

Mozgové štruktúry a mechanizmy riadenia strategickej pozornosti

Rudolf Andoga a Norbert Kopčo

Katedra kybernetiky a umelej inteligencie, FEI TU Košice

Email: andoga@neuron.tuke.sk, kopco@bu.edu

Abstrakt

V každodennom prostredí sú zmysly človeka nepretržite vystavené množstvu zložitých, súčasne prichádzajúcich podnetov. Strategická pozornosť nám umožňuje efektívne sa orientovať v takomto prostredí, a to optimálnym výberom behaviorálne dôležitých podnetov a sústredením limitovanej kognitívnej kapacity mozgu na ich spracovanie. Cieľom tohto príspevku je poskytnúť prehľad o neurálnych štruktúrach podieľajúcich sa na riadení strategickej pozornosti s dôrazom na vizuálne, sluchové, a krosmodálne mechanizmy riadenia pozornosti. Článok taktiež sumarizuje existujúce modely strategickej pozornosti, zdôrazňuje otvorené otázky a navrhuje alternatívne hypotézy, ktoré môžu byť východiskom pre budúce experimentálne a teoretické štúdie riadenia strategickej pozornosti.

1. Úvod

Vnímanie a reprezentácia priestoru sú dôležitými kognitívnymi schopnosťami živých a umelých organizmov (Colby, 1999), nevyhnutnými pre plynulú orientáciu, pohyb a interakciu organizmu s okolím. Človek používa na orientáciu v priestore primárne zrak. Ale aj ostatné zmysly (sluch, hmat atď.) prispievajú k priestorovému vnímaniu a spôsob, akým si vyberáme, ktorej modalite dáme prednosť, alebo ako informácie z rôznych modalít kombinujeme, je predmetom intenzívneho výskumu.

Vnímanie a správanie sa živočíchov a človeka nezáleží len na priamej stimulácii jednotlivých sensorických orgánov, ale aj na tom, ktorým aspektom tejto stimulácie v tom ktorom momente venujeme pozornosť (Spence a Driver, 2004). Pojem „selektívna pozornosť“ popisuje tie procesy, ktoré nám umožňujú selektívne spracovávať niektoré z prichádzajúcich sensorických podnetov tak, aby sa informácie dôležité pre našu aktivitu, alebo podnety s veľkým biologickým významom, spracovali detailnejšie než ostatné prichádzajúce podnety.

Tradične sa výskum selektívnej pozornosti zameriaval zvlášť na jednotlivé sensorické modality, počnúc klasickými experimentmi študujúcimi selektívne počúvanie (Cherry, 1953; Broadbent, 1958) a pokračujúc štúdiom vizuálneho vnímania (napr. LaBerge, 1995). V reálnych prostrediach je ale veľká väčšina podnetov vnímaná multimodálne. Napríklad aj v klasickom príklade sluchovej selektívnej pozornosti – v tzv. koktail párty situácii, keď sa snažíme počúvať jednu z niekoľkých súbežných konverzácií (Cherry, 1953) máme k dispozícii nie len sluchové podnety, ale aj vizuálne podnety (napr. pohyb perí, výraz tváre a gestá hovoriaceho, na ktorého sa sústredíme). Navyše, informačná hodnota jednotlivých modalít sa môže dynamicky meniť, napr. keď sa zvýši hlučnosť v miestnosti (kedy by sme sa mali viac sústrediť na vizuálne podnety) alebo keď sa tvár hovoriaceho na chvíľu stratí (kedy sa zas musíme sústrediť na to, čo počujeme). Napriek tomu ale prevládajúci model selektívnej pozornosti sedemdesiatych a osemdesiatych

rokokoch uvažoval o každom vnemovom systéme ako o samostatnom výpočtovom prostriedku s obmedzenou kapacitou, pričom pozornosť tu slúži na riadenie pridelovania výpočtovej kapacity pre spracovanie najdôležitejšieho z prichádzajúcich podnetov. Čiže, pozornosť je špecifická pre danú vnemovú modalitu a nezávislá od ostatných modalít (napr. Treisman and Davies, 1973; Soto-Faraco and Spence, 2002).

