



UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA  
FACULTATEA DE MATEMATICĂ ȘI INFORMATICĂ

# Metode robuste de detecție a obiectelor cu aplicații în detecția facială

Rezumatul tezei de doctorat

*Conducător științific:*

Prof. univ. dr. Horia F. POP

*Doctorand:*

LEFKOVITS Szidónia

# Cuprinsul tezei de doctorat

<b>1</b>	<b>Introducere</b>	<b>1</b>
	Publicații . . . . .	3
	Contribuții originale . . . . .	5
	Organizarea tezei . . . . .	6
<b>2</b>	<b>Sisteme de detecție. Contextul general</b>	<b>8</b>
<b>3</b>	<b>Sistem de detecție bazat pe aspect global</b>	<b>13</b>
3.1	AdaBoost . . . . .	13
3.2	Funcții Haar ca clasificatori slabi . . . . .	17
3.2.1	Clasificatori slabi . . . . .	19
3.3	Clasificatorul cascadă . . . . .	20
3.4	Crearea clasificatorului propriu . . . . .	23
<b>4</b>	<b>Sistem de detecție bazat pe aspect local</b>	<b>33</b>
4.1	Modelul de obiect deformabil . . . . .	34
4.1.1	Definirea modelului . . . . .	34
4.1.2	Determinarea parametrilor modelului . . . . .	36
4.1.3	Algoritmul de detecție . . . . .	40
4.1.4	Transformata de distanță . . . . .	50
4.1.4.1	Algoritmul propus . . . . .	53
4.2	Descriptori locali cu filtre Gabor . . . . .	61
4.2.1	Filtre 2D Gabor . . . . .	61
4.2.1.1	Construcția unui jet idealizat . . . . .	74
4.2.2	Selectarea caracteristicilor cu algoritmul Gentle AdaBoost . . . . .	80
4.2.3	Crearea experimentală a descriptorului local . . . . .	88
4.2.4	Performanța descriptorului local creat . . . . .	122
<b>5</b>	<b>Concluzii și direcții de cercetare</b>	<b>132</b>
	<b>Anexe</b>	<b>134</b>
A	Prelucrarea imaginilor . . . . .	134
A.1	Egalizarea de histogramă . . . . .	134
A.2	Normalizarea imaginii . . . . .	135

## CUPRINSUL TEZEI DE DOCTORAT

---

B	Puncte de interes LoG – laplacian de gaussian . . . . .	137
C	Algoritmul AdaBoost – Analiza erorii și alegerea parametrilor . . . . .	139
D	Transformata Fourier a filtrelor 2D Gabor . . . . .	144
<b>Bibliografie</b>		<b>147</b>

# Cuprinsul rezumatului

<b>Introducere</b>	<b>1</b>
Publicații . . . . .	2
Contribuții originale . . . . .	4
Organizarea tezei . . . . .	5
<b>Sistem de detecție bazat pe aspect global</b>	<b>7</b>
<b>Sistem de detecție bazat pe aspect local</b>	<b>12</b>
<b>Concluzii și direcții de cercetare</b>	<b>27</b>
<b>Bibliografie</b>	<b>29</b>

# Mulțumiri

Adresez sincere mulțumiri domnului prof. univ. dr. Horia F. Pop, conducătorul științific al acestei teze, pentru ajutorul oferit pe tot parcursul acestui demers științific.

Mulțumesc, de asemenea, părinților mei pentru ajutorul pe care mi l-au oferit. Pasiunea pentru cercetare, încurajările și multe idei au venit din partea tatălui meu, iar mama mea m-a susținut moral pe tot parcursul elaborării acestei teze.

# Introducere

Sistemul vizual uman are capacitatea extraordinară de a recunoaște o varietate extrem de mare de obiecte sau categorii de obiecte din informația vizuală bidimensională sau tridimensională. Rezolvarea acestei sarcini a detecției și a recunoașterii fără efort se datorează mai ales procesării paralele extrem de rapide, a aptitudinii de a se dezvolta, a învățării din experiență și a structurii excepționale al creierului uman.

Scopul principal al vederii artificiale nu este înțelegerea percepției umane, ci modelarea și crearea sistemelor automate robuste de detecție a obiectelor. Vederea artificială este un subdomeniu al detecției generale de obiect, care prelucrează imagini bidimensionale ca o proiecție a spațiului tridimensional. Cercetările în acest domeniu nu au dus la crearea unui sistem general utilizabil pentru soluționarea tuturor aplicațiilor impuse de practică. Fiecare din sistemele existente este creat pentru un scop bine determinat și lucrează în anumite condiții date.

Realizarea unui sistem de detecție automată este o provocare pentru cercetătorii din domeniu. Principala dificultate constă în varietatea extrem de mare în care apar obiectele. Această varietate este indusă de mulțimea largă a obiectelor, iar fiecare obiect la rândul său poate să apară într-o infinitate de ipostaze: de dimensiune, formă, culoare, poziție, iluminare, umbrire, obturare. De asemenea, formele geometrice prezintă ambiguități datorită proiecției spațiului tridimensional în spațiul bidimensional al imaginilor. Un factor perturbator în imaginea obiectelor este modul de iluminare cu apariția umbrelor și a suprafețelor absorbante sau reflectoare de lumină. Dificultatea în detecția obiectelor este indusă și de cantitatea mare de informație conținută de o imagine. Un alt aspect nefavorabil constă în varietatea de prezentare a aceluiași obiect într-o multitudine de contexte de apariție [35].

Într-o serie de domenii de activitate detecția obiectelor în imaginile bidimensionale joacă un rol însemnat. Cele mai importante aplicații se regăsesc în aria medicinei, securității circulației, securității sistemelor de acces, sistemelor de identificare și autentificare, a industriei cu roboți inteligenți, a producției cinematografice etc.

## Publicații

### Jurnale ISI

- **Sz. Lefkovits** [57] "Improvements on Gabor Descriptor Retrieval for Patch Detection" *Computing and Informatics* (ISI IF=0.66, SRI = 0.15) [sub recenzie]

### ISI Proceedings

- **Sz. Lefkovits** [54] "Numerical Computation Method of the General Distance Transform" *KEPT 2011 International Conference Knowledge Engineering Principles and Techniques*, Babeș-Bolyai University, 2011. ISSN: 2067-1180, pp. 143–152, indexed MathSciNet, ZMath.
- **Sz. Lefkovits** [48] "Performance Analysis of Face Detection Based on Haar-features" *Proceedings of the International Conference on Complexity and Intelligence of the Artificial and Natural Complex Systems: Medical Applications of the Complex Systems: Biomedical Computing*, Petru Maior University, 2008, ISBN: 978-973-7794-76-5, pp.184–192.

### Jurnale internaționale indexate BDI

- **Sz. Lefkovits** [58] "Novel Gabor Filter Based Patch Descriptor" *10th Jubilee International Symposium on Intelligent Systems and Informatics, IEEE Conference*, Serbia, 2012 (lista Australian Research Council în clasa C) [acceptat]
- **Sz. Lefkovits** [55] "Numerical Computation Method of the General Distance Transform" *Studia Informatica Universitatis Babeș-Bolyai*, vol. 56, no. 2, 2011, pp. 68–74. indexed MathSciNet, ZMath.
- **Sz. Lefkovits** [49] "Assessment of Building Classifiers for Face Detection" *Acta Universitatis Sapientiae, Electrical and Mechanical Engineering*, no. 1, 2009, ISSN: 2065-5916, pp. 175–186, indexed EBCSO Databases, EZB – Karlsruhe Institute of Technology

### Jurnale internaționale recunoscute CNCSIS categoria B

- **Sz. Lefkovits** [53] "Combining Boosted Global- and Part-aspect Face Detectors" *Scientific Bulletin of the Petru Maior University*, vol.8, no. 1, 2011, ISSN: 1841-9267, pp. 35–40.

**Conferințe internaționale**

- **Sz. Lefkovits** [56] "Hybrid Face Detector Based on Boosted Classifiers" *The Eighth Conference of PhD Students in Computer Science*, Optimization session, University of Szeged, Hungary, Institute of Informatics, 2012 „**Best talk of the Session**” award
- **Sz. Lefkovits** [50] "Teaching Improvements on Haar-based Classifiers" *International Conference IETM-3*, Mathematics and Computer Science section, Petru Maior University 2009, ISSN: 1844-2048, pp. 1050–1060.
- **Sz. Lefkovits**, C. Enăchescu [59] "Face Detection System Based on Artificial Intelligence", *International Conference ICELM*, Decision Systems, Mathematical Modelling and Statistics section, Petru Maior University, 2008, ISBN: 973-7794-00-10, pp. 112–121

**Conferințe naționale**

- **Sz. Lefkovits** [52] "Classification Mixture for Object Detection" *Zilele Academice Clujene*, 2010, ISSN: 2066-5768, pp. 5–10



## Contribuții originale

- abordarea detecției în două faze: restrângerea domeniului de interes cu un detector rapid bazat pe aspect global urmat de o prelucrare detaliată cu un detector bazat pe aspect local [51]
- realizarea bazei de date pentru detecția bazată pe aspect global (subcapitolul 3.4, [48])
- îmbunătățirea algoritmului de creare a clasificatorului cascadă, scăderea semnificativă a detectărilor false prin selectarea de imagini negativ semnificative (subcapitolul 3.4, [59], [49])
- construirea clasificatorului facial propriu și determinarea performanțelor acestuia (subcapitolul 3.4, [49], [50])
- definirea unui descriptor local bazat pe filtre 2D Gabor (subcapitolul 4.2, [56])
- analiza și proiectarea teoretică a filtrelor 2D Gabor (subcapitolul 4.2.1.1, [52])
- experimentarea și evaluarea parametrilor filtrelor 2D Gabor definatorii descriptorului local (subcapitolul 4.2.3, [56])
- restrângerea domeniului de valori utile pentru cei 9 parametri ai filtrelor 2D Gabor (subcapitolul 4.2.3, [56])
- realizarea sistemului de selectare și evaluare a filtrelor caracteristice obiectului (subcapitolul 4.2.3, [58])
- crearea, determinarea și evaluarea experimentală a detectorului local propus (subcapitolul 4.2.3, [58])
- elaborarea unui algoritm original pentru evaluarea transformatei generalizate de distanță (subcapitolul 4.1.4.1, [54], [55])
- crearea sistemului de detecție de obiecte și implementarea acestuia cu procesare distribuită multisistem (subcapitolul 4.2.4, [52], [53], [56], [58])

## Organizarea tezei

**1. Introducere.** În acest capitol sunt prezentate scopul, sarcina, problematica și principalele aplicații ale detecției de obiecte. Se continuă cu enumerarea publicațiilor aferente tezei și a contribuțiilor originale aduse în lucrare.

**2. Sisteme de detecție. Contextul general.** Acest capitol conține aspectele generale și structura sistemelor de detecție a obiectelor. Sunt prezentate, pe scurt, principalele metode și sistemele cele mai semnificative din domeniu. Pentru fiecare parte a sistemelor de detecție sunt amintite tehnicile utilizate. La sfârșitul capitolului se descrie sistemul de detecție propus ce va fi analizat în detaliu în cursul lucrării.

**3. Sistem de detecție bazat pe aspect global.** În acest capitol prezentăm un sistem bazat pe aspect global. Subcapitolele 3.1, 3.2 și 3.3 conțin succint bazele teoretice.

Subcapitolul 3.1 descrie algoritmul AdaBoost elaborat de Y. Freund și R. Schapire [32, 78]. În subcapitolul 3.2 este analizat modelul de construire a clasificatorilor slabi din funcții Haar, creat de P. Viola și M. Jones [92]. Subcapitolul 3.3 prezintă ideea lui R. Lienhart et al. [61] de îmbunătățire a performanțelor sistemului prin clasificatorul cascadă.

