

SZEMLE

FISER JÓZSEF

Budapesti Műszaki Egyetem

ALAKPERCEPCIÓ: NEUROFIZIOLÓGIAI, PSZICHOFIZIKAI EREDMÉNYEK ÉS HATÁSAIK A GÉPI LÁTÁSKUTATÁSRA

1. BEVEZETÉS

Jelen tanulmány célja, hogy tömören áttekintse azokat az eredményeket, melyek a látás neurofiziológiájának és pszichológiájának kutatása során az elmúlt 15 évben összegyűltek, valamint hogy rámutasson, hogyan lehet ezeket a tényeket felhasználni hatékonyabb számítógépes látórendszerek kifejlesztésére. Az áttekintés vázlatos, és nem terjed ki minden területre. Céлом azoknak az eredményeknek az összegyűjtése volt, melyek egyrészt több forrásból származnak, másrészt valamennyire egységes szemléletmód-ba ötvözhetőek. A végső cél nem az emberi látás tökéletes modellezése, hanem a gépi látáskutatás segítése új gondolatokkal. Emiatt néhány tény felhasználása talán egyszerűsített lesz. Remélem, nem túl egyszerűsített.

2. NEUROFIZIOLÓGIAI EREDMÉNYEK

2.1. A CORTEX SZERKEZETE

Az evolúció során a primitív kérgek visszafejlődtek, és a neocortex (a továbbiakban cortex) lett az agy abszolút domináns része. Ez kezeli az asszociációkat, a szenzoros-kiértékelő „ösvényeket” (paths), a látóösvényt is beleértve. A cortex struktúrája annak teljes területén ugyanaz, a felszínen párhuzamos 6 sejtrétegből áll, melyekre a felszíntől lefelé növekvő (I—VI.) számokkal hivatkoznak.

A sejtek alapvetően két fő csoportba oszthatók: piramis- és szemcse-(stellate)-sejtekre (CHANGEUX, 1986). A piramis sejtek fenyőfa formájú sejtek, csúcsukkal — a csúcsi vagy apikális dendritjeikkel — a felszín felé mutatnak, sok bazális dendritjük ágazik ki a legvastagabb részükből, és axonjuk a fehérállományba fut. A szemcse sejtek erősebben körszimmetri-

1. ábra

A cortexben található főbb sejtípusok

A két szaggatott vonal között a IV. réteg, felette az I-II-III. rétegek, alatta az V-VI. rétegek láthatóak. A piramis sejtek A-tól G-ig kúpos sejtestestükről, a függőlegesen a felszín felé tartó apikális dendritjeiről és gyökérszerű basal-dendritjeikről ismerhetők fel. Az axon (a) a cortex mély rétegei felé ereszkedik, majd végül távozik a cortex-ből. A stellate-sejtek (H-től M-ig) gyakran kapnak olyan neveket amelyek utalnak fa-formájuk különbözőségére (H — "dupla csokor" sejt, L — "rövid axonú" sejt stb). (Changeux 1986).



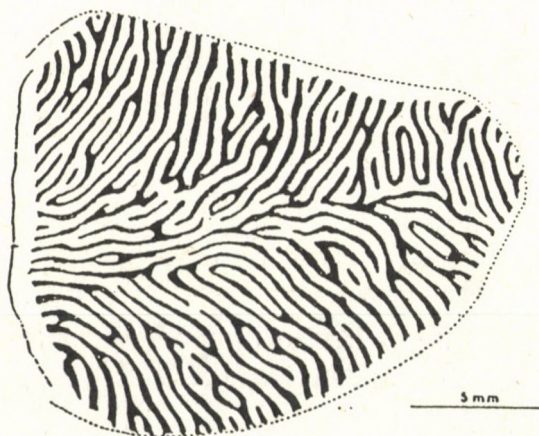
kusak, ha van egyáltalán valamilyen fő orientációjuk, az párhuzamos a rétegekkel, axonjaik sose mennek a fehérállományba. Ezek a sejtek a piramis-sejtek közötti kommunikációt segítik elő. Az első rétegben nincsenek piramis-sejtek, viszont a II—VI. rétegekben fokozatosan egyre sűrűbben helyezkednek el. Szemcsesejtek csak a IV. rétegben találhatók. A látó-cortexben 3 sűrű sávot képeznek párhuzamosan a felszínnel. Az 1. ábrán tipikus piramis- és szemcsesejtek láthatóak.

A szenzoros ösvények a IV. rétegnél csatlakoznak be, ez a cortex igazi bejárata. A kimenetek a VI., V., illetve a II., III. rétegekből indulnak a thalamus, a colliculus superior és egyéb célpontok felé.

2. ábra

Fent: A felülettől párhuzamosan elmetezett látókéreg képe felülnézetből egy makákó majomnál

Lent: Egy ujjlenyomat képe megfelelő módon kicsinyítve (Changeux 1986).



Ugyanekkor a cortexnek oszlopos szerveződése is van. Ha a felszínen egy ponton ingereljük a cortexet, merőlegesen az összes sejt aktiválódik. Ezek nem valódi oszlopok, hanem például a látókéregben a felszín felől nézve egy „ujjlenyomat”-struktúrát mutatnak (2. ábra), ahol a fekete részek az egyik szemhez tartoznak, míg a fehérek a másikhoz. Összehasonlításképpen egy valódi ujjlenyomat képe is látható az ábrán, megfelelően kicsinyítve.

2.2. A LÁTÓKÉREG FELOSZTÁSA

A látókéreg most következő leírása alapvetően makákó majmok cortexének vizsgálataiból származik Van ESSEN (1985) alapján, de az eredmények nagy valószínűséggel igazak az emberi cortexre is.

Az elmúlt 15 évben az elmélet, miszerint a látókéreg elkülöníthető részekre (ún. areákra) osztható, véglegesen bizonyítottá vált. Ezen areák közül körülbelül huszat azonosítottak, és még néhány új azonosítás várható. A cortex 3. ábrán szereplő kiterített ábrázolásmódjában látható az areák relatív aránya. A cortex felosztása az alábbi módszerek egyikének vagy néhányának felhasználásából jött: (1) strukturális vizsgálat, (2) kapcsolódási viszonyok vizsgálata, (3) topografikus feltérképezés, (4) fiziológiai válaszfüggvények különböző ingerekre, (5) agyi sérülések hatása a viselkedésre.

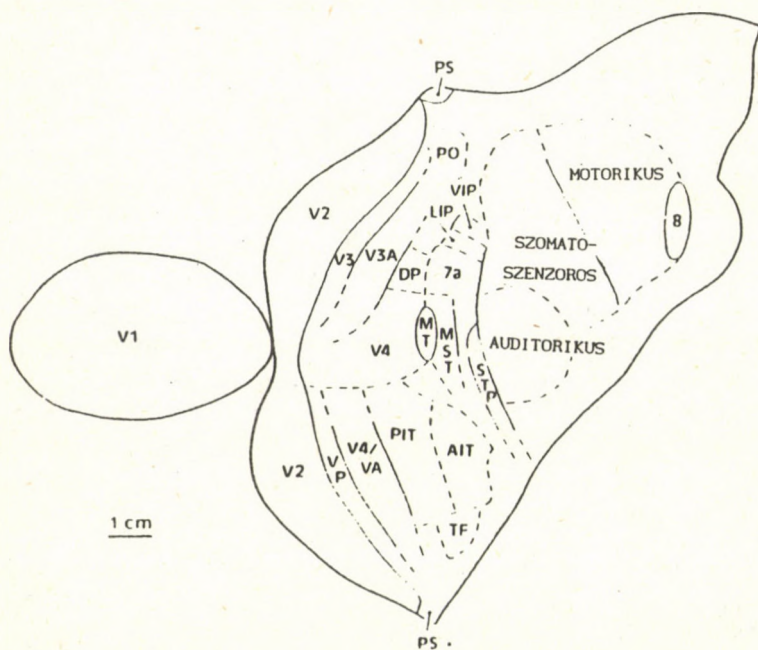
Az areák tulajdonságai erősen változóak, de két általános tendencia megfigyelhető, amint egyre feljebb haladunk a látópályán (a 3. ábrán ez durván a balról jobbra haladással ekvivalens): a receptív mezők mérete növekszik, a receptív mezők széleinek élessége és topografikus szervezettségük viszont fokozatosan csökken. A következőkben nagyon tömören végigvesszük a főbb areákat. Csupán a közöttük levő legfőbb kapcsolatokat emeljük ki.

A három legjobban azonosított és megvizsgált area a V1 (elsődleges látókéreg v. striátum), a V2 (másodlagos látókéreg) és az MT. V1 az első area a látókéregben, inputját az oldalsó kérgestestből (corpus geniculare laterale v. CGL) kapja. Ebbe az areába a vizuális látvány teljesen hű mása képződik le. Ez a legnagyobb area, és receptív mezőinek tulajdonságai jól definiáltak.

V2 hosszú közös határral bír V1-gyel. Vizuális reprezentációja tükörképe a V1-ben levőnek, kivéve, hogy középen egy hosszú törésvonal húzódik benne. Egyenetlen foltokban kapja inputját V1-ből, struktúrája heterogén, de valamennyire topografikusan szervezett.

Egy jobb félteke teljes cortexének kiterített képe

A látókéreg a bal oldalon van, és több mint felét elfoglalja a cortexnek (Maunsell and Newsome 1987).



MT — apró area sok irányszelektív sejttel. V1-gyel áll kapcsolatban.

V3 és VP abszolút szimmetrikus helyzetben van, mégis jelentősen eltér egymástól. A dorsalis (hátsó) V1 V2-be, V3-ba és MT-be vetít, de a ventralis (hasi oldali) V1 csak a ventralis V2-be és V3-ba, de VP-be nem. VP-ben háromszor annyi színszelektív sejt van, mint V3-ban, viszont V3-ban háromszor annyi irányszelektív sejt, mint VP-ben. Végül VP-ben kétszer akkora a receptív mezők (RM-ek), mint V3-ban.

V4 fő inputját V2-ből kapja. A VA—V4 gyengén felderített area, lehet hogy V4 része, de az is lehet, hogy nem.

V3A, PO és DPL fő inputját V1-ből kapja. Ezek kisméretű areák, és komplex topografikus szerveződésűek.

Az MST és VIP fő inputját MT-ből kapja. Az MST-ben néha nagyon nagy RM-ek vannak mindenféle topografikus szerveződés nélkül. Megesik, hogy néha egy ilyen RM inputját az MT-nek teljesen elkülönülő foltjaiból kapja.

A PIT inputját V4-ből, míg az AIT inputját PIT-ből kapja. Ebben a két areában további felosztások valószínűek.

A fennmaradó areák — 8, 7a, LIP, STP, TF, PS — vizuomotoros és poliszenzoros areák. Válaszaikban a vizuális input szerepet játszik, de a legkülönbébb egyéb területről is kapnak inputot.

Befejezésül két tényt érdemes megemlíteni. Az első, hogy a makákó majmok areáinak felderítése még nem lezárt, vagyis néhány újabb area felfedezése várható. A másik, hogy az emberi cortex a maga agyféltekénkénti 800 cm^2 -es felületével mintegy tízszer nagyobb egy majoménál. Habár a cortex felülete nincs lineáris arányban az areák számával, mégis az ember esetében sok további area létezése valószínű.

2.3. A VIZUÁLIS AREÁK KAPCSOLATA

Jelen pillanatban mintegy 100 bizonyított kapcsolatot tártak fel a vizuális areák között. Ezekre a kapcsolatokra egy majdnem univerzális törvény igaz, a reciprok párok törvénye. Ha egy area inputot szolgáltat egy másiknak, akkor abból az areából inputot is kap. Emellett feltűnő aszimmetria is megfigyelhető a kapcsolatokban. Az egyik irányba a kapcsolatok majdnem kizárólagosan a felszíni rétegekből indulnak, és legsűrűbben a IV. rétegben landolnak. Az ellenkező irányban viszont a mély rétegekből indulnak, és többnyire a IV. rétegen kívül landolnak. Az általános elképzelés szerint ezek a kapcsolatok az előre irányuló („felszálló”) és a feedback („leszálló”) információáramlásokat tükrözik. A 4. ábra a kapcsolatok gyors áttekintésére ad lehetőséget. A kevert felszálló és leszálló kapcsolatok közbülső tulajdonságokat mutató csatolásokat jelölnek, vagyis ezek a kapcsolatok néha sem el nem kerülnek, sem meg nem célozzák kizárólagosan a IV. réteget.

A legalaposabban tanulmányozott areákban (V1, V2, V3, V4, VP, MT) a kapcsolatok átlagos száma 6 főbb és 2 jelentéktlenebb pár. Ha a magasabb areák nem bonyolultabbak, mintegy 100 főbb és néhány tucat jelentéktlenebb kapcsolatpár létezése feltételezhető a majmok esetében.

A 4. ábrán az areák oly módon vannak felsorolva, hogy a diagonál felett a nyilak mind felfelé, alatta lefelé mutatnak. Ez a rendezés lehetővé

Az areák közötti kapcsolatok táblázata

Magyarázat: nyíl fel= felszálló kapcsolat; nyíl le= leszálló kapcsolat; nagy nyíl=erősebb; kis nyíl= jelentéktelenebb; dupla nyíl= kevert felszálló és leszálló; ?= kétséges; += meg nem határozott (Van Essen 1985).

HOVÁ:

	V1	V2	V3	VP	V3A	V4	VA/V4	MT	PO	VIP	7a	MST	LIP	DPL	PIT	TF	AIT	8	STP
HONNAN:	V1	— ↑	↑		↑	↑		↑	+										
	V2	↑ —		↑	↑	↑		↑	+	↑		+							
	V3	↑	↑ —		↑	↑		↑	+				?						
	VP		↑	—	↑	↑		↑	+	↑		↑		↑		↑			
	V3A		↑	↑	—	↑			?								?		
	V4	↑	↑	↑	↑	↑ —		↑				↑	↑	↑	↑	↑	?	+	
	VA/V4		↑	↑			—								↑			+	
	MT	↑	↑	↑	↑	↑	↑	—	?	↑		↑						?	
	PO								—	?									
	VIP			+				↑	?	—									
	7a										—	↑?				↑?			?
	MST	+	↑	↑	↑	↑		↑			↑	—						+	
	LIP		?										—						
	DPL													—					
	PIT					↑									—	+	↑		
	TF			↑	↑	↑				↑?						—			
	AIT					?									↑		—		
	8					↑		?				↑?						—	
	STP																		—

teszi egy, csak a kapcsolódásokon alapuló hierarchia felállítását. Ebben a hierarchiában minden area olyan pozícióban van, hogy felszálló kapcsolódást csak tőle lejjebb helyezkedő, leszálló kapcsolódást csak felette álló areán az ábrán 7 jól definiált area (téglalapok) és 10 kevésbé felderített area (ellipszisek) van egy hétszintű rendszerbe szervezve, 33 főbb kapcsolattal, melyből 28 bizonyíthatóan és további 5 feltételezhetően reciprok. A jelentéktelenebb vagy inkonzisztens kapcsolatok többsége nincs ábrázolva. Ez a hierarchia teljesen mellőzi a cortex és a kéreg alatti struktúrák (colliculus superior, pulvinar stb.) közötti kapcsolatokat. Ezeken a pályákon jelentős vizuális és nem vizuális információ folyik, de funkciójuk eléggé homályos.

2.4. A KÉRGAI KAPCSOLATOK FUNKCIONÁLIS RENDSZERE

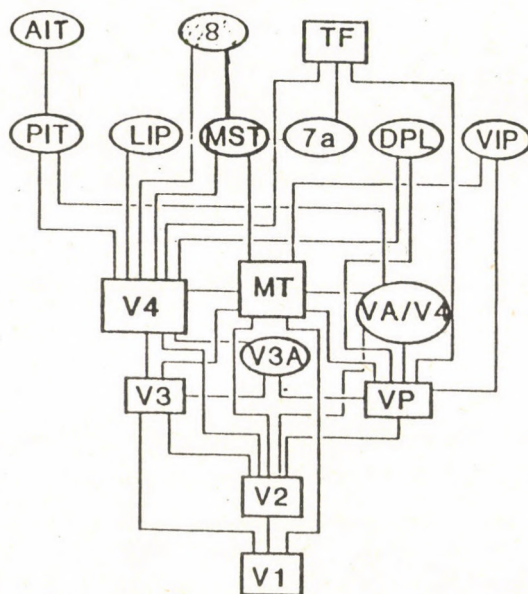
Mostanában vált világossá, hogy a látás folyamata tovább strukturálható, ha további tényeket is felhasználunk. A fő állítás az, hogy különböző soros kapcsolódású „pályák” léteznek, melyek meglehetősen függetlenül haladnak és működnek egymással párhuzamosan a cortexben (MAUNSELL

és NEWSOME, 1987; ZEKEI és SHIPP, 1988). Sok tény támasztja alá azt az elképzelést, hogy két fő áramlat van a látókéregben: az első a fali lebenybe (parietal lobe) vezet és a térbeli elrendeződések, viszonyok vizuális becslésében és a mozgásban játszik szerepet; a másik a halántéki lebenybe (temporal lobe) vezet, és a forma-, illetve színanalízisben játszik szerepet. Ez a felosztás természetesen nem kizárólagos két okból sem. Egyrészt mert a mozgás- és szín-forma pályák valószínűleg csak egy részét reprezentálják az információnak, ami a fali és halántéki lebenyeket eléri, és így valószínűtlen, hogy meghatároznák az összes vizuális reakciót, melyet ezek a régiók végeznek. Másrészt mert például a mozgáspálya lehet, hogy nem szigorúan csak mozgásanalizálással foglalkozik, és lehet, hogy nem is dolgozza fel a mozgás minden aspektusát.

Gyanítható, hogy a két pálya szétválasztódása már rögtön a retinában megkezdődik. A retina ganglion sejtjei durván két osztályba sorolhatóak. Ezek elkülönülve a CGL két különböző rétegébe kapcsolódnak, és onnan a V1 két különböző részébe vetítenek. Az CGL két rétegében a magno-

5. ábra

A makákó majom vizuális areáinak hierarchiája a cortexen belüli kapcsolatok jellegére alapozva (Van Essen 1985)



celluláris és parvocelluláris rétegekben a sejtek (legyenek M-sejtek, illetve P-sejtek) hasonlóak a retinában találhatóakhoz. A P-sejtek hullámhossz-szelektívek, és lassú, erősödő választ adnak. Az M-sejtek nem hullámhossz-szelektívek, gyorsabb tranziens választ adnak, és nagy RM-jük van. Másként fogalmazva a P-sejtek finom térbeli és durva időbeli felbontásúak, míg az M-sejtek finom időbeli és durva térbeli felbontásúak. Az 1. táblázat a két típus közötti különbségeket mutatja.