Že sa o pozornosti nedá rozmýšľať ako o súbore nezávislých mechanizmov pracujúcich oddelene pre každú vnemovú doménu ukázali ako prví Driver a Spence (1994). Vo svojom experimente prezentovali súbežne cez dva monitory dva rôzne záznamy rozprávajúceho človeka, pričom zvuky boli prezentované v jednom meraní z reproduktorov zodpovedajúcich obrazu a v druhom meraní z navzájom vymenených reproduktorov. Úlohou subjektu bolo stále počúvať čo sa hovorí na jednom z monitorov, pričom o experimentálnej manipulácii so zámenou monitorov bol subjekt informovaný. Takže jeho úlohou v prvom meraní bolo sústrediť svoju vizuálnu aj sluchovú pozornosť na to isté miesto, zatiaľ čo v druhom meraní potreboval sústrediť vizuálnu pozornosť na jeden monitor a sluchovú pozornosť na polohu reproduktora druhého monitora. Presnosť identifikovaných slov bola v druhom meraní signifikantne nižšia než v prvom meraní, čo značí, že schopnosť človeka sústrediť svoju vizuálnu pozornosť na jedno miesto a sluchovú pozornosť na iné miesto je obmedzená a že mechanizmy strategického riadenia zrakovej a sluchovej pozornosti nie sú navzájom nezávislé.

Nasledujúce kapitoly podávajú prehľad o súčasnom stave poznatkov o priestorovom riadení endogénnej (strategickej, vedomím riadenej) selektívnej pozornosti, a o jej význame pre integráciu a interakcie medzi jednotlivými zmyslovými vnemami.

2. Rozdelenie pozornostných mechanizmov

Jednou z hlavných téz o pozornosti je, že pozornosť nie je jedna entita, ale súbor procesov, ktoré prebiehajú v neurálnom systéme. Tieto procesy môžu navzájom interagovať s inými procesmi v mozgu a podieľať sa tak na rôznych perцепčných, kognitívnych a motorických úlohách. Pozornosť akéhokoľvek podsystému neurálneho systému nie je možné jednotne definovať. Je možné sa na ňu pozeráť z hľadiska procesov redukcie informácii na vstupe systému, alebo ako na selektívny proces pre prácu s cieľovým objektom v prostredí s distraktormi. Je možné ju chápať ako proces vyhľadávania a v konečnom dôsledku aj selektívny klasifikačný proces, alebo ako proces tvorby konceptov (vizuálnych, senzomotorických). Pozornosť je možné rozdeliť do dvoch základných skupín a to na:

- **automatickú** pozornosť, vyvolanú podnetom z vonkajšieho prostredia, tiež nazývanú exogénnou alebo senzoricou pozornosťou. Príkladom fungovania automatickej pozornosti je ilúzia vykresľovanej čiary (anglicky line motion illusion, [23])
- **strategickú** (vedomím riadenú, dobrovoľnú, endogénnu) pozornosť, ktorá charakterizuje proces, keď sa človek vedome rozhoduje či sa sústredí na vonkajšie podnety, vnútorné myšlienky, alebo na iné kognitívne procesy a aktivity.

Toto rozdelenie, akceptované aj dnešnými teóriami prezentoval už James v roku 1890. Ďalšie a všeobecnejšie rozdelenie pozornostných procesov môže byť v zmysle