În subcapitolul 3.4 sunt prezentate problemele ascunse ce trebuie soluționate pentru a crea un astfel de sistem funcțional. Aici se descriu metodele de creare ale bazelor de date precum și un algoritm propriu cu ajutorul căruia am reușit să construim clasificatori [48, 59] cu performanțe comparabile cu cele ale clasificatorilor existenți. Acest capitol se încheie cu prezentarea experimentelor cu clasificatorul creat [49, 50] și concluzionează avantajele și dezavantajele unui astfel de sistem, punctând către următorul capitol în care se creează a doua parte a detectorului.

**4. Sistem de detecție bazat pe aspect local.** Acest capitol detaliază aspectele teoretice și practice ce alcătuiesc un sistem bazat pe aspectul părților de obiect. Subcapitolul 4.1 prezintă un model de obiect deformabil [28]. În subcapitolul 4.1.3 am detaliat algoritmul de detecție și am definit tabelul de stare, ceea ce înlesnește implementarea efectivă a modelului propus. În subcapitolul 4.1.4 propunem un algoritm propriu [54, 55] pentru evaluarea eficienței a transformatei de distanță generalizată care este determinant în cazul evaluării funcției de cost al modelului deformabil. Acest algoritm este comparat cu algoritmul înfășurătoarei minime [70]. Experimentele efectuate scot în evidență avantajele algoritmului propus. Dacă înfășurătoarea minimă se poate calcula analitic, atunci timpii de execuție ai celor doi algoritmi sunt comparabili, dar dacă înfășurătoarea minimă se determină numeric, atunci algoritmul propus este mult mai eficient.

În subcapitolul 4.2 sunt expuse bazele teoretice ale filtrelor 2D Gabor [72]. În subcapitolul 4.2.1.1 propunem o metodă de proiectare a filtrelor Gabor bidimensionale [52, 56, 58], prin care se reduce domeniul de valori ai celor 9 parametri. Din constrângerile teoretice și experimentale spațiul parametrilor filtrelor s-a redus, dar nu suficient pentru a le utiliza ca descriptori locali. Se impune o altă metodă de reducere a numărului filtrelor cu algoritmul Gentle AdaBoost [33]. În subcapitolul 4.2.2 prezentăm implementarea acestui algoritm [56, 58]. În subcapitolul 4.2.3, prin utilizarea algoritmului Gentle AdaBoost și evaluarea rezultatelor cu validare pe un set de imagini, s-au putut alege filtrele cele mai caracteristice pentru pata de imagine analizată. Experimentele expuse determină influența parametrilor analizați. Subcapitolul 4.2.4 descrie performanțele descriptorului local creat [53, 56, 58]. Rezultatele practice se concentrează asupra utilizării acestuia în detecția ochiului uman. Având în vedere necesitatea de reducere a timpilor de execuție, se detaliază implementarea paralelă multisistem a proceselor de învățare și de detecție ale acestui descriptor.

## Cuvinte cheie

detecție a obiectelor, aspect-global, aspect-local, funcții Haar, undine Gabor bidimensionale, analiza în domeniul frecvență, proiectarea filtrelor, optimizarea parametrilor, baze de date marcate, învățare supervizată, AdaBoost, Gentle AdaBoost, generator de fundaluri, clasificator, reducerea detectărilor false, descriptor local, selectarea filtrelor, harta răspunsurilor, transformată de distanță generalizată, implementare paralelă.

# Sistem de detecție bazat pe aspect global

Sistemele bazate pe aspect global prezintă obiectele ca o singură entitate indivizibilă. În general, procesul de detecție folosește tehnica piramidei de imagini, adică imaginile sunt re-dimensionate și scanate cu o imagine de dimensiune standard. În acest capitol prezentăm un astfel de sistem [48, 49, 50, 53, 59]. Subcapitolele 3.1, 3.2 și 3.3 descriu succint bazele teoretice, ce compun sistemul.

Subcapitolul 3.1 prezintă algoritmul AdaBoost elaborat de Y. Freund și R. Schapire [32, 78]. Algoritmul construiește un clasificator puternic, alegând dintr-o mulțime de clasificatori slabi o serie de clasificatori cu pondere bine determinată, adaptată problemei de decizie. În subcapitolul 3.2 este analizat modelul de construire a clasificatorilor slabi din funcții Haar. Ideea folosirii funcțiilor dreptunghi în domeniul detecției de obiecte aparține lui C. Papanicolaou et al. [76]. Utilizând algoritmul AdaBoost și funcțiile Haar ca clasificatori slabi, P. Viola și M. Jones [92] creează un sistem de detecție de obiecte cu performanțe deosebite de funcționare în timp real. Subcapitolul 3.3 prezintă ideea lui R. Lienhart et al. [61] de îmbunătățire a performanțelor sistemului prin modificarea clasificatorului monolitic într-un clasificator cascada.

În studiile efectuate până în momentul de față, nu am găsit nicio referință despre metodologia de creare a clasificatorilor. Cercetătorii publică numai rezultatele experimentale, fără a divulga baza de date și tehnologia utilizată. În subcapitolul 3.4 sunt prezentate problemele ascunse [50], ce trebuie soluționate pentru a crea un sistem funcțional. Aici descriem metodele de creare ale bazelor de date [49] precum și un algoritm propriu [48], cu ajutorul căruia am reușit să construim clasificatori cu performanțe comparabile cu cele ale clasificatorilor existenți.

## Crearea clasificatorului propriu

Experimentele prezentate în continuare sunt executate cu programul Haartraining, parte integrantă a colecției OpenCV [11]. Ne-am propus crearea unui clasificator propriu pentru a putea detecta orice tip de obiect.

Pentru a ne ușura munca am creat programe auxiliare pentru marcarea și decuparea imaginilor de diferite extensii, folosind pachetul de programe Boost C++ și OpenCV. Acest program auxiliar facilitează marcarea punctelor de interes ale feței și decuparea la dimensiuni date. Punctele marcate folosite la decupare sunt stocate într-un fișier text și se vor folosi și

în procesul de învățare. Baza de date este compusă din 2893 de imagini faciale distincte.

Pe parcursul experimentelor am pornit de la o serie de întrebări la care am primit răspuns realizând propriul nostru clasificator.

1. Care este cea mai bună mărime a imaginilor de intrare? Cele mai bune rezultate au ieșit din măsurările noastre pentru o dimensiune de  $18 \times 24$ , având mai puține detectări false la aceeași rată de detecție.
2. Cum să decupăm fețele? Cu cât imaginea conține mai multe detalii semnificative, cu atât dimensiunea imaginii este mai mare. Rezultate considerabil mai bune le-am obținut decupând nu numai fața, ci și întreg conturul capului.
3. Care sunt imaginile semnificative? Definim imaginile pozitiv-semnificative acele imagini faciale care nu sunt detectate. Aceste tipuri de imagini ar trebui să fie introduse în baza de date a imaginilor pozitive. Trebuie găsite acele tipuri de fețe care acoperă lipsurile setului de antrenare. Procesul este anevoios deoarece după modificarea setului de antrenare presupune reluarea procesului de învățare.

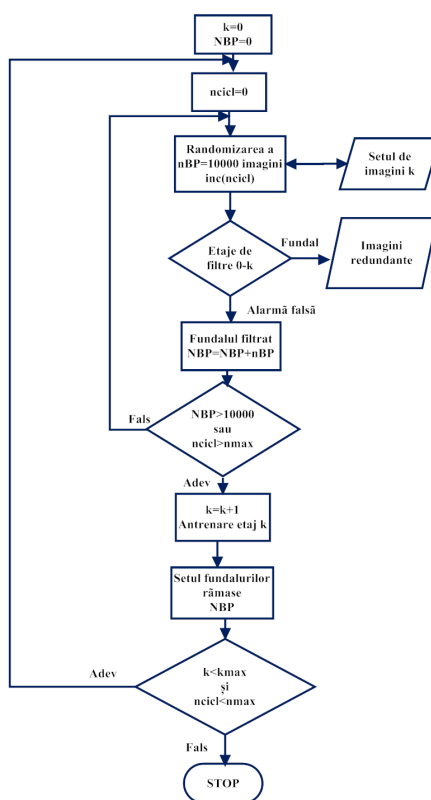


Figura 1: Generator de fundale semnificative

© Sz. Lefkovits

## 4. Cât este dimensiunea optimă a setului de antrenare?

Rezultatele au demonstrat că, cu cât este mai mare și variată setul de învățare, cu atât crește performanța de detecție, dar totodată se mărește timpul procesului de antrenare. Setul de antrenare a constat, în final, din 2893 de imagini faciale și milioane de imagini fundal.

5. Cum atingem o valoare mai mică decât  $5 \cdot 10^{-6}$  pentru rata detectărilor false?

În procesul de detecție se folosește tehnica piramidei de imagini. Imaginile piramidei (aceeași imagine la diferite scale) sunt parcurse de o fereastră glisantă de dimensiune dată (dim.  $18 \times 24$ ). Numărul de imagini de evaluat este de ordinul sutelor de mii, sau chiar de ordinul milioane. Acest lucru impune ca rata de detecție pozitiv falsă să fie sub  $5 \cdot 10^{-6}$ , performanțe, ce se pot obține, numai dacă numărul imaginilor negative din setul de antrenare este de ordinul milioane. Un număr mare de imagini negative ar prelungi timpul de antrenare și ar presupune un spațiu de memorare excesiv de mare. Am ajuns la concluzia, că fiecare etaj trebuie antrenat cu 3000-5000 de imagini faciale și 10000 de imagini fundal semnificative.

Definim imaginile negativ-semnificative acele imagini care nu reprezintă fețe și totuși sunt indicate ca imagini faciale, adică imaginile fals pozitive. Pentru marcarea și decuparea imaginilor negativ-semnificative am creat algoritmul generator de fundale (figura 1). Acest generator este utilizat înainte de crearea unui etaj nou pentru a dispune de un set suficient de imagini negative. Generatorul de fundale semnificative are sarcina de a găsi peste 10000 de imagini fals pozitive cu care să se antreneze următorul etaj. Bazându-ne pe experimentele efectuate, putem afirma că durata procesului de selecție a imaginilor negativ-semnificative pentru etajele 15 – 20 era mult mai mare ca efectiv procesul de învățare prin algoritmul AdaBoost.

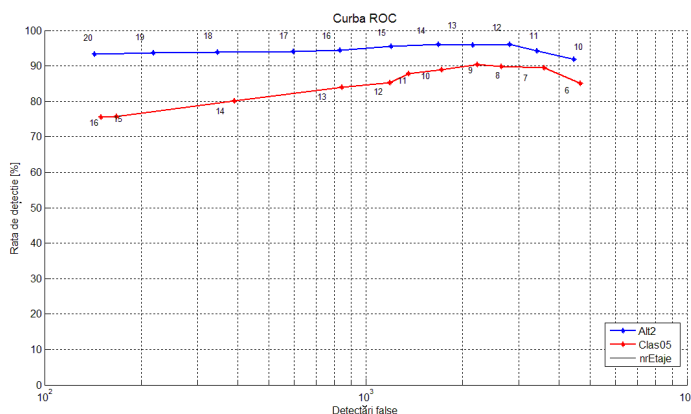


Figura 2: Curba ROC pe etaje

© Sz. Lefkovits

## Performanța clasificatorului propus

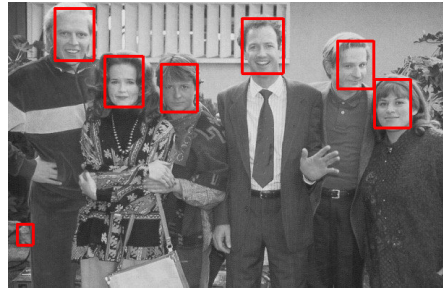
După obținerea clasificatorului ce satisface condițiile anterioare, am comparat performanțele acestuia cu clasificatorii existenți, folosind setul de imagini de testare CMU [94] set de imagini de test de referință în domeniu. În scopul comparării performanței clasificatorilor am reprezentat rata de detecție în funcție de numărul detectărilor false, adică o curbă ROC a clasificatorului. Un rezultat important este reducerea imaginilor fals detectate.