1. táblázat

A nagysejtes és kis sejtes geniculate részek fiziológiai tulajdonságai (Livingstone és Hubel, 1987)

Tulajdonságok	Parvocelluláris CGL rész	Magnocelluláris CGL rész
Színérzékenység	Van (szín-opponens s.)	Nincs (szélessávú s.)
Kontraszt érzékenység	Alacsony (küszöb > 10%)	Magas (küszöb < 2%)
Térbeli felbontás	Nagy	Kicsi (fele, harmada a másíknak egy adott távra a nézés középpontjától)
Időbeli felbontás	Lassú (kitartott válasz, kis vezetési sebesség)	Gyors (tranziens válasz, nagy vezetési sebesség)

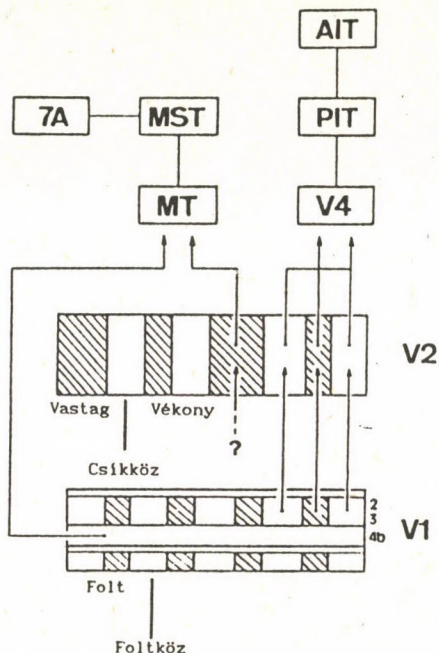
Egy bizonyos festékanyagot (mitochondriális enzim cytochrom oxidáz) használva a V1-ben szabályos struktúrákat, sötét foltokat és világos foltközöket választhatunk szét. A foltközök az area 90%-át foglalják el. A cortexen belül a foltok elsősorban foltokkal, foltközök elsősorban foltközökkel létesítenek kapcsolatot. Ugyanazt a festékanyagot használva a V2-ben vastag és vékony sötét csíkok és világos csíkközök választhatók szét. A V1 foltjaiban sok hullámhossz-szelektív sejt és kevés orientáció-szelektív sejt van, így nagyon alkalmas a színlátásra. A foltközökben a sejtek erősebben orientáció-szelektívek, kiválóan reagálnak a kontúrok által keltett hullámhossz- és fényességkülönbségekre. A helyzet rendkívül hasonlít a V2 sötét csíkjaira és világos csíkközeire.

Az CGL P-sejtek V1 IVc β rétegébe kapcsolódnak (a rétegsorszám utáni betűk az adott réteg belső tovább-bontásából erednek), mely erős kapcsolatban áll a V1 felszíni rétegeivel (II. és III.). Az M-sejtek V1 IVc α -ba kapcsolódnak, ami V1 IVb-vel van erős kapcsolatban. V1-gyel kezdve a struktúra egyszerűsített ábrázolása látható a 6. ábrán.

A két pálya szétválása evidens V1-ben, ahol a szín—forma pálya a II. és III. réteg folt- és foltközrégióiból indul. Ezek a régiók V2 vékony csíkjaira és csíkközeibe kapcsolódnak, melyek tovább vetítenek V4-be. V4

A mozgáspálya és a szín—forma pálya főbb komponensei

A rajz sematikusan ábrázolja V1 rétegeit (2, 3, 4b), foltjait (blob) és foltközöket (interblob), V2 vastag (thick) és vékony (thin) sötét csíkjaikat, valamint csíkközöket (interstripe) és a két pálya magasabb areáit (Maunsell and Newsome 1987).



outputja PIT-be, majd AIT-be vezet. A mozgáspálya V1 IVb-ből indul, és egyenesen MT-be kapcsolódik. MT V2 vastag csíkjaiból is fogad inputot, amelyek inputja viszont nem azonosított még. Hubel és Livingston nemrégiben (1987a) publikálták feltevésüket, miszerint a vastag csíkok — melyek inputjánál ? áll a 6. ábrán — a V1 IVb rétegéből kapják inputjukat mókusmajmok esetén. Így egy részletesebb felépítés a két teljes pályáról a 7. ábrán látható.

2.5. A KÉT PÁLYA ÁLTALÁNOS TULAJDONSÁGAI

A két pálya tulajdonságainak hasonlósága arra mutat, hogy mindkettőnél azonos számítási stratégia érvényesül. Mostanra a prestriátrális areák szerepe tisztázódott. A szín—forma pálya csak néhányat használ ezekből az areákból, a mozgáspálya sokat. A mozgáspályát szélesebb körben vizsgálták.

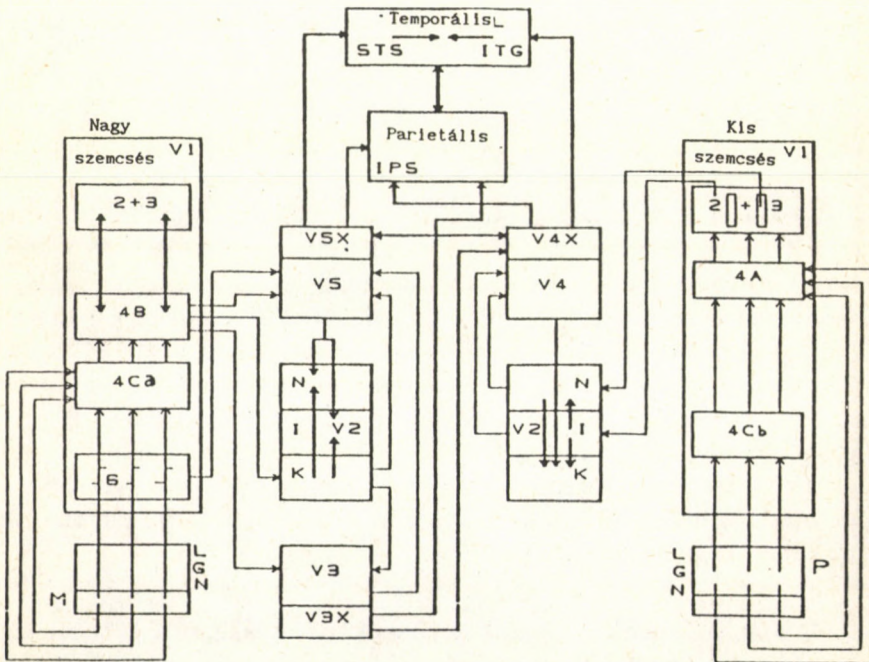
Ahogy egyre feljebb jutunk a hierarchiában, nem találunk egyre éle-
sebben hangolt sejteket, például az MT-ben a sávszélesség egy kicsit na-
gyobb, mint a V1-ben. Az RM-ek átlagos mérete meredeken változik:
már az MT-ben is nagyobb, de a V4-ben 36-szor, az AIT-ben mintegy
100-szor akkora, mint a V1-ben. A tendencia folytatódik, az MST-ben
olyan sejtek is vannak, melyek RM-je egy teljes negyed képsíkot lefed.

A vizuális információ magasabb szintű feldolgozása erősebb szelekti-
vítás alkalmazása helyett a vizuális input bonyolultabb tulajdonságainak
kezelése által történik. Például ha egy 45°-os szögben felfelé mozgó szin-
uszrácsra egy 45°-os szögben lefelé mozgó szinuszrácsot szuperponálunk,

7. ábra

Az M- és P-pálya részletesebb ábrázolása

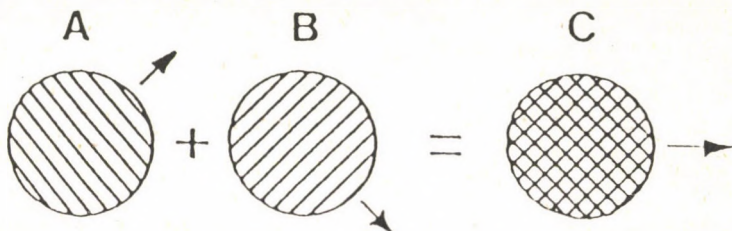
Az ábra a CGL-t magnocelluláris (M) és parvocelluláris (P) rétegre felosztva mutat-
ja. V1 és V2 szintén fel van osztva nagysejtes és kis sejttes részre. V1 fel van oszt-
va IV, IVCa, IVCb alrétegekre is (melyek a szövegben IV, IVc alfa és IVc béta-
ként szerepeltek), továbbá IVb, IVa, II és III-ra, bennük a foltok kis hengerekkel
vannak jelölve a kis sejttes részen. V2 fel van osztva vastag (K), vékony (N) csíkok-
ra és csíklóközökre (I). Az X-ek az adott area komplex részét jelölik, V5 pedig MT-
re utal (Zeki and Shipp 1988).



az észlelt kép vízszintesen mozgó rács képe lesz (8. ábra). V1-ben a sejtek csak a komponensek mozgására reagálnak, míg az MT-ben a sejtek 20%-a a komplex mozgásra reagál. Egy másik példa a rotáció-szelektív sejt, amely csak rotációra, esetleg nyírómozgásra reagál, de lineáris mozgásra nem. Ilyen sejteket az MST-ben (14%) és 7a-ban lehet találni, de az MT-ben vagy előbb nem (MAUNSELL és NEWSOME, 1987).

8. ábra

Egy szinuszoid fényességű vonalkázás jobbra felfelé mozogva (A) és egy másik jobbra lefelé mozogva (B) egy jobbra mozgó rács érzetét kelti (C) (Maunsell and Newsome 1987)



Hosszú ideig úgy gondolták, hogy a sejtek az információt csupán serkentő (excitációs) válaszokból gyűjtik. Az volt az elképzelés, hogy azok az ingerválaszok, amelyekhez a látótér nagy részéről kell az információt integrálni vagy olyan neuroncsoportok összmunkájának az eredményei, melyek lefedik az adott területet, vagy olyan sejtek állítják elő őket, melyeknek nagyon nagy az RM-jük. ALLMAN és munkatársai (1985) bebizonyították, hogy ez nem így van. Sokféle inger van, ami erősen és szelektív módon befolyásolja a sejt választ a klasszikus receptív mezőn (KRM) kívülről. Ezek a környezeti ingerek képesek mind gátolni (akár a teljes elnyomásig), mind erősíteni a válaszokat, de önmagukban nem képesek előidézni azokat. A V1-ben ez a környezet mintegy háromszorosa a KRM-nek, az MT-ben 7—20-szorosa. Minél távolabb van a KRM a látótér közepétől, annál nagyobb ez a környezet. A KRM-en belül feldolgozásra kerülő összes ingerfajtát bizonyítottan befolyásolja ez a hatás (kivételet a középpont és szél közötti relatív mélység ingere, melyet még nem vizsgáltak meg ebből a szempontból).

Különböző mechanizmusok segítenek az areákban létrehozni ezt a hatást. Az areákon belül az axonok rendkívül távolra kinyúlnak, és a V1, V2, V4-ben egy rácsszerű szerkezetet mutatnak. Ez a struktúra feltűnően

egybeesik az orientáció—preferencia eloszlással, és valószínűleg gátló hatások érvényesülnek benne. Az areák között három módja van a környezeti hatások közvetítésének (1) az átellenes félteke reprezentációjának felhasználása, (2) nem cortexbeli areák felhasználása, és (3) feedback felhasználása egy magasabb areából, ahol egy KRM már lefedti az alacsonyabb area környezetét is.

A jelen pillanatig a látókéreg információfeldolgozásának vizsgálatában a szigorúan szenzoros megközelítések domináltak, habár úgy tűnik, hogy bizonyos robosztus extraretinális jelek erős kölcsönhatásban vannak a vizuális jelekkel (MAUNSELL és NEWSOME, 1987). Az eredmények arra mutatnak, hogy e jelek hatása V1-ben nem erős, de a magasabb areákban (például V4) a sejtek mintegy 80%-a befolyásolódik ily módon. Ezen hatások eredményeképpen az adott sejtek más válaszokat adnak a kísérleti állat különböző figyelmi állapotában vagy különböző feladatok során.

2.6. A SZÍN—FORMA PÁLYA AREÁINAK TULAJDONSÁGAI ÉS FUNKCIÓI

A 9. ábra emlékeztet a szín—forma pálya fő areáira. V1 az első ilyen area a látókéregben és az IT (mely magába foglalja az AIT és PIT areákat) tűnik az utolsónak az alakfeldolgozás láncolatában. Ez utóbbi ugyanis olyan areákba vetít tovább, melyek már főleg nem vizuális információval foglalkoznak. Az ábrán a nyilak a felszálló kapcsolatok irányát mutatják. A kis „d” jelöli, hogy az adott vetítés csak V3 hátsó (dorsalis) részére korlátozódik. Az alábbiakban ezen areák alaposabb vizsgálata következik: Milyen információt és milyen módon dolgoznak fel?

2.6.1. V1

A teljes vizuális információ döntő többsége átmegy a V1-en, így érdemes egy kicsit alaposabban megvizsgálni. Schiller és kollégái (SCHILLER és mts., 1976a, 1976b, 1976c) hétféle egyszerű (S) és egy komplex (CX) típusú sejtet azonosítottak V1-ben. Az S- és CX-sejtek ilyen osztályozását más kutatók eredményei is alátámasztották. Az S-sejtek két vagy több elkülönülő almezővel rendelkeznek, világos és sötét élekre reagálnak, almezőik ellentétes tulajdonságú környezetekkel rendelkeznek. A CX-sejteknek nincsenek elkülönülő almezőik, sem ellentétes környezetük, RM-jük mintegy kétszerese az S-sejtek RM-jének. A 10. ábra az S-sejtek típusait és tipikus CX-sejtet mutat.

Az S-sejtek a CGL sejtjeire emlékeztetnek, de orientáció- és irány-szelektívek, míg amazok nem. A világos és sötét élekre válaszoló sejtek

száma megközelítőleg azonos. Az RM-ek kör- vagy ellipszoid alakúak, a két almező közötti táv kb. további egy almezőnyi. Megközelítőleg kétszer annyi CX-sejt van, mint S-sejt.

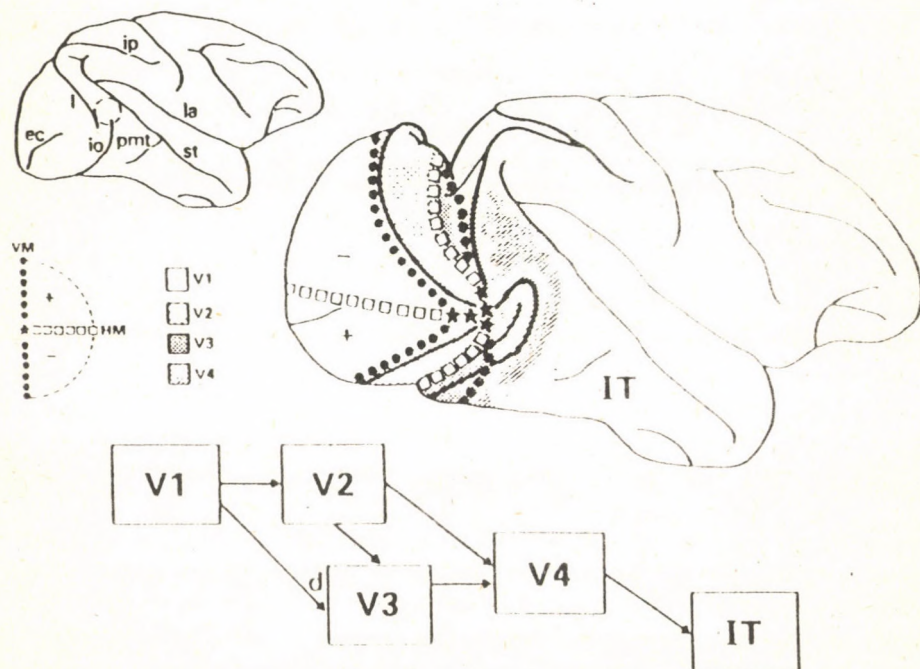
A V1-ben a sejtek legalább tízféle egymásra ortogonális információt dolgoznak fel: orientáció, irány, sebesség, binokularitás, binokuláris diszparitás, diszparitásváltozás, hossz, térbeli frekvencia/szélesség, szín/hullámhossz, fényerő/fényesség. Az ortogonalitás elkülönülő feldolgozó mechanizmust jelent mindegyik számára. A lista nem lezárt, könnyen lehet, hogy a görbület és a textúra feldolgozása is V1-ben kezdődik.

Az orientáció oszlopos szerveződésben kerül feldolgozásra, ahol minden oszlopnak saját optimális orientációja van. Egy ilyen oszlopon belül a szomszédos sejtek gyakran különböző irányselektivitással rendelkeznek.

9. ábra

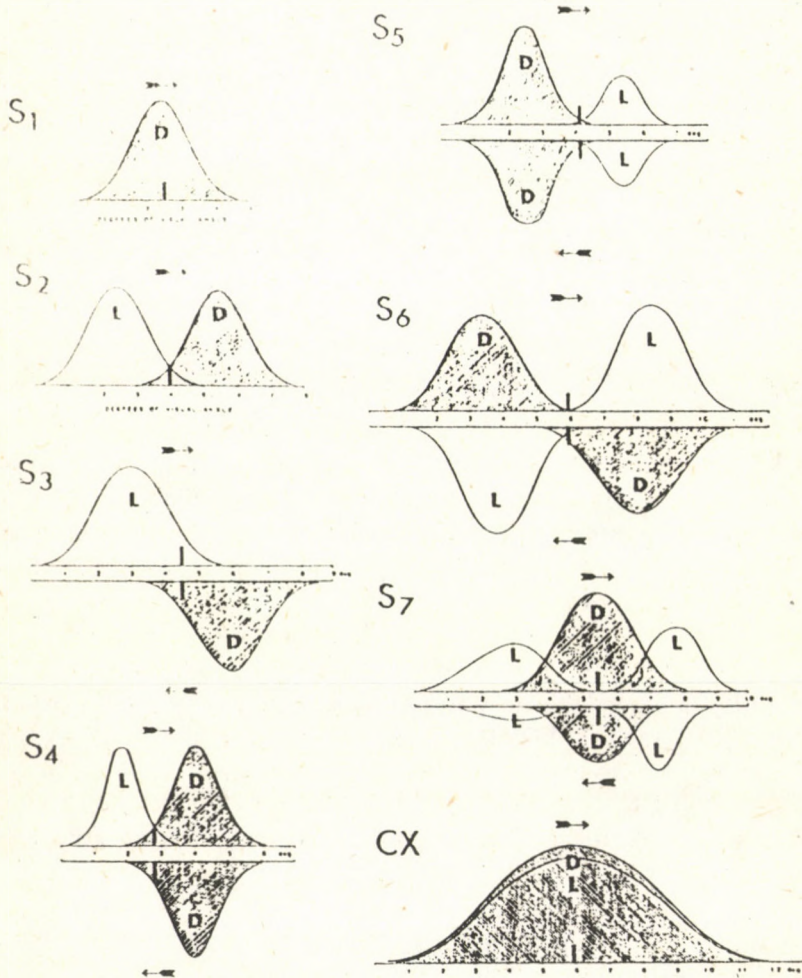
A cortexen belüli ösvény a striate cortextől a halántéki lebenyig

Fent: A V1, V2, V3, V4 és IT ábrázolása egy részlegesen bevágott féltekén oldalnézetből. Lent: A szín—forma ösvény diagramos ábrázolása (Desimone et al., 1985).