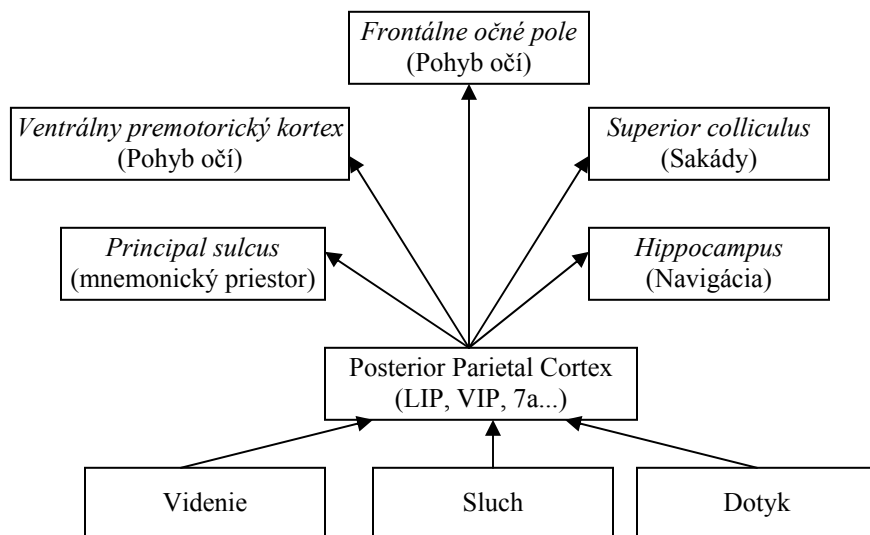
následovných troch komponentov: *selekcie, sledovania a riadenia*. Selekcia je zrejme zásadný a základný proces pozornosti, pričom existujú tri hlavné hypotézy o jej fungovaní. Podľa nich je selekcia založená buď na *umiestnení objektu – location based* (Cave, Pashler, 1995), na *objekte samotnom*, t.j., že je *objektovo orientovaná – object oriented*, (Duncan, 1984), alebo že je *objektovo znakovito založená – object token based* (Kanwisher, Driver, 1992). *Sledovanie* alebo udržiavaná pozornosť je dôležitým aspektom pri cielej interakcii s objektom v prostredí, v ktorom sa nachádzajú aj rušivé „distraktor“ objekty. Sledovanie zaručuje, že ciele sú v čase udržiavané. *Riadenie* pozornosti je proces, pri ktorom je systém schopný zmeniť stálu pozornosť z jedného objektu na iný objekt, ktorý sa stáva novým cieľom interakcie.

Z hľadiska modelovania priestorového vnímania je dôležité poznať kedy (za akých okolností) je pozornosť založená na polohe objektu, a za akých na objekte samotnom. Taktiež je základnou otázkou ako sa zameranie priestorovej pozornosti riadi.

3. **Modality pozornostných procesov**

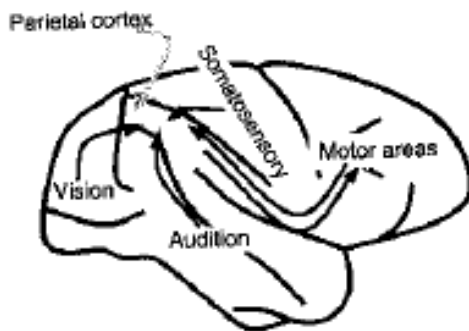
Množstvo každodenne vnímaných objektov a udalostí generuje pozornostné vodítka (anglicky cues), ktoré sa objavujú buď v jednej modalite, alebo nejakým stupňom pokrývajú viacero zmyslových modalít. Kapacita mozgu určená pre koordináciu rôznych sensorických signálov prichádzajúcich z jedného alebo viacerých zdrojov je základná vlastnosť, ktorá nám vo výsledku poskytuje jednotný vnem okolitého sveta. Táto koordinácia je tiež základom pre smerovanie (riadenie) pozornosti a riadenie pohybu v rámci nej. Predošlým výskumom bolo jasne preukázané, že neurálna syntéza multisenzorických vnemov môže zlepšiť pravdepodobnosť detekcie a odpovede na nejakú udalosť, ale tiež aj schopnosť identifikovať a lokalizovať danú udalosť. Na druhej strane, ak je rôznym modalitám prezentovaná konfliktná sensorická informácia, pravdepodobnosť správnej lokalizácie a identifikácie sledovanej udalosti sa môže znížiť. Na ilustráciu takejto situácie môže poslúžiť napr. známy McGurkov efekt, ktorý ukazuje, že aj keď porozumenie počutej reči môže sa môže čítaním z pier značne zlepšiť, počúvanie človeka vyslovujúceho jednu hlásku a dívanie sa na toho istého človeka pri vyslovovaní inej hlásky vedie k vnemu, ktorý nezodpovedá ani videnému ani počutému podnetu.

