

Nr. Inv.

Dulap Lit.

TEZA DE DOCTORAT

Diana Lungeanu

**MODELAREA PROCESELOR VIZUALE PRIMARE.
PERCEPTIA PROFUNZIMII SI INTEGRAREA CONTURURILOR**

Conducatori stiintifici:

Prof. dr. Gh. I. Mihalaș

Prof. dr. ing. Ștefan Holban

Timisoara
2000

CUVANT DE MULTUMIRE

As vrea acum, la ceas de bilant, sa-mi exprim recunostinta stinta si cele mai alese multumiri dascalilor mei, care m-au indrumat si m-au format ca om si ca cercetator in anii de scoala si dupa aceea.

Mai ales, le multumesc din suflet celor care mi-au condus cu profesionalism, rabdare si intelegere munca de pregatire a prezentei teze: profesor dr. Gh. I. Mihalaş si profesor dr. ing. Ştefan Holban.

Domnul Profesor Holban mi-a indrumat pasii inca de cand eram studenta la Facultatea de Automatica si Calculatoare in anii '80 si l-am simtit mereu aproape in toti anii care au urmat, incurajandu-ma si sprijinindu-ma.

Domnul Profesor Mihalaş mi-a acordat increderea dumnealui chiar de la inceput, cand am venit la Universitatea de Medicina si mi-a sustinut cu rabdare pasii timizi in lumea stiintifica si cea universitara.

Trebuie sa multumesc colegilor mei mai tineri, absolventi ai Facultatii de Automatica si Calculatoare, Carmen Popa-Lile, Viviana Cotîrlea si Sergiu Hotca. Ei au fost cei care m-au ajutat sa pun in practica algoritmii despre care este vorba in aceasta lucrare. Pe scurt, ei au fost colegii mei de echipa.

Totodata, trebuie sa multumesc tuturor prietenilor si colegilor care m-au sustinut cu sugestii si discutii, dar si cu sprijin efectiv sau, pur si simplu, prin exemplul personal. Ei sunt asa de multi ca nu-i pot aminti aici, dar le multumesc din toata inima.

Omului care mi-a fost alaturi in toti acesti ani, sotului meu, nici nu stiu cum sa-i multumesc...

BIBLIOTECA CENTRALĂ
UNIVERSITATEA "POLITEHNICA"
TIMIŞOARA

627699
781 E

CUPRINS

Capitolul 1: INTRODUCERE. OBIECTUL SI CONTINUTUL TEZEI	3
1.1 Obiectul si continutul tezei	6
Capitolul 2: SISTEMUL VIZUAL UMAN. ASPECTE ANATOMICE SI FIZIOLOGICE	8
2.1 Ochiul uman	8
Celulele ganglionare	11
Consecinte ale comportamentului campurilor receptoare ganglionare	13
Sensibilitate si rezolutie	13
2.2 Caile vizuale centrale	14
Coliculi superiori	15
Nucleii geniculati laterali	16
Functiile nucleilor geniculati laterali	17
2.3 Cortexul vizual	17
Hartile ariei vizuale primare	18
Proprietatile functionale ale celulelor din aria vizuala primara	18
Celule simple si celule complexe	20
Celulele din V1 ca "feature detectors"	20
Capitolul 3: MODELAREA VAZULUI. REPREZENTAREA SI PROCESAREA INFORMATIEI VIZUALE	21
3.1 Nivele in sistemul nervos. Nivele in modelare	21
Nivele de analiza	21
Nivele de organizare	21
Nivele de procesare	22
3.2 Organizarea cortexului striat	22
Aspecte anatomice si fiziologice	22
Modelarea organizarii pe coloane	24
3.3 Reprezentarea si procesarea informatiei vizuale	25
Codificare locala sau reprezentare distribuita?	25
Reprezentarea informatiei vizuale	26
Procesarea informatiei vizuale in cortexul striat	27
Ce fel de imagini se prelucreaza la nivelul ariei vizuale primare?	28
Ce fel de imagini sunt prelucrate la nivele mai inalte?	28
Capitolul 4: PERCEPTIA ADANCIMII IN VEDEREA UMANA. MODELAREA VEDERII BINOCULARE	30
4.1 Perceptia adancimii in vederea umana	30
Vederea stereoscopica	30
Imagistica stereo. Considerente tehnice	31
Disparitatea binoculara. Evidente experimentale	33
4.2 Modelarea vederii binoculare	34
Modelul Marr&Poggio	35
4.3 Model "biologic" pentru harta de disparitate binoculara	37
Prezentarea principiului	37
Parametrii modelului	38
4.4 Rezultate. Discutii	39
Rezultate obtinute pe RDS-uri	39
Rezultate obtinute pe imagini "reale"	44
4.5 Concluzii	48

Capitolul 5: DE LA IMAGINI LA SUPRAFETE. MODELAREA PERCEPTIEI CONTURURILOR	50
5.1 Coloane de orientare in aria vizuala primara. Modelarea perceptiei contururilor.	50
Contururi in aria vizuala primara	50
Modele de perceptie a contururilor	50
5.2 Model pentru intarirea contururilor pornind de la muchii elementare	51
Descrierea modelului	51
Alegerea parametrilor retelei	53
5.3 Rezultate. Discutii	55
Imagine-test cu zgomot si fara zgomot	55
Prelucrarea unei imagini Mondrian	57
Alte rezultate	58
5.4 Concluzii	61
Capitolul 6: CONCLUZII GENERALE	63
Anexa 1	65
Date despre sistemul nervos uman	
Anexa 2	66
Organizarea cortexului vizual la maimuta	
Anexa 3	67
Ariile corticale Brodmann	
Anexa 4	68
Prezentarea aplicatiei Stereo-Vision	
Anexa 5	78
Prezentarea aplicatiei Vision-Visual Cortex	
BIBLIOGRAFIE SI REFERINTE	88

Capitolul 1: INTRODUCERE. OBIECTUL SI CONTINUTUL TEZEI

CUNOASTEREA, PRECUM SI REPREZENTAREA ACESTEIA IN CREIER – "CEL MAI MISTERIOS LUCRU DIN LUME" *William James (1842-1910)*

De ce arata lucrurile asa cum arata? Intrebarea pusa de psihologii curentului Gestalt (Albright 1994, Jordan&Russell 1999, Weintraub&Walker 1966) ramane extrem de incitanta. Lumea este plina de obiecte si evenimente care ofera o imensa bogatie de informatii. Fiintele umane sunt dotate cu "echipamente" specializate in colectarea acestor informatii si prelucrarea lor. In acest lant, al carui inceput il constituie conversia energiei fizice in impuls nervos si al carui final este perceptia, au loc o serie de transformari ce constituie obiectul de studiu al unor ramuri stiintifice foarte diverse: psihologie, psihofizica, fizica, biologie, neurologie, neurofiziologie, dar si inteligenta artificiala.

Pentru a fi perceputa, orice informatie din lumea inconjuratoare trebuie sa fie inregistrata de sistemul senzorial. Vernon Mountcastle sublinia ca "suntem prizonierii propriilor noastre creiere" (Sekuler&Blake 1990). De la creier sunt proiectate spre exterior milioane de fibre nervoase specializate pentru a achizitiona stari energetice ce descriu lumea exterioara: lumina, caldura, forta, compozitie chimica. Asta e tot ceea ce stim in mod direct, restul este "inferenta logica". Pe scara evolutiei, sistemul nervos (apoi creierul) a aparut odata cu necesitatea de miscare a organismului, atunci cand au aparut animalele. Iar creierul are nevoie de informatie despre mediul in care organismul se misca.

Cum are loc toata aceasta prelucrare, care sunt mecanismele, algoritmi implicati? Ceea ce se stie in mod clar este ca noi percepem doar o parte restransa din intregul univers in care ne miscam, ca celelalte fiinte de pe pamant percep "lumi" mai mult sau mai putin diferite de a noastra, functie de nevoile de supravietuire la care s-a adaptat sistemul lor senzorial. Chiar intre fiintele umane apar uneori deosebiri destul de substantiale.

Asa cum arata Thomas Papathomas (Papathomas 1995), luam vederea ca pe ceva de la sine inteles. Desi pare ca are loc instantaneu si fara efort, ea este rezultatul unor prelucrari complexe la nivelul creierului, implicand interactiunea unor grupuri mari de celule nervoase.

Mai mult, oamenii au aflat abia in acest secol ca nu ne nastem cu sistemul vizual asa cum este el la maturitate si ca "invatam" sa vedem lumea inconjuratoare (Held 1987), ca primele luni de viata sunt extrem de importante pentru dezvoltarea capacitatii de a percepe vizual corect mediul inconjurator.

Daca la inceput sistemele vizuale biologice ii interesau cu preponderenta pe cercetatorii din domeniul stiintelor biologice sau inrudite cu acestea, o data cu cercetarile lui David Marr in anii '70 la MIT, preocuparile legate de vederea biologica si-au facut loc in activitatea de cercetare din domeniul stiintelor computationale. Marr pornea de la premisa ca obiectele sunt definite de atributele lor (forma, marime, culoare, textura, etc.) si ca in cortexul uman exista module specializate pentru determinarea acestor atribute ale informatiei venite pe cale vizuala (Finkel 1994). Este acum pe deplin acceptata ideea ca procesele vederii necesita zeci de module specializate care sa ne permita "descoperirea" continutului imaginilor, lucru ce fusese descoperit deja in mod intuitiv de artistii Renasterii, care stiau ca atributele obiectelor pot fi redade in numeroase modalitati subtile, manuind intensitatea luminii, umbrele, contrastele, schimbarile de textura, etc. Marr si continuatorii lui definesc sistemul vizual uman ca un sistem de imagistica activa (A. Blake 1995) care preia si prelucreaza doar ceea ce este "interesant" pentru problema pe care o are de rezolvat la un moment dat. Pentru o teorie completa a structurii si functionarii creierului, trebuie sa se mearga dincolo de interpretarea datelor anatomice si fiziologice, e necesara o analiza a sarcinii pe care el o are de indeplinit intr-o abordare interdisciplinara (Marr 1982).

In domeniul neuro-stiințelor exista o extrem de mare diversitate in abordare și, deci, de opinii privind mecanismele biologice implicate. Ceea ce surprinde inșă este ca cele mai multe dintre modelele construite (fiecare cu mijloacele proprii) de psiho-fizicieni, electrofiziologi, sau specialiști in inteligenta artificiala se suprapun in mare masura, toate converg spre concluzii similare, sau cel puțin nedivergente. Asta ne conduce la speranta ca aceste secrete vor fi curând descoperite. Dar de ce modelare?

DE CE MODELARE?

De ce să ne mai complicăm cu modele? De ce să nu facem experimente și să înregistrăm observațiile? Pentru ca există motive in favoarea ideii de modelare ce merg dincolo de neuroștiințe in particular, ele se aplica științei, in general.

Modelele ne ajuta să organizăm datele și motivează experimentele – ele sugerează modul in care datele se potrivesc pentru a conduce spre o explicare a fenomenului. E bine, de aceea, să avem un model (chiar imperfect) decât să nu avem nici unul. In fapt, oamenii de știință au întotdeauna ipoteze care-i motivează in cercetarea întreprinsă, chiar atunci când acestea nu sunt foarte clare sau bine articulate; un model cantitativ este in acest caz un pas înainte pentru ca implica utilizarea combinată a acestor raționamente sau supozitii disparate și, deci, analiza critica a motivelor pentru care ele ar putea fi adevărate sau nu.

Un alt motiv este acela ca modelele conferă mai multa plauzibilitate presupunerilor dacă ele "supraviețuiesc" testelor experimentale (simularilor). Mai ales in faza de pionierat a unei discipline științifice, când datele sunt disparate, testarea unor modele și ipoteze poate fi un pas important înainte. Dacă modelul trece testele, devine mai plauzibil și poate fi testat cu noi date. Dacă acest lucru nu se întâmplă, dacă modelul se dovedește eronat, el constituie punctul de plecare pentru noi ipoteze, pentru un nou model (Erwin et al. 1995).

O caracterizare frecventă a neuroștiințelor este: "bogate in date, dar sărace in teorie". Acest lucru rămâne valabil atât timp cât nu știm încă cum anume vedem, învățăm, luăm decizii și acționăm. Dar cercetarea in domeniu avansează cu mare repeziție și o mare bogăție de date experimentale asteapta să fie sistematizate și incorporate in modele plauzibile, care să impulsioneze și să direcționeze noi investigații.

In neuroștiințele computaționale, distingem două idei de bază: notiunea de *nivele* și aceea de *co-evoluție* a cercetării la diferite nivele (Churchland & Sejnowski 1992). La nivel cerebral există atât o organizare pe scară largă, cât și una la scară mică. La nivele de jos (canalele de ioni, membrana neuronului, sinapsa nervoasă, etc.) sau la nivele înalte (sisteme cognitive) au loc funcții foarte diferite. Sunt necesare abordări diferite pentru a studia modul in care se sumează semnalele in dendrite, sau pentru a explica interacțiunea neuronilor dintr-o rețea, sau, mai departe, a rețelelor într-un sistem.

Trebuie înșă ca teoriile dezvoltate la un nivel de abordare să fie compatibile cu cele de la nivelele superioare sau inferioare. Inconsistențele din lant sunt create de o greșită înțelegere a unora din fenomene. Creierul este un ansamblu de celule (in cele din urmă) și, in mod sigur, ceva nu este in regula dacă neuronii din unele modele sunt incompatibili cu neuronii din altele.

Organizarea pe nivele este o stare de fapt in cercetare, dar și o strategie ce permite evoluția in paralel pe mai multe direcții și, până la urmă, colaborarea cercetătorilor din diferite domenii: cercetarea in neuro-psihologie și psihofizică (in psihologia experimentală, in general) furnizează caracterizări detaliate pentru ceea ce trebuie explicat la nivele inferioare de abordare. In același timp, cercetarea privind principiile computaționale conduce la progres in neuro-biologie și psihologie prin faptul ca descoperiri relativ abstracte aduc sugestii privind mecanismele biologice, iar acestea, la rândul lor, sugerează soluții in construirea sistemelor artificiale.

MODELE REALISTE SAU ABSTRACTE?

In mod evident, dorim ca un model sa incorporeze cat mai bine realitatea modelata. Totodata, trebuie amintit faptul ca nu intotdeauna un grad ridicat de realism conduce si la o valoare stiintifica ridicata. Diverse modele sunt utile in diverse circumstante.

Cateodata pot fi mai utile modele simplificate, deoarece un exces de realism (pana la nivel biochimic) poate conduce la imposibilitatea unei analize a dinamicii la nivel de sistem, sau pur si simplu la imposibilitatea implementarii si testarii lui pe calculatoarele actuale.

Pe de alta parte, atunci cand se studiaza dinamica la nivelul dendritelor sau sinapselor, cu cat ele sunt redate cu mai mult realism, cu atat mai utile vor fi concluziile trase in final.

Desigur, nu exista reguli universal valabile si fiecare situatie trebuie gandita si abordata in mod specific.

Modelele descrise in aceasta lucrare sunt abordate la nivelul ariei vizuale primare, deci la nivel de retea neuronală.

MOTIVE PRACTICE PENTRU STUDIAREA VEDERII

Pe langa imboldul natural de curiozitate, exista si cateva motive practice pentru care studierea mecanismelor perceptiei (in general) si a perceptiei vizuale (in particular) se justifica.

Cunoscand aceste mecanisme se pot lua masuri pentru a asigura performante optime in ceea ce priveste perceptia in viata de zi cu zi. Exista un optim de stimulare a simturilor umane, atat in privinta naturii acestora, cat si a intensitatii. O lipsa a stimulării, sau o supra-stimulare pot conduce la deteriorari grave, chiar ireversibile ale sistemului nervos. Sistemul nostru nervos a evoluat intr-un mediu substantial diferit de cel al lumii moderne (aflat intr-o extraordinar de ridicata dinamica). Studiind mecanismele perceptiei, putem anticipa si preveni multe din situatiile de interactiune cu mediul care ar putea afecta abilitatea noastra de a lua deciziile corecte.

Mai mult, exista problema persoanelor al caror sistem vizual a fost afectat, maladii degenerative care conduc la pierderea completa a vederii in urma distrugerii fotoreceptorilor. Auzul, cel de-al doilea simt ce permite investigarea mediului la distanta, este deja protezabil (implantul de cohlee s-a realizat din anii 1970). Exista mari sperante ca vazul va fi urmatorul (Brown 1999, Chase 1999, Normann 1995). Universitati importante (Massachusetts Institute of Technology, Johns Hopkins University, Duke University, Universitatea din Bonn, Universitatea din Utah) sunt implicate in programe de cercetare privind implantul retinian. Faptul ca sperantele nu sunt hazardate este atestat si de interesul crescand al firmelor comerciale pentru aceasta problema (pana acum cativa ani programele au fost sustinute doar din fondurile de cercetare academica). Viitorul imediat se anunta optimist pentru cei ale caror celule ganglionare nu au fost afectate. Totusi, pana la implantul la nivel cerebral este inca o cale lunga...

Un alt domeniu de aplicatii este acela al realitatii virtuale. Se prevede ca, intr-un viitor nu prea indepartat, sistemele de realitate virtuala vor beneficia din plin de descoperirile privind perceptia vizuala umana (Pescovitz 1999).

1.1. OBIECTUL SI CONTINUTUL TEZEI

Cercetarea mea s-a focalizat asupra vederii umane primare, anume asupra mecanismelor de detectie a profunzimii si a contururilor in vederea umana. Am dezvoltat doua modele: unul de determinare a disparitatilor binoculare (necesare in vederea stereoscopica) si un model de intarire a contururilor pornind de la muchii elementare.

Fiind o abordare interdisciplinara, pentru a nu fragmenta prezentarea, am renuntat la structura clasica a unei teze de doctorat (cu prezentare generala a realizarilor in domeniu, urmata de contributiile proprii). Am considerat ca o alta organizare aduce un plus de claritate, fara a prejudicia acuratetea stiintifica. Voi rezuma aceste principii de organizare in cele ce urmeaza.

Lucrarea este structurata in sase capitole, cinci anexe si un capitol special de bibliografie si referinte:

Capitolul 1: Introducere. Obiectul si continutul tezei

Capitolul 2: Sistemul vizual uman. Aspecte anatomice si fiziologice

In acest capitol am prezentat foarte pe scurt cateva aspecte privind anatomia si fiziologia sistemului vizual. Scopul lui este de a oferi substratul necesar pentru urmarirea aspectelor biologice modelate. Unele dintre problemele specifice sunt reluate in detaliu in capitolele urmatoare.

Capitolul 3: Modelarea vazului. Reprezentarea si procesarea informatiei vizuale

Se poate considera ca acest capitol este dedicat unor consideratii generale. El defineste contextul general in care se incadreaza modelele realizate de mine si prezentate in aceasta lucrare. Sunt cuprinse cateva consideratii de baza privind modelarea mecanismelor biologice, urmate de o descriere a structurii cortexului striat, cu accentuarea functiei sale: reprezentarea si procesarea informatiei vizuale. Sunt apoi amintite modelele de organizare competitiva a coloanelor de dominanta oculara si orientare, precum si functiile Gabor, ce modeleaza comportamentul celulelor simple (care fac, in fapt, codarea informatiei vizuale). In final, este prezentata schema-bloc a procesarii la nivelul ariei vizuale primare, cu specificarea mecanismelor modelate de mine si descrise apoi in Capitolele 4 si 5.

Trebuie sa mentionez ca atat in acest capitol, cat si in urmatoarele am introdus cate o prezentare a bazelor biologice sau psihofizice pentru mecanismele modelate. Atata timp cat pornim de la premiza modelarii unor mecanisme biologice si a plauzibilitatii biologice, descrierea acestora in detaliu s-a impus de la sine. Am considerat ca gruparea lor intr-un capitol separat ar fi fost nepotrivita, deoarece ele sunt aspecte foarte specifice si nu tot sistemul vizual este prezentat la acest nivel de detaliu.

Capitolul 4: Perceptia adancimii in vederea umana. Modelarea vederii binoculare

In acest capitol sunt descrise aspectele abordate in modelarea vederii binoculare si anume detectia disparitatilor orizontale. La inceput am facut o prezentare a problemei pe care o are de "rezolvat" sistemul nostru vizual pentru detectia adancimii pornind de la disparitatile binoculare si o expunere a datelor experimentale. Sunt amintite principalele tipuri de modele ale vederii stereo, urmate de descrierea modelului realizat de mine, cu rezultate (Lungeanu 1995, Lungeanu et al. 1996, Lungeanu et al. 1998b, Lungeanu et al. 1998c) si discutii pe marginea acestora. Trebuie sa mai mentionez ca pentru implementarea si testarea modelului a fost realizata o aplicatie (descrisa in Anexa 4), ce implementeaza si modelul clasic de vedere stereoscopica al lui David Marr (pentru a avea un termen viabil de comparatie pentru rezultatele mele).

Capitolul 5: De la imagini la suprafete. Modelarea perceptiei conturilor

Acest capitol descrie aspectele abordate in integrarea conturilor pornind de la muchii elementare. Dupa prezentarea problematicii, se face descrierea modelului realizat de mine pentru intarirea conturilor, in fapt o retea recurenta ce modeleaza comportamentul celulelor complexe de la nivelul hipercoloanelor de orientare din cortexul striat. Sunt apoi prezentate rezultatele obtinute (Lungeanu et al. 1998, Lungeanu&Cotirlea 1999, Lungeanu et al. 2000), cu discutii pe marginea lor.

Mentionez faptul ca pentru implementarea si testarea modelului a fost realizata o aplicatie (descrisa in Anexa 5).

Capitolul 6: Concluzii generale

Acest capitol cuprinde concluzii generale privind mecanismele modulate si rezultatele obtinute, privite in perspectiva unor dezvoltari si integrari ulterioare.

Trebuie sa subliniez ca cele doua modele nu constituie abordari disparate. Amandoua se refera la mecanisme din vederea primara implicate in perceptia tri-dimensională a spatiului, domeniu asupra caruia mi-am concentrat munca de cercetare.

Anexa 1 cuprinde cateva date despre sistemul nervos uman.

Anexele 2 si 3 cuprind harti cu ariile corticale ce au fost discutate in lucrare .

Anexele 4 si 5 prezinta aplicatiile dezvoltate pentru implementarea modelelor descrise in teza (discutate in capitolele 4 si 5).

Bibliografie si referinte

Acest capitol cuprinde titlurile lucrarilor consultate, atat lucrari fundamentale pentru modelarea sistemului nervos si a vederii (Churchland&Sejnowski 1992, Marr 1982, Papathomas et al. 1995), cat si enciclopedii de prestigiu (Adelman 1987, Arbib 1995, Wilson&Keil 1999) care m-au ghidat si mi-au dat perspectiva asupra domeniului. Sunt cuprinse, de asemenea, articolele semnificative din diverse reviste (Neural Computation, Neural Networks, Trends in Neurosciences, etc.)

Am inclus aici si rezultatele cercetarii proprii comunicate si publicate, atat ca prim autor (9 lucrari), cat si co-autor (9 lucrari).

Capitolul 2: SISTEMUL VIZUAL UMAN. ASPECTE ANATOMICE SI FIZIOLOGICE

(Dupa Adelman 1987, Despopoulos 1991, Hall&Robinson 1998, Robinson 1998, Sekuler&Blake 1990, Shepherd 1988)

Sistemul vizual consta din trei sectiuni majore (figura 2.1): **ochii** (care capteaza lumina si genereaza impulsurile nervoase) **caile vizuale** (care transmit mesajul de la ochi catre creier) si **centrii vizuali din creier** (care interpreteaza mesajele primite). Deoarece toate contribuie la procesul vazului, pentru a intelege cum anume vedem, trebuie sa amintim pe scurt rolul fiecaruia. De asemenea, trebuie sa avem mereu in vedere conexiunea intima dintre structura si functie.

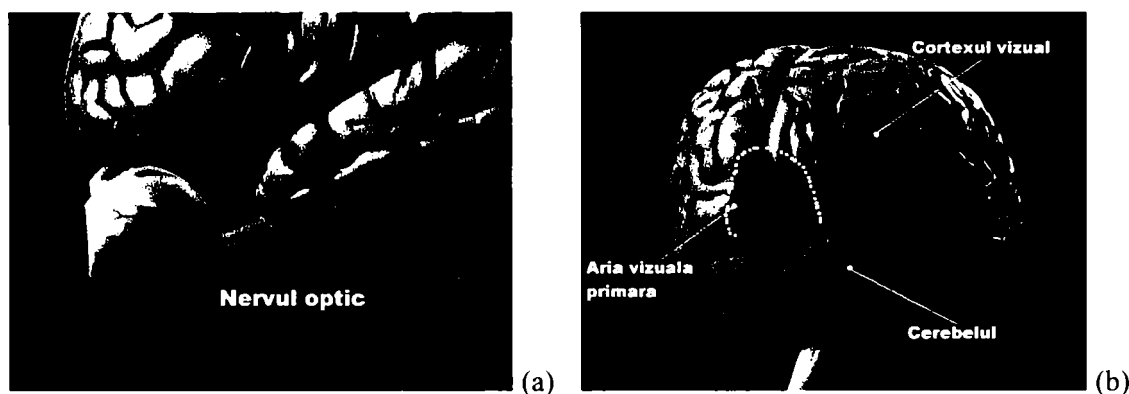


Figura 2.1: Principalele componente ale sistemul vizual uman: (a) ochiul si nervul optic (parte a cailor vizuale); (b) ariile vizuale corticale (Dupa Hall & Robinson 1998)

2.1. OCHIUL UMAN

Din punct de vedere embriologic, ochii sunt o prelungire a creierului – ii putem considera o proiectie a acestuia catre lumea exterioara. Importanta acordata informatiei ce vine pe cale vizuala a crescut continuu pe scara evolutiei, la om simtul vazului ocupand un loc deosebit de important (aproape 50% din suprafata neocorticala este implicata in procesarea informatiei vizuale). Trebuie sa mentionam si faptul ca, atunci cand intra in conflict cu alte simturi, vazul le domina pe celelalte: daca tinem in mana un obiect, iar simtul tactil si vazul "spun" lucruri contradictorii creierului, vom "da crezare ochilor".

Ochii detecteaza radiatia electromagnetica cu $\lambda=400 \text{ nm} \div 750 \text{ nm}$ (intre violet si rosu), ceea ce in mod obisnuit numim "spectrul vizibil". Sensibilitatea este maxima pentru $\lambda=510 \text{ nm}$ in vederea nocturna ("scotopica") si pentru $\lambda=555 \text{ nm}$ in vederea diurna (fotopica). Pentru vederea diurna, care permite distinctia culorilor, situatia este diferita pentru cele trei tipuri de celule cu conuri care au sensibilitati maxime pentru lungimi de unda diferite (functie de fotopigmentul utilizat): $\lambda=440 \text{ nm}$ (violet), $\lambda=550 \text{ nm}$ (verde-galbui), $\lambda=570 \text{ nm}$ (galben).

Ochii sunt asezati frontal (in zona orbitara) si furnizeaza la iesire, pe nervul optic, doua imagini partial suprapuse ale mediului inconjurator. **Campul vizual** al fiecarui ochi este constituit de segmentul vazut din lumea externa atunci cand consideram atat ochiul, cat si capul in pozitie fixa. Campul vizual se masoara prin perimetrul sau, care este, in principiu, o semisfera. Intreruperea partiala a campului vizual poarta numele de **scotom**. Putem considera ochiul ca fiind un aparat optic uimitor de precis creat de natura din materialele pe care le-a avut la dispozitie: apa, proteine, grasimi, zaharuri.

Ochiul este un organ aproape sferic, cu un diametru de aprox. 24 mm – figura 2.2. Muschii extraoculari (trei perechi) ne permit sa miscam globul ocular pentru a fixa obiectele care ne intereseaza. Actiunea lor concertata permite atat miscari de amplitudine mare, cat si miscari foarte

fine de fixatie, sau miscarea de nistagmus (asemanatoare celei in care urmarim peisajul de la fereastra unui tren in miscare).

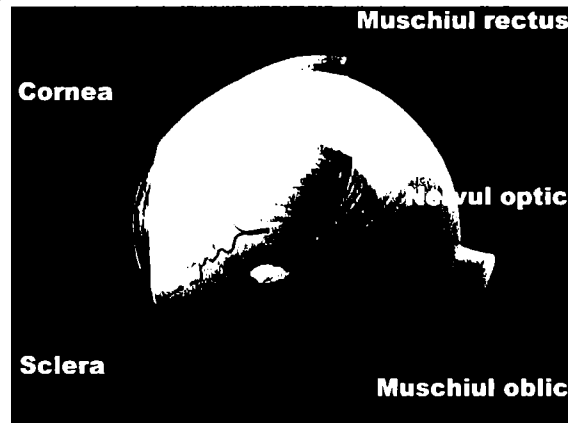


Figura 2.2: Muschii extraoculari ne permit sa miscam globul ocular (Dupa Hall & Robinson 1998)

Ochiul consta din trei straturi concentrice, fiecare cu aspect, structura si functie diferita – figura 2.3: **tunica fibroasa** (care protejeaza ochiul), **tunica vasculara** (care hraneste ochiul) si **retina** (care detecteaza lumina si o transforma in impulsuri nervoase. Ochiul contine doua compartimente: camera anterioara (sau umoarea apoasa) si corpul vitros (sau umoarea sticloasa). Trebuie sa mentionam ca, pentru fiecare componenta in parte, forma, transparenta si netezimea suprafetelor trebuie sa asigure obtinerea unei imagini nedistorsionate.

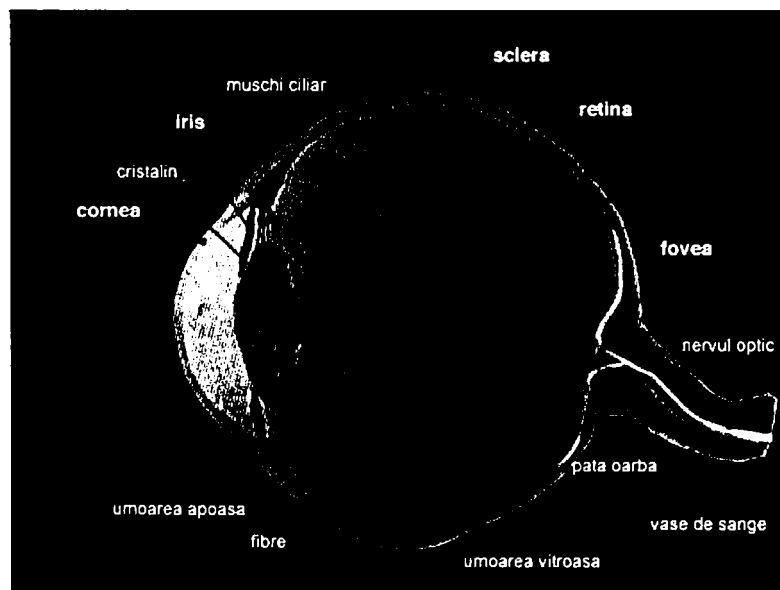


Figura 2.3: Ochiul ca instrument optic, cu invelisurile sale. (Dupa Hall & Robinson 1998)

Cristalinul este o lentila biconvexa de 9 mm ϕ si de 4 mm grosime, formata din proteine si apa. **Sclera**, sau tunica fibroasa, are rol de protectie. In partea frontala este transparenta si formeaza **cornea**. **Coroida**, sau tunica vasculara, are 0.2 mm grosime si este puternic pigmentata. Ea hraneste stratul interior al ochiului, retina.

Retina este formata din tesut de natura neuronal si pe ea este proiectata imaginea bidimensionala redusa si inversata obtinuta prin intermediul instrumentului optic care este ochiul. Asa cum ii spune si numele (*rete* = plasa de pescuit in limba latina), retina are o structura plexiforma si este formata din mai multe straturi de celule nervoase – figura 2.4. In faza embrionara, retina se dezvolta din acelasi tesut ca si creierul. Zona centrala a retinei (*macula lutea*) are in centru o "gropita" (*fovea centralis*) de ≈ 1.8 mm in diametru in care densitatea celulelor receptoare este maxima ($150000 \div 200000 /\text{mm}^2$). Zona foveala este si ea structurata in zone concentrice, zona centrala propriu-zisa (de 300 μm) numindu-se *foveola*. In zona foveala retina are o grosime minima pentru a se asigura o transparenta maxima si, deci, o acuitate vizuala maxima.

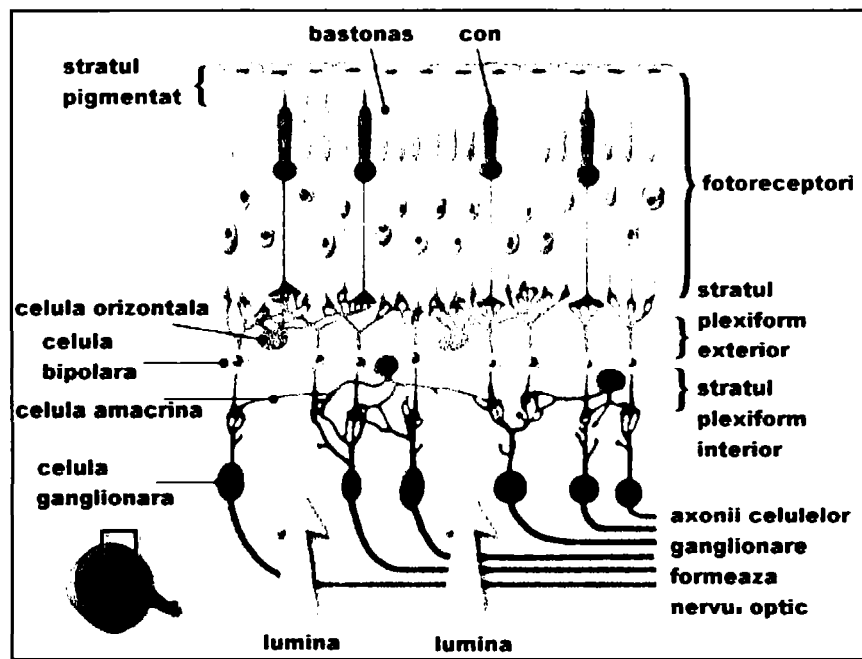


Figura 2.4: Secțiune transversala prin retina. De notat faptul ca celulele receptoare se gasesc in partea exterioara a retinei (spre coroida) si lumina strabate aceasta structura plexiforma pentru a ajunge la ele. (Dupa Hall & Robinson 1998)

Celulele receptoare fac transformarea semnalului luminos in impuls nervos. Ele sunt de doua tipuri: **conuri** (pentru vederea diurna) si **bastonase** (pentru vederea nocturna si cea periferica) – figura 2.5. Ochiul uman contine aprox. 120 milioane de bastonase si aprox. 7-8 milioane de conuri. In zona foveala se gasesc doar conuri, in timp ce densitatea lor scade puternic cu distanta, iar in afara zonei maculare nu mai exista conuri, ci doar bastonase. Densitatea bastonaselor este maxima la aprox. 7 mm de centrul foveal (densitate aproximativ egala cu cea a conurilor din zona foveei) si scade, si ea, spre zonele periferice ale retinei. Acuitatea vizuala in zona foveala este de aprox. 1 minut (un segment de 1.5 mm vazut de la distanta de 5 m) si ea scade rapid catre periferie.

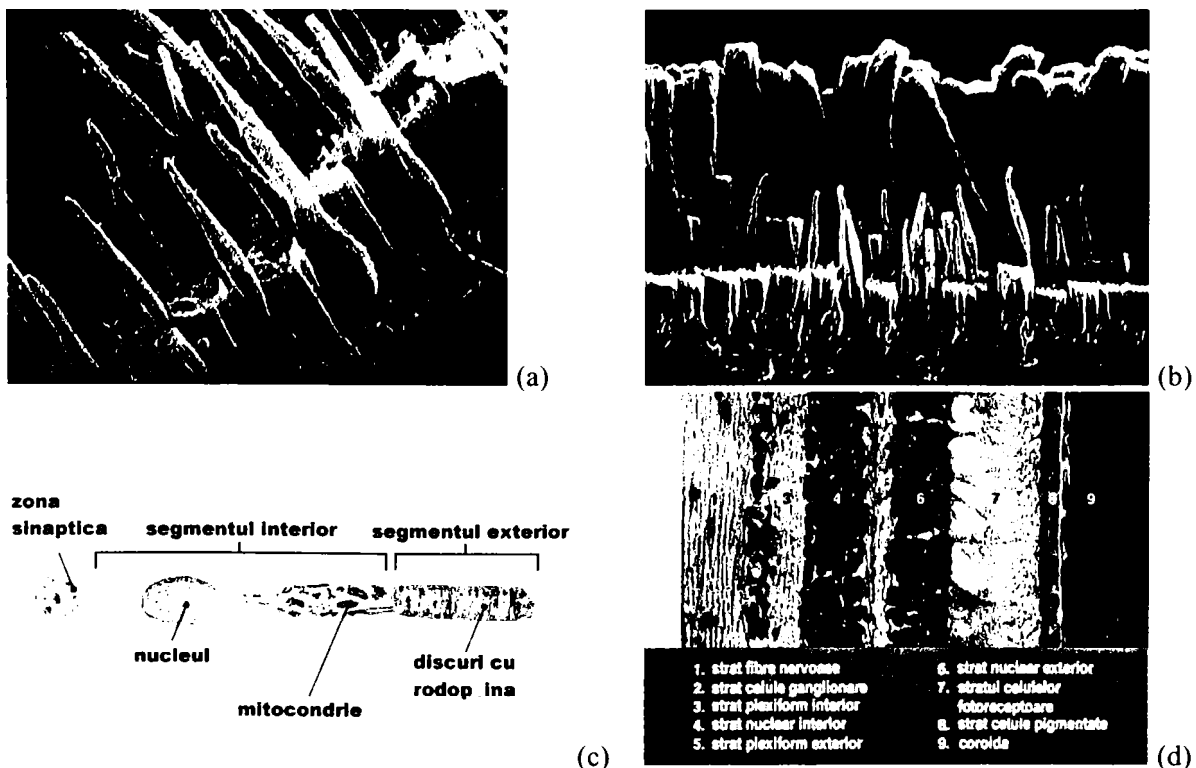


Figura 2.5: Imaginea retinei și ale celulelor receptoare: (a) conurile așa cum arată ele în realitate; (b) conuri și bastonase; (c) un bastonas schematizat, la care s-au pus în evidență segmentele funcționale; (d) aspectul retinei la microscop. (Dupa Hall & Robinson 1998)

La aprox. 20° spre zona nazala, apare un scotom natural – **discul optic**, zona in care axonii celulelor colectoare de informatie parasesc ochiul. Aceasta are ca urmare existenta unei "pete oarbe" in campul vizual. In mod obisnuit, in vederea binoculara, pata oarba a unui ochi este compensata de calalalt ochi.

Aceasta "matrice" de aprox. 130 milioane de fotoreceptori converteste imaginea retiniana intr-o imagine neuronală ce este transmisa creierului prin intermediul nervului optic. Fiecare celula fotoreceptoare cuantifica cantitatea de lumina pe care o primeste, iar celule colectoare de pe celelalte straturi retiniene condenseaza si reorganizeaza aceste mesaje primare. Reorganizarea reflecta adaptarea sistemului vizual la necesitatea de a detecta **diferentele** de iluminare si schimbarile ce survin in campul vizual, fiind mai putin importante zonele iluminate uniform sau in care nu intervine nici o modificare. Condensarea informatiei este substantiala, daca tinem cont de faptul ca exista aprox. 130 milioane de fotoreceptori si doar aprox. 1 milion de celule ganglionare, ai caror axoni formeaza nervul optic. Mai trebuie mentionat ca "decalajul" dintre momentul bombardarii fotoreceptorului cu fotoni si semnalul corespunzator in nervul optic este relativ mare (≈ 25 ms) si se datoreaza atat faptului ca fotoreceptorii sunt lenti, cat si celor doi pasi intermediari din retina (celulele bipolare, amacrine si horizontale) – apar minimum inca doua legaturi sinaptice suplimentare.

Daca ochii constituie "fereastră" noastra catre exterior, ei sunt si "fereastră" prin care putem vedea interiorul. Ochiul este singurul loc unde sistemul nervos si rețeaua sanguina pot fi vazute fara a utiliza metode invazive de examinare (de exemplu, folosind un oftalmoscop, dispozitiv inventat de Hermann Helmholtz in sec. 19). Figura 2.6 ilustreaza imaginea unui fund de ochi.

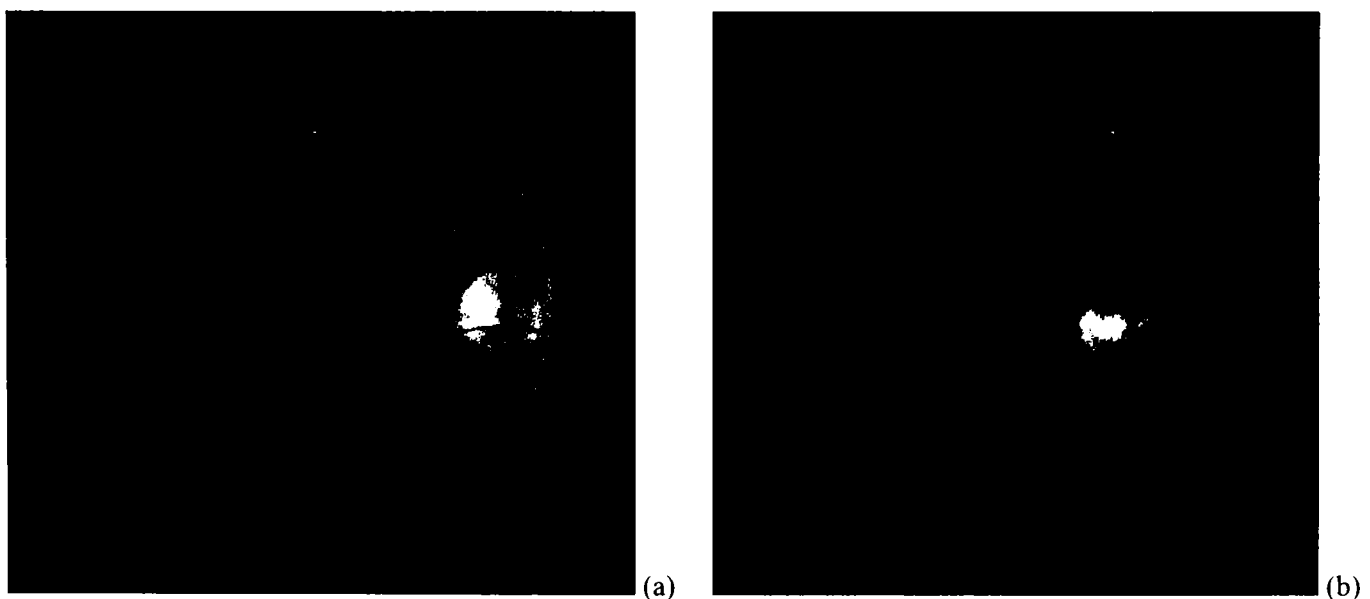


Figura 2.6: Imaginea ce poate fi vazuta utilizand un oftalmoscop. (a) Un fund de ochi normal, la care se poate observa rețeaua de vase ce hranesc retina, precum si faptul ca in zona centrala (macula) acestea lipsesc, pentru a permite o transparenta maxima. (b) Imagine in urma unui traumatism macular cauzat de lumina coerenta (pe aceasta imagine macula poate fi identificata mai usor). (Cabanis et al. 1996)

Retina este transparenta, asa ca practic se vizualizeaza structurile ce se gasesc in fata si in spatele acesteia (coroida). Se pot observa zona centrala (macula) si discul optic, cu vasele care patrund in aceasta zona in interiorul ochiului (aceasta rețea este unica pentru fiecare individ). Se observa ca in zona maculei nu exista vase (pentru o mai buna transparenta), aceasta zona fiind hranita exclusiv de substante nutritive ce sunt furnizate de coroida.

CELULELE GANGLIONARE

Zona retinei in care stimularea luminoasa conduce la raspunsul unui anumit neuron se numeste **camp receptor** al aceluia neuron. Celulele ganglionare au campul receptor de forma

circulara, simetrica, compus dintr-o zona centrala si una periferica, aranjate concentric. Cele doua zone raspund la diferentele de iluminare de o maniera antagonista. Exista celule centrate ON cu periferie OFF (ON-OFF) si celule centrate OFF cu periferie ON (OFF-ON) – figura 2.7. Acest tip de interactiune intre zone antagoniste se numeste **inhibitie laterala** si este rolul celulelor colectoare din straturile intermediare ale retinei sa realizeze acest comportament.

Trebuie precizat faptul ca, datorita simetriei campului receptor, aceste celule raspund la contrast si nu la orientarea stimulului. Centrele campurilor receptoare pentru celulele centrate ON si cele centrate OFF se suprapun, astfel incat mozaicul format acopera intreaga retina. Mai mult, atunci cand o zona limitata a retinei este stimulata, acest lucru afecteaza celule de ambele tipuri, producand efect opus in cele doua cazuri. Asta nu inseamna ca efectele se compenseaza reciproc, informatia de ambele tipuri este transmisa mai departe, spre zonele corticale.

Marimea campurilor receptoare variaza sistematic functie de localizarea retiniana. In zona maculara, campurile receptoare sunt mici, de ordinul a 0.01 mm. Spre periferie dimensiunea creste, astfel incat la aprox. 10 mm distanta de fovee dimensiunea lor ajunge la 0.5 mm. Pe langa aceasta variatie cu excentricitatea fata de fovee, exista o variatie locala a dimensiunii campurilor receptoare, ceea ce face ca pentru aceeasi locatie retiniana sa existe campuri receptoare de diferite dimensiuni (desigur, intre limite nu foarte largi pentru fiecare zona a retinei). Aceasta permite "captarea" unor stimuli de dimensiuni si facturi diferite (muchii mai fine sau mai largi). Acest principiu de organizare sugereaza faptul ca analiza dimensiunii obiectelor incepe la nivelul retinei, chiar daca perceptia propriu-zisa are loc la nivel cortical.