A hét S-sejttípus és a CX-típus sematikus ábrázolása

Az ábrán a nyilak irányába mozgó világos (L) és sötét (D) élekre reagáló almezők elhelyezkedése látható a sejt középpontjától való eltérés függvényében (Schiller et al. 1976a).



Ez is alátámasztja azt az elképzelést, miszerint az orientációt és az irányt különböző mechanizmus kezeli. Jól bizonyított tény, hogy az ember erősebben érzékeny a függőleges és vízszintes irányokra, mint a többire, és ez a különbség teljes egészében a foveális sejtektől és nem a parafoveá-

lisaktól származik (De VALOIS és mts., 1982). Nincs különbség a mozgó és felvillanó stimulusok orientációjára adott válaszokban.

Az okuláris dominancia nagyobb oszlopos szerveződéssel bír. Az orientációoszlopokban jól megkülönböztethető színszelektív clusterek vannak. Keveset tudunk arról, hogy a V1-ben a színinformáció átesik-e bármilyen komolyabb transzformáción.

A sejtek legnagyobb része (S- és CX-sejtek egyaránt) nem különbözteti meg a különböző térfrekvenciájú stimulusokat, ha azokat éles kontúrlépcsőkkel ábrázoljuk, de jól reagálnak, ha szinuszoid fényességátmeneteket alkalmazunk. Úgy tűnik, hogy két mechanizmus van a kontúrorientáció és a térfrekvencia számára. Az utóbbi csak akkor aktivizálódik, ha az előbbi nem tud jól működni. Ezt az eredményt nem lehet csupán izgató válaszfüggvényekkel leírni.

2.6.2. V2

V2 olyan areaként vált ismertté, ahol a különböző típusú információk szétválasztása térbeli clusterek kialakulásával és különböző rákövetkező areákba való vetítéssel is történik. A szétválasztásról már tettünk említést. A 11. ábra egy mókusmajom látókérgi metszetén mutatja a borotvaéles V1—V2 határt, a vékony és vastag sötét csíkokat és a csíkközöket.

A tulajdonságok sok párosítási lehetősége ellenére a V2-ben clusterek alakultak ki, amelyek négy fő sejtípus elkülönítésével jól leírhatók. A négy sejtípus, melyek között szinte alig van átfedés, a következő: (1) nem-orientált, (2) komplex orientált, (3) diszparitás-hangolt és (4) végződésjelző (HUBEL és LIVINGSTON, 1987b). A nem-orientált sejtek élénken reagálnak stacionárius stimulusra, a másik három gyengén vagy szinte nem is.

A nem-orientált sejtek két alcsoportra oszthatóak. A sejtek fele minden tekintetben hasonlít a feltsejtekhez a V1-ben, a másik fél szintén hasonlít, de egy fontos kivétellel. Ezeknek a sejteknek van egy optimális stimulusméretük, ami lényegesen kisebb, mint az RM-jük, és erős választ adnak, bárhol is tűnik fel ez a stimulus az RM-jükben.

A komplex orientált sejtek nagyon hasonlítanak a V1 CX-sejtekhez. Ezek gyakran együtt fordulnak elő akár diszparitás-hangolt, akár végződésjelző sejtekkel, míg e két utóbbi sohasem keveredik egymással.

Végződésjelző sejtek mind S, mind CX típusúak lehetnek. A V2-ben a sejtek vagy végződésjelzőek (tehát teljesen elnyomódik a válaszuk, ha a stimulus hosszabb az RM-jüknél), vagy nem. A V1-ben ez nem ennyire polarizált, az elnyomás mértékének eloszlása közel egyenletes.

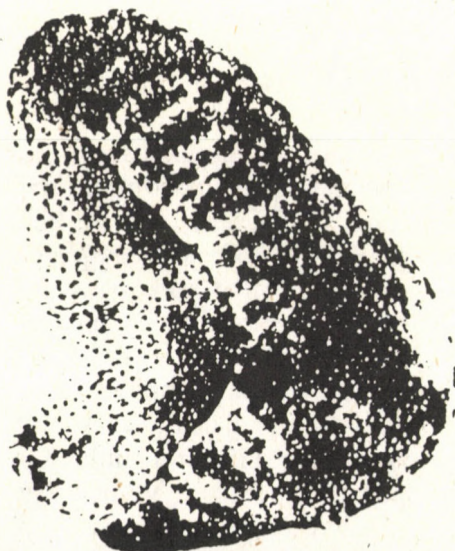
Binokuláris sejtől lényegesen több van V2-ben, mint V1-ben. A diszparitás-hangoltság bimodális: egy sejt vagy erősen, vagy egyáltalán nem hangolható ily módon. A három orientációszelektív sejtcsoport nem szín-szelektív. A négyféle sejtcsoport 3 clusterben található V2-ben: (1) nem-orientált, (2) komplex orientált és diszparitás-hangolt, (3) komplex orientált és végződésjelző. HUBEL és LIVINGSTON (1987b) a következő párosítást ajánlották:

1. Világos csíkközök — végződésjelző domének

11. ábra

Egy squirrel majom cortexének megfestett metszete a V1--V2 határral középen, a csíkokkal és a csíkközökkel V2-ben

A megfestett és megfagyasztott cortexet a felülettel párhuzamosan metszették el, és a felület irányából tekintünk rá (Hubel and Livingstone 1987b).



1 cm

2. Vastag sötét csíkok — diszparitás-hangolt domének

3. Vékony sötét csíkok — nem-orientált domének

Ezek alapján a vastag csíkok főleg sztereopszissal, a vékonyak szín-
nel, a világos csíkközök pedig alakkal foglalkoznak.

A V2 feladata az információfeldolgozásban nem világos. Az egyik feltevés szerint a különböző tényezők komplexitásának fokozódása a feldolgozás során elkülönülve megy végbe. Például a színfeldolgozásban a nem-orientált sejtek gyengén reagálnak diffúz fényre, kontúrra van szükségük. Tehát ezek a sejtek jól reagálnak nagyobb színfoltokra, de csak a kontúrok közelében. Az egzakt pozícióinformáció elvész, de a felbontás megmarad. A végződésjelző sejtek a görbülő kontúrokra reagálnak a legjobban, jobban, mint a hosszú egyenes vonalakra. Tehát ezek a sejtek az alak-analízisben játszanak szerepet. A retina CGL-jében a pontinformációk és a határok kapnak hangsúlyt. A V1-ben az S-sejtek szintén reagálnak a határookra, de csak bizonyos orientációjú határookra. Lehet, hogy néhány sejt reagál a görbületre is, bár Hubel és Livingston szerint a görbület feldolgozása a V2-ben kezdődik. A CX-sejtek és a V2 komplex sejtjei a teljes formát dolgozzák fel, meghatározzák a kontúrokat és a kontúrok egyes görbült részeit.

Egy másik feltevés von der HEYDT-től és kollégáitól ered (1984). Olyan sejteket fedeztek fel a V2-ben (30—35%), amelyek jól reagálnak illuzórikus kontúrokra, de valódi kontúrokra nem. A V1-ben nem találtak ilyen sejteket. Ez egy fontos megfigyelés, mert ezek a sejtek eltérően az összes előzően megfigyelttől akkor válaszolnak, amikor RM-jükben nincs inger, vagyis a környezet kelti a teljes választ. Például ha egy illuzórikus kontúrt két szomszédos sejt végződésjelző sejtjeinek ingerlésével érünk el, a kétoldali inger egyik oldalával próbálkozva sem olyan, sem feleakkora választ nem sikerül elérni az adott sejtnél, mint az illuzórikus kontúrral. Úgy tűnik, a cortex képes extrapolálni és összekapcsolni távoli ingereket, melyek azonos tárgyhöz tartozónak tűnnek. Vannak sejtek, melyek csak illuzórikus, csak valódi vagy mindkét típusú kontúrra reagálnak. Illuzórikus kontúr 3 vagy több vonalvégződéssel kelthető, a sejtválasz erőssége lineárisan nő a végződések számával. A V2-ben egy sejt válaszából különböző típusú kontúrok helyzetére és létezésére gyanakodhatunk, a V1-ben ez csak fényesség- és színkülönbségből adódó kontúrookra igaz. A V2-beli sejtek működésében a sarkok és vonalvégződések szerepe fontosnak tűnik. A V1 sejtjei kiemelik ezeket a tulajdonságokat, és számos sejt válasza konvergálhat egy V2-beli sejtbe.

Egy harmadik megfigyelés szerint V2-beli sérülés esetén a károsodott szelektív módon képtelen lokális-globális megkülönböztetési feladatok végrehajtására (ALLMAN és mts., 1985).

2.6.3. V3

Mint már említettük, a V3 csak az átellenes oldali fél látómező alsó felét ábrázolja, és VP foglalkozik a felsővel. V3-nak meglehetősen topografikus szerveződése van (80%) (FELLEMAN és VAN ESSEN, 1987). Ez azt jelenti, hogy a többségében topografikus areákban lokális komplexitások vannak. V3-nak nagyobb az RM-je, mint V1 és V2-nek, de a középponttól kifelé haladással történő RM-méret változás hasonló. V3 és VP RM-je azonos méretű.

V3 sejtei általában eléggé irányszelektívek (40%), habár a válaszamplitúdók, illetve a preferált irányok erősen változóak. Szomszédos sejtek 30°-os eltéréseken belül maradónan hasonló irányt preferálnak. A V3 sejtei erősen orientációszelektívek, és 90%-uk sebességszelektív is. Az előzetes hiedelmekkel ellentétben a sejtek 20%-a színszelektív. Ha egy sejt többféle szelektivitással bír, nincs korreláció a szelektivitások között. A V1-ben és V2-ben talált mindenféle szelektivitás megtalálható V3-ban is, és egyik sem erősebb. Viszont míg V1-ben a sejtek 30—35%-a irányszelektív, addig V3-ban ez az arány 70—75%. Felleman és Van Essen szerint a V3 erőteljes szerepet játszik a sztereopszisban. Minthogy V3 V1 és V2-ből is kap inputot, V2-ből valószínűleg a szín-input jön. A „magnó-ösvény” feladata (mely szintén magában foglalja V3-at) valószínűleg nem csak a mozgásfeldolgozás minthogy az irányszelektív sejtek száma relatíve alacsony (40%). V3 valószínűleg az alacsony kontrasztú alakanalízisben is részt vesz. V3-nak erős struktúrálatlan kapcsolata van V4-gyel, és úgy tűnik, hogy ez a fő átkötő út a nagysejtes áramlatból a halántéki lebenybe. A multiorientációjú sejtek létezése magasabb szintű funkciókat is sejtet. A V3 és VP közötti különbséget — mely leginkább az irányszelektív (40%—13%) és a színszelektív (21%—64%) sejtek aránybeli eltéréseben látszik — tények támasztják alá, melyek arra mutatnak, hogy vannak csekély eltérések az emberi percepcióban a két fél látótér esetén (FELLEMAN és VAN ESSEN).

2.6.4. V4

V4 az utolsó „vizuotopikus” area a szín—forma pályán. Ez azt jelenti, hogy a vizuális mező észlelhető rajta, de már nem „retinotopikus”, azaz a V4 bármely kis alterületéből véve egy sejtet, az a látótérnek eléggé eltérő pontjaiból kap inputot. V4 főleg V2-ből kap inputot, és kisebb mér-

tékben V3-ból. V4 főleg IT-be vetít. Sok színszelektív sejtje van, és habár a populáció erősen nem irány szelektív, van egy kicsi, de tisztán irány szelektív szubpopulációja. Ez arra utal, hogy bár V4 nem mozgásfeldolgozó area, hasznosít mozgásinformációt.

A sejtek struktúrája a foltos elrendeződés mellett (mely a felületről megfigyelhető) a harmadik dimenzióban monolitikus. Ez azt jelenti, hogy egy oszlopban mondjuk csak hullámhossz-szelektív sejtek vannak. A sejtek a V1 sejtekhez hasonlóan reagálnak az ingerekre, de ennek dacára lényegi különbségek vannak a két area között. Habár az RM-méretekben nagy a különbség ($1:40 = V1:V4$) a V4-beli sejtek majdnem olyan szelektívek, mint a V1-beliek. V4-ben a válasz erőssége nem növekszik a kontraszt-növeléssel, sőt néha csökken. V4-ben a sejtek többségének van elnyomó környezete, amely erősen és szelektíven befolyásolja a sejt viselkedését. V1-ben nincs ilyen hatás. V4 sejtjei képesek akkora méretű tárgy képét kezelni, amit V1-beli sejtek nem. V4-ben a sejtek színkonstansságot mutatnak, ami tipikus az emberi percepcióban. V1-ben nem fedeztek fel ilyen tulajdonságot. A V4 sejtek képesek optimális ingerekre reagálni, bárhol is tűnjék az fel RM-jükben. Ez hasonlít a V2 sejtekre, de a sejtek V4-ben megtartják válaszukat akkor is, ha az inger mozog. Egy másik hasonlóság V2 sejtjeivel, hogy V4-ben a sejtek folytonos spektrumot képviselnek V1 S- és CX-sejtjei között. Úgy tűnik, hogy a nem fázisszelektív magasfrekvenciás sejtek nagy RM-mel a textúrát kezelik, míg a fázisszelektív alacsony frekvenciás sejtek a kontúrok és határok kezelésére alkalmasabbak. Néhány sejt V4-ben lényegesen jobban reagál lépcsőölekre (ez a jelenség V1-ben nincs), és ez azt mutatja, hogy a kontraszt explicit módon reprezentálva van V4-ben (DESIMONE és SCHEIN, 1987; TANAKA és mts., 1986; DESIMONE és mts., 1985; VAN ESSEN, 1985).

2.6.5. IT

Az IT (inferotemporális, azaz alsó halántéki lebeny) nagyon magas szintű alakfelismerési feladatokat végez. Sok cikk számol be sejtekről az IT-ben, amelyek igen speciális ingerekre reagálnak, mint például arcra, kézre stb. (VAN ESSEN, 1985). Az arcszelektív sejtek nem válaszolnak sem egyszerű ingerekre, sem változásokra (méret, szín, orientáció, mozgásirány), melyek nem befolyásolják az arc lényegi tulajdonságait. Ezek a sejtek az arc komplexumára reagálnak. IT-sérült majmok képtelenek formatanulásra, tárgymegkülönböztetésre, viszont könnyedén képesek vizuális-térbeli feladatok elvégzésére, mint például vizuálisan vezérelt elérésére valaminek, vagy két hasonló tárgy közül a közelebbinek a kiválasztására (DESIMONE, és mts., 1985).

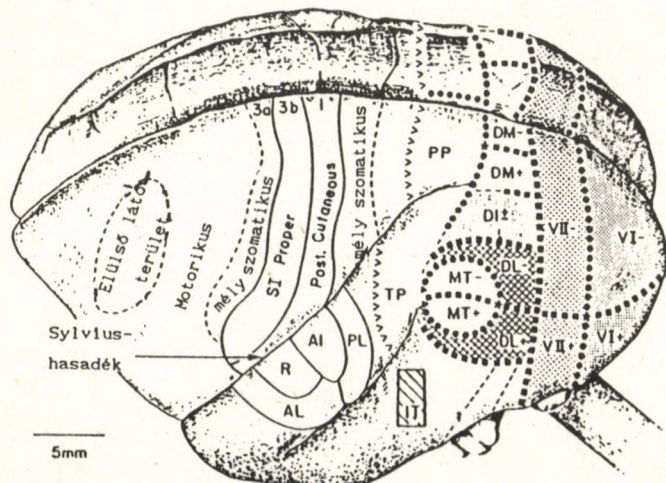
Ismereteink az IT-ről vagy tágabb értelemben a halántéki lebenyről rendkívül vázlatosak. Lényegében egy hatalmas űr tátong a V1 és V2 leg-egyszerűbb ingerprocesszorai és az IT rendkívül specifikus sejtjei között, és semmi elképzelésünk sincs az optimális ingerek típusairól a köztes részekben.

Az IT-ben nem ismert semmilyen retino- vagy vizuotopikus reprezentáció. Az RM-ek mérete az IT-ben igen nagy, többnyire minden RM magában foglalja a látómező középpontját, így az IT általános formaanalízis-re alkalmas. POLLEN és munkatársai (1984) Gábor-függvényeket alkalmaztak sejtek ingerlésére a bevonalkázott téglalapon belül az IT-ben (12. ábra). Fő eredményük az volt, hogy az IT-sejtek nem ingerelhetők azokkal az egyszerű ingerekkel, melyek az alacsonyabb areákban elterjedten használatosak.

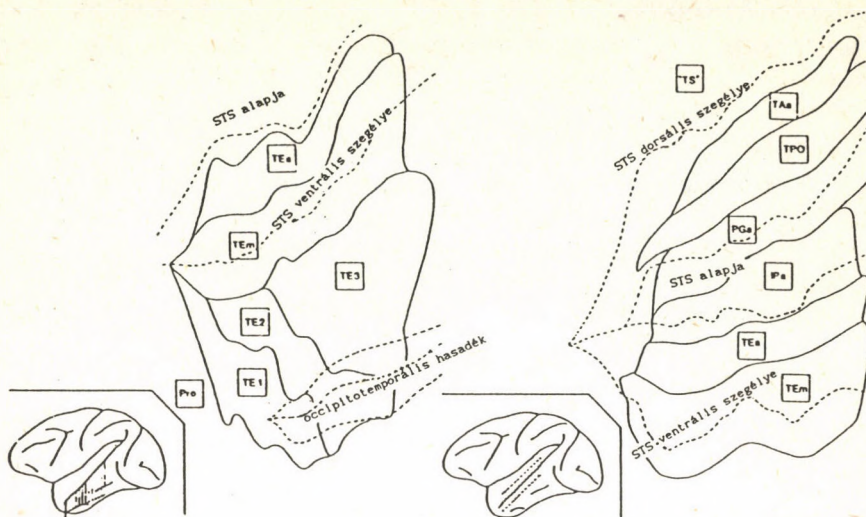
Nemrégiben BAYLES és munkatársai (1987) megvizsgálták a teljes halántéki lebeny neocortexét, és 10 szubareát határoztak meg: TE1, TE2, TE3, TEm, TEa, IPa, PGa, TPO, TAa, TS. Az általuk vizsgált területet a 13. ábra mutatja, míg a 14. ábrán egy metszeti fotó ábrázolja egy majom cortexében az areák elhelyezkedését. Ezek a szubareák erősen különböznek szelektivitásban, válaszidőben és ingerpreferenciában. Az anatómiai

12. ábra

Egy majom cortexének oldalnézete. A csíkozott téglalap területén az IT-ben kísérleteztek Pollenék (Pollen et al. 1984)



Metszeti kép a 13. ábra területeiről. Lépték: 5 mm (Baylis et al. 1987)



adatok és a válaszok alapján 5 cluster alakítható ki: (1) TE1+TE2, (2) TE3+TE_{em}+TE_a, (3) IP_a+PG_a, (4) TPO, (5) TA_a+TS.

