<b><i>VȘ Sinopsis al contribuțiilor originale. Studiul viitor</i></b> .....	<b>54</b>
V.1. Vedere sinoptică a contribuțiilor noastre.....	54
V.2. Studiul viitor .....	57
<b><i>BIBLIOGRAFIE</i></b> .....	<b>59</b>

# Iș Încadrarea problemei CI și descrierea ei în Domeniul Spațial și în Domeniul Probabilistic

## I.1 Introducere

Scopul sistemelor Computer Vision și al CI

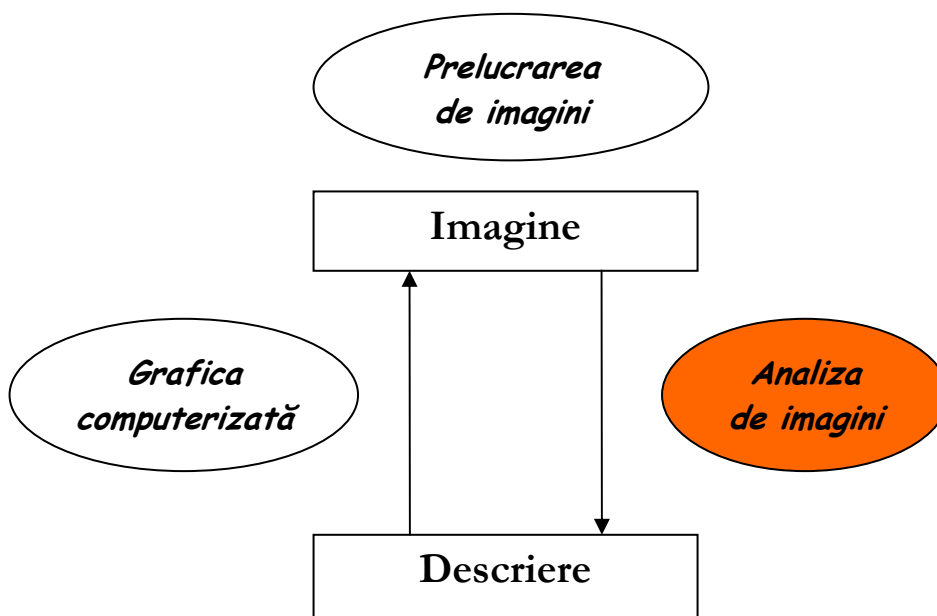


Fig. I.1.1\_3 Schemă sintetică [Pavlidis 1982] a imagisticii computerizate

Imagistica computerizată ('Computer Image'), adică utilizarea calculatorului pentru operațiuni legate de domeniul vizualului (2D sau 3D) sau folosirea imaginilor prin intermediul calculatorului, ocupă o parte dominantă a 'Computer Culture', tendința fiind în creștere, mai ales după răspândirea în public a unor echipamente digitale ca televizorul, camera video, scanner-ul. În literatura de specialitate s-au consacrat ca 3 mari domenii:

## 1. Sinteza imaginilor / Grafica computerizată ('Image Synthesis' / 'Computer Graphics')

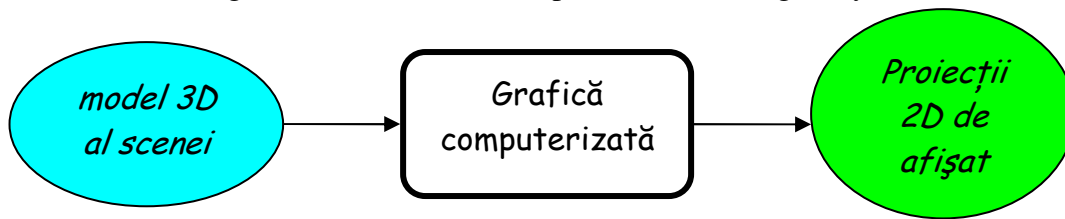


Fig. I.1.1\_1 Intrarea și ieșirea graficii computerizate

Prescurtat, se folosește termenul de Grafică. Nu e aria noastră de interes în această lucrare. Aria mare în care ne încadrăm cu lucrarea de față o numim prescurtat: Imagistică. Să avem în vedere abuzul de limbaj ca efect al prescurtării: “Imagistica” nu e totuna cu “Imagistica computerizată”. În continuare menționăm cele două subarii (vom vedea ca nu și partiții strict delimitate) ale imagisticii:

## 2. Prelucrarea imaginilor ('(Digital) Image Processing') :

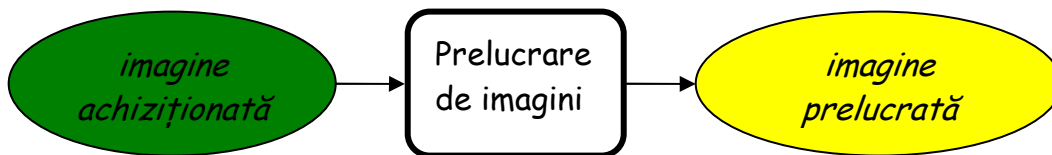


Fig. I.1.1\_2 Intrarea și ieșirea prelucrării de imagini. Imaginea e prelucrată pentru a fi adusă într-o reprezentare cât mai adecvată unei analize de imagini, analiză presupusă de un sistem cu viziune.

Aici suntem într-un domeniu preliminar al problemei CI. Acest domeniu, cu “preproblemele” restaurării și îmbunătățirii imaginilor nu e aprofundat în lucrarea noastră, ci doar trecut în revistă.

3. *Computer Vision*, care s-a definit ca "preluarea de date-imagini brute și, pe baza acestor date, producerea unei acțiuni corespunzătoare categoriilor (claselor) datelor" [wikipedia] Sistemele cu viziune ('Computer Vision') se pot clasifica ca fiind de nivel:

- scăzut: sunt sisteme de prelucrare de imagini - intrarea este o imagine digitală, iar ieșirea o imagine prelucrată prin tehnici de restaurare și îmbunătățire
- intermediar: intrarea este o imagine prelucrată, iar ieșirea o reprezentare simbolică de nivel coborât (e.g., se extrag linii drepte, contururi)
- înalt: intrarea este o reprezentare simbolică de nivel coborât, iar ieșirea o reprezentare simbolică de nivel înalt (de ex., se extrag forme de obiecte).

În general, CI se încadrează la 3., pesupunând și tehnici din 2..

CI își propune să clasifice informațiile obținute din imaginile de intrare pe baza unor cunoștințe apriorice sau pe baza unor informații statistice extrase din date.

## Limitări ale CI

Problema fundamentală în CI e de a determina automat dacă o imagine conține sau nu un obiect, o caracteristică. Aceasta problemă, care poate fi soluționată natural de un om - cortexul vizual uman e, indiscutabil, un inegalabil aparat de analiză de imagini, în special pentru extragerea informațiilor de nivel înalt -, pune de multe ori în impas un computer, mai ales în cazul general: obiecte arbitrare, în situații arbitrare. Metodele existente pot rezolva problema, în cazul cel mai bun, pentru obiecte *specifice* - e.g., obiecte geometrice simple (poliedre), caractere scrise de mână sau imprimate, vehicule, problema devenind nedeterminată și dificilă pentru fețe umane. Chiar și pentru aceste situații relativ simple se impun condiții: iluminare bine definită, mediu etc.

## Denumiri

Taxonomiile nu sunt fixe. Oficial se spune că acest domeniu a fost dezvoltat în anii 1950 în S.U.A. pornind de la robotică și inteligență artificială. ‘Computer Vision’ (CV), termen care “stă” în multe cazuri pentru CI, e acel domeniu larg, atât științific cât și tehnologic, al mașinilor care “văd”. Ca disciplină științifică, CV se preocupă de teoria necesară construcției de sisteme care să obțină informație semnificativă din imagini. Imaginile pot veni în multe forme: secvență video, proiecții de la mai multe camere, tomograf etc. CI e văzută și ca un câmp al inteligenței artificiale sau al științei calculatoarelor în general. Se înrudește cu prelucrarea de imagini, analiza de imagini, termenii suprapunându-se în multe cazuri. Se poate spune că analiza de imagini - termen sinonim cu CI - se încadrează în domeniul larg CV, că e folosită din plin în industrie, că folosește geometria digitală și procesarea de semnal. Chiar și procese interne CI, pot avea o parte CI, e.g., restaurarea de imagine: analizând inițial imaginea în termeni de structură locală (linii, muchii, regiuni) putem obține o metoda mai bună de diminuare a zgomotului.

## Probleme concrete de CI. Aplicații tipice

Să ilustrăm o aplicație software de CI care, în multe cazuri, ne vine odată cu aparatul de fotografiat achiziționat:



Fig. I.1.1\_4 Recunoașterea / detecția feței

Au devenit de uz comun aplicațiile OCR (‘Optical Character Recognition’) și, mai nou, OMR (‘Optical Markup Recognition’ – e.g., corectarea testelor grilă prin recunoașterea marcajelor)

CI, după scop general, prezintă următoarele categorii:

- Recunoaștere: unul sau câteva (clase de) obiecte “știute” se recunosc, împreună cu poziția lor în imaginea 2D/scena 3D
- Identificare: o instanță individuală a unui obiect e recunoscută - e.g., identificarea amprenteii unei persoane
- Detecție: imaginea e scanată pentru o condiție specifică - e.g., detecția unor celule anormale

Tehnicile CI se implementează intensiv în

- Robotică (roboți industriali, vehicule autonome etc.) sau în alte automatizări, prospectări:
- Inspecția vizuală a articolelor
- Recunoașterea și interpretarea obiectelor într-o scenă, urmate de acțiuni corespunzătoare - estimarea poziției/orientării : e.g un braț de robot poate fi ghidat de informații vizuale (robotul poate și citi de pe etichete);
- Evaluarea mișcării cu aspectele ‘*Egomotion*’ - determinarea mișcării 3D a camerei foto/video sau ‘*Tracking*’ - urmărirea traiectoriilor obiectelor
- Controlul proceselor și detectarea evenimentelor (supravegherea vizuală) – e.g., Interacțiunea om ↔ mașină: pornirea robinetului când trecem mâinile pe sub el
- Analiza de imagini medicale (imagistica medicală) – detecție și măsurători ale tumorilor, organelor, ‘*Computer Added Diagnosis*’
- Procesarea imaginilor achiziționate de la distanță (‘*Remote Processing*’): imagistica radar – detecția și identificarea țintei; cartografieri și prognoze meteo pe baza imaginilor satelitare, analize de trafic
- Expertiză (judiciară sau nu): potrivirea amprentelor, recunoașterea scrisului de mână, a vorbirii
- Organizarea informației vizuale (pentru baze de date de imagini) - Recuperarea imaginii după conținut (‘*Content-based image retrieval*’): găsirea tuturor imaginilor cu un conținut specific într-un set larg de imagini. Conținutul poate fi specificat în mai multe moduri, e.g., în termeni de similaritate relativ la o imagine reper sau în termeni de nivel înalt (“toate imaginile care conțin multe case”)
- Modelarea obiectelor și Reconstrucția scenelor: dându-se mai multe imagini ale unei scene se reconstruiește scena 3D; în cazul tomografiei computerizate, imaginile achiziționate sunt secțiuni interne ale corpurilor/obiectelor.

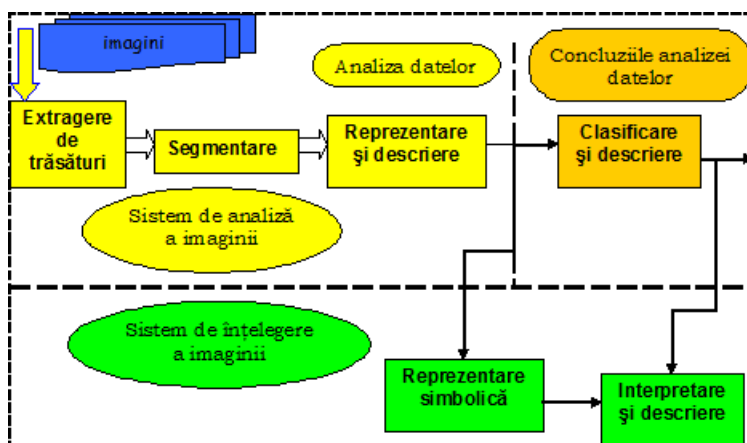
Organizarea unui sistem CV e foarte dependentă de aplicație. Anvergura problemei concrete dictează lungimea lanțului de subprobleme componente. Implementarea unui sistem CV depinde, de asemenea de posibilitatea învățării/modificării în timp real a anumitor funcționalități. Totuși, oricum s-ar prezenta datele problemei, există o succesiune de etape/ funcționalități tipice unui sistem CV. Le redăm în continuare în ordinea cronologică în lanțul procesării.

## Organizarea unui sistem Computer Vision (CV / CI)

La modul general, obiectivul oricărui sistem de CV / CI se poate defini ca: analiză de imagini spre a extrage mai întâi informații cu semnificație și, apoi, producerea de descrieri, interpretări sau înțelegeri ale imaginilor de intrare.

Presupune că avem la intrarea sistemului de CI imagini digitale fără defecte care să afecteze major procesarea lor ulterioară. Sărind etapa de achiziție de imagini (etapă mai puțin accesibilă) și etapa de preprocesare de imagini (etapă sumar detaliată în lucrare, fiind tratată până aproape de epuizare în literatura de specialitate), presupunând că avem la intrarea sistemului de CI imagini digitale fără defecte care să afecteze major procesarea lor ulterioară, următoarele etape/module definesc în principal sistemul de CI:

Fig. I.1.2\_1 Schema unui sistem de CI în etape / module



*Segmentarea* înseamnă partiționarea imaginii în zone relativ omogene. Ea presupune extragerea unor anumite trăsături în vederea acestei împărțiri, e.g., extragerea conturilor. Se poate ca problema de CI să se termine odată cu segmentarea - e cazul unor probleme concrete de detecție, selecție a unei zone de interes. Imaginea segmentată e reprezentată și descrisă într-o formă potrivită pentru a fi dată clasificatorului. Acesta trebuie să mapeze diverse regiuni (în cazul unei reprezentări interne) sau contururi (în cazul unei reprezentări externe) în clase de obiecte. Să reținem că, odată ieșiți din zona analizei datelor (cea marcată cu galben și galben-oranj), intrăm în zona concluziilor analizei datelor (cea marcată cu verde), unde pe baza unor reguli predefinite putem genera descrieri și interpretări utile calculatorului sau omului. De exemplu, în cazul unei probleme de sortare a merelor de pere, obiectele de formă mai rotundă vor fi clasificate ca mere, cele de formă mai alungită, ca pere. Sistemele de înțelegere a imaginii preiau o clasificare sau o reprezentare simbolică și determină relațiile dintre obiecte în scena 'input'. De exemplu, un asemenea sistem poate raporta: <<În scenă se văd 3 pere înconjurate de 7 mere.>> Sistemele de înțelegere nu vor fi abordate în această lucrare – ele s-au consacrat mai degrabă în domeniul inteligenței artificiale, în timp ce noi ne propunem să abordăm CI dinspre partea analizei imaginilor.

Analiza de imagini, înseamnă în principal: extragere de trăsături, segmentare și clasificare.

1. Extragerea de trăsături se referă la: trăsături spațiale, trăsături ale transformatei (e.g., Hough, Fourier), muchii (+ colțuri și contururi), trăsături de formă, texturi.

2. Segmentarea se poate baza pe: praguri ('thresholding'), acumulări ('clustering'), arbori cuaternari, detecție de contururi ('boundary detection'), potrivire de șabloane ('pattern matching'), potrivire de texturi ('texture matching').

3. Clasificarea se poate baza pe: statistici, acumulări, arbori de decizie, arbori de acoperire minimă, măsuri de similaritate.

În acest capitol trecem în revistă aceste operațiuni în domeniul spațial al imaginii pe care o reprezentăm matematic cu o matrice 2-D. Cea mai simplă și probabil cea mai folosită trăsătură spațială e *amplitudinea*. Amplitudinea se regăsește ca proprietate fizică ca atare sau ca nivel de gri/luminozitate (când nu facem alte precizări explicite, presupunem că lucrăm cu imagini în scară de gri – cu imagini color vom lucra începând cu capitolul III). Ca proprietăți fizice ce pot da o amplitudine exemplificăm: reflectivitatea, refractivitatea, valorile tristimulus care dau culoarea. În cazul nivelului de gri vorbim despre trăsăturile de histogramă.

## *1.2. Detecția muchiilor (edge detection)*

Aceasta e o problemă fundamentală în analiza de imagini, deoarece muchiile caracterizează contururile obiectelor, iar identificarea obiectelor, segmentarea și, eventual, alinierea obiectelor sunt scopurile analizei de imagini. Condiția fundamentală a oricărui filtru folosit la detecția muchiilor e ca suma elementelor / ponderilor filtrului să fie 0 (e condiția duală celei de la filtrele de netezire ("blurring"): suma să fie 1): se asigură "ieșirea" 0 pentru zone uniforme.

Ceea ce e important pentru detecția muchiilor e rata de variație a nivelelor de gri, deci derivata funcției imagine - o variație lentă nu poate corespunde unei muchii. În detecția de muchii putem folosi operatori *gradient* sau operatori *busolă (compass)*, care sunt operatori de diferențiere de ordinul 1. Putem folosi și operatori de diferențiere de ordinul 2, ca operatorul Laplace. Se mai folosesc și gradienti stocastici, care iau în considerare prezența zgomotului și se prezintă ca o alternativă mai bună față de mediere.

Având pixeli de muchie, poate urma extragerea conturului - continuăm cu un sumar al acestor tehnici.

## *1.3. Extragerea conturilor*

- *Urmărirea ('following') conturului* – are loc ordonarea pixelilor de muchie succesivi

- *Legarea muchiilor ('edge linking')* și *căutarea euristică în graful muchiilor ('heuristic graph searching')*: legarea muchiilor se face conform unor reguli prestabilite, trebuind maximizată alegerea următoarei muchii în graf

- *Programarea dinamică*: metoda de găsire a optimului global în cazul unui proces în mai multe faze; *principiul optimalității* (Bellman): calea optimă între două puncte date e de asemenea optimă între oricare două puncte de pe cale.

- ***Transformarea Hough*** – în forma elementară găsește liniile drepte, în formă generalizată poate fi folosită, de exemplu, pentru găsirea conturilor. În fapt, **această metodă e reprezentativă pentru tehnicile de vot majoritar (tehnici de care vom face uz și noi în capitolul IV), pretându-se la o abordare prin prisma probabilităților, cu frecvența relativă pe post de probabilitate.**

#### Operatorul optimal de extragere a conturilor

Conform [Vertan & Ciuc, 2007, p.129], problema detecției conturilor pe baza prăguirii unei hărți de variații, harta fiind obținută prin filtrare liniară, poate fi văzută ca o “problemă de proiectare a unui filtru liniar, al cărui răspuns să marcheze cu maxime bine definite pozițiile tranzițiilor importante din imaginea-intrare. Filtrul dorit va fi optimal în sensul de a oferi cea mai bună performanță medie de detecție a muchiilor (pentru orice poziție și magnitudine a variației ce definește muchia și orice putere a zgomotului suprapus acestuia).” Pornind de la cazul unui semnal treaptă 1D, degradat cu zgomot alb, gaussian și aditiv, Canny a sintetizat un filtru FIR (‘Finite Impulse Response’ cu suport finit al răspunsului la impuls), caracterizat de un produs detecție x localizare cât mai apropiat de valoarea optimă: filtrul DoG = derivata (de ordinul întâi) de gaussiană.

### *1.4. Reprezentarea formelor*

Forma fiind o singură componentă conexă, problema reprezentării formei, în vederea recunoașterii ei, poate fi văzută și ca problemă a reprezentării conturilor, dar și a *regiunilor*.

Reprezentări mai simple ale conturilor sunt:

- *Codurile înlănțuite ('chain codes')*: se pleacă cu un pixel de start, indicându-se în continuare direcția pixelului următor

- *Aproximarea poligonală ('fitting line segments')*: se “segmentează” curbele, astfel încât abaterea să fie sub un prag ales conform preciziei cerute

- Curbe B-spline

Dacă vorbim de regiuni, acestea se pot reprezenta ca:

1. ‘Secvențe de coduri’ (‘run length codes’)

- avantaje: se obțin la o singură baleiere a imaginii

- dezavantaje: nu dau contururi

## 2. ‘Arbori cuaternari’ (‘quad trees’)

- avantaje: compresie

- dezavantaje: nu dau măsurători ale formei ca perimetru, arie

## 3. Proiecții

- e, mai ales, cazul specific al tomografiei computerizate

Vom trece în revistă și alte posibile reprezentări ale unei forme, cu adăugări și observații la [Vertan & Ciuc, 2007, p. 157-171]. Menționăm și că, în multe aplicații ‘Computer Vision’, obiectele pot fi caracterizate ca structuri compuse, e.g., din linii și arcuri, în general vorbind, din *elemente structurale (primitive)*. Dând o *sintaxă* pentru elementele primitive, sintaxă bazată pe anumite *reguli de conectivitate*, putem obține o reprezentare unică a scenei într-o *înșirare de simboluri* corespunzătoare primitivelor. Stabilirea regulilor de conectivitate, a unei *gramatici*, vine din *înțelegerea / interpretarea scenei* (‘*image understanding*’) și poate fi dificilă, chiar imposibilă fără cunoștințe *a priori*.

Reprezentarea formelor aduce ca protagonist și prelucrarea morfologică a imaginilor. De așteptat: într-o reprezentare a formei încercăm să “surprindem” structura ei.

*Parametrii de formă* sunt definiți astfel încât să se respecte invarianța la transformările geometrice afine uzuale (translație, rotație, scalare) și să fie asigurată o toleranță la zgomot (în caz de mici variații ale formei). În general, parametrii de formă sunt scalari sau funcții, cu observația ca există reprezentări ale formei obținute prin simplificări ale structurii acesteia, e.g., chiar aproximarea poligonală..

### Simplificarea formelor

1. Anvelopa convexă: cea mai mică formă plană convexă ce include forma dată. Putem spune că anvelopa convexă e forma cea mai asemănătoare formei inițiale, în care concavitățile conturului au fost “umplute.” Anvelopa convexă a oricărei mulțimi discrete e un poligon ale cărui vârfuri sunt o parte din punctele situate pe frontiera formei. Algoritmi clasici: ‘Jarvis march’ și ‘Graham scan’.

2. Skeletonul (sau axa mediană): reprezentare prin linii, regenerativă a unei forme plane. S-a propus un model intuitiv de descriere a skeletonului, model care oferă și o sugestie de implementare a acestuia: modelul “focului de preerie”. Să ne imaginăm o suprafață de preerie acoperită omogen și izotrop cu iarbă uscată. Într-o zi fără vânt, focul izbucnește în fiecare punct de pe marginea suprafeței de preerie. Cum materialul inflamabil, iarba uscată, se găsește numai în interiorul suprafeței, frontul de flacără se propagă numai în interiorul regiunii, consumând iarba uscată. La un moment dat, flăcările venind din direcții opuse, cu aceeași viteză, se întâlnesc și focul se stinge, nemaivând material de combustie. Punctele de întâlnire ale fronturilor de flacără (și respectiv punctele de stingere a focului) se găsesc pe axa mediană a regiunii. Matematic, skeletonul se definește prin noțiunea de disc maximal și poate fi determinat prin operații morfologice.

## Funcții asociate formelor

Funcțiile asociate formelor pot fi utile ca dicriminanți de tip “linie ‘hashcode’”. Exemple clasice de:

1. *Semnătura formei*: reprezentare regenerativă dată de o funcție scalară de o variabilă, definită de distanța de la un punct de referință  $x_0$  fixat (în general centrul de greutate al formei) la fiecare punct de pe conturul formei. Variabila e unghiul la centru  $\theta$  realizat de punctul curent de pe contur cu axa orizontală de referință,  $d_{x_0}(\theta)$ , sau variabila poate fi abscisa curbilinie  $\rho =$  lungimea conturului cuprins între punctul curent de pe contur și punctul în care axa de referință intersectează conturul.

