

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

Použitie pulzných neurónových sietí pre
spracovanie stereovízneho obrazu

Diplomová práca

2014

Bc. Martin Kolesnáč

Technická univerzita v Košiciach
Fakulta elektrotechniky a informatiky

Použitie pulzných neurónových sietí pre spracovanie stereovízneho obrazu

Diplomová práca

Študijný program: Umelá inteligencia
Študijný odbor: Umelá inteligencia
Školiace pracovisko: Katedra umelej inteligencie (KKUI)
Školiteľ: Ing. Marek Bundzel Phd.
Konzultant:

Košice2014

Bc. Martin Kolesnáč

Erráta

Použitie pulzných neurónových sietí pre spracovanie stereovízneho
obrazu

Bc. Martin Kolesnáč

Košice 2014

Strana	Riadok	Chybne	Správne

Abstrakt v SJ

Táto práca sa zaoberá spracovaním stereovízneho obrazu pomocou pulzných neurónových sietí. V práci sú popísané vlastnosti metódy, ako aj jej klady a zápory a v závere sú navrhnuté postupy, ktoré by zlepšili výkon metódy. Následne je neurónová sieť implementovaná a otestovaná na reálnych dátach.

Kľúčové slová

Stereovízny, obraz, 3D, pulzné, neurónové, siete

Abstrakt v AJ

This diploma thesis deals with stereoscopic image processing using spiking neural networks. The paper describes the basic properties of the method, as well as its pros and cons and in the end some techniques that would improve performance of the method are proposed. Subsequently, the neural network is implemented and tested on real-world data sets.

Kľúčové slová v AJ

Stereovision, 3D, spiking, neural, networks

TECHNICKÁ UNIVERZITA V KOŠICIACH
FAKULTA ELEKTROTECHNIKY A INFORMATIKY
Katedra kybernetiky a umelej inteligencie

ZADANIE DIPLOMOVEJ PRÁCE

Študijný odbor: 9.2.8 Umeľá inteligencia

Študijný program: Umeľá inteligencia

Názov práce:

Použitie pulzných neurónových sietí pre spracovanie stereovízneho obrazu

Using Spiking Neural Networks for stereovision

Študent: **Bc. Martin Kolesnáč**

Školiteľ: **Ing. Marek Bundzel, PhD.**

Školiace pracovisko: **Katedra kybernetiky a umelej inteligencie**

Konzultant práce:

Pracovisko konzultanta:

Pokyny na vypracovanie diplomovej práce:

1. Vypracovať prehľad problematiky biologicky inšpirovaných metód pre spracovanie obrazu so zameraním na pulzné neurónové siete (spiking neural networks).
2. Návrh a implementácia systému využívajúceho pulzné neurónové siete na spracovanie stereovízneho obrazu s cieľom vytvorenia 3D modelu priestoru.
3. Navrhnuť a realizovať experimentálne overenie funkčnosti systému.
4. Vypracovať dokumentáciu podľa pokynov vedúceho DP.

Jazyk, v ktorom sa práca vypracuje: slovenský

Termín pre odovzdanie práce: 02.05.2014

Dátum zadania diplomovej práce: 31.10.2013

prof. Ing. Peter Šinčák, CSc.

vedúci garantujúceho pracoviska



prof. Ing. Liberios Vokorokos, PhD.

dekan fakulty

Čestné vyhlásenie

Vyhlasujem, že som diplomovú prácu vypracoval(a) samostatne s použitím uvedenej odbornej literatúry.

Košice 2.5.2014

.....

Vlastnoručný podpis

Pod'akovanie

Ďakujem vedúcemu diplomovej práce Ing. Marekovi Bundzelovi, Phd. za odborné pripomienky, odbornú pomoc a podnetné návrhy pri vypracovaní diplomovej práce.

Ďakujem svojej rodine za podporu počas celého štúdia a za cenné pripomienky k tejto práci.