Se poate observa din curba ROC a clasificatorului propriu că odată cu descreșterea ratei detectărilor false descrește considerabil și rata de detecție. Totodată, numărul clasificatorilor slabi crește considerabil prin adăugarea de noi etaje.

Clasificatorul *Clas05* s-a creat cu o bază de date ce conține fețe umane europene, tinere, fără barbă, mustață sau ochelari. Am comparat clasificatorul nostru cu cel creat de Lienhart *frontalface\_alt2* [11] pe o serie de fotografii de pe internet și cele furnizate de FotoVision Studio [1]. Putem afirma, că am obținut rezultate mai bune, adică numărul detectărilor este similară, în schimb numărul detectărilor fals pozitive s-a redus considerabil. Se poate compara vizual performanța clasificatorului creat cu clasificatorul *frontalface\_alt2* pe setul de test CMU [94] și pe un set de fotografii proprii (figura 3).

O serie de sisteme de detecție bazate pe aspect global au fost elaborate de H. Rowley et al. [80], de H. Schneiderman [83], de P. Viola și M. Jones [92] și de R. Lienhart et al. [63]. Dezavantajul primordial al acestor sisteme este necesitatea unei baze de date de imagini, ce trebuie să acopere cât mai bine aspectele de apariție ale obiectului. În general, se concentrează în detecția unei singure clase de obiecte pentru o singură orientare spațială. Pentru a acoperi diferite orientări spațiale este necesară crearea bazei de date specifice orientărilor, iar pentru a distinge orientarea, se impune specializarea clasificatorului prin structuri paralele sau printr-o structură de arbore cu noduri de decizii specifice. Avantajul semnificativ al acestor sisteme este rata de detecție și timpul de detecție ce permite utilizarea detectoarelor în aplicații în timp real. Din experimentele efectuate putem concluziona, că soluția de detecție ce utilizează aspectul global se bazează pe capacitatea sistemului de a elimina cât mai rapid imaginile negative și fără pierderi de imagini pozitive. Dezavantajul acestor sisteme este, că detectează relativ multe imagini fals pozitive al căror aspect vizual nu seamănă de loc cu cel al imaginilor faciale. Pentru a diminua numărul detectărilor fals pozitive propunem crearea unui sistem care, în prima fază să detecteze rapid zonele de imagine care ar putea conține fețe umane, iar în faza a doua să detecteze caracteristici faciale. Astfel, configurația spațială a caracteristicilor detectate va micșora probabilitatea detectărilor fals pozitive. Bineînțeles performanțele procesului de detecție se îmbunătățesc în dauna timpului de procesare. Pentru detecția feței umane am utilizat sistemul propus de Lienhart et al. [61], iar pentru detecția caracteristicilor faciale vom utiliza un model deformabil bazat pe filtre 2D Gabor ca descriptorii locali.

Clasificatorul propriu *Clas05*



Clasificatorul *frontalface\_alt2*

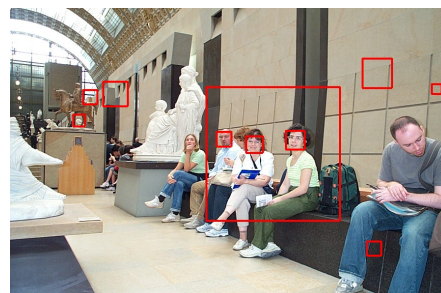
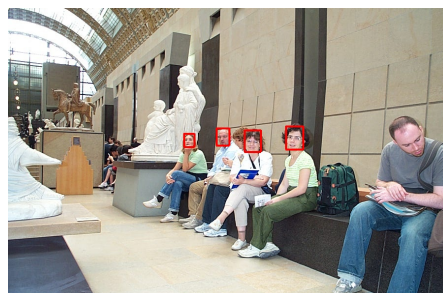
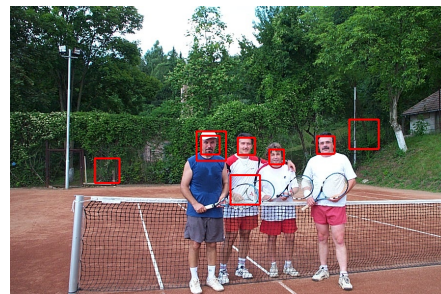
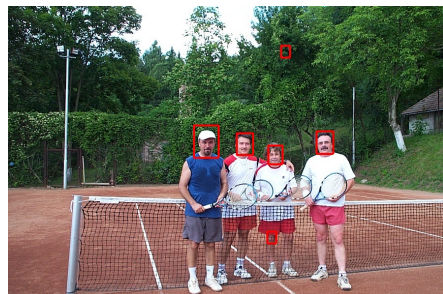
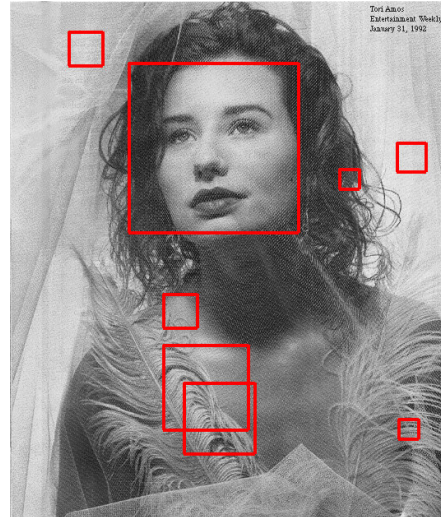
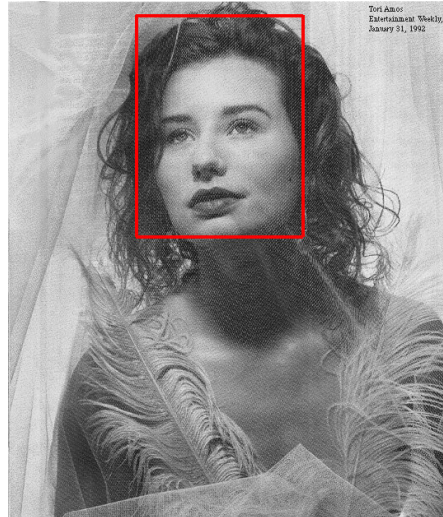
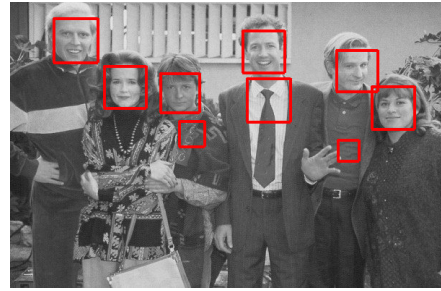


Figura 3: Exemple de detecție

© Sz. Lefkovits



# Sistem de detecție bazat pe aspect local

Cercetările recente în domeniul vederii artificiale tind să depășească detecția individuală de obiecte și se concentrează primordial la elaborarea unor metode de detecție a claselor de obiecte. Se impune crearea unor modele generice bazate pe părți de obiect și pe relațiile spațiale ale acestora. Una din soluțiile posibile sunt modelele deformabile aplicate obiectelor vizuale.

În general, într-un sistem de detecție bazat pe aspect local se pot delimita trei părți:

- punctele de interes – reprezintă mulțimea de puncte, unde se caută părțile locale de obiect.
- descriptorul local – reprezintă o descriere formală a petelor corespunzătoare cu diferite părți de obiect.
- modelul de obiect – reprezintă formalismul în care părțile de obiect detectate alcătuiesc obiectul de interes.

Acest capitol detaliază aspectele teoretice și practice, ce alcătuiesc un astfel de sistem [53]. Subcapitolul 4.1 prezintă un model de obiect deformabil. Partea teoretică de definire și determinare a parametrilor modelului este inspirat din modelul utilizat de Felzenszwalb et al. [28]. În subcapitolul 4.1.3 am detaliat algoritmul de detecție și am definit tabelul de stare, ceea ce înlesnește implementarea efectivă a modelului propus. Un factor determinant al complexității algoritmului este modelul de soluționare a transformatei de distanță. În subcapitolul 4.1.4 propunem un algoritm propriu [54] pentru evaluarea eficientă a transformatei de distanță generalizată. Algoritmul propus [55] este comparat cu algoritmul înfășurătoarei minime a lui Meijster [70] pentru diferite funcții de distanță. Dacă înfășurătoarea minimă se poate calcula analitic, atunci timpurile de execuție ai celor doi algoritmi sunt comparabili, dar dacă înfășurătoarea minimă se determină numeric, atunci algoritmul propus este mult mai eficient.

În subcapitolul 4.2 sunt expuse bazele teoretice ale filtrelor 2D Gabor [72]. Un filtru 2D Gabor este definit de 9 parametri, astfel numărul filtrelor utilizate este extrem de mare. Se impune restrângerea domeniului de variații ale parametrilor filtrelor. Rezultatele cercetărilor biologice și fizice reduc acest spațiu. Astfel, am propus o metodă de proiectare a filtrelor bi-dimensionale [58]. Din constrângerile teoretice și experimentale spațiul parametrilor filtrelor s-a redus, dar nu suficient pentru a le utiliza ca descriptori locali. În subcapitolul 4.2.3, prin utilizarea algoritmului Gentle AdaBoost [33] și evaluarea rezultatelor cu validare pe un set de imagini, s-au putut alege filtrele cele mai caracteristice pentru pata de imagine analizată.

Experimentele expuse [56] determină influența parametrilor analizați din punct de vedere informațional. Scopul este determinarea unui set restrâns de filtre cu o utilizare eficientă în localizarea și detecția unei părți de obiect. Subcapitolul 4.2.4 prezintă performanțele descriptorului local creat [56, 58] și posibilitățile de utilizare ale acestuia în detecția ochiului uman. Având în vedere necesitatea de reducere a timpilor de execuție, se detaliază implementarea paralelă multisistem a proceselor de învățare și de detecție ai acestui descriptor.

## Modelul de obiect deformabil

În cadrul modelelor deformabile obiectele sunt codate prin proprietățile vizuale locale ale părților și prin informația de legătură dintre acestea. În cadrul modelelor deformabile un obiect este definit de o mulțime de părți și de un set de relații de interconectare între anumite părți. Un astfel de model se poate defini printr-un graf neorientat.  $G = (V, M)$  în care  $V = \{v_1, v_2, \dots, v_n\}$  reprezintă mulțimea vârfurilor corespunzătoare părților obiectului și  $M = \{(v_i, v_j) | v_i, v_j \in V\}$  reprezintă mulțimea muchiilor, adică conectarea părților.

În general, soluția detecției în aceste modele este definită prin minimizarea unei funcții energetice [31]. Energia sau costul unei configurații particulare depinde, pe de o parte de măsura asemănării  $\mathbf{a}_i$  a părților  $v_i$  cu partea de imagine în locațiile corespunzătoare  $x_i$ , iar pe de altă parte, de distanțele  $\mathbf{d}_{ij}$  relative ale părților conectate.

Cu ajutorul acestor funcții se poate defini ca măsură de nepotrivire a modelului de obiect, cu imaginea obiectului, funcția de cost

$$C(X) = \left( \sum_{i=1}^n \mathbf{a}_i(x_i) + \sum_{(v_i, v_j) \in M} \mathbf{d}_{ij}(x_i, x_j) \right). \quad (1)$$

Minimizarea acestei funcții va determina o configurație de obiect, care corespunde cel mai bine cu imaginea obiectului.

Felzenszwalb et al. [28] propune în scopul eficienței calculului o limitare a structurii grafului  $G$ , un model adecvat pentru funcțiile  $\mathbf{d}_{ij}$  și un algoritm care ține cont de funcțiile  $\mathbf{d}_{ij}$  și  $\mathbf{a}_i$ , asigurând o minimizare eficientă a funcției de cost. Una din posibilitățile de determinare a parametrilor este dată de formularea statistică a modelului.

Formalismul statistic înlesnește învățarea parametrilor modelului din exemple. În principiu, toți parametrii se pot determina din setul de învățare, utilizând principiul verosimilității maxime. Totodată minimizarea funcției de cost este echivalentă cu maximizarea probabilității a posteriori de găsim a configurației de obiect în imaginea analizată.