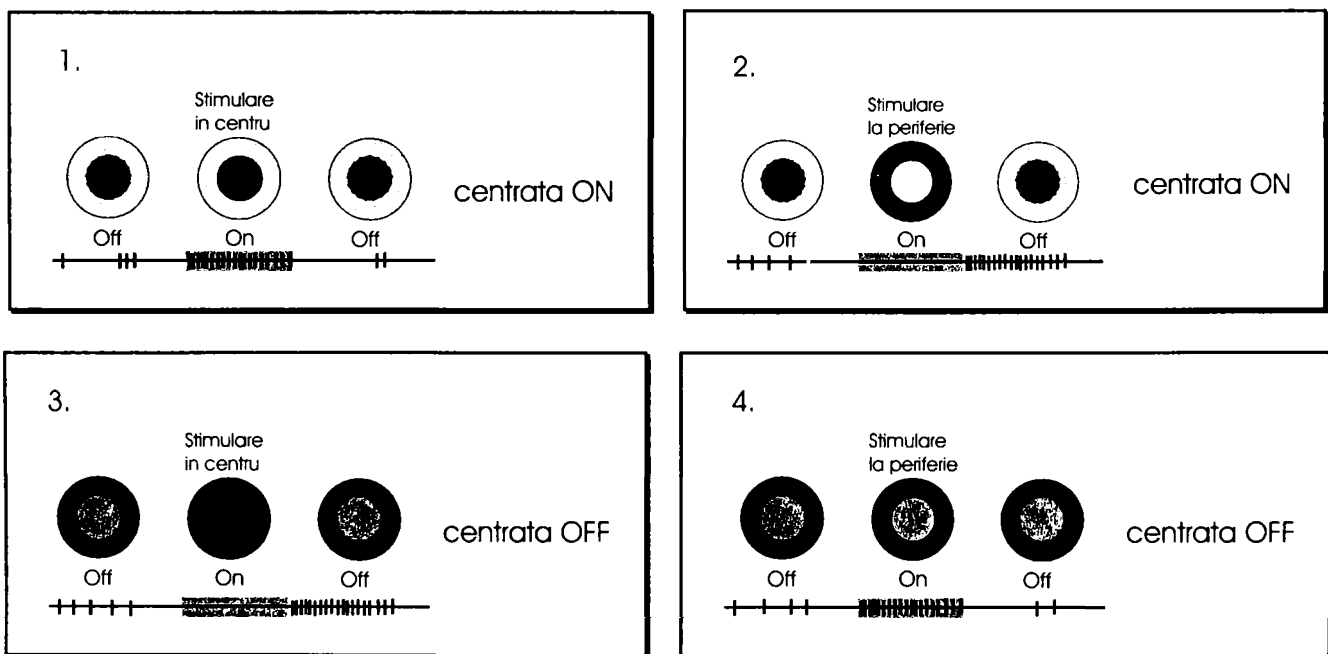


Figura 2.7: Campurile receptoare ale celulelor ganglionare ale retinei raspund la diferente lumina-intuneric sau iluminare-intreruperea acesteia.

Exista doua tipuri de celule ganglionare la nivelul retinei: celule de **tipul M** si celule de **tipul P**. diferentele dintre ele constau in urmatoarele:

- celulele M sunt mai mari decat cele P si au axonul mai gros, ceea ce conduce la o viteza mai mare a impulsurilor nervoase
- celulele P formeaza aprox. 80% din totalitatea celulelor ganglionare la primate
- pentru orice excentricitate, campurile receptoare ale celulelor P sunt de cateva ori mai mici decat al celulelor M, deci ele vor raspunde mai bine la obiecte mai mici
- celulele M raspund bine la diferente mici de iluminare intre zona centrala si periferia campului receptor, in timp ce celulele P necesita un contrast mai puternic
- celulele P raspund la stimuli colorati (cu sensibilitatea acordata pentru anumite lungimi de unda), in timp ce celulele M raspund la fel, indiferent de culoarea stimulului

Celulele P si M se proiecteaza in zone ale creierului "tinere" din punct de vedere filogenetic (aria vizuala corticala), in care are loc perceptia vizuala. Pe langa celule M si P (aprox. 90% din totalul celulelor ganglionare), mai exista celule ai caror axoni sunt proiectati spre zone mai "vechi" ale creierului, care raspund de miscarile ochilor si ale capului (coliculi superiori) – aceste celule nu pot fi categorisite ca fiind M sau P.

CONSECINTE ALE COMPORTAMENTULUI CAMPURILOR RECEPTOARE GANGLIONARE

Ca urmare a antagonismului dintre zona centrala si cea periferica a campului receptor, celulele ganglionare raspund la diferentele de iluminare. De aceea, se poate ca zone cu acelasi nivel de iluminare sa fie percepute diferit, functie de contextul in care se gasesc. Implicit, asta inseamna ca stimuli de intensitati diferite pot fi perceputi identic functie de contextul in care ei apar.

Un astfel de efect este pus in evidenta de "grilele Hermann", descrise pentru prima data de Ludimar Hermann in sec. 19 – figura 2.8. Se observa ca la intersecții apar "buline" cu contrast mai scazut, care dispar daca le fixam foarte atent (daca se proiecteaza exact in centrul foveale).

O alta situatie este cea din figura 2.9, in care zona centrala are aceeasi culoare, dar este perceputa diferit, functie de fondul pe care se gaseste.

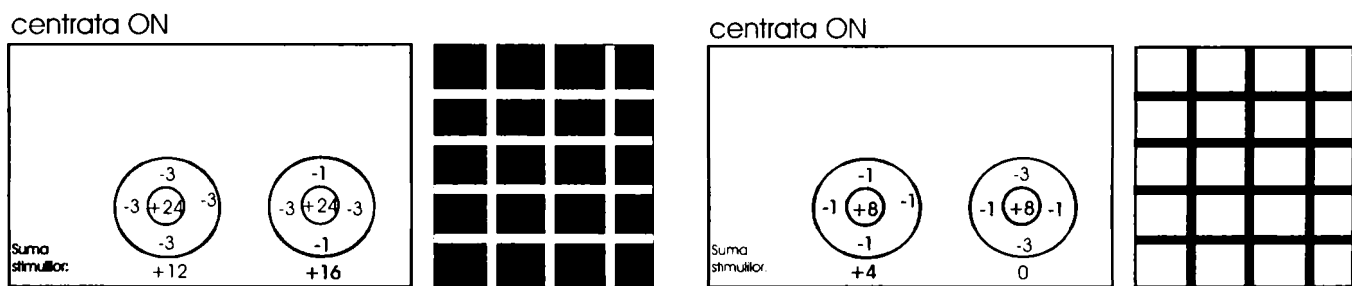


Figura 2.8: Ilustrarea efectului "grilelor Hermann" si a mecanismului prin care se ajunge ca diferenta de stimulare dintre periferie si centru sa fie diferita la celulele ce "raspund" de liniile verticale/orizontale fata de celulele "raspuzatoare" de intersecții Stanga: totalul este mai scazut la intersecții. Dreapta: totalul este mai ridicat la intersecții.

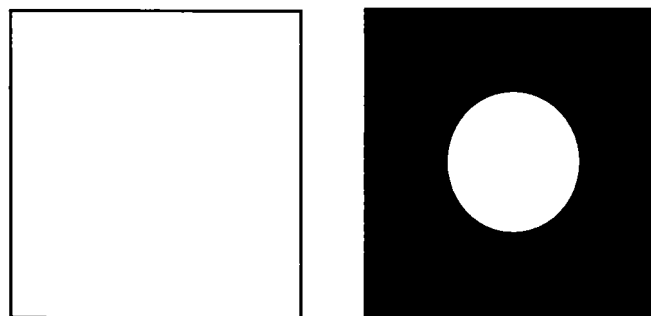


Figura 2.9: Ilustrarea faptului ca doua zone cu acelasi nivel de iluminare (centrul) sunt percepute ca avand luminozitate diferita, functie de context.

SENSIBILITATE SI REZOLUTIE

In medie, la o examinare globala, exista o convergenta de 130 milioane de celule fotoreceptoare la 1 milion de celule ganglionare (deci de fibre in nervul optic). Trebuie insa sa mentionam ca aceasta corespondenta nu este uniforma. In zona foveala corespondenta este de 1:1, in timp ce la periferie corespondenta ajunge la o rata de cateva sute de celule fotoreceptoare la o celula ganglionara. O astfel de celula ganglionara, cu campul receptor in zona periferica a retinei, va suma intrarile primite – fenomenul de **sumare spatiala**. Ea va avea, deci, o sensibilitate ridicata, putand fi excitata si de stimuli relativ scazuti ca intensitate. Celulele ganglionare din zona foveala permit o rezolutie mult crescuta, dar au o sensibilitate scazuta.

Aceasta distributie a campurilor receptoare ganglionare se combina cu proprietatile celulelor receptoare (bastonasele sunt mult mai sensibile decat conurile) si cu distributia acestora pe suprafata retinei (mai dense in centru si tot mai rare spre periferie). Toate acestea conduc la o sensibilitate ridicata la periferia campului vizual (cu o rezolutie slaba in aceasta zona) si o rezolutie foarte buna in zona de focalizare a privirii.

Pe langa sumarea spatiala, exista si un fenomen de **sumare temporală**, exprimata de **Legea lui Bloch**:

$$I \times T = C$$

unde s-a notat: I – intensitatea stimulului, T – durata stimulului, C – efectul vizual constant.

Legea lui Bloch poate fi sintetizata in cateva cuvinte: marind durata unor stimuli de slaba intensitate, ei devin la fel de luminosi ca si cei de intensitate mai ridicata, dar prezentati pe o durata mai scurta (daca doi stimuli contin aceeasi cantitate de energie ei vor detectati la fel). Trebuie insa mentionat faptul ca acest lucru este valabil in intervale relativ scurte (ms), deoarece "memoria" celulelor receptoare este foarte scurta, acest lucru ajutand la discriminarea rezonabila a diversilor stimuli vizuali.

2.2. CAILE VIZUALE CENTRALE

Nervul optic este format din axonii celulelor ganglionare ale retinei (aprox. 1 milion de fibre). De la cei doi ochi, creierul primeste informatie vizuala prin 2 milioane de canale separate.

Dupa parasirea ochiului, fibrele nervoase sunt invelite in teaca de mielina creata de celulele gliale, ceea ce asigura izolarea acestor si maresc considerabil viteza de transmisie a impulsurilor nervoase. Aceasta teaca dubleaza insa diametrul fiecarui axon si, de aceea, in interiorul ochiului axonii celulelor ganglionare nu sunt mielinizati (pentru a reduce cat mai mult grosimea stratului pe care-l are lumina de strabatut).

Nervii optici ce vin de la fiecare ochi converg in **chiasma optica**, unde are loc o incrucisare a fibrelor: aproximativ jumătate din fibrele ce vin de la fiecare ochi (jumătatea nazala) traverseaza linia de simetrie a corpului si, unindu-se cu jumătate din fibrele celuilalt ochi (jumătatea temporală), formeaza tractul optic, ce intra in nucleii talamusului. Aşa cum se vede din figura 2.10 (b), fibrele ipsilaterale sunt cele care aduc informatia din jumătatea temporală a retinei, iar cele contralaterale aduc informatia din jumătatea nazala.

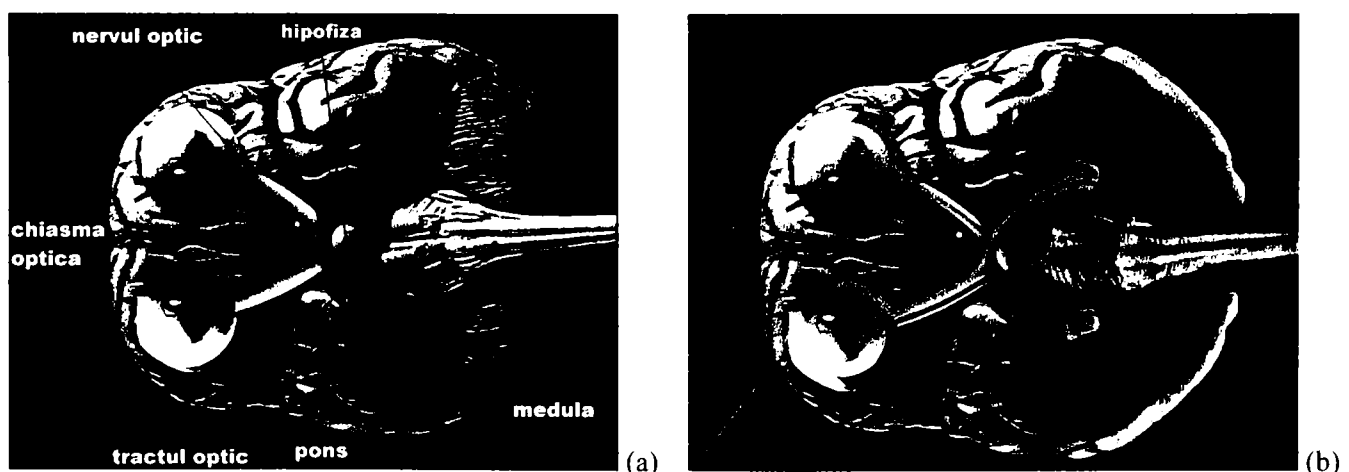


Figura 2.10: Caile vizuale si proiectia fibrelor nervoase in aria vizuala primara: (a) nervul optic si apoi tractul optic dupa chiasma; (b) evidentierea modului in care se face incrucisarea fibrelor in chiasma optica si proiectia acestora in corpii geniculati laterali si apoi in cortexul striat. De notat faptul ca fibrele corespunzatoare jumatatii stangi, respectiv drepte a campului vizual au culori diferite. (Dupa Hall & Robinson 1998)

Datorita modului in care se face proiectia informatiei vizuale pe retina, imaginea retiniana este o copie inversata a scenei din campul vizual. De asemenea, in cazul primatelor, cei doi ochi

sunt asezati frontal, asa ca imaginile retiniene contin, in mare parte, aceeasi informatie vizuala, putin decalata pe orizontala. Acest fapt are doua consecinte importante:

- imaginea din zona temporală a ochiului stang si din zona nazala a ochiului drept contin informatie despre aceeasi zona a campului vizual, anume partea dreapta a acestuia; pentru ochiul drept, lucrurile sunt similare. Inseamna ca fibrele nervoase ce aduc informatie despre aceeasi parte a campului vizual se "imperecheaza" in chiasma optica si vor fi procesate impreuna in continuare.
- informatia vizuala pastreaza regula careia i se supun celelalte informatii privind mediul inconjurator si propriul nostru corp – este procesata in partea opusa a creierului, fata de cea din care provine (jumătatea contralaterală). Trebuie sa mentionam faptul ca, la fel ca in cazul celorlalte tipuri de informatie, acest lucru nu conduce la o fractura a campului vizual, deoarece exista permanente interactiuni intre cele doua jumătati cerebrale.

Aproximativ 80% din fibrele tractului optic se proiecteaza in **nucleii geniculati laterali (LGN)** ai talamusului, care sunt "poarta de intrare" in cortexul vizual. Celelalte fibre ajung in mai multe zone mai "vechi" din punct de vedere filogenetic, dintre care cei mai importanti sunt **coliculii superiori** (aprox. 10% din fibre) asezati sub talamus – figura 2.11.

COLICULII SUPERIORI

Coliculii superiori (figura 2.11) se gasesc in partea superioara a trunchiului cerebral si sunt componenta a mezencefalului. Rolul lor este foarte important la animalele inferioare, deoarece ei sunt cei care controleaza raspunsul organismului la orice fel de stimul vizual ce ajunge in campurile lor receptoare, neavand nici o importanta forma sau orientarea acestuia.

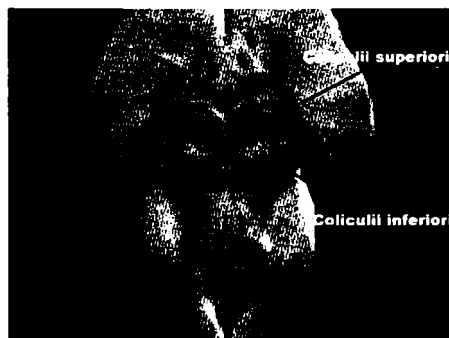


Figura 2.11: Coliculii superiori si cei inferiori. (Dupa Hall & Robinson 1998)

In coliculii superiori se creaza **trei harti senzoriale**, corespunzatoare stimulilor vizuali, auditivi si somatici. Combinarea acestora conduce la crearea unei **harti motorii**, ce permite controlul miscarilor ochilor si ale capului in directia stimulului (indiferent de natura acestuia).

Coliculii superiori contin asa-numite "celule multisenzoriale", ceea ce le permite ca o stimulare combinata cu stimuli nu foarte puternici, de natura diferita, dar care provin din aceeasi regiune a spatiului, sa conduca la un raspuns puternic.

Coliculii inferiori primesc informatie acustica, proiectata apoi in cortexul auditiv primar (tot prin nucleii talamusului).

Coliculii superiori au conexiuni extinse cu cortexul cerebral (prin intermediul talamusului) si cu cerebelul (prin punte – *pons*).

In timp ce coliculii superiori fac o procesare primitiva a informatiei vizuale, celelalte 80% din fibrele tractului optic ajung in zonele mai "tinere" ale encefalului: nucleii geniculati laterali si apoi cortexul cerebral, pentru o procesare detaliata.

NUCLEII GENICULATI LATERALI

Nucleii geniculati laterali (LGN) sunt parte a talamusului si sunt in numar de doi, fiecare primind informatie vizuala dintr-o ramura a tractului optic – figura 2.10 (b) si figura 2.12 (b)– structura si functiile fiind similare pentru cei doi nuclei. La primate, LGN este format din 6 straturi formate din celule nervoase, ceea ce-i confera o aparenta striata – figura 2.12 (a). straturile sunt numerotate de la 1 la 6, incepand cu partea inferioara a acestei structuri.

Intrarile in LGN pastreaza aranjarea topografica a informatiei din imaginea retiniana, formand o asa-numita **harta retinotopica** a informatiei vizuale. Straturile 1, 4 si 6 primesc ca intrari fibrele contralaterale ale tractului optic, iar straturile 2, 3 si 5 pe cele ipsilaterale, astfel incat se ajunge la o aranjare in pereche a informatiei corespunzatoare acelorasi puncte din campul vizual (dar provenite de la cei doi ochi).

Straturile 1 si 2 primesc intrari de la celulele ganglionare de tip M si sunt, si ele, mai mari. Aceste straturi se numesc **straturi magnocelulare**. Straturile 3, 4, 5 si 6 primesc intrari de la celulele ganglionare de tip P. contin celule mai mici si se numesc **straturi parvocelulare**.

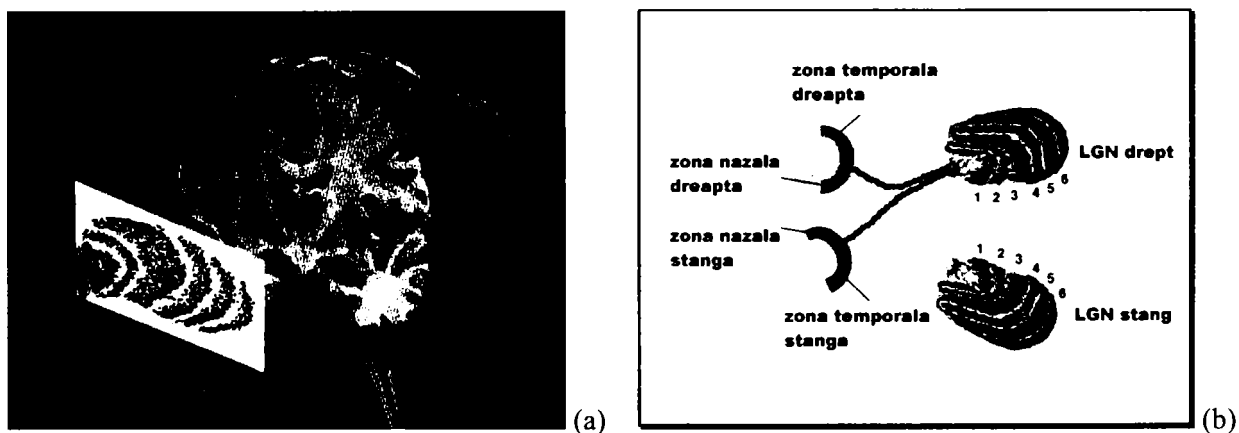


Figura 2.12: Corpui geniculati laterali (LGN): (a) asezarea si aspectul LGN; (b) intrarile tractului optic in LGN si numerotarea straturilor. (Dupa Hall & Robinson 1998)

Celulele din LGN (atat din straturile magnocelulare, cat si de cele parvocelulare) au campuri receptoare circulare, formate din doua zone concentrice cu comportament antagonist, asemanatoare celulelor din retina (celule ON-OFF si celule OFF-ON). Diferenta fata de celulele ganglionare consta in faptul ca efectul inhibitor al zonei periferice este mult mai puternic la celulele LGN, de aceea ele amplifica si accentueaza diferentele de iluminare.

Majoritatea celulelor P (straturile 3÷6) sunt sensibile la culoare; mai mult, zona centrala si cea periferica raspund la culori opuse din spectru (de exemplu, rosu-verde). Celulele M (straturile 1 si 2) si o mica fractiune din celule P raspund la toate culorile.

O alta diferenta intre cele doua tipuri de celule consta in acuitatea diferita (asemanatoare situatiei de la nivelul retinei): celulele M au campuri receptoare mult mai mari, in timp ce celulele P au campuri receptoare mici si fac o analiza in detaliu. Pe de alta parte, celulele M raspund puternic la contururi ce se deplaseaza rapid prin campul lor receptor, in timp ce celulele P prefera miscarea lenta a stimulilor si nu raspund la miscari rapide.

Exista si un **strat coniocelular**, ce primeste intrari de la celulele ganglionare ce nu pot fi clasificate in nici una din cele doua categorii, M sau P. Rolul acestuia este, deocamdata, incert.

Iesirile din nucleii geniculati laterali se proiecteaza in cortexul vizual, aria vizuala primara (V1) si, de acolo, spre ariile mai inalte (V2, V3 si V4), precum si alte zone de prelucrare corticala. **Caile de proiectie vizuala** sunt paralele (coniocelulara, magnocelulara si parvocelulara) si ajung separat in stratul IV C al cortexului vizual. Abia in celelalte straturi informatia este mixata.

FUNCTIILE NUCLELOR GENICULATI LATERALI

Nucleii geniculati laterali nu constituie doar o "statie intermediara" a tractului optic (cum a fost sugerat pana de curand), in care s-ar face doar o ordonare si o accentuare a informatiei existente deja, extrase de celulele retinei.

LGN primesc intrari de la **sistemul reticulat activator ascendent** din trunchiul cerebral, sistem care guverneaza starea generala de activare a organismului. Practic, informatia vizuala este modulata de starea de activare, care poate creste sau scadea nivelul intrarilor corticale.

Pe langa aceste intrari, LGN primesc intrari dinspre cortexul vizual – un *feed-back* al carui rol exact nu se cunoaste inca. Se fac presupuneri si speculatii pornind de la similitudini cu sistemele de reglare din tehnica, la care astfel de legaturi inverse sunt indispensabile pentru buna functionare. Oricum, majoritatea modelelor privind procesarea informatiei vizuale prevad astfel de legaturi de *feed-back*.

Lucrarea de fata revine la acest aspect in capitolul ce trateaza modelarea integrarii contururilor.

2.3. CORTEXUL VIZUAL

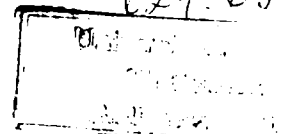
Cortexul vizual este situat in lobul occipital si primeste intrari de la nucleii geniculati laterali – figura 2.1 (b) si figura 2.10 (b). El este format din mai multe zone concentrice, ariile vizuale (V1, V2, V3 si V4). **Aria vizuala primara** este notata cu **V1** de catre fiziologi (Anexa 3) si se identifica cu **aria 17** potrivit hartilor corticale preferate de anatomisti (Ariile Brodmann – Anexa 2). Aceeasi zona mai poarta si numele de **cortex striat**, din cauza aspectului sau "dungat", datorat coloanelor de dominanta oculara – figura 2.13.



Figura 2.13: Aspectul cortexului striat, dat de coloanele de dominanta oculara – montaj dupa mai multe sectiuni din stratul IVC (evidentiere obtinuta cu injectare de 3H -prolina) Bara de scalare din dreapta, 5mm. (Cynader 1987)

Ca intreg cortexul cerebral, aria vizuala primara consta dintr-un strat de celule nervoase de aproximativ 2 mm grosime, care are la randul sau alte 6 straturi – figura 2.14. Numerotarea incepe de la suprafata corticala, sub stratul 6 gasindu-se materia alba.

Aria 17 a fiecarei emisfere cerebrale contine aprox. 100 milioane de celule, ce difera ca aspect si densitate de la un strat la altul.



Intrările ce provin de la LGN sunt proiectate în **stratul IVC**: cele provenite de la straturile magnocelulare în partea superioară, iar cele de la straturile parvocelulare în partea inferioară. De aici se fac conexiunile cu straturile din partea inferioară și superioară a cortexului striat și apoi cu alte zone ce procesează informația vizuală. În toată această rețea de conexiuni, caile magnocelulare și cele parvocelulare rămân separate, întărind ideea că ele reprezintă canale distincte de procesare a informației.

Dincolo de aspectele anatomice (prezentate pe scurt aici), foarte importante și interesante sunt cele legate de transformările informației care au loc în această zonă corticală. Acestea au fost aduse în atenția lumii științifice de cercetările efectuate de neurobiologii David Hubel și Torsten Wiesel, pentru care ei au și primit Premiul Nobel în 1981. Descoperirile lor privind organizarea cortexului striat au avut un impact imens asupra cercetărilor din domeniul neuroștiințelor și asupra colaborărilor interdisciplinare în cercetarea privind procesarea informației vizuale.

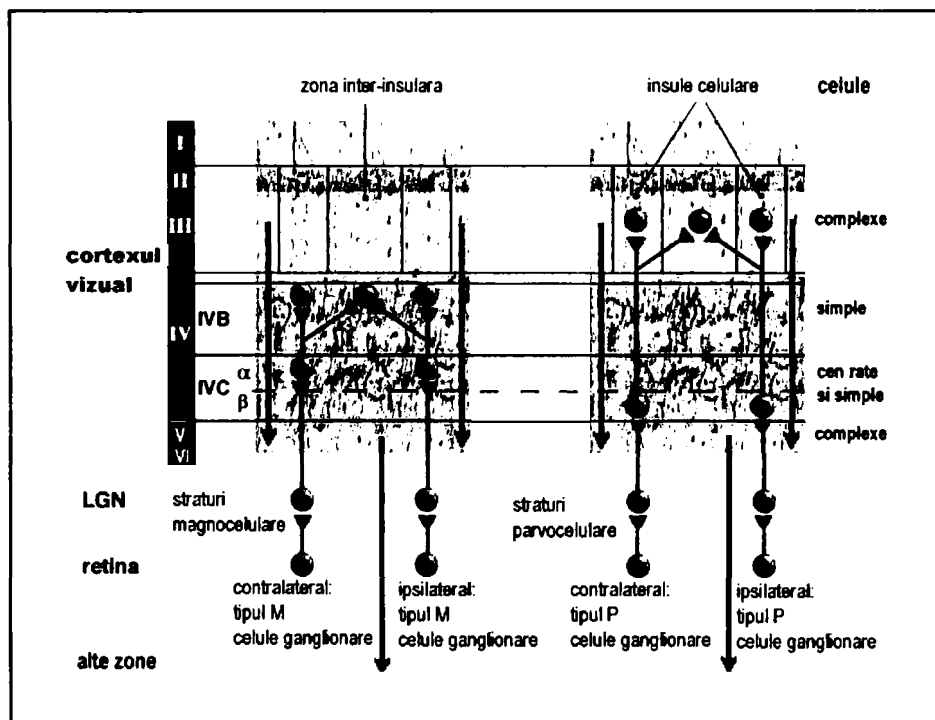


Figura 2.14: Cortexul striat, cu evidențierea conexiunilor în și dinspre cortex, precum și a celor intracorticele. Observați că intrările de la LGN au loc în stratul IVC, iar caile magnocelulare și cele parvocelulare se mențin tot timpul separate. (După Hall & Robinson 1998)

HARTILE ARIEI VIZUALE PRIMARE

Aria V1 are o **organizare retinotopică**, ca o continuare a schemei de procesare începută la nivelul retinei și păstrată de nucleii geniculati laterali. La toate aceste stadii de prelucrare, celulele individuale analizează informația dintr-o zonă limitată a retinei. Campurile receptoare ale celulelor din V1 formează o hartă topografică a jumătății contralaterale a campului vizual.

La fel ca organizarea retinotopică, există și o alta caracteristică "mostenită" de la procesarea retiniană, însă mult exagerată la nivel cortical: zona foveală este foarte bine reprezentată ($\approx 80\%$), în timp ce zonelor periferice li se alocă o reprezentare mult redusă ($\approx 20\%$).

PROPRIETĂȚILE FUNCȚIONALE ALE CELULELOR DIN ARIA VIZUALĂ PRIMARĂ

La fel ca și celulele retiniene și cele din LGN, celulele din V1 nu sunt afectate de modificări ale iluminării globale, ele răspund la diferențe de iluminare produse de muchii, granitele dintre obiecte, etc.

Una din caracteristicile foarte importante ale celulelor din V1 este **selectivitatea lor la orientarea stimulilor**. Aceasta proprietate, care le distinge de celulele retiniene si cele din LGN, se datoreaza faptului ca au campurile receptoare de forma alungita – figura 2.15 (a) – pastrand in sa zonele excitatorii/inhibitorii, ceea ce le permite sa raspunda optim pentru muchii de o anumita orientare. Interesant este faptul ca o celula "acordata" pentru o anumita orientare va raspunde totusi (mai slab in sa) si pentru orientarile invecinate, raspunsul devenind nesemnificativ la o distanta mai mare de 15°-20° de orientarea optima. Astfel de celule sensibile la orientare se gasesc in toate straturile, cu exceptia partii inferioare a stratului IV (in care vin intrarile de la LGN) si in zone bine conturate (insulare) aproape de suprafata corticala.

Mai mult decat atat, celulele ce au campurile receptoare in aceeasi zona a campului vizual, dar "raspund" de orientari diferite sunt grupate in asa-numite **hipercoloane de orientare**. Aceste hipercoloane cuprind coloane de orientare acoperind tot spectrul de orientari posibile si se extind pe toata grosimea scoartei cerebrale – figura 2.15 (c).

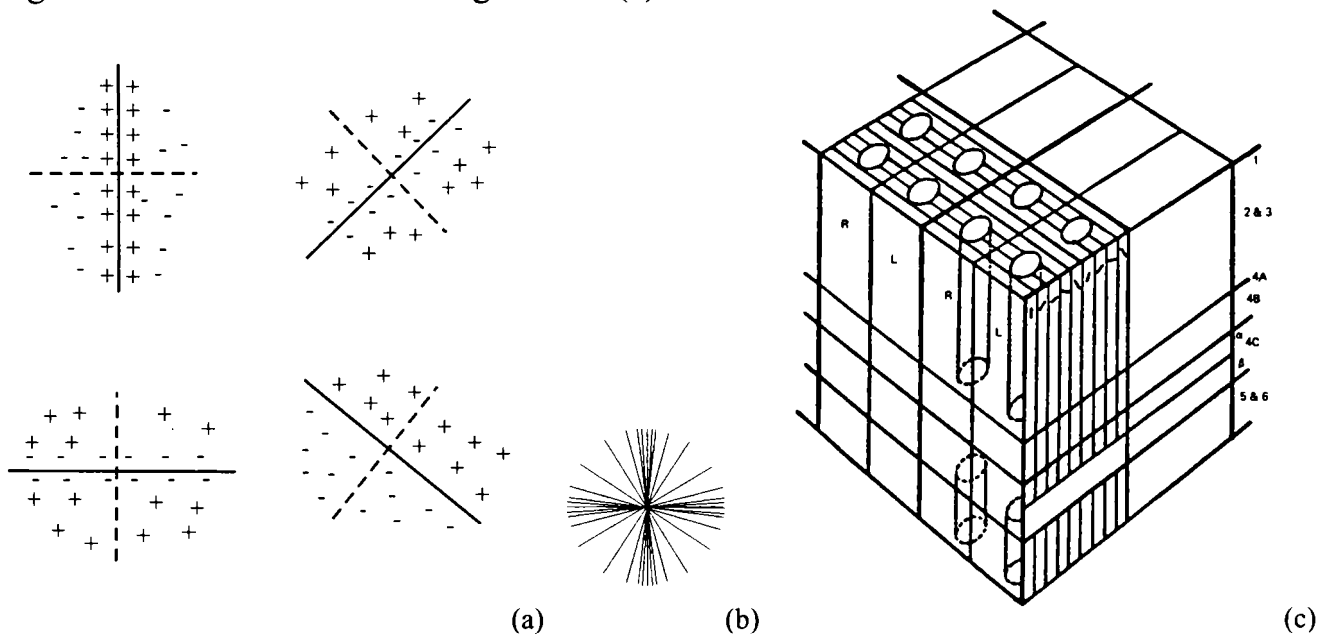


Figura 2.15: Celulele simple din cortexul striat prezinta selectivitate la orientare: (a) forma alungita a campurilor receptoare, pastrand zone cu comportament antagonist; (b) pentru o aceeaasi locatie spatiala, distributia orientarilor nu este uniforma, ci se prefera directiile verticala si orizontala; (c) cortexul striat este organizat in hipercoloane de orientare ce se intrepatrund cu cele de dominanta oculara si se prelungesc in toata profunzimea scoartei. (Dupa Shiller 1987)

"Fasiile" care raspund de o anumita orientare sunt de ≈ 0.05 mm, iar la distanta de ≈ 1 mm pe cortexul vizual, se constata ca s-a facut o rotatie completa (incepe o noua hipercoloana). O astfel de hipercoloana contine zeci de mii de celule, ale caror campuri receptoare acopera aceeaasi zona retiniana. Hipercoloanele sunt de aproximativ aceeaasi latime pe cortex (≈ 1 mm), dar cele "raspuzatoare" de zona foveala au campurile receptoare mult mai mici in dimensiuni. Trebuie sa mai mentionam ca cercetarile au dovedit ca distributia orientarilor preferate nu este uniforma si se constata o mai buna perceptie a orientarilor verticale si orizontale – figura 2.15 (c) da o reprezentare polara aproximativa a acuitatii in perceptia orientarilor la nivelul celulelor din V1.

Celulele din cadrul unei coloane de orientare raspund toate de o aceeaasi localizare spatiala si prezinta o aceeaasi orientare preferata (chiar daca ele au dimensiuni diferite, sau campuri receptoare de dimensiuni diferite, sau mai mult, unele "prefera" muchiile, in timp ce altele "prefera" barele, etc.).

Putem spune ca hipercoloanele contin mecanismele neuronale pentru a analiza informatia vizuala din campurile receptoare respective. Lucrand simultan, in paralel, hipercoloanele analizeaza intreaga imagine proiectata pe retina.

Mai trebuie sa mentionam ca vederea noastra este binoculara, cei doi ochi "culegand" imagini putin deplasate pe orizontala ale unei aceleiasi realitati sau scene. Asa cum exista coloane

de orientare in cadrul cortexului vizual, exista si **coloane de dominanta oculara**, in cadrul carora celulele receptoare au campurile dedicate unuia dintre cei doi ochi (figura 2.13). Pentru o persoana normala, suprafata cortexului striat este egal impartita intre cei doi ochi.

Coloanele de orientare si cele de dominanta oculara se intrepatrund in cortexul vizual. Vom reveni asupra implicatiilor acestei organizari si asupra incercarilor de modelare in capitolele urmatoare.

CELULE SIMPLE SI CELULE COMPLEXE

Celule corticale din V1 se impart in doua categorii: *celule simple* si *celule complexe*.

Celulele simple sunt cele care au campuri receptoare exact asa cum au fost prezentate in fig. 2.15 (a), ele raspunzand cu mare acuratete la stimulii de orientare potrivita si cu positionarea corespunzatoare in spatiu – mai exact, ele “semnaleaza” orientarea unui stimul ce cade intr-o anumita zona particulara a campului vizual.

Celulele complexe nu au asa de clar delimitate zonele de ON/OFF, desi au, si ele, “preferinta” pentru anumite orientari. Raspunsul lor este insa mai complicat si nu se poate prezice cu exactitate ce tip de stimul individual le-ar activa optim – ele raspund la stimuli in miscare (prefera contururile in miscare celor stationare), sunt selective la directia de miscare, la culoare, la disparitatea binoculara, la "muchii" care se incadreaza in contururi mai largi, etc.

Detectia culorii se face prin intermediul unor grupari de celule concentrate in "insule celulare" dispuse regulat pe suprafata corticala, in care sunt proiectate intrari de la celulele din stratul IV C, inervate de caile parvocelulare – figura 2.14. Aceste celule nu au nici o preferinta pentru orientare, dar prezinta preferinta clara de culoare, efectul inhibitor sau excitator al culorii depinzand de zona campului receptor in care se situeaza.

CELULELE DIN V1 CA "FEATURE DETECTORS"

Dincolo de toate aceste observatii neuro-fiziologice, care demonstreaza mai mult faptul ca anumite zone ale creierului sunt active in timpul perceptiei vizuale, o problema importanta ramane: *cat de mult din ceea ce vedem poate fi explicat de activitatea unor celule corticale care “prefera” un anumit tip de stimuli?*

David Marr (Marr 1982) sublinia faptul ca a avea cortexul vizual in mod elegant proiectat sa extraga informatie privind formele si localizarea contururilor in imaginea retiniana este doar un prim pas in procesul perceptiei vizuale, dar unul de o importanta covarsitoare.

Celulele corticale sunt considerate a fi *feature detectors*, dar ele nu sunt chiar detectoare de caracteristici in sensul strict al cuvantului – ele nu detecteaza cu certitudine un anumit tip de stimul, ci doar isi moduleaza activitatea functie de prezenta/absenta stimulilor “preferati”. In fapt, este vorba de o activitate de grup, la nivelul retelei neuronale, o “colaborare” intre populatii intinse de neuroni.

Trebuie sa mai mentionam faptul ca aria vizuala primara face, asa cum ii spune si numele, o procesare primara a informatiei vizuale si constituie o prima etapa intr-un lant complex de prelucrari, in care sunt implicate multe din zonele corticale pentru a se ajunge la perceptia constienta. In aceste stadii ulterioare se pastreaza insa organizarea retinotopica in paralel cu o tot mai accentuata "specializare" in detectia unor anumite caracteristici sau informatii: analiza culorilor, analiza miscarii, perceptia formelor, etc.

De ce a ales natura o prelucrare distribuita a informatiei? La aceasta intrebare nu s-a gasit inca un raspuns cert, dar se presupune ca ratiunea a fost aceeaasi pentru care se prefera si in tehnica sistemele distribuite: sunt mai usor de depanat si de imbunatatit in decursul evolutiei naturale.

Important este ca la nivelul ariei V1 se realizeaza o serie de prelucrari care nu presupun o implicare a constientului si pe o parte dintre acestea am incercat sa le abordez in lucrarea de fata.

Capitolul 3: MODELAREA VAZULUI. REPREZENTAREA SI PROCESAREA INFORMATIEI VIZUALE

3.1. NIVELE IN SISTEMUL NERVOS. NIVELE IN MODELARE

Asa cum am sugerat si in capitolul introductiv, abordarea fenomenelor psihologice si a celor neurobiologice se face intotdeauna pe nivele (Churchland&Sejnowski 1992). Exista trei clase de nivele: nivele de analiza, nivele de organizare si nivele de procesare.

NIVELE DE ANALIZA

Nivelele de analiza sunt conceptuale si se refera la diversele tipuri de intrebari privitoare la functiile creierului, la modul in care acestea pot fi divizate in sub-functii, care sunt pasii necesari pentru rezolvarea acestor sub-functii, ce structuri pot efectua acesti pasi.

David Marr (1982) a enuntat pentru prima data importanta si principiile rationamentului pe nivele in contextul unei teorii computationale privind mecanismele vederii. Marr a formulat necesitatea abordarii pe 3 nivele:

- (1) **nivelul computational**, la care se face analiza problemei si descompunerea ei in sub-functii (de exemplu, principiul de determinare a adancimii pornind de la imaginile 2-D retiniene);
- (2) **nivelul algoritmului**, la care se face specificarea unei proceduri formale pentru ca sub-functiile sa poata fi realizate, adica pentru o intrare data, sa se obtina o iesire corecta;
- (3) **nivelul implementarii fizice**, la care, utilizand o anumita tehnologie, se trece la realizarea unui dispozitiv (natural sau artificial) ce implementeaza algoritmul stabilit la nivelul 2.

Marr a considerat ca cele trei nivele sunt independente, sarcinile fiecarui nivel putand fi rezolvate fara nici o influenta a celorlalte – o abordare tipica *top-down*. In practica, el nu a procedat exact asa atunci cand a dezvoltat propriile modele ale mecanismelor vederii, ci chiar a fost puternic influentat de considerentele neuro-biologice (dovada ca lucrarile lui raman inca puncte de referinta in modelarea vazului).

Pentru sistemele naturale, este evident ca nu se poate vorbi de o independenta a algoritmului fata de aspectele de implementare si aceasta nici nu este o strategie viabila in construirea modelelor mecanismelor biologice. Sistemele naturale au evoluat in timp si au fost realizate cu materialele avute la dispozitie (apa, proteine, lipide, etc.), iar cerintele impuse de nivelul 3 au fost cele care au determinat alegerea unui algoritm sau al altuia. Natura nu a ajuns totdeauna la cea mai buna solutie posibila, ci la una satisfacatoare din punct de vedere al supravietuirii si reproducerii. De multe ori, pentru aceeasi functie, au fost dezvoltati algoritmi diferiti la specii diferite. De aceea, cercetarile actuale pornesc de la premisa ca se ajunge la un algoritm plauzibil biologic daca se foloseste o strategie combinata, adica o analiza *top-down*, dar si *bottom-up* (Jordan&Russel 1999).

Pentru sistemele artificiale, acolo unde considerentele de viteza, eficienta, eleganta, etc. primeaza, independenta formala a algoritmului fata de arhitectura si implementare poate fi o realitate.

NIVELE DE ORGANIZARE

Nivelele de organizare sunt, in esenta, anatomice si se refera la ierarhia componentelor si la structurile in care acestea se integreaza. Daca se face un paralelism intre nivelele de analiza formulate de David Marr si nivelele de organizare a sistemului nervos, imaginea poate fi destul de confuza. Exista insa o structura organizata la scari diferite: molecule, sinapse, neuroni, retele de neuroni, straturi, harti corticale, sisteme. La fiecare astfel de strat structural se pune problema functiei pe care acesta o indeplineste. De aceea, pentru fiecare nivel de organizare, putem vorbi de cele trei nivele de analiza formulate de Marr.

NIVELE DE PROCESARE

Nivelele de procesare sunt fiziologice si se refera la localizarea unui proces – ele constituie legatura dintre anatomie si functiile acesteia. Aceste nivele pot fi considerate pe toata scara de procesare, de la informatia data de simturi, pana la raspunsul pe care organismul il da la anumite intrari.

In general, se considera ca cu cat este mai mare distanta sinaptica a sistemului considerat fata de intrarea senzoriala, cu atat mai inalt este nivelul de procesare respectiv. Pe aceasta scara, celulele din aria vizuala primara sunt pe un nivel mai ridicat decat cele din nucleii geniculati laterali, care (la randul lor) sunt la un nivel mai inalt decat celulele ganglionare din retina. Deocamdata insa, nu se cunoaste foarte mult despre aceste ierarhii si de aceea putem spune doar ca un nivel este mai ridicat decat altul, dar nu le putem ordona in sensul strict: primul, al doilea, etc. Acest lucru nu este in totalitate posibil nici macar la nivelul V1, iar mai departe lucrurile se complica si mai mult.

Ceea ce se stie insa clar este ca nu exista o procesare strict secventiala si ca exista legaturi de *feed-back* intre nivelele asa-zise inalte si cele joase, primare (Nelson 1995). Anatomia cortexului frontal si a altor arii dincolo de ariile senzoriale primare sugereaza o “organizare ce seamana mai mult cu o democratie ateniana, decat cu o linie de asamblare Ford” (Churchland&Sejnowski 1992).

Problema cercetatorului care lucreaza la un anumit nivel este aceea ca trebuie sa faca apel la nivelele adiacente: in jos (pentru a gasi mecanismele functiilor studiate) si in sus (pentru a intelege rolul jucat de functiile studiate in schema mecanismului de ansamblu). Acest lucru transpare si in aceasta lucrare: desi am abordat doar doua din mecanismele de la nivelul ariei vizuale primare, in prezentare voi incerca mereu sa descriu cadrul de ansamblu.

Asa cum am specificat la nivelele de analiza, aceasta abordare pe nivele este utila din punct de vedere al teoriei, in practica insa trebuie sa nu pierdem din vedere legatura intrinseca dintre structura si functie la nivelul mecanismelor biologice (Intrator&Cooper 1995, Jordan&Russel 1999). Mai mult, atunci cand vorbim despre modelarea unor functii biologice, trebuie sa avem mereu in vedere considerentul biologic in ceea ce priveste implementarea (inclusiv datele experimentale disponibile). De aceea, in continuarea acestui capitol si in urmatoarele doua, am abordat urmatoarea **strategie de prezentare**:

- am specificat datele anatomice, fiziologice si psihofizice de la care se porneste, in paralel cu functia ce urmeaza a fi modelata;
- am amintit, pe scurt, principalele tipuri de modele dezvoltate pentru functia si/sau structura respectiva;
- pentru capitolele 4 si 5, am descris modelele propuse de mine pentru functia respectiva, cu discutie pe marginea rezultatelor.

3.2. ORGANIZAREA CORTEXULUI STRIAT

ASPECTE ANATOMICE SI FIZIOLOGICE

Asa cum am aratat in Capitolul 2, cortexul striat (sau aria vizuala primara – V1) este zona corticala ce ocupa o larga portiune a lobului occipital (Anexa 1). El este organizat in 6 straturi (impartite in sub-straturi), ce se intind pe o grosime de $\approx 2\text{mm}$ de materie cenuie – figura 3.1.

Cei mai largi corpi neuronali sunt ai celulelor piramidale din straturile 5 si 6 (pusi in mod deosebit in evidenta prin metoda Golgi, care evidentiaza $\approx 5\%$ dintre celule). Aceste celule au dendrite foarte lungi, ce se intind atat in adancimea cortexului (pe toate straturile), cat si lateral

(pana la 0.5 mm). Ele constituie principalele iesiri de la nivel cortical, mergand spre celelalte portiuni (ne-corticale) ale creierului.

In straturile 2 si 3 se gasesc celule piramidale mai mici, iar axonii lor merg cu precadere spre alte zone corticale – de aceea se numesc “axoni de asociere” (cei mai lungi se gasesc in corpul calos).