Predhovor

Problematika spracovania stereovíznej obrazovej informácie je predmetom skúmania už niekoľko desiatok rokov. Avšak ani jedna z metód nedosahovala efektivitu porovnateľnú s efektivitou ľudského vizuálneho kortexu. Väčšina navrhnutých metód bola aplikovateľná len na úzku oblasť spracovania stereovízneho obrazu pri vopred definovaných podmienkach. Preto začalo vznikať množstvo metód inšpirovaných vizuálnym kortexom, od modelov s vysokou úrovňou abstrakcie ako napríklad PCNN (Eckhorn et al., 1990), až po špeciálne typy konvolučných neurónových sietí ako LIS-SOM (Miikkulainen et al., 2005). Model použitý v tejto práci predstavuje typ pulznej neurónovej siete, ktorá zavádza laterálne prepojenia v rámci jednej vrstvy .

Mojou ambíciou je poukázať v tejto práci na aplikačné možnosti pulzných neurónových sietí v oblasti spracovania obrazu ako aj na ich potenciálne nevýhody, ktoré môžu predstavovať predmet ďalšieho skúmania.

Obsah

Úvod	1
1 Formulácia úlohy	3
2 Spracovanie obrazu v biologických systémoch	5
2.1 Projekcie na gangliové bunky sietnice	5
2.2 Retinotopická reprezentácia vizuálnej informácie	6
2.3 Priestorovo-časové vlastnosti neurónov v primárnej zrakovej kôre	9
2.4 Architektúra primárnej zrakovej kôry	10
2.5 Spájanie vstupu z dvoch očí	13
2.6 Rozdelenie činností v rámci zrakovej dráhy	15
2.7 Rozdelenie činností v sekundárnej zrakovej kôre	17
2.8 Laterálne prepojenia	17
3 Základy výpočtovej neurovedy	19
3.1 Výpočtové modely	19
3.1.1 Kompartmentálny model	20
3.1.2 Coupled oscillators	22
3.1.3 Integrate-and-Fire neuróny	23
3.1.4 Firing-rate neróny	25
4 Použitie pulzných neurónových sietí pri spracovaní obrazu	28
4.1 LISSOM	28
4.2 PCNN	29
5 Výpočtový model	30
5.1 Architektúra	30
5.1.1 Konektivita v LGN vrstve	32
5.1.2 Konektivita v kortikálnych vrstvách	34

5.2	Generovanie aktivácií neurónov	34
5.2.1	Aktivácie v LGN	35
5.2.2	Aktivácie v PVA	36
5.2.3	Aktivácie v EVA	38
6	Experimenty	40
6.1	Použitý model	40
6.2	Výsledky experimentov	41
7	Záver	47
7.1	Splnenie pokynov na vypracovanie	47
7.2	Zhodnotenie, výstupy a ďalší výskum	47
	Zoznam príloh	54

Zoznam obrázkov

2–1	Projekcie gangliových buniek sietnice	7
2–2	Projekcia binokulárnej časti zorného poľa	8
2–3	Reprezentácia obrázku uhlovo selektívnymi neurónmi	11
2–4	Štruktúra primárnej zrakovej kôry	12
2–5	Kortikálne mapy primárnej zrakovej kôry	13
2–6	Magno-, parvo- a koniocelulárne dráhy	16
3–1	Výpočtové modely neurónov a neurónových sietí	27
5–1	Architektúra modelu BSNN	31
5–2	Receptívne polia ON a OFF neurónov	33
5–3	Aktivačná funkcia modelu BSNN	36
6–1	Disparitná mapa obrázkov z reálneho prostredia	42
6–2	Disparitná mapa v prípade absencie laterálnych spojení v LGN	43
6–3	Disparitná mapa v prípade použitia Canny detektoru	44
6–4	Disparitná mapa v prípade absencie LGN neurónov	45

Zoznam symbolov a skratiek

LGN	Lateral geniculate nucleus
V1	Primary visual cortex
PVA	Primary visual area
EVA	Extrastriate visual area
PCNN	Pulse-coupled neural network
LISSOM	Laterally Interconnected Synergetically Self-Organizing Map
BSNN	Binocular spiking neural network
AP	Action potential

Slovník termínov

Aferentné spojenia sú dopredné spojenia neurónov medzi jednotlivými vrstvami.