După găsirea modelului, detecția constă în minimizarea funcției de cost, definită de relația (1) și determinarea simultană a minimumului. Această sarcină este rezolvată prin tehnica

programării dinamice, în cadrul căreia am detaliat tabelele de stare și algoritmi utilizați.

Complexitatea algoritmului de calcul este  $\mathcal{O}(h^2n)$ . Deficiența semnificativă a algoritmului este procesul de minimizare, ce trebuie efectuat pentru toate pozițiile posibile  $h$ , care sunt de ordinul zecilor de mii. Funcția de minimizat este o transformată de distanță generalizată, pentru care prezentăm un algoritm de calcul original în subcapitolul 4.1.2. Acest algoritm are o complexitate  $\mathcal{O}(h \log h)$  care trebuie executat în fiecare nod, parte de obiect. Rezultă, astfel, o complexitate globală pentru minimizarea funcției de cost  $\mathcal{O}(n \cdot h \log h)$ .

## Transformata de distanță

Transformata de distanță este un operator matematic, care calculează harta de distanță a unei imagini. Această hartă reprezintă tot o imagine.

Definim transformata de distanță pe mulțimea de puncte  $\mathcal{P}$ ,  $D_{\mathbf{a}} : \mathcal{P} \rightarrow \mathbb{R}_+$  astfel

$$D_{\mathbf{a}}(y) = \min_{x \in \mathcal{P}} M(x, y) = \min_{x \in \mathcal{P}} (\mathbf{d}(x, y) + \mathbf{a}(x)). \quad (2)$$

### Algoritm propus

În acest subcapitol se propune un algoritm original pentru calcularea transformatei generale de distanță conform definiției (2). În contrast cu algoritmi cunoscuți, această metodă poate fi folosită în cazul cel mai general, luând în considerare două restricții definite de Paglieroni [75]

$$\begin{aligned} \mathbf{d}(x, y) &= f(|x_1 - y_1|, |x_2 - y_2|, \dots, |x_n - y_n|) \\ |x_i - y_i| < |z_i - t_i| &\Rightarrow \\ \Rightarrow f(m_1, m_2, \dots, |x_i - y_i|, \dots, m_n) &< f(m_1, m_2, \dots, |z_i - t_i|, \dots, m_n). \end{aligned} \quad (3)$$

Ideea algoritmului se bazează pe următoarea observație. În scopul găsirii minimului funcției de măsură (2), într-un punct  $y$ , nu trebuie să se calculeze distanța  $\mathbf{d}(x, y)$  pentru toate punctele  $x$ . Presupunând că în procesul de minimizare s-a găsit un minim actual în  $x_0$ , punctele  $x$ , pentru care nu este necesar calculul efectiv al funcției de măsură sunt punctele, în care valoarea funcției de disimilaritate  $\mathbf{a}(x)$  este mai mare decât valoarea actuală și este la o distanță mai mare decât punctul actual.

$$\mathbf{a}(x) > \mathbf{a}(x_0), \quad (4)$$

$$\mathbf{d}(y, x) > \mathbf{d}(y, x_0), \quad \forall y. \quad (5)$$

Suma celor două funcții nu poate să fie sub nicio formă mai mică, decât valoarea funcției de măsură în punctul actual.

Algoritmul se bazează pe ordinea disimilarității și a distanței și definește punctele de minim potențial. Punctele de minim potențial sunt punctele, în care are sens să evaluăm valoarea funcției de măsură. Transformata de distanță va fi egală cu cea mai mică valoare a funcției de măsură în punctele de minim potențial.

Acest algoritm ne sugerează o implementare recursivă. Metoda prezentată poate fi concepută cu ajutorul unui arbore binar de căutare. În acest arbore elementele se introduc în ordinea crescătoare disimilarității (ordinea crescătoare a lui  $\mathbf{a}$ ). Poziția elementelor în arborele binar determină poziția lor față de poziția părintelui. Dacă s-a inserat în arbore fiecare disimilaritate  $\mathbf{a}(1), \dots, \mathbf{a}(n)$ , atunci printr-o parcurgere în preordine se soluționează transformata de distanță.

Algoritmul este implementat conform pseudocodului (algoritmul 1)

---

**Algoritmul 1** TRANSFORMATA DE DISTANȚĂ( $\mathbf{a}, d$ ) – © Sz. Lefkovits

---

**Date:**  $a, d$

**Rezultate:**  $C_a, \text{argmin}(C_a)$

- 1: SORTARE( $[a, d]$ ) {sortare crescătoare după  $a$ }
  - 2: arbore:=CREARE ( $[a, d](1)$ )
  - 3: **pentru**  $i := 2..n$  **execută**
  - 4:   INSERARE(arbore, $[a, d](i)$ ) {inserare în ordinea crescătoare a lui  $a$ ,  
    cu relația de ordonare a nodurilor după  $d$ }
  - 5: **sfârșit pentru**
  - 6: (argmin, min):=MINIMTD(rădăcină\_arbore, 1, dim( $[a, d]$ )) {traversare preordine}
  - 7: **returnare**  $C_a, \text{argmin}(C_a)$
- 

Am comparat acest algoritm (1) propus cu algoritmul prezentat de Meijster [70] și implementat de Felzenszwalb et al. [28]. Experimentele comparative pentru distanțele  $l_1$  și  $l_2$  pledează în favoarea algoritmului parabolilor.

În schimb, pentru un număr mare de puncte/noduri ( $n > 5000$ ) algoritmul propus găsește mai repede transformata de distanță. Experimentul infirmă timpul de execuție, „aproape liniar” [27] al algoritmului parabolilor  $\theta(n)$ , care devine mai lent față de algoritmul propus  $\theta(n \log n)$ . În cazuri generale, se utilizează distanțe complexe.

Algoritmul lui Meijster presupune determinarea înfășurătoarei minime, ce este un factor important în durata algoritmului. Algoritmul elimină acest inconvenient, devenind astfel mai eficient, reducând durata de execuție cu 1-2 ordini de mărime.

## Descriptori locali cu filtre Gabor

Reformulând ideile expuse în capitolele anterioare, se impune realizarea unui descriptor local, ce poate fi utilizat într-un proces de detecție robust. Cum am mai expus, procesul de învățare

este una supervizată, în care sunt marcate părțile semnificative de obiect. În prima fază, se creează un clasificator robust, ce se bazează pe aspectul global al obiectului, iar în faza a doua, se creează un clasificator, ce se bazează pe aspectul local având rolul de reducere, de eliminare a detectărilor false. Se dorește construirea unui descriptor de aspect local pe baza răspunsului filtrelor Gabor bidimensionale.

Acest subcapitol este structurat în trei părți. În prima parte se prezintă teoria filtrelor Gabor bidimensionale [72], subliniind relațiile teoretice ce stau la baza proiectării filtrelor. În partea a doua se prezintă algoritmul de clasificare Gentle AdaBoost, justificarea alegerii, modul de implementare al acestuia. În ultima parte, pe baza teoriei expuse, se determină experimental un descriptor local și se analizează comportarea și posibilitățile de utilizare ale acestuia în modelul de obiect deformabil propus.

## Filtre 2D Gabor

Reprezentarea imaginilor prin coeficienții de descompunere cu undine 2D Gabor ne sugerează filtrarea imaginii cu filtru Gabor bidimensional într-o serie de puncte bine stabilite. De fapt, determinarea coeficienților descompunerii este rezultatul procesului de convoluție dintre imaginea  $I(x, y)$  și filtrul Gabor  $g(x, y)$ . Formula generală a filtrului Gabor bidimensional este identică cu cea a unei spațiale propusă de Daugman [19, 20]

$$g(x, y) = \frac{1}{k} e^{-\pi \left[ \frac{(x-x_0)_r^2}{\alpha^2} + \frac{(y-y_0)_r^2}{\beta^2} \right]} e^{i[(\xi_0(x-x_0) + \nu_0(y-y_0)) + P]}, \quad (6)$$

unde

$$(x - x_0)_r = (x - x_0) \cos \theta + (y - y_0) \sin \theta, \quad (7)$$

$$(y - y_0)_r = -(x - x_0) \sin \theta + (y - y_0) \cos \theta. \quad (8)$$

Astfel, filtrul Gabor bidimensional este determinat prin 9 parametri.

- $\frac{1}{k}$ : amplitudinea înfășurătoarei gaussiene;
- $(\alpha, \beta)$ : mărimea axelor înfășurătoarei gaussiene;
- $\theta$ : unghiul de rotație al înfășurătoarei gaussiene;
- $(x_0, y_0)$ : punctul de maxim al înfășurătoarei gaussiene;
- $(\xi_0, \nu_0)$ : frecvența spațială a unei plane sinusoidale;
- $P$ : faza unei sinusoidale.

(9)

În alegerea filtrului Gabor avem 9 grade de libertate, factor ce îngreunează alegerea filtrelor de utilizat. Se impune o reducere rațională a spațiului parametrilor în vederea utilizării eficiente a filtrului propus. Analiza teoretică în domeniul frecvență prezentată în subcapitolul 4.2 s-a realizat în scopul găsirii anumitor legături și restricții între parametrii, ceea ce duce implicit la reducerea dimensiunii spațiului de definiție și la limitarea teoretică a mulțimii de valori a fiecărui parametru. Relațiile, care s-au luat în considerare, sunt:

1.  $\frac{1}{k}$  amplitudinea înfășurătoarei gaussiene este un factor multiplicativ, aparent fără implicație semnificativă. Se poate alege empiric o valoare constantă arbitrară. Deoarece aceste undine sunt folosite cu scopul de filtrare a imaginilor și utilizate ca funcții de clasificare, se impune compararea răspunsului acestora. Pentru a realiza o comparare reală, trebuie ca energia maximă, care trece prin filtre să fie identică. Pe baza teoriei filtrării s-a considerat util ca integrala înfășurătoarei gaussiene să fie egală cu unitatea, de unde  $\frac{1}{k} = \frac{1}{\alpha\beta}$ .
2. Pe baza considerentelor neurobiologice, orientarea înfășurătoarei gaussiene, direcția axei principale  $\theta$  se consideră identică cu direcția de propagare a undinei plane  $\theta = \omega_0$ .
3. Raportul axelor  $\alpha$  și  $\beta$  determină aspectul  $S = \frac{\beta}{\alpha}$ . În general, aspectul  $S \geq 1$ , ceea ce implică în domeniul spațial că axa principală este mai mică ca axa secundară a înfășurătoarei.
4. Punctul  $(x_0, y_0)$  s-a considerat a fi nul, deoarece este punctul central al undinei Gabor. Acest punct este luat ca referință în expresia filtrelor. Punctul de referință va coincide cu punctul de imagine unde se aplică filtrarea.
5. Faza  $P$  a undinei plane poate fi considerată nulă,  $P = 0$  deoarece punctul de referință  $(x_0, y_0)$  parcurge toate punctele de interes.
6. Relația de legătură a parametrilor de filtrare (banda de frecvență  $bw$ , lungimea de undă  $\lambda_0$  și axa  $\alpha$ ) este exprimată de

$$K_\alpha = \frac{2^{bw} - 1}{2^{bw} + 1}, \quad (10)$$

$$\frac{K_\alpha}{\lambda_0} = \frac{C}{\alpha}. \quad (11)$$

Obținem

$$\frac{\alpha}{\lambda_0} = \frac{C}{K_\alpha}. \quad (12)$$

Această relație exprimă faptul că pentru o lărgime de bandă raportul dintre axa principală și lungimea de undă  $\lambda$  este o constantă.