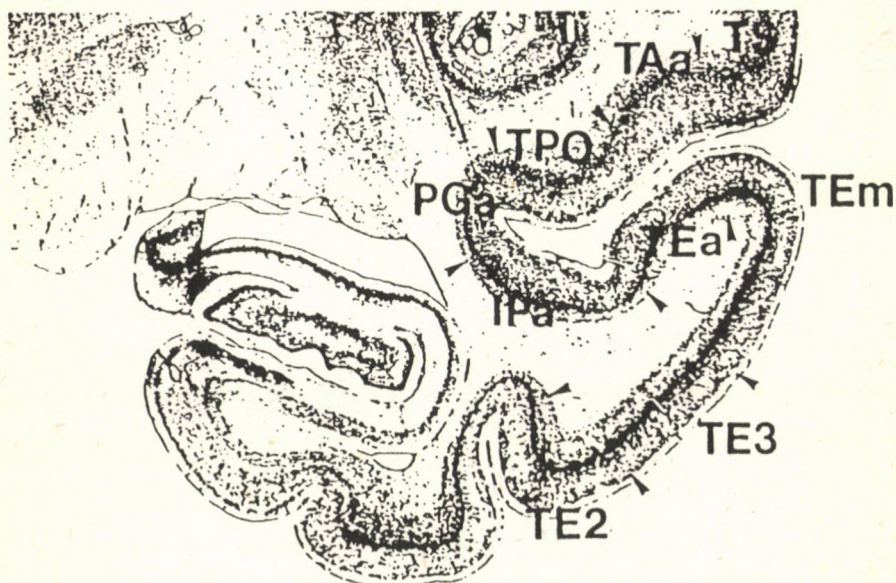
Az első clustert lehetett a legkevésbé ingerelni bármivel is, beleértve vizuális, auditorikus, szomatoszenzoros, orális és egyéb ingereket. A második cluster kétségkívül vizuális areákból áll (60% reagál vizuális ingerekre), de csak TE_a- és TE_{em}-sejtjei reagálnak arcokra. A harmadik cluster főleg auditorikus areákból áll, de ezek meglepően sok vizuális sejtet is tartalmaznak (27%). A hosszú válaszidő késleltetés azt sejteti, hogy a vizuális információ indirekt úton érkezik.

Az IT-nek a rövidtávú feladatokban van szerepe, de a hosszútávúakban nincs. Lehet, hogy közvetít asszociatív területek felé, de az IT maga nem asszociatív area. Az IT feladata valószínűleg az, hogy a különböző areák által szállított szubinformációkat összeillessze.

IWAI (1985) a majmok formalátását vizsgálta. Megfigyelte, hogy a mintatanulási folyamat két elkülönülő alfolyamatra bontható: Tanulás₁, ahol a helyes válasz véletlenszerűen fordul elő és Tanulás₂, ahol a helyes válasz aránya viharosan nő egészen 100%-ig (15. ábra). Állítása szerint

A temporális lebeny Baylisék által vizsgált részei

Balra: A cortex ITG-areájának síkábrázolása, amely megtartja a topológiát. Jobbra: A cortex STS-areájának sík ábrázolása, amely megtartja a topológiát (Baylis et al. 1987).



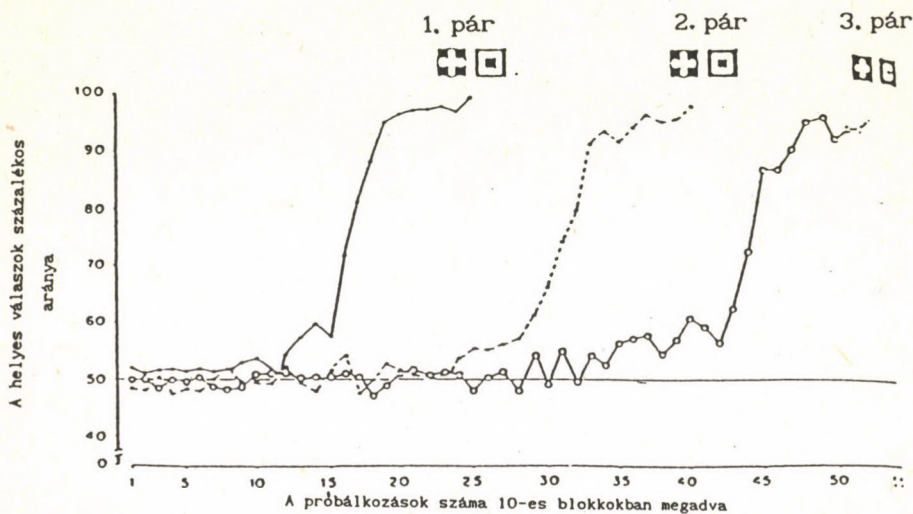
Tanulás1 az a periódus, ahol a majom a megkülönböztető vonásokat megtalálja, míg Tanulás2 a tényleges diszkriminatív tanulása az adott vonásnak. Következtetése szerint az IT a Tanulás2 folyamatába van szorosan bevonva, méghozzá AIT a „mintafelismerés” (vagy memória), míg PIT a „mintaészlelés” (pattern perception) helye. Az alsó fal lebeny (IP) feladata a háttértől elkülönülő forma vonásainak szelektív vizsgálata (Tanulás1) (16. ábra).

3. AZ ALAKPERCEPCIÓ PSZICHOFIZIKAI VIZSGÁLATAI

Befejeztük a szín—forma pálya neurofiziológiai vizsgálatainak eredményeit áttekintő részt. Ez a fejezet néhány pszichofizikai és pszichológiai tény összegyűjtését és strukturálását kísérli meg. Először megnézzük, hogy ezen a területen milyen alátámasztást nyerhet az elkülönülő pályák elmélete. Másodszor megnézzünk néhány javaslatot, hogy mik is lehetnek az

Általános tanulási görbék megkülönböztetési feladatoknál 3 hasonló konfigurációjú mintapár esetén

A görbék 55, 29, illetve 23 IT-szerűt majom eredményeiből átlagolták. A függőleges tengelyen a helyes válaszok százalékos aránya látható, a vízszintes tengelyen a próbák száma 10-es blokkokban (Iwai 1985).

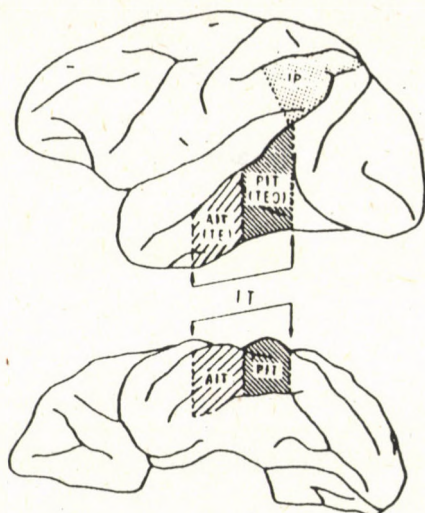


ún. „primitívek” az alakpercepcióban. Ejtünk pár szót a szimmetria kérdéséről. Ezek után áttekintjük, hogy milyen elképzelések születtek néhány ilyen primitívet kezelő mechanizmusra. Végül vetünk egy pillantást a magasabb folyamatokra, konkrétan a figyelem nélküli, illetve a figyelemirányult látás kérdésére.

3.1. AZ ELKÜLÖNÜLŐ CSATORNÁKAT ALÁTÁMASZTÓ PSZICHOFIZIKAI TÉNYEK

LIVINGSTON és HUBEL (1987) rengeteg illúziót és egyéb jelenséget vizsgáltak meg annak érdekében, hogy alátámasszák az elkülönülő csatornák elméletét. Mint már említettem, három különálló csatorna létezését feltételezték, melyeket hívjunk „nagysejtes pályá”-nak (magno-path), „kis-sejtes pályá”-nak (parvo-path) és „foltpályá”-nak (blob-path), ahol az utóbbi a kis-sejtes részben található. A 2. táblázat mutatja az ösvényeket és a sejtek tulajdonságait mindegyik ösvényen belül Livingston és Hubel szerint. (Area 17 V1-et, area 18 V2-t jelöli.)

**A vizsgált IT- és IP-areák helye és kiterjedése sematikusan
ábrázolva (Iwai 1985)**



Az ember a színek váltakozását csak sokkal alacsonyabb frekvenciáig képes követni, ha az egyenfényességű, mint ha nem. (Egyenfényességről akkor beszélünk, ha két vagy több felület fényességintenzitása megegyezik, dacára annak, hogy más paramétereikben, pl. szín, eltérnek.) A színlátás kevésbé járul hozzá a mozgáspercepcióhoz, ha egyáltalán hozzájárul, míg a fényességkontraszt előjele lényegtelen a mozgás szempontjából.

Az alaklátás nem járul hozzá a mozgás észleléséhez. A kis sejtes pálya érzékenysége a mozgásra az adaptáció kezeléséhez szükséges, és nincs semmi köze a nagy sejtes pályához.

A szín nem segít a sztereopszisban. Habár némi alakinformáció hozzájárul a sztereopszishoz, ez az információ függetlenül érkezik a fő alakfeldolgozó áramlattól. Ezt jól alátámasztja az a tény, hogy a sztereohanagolt sejtek többnyire orientációszelektívek is. A sztereopszisban az orientáció lényegesebb, mint a fényesség, majdnem annyira számít, mint a mozgás. A sztereopszis eltűnik nagy térfrekvenciák és egyenfényesség esetén. Ez arra utal, hogy a sztereopszis a nagy sejtes pályához tartozhat.

Ha az orientációszelektivitást tekintjük az alaklátás alapjának, látnunk kell, hogy mind a nagy-, mind a kis sejtes pálya foglalkozik a formá-

val, de különböző aspektusokból. A nagysejtes rész nem képes színt-contrast-információ hasznosítására, a kis sejtes rész viszont igen. Az ember még nagyon alacsony kontrasztra is érzékeny. Ez egy nagysejtes és nem kis sejtes tulajdonság. Másrészt viszont a kis sejtes rész színt-contrast-érzékenység függvénye unimodális, a fényességkontraszt-érzékenység függvénye bimodális. Livingston és Hubel szerint a fényességérzékenység magas kontrasztú modusát a nagysejtes rész okozza.

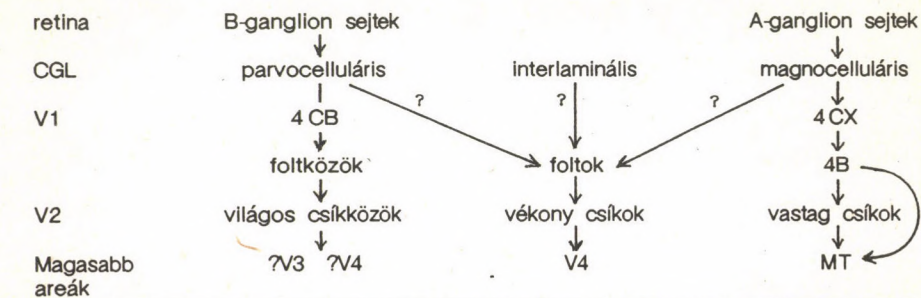
A szín- és alakfeldolgozás szintén két különböző csatornán megy, habár ez a szétválás nem olyan látványos, mint a szín és a mozgás esete. A legerősebb bizonyíték azon betegek léte, akik sérült vizuális areával formákat látnak, de szín nélkül, valamint az a tény, hogy kis tárgyak esetén

2. táblázat

A főbb egységek és kapcsolataik összegzése a főemlősök látórendszerének kezdeti szakaszán Livingstone-ék szerint (Livingstone és Hubel, 1987)

Nagysejtes pálya:	Nagy időbeli és kontrasztfelbontás, valószínűleg színvaktság, van orientációszelektivitás, de nem világos, hogy mire való.
Kissejtes pálya:	Valószínűleg nagyfelbontású alakpercepció, nem járul hozzá a színtfeldolgozáshoz.
Foltpálya:	Nincs orientációszelektivitás, alacsony térbeli felbontás, színtfeldolgozás.

A következőkben az ő eredményeket és állításait tekintjük át a kísérletek részletes taglalása nélkül.



Tulajdonságok:

Szín	igen/nem	igen	nem
Kontrasztérzékenység	kicsi	nagy	nagy
Térbelifelbontás	nagy	kis	kicsi
Orientációszelektivitás	igen	nem	igen
Mozgásérzékenység	igen	nem	igen
Írányérzékenység	nem	nem	igen
Sztereopszis	nem	nem	igen

az alakfeldolgozás még működik, de egy keveredési hatás lép fel a színek között. A foltoknak ugyanis alacsonyabb felbontási képességük van.

A színlátás határokat igényel -- ez jól bizonyított tény. Ellentétben a mozgással és a sztereopszissal, ahol egyenfényességű határok esetén a percepció eltűnik, alakpercepcióban ez csak kisebb minőségcsökkenést okoz.

Jól ismert, hogy a retinán mesterségesen rögzített képek elveszítik fényességkontrasztjukat. Hasonlóképpen, ha rögzítjük nézőpontunkat egy kép bármely pontján, az egyenfényességű színkontrasztok eltűnnek. Livingston és Hubel szerint csak a kisfelbontású pályának van megfelelő térbeli felbontóképessége, amely meggátolja az elmosódást, mégpedig azon minimális szemmozgás felhasználásával, amely az önkéntes fixálás alatt előfordul.

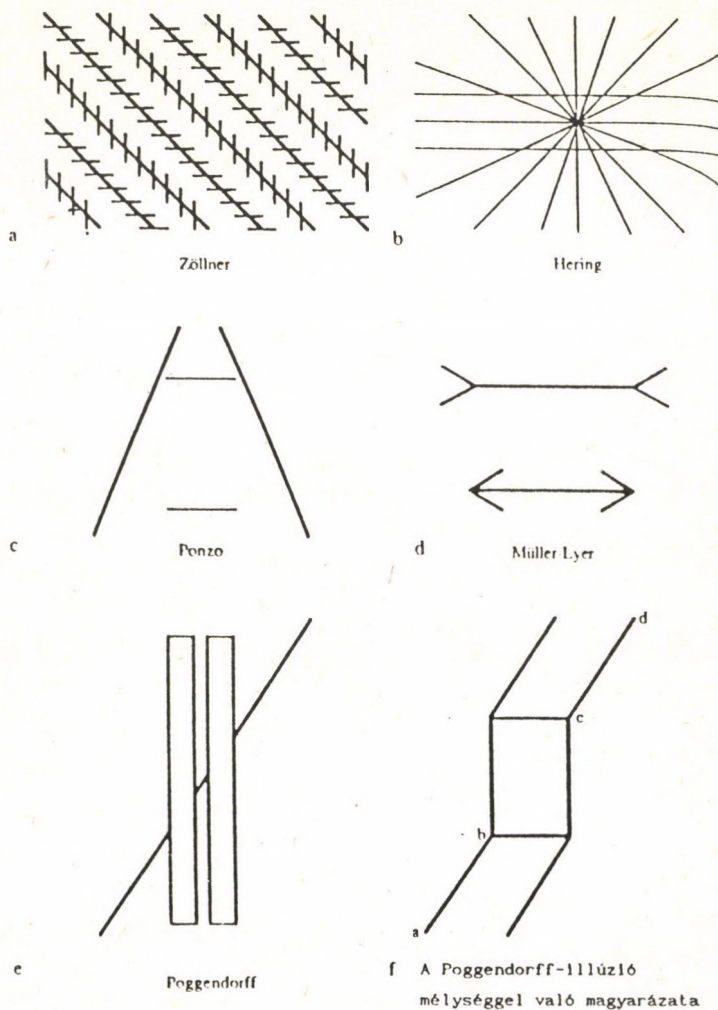
A klasszikus illúziók többsége eltűnik egyenfényesség esetén, habár a vonalak tökéletesen láthatóak. A 17. ábra bemutat néhányat ezen illúziók közül: (a) ZÖLLNER (1860): a hosszú vonalak az ábrán nem tűnnek párhuzamosnak; (b) HERING (1894): a horizontális vonalak nem tűnnek egyenesnek; (c) PONZO (1913) és (d) MÜLLER-LYER (1894): a horizontális vonalak különböző hosszúságúnak tűnnek; (e) Pogendorff (DEBLOEF, 1893): úgy tűnik, mintha egy elcsúszás lenne a ferde vonalban; (f) a Pogendorff-illúzió mélységi magyarázata: a, b és c párhuzamosak, de 3D-ben úgy értelmeződnek, mintha nem lennének azok.

Ezek az illúziók eltűnnek egyenfényességnél, de fennmaradnak alacsony fényességkontraszt esetén. Most nem lényeges, hogy mi a pontos magyarázatuk ezen illúzióknak. A fontos, hogy közös vonásuk az, hogy a virtuális mező különböző részei közötti összehasonlításon alapulnak. Gyanítható, hogy mindegyik a 2D információ és a 3D világ közötti korrelációképzési folyamatokban felmerülő hibákból ered. Az egyenfényességű hatás arra mutat, hogy a nagysejtes rendszer egyedül hozza létre ezeket az illúziókat, vagyis a nagysejtes rendszer felelős a mélységpercepcióért. A mélységérzet különböző nyomokból épül fel sztereopszis, perspektíva, árnyék, textúra, mozgás, takarások, illuzórikus kontúrok. Mindezen hatások eltűnnek egyenfényesség esetén, így mindezen nyomokat kezeli a nagysejtes rendszer.