Laterálne spojenia sú spojenia neurónov v rámci jednej vrstvy.

OFF neuróny sú neuróny nachádzajúce sa v sietnici a v LGN preferujúce tmavé oblasti obklopené svetlými oblasťami.

ON neuróny sú neuróny nachádzajúce sa v sietnici a v LGN preferujúce svetlé oblasti ohraničené tmavými oblasťami.

Receptívne pole je množina neurónov, ktoré sú s daným neurónom spojené.

Kortikálne oblasti sú zrkové centrá, ktoré sú v smere spracovania zrkového signálu vyššie ako primárna zrková kôra.

Subkortikálne oblasti sú zrkové centrá, ktoré sú v smere spracovania zrkového signálu nižšie ako primárna zrková kôra

Disparita je rozdiel medzi pozíciami dvoch korešpondujúcich bodov stereovízneho obrazu, definuje hĺbku daného bodu

Úvod

Trojrozmerná rekonštrukcia predmetov je dlhodobý problém v oblasti počítačovej grafiky a videnia. Napriek tomu, že počas rokov bolo vyvinutých a úspešne aplikovaných množstvo algoritmov, spoľahlivý proces vytvárania presných trojrozmerných (3D) modelov je stále je stále nevyriešený problém.

Digitálny 3D model je numerická reprezentácia reálneho objektu. Existujú dve veľké skupiny typov spracovaných modelov: v objemových modelov sú popísané lokálne vlastnosti v rámci snímaného objektu, zatiaľ čo pri povrchových modeloch je popísaný len povrch a vzhľad objektov. Táto práca sa zaoberá hlavne povrchovými modelmi. Takéto modely môžu byť zobrazené na počítači z akéhokoľvek pohľadu ako 2D obrázkov pomocou trojrozmernej projekcie a renderovania.

Aplikácie založené na trojrozmernom spracovaní objektov sú čoraz viac populárnejšie, najmä vďaka väčšej dostupnosti 3D grafických zariadenia a čoraz nižšou cenou výpočtového výkonu. Popularita 3D digitálnych modelov pramení z možnosti ich digitálne spracovávať: možnosti vizualizácie a modifikácie ich interakcie s inými digitálnymi objektami je ich hlavnou výhodou oproti ich fyzickým kópiám. Príklady typických aplikácií algoritmov 3D modelovania zahŕňajú oblasti architektúry, strojárstva, alebo virtuálnej reality.

Jedna z hlavných limitácií použitia 3D modelov je ich vysoká cena, keďže väčšinou musia byť vytvorené ručne. Obzvlášť pri rekonštrukcii existujúcich objektov je skenovací proces, za účelom získania presnej geometrie objektu, zdĺhavý a technicky náročný. Tradičné prístupy k 3D skenovaniu zahŕňajú fotogrametriu, alebo laserové skenovanie. Takéto skenovacie zariadenia sú často veľmi drahé, vyžadujú opatrné zaobchádzanie a vykonávanie zložitých kalibračných postupov.

Algoritmus navrhnutý v tejto práci rieši všetky vyššie spomenuté problémy. 3D rekonštrukcia videnej scény je realizovaná len na základe dvoch navzájom posunutých

obrazov danej scény. Teda na vytvorenie trojrozmerného modelu postačuje akékoľvek obrazovo snímacie zariadenie, napríklad digitálna videokamera.