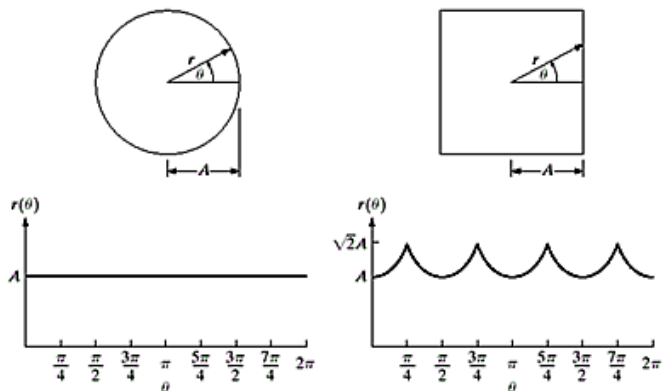


Fig.I.4.2.\_1 Exemplu de semnătură pentru discuri cu defecte la margine – defectul fiind reprezentat ca abaterea de la “linie” [Gonzalez, 2008, p. 863]

2. *Curba granulometrică*: Reprezintă o funcție scalară de o variabilă reală pozitivă, definită de măsurarea ariei formei prin “umplerea” formei printr-o operație morfologică convenabil aleasă. Variabila reală ce indexează funcția granulometrică e dimensiunea elementului structurant folosit în umplerea morfologică. Preluăm o descriere sugestivă din domeniul compoziției materialelor : “*curba granulometrică a pământului* e reprezentarea grafică a cantității de pământ (în procente din greutate) dintr-o anumită probă, funcție de diametrul particulelor care alcătuiesc proba. Se determină prin

cernerea probei succesiv prin site cu dimensiunile ochiurilor din ce în ce mai mici și cântărindu-se greutatea materialului care nu a trecut prin ochiurile sitei.” Alura curbei granulometrice ne dă indicații asupra gradului de uniformitate a materialului.

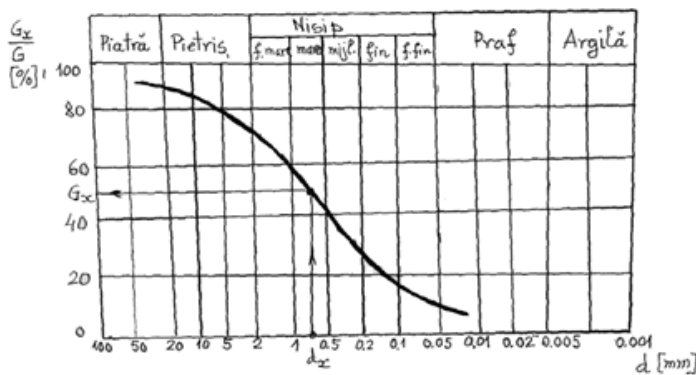


Fig. I.4.2\_2 Curbă granulometrică a pământului [crb grn]

## Parametrii de formă scalari

Aceștia pot funcționa ca dicriminanți de tip “număr ‘hashcode’” Exemple:

1. *Parametri geometrici (rapoarte de aspect)*: se bazează pe măsurarea unor atribute geometrice simple ca: aria S, perimetrul P, numărul de găuri, numărul lui Euler (diferența dintre numărul de componente

conexe și numărul de găuri), dimensiunile maxime ale formei pe direcție orizontală (Dh) sau verticală (Dv), dimensiunile cercurilor înscrise (r) și circumscrise (R) formei. Evident, atribute ca S, P, r, R nu pot fi parametri de formă invarianți la scalare, dar se pot construi rapoarte invariante la scalare, în intervalul [0,1]. E.g., raportul de compacitate (numit și factor de formă) e definit ca:  $k = \frac{4\pi S}{P^2}$ .

2. Momente statistice și invarianți: interpretând funcția caracteristică a formei ( $f(x,y):R \rightarrow \{0..L-1\}$  normalizat la [0,1], R fiind regiunea care definește suportul formei) ca pe o funcție de densitate de probabilitate 2D, putem defini *momentele statistice de ordin p+q sau p,q* pentru cele două variabile aleatoare corespunzătoare coordonatelor punctelor formei:  $m_{pq} = \int \int_R f(x,y) x^p y^q dx dy$

Teorema reprezentării cu momente afirmă că mulțimea infinită de momente  $m_{pq}$  determină unic  $f(x,y)$  și reciproc. Caracterizarea formei printr-o serie infinită de numere fiind imposibilă, în practică se folosesc pentru aproximare serii de momente până la un anumit ordin:  $p+q \leq N$ . De remarcat că, folosirea momentelor statistice definite ca mai sus pentru caracterizarea unei forme nu asigură îndeplinirea nici unuia dintre principiile de invarianță căutate; de aceea, au fost introduse momente statistice invariante. Inițial, în anii '60, acești invarianți au fost folosiți pentru recunoașterea caracterelor mari de tipar, cu rezultate modeste. Eficiența lor constă însă în modul rapid de calcul și posibilitatea de a le utiliza cu succes pentru recunoașterea formelor geometrice **convexe**.

3. Descriptori Fourier – pleacă de la ideea că un contur închis poate fi văzut ca o funcție complexă periodică, de perioadă N ce reprezintă lungimea conturului. Observații: amplitudinile sunt invariante la transformările de rotație și schimbare a punctului de start, iar faza e invariantă la scalare. DF sunt caracteristici de formă regenerative, care se folosesc pentru recunoașterea formelor, mergând cu calculul DF până la precizia dorită - precizia e dată de rangul ultimului DF calculat. DF se mai pot folosi pentru: analiza curbilor / liniilor - le putem considera contururile unor obiecte cu grosimea de un pixel, calculul ariei unei suprafețe, skeletizare și '*potrivirea conturilor*' ('*boundary matching*'). Putem folosi DF pentru a potrivi forme similare, chiar dacă au dimensiuni și orientări diferite; se poate stabili un prag de eroare între potrivirea coeficienților DF ai celor două contururi ale formelor de comparat.

### Concluzii

Reprezentarea formelor poate fi folosită la: recunoașterea formelor, potrivirea și măsurarea formelor. Măsurarea formelor poate fi văzută ca extragerea de caracteristici bazate pe geometria sau momentele formei. Pentru potrivirea și recunoașterea formelor, reprezentările cu adevărat "sigure", care garantează identificarea formei, admițând transformare inversă (de la reprezentare la formă), sunt reprezentările regenerative, cum ar fi conturul, regiunea, reprezentarea structural sintactică, skeletonul, semnătura formei, descriptorii Fourier, aproximările cu momente. Ceea ce decide metoda optimă de reprezentare folosită e impactul asupra acurateței, vitezei și costului aplicației..

## *1.5. Tehnici de segmentare*

Segmentarea e procesul care partiționează imaginea în regiuni relativ omogene. Îi putem spune și recunoaștere a regiunilor, fiind în fond o clasificare de nivel jos, care pregătește clasificarea de nivel înalt ce va avea loc la finalul întregului proces de CI. Tipuri de segmentare:

- Limitarea intensității ('amplitude thresholding / window slicing')
- Etichetarea componentelor ('component labeling')
- bazată pe: contururi ('boundary based approaches') sau regiuni ('region based approaches & clustering') sau potrivire de șabloane ('image matching').

De segmentare cu șabloane putem vorbi în cazul separării textului de grafice într-o pagină, fiecare zonă necesitând apoi tehnici specifice de prelucrare. La această categorie de segmentare includem și *segmentarea bazată pe texturi*, abordare care își poate dovedi valoarea în cazul imaginilor cu o *densitate mare de muchii*, caz în care abordarea bazată pe contur eșuează. Ca procedee de prelucrare, menționăm prelucrările morfologice, ca închiderea și deschiderea – putem intra aici în domeniul *granulometriei*, al cărui principiu general ar fi că "au efect" prelucrările morfologice cu elemente structurante de dimensiuni similare cu obiectele.

Cazul mai rafinat al acestui segmentării cu șabloane e potrivirea imaginilor, despre care vorbim în subcapitolul următor. Prezentăm separat acest caz, trecând în revistă o metodă generală de potrivire, importantă și prin felul în care pune în lumină conceptul de **corelare**, concept asupra căruia vom reveni pentru a face o sinteză din punct de vedere matematic asupra recunoașterii formelor în secțiunea II.1.5, "O viziune abstractă asupra recunoașterii formelor, în jurul conceptului de corelare".

## *1.6. Clasificarea imaginilor prin potrivire ('Image Matching')*

Detecția prezenței sau schimbării poziției unui obiect într-o scenă dată e o problemă de mare importanță în analiza de imagini. Domenii concrete de interes sunt robotica, prognozele meteo conform imaginilor obținute din sateliți, diagnosticele medicale conform imaginilor obținute prin radiologie sau tomografie computerizată etc. Detecția schimbării e utilă și în *aliniera* ('*alignment or spatial registration*') imaginilor "luate" la moment / din poziții sau cu aparate foto diferite. Cea mai simplă metodă de detecție a schimbărilor e *scăderea imaginilor*. Deși rudimentară, această metodă dă rezultate bune în cazul unor condiții controlate de formare a imaginilor, e.g., făcând diferența între imaginile medicale ale unui organ înainte și după injectarea unei substanțe la radiologie, în cadrul sistemelor de securitate bazate pe monitorizare, la segmentarea unei părți dintr-un ansamblu complex.

Potrivirea pe baza corelării ('Template Matching and Area Correlation')

În esență, “potrivirea” unui model într-o imagine constă în căutarea maximului corelației între imagine și model (șablon). Presupunem că vrem să găsim / potrivim într-o imagine  $f(x,y)$  un șablon  $w$  de dimensiune semnificativ mai mică. Definim coeficientul normalizat de corelație:

$$\gamma(x, y) = \frac{\sum_s \sum_t w(s,t) - \bar{w} [f(x+s, y+t) - \bar{f}]}{\sqrt{\left( \sum_s \sum_t [w(s,t) - \bar{w}]^2 \right) \left( \sum_s \sum_t [f(x+s, y+t) - \bar{f}]^2 \right)}}$$

unde  $\bar{w}$  e media intensităților pixelilor șablonului, iar  $\bar{f}$  se calculează ca medie a intensităților pixelilor imaginii, în dreptunghiul de suprapunere a acestuia cu  $w$ : corelăm imaginii  $f$  cu masca  $w$ . Conform inegalității lui Cauchy,  $\gamma(x, y) \in [-1, 1]$ : ne asigurăm că avem o măsură normalizată a potrivirii șablonului în imagine, măsură independentă de diferența de scală între cei doi termeni de comparat. Numărătorul ar putea fi calculat ca produs, trecând în domeniul Fourier, pe baza teoremei convoluției. Coeficientul de corelație nu poate fi calculat în totalitate cu aceasta teoremă, din cauza termenilor neliniari (ridicare la pătrat și împărțire). Conform aceleiași inegalități a lui Cauchy,  $\gamma$  atinge valoarea extremă ( $\pm 1$ ) când  $w$  centrat e o scalare a lui  $f$  centrat:  $w(s,t) - \bar{w} = k[f(x+s, y+t) - \bar{f}]$ . Corelația extremă ( $\gamma$  maxim în modul) înseamnă cea mai bună potrivire posibilă. Rezultă imediat algoritmul de potrivire: parcurgerea incrementală a imaginii pentru corelația ei cu șablonul și reținerea poziției maximului (sau maximelor) ca poziție centrală de potrivire a șablonului în imagine.

Conform [Bulea 1, 2003, p. 165]: “Volumul mare de calcule necesare pentru obținerea funcției de intercorelație (coeficientului normalizat de corelație) poate fi diminuat aplicând unele strategii.”: reducerea rezoluției imaginilor sau simpla selecție a unui subset de pixeli pentru o localizare grosieră a maximelor → restrângerea semnificativă a dimensiunilor ferestrei de corelație.

Am arătat că e posibilă normalizarea corelației pentru schimbarea intensității imaginilor potrivite. Normalizarea față de redimensionarea imaginilor și față de rotația lor sunt probleme mai complicate. Redimensionarea implică cunoașterea dimensiunii comune la care să fie aduse imaginile. Dacă luăm exemplul unei aplicații controlate la distanță, această dimensiune se poate deduce știind geometria senzorilor și poziția lor față de obiectele “fotografiate”. În mod asemănător, s-ar putea rezolva și normalizarea față de rotație, care presupune cunoașterea unghiului după care trebuie rotite imaginile.

Dar, în general, potrivirea în spațiul parametrilor  $\left( \frac{x-s}{\gamma_1}, \frac{y-t}{\gamma_2}, \theta \right)$  unde numitorii și al treilea

parametru corespund normalizării față de redimensionare și, respectiv, rotire, devine imposibil dacă nu avem estimări rezonabile și dacă detecția trăsăturilor imagistice nu ne ajută îndeajuns. Generalizarea metodei de potrivire bazată pe calculul corelației constă în înlocuirea funcțiilor de imagine cu trăsături extrase din imagini, cum ar fi momentele invariante sau coeficienții unor transformări. Astfel, putem încerca potrivirea pe baza unor puncte cheie (‘control points’), detectate automat sau alese manual.

Căutarea logaritmică

Dacă nu există informații *a priori* despre variația abaterii potrivirii șablonului în diferitele locații ale imaginii, suntem nevoiți să căutăm în toată imaginea, pixel cu pixel. Dacă știm, de exemplu, că abaterea crește monoton - definim o funcție de abatere:  $A(i, j) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N p[f(i-m, j-n) - w(m, n)]$ ,

$p(x)$  fiind o funcție crescătoare, pozitivă (în general  $|x|$  sau  $x^2$ ) - pe măsură ce ne îndepărtăm de locația potrivirii, putem reduce costul căutării de la complexitatea  $O(n^2)$  de la  $O(\log n)$ , astfel

Inițial) căutarea în cinci locații = centrul imaginii + mijloacele segmentelor definite de centru și mijloacele laturilor imaginii;

Apoi) căutare în cinci locații în jurul locației cu cea mai mică abatere și tot așa până putem ajunge la o căutare secvențială într-o zonă de 3x3 pixeli, dacă nu găsim potrivire pe parcurs.

## 1.7. Tehnici de clasificare

CI își propune să obțină o descriere simbolică de nivel înalt a conținutului imaginii. Un sistem complet de recunoaștere a formelor constă într-un senzor care adună datele brute, un mecanism de extragere a caracteristicilor care calculează numeric sau simbolic informația din aceste date și o schemă de clasificare sau descriere a datelor pe baza caracteristicilor extrase. Datele de clasificat sunt de obicei măsurători și / sau observații, ce se pot asocia cu puncte dintr-un spațiu multidimensional definit adecvat. O varietate largă de algoritmi poate fi aplicată pentru recunoașterea de forme, de la *clasificarea prin potrivire* la mult mai complexe *rețele neurale*.

În general, schema de clasificare folosește două abordări: *statistică* (sau *decizional teoretică*) – bazată pe caracteristicile statistice ale datelor – și *sintactică* (sau *structurală*) – bazată pe relațiile structurale între *caracteristici* ('features'). Prima categorie lucrează cu *descriptori cantitativi*, e.g., lungime, arie, textură. A doua categorie lucrează cu *descriptori calitativi (relații)*, e.g., șiruri pentru conectivități simple între primitive (o scară se poate reprezenta ca un șir (ab)+ unde a semnifică ,\_' și b semnifică ,l'). Descriptorii calitativi pot însemna gramatici - o producție poate însemna adunarea / multiplicarea de vectori sau de structuri compuse (rezultate în urma adunării / multiplicării de vectori).

Tema centrală a clasificării e *învățarea delimitării claselor de forme* ('pattern classes'). Clasa e o mulțime de forme cu proprietăți comune. Înțelegem prin *prototipul unei clase* o formă de referință a unei clase, o reprezentare a clasei: se calculează pe baza unor (forme) mostre (forme a căror apartenență la clasă e cunoscută). e.g., poate fi centrul de masă pentru reprezentarea formelor prin puncte în spațiul N-dimensional al caracteristicilor / trăsăturilor formelor. Clasificarea poate fi fără învățare (*nesupervizată sau 'clustering'*): nu se folosește un set de forme mostră sau prototipuri (*setul de învățare*) pentru antrenamentul clasificatorului, ci se încearcă identificarea de grupări.

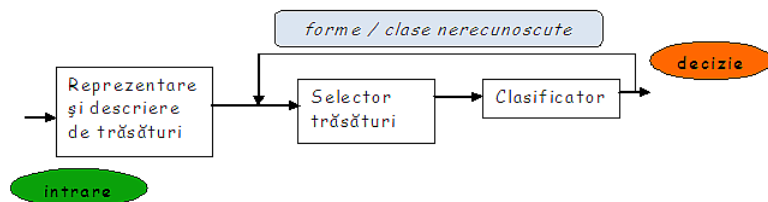


Fig. I.7\_2 Schema unui sistem de recunoaștere a formelor cu învățare. Săgeata înapoi poate însemna: formele nerecunoscute se încearcă a fi învățate sau se pot înlătura din setul de învățare forme

care nu au avut impact semnificativ asupra învățării și se pot adăuga forme de intrare noi.

### Clasificarea supervizată

- *statistică*: o abordare probabilistică a CI – clasele sunt văzute ca variabile aleatoare N-dimensionale, care se bazează pe modele de distribuție parametrică sau neparametrică, e.g., clasificarea de risc minimizat tip Bayes (vezi secțiunea I.8.3., “Clasificarea statistică”);

- *fără distribuții*: euristică ce nu necesită cunoștințe *a priori* despre probabilitățile claselor

#### Clasificarea fără distribuții

Fie  $k$  clase de obiecte (‘pattern classes’).  $X$  fiind vectorul de trăsături al formei de intrare, se definește funcția discriminant pentru clasa  $k$ :  $d_k(X) > d_i(X), \forall k \neq i \Leftrightarrow X \in C_k$ . Funcția separator între clasele  $i$  și  $j$ :  $s_{ij}(X) = [d_i(X) - d_j(X) = 0]$ . Pentru o problemă cu  $K$  clase trebuie  $K-1$  funcții separator: trebuie divizat spațiul  $N$ -dimensional al caracteristicilor în  $K$  regiuni diferite. Aceste regiuni sunt hiperplane dacă funcția discriminant / separator e liniară:  $d_k(X) = A_k^T X + b_k$ ,  $A_k$  fiind un vector (coloană, de dimensiune  $N \times 1$ ), iar  $b_k$  un scalar. În acest caz, vorbim de *clase liniar separabile*.

Dacă algoritmi liniari, ca în cazul algoritmului perceptronului, sunt cel mai puțin costisitori și mai populari, ei nu sunt “acoperitori” – există multe situații când clasele nu sunt liniar separabile:

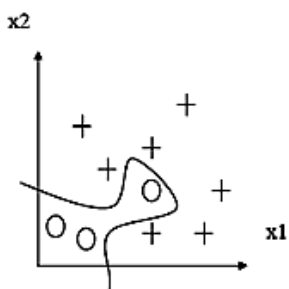


Fig. I.7.1\_5 Clasele neseperabile liniar

Rezolvarea problemei de clasificare când spațiul intrărilor nu e liniar separabil are mai multe variante:

- clasificator liniar, acceptând unele erori
- clasificator neliniar, e.g. *rețea backpropagation*
- clasificator liniar, într-un spațiu de dimensiune mai mare în care au fost transferate (transformate) datele de instruire (de intrare).

O metodă de clasificare neliniară foarte eficientă e *clasificarea cu arbore de decizie* ('*decision tree classification*'). Avantaje importante: 1. algoritmul converge chiar dacă spațiul trăsăturilor nu e liniar separabil și 2. nu trebuie testată fiecare clasă --avantaj mare când avem multe clase.

## Rețele neurale

Clasificatorul de tip rețea neurală nu face estimări inițiale ale funcțiilor de decizie, obținându-le în urma unui proces laborios, folosind în paralel mai multe *elemente nonliniare de procesare: neuroni*.

S-a dovedit matematic că aceste mașini, numite și perceptroni, antrenate cu seturi liniar separabile, adică separabile printr-un hiperplan, converg către soluție după un număr finit de pași. Valabilitatea teoremei învățării perceptronilor, a lui Rosenblatt (1962), demonstrează că perceptronii pot învăța orice poate fi reprezentat. A urmat o dezamăgire în urma nepotrivirii perceptronului cu majoritatea problemelor practice de clasificare. În 1984 Simon a publicat o lucrare despre perceptron care se intitula *Nașterea și moartea unui mit*. Dar în 1986, Rumelhart, Hinton și Williams au "resuscitat" perceptronul cu mai multe straturi ('multilayer'): deși algoritmul de antrenare pentru acest tip de rețea neurală nu converge, ca în cazul perceptronului cu un singur strat, o metodă generală, numită '*generalized delta rule for learning by backpropagation*', a fost folosită cu succes în multe probleme practice. Totuși, și arătăm aceasta în secțiunea I.7.1.2., "Rețele neurale", metodele de clasificare cu rețele neurale sunt relativ limitate, în ciuda faptului că încearcă să imite rețele neurale prezente la om. Practica a dovedit că succesul major al inteligenței artificiale se leagă de 'soft computing', de teoria fuzzy - vezi subcapitolul II.2, aceasta fiind o "imitare" mai realistă a gândirii umane.

## Clasificarea nesupervizată ('clustering')

Această abordare încearcă să identifice clustere = grupări dense de forme în spațiul trăsăturilor. Gruparea formelor în clase pleacă de la definirea în spațiul N-dimensional al formelor a unei măsuri de similaritate, e.g., produsul scalar, distanța euclidiană. Metode tipice (generale) de clasificare:

1. Metoda lanț ('chain'): Prima formă mostră e desemnată ca reprezentând cluster1. Măsura de similaritate a următoarei forme se va calcula față de acest reprezentant. Dacă această distanță e sub un anumit prag, forma se include în cluster, altfel ea va reprezenta al doilea cluster. Procesul continuă până când au fost epuizate toate formele. Avantaj: clasificare într-un singur pas; dezavantaj: trebuie știut a priori pragul ales pentru a clasifica toate formele.

2. Metoda iterativă ('c-means'): presupunem că știm numărul K de clase / clustere / partiții.

Partiționarea se face a.î. împrăștierea medie sau varianța în partiție sunt minimizezate. Fie  $\mu_k(n)$  centru cluster-ului k la a n-a iterație. Inițial,  $\mu_k(0)$  primește valori arbitrare. La a n-a iterație forma  $X_i$  e asimilată în clusterul de al cărui centru e cea mai apropiată – în sensul maximizării măsurii de similaritate, e.g., a minimizării distanței euclidiene  $d$  : 
$$x_i \in C_k \Leftrightarrow d(x_i, \mu_k(n)) = \min_{j=1..K} [d(x_i, \mu_{kj}(n))]$$

. Se recalculează centrul clusterului care a primit  $x_i$  alegându-se punctul (forma) care minimizează suma distanțelor de la el la celelate puncte / forme din cluster. Procedura se repetă până când clusteretele

rămân neschimbate. Avantaje: simplitate, viteză, se minimizează varianța intracluster. Dezavantaje: rezultatele depind de partiționarea inițială, NU se garantează că se minimizează varianța globală. Despre această tehnică vom mai vorbi la secțiunea II.2.3.1., “Matematica fuzzy”.

## *I.8. Domeniul probabilistic*

Aici vorbim de Statistici de ordinul 1 (momente, covarianță și corelație, transformarea Hotelling), de Statistici de ordinul 2 (matricea de coocurență) și de clasificarea statistică, în special de clasificatori statistici optimi (subsecțiunea I.8.3.1.) bayesieni. Încheiem acest capitol remarcând că domeniul probabilistic, deși poate “veni” cu informații în plus față de cel spațial, nu încurajează folosirea intuițiilor, a adaptărilor pentru o rezolvare cât mai adecvată a unor probleme complexe de clasificare.