Valódi kontúrok nem tűnnek el egyenfényesség esetén, illuzórikus kontúrok viszont igen. Ez arra utal, hogy a nagysejtes rendszer foglalkozik az illuzórikus kontúrokkal, a kisfelbontású rész nem köt össze megszakított vonalakat. Ez a sejtés három ténnyel támasztható alá. Az első, hogy a kis-

Különböző illúziók, melyek Livingstonék szerint elvesztik illuzórikus torzulásukat egyenfénység esetén

Ezek az illúziók akkor is elvesztik torzulásukat, ha nézésük közben egy pontra rögzítjük tekintetünket (Livingstone and Hubel 1987).



sejtes rendszerben sok végződésjelző sejt van, a nagysejtesben nincs. A második, hogy illuzórikus kontrasztok alacsony kontraszt esetén is észlelhetők. A harmadik, hogy a sztereo és a mozgás jelentősen erősítik az illuzórikus kontrasztokat, és ez azért lehet, mert mindhármát egyazon rendszer, a nagysejtes kezeli.

Sok kutató szerint a látórendszer bizonyos tulajdonságokat (a mozgás iránya és sebessége; mélység a sztereopsziszból, takarásból stb.; kolinearitás; fényesség; szín; textúra) arra használ fel, hogy általuk csoportosítson részeket a képen („kapcsoló tulajdonságok”). Livingston és Hubel eredményei szerint a szín csupán a színhatárok menti fényességkülönbség által kerül felhasználásra. A többi tulajdonság jól illik a nagysejtes rendszerhez. Ezt az elképzelést alátámasztja az a kísérlet, ahol egy geometriai elemekből álló kontúrrajz elveszíti 3D hatását, ha a fényességkontraszt 10% alá süllyed (18. ábra).

Formamegkülönböztetés azonban alacsony kontraszt, sőt egyenfényesség esetén is lehetséges. Ez azt jelenti, hogy mind a nagysejtes, mind a kis sejtes rész hozzájárul a forma-percepcióhoz. Mindazonáltal a nagysejtes rész a mozgással, mélységgel és a 3D struktúrával törődik, vagyis a HOL kérdésével. A kis sejtes rendszer képes a nagyfelbontású formamegkülönböztetésre vagyis a MI kérdésének megválaszolására. A 3. táblázat Livingston és Hubel kísérleteinek tömör összefoglalását mutatja. Ezek az eredmények alátámasztják a mozgás, szín, mélység és forma feldolgozására létező elkülönülő csatornák elméletét.

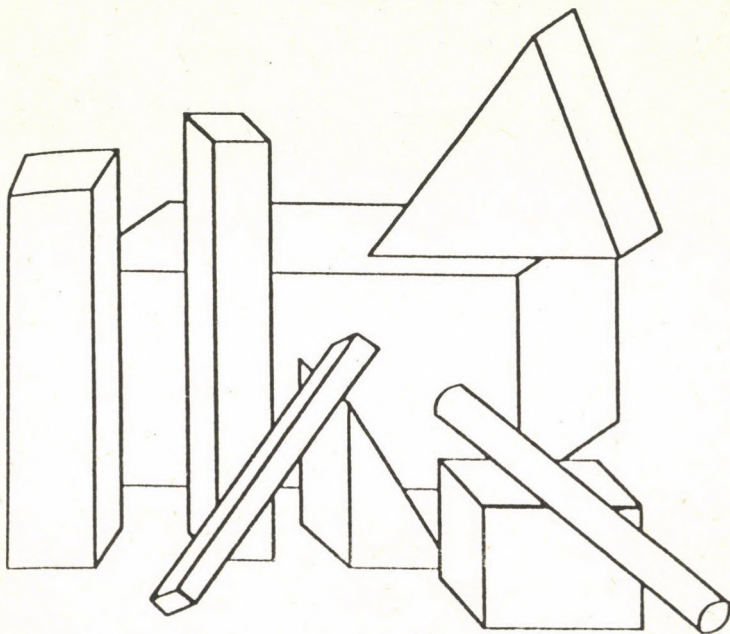
Azonban a szituáció egyáltalán nem olyan tiszta, mint azt Livingston és Hubel javaslatából az ember elképzelné. Az alábbiakban SCHILLER és munkatársainak eredményeit tekintjük át az elkülönülő csatornák kérdésével kapcsolatban (1989, 1990).

Azzal egyetértenek, hogy alapvetően két elkülönülő sejtípus található az CGL-ig: szín-opponens és szélessávú sejtek. Ezek elkülönülve vetítenek a retinából a parvocelluláris (P-) rétegbe és a magnocelluláris (M-) rétegbe. Viszont az CGL-ben vizsgálva ezeket a sejteket azt tapasztalták, hogy az M-réteg sejtjei igenis reagálnak egyenfényességű ingerekre, habár válaszuk mintegy 50%-kal kisebb amplitúdójú volt. Ezért ők a látómező két különböző részén apró roncsolásokat végeztek a majom M-, illetve P-rétegében, és pszichofizikai kísérleteket végeztek az elkülönülő csatornák elképzelésének tesztelésére. A 4. táblázat mutatja eredményeiket.

Van némi egyezés ezen eredmények és Livingstonéké között. A sztereopszis és a kontraszt azonban súlyos és lényeges eltérésnek tűnik. Schil-

**Egy ábra annak demonstrálására, hogy a kolinearitást
felhasználjuk összekapcsoló tulajdonságként**

Livingstonék szerint egyenfényesség esetén a 3D tárgyak érzete eltűnik, és csak vonalak kusza tömegét látjuk (Livingstone and Hubel 1987).



lerék szerint egy apró random pontokból álló céltárgy esetén, amely egy randompont háttérből mélységben kiugrik, a kissejtes rendszer az, amelyik a feldolgozásban részt vesz. Kis felbontású („durva”) sztereopsziszfeladatoknál úgy találták, hogy a kissejtes és a nagysejtes rendszer is részt vesz a feldolgozásban. Eredményeik szerint szintén mindkét rendszer részt vesz a fényességmegkülönböztető feladatokban.

Ezt követően Schillerék a V4-et kezdték tesztelni, hogy megállapítsák, valóban csak a kissejtes részből kap-e inputot. Ahogy elkezdték a tesztelést színnel, meglepődve tapasztalták, hogy a majomnak semmi nehézséget sem okoz a színdiszkriminációs feladat. Alaposabb vizsgálat ki derítette, hogy mégiscsak okoz, de messze nem olyan mértékűt, mint a másik majomnál, amely CGL parvocelluláris roncsolást szenvedett. Az 5. táblázat néhány összehasonlító eredményt mutat. A „kisebb” diszkriminá-

**Az emberekkel végzett pszichofizikai eredmények és a főemlősök látórendszerében feltételezett három alegység fiziológiai tulajdonságai közötti korrelációk összegzése
(Livingstone és Hubel, 1987)**

A * a pszichofizikai és neurofiziológiai adatok egybehangzását jelzi.
Az üresen hagyott helyeknek megfelelő releváns kísérletet ne végeztek.

MAGNO RENDSZER

<i>Fiziológia:</i>	színszelektivitás nincs	kontrasztérzékenység nagy	idői felbontás gyors	téri felbontás kicsi
<i>Humán látás:</i>				
Mozgáspercepció				
Mozgásdetekció	*	*	*	*
Látszólagos mozgás		*	*	*
Mélység-cue-k				
Sztereoopszis	*	*	*	*
Rivalizáció	*			*
Mozgás	*	*		
Parallaxis	*	*		
Árnyékolás	*	*		
Kontúrok	*			*
Takarás	*			
Perspektíva	*	*	*	*
Kapcsoló tulajdonságok				
Kapcsolási mozgás alapján	*	*		
Kolineáris kapcsolás				
(látszólagos kontúrok)	*	*	*	*
Ábra-alap elkülönülés	*			

PARVO-FOLTKÖZ RENDSZER

<i>Fiziológia</i>	színszelektivitás van	kontrasztérzékenység alacsony	idői felbontás lassú	téri felbontás nagy
<i>Humán látás</i>				
Alakdiszkrimináció				
Orientációdiskrimináció	*	*	*	*
Alakdiszkrimináció	*	*	*	*

PARVO + MAGNO-FOLT RENDSZER

<i>Fiziológia</i>	színszelektivitás van	kontrasztérzékenység nagy	idői felbontás lassú	téri felbontás kicsi
<i>Humán látás</i>				
Színpercepció				
Szín-diszkrimináció	*		*	*
„Fliker fotometria	*		*	

CGL-jükben nagysejtes, illetve kissejtes roncsolást szenvedett majmokkal végzett pszichofizikai kísérletek eredményei (Schiller, 1989)

	nagysejtes roncsolás	kissejtes roncsolás
kontrasztérzékenység		
magas térbeli frekvencia esetén	nincs hatás	súlyos hatás
színdiszkriminációs képesség	nincs hatás	súlyos hatás
alak és textúraérzékelés		
nagy térbeli frekvencia esetén	nincs hatás	súlyos hatás
sztereopszis kis céltárggyal	nincs hatás	súlyos hatás
flicker	súlyos hatás	nincs hatás
mozgás	súlyos hatás	nincs hatás

CGL-jükben, illetve V4 areájukban kissejtes roncsolást szenvedett majmokkal végzett pszichofizikai kísérletek eredményei (Schiller, 1989)

	V4 roncsolás	kissejtes roncsolás
színdiszkrimináció	mérsékelt	súlyos
alak- és textúraérzékelés		
nagy térbeli frekvenciákon	mérsékelt	súlyos
sztereopszis kis céltárggyal	mérsékelt	súlyos
„kisebb” diszkrimináció	súlyos	mérsékelt

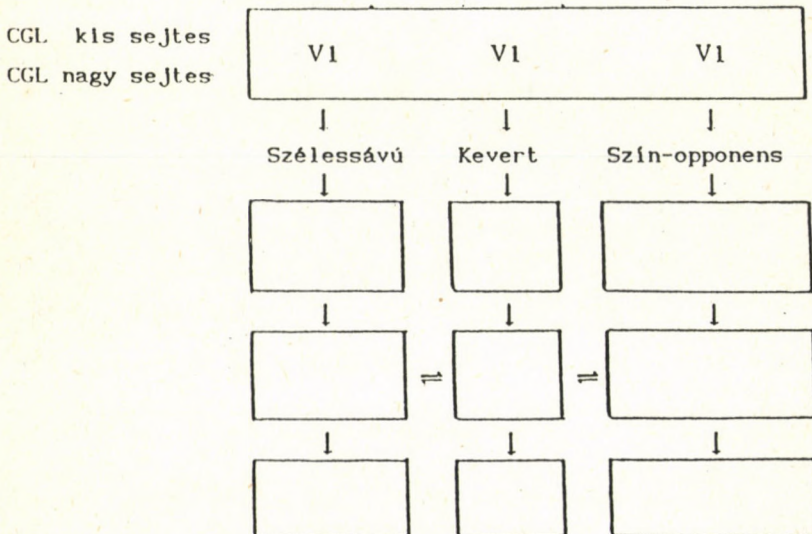
ciós feladat olyan, hogy a majomnak a kevésbé fényes, kevésbé kontrasztos, a kisebb méretű stb. tárgyat kellett választania. A V4 nagy fontossága az ilyen kísérletekben arra utal, hogy a V4 főszerepet játszik a „figyelem-jelenség”-ben, vagyis a figyelem leszűkítésében a paraméterek egy kisebb tartományára. További kísérletekkel Schillerék azt is kiderítették, hogy a V1-ben a sejtek 66%-a kizárólag az CGL M- vagy P-rétegéből nyeri az információt, viszont a maradék mindkét rétegből.

Schiller több következtetést vont le. Az első, hogy az egyenfényesség nem jó szétválasztó szempont a nagysejtes és a kis sejttes részek között, vi-

szont a térbeli frekvencia megfelelőnek tűnik. A második, hogy a nagysejtes rendszer a látvány durva és időbeli, a kis sejtes rendszer pedig a finom részleteivel foglalkozik. A harmadik, hogy az abszolút elkülönülő csatornák elmélete rossz. Schiller egy kevésbé ambíciózusabb sémát javasol (19. ábra). Eszerint a sztereopszis például kevert inputot kap, durva információt a nagysejtes, és részleteset a kis sejtes rendszerből. Ha például a kis sejtes rész hiányzik, a sztereopszis működése még mindig lehetséges, dVégso következtetésünk az lehet, hogy dacára a két javaslat ellentmondásainak (ami elkerülhetetlen az új eredmények legfrissebb értelmezésében) azok segítik a látás folyamatának strukturálását a magunk számára. Világos, hogy abszolút elkülönülő csatornák feltételezése ilyen komplex jelenség esetén, mint az alakpercepció, nem ésszerű. Ugyanekkor konszenzus látszik kialakulni az ösvények és areák azonosításában, melyek a jelenség egyik vagy másik aspektusába jobban be vannak vonva.

19. ábra

A látási folyamatban elkülönülő csatornák között egy lehetséges kapcsolat sematikus vázlata (Schiller 1989)



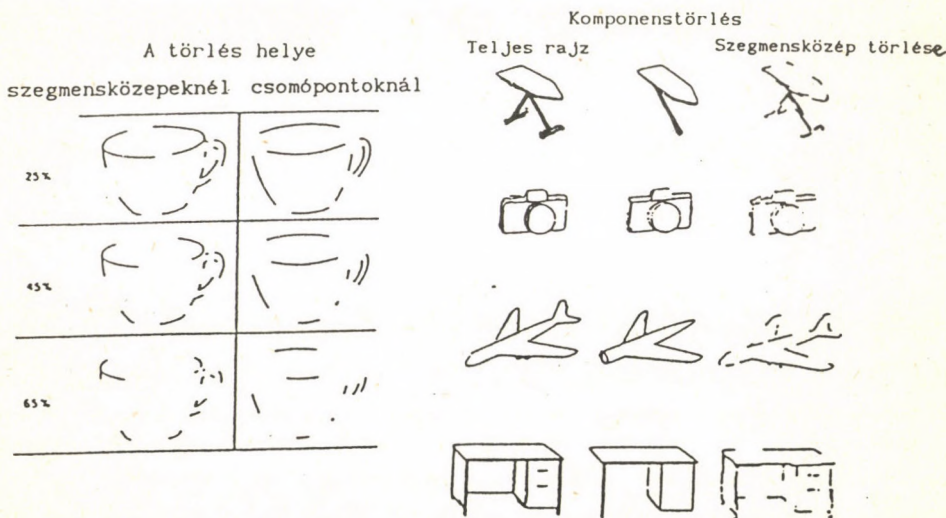
Egy utolsó megjegyzés, hogy az evolúció folyamatában a nagysejtes rész valószínűleg az öregebb. Ily módon nem a „feladatok felosztása” szempontjából kell a teljes folyamatot vizsgálni, hanem azt kell nézni, hogy mi az, amivel a kissejtes rész hozzájárul a hatékonyság fokozásához. A színlátás nagyon későn fejlődött ki, azután hogy a formalítás a mozgáspercepciótól elvált. Ez lehet az ok, amiért a színlátás csak a formalitással dolgozik össze.

3.2. MIK LEHETNEK A PRIMITÍVEK?

Most ama kérdés felé fordulunk, hogy: Mik lehetnek azok az alap- és magasabb szintű primitívek, melyeket az alakpercepció használ? Erős bizonyítékok vannak arra, hogy látórendszerünk gyorsan felhasználja és feltételezi a párhuzamosságot, görbületet, szimmetriát és együttes végződéseket (BIEDERMAN, 1985). Azt is bizonyították, hogy az ember gyenge egy fizikai jellemző abszolút kvantitatív becslésében, és még ez a szerény teljesítmény is gyorsan csökken többszörös fizikai jellemzők esetén. Kér-

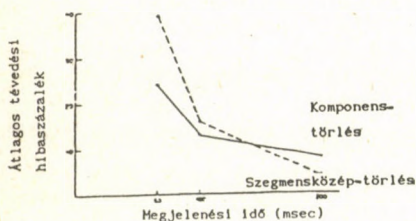
20. ábra

Balra: Egy tárgy, amely kontúrjának 25, 45 és 65%-a törölve van a szegmensközepeknél, illetve a csomópontoknál. Jobbra: Példák olyan rajzokra, ahol ugyanannyi kontúr van kitörölve szegmensközepeknél, illetve teljes komponensek formájában (Biederman 1985)

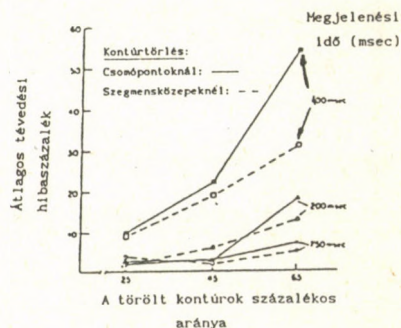


A 20. ábra rajzaival végzett kísérletek eredményei:

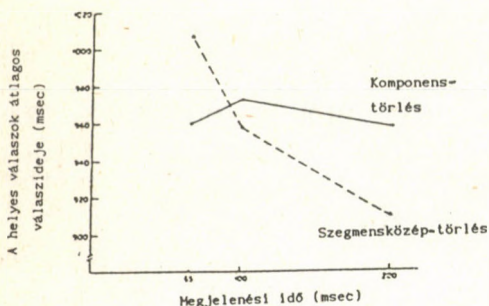
- A tárgy megnevezésének átlagos hibaszázaléka a megjelenési idő függvényében részek (component), illetve vonalközepek (midsegment) törlése esetén;
- A tárgy megnevezésének átlagos hibaszázaléka a törölt rész százalékának függvényében különböző megjelenési időkre, csomópontok (vertex), illetve vonalközepek (midsegment) törlése esetén;
- Átlagos reakcióidő (ms) helyes tárgymegnevezés esetén a megjelenési idő függvényében részek (component), illetve vonalközepek (midsegment) törlése esetén;
- Átlagos reakcióidő (ms) helyes tárgymegnevezés esetén a törölt rész százalékának függvényében különböző megjelenési időkre, csomópontok (vertex), illetve vonalközepek (midsegment) törlése esetén (Biederman 1985).



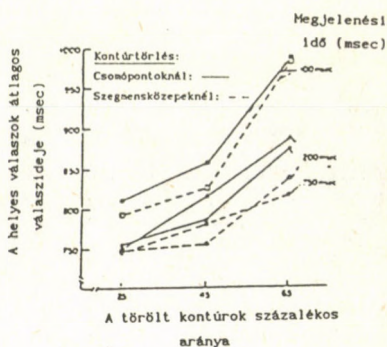
a.



b.



c.



d.

dés, hogy a fent említett jellemzők primitívek-e önmagukban, vagy már egy magasabb szintű reprezentáció elemei.