Kapitola 2 predstavuje stručný úvod do problematiky spracovania obrazu z pohľadu neurobiológie. Kapitola 3 sa zaoberá základmi výpočtovej neurovedy, ktoré sú nevyhnutné pre návrh biologicky inšpirovanej neurónovej siete. Kapitola 4 poskytuje stručný prehľad biologicky inšpirovaných neurónových sietí, ktoré do určitej miery poskytnú základ pre model BSNN, ktorý je použitý v tejto práci. Kapitola 5 sa zaoberá základnými vlastnosťami a architektúrou použitého modelu. V kapitole 6 sú následne popísané vykonané experimenty.

1 Formulácia úlohy

Keďže názov tejto diplomovej práce je Použitie pulzných neurónových sietí pre spracovanie stereovízneho obrazu, jedna časť práce bude zameraná na popis a implementáciu vhodnej metódy spracovania stereovízneho obrazu a druhá časť otestovaním zvolenej metódy na vizuálnych podnetoch z reálneho prostredia. Je dôležité je realizovať tieto testy čo najlepšie, aby prípadne odhalili nedostatky zvolenej metódy, alebo naopak preukázali jej kvalitu.

Hlavné časti tejto diplomovej práce môžeme zhrnúť do splnenia týchto úloh:

1. **Vypracovať prehľad problematiky biologicky inšpirovaných metód pre spracovanie obrazu so zameraním na pulzné neurónové siete (spiking neural networks)**

Táto úloha zahŕňa stručný úvod do problematiky biologicky inšpirovaných inšpirovaných modelov ako aj ich využitie v procese spracovania obrazu v praxi.

2. **Návrh a implementácia systému využívajúceho pulzné neurónové siete na spracovanie stereovízneho obrazu s cieľom vytvorenia 3D modelu priestoru**

Úloha sa venuje podrobnému teoretickému popisu modelu založeného na pulzných neurónových sieťach, ktorý bude použitý v samotnom procese spracovania stereovízneho obrazu.

3. **Navrhnuť a realizovať experimentálne overenie funkčnosti systému**

V tejto úlohe bude navrhnutá a realizovaná implementácia zvoleného modelu a následne bude model experimentálne overený na reálnych dátach.

4. **Vypracovať dokumentáciu podľa pokynov vedúceho diplomovej práce**

Úloha definuje potrebu vypracovať diplomovú prácu spolu so systémovou a

používateľskou príručkou.

Vyššie uvedené úlohy spoločne s nadpisom tejto kapitoly sú súčasne aj pokynmi na vypracovanie tejto diplomovej práce.

2 Spracovanie obrazu v biologických systémoch

Spracovanie obrazu v zrakovom systéme človeka a zvierat začína detekciou svetla svetlocitlivými fotoreceptormi na sietnici oka. Získaná informácia iniciuje interakcie medzi časťami mozgu; tieto interakcie v konečnom dôsledku vedú k vedomému vnímaniu videnej scény, stimulácií bežných reflexov, ako napríklad zmena šírky zreničky, nasmerovaniu očí na bod záujmu a k regulácií homeostatických mechanizmov, ktoré sú viazané k cirkadiálnemu cyklu. Vzhľadom na rozmanitosť spomenutých funkcií, musí byť vizuálna dráha a štruktúry podieľajúce sprostredkovaní týchto funkcií značne odlišné. Z týchto je najdôležitejšia primárna vizuálna dráha vedúca zo sietnice do dorsálnej časti periférnej vizuálnej dráhy (LGN) a následne do primárnej zrakovej kôry (V1). Rôzne druhy neurónov pozdĺž tejto trasy kódujú rôzne druhy vizuálne informácie - osvetlenie, spektrálne rozdiely, orientácia a pohyb. Paralelné spracovanie rôznych druhov vizuálnej informácie pokračuje pozdĺž kortikálnych dráh až do vizuálnych oblastí okcipitálneho, parietálneho a temporálneho laloku. Vizuálne oblasti v temporálnom laloku zabezpečujú rozpoznávanie objektov, zatiaľ čo oblasti v parietálnom laloku sa zaoberajú detekciou pohybu. Normálne videnie vyplýva z integrácie informácií zo všetkých kortikálnych oblastí. Procesy zabezpečujúce zrkové vnímanie sú stále neobjasnené a sú predmetom výskumu modernej neurovedy.