# **IIȘ Separabilitatea claselor**

## *II.1. Descompunere ortogonală cu Unde și Undine*

Putem vedea descompunerea ortogonală a semnalelor în Domeniul Frecvență ca pe o clasificare conform asemănării / compatibilității semnalului cu fiecare din undele în care descompunem semnalul – cu cât coeficientul din descompunerea Fourier este mai mare, cu atât semnalul este mai compatibil cu “clasa” undei corespunzătoare. Aceasta e prima din observațiile, credem interesante, pe care le facem în acest subcapitol. Descompunerea ortogonală în unde poate fi văzută, abstractizând, ca o partiționare / clasificare fuzzy a semnalului printr-o metodă de ‘clustering’. A doua observație a acestui subcapitol e legătura intrinsecă dintre corelație și ortogonalitate: acestea sunt antonime perfecte – vezi [Georgescu 2, 2012]. Dacă vorbim de clasificare, vorbim de (gradul de) corelație și, deci, (de gradul) de descompunere ortogonală. A treia observație se referă la semnificația principiului incertitudinii (al lui Heisenberg) privind invers proporționalitatea rezoluției din domeniul Timp de cea din domeniul Frecvență; vom menționa teorema eșantionării și ‘aliasing’-ul ca efect nedorit ce afectează o partiționare în unde a semnalului astfel încât să îl putem reconstrui cu acuratețe.

Compromisul inteligent între acuratețea în domeniul temporal (sau spațial) și acuratețea în domeniul frecvență e Transformarea Undină (Wavelet): TW. Sau, cu alte cuvinte, analiza multirezoluție.

Principiul nedeterminării a fost ilustrat grafic în [Gonzalez & Woods, 2008, p. 501], atribuind fiecărei funcții din baza de reprezentare o cutie Heisenberg ('Heisenberg box') în planul timp-frecvență:

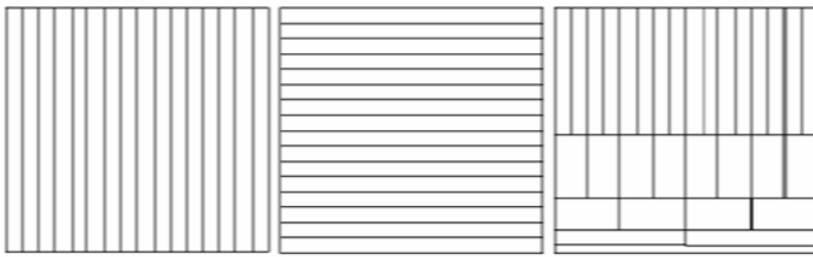


Fig.II.1.3\_11 "Acoperirea" prin trei tipuri de funcții de bază: impulsuri (domeniu temporal), sinusoidale (domeniu frecvență) și TW (aproximare / scalare + detalii în domeniu frecvență-timp)

Aria fiecărei cutii Heisenberg e aceeași. TW e o bună alegere în cele mai multe situații practice: semnale nestaționare, adică cu frecvențe variind în timp. În plus, e o reprezentare mai flexibilă, putând "jongla" în interiorul principiului nedeterminării. Putem reprezenta conceptual figura de mai sus

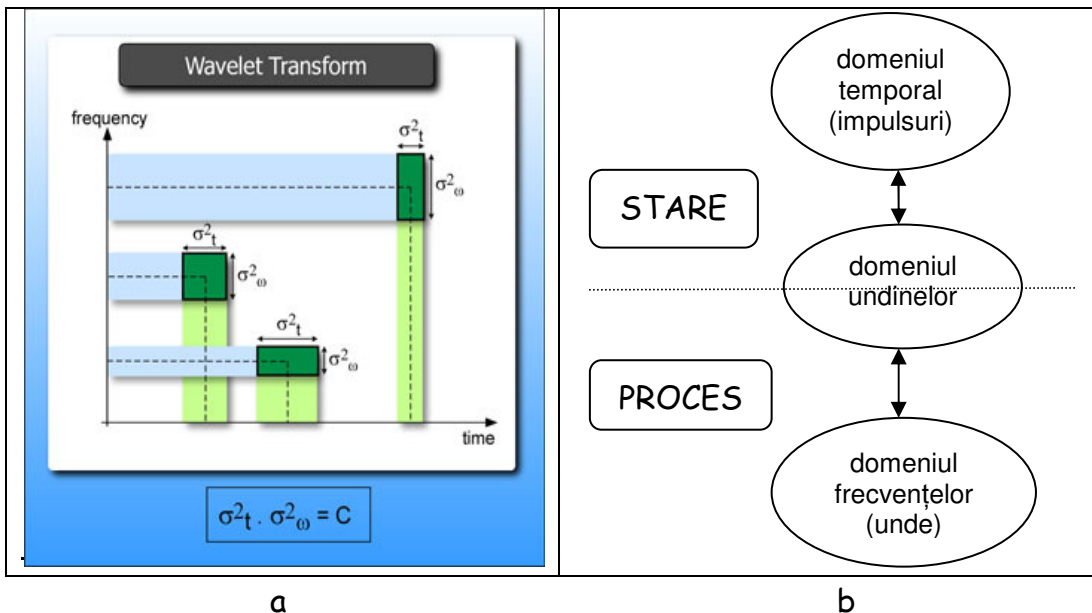


Fig.II.1.3\_12 Posibilitatea de a "jongla" în interiorul principiului nedeterminării (am folosit termenii din engleză) [wv hsn] și o reprezentare conceptuală originală a relațiilor dintre domeniile: temporal, al undinelor și al undelor.

Remarcăm că analiza multirezoluție e un puternic concept general, care se pretează la multe adaptări practice și adaptări în ideea optimizării atât a acurateței, cât și a vitezei de execuție. **Am derivat întregul arsenal de analiză multirezoluție, pornind de la conceptul central de ortogonalitate.** Remarcăm că unele tipuri de undine nu mai operează în domeniul analizei în frecvență (domeniul Fourier), pentru a oferi o adaptabilitate și viteză maximă. Înainte de a intra în domeniul fuzzy, pentru a introduce acest domeniu, vom face o sinteză.

O viziune abstractă asupra recunoașterii formelor, în jurul conceptului de corelare

Ne propunem să captăm într-o viziune unitară, într-o singură imagine, esența (matematică) a problematicii recunoașterii formelor ('pattern recognition'), în principalele domenii de reprezentare a semnalelor: domeniul spațial (sau temporal, sau de altă natură), domeniul frecvență (frecvențele pot fi spațiale, sau temporale, sau de altă natură) și domeniul probabilităților. Se spune că oamenii dobândesc informații de cea mai înaltă calitate din imagini și concepte. Iată o **schiță originală care grupează cele trei domenii în jurul conceptului de corelație**:

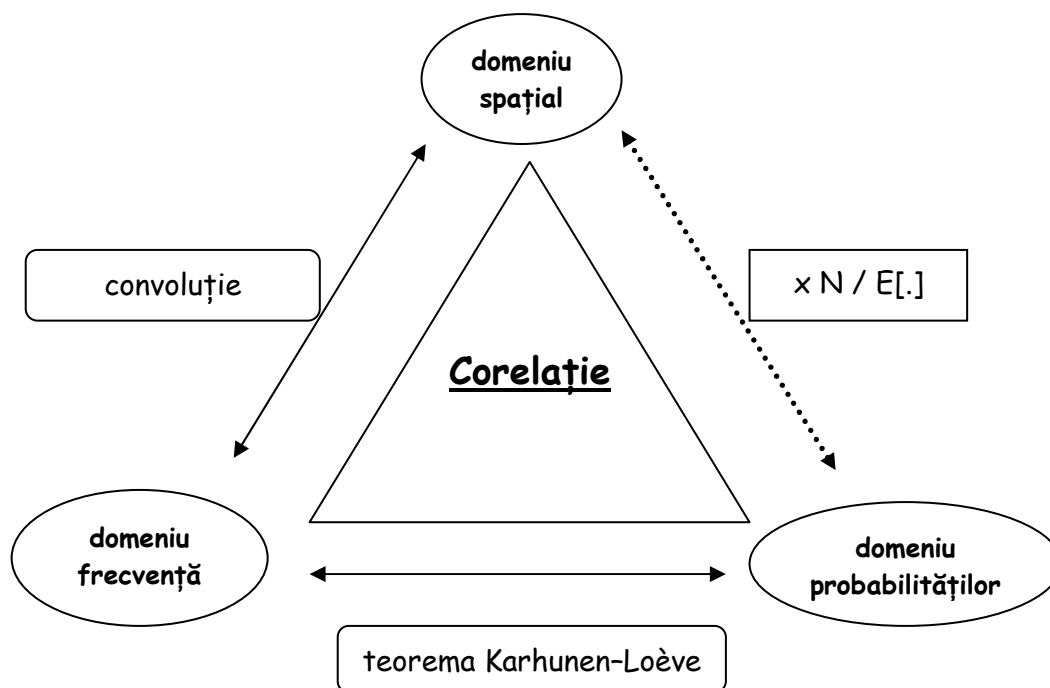


Fig. II.1.5\_1 Imagine de ansamblu a relațiilor fundamentale între domeniile: spațial / temporal, frecvență și al probabilităților

Relații fundamentale între cele trei domenii:

- Teorema de convoluție (sau de corelație) stabilește o relație între domeniul spațial și domeniul frecvență.
- Prin Teorema Karhunen–Loève (sau KLT) un process stochastic poate fi descompus ortogonal, similar descompunerii Fourier. Importanța acestei teoreme e că descompunerea e optimă în ce privește minimizarea erorii pătratice medii. Sau, cu alte cuvinte, în cazul acestei transformări în domeniul frecvență, compactarea energiei e maximă → putem reconstrui cu o bună aproximație semnalul inițial luând în considerare numai primii coeficienți ai transformatei.
- Relația între domeniul spațial și domeniul probabilistic e asemenea celei dintre ceea ce vedem la un moment dat pe o scenă și istoria evenimentelor petrecute pe acea scenă. Dacă 'x N' înseamnă acumularea de «imagini» spațiale care ar putea genera o statistică suficientă, E[.] înseamnă așteptarea

(‘expectation’) sau media, operator care estimează cea mai probabilă realizare conform cu datele probabilistice.

În esență, problema recunoașterii formelor (‘pattern recognition’) sau clasificarea bazată pe forme / regiuni / șabloane e o problemă de corelare. O observație generală, ce se degajă din rezultatele experimentale ale unei largi categorii de cercetători, e aceea că performanța unei metode vine odată cu adaptabilitatea metodei la spațiul concret al problemei. Vezi cercetarea adaptivă (‘adaptive research’) din sectorul clinic - conform [Rosenberg, 2010],[Kruschke, 2011] ș.a. - sau planificarea cultivării terenurilor (‘land use planning’) [N. van Lier, 2002]. În capitolul IV vom numi aceasta “principiul specializării și al calibrării”. KLT e un prototip din acest punct de vedere: interpretând geometric o transformare unitară ca rotație a axelor de coordonate, rotație caracterizată de o matrice de rotație. Putem vedea KLT ca transformarea unitară care reprezintă forma de intrare / semnalul în sistemul de coordonate al vectorilor săi proprii – lucrând în acest sistem poate fi foarte folositor pentru recunoașterea formelor: putem aproxima optimal.

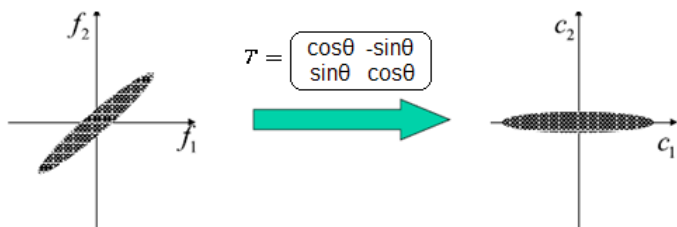


Fig. II.1.5\_2 KLT ca transformare în sistemul de coordonate al vectorilor proprii ai semnalului de intrare: de la variabile aleatoare puternic corelate (energii egale) la variabile aleatoare decorelate, cu majoritatea energiei în primele coeficienți.

În marea lor majoritate, problemele de clasificare contemporane sunt probleme complexe, de clasificare neliniară, cu granițe imprecise între clase (clasele se pot suprapune). Abordarea lor riguroasă se face în cadrul teoriei fuzzy. La Fig. II.1.5\_1 am mai putea adăuga punctat:

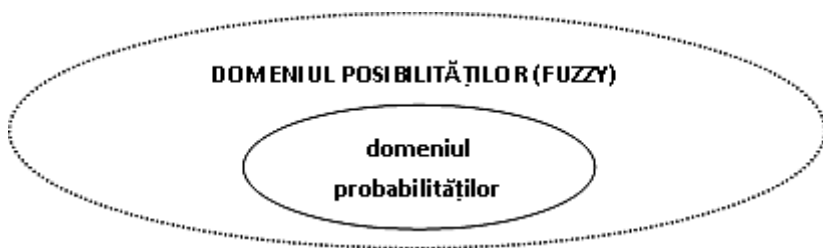


Fig. II.1.5\_2 Relația între domeniul probabilităților și cel al posibilităților (fuzzy) vezi (mai jos) subsecțiunea “II.2.2.1. Teoria fuzzy și teoria posibilităților”

La finalul acestei secțiuni putem da câteva grupuri de concepte sinonime sau, cel puțin, “înrudite”, din diferite puncte de vedere:

\* algebra lineară: produs intern peste un spațiu vectorial, nucleu simetric și separabil, bază ortogonală, proiecție și decompoziție, factorizare, valori și vectori proprii, transformări unitare;

\* principiul nedeterminării al lui Heisenberg: stare-proces, poziție (sau timp)-frecvență, principiul incompatibilității (vezi mai jos subsecțiunea II.2.1.2., “Logica fuzzy, utilitatea și legătura ei cu problema clasificării”) în teoria fuzzy;

\* teoria probabilităților: varianță, covarianță, corelație (=covarianță normalizată),

\* procesarea și analiza imaginilor: corelație spațială (= convoluție “în oglindă”), produs scalar în domeniul frecvență, potrivirea șabloanelor (‘template matching’), ‘pattern recognition’, analiza componentelor principale

În continuare ne vom ocupa cu precădere de domeniul fuzzy. Sumar spus, dacă domeniul probabilistic exprimă o luare în calcul a statisticilor, domeniul «fuzzy» exprimă o luare în calcul a intuițiilor.

## II.2. Granularitate și gradație în Domeniul Fuzzy

### Logica fuzzy și “oficializarea” unei alte paradigme de cunoaștere

L.A. Zadeh, inițiatorul logicii fuzzy, spune [fz lz]: “încă din anii ’40, când singurele calculatoare existente erau cele cu relee, se iveau titluri răsunătoare ca: <<Electric Brain Capable of Translating Foreign Languages is Being Built>>. Nici în zilele noastre, calculatoarele nu s-au apropiat măcar de nivelul general de inteligență umană”. Zadeh crede că această evoluție a fost condiționată în principal de evoluția hardware și software a calculatoarelor și senzorilor, paradigma computațională prin care s-a avansat fiind aceea de ‘soft computing’ - calcul care se poate adapta domeniului impreciziei problemei. În [Balaș & Balaș, 2007] se asociază apariția Inteligenței Artificiale (‘AI’) cu apariția Electronicii, dar și a Postmodernismului. Astfel, Electronica ar fi prima tehnologie care permite controlul energiei și informației. Se mai spune că, deși “moderniștii” contestă abordarea “fuzzy” a “postmoderniștilor”, singurul avans real în AI a avut loc în domeniul “*soft computing, i.e. clustering fuzzy logic, neural networks, genetic algorithms and evolutionary computing*”; dacă modernii puneau accentul pe materie și energie, postmodernii pun accentul pe formă și model. De unde și **euristica sistemelor expert de tip fuzzy**: “în loc să rezolvăm ecuații diferențiale, integrăm numeric prin simulare pe calculator” (e.g. metoda Monte Carlo). Desigur, logica are avantaje, dar și dezavantaje. De multe ori avem și vrem să valorificăm informații imprecise, vagi, imposibil de exprimat cu exactitate, e.g., ce se pot cere imaginiile cu persoanele de rasă caucaziană (ne aflăm în domeniul sistemelor expert de tip fuzzy). Să reținem că **logica fuzzy, deși utilă în rezolvarea multor probleme de clasificare complexe, cu date imprecise, trebuie aplicată cu precauții** - vezi paradoxul ‘sorites’.

În 1973, Zadeh a enunțat un adevăr pe cât de simplu, pe atât de general și fertil: “Pe măsură ce complexitatea unui sistem crește, capacitatea noastră de a face afirmații precise și semnificative despre comportarea sistemului respectiv se diminuează până e atins un prag, dincolo de care precizia și semnificația (sau relevanța) devin caracteristici mutual exclusive” (**Principiul Incompatibilității**). Principii înrudite: posterior, ‘**pull-back**’ [pb VN] și, anterior, cel al nedeterminării, al lui Heisenberg.

Când a introdus pentru prima dată termenul de ‘fuzzy’, în 1965, profesorul Zadeh l-a folosit pentru a descrie mulțimile de elemente multivalente. Mai târziu, el va spune: “Articolul meu din 1965 despre seturile Fuzzy reflecta concluzia mea că era nevoie de o teorie a impreciziei și incertitudinii care să nu

se bazeze pe teoria probabilităților”). În lucrarea sa, [Zadeh, 1965], pleacă de la o logică cu mai multe valori, concept uzitat în anii '20 în legătură cu Principiul Nedeterminării formulat de Heisenberg în mecanica cuantică. Logica cu mai multe valori a avut un pionier în Jan Łukasiewicz, care a formulat logica trivalentă - logica de la care a plecat și Zadeh. În ce privește alegerea cuvântului ‘fuzzy’ - în engleză: vag, cețos, rufos – Zadeh spune: “Motivul pentru care am ales cuvântul fuzzy e că ceea ce am avut în minte sunt **clasele care nu au granițe precis delimitate**”. Conceptul de mulțime fuzzy, a apărut ca o urmare firească a necesității de a modela un sistem imposibil de definit exact (‘ill-defined’), adică cu ajutorul unor instrumente matematice precise. Imperfecția definiției unor astfel de sisteme – majoritare în lumea reală – se poate vedea sub forma falsității sau impreciziei datelor. În lucrarea noastră avem în vedere mai ales imprecizia datelor.

Rezumăm câteva situații care se pretează la folosirea logicii fuzzy:

- logica fuzzy devine necesară în situațiile când **percepția umană**, judecata umană și luarea de decizii sunt implicate inextricabil.
- întrebare: cum inserăm **intuiția** în rezolvarea unor probleme? Teoria fuzzy își propune să răspundă și la această întrebare.
- modelarea prin **expresii lingvistice** (prin natura lor, vagi, imprecise), dată fiind creșterea interesului pentru modelarea prin cuvinte în diferite arii științifice (în probleme de decizie, științe sociale, teoria controlului etc.). Acesta e un atu foarte important al logicii fuzzy.
- reducerea costurilor și adaptabilitate. Precizia ridicată nu implică doar costuri mari, ci și maleabilitate scăzută în rezolvarea multor probleme.
- logica fuzzy poate funcționa integră / fuziona informații imprecise. Cunoașterea fuzzy poate fi folosită în combinație și cu cea probabilistică, e.g., s-a proiectat luarea deciziilor în sistem Bayesian.

Deschiderea de orizont dată de conceptul de set fuzzy, a fost remarcată de Moisil, care, în prefața de la [Moisil, 1975] (una din primele apărute în România, despre logica fuzzy), spunea: “această idee putea fi înțeleasă ca extensia unui predicat în logica cu o infinitate de valori[...] **identitatea, în logica cu mai multe valori, nu e decât o clasificare**. Or, aceasta ne-a condus la înțelegerea dificultăților întâlnite de aceia care ar fi vrut să întrebuițeze matematica, dar care aveau de a face cu concepte prost definite: *fuzzy*, nuanțat, precum și dificultățile filosofice de care se loveau cei care întrebuițau statistica în studiul fenomenelor necantitative.” Să remarcăm distincția între statistica fenomenelor cantitative și logica fuzzy, care se pretează fenomenelor necantitative, distincția între “*randomness*” și “*fuzziness*.”

Acceptarea principiului incompatibilității s-a tradus în realismul abordării unor probleme nedeterminate - majoritatea acestor probleme fiind de clasificare. Zadeh a definit teoria fuzzy ca “precizie a impreciziei (n.n.: precizie maximală în condiții de imprecizie)”. S-a vorbit de o schimbare oficială de paradigmă epistemologică, în sensul acceptării datelor imprecise și / sau de natură subiectivă, date, în fapt, inerente multor probleme reale, sisteme complexe etc., și care, până mai înainte, nu puteau intra în corpusul “riguros” al științei. În câteva decenii, lumea științifică, chiar și cea a științelor “exacte”, a devenit conștientă de faptul că, practic, întotdeauna operăm cu informații imprecise, cu aproximații într-un anumit GRAD, cu informații subiective (conform perspectivei din care privim). În logica fuzzy, în loc de valori de adevăr, există “*grade de adevăr*”. Folosind conceptul creat de Virgil Negoită, am putea spune că atitudinea omului de știință a “suferit” un proces de “pullback”, de desituare, sau, mai precis, de resituare față de imperfecția datelor.

## Domeniul posibilului: o fuziune fuzzy a diverselor tipuri de cunoștințe

### Teoria fuzzy și teoria posibilităților

Spre deosebire de teoria probabilităților, teoria fuzzy își propune captarea informațiilor mai degrabă calitative, decât cantitative. Așa cum am menționat, conceptul de fuzzy are o legătură privilegiată mai degrabă cu domeniul plauzibilului / posibilului decât cu cel al probabilului, probabilul fiind văzut ca o restrângere / probă a posibilului. Ceea ce a fost evidențiat și în [Zadeh, 1978]. *Teoria posibilităților*, pe care o introduce Zadeh, e văzută ca o generalizare a teoriei probabilității, diferențele majore față de aceasta fiind că:

1.  $p(A \cup B) = \max[p(A), p(B)]$  față de  $P(A \cup B) = P(A) + P(B)$ ,

pentru A și B disjuncte (ceea ce nu e o condiție necesară pentru prima egalitate),  $p(A)$  fiind posibilitatea asociată evenimentului A.

2. Posibilitatea unui eveniment nu condiționează posibilitatea complementului său:

$p(\bar{A}) + p(A) \geq 1$ . Aceasta e o relație care, în domeniul clasificării se traduce într-o **eventuală suprapunere a claselor**. Conceptul de posibilitate îl induce pe cel de *necesitate*:  $n(A) \stackrel{\text{def}}{=} 1 - p(\bar{A})$ , inegalitatea de mai sus devenind  $n(A) \leq p(A)$

Și în [Kosko, 1991] a arătat că probabilitatea e un caz special al posibilității. El a folosit o formulare ușor diferită pentru seturile fuzzy, în care funcția caracteristică a unui set fuzzy era înlocuită cu gradul în care un set poate fi considerat subset al altui set – putem formula imediat acest lucru în termeni de clasificare.