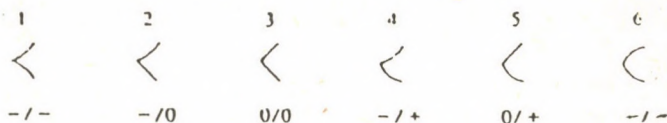
Biederman megvizsgálta a köztes vonalak és csomópontok, illetve a köztes vonalak és komponensek kitörlésének hatását (20. ábra).

Az eredményeket a 21. ábra mutatja. Ezek két következtetést kínálnak. Az első, hogy a csomópontok törlése nagyobb információvesztéssel jár, mint a köztes vonalaké. A második, hogy a kitörölt köztes részek kitöltése időt igényel, de ha egyszer már elegendő idő áll rendelkezésre a lyukak kitöltésére, a kinyerhető információ több, mint komponensterlés esetén. Ezek alátámasztják, hogy a végpontok és a görbületi diszkontinuitások fontos szerepet játszanak az alakpercepcióban.

22. ábra

A Stevens és Brookes kísérleteiben használt 6-féle saroktípus:

(1) negatív/negatív, (2) negatív/zéró, (3) zéró/zéró, (4) negatív/pozitív, (5) zéró/pozitív, (6) pozitív/pozitív görbületek (Stevens and Brookes 1988).



Ha átlátszatlan konvex tárgyak összeérnek vagy részben takarják egymást, a sziluettek közötti kapcsolódási pontok „sarkokat”, kontúrbeli törésvonalakat képeznek. Az alak—háttér döntés nemcsak a kontúr két oldala között tesz különbséget, hanem attól kezdve csak az alak kap további attribútumokat. Így bár a javaslat, hogy a sziluet szegmentálása a görbület negatív minimumainál kell hogy történjen, ígéretes, előfeltételezi az alak—háttér szétválasztási folyamatot. STEVENS és BROOKES (1988) azt vizsgálták, hogy milyen primitívek segítségével képes a látórendszer alak—háttér döntést hozni a feldolgozás korai szakaszában. Véleményük szerint a konkáv és konvex tulajdonságon túl a sarkok önmagukban is független, figyelem nélküli meghatározói az alak—háttér döntésnek. A 22. ábra azokat a különböző saroktípusokat mutatja be, melyeket a fenti eredményt hozó vizsgálataikban használtak.

ZUCKER és DAVIS (1988) pontokat és végpontokat vizsgáltak. Szerintük valódi retinaképek esetén nincs értelme folytonos kontúrokról beszélni, csak sűrű és ritka ponthalmazokról. Jól tudott, hogy egydimenziós ponthalmazok azonnal kontúrokba csoportosulnak (MARR, 1982). Zucker és Davis szerint van egy hirtelen átmenet a méret/távkoz arányban — valahol 1:5 környékén —, ami után bizonyos szubjektív illúziók jelentkeznek (például az ún. Nap-hatás, a Kanizsa-illúzió), de előtte soha.

Ezzel ellentétben a Necker-kocka nem veszi el mélységillúzióját sem sűrűbb, sem ritkább arány esetén. Zuckerék szerint a csoportosulásnak ez a komplex tulajdonsága egy komplex mechanizmusból ered, azaz legalább kétféle csoportosító folyamat létezik, melyek különböző szinten operálnak. A Necker-kocka esetén az output ugyanaz sűrű és ritka pontsor esetén (mélységet érzünk), habár a gerjesztő mechanizmusok különbözőek lehetnek. A Nap-hatás esetében viszont a két csoportosító mechanizmus azon szint előtt, illetve mögött van, ahol az illúzió fejlődik, így mire a ritka pontok csoportosítása megtörténik, az illúziót keltő mechanizmusokon már túljutott a jel. Most elsősorban azt hangsúlyozzuk, hogy a pontok és azok sűrűsége igen lényeges, valamint konzisztens és komplex viselkedésű azon folyamatok triggerelésében, melyek „végpont-folyamatok”-ként ismertek. Figyelembe véve, hogy az ember közismerten sosem csoportosít egybe különböző kontrasztú pontokat, a kontraszt valószínűleg prioritást élvez a csoportosító mechanizmusokkal szemben.

Habár JULESZ (1981) textúrákkal foglalkozott, munkáját érdemes itt megemlíteni abból a szempontból, hogy mi lehet és mi nem lehet primitív („texton”). Ő arra a végkövetkeztetésre jutott, hogy a textúra-analízis során a látórendszer nem tud harmad- és magasabbrendű statisztikákat felhasználni, sőt lehet, hogy már a másodrendű statisztikák sem fontosak. A látófolyamat tehát néhány lokális tulajdonság elsőrendű statisztikáit használja. Julesz állítása szerint a figyelem nélküli textúra-diszkrimináció esetén a szín mellett két egyéb texton létezik: a hosszúkas foltok (adott orientáció és hossz/szél aránnyal) és a végződésesek.

A szimmetria szerepe a látásban mindig is izgalmas kérdés volt. Bár sok kísérletet végeztek e tárgyban, azok eredményei és főleg a következtetések gyakran megkérdőjelezhetők. Mint már említettük, az ember gyorsan észleli a szimmetriát, érzékenyebb a függőleges és vízszintes tengelyű szimmetriákra, mint a ferdekre. Tudjuk, hogy a többszörös szimmetria felgyorsítja a tárgy percepcióját. LOCHER és NODINE (1973) a szemmozgást vizsgálták, midőn a kísérleti alany szimmetrikus és nem szimmetrikus árnyképeket nézett. Azt figyelték meg, hogy a fixációs pontok főleg sarkoknál vagy görbületi találkozáspontoknál voltak, és az alany a szimmetrikus tárgyakat majdnem kizárólagosan csak a szimmetriatengely egyik oldalán pásztázta végig, míg a nem szimmetrikusakat körben a kontúr mentén. Ebből arra következtettek, hogy az ember a szimmetriát fixáció előtt, és anélkül képes észlelni. PALMER és HEMENWAY (1978) azonban megfigyelték, hogy majdnem szimmetrikus tárgyak esetén a percepció gyorsabb, mint bilaterális szimmetria esetén, valamint hogy ha a szimmetriatengely is látható, a percepció időtartama lényegesen lecsökken. Mikor

FISHER és FRACASSO (1987) a gyermekek szimmetriaítéletének környezeti, retinális és fenomenológiai befolyásolhatóságát vizsgálták, úgy találták, hogy a jelenség nagyon összetett. Szerintük habár az ember érzékeli a szimmetrikus struktúra információit, a döntés, hogy valami „szimmetrikus”, az élet során fejlődik, és ezalatt eléggé sokat változik.

Összességében úgy gondolom, hogy nem a szimmetriát érzékeljük *per se* primitívként, hanem az inkább valódi primitívek (sarkok, csatlakozási pontok stb.) bizonyos konstellációjaként jelentkezik. Egy analógiát kínálva habár rendkívül érzékenyek vagyunk egy UV-fényes sötét szobában a fehér foltokra, mégsem állítaná senki, hogy akár az UV-fény, akár a fehér foltok lényegi szerepet játszanának látórendszerünkben.

3.3. ELKÉPZELÉSEK AZ ORIENTÁCIÓT FELDOLGOZÓ MECHANIZMUSSEL KAPCSOLATBAN

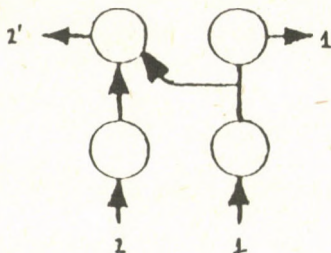
Ellentétben a szimmetriával, az orientáció nagyon lényeges alapprimítívnek tűnik az alaklásban. A fő probléma, hogy az orientációfeldolgozás bonyolultságát hosszú ideig alulbecsülték. Most 3 olyan újabb megközelítést vázolunk fel, melyek a kérdést körültekintőbben tárgyalják.

DICK és HOCHSTEIN (1988) információelméleti oldalról közelítették meg az orientáció és hosszbecslés és -diszkrimináció kérdését. Azt az ismert információelméleti tényt használták fel, hogy ha két zajos csatorna kapcsolatban áll egymással, akkor az átvitt információ mennyisége a két csatorna közötti korreláció függvénye. Bizonyították, hogy az orientáció és a hossz között aszimmetrikus reláció van. Az orientáció primitív, és hozzájárul a hosszdiszkriminációhoz és -becsléshez. A hossz nem primitív, a hosszfeldolgozás valószínűleg végpont-, mélység- és orientációinformációt hasznosít. A 23. ábra mutatja a hossz és az orientáció közötti feltételezett kapcsolatot. Az 1. dimenzió az orientáció, amely függetlenül és így elkülönülve dolgozódik fel. A hossz a 2' output, amely több dimenzióból integrálódik.

VOGELS és ORBAN (1987) vizsgálataikat von der HEYDT (1984) eredményeire alapozták. Feltételezték, hogy a valódi kontúrok orientáció-diszkriminációját V1 S-sejtjei végzik, minthogy az S-sejtek sávszélesség, orientáció-anizotrópia és excentritás tulajdonságai jól egybeesnek az emberi percepció hasonló tulajdonságaival. Eme feltételezés ellenőrzésére épphogy észrevehető különbségek (ÉÉK) mérését alkalmazták illuzórikus kontúrok orientációjában. Ez azt jelenti, hogy „pszichofizikailag kivágták” a V1 S-sejtjeit, minthogy ezekről tudjuk, hogy nem reagálnak illuzórikus kontúrokra. Két lényegi eredményt értek el. Az első, hogy az ÉÉK-ra

A hossz és az orientáció feldolgozását mutató aszimmetrikus modell

A kísérleti alany nem képes figyelmen kívül hagyni az orientációt a hossz megítélése közben (Dick and Hochstein 1988).



adott válaszfüggvények eléggé hasonlítottak valódi és illuzórikus kontúrok esetén, és az amplitúdóban is kevesebb mint egy kettes faktor volt az eltérés. A második, hogy valódi, kontúrokkal való gyakoroltatás után csak a valódi kontúrok esetében javult a felbontási képesség az ÉÉK-k észlelésében, míg illuzórikus kontúrokkal történő gyakoroltatással mind valódi mind illuzórikus kontúrok esetén javult e képesség. Vogelsék állítása az, hogy ezen eredmények fényében az a legegyszerűbb mechanizmus, ami legalább szükséges az orientációsdiszkriminációhoz, a fent említett S-sejteket tartalmazó pályán kívül egy másikat is kell hogy tartalmazzon. Ez a második pálya a V1 végződésjelző és komplex sejtjeit felhasználva a V2-be megy, ahol ismét valamilyen orientációszelektív szűrése megy keresztül (24. ábra). Mindkét pályán magasabb szintű mechanizmusok számítják az orientációt, de a második pálya (II) is képes valódi kontúrokat kezelni. A III-as jelű magasabb szintű mechanizmus az egyéb apróbb észlelt cross-talking hatások miatt van bejelölve.

WENDEROTH és JOHNSTONE (1987) az orientációanalízis kérdését a vonal- és keret-, (VKI), illetve a dőlés-illúzió (DI) oldaláról közelítették meg. VKI akkor lép fel, ha egy függőleges vonal egy dőlt négyzetes keretben van ábrázolva, DI akkor lép fel, ha egy függőleges vonal vagy rács van egy dőlt vonalra vagy rácsra, esetleg mellé helyezve. A függőleges vonal egyik esetben sem tűnik függőlegesnek. Wenderothék egyetértének Vogelsék feltételezésével, de még széleskörűbben kezelik a problémát. Állításuk az, hogy éppúgy, ahogy az MT-ben és más areákban léteznek neuronok, melyek a globális minta mozgására reagálnak, kell hogy legyenek neuronok, melyek a minta orientációjára reagálnak, és nem

csupán a komponensvonalak irányára. Ezek a neuronok kell, hogy befolyásolják (AND-kapuzáson keresztül) a V1 sejteinek eredményeit, és ezek lennének felelősek a szimmetria hatásának jelentkezéséért. A szimmetriával ugyanis jól magyarázható a két fent említett illúzió, mégpedig úgy, hogy a vonal mindig elfelé dőlőnek tűnik attól a szimmetriatengelytől, amely a függőlegeshez a legközelebb van, és nincs torzítás, ha az adott tengelypont függőleges.

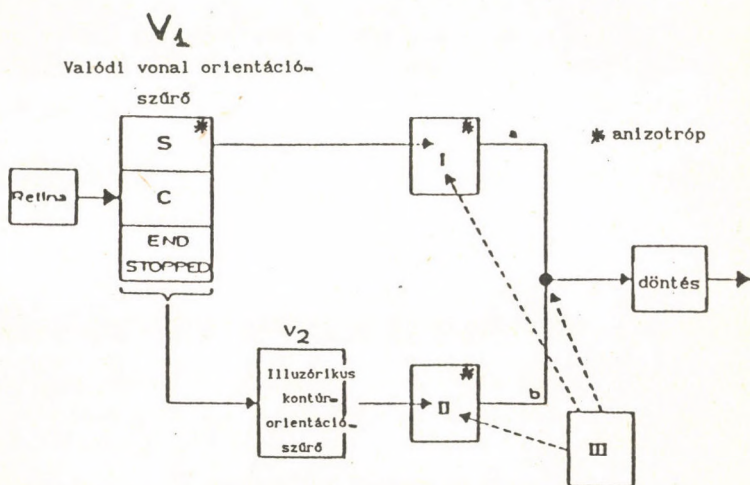
3.4. A FIGYELEM KÉRDÉSE

Mind ez ideig a vizuális információ feldolgozásának kezdeti lépéseiről beszéltünk, holott a magasabb szintekkel kapcsolatban is léteznek pszichológiai elképzelések. Pontosabban megfogalmazva a kérdéseket: Van-e magasabb szint, tovább hierarchizálható struktúra, azaz például feloszthatjuk-e a vizuális percepció folyamatát egy figyelem nélküli (preattentív) és egy figyelemirányított (attentív) részre? Beszélhetünk-e információáramlásról ezeken a szinteken? Hogyan kezeljük a különböző tulajdonságokat együttesen?

24. ábra

A Vogels és Orban által javasolt modell, amely magyarázza kísérelti eredményeket

Az anizotróp mechanizmusokat csillagok jelzik. S: S-sejtek, C: CX-sejtek, end-stop ped végződésjelző sejtek a V1-ben (Vogels and Orban 1987).



Nem tekintjük át a figyelem témakörében felmerülő összes problémát és nézetet. Ehelyett kiemelünk egy jellegzetes vonulatot, amely kétségtelen kapcsolatban van az eddig említett fiziológiai és pszichofizikai eredményekkel, és talán képet ad a figyelemmel foglalkozó elképzelések tipikus vonásairól.

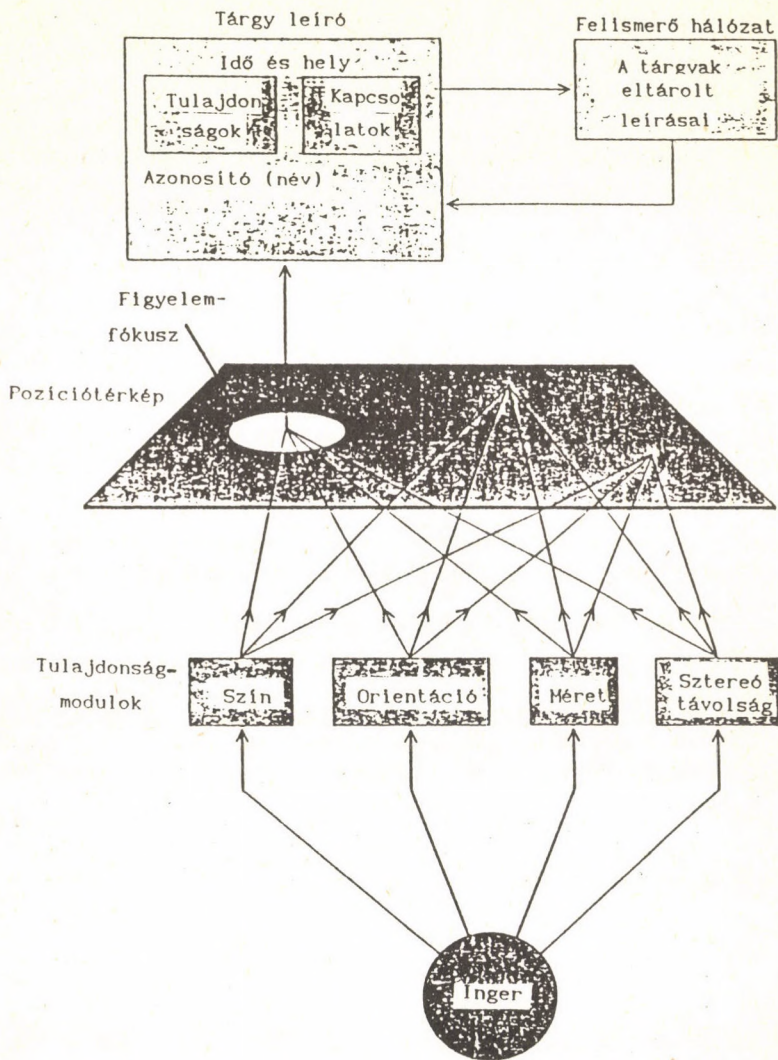
TREISMAN (1985) azt figyelte meg, hogy látórendszerünk bizonyos tulajdonságokat függetlenül és párhuzamosan dolgoz fel. „Pop-out” kísérleteiben a színt, a fényességet, a végződést és a görbületet találta ilyen tulajdonságnak. Állítása szerint viszont a metszéspontok, csatlakozási pontok és a számosság nem ilyen figyelem nélküli tulajdonságok. Megfigyelése szerint látórendszerünk kódolja a dőlést és a görbületet, de nem a függőlegességet és az egyenességet, mintha ezek „zéró-érték” voltak miatt nem ingerelnék a detektorokat. Így az egyéb esetek a referenciától való relatív eltérésként lehetnek kódolva.

Treisman másik eredménye, hogy ha két vagy több figyelem nélküli tulajdonságot összekapcsolt egy céltárgyban, a céltárgy kiválasztásához szükséges idő lineárisan nőtt az elterelő tárgyak számával, míg egy tulajdonság esetén ez az idő független volt ettől a számtól. Egy másik különbség az volt, hogy egy tulajdonság esetén előfordult, hogy a döntés jó volt (van, illetve nincs céltárgy a felvillanó képen), de a tárgy helyének azonosítása hibás volt. A kapcsolt esetben a helyes döntés teljes mértékben a jó helyazonosítás függvénye volt.