Pe baza acestor considerente teoretice numărul parametrilor independenți s-a redus la patru, și anume:

- $\lambda_0 = \frac{1}{F_0}$  lungimea de undă
- $S$  aspectul înfășurătoarei
- $bw$  lărgimea de bandă
- $\theta$  direcția de propagare a undei plane

## Crearea experimentală a descriptorului local

Majoritatea aplicațiilor, care utilizează filtrele Gabor, folosesc un set de filtre cu parametri aleși pe criterii empirice și insuficient argumentate științific. Fiecare autor descrie modul de utilizare al filtrelor, fără a da date concrete detaliate despre dimensiunea și parametri filtrelor utilizate. În general, se utilizează un set de 40 de filtre, având 8 orientări și 5 frecvențe. Aceste informații nu sunt suficiente pentru a reproduce teoriile expuse sau de a le dezvolta și aplica în scopuri de cercetare proprii. Optimizarea numărului de filtre Gabor și implicit determinarea parametrilor acestora este un domeniu de cercetare abordat de mai mulți autori [40, 45, 85, 93]. Această optimizare constă în reducerea numărului de filtre, alegerea unui set minim de parametri și determinarea locațiilor semnificative, unde se aplică filtrarea.

Se prezintă, în continuare, modul de realizare a descriptorului de aspect local, luând considerare o anumită parte de imagine. Procesul de creare a clasificatorului local presupune două etape distincte. În prima etapă trebuie creată baza de date cu imaginile părții locale, iar în etapa a doua se construiește clasificatorul aspectului local corespunzător.

Procesul de învățare supervizat presupune existența unui număr mare de imagini obiect, care trebuie să acopere cât mai bine toate aparițiile posibile ale obiectului. La crearea descriptorului local am folosit 730 de imagini de obiect și 2000 de imagini fundal în setul de antrenare și 159 de imagini obiect și 500 de imagini non-obiect în setul de test.

În experimentele practice trebuie să stabilim valori concrete pentru parametrii. Cercetările neurologice limitează domeniul acestor valori, astfel, aspectul  $S$  ia valori între 1 și 2, iar lungimea de bandă  $bw$  ia valori între 1, 2 și 1,7 octave.

Unghiurile de propagare a undei Gabor se determină din relațiile, ce stabilesc posibilitățile de reconstrucție a imaginii răspunsului filtrelor. Numărul maxim de orientări, ce merită a se lua în considerare este de 24. Astfel, orientarea  $\theta_l$  se definește ca un multiplu întreg al unghiului de rotație minimă

$$\theta_0 = \frac{2\pi}{2L} = \frac{\pi}{L}, \text{ unde } L \in \{6, 8, 12\}. \quad (13)$$

Obținem  $\theta_l = l \cdot \theta_0$  un spațiu de maxim 12 valori distincte. Orientările luate în considerare sunt doar în cadranul 1 și 2 deoarece filtrele cu orientările din cadranele 3 și 4 sunt complex conjugatele acestora.

Pentru lungimile de undă se poate considera tot spațiul discret delimitat superior de mărimea filtrului,  $R = 16$  și având eroarea de discretizare acceptată

$$\frac{\lambda}{R} < \frac{\pi}{\sqrt{-\ln 2 \cdot \ln(\varepsilon)}} \cdot \frac{1}{S} \cdot \frac{2^{bw} - 1}{2^{bw} + 1}, \quad (14)$$

iar inferior de teorema de eșantionare, adică

$$F_{max} < \frac{F_s}{2} = \frac{1}{2} = 0,5 \text{ pixel/ciclu}, \text{ astfel } \lambda > 2. \quad (15)$$

În experimentele efectuate pentru determinarea unui clasificator de aspect local am considerat mărimea filtrului Gabor ca fiind  $33 \times 33$  pixeli, iar domeniul valorilor parametrilor este:

$$\begin{aligned} \lambda &\in \{4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 16, 18, 22\} && - 14 \text{ valori} \\ S &\in \{0, 5; 0, 66; 1; 1, 2; 1, 5; 1, 8\} && - 6 \text{ valori} \\ bw &\in \{1; 1, 5; 2\} && - 3 \text{ valori} \\ \theta &\in \left\{ \frac{l\pi}{12} \mid l = \overline{0, 11} \right\} && - 12 \text{ valori.} \end{aligned}$$

Astfel, într-un singur punct se pot defini un număr de  $14 \times 6 \times 3 \times 12 = 3024$  filtre diferite, iar răspunsul fiecărui filtru este un număr complex. Domeniul de valori este cel al numerelor complexe, unde relația de ordine nu există. Se va impune analiza tipului valorii, ce trebuie utilizat pentru crearea clasificatorilor bazați pe răspunsul filtrelor Gabor. În experimentele efectuate am considerat utilă analiza informațională a 5 valori definite a răspunsului filtrelor: partea reală, partea imaginară, modulul, argumentul și distribuția statistică.

În concluzie, spațiul analizat pentru un singur punct de imagine are o dimensiune de  $15120 = 3024 \times 5$ . Scopul primordial al experimentelor este de reducere a acestui spațiu prin alegerea caracteristicilor semnificative.

Pe lângă spațiul parametrilor de filtrare sunt încă o serie de factori care determină procesul de învățare și de prelucrare a imaginilor. Rezultatele experimentale vor delimita în mod coerent influența fiecărui factor definit în procesul de selectare și de decizie al creerii descriptorului local optim.

Pentru analiza și selecția caracteristicilor s-a utilizat algoritmul de clasificare Gentle Ada-Boost [33]. Acest algoritm, pe lângă proprietatea de selecție a caracteristicilor semnificative,



ne furnizează și nivelul de încredere al deciziei, utilizat în evaluarea clasificatorului obținut.

Pentru a decide influența fiecărui factor asupra performanțelor descriptorului local, experimentele s-au realizat în aceleași condiții, modificându-se numai factorul de analizat. Criteriul de evaluare a factorilor este evoluția erorilor în funcție de numărul de clasificatori. Evaluarea erorilor s-a realizat atât pe setul de antrenare cât și pe setul de test.

În experimente s-au evaluat următoarele erori:

- Eroarea de detecție este raportul exprimat în procente între numărul detectărilor false și numărul total de detecție.
- Rata detectărilor fals negative este raportul exprimat în procente între numărul detecțiilor eronate de imagini pozitive și numărul total al imaginilor pozitive.
- Rata detectărilor fals pozitive este raportul exprimat în procente între numărul detecțiilor eronate de imagini negative și numărul total de imagini negative.

Experimentele detaliate s-au realizat pentru următorii factori:

- a) analiza informațională a răspunsului filtrelor. Aceste experimente au scopul de a decide modul de utilizare a răspunsurilor filtrelor. Partea reală, partea imaginară, modulul și argumentul sunt valorile analizate din răspunsul complex al filtrelor Gabor.
- b) aspectul filtrelor  $S$ . Raportul axelor înfășurătoarei gaussiene definește aspectul filtrelor. În cele mai multe lucrări de specialitate aspectul este considerat 1. Pentru a alege o valoare optimă a aspectului  $S = \frac{\beta}{\alpha}$  ( $\alpha$  fiind dimensiunea axei în direcția de propagare a undei plane,  $\beta$  fiind dimensiunea axei perpendiculare pe această direcție), am experimentat prin procesul de învățare și apoi cel de testare comportarea clasificatorilor finali.
- c) banda de frecvență  $bw$ . Lungimea de undă în direcția de propagare, raportat la lungimea de undă, este invers proporțională cu lățimea benzii, adică dacă crește lățimea de bandă, descrește lungimea de undă. La aceeași lungime de undă și lățime de bandă, odată cu creșterea aspectului, crește și lățimea filtrului.
- d) lungimile de undă  $\lambda$  și numărul de orientări  $\theta$ . Ca clasificatori slabi se pot considera mai multe filtre  $2D$  Gabor, cu mai multe lungimi de undă  $\lambda$  și mai multe direcții de propagare  $\theta$  decât cele calculate din condiția de acoperire a domeniului de frecvență cu filtre tangențiale. În experimente am considerat un număr mai mare de lungimi de undă și orientări, ce se suprapun în domeniul frecvență. Procesul de învățare, folosind algoritmul Gentle AdaBoost, are sarcina de a alege clasificatorii semnificativi, cu care se poate detecta eficient aspectul local impus.

- e) parametrii procesului de învățare. În procesul de învățare, dacă rata ponderii imaginilor pozitive și celor negative este 3:1, se constată efectul inegalității ponderilor imaginilor prin scăderea considerabilă a ratei detectărilor fals negative, și în dauna creșterii ratei detectărilor fals pozitive. Același fenomen se poate observa și în procesul de testare, dar efectul ponderilor este diminuat.
- f) prelucrarea de normalizare și egalizare a imaginii. Rezultatele prezentate indică un rol semnificativ al funcției de egalizare. Funcția de normalizare are un efect mai puțin eficient, deoarece imaginile setului de test prezintă o dispersie a intensității punctelor cu variație relativ mică. Totuși, pentru realizarea unui detector, care este utilizat în condiții reale, această funcție contribuie la păstrarea performanțelor detecției.
- g) alegerea unor puncte de interes (LoG). Mulțimea punctelor, în care trebuie evaluate răspunsurile filtrelor, reprezintă un factor determinant pentru performanța detectorului de obiecte. Timpul de procesare este direct proporțional cu numărul de puncte luate în considerare. Una din metodele de reducere a cardinalității mulțimii punctelor de procesare este de a utiliza doar mulțimea punctelor de interes. Am constatat, că atât procesul de învățare, cât și cel de testare este mai performant pentru punctele marcate. Utilizând detectorul de puncte de interes LoG, eroarea de detecție crește, în schimb, timpul de detecție scade considerabil.
- Performanțele de detectare devin astfel considerabil mai slabe, dar scăderea semnificativă a timpului de detecție recomandă utilizarea acestui procedeu de detecție pentru secvențe de imagini.

## Performanța descriptorului local creat

Prin utilizarea metodologiei expuse în capitolul anterior, obținem descriptorul local cel mai performant pentru setul de imagini de învățare și de test din baza de date creată. Procesul de învățare utilizează imaginile pozitive cu pondere inițială de 3 ori mai mare ca ponderea celor negative. Rezultatele cele mai bune le-am obținut, pentru puncte marcate, cu pete de imagini uniformizate de procesele de normalizare și egalizare. Clasificatorul performant obținut este compus din 40 de filtre, având aspectul  $S = 1,8$  și având configurația parametrilor dată în tabelul 1.

În procesul de învățare, rata detectărilor false devine 0 cu numai 28 de clasificatori slabi, iar imaginile pozitive ale setului de învățare sunt complet separate de cele negative cu 68 de clasificatori (figura 4). Pentru imaginile de test, rata erorilor fals pozitive devine chiar nulă, iar eroarea globală scade sub 2% cu 32 de clasificatori (figura 5).

	Clasificator	$\lambda$	$\theta[^\circ]$	$bw$		Clasificator	$\lambda$	$\theta[^\circ]$	$bw$
1.	GF_02718	18	15	1,5	21.	GF_02086	9	45	1,0
2.	GF_02248	10	135	1,0	22.	GF_01987	8	60	1,0
3.	GF_01574	4	90	2,0	23.	GF_01708	5	135	1,0
4.	GF_01636	5	15	1,0	24.	GF_02933	22	15	2,0
5.	GF_02493	14	0	1,5	25.	GF_01555	4	60	1,0
6.	GF_01555	4	60	1,0	26.	GF_02167	10	0	1,0
7.	GF_02313	11	60	1,5	27.	GF_01592	4	120	2,0
8.	GF_02932	22	15	1,0	28.	GF_01645	5	30	1,0
9.	GF_01825	6	150	1,0	29.	GF_02645	16	75	2,0
10.	GF_01772	6	60	2,0	30.	GF_02140	9	135	1,0
11.	GF_01592	4	120	2,0	31.	GF_02113	9	90	1,0
12.	GF_02275	11	0	1,0	32.	GF_02941	22	30	1,0
13.	GF_01618	4	165	1,0	33.	GF_01681	5	90	1,0
14.	GF_02437	12	90	1,0	34.	GF_02626	16	45	1,0
15.	GF_02932	22	15	1,0	35.	GF_02185	10	30	1,0
16.	GF_02734	18	45	1,0	36.	GF_02815	20	0	1,0
17.	GF_01979	8	45	2,0	37.	GF_02338	11	105	1,0
18.	GF_02347	11	120	1,0	38.	GF_02176	10	15	1,0
19.	GF_02609	16	15	2,0	39.	GF_01915	7	120	1,0
20.	GF_01834	6	165	1,0	40.	GF_02078	9	30	2,0

Tabelul 1: Clasificatorul cel mai performant  $S = 1,8$ 

© Sz. Lefkovits

Descriptorul local creat pe lângă funcționalitatea de descriptor are și rolul de detector al petei locale. Prin aplicarea descriptorului local, în toate punctele unei imagini, obținem imaginea-hartă a răspunsului descriptorului. În prima fază, harta răspunsurilor este suprafață a de valori reale furnizate de funcția de clasificare Gentle AdaBoost ( $R > 0$  obiect,  $R < 0$  non-obiect).