### Teoria fuzzy și bazele de cunoștințe (sistemele expert) cu informații incerte / “granulate”

Plecând de la regula bayesiană, *regula de combinare Dempster-Shafer* schimbă presupunerile în lumina unor noi evidențe. Această teorie diferențiază între *credibilitate*, care corespunde numărului de evidențe în favoarea unei propoziții, care ar corespunde probabilității, și *plauzibilitate*, care ar corespunde posibilității. Analog definiției posibilității și necesității din teoria posibilității, se definesc plauzibilitatea și credibilitatea. În [Zadeh, 1986] se “luminează” simplu și clar teoria Dempster-Shafer (TDS), din perspectiva bazelor de date relaționale, ca o inferență asupra relațiilor de ordinul 2, adică a relațiilor ale căror intrări sunt relații de ordinul 1. În același articol, Zadeh, face o analogie prin care **necesitatea și posibilitatea corespund operațiilor morfologice de eroziune și respectiv dilatare, cu același tip de dualitate**.

Atuurile majore ale logicii fuzzy sunt date de capacitatea de a opera într-un cadru unitar, cu mai multe tipuri de informație: probabilistică / posibilistică, gradată și granulată.

Gradația și granularitatea sunt două concepte înrudite, corespunzând modului în care se generalizează noțiunea de egalitate în cazul mulțimilor fuzzy: două obiecte sunt considerate similare, într-un anumit grad, dacă aparțin aceleiași clase (granule). În logica fuzzy orice element / propoziție poate avea un grad de apartenență / compatibilitate / credibilitate la o clasă / concept / variabilă lingvistică sau

asertiune. Variabila fuzzy are un domeniu granulat - prin granulă înțelegem o grupare de valori în interiorul căreia nu putem discerne aceste valori: gruparea se face după criterii de indiscernabilitate, similaritate sau proximitate.

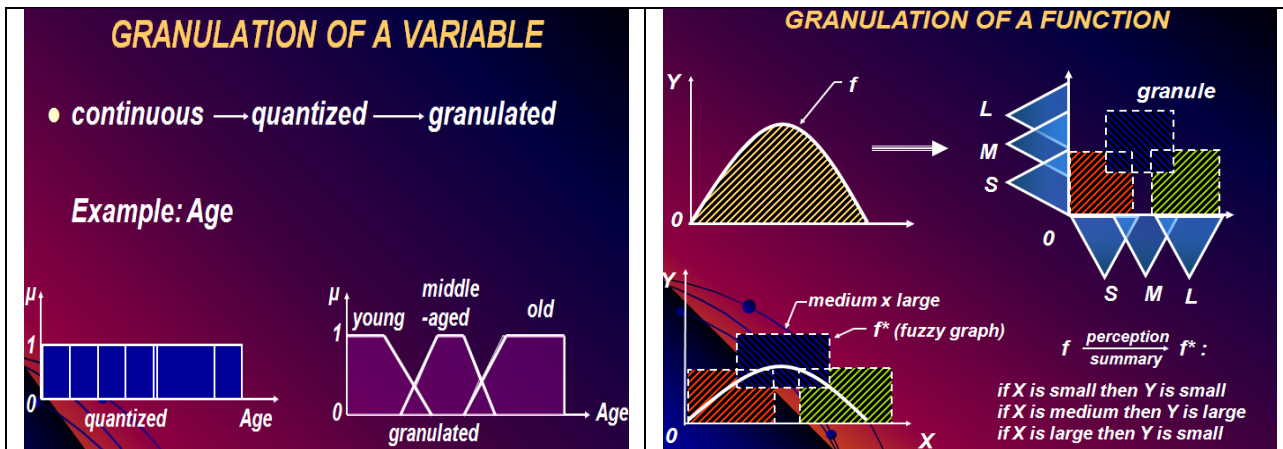


Fig. II.2.2\_1 [fz lz] Stânga: variabila lingvistică “Vârstă” poate lua “tânăr”, “între două vârste” și “bătrân”. Dreapta: se poate granula și o funcție - prin generalizare obținem o schemă de inferență:

**Mulțimile fuzzy adaugă o nouă dimensiune evaluării noastre prin ordonare, gradare,** mulțimea fuzzy putând fi văzută ca o mulțime de mulțimi ‘crisp’ (teorema de prezentare Negoită-Ralescu). Logica booleană clasică e un caz limită al logicii fuzzy, probabilitatea, un caz particular al posibilității.

Teoria fuzzy oferă un instrument matematic pentru reprezentare cunoștințelor umane de forma regulilor, pentru formularea lingvistică a informației, informație de multe ori imposibil de cuantificat numeric. Cunoștințele sunt interpretate ca o mulțime de constrângeri elastice - în cazul bazelor de date relaționale, menționat mai sus, acestea înseamnă relații de ordin 2 (intervale) - asupra unei mulțimi de variabile, *inferența fiind un proces de propagare a constrângerilor elastice.*

### Conceptele de bază ale logicii fuzzy

Se poate spune că matematica fuzzy a ajuns la o maturitate care a deschis și poate deschide părții largi aplicațiilor ingineresti, practica fuzzy consacrand această matematică.

În [Flondor, 2003] se remarcă: “ceea ce algebra booleană e pentru logica clasică, e algebra MV (*many-valued*) pentru logica Łukasiewicz, raționamentul aproximativ putând fi modelat în algebrele MV [...] Prin funcțiile Mc. Naughton algebrele MV pot fi aplicate în studiul rețelelor neurale”. În articol se găsește o conexiune între algebrele MV libere și haosul deterministic, văzut ca dinamică prin substituții între stări modelate printr-o algebră MV, și se evidențiază **caracterul topologic** al raționamentului aproximativ și **caracterul trunchiat al operației de adunare** în algebrele MV, caracter prin care se pot. descrie mai compact automatele finite deterministe și rețelele neurale.

Din punct de vedere matematic, clasele corespund definiției unei relații de echivalență: reflexivă, simetrică și tranzitivă. Aceste relații se pot relaxa prin eliminarea condiției de tranzitivitate – raționamentul aproximativ (inferența fuzzy) nu poate fi aplicat iterativ (vezi ‘Sorites paradox’).

Obținem o *relație de similaritate*. Acest tip de relație corespunde clasificării fuzzy, care permite suprapunerii parțiale ale claselor.

Prin intermediul funcției de apartenență, o generalizare a funcției caracteristice din cadrul teoriei clasice a mulțimilor, se definesc mulțimile fuzzy. Tăieturile (' $\alpha$ -cut') / nivelele (închise) ale unui set fuzzy sunt fundamentale în teoria seturilor fuzzy. Caracterizarea seturilor prin tăieturi, adică  $A_T = \{x/\mu_A(x) \geq T\}$ , o dă teorema de reprezentare Negoită-Ralescu. Conform acestei teoreme o aplicație descrescătoare de la  $[0,1]$  la  $P(X)$  care satisface convergența  $A_t = \bigcap_{s < t} A_s$  corespunde unui ! unic set fuzzy în  $X$ . Și reciproc: nivelele  $A_{(t)}$  ale unui set fuzzy sunt o aplicație descrescătoare de la  $[0,1]$  la  $P(X)$  care satisface convergența. Observații:

- 1) Putem spune că teorema de reprezentare de mai sus leagă matematica clasică / mulțimile clasice de cea / cele fuzzy. Mulțimea fuzzy poate fi văzută ca o mulțime de mulțimi 'crisp'.
- 2) Pe baza  $\alpha$ -cut se definesc operații între numere fuzzy.
- 3)  $\alpha$  din  $\alpha$ -cut poate fi văzut ca un nivel de apoximare al setului (sau numărului) fuzzy.

Teorema de separare pentru seturi fuzzy convexe [Zadeh, 1965]: Fie  $A$  și  $B$  seturi fuzzy n-D, cu suport mărginit și fie  $M_A$  și  $M_B$  gradele lor maxime,  $M$  fiind gradul maximal al setului  $A \cap B$ . Atunci gradul maxim de separare între  $A$  și  $B$  e  $D = 1 - M$ .

Această teoremă are o importanță specială: se poate demonstra astfel utopia unor probleme de clasificare – clasificare – vom exemplifica în finalul secțiunii IV.3.5.

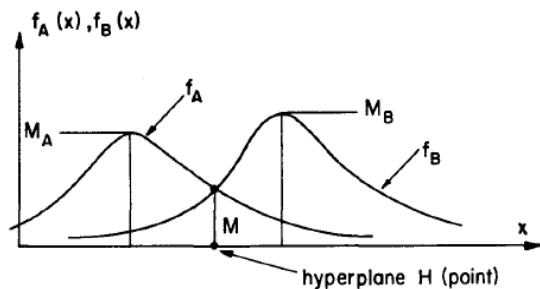


Fig. II.2.3\_9 Un grad mai mare de separare între  $A$  și  $B$  nu se poate realiza. Exemplificare pentru cazul 1-D.

### Reguli și sistem expert fuzzy

În aplicațiile teoriei fuzzy, informația dată prin descriptorii lingvistici, informația incompletă, imprecisă sau calitativă, e.g., informații de expertiză sau reguli / cerințe de proiectare, sisteme "laxe", e procesată conform unor reguli de inferență ale sistemului expert / fuzzy. Aceste reguli de inferență sunt de forma: DACĂ  $x_1$  ȘI  $x_2$  ATUNCI  $y$ . Remarcăm în cazul acestui sistem expert:

- Folosirea variabilelor lingvistice, atât ca intrări cât și ca ieșiri
- Conexiuni logice clasice (ȘI, SAU ș.a.m.d.) ale variabilelor lingvistice și calificatori fuzzy (modificatori generali: foarte, extrem de; cuantificatori: majoritatea, câțiva, puțini; posibilități: aproape imposibil, foarte probabil / posibil etc.).

## “Fuzzy logic = calculul cu cuvinte” [Zadeh, 1996]

În [Zadeh, 2002], se pleacă de la ideea că, pentru a înzestra conceptul de probabilitate cu abilitatea de a opera cu informații bazate pe percepții, trebuie să generalizăm însuși conceptul de probabilitate. Generalizarea se poate face la trei niveluri:

- 1) PT+ (‘Probability theory’) corespunde fuzificării (f) – la nivelul PT+ seturile ‘crisp’ se generalizează la seturi fuzzy, prin introducerea conceptului de apartenență într-un anumit grad
- 2) PT++ corespunde fuzificării granulate (fg) – prin introducerea conceptului de granulă (vezi figura de mai sus)
- 3) PTP corespunde adăugării nivelului PT++ a capacității de a descrie percepțiile prin propoziții într-un limbaj natural (NL = Natural Language). ! NL nu e orice limbaj natural – se impune condiția ca înțelesul unei propoziții exprimate în NL să poată fi reprezentat sub forma unei constrângeri generalizate (GC) a unei variabile.

Primul pas al “raționamentului probabilistic pe baza percepțiilor” e traducerea descrierii percepțiilor în Generalized Constraint Language (GCL) folosind semanticile centrate pe constrângeri ale limbajului natural (CSNL). În pasul al doilea, conform inferenței date de regulile fuzzy, constrângerile generalizate (GC) se propagă de la nivelul premizelor la cel al consecințelor – regulile fuzzy coincid cu regulile de propagare în GCL. Principala regulă de inferență e aceea a generalizării principiului extensiei. O propoziție dintr-un limbaj natural (NL), se zice că e admisibilă dacă se poate traduce în GCL. Propozițiile admisibile din NL alcătuiesc NL precizat. Conceptul de limbaj natural precizat (PNL) deschide cale largă limbajelor naturale în procesarea informației, decizie și control.

În ceea ce privește scalele fuzzy, de care vorbim în subsecțiunea II.2.4.2., remarcăm că folosirea unei astfel de scale metrice e o îmbunătățire, dacă astfel se ia în considerare o relație de proximitate pe setul de reprezentare, care are acoperire într-o relație de proximitate pe setul reprezentat. E.g., în spațiile de culoare perceptual uniforme putem vorbi de o distanță între culori, deci și între denumirile lor.

### PROTOTIPUL = clasa fuzzy

Psihologul Eleanor Rosch a arătat [Rosch, 2010] că anumite cuvinte – pe care le numește “prototipuri” – reprezintă mai bine o clasă decât altele, și că ordonarea acestor prototipuri e conformă cu înțelegerea noastră intuitivă. Rosch a mai afirmat că necesitatea conceptului de clasă ține de randamentul informațional: clasele (n.n.: cuvintele / prototipurile) sunt necesare întrucât furnizează maximum de informație “with the least cognitive effort” [fz ext]. După cum am văzut “legile fuzzy” permit unui obiect să aparțină în același timp mai multor clase: clasele fuzzy se pot suprapune. Dacă pentru clasele / partițiile definite precis (clasic), adică prin relații de echivalență avem termenul de tipuri, pentru clasele / partițiile fuzzy definite imprecis, prin relații de similaritate, propunem termenul de PROTOTIPURI.

Prin PROTOTIP înțelegem un “sâmbure” în care se centrează clasa fuzzy, așa cum un set fuzzy, care poate fi văzut și el ca o clasă fuzzy, are un nucleu și o distribuție de posibilitate în jurul nucleului. Această semnificație a prototipului ni se pare mai potrivită pentru a descrie o clasă fuzzy decât cea de

granulă: granula se construiește în jurul prototipului. Aceasta e ideea esențială de set fuzzy. În secțiunea anterioară am avut pe post de prototip termenul lexical.

Prin PROTOTIP înțelegem un “sâmbure” în care se centrează clasa fuzzy, așa cum un set fuzzy, care poate fi văzut și el ca o clasă fuzzy, are un nucleu și o distribuție de posibilitate în jurul nucleului. Această semnificație a prototipului ni se pare mai potrivită pentru a descrie o clasă fuzzy decât aceea de granulă: granula se construiește în jurul prototipului. Aceasta e, de fapt, ideea esențială de set fuzzy. Putem vedea și variabila lingvistică pe post de prototip. Și o sursă de potențial e un prototip - vezi metoda funcțiilor de potențial de la subsecțiunea II.2.5.1).

Generalizând conceptul de prototip, în domeniul frecvență, prototipul corespunde undei / undinei, iar în domeniul fuzzy, prototipul corespunde clasei. Cu precizarea că, dacă în primul caz, separarea semnalului în unde / undine se face printr-o descompunere ortogonală, în cazul al doilea, separarea în clase admite o suprapunere a claselor. Dacă în domeniul frecvență separabilitatea e controlată conform principiului incertitudinii (vezi II.1.3.2.: Interconținerea rezoluției reprezentărilor în Frecvență și în Frecvență-Timp: principiul nedeterminării), în domeniul fuzzy ea e controlată conform teoremei de separare pentru seturi fuzzy convexe (vezi II.2.2.2.: Mulțimile / partițiile fuzzy). Această teoremă, în “spirit” fuzzy, fără a stabili niște valori precise, vorbește despre un grad maxim de separare.

## **III§ Domeniul multispectral al imaginilor color**

În acest capitol, deși ne vom referi numai la cazul domeniului multispectral al culorilor, adică la receptarea radiației electromagnetice pe trei componente, discuția noastră se poate generaliza la reprezentări spectrale pentru mai multe componente. Practic, în acest capitol prezentăm pe scurt noțiunile teoretice necesare capitolului de contribuții practice (IV§), aducând în același timp argumente în ce privește “fuzitatea” sau caracterul perceptual al culorii.

### *III.1. Introducere în colorimetrie*

Aici prezentăm principiul tricromaticității și diagrama de cromaticitate CIE

### *III.2. O trecere în revistă a spațiilor de culoare*

Aici prezentăm spațiile de culoare primare, derivate, perceptuale și perceptual uniforme.

În practică, e mai la îndemână evaluarea diferenței dintre culori asemănătoare decât dintre culori mult diferite. Astfel, ca unitate de măsură a diferenței dintre culori, apare “diferența abia perceptibilă” (‘Just Noticeable Difference’). JND în sistemele de reprezentare primare e foarte neregulată. În planul de

cromaticitate xy, JND corespunde elipselor Mac Adams – în interiorul acestora culorile nu pot fi diferențiate vizual. Pentru că mărimea și excentricitatea acestor elipse variază puternic cu poziția centrului lor, s-au introdus spațiile de culoare perceptual uniforme.

### *III.3. Caracterul perceptual al culorii*

#### Metamerismul sau granularitatea spectrului vizibil

Potrivirea metamerică (a două culori) sau metamerismul e situația în care două spectre electromagnetice de puteri diferite apar ca aceeași culoare. Metamerismul survine faptului că există numai trei tipuri de conuri care trebuie să încarce cumulativ energie venind din partea unui interval larg de lungimi de undă. În esență metamerismul e aliasing de culori și poate fi descris generalizând teorema de eșantionare a lui Shannon. Metamerismul depinde de:

- iluminare, pentru că apariția culorii depinde de aceasta. Astfel, dacă schimbăm sursa de lumină, culorile pot să nu mai aibă aceeași apariție
- geometrie – schimbarea unghiului de vedere poate “strica” metamerismul
- observator – un observator poate sesiza metamerism acolo unde altul nu-l “vede”
- dimensiune.

Caracterul perceptual al culorii mai reiese și din:

- Constanța culorii și vederea prin ‘retinex’: vedem “nu numai prin retină dar și prin cortex, adică prin ‘retinex’”, după cum a numit Edwin H. Land acest sistem
- Teoria culorilor opuse: stabilește o dihotomie presupunând că orice culoare poate fi descrisă ca fie roșie (galbenă) fie verzuie (albastră). O observație experimentală interesantă e aceea că o linie verde pe un fond roșu apare ca fiind gri. Explicația ține de capacitatea omului de a integra și însuma imediat informațiile pe care le primește. Dacă tricromaticitatea e un model al retinei, teoria culorilor opuse e un model al unui mecanism central din creier.

#### Fuzitatea culorii și procesarea multi-canal

În [Vertan & Buzuloiu, 2000] se observă că procesarea multi-canal a semnalelor:

- 1) a devenit o necesitate.
- 2) procesarea independentă a fiecărui canal s-a dovedit inadecvată deoarece neglijează corelația între canale; pe de altă parte tehnicile de decorelare, e.g., KLT, s-au dovedit ineficiente computațional și adecvate doar pentru clasa semnalelor cu proprietăți statistice stabile în timp.

S-au exploatat însă cu succes, prin reguli fuzzy pentru componente, vezi și [Carron & Lambert, 1996], câteva proprietăți interesante ale spațiilor de culoare perceptuale cum ar fi:

- în imaginile naturale saturația e relativ scăzută și e proporțională într-un grad semnificativ cu nuanța
- zgomotul independent per canalele R / G / B e redus în componenta I (luminanță).

Imaginile digitale sunt mapări 3D → 2D ale scenelor naturale, eșantionate, cuantizate, înglobând astfel multă incertitudine, atât în ceea ce privește valoarea cât și locația. Abordarea fuzzy a procesării multi-canal a venit natural - *filtrele fuzzy* sunt o categorie aparte de filtre, numite și “*granular adaptive*” și bazate pe experiența unui **expert (uman) capabil să furnizeze reguli de clasificare**.

## IVȘ Contribuții practice

### *IV.1 Transformări eficiente între spații de culoare*

Utilitatea trecerii dintr-un spațiu de culoare în altul e evidentă din moment ce o problemă se poate rezolva mai bine (am construit deja un algoritm e.g., un algoritm de “reducere a ochilor roșii”, optim, calibrat) într-un spațiu de culoare, X, decât în altul, Y. Portarea rezolvării problemei din X în Y presupune etapele: aplicarea unei transformări directe  $X \rightarrow Y$ , rezolvarea problemei în Y și transformarea inversă  $Y \rightarrow X$ . În ce privește problema generală pe care o tratăm în această lucrare, considerăm că alegerea unui spațiu de culoare potrivit poate fi decisivă pentru rezolvarea unei probleme de clasificare în imagini.

Contribuția noastră constă în faptul că introducem o metodă nouă prin care obținem performanțe mai bune în ce privește acuratețea și viteza de calcul a unei conversii oarecare între spații de culoare pentru echipamente – vezi și [Georgescu 1, 2009].

#### Implementarea conversiei între două spații de culoare pentru echipamente

În acest subcapitol vom folosi două spații de culoare echipament (pentru echipamente), aditive, care sunt un standard de facto în industrie: sRGB (standard RGB) și Adobe RGB (1998) - vezi [ICC]. sRGB e un standard promovat de corporațiile HP și Microsoft, fiind cel mai cunoscut spațiu de culoare pentru monitoarele PC-urilor și spațiul de culoare presupus de cele mai multe aplicații (inclusiv browser-ele web), deși gama sRGB e una relativ redusă. Adobe RGB (1998) e spațiul de culoare cel mai folosit în publicații de designeri, graficieni și, în general, lucrători profesioniști cu imagini digitale.

Ilustrăm procesul conversiei între spațiile de culoare echipament sRGB (iluminant  $D_{65}$ ) și ProPhoto (iluminant  $D_{50}$ ). Chiar dacă e vorba de aceste spații de culoare, procedura ilustrată e una generală, adică de conversie între oricare două spații de culoare (pentru echipamente). Principiul conversiei: transformarea trebuie să “treacă” prin stimulii absoluți (‘absolute tristimulus’ conform ICC). Etape:

1.  $(R', G', B')$  echipament  $\rightarrow (R, G, B)$  liniar după corecția gamma sRGB
2.  $(R, G, B)$  liniar  $\rightarrow (X, Y, Z)$  : tristimulus relativ sRGB
3.  $(X, Y, Z)$ , adică tristimulus relativ sRGB  $\rightarrow (X, Y, Z)$  tristimulus absolut în cadrul conversiei între iluminanți diferiți (‘chromatic adaptation transform’)
4. și 5. reiau în sens invers 2. și 1.

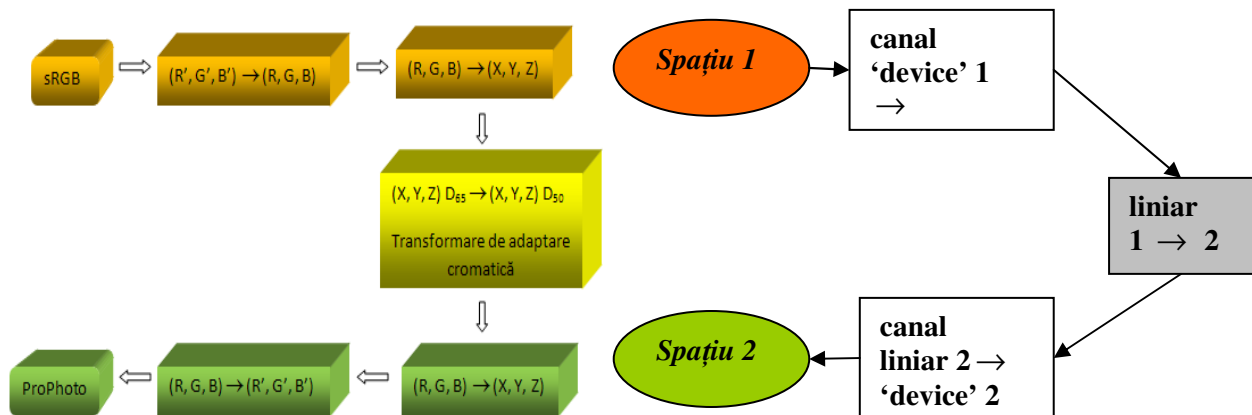


Fig. IV.1.1\_2

Fig. IV.1.2\_1

Fig. IV.1.1\_2 Etapele și schema (Fig. IV.1.2\_1) procesului de conversie între cele două spații de culoare

Testarea acurateții și vitezei de calcul. Surse de erori (puncte slabe)

Prin acuratețe înțelegem impresia perceptuală și reversibilitatea transformării, adică cât mai puține variații (erori) pe parcursul unei transformări dus-întors. e.g., dacă plecăm de la o imagine  $A$  în Adobe RGB și facem conversiile  $A(AdobeRGB) \rightarrow S(sRGB) \rightarrow AA(AdobeRGB)$ . Ideal,  $AA = A$ .