Ezen tényekre támaszkodva Treisman javaslata szerint a látás folyamatának 2 szintje van (25. ábra). Az első szinten a tulajdonságok feldolgozása függetlenül és elkülönülve történik, de nincs pontos információ arról, hogy hol is van az adott tulajdonság. A második szint -- a „mastermap” -- összehozza az összes diszkontinuitást, és a figyelem ezt a térképet hasznosítja, kiválogatva az összes tulajdonságot, ami adott pillanatban egy adott (a figyelem által) kiválasztott területen van. Még magasabb szinten a percepció a korai látási információkat egy időleges tárgyspecifikus reprezentációba integrálja.

NAKAYAMA és SILVERMAN (1986) igazolták Treisman eredményeit, és ki is terjesztették a szín és mozgás kapcsolatára. Egy kivételt is felfedeztek: ha a kapcsolt tulajdonságok egyike a sztereopikus diszparitás, a másik, legyen bár mozgás vagy szín, párhuzamosan kereshető. Úgy tűnik, hogy képesek vagyunk kutatásunkat egy sztereoszkopikus mélységre korlátozni. Az a tény, hogy az MT-ben vannak mozgásra és diszparitásra együttesen hangolható sejtek, viszont még nem találtak hasonló sejteket mozgásra és színre, szerintük támogatja ezt az elképzelést.

A Treisman által javasolt hipotetikus modell, amely a vizuális információ kétszintű feldolgozását feltételezi (Treisman 1986)



STEINMAN (1987) olyan vizuális keresések reakcióidejét vizsgálta, ahol a céltárgyak térbeli pozíció-információk kapcsolatait tartalmazzák, mint például horizontális offset, sztereopikus diszparitás, két részre különülés és orientáció (26. ábra).

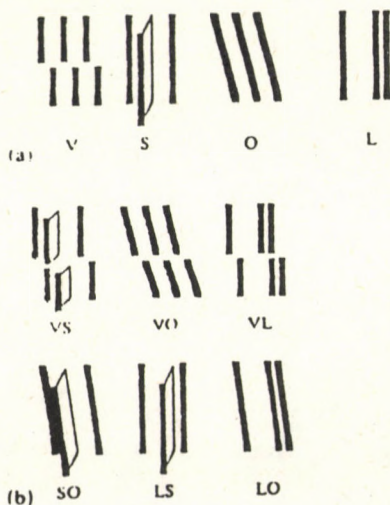
A képen a vékony vonalak a mélységbeli eltérésre utalnak. Steinman igazolta Nakayamáék sztereopszisra vonatkozó eredményét (27. ábra). Ha a kutatás párhuzamos volt, a reakcióidők meglehetősen állandók maradtak az elterelő tárgyak számbeli növekedésének dacára, viszont arányosan növekedtek soros kutatás esetén. A két részre különítést és a sztereopszist párosítva ellenben kiderült, hogy a sztereopszis mint „elkülönítő” tulajdonság működik, azaz: bár egy mélységben egy tulajdonság keresése párhuzamosan történik, különböző mélységű síkok közötti összevetés lehetetlen párhuzamosan. A mozgás hasonló elkülönítő tulajdonságnak tűnik.

Steinman második eredménye, hogy gyakorlással minden (!) inger párhuzamos kezelése megtanulható (28. ábra). Az összes tesztelt ingerpárosítás reakcióidő-görbéje ellaposodott a 27. ábrához viszonyítva, arra

26. ábra

A Steinman által használt ingerkonfigurációk:

- (a) a négyféle használt térbeli lokalizációs dimenzió: vízszintes offset (V), sztereopikus diszparitás (S), orientáció (O), kettéválasztás (L);
 (b) Két ingerdimenziót tartalmazó céltárgyak (Steinman 1988).

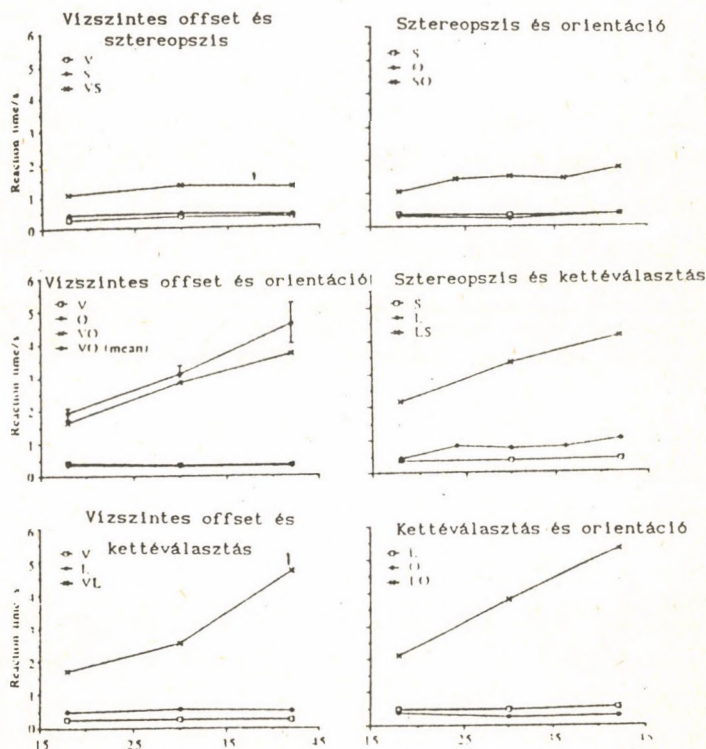


utalva, hogy a kutatási stratégia sorosból párhuzamosba tolódott el. Egy tulajdonság esetén ez gyorsan kialakult, néhány párosítás esetén csak több ezer próbálkozás után, de az eredmény nem tűnt el több hónap szünet után sem. Ezeket az eredményeket szín, méret és egyéb tulajdonságok bevonásával is ellenőrizték, hasonló eredménnyel. Steinman szerint az egyik lehetséges magyarázat az, hogy a kísérleti alany megtanulja, hogyan bontsa fel a vizuális mezőt egy adott tulajdonság szempontjából éppúgy, ahogy képes ezt megtenni tanulás nélkül sztereoszkopikus és mozgó ingerek esetében.

27. ábra

Átlagolt céltárgy-detektálási reakcióidők az elterelő tárgyak számának függvényében

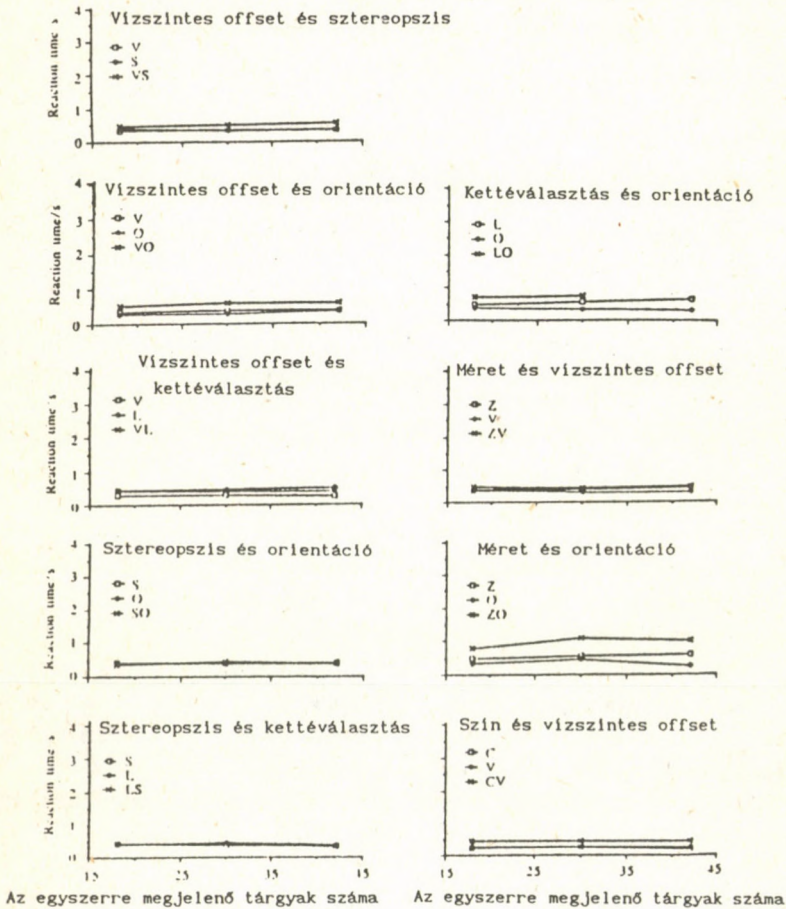
Ingerek rövidítése mint a 31. ábránál, a vízszintes tengelyen az egyszerre megjelenő tárgyak száma, a függőleges tengelyen a reakcióidő (Steinman 1988).



Az egyszerre megjelenő tárgyak száma Az egyszerre megjelenő tárgyak száma

Átlagolt céltárgy-detektálási reakcióidők az elterelő tárgyak számának függvényében hosszabb gyakorlás után

Ha színt (C), illetve méretet (Z) párosítottak egyéb dimenziókkal, az ellaposodási effektus ugyanúgy fellépett (Steinman 1988).



Összességében úgy tűnik, hogy ezen eredményeknek még nem nagyon körvonalazódott egy olyan közös magja, amely a fiziológiai tényekkel összhangban van és emellett önmagában is megalapozott gondolatmenettel rendelkezik. Például a Treisman által javasolt struktúrának a maga egyszerűségében még néhány jól ismert alapösszefüggés-magyarázata is nehézséget okoz.

4. KÖVETKEZMÉNYEK A GÉPI LÁTÁSKUTATÁS TERÜLETÉN

Ebben a fejezetben először összegzem azokat a tényeket, melyek az előző fejezetekből következően a látórendszeréről tudunk. Ezek után rátérek arra, hogy mindezen tudás birtokában véleményem szerint milyen problémák milyen megközelítése indokolt annak érdekében, hogy jobb minőségű alakfeldolgozó rendszerhez jussunk.

4.1. A NEUROFIZIOLÓGIAI ÉS PSZICHOFIZIKAI TÉNYEK ÖSSZEFOGLALÁSA

Azt mondhatjuk, hogy nem túl sok, ugyanakkor rengeteg dolog történt MARR (1982) alapművének megjelenése óta. Nem túl sok, mert ő az alapvető primitíveket majdnem teljesen definiálta könyvében, mert ezeknek a következő szinten történő összefonódásáról azóta sem tudunk egyértelműen újat mondani, és éppen ezért (sőt ettől függetlenül sem) azóta sem tudtunk az éldetektálás szintjénél feljebb lépni. És mégis rengeteg, mert a látás folyamatának struktúrájáról és a struktúrát alkotó részek tulajdonságairól gyűjtött információink megtöbbszöröződtek, és mert éppen ebből fakadóan a probléma megfogalmazásában folytonos szemléletbeli változáson megyünk keresztül. Az előző fejezetekben leírtakból következő főbb gondolatokat 4 csoportban foglalom össze.

Az első, amit ezen eredményekből kiemelnék, hogy a látórendszer egy igen bonyolult mechanizmus, és éppen ezért sem modellezésében, sem az általa elvégzett feladatot megoldó rendszer tervezésében az egyszerű „huszáros” megoldások nem segítenek. Nincs esélyük az olyan elméleteknek, amelyek némi szűrés után egyenlőséget vagy bármilyen hasonlót alkalmaznak, és utána helyes alakpercepciót várnak. A sejtek az inger igen sok tulajdonságával foglalkoznak, és habár tudásunk igen töredékes, az máris kiderült, hogy már a legalacsonyabb szinten is igen bonyolult módon teszik ezt. Kétségtelen, hogy nem lehet és nem is szabad minden kísérleti eredményt beépíteni a modellbe. Például Treisman, Nakayama és Steinman kísérleteiben az ellentmondások okai nagyon távolra is mutathatnak. Mégis úgy tűnik, hogy lassan tisztázódik egy olyan bonyolultsági fok, amely óhatatlanul is minden olyan rendszer sajátja lesz, amely sikerrel teljesíti azt, és csakis azt a feladatot, amit formalításnak hívunk.

A második gondolatkör, hogy dacára annak, hogy „minden mindenel összefügg”, mégis világos elkülönítés lehetséges bizonyos részek között. A különböző információs áramok csak különböző pontokon, definiált csatornákon kommunikálnak egymással. Ezen csatornáknak sem definíciójuk, sem az elkülönültségi fokuk, sem funkciójuk nincs megnyugtatóan tisztázva még. Mindenesetre a vizuális folyamatnak az általuk sugallt

strukturális lebontása jelenti a megértés első fokozatát. A folyamat alapjainál nincs „rádöbbenés”, nincs „Gestalt-érzet”, hanem immár jól definiált tulajdonságok fokozódó bonyolultságú feldolgozása. A bonyolultabb percepciók jelenségek kielégítő funkcionális magyarázata még jócskán odébb van, de kétségbe vonható, hogy egy homogén struktúrájú cortexben ez a működési elv valahol megváltozna.

A harmadik gondolatcsoport azon eredményekre vonatkozik, melyek az egyes pályák mindegyik szintjén az ottani sejtek fontos tulajdonságairól, vagy az adott szint absztrakt tulajdonságairól szólnak. Tudásunk itt a leg-töredékesebb, a 2.6.5. fejezetben már említett hatalmas úr a V1—V2 és az IT szintjei között komoly információhiányról tanúskodik. Érdekes kiemelni az orientáció meghatározó szerepét a formalításban. Az illuzórikus kontúrok felhasználása és az őket feldolgozó mechanizmus talán az első komolyabb információ, ami az egyik szintről a másikra lépésnél felmerülő változásokra mintaként szolgálhat. A különböző „primitív-jelöltekkel” (mint például pontok, végpontok, sarkok és görbület-szélsőértékek) az alapvető probléma az, hogy mind térben, mind összekapcsolódásban lényegesen szélesebb kontextusban kerülnek feldolgozásra, mint azt sejtettük. Egyéb jelöltek, mint a szimmetria, a párhuzamosság, úgy tűnik, hogy nem primitívek ugyanolyan értelemben, mint az első csoport elemei. Ezek valószínűleg kiemelnek olyan tényezőket, amelyek fontosak a vizuális folyamatban. Éppen ezért az alapvető erőfeszítést nem annak érdekében kell tenni, hogy ezeket egy képen képesek legyünk detektálni, hanem hogy megértsük azokat a tényezőket, melyek kihangsúlyozása miatt érzékenyek vagyunk rájuk.

A negyedik csoport a gyakorlati technikai ötletek, melyek a funkciók megvalósításában játszanak szerepet. Az egyik ilyen az RM-ek mérete és a környezeti mechanizmusok, melyek valószínűleg lényegiek azon komplex válaszokban, melyeket mind ez idáig sem értelmezni, sem mesterséges rendszerben előállítani nem tudtunk. A másik a reciprok kapcsolatok az areák között, melyek az sugallják, hogy a „vedd az első szűrőt, alkalmazd az I tömbön, vedd a másodikat és alkalmazd az eredménytömbön...” típusú megközelítések nem alkalmasak ilyen folyamatok modellezésére. Több egyidőben létező reprezentáció szükséges erős interakciókkal közöttük.

4.2. A GÉPI LÁTÁS KÉRDÉSÉNEK MEGKÖZELÍTÉSE

Az alapvető probléma a gépi látásban, hogy nem tudjuk, mi is a látás valójában. Amikor megfogalmazunk egy úgymond látással kapcsolatos feladatot, általában két szélsőséges eredmény valamelyikét kapjuk. Az egyik

egy olyan szűkített feladat, amelyben kihasználva az eset specifikus tulajdonságait egy olyan feladatba transzformálható, mely már jól ismert, standard eszközökkel megoldható, viszont vajmi kevésbé vetekedhet képességeiben az emberi látás eredményeivel. A másik egy olyan nagyigényű általános feladat, mely aztán dominószerűen magával rántja a tanulás, a gondolkodás és a memória problémáit, és így megoldhatatlan. Átmeneti megfogalmazás mindaddig nem nagyon sikerült, és ez erősíti a gyanút, hogy ez nem véletlen.

Tudásunk a látás folyamatából töredékes, nem elégséges. Ugyanekkor a célunk egy teljes és működő látó rendszert alkotni. Ez az ellentmondás arra csábít, hogy a kérdést „műszakilag” közelítsük meg és a tudás, valamint a cél közötti lyukat bármilyen alkalmasnak látszó eszközzel betömjük. Csakhogy már bizonyítást nyert (TSOTSOS, 1988), hogy a látásnak akár alproblémái is NP-teljesek, vagyis bonyolultságuk miatt a „nyers erőszak” megközelítései szükségszerűen csődöt mondanak. Ugyanitt Tsotsos bizonyítja, hogy a párhuzamos rendszerek, melyek jelen pillanatban az érdeklődés homlokterében állnak (például BALLARD és mts., 1983; LIPPMAN, 1987), bár feltétlen szükségesek, önmagukban nem elégségesek a számítási komplexitás olyan mérvű csökkentésére, ami szükséges egy látási probléma megoldásához. Algoritmikus és strukturális hiányosságok vannak. Végül tisztán kell látni, hogy a „különálló csatornák” vizsgálatai nem a látásfolyamat elkülönülő, teljesen független részekre való „vagdosásához” segítenek hozzá, hanem a részt vevő változók együttműködésének és eme együttműködés részeredményeinek, részleteinek jobb megértéséhez. A különböző paraméterek együttműködésének magas szintje már a folyamat legelején is teljesen nyilvánvaló.

Két gondolat következik a fentiekből. Az első, hogy a látás problémájának definiálásában a külső jegyek helyett (mint például képekkel foglalkozik, kiválasztja a pirosat, sztereó kamera az input eszköz) a feladat bonyolultsága kell hogy domináljon. Azok valódi látásproblémák, melyek egy adott komplexitáosztályba sorolhatóak. Ezek alapján - bár abszurdul hangzik -- a hallás és a beszédértés inkább látásprobléma, mint az ipari kamera segítségével történő selejtválogatás. A második gondolat, hogy látási folyamatnak egy „meglehetősen” teljes modellezése szükséges. A homályos „meglehetősen” szó pontosan a feladat bonyolultsági szintjének meghatározásával tehető explicitté. Az alacsonyabb bonyolultságú, ún. computer vision célfeladatok megoldására alkotott eljárások tapasztalatai mind ez ideig többségükben kívül maradtak a látás problémájában hasznosítható tapasztalatok körén.

Egyes vélemények szerint felesleges az orientációval, a sztereopszissal, a színnel és mozgással mind egyszerre foglalkozni, hisz az ember álló kontúrrajzokat fél szemmel is jól lát. Olyan értelemben igazuk van, hogy álló kontúrrajzokra is megfogalmazható egy a látás problémájával azonos bonyolultságosztályú feladat (anélkül, hogy ezt a bonyolultságot pontosan ismernénk). Hogy szerintem miért érdemes a feladat megfogalmazásában mégis a valóság közelében maradni, a következőkben indoklom.