Aplicând algoritmul prezentat în subcapitolul 4.2.4, pentru o imagine obținem harta răspunsurilor descriptorului. Procedeele de filtrare în  $4 \times 4$  pixeli este suficient pentru detecția petei și necesar pentru reducerea de aproape 16 ori a timpului de detecție. Această hartă va fi utilizată în modelul de obiect în scopul detecției obiectului. Modelul de obiect determină configurațiile posibile din răspunsul clasificatorilor, eliminând petele pozitiv detectate, ce nu pot fi atribuite obiectului. Astfel, harta răspunsurilor trebuie evaluată conform performanțelor descriptorului local.

Una din deficiențele majore ale descriptorului propus este durata mare de procesare, pentru că în procesul de învățare fiecare filtru trebuie să fie evaluat pentru toate imaginile din setul de antrenare. În procesul de detecție numărul convoluțiilor depinde de numărul de filtre din clasificatorul puternic și de numărul de puncte de interes de pe o imagine. Am

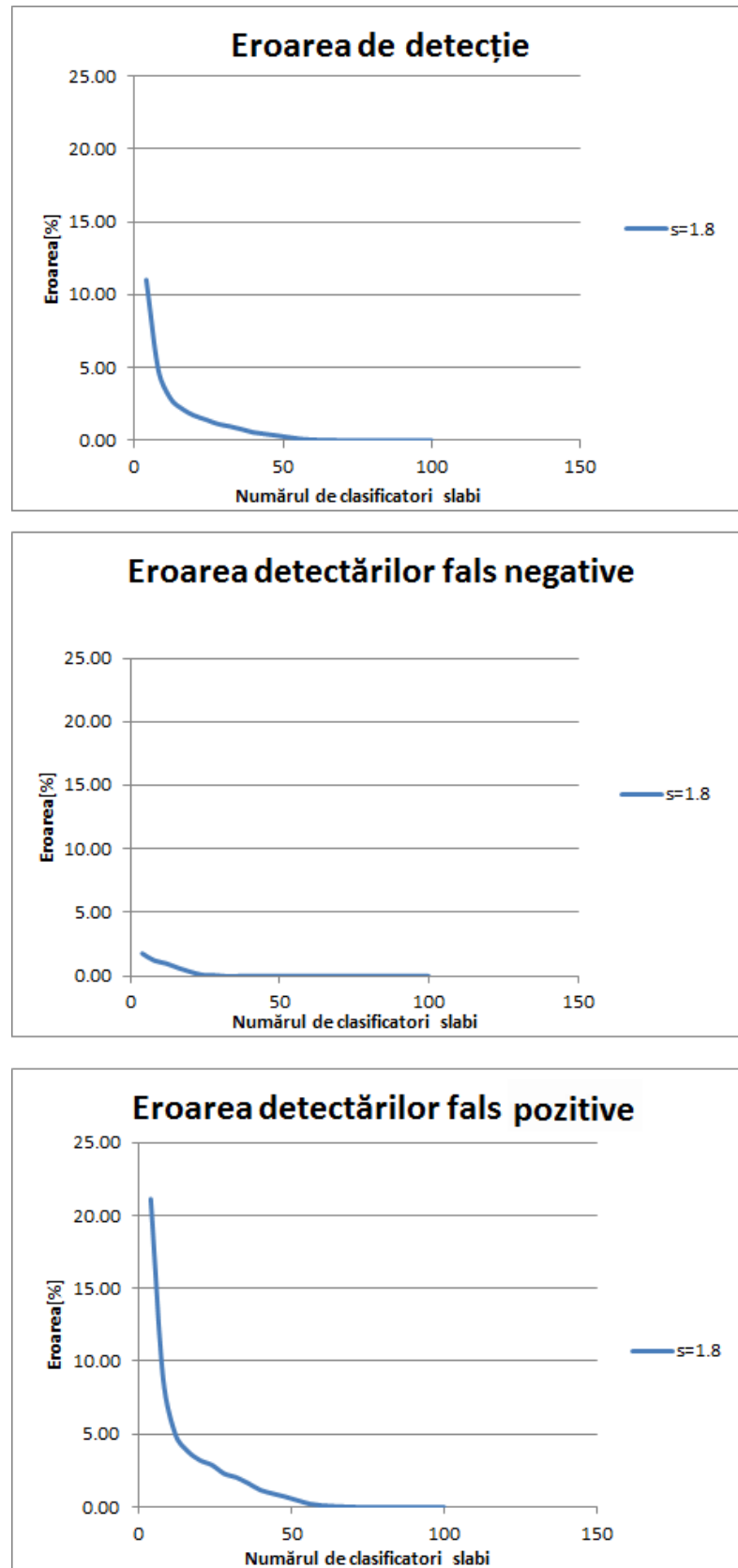


Figura 4: Erorile de detecție pe setul de antrenare ale celui mai performant clasificator  
 © Sz. Lefkovits

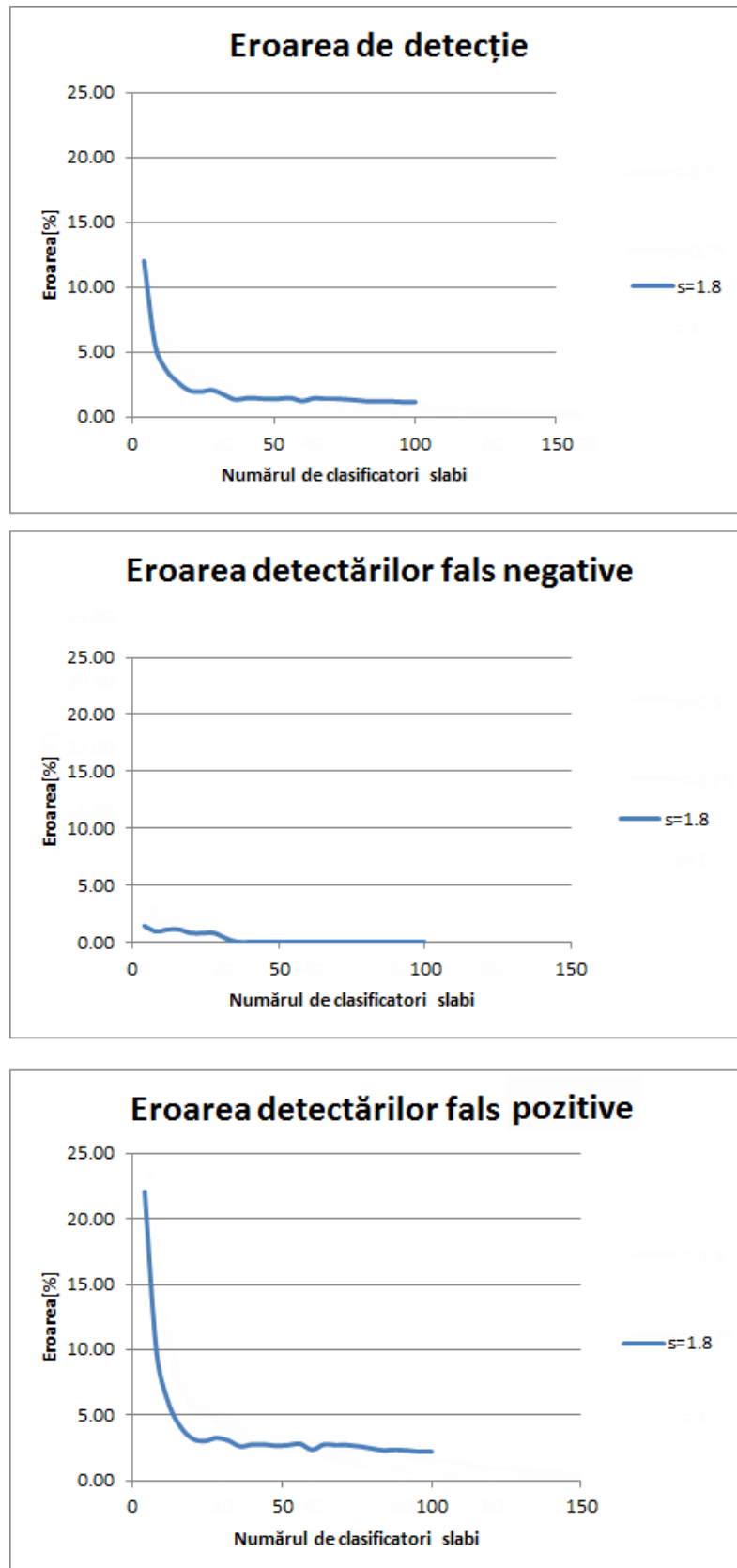


Figura 5: Erorile de detecție pe setul de test ale celui mai performant clasificator  
 © Sz. Lefkovits

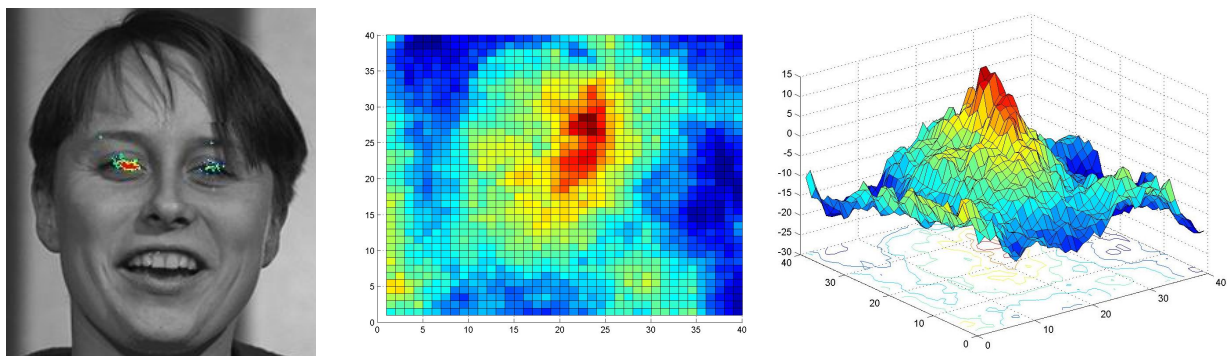


Figura 6: Harta răspunsurilor în vecinătatea obiectului

© Sz. Lefkovits

demonstrat, că pentru o detecție corectă este suficientă evaluarea descriptorului într-o rețea cu pasul de 4 pixeli. În scopul reducerii timpului de calcul am implementat descriptorul pe un sistem multiprocesor folosind un cluster. Aplicația creată utilizează pachetul de programe Matlab [2]. Acest pachet de programe asigură procesarea paralelă prin serverul distribuit MDCS (Matlab Distributed Computing Server).

În figura 7 am reprezentat câteva exemple ale celui mai performant descriptor creat (tabelul 1 și figurile 4 și 5) cu acest sistem pentru detecția ochiului uman.