Numim cele trei blocuri din IV.1.2\_1.a: Cdev\_lin, 01 (acest bloc operează cu numere pozitive subunitare) și Clin\_dev. Mai mult, Cdev\_lin și Clin\_dev fiind transformări gamma directă per canal procesul unei conversii devine:  $\text{gamma} \rightarrow 01 \rightarrow \text{gamma}^{-1}$ . De asemenea, în loc de (R', G', B') echipament în spațiul 1, vom scrie RGBdev1. În ce privește blocul 01, de conversie de la coordonatele liniare într-un spațiu de culoare la coordonatele liniare într-un alt spațiu de culoare, nu apar erori: într-o aplicație scrisă în limbajul C, optimizată, erorile apar la a opta zecimală după virgulă. Motivul e simplu: putem reduce toate operațiile care au loc în acest bloc la o înmulțire cu o matrice (3 x 3).

Erorile apar inerent pentru transformările gamma (directă și inversă), iar noi trebuie să interpolăm pentru a efectua această transformare dată, prin puncte, și nu printr-o formulă matematică (deși știm din teorie și vom folosi că transformarea (sau corecția) gamma e o funcție putere de forma  $O = I^\gamma$ ,  $\gamma$  fiind coeficientul gamma, intrarea I și ieșirea O situându-se în intervalul [0, 1]).

### Tratarea erorilor. Optimizări

Întrebarea principală care se pune în ce privește acuratețea conversiei e: cum să interpolăm între punctele de control care specifică transformarea gamma pentru a obține cea mai "exactă" transformare gamma? Conform teoriei, transformarea gamma e  $y = x^{\text{exp}}$  (g), unde x e valoarea unei componente R/G/B 'device' normalizată la [0,1], y e R/G/B liniar și exp e un număr real. E o dependență nonlineară. Interpolarea liniară ar fi cea mai bună metodă, prin simplitate și viteză de calcul. Să aplicăm logaritmul relației (g). Obținem o dependență liniară [5]:  $\lg(y) = \text{exp} \lg(x) \rightarrow$  pașii:

1. Aplicăm logaritmul coordonatelor punctelor de control (la o intrare 0, ieșirea va fi 0: tratăm special acest caz pentru a ne asigura condițiile de existență pentru logaritmare):
2. Interpolează liniar pentru a calcula  $\lg(\text{R/G/B liniar}) = \text{R/G/B } i$
3. Calculează  $\text{R/G/B linear} = (\text{R/G/B } i)^2$  (folosim logaritmul în baza 2 pentru viteză de calcul – împărțirile la 2 se realizează ca deplasări la dreapta ale biților).

Toate acestea le facem într-o etapă de preprocesare - construim un LUT ('lookup table') pentru procesarea efectivă. Mai mult, să ne amintim că trecerea de la un RGB linear la altul (blocul 01) e, în fapt, înmulțire cu o matrice (3 x 3). Inserăm ȘI această informație (înmulțirea cu o constantă) în LUT. Astfel, pentru o acuratețe și o viteză de calcul mai mari avem

- pentru blocul gamma: lutTRC R/G/B : [0..255]  $\mapsto$  [0, 1]
- pentru blocul 01: lutMatrix R/G/B : [0..255]  $\mapsto$  [0,1]. Acest LUT folosește lutTRC și precalculează termenii conform relației (3), astfel încât trecem un pixel din RGBdev1 în RGBlin2 prin trei operații de 'lokup' cu lutMatrix și două adunări.
- pentru blocul  $\text{gamma}^{-1}$  folosim lutTRC<sup>inv</sup> [0,1]  $\mapsto$  [0..255].

Cum interpolăm pentru blocul  $\gamma^{-1}$ ? Avem de ales o singură valoare ieșirea corespunzătoare componentei R/G/B a unui pixel. Vom folosi această restricție și faptul că transformarea gamma (directă sau inversă) e monotonă. Astfel, calculăm la preprocesare conform coordonatelor punctelor de control (pentru acuratețe folosim aceeași abordare ca în cazul transformării gamma directe - logaritmare, interpolare liniară și ridicare la putere):  $\text{lutTRC}^{\text{inv}} \text{ R/G/B} : [0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots, 255] \mapsto [0, 1]$ , adică luăm și mijloacele intervalelor definite de întregi în domeniul acestui LUT. Acest lucru ne va fi deosebit de util în a implementa o transformare gamma inversă eficientă. Fie  $x$  valoarea unei componente liniare R/G/B și, prin transformarea gamma inversă, vrem să aflăm valoarea componentei echipament  $y \in [0..255]$ . Folosim următoarea teoremă, valabilă (cu adaptarea inegalității) pentru orice  $f$  monotonă:

Teoremă: Fie o funcție  $f : \mathbb{R} \mapsto \mathbb{R}$  monoton crescătoare. Pentru a decide dacă  $y$  e una din două valori consecutive pe ordonată ( $y_1$  și  $y_2 = y_1 + 1$ ), dându-se o valoare  $x$  pe abscisă, cea mai eficientă metodă e aceasta:  $y > f(x_m) ? x_2 : x_1$ , unde  $x_m = x_1 + \frac{1}{2}$ , i.e. alege  $x_2$  dacă  $y > f(x_m)$ , altfel alege  $x_1$ .

Funcția gamma e monoton crescătoare. Astfel, conform teoremei de mai sus, putem folosi un LUT ( $\text{lutTRC}^{-1}$ ) și pentru blocul  $\gamma^{-1}$ . Operația de lookup, făcându-se pe un vector ordonat, are complexitatea  $O(\lg N)$ , aceasta fiind și complexitatea trecerii unui pixel dintr-un spațiu de culoare echipament într-altul.

Mai putem optimiza viteza de calcul folosind:

- operatori de deplasare pe biți pentru implementarea înmulțirilor și împărțirilor
- o abordare ‘run length’ – în general, într-o imagine sunt regiuni de mărime considerabilă în interiorul cărora pixelii au culori cel puțin similare (e o problemă foarte interesantă de clasificare) => într-un proces de conversie ar fi util să reținem valorile deja calculate. Există și CLUT (color look-up table): un mecanism folosit pentru a transforma un interval de culori în alt interval de culori. CLUT poate fi implementat hardware sau software.

### Rezultatele conversiei

Pentru testarea am folosit un ‘Color Checker’ fotografiat în format .bmp, la o rezoluție de 7 M px:

Așa cum am spus la secțiunea IV.1.2.2 testăm acuratețea conversiei prin impresia perceptuală și reversibilitatea transformării, adică gradul în care apar cât mai puține variații (erori) pe parcursul unei transformări dus-întors între cele două spații de culoare. Impresie perceptuală pe care am încercat să o redăm în figura de mai jos, chiar dacă, restricționați de suportul acestei lucrări, prezentăm imagini nu numai comprimate de la formatul .bmp la formatul .jpg, dar și micșorate:



Fig. IV.1.3\_1 și IV.1.3\_2: prima imagine corespunde unei conversii Adobe RGB → sRGB cu logaritm (metoda noastră), următoarea corespunde aceleiași conversii cu metoda clasică (interpolare liniară)

La o privire atentă, câteva diferențe între cele două imagini pot fi observate, în special, în regiunile unde negru e predominant, e.g. sub masă, la obiectele negre și în zonele de umbră. În conversia fără logaritm, pare că un praf foarte fin se întinde peste imagine. E predictibilă această aparență – iată un argument grafic:

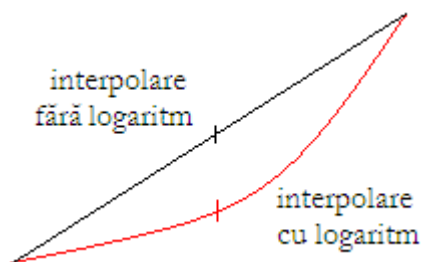


Fig. IV.1.3\_3 Argument grafic al acurateții superioare a metodei noastre, dar și al aparenței imaginilor anterioare

Reversibilitatea transformării a fost măsurată în tabela de diferențe:

<i>DIFFs %</i>	<i>AAlog_A</i>	<i>AAdll_A</i>	<i>Sfoto_Slog</i>	<i>Sfoto_Sdll</i>	<i>Slog_Sdll</i>
0	0.2540932	73.9935169	77.5440351	77.6912546	64.4213062
1	0.0842800	0.0898927	24.5896225	36.5294016	0.4897142
2	0.0518853	0.0539477	2.8738575	5.5211811	0.0000000
3	0.0479098	0.0487633	1.2290542	1.1713816	0.0000000
4	0.0437978	0.048417	0.6488032	0.7836487	0.0000000
5	0.0167462	0.0228856	0.3099478	0.4948573	0.0000000
6	0.0004093	0.0014532	0.1009818	0.2982464	0.0000000
<i>Max difference</i>	34	36	52	59	8
<i>channel</i>	B	B	R	R	R
<i>(R,G,B)pixel</i>	(186,166,3)	(186,166,3)	(52,92,56)	(0,90,49)	(9,15,11)

Tabela arată diferențele între imaginile rezultate în urma procesării unei imagini inițiale: A.bmp = imaginea cu ColorChecker luată cu o cameră foto Nikon în spațiul de culoare Adobe RGB. Acesta e înțelesul lui ‘A’. ‘AA’ înseamnă imaginea trecută în sRGB (imaginea ‘S’) și înapoi în Adobe RGB. ‘log’ se referă la conversia cu logaritm, ‘dll’ se referă la soft-ul Nikon pe care l-am avut la dispoziție

pentru a avea un termen de comparație relevant, iar ‘foto’ se referă la o cameră Nikon performantă care putea achiziționa aceeași scena atât în sRGB, cât și în Adobe RGB.

Prima coloană indică:

- $\langle i \rangle$  : =  $i * 5$  prag pentru compararea imaginilor octet cu octet, e.g. ‘3’ se referă la diferențele peste 15. Diferențele sunt calculate ca procent din numărul total de octeți din imagine.
- ‘Max difference’: diferența maximă per octet între imagini
- ‘Channel’: componenta R/G/B pentru care am obținut ‘Max diff’
- ‘(R,G,B) pixel’: pixelul care conține ‘Channel’.

Cifrele din tabela arată nu numai că metoda noastră e comparabilă cu cea a lui Nikon – vezi coloana Sfoto\_Sdll, coloană care mai arată că softul produs de Nikon nu interpolează în domeniul logaritmic (și, știm, nici liniar) – ci și că, în unele privințe, chiar îi e superioară. Astfel, se confirmă impresia perceptuală că varianta ‘log’ e mai precisă, comparând coloanele ‘AAlog\_A’ și ‘AAdll\_A’ din tabelă, coloanele care numără erorile introduse de o conversie dus-întors. E interesant de observat că ‘(R,G,B) pixel’ e același pentru aceste coloane. Coloanele Sfoto\_Slog și Sfoto\_Sdll ne arată că, prin varianta noastră, ne apropiem mai bine decât însuși softul produs de Nikon de fotografia obținută cu camera Nikon, fotografie care beneficiază și de niște corecții suplimentare.

În ce privește viteza de calcul a conversiei, considerăm că am argumentat suficient. Mai afirmăm numai că, folosind LUT-uri, timpul de procesare se reduce de la minute la secunde.

### Concluzii

Termenul de comparație a metodei noastre a fost o metodă de conversie (între spații de culoare pentru echipamente) “de top” din industrie, anume cea folosită de firma Nikon, metodă a cărei implementare (cod executabil, dar și cod sursă) am avut-o la dispoziție. Am lucrat chiar cu profilele de culoare sRGB și Adobe RGB pentru Nikon. Implementarea noastră a fost propusă firmei Nikon, nu știm ce s-a întâmplat mai departe, întrucât am schimbat serviciul. Avantajele majore ale metodei noastre sunt:

1. viteza de calcul superioară grație unei “lut-ificări” ingenioase, fără a folosi multă memorie suplimentară pentru aceasta. În fapt, construim în etapa de preprocesare următoarele 3 LUT:

$$\text{lutTRC R/G/B} : [0..255] \mapsto [0, 1]$$

$$\text{lutMatrix R/G/B} : [0..255] \mapsto [0, 1]$$

$$\text{lutTRC}^{-1} \text{ R/G/B} : [0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots, 255] \mapsto [0, 1]$$

Prima și a treia sunt destinate transformării / corecției gamma, directe și, respectiv, inverse, într-un spațiu de culoare (pt.) echipament(e) – de fapt, transformarea care pune “la încercare” viteza de calcul și acuratețea metodei; cea de-a doua e destinată trecerii de la “coordonatele” liniare într-un spațiu de culoare echipament la “coordonatele” liniare într-un alt spațiu de culoare echipament.

“Lut-ificarea” e bazată atât pe faptul că reprezentarea componentelor unei culori se face într-un domeniu limitat, darși pe analiza pe componente a procesului conversiei între cele două spații de culoare. Pentru transformarea gamma inversă lut-ificarea” e bazată pe teorema pe care am enunțat-o la sfârșitul subsecțiunii IV.1.2.3.

2. acuratețe grație observării alurii graficului transformării gamma conform punctelor de control specificate de profilele de culoare și aplicării unei interpolări potrivite, adică a interpolării liniare în domeniul logaritmic – aceasta a accelerat ȘI viteza de calcul

3. generalitatea abordării: adresăm problema oricărei conversii între spații de culoare pentru echipamente. Așa cum am subliniat mai sus, zona slabă a unei conversii între două spații de culoare, transformarea gamma o tratăm într-o manieră care se pretează și la alte transformări specificate prin puncte de control, ideea fiind că interpolăm conform alurii curbei transformării, acest fapt dându-ne posibilitatea să avem performanțe atât în ceea ce privește acuratețea, cât și viteza de calcul.

Chiar și procesul de “lut-ificare” poate fi generalizat – de aceea am simțit nevoia să precizez la sfârșitul subsecțiunii IV.1.2.3. acea teoremă.

## *IV.2 Definirea formei spațiului unei probleme de recunoaștere nedeterminate și modelarea acestuia. Cazul ‘skin detection’*

Chiar dacă vom trata problema specifică a determinării eficiente a unui “spectru” de culori conform căruia să putem detecta zonele de piele (umană) dintr-o imagine digitală, adică, ne propunem să determinăm un ‘skin locus’ (SL) în vederea detecției zonelor de piele (‘skin detection’: SD), metoda noastră de extragere maximală (în sensul eficienței menționate) a informației din date poate fi generalizată și pentru alte probleme nedeterminate de clasificare în imagini. Această “extragere maximală” presupune nu numai contorizarea pixelilor care aparțin unei clase (în cazul nostru, clasa pixelilor de skin (S)) și, mai departe, stabilirea unei măsuri probabilistice de a fi ‘skin pixel’ (SP) pentru fiecare pixel dintr-un anumit spațiu de culoare. Se rețin ȘI condițiile de achiziție a imaginilor. Vom stabili o taxonomie a bazei de imagini de calibrare ȘI după tipul și setările camerei foto. Astfel, vom furniza un cadru de lucru adaptiv și dinamic, capabil să îmbunătățească orice metodă de clasificare bazată pe clasificarea pixelilor. Această flexibilitate plus cele două ingrediente majore în industrie, acuratețea și viteza de calcul, vor configura față cu ‘the state of the art’-ul existent o metaeuristică mai eficientă de determinare a formei spațiului unei probleme de recunoaștere

nedeterminate - vezi și [Georgescu, 2010]. Tot în acest subcapitol vom trata și problema modelării acestui spațiu - vezi și [Georgescu, 2011].

Așa cum e arătat într-un articol recapitulativ, [Kakumanu et al., 2006] “recent, metodele SD bazate pe informația culorii pielii au câștigat o mare atenție din punct de vedere al efortului computațional, al robusteții la rotații și scalări, și ocluzii parțiale”. Aceste metode s-au dovedit a fi optime în termeni de robustețe și, vom vedea, în cazul metodei noastre, modularitate. Scopul nostru e de a găsi cel mai “generos” cadru de lucru pentru aceste metode - stabilim o mataeuristică care permite înglobarea eficientă (‘plug and play addition’) într-o metodă de SD a oricărui criteriu nou (e.g. un criteriu de mișcare, sau un criteriu de vecinătate, cum ar fi textura pielii, sau prototipuri de culori ale pielii, sau caracteristici de formă pentru fețe / ochi), în vederea creșterii selectivității metodei (reducerii fals pozitivelor). E o metodă dinamică de detecție (bazată pe pixeli) și în sensul că putem îngloba informații noi pentru detecție prin adăugarea la Skin Locus-ul existent de noi pixeli de piele, sau prin modificarea informației de probabilitate (vezi mai jos) pentru pixeli din SL-ul existent. Dinamicitatea metodei permite și calibrarea ei, e.g., putem calibra metoda pentru detecția pielii în imagini cu persoane de rasă caucaziană.

Pentru a “profita” cât mai mult de pe urma oricărei informații, vom avea în vedere că SD trebuie să țină cont și de caracteristicile de conectivitate a pixelilor de piele - în faza de modelare a SL vom face prelucrări morfologice corespunzătoare asupra SL-ului obținut anterior prin culegerea informației pixelilor. Această “intuiție” topologică, pusă în serviciul metodei noastre, împreună cu informația de probabilitate asociată fiecărui SP ne situează în domeniul tehnicilor fuzzy.

Abordarea e centrată pe datele experimentale, “căzând” abordările care încearcă să deducă ecuații matematice, prin care am putea modela spațiul soluției problemei. De altfel, noi, ca și alți cercetători, am experimentat faptul că o asemenea abordare e mai eficientă decât a încerca să “prindem” SL în formule explicite.

Pentru a ajunge la implementarea de SD, urmăm o succesiune de etape general acceptată – conform [Vezhnevets et al., 2003] apar trei mari subprobleme în construcția unui sistem de SD:

1. ce spațiu de culoare să alegem
2. cum vom modela distribuția culorilor de piele (‘Skin Pixels’) în acest spațiu
3. algoritmul efectiv de SD (ce se întâmplă la preprocesarea unei imagini)

#### Alegerea spațiului de culoare și colectarea pixelilor de piele

Apariția pielii în imagini e, mai degrabă, controlată de cromaticitate decât de luminanță: luminanța poate varia mai amplu decât cromaticitatea. Multe lucrări de determinare a SL într-un spațiu de culoare [Vezhnevets et al., 2003] au renunțat la componenta de luminanță. Această *reducere a dimensionalității* SL (de la 3D la 2D) a eficientizat foarte mult procesarea ulterioară. În plus, SL-ul 2D se poate

“generaliza” la unul 3D, asigurându-ne astfel de reducerea fals negativelor. De altfel, în general, se știe că pentru sistemele SD existente, e mai puțin probabil să ratăm un pixel de piele (SP), decât să confundăm non-SP cu SP. “Culorile de piele” fiind des întâlnite (la diverse obiecte), odată ce am mai și generalizat adecvat SL-ul, putem aborda doar problema reducerii fals pozitivelor. Cum generalizăm SL-ul? Punând praguri globale pentru luminanța și aducând la o formă “cât mai” conexă și convexă SL-ul – pe asemenea componente se pot rula algoritmi foarte eficienți de SD: de tip ‘run-length’.

Mai întâi, devine foarte utilă o decorelare cât mai bună între crominanță și luminanță. E unul dintre motivele principale pentru care am ales spațiul de culoare YCbCr pentru metoda noastră de SD. YCbCr e un spațiu de culoare care se mapează foarte bine și pe cerințele industriei fiind și foarte popular în ‘pixel based SD’ [Gonzales & al., 2010].

### Colectarea pixelilor de piele

Pixelii de piele (SP) îi vom culege dintr-o bază cu fotografii, în care fețele au fost marcate (selectate) prin poligoane, iar benzile (dreptunghiurile) gurii și ochilor au fost îndepărtate. Marcarea nu trebuie să fie foarte exactă. Punând unele praguri pe ordonata graficului histogramei cumulative a luminanței putem elimina din selecție culorile minoritare, care nu pot fi luate drept culori de piele, e.g. fire de păr negre (prag inferior) sau albul ochilor (prag superior). Aceste praguri le punem procentual – între 5 și 10%. Mai reținem pentru fiecare SP: Y, Cb, Cr numărul de apariții (**informație de probabilitate**). Informația de probabilitate, în fapt frecvența relativă a SP, ne dă posibilitatea construirii unui cadru “generos”, cu posibilități de extindere și înglobare a diverse informații, nu numai de natură probabilistică, dar și fuzzy, cu toate avantajele lor menționate la capitolul “Domeniul impreciziei”.

Dacă un om recunoaște foarte ușor o zonă de piele, e nu numai datorită capacităților încă neexplicate ale sistemului său nervos, dar și datorită faptului că această clasificare o face dispunând și de o informație 3D și de multe alte informații. Putem elabora un model de calcul satisfăcător al SD numai dacă ne delimităm cât mai bine condițiile în care clasificăm. E important de avut în vedere că lucrăm cu baze de date de imagini cât mai mari, mai complete, dar și mai “specializate”. Acesta credem că e bine să poarte numele de **principiul specializării și al calibrării**: vom lucra cu seturi de imagini care au fost achiziționate în diverse condiții, cu diverse tipuri și parametri de camere foto, cu persoane de diverse etnii și, deci, culori ale pielii ș.a.m.d.. De fapt, cu factori care ar putea fi știuți înaintea procesării efective a imaginii pentru SD, a.î. să putem alege setul corespunzător.

### Reprezentarea Skin Locus-ului în vederea eficienței Skin Detection

Să tranșăm în vederea eficienței SD, mai exact a aplicabilității SD în timp real: observăm că reprezentarea 3D de la Fig. IV.2.2\_1 seamănă destul de mult cu un cilindru. Se confirmă percepția de la subsecțiunea IV.2.2.1: pentru componenta de luminanță putem reține doar două praguri globale:

unul inferior și unul superior. Astfel, pentru un SP vom reține coordonata Cb, coordonata Cr și numărul de apariții. Să avem în vedere că proiecția în planul CrCb a SL amestecă clasa pixelilor “de” piele SP, cu clasa non-SP. Cu atât mai importantă devine informația de probabilitate, care ne ajută să partiționăm cât se poate (fuzzy) în cele două clase.

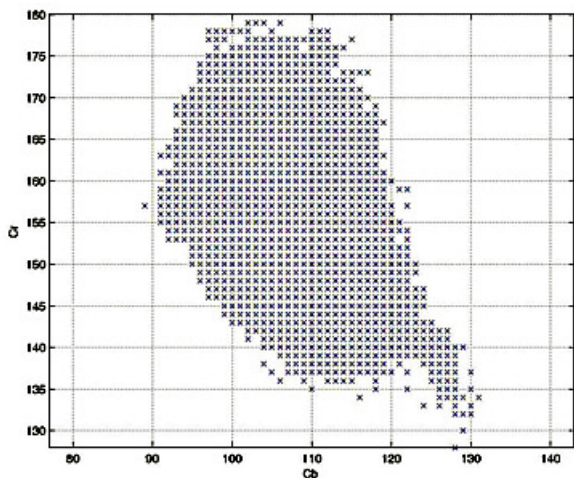
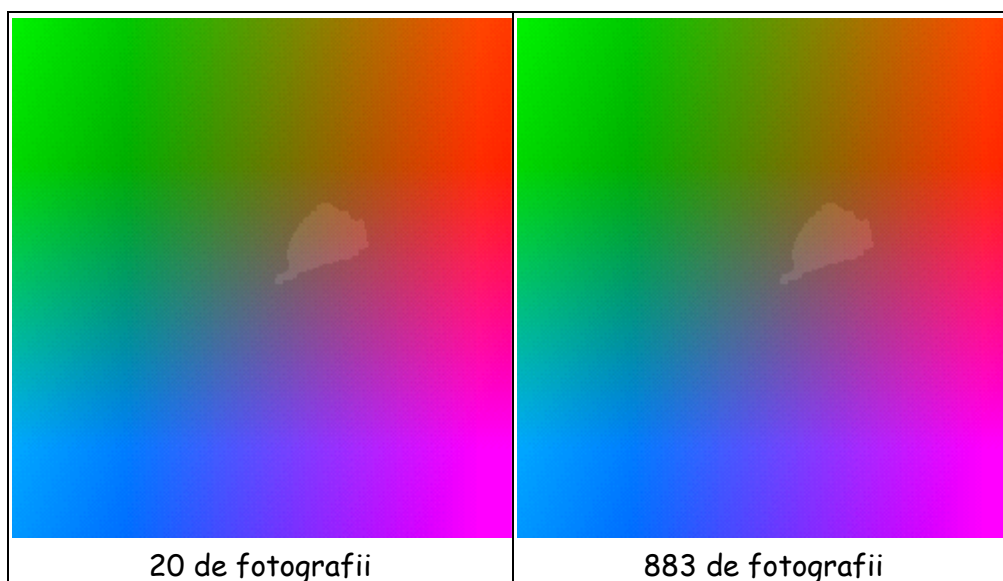


Fig. IV.2.3\_1 Reprezentarea 2D a Skin Locus-ului - în planul CrCb conform [Georgescu, 2010]

### Modelarea distribuției SL



Cât de importantă e completitudinea bazei de imagini din care culegem SP (pentru SL)?