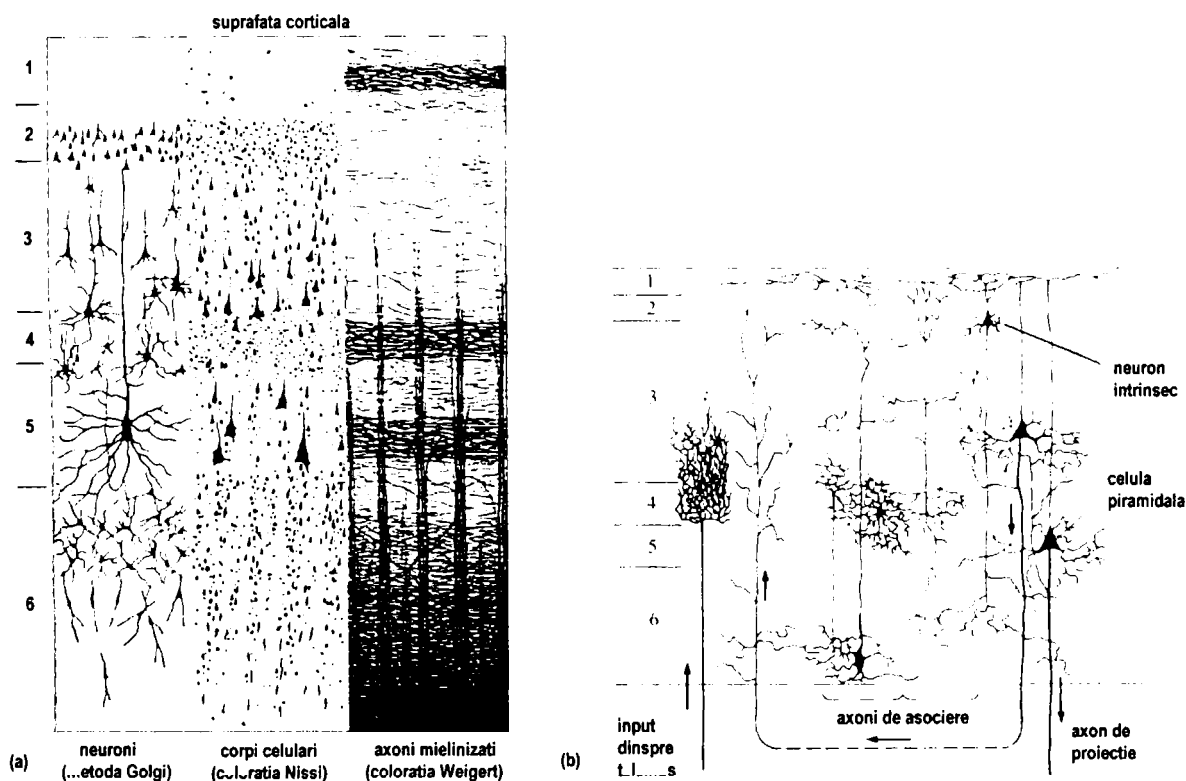


Figura 3.1: Structura corticala: (a) sectiuni prin cortexul cerebral uman, utilizand trei metode diferite de coloratie, pentru a pune in evidenta corpii neuronali si distributia axonilor mielinizati; (b) diagrama ce prezinta formele celulelor si ale conexiunilor inter-neuronale. Sagetile arata directia in care circula potentialul de actiune. (Dupa Robinson 1998)

Neuronii intrinseci comunica vertical (intre straturi) si orizontal (in cadrul aceluiasi strat). Acestia sunt cei care contribuie la organizarea pe coloane.

Asa cum am aratat in Capitolul 2, campul vizual este reprezentat pastrand organizarea topografica, zona foveala avand alocata o suprafata considerabil mai larga decat zonele periferice. **Celulele corticale** sunt mult mai specifice in raspunsul lor la stimuli si se caracterizeaza prin urmatoarele **atribute**:

- selectivitate la orientare si la muchii;
- selectivitate la directia si viteza de miscare a stimulilor;
- selectivitate la frecvente spatiale pentru stimuli repetitivi (texturi);
- selectivitate la culoare;
- sensibilitate la increment/decrement de iluminare;
- activitate binoculara, chiar sensibilitate la disparitatile retiniene.

Pe langa aceasta organizare laminara, cortexul striat este organizat si in **coloane** (descrise, de asemenea, in Capitolul 2). Ceea ce trebuie accentuat aici este faptul ca ele se intind in toate straturile corticale si a fost demonstrata existenta unor coloane cu diverse proprietati: coloane si hipercoloane de orientare, coloane de dominanta oculara, coloane de frecvente spatiale. In cadrul unei aceleiasi coloane, in straturi diferite, exista celule cu raspuns divers (diferite intre ele), dar toate prezinta si proprietati care le fac sa apartina acelei coloane (selectivitate la o anumita orientare, raspund la stimularea aceluiasi ochi, etc.) Din punctul de vedere al modului in care raspund la stimuli, celulele din V1 se impart in doua categorii mari: **celule simple** (au campuri receptoare alungite si permit detectia muchiilor lumina/intuneric in campul vizual) si **celule**

complexe (primesc, in general, intrari de la ambii ochi, au un raspuns mai complex si o selectivitate mai ridicata in privinta tipului de stimulare la care raspund).

Evidentele experimentale atesta faptul ca organizarea cortexului striat este modulara, fiecare modul analizand o anumita zona a spatiului vizual. Un astfel de modul se intinde pe aprox. $1 \times 1 \times 2 \text{mm}$ si contine un ciclu complet de orientari si dominante oculare.

Organizarea de baza a cortexului striat pare sa fie determinata genetic (Schiller 1997), dar o mare influenta o are activitatea vizuala din primele luni de viata (Goodhill 1998, Held 1987, Horton&Hocking 1996, Miller 1995, Schiller 1995, Sekuler 1990). Atat distributia coloanelor de dominanta oculara, cat si a celor de orientare este puternic influentata de intrarile vizuale in aceasta perioada critica. Privarea vizuala a unuia dintre ochi conduce la "acapararea" spatiului cortical de ochiul normal, iar fenomenul este ireversibil (ochiul initial privat nu-si mai poate, ulterior, "castiga" spatiul pierdut), ceea ce conduce la pierderea vederii binoculare. De asemenea, un mediu in care stimulii nu acopera un spectru suficient de larg (apar doar anumite frecvente spatiale, de exemplu) conduce la imposibilitatea perceptiei ulterioare a stimulilor ce au lipsit in perioada critica de dezvoltare.

Aceste evidente experimentale au condus, in mod firesc, la intrebari legate de mecanismele prin care iau nastere aceste structuri functionale – coloanele de organizare de la nivelul ariei vizuale primare.

MODELAREA ORGANIZARII PE COLOANE

Modelele organizarii pe coloane au simplificat putin problema, considerand-o in spatiu bi-dimensional (au pornit de la premiza ca in profunzimea cortexului proprietatile acestor coloane sunt constante, din punctul de vedere al aspectului modelat). Modelele au utilizat, in general, competitia care apare in cazul retelelor cu auto-organizare (Bauer 1995, Grossberg&Olson 1994, Kohonen 1988, Malsburg 1995, Miller 1995, Mitchinson&Swindale 1999, Riesenhuber et al. 1998, Ritter 1995, Wiskott&Sejnowski 1998, Yuille 1995, Yuille et al.1996). Rezultatele obtinute au confirmat au fost, in general, destul de bune obtinandu-se harti corticale fidele realitatii – figura 3.2.

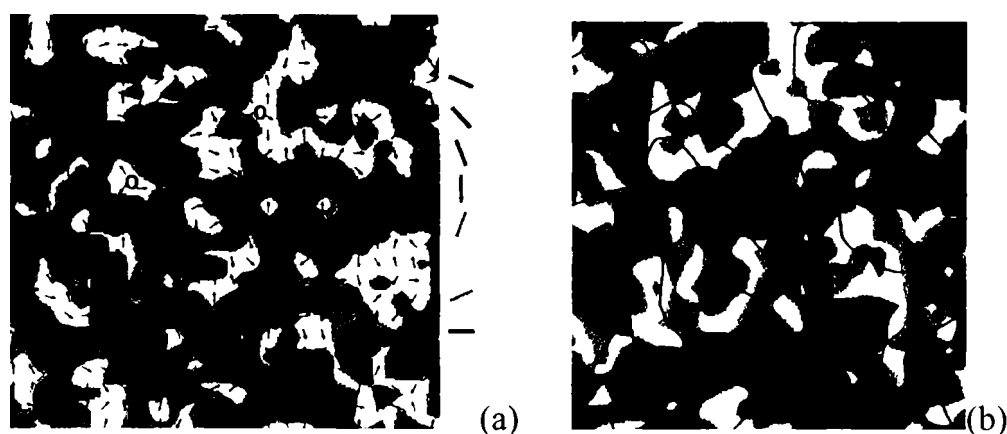


Figura 3.2: Rezultate obtinute cu o retea de tip *self-organizing feature map*: (a) coloane de orientare – prin culori diferite sunt indicate orientarile, iar cu sageti sunt indicate fazele campurilor receptoare; (b) coloane de dominanta oculara (cu linie continua) suprapuse peste cele de orientare (in culori). De notat modul cum coloanele de dominanta se intrepatrund cu cele de orientare (foarte similar realitatii) si ca se ajunge la o reprezentare topografica a intrarilor (acest lucru nu se vede in figura). (Dupa Riesenhuber et al. 1998)

O alta intrebare care se ridica este cea referitoare la transformarile pe care le sufera informatia vizuala ce vine pe nervul optic, astfel incat ea sa poata fi utilizata ulterior, in ariile inalte de procesare.

3.3. REPREZENTAREA SI PROCESAREA INFORMATIEI VIZUALE

Cum se face reprezentarea si codificarea informatiei in creierul uman? Care a fost solutia aleasa de natura? Am atins, foarte scurt, acest subiect si in Capitolul 2, voi incerca acum sa sistematizez putin.

CODIFICARE LOCALA SAU REPREZENTARE DISTRIBUITA?

Codificarea locala la nivelul perceptiei presupune unicitatea raspunsului pentru un anumit stimul. Celebra este sintagma “celulelor-bunica” (celulele care se vor activa doar la vederea bunicii) sau ideea potrivit careia in reprezentare exista “o celula pentru un cuvânt”.

Alternativa codificarii locale o constituie **codificarea vectoriala** sau **reprezentarea distribuita** (reprezentare in spatiul starilor, reprezentare multidimensionala). Criticile care se aduc codificarii locale (pe care codificarea vectoriala le inlatura) sunt:

- (1) codificarea locala este rejectata de chiar evidenta biologica, care este perfect in concordanta cu reprezentarea distribuita;
- (2) recunoasterea unor lucruri noi si raspunsul motor la noi stimuli devin problematice in cazul codificarii locale, ceea ce nu se intampla in cea distribuita;
- (3) formele recunoscute vizual intr-o viata de om depasesc de departe numarul de neuroni din sistemul nervos uman.

Mai mult, reprezentarea distribuita apare mult mai plauzibila din considerente computationale, asa cum retelele neuronale au dovedit-o (Churchland&Sejnowski 1992, Meunier&Nadal 1995, Thorpe 1995).

In ultimul timp insa, tot mai multi cercetatori sunt de acord ca lucrurile nu trebuiesc impinse atat de departe: nu trebuie sa vedem codificarea locala ca fiind strict “un neuron pentru un obiect”, sau ca cea vectoriala implica “toti neuronii unui sistem participa la toate reprezentarile”. Mai degraba, se considera ca cele doua alternative nu sunt mutual exclusive si ca diferitele structuri neuronale utilizeaza strategii de reprezentare a informatiei diferite (Churchland&Sejnowski 1992, Field 1994, Field 1995). Exista strategii mixte chiar in cadrul aceleiasi structuri functionale (hiperacuitatea in detectia disparitatilor binoculare, de exemplu, permite detectia unor diferente mai mici decat diametrul unei celule fotoreceptoare). Reprezentarile valorilor scalare este mai eficienta utilizand o codificare locala, in timp ce pentru valori vectoriale codificarea distribuita apare ca fiind o solutie mai buna.

La nivelul ariilor senzoriale primare apare evidenta o codificare “grosiera” a informatiei, in contrast cu o crestere a selectivitatii in raspuns pe masura ce se ajunge la nivele de procesare mai inalte. Aceasta nu inseamna neaparat existenta “celulelor-bunica”, ci faptul ca **vectorii** reprezinta stimulii cu tot mai multa specificitate (Field 1994, Field 1995).

In fapt, sistemul nervos nu utilizeaza populatii neuronale largi pentru a genera raspunsuri la stimuli, ci “proiecteaza” aceste reactii de raspuns catre grupuri neuronale restranse, in care celulele individuale codifica informatia – aceste “celule-bunica” sunt cele implicate in controlul raspunsului (Mesulam 1998).

Lucrarea de fata abordeaza mecanisme la nivelul ariei vizuale primare si, de aceea, reprezentarea distribuita a informatiei a fost cea preferata. Totusi, asa cum se va vedea, chiar la nivelul VI apare o specializare a grupurilor de celule in detectia disparitatii binoculare (dar si a altor proprietati).

Cod compact sau imprastiat?

In aceasta privinta, parerile sunt mai unanime: daca se pune problema reprezentarii distribuite, atunci codul imprastiat (*sparse coding*) este mai plauzibil din punct de vedere biologic, mai fiabil (Field 1995, Foldiak&Young 1995, Meunier&Nadal 1995).

REPPREZENTAREA INFORMATIEI VIZUALE

Celulele receptoare si, mai apoi, celulele din LGN si din V1 trebuie sa permita o codificare utila din punctul de vedere al teoriei informatonale (Field 1995).

David Marr (1982) a scos in evidenta importanta profilului campurilor receptoare ale celulelor la diferitele nivele de procesare. Pentru modelarea comportamentului celulelor ganglionare si a celor din LGN, el a propus o convolutie a imaginii cu Laplacian-ul unei functii de forma unei distributii Gauss bi-dimensionale (figura 3.3):

$$h(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}}; \quad \nabla^2 h = \frac{r^2 - \sigma^2}{\sigma^4} e^{-\frac{r^2}{2\sigma^2}} \quad \text{cu } r^2 = x^2 + y^2 \quad (3.1)$$

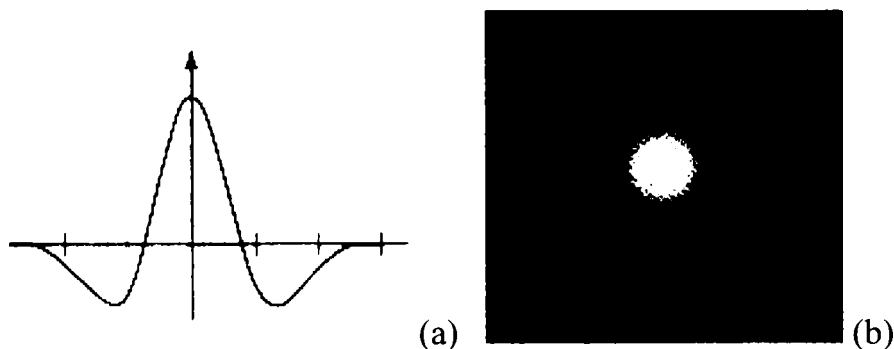


Figura 3.3: Functia propusa de Marr pentru profilul unui camp receptor: (a)sectiune transversala prin $\nabla^2 h$ (functie 1-D); (b) $\nabla^2 h$ ca imagine (functie 2-D). (Dupa Marr 1982 si Gonzalez&Woods 1993)

Marr chiar argumenteaza faptul ca se pot utiliza combinatii de siruri de astfel de profile ON-OFF si OFF-ON pentru detectia muchiilor.

Pentru modelarea profilului campurilor receptoare ale celulelor simple din V1, cercetatorii sunt cvasi-unanimi in a considera functiile Gabor ca fiind cele care "imita" cel mai bine comportamentul acestora (Daugman&Downing 1995, Field 1995, Pollen&Gaska 1987, Qian 1994, Qian&Mikaelian 2000, Salinas&Abott 2000, Tsao&Chen 1994). Au mai fost propuse si alte solutii, cum sunt filtrele *Cepstrum* (Ludwig 1994) si DOOG – *Difference Of Off-set Gaussians* (Marr 1982, Pollen&Gaska 1987), dar familiile de filtre Gabor uni- si bi-dimensionale au castigat teren net.

O familie de filtre Gabor se exprima printr-o functie complexa de tipul:

$$G(x, y) = e^{-\pi \left[\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2} \right]} e^{-2\pi i [u_0(x-x_0) + v_0(y-y_0)]} \quad (3.2)$$

unde:

(x_0, y_0) specifica pozitia in imagine

(α, β) specifica latimea si lungimea filtrului

(u_0, v_0) precizeaza vectorul-unda de modulare specific filtrului.

- frecventa spatiala $\omega_0 = \sqrt{u_0^2 + v_0^2}$
- orientarea sau directia $\theta_0 = \arctan\left(\frac{v_0}{u_0}\right)$

Partea reala, respectiv cea imaginara a acestei functii complexe (care sunt valori reale) descriu fiecare comportarea unei celule simple si formeaza o "pereche in cuadratura" (la diferenta de 90°), ceea ce corespunde datelor experimentale si este aratat in figura 3.4.

Transformarea Fourier 2-D ($F(u, v)$) a unui filtru Gabor are aceeasi forma ca si functia in domeniul spatial - filtrele Gabor sunt functii "self-Fourier":

$$F(u, v) = e^{-\pi \left[(u-u_0)^2 \alpha^2 + (v-v_0)^2 \beta^2 \right]} e^{-2\pi i [x_0(u-u_0) + y_0(v-v_0)]} \quad (3.3)$$

De aici, raspunsul maxim apare pentru orientarea θ_0 si frecventa spatiala ω_0 , asa cum am precizat anterior. Acest maxim corespunde structurii de excitatie/inhibitie a campurilor receptoare.

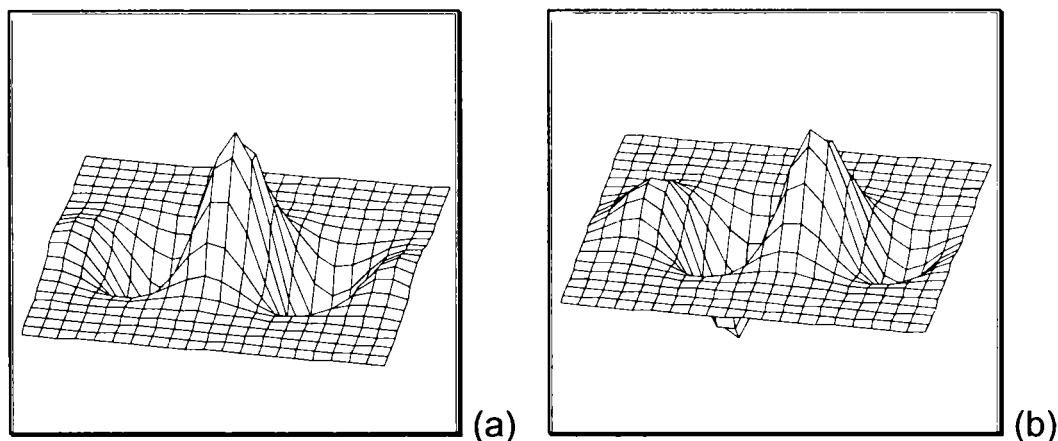


Figura 3.4: O pereche de filtre Gabor in cuadratura.

Desi ele nu constituie un set de vectori ortogonali, s-a aratat ca sunt solutia optima aleasa de natura pentru codificarea informatiei vizuale (Field 1995, Salinas&Abbott 2000).

PROCESAREA INFORMATIEI VIZUALE IN CORTEXUL STRIAT

Din punct de vedere al procesarii informatiei, la nivelul ariei vizuale primare au loc o serie de transformari ce contribuie substantial la perceptia spatiului tridimensional (inclusiv limite de obiecte) – figura 3.5.

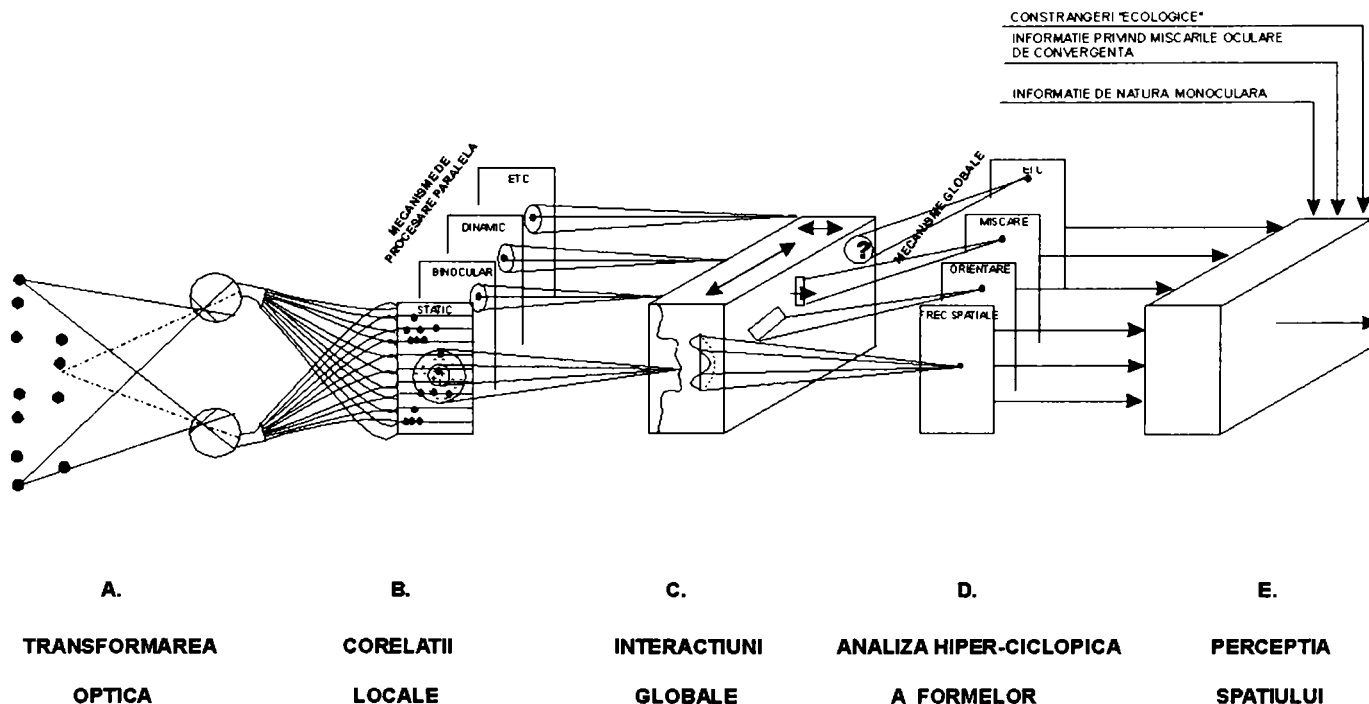


Figura 3.5: Procesarea in V1 (pasii B si C) contribuie substantial la perceptia spatiului 3-D, prin mecanisme paralel-ierarhizate. (Adaptata dupa Tyler 1995)

Figura 3.5 scoate in evidenta faptul ca informatia este preluata si procesata pana la un punct pe canale de procesare paralele, este apoi mixata si combinata pentru a se ajunge la perceptia spatiului 3-D. Reprezentarea lumii in ansamblu se obtine prin interactiunea reprezentarilor pariale (muchii, contururi, miscare, disparitati, constrangeri ecologice, etc.).

Principiul organizarii ierarhice (reprezentat in schema prin etapele distincte de procesare) deriva din organizarea pe straturi a cortexului – la nivelul fiecarui strat apare o condensare a informatiei, dar si o reorganizare a acesteia. Exista un mare grad de redondanta a informatiei, care se accentueaza pe masura ce procesarea avanseaza (Field 1995) – procesele paralele se suprapun in tot mai mare masura.

Efectele hiper-ciclopice se obtin din combinarea informatiilor privind frecventele spatiale, rezolutii diferite, orientari si contururi, miscarea, etc. Aceasta analiza se face in fazele medii de prelucrare a informatiei (ariile V2, V3), iar apoi se combina cu informatii din alte surse senzoriale si din memorie pentru a se obtine o corecta perceptie a spatiului 3-D (Tittle&Todd 1995).

Trebuie sa mentionez faptul ca inca de la nivelul cortexului striat exista mecanisme pre-attentionale (Amit&Geman 1999, Kowler 1995, Nakayama& Zijiang 1995), desi mecanismele de atentie propriu-zise apar abia in ariile mai inalte. Aceste mecanisme pre-attentionale previn supraincarcarea centrelor inalte de procesare si depind de istoria recenta a stimulării (Ahisar&Hochstein 1995). De asemenea, ele pot influenta timpul de reactie la stimuli.

Ramachandran (1995) chiar sugereaza ca este posibil ca dimensiunea campurilor receptoare sa nu fie constanta, ea putand sa se modifice functie de istoria stimulării, in paralel cu o reorganizare a distributiei lor in campul vizual. Interesante sunt si sugestiile lui cu privire la interpolarea suprafetelor – completarea obiectelor partial acoperite. Fenomenul de “umplere” a scotoamelor pare sa aiba loc intr-o faza primara de prelucrare si utilizeaza informatia pe o scara relativ larga.

In lucrarea de fata am abordat doua dintre mecanismele implicate in perceptia spatiului: detectia disparitatilor orizontale si intarirea muchiilor ce apartin unor contururi continue.

CE FEL DE IMAGINI SE PRELUCREAZA LA NIVELUL ARIEI VIZUALE PRIMARE?

V1 este, practic, poarta de intrare a informatiei vizuale in cortex, deci toate imaginile sunt pre-procesate la acest nivel. Asa cum se accentueaza in schema de procesare (figura 3.5), exista caracteristici care sunt detectate si/sau intarite in V1 in aceasta ierarhie de procese paralele (Albright&Neville 1999): se detecteaza muchiile (directia si intensitatea), se determina viteza si directia miscarii, se detecteaza disparitatile, etc.

CE FEL DE IMAGINI SUNT PRELUCRATE LA NIVELE MAI INALTE?

Toate imaginile ce contin contururi iluzorii, sau cele la care trebuie sa se faca completarea de contururi ca urmare a unor obiecte suprapuse sunt “intelese” (adica completate) in ariile mai inalte.



Figura 3.6: Scene a caror procesare se face la nivele mai inalte: (a) figura Kanisza cu contururi iluzorii; (b) “Dalmatianul” - scena la care este nevoie de implicarea sferei constiente; (c) figura dubla (oscileaza) – doua capete sau un piedestal.

De asemenea, scenele la care trebuie sa utilizam informatii din sfera constienta pentru a completa contururile (informatii legate de continutul/semnificatia imaginii), sau obiectele “duble” la care apare o oscilatie intre doua interpretari care se exclud reciproc – figura 3.6.

Trebuie totodata sa mentionez ca putem prelucra optim informatii pe care creierul nostru a fost pregatit (genetic sau prin “antrenarea” din primii ani de viata) sa o primeasca (Field 1995, Hertz 1995, Marr 1982). De aceea, scenele care nu se “conformeaza” sunt dificil de prelucrat – figura 3.7 prezinta cateva “obiecte imposibile”.

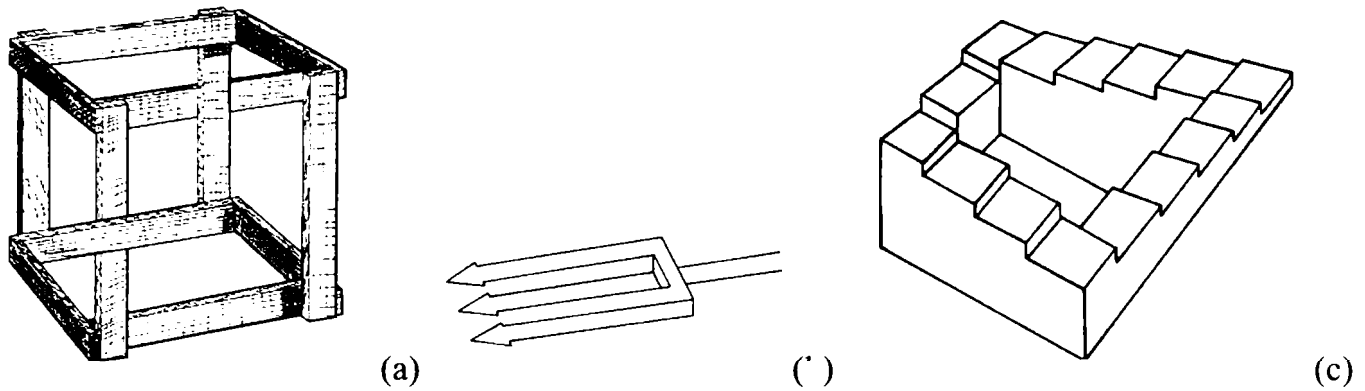


Figura 3.7: Obiecte imposibile

Utilizarea unor stimuli artificiali special conceputi pentru a “deruta” sistemul vizual constituie un instrument de studiu in experientele de psihofizica. Utilizand aceasta metoda de investigare a reactiilor noastre in conditii neobisnuite s-au tras concluzii foarte importante privind perceptia vizuala (Dayan 1998). Un exemplu celebru il constituie stereogramele cu puncte aleatoare pe care le voi prezenta in capitolul urmator.

Trebuie sa mai amintim de importanta contributie pe care au adus-o artistii la intelegerea unora dintre mecanismele perceptiei vizuale prin dezvoltarea tehnicilor de redare a realitatii 3-D pe suprafete 2-D (Cabanis et al. 1996, Cavanach 1999). Incepand cu Renasterea, tehnicile bazate pe perspectiva, ocluzii, gradient de textura si umbre s-au perfectionat continuu. Mai mult, s-au utilizat linii ce nu au ca si corespondent muchii reale, precum si umbre imposibile (admisibile local, dar in contradictie cu alte zone ale panzei) – faptul ca ele nu deranjeaza, ci dimpotriva, sunt percepute ca o imbogatire a imaginii, dovedeste ca procesarea lor se face la nivel local si nu la nivel global.

Capitolul 4: PERCEPTIA ADANCIMII IN VEDEREA UMANA. MODELAREA VEDERII BINOCULARE

4.1. PERCEPTIA ADANCIMII IN VEDEREA UMANA

Asa cum am aratat in schema de procesare a informatiei vizuale din figura 3.5, la perceptia spatiului contribuie o serie de informatii, atat de natura vizuala, cat si non-vizuala. Una dintre sursele importante de informatie vizuala este vederea binoculara, dar exista si o serie de alte **informatii vizuale monoculare**: miscarea (atat miscarea obiectelor, cat si miscarea noastra), perspectiva, diferentele de iluminare, efectul de incetosare la distante mai lungi, diferentele de textura, interpozitiile. La distante mai mari, axele celor doi ochi devin practic paralele si percepem adancimea utilizand aceste informatii monoculare. Dintre **informatiile non-vizuale**, putem aminti: miscarile de convergenta oculara si de acomodare (muschii “transmit” creierului informatii despre intinderea lor), familiarizarea privitorului cu obiectele, informatia auditiva.

Cercetarea mea s-a focalizat asupra informatiilor de natura binoculara – **vederea stereoscopica** (sau simplu *stereo*).

VEDEREA STEREOSCOPICA

Vederea stereo exploateaza faptul ca un punct din spatiul tridimensional al campului vizual se proiecteaza, in general, in pozitii diferite in cele doua imagini retiniene bi-dimensionale. Aceste diferente se numesc **disparitati binoculare**. Problema corespondentei stereo consta in obtinerea unei singure reprezentari a campului vizual 3-D pornind de la cele doua reprezentari 2-D. Aceasta “*image ciclopica*” utilizeaza in principal disparitatile si, de aceea, o importanta deosebita o are repartitia acestora in campul vizual – **harta de disparitate**. Problema obtinerii acestei harti (pornind de la cele doua imagini 2-D) a suscitat mult interes, atat pentru cei preocupati de intelegerea mecanismelor vederii, cat si pentru cei ce doresc sa utilizeze principiul in tehnica (pentru diverse masuratori terestre si aeriene, pentru echipamentele de vedere artificiala, etc.).

Problema corespondentei stereo este ilustrata in figura 4.1. Se observa ca in cazul in care cele doua camere sunt aliniate (asa cum este cazul ochilor nostri), problema se simplifica putin: punctele ce sunt “candidate” la corespondenta se gasesc de-a lungul celor doua **linii epipolare**.

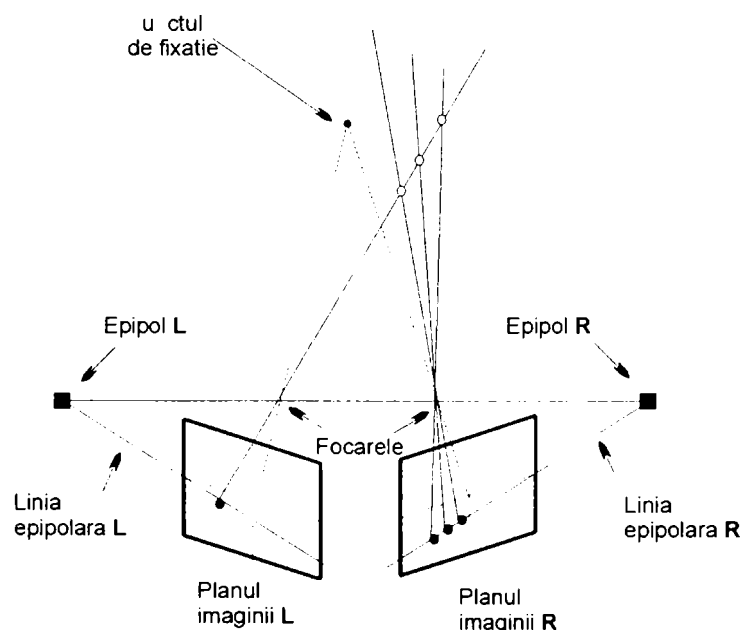


Figura 4.1: Ilustrarea problemei vederii stereo – gasirea corespondentelor (Dupa Frisby 1995).

Fiecare linie epipolara intersecteaza axa ce uneste focarele celor doua camere intr-un punct numit **epipol**. Importanta evidentierii liniilor epipolare in cele doua imagini consta in faptul ca ea permite reducerea cautarii corespondentelor la o singura dimensiune (de-a lungul acestor linii). Cele mai multe dintre modelele vederii stereo fac uz de aceasta "geometrie epipolara" si chiar simplifica mai mult problema: desi aceste linii sunt usor inclinate (in functie de distanta la punctul de focalizare) majoritatea modelelor fac cautarea pe linii paralele, orizontale in cele doua imagini. Totusi, chiar daca din punct de vedere al formalizarii teoretice este mai usor de descris problema asa (strategie abordata si de mine in continuare), din punct de vedere practic aceasta poate conduce la erori mari si la imposibilitatea rezolvarii situatiilor din lumea reala.

IMAGISTICA STEREO. CONSIDERENTE TEHNICE

Proiectarea unei scene 3-D pe un plan-imagini este o transformare descrisa de o functie care nu este injectiva: un punct din imaginea 2-D nu determina in mod unic un punct corespunzator in scena 3-D reala (figura 4.1). Tehnicile stereo de prelucrari de imagini implica obtinerea a doua imagini separate ale unui obiect sau scene de analizat – figura 4.2 (punctul **w**).

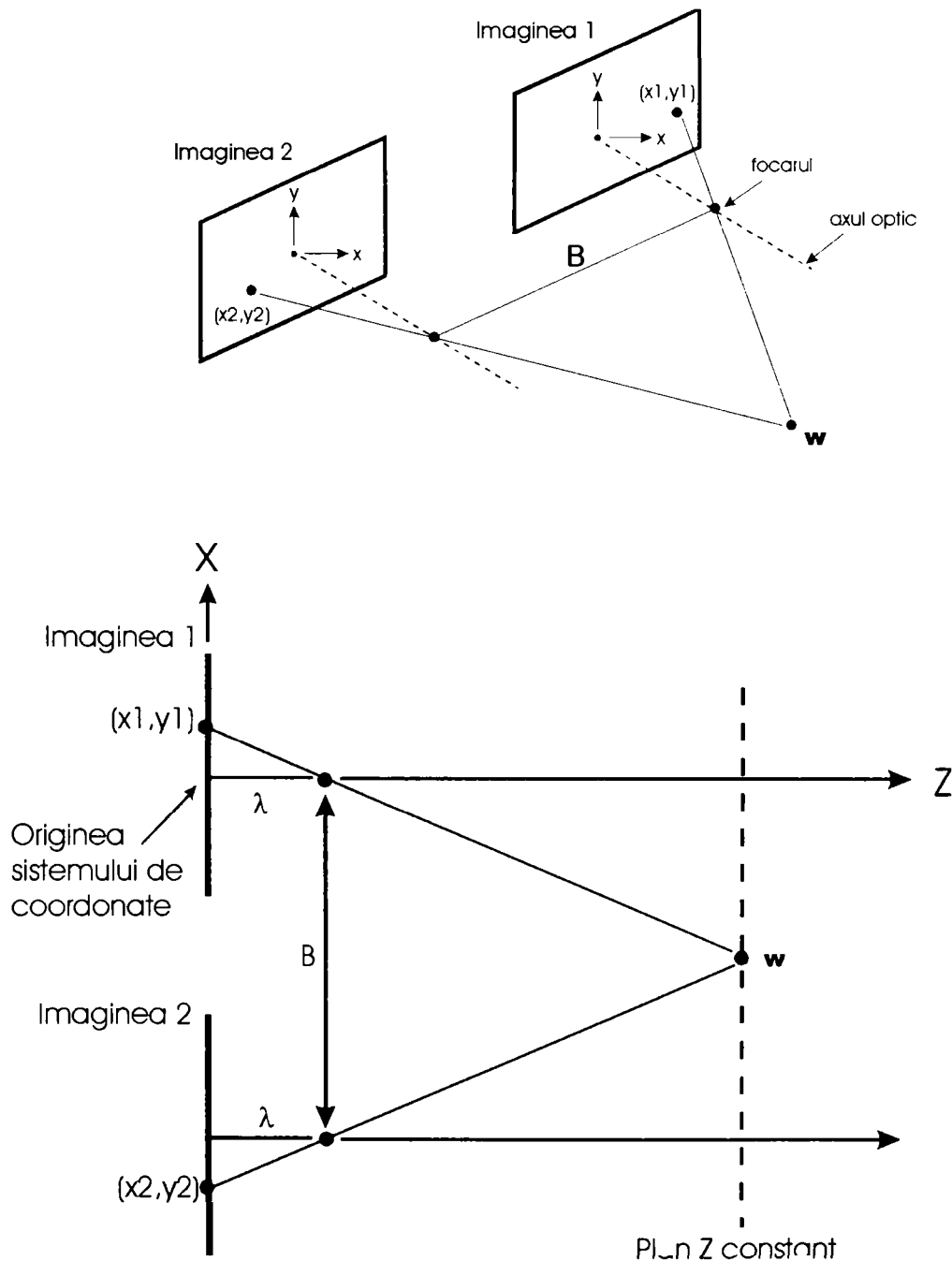


Figura 4.2: Ilustrarea procesarii stereo (Dupa Gonzalez&Woods 1993).

Dreapta ce uneste cele doua obiective este numita **linie de baza**. Scopul final este de a determina coordonatele (X, Y, Z) ale punctului w care este proiectat in punctele (x_1, y_1) si (x_2, y_2) pe cele doua imagini planare.

Pornim de la presupunerea ca cele doua camere sunt identice si ca sistemele de coordonate ale ambelor camere sunt perfect aliniata, diferind doar prin localizarea originilor (conditii de obicei respectate in practica), deci coordonata Z a punctului w este aceeaasi pentru ambele sisteme de coordonate (figura 4.2).

Sa presupunem ca alegem sistemul de coordonate reale astfel incat el sa coincida cu sistemul de coordonate al primei camere (figura 4.2 – vederea luata de sus). Atunci, w se va gasi la:

$$X_1 = \frac{x_1}{\lambda}(\lambda - Z), \quad \text{unde } \lambda \text{ este distanta focala.}$$

Daca (in mod similar) se alege originea sistemului de coordonate reale in cea de-a doua camera, w se va gasi la :

$$X_2 = \frac{x_2}{\lambda}(\lambda - Z).$$

Distanta dintre cele doua camere este B (linia de baza), deci:

$$X_2 = X_1 + B$$

$$Z_2 = Z_1 = Z$$

Din sistemul

$$\begin{cases} X_1 = \frac{x_1}{\lambda}(\lambda - Z) \\ X_1 + B = \frac{x_2}{\lambda}(\lambda - Z) \end{cases} \quad \text{rezulta} \quad Z = \lambda - \frac{\lambda B}{x_2 - x_1}$$

Rezultatul obtinut indica faptul ca, daca diferenta dintre coordonatele x_2 si x_1 in cele doua imagini planare poate fi determinata, iar linia de baza (B) si distanta focala (λ) sunt cunoscute, calcularea coordonatei Z (adancimea punctului w) este o problema simpla:

$$X = \frac{x_1}{\lambda}(\lambda - Z)$$

$$Y = \frac{y_1}{\lambda}(\lambda - Z)$$

Problema dificila consta in gasirea a doua puncte corespondente in cele doua imagini. Trebuie adaugat ca, datorita faptului ca dimensiunea proiectiei unui obiect pe retina este dependenta de distanta la acesta, disparitatile se exprima, in general, in unitati de arc.

Problema, asa cum am prezentat-o pare simpla, dar in realitate ea se complica destul de mult. Mai adaugam ca scenele care ne intereseaza nu sunt formate din puncte disparate, ci din suprafete inclinate (figura 4.3), curbe, etc. si problema determinarii corespondentelor si a celei de-a treia dimensiuni nu mai apare ca banala.

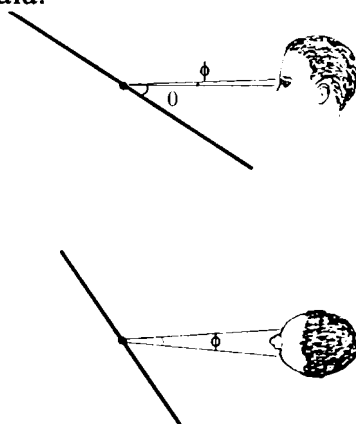


Figura 4.3: Perceptia suprafetelor complica problema stereo (Dupa Marr 1982).

Deoarece aceste puncte sunt, in general, in aceeasi vecinatate, o abordare uzuala in imagistica este selectarea unui punct intr-o regiune limitata intr-una din imagini si apoi gasirea corespondentului prin tehnici de corelare. Cand scena analizata contine caracteristici distincte, tehnicile de potrivire pe baza analizei caracteristicilor conduce la o rezolvare rapida.

Dar ce caracteristici sunt mai potrivite pentru a fi puse in corespondenta intre cele doua imagini 2-D? Muchiile si colturile au constituit alegerea unora dintre cercetatori, dar ele produc harti de disparitate imprastiate si necesita o interpolare intre puncte pentru a obtine o harta densa (adica pentru a determina valoarea de disparitate in fiecare punct al campului vizual), ceea ce ridica alte probleme. De obicei, interpolarea se face pe baza granitelor de textura.

Care a fost solutia aleasa de natura pentru sistemul nostru vizual?

DISPARITATEA BINOCULARA. EVIDENTE EXPERIMENTALE

Mecanismele din vederea stereo au inceput sa-i preocupe pe oameni abia in secolul trecut, printre pionierii in domeniu fiind Hermann Helmholtz si Charles Wheatstone. In 1838, Wheatstone a inventat **stereoscopul**, dispozitivul care permite obtinerea senzatiei de adancime utilizand perechi de imagini putin decalate pe orizontala. El a permis evidentierea faptului ca disparitatea binoculara constituie o informatie foarte puternica pentru sistemul nostru vizual.

Un alt moment important, de cotitura in domeniu a fost inventarea **stereogramelor cu puncte aleatoare** (**Random Dot Stereograms – RDS**) de catre Bela Julesz in 1960. Ele au permis izolarea disparitatii binoculare de celelalte indicii privind adancimea si au demonstrat ca recunoasterea monoculara a formelor nu este o cerinta a fuziunii binoculare, ea se petrece mult mai devreme in lantul de procesare a informatiei vizuale.

Stereogramele cu puncte aleatoare (figura 4.4) constau din puncte albe si negre generate aleator (cu distributie uniforma pentru "varianta clasica").

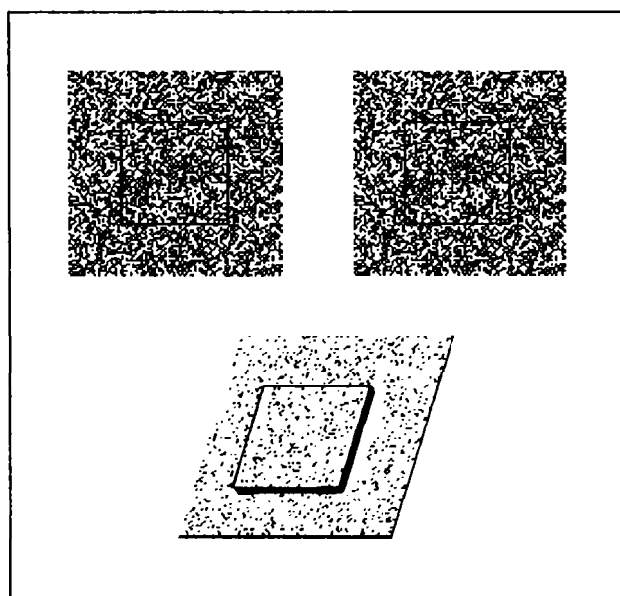


Figura 4.4: Principiul de creare a unei RDS

Imaginea dreapta se formeaza din cea stanga (sau invers) prin deplasarea laterala a unor subseturi din imagine (in figura 4.4 s-a indicat zona patratica din centru care a fost deplasata spre stanga cu 2 pixeli).

Daca privim o astfel de stereograma printr-un stereoscop (pentru a putea face mai usor fuziunea) vom percepe zonele deplasate lateral ca fiind deplasate in profunzime fata de fondul imaginii (in figura 4.4 zona centrala "vine" spre privitor). Senzatie de profunzime depinde de sensul si marimea deplasarii pe orizontala a ariilor "decupate".

In experimentele de psihofizica, se utilizeaza RDS-uri cu diferite distributii si densitati de umplere, cu nuante de gri, cu diverse zgomote suprapuse sau cu diferente de luminozitate, etc.