Aj keď náš primárny záujem sa sústreďuje na fungovanie pozornosti pri vizuálnom a sluchovom priestorovom vnímaní, pre pochopenie týchto interakcií je dôležité aj skúmanie interakcií s dotykovými vnemami. Neurálne systémy zodpovedné za základné spracovanie vnemov jednotlivých zmyslových modalít sú v mozgu priestorovo i funkčne oddelené. Pre behaviorálne skúmanie ich fungovania a interakcie medzi nimi sa používajú psychofyzikálne merania na ľudských subjektoch (Driver and Spence, 2000). Pre identifikáciu fyziologických závislostí procesov multimodálnej pozornosti sú používané elektrofyziologické štúdie na zvieratách. Obrázok 3.1 ukazuje závislostí jednotlivých modalít a neuroanatomické oblasti zodpovedné za ich spracovanie.



Obrázok 3.1. Schématická reprezentácia štandardnej teórie multisenzorickej integrácie (Gross, Graziano 1995)

Centrálным miestom v mozgu u primátov, kde sa integrujú signály jednotlivých modalít je oblasť posteriórneho parietálneho kortexu. Rozloženie týchto oblastí v mozgovej kôre je ilustrované na Obrázku 3.2).



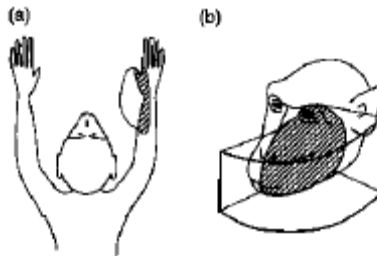
Obrázok 3.2. Schématické zobrazenie prepojení posteriórneho parietálneho kortexu s inými mozgovými oblasťami (Spence a Driver, 2004)

Pre pochopenie funkcií jednotlivých subsystémov je nutné pozerať sa na ich anatómiu a fyziológiu a to akým spôsobom sú realizované pozornosťné mechanizmy na úrovni jednotlivých modalít a osobitne na interakciu a spájanie týchto procesov do jedného celku.

4. Neurálna báza krosmodálnych pozornostných vplyvov

Pre tvorbu modelov a skúmanie neurofyziologických závislostí je potrebné poznať štruktúry a anatómiu oblastí zodpovedných za multimodálne spracovanie vnemov. Oblasťou mozgu zodpovednou za spracovanie vnemov rôznych modalít je posteriórny parietálny kortex (PPC), ktorý predstavuje rozhranie medzi vnemami a následnou akciou (Critchley 1966; Mountcastle et al. 1975; Andersen et al. 1997). Experimentálne boli u neľudských primátov zistené multimodálne oblasti PPC a to, ventrálny interparietálny kortex (VIP), laterálny intraparietálny kortex (LIP), polysenzorická zóna (PZ). Neuróny v týchto oblastiach odpovedajú na stimuly rôznych modalít.

Neuróny oblasti PZ sú citlivé hlavne na dotykové a vizuálne vnemy z blízkeho priestoru hlavy a hornej časti tela. Pre 95% buniek oblasti PZ platí, že ich receptívne pole je iba v časti tela kontralaterálnej k hemisfére, ktorú skúmame. Pre temer všetky bunky, vizuálne receptívne pole je obmedzené hĺbkou a vonkajšia vzdialenosť receptívneho poľa je v tejto oblasti obyčajne menšia ako 20 cm. Všetky pohybové smery sú reprezentované, rôzne bunky preferujú pohyb z ľava, prava, hora, dola a taktiež pohyb objektov oproti a od opice. Smerová preferencia pohybu vizuálnej a dotykovej modality je rovnaká pre jednotlivé neuróny pre 80% neurónov oblasti PZ.