Sistemul de detecție de obiect definit în acest capitol este nou prin utilizarea filtrelor 2D Gabor, ca descriptor local. Studiul amănunțit demonstrează, că un set de filtre alese pe criterii informaționale alcătuiește un descriptor suficient de general și totodată specific. În schimb, prin delimitarea judicioasă a domeniului de valori al parametrilor filtrelor, procesul de determinare al descriptorului local se poate automatiza fără evaluarea intermediară a performanțelor. În procesul de creare al descriptorului este nevoie de o bază de date de imagini cu părțile obiectului de interes marcate. Se impune, ca în procesul de învățare petele de imagine corespunzătoare părților, să fie rescalate la o dimensiune standard. În procesul de detecție bazat pe aspect local utilizăm imagini de aceeași dimensiune. Acest aspect se poate rezolva cu procesarea multi-rezoluție, sau cu un detector de aspect global. Sistemul propus utilizează, în prima fază, un detector rapid de aspect global, deci în procesul de detecție este suficientă doar o simplă redimensionare a imaginii.

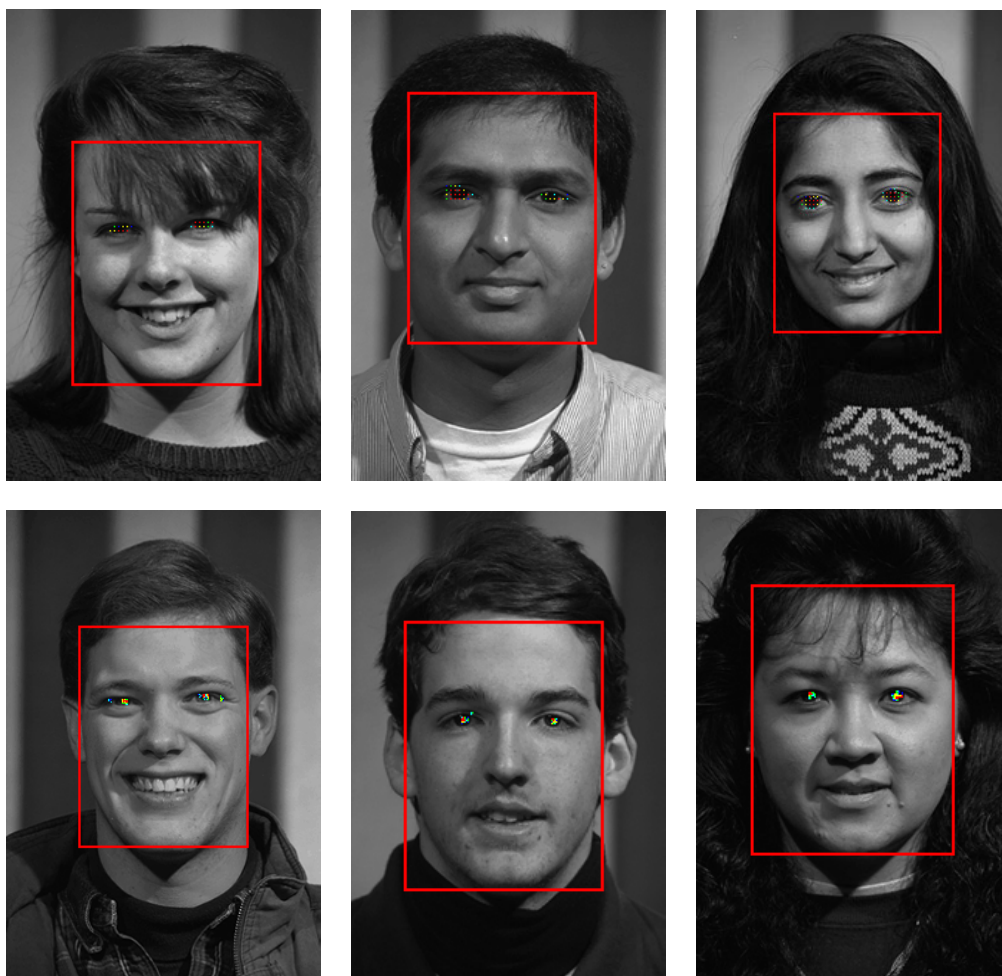


Figura 7: Exemple de imagini de test din baza de date FERET  
© Sz. Lefkovits

Durata procesului de detecție este determinat de numărul de filtre utilizate și de numărul punctelor, în care ele se evaluează. Pentru procesul de detecție am propus o procesare paralelă pentru calculul răspunsurilor filtrelor. Prin procesare paralelă timpul de detecție se poate reduce în limitele dorite. Descriptorul local creat prezintă proprietățile filtrelor Gabor de a fi invariante la translații, rotații și scalări.

Experimentele făcute în detectarea ochiului uman cu descriptorul creat, demonstrează performanțe de detecție remarcabile. Un astfel de descriptor este adecvat utilizării în orice model bazat pe părți de obiect.

# Concluzii și direcții de cercetare

Lucrarea de doctorat prezentată și-a propus studiul și realizarea unui sistem robust de detecție de obiecte. Caracteristica de robust a unui detector se referă la performanța de a fi sigur de obiectul detecției și de a fi sigur, că s-au detectat toate obiectele de interes. În termeni tehnici, această caracteristică înseamnă o rată de detecție mare și o rată a detectărilor false mică, adică aproape nulă. Pentru a atinge aceste performanțe, s-a realizat un sistem format din două părți.

Prima parte este un sistem de detecție bazat pe aspect global, ce scanează imaginea de interes și utilizează tehnica piramidei de imagini. Caracteristica acestui sistem este o detecție foarte rapidă, cu o rată de detecție ridicată, dar cu un număr mare de detectări false. Rata de detecție ridicată s-a putut atinge prin crearea unei baze de date de imagini pozitive (subcapitolul 3.4), care ia în considerare aproape toate aspectele de apariție a obiectului de interes. Prin metodologia expusă (subcapitolul 3.4) s-a creat un clasificator propriu, cu performanțe mai bune decât ale clasificatorilor publici existenți. S-a obținut o rată foarte mică a detectărilor false, de aproximativ  $5 \cdot 10^{-6}$ , care totuși nu este suficient pentru o detecție robustă. Se detectează încă destul de multe zone false. Analizând imaginile false, ajungem la concluzia, că aceste imagini nu seamănă de loc cu obiectul de interes. Acesta este motivul, pentru care s-a propus, ca în faza a doua de detecție, să se utilizeze un detector mai lent, care să analizeze detaliile.

Partea a doua este un sistem de detecție bazat pe aspect local. Acesta este un sistem mult mai lent, dar prelucrează imaginea în detaliu. Sistemul analizat este alcătuit din două părți: modelul de obiect și descriptorul local. Modelul de obiect este un model deformabil cunoscut (subcapitolul 4.1), dar care se bazează pe informația descriptorului local propus. Modelul de obiect deformabil are soluționare eficientă prin algoritmul descris în subcapitolul 4.1.3. Eficientizarea algoritmului constă în aplicarea corectă a programării dinamice și evaluarea rapidă a transformatei de distanță generalizată cu un algoritm original (subcapitolul 4.1.4.1). Factorul primordial care determină performanța unui sistem de detecție bazat pe aspect local este descriptorul local și modul de alegere a punctelor de interes. Am pornit de la ideea de funcționare al câmpului receptiv al sistemelor biologice. Am propus un descriptor local bazat pe filtre 2D Gabor. În subcapitolul 4.2.1 am analizat atât teoretic, cât și experimental modul de creare al unui astfel de descriptor. Problema definirii limitelor filtrelor, modul de evaluare și realizarea efectivă a descriptorului local am realizat printr-o succesiune de experimente. Metodologia de abordare, munca sistematică și migăloasă s-a concretizat într-o serie de experimente reușite, expuse în subcapitolul 4.2.3. Rezultatele experimentale sunt concludente



și conturează posibilitățile de utilizare a descriptorului propus într-un model deformabil.

Sistemul de detecție realizat este un sistem supervizat, ce necesită o bază de date de imagini acoperitoare ipostazelor de apariție a obiectului de detectat. Prin modul de abordare sistemul încearcă să găsească un compromis optim între aspectul global și local, între prelucrarea globală și cea detaliată, între viteza și performanța de detecție. Analiza teoretică și experimentală, prezentată în lucrare, este o bază solidă pentru crearea unui sistem de detecție performant.

Parametrii procesului de învățare și ai procesului de detecție se pot îmbunătăți prin dezvoltări ulterioare. Se impune reducerea timpului de procesare determinat de numărul mare de convoluții și durata acestora. Durata unei convoluții se poate reduce prin dimensionarea optimă a filtrelor, prin utilizarea de coeficienți de tip întreg cu un nivel de cuantizare optimizat. Totodată, se impune utilizarea unei procesări multisistem cu sarcini de calcul distribuite judicios. Odată cu eficientizarea calculelor, se va putea aborda realizarea unui sistem de detecție mai complex, ce poate trata obiectele din diferite unghiuri de vedere sau cu acoperiri parțiale, într-o varietate infinită de ipostaze și contexte.

În afară de dezvoltările imediate ale acestei lucrări se pot aborda o serie de subiecte de cercetare viitoare:

- crearea unor modele de detecție ce au la bază sisteme de învățare semisupervizate și nesupervizate;
- crearea unor descriptori locali ce sunt alcătuiți din informația vizuală a petelor;
- studierea complexă a filtrelor 2D Gabor în vecinătatea punctelor de interes;
- crearea unor sisteme de recunoaștere utilizabile în identificare și autentificare.

# Bibliografie

- [1] [www.fotovision.ro](http://www.fotovision.ro). [Online; accessed 2012]. (document)
- [2] <http://www.mathworks.com/>. [Online; accessed 2012]. (document)
- [3] [http://www.itl.nist.gov/iad/humanid/feret/feret\\_master.html](http://www.itl.nist.gov/iad/humanid/feret/feret_master.html). [Online; accessed 2012].
- [4] <http://vision.ucsd.edu/~leekc/ExtYaleDatabase/ExtYaleB.html>. [Online; accessed 2012].
- [5] <http://coreldraw.com/media/>. [Online; accessed 2012].
- [6] Bioid database. <https://support.bioid.com/downloads/facedb/index.php>. [Online; accessed 2012].
- [7] P. Addison, J. Walker, and R. Guido. Time–frequency analysis of biosignals. *Engineering in Medicine and Biology Magazine, IEEE*, 28(5):14–29, september–october 2009.
- [8] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1st ed. 2006. corr. 2nd printing edition, October 2007.
- [9] G. Borgefors. Hierarchical Chamfer matching: A parametric edge matching algorithm. *IEEE Trans. Pattern Anal. Mach. Intell.*, 10:849–865, November 1988.
- [10] G. Borgefors and I. Nyström. Efficient shape representation by minimizing the set of centres of maximal discs/spheres. *Pattern Recognition Letters*, 18(5):465–471, 1997.
- [11] G. Bradski and A. Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O’Reilly, Cambridge, MA, 2008. (document)
- [12] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.*, 2(2):121–167, 1998.
- [13] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, June 1986.
- [14] O. Cuisenaire. *Distance Transformations: Fast Algorithms and Applications to Medical Image Processing*. PhD thesis, Universite Chatolic de Louvain, 1999.

- [15] M. N. Dailey, G. W. Cottrell, C. Padgett, and R. Adolphs. Empath: A neural network that categorizes facial expressions. *J. Cognitive Neuroscience*, 14:1158–1173, November 2002.
- [16] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893, 2005.
- [17] P. E. Daniellson. Euclidian distance mapping. *Computer Vision Graphics and Image Processing*, 14:227–248, 1980.
- [18] J. Daugman. Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proceedings of the IEEE*, 94(11):1927–1935, nov. 2006.
- [19] J. G. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20(10):847–856, 1980. ([document](#))
- [20] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America A: Optics, Image Science, and Vision*, 2(7):1160–1169, 1985. ([document](#))
- [21] J. G. Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics Speech and Signal Processing*, 36(7):1169–1179, 1988.
- [22] R. L. De Valois and K. K. De Valois. Spatial vision. *Annual Review of Psychology*, 31(1):309–341, 1980.
- [23] H. Eggers. Two fast euclidean distance transformations in z2based on sufficient propagation. *Computer Vision and Image Understanding*, 69(1):106–116, 1998.
- [24] R. Fabbri, L. D. F. Costa, J. C. Torelli, and O. M. Bruno. 2D euclidean distance transform algorithms: A comparative survey. *ACM Comput. Surv.*, 40:2:1–2:44, February 2008.
- [25] H. Farzin, H. A. Moghaddam, and M.-S. Moin. A novel retinal identification system. *EURASIP J. Adv. Sig. Proc.*, 2008, 2008.
- [26] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99(PrePrints), 2009.