(Julesz 1995). De asemenea, se utilizeaza cinematograme cu puncte aleatoare - RDS-uri la care anumite zone se gasesc in miscare (Miclea 1999, Wilson&Keil 1999).

Introducerea tehnicilor ce utilizeaza RDS-uri a permis studierea rolului disparitatilor binoculare in perceptia adancimii, precum si robustetea mecanismelor vederii stereo. Ele sunt, totodata, o metoda prin care neurofiziologii au putut identifica zonele acordate pentru prelucrarea de disparitate in cortexul vizual.

Modelele computationale ale vederii stereo se bazeaza intr-o masura destul de mare pe cercetarile cu RDS-uri si utilizeaza aceste stereograme ca un etalon standard pentru testarea algoritmilor stereo.

Daca sistemul nostru vizual nu potriveste muchii si colturi, atunci care sunt primitivele pe baza carora se face totusi potrivirea? Pentru ca s-au pus in evidenta celule sensibile la disparitate la nivelul celulelor complexe din V1 – in 1967 s-au facut primele inregistrari experimentale la pisica, apoi la maimuta (Ramachandran 1995). Sistemele tehnice fac de multe ori uz de obiecte pe care le pun in scena si pe baza carora fac mai apoi potrivirea imaginilor 2-D. Despre sistemul nostru vizual s-a spus ca ar face uz de disparitatile verticale si de informatiile date de sistemul oculomotor pentru calibrare (Frisby 1995). Probabil ca se utilizeaza si aceste mecanisme (pentru ajustarile fine ale perceptiei), dar RDS-urile arata clar ca disparitatea binoculara orizontala in sine constituie un indiciu foarte important pentru sistemul vizual si ca fuziunea se petrece foarte devreme in lantul de procesare vizuala. Pornind de la aceste intrebari, cercetarea in domeniul vederii stereo si-a gasit multi adepti.

4.2. MODELAREA VEDERII BINOCULARE

Asa cum am aratat deja, extragerea informatiei 3-D din disparitate binoculara consta din doua etape: stabilirea corespondentelor si interpretarea geometrica a disparitatii (Papathomas et al. 1995). Tehnicile actuale de stabilire a potrivirilor dintre imaginea stanga si cea dreapta se pot clasifica in doua categorii: continue si discrete.

Tehnicile continue sau *area-based techniques* (Marr&Poggio 1976, Marr et al. 1978, Hsieh et al. 1992, Ludwig et al. 1994, Qian 1994, Qian&Mikaelian 2000) obtin o harta de disparitate continua pe tot campul vizual, de obicei prin tehnici de corelare a imaginilor, aplicand diverse tipuri de filtre pe imagini (pentru a extrage diversele orientari spatiale si frecvente spatiale sau temporale).

Tehnicile discrete sau *feature-based techniques* (Hsieh 1992, Marr&Poggio 1979, Nasrabadi&Choo 1992, Cruz et al. 1995, Hongo et al. 1996) selecteaza caracteristici discrete in cele doua imagini si apoi incearca sa le potriveasca, obtinand o harta de disparitate discontinua (*spase*). In urma evidentelor experimentale, este clar ca vederea umana nu utilizeaza aceste mecanisme.

Pe langa constrangerea de epipolaritate, in modelarea stereoscopiei au mai fost formulate cateva **constrangeri** impuse de lumea in care traim (Marr & Poggio 1976, Marr 1982, Frisby 1995):

- **Constrangerea de compatibilitate a primitivelor puse in corespondenta.** Problema ambiguitatilor poate fi substantial redusa daca se considera in problema corespondentei doar caracteristicile cu forme si contraste similare (aceasta strategie a fost deja amintita la problema "colturilor" din sistemele artificiale). Ea nu elimina cu totul ambiguitatile din scenele naturale. David Marr a exprimat aceasta conditie ca: *se pun in corespondenta doar acele caracteristici din imaginea stanga si cea dreapta care ar fi putut lua nastere din aceeasi entitate a scenei.*
- **Constrangerea de coeziune.** David Marr si Tomaso Poggio au argumentat faptul ca lumea pe care o percepem vizual este formata din materie compacta, dar separata in obiecte si este rezonabil sa presupunem ca suprafata acestora este, in general, neteda comparativ cu distanta pana la privitor (adica lumea nu este formata din nori de particule de praf sau fulgi de zapada). Ei au formulat conditia astfel: *se prefera corespondentele care ar putea proveni de la suprafete*

netede. Ca un corolar, disparitatea variaza lin aproape peste tot, cu exceptia granitelor. Adeseori, implementarea acestei constrangeri conduce la fortarea fronto-parallelismului suprafetelor.

- **Constrangerea de unicitate.** David Marr si Tomaso Poggio au remarcat si faptul ca o entitate a scenei nu se poate gasi in doua locuri in acelasi timp si au formulat urmatoarea regula de punere in corespondenta: *fiecarei primitive din fiecare imagine i se poate asigna cel mult o valoare de disparitate.* Aceasta constrangere este valabila atunci cand suprafetele sunt presupuse opace.
- **Constrangerea de continuitate figurala.** Aceasta conditie impune urmatoarea regula: obiectele coezive genereaza muchii si delimitatori ai suprafetelor care sunt continui spatia; de aceea, se prefera corespondentele care pastreaza continuitatea figurala. Conditia a fost formulata de Mathew si Frisby in 1981 (Frisby 1995).
- **Constrangerea de ordonare.** Daca se presupune opacitatea suprafetelor, atunci se poate exploata faptul ca ordinea primitivelor de-a lungul unei linii epipolare intr-una dintre imagini este pastrata si in cealalta imagine. Se pare ca sistemul vizual uman nu incalca niciodata aceasta constrangere (Frisby 1995).

Dintre modelele care au reprezentat jaloane importante in cercetarea mecanismelor vizuale, cel propus de David Marr si Tomaso Poggio (Marr & Poggio 1976; Marr et. al. 1978) a ramas un punct de referinta. De aceea, il voi descrie pe scurt aici.

MODELUL MARR&POGGIO

Modelul propus de ei consta dintr-o retea de tipul celei din figura 4.5 (imaginea reprezinta o sectiune printr-o retea corespunzatoare unei imagini bidimensionale, dar formata din 9 plane de disparitate).

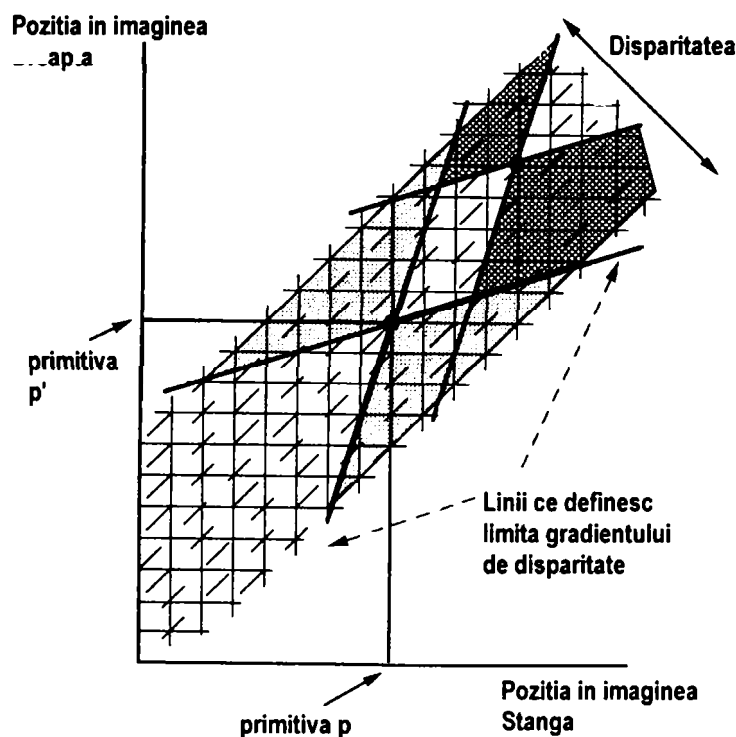


Figura 4.5: O retea de tipul celei propuse initial de D. Marr si T. Poggio pentru determinarea disparitatilor binoculare. (Dupa Frisby 1995)

Liniile continui verticale si orizontale reprezinta pozitiile primitivelor in imaginea stanga, respectiv cea dreapta, iar punctele lor de intersectie corespund unor valori posibile de disparitate (potriviri potentiale). Liniile intrerupte diagonale conecteaza punctele cu aceeasi disparitate. In fiecare nod al retelei se gaseste cate o unitate de procesare (o celula).

Constrangerea de unicitate impune ca toate conexiunile reprezentate cu linie continua (pe orizontala si pe verticala) sa fie inhibitorii - fiecare primitiva (sau element de imagine) poate avea o singura valoare de disparitate.

Constrangerea de coeziune impune ca de-a lungul diagonalelor conexiunile sa fie excitatorii.

Pentru imagini bi-dimensionale, liniile din figura 4.5 devin plane. Structura vecinatatii va fi cea din figura 4.5', in care cerculetele goale reprezinta vecinii din zona inhibitoare, iar cerculetele pline vecinii din zona de excitatie.

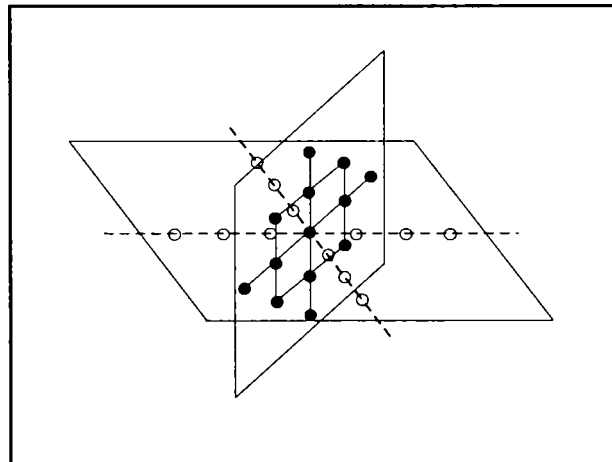


Figura 4.5': Vecinatatea unui neuron din retea Marr&Poggio: cerculetele pline reprezinta conexiuni excitatorii, iar cele goale sunt conexiuni inhibitorii

Algoritmul de rezolvare a corespondentelor este iterativ, retea atinand o stare stabila atunci cand sunt indeplinite conditiile de unicitate si coeziune (Marr et al. 1978). Algoritmul consta in urmasorii pasi:

- se noteaza cu $L_{x,y}$ si $R_{x,y}$ pozitiile unei primitive si cu $C_{x,y;d}^{(t)}$ starea unei celule corespunzand coordonatelor (x,y) la momentul t
- algoritmul de functionare a retelei va fi

$$C_{x,y;d}^{(0)} = L_{x,y} \cdot R_{x+d,y}$$

$$C_{x,y;d}^{(t+1)} = \sigma \left\{ \sum_{x',y',d' \in S(x,y,d)} C_{x',y';d'}^{(t)} - \varepsilon \sum_{x',y',d' \in O(x,y,d)} C_{x',y';d'}^{(t)} + C_{x,y;d}^{(0)} \right\} \quad (4.1)$$

unde

d este disparitatea (in domeniul acceptat de disparitati)

ε este constanta de inhibitie

σ este functia sigmoida in domeniul $[0,1]$

$S(x,y,d)$ este vecinatatea excitatorie a lui $C_{x,y;d}$

$O(x,y,d)$ este vecinatatea inhibitorie a lui $C_{x,y;d}$

- Marr si Poggio au testat retea utilizand RDS-uri.

Algoritmul se dorea plauzibil biologic atunci cand a fost formulat. Experimentele si studiile care s-au facut de atunci au lamurit multe din necunoscute si s-a dovedit ca unele din presupunerile lui Marr si Poggio nu erau corecte. Algoritmul ramane foarte important prin faptul ca a deschis o abordare noua in problema modelarii proceselor vizuale si, de asemenea, prin robustetea lui. El este un termen de referinta pentru abordarile si extinderile mai recente (Frisby 1995, Khotanzad 1993).

Alte modele utilizeaza retele neuronale mai simple (Churchland&Sejnowski 1992), dar ele prezinta serioase limitari si nu fac concurenta serioasa modelului Marr&Poggio.

Trebuie sa mentionez si modele tehnice care sunt de obicei *feature-based* (Cruz 1995, Hongo et al. 1996, Hsieh&McKeon 1992, Nasrabadi&Choo 1992, Weng 1992) si nu sunt interesante din punctul de vedere al modelarii plauzibile biologice. Totusi, o harta de disparitate obtinuta de Nasrabadi &Choo am utilizat-o ca si termen de comparatie (a fost disponibila pe WEB).

Indiferent de tehnica abordata in reconstituirea informatiei 3-D, sunt aspecte care inca nu au fost cu adevarat rezolvate si una delicata este aceea a **suprafetelor transparente**, desi vederea umana rezolva foarte bine aceasta problema. Constrangerea unicitatii, care ajuta mult in modelele de vedere stereo, este prea restrictiva – Frisby (1995) a propus solutii, dar problema ramane nerezolvata in mod real.

4.3. MODEL “BIOLOGIC” PENTRU HARTA DE DISPARITATE BINOCULARA

PREZENTAREA PRINCIPIULUI

Modelul pe care l-am gandit si l-am implementat se bazeaza pe caracteristicile celulelor simple si complexe din V1 si utilizeaza un principiu prezent si la alte modele, sugerat si de Marr (Marr 1982, Qian 1994, Qian&Mikaelian 2000): filtrarea imaginilor cu un set de filtre care sa extraga caracteristicile imaginii in mod similar celui in care celulele receptoare din V1 o fac. Am utilizat in acest scop filtrele Gabor, unanim acceptate ca un model bun pentru campurile receptoare ale celulelor corticale simple (asa cum am aratat in Capitolul 3).

In prezentarea principiului de baza, voi face o serie de simplificari, considerand **conditia de epipolaritate satisfacuta** si voi generaliza dupa aceea.

Deci voi considera ca eventuala “potrivire” dintre imaginile unei stereograme trebuie cautata de-a lungul liniilor epipolare, iar pe acestea le consider horizontale (practic, pixelii de pe randuri). Campurile receptoare vor fi deci functii Gabor unidimensionale:

$$\begin{aligned} f_l &= e^{-\frac{x^2}{2\sigma^2}} \cos(\omega x + \varphi_l) \\ f_r &= e^{-\frac{x^2}{2\sigma^2}} \cos(\omega x + \varphi_r) \end{aligned} \quad (4.2)$$

unde σ este marimea campului receptor, iar ω este frecventa spatiala preferata

O celula simpla va avea raspunsul dependent de produsul de convolutie al campului ei receptor pe imaginea analizata

$$r_s = \int_{-\infty}^{\infty} [f_l(x) \cdot I_l(x) + f_r(x) \cdot I_r(x)] dx \quad (4.3)$$

unde $I_l(x)$ si $I_r(x)$ reprezinta imaginea retinala stanga/dreapta pentru un acelasi stimul.

Disparitatea binoculara D pentru acelasi stimul se poate scrie:

$$\begin{aligned} I_l(x) &= I(x) \\ I_r(x) &= I(x + D) \end{aligned} \quad (4.4)$$

Atunci, in conditii in care largimea campului receptor este mult mai mare decat disparitatea, se poate spune ca raspunsul celulei simple va fi

$$r_s \approx 2\rho \cos\left(\theta + \frac{\varphi_l + \varphi_r}{2} - \frac{\omega D}{2}\right) \cos\left(\frac{\varphi_l - \varphi_r}{2} + \frac{\omega D}{2}\right) \quad (4.5)$$

unde ρ si θ reprezinta amplitudinea, respectiv faza transformatei Fourier a imaginii $I(x)$ la frecventa ω . Daca se combina raspunsurile mai multor celule simple, ar trebui sa se obtina raspunsul unei celule complexe.

Ne intereseaza ca raspunsul celulei complexe sa fie dependent de diferenta de faza dintre campurile receptoare si independent de parametrii transformatei Fourier (neplauzibil a fi efectuata la nivel cortical). Daca vom utiliza doua celule simple cu aceleasi σ , ω , $(\phi_l - \phi_r)$, dar cu $(\phi_l + \phi_r)/2$ la

diferenta de faza de $\pi/2$ si raspunsul unei celule complexe ca fiind $R_c = r_{s1}^2 + r_{s2}^2$ (ceea ce este plauzibil biologic), atunci raspunsul acesteia va deveni:

$$R_c \approx 4\rho^2 \cos^2\left(\frac{\varphi_l - \varphi_r + \omega D}{2}\right) \quad (4.6)$$

si se vede ca el este dependent doar de disparitate si de diferenta de faza. Functia \cos^2 este periodica cu perioada π , ceea ce conduce la un interval restrictiv pentru $\omega D/2$. Intervalul de disparitate pe care il va putea detecta celula va fi $2\pi/\omega$ si ea va raspunde maxim pentru o disparitate preferentiala data de:

$$D_{pref} = \frac{\varphi_r - \varphi_l}{\omega} \quad (4.7)$$

Daca vom considera o familie de celule complexe pentru fiecare locatie retiniana, cu $(\phi_l - \phi_r)/2$ acoperind domeniul $[-\pi/2, \pi/2]$, aceste celule vor forma o reprezentare distribuita a disparitatii, fiecare fiind "acordata" pentru a sesiza daca in locatia de care este ea "raspuzatoare" apare valoarea respectiva de disparitate. Aceste concluzii sunt in concordanta cu descoperirile biologice pe care le-am prezentat in capitolele precedente.

Trebuie sa mentionez ca in modelul "biologic" am folosit familii de functii Gabor bi-dimensionale:

$$G(x, y) = e^{-\pi\left[\frac{(x-x_0)^2}{\alpha^2} + \frac{(y-y_0)^2}{\beta^2}\right]} e^{-2\pi i[u_0(x-x_0) + v_0(y-y_0)]} \quad (4.8)$$

unde:

(x_0, y_0) specifica pozitia in imagine

(α, β) specifica latimea si lungimea filtrului

(u_0, v_0) precizeaza vectorul-unda de modulare specific filtrului

frecventa spatiala $\omega_0 = \sqrt{u_0^2 + v_0^2}$; orientarea sau directia $\theta_0 = \arctan\left(\frac{v_0}{u_0}\right)$.

PARAMETRII MODELULUI

Pentru implementarea si testarea modelului am realizat o aplicatie in Visual C++, descrisa pe scurt in Anexa 4, care permite crearea de RDS-uri si testarea modelului.

Pentru fiecare locatie spatiala, am utilizat o **familie de celule complexe** acordate pe o anumita valoare de disparitate (intre -4 si 4), care se vor constitui intr-o harta de disparitate pentru stereograma procesata.

In mod normal, rezolutia in sistemul vizual (deci si rezolutia de disparitate) se exprima in arc-grade, dimensiunea imaginii proiectate pe retina fiind dependenta de distanta privitor - punct de focalizare. Rezolutia vizuala obisnuita este de 1 arc-min (1/60 grade). Un fenomen remarcabil este insa hiperacuitatea in detectia muchiiilor, dar si a disparitatilor, care ajung pana la 2 arc-sec ($\approx 25 \mu\text{m}$ vazuti de la 25 cm), mult sub diametrul celulelor receptoare. De asemenea, fuziunea binoculara are loc (chiar daca cu putin efort) pana la disparitati extrem de largi, functie de tipul de stimuli si de istoria stimularii.

De asemenea, functie de imaginea analizata, o aceeaasi rezolutie spatiala reala se "traduce" in una unghiulara diferita, deci intr-o disparitate a localizarii retiniene diferita. Acest lucru inseamna ca nu putem stabili o relatie directa intre disparitatea unghiulara si cea din imaginile analizate.

Deoarece experientele cu RDS-uri au dovedit ca fuziunea se produce in mod normal intre limitele $[-4, 4]$, aceasta a fost plaja de disparitati spatiale considerate in implementarile facute (care apar in rezultatele obtinute). Trebuie subliniat insa ca acest fapt constituie un aspect secundar - algoritmul ramane valabil pentru orice alta plaja de disparitati (care trebuie aleasa functie de criterii ce includ rezolutia).

Am utilizat $(\alpha, \beta) = (4, 8)$, iar numarul de celule complexe pentru fiecare locatie a fost 9×5 (orizontal \times vertical).

Am testat diferite combinatii (u_0, v_0) , cu $\omega_0 \approx 1/8$ (in acceptiunea $\omega_0 = \sqrt{u_0^2 + v_0^2}$) – de la filtre Gabor uni-dimensionale ($u_0=1/8, v_0=0$) la diferite filtre bi-dimensionale.

Harta de disparitate a fost reprezentata sub forma de grafice de tip *wireframe* si de imagini in nuanțe de gri (nuanțe deschise pentru planele de disparitate apropiate si inchise pentru cele departate). Am preferat hartile in nuanțe de gri pentru prezentarea rezultatelor in aceasta lucrare.

Am implementat si rețeaua Marr&Poggio pentru a avea un termen de comparatie. Pentru aceasta rețea am considerat 9 plane de disparitate, in plaja $[-4, 4]$, similare cu cele ale modelului “biologic”.

La crearea RDS-urilor, aplicatia realizata creaza automat o harta de disparitate “ideala” pe care am utilizat-o in evaluarea performantelor modelului. Ca si criterii de performanta am considerat procentul de pixeli eronati si eroarea medie patratica la estimarea valorii de disparitate pentru punctele unei imagini. Aplicatia realizata mi-a permis sa aplic si diferite tipuri de filtrare, precum si zgomot de natura si nivel diferite.

4.4. REZULTATE. DISCUTII

REZULTATE OBTINUTE PE RDS-URI

Primele teste le-am facut pe stereograme cu puncte aleatoare negre si albe (RDS-uri B&W) cu distributie spatia uniforma si Gaussiana (centrata in centrul imaginii si $\sigma=1$).

Stereogramele de test au fost de 200×200 pixeli – tabelul 1 sintetizeaza rezultatele obtinute pentru diferite densitati de umplere ale imaginilor.

Mentionez ca am aplicat si diferite tipuri de filtre ulterioare obtinerii hartii de disparitate brute: filtrare uniforma, filtrare mediana, filtrare prin metoda celei mai frecvente valori (notata *cmfv* in tabele si figuri), precum si un tip de filtrare adaptiva care mi s-a parut potrivit (Saint-Marc et al. 1991), deoarece face o “evaluare” a caracteristicilor locale ale imaginii, compatibile cu prelucrarea in V1.

TABEL 1
 Comparatie intre algoritmi aplicati pe RDS de 200×200 pixeli B&W, in termenii procentului de eroare (%) si ai erorii medii patratic

Distributia pixelilor	Model	Densitatea RDS (puncte negre)				
		30%	40%	50%	60%	70%
uniforma	Marr&Poggio	11.9950%	7.4100%	7.3100%	9.3475%	13.0580%
		0.6510	0.4476	0.4204	0.4722	0.6743
	"biologic" fara filtrare	30.2630%	27.1930%	27.9750%	27.1030%	27.8400%
		3.5979	3.2878	2.7989	2.9540	1.9602
	"biologic" +filt. median	26.8630%	23.6330%	21.7080%	22.4100%	18.3930%
		1.4256	1.5318	1.0861	1.0852	0.6060
	"biologic" +filt. cmfv	17.8650%	16.7900%	15.9130%	16.2150%	15.4950%
		1.8103	1.8996	1.3368	1.4317	0.9044
	"biologic" + filt. adapt.	34.9580%	31.0750%	29.8400%	30.0830%	26.2080%
		1.8459	1.7905	1.3747	1.5369	0.9186
"biologic" +filt. cmfv+filt. adapt.	21.7180%	20.4850%	18.5430%	19.6100%	16.1780%	
		1.0147	1.179	0.7078	0.8162	0.5148
Gauss	Marr&Poggio			7.8975%		
				0.3042		
	"biologic" +filt. cmfv+filt. adapt			18.6030%		
				0.8722		

Mentionez ca am utilizat filtre Gabor uni-dimensionale in aceste prelucrari, alternativa ce a fost sugerata de modul in care iau nastere RDS-urile (deplasari strict pe orizontala).

Figurile 4.6 si 4.7 prezinta graficele de eroare pentru cele doua criterii considerate, in cazul prelucrarilor pe RDS-uri cu distributie uniforma cuprinse in tabelul 1.

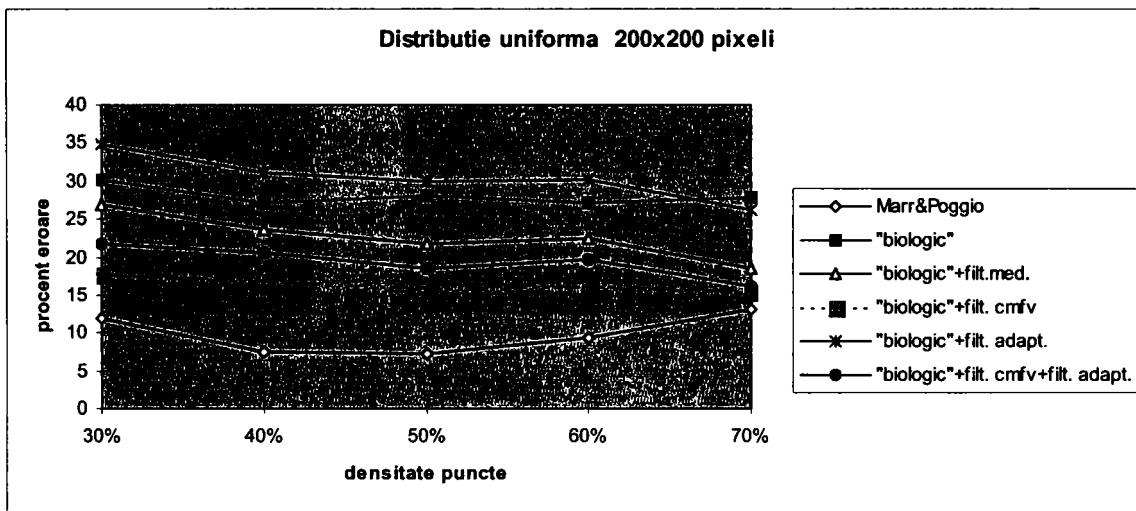


Figura 4.6: Comparatie a performantelor celor doi algoritmi implementati, utilizand RDS-uri B&W cu distributie uniforma. Criteriu utilizat - procentul de pixeli eronati

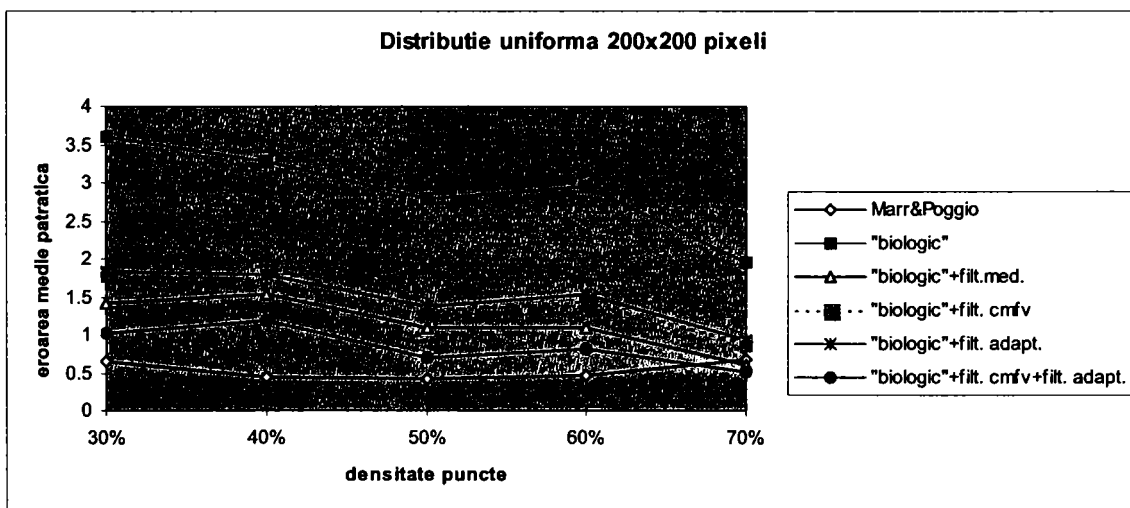


Figura 4.7: Comparatie a performantelor celor doi algoritmi implementati, utilizand RDS-uri B&W cu distributie uniforma. Criteriu utilizat -eroarea medie patratica a nivelului de disparitate

Valorile din tabel si graficele par sa sugereze ca algoritmul Marr&Poggio are performante mai bune (atat pentru distributia uniforma, cat si pentru cea normala). Figura 4.8 prezinta cateva harti de disparitate rezultate in urma prelucrarilor si se poate observa ca aspectele calitative sugereaza exact contrariul: algoritmul Marr&Poggio deterioreaza puternic formele. Este adevarat ca pastreaza un nivel de disparitate constant (in timp ce algoritmul "biologic" obtine o harta cu un nivel de zgomot ridicat), dar efectele asupra "continutului" imaginilor este negativ.

Algoritmul biologic pastreaza bine formele, iar alegerea unor filtrari potrivite poate imbunatati considerabil performantele. Cea mai buna combinatie pare sa fie filtrarea prin metoda cea mai frecventa valoare, urmata de filtrarea adaptiva.

Se mai poate observa ca nici unul dintre algoritmi nu este influentat substantial de tipul de distributie, ci au comportamente similare in cele doua situatii considerate.

Densitatile mici influenteaza negativ performantele ambelor modele, dar cel "biologic" pare sa le tolereze mai bine pe cele ridicate.

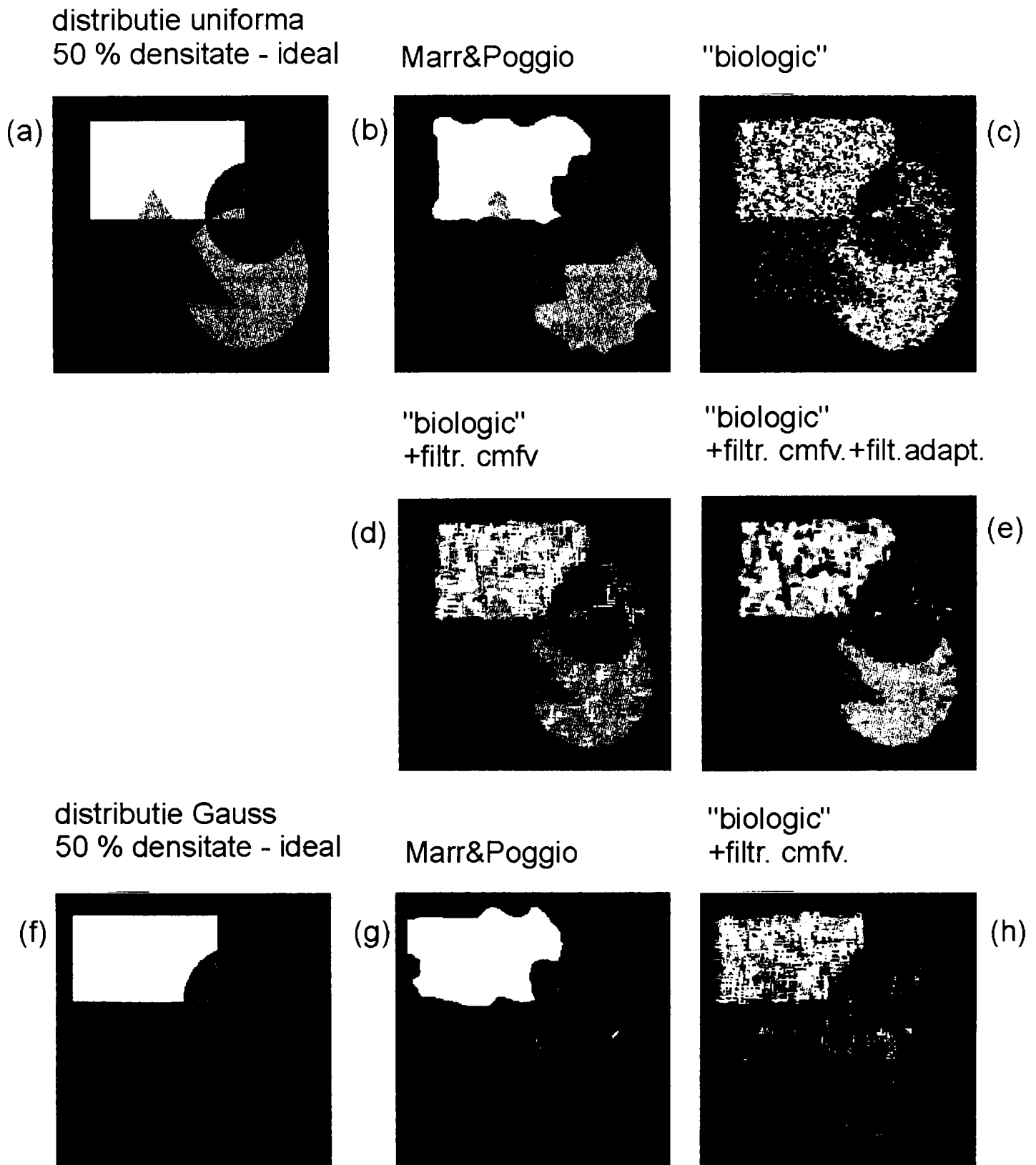


Figura 4.8: Harti de disparitate obtinute pe RDS-uri B&W de 200×200 puncte, uniform si normal distribuite: (a) si (f) sunt hartile "ideale" ce constituie termenele de comparatie; "filtr. cmfv" se refera la filtrare de tip cea mai frecventa valoare. Rezultatele au fost obtinute pentru $u=1/8$ si $v=0$ (filtrare Gabor uni-dimensionala).

Am testat modelul "biologic" si pentru distributii invers-Gaussiene, precum si pentru diferite valori σ , iar rezultatele au fost comparabile cu cele prezentate. Aceeasi a fost situatia pentru RDS-uri cu puncte colorate. Performantele s-au inrautatit dramatic in apropierea limitelor de disparitate considerate, dar acest lucru se datoreaza valorilor considerate pentru ω . Daca am fi considerat familii de celule complexe care sa acopere o plaja de frecvente spatiale, acest neajuns ar fi fost inlaturat.

In vederea biologica, atunci cand se ajunge la limita de disparitate nu se mai face fuziunea binoculara si se obtin imagini diplopice.

Tabelul 2 prezinta rezultate obtinute cu modelul "biologic" atunci cand am utilizat filtre Gabor bi-dimensionale pentru modelarea comportamentului celulelor simple si am prelucrat RDS-uri. Se observa ca am pastrat frecventa spatiala aproximativ aceeasi, dar orientarea undelor a variat.

TABEL 2
 Comportamentul algoritmului "biologic" pe RDS de 200x200 pixeli B&W (50% densitate),
 in termenii procentului de eroare (%) si ai erorii medii patratice

valorile u v	valorile θ ω	Distributia pixelilor	Modelul "biologic" urmat de filtrare		
			fara filtrare	filt. cmfv	filt. cmfv+filt. adapt.
u=1/8 v=0	$\theta=0$ $\omega=0.125(1/8)$	uniforma	27.9750%	15.9130%	18.543%
			2.7989	1.3368	0.7078
		Gauss ($\sigma=1$)	27.0100%	16.3650%	18.6030%
			2.6640	1.4924	0.87223
u=1/8 v=1/8	$\theta=45^\circ$ $\omega=0.118$	uniforma	35.7680%	34.5800%	34.8480%
			2.1917	1.2212	1.0025
		Gauss ($\sigma=1$)	30.6500%	27.6330%	27.7400%
			2.0710	1.1874	0.98223
u=1/12 v=1/12	$\theta=45^\circ$ $\omega=0.177$	uniforma	46.7300%	46.3450%	46.3580%
			2.6802	2.3221	2.3163
		Gauss ($\sigma=1$)	42.1650%	41.3620%	42.0930%
			2.2015	2.02	2.0881
u=1/8 v=1/12	$\theta=37^\circ$ $\omega=0.15$	uniforma	31.5630%	25.3600%	25.4450%
			2.2029	0.9179	0.6798
		Gauss ($\sigma=1$)	29.2250%	22.2280%	21.2080%
			2.1283	0.9520	0.7002
u=1/10 v=1/12	$\theta=44^\circ$ $\omega=0.13$	uniforma	44.1650%	43.8705%	44.7880%
			2.8890	1.9050	1.7445
		Gauss ($\sigma=1$)	41.6730%	40.9050%	41.8700%
			2.3156	2.0050	1.9395
u=1/12 v=1/10	$\theta=55^\circ$ $\omega=0.13$	uniforma	46.6530%	46.0050%	46.3380%
			2.7389	2.2030	2.1900
		Gauss ($\sigma=1$)	42.1530%	42.0890%	42.0930%
			2.2208	1.0930	2.0797

Rezultatele s-au inrautatit considerabil in situatia utilizarii filtrelor bi-dimensionale, ele introducand factori perturbatori suplimentari. Acest lucru era de asteptat: aceste imagini RDS sunt artificiale, create prin chiar deplasare laterala stricta.

Atunci cand privim pentru prima data o astfel de stereograma cu un stereoscop (mai ales pentru cele mai complexe, cu multe figuri geometrice, ca cele utilizate in testele facute), avem nevoie de putin timp pentru acomodare. Cred ca acest lucru este o dovada ca si pentru sistemul vizual uman RDS-urile constituie o problema – imaginile reale le fuzionam instantaneu. S-ar putea sa "ajute" la aceasta fuzionare rapida indiciile suplimentare pe care le avem (vizuale sau nu), iar atunci cand suntem restrictionati la disparitatea binoculara sa facem fuzionarea mai dificil. S-ar putea ca aceasta harta de disparitate obtinuta strict din disparitatile orizontale sa aiba zgomot ridicat si pentru sistemul biologic.

De aceea, rezultatele nu surprind, ci vin ca o confirmare ca presupunerile privind mecanismele utilizate sunt plauzibile.

Tabelul 3 prezinta influenta zgomotului adaugat RDS-urilor – se observa ca efectul nu este dramatic pana la un nivel de aprox. 5%, aspect ilustrat si de figurile 4.9 si 4.10.

TABEL 3

Comparatie intre algoritmii aplicati pe RDS de 200x200 pixeli B&W uniform distribuiti (50% densitate), cu zgomot adaugat, in termenii procentului de eroare (%) si ai erorii medii patratice

Distributia zgomotului	Model	Nivelul de zgomot			
		fara zgomot	5% pe Stanga	10% pe Dreapta	5% pe Stg. & 5% pe Dr. (indep.)
uniforma	Marr&Poggio	7.3100%	9.4425%	14.7080%	
		0.4204	0.57168	0.8053	
	"biologic" +filt. cmfv	27.9750%	25.6950%	37.4100%	36.5600%
		2.7989	1.5255	2.3487	2.3577
	"biologic" +filt. cmfv+filt. adapt.	18.5430%	24.9100%	36.5300%	35.9650%
		0.7078	0.8585	1.4305	1.4166
Gauss	"biologic" +filt. cmfv		19.8530%	23.5050%	
			1.6255	1.7319	
	"biologic" +filt. cmfv+filt. adapt.		21.0300%	24.5050%	
			0.9196	0.9801	

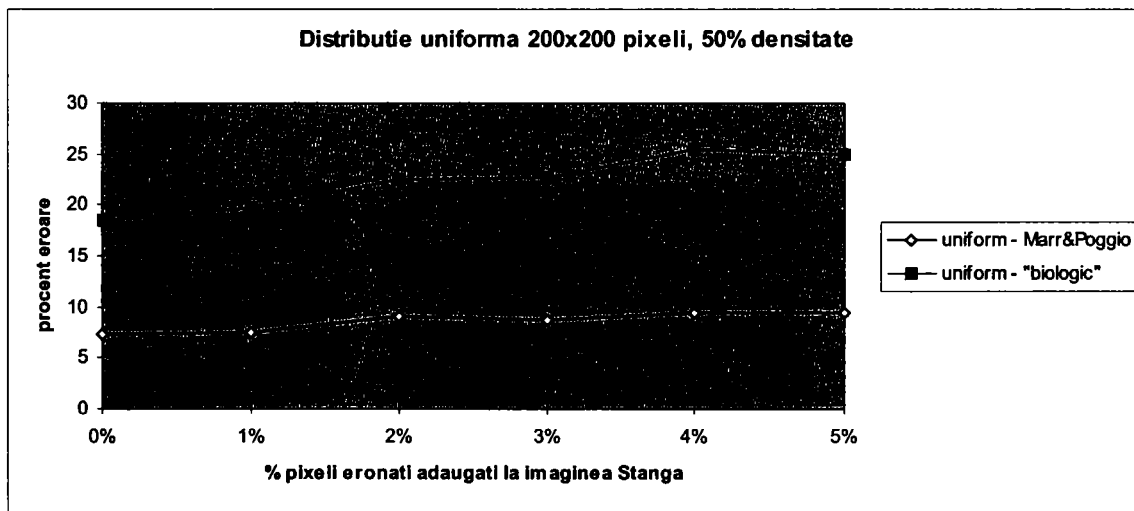


Figura 4.9: Comparatie a performantelor celor doi algoritmi implementati, utilizand RDS-uri B&W cu distributie uniforma si densitate de 50%, atunci cand a fost adaugat zgomot pe una dintre imagini. Se utilizeaza ca si criteriu procentul de pixeli eronati.

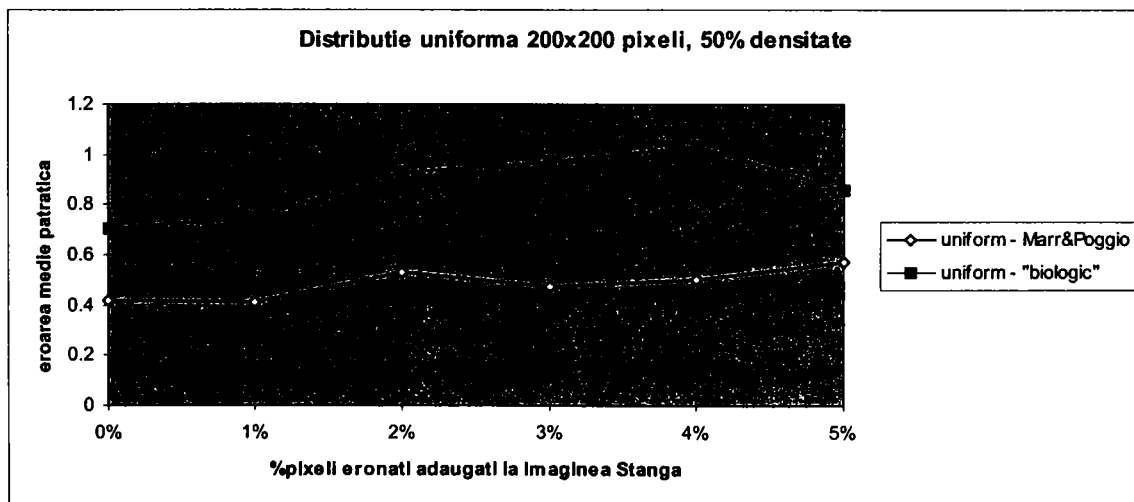


Figura 4.10: Similar cu figura 4.9, dar se utilizeaza ca si criteriu eroarea medie patratice.

Figura 4.11 prezinta harti de disparitate obtinute cu cele doua modele pe RDS-uri cu zgomot adaugat. Se observa ca algoritmul Marr&Poggio pare mai mult afectat de zgomot decat algoritmul "biologic", in mod special in privinta pastrarii muchiilor, care ajung complet alterate.

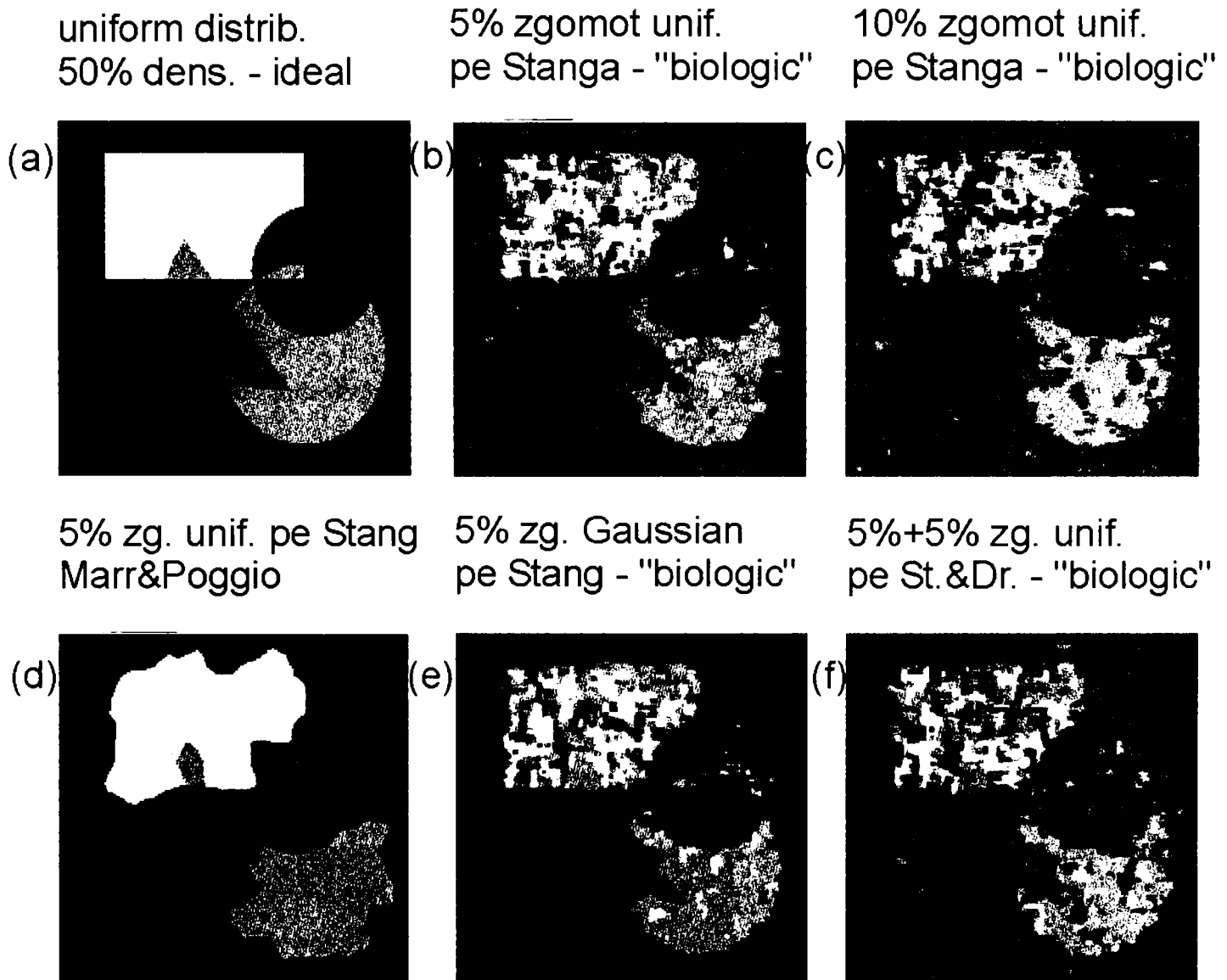


Figura 4.11: Harti de disparitate obtinute pe RDS-uri B&W de 200×200 puncte, uniform distribuite (densitate 50%) la care a fost adaugat zgomot: (a) este harta "ideala" ce constituie termenul de comparatie. Rezultatele au fost obtinute pentru $u=1/8$ si $v=0$ (filtrare Gabor uni-dimensionala).