Fig.IV.2.4\_1 Reprezentarea în planul CrCb a SL cules din 20 (în stânga) și din 883 fotografii

Cele două forme de SL de mai sus au câteva asemănări importante:

- sunt regiuni relativ compacte
- acoperă cam aceleași intervale Cb și Cr
- apar similare nu numai topologic, dar și ca formă.

Aceste asemănări majore ne duc la două concluzii importante:

C1) **nu avem nevoie de foarte multe imagini pentru a configura / culege un Skin Locus** (mai degrabă avem nevoie de câteva scene fotografiate în diverse situații (câteva fețe, acestea conținând mulți SP), ceea ce nu înseamnă abandonarea principiului specializării și al calibrării, ci doar că, între anumite limite putem interpola.

C2) SL e relativ **compact**

Să coroborăm aceste informații despre formă cu cele de probabilitate.

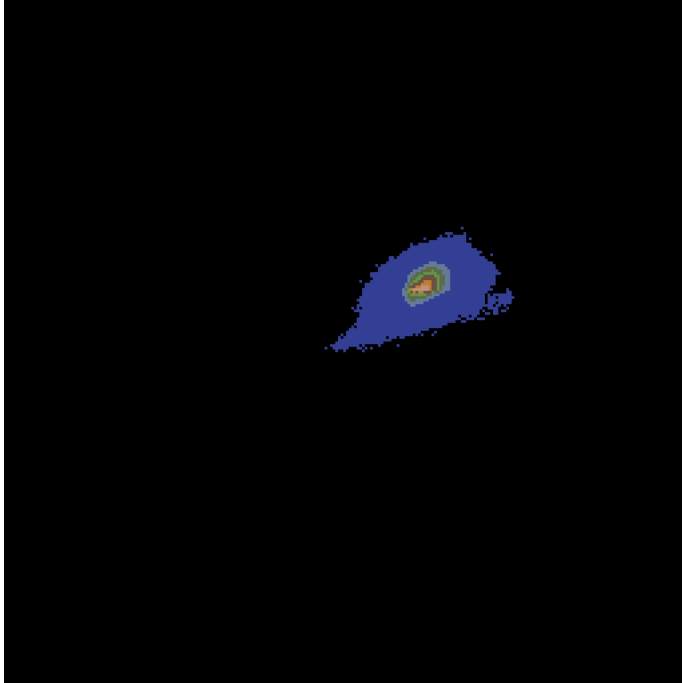


Fig. IV.2.4\_2 Nivele de probabilitate ale SL (cules din 883 fotografii) în planul CrCb. Am reprezentat șapte nivele de probabilitate pseudocolorate (după o convenție asemănătoare cu cea pentru relief, adică albastru închis corespunde SP cu cea mai mică probabilitate, iar cea mai deschisă culoare corespunde SP cu cea mai mare probabilitate).

Observăm **dispunerea concentrică** a nivelelor de probabilitate ale SL, în centru fiind pixelii cu cea mai mare probabilitate. Această observație împreună cu C1) și C2) de mai sus ne face să ne gândim la un SL compact, care poate fi partiționat în componente convexe și conexe, foarte ușor de parcurs de un **algoritm de SD de tip ‘run-length’ – cel mai rapid algoritm pentru problema noastră**: testul dacă un pixel de coordonate [cr] [cb] e SP se

reduce la două comparații (testăm dacă un punct aparține unui segment). Pentru a compacta adecvat SL, ca să aplicăm un astfel de algoritm, putem prelucra morfologic, ținând cont și de informația de probabilitate. Pentru compactare sunt două obiective: întâi conectivitatea și apoi convexitatea.

### Conectarea SL

SL reprezentat în Fig. IV.2.4\_2 prezintă la margine pixeli izolați care ar putea fi zgomote – pixeli care apar accidental ca fiind “de” piele. Putem înlătura acești pixeli prin câteva prelucrări morfologice simple, cum ar fi închiderea cu un mic element structurant. Sau, mai eficient după regulile:

- 1) doi pixeli de același tip (SP sau non-SP) “forțează” pixelul aflat direct între ei la tipul lor, probabilitatea în cazul unui SP creat astfel, fiind media aritmetică a vecinilor care l-au “forțat” să apară
- 2) un SP e făcut non-SP, dacă nu are cel puțin doi vecini

Aplicăm prelucrările 1) și 2) până obținem o singură regiune conexă pentru SL: CN. În cazul nostru, după doar două iterații ale acestor prelucrări, obținem un SL conex și aproape convex – ne punem acum problema obținerii convexității. Putem ajusta local SL (fie și manual). Sau putem aplica o transformare morfologică de tip *acoperire convexă* ('*convex hull*'). **Dacă obținem un CN cu concavități accentuate, îl partiționăm în componente convexe**; altfel, vom pierde mult din selectivitatea metodei de SD.

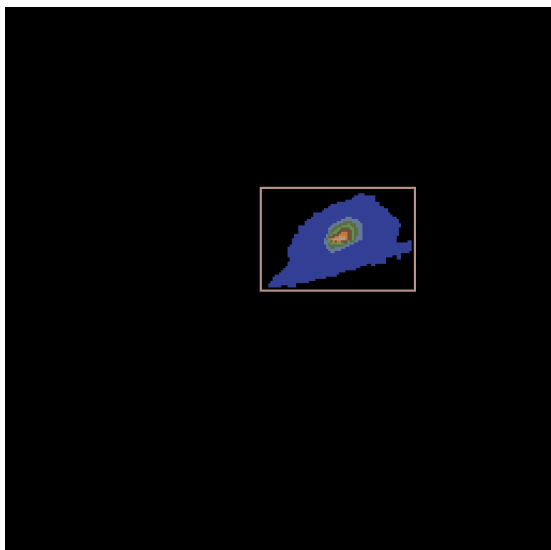


Fig. IV.2.4\_4 Rezultatul aplicării pașilor 1) și 2) e un SL conex. În figură, am reprezentat și dreptunghiul încadrator al SL înainte de a aplica 1) și 2)

### Concluzii

Am abordat dintr-o perspectivă *experimentală, destinată să valorifice optimal informația care poate fi extrasă dintr-o bază de date*, proiectarea unei metaheuristică care să clasifice, în caz concret, pixelii "de" piele în imagini digitale. Avantajele metodei noastre constau în:

- modularitate – etapa de calibrare înglobează dinamic noi informații și se desfășoară decuplată de etapa de procesare efectivă
- viteză de calcul – prin Lut-ificare, compactare a.î. să rulăm un algoritm 'run-length' pe componente conexe și convexe
- acuratețe grație unei adaptabilități superioare conform a ceea ce am numit "principiul specializării și al calibrării"
- generalitate – am proiectat de fapt o metaheuristică
- deschidere – prin informația de probabilitate ne putem extinde metoda în domeniul "impreciziei", adică putem îngloba cunoștințe probabilistice sau fuzzy (lingvistice, perceptuale).

### IV.3 Folosirea subclasării și a marcajelor pentru reducerea fals pozitivelor.

#### Cazul 'skin detection'

##### Abordare

Se pune problema creșterii selectivității unei metode de clasificare, adică a reducerii fals pozitivelor (FP). Ideea principală, pe care se bazează contribuția noastră, e aceea de a împărți / **partiționa clasele furnizate de metoda de clasificare dată (inițială, de intrare) în subclase gășind condiții suplimentare de rafinare a clasificării în interiorul fiecărei clase**. Pentru a ne asigura de corectitudinea acestor condiții suplimentare vom folosi informații perceptuale, adică **marcaje de clasificare furnizate de expertul uman** – vezi și [Georgescu 1, 2012].

Dacă numim prototipuri ale unei clase, subclasele obținute în interiorul fiecărei clase prin partiționare, schema de selecție e dată de relația:  $zona\_de\_interes = zona\_prototipurilor \cap zona\_marcajelor$

Pentru punerea în lumină a acestei metode de creștere a selectivității unei clasificări, vom analiza cu date experimentale, cazul detecției zonelor de piele (Skin Detection) în fotografii – ceea ce corespunde faptului că pornim de la o clasificare cu două clase: 'skin' și 'non-skin'. Condițiile suplimentare le punem pentru a subclasa în interiorul clasei 'skin'. Zona corespunzătoare clasei 'non-skin' nu ne mai interesează întrucât se consideră că, pentru metoda SD dată – acesta e cazul general, fals negativele nu creează probleme, fiind neglijabile. În general, problema majoră pentru SD e dată de faptul că se detectează prea multe fals pozitive, adică se detectează ca 'skin' și zone care în fapt nu sunt de 'skin'.



Fig. IV.3.1\_1 SD tipic. Observație: există multe culori cu o cromaticitate ca a pielii (mai ales în interioare)

Pentru luminanță s-a impus un anumit interval, un prag superior și unul inferior – de aceea nu e selectată zona de piele (S) din depărtare. Dacă ne interesează să selectăm numai S din plan apropiat, mai precis, cu luminanța peste un anumit prag, abordarea e corectă. E și cazul nostru, **întrucât am**

**folosit SD ca pas intermediar pentru problema corecției ochilor roșii** ('Red Eyes Reduction') – e improbabil să obținem ochi roșii în depărtare, deoarece 'flash'-ul, cauza apariției ochilor roșii, are o rază de acțiune destul de mică. A se vedea influența mare a direcției 'flash'-ului: dacă 'flash'-ul ar fi fost "pe" fața din prim plan, această față ar fi fost selectată mai compact - în situația figurată mai sus, pentru o selecție de zone S conexe și convexe, cum ne așteptăm să fie zonele de piele, ca și în cazul metodei de determinare a SL-ului (vezi subcapitolul anterior), pot fi utile anumite prelucrări morfologice. Remarcăm și că o **mulțime de alți factori apar odată cu 'flash'-ul**, fie și numai dacă menționăm numai artefacte ca umbre sau specularități.

Unul dintre scopurile principale ale proiectului la care am lucrat era acela de a crește selectivitatea unei metode de SD bazată pe clasificarea pixelilor (ca fiind de 'skin' sau nu), luând în calcul și informații legate de diferența de luminanță Dif între o fotografie făcută cu 'flash' și una fără 'flash', asupra aceleiași scene, ideea fiind că în zona feței (zona S pe care îmi propun să o selectez cu cât mai puține FP) Dif are o valoare specifică pentru reflectanța pielii (din acea zona a feței). Softul care permitea achiziția fotografiei fără 'flash' era un firmware Nikon. În fapt, fotografia fără 'flash' e un 'preview' (P), adică acea imagine care se formează în timp real pe vizorul camerelor foto digitale de generație mai nouă. Pentru concizia exprimării vom folosi F în loc de "fotografia făcută cu 'flash'". Ne vom raporta în continuarea subcapitolului la relația: **Dif** (diferența pe canalul de luminanță)  $\stackrel{\text{def}}{=} \mathbf{F}$  (luminanța fotografiei cu 'flash') – **P** (luminanța fotografiei fără 'flash')

### Clasificarea skin / non skin

SD făcându-se într-un spațiu perceptual uniform, se justifică definirea de subclase (prototipuri) sub forma unor discuri  $D_i$ ,  $i = 1 \dots n$ , de raze aproximativ egale (6) care acoperă un SL-ul.

În ciuda situației ilustrate în Fig.4.3.1\_1 pentru a pune în evidență problemele majore ale SD, această metodă de definire de SD nu e una cu performanțe scăzute.



Fig. IV.3.2\_2 Selecția zonelor de piele conform metodei noastre

Selecția noastră corespunde unui SL intermediar între cel definit conform [Hsu et al., 2002] și cel definit, precum am menționat în subcapitolul anterior, prin inegalități.

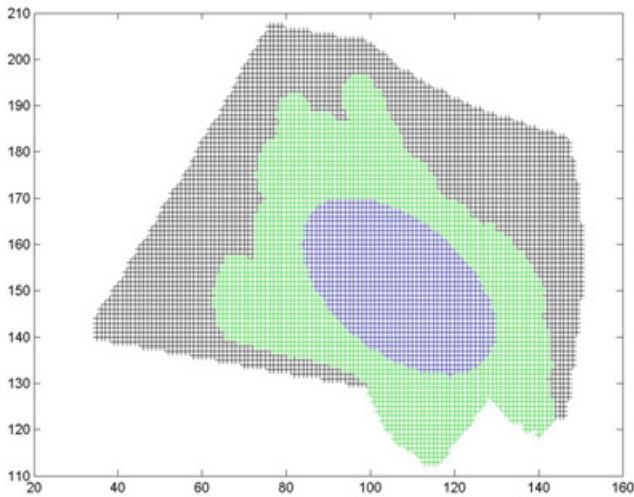


Fig. IV.3.2\_3

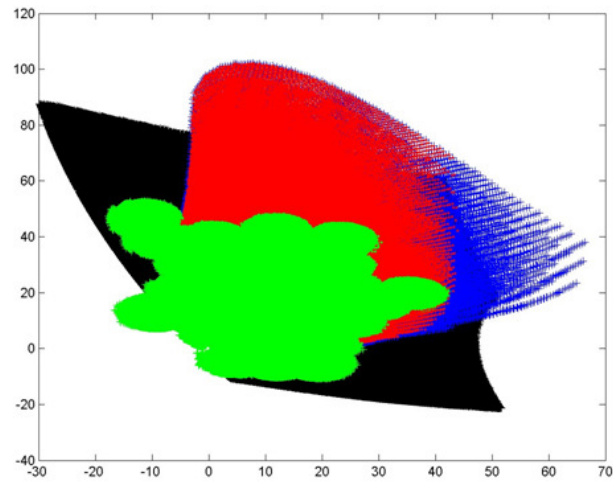


Fig. IV.3.2\_4

Fig. IV.3.2\_3 ‘Skin gamut’-uri în planul CrCb, definite prin trei metode: albastru – metoda conform [Hsu et al., 2002], verde – metoda noastră, negru – metoda cu inegalități; (b) aceleași ‘skin locus’-uri în planul ab (din Lab), roșu apare pentru o limitare inferioară globală a luminanței a modelului conform [Hsu et al., 2002].

Fig. IV.3.2\_4 Observăm că în spațiul de culoare Lab, reprezentarea cu discuri e cea mai compactă. Vom compara cele trei metode și aplicându-le pentru o imagine care conține mostrele de piele umane:

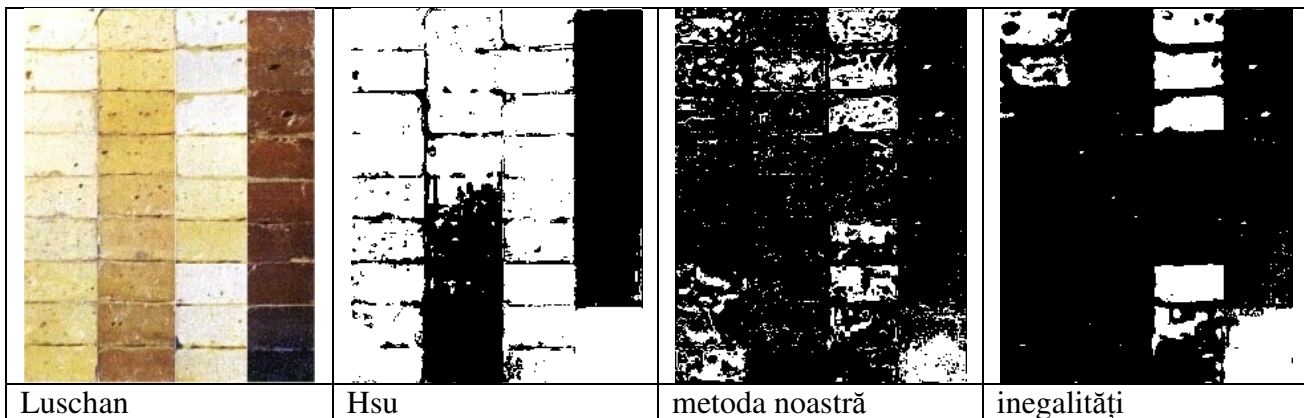


Fig. IV.3.2\_5 De la stânga la dreapta și de sus în jos: harta antropologică a culorilor după von Luschan (‘the skin color chart of Felix von Luschan’), SD conform [Hsu et al., 2002], conform metodei noastre și conform metodei care se bazează pe inegalități.

Se confirmă că metoda noastră pare să aibă cea mai bună ‘acoperire’ – prin prelucrări morfologice cu mici elemente structurante putem acoperi aproape integral zonele de piele. Astfel, putem neglija problema fals negativelor.

Putem spune că subclasarea / prototipizarea în interiorul clasei ‘Skin Locus’ (SL) o facem prin partiționarea SL., cu observația că mulțimile care formează partiția nu sunt disjuncte – discurile  $D_i$  (discuri descrise în secțiunea anterioară, care definesc prototipurile) se suprapun într-o anumită măsură. În fapt, e vorba de o **partiție fuzzy** – în planul de cromaticitate ab; în interiorul SL putem acorda grade de apartenență la fiecare disc / prototip  $D_i$  pentru un pixel.

Algoritm bazat pe variația specifică de luminanță pentru fiecare prototip al pielii

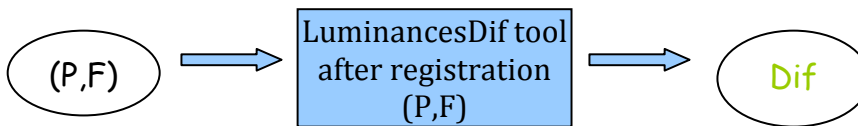
Redăm metoda folosită conform observațiilor de mai sus:

**Etapa I.** Detectarea prototipurilor de piele  $P_i$  în fotografia obținută cu ‘flash’ (F):

Notăm  $R_i$  = regiunea corespunzătoare prototipului  $P_i$ , fiecare pixel din F va fi etichetat cu ‘ $P_i$ ’, adică prototipul / discul  $D_i$  cu centrul cel mai apropiat de pixel în planul cromaticității ab:



**Etapa II.** Calculul  $Dif = F - P$ . Pentru acest calcul trebuie să aliniem (‘registration’) F și P.



**Etapa III.** Determinarea pragurilor pentru fiecare prototip  $P_i$  – pseudocod:

1) Inițializare algoritm de antrenare:

Pentru fiecare prototip de piele ( $P_i$ )

- inițializează pragurile de luminanță - inferior (t) și superior (T) - la valorile maxim, respectiv minim posibil.

2) Algoritm de antrenare pentru determinarea pragurilor de luminanță pentru fiecare prototip:

- pentru fiecare pereche  $\langle P, F \rangle$  (corespondență ‘preview’ – fotografie cu ‘flash’)

- *interestRegion* (vezi observația de după prezentarea acestui algoritm) = în interiorul (măștii)  $Z1$  marchează manual zonele de piele

- pentru fiecare pixel  $\in$  *interestRegion* modifică pragurile astfel:

$Fie P_i$  eticheta de prototip setată în cadrul etapei I și fie  $L$  = luminanța aceluiași pixel în *Dif*.

if (  $L < P_i.t$  )  $P_i.t = L$ ; //  $P_i.t$  și  $P_i.T$  se manifestă ca “praguri regionale”,

if (  $L > P_i.T$  )  $P_i.T = L$ ; // ca măsuri (praguri) ale prototipului  $P_i$

După rularea acestui algoritm vom avea pragurile pentru fiecare prototip:  $[P_i.t, P_i.T]$ . Acesta e algoritmul de antrenare. În cadrul fiecărui prototip  $P_i$ , metoda de clasificare va selecta pixelii cu:

Criteriu: pixel chromaticity  $\in$  skin (prototype  $P_i$ ) AND Dif (pixel)  $\in [P_i.t, P_i.T]$

Determinarea interestRegion nu trebuie să fie exactă ca să contribuie la creșterea selectivității metodei (de SD în cazul nostru). Mai mult, prin faptul că folosim masca  $Z1$ , ne asigurăm că selectivitatea metodei nu poate să scadă. Altfel spus, **creșterea selectivității metodei e tolerantă la defecte.**

Eliminarea zgomotelor și alinierea imaginii obținută fără ‘flash’ cu imaginea obținută cu ‘flash’

Având în vedere că imaginea fără ‘flash’ (P) e, mai ales în cazul nostru, de calitate inferioară (să nu uităm că asupra ‘preview’-ului nu se aplică corecțiile, cum ar fi ‘white balace’, care se aplică pentru o fotografie normală), mai întâi vom încerca să eliminăm cât mai mult din zgomotul ce vine cu această imagine. Urmărim să facem această imagine cât mai “netedă”, mai comparabilă cu imaginea obținută cu ‘flash’ (F):



Fig. IV.3.3\_1 Înainte și după eliminarea zgomotelor pentru imaginea ‘preview’ în scară de gri (luminanță). Deși avem o dimensiune mică a imaginii, se observă eliminarea zgomotului.

Am testat pe sute de poze făcute în condiții cât mai diferite (orientare, distanță, iluminare, indoor etc). Zgomotul apare ca fiind mai degrabă uniform decât neuniform, fără a manifesta însă o regularitate clară. Apar frecvent în P, pentru obiectele cu o reflectivitate mai mare, ‘hot pixels’, adică pixeli cu exces de lumiananță. Chiar dacă după eliminarea zgomotelor nu putem spune că din P au fost eliminați toți ‘hot pixels’, și putem pierde unele detalii fine din P, am remarcat că prin eliminarea zgomotelor, în zonele cu piele, obținem o “netezime” relativ bună. Eliminarea zgomotelor: aducem P la o dimensiune comună cu F (pentru a putea face diferența lor) și aplicăm un filtru de rang pornind de la datele ce ne

sunt oferite de aparatul foto: dintr-o secvență de patru valori am ales-o pe a doua (am **observat vizual**) că astfel imaginea obținută e mai curată, decât dacă aș lua alt index - din cei patru - sau media între cele patru valori.

Desigur, există și alte metode de eliminare a zgomotelor. Mai general, se pot elimina și alte artefacte – e.g., în afara ochilor roșii, mai pot surveni ca artefacte umbre și specularități - care apar odată cu ‘flash’-ul, pornind de la perechea <P,F>. În esență, se poate spune că aceste metode fac o estimare a imaginii care ar fi trebuit achiziționată prin camera foto, în condițiile unor condiții bune pentru fotografiat. Estimarea se face pe baza imaginilor P și F ale aceleiași scene, ideea fiind că P ne dă informații despre ambient, iar F despre detalii. Menționăm unele metode destul de populare: aceea a *filtrului bilateral* destinat să producă o imagine ‘large scale’ a imaginii pe care o prelucrează prin medierea pixelilor învecinați și cu luminanțe similare – vezi [Eisemann et al., 2004] și [Georg Petschnigg et al., 2004], aceea a *tehnichilor de gradient* – vezi [Agrawal et al., 2005], aceea a *estimării reflectanței* – vezi [J. M. DiCarlo et al., 2001]. Am încercat să ne inspirăm din aceste metode, dar ele prezintă un cost suplimentar prea mare pentru un firmware (presupunem că dorim să facem Red Eyes Reduction ‘in camera’) și, mai ales, nu oferă un suport satisfăcător pentru calibrare. Pentru a alinia P cu F, am încercat să aplicăm un model “factorial” pentru expunere:  $E = k g T ISO$ , unde k e o constantă dependentă de cameră, g e câștigul (‘gain’) ‘CCD’, T e timpul de expunere și ISO caracterizează sensibilitatea camerei. Dar această formulă nu s-a dovedit realistă.