- [27] P. F. Felzenszwalb and D. P. Huttenlocher. Distance transforms of sampled functions. Technical report, Cornell Computing and Information Science, September 2004. ([document](#))
- [28] P. F. Felzenszwalb and D. P. Huttenlocher. Pictorial structures for object recognition. *Int. J. Comput. Vision*, 61(1):55–79, 2005. ([document](#))
- [29] R. Fergus. Object class recognition by unsupervised scale-invariant learning. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 264–271, 2003.
- [30] R. Fergus. *Visual Object Category Recognition*. PhD thesis, Robotics Research Group Department of Engineering Science, University of Oxford, 2005.
- [31] M. A. Fischler and R. A. Elschlager. The representation and matching of pictorial structures. In *Computers, IEEE Transactions on*, volume C-22, pages 67–92, 1973. ([document](#))
- [32] Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, vol. 55, pp. 119-139, Art. No. SS971504,, 1997. ([document](#))
- [33] J. Friedman, T. Hastie, and R. Tibshirani. Additive Logistic Regression: a Statistical View of Boosting. *The Annals of Statistics*, 38(2), 2000. ([document](#))
- [34] D. Gabor. Theory of communication. part 1: The analysis of information. *Electrical Engineers - Part III: Radio and Communication Engineering, Journal of the Institution of*, 93(26):429 –441, 1946.
- [35] C. Galleguillos and S. Belongie. Context based object categorization: A critical survey. Technical Report CS2008-092, 08/2008 2008. ([document](#))
- [36] D. Gao. *A Discriminant Hypothesis for Visual Saliency: Computational Principles, Biological Plausibility and Applications in Computer Vision*. PhD thesis, University of California, 2008.
- [37] C. Harris and M. Stephens. A Combined Corner and Edge Detection. In *Proceedings of The Fourth Alvey Vision Conference*, pages 147–151, 1988.
- [38] C. Huang, B. Wu, H. Ai, and S. Lao. Omni-directional face detection based on real adaboost. In *In International Conference on Image Processing*, 2004.

- [39] D. Huttenlocher, G. Klanderman, and W. Rucklidge. Comparing images using the hausdorff distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(9):850–863, sep 1993.
- [40] J. Ilonen, J.-K. Kamarainen, P. Paalanen, M. Hamouz, J. Kittler, and H. Kälviäinen. Image feature localization by multiple hypothesis testing of Gabor features. *IEEE Transactions on Image Processing*, 17(3):311–325, 2008. ([document](#))
- [41] Y. Jie, Y. Y. fang, Z. Renjie, and S. Qifa. Fingerprint minutiae matching algorithm for real time system. *Pattern Recognition*, 39(1):143–146, 2006.
- [42] J. Jones and L. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in the cat striate cortex. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 58:1,233–1,258, 1987.
- [43] A. Jorgensen. Adaboost and histograms for fast face detection. Master’s thesis, University of Stockholm, 2006.
- [44] T. Kadir and M. Brady. Saliency, scale and image description. *Int. J. Comput. Vision*, 45(2):83–105, 2001.
- [45] P. Kalocsai. Face recognition by statistical analysis of feature detectors. *Image and Vision Computing*, 18(4):273–278, 2000. ([document](#))
- [46] Y. Ke and R. Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors. In *Computer Vision and Pattern Recognition, CVPR*, pages 506–513, 2004.
- [47] T. S. Lee. Image representation using 2D Gabor wavelets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(10):959–971, Oct. 1996.
- [48] S. Lefkovits. Performance analysis of face detection systems based on haar features. In *Complexity and Intelligence of the Artificial and Neural Complex Systems*, volume 1, pages 184–192, 2008. ([document](#))
- [49] S. Lefkovits. Assessments of building classifiers for face detection. In *International Conference On Recent Achievements in Mechatronics, Automation, Computer Sciences and Robotics MACRO*, pages 175–186, 2009. ([document](#))
- [50] S. Lefkovits. Teaching improvements on Haar-based classifiers. In *International Conference IETM-3, Petru Maior University, Tîrgu Mureş*, pages 1050–1060, 2009. ([document](#))

- [51] S. Lefkovits. Classification mixture for object detection. In *National Conference ZAC Cluj Academic Days*, pages 5–10, 2010. ([document](#))
- [52] S. Lefkovits. Algoritmi de detecție facială – studiu comparativ. Technical report, Babeș-Bolyai Cluj-Napoca, 2011. ([document](#))
- [53] S. Lefkovits. Combining boosted global- and part-aspect face detectors. *Scientific Bulletin of the Petru Maior University*, 8(1):35–40, 2011. ([document](#))
- [54] S. Lefkovits. Numerical computation method of the general distance transform. *KEPT 2011 International Conference Knowledge Engineering Principles and Techniques*, 1(1):143–152, 2011. ([document](#))
- [55] S. Lefkovits. Numerical computation method of the general distance transform. *Studia Informatica Universitatis Babeș-Bolyai*, 53(2):68–74, 2011. ([document](#))
- [56] S. Lefkovits. Hybrid face detector based on boosted classifiers. In *The Eighth Conference of PhD Students in Computer Science, University of Szeged*, volume 1, 2012. ([document](#))
- [57] S. Lefkovits. Improvements on Gabor descriptor retrieval for patch detection. *Computing and Informatics*, 2012. ([document](#))
- [58] S. Lefkovits. Novel Gabor filter based patch descriptor. In *10th Jubilee International Symposium on Intelligent Systems and Informatics SISY, IEEE Conference*, volume 1, 2012. ([document](#))
- [59] S. Lefkovits and C. Enăchescu. Face detection system based on artificial intelligence. In *International Conference ICELM, Petru Maior University Tîrgu-Mureș Computer Science section, Decision Systems, Mathematical Modelling and Statistics*, pages 112–121, 2008. ([document](#))
- [60] B. Leibe. *Interleaved Object Categorization and Segmentation*. PhD thesis, Swiss Federal Institute of Technology Zurich, 2004.
- [61] R. Lienhart, A. Kuranov, and V. Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. pages 297–304. 2003. ([document](#))
- [62] R. Lienhart, L. Liang, and E. Kuranov. A detector tree of boosted classifiers for real-time object detection and tracking. In *IEEE ICME2003*, pages 277–280, 2003.
- [63] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In *IEEE ICIP 2002*, pages 900–903, 2002. ([document](#))

- [64] C. Liu. Gabor-based kernel PCA with fractional power polynomial models for face recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(5):572–581, may 2004.
- [65] P. M. Lopez. *Object Component Models using Gabor Filters for Visual Recognition*. PhD thesis, Technical University Lisbon, 2008.
- [66] R. A. Lotufo and F. A. Zampiroli. Fast multidimensional parallel euclidean distance transform based on mathematical morphology. *Graphics, Patterns and Images, SIB-GRAPI Conference on*, 0:100, 2001.
- [67] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.
- [68] S. Marcelja. Mathematical description of the responses of simple cortical cells. *J Opt Soc Am*, 70(11):1297–1300, Nov. 1980.
- [69] C. R. Maurer, Jr., R. Qi, and V. Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25:265–270, February 2003.
- [70] A. Meijster, J. B. T. M. Roerdink, and W. H. Hesselink. A general algorithm for computing distance transforms in linear time. In J. Goutsias, L. Vincent, and D. S. Bloomberg, editors, *Mathematical Morphology and its Applications to Image and Signal Processing*, volume 18 of *Computational Imaging and Vision*, pages 331–340. Springer US, 2002. ([document](#))
- [71] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark*, pages 128–142. Springer, 2002. Copenhagen.
- [72] J. R. Movellan. Tutorial on Gabor filters. ([document](#))
- [73] J. Mundy. Object recognition in the geometric era: A retrospective. pages 3–28. 2006.
- [74] M. J. P. Viola. Fast multi-view face detection. Technical Report TR2003-096, Mitsubishi Electric Research Laboratories, Cambridge, July 15, 2003.
- [75] D. W. Paglieroni. Distance transforms: Properties and machine vision applications. *CVGIP: Graphical Models and Image Processing*, 54(1):56 – 74, 1992. ([document](#))

- [76] C. P. Papageorgiou, M. Oren, and T. Poggio. 1998, a general framework for object detection. *Computer Vision, 1998. Sixth International Conference on*, pages 555–562, 1998. ([document](#))
- [77] D. Pollen and S. Ronner. Phase relationship between adjacent simple cells in the visual cortex. *Science*, 212:1409–1411, 1981.
- [78] E. Robert. A Brief Introduction to Boosting. *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 1999. ([document](#))
- [79] A. Rosenfeld and J. Pfaltz. Distance functions on digital pictures. *Pattern Recognition*, 1(1):33 – 61, 1968.
- [80] H. A. Rowley, S. Baluja, and T. Kanade. Rotation invariant neural network-based face detection. pages 38–44, 1998. ([document](#))
- [81] Y. Sato. A Gabor wavelet pyramid based object detection algorithm. In *International Symposium on Neural Networks ISNN*, 2011.
- [82] R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.*, 37:297–336, December 1999.
- [83] H. Schneiderman. *A Statistical Approach to 3D Object Detection Applied to Faces and Cars*. PhD thesis, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, May 2000. ([document](#))
- [84] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:411–426, 2007.
- [85] L. Shen, L. Bai, D. Bardsley, and Y. Wang. Gabor feature selection for face recognition using improved Adaboost learning. In *Proceedings of International Workshop on Biometric Recognition System, in conjunction with ICCV'05*, pages 39–49, 2005. ([document](#))
- [86] F. Shih and Y.-T. Wu. The efficient algorithms for achieving euclidean distance transformation. *Image Processing, IEEE Transactions on*, 13(8):1078 –1091, 2004.
- [87] T. Tamminen and J. Lampinen. Sequential Monte Carlo for bayesian matching of objects with occlusions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(6):930 –941, june 2006.



- [88] M. Toews and T. Arbel. Detection over viewpoint via the object class invariant. In *ICPR '06: Proceedings of the 18th International Conference on Pattern Recognition*, pages 765–768, Washington, DC, USA, 2006. IEEE Computer Society.
- [89] M. Toivanen and J. Lampinen. Incremental bayesian learning of feature points from natural images. *Computer Vision and Pattern Recognition Workshop*, 0:39–46, 2009.
- [90] T. Tuytelaars and K. Mikolajczyk. *Local Invariant Feature Detectors: A Survey*. Now Publishers Inc., Hanover, MA, USA, 2008.
- [91] M. Vidal-Naquet and S. Ullman. Object recognition with informative features and linear classification. In *ICCV*, pages 281–288, 2003.
- [92] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision - to appear*, 2002. ([document](#))
- [93] D. Vukadinovic and M. Pantic. Fully automatic facial feature point detection using Gabor feature based boosted classifiers. In *IEEE International Conf. on Systems, Man and Cybernetics 2005*, pages 1692–1698, September 2005. ([document](#))
- [94] B. Wang. CMU VASC image database. [http://vasc.ri.cmu.edu/idb/images/face/frontal\\_images/](http://vasc.ri.cmu.edu/idb/images/face/frontal_images/). [Online; accessed 2012]. ([document](#))
- [95] X. Wang and H. Qi. Face recognition using optimal non-orthogonal wavelet basis evaluated by information complexity. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 164 – 167 vol.1, 2002.
- [96] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:775–779, 1997.
- [97] Y. Zhou, Y. Li, Z. Wu, and M. Ge. Robust facial feature points extraction in color images. *Eng. Appl. Artif. Intell.*, 24(1):195–200, Feb. 2011.