De asemenea, un nivel de zgomot de 10% (fie doar pe una din imagini, fie cate 5% pe fiecare, zgomot independent) conduce la deteriorarea muchiilor si in cazul algoritmului "biologic".

REZULTATE OBTINUTE PE IMAGINI "REALE"

Asa cum deja am mentionat, imaginile pe care sistemul nostru vizual le prelucreaza curent nu sunt de genul RDS-urilor. Era deci firesc sa testez comportamentul celor doua modele si in cazul unor imagini din lumea reala.

Figura 4.12 prezinta hartile de disparitate obtinute pe doua astfel de imagini. S-au folosit filtre Gabor uni-dimensionale ca si model de camp receptor pentru celulele simple. Se observa ca rezultatele par destul de bune.

Figura 4.13 prezinta hartile de disparitate obtinute pentru o imagine la care am gasit o harta de disparitate de referinta (doar din punct de vedere pictografic). Harta de disparitate a fost obtinuta printr-o metoda *feature-based*, utilizand o retea Hopfield (Nasrabadi&Choo 1992). In aceeasi imagine am prezentat si rezultatul obtinut cu algoritmul Marr&Poggio si se observa ca, din punct de vedere practic, harta este neutilizabila.

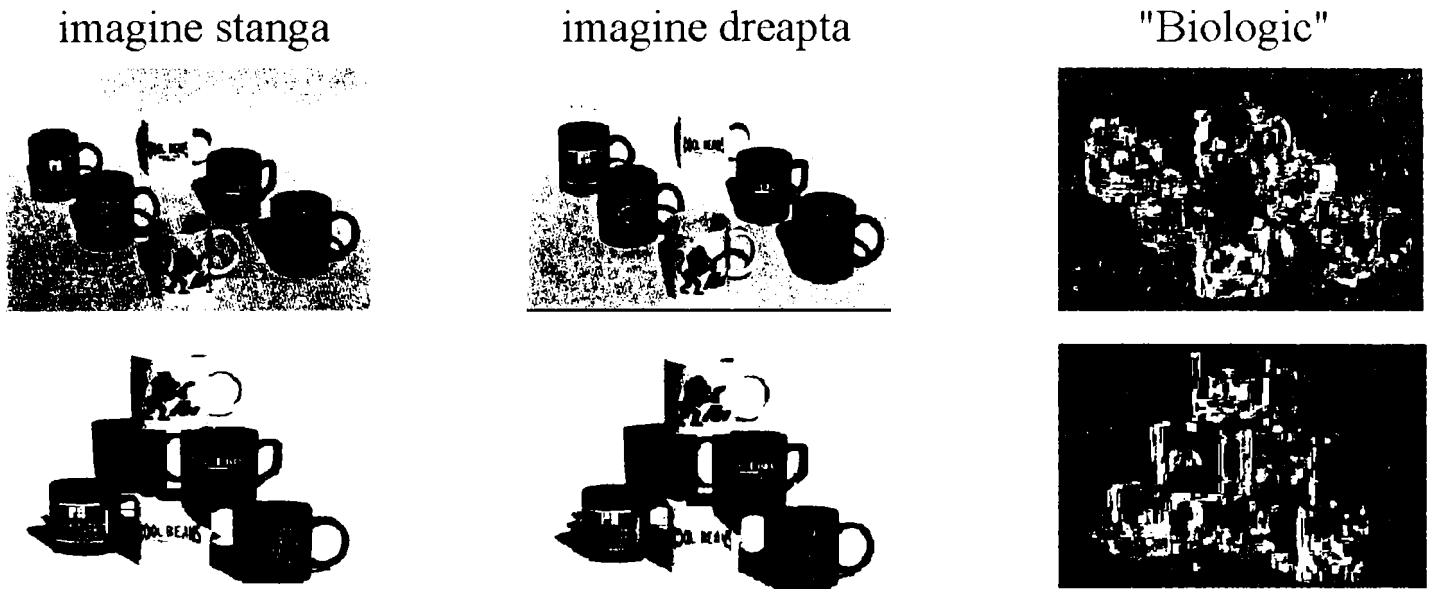


Figura 4.12: Rezultate obtinute pe imagini "reale" color cu modelul "biologic", utilizand filtrare Gabor uni-dimensionala ($u=1/8, v=0$).

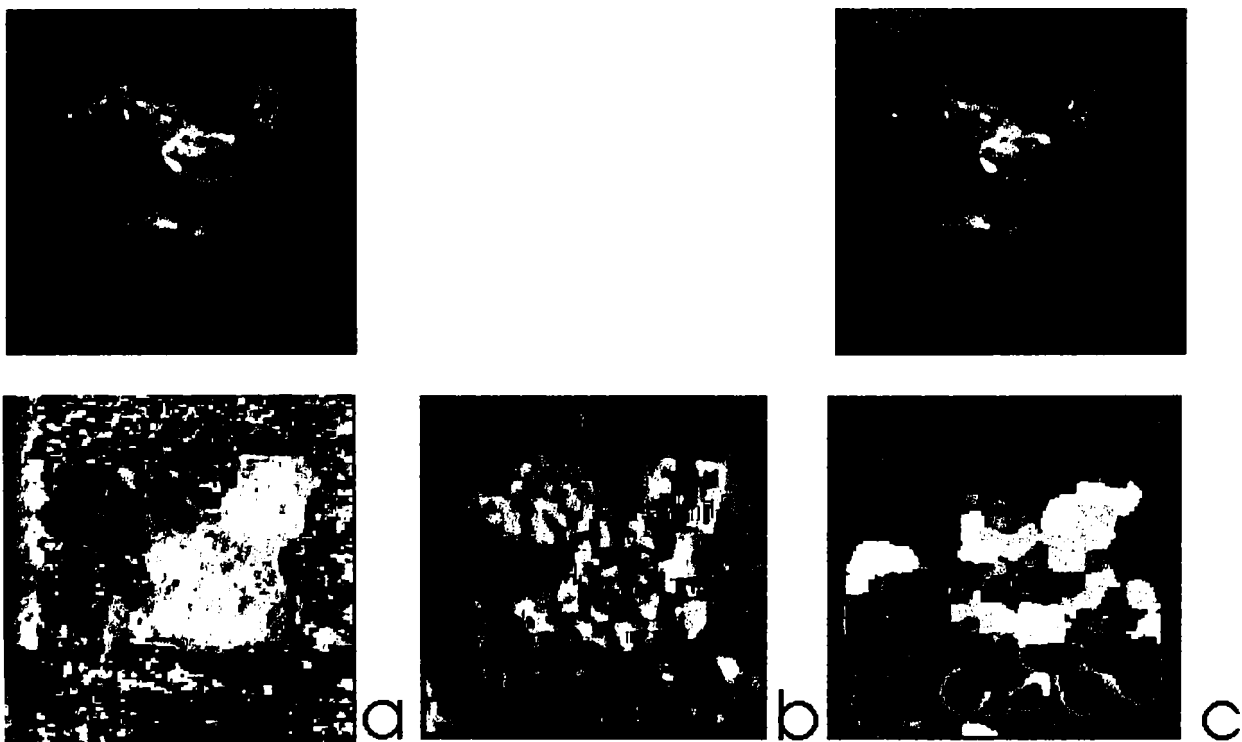


Figura 4.13: Rezultate comparative: (a) harta obtinuta de Nasrabadi&Choo (1992) cu o retea Hopfield; (b) harta obtinuta cu modelul "biologic" aplicand filtrare Gabor uni-dimensionala; (c) harta obtinuta prin metoda Marr&Poggio.

Rezultatele obtinute vin sa confirme faptul ca modelul "biologic" se comporta bine pe imagini asemanatoare scenelor reale prelucrate si faptul ca constrangerile foarte dure (mai ales cea de coeziune) impuse de modelul lui Marr&Poggio conduc la imposibilitatea utilizarii lui in afara RDS-urilor.

Am incercat sa aplic zgomot si pe imaginile "reale" si in figura 4.14 arat cateva harti de disparitate obtinute in aceste conditii. Asa cum era de asteptat, modelul Marr&Poggio nu poate "rezolva" astfel de stereograme, chiar in conditii in care nu este adaugat zgomot. De asemenea, harta de disparitate este deosebit de "zgomotoasa" in cazul utilizarii filtrelor Gabor uni-dimensionale la un nivel de zgomot de 10% (asa cum au anticipat, de altfel, rezultatele pe RDS-uri cu zgomot). In schimb, filtrele Gabor bi-dimensionale reusesc sa obtina harta de disparitate corecta.

imagine Stanga

imagine Dreapta



"biologic" $u=1/8$ $v=0$

Marr&Poggio



imagine Stanga

imagine Dreapta

10%
zgomot unif.
pe Stanga



"biologic" $u=1/8$ $v=0$

"biologic" $u=1/10$ $v=1/12$

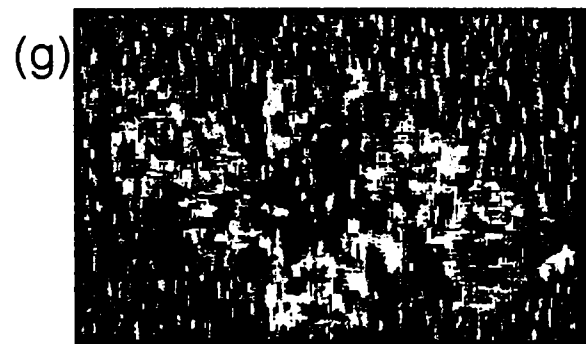


Figura 4.14: Rezultate obtinute pe imagini din "lumea reala", cu si fara zgomot adaugat. Se observa ca, in aceasta situatie, filtrele Gabor bi-dimensionale sunt singurele care reusesc sa "rezolve" problema gasirii disparitatilor orizontale din stereograma:(g) si (h) contin harti de disparitate obtinute pe stereograma cu 10% zgomot pe imaginea stanga.

In figura 4.15 sunt prezentate alte rezultate obtinute cu modelul "biologic" utilizand filtre Gabor uni-dimensionale si bi-dimensionale, toate cu aceeasi frecventa spatiala, dar cu orientari diferite.

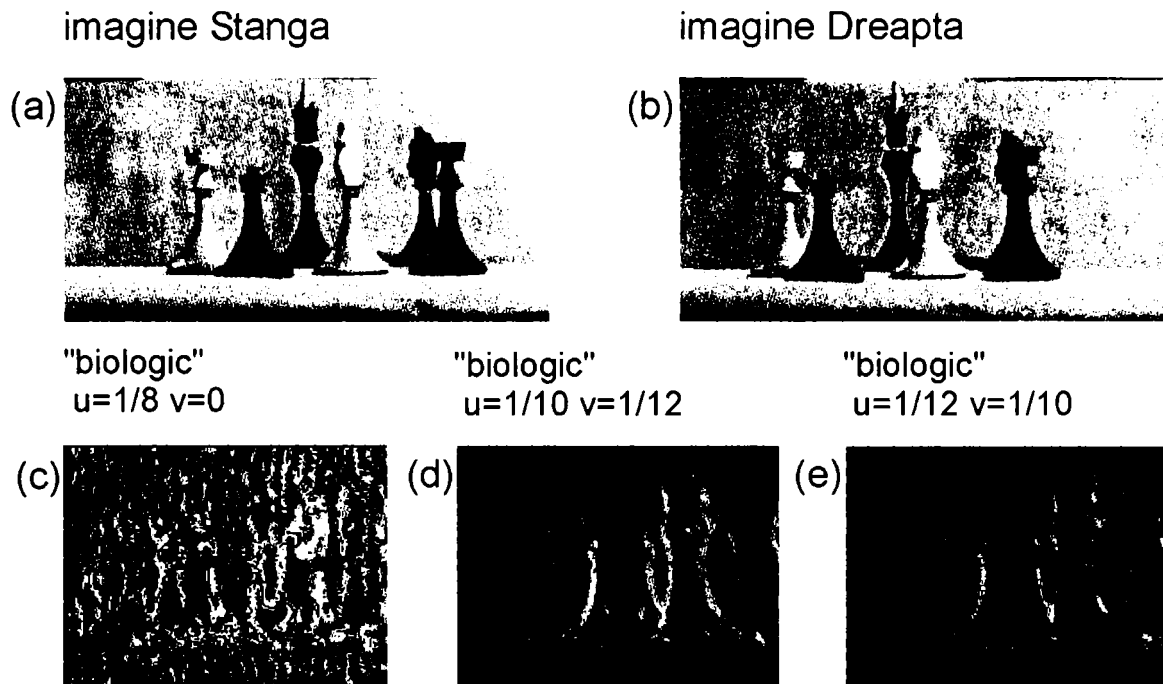


Figura 4.15: Rezultate obtinute pe imagini color cu obiecte subtiri si ocluzii partiale. Se observa ca doar utilizarea filtrelor Gabor bi-dimensionale rezolva astfel de stereograme

Am incercat efectul utilizarii filtrelor Gabor bi-dimensionale si asupra stereogramei lui Nasrabadi&Choo – in figura 4.16 se poate vedea rezultatul intr-o imagine "capturata" in timpul executiei programului cu care am efectuat testele si evaluarile. Trebuie sa mentionez ca cei doi autori (Nasrabadi&Choo 1992) "cauta" eventualele potriviri (intre caracteristicile identificate in imagini) atat pe orizontala si verticala , cat si pe cele doua diagonale si fac apoi o interpolare intre punctele disparate ale hartii.

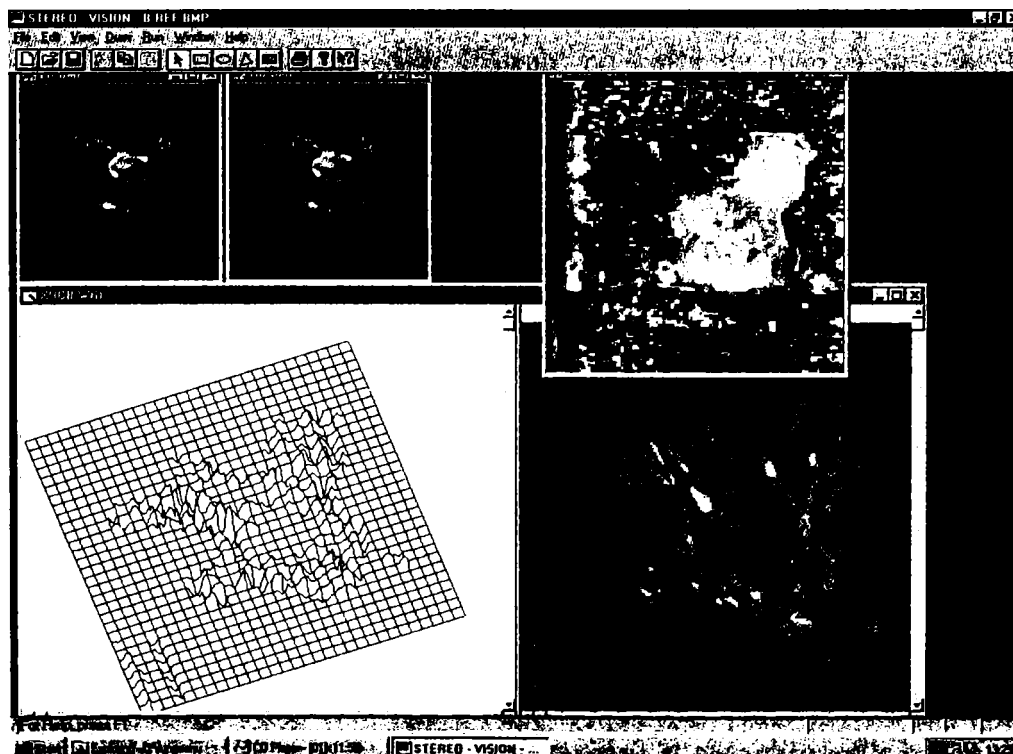


Figura 4.16: Rezultate obtinute cu modelul "biologic" bi-dimensional pe imaginea de test a lui Nasrabadi&Choo. In imagine se mai poate observa interfața programului cu care am facut testele si reprezentarea hartii de disparitate sub forma de *wireframe*.

4.5. CONCLUZII

Cercetarile arata ca in aria vizuala primara se gasesc celule complexe sensibile la disparitatea binoculara, sesizarea acesteia fiind unul din primii pasi catre perceptia profunzimii.

Pomind de la considerentele experimentale, am conceput si am implementat un model de determinare a hartii de disparitati binoculare orizontale ce utilizeaza proprietatile cunoscute ale celulelor simple si complexe din cortexul striat. Aceste premize initiale, precum si rezultatele obtinute (Lungeanu 1995, Lungeanu et al. 1996, Lungeanu et al. 1998b, Lungeanu et al. 1998c) confera plauzibilitate biologica modelului realizat.

Am utilizat filtre Gabor pentru simularea comportamentului celulelor simple din V1 si am determinat raspunsul unei celule complexe ca fiind insumarea (pe cale neuronală) a iesirilor patraticice a doua celule simple in cuadratura. Aceasta relatie celula simpla ↔ celula complexa permite ca raspunsul celulei complexe sa fie dependent doar de diferenta de faza dintre campurile celor doua celule simple si de disparitate. Mai mult, ea este in concordanta cu dovezile actuale privind celulele din cortexul striat, in sensul ca raspunsul celulei complexe este combinatia intrarilor de la mai multe celule simple.

Am aratat ca daca, pentru fiecare locatie spatiala, exista o familie de astfel de celule complexe acordate pentru disparitati diferite, ele vor fi in stare sa “detecteze” disparitatile binoculare si se vor constitui intr-o harta distribuita de disparitati. Acest lucru este, de asemenea in concordanta cu organizarea cunoscuta a cortexului striat, pe coloane specializate in procesarea anumitor caracteristici, coloane ce pastreaza o distributie retinotopica (Capitolul 3 din prezenta teza).

Am realizat o aplicatie ce a permis implementarea si testarea modelului, precum si compararea rezultatelor cu cele ale modelului clasic al lui David Marr si Tomaso Poggio (si acest model a fost implementat).

Am testat modelul pe RDS-uri cu diferite densitati, distributii, nivele de zgomot si tipuri de filtrare ulterioara determinarii disparitatilor (filtrare pe harta de disparitate). Aceste teste mi-au permis sa evaluez in mod obiectiv performantele modelului prin procentul de locatii cu eroare de disparitate si prin eroarea medie patratica a nivelului de disparitate determinat.

Desi conditia de epipolaritate este, in general, asumata in alte modele, modelul “biologic” nu este restrictionat de aceasta. In testele efectuate am aratat ca RDS-urile (imagini artificiale create pe baza deplasarii laterale a sub-seturilor de pixeli) se comporta mai bine la filtrarea Gabor uni-dimensională, realitate sugerata si de experientele psihofizice cu RDS-uri si explicata prin faptul ca aceste imagini sunt formate din plane fronto-paralele.

Am abordat o problema “ocolita” de alti cercetatori – situatiile in care geometria epipolara nu poate fi strict asumata si, deci, cautarea “potrivirilor” limitata la o singura dimensiune. Prelucrarea imaginilor “reale” a dovedit clar ca in aceasta situatie utilizarea filtrelor Gabor bi-dimensionale se impune de la sine. Mai mult, filtrele Gabor bi-dimensionale reduc extrem de performant zgomotul din stereograme.

Rezultatele s-au dovedit incurajatoare – ele au aratat clar ca modelul “biologic” pastreaza bine formele (colturi, muchii, etc.) si ca este robust la zgomote. Daca privim din punctul de vedere al evaluarii performantelor prin procentul de pixeli eronati si eroarea medie patratica a disparitatii, s-ar parea ca performantele algoritmului “biologic” sunt mai slabe decat ale modelului Marr&Poggio. Chiar filtrate, hartile de disparitate se mentin la un nivel de eroare mai ridicat. In momentul in care analizam harta de disparitate pictografica insa, apar evidente alte aspecte:

- algoritmul Marr&Poggio obtine o harta uniforma (datorita constrangerii de coeziune foarte puternica la acest model), dar asta conduce la deteriorarea grava a formelor, deci a continutului imaginilor;
- algoritmul “biologic” obtine o harta “zgomotoasa”, dar reda cu fidelitate formele.

Trebuie sa mentionez ca nu se stie cu exactitate ce se intampla la nivelul V1 si nu s-au putut determina profilele campurilor receptoare ale celulelor complexe (nici pentru contururi, nici pentru disparitati). Se stie doar ca exista aceste celule ce sunt sensibile la disparitate si ca mai exista astfel de celule in aria vizuala secundara (V2) si aria temporală mediala (MT). Acestea din urma au sensibilitate ridicata la directia si viteza de miscare, precum si la orientarea suprafetelor (Liu&Hulle 1998). Se fac supozitii asupra rolului lor si probabil ca perceptia adancimii se obtine combinand informatiile prelucrate de canalele paralele (inclusiv informatia ne-vizuala). S-ar putea ca, asa cum am sugerat deja, si harta de disparitate biologica sa fie “zgomotoasa”, ceea ce sa explice relativa intarziere la fuzionarea RDS-urilor (adica atunci cand sistemul vizual este obligat sa se limiteze la disparitatile orizontale). Combinarea informatiei de diverse tipuri pentru a putea obtine rezultate solide este reclamata si de sistemele tehnice (Overington 1992, Wörgötter et al. 1999). Sunt cercetatori care sugereaza chiar existenta unei organizari paralel-distribuite a perceptiei de adancime, sugerandu-se existenta mai multor subsisteme paralele cu specializare ingusta (Fine&Jacobs 1999, Tyler 1994).

Toate acestea dovedesc plauzibilitatea obtinerii unei harti de disparitate grosiere in prima faza, urmata de o rafinare a ei in etapele ulterioare. Mai mult, exista dovezi experimentale potrivit carora campurile receptoare constituie entitati spatio-temporale (desi pana acum cercetarile s-au concentrat doar asupra aspectelor spatiale) si ca au comportamentul dependent de context si de activitate, modificandu-se atat dimensiunea lor, cat si natura sinaptica (DeAngelis et al. 1995, Miller 1998, Ramachandran 1995).

In concluzie, cred ca rezultatele dovedesc clar ca in V1 se poate obtine o harta distribuita a disparitatilor orizontale utilizand doar proprietatile cunoscute ale celulelor simple si complexe. Chiar daca aceasta nu este suficient pentru reconstituirea informatiei 3-D din imaginile analizate, disparitatea constituie un indiciu puternic pentru sistemul vizual uman, asa cum experimentele cu stereograme cu puncte aleatoare o dovedesc.

Capitolul 5: DE LA IMAGINI LA SUPRAFETE. MODELAREA PERCEPTIEI CONTURURILOR

5.1. COLOANE DE ORIENTARE IN ARIA VIZUALA PRIMARA. MODELAREA PERCEPTIEI CONTURURILOR.

Deoarece modelarea percepției contururilor în aria vizuală primară abordează aspecte de arhitectura corticală și de neurofiziologie deja prezentate în Capitolele 2 și 3, voi trece doar în revistă, pe scurt, principalele caracteristici ale cortexului striat din punctul de vedere al întăririi contururilor. De asemenea, în paralel, voi aminti și principalele strategii de modelare neuronală a acestor mecanisme: în principiu, ele se bazează pe rețele neuronale artificiale, mai mult sau mai puțin fidele realității, ca arhitectura și dinamică.

CONTURURI IN ARIA VIZUALA PRIMARA

În stadiile primare ale prelucrării vizuale, neuronii răspund la stimuli ce apar în câmpul lor receptor și corespund profilului acestora (celulele ganglionare, celulele din LGN și celulele corticale simple), devenind tot mai selective pe măsură ce se avansează în lanțul de procesare vizuală. Sistemul vizual face detectia elementelor locale și le grupează în caracteristici globale din punctul de vedere al informației din câmpul vizual – realizează așa-numita **schita primară** (Marr 1982).

Evidențele experimentale arată că la nivelul V1 există o selectivitate pronunțată la orientare în cadrul hipercoloanelor de orientare și există tendința ca, drept răspuns la stimularea vizuală cu linii, bare, contururi mai lungi (eventual închise), să apară o sincronizare a răspunsului unor grupuri mari de neuroni, ce oscilează la scara de msec (Engel et al. 1992, Li 1998, Singer 1995, Vershure&Konig 1999).

Una din principalele probleme în integrarea contururilor este faptul că muchiile elementare sunt sesizate de celulele corticale simple și imaginea “abundă” în prima fază în astfel de muchii, multe dintre ele fiind, de fapt, zgomot. De aceea, sunt interesante mecanismele prin care, la nivelul rețelei/rețelelor de hipercoloane din cortexul striat, se ajunge la o întărire selectivă a acelor muchii care aparțin contururilor (Li&Atick 1994, Marr 1982, Miclea 1999, Vidyasagar et al. 1996, Wilson&Keil 1999). Aceasta întărire pare să fie influențată și de combinarea cu alte caracteristici detectate la nivelul V1, precum și de mecanisme pre-attentionale (Ahissar&Hochstein 1995, Kowler 1995).

MODELE DE PERCEPTIE A CONTURURILOR

Abordări care au condus la rezultate similare celor observate de neurofiziologi au utilizat arhitecturi de rețele neuronale ce, asemenea cortexului striat, au încercat să pastreze organizarea pe coloane și/sau comportamentul cunoscut al celulelor simple și complexe (Hirakura et al. 1996, Marr 1982, Miller&Zucker 1999, Okajima 1996). În general, această organizare în coloane se datorează inter-neuronilor existenți în V1 și a legăturilor de *feed-back* local, așa cum o dovedesc atât evidențele biologice (Sagawa 1997, Vidyasagar et al. 1996), cât și necesitățile computaționale (Blais 1998, Kitajima&Hara 1997, Li 1998).

În fapt, modelarea întăririi contururilor în V1 este o problemă dificilă dacă se dorește respectarea comportamentului celulelor dovedit a exista la acest nivel. Foarte multe modele, pentru a reuși să integreze contururile, au introdus *feed-back* de la nivele mai ridicate (una dintre problemele dificile a fost realizarea stabilității rețelei).

În mod firesc, se pune întrebarea dacă poate fi realizată o întărire de contururi la nivelul hipercoloanelor de orientare din V1 (așa cum sugerează datele experimentale) doar pe baza

conexiunilor locale (fără legături *top-down* de la arii vizuale mai înalte) și doar utilizând proprietățile celulare cunoscute în aria vizuală primară.

Trebuie menționat că la nivelul V1 nu există încă celule “responsabile” de joncțiunile de tip T, L sau X (Li 1998), care pot să contribuie la o integrare mai fidelă a contururilor, eventual a contururilor iluzorii – ele apar abia în V2. De aceea, nu ne putem aștepta că, la acest nivel, să se producă o detecție de contururi, ci mai degrabă o întărire a muchiilor care aparțin unor contururi.

5.2. MODEL PENTRU ÎNTĂRIREA CONTURURILOR PORNIND DE LA MUCHII ELEMENTARE

DESCRIEREA MODELULUI

Modelul de întărire de contururi la nivelul V1, pe care l-am gândit și l-am implementat, constă dintr-o **rețea recurentă** bazată pe o grilă hexagonală de **hipercoloane de orientare** (pastrand, astfel, distribuția retinotopică a hipercoloanelor) – figura 5.1.

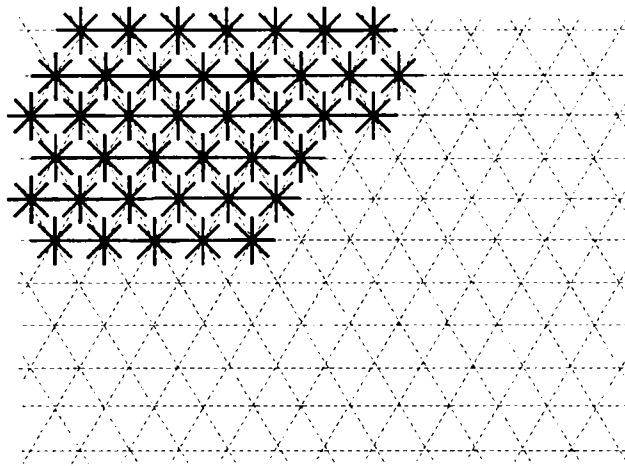


Figura 5.1: Modelul utilizează hipercoloane de orientare aranjate pe o grilă hexagonală

Am ales această organizare pentru o cât mai bună “acoperire” a spațiului vizual și o bună descriere a caracteristicilor primare ale scenei, fără ca aceste cerințe să ducă la o complicare suplimentară a calculelor (Li 1998).

Intrările în rețea ajung la poziții spațiale discrete – în fiecare poziție se găsește câte o hipercoloană compusă din **muchii elementare orientate**. Se poate observa că în felul acesta se păstrează atât distribuția spațială, cât și orientarea muchiilor.

Procesarea parcurge două **etape** (figura 5.2), descrise în continuare.

Etapa 1

Pornind de la imaginea inițială (scena analizată) se face detecția muchiilor elementare și trecerea de la grila ortogonală la cea hexagonală – în implementarea realizată, detecția muchiilor am realizat-o cu filtrul gradient (Marr 1982, Gonzalez & Woods 1993) și am obținut modulul (considerat taria muchiei) și orientarea muchiei.

Am ales această soluție pentru simplitatea ei, considerând că filtrul gradient “mimează” satisfactor comportamentul celulelor receptoare de la nivelul V1. Alegerea unei alte soluții (filtre Gabor, de exemplu) ar fi însemnat o complicare nejustificată a calculelor în această etapă.

Suplimentar, la schimbarea grilei, am făcut și o reducere dimensională.

Etapa 2

Modelul propriu-zis face procesarea în cadrul hipercoloanelor de orientare. Există conexiuni în cadrul fiecărei coloane și conexiuni între coloane. Iesirea rețelei merge spre ariile vizuale mai

inalte. Nu am prezentat în figura toate conexiunile (laterale și de *feed-back*) pentru ca ar fi încărcat desenul, dar ele vor rezulta din dinamica rețelei.

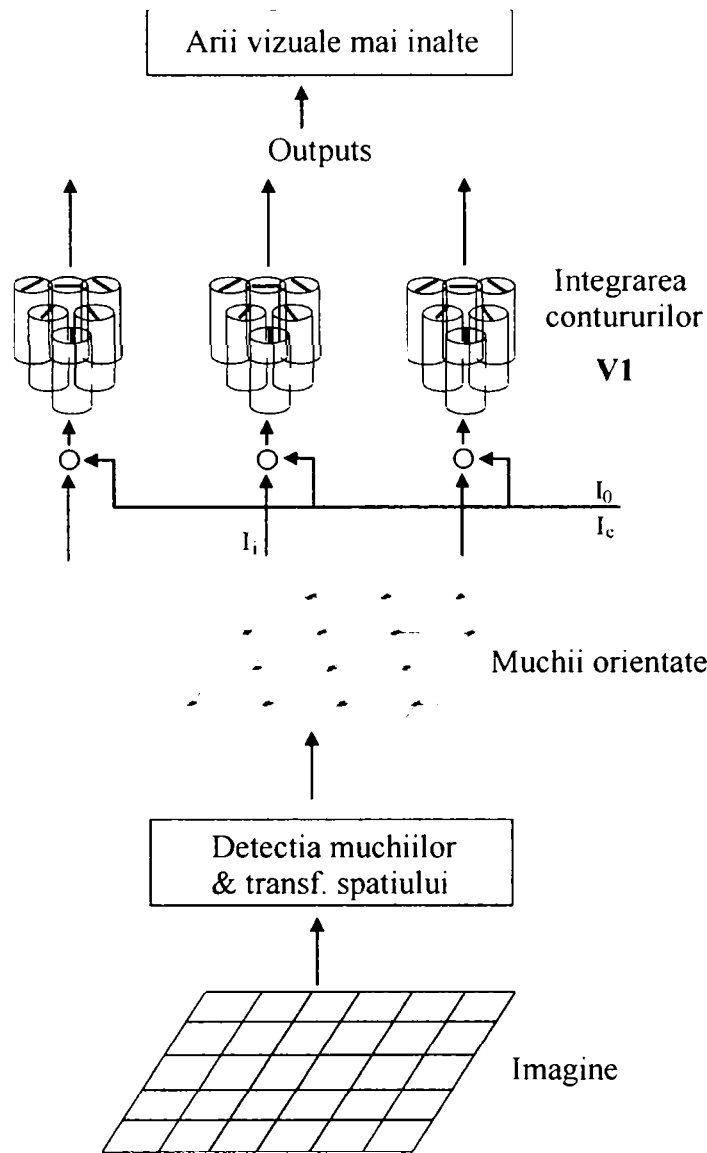


Figura 5.2: Etapele de procesare a informației pentru integrarea contururilor în V1

Fiecare hipercoloană constă din elemente de tip muchie, în fapt oscilatori neuronali formați din perechi de neuroni excitatori-inhibitori (figura 5.3). Așa cum am mai spus, pentru fiecare locație spațială există o familie de astfel de perechi, care formează hipercoloane de orientare.

Există conexiuni în cadrul fiecărei hipercoloane și între vecinii imediați, dar și conexiuni pe distanțe mai lungi, similare celor din cortexul striat (Capitolul 3).

Dinamica rețelei este descrisă de următoarele ecuații:

$$\begin{aligned} \dot{x}_{i\theta} &= -\alpha_x x_{i\theta} + I_{i\theta} - \sum_{\Delta\theta} V(\Delta\theta) g_y(y_{i,\theta+\Delta\theta}) + J_0 g_x(x_{i\theta}) + \sum_{j \neq i, \theta'} J_{i\theta, j\theta'} g_x(x_{j\theta'}) + I_0 \\ \dot{y}_{i\theta} &= -\alpha_y y_{i\theta} + g_x(x_{i\theta}) + \sum_{j \neq i, \theta'} W_{i\theta, j\theta'} g_x(x_{j\theta'}) + I_c \end{aligned} \quad (5.1)$$

unde:

$x_{i\theta}$ și $y_{i\theta}$ reprezintă potențialul de membrană pentru neuronul excitator, respectiv inhibitor din fiecare pereche ce modelează muchia de orientare θ , aflată la locația i ;

g_x și g_y sunt funcțiile de transfer ce generează potențialul de acțiune (liniare pe porțiuni, cu prag și saturatie);

$I_{i\theta}$ este intrarea pentru muchia respectivă (obținută pe baza funcției gradient);

$I_{i\theta}$ este intrarea pentru muchia respectiva (obtinuta pe baza functiei gradient);
 $1/\alpha_x$ si $1/\alpha_y$ sunt constantele de membrana;
 $V(\Delta\theta)$ modeleaza inhibitia din cadrul aceleiasi hipercoloane si descreste cu $\Delta\theta$;
 J_0 modeleaza conexiunea de auto-excitatie;
 $J_{i\theta}$ modeleaza influenta excitatorie a vecinilor cu orientare similara, aflati pe un (eventual) contur comun;
 I_0 si $I_{c\ back}$ modeleaza activitatea de fond a retelei;
 $W_{i\theta, j\theta}$ modeleaza influenta inhibitorie a vecinilor cu orientare similara, dar aflati pe contururi alternative, in competitie cu muchia ($i\theta$);
 $I_{c\ control}$ permite controlul dinspre ariile vizuale mai inalte (intrare de *feed-back*, care poate fi atat inhibitorie, cat si excitatorie). In principiu, aceasta influenta s-ar putea exercita asupra ambilor neuroni din cadrul unei muchii – am ales aceasta solutie propusa si in alte modele (Li 1998) pentru a separa intrarea propriu-zisa in retea de reactia de *feed-back*, solutie aleasa si de natura (in stratul IV C din cortexul striat nu vin reactii de *feed-back*). Mai trebuie sa mentionez si faptul ca in acest model nu am abordat, deocamdata, aspecte legate de determinarea acestei reactii.

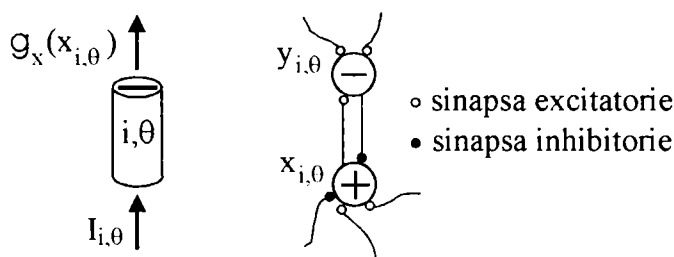


Figura 5.3: Muchiile elementare sunt oscilatori neuronali, caracterizati de $i\theta$

Modelul a fost implementat si testat, iar aplicatia realizata in acest scop este descrisa in Anexa 5.

ALEGEREA PARAMETRILOR RETELEI

Dimensiunea grilei hexagonale

Am considerat o esantionare uniforma a spatiului, iar dimensiuna a fost redusa pe ambele directii cu un factor de $1/2$. In ecuatie ce descrie dinamica retelei am utilizat un singur indice (i) pentru a descrie localizarea din motive de claritate. In fapt, am utilizat doi indici ce variaza cu un pas de 2 (practic, doua grile intretesute).

Functiile de transfer pentru neuronii ce formeaza o muchie elementara

Functiile de transfer pentru generarea potentialului de actiune la iesirea neuronului sunt mult simplificate:

$$g_x(x) = \begin{cases} 0 & x < T_x \\ (x - T_x) & \text{daca } T_x \leq x \leq T_x + 1 \\ 1 & x > T_x + 1 \end{cases} \quad (5.2)$$

$$g_y(y) = \begin{cases} 0 & y > 0 \\ g_1 y & \text{daca } 0 \leq y \leq L_y \\ g_1 L_y + g_2 (y - L_y) & 0 < L_y \leq y \end{cases}$$

Aplicatia permite modificarea valorilor T_x , L_y , g_1 si g_2 in mod interactiv.

Numarul de muchii elementare din cadrul unei hipercoloane poate, de asemenea, fi stabilit in functie de dorinta (cu consecintele corespunzatoare asupra timpului de prelucrare). In simularile facute, am lucrat cu 2, 4 si 8 muchii. Muchiile sunt neorientate si am considerat acelasi numar de muchii pentru toate hipercoloanele corespunzatoare unei imagini prelucrate.

Unghiul θ variaza intre $[0, \pi)$. Daca numarul de muchii din cadrul unei hipercoloane va fi N , atunci muchiile vor avea orientarile $\theta = n * \pi / N$, cu $n = 0, 1, \dots, N-1$

Inhibitia in cadrul aceleiasi hipercoloane

Funcția care modeleaza inhibitia este distribuita pe patru trepte:

$$V(\Delta\theta) = \begin{cases} Maxim & \text{pt. } \Delta\theta = 0 \\ Mediu & |\Delta\theta| = \pi/12 \\ Mic & |\Delta\theta| = \pi/6 \\ Minim & \text{in rest} \end{cases} \quad (5.3)$$

Aceasta inhibitie are rolul de a "cupla" neuronii inhibitori cu corespondentii lor excitatori, sau cu cei de orientare "similari". Cele patru valori pot fi modificate la simulare.

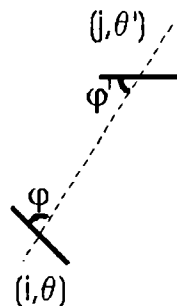
Interactiunile intre hipercoloane

Scopul conexiunilor inter-coloana, $J_{i\theta}$ si $W_{i\theta, j\theta'}$, este acela de a intari muchiile din cadrul unor posibile contururi netede, regulate, fara ondulatii frecvente.

Am putea rezuma aceste cerinte in felul urmatoar:

- doua muchii vecine ar trebui sa se intareasca reciproc daca se poate trasa un contur neted ce trece prin amandoua;
- daca, in cadrul unui contur, exista o portiune lipsa de dimensiuni mici, ea ar trebui sa fie compensata (retea sa reconstituie conturul ca fiind continuu);
- retea nu are voie sa mareasca sau/si sa ingroase contururile de lungime si grosime finita existente.

Un alt aspect avut in vedere a fost acela al unei posibile (sau probabile) alinieri a muchiilor (i, θ) si (j, θ') la un contur ce trece prin i si j



Am considerat (ca si conditie de conexiune excitatorie/inhibitorie) diferenta dintre unghiurile φ si φ' si, in acelasi timp, diferenta dintre θ si θ' (orientarile celor doua muchii aflate in coloane diferite).

Similaritatea de orientare am considerat-o pentru $|\Delta\theta| = |\theta - \theta'| < \pi/4$.

De asemenea, am considerat distanta maxima de influenta de 10 pasi pe grila hexagonala, adica $d = \max(|u|, |v|) \leq 10$. Figura 5.4 arata cum sunt distribuite aceste influente (cu alura Gaussiana) considerate pentru o muchie cu $\theta = 0$ (practic, o muchie orizontala). Se observa ca am favorizat contururile line (nu schimbari bruste de directie), dar influentele inhibitorii le-am considerat mai mici ca si pondere decat pe cele excitatorii.

Activitatea de fond a retelei

Activitatea de fond are doua componente: I_0 care apare ca intrare pentru neuronul excitator si $I_{c \text{ back}}$ prevazut ca intrare pentru neuronul inhibitor.

Intrarea I_0 este compusa dintr-o influenta excitatorie $I_{e \text{ back}}$ similara celei inhibitorii $I_{c \text{ back}}$ date de sistemul reticulat ascendent activator (care influenteaza starea generala de activare a organismului) si o componenta ce normalizeaza activarea retelei-model pentru o vecinatate $S \leq 2$ calculata pentru toate coloanele din grila.

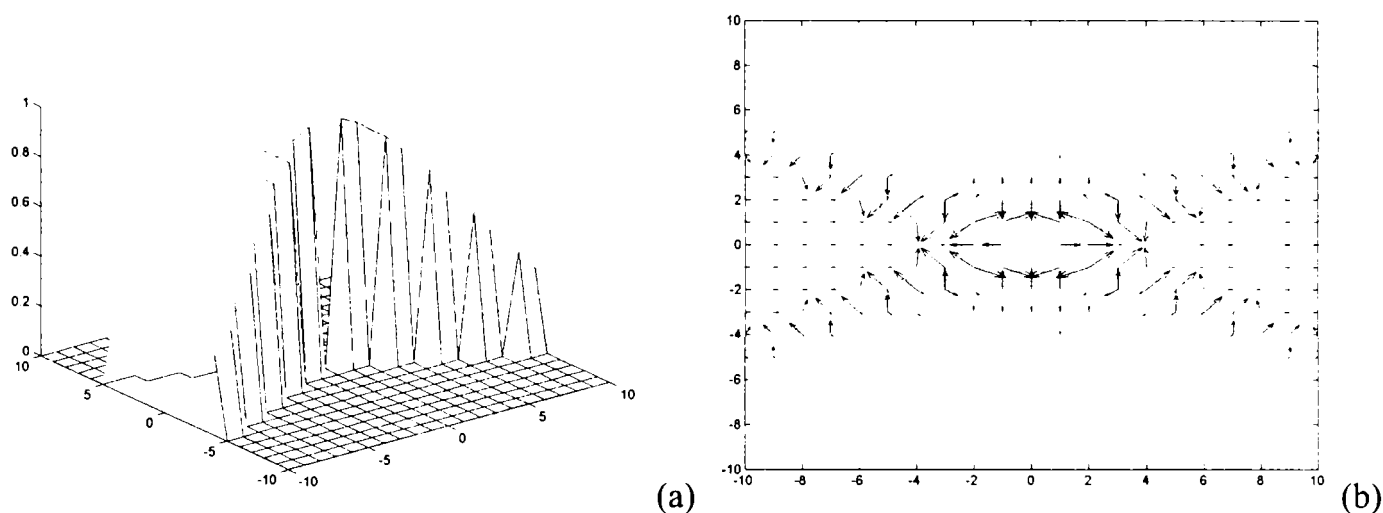


Figura 5.4: Distributia influentelor excitatorii si inhibitorii intre coloane pentru $d \leq 10$: (a) influenta excitatorie aratata ca "meshgrid"; (b) influentele excitatorie si inhibitorie prezentate cu gradientul pe vecinatatea considerata.

Asta inseamna ca o activitate de orice natura in apropierea unei muchii o va influenta in sensul scaderii potentialului de membrana:

$$I_0 = I_{eback} + I_{normaliz}$$

$$I_{normaliz}(i\theta) = -2.0 \left[\frac{\sum_{j \in S} \sum_{\theta'} (x_{j\theta'})}{\sum_{j \in S} 1} \right]^2 \quad (5.4)$$

$I_{normaliz}$ modeleaza legaturile inhibitorii interneuronale pe distanta scurta din cortexul striat.

Mentionez ca aplicatia dezvoltata pentru implementarea modelului permite setarea parametrilor retelei, deci testarea comportamentului acesteia pentru diferite combinatii.

5.3. REZULTATE. DISCUTII

Am testat modelul de intarire a contururilor cu diverse combinatii ale parametrilor si pe diferite imagini, toate fisiere *.BMP cu 256 culori (nuante de gri). Am constatat o mare dependenta a performantelor de continutul imaginilor si nu am gasit o solutie general valabila.

Rezultatele pe care le prezint in continuare sunt obtinute in urma medierii iesirilor retelei pe $2 \div 12$ iteratii (practic pe distanta mai multor oscilatii).

IMAGINE-TEST CU ZGOMOT SI FARA ZGOMOT

Figura 5.5 prezinta o imagine simpla, de test ($200 \times 100 \times 256$ BMP), care dezvaluie foarte bine comportamentul retelei. Subliniez ca nu este o imagine binara, am creat-o in alb-negru pentru teste (toate imaginile sunt cu 256 de culori) si are o rezolutie mica, ceea ce permite sa se vada clar modul in care conturul este pastrat sau nu. In prelucrare am utilizat 4 muchii/hipercoloana.