Úgy gondolom, hogy érdemes a valódi látórendszer struktúráját követni. Nem az idegrendszer, az agy jelenségeinek kritika nélküli utánzása okán, hanem mert nem ismerjük a látást. Ennek megfelelően a kívánt output definiálásában a definiáló által fontosnak tartott néhány kritérium teljesítésén túl szabad a vásár. (Az output olyan pontos és reális definíciója, mely az ismert kritériumok túlnyomó többségét — valós időben elérhetőség, általánosító készség, különböző típusú információk együttes adaptív kezelése stb. — teljesíteni tudná, még nincs). Egy ilyen modellben, ha az inputnak, az eseményeknek és szabályoknak, valamint a végeredménynek csak erős áttételeken keresztül van kapcsolata a modellezendő folyamattal, sokkal több esély van az elkalandozásra és sokkal kevesebb a rejtett összefüggések megértésére a valósággal észlelt párhuzam kapcsán.

Jelen pillanatban úgy tűnik, a hierarchikus felépítés, a párhuzamos feldolgozás és egyidejű reprezentálás különböző, ún. „feature-map”-ekben, az erős visszacsatolás, a receptív mezők növekedése és a „feature”-ök növekvő absztrakciója a hierarchiában felfelé haladva — ezek azok a fő ismérvek, melyeknek szerepelniük kell egy megfelelő modellben. Tsotsos továbbmegy ennél az állításnál. Két alapelvet vezet be. Az első a modellek „komplexitási szint” szerinti analízise, azaz annak vizsgálata, hogy az adott modell kielégít-e bizonyos térbeli és működésbeli alapkövetelményeket. A második a maximális teljesítmény/minimális ár elv, amely rangsorolja az első elvet teljesítő modelleket. Tsotsos szerint az ezen szempontok szerint kiválasztható legjobb architektúra teljesen kompatibilis az emberi látórendszer megismert architektúrájával (TSOTSOS, 1988). A 29. ábrán, amely ezt az architektúrát mutatja, a prototípusok bináris fába szervezett tudásbázisa látható felül. A középen lévő processzorréteg minden processzora egy prototípussal van kapcsolatban, és egy vagy több „map”-ben receptív mezőkkel bír. Az ábra alsó része az absztrakciós hierarchia kezdeti szakaszát mutatja logikailag elkülönülő reprezentációkkal és felfelé csökkenő tendenciájú felbontási tulajdonsággal. Tsotsos modellje nem feltétlenül tökéletes, ezen kívül az architektúra mellett a működési elvről keveset mond. Mégis jól kezelhető keretet ad a továbblépéshez,

mert a struktúra jól ellenőrizhetően elvi lehetőséget hagy egy működőképes rendszer fejlesztésére.

Elérkeztünk ama végső kérdéshez, hogy mi módon érdemes figyelembe venni az idáig megemlített eredményeket a gépi látáskutatás problémafelvetéseiben. Úgy gondolom, a választ a mit?, mivé? és a hogyan? problémakörök köré lehet csoportosítani. A hogyan? kérdését részben már körüljártuk a feladat komplexitásának és az ebből fakadó strukturális megkötések taglalásával. Az e cikkben is megemlített fiziológiai eredmények alátámasztják a hierarchikus architektúra feltevését. A másik feladat ebben a körben a dinamikus, visszacsatolt nem lineáris működés vizsgálata ilyen struktúrákban. Csak annyit tudunk, hogy az agy dinamikusan működik. Hogy ennek a működésnek milyen aspektusai lényegesek, és azok hogyan tehetők át esetlegesen egy nem dinamikus rendszerre, nyitott kérdés.

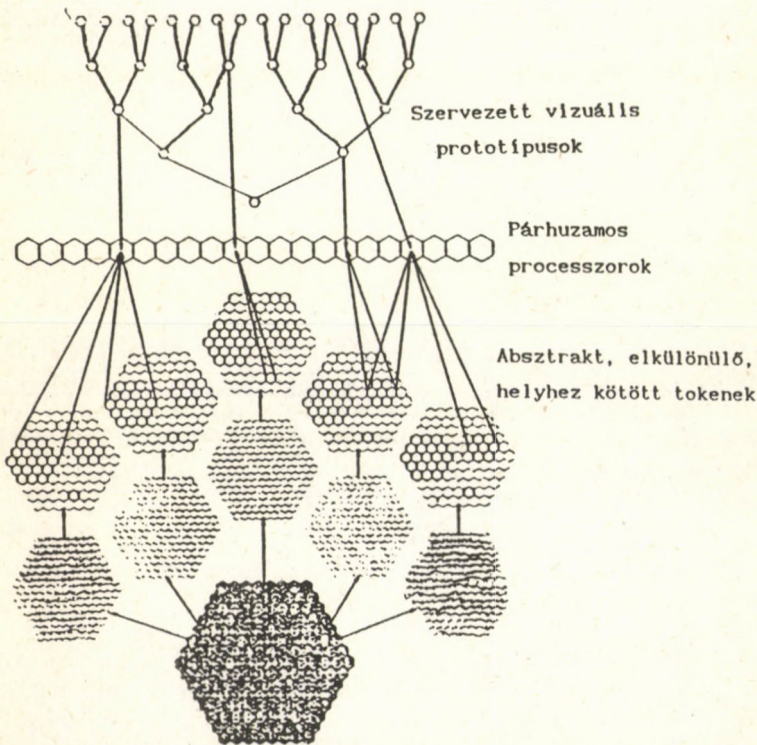
A mivé? kérdésköre a legködösebb a három közül. A feladat itt a tudás, az információ olyan rendszerbe szervezése, amelyből egy definiált korlátok között minden esetben az adott szituációnak megfelelő válasz kinyerése rövid idő alatt lehetséges. A fiziológia és pszichofizika itt szolgáltatja a legkevesebb használható kísérleti eredményt. Az ide tartozó tanulással, memóriával, kognitív modellekkel kapcsolatos szerteágazó kutatások taglalása nem e cikk feladata. Két gondolat van, amit a gépi látás szempontjából fontosnak tartok. Az első, hogy amíg ezen a területen meggyőző eredményeket nem érünk el, addig valódi értelemben vett látásról nem beszélhetünk. A látás csak a „magasabb” agyi folyamatokkal az, ami. Éles szétválasztó vonal az ún. early vision és a kognitív rész között úgy tűnik, nincs, mint erre MAUNSELLék (1987) kísérletei is utalnak (ti. hogy egészen a V1-ig visszamenőleg befolyásolódik a sejtek viselkedése különböző extrastriális jelek hatására). A második, hogy az előző gondolat egyáltalán nem jelenti azt, hogy a látás problémája kizárólag ennek a problémakörnek a feldolgozásával megoldottá válik. A hiányzó aspektus pontosan a mit? kérdéskör.

Úgy vélem, hogy téves az az elképzelés, mely szerint elegendő egy megfelelő tanulóalgoritmust kutatni, mert ha az meglesz, akkor majd az inputnak képeket adunk, és az algoritmus kifejleszti az összes köztes feature-t és reprezentációt a végsővel egyetemben. Az általam említett „teljes” modellezést nem ebben a megközelítésben látom. Szerintem a látás modellezésében a mivé? területén nem lesz lényegi elmozdulás addig, míg a mit? területén további ismeretek birtokába nem jutunk. Egyrészt mert ha a látást akarjuk modellezni, a látvány által hordott információ ugyanazon aspektusait kell kezelnünk, mint amit a valós látórendszer, ezeket viszont nem ismerjük. Nem hiszem, hogy az információelméleti értelemben

vett információ veszteségének valamilyen általános (négyzetes vagy egyéb) szempontú minimalizálása megfelelő vezérelvként behelyettesíthető lenne. Másrészt nincs semmilyen megfelelő gondolati modellünk sem a mivé? kérdéskörben. Nem tudjuk, hogy egy kabát észlelésekor annak konkrét lá-
tott tulajdonságai vagy ezen kívül használati tulajdonságai vagy minden kabáttal kapcsolatos valamikori információnk aktivizálódik, és ha bármelyikre a válasz igen, akkor az milyen összetevőket is jelent. Ha a látás le-
galacsonyabb szintjén az inputok egy magasabb (de nem legmagasabb) szinten történő feldolgozására modelleket kapunk, az példaként, útmuta-
tóként használható fel ezeken a még magasabb szinteken.

29. ábra

**Egy architektúra a figyelem nélküli látás modellezésére, amely
teljesíti az idő-megkötéseket egyszeri bottom-up feldolgozás
esetén (Tsotsos 1988)**



Tehát a mit? kérdéskörben a feladat a hogyan?-ban meghatározott kötések kielégítő modellstruktúrában annak vizsgálata, hogy a már ismert „primitívek” (orientáció, fényesség, szín, térfrekvencia stb.) milyen kapcsolódásban alkotnak feature-öket magasabb szinten. Az alaklátásban a problémák a V1—V2 szintjén vannak. Magasabb szintre lépni addig nem lehet, míg a kapcsolatok itt nincsenek tisztázva és megfelelő módon a modellbe építve. Például az orientáció információfeldolgozásáról már most is sokkal többet tudunk annál, mint amennyi a használt képfeldolgozó rendszerek többségébe be van építve, bár néhány részletbemenőbb megközelítés létezik (például ZUCKER, 1985, GROSSBERG és TODOROVIC, 1988). Ezen a területen lehet az idegtudományok eredményeit a leginkább hasznosítani. Az alapprimítívek köre már megbízhatóan stabilizálódott. A következő szintek alkotóelemeinek meghatározásán munkálkodást tartom a gépi látáskutatás olyan alapvető feladatának, mellyel a többi terület kutatásai nem foglalkoznak. Hangsúlyozom azonban, hogy a másik két kérdés (mivé? hogyan?) területeivel és eredményeivel a kapcsolat legalább olyan szoros, mint a kissejtes és a nagysejtes részek között a látás agyi folyamatában.

Köszönetnyilvánítás:

Köszönettel tartozom Rufin Vogelsnek, amiért a neurofiziológia területén hatékony irányítást nyújtott és fontos észrevételeivel segítette ezen cikk megszületését.

A kézirat elfogadva: 1990. május

IRODALOM

- ALLMAN, J., MIEZIN, F., MCGUINESS, E., 1985, Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons, *Annual Review of Neurosci.*, vol. 8, 407—430.
- BALLARD, D. H., HINTON, G. E., SEJNOWSKI, T. J., 1983, *Parallel Visual Computation Nature*, vol. 306, 21—26.
- BIEDERMAN, I., 1985, Human Image Understanding: Recent Research and a Theory, *Comp. Vision, Graphics and Image Proc.*, vol. 32, 29—73.
- CHANGEUX, J. P., 1986, *Neural Man*, Oxford Univ. Press, 348.
- CREUTZFELDT, O. D., WEBER, H., TANAKA, M., LEE, B. B., 1987, Neural representation and spatial stimulus aspects in foveal and parafoveal area 17 of the awake monkey, *Exp. Braines.*, vol. 68, 541—564.
- DICK, M., HOCHSTEIN, S., 1988, Interactions in the discrimination and absolute judgement of orientation and length, *Perception*, vol. 17, 1988, 177—189.

- DESIMONE, R., SCHEIN, S., 1987, Visual properties of neurons in Area V4 of the macaque Sensitivity to stimulus form, *Jou. of Neurophysiology*, vol. 57, No. 3, 835—868.
- DESIMONE, R., SCHEIN, S. J., MORAN, J., UNDERLEIDER, L. G., 1985, Contour, color and shape analysis beyond the striate cortex, *Brain Res.*, vol. 25., No. 3, 441—452.
- DE VALOIS, R. L., YUND, E. W., HEPLER, N., 1982, The orientation and direction selectivity of cells in macaque visual cortex, *Vision Res.*, vol. 22, 531—544.
- GROSSBERG, S., TODOROVIC, D., 1988, Neural dynamics of 1D and 2D brightness perception: A unified model of classical and recent phenomena, *Perception and Psychophysics*, vol. 43, 241—277.
- FISHER, C. B., FRACASSO, M. P., 1987, The Goldmeier effect in adults and children environmental, retinal, and phenomenal influences on judgements of visual symmetry, *Perception*, vol. 16, 29—39.
- VON DER HEYDT, R., PETERHANS, E., BAUMGARTNER, G., 1984, Illusory Contours and Cortical Neuron Responses, *Science*, vol. 224, 1260—1262.
- HUBEL, D. H., LIVINGSTONE, M. S., 1987, Connections Between Layer 4B of Area 17 and the Thick Cytochrome Oxidase Stripes of Area 18 in the Squirrel Monkey, *The Jou. of Neurosci.*, vol. 7. No. 11, 3371—3377.
- HUBEL, D. H., LIVINGSTONE, M. S., 1987, Segregation of Form, Color and Stereopsis in Primate Area 18, *The Jou. of Neurosci.*, vol. 7. No. 11, 3378—3415.
- IWAI, E., 1985, Neurophysiological basis of pattern vision in macaque monkeys, *Brain Res.*, vol. 25. No. 3, 425—439.
- JULESZ, B., 1981, Textons, the elements of texture perception, and their interactions, *Nature*, vol. 290, 1981, 91—97.
- LIVINGSTONE, M. S., HUBEL, D. H., 1987, Physophysical Evidence for Separate Channels for the Perception of Form, Color, Movement and Depth, *The Jou. of Neurosci.*, vol. 7. No. 11, 3416—3468.
- LIPPMANN, R. P., 1987, An Introduction to Computing with Neural Nets, *IEEE ASSP Magazine*, vol. 3. No. 4, 4—22.
- LOCHER, P., NODINE, C., 1973, Influence of Stimulus Symmetry on Visual Scanning Patterns, *Perception and Psychophysics*, 73., vol. 13, No. 3, 408—412.
- MARR, D., 1982, *Vision*, Freeman and Co., San Francisco, Ca.
- MAUNSELL, J. H. R., NEWSOME, W. T., 1987, Visual processing in monkey extrastriate cortex, *Ann. Review of Neurosci.*, vol. 10, 363—401.
- K. NAKAYAMA, G. H. SILVERMAN, 1986, Serial and parallel processing of visual feature conjunctions, *Nature*, vol. 320, 264—265.
- PALMER, S. E., HEMENWAY, K., 1978, Orientation and Symmetry Effects of Multiple, Rotational, and Near Symmetries, *Jou. of Exp. Psychology Human Perc. and Perf.*, vol. 4., No. 4, 691—702.
- SCHILLER, P. H., 1989, *Talk held at Dept. Neurophysiology, Gasthuisberg, Catholic University of Leuven, Belgium, on 04. 10. 1989.*

- SCHILLER, P. H., FINLAY, B. L., VOLMAN, S. F., 1976a, Quantitative studies of single-cell properties in monkey striate cortex I., *Jou. of Neurophysiology*, vol. 39, No. 6, 1288—1319.
- SCHILLER, P. H., FINLAY, B. L., VOLMAN, S. F., 1976b, Quantitative studies of single-cell properties in monkey striate cortex II., *Jou. of Neurophysiology*, vol. 39., No. 6, 1320—1333.
- SCHILLER, P. H., FINLAY, B. L., VOLMAN, S. F., 1976c, Quantitative studies of single-cell properties in monkey striate cortex III., *Jou. of Neurophysiology*, vol. 39., No. 6, 1334—1351.
- SCHILLER, P. H., LOGOTHETIS, N. K., CHARLES, E. R., 1990, Functions of the colour-opponent and broad-band channels of the visual system, *Nature*, vol. 343, 68—70.
- STEINMAN, S. B., 1987, Serial and parallel search in pattern vision?, *Perception*, vol. 16, 389—398.
- STEVENS, K. A., BROOKES, A., 1988, The concave cusp as a determiner of figure-ground, *Perception*, vol. 17, 35—42.
- TANAKA, M., WEBER, H., CREUTZFELDT, O. D., 1986, Visual properties and spatial distribution of neurones in the visual association area on the prelunate gyrus of the awake monkey, *Exp. Brain Res.*, vol. 65, 11—37.
- TREISMAN, A., 1985, Preattentive Processing in Vision, *Computer Vision, Graphics and Image Processing*, vol. 31., 156—177.
- TREISMAN, A., 1986, Features and Objects in Visual Processing, *Scientific American*, 106—115.
- TSOTSOS, J. K., 1988, A 'Complexity Level' of Immediate Vision, *Int. Jou. of Computer Vision*, vol. 1. No. 4, 303—320.
- VAN ESSEN, D. C., 1985, Functional Organization of Primate Visual Cortex, in: PETERS, A., JONES, E. G., *Cerebral Cortex*, N. Y. Plenum Press, 259—329.
- VOGELS, R., ORBAN, G. A., 1987, Illusory contour orientation discrimination, *Vision Res.*, vol. 27, No. 3, 453—467.
- WENDEROTH, P., JOHNSTONE, S., 1987, Possible neural substrates for orientation analysis and perception, *Perception*, vol. 16, 693—709.
- ZEKI, S., SHIPP, S., 1988, The functional logic of cortical connections, *Nature*, vol. 335, 22. Sept., 311—317.
- ZUCKER, S. W., DAVIS, S., 1988, Points and endpoints a size/spacing constraint for dot grouping, *Perception*, vol. 17, 229—247.
- ZUCKER, S. W., 1985, Early orientation selection Tangent fields and the dimensionality of their support, *Computer Vision, Graphics, and Image Proc.*, vol. 32, 74—103.

FORM-PERCEPTION: NEUROPHYSIOLOGICAL AND PSYCHOPHYSICAL EVIDENCE AND THEIR EFFECT ON COMPUTER VISION

We consider recent results both neurophysiological and psychophysical on human form-perception and suggest what strategies should be applied to make use of these results in order to develop more efficient computer vision systems.

In the first section after a very brief review of the general properties of the cerebral cortex the visual cortex is considered in details. We examine the division of the areas, their connections to each other and the functional hierarchy that can be drawn from these results. Dividing the areas into different pathways we take a closer look at the areas which are concerned primarily with the color-form pathway. We investigate the properties of the cells, their organization and their possible function starting with V1 up to the subareas of AIT and PIT.

In the second section first we attempt to summarize psychophysical evidence in support of the idea of separate channels in the vision pathway. Next we consider some suggestions for „primitives” in form-perception, and some proposals for possible extracting mechanisms of those. Finally we mention some open questions of higher processes in vision and some initial proposals solving them.

In the last section we sum up those evidence we feel to be important to influent the computer vision researchers' mind and argue for the necessity of somewhat more complex approaches in this area.