O concluzie generală asupra posibilității de a compara” P cu F ni s-a impus: comparația a două imagini achiziționate în condiții atât de diferite e dificilă. Am exclus problema care ar apărea datorită nealinierei spațiale a celor două imagini, deși e o problemă care poate apare și poate împiedica creșterea selectivității detecției zonelor de piele - exemplu:

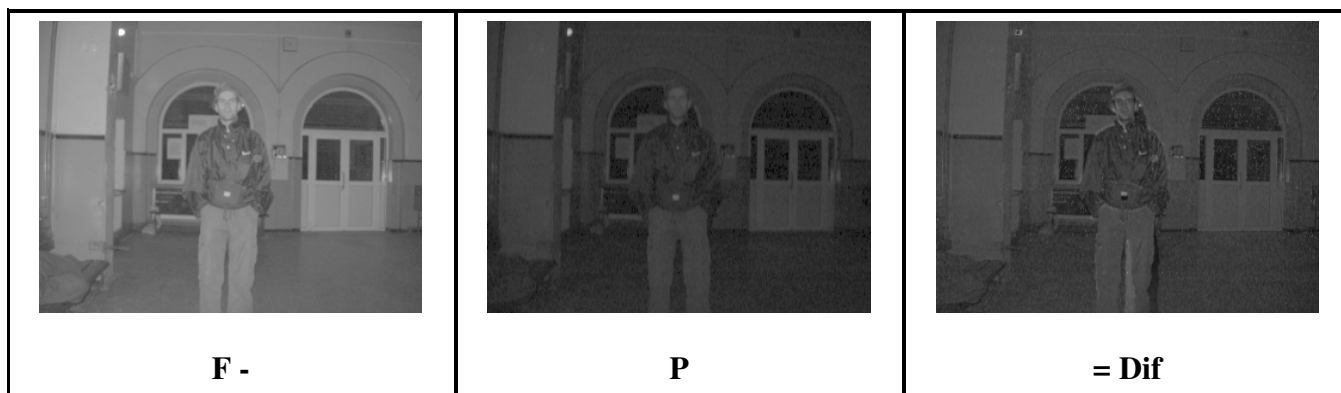


Fig. IV.3.3\_2 Mișcarea persoanei fotografiate în momentul fotografierii cu ‘flash’ (F) față de momentul achiziției imaginii scenei fără ‘flash’ (P, care, astfel, apare “întunecată”) poate împiedica creșterea selectivității detecției zonelor de piele / față prin apariția unor neregularități suplimentare.

Vom presupune că problema mișcării scenei sau în scenă nu va împiedica alinierea spațială a lui P cu F, aliniere pe care o facem prin procedeul de *corelare a fazei* (‘*phase correlation*’) în domeniul frecvență. Dacă această diferență e prea mare, pur și simplu în algoritmul de mai sus nu mai punem și condiția  $AND Dif (pixel) \in [Pi.t, Pi.T]$ , nemaiputându-ne permite creșterea selectivității metodei de

SD. Același lucru îl vom face și dacă P e o imagine “prea întunecată” în care nu distingem forme, contururi, adică luminozitatea globală a lui P e sub un anumit prag.

În plus, ‘Flash’-ul introduce niște umbre diferite de cele ale scenei inițiale - vom exemplifica în figura de mai jos cu o scenă tipică, des întâlnită, în care lumina vine din lateral. Excludem și această problemă având posibilitatea să achiziționăm un ‘preview’ cu ‘preflash’. ‘Preflash’-ul e un mic ‘flash’, care, anunțând din timp ‘flash’-ul, dă mai multe șanse pupilei oculare de a se închide înainte de a se produce ‘flash’-ul, prevenind astfel, în mod hardware, apariția ochilor roșii. Cum ‘preflash’-ul iluminează scena din același “punct” din care o face și ‘flash’-ul, nu se mai pune problema umbrelor modificate între P și F. În plus, dacă e suficient de puternic, ‘preflash’-ul rezolvă P-urilor foarte întunecate și, deci, cu mult zgomot. Facem toate simplificările menționate mai sus pentru a avea șanse reale de a crește selectivitatea metodei noastre de SD. Comparația lui P cu F se face în “condiții grele”, când pare că singura comparație care se poate face e aceea referitoare la luminanță, în condițiile pe care le-am specificat. Chiar și așa, comparația lui P cu F nu pare să aducă o condiție de discriminare foarte puternică – în figura următoare se observă dependența  $\Delta I$  - a lungimii intervalului de variație luminanței în Dif pentru zonele cu fețe - de mediul ambiant și direcția ‘flash’-ului, pentru fețe ale unor persoane de rase diferite, dar și asemănătoare:

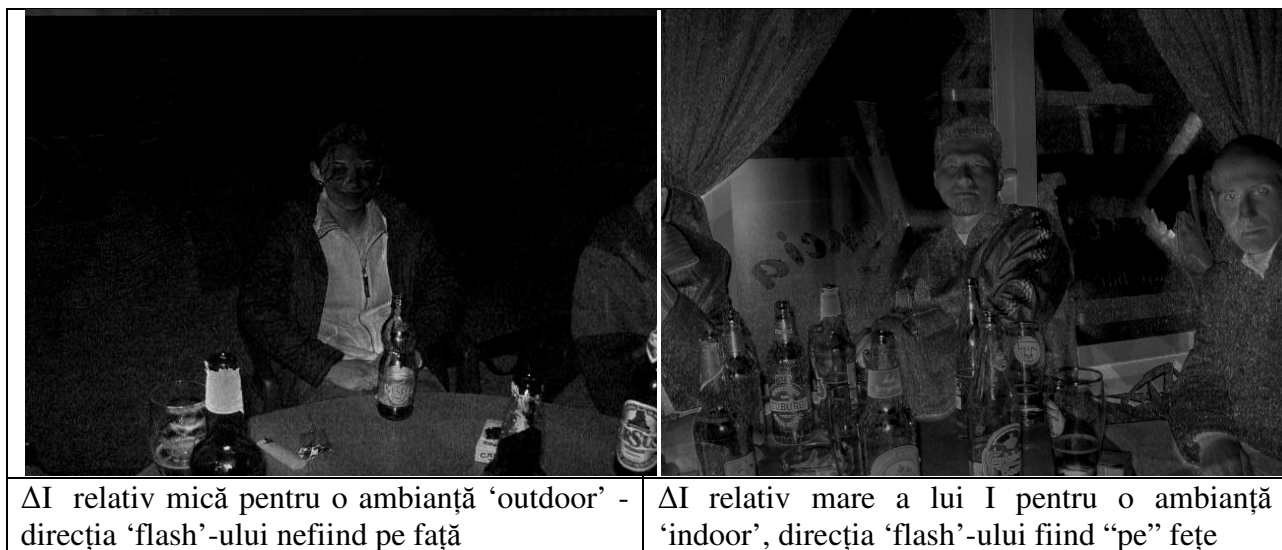


Fig. IV.3.3\_4 Variația luminanței în Dif pentru zonele cu fețe

Am încercat să folosim un alt criteriu de comparație: în loc de diferența în luminanță Dif, am calculat rapoarte, am încercat să deducem diferențe adaptive pentru a elimina neregularități de genul obținerii unei valori mai mari în P decât în F, pentru același pixel. Nu am găsit o metodă care să acopere toate cazurile - aceste neregularități, inevitabile (am arătat anterior), pot fi detectate pentru a fi “sărite”.

## Discuție în urma experimentelor

### Robustețea metodei



Fig. IV.3.4\_2 Pentru acest Dif dacă, în etapa III a algoritmului, marcăm și zonele de față “înnegrite” în Dif vom afecta selectivitatea optimă (din punct de vedere al metodei noastre) a metodei, prin afectarea intervalelelor de luminanță  $[Pi.t, Pi.T]$ , calculate de algoritm pentru fiecare prototip (“înnegrirea” corespunde unui  $Pi.t$  mai mic, iar “albirea” corespunde unui  $Pi.T$  mai mare), ca și condiție suplimentară de selecție, aceste intervale ajungând să fie “prea” largi. Totuși selectivitatea va crește. Dar, precum am spus, metaeuristica noastră are un succes cu atât mai mare cu cât ne asigurăm că lucrăm cu un Dif care să nu prezinte neregularități prea mari –

incluând aici: zgomotele din P, umbrele (pe fețe) din P, deplasările prea mari ale camerei sau în scena fotografiată, între momentul achiziției lui F și momentul achiziției lui P etc. Succesul ar fi și mai mare dacă am reuși să obținem o partiționare cât mai clară (fie ea fuzzy) a SL și dacă am realiza o anumită taxonomie a scenei după: tipul de ambient în care se află scena (înăuntru sau afară), ce fel de iluminare e folosită și toată gama menționată sub numele de principiul specializării și al calibrării în subcapitolul anterior. Ilustrăm un exemplu de reducere a fals pozitivelor, chiar în condițiile în care nu ne-am luat precauțiile menționate la subsecțiunea anterioară:

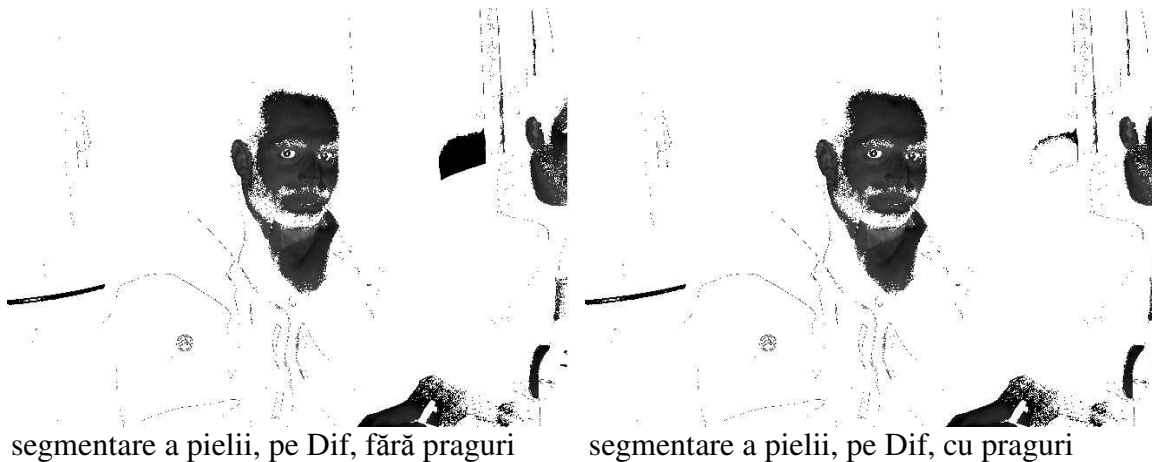


Fig. IV.3.4\_4 O situație în care aplicarea pragurilor pentru luminanțele prototipurilor are ca efect restrângerea selecției zonelor de piele – observăm că porțiunea dintre cele două fețe a fost scoasă din selecție.



Fig. IV.3.4\_5 Fotografia cu ‘flash’ ne confirmă că această porțiune menționată la Fig. IV.3.4\_4 putea fi luată relativ ușor drept zonă de piele

### Concluzii

În această subcapitol a dăm o *metaeuristică* de reducere a fals pozitivelor (FP) prin folosirea subclasării și a marcajelor *perceptuale*, adică a marcajelor de clasificare furnizate de expertul uman (putem spune că metoda de Skin Detection e antrenată conform observațiilor expertului uman). Subclasarea presupune partiționarea în subclase a claselor furnizate de o metodă oarecare de clasificare inițială (care constituie intrarea metaheuristicii) și găsirea de condiții suplimentare de rafinare a clasificării în interiorul fiecărei clase. Dacă numim *prototipuri* ale unei clase, subclasele obținute în interiorul fiecărei clase prin partiționare, creșterea selectivității unei clasificări, adică reducerea FP, e dată de relația:  $zona\_de\_interes = zona\_prototipurilor \cap zona\_marcajelor$  (\*)

Abordarea e “pixel based” datorită avantajelor oferite de acest tip de metode (vezi secțiunea IV.2.1.). Căutăm să selectăm cât mai precis ‘Skin Pixels’. Presupunem că într-o etapă anterioară, de antrenare în vederea recunoașterii SP (vezi IV.2.), am colectat (dintr-o bază de date cu imagini) SP și avem o **reprezentare și, poate, chiar o partiționare a ‘Skin Locus’-ului.**

Subclasarea corespunde partiționării SL. Pentru că ne propunem să partiționăm Skin Locus-ul, alegem ca spațiu de reprezentare a culorilor un spațiu de culoare perceptual uniform – desigur, am ales Lab și datorită generalității sale (capacității sale de reprezentare a culorilor) și decorelării în cel mai înalt grad a luminanței de cromaticitate. În plus, SD făcându-se într-un spațiu perceptual uniform, se justifică definirea de subclase (prototipuri) sub forma unor discuri Di. Di nu sunt disjuncte: se suprapun într-o anumită măsură – e vorba de o *partiție fuzzy*. Pentru fiecare prototip Pi, metoda va selecta pixelii cu:

$$cromaticitate\ pixel \in\ prototip\ Pi) \quad \text{ȘI} \quad Dif(pixel) \in [Pi.t, Pi.T] \quad (**)$$

, unde pragurile inferior și respectiv superior Pi.t și Pi.T se determină printr-un algoritm iterativ de antrenare care rulează peste imagini cât mai “apropiate” de cele pentru care ne propunem SD. Acest tip

de **antrenare adaptivă** e foarte important pentru creșterea selectivității metodei. Portarea SD la altă categorie de imagini va însemna doar rerularea acestei preprocesări.

Un alt avantaj important al metaheuristicii noastre e că, practic, e o unealtă care poate fi foarte utilă pentru cercetare. **Se poate demonstra astfel, în mod practic și foarte benefic, utopia unor probleme de clasificare** – așa cum teorema de separare pentru seturi fuzzy convexe (vezi și II.2.3.2. “Mulțimi și partiții fuzzy”) vorbește de un grad maxim de separabilitate între două mulțimi fuzzy (putem proiecta prototipurile  $P_i$  ca mulțimi fuzzy).

Idei de îmbunătățire a metaheuristicii:

- creșterea granularității prototipurilor
- creșterea “naturalității” / separabilității prototipurilor – evident, subclasarea pe care am aplicat-o (prin stabilirea discurilor  $D_i$  care partiționează SL) poate fi rafinată, “clusterizând”
- prin prelucrări morfologice putem elimina “accidentele” (situațiile necorespunzătoare) din faza de antrenare (în faza de procesare presupunem că nu e timp pentru asemenea precauții)
- proiectarea prototipurilor  $P_i$  ca mulțimi fuzzy – dacă partiționarea în prototipuri se face într-o manieră fuzzy (ele se suprapun), totuși ele nu au fost gândite ca mulțimi fuzzy. Aceasta ar fi dus la un control mai bun al prototipurilor și, deci, la o creștere suplimentară a selectivității clasificării

## **VȘ Sinopsis al contribuțiilor originale. Studiul viitor**

### *V.1. Vedere sinoptică a contribuțiilor noastre*

În aceasta lucrare, teoria a fost abordată din punct de vedere pragmatic – chiar și în cazul explicațiilor care s-au dat pentru unele noțiuni, unele matematice, procedee folosite des în practică, dar mai rar “deconstruite” pentru a putea fi înțelese în profunzime, aplicate / adaptate creativ de către cei care caută rezolvări unor probleme nedeterminate concrete de Clasificare a Imaginilor (CI), de către cei care caută noi criterii de clasificare pentru rezolvarea unor asemenea probleme. Rezumăm în următoarele două secțiuni (V.1.1. și V.1.2.) contribuțiile noastre de sinteză / observație teoretică, contribuții la care

am ajuns în urma investigațiilor cu ocazia acestor probleme, investigații care ne-au dus cercetarea teoretică spre domeniul fuzzy.

### V.1.1. Contribuții teoretice (per capitole)

În această lucrare încadrăm în mod original problema Clasificării în Imagini într-o viziune unitară, per domenii: Domeniul Spațial, Domeniul Probabilistic, Domeniul Frecvență și Domeniul Fuzzy:. Secțiunea II.1.5., “O viziune abstractă asupra recunoașterii formelor, în jurul conceptului de corelare” întărește această legătură între domenii, în jurul conceptului de corelare, concept care, din punct de vedere matematic, se bazează pe produsul intern într-un spațiu vectorial. Multe alte informații sunt integrate în această viziune, având astfel o lucrare legată, coerentă.

**Iș Încadrarea problemei CI și descrierea ei în Domeniul Spațial și în Domeniul Probabilistic – capitol introductiv, în care arătăm limitarea teoretică a unor tehnici clasice, e.g., a rețelelor neuronale, în ce privește rezolvarea unor probleme concrete de CI. Încheiem acest capitol remarcând că **domeniul probabilistic, deși poate “veni” cu informații în plus față de cel spațial, nu încurajează folosirea intuițiilor, a adaptărilor pentru o rezolvare cât mai adecvată a unor probleme mai complexe de clasificare.****

IIș “Separabilitatea claselor” – clasificarea fiind în esență o problemă de separare

- II.1.3. Eșantionarea semnalului: precizia necesară și intercondiționarea reprezentărilor în Domeniul Frecvență și în Domeniul Frecvență-Timp – imprecizia ca aliasing (II.1.3.1. Precizia necesară reconstruirii semnalului și efectul de aliasing) și ca principiu al nedeterminării (II.1.3.2. Intercondiționarea rezoluției reprezentărilor în Frecvență și în Frecvență-Timp: principiul nedeterminării). **Fig.II.1.3\_12 prezintă o imagine sintetică originală. Secțiunea II.1.5., “O viziune abstractă asupra recunoașterii formelor, în jurul conceptului de corelare”, propune o sinteză asupra problemei clasificării, în domeniile spațial, frecvență și probabilistic, problemă care poate fi văzută ca una de corelare / de-corelare sau de descompunere.**

- secțiunea II.2.2., “Domeniul posibilului: o fuziune fuzzy a diverselor tipuri de cunoștințe” dă seama de capacitatea de integrare a informațiilor precise sau imprecise, cu diverse grade de credibilitate (probabilitate) de către domeniul fuzzy, un domeniul față de care domeniul probabilistic e un caz particular. Domeniu care poate fi văzut și ca “loc de trecere” între știința calculatoarelor și alte discipline de cunoaștere, cum ar fi cele sociale. Ceea ce dă **posibilitatea extinderii raționamentului fuzzy (aproximativ) în științele “inexacte”** e cuvântul / termenul lexical / variabila lingvistică. **În secțiunea II.2.4., <<“Fuzzy logic = calculul cu cuvinte” [Zadeh, 1996]>>, sintetizăm concluziile pe care le-am desprins din literatura de specialitate pe această temă, propunând ca legătura între**

**științele “exacte”** (fizica a arătat până unde poate merge acest “exact” – vezi principiul incertitudinii al lui Heisenberg, enunțat încă din 1927) **și cele “inexacte” să se facă în jurul criteriului de proximitate, pe baza claselor fuzzy.** În esență facem următoarele observații:

- II.2.4.1. Termenul lexical și abilitatea de a opera cu informații bazate pe percepții sau fler – teoria fuzzy are o legătură profundă cu semantica

- II.2.4.2. Scale fuzzy – e interesant de observat că **scațele metrice ordinale** au fost introduse în contextul psihofizicii (deci, iată, putem vorbi de o contribuție la teoria fuzzy din partea psihologiei, o disciplină de care, acum un secol, științele exacte s-ar fi “ferit”) ca metodă de a măsura date multidimensionale (în capitolul III vom vorbi de procesarea multi-canal, cu aplicație concretă în procesarea imaginilor color, culoarea fiind, în fapt, o percepție psiho-fizică). Folosirea unei scale metrice fuzzy e o îmbunătățire dacă se ia în considerare o relație de proximitate pe setul de reprezentare, relație de proximitate care are acoperire într-o relație de proximitate pe setul reprezentat.

- în secțiunea II.2.5. **propunem conceptul de PROTOTIP pe post de clasă fuzzy:** “sâmbure” în jurul căruia se centrează clasa fuzzy conform unei distribuții de posibilitate. Acest concept, pe care l-am utilizat efectiv în contribuția practică IV.3, “Folosirea subclasării și a marcajelor pentru reducerea fals pozitivelor”, apare ca o generalizare și ca o tratare mai adecvată a claselor fuzzy.

III§ Domeniul multispectral al imaginilor color – deși ne vom referi numai la cazul domeniului multispectral al culorilor, discuția noastră se poate generaliza în ce privește procesarea informației care ne vine pe mai multe canale. Aici arătăm cum ceea ce poate integra - matematic, pentru descrierea sensibilităților de culoare, folosim produsul scalar - aceste informații multi-canal este percepția, percepție care stă la baza teoriei fuzzy, dar și a colorimetriei - spațiile de culoare uniform perceptuale au fost concepute plecând de la conceptul / percepția de proximitate (vezi Just Noticeable Difference).

În subcapitolul III.3., “Caracterul perceptual al culorii”, **remarcăm că acolo unde filtrarea liniară nu a avut succes (e.g, pentru procesarea multi-canal), poate avea succes o abordare fuzzy prin caracterul său adaptiv, trunchiat (vezi operațiile pe mulțimi fuzzy), asemănător unei logici a percepțiilor.**

CI, prin chiar obiectul ei, “ilustrează”, poate, cel mai bine multe dintre conceptele clasificării. Așa cum am spus și în rezumat-ul lucrării, discuția noastră se aplică și altor probleme de clasificare, mai ales în ceea ce privește metodele bazate pe elemente componente – care corespund în cazul CI metodelor de clasificare bazate pe pixeli.

## V.1.2. Contribuții practice

Rezumăm contribuțiile noastre practice, contribuții ivite, ca și cele teoretice, în urma confruntării cu probleme nedeterminate concrete de CI:

IV.1 Transformări eficiente între spații de culoare – **am proiectat și implementat o metodă de interpolare cu un grad ridicat de generalitate, superioară ca acuratețe și viteză de execuție.**

IV.2 “Definirea formei spațiului unei probleme de recunoaștere nedeterminate și modelarea acestuia. Cazul ‘skin detection’” – **am proiectat și implementat o metaeuristică adaptivă și dinamică în vederea rulării unor algoritmi ‘run-length’, superiori ca acuratețe și viteză de execuție.** Prin folosirea de **informație probabilistică** ne putem extinde în domeniul impreciziei, probabilistic / fuzzy, putem prelua și furniza cunoștințe probabilistice / fuzzy.

IV.3 “Folosirea subclasării și a marcajelor pentru reducerea fals pozitivelor. Cazul ‘skin detection’” – **am proiectat și implementat o metaeuristică adaptivă și dinamică a cărei noutate se bazează pe folosirea subclasării și a marcajelor perceptuale.** Conform teoriei fuzzy, **se poate demonstra în mod practic utopia unor probleme de clasificare, ceea ce poate scuti de foarte costisitoare investigații** – vezi și unele dintre concluziile din finalul secțiunii IV.3.5, între care și ideea de a găsi un criteriu de subclasificare care să crească (față de ‘Dif’ și partiționarea (fuzzy) a ‘Skin Locus’-ului în discuri de raze de valoare constantă) separabilitatea claselor. Din punct de vedere teoretic, teorema de separare pentru seturi fuzzy convexe, teoremă care încheie novatorul articol [Zadeh, 1965], este, și ea, grăitoare în acest sens.