Zgomotul din imaginea (e) este de nivel nu foarte ridicat (5%) si uniform distribuit spatial. Are insa o caracteristica care-l face foarte defavorabil: pixelilor afectati de zgomot li s-a inversat culoarea (alb \leftrightarrow negru).

Rezultatele obtinute dovedesc ca, in conditii de imagine cu contururi continui, fara nici un fel de zgomot, modelul poate sa pastreze liniile, fara a avea tendinta sa le ingroase sau sa le deformeze (c), ba chiar intarind muchiile. Daca sistemul nu este "reglat" corespunzator, pot apare efecte de deformare, asa cum se vede in (d). Subliniez ca, datorita modului de distribuire a hipercoloanelor pe imagine si a muchiiilor elementare in hipercoloana, ele pot apare ca discontinui in reprezentarea picturala (cu exceptia muchiiilor orizontale) la o privire rapida.

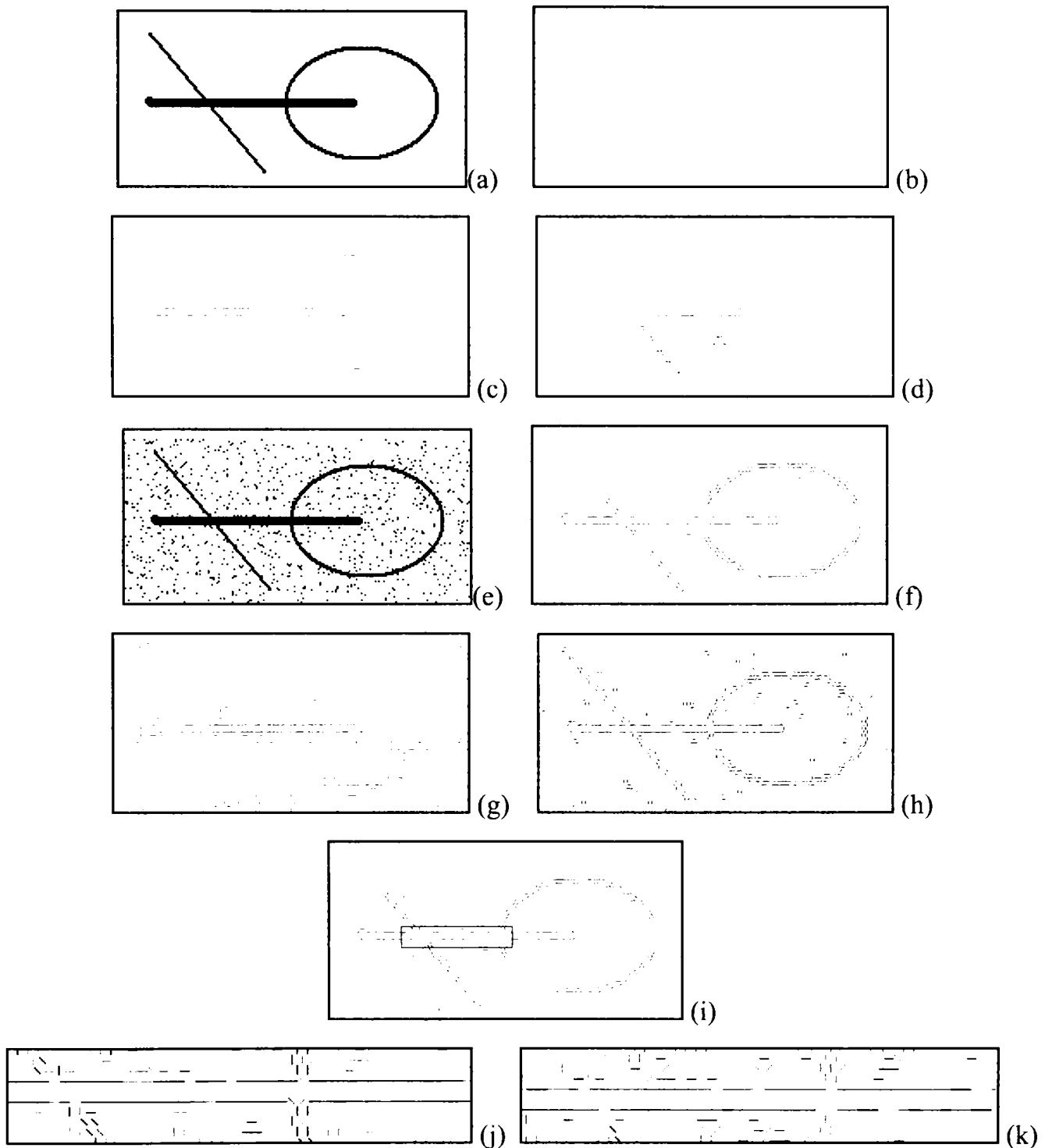


Figura 5.5: Rezultat obtinut pe o imagine foarte simpla (200×100×256 BMP), utilizand 4 muchii/hipercoloana: (a) imaginea initiala; (b) intrarile in hipercoloane stabilite pe baza gradientului; (c) iesirea dupa 14 iteratii; (d) aceeași imagine a – iesirea dupa 14 iteratii, dar cu alti parametri; (e) imaginea a – dupa ce s-a adaugat zgomot 5% - detalii in text; (f) intrarile in hipercoloane pentru imaginea e; (g) imaginea e – iesirea dupa 6 iteratii; (h) aceeași imagine e – iesirea dupa 14 iteratii, dar cu alti parametri; (i) imaginea f – pe care s-a pus in evidenta zona ce a fost marita in imaginile j si k; (j) detaliu din i – inainte de prelucrare; (k) aceeași zona ca in imaginea j – dupa 2 iterari cu aceeasi parametrii folositi la imaginea h (detaliu din h, dar in faza intermediara).

In situatia unei imagini cu zgomot, lucrurile se prezinta analog: functie de parametri, muchiile pot fi intarite si se pot obtine contururi destul de bine evidentiata (h), sau contururile se pot “pierde” in zgomotul de fond care este intarit (g). In imaginile (j) si (k) am evidentiat faptul ca muchiile din contur au tendinta de a ramane cu nivel mai ridicat decat cele provenite din zgomot. In (k) ele par sa piarda din intensitate, dar trebuie subliniat ca este abia a doua iteratie si ele se intaresc apoi, asa cum dovedeste imaginea din (h) obtinuta pentru 14 iteratii – in fapt (k) este un detaliu din (h) in timpul prelucrării.

Ca o concluzie, putem spune că, printr-o alegere judicioasă a parametrilor, zgomotul poate fi înlăturat. Se ridică însă problema alegerii acestor parametri, astfel încât ei să fie cei “potriviti” pentru o imagine dată. Cred că în rezolvarea unor astfel de aspecte ar putea să conștientizăm rolul conexiunilor de *feed-back* din zonele mai înalte – să “ajute” la alegerea soluției optime din mai multe alternative. S-ar putea să nici nu fie nevoie de influențe *top-down*, ci această “reglare de parametri” să rezulte ca urmare a interacțiunilor cu canalele paralele de procesare a altor informații (mişcare, disparitate, textură, culoare) chiar la nivelul ariei vizuale primare. Detaliile (i) și (k) arată destul de clar faptul că există tendința ca muchiile ce sunt “false candidate” la contururi să fie înlăturate, combinarea cu alte informații chiar în acest stadiu primar de prelucrare ar putea fi benefică.

PRELUCRAREA UNEI IMAGINI MONDRIAN

În figura 5.6 sunt prezentate rezultate obținute pe o imagine de tip Mondrian (137×166×256 BMP). Se observă că, așa cum era de așteptat, modelul “descoperă” contururi diferite, funcție de setarea parametrilor. În imaginea (b) sunt tratate ca și contururi și nuanțele de gri de pe margini, în timp ce în imaginea (c) acestea au fost eliminate. În imaginea (d) am suprapus contururile detectate peste imaginea inițială și se observă că ele au fost integrate cu destul de bună fidelitate. Trebuie să menționez că există o ușoară deplasare spre dreapta a contururilor datorită aplicării filtrului gradient pentru detectia muchiilor elementare din imagine, dar se observă că informația intrinsecă din imaginea inițială este păstrată.

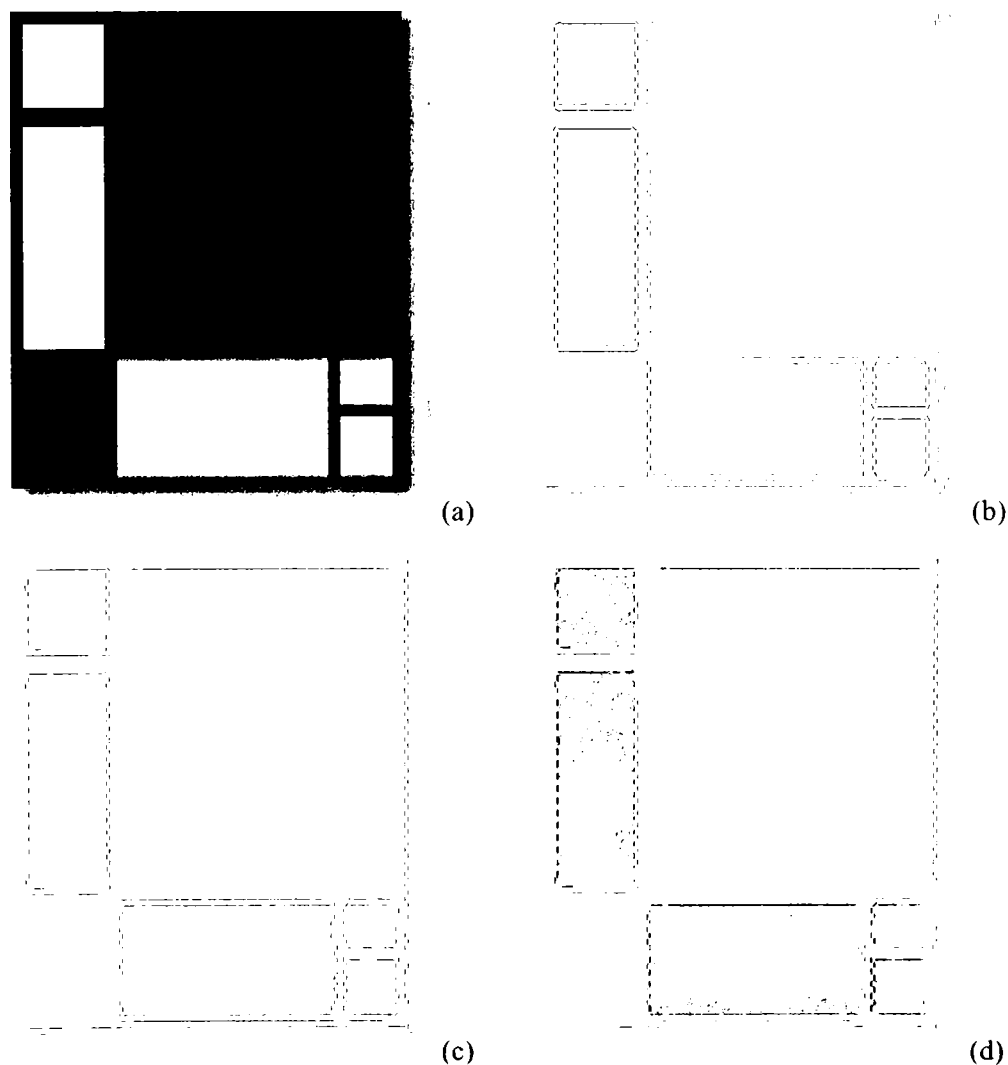


Figura 5.6: Rezultate obținute pe o imagine de tip Mondrian (cu 4 muchii/hipercoloana): (a) imaginea inițială; (b, c) rezultate obținute pentru diversi parametri; (d) imaginea c a fost suprapusă peste imaginea inițială a (la care culorile au fost deplasate spre plaja superioară pentru a putea distinge rezultatul).

ALTE REZULTATE

În figura 5.7 este prezentată o situație în care rezultatele obținute au fost destul de descurajante (imagine $121 \times 121 \times 256$ BMP). Contururile rezultate permit cu greutate recunoașterea conținutului inițial al imaginii.

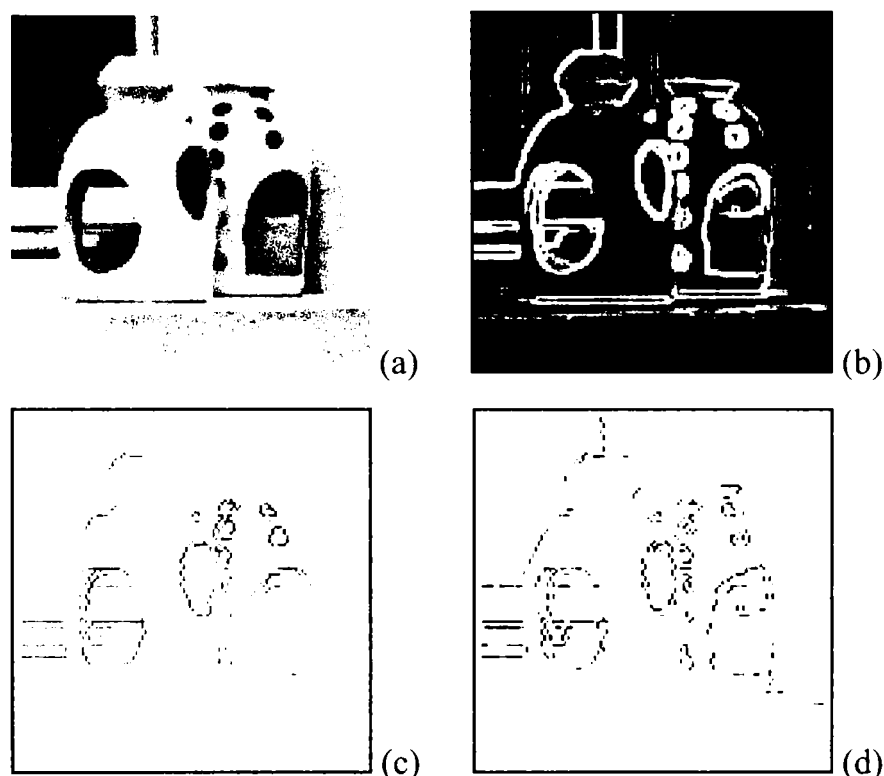


Figura 5.7: Rezultate în care contururile nu permit nici măcar recunoașterea (umana) a obiectelor (cu 8 muchii/hiperculoana): (a) imaginea inițială; (b) rezultatul aplicării filtrului gradient - modulul (orientarea nu apare aici); (c) intrarea în hipercoloanele de orientare; (d) ieșirea după 2 iterații.

Cred că eșecul poate fi pus pe seama contrastului scăzut din imagine pe cea mai mare parte a ei și ridicat doar în anumite zone – muchiile rezultate în urma aplicării gradientului au fost “tari” în acele zone de contrast ridicat, iar în rest au fost așa de slabe încât au fost tratate ca și zgomot. Cred că o tratare locală a problemei alegerii parametrilor ar rezolva situații de genul acesta, iar natura exact așa a rezolvat: imaginile sunt tratate local, nu numai în detectia primitivelor inițiale, dar și în mai toate prelucrările preconștiente (Cavanach 1999).

Rezultatele din figura 5.8 (imagine $127 \times 118 \times 256$ BMP) le-am inclus pentru a scoate în evidență faptul că în situația în care nu apar diferențe substanțiale de iluminare în zona care ne interesează (statuia), ci chiar dimpotrivă, ea se detasează net de fondul mai închis, rezultatele sunt relativ multumitoare. Am prezentat modulul gradientului pe imagine pentru a scoate mai bine în relief și multitudinea de muchii false de mică intensitate, ce sunt imediat eliminate. Desigur că o filtrare obișnuită pe imagine ar scoate aceste zgomote, dar am vrut să arăt că și rețeaua de integrare a contururilor le elimină cu succes. Detaliile din (d)-inițial și (f)-final arată că muchiile importante din punct de vedere al informației picturale au fost relativ bine pastrate.

Rezultatele din figura 5.9 (imagine $348 \times 447 \times 256$ BMP) sunt prezentate în aceeași idee, de contururi clar definite și ușor de identificat. Deși la prima vedere rezultatele sunt banale, le-am inclus și ca o dovadă că modelul propus trece testele de bază. În situația mecanismelor vederii naturale există date experimentale (Held 1987, Wilson&Keil 1999) care sugerează faptul că, în timp ce vederea binoculară, prelucrarea texturii și a perspectivei tri-dimensionale trebuie să “învăte” și să se facă ușor de către copii foarte mici și de către animale.

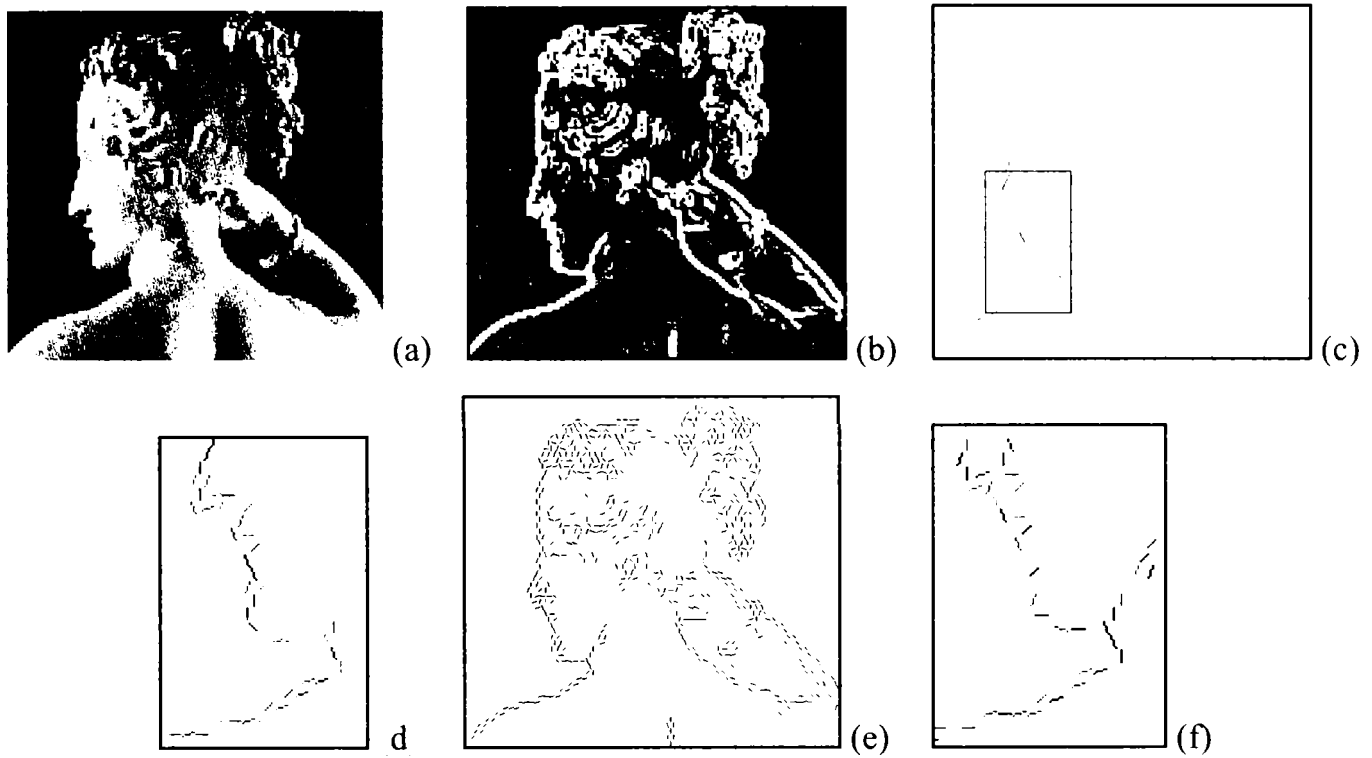


Figura 5.8: Imagine cu rezultate acceptabile (8 muchii/hipercoloana): (a) imaginea initiala; (b) rezultatul aplicarii filtrului gradient (modulul); (c) intrarile initiale in hipercoloane, cu evidentierea zonei care este apoi marita in d si f; (e) rezultatul obtinut dupa 4 iteratii; (d) detaliu din imaginea c; (f) detaliu din imaginea e.

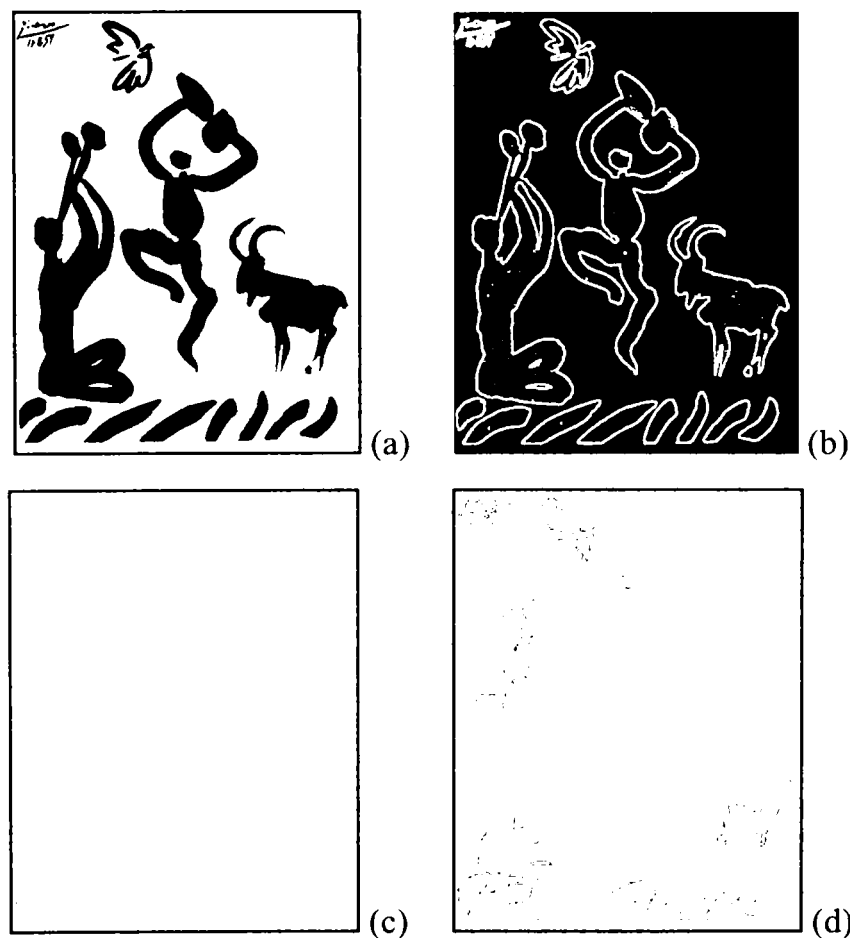


Figura 5.9: Imagine cu contururi clar delimitate (8 muchii/hipercoloana): (a) imaginea initiala; (b) rezultatul aplicarii filtrului gradient pe imaginea a (modulul); (c) intrarile in hipercoloane; (d) iesirea dupa 2 iteratii.

Figura 5.10 prezinta rezultatul obtinut in urma prelucrarii celebrei imagini Lenna si se poate observa ca sunt pastrate destul de bine contururile esentiale din imagine, daca tinem cont de faptul ca am redus mult rezolutia initiala a imaginii (300×281×256 BMP).

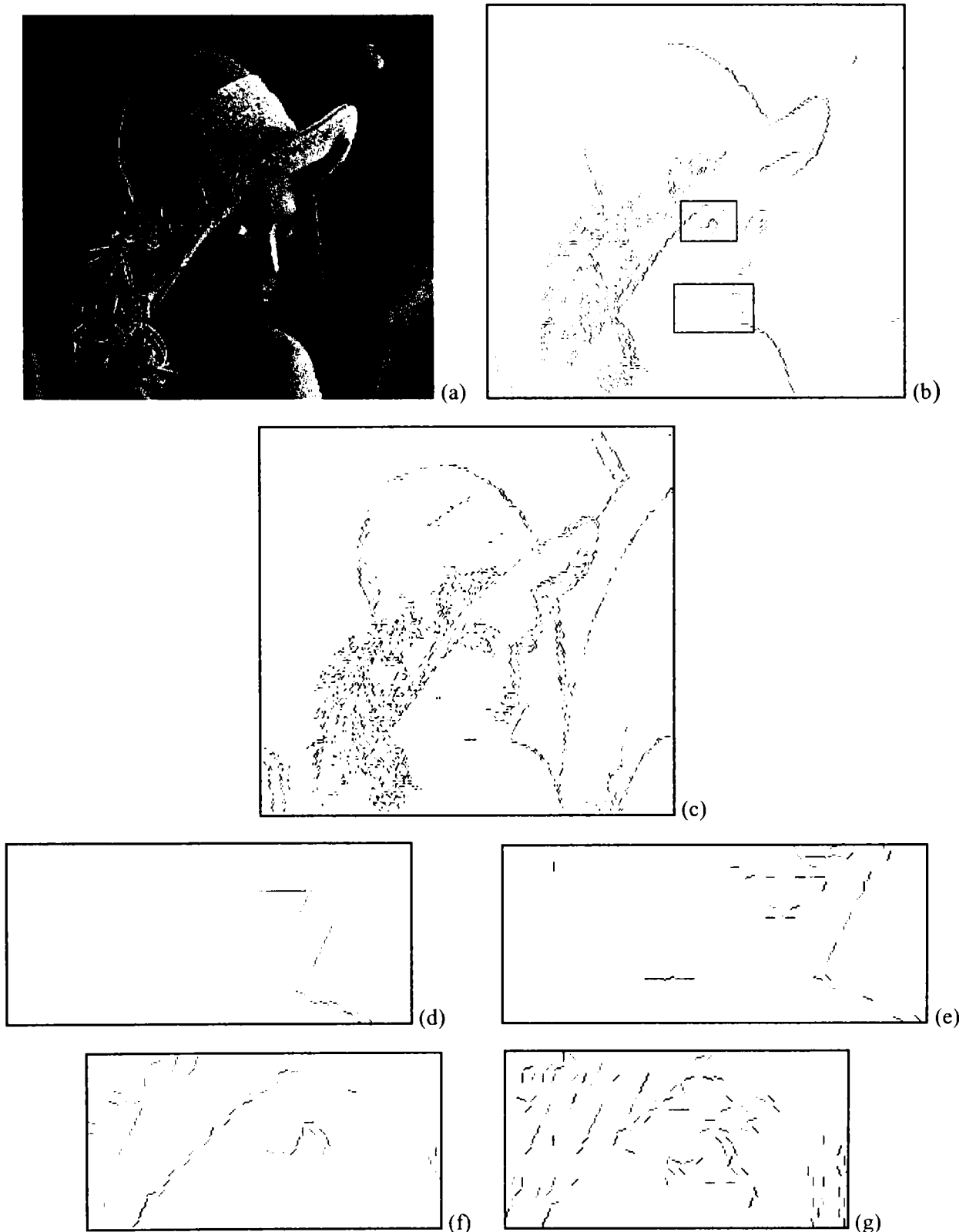


Figura 5.10: Rezultate obtinute in urma prelucrarii imaginii Lenna (8 muchii/hiperculoana): (a) imaginea initiala; (b) intrarea in hipercoloanele de orientare, cu evidentierea zonelor detaliate in **d**, **e**, **f**, si **g**; (c) rezultatul obtinut dupa 4 iteratii; (d) detaliu din **b**; (e) detaliu din **c**; (f) detaliu din **b**; (g) detaliu din **c**.

5.4. CONCLUZII

Am conceput, am implementat și am testat un model de întărire a contururilor pornind de la muchii elementare, bazat pe considerente biologice (Lungeanu et al. 1998, Lungeanu&Cotirlea 1999, Lungeanu et al. 2000).

Modelul este realizat printr-o rețea recurentă formată din hipercoloane de orientare. Arhitectura și dinamica rețelei se bazează pe datele privind procesarea informației vizuale primare la nivelul hipercoloanelor de orientare din cortexul striat și pe proprietățile cunoscute ale celulelor simple și complexe din V1. Comportamentul caracteristic de integrare a contururilor în această rețea se datorează în principal legăturilor excitatorii și inhibitorii atât pe distanța scurtă, cât și pe distanțe lungi, dovedite a avea un rol important în funcționarea hipercoloanelor din V1 (Sagawa 1997, Singer 1995, Vidyasagar et al. 1996). Funcționarea rețelei nu presupune intervenția unui *feed-back* din exterior (sau *top-down* din alte arii).

Am implementat rețeaua realizând o aplicație descrisă în Anexa 5 și am testat modelul pentru diferite tipuri de imagini. Din considerente practice, am recurs la o serie de **simplificări**: am utilizat o grilă hexagonală pentru repartizarea hipercoloanelor în spațiul vizual (ceea ce a simplificat foarte mult calculul distanțelor); la trecerea de la grilă ortogonală la cea hexagonală am redus cu un factor de 1/4 dimensiunea imaginii; am utilizat un filtru gradient de determinare a muchiilor elementare (tarie și orientare), am utilizat un număr redus de muchii elementare în fiecare coloană (4 sau 8 în cazurile prezentate aici) distribuite uniform unghiular.

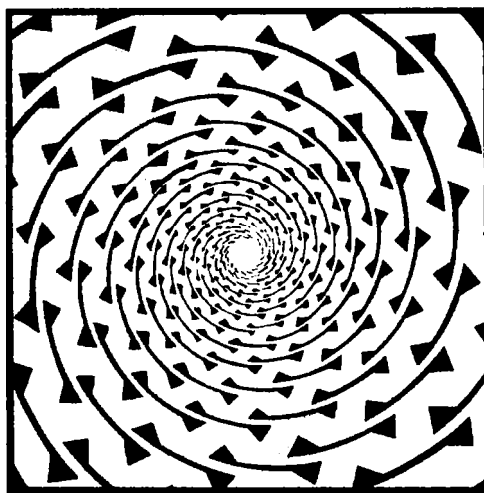
Așa cum era de așteptat, simularile pe diverse tipuri de imagini au evidențiat faptul că performanțele modelului depind nu numai de parametrii săi, ci și de conținutul imaginii procesate. Acest lucru este perfect compatibil cu datele experimentale: la nivelul cortexului striat există un grad mare de redundanță informațională și de paralelism în procesare, dat de utilizarea unor strategii combinate (Churchland&Sejnowski 1992, Li&Atick 1994, Marr 1982). Acest fapt sugerează necesitatea includerii unor criterii de “ajustare” a parametrilor sau de combinare a rezultatelor paralele în cadrul modelului, soluție adoptată și în alte modele (Miller&Zucker 1999).

Simplificările făcute s-au reflectat în aspectele strict tehnice ale performanțelor obținute în simulare. Totuși, datorită faptului că mi-am propus realizarea unui model care să testeze plauzibilitatea unor ipoteze privind mecanismele biologice de integrare a contururilor la nivelul ariei vizuale primare, cred că aceste rezultate nu trebuie privite strict prin prisma performanțelor în ceea ce privește prelucrarea imaginilor (care nu sunt competitive cu cele obținute prin tehnicile clasice din domeniu). De aceea nici nu am considerat necesar să caut neapărat un criteriu obiectiv de evaluare a performanțelor modelului și, în această fază, am preferat o evaluare vizuală a lor (chiar dacă ea este implicit subiectivă).

În evaluarea rezultatelor obținute, mai trebuie avut în vedere și faptul că cercetarea experimentală a dovedit existența coloanelor de orientare și a întăririi contururilor la nivelul V1. Însa nici din punctul de vedere al științelor biologice nu există încă înregistrări care să dea o măsură cantitativă a gradului de întărire a muchiilor și de integrare a contururilor. Ceea ce se știe este că există o “cablare” a rețelei neuronale, la nivelul V1, pentru integrarea contururilor (hipercoloanele de orientare și conexiunile dintre acestea). Se mai știe și că activitatea rețelei și a neuronilor individuali sunt dependente de context, chiar până la modificarea dimensiunilor și profilului câmpurilor receptoare, în anumite limite (DeAngelis et al 1995, Engel et al. 1992, Ramachandran 1995, Singer 1995). Mai mult, apare nu numai ca plauzibilă, dar chiar ca probabilă o combinație a informației vizuale de pe canale paralele de procesare de natură diferită pentru a obține senzația de ansamblu (Ernst et al. 1999, Vershure&Konig 1999).

Ca o concluzie generala, se poate spune ca, din punct de vedere biologic, mecanismele implementate in acest model sunt plauzibile si ca el conduce intr-adevar la intarirea muchiilor ce apartin unor contururi, eliminandu-le pe cele disperate din imagine, fara vreo interventie *top-down*.

O directie in care s-ar putea aduce imbunatatiri prezentului model ar fi aceea a utilizarii unor retele adaptabile la context. Dificultatea abordarii consta in faptul ca, dincolo de performantele tehnice urmarite, trebuie mereu mentinuta legatura cu aspectele neurofiziologice, cu evidentele si dovezile experimentale. Nu trebuie pierdut din vedere ca se incearca modelarea de mecanisme biologice.



Capitolul 6: CONCLUZII GENERALE

Preocuparile mele privind modelarea proceselor vizuale s-au concentrat asupra ariei vizuale primare deoarece exista o mare bogatie de date experimentale privind mecanismele ce actioneaza la acest nivel. Totodata, au fost dezvoltate multe modele care au incercat sa conecteze informatiile ce provin dintr-un domeniu larg de preocupari stiintifice (neurostiinte, biologie, psihofizica, stiinte cognitive, etc.) si sa “umple” oarecum zonele ce lipsesc din punctul de vedere al datelor obiective. Toate acestea au facut ca imboldul initial de curiozitate sa fie sustinut pe fundamente stiintifice.

In acest sens, cred ca rezultatele obtinute pana acum sunt incurajatoare (Lungeanu 1995, Lungeanu et al. 1996, Lungeanu et al. 1998, Lungeanu et al. 1998b, Lungeanu et al. 1998c, Lungeanu&Cotirlea 1999, Lungeanu et al. 2000). Voi incerca sa le sintetizez in cele ce urmeaza.

1. Am abordat modelarea a doua dintre mecanismele implicate in perceptia 3-D a spatiului: perceptia profunzimii in vederea stereoscopica si integrarea primara a contururilor. Cele doua mecanisme modelate se desfasoara in paralel la nivelul cortexului striat, pentru a fi apoi mixate si, modulate de mecanismele pre-atentionale, trimise spre ariile mai inalte de prelucrare.
2. Am realizat doua aplicatii care implementeaza cele doua modele si permit testarea acestora. Pentru vederea stereo am implementat si un al doilea model, al lui David Marr si Tomaso Poggio, care este clasic si pe care l-am folosit ca termen de referinta in testele facute.
3. Concluziile specifice privind comportamentul fiecaruia dintre modele au fost formulate la capitolele respective. Totusi, cred ca se pot trage cateva concluzii privind ambele modele:
 - ambele modele s-au dovedit atat viabile, cat si plauzibile biologic;
 - performantele obtinute sunt incurajatoare din punctul de vedere al modelarii mecanismelor biologice;
 - din punctul de vedere al tehnicilor clasice de prelucrari de imagini, aceste modele nu constituie un concurent (sau, cel putin, nu deocamdata). Acest lucru se poate datora unor cauze de natura diferita: atat simplificari facute in modelare, cat si faptului ca s-ar putea ca reprezentarea acestor informatii la nivelul V1 sa fie “zgomotoasa” in fapt.

Este plauzibil sa se obtina o reprezentare grosiera a informatiei in prima faza, pentru ca apoi ea sa fie rafinata pe baza interactiunilor dintre populatiile de neuroni ce sunt specializate pentru prelucrari de natura diferita: stereo, contururi, textura, miscare. S-ar putea, de asemenea, ca in aceasta “rafinare” a informatiei sa apara rolul legaturilor *top-down*. Ramane ca viitoarele cercetari experimentale sa lamureasca aceste supozitii.
4. In ceea ce priveste modelarea hartilor de disparitate, modelul “biologic” a permis evidentierea unor aspecte interesante, oarecum “ocolite” ca abordare de unii cercetatori care prefera sa-si testeze ipotezele si modelele doar in conditiile care avantajeaza mai mult:
 - obtinerea unor rezultate bune pe RDS-uri nu implica neaparat aceleasi performante si in cazul unor imagini obisnuite;

- testarea si evaluarea performantelor modelelor stereo trebuie sa se faca atat pentru RDS-uri, cat si pentru imagini “reale” – sistemul vizual uman trece ambele teste;
 - pentru stereograme cu imagini “reale”, asumarea conditiei de epipolaritate in sensul strict este o greseala si ea se reflecta in performantele obtinute cu astfel de modele sau sisteme.
5. In ceea ce priveste modelul de integrare a conturilor, el dovedeste plauzibilitatea intaririi conturilor la nivelul hiper-coloanelor de orientare din cortexul striat. Acest comportament are la baza numeroasele conexiuni intra-corticale pe distante scurte si lungi si nu implica legaturi de *feed-back* din alte zone corticale. Chiar daca nu rezolva in fapt integrarea propriu-zisa a conturilor (ea doar incepe la nivelul V1), el dovedeste ca intarirea muchiilor ce apartin unor contururi are loc in cortexul striat, prin chiar organizarea acestuia in coloane de orientare.
 6. Ambele modele realizeaza, in fapt, o reprezentare distribuita a informatiei, analoaga organizarii retinotopice de la nivelul ariei vizuale primare.
 7. In concluzie, cred ca rezultatele dovedesc plauzibilitatea presupunerilor initiale. De asemenea, cred ca in viitorul nu prea indepartat voi reusi sa integrez informatiile deocamdata dispartate privind scenele tri-dimensionale.

ANEXA 1

DATE DESPRE SISTEMUL NERVOS UMAN

CREIERUL

Greutatea creierului adult – 1300 ÷ 1400 grame
Greutatea creierului unui nou nascut – 350 ÷ 400 grame
% din greutatea totala a corpului – 2%
Numar mediu de neuroni – 100 miliarde
Numar mediu de celule gliale – de 10 ÷ 50 ori mai multe decat numarul de neuroni
Numarul de fibre optice in nervul optic – 1.200.000
Numarul de fibre in nervul oculomotor – 25.000 ÷ 35.000

VAZUL

Lungimea globului ocular – 24.5 mm
Timpul mediu intre clipiri – 2.8 s
Durata medie a unei clipiri – 0.3 s
Grosimea corneei – 0.54 mm in centru; 0.65 la periferie
Diametrul corneei – 11.5 mm
Grosimea cristalinului – 4 mm
Diametrul cristalinului 9 mm
Compozitia cristalinului – 65% apa si 35% proteine

Numarul de celule receptoare – 6 milioane de conuri si 120 milioane de bastonase
Densitatea conurilor in fovee – 200.000 / mm²
Densitatea maxima a bastonaselor – 175.000 / mm² (la 20 de grade de fovee)
Diametrul foveei ≈ 1.8 mm

Acuitatea vizuala – 1 arc min (1/60 grade), adica ≈ 1.5 mm de la distanta de 5 m
Hiperacuitatea – 4 ÷ 6 arc sec pentru segmente, 2 arc sec pentru disparitati binoculare

Numarul de celule in LGN – 570.000
Numarul de celule in V1 – 538.000.000

Lumina vizibila – lungimi de unda 400 ÷ 700 nm
Cantitatea de lumina necesara pentru a excita un bastonas – 1 foton
Cantitatea de lumina necesara pentru a excita un con – 100 fotoni
Presiunea intraoculara - 10 ÷ 20 mm Hg
Volumul orbitei – 30 ml

NEURONII

Numarul tipic de sinapse – 1.000 ÷ 10.000
Viteza de conductie a potentialului de actiune – 0.6 ÷ 120 m/s
% de neuroni vizualizati prin metoda Golgi – 5%
Compozitia mielinei – 70 ÷ 80% lipide, 20 ÷ 30% proteine

% oxigen utilizat de creier – 20%
% flux sanguin de la inima la creier - 15 ÷ 20%

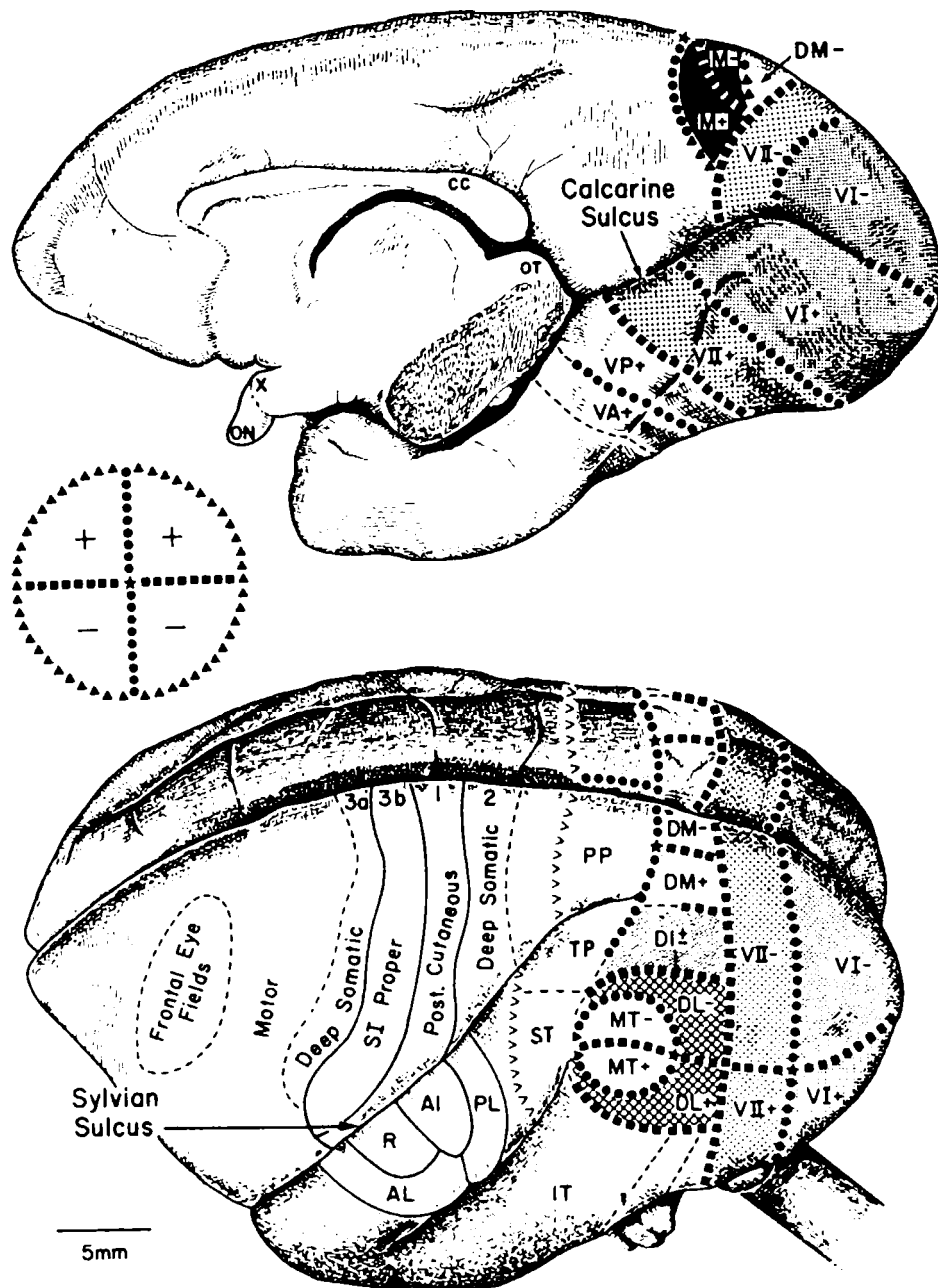
ANEXA 2

ORGANIZAREA CORTEXULUI VIZUAL LA MAIMUTA

(Allman 1987)

Reprezentarea domeniilor senzoriale ale cortexului cerebral la maimuta:

- sus: vedere ventro-mediala;
- jos: vedere dorso-laterala.
- in stanga: o harta a perimetrului campului vizual (simbolurile sunt suprapuse pe harta cortexului vizual)
 - plusurile indica reprezentari ale cadranelor superioare
 - minusurile - cadranele inferioare
 - linii intrerupte - granitele ariilor ce corespund unor reprezentari ale partilor relativ periferice ale campului vizual
 - siruri de "V"-uri - granita aproximativa a cortexului vizual
 - linia punctata (cu semne "?") - granita incerta

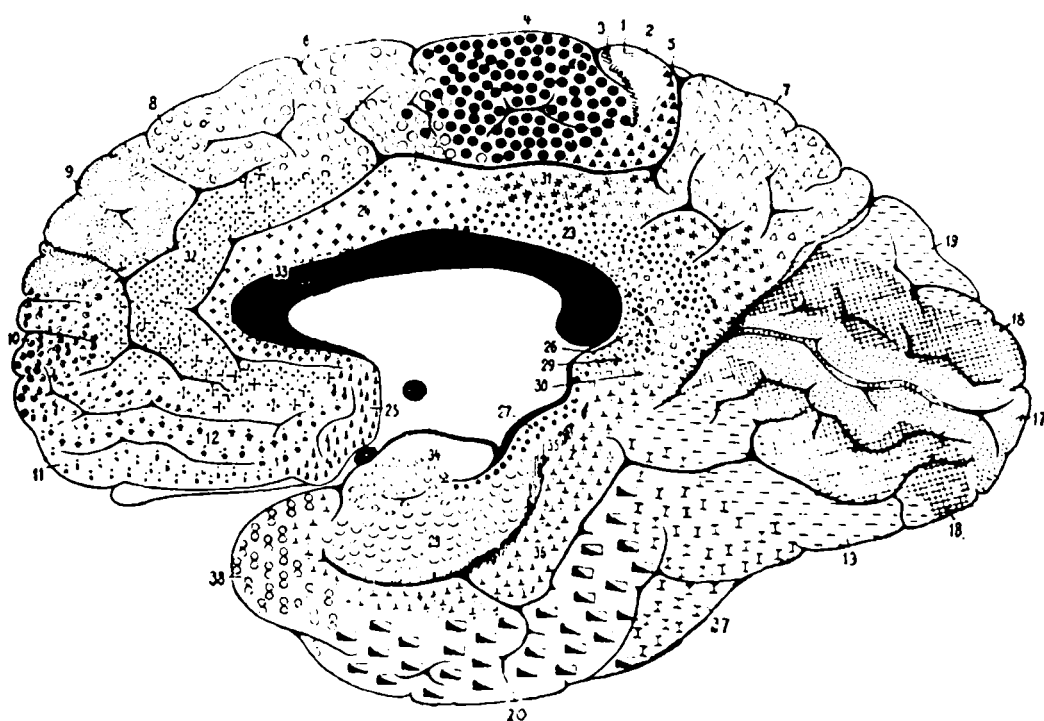
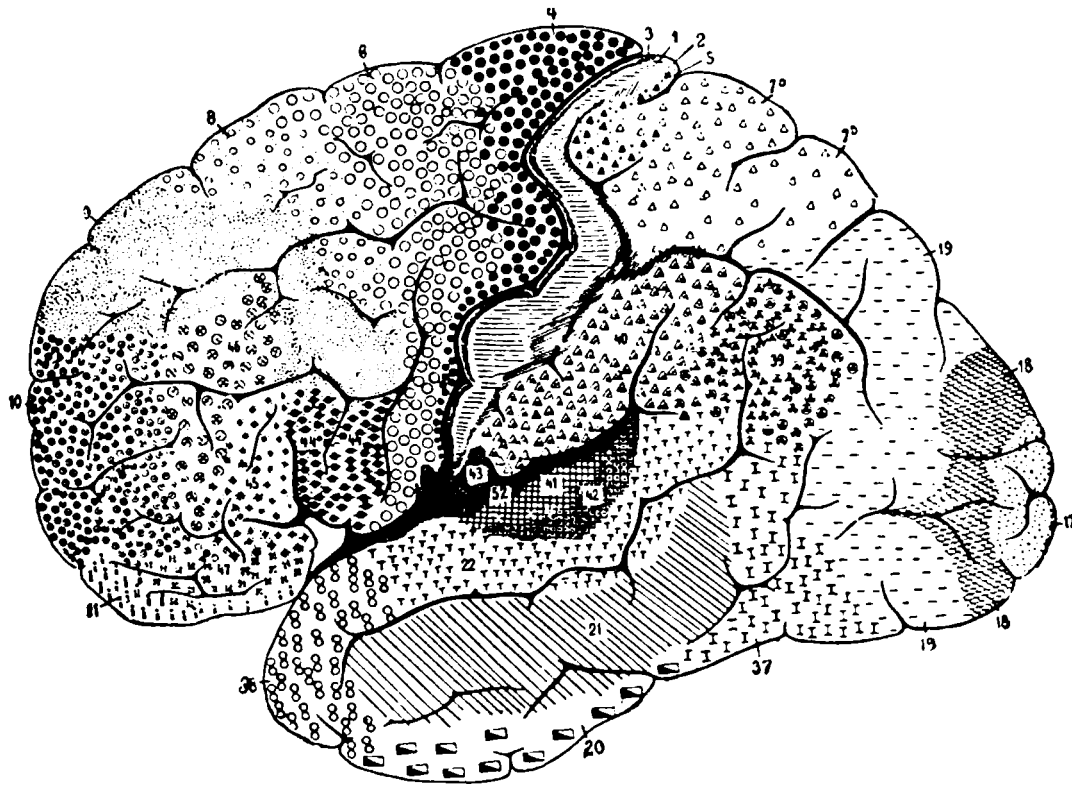


ANEXA 3

ARIILE CORTICALE BRODMANN

(Jones 1987)

Zonele propuse de K. Brodmann in 1909 (ariile preferate de "anatomisti"): vedere laterala si vedere mediala.