Obrázok 4.1. Príklady bimodálnych vizuálne dotykových receptívnych polí neurónov (Spence a Driver, 2004)

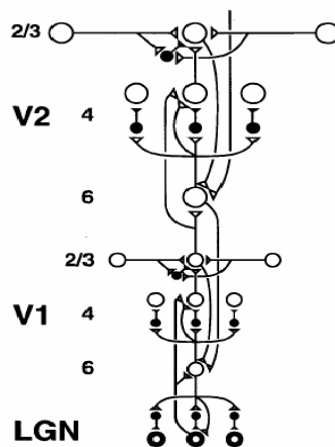
Neuróny oblasti VIP sú bimodálne a reagujú hlavne na dotykové a podnety na tvári a vizuálne podnety blízko tváre (Colby et. al 1993; Duhamel et. al 1998). Väčšina týchto buniek je podobne ako v oblasti PZ orientačne selektívna. Niektoré z týchto buniek odpovedajú na blízke podnety vzdialené iba pár centimetrov od tváre, iné zas aj na vzdialenejšie podnety. Podobne ako v oblasti PZ orientačná citlivosť pre jednotlivé modalít je zhodná pre jednotlivé neuróny. Niektoré neuróny tejto oblasti sú dokonca trimodálne a reagujú aj na sluchové podnety. Odlišnosťou týchto neurónov sú rôzne súradnicové systémy, na ktoré sú ukotvené jednotlivé modalít neurónov. V oblasti PZ sú všetky neuróny ukotvené na súradnicový systém spojený s časťou tela, ktorá je reprezentovaná danými receptívnymi poľami. V oblasti VIP je zraková modalita ukotvená na súradnicový systém so stredom na sietnici, to znamená, že ak sa pohne oko, zmenia sa aj koordináty objektov vzhľadom na tento pohyb. Dotyková modalita je ukotvená na tú oblasť, ktorú pokrýva receptívne pole daného neurónu (časti tváre).

Z hľadiska predmetu výskumu krosmodálnych vplyvov sluchovej a vizuálnej pozornosti je dôležité práve skúmanie oblasti LIP. Neuróny tejto oblasti bimodálne reagujú na vizuálne a sluchové podnety, pričom táto oblasť je zodpovedná za plánovanie

pohybu očí. Oblasť LIP parietálneho kortexu reprezentuje prepojenie medzi vizuálnym systémom a sluchovým systémom. Problémom pri spracovaní bimodálnych vstupov sú v tomto prípade rôzne súradnicové systémy, ktoré reprezentujú priestor. Z hľadiska modelovania týchto vplyvov je dôležité skúmanie aspektu transformácie reprezentácie sluchovej priestorovej informácie, ktorá je ukotvená v súradnicovej sústave vzhľadom na hlavu, a reprezentácie vizuálnej priestorovej informácie, ktorá je ukotvená v súradnicovej sústave vzhľadom na stred oka. Výsledky predošlého výskumu nasvedčujú, že priestorové koordináty sluchového podnetu sú prekódované do súradnicovej sústavy so ukotvenej na stred oka (Cohen; Andersen 2003).