## V.2. Studiul viitor

Într-un articolul recapitulativ, [Vertan & Buzuloiu, 2000], s-a observat că pentru imagini gri au avut succes filtrările adaptive și, deci, neliniare. Pentru imagini color, extinderea acestor filtre prin prelucrarea individuală a componentelor de culoare s-a dovedit inadecvată în ce privește păstrarea topologiei. Ori păstrarea topologiei e un criteriu important de validare a reușitei multor procesări de imagini – de fapt, și noi am aplicat un raționament topologic când am configurat ‘Skin Locus’-ul, considerând că probabilitatea unui ‘Skin Pixel’ crește spre centrul SL – altfel, SL ar fi contrar percepției noastre. **Susținem: criteriul păstrării topologiei e unul determinant în găsirea soluțiilor majorității problemelor de clasificare nedeterminate în imagini (putem avea și alte reprezentări decât imagini digitale). Raționamentul fuzzy, dovedindu-se până în zilele noastre cadrul teoretic**

cel mai “integrator” de abordare a acestora, remarcăm caracterul său intrinsec topologic (fundamentală fiind relația de proximitate). Ne propunem să continuăm investigația CI pe partea de procesare neliniară a semnalelor (multi-canal), apelând la fundamentele teoriei fuzzy, în direcția prelucrărilor morfologice. Din punct de vedere morfologic, conceptul de PROTOTIP, pe care l-am propus, poate fi văzut ca un schelet (‘skeleton’) / nucleu de clasă (fuzzy) – noțiunea de ‘skeleton’ e prezentată în lucrare la subsecțiunea I.4.1.2., “Skeletoane”, în cadrul secțiunii I.4.1., “Simplificarea formelor”. Astfel, “fuzificarea” ar corespunde extrapolării de la un ‘skeleton’, iar “defuzificarea”, care, în general, ia forma mapării unui vector într-un scalar și e asociată transformării raționamentelor umane calitative în expresii numerice (cantitative), ar corespunde calculării ‘skeleton’-ului. În cazul GIS, operația de generalizare poate corespunde calculării ‘skeleton’-ului cu păstrarea topologiei [Gold & Shankar, 2001]. În [Georgescu 2, 2009] arătăm cum, în vederea reconstrucției 3D din secțiuni paralele, pentru o partiționare cât mai corectă în cazul ‘Double Branching Problem’ e nevoie să găsim un skeleton “cât mai adecvat” între cele două contururi în discuție - evident, problema având mai multe soluții, alegem soluția care se mulează cel mai bine criteriilor soluției (se știe că pentru o decizie multicriterială putem folosi o integrală fuzzy), dar și de acuratețe și viteză ale problemei concrete. În general, raționamentul fuzzy, chiar dacă presupune îndeplinirea unor condiții topologice, se poate adapta problemei în cauză.

Pe de altă parte, în subsecțiunea II.2.2.2., <<Teoria fuzzy și bazele de date sau de cunoștințe (sistemele expert) cu informații incerte / “granulate”>>, menționăm exemplarul articol [Zadeh, 1986], în care se arată că dilatarea și eroziunea, sau închiderea și deschiderea, ca operațiuni morfologice de bază, corespund necesității și posibilității, având, în plus, același tip de dualitate. Necesitatea și posibilitatea fiind conceptele de bază ale teoriei posibilității, iar această teorie, după cum am evidențiat în subsecțiunea II.2.2.1., “Teoria fuzzy și teoria posibilităților”, corespunzând teoriei fuzzy.

Teorema de separare pentru seturi fuzzy convexe [Zadeh, 1965] e un alt indiciu în favoarea studiului CI în privința prelucrărilor morfologice care pot ajuta clasificarea. În general, clasificarea se face în clase / mulțimi **convexe** – vezi și reprezentarea Skin Locus-ului de la contribuția practică IV.2.

În subsecțiunea II.2.3.2., “Mulțimi și partiții fuzzy”, am accentuat importanța teoremei de reprezentare Negoită-Ralescu a seturilor fuzzy ca tăieturi  $\alpha$ . E o teoremă care ne poate furniza o intuiție clară asupra unei posibile aproximări a unei mulțimi fuzzy: “ $\alpha$  din ‘ $\alpha$ -cut’ poate fi văzut ca un **nivel de aproximare** al setului sau numărului fuzzy” - “pe baza ‘ $\alpha$ -cut’ se definesc operații între numere fuzzy”.

Ordonarea cât mai “inspirată” a numerelor fuzzy se poate face în lumina unor intuiții geometrice [Rao & Shankar, 2011]. În [Vertan, 2009], se vorbește de “coordonate paralele” – fiecare culoare are asociat un poligon, calculul distanțelor între culori reducându-se la calculul intersecțiilor acestor poligoane; și aici  $\alpha$  e pe post de grad de aproximare.

Toate acestea sunt, credem, suficiente motivații pentru studiul utilității prelucrărilor morfologice pentru problemele de CI, sau, mai general spus, pentru înțelegerea imaginilor. Mai credem că, și am

susținut în lucrarea de față această aserțiune prin multiple argumente, cadrul teoretic mare care permite abordarea riguroasă a problemelor complexe de CI e ceea ce am numit “domeniul fuzzy”.

## BIBLIOGRAFIE

Articole, cărți

[Agrawal et al., 2005] A. Agrawal, R. Raskar, S. K. Nayar and Y. Li, “Removing Photography Artifacts using Gradient Projection and Flash-Exposure Sampling”, Proceedings of ACM SIGGRAPH 2005, Vol. 24, Issue 3, August, New York, USA.

[Albiol, Torres & Delp 2001] A. Albiol, L. Torres & E.J.Delp, “Optimum color spaces for skin detection, International Conference in Image Processing, vol.1, pp. 122-124.

[Balaș & Balaș, 2007] V. E. Balaș & M.M. Balaș, “Postmodernism and control engineering”, EUSFLAT.

[Benoit & Foulloy, 2002] E. Benoit & L. Foulloy, "Fuzzy nominal scales", Proc. of IMEKO TC7 Symposium, Poland, Cracow, pp. 21-25.

[Benoit et al., 1994] E.Benoit, G. Mauris and L. Foulloy, "A fuzzy colour sensor", Proc. of XIII IMEKO World congress, Torino, pp. 1015-1020.

[Benoit & Foulloy, 2003] E. Benoit & L. Foulloy, “Towards fuzzy nominal scales Measurement”, ed. Amsterdam: Elsevier, ISSN 0263-2241, vol. 34(1) (Fundamental of Measurement), pp. 49-55.

[Berka, 1983] Berka K., 1983, “Scales of measurement; a critical analysis of the concept of scales and of their function in the theory of measurement”, Language, Logic and Method, ed. Dordrecht: D. Reidel Publishing Company, pp.1-73, apud. [Benoit, 2010].

[Benoit, 2010] E. Benoit , “An ordinal metrical scale built on a fuzzy nominal scale”, 13th IMEKO TC1-TC7 Joint Symposium IOP Publishing, Journal of Physics: Conference Series 238.

[Beu, 1999] T.A.Be u , “Calcul numeric in C”, capitolul “Aproximarea funcțiilor tabelate”, p. 202-242, Ed. Microinformatica, Cluj-Napoca

[Blaga, 1937] L. Blaga, Trilogia culturii. III, Geneza metaforei și sensul culturii, p. 31-36, Ed. Humanitas, 1994

[Bulea 2003] M. Bulea, “Prelucrarea imaginilor și recunoașterea formelor (teorie și aplicații)”, ed. Academiei Române, București

[Chang & Zhiqiang 2007] Chang, F & Zhiqiang, T, “A region-based skin color detection algorithm”, Proceedings of the 11-th Pacific-Asia conference on Advances in knowledge discovery and data mining, Nanjing, China, pp. 417-424.

[Coombs, 1951] C. H. Coombs, “Mathematical models in psychological scaling”, Journal of the American Statistical Association, Alexandria, American Statistical Association, vol. 46 (256) pp. 480-489, apud. [Benoit, 2010].

[Carron & Lambert, 1996] T. Carron & P.Lambert, ”Symbolic Fusion of Hue-Chroma-Intensity Features for Region Segmentation”, proc. IEEE Conference on Image Processing, Lausanne, Switzerland, apud [Vertan & Buzuloiu, 2000].

[Dimitrov, 2002] V. Dimitrov, “Introduction to fuzziology”, Fuzzy logic: a framework for the new millennium, Studies in fuzzyness and soft computing, vol. 81, Eds. Vladimir Dimitrov & Victor Korotkich, ed. Physica-Verlag Heidelberg New York.

[de Dios & Garcia 2003] de Dios, J. and Garcia N., “Face detection based on a new color space YCgCr”, Proceedings of IEEE International Conference on Image Processing, Vol. 3, 14-17 Sept., IEEE Computer Society, Barcelona, Spain, pages 909-912.

[Do & Vetterli 2005] M. N. Do & M. Vetterli, “The contourlet transform: an efficient directional multiresolution image representation”, proc. IEEE Transactions Image on Processing, vol. 14 (12), pp.2091-2106. [Donoho, 1999] D. L. Donoho, ”Wedgelets: nearly-minimax estimation of edges”, Ann. Statist, pp.859—897.

[Donoho et al., 2001] D. L. Donoho, X. Huo, I. Jermyn, P. Jones, G. Lerman, O. Levi and F. Natterer, ”Beamlets and Multiscale Image Analysis”, proc. Multiscale and Multiresolution Methods, ed. Springer, pp.149—196 .

[Eisemann & Durand 2004] Elmar Eisemann and Fredo Durand, “Flash Photography Enhancement via Intrinsic Relighting”, Proceedings of ACM SIGGRAPH 2004, Vol. 23, Issue 3, August, New York, USA.

[Finkelstein, 1975] L. Finkelstein, “Representation by symbol systems as an extension of the concept of measurement”, Kybernetes , vol. 4, pp. 215-223, apud. [Benoit, 2010].

[Flondor, 2003] Paul Flondor, “Dynamics and approximate reasoning”, IAFA, București.

[Foulloy & Benoit, 2006] L. Foulloy & E. Benoit, “Building a class of fuzzy equivalence relations”, Fuzzy Sets and Systems, no. 157, pp. 1417- 1437.

[Georgescu 1, 2009] O.F. Georgescu, “Accuracy and Speed for Device Color Spaces Conversions, CSCS17 at Politehnica University of Bucharest, May 26-29.

[Georgescu 2, 2009] O.F. Georgescu, C.Tudose, C. Odubășteanu and C.Munteanu, “Reconstruction of Three-Dimensional Objects from Parallel Planar Sections - Analysis, Solutions and Implementation of Double Branching Problem”, CSCS17 at Politehnica University of Bucharest, May 26-29.

[Georgescu, 2010] O.F. Georgescu, “Determining the solution space form for an indeterminate pattern recognition problem. The ‘skin detection’ case”, *Annals of DAAM for 2010 & Proceedings*, pp. 1491-1492.

- [Georgescu, 2011] O.F.Georgescu, “Modelling the solution space form for an indeterminate pattern recognition problem. The ‘skin detection’ case”, CSCS18 at Politehnica University of Bucharest, May 24-27.
- [Georgescu & Moldoveanu, 2012] Octavian Georgescu, Florica Moldoveanu, “An abstract view on pattern recognition based on correlation”, accepted for publication in Scientific Bulletin of UPB, C series.
- [Georgescu 2, 2012] O.F.Georgescu, “Reducing false positives by marking and overclassifying”, accepted for publication in Scientific Bulletin of UPB, C series.
- [Grabisch et al., 2000] M. Grabisch, “Fuzzy Integral for Classification and Feature Extraction”, eds. M. Grabisch, T. Murofushi, M. Sugeno, and J. Kacprzyk, Fuzzy Measures and Integrals - Theory and Applications, pp.415-434. Physica Verlag.
- [Gomez, 2000] G. Gomez, “On selecting color components for skin detection”, Proc of ICPR, Volume no 2, pp. 961-964.
- [Gomez & Morales 2002] G. Gomez & E. Morales, “Automatic feature construction and a simple rule induction algorithm for skin detection”, ICML Workshop On Machine Learning in computer Vision, pp. 31-38.
- [Gonzalez et al. 2010] J. Gonzalez, M. Rodriguez, J. Pulido & J. Perez, “Detecting skin in face recognition systems: A colour spaces study”, Digital Signal Processing, Vol. 20, May 2010.
- [Gonzalez, 2008] R.C. Gonzalez, “Digital Image Processing (3rd ed.)”, ed. Pearson Prentice Hall, New Jersey
- [Gold & Shankar, 2001] C. Gold & D. Thibault, “Map generalization by skeleton retraction”, in proc. 20th Int. Cartographic Conf. (ICC 2001), pp. 2072-2081.
- [Hsu et al, 2002] R. Hsu, M. Abdek-Mottaleb, and A.K. Jain, “Face detection in color images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, Issue 5, May, pp.696–706.
- [J. M. DiCarlo et al, 2001] J. M. DiCarlo, F. Xiao & B. A. Wandell, “Illuminating Illumination”, Ninth Color Imaging Conference, pp.27-34.
- [Jones & Rehg, 2002] M.J. Jones and J.M. Rehg, “Statistical color models with application to skin detection”, in International Journal of Computer Vision, (IJCV) 46(1), pp. 81–96.
- [Kakumanu et.al, 2006] P. Kakumanu, S. Makrogiannis and N. Bourbakis, “A survey of skin-color modeling and detection methods”, ITRI/Department of Computer Science and Engineering, Wright State University, Dayton OH 45435, USA.
- [Kolata, 1991] G. Kolata, "Math Problem, Long Baffling, Slow Yields", The New York Times.
- [Kosko, 1991] B. Kosko, “Neural Networks and Fuzzy Systems”, ed. Prentice Hall, Englewood Cliffs, NJ.
- [Kruschke, 2011] John K. Kruschke, "Doing Bayesian Data Analysis. A Tutorial with R and BUGS", ed. Academic Press / Elsevier, 2011

- [N. van Lier, 2002] Hubert N. van Lier, "Land Use Planning: A Key to Sustainable Development", XXX International Symposium Actual Tasks on Agricultural Engineering, Croatia, 2002
- [McNeill & Freiberger, 1993] D. McNeill & P. Freiberger, "Fuzzy Logic", ed. Simon & Schuster, New York, pp. 46-48.
- [Meyer1 & Coifman, 1997] F.G. Meyer1 and R. R. Coifman, "Brushlets: a tool for directional image analysis and Image Compression", Applied and computational harmonic analysis, vol.4, pp.147–187.
- [Moisil, 1975] Grigore Moisil, "Leccións despre logica ratiunamentului nuanat", ed. Științifică și enciclopedică.
- [Negoiță, 2002] C. V. Negoiță, "Postmodernism, Cybernetics and Fuzzy Set Theory", Proc. of IPMU, Annecy, France, pp. 681-686, apud [Balaș & Balaș, 2007].
- [Negoiță, 2006] C. V. Negoiță, "Concert la Carnegie Hall", p. 118, Ed. Paralela 45, 2006
- [Negoiță & Ralescu, 1974] C. V. Negoiță & D. A. Ralescu "Mulțimi vagi și aplicațiile lor", ed. Tehnică.
- [Neumaier, 2003] A. Neumaier, "Fuzzy modeling in terms of surprise", Fuzzy Sets and Systems, Vol. 135, Issue 1, Elsevier North-Holland, Inc. Amsterdam, The Netherlands.
- [Pavlidis, 1982] T. Pavlidis, "Algorithms for Graphics and Image Processing", ed. Computer Science Press, apud M. Bulea, Prelucrarea imaginilor și recunoașterea formelor, ed. Academiei Române, p.15.
- [Pesquet-Popescu and Popescu, 2010] Pesquet-Popescu B. and Popescu S., "Analiza cu undine (wavelets). Fractali – modele stocastice", ed. Politehnica Press
- [Peyré & Mallat, 2005] G. Peyré & S. Mallat, "Surface Compression With Geometric Bandelets", ACM Transactions on Graphics, Vol. 24(3), (Proc. of SIGGRAPH'05), pp.601-608.
- [Petschnigg et al, 2004] G. Petschnigg, M. Agrawala, H. Hoppe, R.Szeliski, M. Cohen and K. Toyama, "Flash Photography Enhancement via Intrinsic Relighting", Proceedings of ACM SIGGRAPH 2004, Vol. 23, Issue 3, August, New York, USA.
- [Rao & Shankar, 2011] P.B. Rao & N. R. Shankar, "Ranking Fuzzy Numbers with a Distance Method using Circumcenter of Centroids and an Index of Modality", Advances in Fuzzy Systems, ed. Hindawi Publishing Corporation.
- [Rosch, 2010] E. Rosch, "Principles of Categorization", Cognition and categorization pp. 27-48, University of California, Berkeley.
- [Rosenberg, 2010] Michael J. Rosenberg, "The Agile Approach to Adaptive Research: Optimizing Efficiency in Clinical Development", ed. Wiley, 2010
- [Shakhatreh & Hayajneh 2010] M. A. Shakhatreh & M. A. Hayajneh, "S-H Fuzzy Partition and Fuzzy Equivalence Relation", Int. J. Contemp., Math. Sciences, Vol. 5 (8), online ed., pp.381 – 392.
- [Starck et al., 2002] J. L. Starck, E. J. Candes and D. L. Donoho, "The curvelet transform for image denoising", proc. IEEE Trans. Image Process, vol. 11 (6), pp.670-684.

[Stein & R. Shakarchi, 2003] E. Stein & R. Shakarchi, *Fourier Analysis: An introduction*, ed. Princeton University Press.

[Sweldens, 1997] W. Sweldens, "The lifting scheme: a construction of second generation wavelets", *Siam J. Math. Anal.*, Vol. 29 (2).

[Terrillon et.al. 2000] J.C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance for the automatic detection of human faces in color images", *Proceedings of the IEEE International conference of Face and Gesture Recognition*, pp.54–61.

[Teodorescu, 2011] H.-N. L. Teodorescu, "On the meaning of approximate reasoning – An unassuming subsidiary to Lotfi Zadeh's paper dedicated to the memory of Grigore Moisil –", *Int. J. of Computers, Communications & Control*, Vol. VI, No. 3, pp. 577-580.

[Vertan, 2009] C. Vertan, "Parallel Coordinates-based Fuzzy Color Distance and its Applications", *IFSA-EUSFLAT*.

[Vertan & Boujemaa, 2000] C. Vertan and N. Boujemaa, "Embedding fuzzy logic in content based image retrieval", *proc. AFIPS, Atlanta, Georgia*, pp. 85-90.

[Vertan & Buzuloiu, 2000] C. Vertan & V. Buzuloiu, "Fuzzy Nonlinear Filtering of Color Images: A Survey", *Fuzzy Technique in Image Processing, Studies in fuzzyness and soft computing*, vol. 81, Eds. Etienne E. Kerre & Mike Nachtgeael, ed. Physica-Verlag Heidelberg New York pp. 248-266.

[Vertan et al., 2000] C. Vertan, N. Boujemaa and V. Buzuloiu, "A fuzzy color credibility approach to color image filtering", *proc. International Conference on Image Processing*, vol. 2, pp. 808-811.

[Vertan & Ciuc, 2007] Constantin Vertan & Mihai Ciuc, "Tehnici fundamentale de prelucrarea și analiza imaginilor", ed. Matrix Rom, București.

[Vezhnevets et al. 2003] V. Vezhnevets, V. Sazonov & A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques", *Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Russia*.

[Zadeh, 1965] L. A. Zadeh, "Fuzzy Sets", *Department of Electrical Engineering and Electronics Research Laboratory, University of California, Berkeley*.

[Zadeh, 1971] L.A. Zadeh, "Quantitative fuzzy semantics", *Information Sciences*, Vol. 3, 1971, pp. 159-176

[Zadeh, 1978] L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems*, Vol. 1, pp.3-28.

[Zadeh, 1979] L. Zadeh, "On the validity of Dempster's rule of combination", *Memo M79/24, Univ. of California, Berkeley, USA*.

[Zadeh, 1984] L. Zadeh, Book review: "A mathematical theory of evidence," *The AI Magazine*, Vol. 5(3), pp. 81-83.

[Zadeh, 1986] L. Zadeh, “A simple view of the Dempster-Shafer Theory of Evidence and its implication for the rule of combination”, The AI Magazine, vol. 7 (2), pp. 85-90.

[Zadeh, 1996] L.A. Zadeh, “Fuzzy logic = computing with words”, Fuzzy Systems, IEEE Transactions on, Vol. 4 Issue 2, pp. 103–111.

[Zadeh, 2002] L.A. Zadeh, "Toward a perception-based theory of probabilistic reasoning", Fuzzy logic: a framework for the new millenium, Studies in fuzzyness and soft computing, vol. 81, Eds. Vladimir Dimitrov & Victor Korotkich, ed. Physica Verlag New York.

[Zemankova & Kandel, 1984] M. Zemankova-Leech & A. Kandel , “Fuzzy Relational Databases—a Key to Expert Systems”, Interdisciplinary Systems Research, Verlag TUV: Rheinland, Koln, p. 178.

[Zheng, Zhang & Wang, 2004] Q.F. Zheng, M.J. Zhang and W.Q. Wang, “ A hybrid approach to detect adult web images”, PCM (2), pp. 609–616.

[Yang & Ahuja, 1999] M.Yang & N. Ahuja, “Detecting human faces in color images”, International Conference in Image Processing, Vol. 1, pp. 127-130.

[Yang & Ahuja, 1999] M. Yang and N. Ahuja, “Gaussian mixture model for human skin color and its applications in image and video databases”, SPIE Proceedings Storage and Retrieval for Image and Video Databases VII, San Jose, USA, pp.458–466.

Link-uri

[AdobeRGB1998] <http://www.adobe.com/digitalimag/pdfs/AdobeRGB1998.pdf>

[ICC] <http://www.color.org>

[if fl] <http://egovbus.net/rdl/articole/No1Art50.pdf>

[cl cr] <http://www.exo.net/~pauld/workshops/Color2011.html>

[crb grn] [http://www.hydrop.pub.ro/vn\\_cap361.pdf](http://www.hydrop.pub.ro/vn_cap361.pdf)

[DSmT] <http://fs.gallup.unm.edu/DSmT.htm>

[eig img] [http://www.stanford.edu/class/ee368/Handouts/10-EigenImagesMay25\\_27.pdf](http://www.stanford.edu/class/ee368/Handouts/10-EigenImagesMay25_27.pdf)

[fz ap] <http://www.scribube.com/tehnica-mecanica/Fuzzy-sets102101833.php>

[fz hedge] <http://statpac.org/walonick/systems-theory.htm>.

[fz lz] [www.nvc.cs.vt.edu/ictai06/Zadeh.ppt](http://www.nvc.cs.vt.edu/ictai06/Zadeh.ppt)

[fz prot] <http://statpac.org/walonick/systems-theory.htm>

[fl ppt] [www.comp.dit.ie/rlawlor/KE/notes/Fuzzy%20Logic.ppt](http://www.comp.dit.ie/rlawlor/KE/notes/Fuzzy%20Logic.ppt)

[fs ppt] [www.cs.ru.ac.za/courses/Honours/ai/FuzzySystems/FuzzySystemsP2.ppt](http://www.cs.ru.ac.za/courses/Honours/ai/FuzzySystems/FuzzySystemsP2.ppt)

[pb VN] <http://convorbiri-literare.dntis.ro/NEGOITAMar6.htm>

[wv hsn] <http://lumenosys.com/wavelet-transforms.html>