ANEXA 4

PREZENTAREA APLICATIEI *STEREO-VISION*

Aplicatia *Stereo-Vision* a fost realizata pentru implementarea si testarea algoritmilor de determinare a hartii de disparitate binoculara descrisi in Capitolul 4: algoritmul clasic al lui Marr&Poggio si algoritmul "biologic".

Aplicatia a fost dezvoltata in Microsoft Visual C++ v 1.50 si a fost rulata pe un PC sub Windows 95.

INTERFATA CU UTILIZATORUL

Stereo-Vision are sase functii:

- crearea stereogramelor cu puncte aleatoare pentru testarea algoritmilor
- aplicarea algoritmilor de determinare a hartii de disparitate
- vizualizarea rezultatelor
- evaluarea performantelor
- posibilitatea adaugarii de zgomot pe imagini
- filtrarea imaginilor

Crearea streogramelor cu puncte aleatoare. Intrarile si iesirile

Figura 1 prezinta *main-frame*-ul aplicatiei, cu urmatoarele elemente:

- Bara de titlu (*Title Bar*)
- Bara de meniu (*Menu Bar*)
- Bara de unelte a aplicatiei (*Toolbar*)
- Suprafata de lucru a utilizatorului (*Client Area*), impreuna cu doua fereastre copil ale aplicatiei (*Child frame window*)
- Bara de stare (*Status Bar*)

Funcție de contextul de executie, meniul se modifica, oferind functii specifice. Optiunile puse la dispozitia utilizatorului pe parcursul executiei aplicatiei sunt:

Meniul *File*:

1. *New* → ofera posibilitatea crearii unui nou bitmap generat aleator (cu puncte albe si negre, sau colorate) – figura 1. Utilizatorul este solicitat sa specifice tipul bitmap-ului (alb-negru sau color), dimensiunea (limitele sunt intre 10 si 300 pixeli pe fiecare dimensiune), gradul de umplere (densitatea) si distributia punctelor – figura 2.

Conventia de atribuire a numelor este: Imaginea LEFT trebuie sa aiba numele **L*.bmp**, iar imaginea pereche RIGHT, **R*.bmp** (figura 1). Nerespectarea acesteia va avea ca si consecinta functionarea incorecta a programului (rezultate eronate).

2. *Open* → permite selectarea unui tip de fisier din cele doua existente (bitmap sau *wireframe*) si incarcarea acestuia in memorie.
3. *Lista cu cele mai recente 4 fisiere deschise*. Selectarea unui nume din aceasta duce la selectarea fisierului corespunzator si incarcarea lui in memorie.
4. *Exit* → parasirea programului.

Meniul *Edit*: se activeaza ca urmare a incarcarii unui fisier de tip bitmap si ofera functiile standard: *Clear, Copy, Cut i Paste*.



Figura 1

Meniul View:

1. Toolbar → selectarea / deselectarea acestuia permite afisarea, respectiv ascunderea barei de unelte.
2. Status bar → selectarea / deselectarea acestuia are ca efect vizualizarea / ascunderea barei de stare (linia pe care se afiseaza mesaje scurte catre utilizator).
Urmatoarele trei optiuni se activeaza doar in cazul in care exista un fisier deschis:
3. Normal → aceasta este implicit selectata, reprezentand fisierul in dimensiune reala.
4. Zoom In → scaleaza imaginea respectiv: 2:1, 3:1, 4:1
5. Zoom Out → este optiunea complementara celei anterioare realizand scalari de 1:2, 1:3 i 1:4.

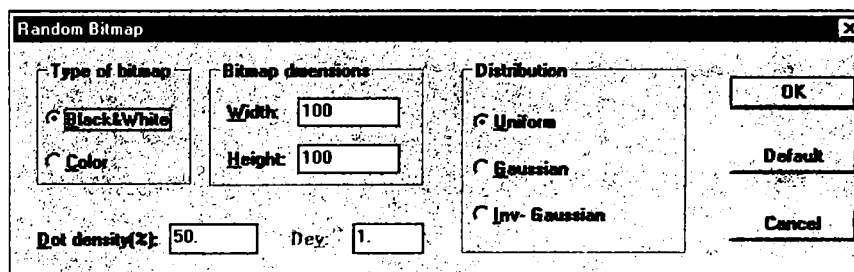
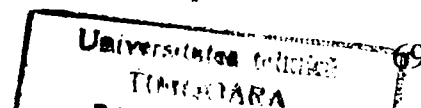


Figura 2

Meniul Draw apare doar in cazul incarcarii unui fisier bitmap. Actiunile oferite de acesta au corespondent in bara de unelte (figura 1):

1. Select → este implicit activ, in acest context nefiind posibila desenarea.
2. Rectangle, Triangle, Ellipse → permit desenarea figurilor geometrice corespunzatoare. Dupa terminarea unui desen, utilizatorul poate deplasa pe orizontala (stanga-dreapta) suprafata delimitata de conturul anterior marcat, tinand butonul stang al mouse-ului apasat. Stergerea unui desen se realizeaza actionand butonul drept al mouse-ului.

Atunci cand se "decupeaza" zone de diferite forme si se deplaseaza pe orizontala, se creeaza automat un fisier cu denumirea Y*.wrf, ce contine harta de disparitate "ideala"



(utilizata apoi pentru evaluarea performantelor). Fisierul *.wrf este un fisier de tip text, specific aplicatiei, care contine valorile de disparitate pentru fiecare punct al imaginii (figura 5, sus).

3. *Noisy* → permite adaugarea de zgomot pe imaginea în lucru. Utilizatorul trebuie sa specifice densitatea zgomotului, precum si tipul distributiei acestuia – figura 3. Zgomotul diagonal este adaugat dupa un algoritm propus de Bela Julesz (1995).

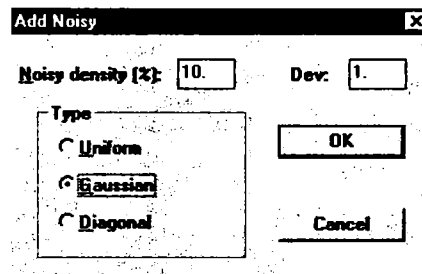


Figura 3

Aplicarea algoritmilor. Evaluarea performantelor

Au fost implementati cei doi algoritmi descrisi in Capitolul 4, care pot fi alesi dintr-un meniu *Run* indicandu-se informatiile necesare.

La executie, harta de disparitate este generata automat si este salvata pe disc sub forma de fisier *wireframe* cu numele *X*.wrf* in directorul de origine a fisierelor ce formeaza stereograma (figura 5, jos).

Meniul *Run*: ofera utilizatorului selectarea si executarea unuia dintre modelele implementate, aprecierea rezultatelor obtinute, precum si aplicarea unei medieri suplimentare pe un anumit rezultat:

1. *Models*: la alegerea acestei optiuni se deschide o fereastră care permite utilizatorului sa aleaga modelul dorit. In cazul selectarii modelului lui Marr&Poggio, este necesar sa se specifice un numar de iteratii de executat (care implicit este 15), iar in cazul modelului “biologic” se ofera psibilitatea alegerii unei variante de mediere (Gauss, mediere la valoarea cea mai frecventa, la valoarea mediana sau o mediere adaptiva). Ambele modele necesita o pereche de fisiere de intrare reprezentand stereograma de prelucrat.
2. *Quality assessment* → furnizeaza utilizatorului o apreciere a rezultatelor obtinute in urma aplicarii unuia dintre modele. In cazul in care s-a lucrat pe imagini create de utilizator (adica RDS-uri), este necesara specificarea unui fisier rezultat in urma unei modelari/simulari si a fisierului “ideal” de referinta. Daca se doreste aprecierea obiectiva a performantelor unuia dintre modele aplicat pe o imagine reala, se poate face o comparatie intre doua rezultate obtinute pe aceeasi imagine. Utilizatorului i se cere sa specifice cele doua nume de fisiere in care sunt memorate rezultatele.

Evaluarea performantelor se face prin numarul de pixeli in care valoarea de disparitate este diferita in cele doua harti comparate si prin eroarea medie patratica a valorii de disparitate – figura 4.

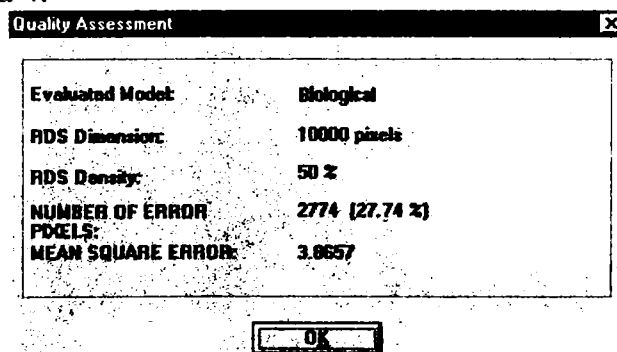


Figura 4

Ca rezultat al aplicarii unei medieri, rezulta un fisier al carui nume are urmatorul tipar: $xTq*.wrf$, unde $T = '0'$ daca se paraseste dialogul prin "Cancel", $T = 'g'$ daca este aleasa medierea de tip gauss, $T = 'f'$ in cazul medierii la valoarea cea mai frecventa, $T = 'm'$ daca se alege medierea la valoarea mediana, si respectiv $T = 'a'$ in cazul medierii adaptive.

Mai trebuie mentionat ca nu exista restrictii privitoare la numarul de ferestre "decupate" la realizarea RDS-rilor, la forma acestora, sau la combinatiile si suprapunerile facute. Acest lucru permite testarea unor forme relativ complexe. Singura restrictie se refera la deplasarea strict pe horizontala (impusa prin program) in plaja $[-4, +4]$.

Se pot prelucra si testa si imagini "reale", pentru care se obtin acelasi tip de rezultate ca si pentru RDS-uri (harta de disparitate sub forma de *wireframe* si in nuante de gri) si pe care se pot face acelasi tip de prelucrari ulterioare (adaugare de zgomot, filtrari, evaluari rezultate) – figura 7.

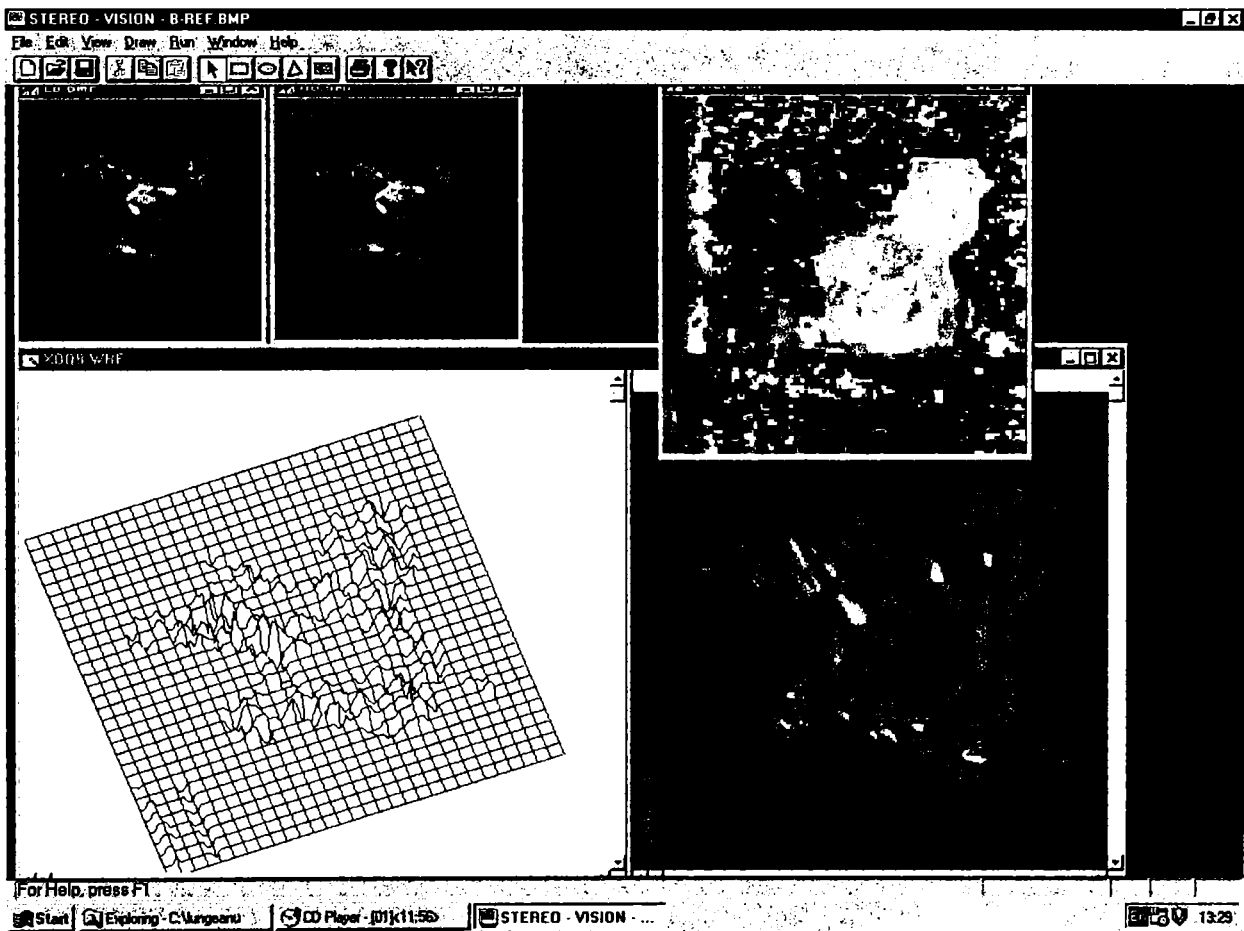


Figura 7

IMPLEMENTAREA APLICATIEI STEREO-VISION

Aplicatia a fost scrisa in C++, utilizand biblioteca de obiecte MFC (*Microsoft Foundation Classes*). Biblioteca ofera o aplicatie schelet (*application framework*) pe care programatorul o dezvolta. Schema din figura 8 prezinta modul in care clasele programatorului se leaga de cele standard MFC.

Clasele de dialog ofera facilitati pentru o comunicare interactiva cu utilizatorul. Deoarece acestea sunt derivate din clasa *CWnd*, ele primesc mesaje, pot fi deplasate sau inchise.

Pentru fiecare dialog cu utilizatorul, este prevazuta o clasa asociata. Toate dialogurile existente sunt modale: este necesara terminarea unui dialog prin **OK** sau **CANCEL**, pentru continuarea executiei programului.

Celelalte dialoguri sunt:

CRandomDialog
CAveragesDialog,
CModelsDialog, *CResultsDialog*
CComparesDialog
CAddNoisy

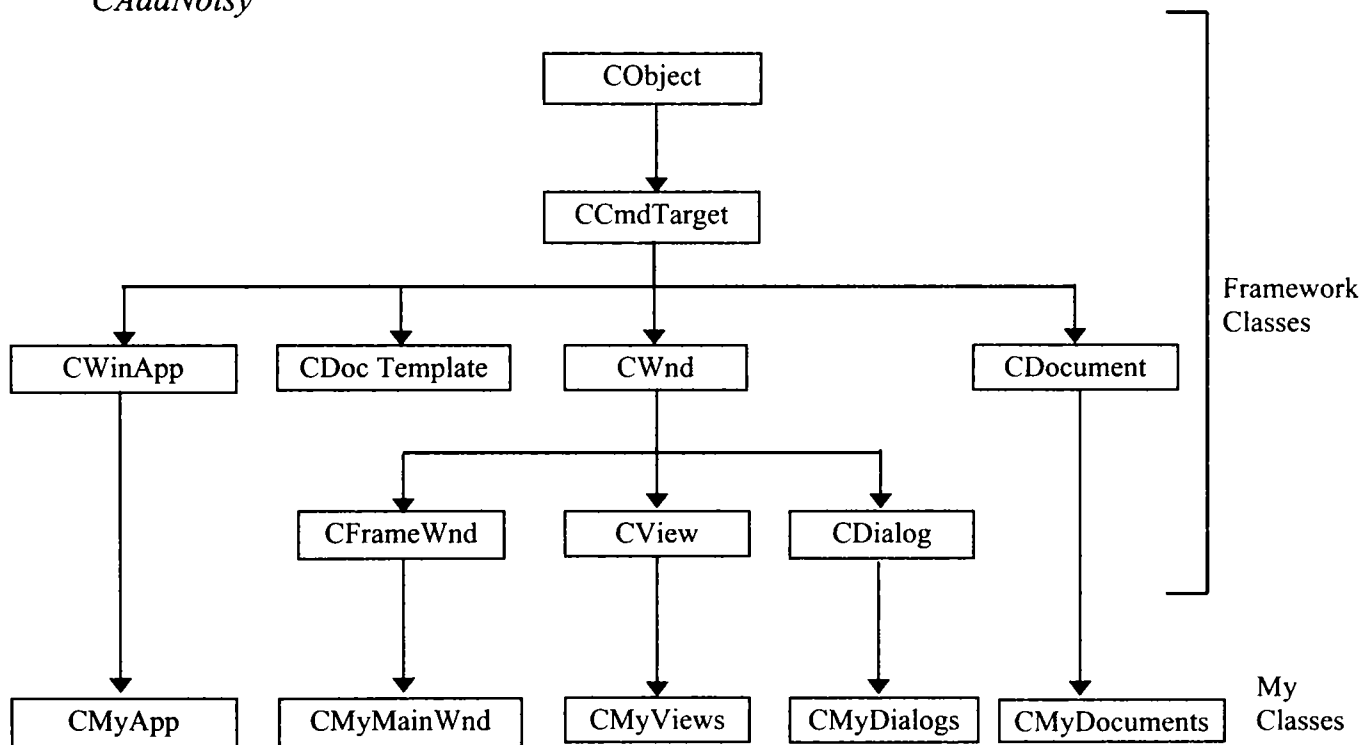


Figura 8

Clasa CAlgorithm este cea care realizeaza implementarea modelelor propuse:

```

class CAlgorithm
{
private:
    int m_nROWS, m_nCOLS;
    MATRIX_BYTE* m_RezMatrix;
    CString      m_RezFileName;
    CProgressBar m_ProgressionBar;
private:
    void marr_Cycle(MATRIX_3D& , int);
    void Med(MATRIX_FLOAT&, int, CString);
    BOOL WriteFile(CString);
protected:
    void marr (int, MATRIX_BYTE& , MATRIX_BYTE&); //nr. iteratii, img. stanga, img. dreapta
    void qian (int, MATRIX_BYTE& , MATRIX_BYTE&, CString);
public:
    CAlgorithm();
    ~CAlgorithm();
    void method(int , int, MATRIX_BYTE& , MATRIX_BYTE&, CString);
    void Med1(MATRIX_FLOAT&, int);
    MATRIX_BYTE& GetResult() { return *m_RezMatrix; }
    CString GetRezFileName(){ return m_RezFileName;}
}
    
```

Implementarea modelului Marr&Poggio

// MARR & POGGIO's model

```

void CAlgorithm:: marr_Cycle(MATRIX_3D &N, int nr)
{
MATRIX_3D rez( m_nROWS, m_nCOLS, DISP );
int x,y,d,i;
int inh, exc;
double dif;

    for (int c=0; c<nr; c++)
    {
        for (d=0; d<DISP; d++)
            for (x=2; x<m_nROWS-2; x++)
                for (y=DISP; y<m_nCOLS-DISP; y++)
                {
                    inh=0;
                    exc=N(x,y,d);
                    for(i=d-1; i>=0; i--)
                        if (i>=0){
                            inh+=N(x,y-(d-i),i) + N(x,y+d-i,i);
                        }
                    for (i=d+1; i<DISP; i++)
                        if(i<DISP){
                            inh+=N(x,y-(i-d),i) + N(x,y+i-d,i);
                        }
                    inh*=EPS;
                    exc+=N(x-1,y,d) +N(x-2,y,d) +N(x+1,y,d)+
                        N(x+2,y,d)+ N(x-1,y-1,d) +N(x-1,y+1,d) +
                        N(x+1,y-1,d) +N(x+1,y+1,d) +N(x,y-1,d) +
                        N(x,y-2,d) +N(x,y+1,d) +N(x,y+2,d);
                    dif = exc - inh;
                    if ( dif >= TETA )
                        rez(x,y,d) = 1;
                    else
                        rez(x,y,d) = 0;
                }
            } // end for(y<m_nCOLS)
        m_ProgressionBar.UpdateProgressionBar(c+1);
        N = rez;
    } // end for(c<nr)
}
//nr. iteratii, img. stanga, img. dreapta
void CAlgorithm::marr (int i, MATRIX_BYTE& MLeft, MATRIX_BYTE& MRight)
{
MATRIX_3D image(m_nROWS, m_nCOLS, DISP);
image.Initialize(MLeft,MRight);
marr_Cycle(image,i);
m_RezMatrix = new MATRIX_BYTE(m_nROWS, m_nCOLS);
image.Conv_3Dto2D(m_RezMatrix);
}

```

Funcția de bază a modulului este *marr()* ai cărei parametri au următoarea semnificație. Numărul de iterații *i*, furnizat de utilizator, este folosit pentru antrenarea rețelei (pentru a o conduce către o stare stabilă). Perechea de imagini de intrare este dată de matricile *MLeft* și *MRight*, de tip *MATRIX_BYTE* (matrici 2D cu elemente de tip byte). În primul pas al modelării se construiește un volum *image* din cele două matrici de intrare, având două dimensiuni date de *ROWS* și *COLS* ale

imaginilor de intrare si inaltimea fiind data de numarul nivelelor de disparitate luate in considerare, *DISP*. Regula dupa care se determina valorile de pornire in algoritm:

$$image(x,y,d) = MLeft(x,y) * MRight(x+d,y),$$

$$x=0 \div ROWS, y=0 \div COLS, d=0 \div DISP,$$

exprimand conditia initiala a retelei.

Elementele volumului anterior creat *image* reprezinta noduri ale retelei. In functia *marr_Cycle()* pentru fiecare nod se calculeaza contributia vecinilor *inhibitori* si a celor *excitatori* si folosind functia prag se determina starea celulei corespunzatoare iteratiei curente. Procesul se repeta de *i* ori, de fiecare data actualizandu-se volumul. Daca numarul iteratiilor este suficient de mare, se ajunge intr-o stare stabila.

Am considerat pentru valoarea constantei de inhibitie EPS=2, iar pentru valoarea de prag TETA=3.

Rezultatul funciei *marr_Cycle()* fiind un volum, este necesar sa fie convertit la o matrice 2D, reprezentand harta de disparitati. Initial, elementele matricii *RezMatrix* au valoarea corespunzatoare disparitatii ZERO. Pentru realizarea conversiei se ia in considerare fiecare plan de disparitate *d* constanta si se aplica urmatoarea formula:

daca $image(x,y,d) = 1$ atunci $RezMatrix(x,y) = d$
altfel valoarea lui $RezMatrix(x,y)$ ramane nemodificata.

Volumul *image* se parcurge incepand cu nivelele cele mai indepartate (disparite +4) catre nivelele cele mai apropiate de privitor (disparitate -4), solutie considerata fireasca: obiectele mai apropiate le obtureaza pe cele indepartate.

Implementarea modelului "biologic"

```
// Biological model
```

```
int maxd(double* m)
{
    int i, indice = 0;
    double mx=m[0];
```

```
    for(i=0; i<N; i++) {
        if(mx<m[i]) {
            mx=m[i];
            indice=i;
        }
    }
    return indice;
}
```

```
void CAlgorithm::qian (int avg_mth, MATRIX_BYTE& MLeft, MATRIX_BYTE& MRight, CString FileName)
{
    MATRIX_FLOAT map(m_nROWS, m_nCOLS);
    double rc[NY], rs1[NY], rs2[NY], disp;
    int x, y, i, j, d, i_disp, iq, jq;
```

```
    int level = (int) NX/2;
    m_ProgressionBar.Create((long)m_nROWS);
    for(x=level; x<m_nROWS-level; x++) {
        for(y=DISPZ; y<(m_nCOLS-DISPZ); y++) {
            for(d=0; d<NY; d++)
```

```

{
    rs1[d]=0; rs2[d]=0;
    for(i=0; i<NY; i++) {
        iq=i-(NY)/2;
        for(j=0; j<NX; j++)
        {
            jq = j - (NX)/2;
            rs1[d]+=gaussy[i] * gaussx[j] * ( cos(2*PI* (U * iq + V * jq) ) *
MLeft(x+jq,y+iq)+ cos(2*PI* (U * iq + V * jq) + PhiR[d]) * MRight(x+jq,y+iq) );
            rs2[d]+=gaussy[i] * gaussx[j] * ( sin(2*PI* (U * iq + V * jq) ) *
MLeft(x+jq,y+iq)+ sin(2*PI* (U * iq + V * jq) + PhiR[d]) * MRight(x+jq,y+iq) );
        } //for j
    } //for i
    rc[d]=rs1[d]*rs1[d]+rs2[d]*rs2[d];
} //for d

i_disp=maxd(rc);
disp=PhiR[i_disp]/omega;
map(x,y)=(float)(-disp);
} //for y
m_ProgressionBar.UpdateProgressionBar(x+1);
} //for x
m_ProgressionBar.DestroyWindow();
Med(map, avg_mth, FileName);
for(int r=0;r<m_nROWS;r++)
    for(int p=0;p<m_nCOLS;p++)
        map(r,p)+=DISPZ;

m_RezMatrix= new MATRIX_BYTE(map);
} //end Biological

```

Functia de baza a acestui modul este $qian()$, ai carei parametri au urmatoarea semnificatie. Perechea de imagini de intrare este data de matricile $MLeft$ si $MRight$, de tip $MATRIX_BYTE$ (matrici 2D cu elemente de tip byte). Parametrul de tip intreg avg_mth reprezinta tipul de mediere folosit si este specificat de catre utilizator.

Pentru fiecare locatie spatiala se determina raspunsurile a 9 celule complexe, avand fazele $(\phi_r - \phi_i)$ uniform distribuite in intervalul $[-\pi/2, \pi/2]$, si acestea se combina in raspunsul complex. Raspunsurile celulelor simple care formeaza o astfel de familie sunt reprezentate la nivelul algoritmului prin sirurile de numere reale $rs1[NY]$, $rs2[NY]$, ($NY=9$ reprezinta numarul de celule considerate pe orizontala), iar raspunsurile celulelor complexe prin $rc[NY]$.

Raspunsul unei celule simple este calculat aplicand anvelopa gaussiana asupra celulelor simple care formeaza vecinatatea acesteia. Vecinatatea este considerata ca fiind compusa din NY (numarul de celule complexe cu o aceeasi localizare spatiala pe orizontala) si NX (numarul de celule cu o aceeasi localizare spatiala pe verticala). Raspunsul este dat de urmatoarea formula:

$$rs1[d] = 0 \text{ (initial)}$$

$$rs1[d] += gaussy[i] * gaussx[j] * (\cos(2*PI* (U * iq + V * jq)) * MLeft(x+jq,y+iq) + \cos(2*PI* (U * iq + V * jq) + PhiR[d]) * MRight(x+jq,y+iq));$$

unde:

- x si y sunt coordonatele celulei in perechea de imagini,
- $iq = -4 \div +4$, $jq = -2 \div +2$.
- $U = 1/8$ si $V = 0$ (pentru filtrare Gabor uni-dimensionala) sau alte valori pentru filtrare Gabor bi-dimensionala.

Prin combinarea a doua celule simple se formeaza o celula complexa, al carei raspuns este dat de:

$$rc[d]=rs1[d]*rs1[d]+rs2[d]*rs2[d];$$

Pentru a calcula explicit disparitatea imaginii la o anumita locatie, se determina celula cu cel mai puternic raspuns din familia corespunzatoare locatiei respective (functia *maxd()*).

Pentru a imbunatati rezultatele obtinute, se aplica o mediere peste harta de disparitati, mediere de tipul specificat prin parametrul *avg_mth*. Aplicatia pune la dispozitie 4 tipuri de medieri:

1. **mediere gaussiana unidirectionala** pe orizontala (2 vecini).
2. **mediere gaussiana pe cele 4 directii** (8 vecini).

Pentru cele doua medieri prezentate anterior se aplica anvelopa gaussiana pe una, respectiv patru directii.

3. **mediere la valoarea cea mai frecventa** - pentru fiecare locatie (x,y) se determina valoarea cea mai frecventa dintr-un sir de valori considerate pe cele doua axe OX si OY in intervalele $(x-4, x+4)$ si respectiv $(y-4, y+4)$.

4. **mediere la valoarea mediana** - pentru fiecare locatie (x,y) se considera valorile elementelor din harta de disparitati de la locatiile din intervalele $(x-4, x+4)$ si respectiv $(y-4, y+4)$. Acest sir se ordoneaza crescator si se ia valoarea din mijloc.

5. **mediere adaptiva** - se realizeaza printr-o compunere repetata a semnalului cu o masca foarte mica ponderata de o masura a continuitatii semnalului in fiecare punct. Pasii care se realizeaza sunt urmasorii (Saint-Marc et al. 1991):

- 2) calcularea gradientilor $G_x(x,y)$ si $G_y(x,y)$:

$$G_x(x,y) = \frac{1}{2}(I^{(1)}(x+1,y) - I^{(1)}(x-1,y))$$

$$G_y(x,y) = \frac{1}{2}(I^{(1)}(x,y+1) - I^{(1)}(x,y-1))$$

- 2) calcularea coeficientului de continuitate $w(x,y)$:

$$w(x,y) = e^{-\frac{G_x^2(x,y)+G_y^2(x,y)}{2k^2}}$$

- 3) calcularea mediei ponderate pentru a actualiza $I(x,y)$:

$$I^{(n+1)}(x,y) = \frac{\sum_{i=-1}^{+1} \sum_{j=-1}^{+1} I^{(n)}(x+i,y+j)w^{(n)}(x+i,y+j)}{\sum_{i=-1}^{+1} \sum_{j=-1}^{+1} w^{(n)}(x+i,y+j)}$$

ANEXA 5

PREZENTAREA APLICATIEI *VISION - VISUAL CORTEX*

Aplicatia *Vision-Visual Cortex* a fost realizata pentru implementarea si testarea modelului de integrare a conturilor descris in Capitolul 5.

Aplicatia a fost dezvoltata in Microsoft Visual C++ v 5.0 si a fost rulata pe un calculator PC sub Windows NT v 4.0.

INTERFATA CU UTILIZATORUL

Vision - Visual Cortex indeplineste urmatoarele functii:

- determinarea muchiiilor elementare din imagini *.bmp prin aplicarea filtrului gradient
- schimbarea coordonatelor spatiale – pregatirea datelor ca intrari pentru retea ce modeleaza arhitectura corticala pe hipercoloane
- setarea parametrilor retelei ce modeleaza integrarea conturilor la nivelul V1
- aplicarea algoritmului ce descrie dinamica retelei
- vizualizarea rezultatelor intermediare si a celor finale
- salvarea rezultatelor intermediare si finale

Functie de contextul de executie, meniul se modifica, oferind functii specifice. Optiunile puse la dispozitia utilizatorului sunt prezentate in continuare.

Intrarile si iesirile

Figura 1 prezinta *main-frame*-ul aplicatiei, cu urmatoarele elemente:

- Bara de titlu (*Title Bar*)
- Bara de meniu (*Menu Bar*)
- Suprafata de lucru a utilizatorului (*Client Area*), împreun cu o fereastra copil a aplicatiei (*Child frame window*)
- Meniul *File* deschis

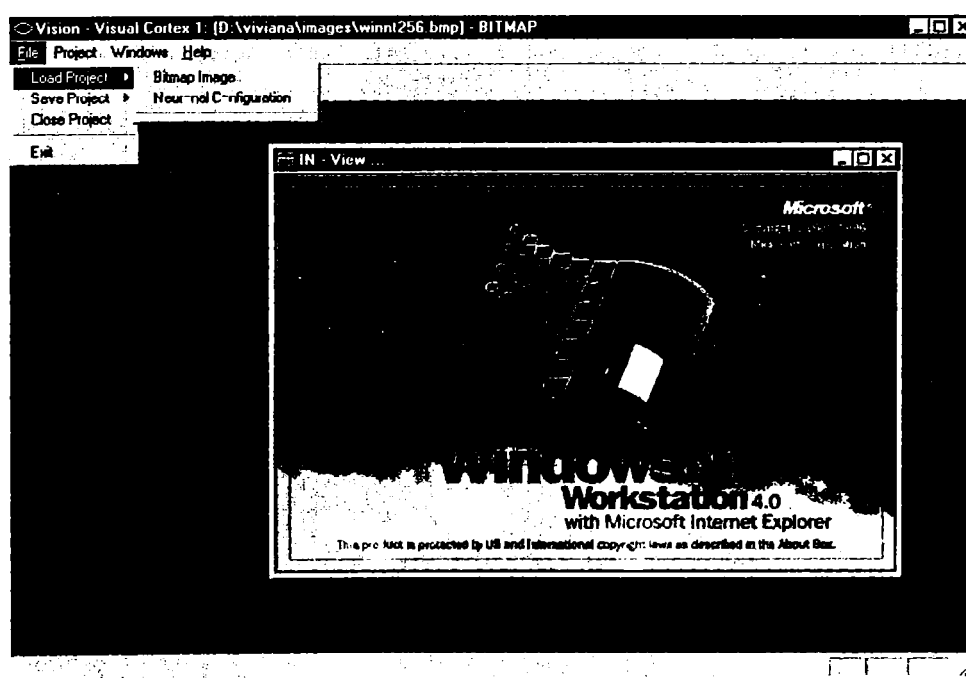


Figura 1

Meniul *File*:

1. *Load Project* → permite incarcarea unui fisier ***.bmp** (*Bitmap Image*) sau a configuratiei de lucru salvate la o sesiune de lucru anterioara, intr-un format specific aplicatiei (*Neuronal Configuration*) – figura 1.
2. *Save Project* → permite salvarea rezultatelor din fereastra **Out Bmp** sub forma de fisier ***.bmp** (*As a Bitmap*), sau a intregii configuratii de lucru pentru o reluare ulterioara a prelucrarii (*As a Neuronal Configuration*).
3. *Close Project* → permite inchiderea prelucrarii curente, fara parasirea programului.
4. *Exit* → parasirea programului.

Setarea parametrilor retelei. Executia algoritmului si vizualizarea rezultatelor

Meniul *Project*:

1. *Settings* → *General Parameters* ofera posibilitatea setarii parametrilor ce descriu dinamica retelei – figura 2. Exista valori implicite pentru acesti parametri, dar ei pot fi schimbati dupa dorinta.

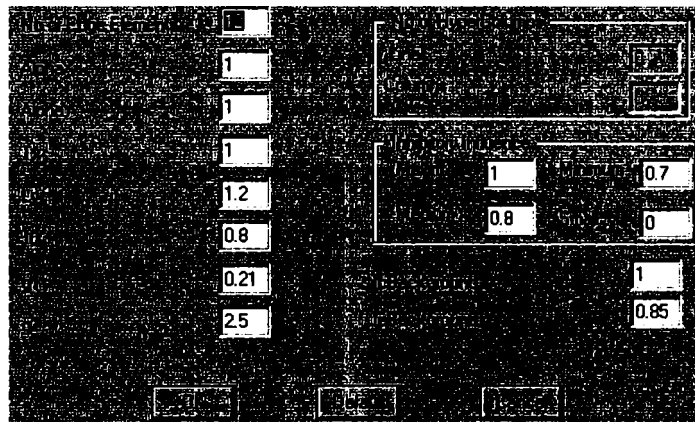


Figura 2

2. *Settings* → *Workspace Settings* ofera patru categorii de setari referitoare la spatiul de lucru si la modul in care se face vizualizarea rezultatelor.
 - Optiunea *OUT-View* permite specificarea dimensiunii hipercoloanelor (in pixeli), asa cum vor apare afisate pe ecran si, eventual, salvate in fisierul ***.bmp** rezultat – figura 3. Tot in cadrul acestei optiuni, se poate specifica daca se doreste actualizarea iesirii dupa fiecare iteratie sau doar la sfarsitul seriei de iteratii specificate.

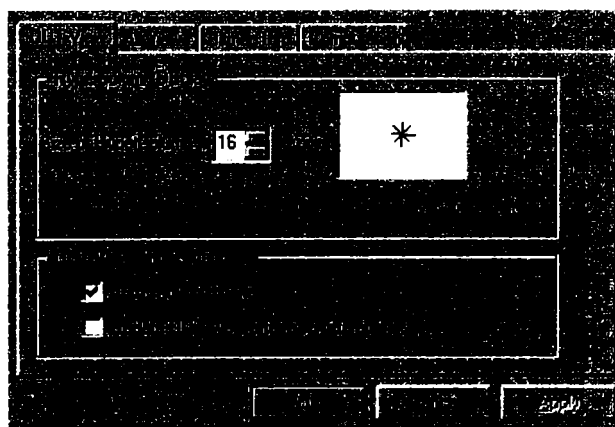


Figura 3

- Optiunea *All Views* permite specificarea rezultatelor ce se doresc vizualizate pe ecran, deci a ferestrelor ce vor fi deschise automat in timpul executiei – figura 4.

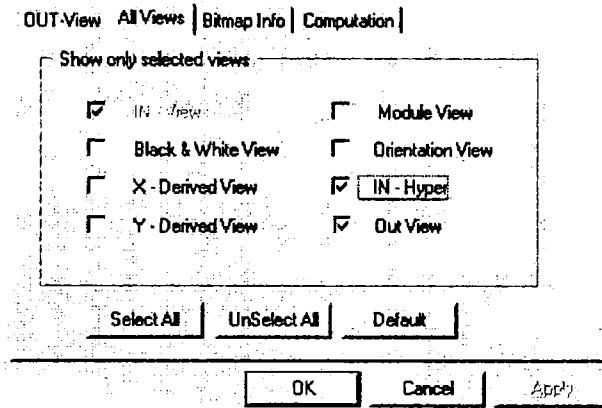


Figura 4

Se vede ca se pot vizualiza: *IN-View* (imaginea *.bmp ce se prelucreaza), *Black&White View* (imaginea monocroma), *X-Derived View* si *Y-Derived View* (rezultatul aplicarii filtrului gradient pe cele doua directii), *Module View* (modulul obtinut in urma aplicarii gradientului, deci taria muchiiilor elementare), *Orientation View* (orientarea muchiiilor elementare), *IN-Hyper* (intrarea in retea, sub forma in care ele vor fi afisate si ca rezultat de iesire din retea), *OUT View* (iesirea din retea). Figura 5 prezinta modul in care se face vizualizarea acestor rezultate intermediare.

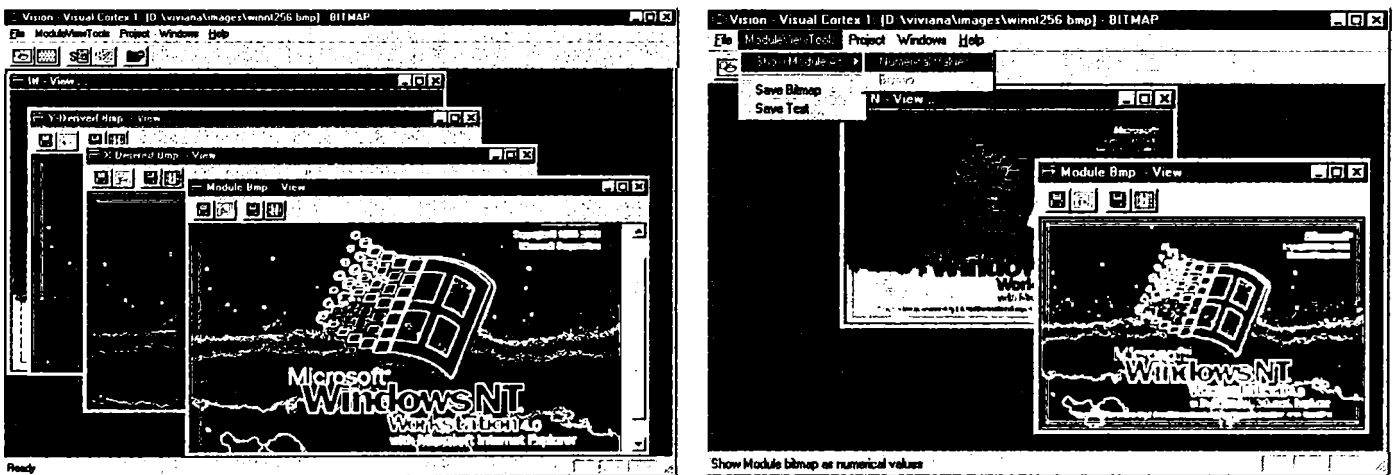


Figura 5

Se poate observa ca vizualizarea se face fie sub forma de imagine (ca in figura), fie sub forma de valori numerice (figura 5 dreapta, optiunea din meniul). Totodata, asa cum am precizat la meniul *File*, aceste rezultate se pot salva sub oricare din cele doua forme (imagine *.bmp sau text), uneltele de salvare fiind disponibile atat ca butoane (la fiecare fereastră), cat si ca optiuni din meniul. Intrarea in retea, respectiv iesirea se pot vizualiza sub forma de imagini, in format video-normal (figura 7) sau video invers, sau in format text si se pot salva similar.

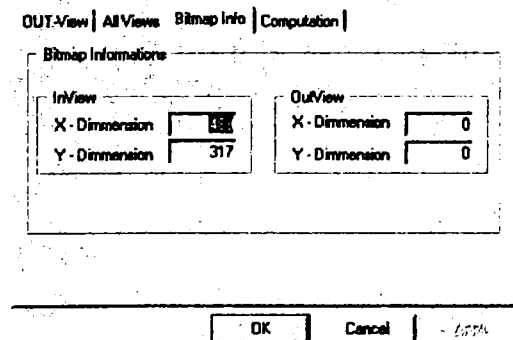


Figura 6

- Optiunea *Bitmap Info* permite specificarea informatiilor privind dimensiunea imaginii de intrare, respectiv a celei de iesire (in cazul salvarii sub forma de *.bmp) – figura 6. Dimensiunea imaginii **Out** este dependenta de dimensiunea specificata a hipercoloanelor (*Out-View*).