5. Existujúce modely a možnosti ich rozšírenia

Pre modelovanie krosmodálnych vplyvov riadenia pozornosti je treba uvažovať výpočtové modely reprezentujúce priestorovú informáciu o podnete spolu s pozornosťnými mechanizmami. Za posledných dvadsať rokov dodalo veľké množstvo experimentálnych a teoretických prác zvyšujúcu sa podporu prezentovanej Helmholtzovej myšlienke, že vizuálny kortex venuje podstatné procesné zdroje generovaniu 3-D reprezentácií perceptuálnych hraníc a povrchov, obzvlášť reprezentáciám, ktoré dokážu separovať figúry od ich pozadia a doplnia reprezentácie čiastočne zakrytých objektov. Bolo navrhnuté, že tieto reprezentácie hraníc a povrchov sú formované v prúdoch, ktoré premietajú medzi kortikálnymi oblasťami V1 a V4 (Grossberg 1994). Tieto reprezentácie potom projektujú do vyšších oblastí mozgu, najmä inferotemporálneho kortexu (IT) vo ventrálnej (ČO) dráhe, kde sú kategorizované alebo zjednotené do objektových reprezentácií. Navyiac, všetky kortikálne oblasti sú navzájom prepojené spätnými väzbami a väzbami na ďalšie oblasti. V tomto ohľade sa upiera naša pozornosť práve na modely reprezentujúce dopredné a spätné interakcie medzi oblasťami V1 a V4. Konkrétne modelovanie pozornostného riadenia môže byť založené na Teórii adaptívnej rezonancie (Grossberg, 1976; Kopčo and Carpenter, 2000) a špecificky na modeli vizuálnych štruktúr založenom na tejto teórii, nazývanom LaminART (Grossberg 1994).



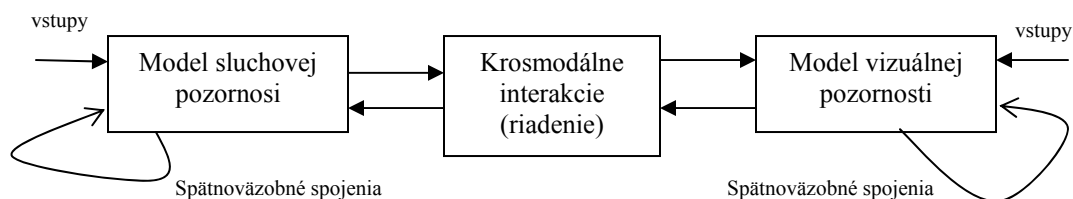
Obrázok 5.1. Schématické zobrazenie štruktúry siete LaminART (Grossberg, 1994)

Tento model navrhuje neurálne mechanizmy umožňujúce vyvíjajúcim sa kortikálnym obvodom stabilizovať sa, a to využitím vlastností ich samo-organizovaných obvodových interakcií spolu s implementáciou komplexnej dynamiky neurónov.

Ďalším jednoduchším existujúcim modelom vizuálnej pozornosti, ktorý je možné využiť pre skúmanie krosmodálnych pozornostných vplyvov vo vizuálnom kortexe je model navrhnutý vrámci KKUI (Andoga, Dobeš 2003) zohľadňujúci interakcie medzi vrstvami V1 – V4 a inferotemporálnej oblasti (IT) a získanie priestorovej informácie o objekte pozornosti. Tento model sa prikláňa k hypotéze objektovo orientovanej pozornosti, pričom je schopný simulovať jednak selekciu v zmysle objektovo založenej teórie a jednak proces sledovania objektov.

Oba modely reprezentujú mechanizmy exogénnej pozornosti ako pár uzavretých spätnoväzobných slučiek. Kľúčovou schémou pre vyššie kognitívne funkcie je však riadenie. Práve na exaktnejšie modelovanie procesov spojených s riadením vo vyšších vrstvách vizuálneho kortexu môže byť jedným zo smerov budúceho výskumu.

Na popis aktivity sluchovej dráhy môže byť použitý model periférneho sluchového systému (Zhang et al., 2001) skombinovaný s modelom binaurálneho sluchového spracovania v mozgovom kmeni od Colbulrna (1977). Podobné dostupné submodely pre popis ďalších častí sluchovej dráhy, ako aj pre popis vizuálnej dráhy (napr. Grossbergov BCS-FCS model) môžu slúžiť ako senzorické vstupy pre model krosmodálneho strategického pozornostného riadenia, ktorého možná štruktúra je uvedená na Obrázku 5.2.



Obrázok 5.2. Schématické zobrazenie štruktúry prepojenia modelov

6. Záver

Z uvedeného prehľadu vyplýva, že napriek jednoznačnému významu strategického riadenia priestorovej pozornosti pre každodenné správanie sa človeka, sú dnes jej mechanizmy pochopené len čiastočne, a to ako na neurofyziologickej tak aj na behaviorálnej úrovni. Taktiež existuje len niekoľko málo výpočtových modelov, ktoré sa snažia tieto mechanizmy teoreticky popísať. Keďže podobné mechanizmy môžu byť užitočné aj pre umelé systémy, ktoré by mali byť schopné orientovať sa v priestore na základe integrácie vizuálnych a sluchových podnetov, je porozumenie riadeniu strategickému pozornosti jednou z dôležitých výziev pre kognitívnu neurovedu dneška.

Literatúra

- [1] Asemi, N., Sugita, Y., Suzuki, Y. (2003) Auditory search asymmetry between pure tone and temporal fluctuating sounds distributed on the frontal-horizontal plane. In:

- Acta acustica united with acustica. 89 (2): 346-354.
- [2] Andoga R (2003) Modelovanie objektovo založenej pozornosti použitím umelých neurónových sietí. Diplomová práca. FEI TU Košice.
 - [3] Blauert, J (1997) Spatial Hearing. MIT Press
 - [4] Broadbent DE (1958) Perception and Communication. Pergamon, Elmsford, New Jersey.
 - [5] Cherry EC (1953) Some experiments upon the recognition of speech with one and two ears. Journal of the Acoustical Society of America, 25, 975-979.
 - [6] Colby, C. L. (1999) Spatial Perception. In: R. A. Wilson and F. Keil (eds.) MIT Encyclopedia of Cog. Sciences. MIT Press
 - [7] Critchley M (1966) The parietal lobes. New York: Hafner.
 - [8] Driver J a Spence Ch. (1994) Spatial synergies between auditorz and visual attention. In Attention and performanc: conscious and nonconscious information processing (ed. C. Umilta and M. Moscovitch), Vol. 15, pp. 311-31. MIT Press, Cambridge, MA.
 - [9] Gilkey R, Anderson T (1997) Binaural and Spatial Hearing in Real and Virtual Environments. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.
 - [10] Grossberg S (1976). Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. Biological Cybernetics, 23:121-134.
 - [11] Grossberg, S. (1980). How does a brain build a cognitive code? Psychological Review, 87, 1-51.
 - [12] Grossberg, S. (1994). 3-D vision and gure-ground separation by visual cortex. Perception and Psychophysics, 55, 48-120.
 - [13] Kopco N and GA Carpenter (2000) "Graded Signal Functions for ARTMAP Neural Networks," In Sincak et al. (Eds.) The Stane of the Art in Computational Intelligence (Collection of papers presented at the European Symposium on Computational Intelligence, Kosice, Slovakia, Aug 30 - Sept 1, 2000). Physica-Verlag. pp. 9-14.
 - [14] Kopčo, N, A Ler, and BG Shinn-Cunningham (2001) Effect of auditory cuing on azimuthal localization accuracy. J Acoust Soc America, 109. p. 2377
 - [15] LaBerge D (1995) Attentional processing. Harvard University Press, Cambridge, Massachusetts.
 - [16] Mountcastle, V. B., Lynch, J. C. G. A., Georgopoulos, A., Sakata, H., & Acuna, C. (1975). Posterior parietal association cortex of the monkey: Command functions for operations within extrapersonal space. Journal of Neurophysiology, 38, 871-908.
 - [17] Soto-Faraco S and Spence C (2002). Modality-specific auditory and visual temporal processing deficits. Quarterly Journal of Experimental Psychology A 55, 23-40.
 - [18] Spence Ch a Driver J, (2004) Crossmodal Space and Crossmodal Attention. Oxford University Press.
 - [19] Treisman AM, Davies A (1973). Divied attention to ear and eye. In Attention and Performance (ed. S. Kornblum), Vol. 4, pp. 101-117. Academic Press, New York.
 - [20] Shimojo, S., Miyauchi, S. and Hikosaka, O. (1992). Visual attention field can be assessed by illusory line motion sensation." IOVS, 33:1354.