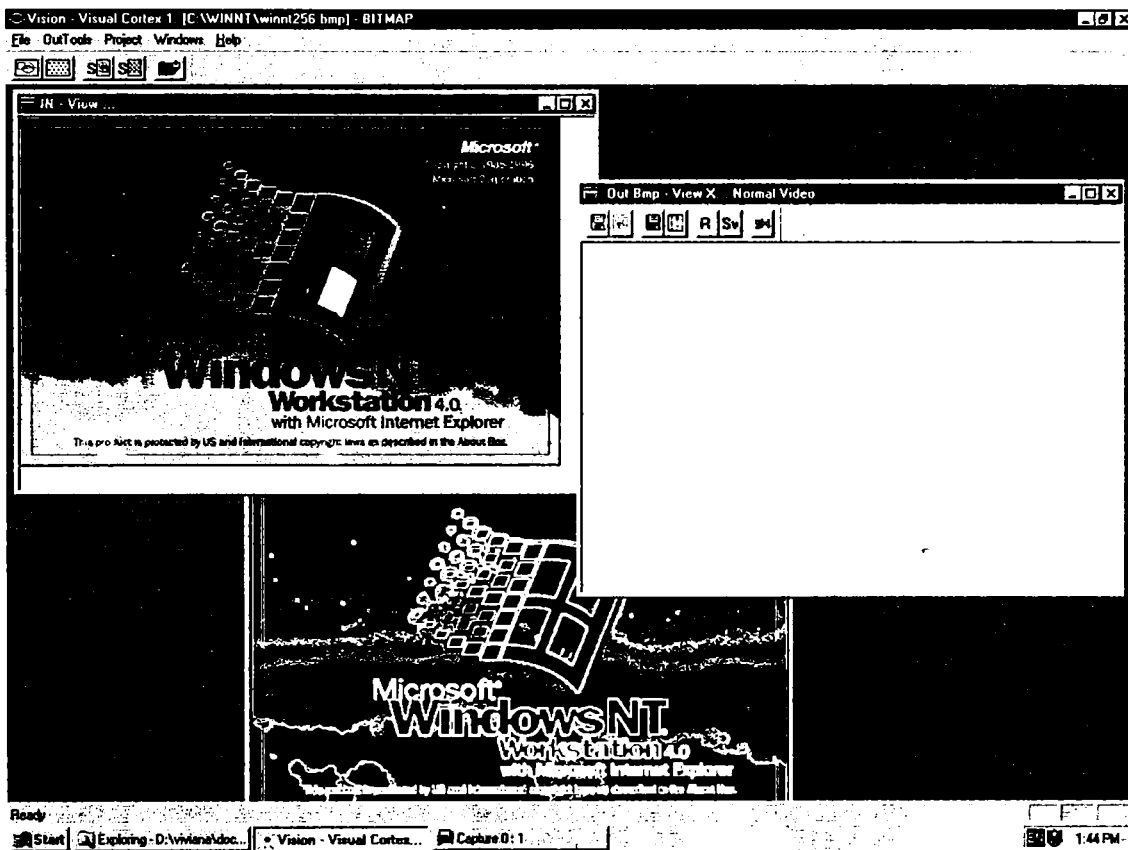


Figura 7

- Optiunea *Computation* permite specificarea numarului de iteratii ce urmeaza a fi efectuate si daca se doreste sau nu medierea rezultatului – figura 8.

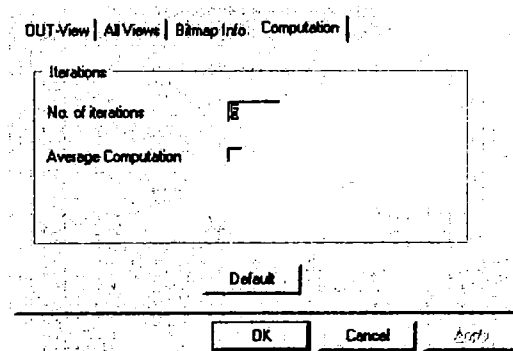


Figura 8

Meniul *Project*:

3. *Run* permite executia algoritmului: fie pas cu pas, cu posibilitatea vizualizarii rezultatelor intermediare (*Step by Step*), fie executia automata a unui numar specificat de iteratii (*Run away...*) – figura 9.

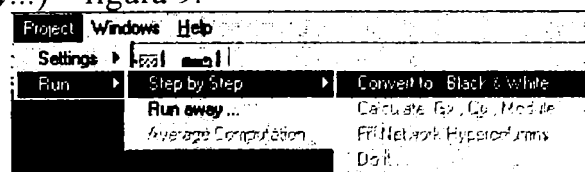


Figura 9

In timpul executiei este vizualizat stadiul in care se gaseste aceasta intr-un *Progress Dialogue* – figura 10. Executia poate fi intrerupta temporar si reluata apoi (eventual dupa vizualizarea unor rezultate intermediare).

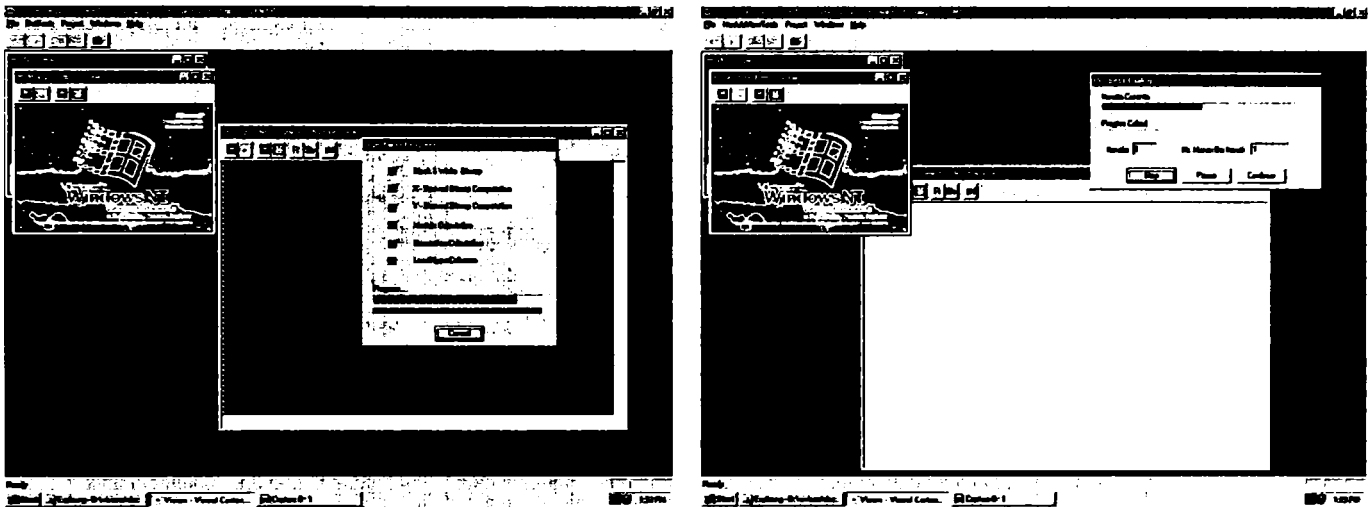


Figura 10

Meniul *Windows* permite alegerea ferestrelor active (figura 11), ferestre ce au aceeași semnificație cu cea descrisă la opțiunea *All Views* de la *Project Workspace* → *Settings*. Se pot deschide oricare dintre ferestrele ce conțin informație/rezultate, fără nici o condiționare legată de opțiunea de la vizualizare automată.

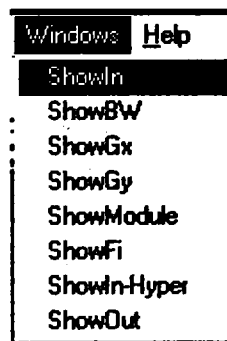


Figura 11

Meniul *Help* oferă câteva informații despre program.

IMPLEMENTAREA APLICATIEI *VISION - VISUAL CORTEX*

Din punct de vedere tehnic, aplicația este împărțită în două module:

- Un modul care implementează o clasă ce încapsulează primitive de prelucrare a fișierelor format *bitmap* (*.bmp), numit **MyBitmapDll**, care extinde funcțiile clasei *CBitmap*. Acest modul este o parte auxiliară independentă a aplicației și furnizează funcții ce pot fi reutilizate și de alte programe: transformarea unei imagini color în imagine monocromă, aplicarea filtrului gradient pe imagine, calculul modulului și orientării, afișarea imaginii încapsulate în această clasă într-un DC (*Device Context*), operatorii = și +, calculul dimensiunii imaginii încapsulate.
- Un modul **Prj**, în care sunt implementate funcțiile aplicației:
 - nucleul – rețeaua propriu-zisă
 - interfața utilizator
 - operațiile de I/O cu mediile de stocare

Descrierea de ansamblu a modului Prj

Modulul principal al aplicatiei este o aplicatie MDI (*Multiple Document Interface*). **Nucleul** este organizat in trei straturi: definirea structurilor de date, constructia retelei neuronale, algoritmul care descrie dinamica retelei. Aceste straturi asigura interfata de comunicare cu celelalte module si sub-module ale aplicatiei.

Interfata utilizator asigura comunicarea dintre componentele aplicatiei si utilizator, facand posibila: introducerea de date, setarea parametrilor, setarile referitoare la conditiile de executie, controlul executiei programului, vizualizarea datelor de intrare, intermediare si finale sub forma numerica sau de imagine *.bmp.

Operatiile de I/O cu mediile de stocare nu constituie un modul propriu-zis, ele sunt distribuite, de fapt, la toate nivelele aplicatiei.

Aplicatia utilizeaza mecanismele furnizate de obiectele MFC (*Microsoft Foundation Classes*) si de sistemul de operare Windows NT - biblioteca MFC acopera functiile si structurile de date SDK-API (*Software Developer Kit – Application Programming Interface*).

Aplicatia utilizeaza atat ferestre modale (care necesita inchiderea pentru continuarea executiei programului), cat si ferestre nemodale.

Structurile de date folosite

Reteaua a fost implementata ca o matrice ce descrie organizarea retinotopica de la nivelul cortexului striat. Ea este formata din hipercoloane (distribuite uniform pe imagine), iar acestea sunt formate, la randul lor, din neuroni-pereche, fiecare neuron-pereche fiind caracterizat de doi neuroni primari si o orientare. In continuare sunt prezentate clasele ce implementeaza aceasta structura.

Clasa CNeuronPrimar

```

////////////////////////////////////
// header file for class CNeuronPrimar

#ifndef __NeuronPrimar_h_
#define __NeuronPrimar_h_

class CNeuronPrimar
{
private:
    double m_Potential;
public:
    //Constructori
    CNeuronPrimar();
    CNeuronPrimar(double);
public:
    //Metode
    double& Potential();
public:
    // Operators
    CNeuronPrimar& operator = (CNeuronPrimar& ob);
};

#endif

```

Neuronul primar este caracterizat de o valoare reala, numita *m_Potential*. Ca metode, se poate obtine potentialul unui neuron primar, sau se poate atribui un potential unui neuron primar.

Clasa COrientare

```

////////////////////////////////////

```

```
// header file for class COrientare

#ifndef __Orientare_h_
#define __Orientare_h_

class COrientare
{
private:
    int m_ID;
    double m_Unghi;
public:
    //Constructori
    COrientare();
    COrientare(int max_ID);
    //Metode
    double CalculUnghi();
    double CalculUnghiID(int arg_ID);
    double CalculUnghiRadiani();
    double CalculUnghiIDRadiani(int arg_ID);
    int GetID();
    double GetUnghi();
    void SetID(int);
    COrientare& operator = (COrientare& ob);
};

#endif
```

Orientarea este folosita in doua moduri: ca numar intreg (m_ID) ce indica la al catelea neuron-pereche din cadrul unei hipercoloane se refera orientarea, ca numar real (m_Unghi) ce indica unghiul pe care-l face muchia respectiva cu orizontala. Metodele implementate:

- *CalculUnghi()* calculeaza unghiul dintre doi neuroni-pereche;
- *CalculUnghiID(int arg_ID)* calculeaza unghiul facut de neuronul-pereche *arg_ID* cu orizontala;
- *CalculUnghiRadiani()* calculeaza unghiul in radiani dintre doi neuroni-pereche;
- *CalculUnghiIDRadiani(int arg_ID)* calculeaza in radiani unghiul facut de neuronul-pereche *arg_ID* cu orizontala.

Clasa CNeuronPereche

```
////////////////////////////////////
// header file for class CNeuronPereche

#ifndef __NeuronPereche_H_
#define __NeuronPereche_H_

#include "NeuronPrimar.h"
#include "Orientare.h"

class CNeuronPereche
{
private:
    // Data Members
    CNeuronPrimar m_Excitator;
    CNeuronPrimar m_Inhibitor;
    COrientare m_Orientare;
public:
    //Constructori
    CNeuronPereche();
```

```

CNeuronPereche(CNeuronPrimar&,CNeuronPrimar&,COrientare);
CNeuronPereche(CNeuronPereche&);
public:
//Operatorul egal
CNeuronPereche& operator = (CNeuronPereche&);
public:
//Metode
COrientare GetOrientare();
void SetOrientare(COrientare);
CNeuronPrimar& Excitator();
CNeuronPrimar& Inhibitor();
};
#endif
    
```

Neuronul-pereche este format din doi neuroni primari (unul excitator si unul inhibitor) si o orientare. Metodele implementate permit setarea/obtinerea valorii pentru cei doi neuroni primari si setarea/obtinerea valorii pentru orientare.

Clasa CHiperColoana

```

////////////////////////////////////
// Header file for class CHiperColoana

#ifndef __hiperColoana_H_
#define __hiperColoana_H_

class CNeuronPereche;
//-----BEGIN--
// class CHiperColoana definition
//-----
class CHiperColoana
{
private:
    CArray<CNeuronPereche,CNeuronPereche&> m_HiperColoana;
public:
    //Constructor
    CHiperColoana();
    CHiperColoana(CHiperColoana&);
public:
    //Operatorul egal
    CHiperColoana& operator = (CHiperColoana&);
    // Operatorul []
    CNeuronPereche& operator [] (int);
public:
    //Metode
    void Initializare();
};
//-----END---
#endif
    
```

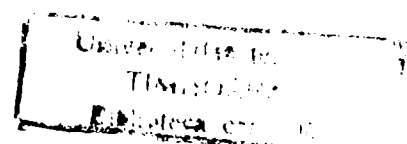
HiperColoana e formata dintr-un numar precizat de catre utilizator de muchii (neuroni-pereche), deci este un sir de neuroni-pereche. Metodele implementate permit setarea/obtinerea neuronilor-pereche din cadrul hiper-coloanei.

Clasa CCellula

```

////////////////////////////////////
// Header file for class CCellula

#ifndef __celula_h_
    
```



```
#define __celula_h_
#include "HiperColoana.h"

class CCelula
{
private:
    //Class members
    int m_linie;
    int m_coloana;
    CHiperColoana m_HiperColoana;

public:
    //Constructorii
    CCelula();
    CCelula(int,int,CHiperColoana);
    CCelula(CCelula&);

public:
    //Operatorul egal
    CCelula& operator = (CCelula&);

public:
    //Metode
    CHiperColoana& HiperColoana();
    int GetLinie();
    void SetLinie(int);
    int GetColoana();
    void SetColoana(int);
};

#endif // __celula_h
```

Celula unei matrici este caracterizata de localizarea spatiala (numarul liniei, numarul coloanei) si contine o hipercoloana. Metodele implementate permit: setarea/obtinerea coordonatelor, respectiv a valorilor pentru hipercoloana.

Clasa CMatrice

```
////////////////////////////////////
// header file for class CMatrice

#ifndef __matrice_h_
#define __matrice_h_

class CCelula;
class CHiperColoana;
class CNeuronPrimar;
class CGlobale;
class CMatrice
{
private:
    //Data members
    CArray<CCelula,CCelula> m_Matrice;

public:
    //Constructorii
    CMatrice();

public:
    //Metode
    long PozitieI(long arg_i,long arg_j);
    long PozitieJ(long arg_i,long arg_j);
```



```

long IndexMatrice(long i_nou,long j_nou);
CCelula& Celula(long , long );
void InitializeazaIndiciMatrice();
void SetNeuronExcitator(long arg_i,long arg_j,long arg_orientare,double arg_potentialNeuronExcitator);
double& GetNeuronExcitator(long arg_i,long arg_j,long arg_orientare);
void SetNeuronInhibitor(long arg_i,long arg_j,long arg_orientare,double arg_potentialNeuronInhibitor);
double& GetNeuronInhibitor(long arg_i,long arg_j,long arg_orientare);
public:
    // Operators
    CMatrice& operator = (CMatrice&);

public:
    // More methodes
    double CalculXi(long i,long j,long teta);

//f-ctii pt calculul formulei
    //pt x
    double gx(double);
    double gy(double);
    double Primul_TermenX(long ,long ,long );
    double Al_Treilea_TermenX(long ,long ,long );
    double Al_Doilea_TermenX(long ,long ,long );
    double corecturi(double);
    double putere_func(double,double,long );
    double J(long ,long ,long ,long ,long ,long );
    double Al_Patrulea_TermenX(long ,long ,long );
    double INormalizare(long ,long ,long );
    double Al_Cincelea_TermenX(long i,long j,long teta);
    double Al_Saselea_TermenX(long i,long j,long teta);

    //pt y
    double Primul_TermenY(long i,long j,long teta);
    double Al_Doilea_TermenY(long i,long j,long teta);
    double Al_Patrulea_TermenY();
    double W(long ,long ,long ,long ,long ,long );
    double Al_Treilea_TermenY(long ,long ,long );
    double CalculYi(long i,long j,long teta);

};

#endif

```

Matricea descrie organizarea retinotopica a hiper-coloanelor de orientare si este definita ca un sir de celule. Metodele implementate sunt pentru calculul formulelor ce descriu dinamica retelei, formule prezentate in Capitolul 5 (x – neuronul excitator, iar y – neuronul inhibitor).

In simularile prezentate in Capitolul 5, am utilizat:

Numar muchii elementare/hiper-coloana: 4 sau 8

$\alpha_x=1\div 5$; $\alpha_y=1\div 5$

$T_x=0.7\div 1$

$g_1=0.21$; $g_2=1.5\div 2.5$; $L_y=1.2$

$J_0=0.8\div 1.2$

$I_{e\ back}=0\div 1$; $I_{c\ back}=0\div 1$

$W_{\max}=0.5\div 1$; $J_{\max}=0.8\div 2$

Nr. iteratii: $2\div 12$

BIBLIOGRAFIE SI REFERINTE

- Adelman, G. (Ed.), (1987). *Encyclopedia of Neuroscience (I+II)*, Birkhauser, Boston, Massachusetts.
- Ahissar, M., Hochstein, S., (1995). *How early is the early vision? Evidence from perceptual learning*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E.Kowler, Eds.), The MIT Press, Cambridge, 199-206.
- Ahuja, N., (1996). *A transform for multiscale image segmentation by integrated edge and region detection*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 12, 1211-1235.
- Albright, T.D., (1994). *Why do things look as they do*, TINS, Vo.17, No.5, 175-177.
- Albright, T.D., Neville, H.J., (1999). *Neurosciences*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, li-lxxii.
- Allman, J., (1987). *Visual system, organization*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1283-1287.
- Aloimonos, J., (1989). *Integration of visual modules*, Academic Press, London.
- Amit, Y., Geman, D., (1999). *A computational model for visual selection*, Neural Computation 11, 1691-1715.
- Andrade, M.A., Moran, F., (1997). *Receptive field map development by anti-hebbian learning*, Neural Networks, Vol. 10, 1037-1052.
- Arbib, M.A., (Ed.), (1995). *The Handbook of Brain Theory and Neural Networks*, The MIT Press, Cambridge, Massachusetts.
- Azimi-Sadjadi, M.R., Xiao, R., Yu, X., (1999). *Neural network decision directed edge-adaptive Kalman filter for image estimation*, IEEE Trans. on Image Processing, Vol. 8, No. 4, 589-592.
- Bauer, H.U., (1995). *Development of oriented ocular dominance bands as a consequence of areal geometry*, Neural Computation 7, 36-50.
- Bengtsson, M. (1993). *A neural system as a dynamical model for early vision*, Neural Networks, Vol. 6, 313-325.
- Blais, B.S., Intrator, N., Shouval, H., Cooper, L.N., (1998). *Receptive field formation in natural scene environments: comparison of single-cell learning rules*, Neural Computation 10, 1797-1813.
- Blake, A., (1995). *Active vision*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 61-63.
- Blake, R., (1995). *Psychoanatomical strategies for studying human visual perception*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 17-25.
- Bojkovic, Z.S., Toma, C.I., Gui, V., Vasiu, R., (1997). *Advanced topics in digital image processing*, Editura Politehnica, Timisoara.
- Boykov, Y., Veksler, O., Zabih, R., (1998). *A variable window approach to early vision*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, No. 12, 1283-1294.
- Brown, K.S., (1999). *Are you ready for a new sensation?*, Scientific American, Vol. 10, No. 3, 38-43.
- Cabanis, E.A., Bourgeois, H., Iba-Zizen, M.-T., (1996). *L'imagerie en ophtalmologie*, Société Française d'Ophtalmologie et Masson, Paris.
- Caelli, T. (1995). *A brief overview of texture processing in machine vision*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 79-87.

- Calvin, W.H., (1999). *Columns and modules*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, 148-150.
- Cavanach, P., (1999). *Pictorial art and vision*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, 648-650.
- Chase, V.D., (1999). *Seeing is believing*, Technology Review 1999, <http://www.web.mit.edu/resources/publications>
- Chechik, G., Meilijson, I., Ruppim, E., (1999). *Neuronal regulation: a mechanism for synaptic pruning during brain maturation*, Neural Computation 11, 2061-2080.
- Churchland, P.S., Grush, R., (1999). *Computation and the brain*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, 155-157.
- Churchland, P.S., Sejnowski, T.J., (1992). *The computational brain*, The MIT Press, Cambridge.
- Cohen, D., (1987). *Magneto-electric stimulation of the brain*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 600-601.
- Cook, J., Murray, J.D., (1995). *Pattern formation, biological*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 705-710.
- Crivianu-Gai, D., Miclea, F., Holban, S., Lungeanu, D., Hermann, I., (1998). *Automatic classification of prostate ultrasound images using neural networks*, Current perspectives in healthcare computing, Harrogate 23-25 March 1998 (B. Richards, H. Wood, Eds), Weybridge: BJHC Ltd, UK, 122-126.
- Crivianu-Gai, D., Miclea, Fl., Lungeanu, D., (1997). *Choosing the modality of 3D representation*, Vol. rez. Conferina Naional de Informatic Medical, Satu-Mare, 22-24 oct., 11-12.
- Cruz, J.M., Pajares, G., Aranda, J., (1995). *A neural network model in stereovision matching*, Neural Networks, Vol. 8, No. 5, 805-813.
- Cynader, M., (1987). *Visual cortex*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1267-1271.
- Daugman, J., Downing, C., (1995). *Gabor wavelets for statistical pattern recognition*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 414-420.
- Dayan, P., (1998). *A hierarchical model of binocular rivalry*, Neural Computation 10, 1119-1135.
- DeAngelis, G.C., Ohzawa, I., Freeman, R.D., (1995). Receptive-field dynamics in the central visual pathways, TINS, Vo.18, No. 10, 451-458.
- Desimone, R., Chelazzi, L., Miller, E.K., Duncan, J., (1995). *Neuronal mechanisms of visual attention*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 219-226.
- Despopoulos, A., Silbernagl, S., (1991). *Color atlas of physiology*, Georg Thieme Verlag, Stuttgart.
- Dimitrov, A., Cowan, J.D., (1998). *Spatial decorrelation in orientation-selective cortical cells*, Neural Computation 10, 1779-1795.
- Edelman, S., Weiss, Y, (1995). *Vision: hyperacuity*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 1009-1012.
- Engel, A.K., König, P., Kreiter, A.K., Schillen, T.B., Singer, W., *Temporal coding in the visual cortex: new vistas on integration in the nervous system*, TINS, Vol. 15, No. 6, 218-226.
- Ernst, U., Pawelzik, K., Tsodyks, M., Sejnowski, T.J., (1999). *Relation between retinotopical and orientation maps in visual cortex*, Neural Computation 11, 375-379.
- Erwin, E., Obermayer, K, Schulten, K., (1995). *Models of orientation and ocular dominance columns in the visual cortex: a critical comparison*, Neural Computation 7, 425-468.
- Fain, G.L., Hubbell, W.L., (1987). *Visual transduction*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1287-1300.
- Farah, M.J., (1998). *Why does the somatosensory homunculus have hands next to face and feet next to genitals? A hypothesis*, Neural Computation 10, 1983-1985.
- Faugeras, O., Keriven, R., (1998). *Variational principles, surface evolution, PDE's level set methods, and the stereo problem*, IEEE Trans. on Image Processing, Vol. 7, No. 3, 336-344.

- Field, D.J., (1994). *What is the goal of sensory coding?*, Neural Computation 6, 559-601.
- Field, D.J., (1995). *Visual coding, redundancy, and feature detection*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 1012-1016.
- Fine, I., Jacobs, R.A., (1999). *Modeling the combination of motion, stereo, and vergence angle cues to visual depth*, Neural Computation 11, 1297-1330.
- Finkel, L.H., Sajda, P., (1994). *Constructing visual perception*, American Scientist, Vol. 82, 224-237.
- Fohlmeister, C., Gerstner, W., Ritz, R., Van Hemmen, J.L., (1995). *Spontaneous excitations in the visual cortex: stripes, spirals, rings, and collective bursts*, Neural Computation 7, 905-914.
- Foldiak, P., Young, M.P., (1995). *Sparse coding in the primate cortex*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 895-898.
- Frisby, J.P., (1987). *Stereopsis, binocular perception*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1142-1143.
- Frisby, J.P., (1995). *Stereo correspondence and neural networks*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 937-941.
- Goltsev, A., (1996). *An assembly neural network for texture segmentation*, Neural Networks, Vol. 9, No.4, 643-653.
- Gonzalez, R.C., Woods, R.E., (1993). *Digital image processing*, Addison-Wesley, Reading, Massachusetts.
- Goodhill, G.J., (1998). *The influence of neuronal activity and intracortical connectivity on the periodicity of ocular dominance stripes*, Network: Computation in Neural Systems, Vol. 9, No. 3, 419-432.
- Gorea, A., (1995). *Visual texture*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 55-57.
- Greenfield, S., (1998). *The human brain: a guided tour*, Phoenix, Orion Books, London.
- Grossberg, S., (1993). *A solution to the figure-ground problem for biological vision*, Neural Networks, Vol. 6, 463-483.
- Grossberg, S., (1995). *Figure-ground separation*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 395-399.
- Grossberg, S., Olson, S.J., (1994). *Rules for the cortical map of ocular dominance and orientation columns*, Neural Networks, Vol. 7, No. 6/7, 883-894.
- Grossberg, S., Olson, S.J., (1998). *A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance*, Neural Networks, Vol. 11, 189-208.
- Hall, M., Robinson, D., (1998). *The human brain* (CD-ROM), The Open University, BBC & Springer Electronic Media.
- Haring, S., Viergever, M.A., (1994). *Kohonen networks for multiscale image segmentation*, Image and Vision Computing, Vol. 12, No. 6, 339-344.
- Harth, E., (1997). *From brains to neural nets to brains*, Neural Networks, Vol. 10, 1241-1255.
- Hayakawa, H., Nishida, S., Wada, Y., Kawato, M., (1994). *A computational model for shape estimation by integration of shading and edge information*, Neural Networks, Vol. 7, No. 8, 1193-1209.
- Held, R., (1987). *Visual development, infant*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1273-1275.
- Hertz, J., (1995). *Sensory coding and information theory*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 864-867.
- Hildreth, E., (1999). *Computational vision*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, 172-173.
- Hirakura, Y., Yamaguchi, Y., Shimizu, H., Nagai, S., (1996). *Dynamic linking among neural oscillators leads to flexible pattern recognition with figure-ground separation*, Neural Networks, Vol. 9, No. 2, 189-209.

- Hongo, S., Sonehara, N., Yoroizawa, I., (1996). *Edge-based binocular stereopsis algorithm - a matching mechanism with probabilistic feedback*, Neural Networks, Vol. 9, No. 3, 379-395.
- Horton, J.C., Hocking, D.R., (1996). *Pattern of ocular dominance columns in human striate cortex in strabismic amblyopia*, Visual Neuroscience 13, 787-795.
- Hotca S., Popa C., Luncan D., Lungeanu D., Holban S., (1997). *A computer application for human depth perception models: computation of stereo disparities*, Proceedings of the 4th International Conference on Computers in Medicine (E. Kacki, Ed.), Polish Society of Medical Informatics, Lodz, 151-156.
- Hsieh, Y.C., McKeown, D.M., (1992). *Performance evaluation of scene registration and stereo matching for cartographic feature extraction*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, 214-238.
- Hwang, J.N., (1995). *Textured images: modeling and segmentation*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 484-487.
- Intille, S.S., Bobick, A.F., (1994). *Incorporating intensity edges in the recovery of occlusion regions*, MIT Media Lab Perceptual Computing Group Technical Report No. 246.
- Intrator, N., Cooper, L.N., (1995). *Information theory and visual plasticity*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 864-867.
- Jackson, H., (1987). *Neural specificity*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 758-760.
- Jolion, J.M., Rosenfeld, A., (1994). *A pyramid framework for early vision*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Jones, E.G., (1987). *Brodman's areas*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 180-181.
- Jordan, M.I., Russell, S., (1999). *Computational intelligence*, The MIT Encyclopedia of the Cognitive Sciences (R.A. Wilson, F.C. Keil, Eds.), The MIT Press, Cambridge, Massachusetts, lxxiv-xc.
- Julesz, B., (1987). *Texture preception*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1200-1202.
- Julesz, B., (1995). *Some afterthoughts*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 249-257.
- Kayakawa, H., Nishida, S., Wada, Y., Kawato, M., (1994). *A computational model for shape estimation by integration of shading and edge information*, Neural Networks, Vol. 7, No. 8, 1193-1209.
- Khotanzad, A., Bokil, A., Lee, Y.W., (1993). *Stereopsis by constraint learning feed-forward neural networks*, IEEE Trans. on NN, Vol. 4, No. 2, 332-342.
- Kilis, D., Gatewood, L., Zhuo, Z., Hatfield, G., Seaholm, S., Ackerman, E., (1993). *Simulation of stochastic micropopulation models - III. COGNET: an artificial neural network for visual recognition*, Computers in Biol. and Med., Vol. 23, No. 3, 215-255.
- Kitajima, T., Hara, K.-I., (1997). *An integrated model for activity-dependent synaptic modification*, Neural Networks, Vol. 10, No. 3, 413-421.
- Kohonen, T., (1988). *Self-organization and associative memory*, 2nd Ed., Springer Verlag, Berlin.
- Kowler, E., (1995). *Attention*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 177-179.
- Kube, P., Perona, P., (1996). *Scale-space properties of quadratic feature detectors*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 10, 987-999.
- Kulikowski, J.J., Walsh, V., (1995). *Demonstration of binocular fusion of color and texture*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 27-32.
- Laar van der, P., Heskes, T., Gielen, S., (1997). *Task-dependent learning of attention*, Neural Networks, Vol. 10, 981-992.

- Lane, R.a., Thacker, N.A., Seed, N.L., (1994). *Stretch-correlation as a real-time alternative to feature-based stereo matching algorithms*, Image and Vision Computing, Vol. 12, No. 4, 203-212.
- Law, T., Itoh, H., Seki, H., (1996). *Image filtering, edge detection, and edge tracing using fuzzy reasoning*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 5, 481-491.
- Lee, T.S., (1996). *Image representation using 2D Gabor wavelets*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No. 10, 959-971.
- Leshner, G.W., Mingolla, E., (1995). *Illusory contour formation*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 481-483.
- Lewis, E.R., (1987). *Neural nets, modeling*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 756-758.
- Li, Z., (1998). *A neural model of contour integration in the primary visual cortex*, Neural Computation 10, 903-940.
- Li, Z., Atick, J.J., (1994). *Toward a theory of the striate cortex*, Neural Computation 6, 127-146.
- Liu, L., Hulle, van, M.M., (1998). *Modeling the surround of MT cells and their selectivity for surface orientation in depth specified by motion*, Neural Computation 10, 295-312.
- Ludwig, K.O., Neumann, H., Neumann, B., (1994). *Local stereoscopic depth estimation*, Image and Vision Computing, Vol. 12, No. 1, 16-35.
- Lund, J.S., Wu, Q., Levitt, J.B., (1995). *Visual cortex cell types and connections*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 1016-1021.
- Lungeanu, D., (1995). *Simulation of binocular depth perception : computing the stereo disparity*, Vol. rez. Conferinta Naional de Informatic Medical, Bucureti, 3-5 noiembrie, 222.
- Lungeanu, D., Cotirlea, V., (1999). *Modelarea perceptiei contururilor in vederea umana primara*, Vol. rez. Zilele Academice Timisene (Medicina si Farmacie), Timisoara, 27-28 mai, 12.
- Lungeanu, D., Cotirlea, V., Hanga, L. (2000). *From primary edges to contours: modelling contour integration in human early vision*, Lucrare acceptata la Conferinta Europeana de Informatica Medicala MIE2000 (august 2000), Hanovra, Germania.
- Lungeanu, D., Mihalas, G.I., (1995). *Simularea perceptiei de adâncime în vederea binocular*, comunicat la Zilele Academice Timiene (Medicina si Farmacie), Timisoara, mai 1995.
- Lungeanu, D., Popa, C., Hotca, S., (1998). *From images to surfaces in human early vision - a computational perspective*, A XV-a Conf. Nat. de Fiziologie, Arad 25-26 sept. 1998; rez. publ. in *Fiziologia*, Vol. 8, Nr. 3(19), 32.
- Lungeanu, D., Popa, C., Hotca, S., Luncan, D., Holban, S., (1996). *Modelling human depth perception: computing of binocular horizontal disparity*, Joint Meeting Hungarian Phys. Soc.& Romanian Soc. Phys. Scien., Szeged-Timisoara 30 iunie - 4 iulie; rez. publ. in *Fiziologia*, Vol. 6, Nr. 2(10), 30.
- Lungeanu, D., Popa, C., Hotca, S., Luncan, D., Holban, S., (1996b). *Modelling the depth perception in stereopsis: models for the stereo horizontal disparities computation*, Vol. rez. Conferinta Naional de Informatic Medical, Cluj, 6-9 noiembrie, 52-53.
- Lungeanu, D., Popa, C., Hotca, S., Macovievici, G., (1998b). *Modelling biological depth perception in binocular vision: the local disparity estimation*, Medical Informatics (Taylor&Francis), Vol. 23, No. 2, 131-143.
- Lungeanu, D., Popa, C., Hotca, S., Macovievici, G., (1998c). *Modelling human depth perception in binocular vision: obtaining the horizontal disparity map*", Proceedings of the 9th World Congress on Medical Informatics MEDINFO'98 (B. Cesnik, A.T. McCray, J.R. Sherrer, Eds.), IOS Press, Amsterdam: Netherlands, 380-384.
- Malsburg, von der, C., (1995). *Self-organizing and the brain*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 840-843.

- Marr, D., (1974). *A note on the computation of binocular disparity in a symbolic low-level visual processor*, MIT A.I. Lab. Memo 327. 1991, From the retina to the neocortex: selected papers of David Marr (L.M. Vaina, Ed.), Birkhauser, Boston, 231-238.
- Marr, D., (1982). *Vision. A computational investigation into the human representation and processing of visual information*, W.H. Freeman, New York (13th printing 1996).
- Marr, D., Pal, D., Poggio, T., (1978). *Analysis of a cooperative stereo algorithm*, Biol. Cybernetics, 28, 223-236. 1991, From the retina to the neocortex: selected papers of David Marr (L.M. Vaina, Ed.), Birkhauser, Boston, 245-262.
- Marr, D., Poggio, T., (1976). *Cooperative computation of stereo disparity*, Science, Vol. 194, 283-287. 1991, From the retina to the neocortex: selected papers of David Marr (L.M. Vaina, Ed.), Birkhauser, Boston, 239-244.
- Marr, D., Poggio, T., (1979). *A computational theory of human stereo vision*, Proc. Roy. Soc. London, B. 204, 301-328. 1991, From the retina to the neocortex: selected papers of David Marr (L.M. Vaina, Ed.), Birkhauser, Boston, 263-290.
- Masters, B.R. (Ed.), (1990). *Noninvasive diagnostic techniques in ophtalmology*, Springer-Verlag, New-York.
- Matalas, I., Benjamin, R., Kitney, R., (1997). *An edge detection technique using the facet model and parameterized relaxation labelling*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 4, 328-341.
- Mesulam, M.M., (1998). *From sensation to cognition*, Brain 121, 1013-1052.
- Meunier, C., Nadal, C. (1995). *Sparsely coded neural networks*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 899-901.
- Miclea, M., (1999). *Psihologie cognitiva. Modele teoretico-experimentale* (Ed. a 2-a), Polirom, Iasi.
- Mihalas, G. I., Lungeanu, D., Macovievici, G., Gruia, R., Neagu, M., (1999). *SIM – A computer program for simulation of regulatory processes in biosystems*, XIII International Biophysics Congress Abstracts, Indian Academy of Sciences, Bangalore, India, 172.
- Mihalas, G. I., Macovievici, G., Simon, Z., Mihalas, G., Balea, G., Lungeanu, D., (1999b). *MOBISIM-Package for simulation in molecular biology*, Medical Informatics Europe '99 (P. Kokol et al. Eds.), IOS Press, Amsterdam, 617-620.
- Mihalas, G., Lungeanu, D., Macovievici, G., Vernic, C., Gruia, R., (1999c). *Modulul de diagrama a fazelor pentru programele de simulare a proceselor biologice*, Vol. rez. Zilele Academice Timisene (Medicina si Farmacie), Timisoara, 27-28 mai, 11.
- Mihala, G.I., Lungeanu, D., Kigyosi, A., Vernic, C., (1995). *Criteria for classification of computer simulation programs*, Proceedings of the 8th World Congress on Medical Informatics MEDINFO'95 (R.A. Greenes, M.E. Peterson, D.J. Protti, Eds.), IOS Press, Amsterdam: North-Holland, 1209-1213.
- Mihalas, G.I., Lungeanu, D., Kigyosi, A., Vernic, C., eicu C., Gherasim, M., (1995b). *Criteria de clasificare a programelor de simulare utilizate în învmântul medical*, Timioara Medical, (ediie omagial 1945-1995), 30-33.
- Miller, D.A., Zucker, S.W., (1999). *Computing with self-excitatory cliques: a model and an application to hyperacuity-scale computation in visual cortex*, Neural Computation 11, 21-66.
- Miller, K.D., (1995). *Ocular dominance and orientation columns*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 660-665.
- Miller, K.D., (1998). *Equivalence of a sprouting-and-retraction model and correlation-based plasticity models of neural development*, Neural Computation 10, 529-547.
- Mitchison, G., Swindale, N.V., (1999). *Can Hebbian volume learning explain discontinuities in cortical maps?*, Neural Computation 11, 1519-1526.
- Moellenhoff, M.S., Maier, M.W., (1998). *Transform coding of stereo image residuals*, IEEE Trans. on Image Processing, Vol. 7, No. 6, 804-812.

- Nagy, I.I., Mihalas, G. I., Neagu, M., Balaceanu, L., Neagu, A., (1999). *Curs de biofizica*, Eurobit, Timisoara.
- Nakayama, K., He, Z.J., (1995). *Attention to surfaces: beyond a Cartesian understanding of focal attention*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 181-188.
- Nasrabadi, N.M., Choo, C.Y., (1992). *Hopfield network for stereo vision correspondence*, IEEE Trans. on NN, Vol. 3, No. 1, 5-13.
- Nelson, J.I., (1995). *Binding in the visual system*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 157-159.
- Neumann, H., (1996). *Mechanisms of neural architecture for visual contrast and brightness perception*, Neural Networks, Vol. 9, 921-936.
- Normann, R.A., (1995). *Visual neuroprosthetics – functional vision for the blind*, IEEE Engineering in Medicine and Biology, 77-83.
- Ögmen, H., (1993). *A neural theory of retino-cortical dynamics*, Neural Networks, Vol. 6, 245-273.
- Okajima, K., (1996). *A model visual cortex incorporating intrinsic horizontal neuronal connections*, Neural Networks, Vol. 9, No. 2, 211-222.
- Overington, I., (1992). *Computer vision. A unified, biologically-inspired approach*, Elsevier, Amsterdam.
- Papathomas, T.V., (1995). *Binocular vision and stereopsis*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 1-3.
- Papathomas, T.V., Chubb, C., Gorea, A., Kowler, E., (1995). *Early vision and beyond*, The MIT Press, Cambridge, Massachusetts.
- Papilinski, A.P., (1998). *Directional filtering in edge detection*, IEEE Trans. on Image Processing, Vol. 7, No. 4, 611-615.
- Pentland, A.P., (1987). *A new sense for depth of field*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-9, No.4, 523-531.
- Perkel, D.H., (1987). *Neural coding*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 751-752.
- Pescovitz, D., (1999). *Getting real in cyberspace*, Scientific American, Vol. 10, No. 3, 48-51.
- Peters, A., (1987). *Cortical neurons*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 281-284.
- Pollen, D.A., Gaska, J.P., (1987). *Vision, visual cortex, and frequency analysis*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1262-1263.
- Popa, C., Hotca, S., Lungeanu, D., Luncan, D., Holban, S., (1996). *STEREO-VISION: a computer application for implementing and testing the models of depth perception in stereopsis*, Vol. rez. Conferinta Nationala de Informatic Medical, Cluj, 6-9 noiembrie, 54-55.
- Postma, E.O., Herik van der, H.J., Hudson, P.T.W., (1997). *SCAN: a scalable model of attentional selection*, Neural Networks, Vol. 10, 993-1015.
- Qian, N., (1994). *Computing stereo disparity and motion with known binocular cell properties*, Neural Computation 6, 390-404.
- Qian, N., Mikaelian, S., (2000). *Relationship between phase and energy methods for disparity computation*, Neural Computation 12, 279-292.
- Ramachandran, V.S., (1995). *Perceptual correlates of neural plasticity in the adult human brain*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 127-147.
- Riesenhuber, M., Bauer, H.U., Brockmann, D., Geisel, T., (1998). *Breaking rotational symmetry in a self-organising map model for orientation map development*, Neural Computation 10, 717-730.
- Ritter, H., (1995). *Self-organizing feature maps: Kohonen maps*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 846-851.

- Robinson, D. (Ed.), (1998). *Neurobiology (Biology: Brain & Behaviour series)*, Springer-Verlag & The Open University, Berlin.
- Rolls, E.T., (1997). *Consciousness in neural networks?*, Neural Networks, Vol. 10, 1227-1240.
- Rumelhart, D.E., McClelland, J.L., and the PDP Research Group, (1986). *Parallel distributed processing: explorations in the microstructure of cognition (I+II)*, The MIT Press, Cambridge, Massachusetts.
- Sagawa, Y., (1997). *A function of neural networks generating end- and side-inhibition in the primary visual cortex*, Neural Networks, Vol. 10, No. 3, 423-429.
- Sagi, D., (1995). *The psychophysics of texture segmentation*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 70-78.
- Saint-Marc, P., Chen, Y.S., Medioni, G., (1991). *Adaptive smoothing: a general tool for early vision*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 13, No.6, 514-529.
- Salinas, E., Abbott, L.F., (2000). *Do simple cells in primary visual cortex form a tight frame?*, Neural Computation 12, 313-335.
- Sekuler, R., Blake, R., (1990). *Perception*, 2nd Ed., Mc-Graw-Hill, New-York.
- Shepherd, G.M., (1988). *Neurobiology*, 2nd Ed., Oxford University Press, New-York.
- Shiller, P.H., (1987). *Striate cortex*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1148-1149.
- Singer, W., (1995). *Synchronization of neural responses as a putative binding mechanism*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 960-964.
- Siper, B.J., Klarquist, W.N., (1997). *Patched-based stereo in a general binocular viewing geometry*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No.3, 247-253.
- Sokoloff, L., (1987). *Mapping cerebral functional activity with radioactive deoxyglucose*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 604-609.
- Solomon, F., Ikeuchi, K., (1996). *Extracting the shape and roughness of specular lobe objects using four light photometric stereo*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 18, No.4, 449-454.
- Stavenga, D.G., (1992). *Eye regionalization and spectral tuning of retinal pigments in insects*, TINS, Vol. 15, No. 6, 213-218.
- Taylor, J.G., (1997). *Neural networks for consciousness*, Neural Networks, Vol. 10, 1207-1225.
- Thorpe, S., (1995). *Localized versus distributed representations*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 549-552.
- Tittle, J.S., Todd, J.T., (1995). *Perception of three-dimensional structure*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 715-718.
- Tsao, T.R., Chen, V., (1994). *A neural scheme for optical flow computation based on Gabor filters and generalized gradient method*, Neurocomputing 6, 305-325.
- Tsotsos, J.K., (1987). *Analyzing vision at the complexity level: constraints on an architecture, an explanation for visual search performance, and computational justification for attentive processes*, Technical Reports on Research in Biological and Computational Vision at the University of Toronto, RCBV-TR-87-20.
- Tsotsos, J.K., (1995). *Toward a computational model of visual attention*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 207-218.
- Tyler, C.W., (1995). *Cyclopean riches: cooperativity, neuroentropy, hysteresis, stereoattention, hyperglobality, and hypercyclopean processes in random-dot stereopsis*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 5-15.
- Ullman, S., (1987). *Visual motion perception*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1280-1281.

- Verschure, P.F.M.J., König, P., (1999), *On the role of the biophysical properties of cortical neurons in binding and segmentation of visual scenes*, Neural Computation 11, 1113-1138.
- Vidyasagar, T.R., Pei, X., Vogushev, M., (1996). *Multiple mechanisms underlying the orientation selectivity of visual cortical neurones*, TINS, Vol. 19, No. 7, 272-277.
- Watt, R.J., (1995). *Some speculations on the role of texture processing in visual perception*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 59-67.
- Weinshall, D., Malik, J., (1995). *Review of computational models of stereopsis*, Early vision and beyond (T.V. Papathomas, C. Chubb, A. Gorea, E. Kowler, Eds.), The MIT Press, Cambridge, 33-41.
- Weintraub, D.J., Walker, E.L., (1966). *Perception*, Wadsworth, Brooks/Cole Publishing Co., Belmont, California
- Weng, J., Ahuja, N., Huang, T.S., (1992). *Matching two perspective views*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 8, 806-825.
- Wilson, R.A., Keil, F.C., (Eds.), (1999). *The MIT Encyclopedia of the Cognitive Sciences*, The MIT Press, Cambridge, Massachusetts.
- Wiskott, L., Sejnowski, T., (1998). *Constrained optimization for neural map formation: a unified framework for weight growth and normalization*, Neural Computation 10, 671-716.
- Wolfe, J.M., (1987). *Visual aftereffects*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1266-1267.
- Wolfe, J.M., (1987b). *Visual perception*, Encyclopedia of Neuroscience (G. Adelman, Ed.), Birkhauser, Boston, 1282-1283.
- Wörgötter, F., Cozzi, A., Gerdes, V., (1999). *A parallel noise-robust algorithm to recover depth information from radial flow fields*, Neural Computation 11, 381-416.
- Yuille, A.L., (1995). *Constrained optimization and the elastic net*, Handbook of Brain Theory & NN (M.A. Arbib, Ed.), The MIT Press, Cambridge, 250-255.
- Yuille, A.L., Kolodny, J.A., Lee, C.W., (1996). *Dimension reduction, generalized deformable models and the development of ocularity and orientation*, Neural Networks, Vol. 9, No. 2, 309-312.

96 p.
tab