2.1 Projekcie na gangliové bunky sietnice

K prvému rozsiahlejšiemu spracovaniu obrazu dochádza už v gangliových bunkách sietnice, ktoré vykonávajú určitý typ detekcie okrajov. Existujú dva hlavné typy týchto buniek, takzvané ON-center a OFF-center bunky. ON-center gangliové bunky excitujú ak sa v ich receptívnom poli nachádza podnet skladajúci sa zo svetlej oblasti obklopenej tmavou oblasťou. OFF-center gangliové bunky naopak preferujú tmavé oblasti ohraničené svetlými oblasťami.

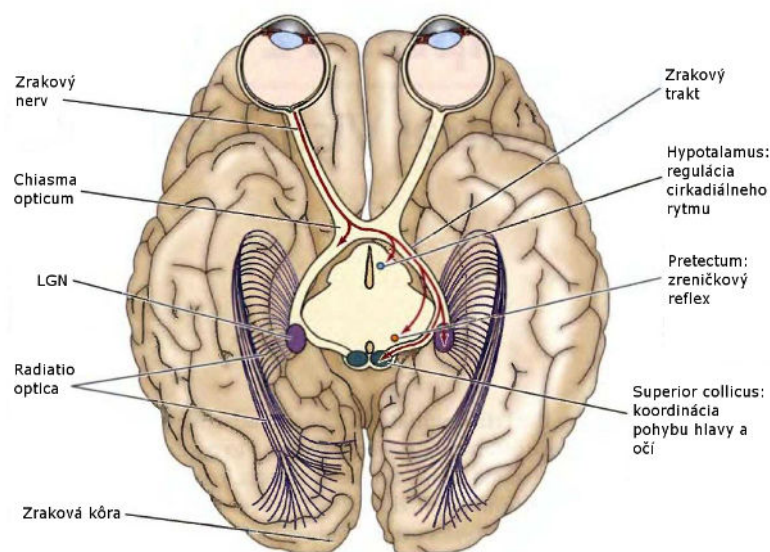
Axóny gangliových buniek vychádzajú zo sietnice cez kruhovú oblasť v jej nasálnej časti nazvanej zrkový disk, kde formujú optický nerv. Axóny v optickom nerve prechádzajú cez zrkový chiasmu (optic chiasm). U ľudí približne 60P týchto vlákien križuje chiasmu, ostatných 40P pokračuje až do oblastí talamusu a stredného mozgu na tej istej strane.

Za zrkovou chiasmou formujú axóny gangliových buniek z každej strany zväzok zvaný zrkový trakt. Teda zrkový trakt obsahuje, na rozdiel od zrkového nervu, vlákna z oboch očí. Čiastočné prekríženie axónov gangliových buniek v zrkovej križovatke umožňuje, aby bola informácia z dvoch korešpondujúcich bodov na dvoch sietniciach spracovaná na približne rovnakej kortikálnej strane v každej hemisfére.

Axóny gangliových buniek v zrkovom trakte smerujú do rôznych štruktúr medzmozgu 2–1. Hlavným cieľom v medzmozgu je párová štruktúra LGN. Gangliové bunky z temporálnej časti sietnice (tá časť, ktorá je ďalej od nosa) premietajú svoju aktivitu do ipsilaterálnej časti LGN a bunky z nasálnej časti sietnice do kontralaterálnej časti LGN. Neuróny v LGN smerujú svoje axóny ďalej cez vnútorné puzdiečko až do primárnej zrkovej kôry (V1). Medzi ďalšie ciele axónov gangliových buniek patria napríklad štruktúry ako pretectum, ktorá je koordinačné centrum pre zreničkový reflex a superior colliculus, ktorá je zodpovedná za koordináciu pohybu hlavy a očí na cieľové body v zrkovom poli (Wallace et al., 1997).

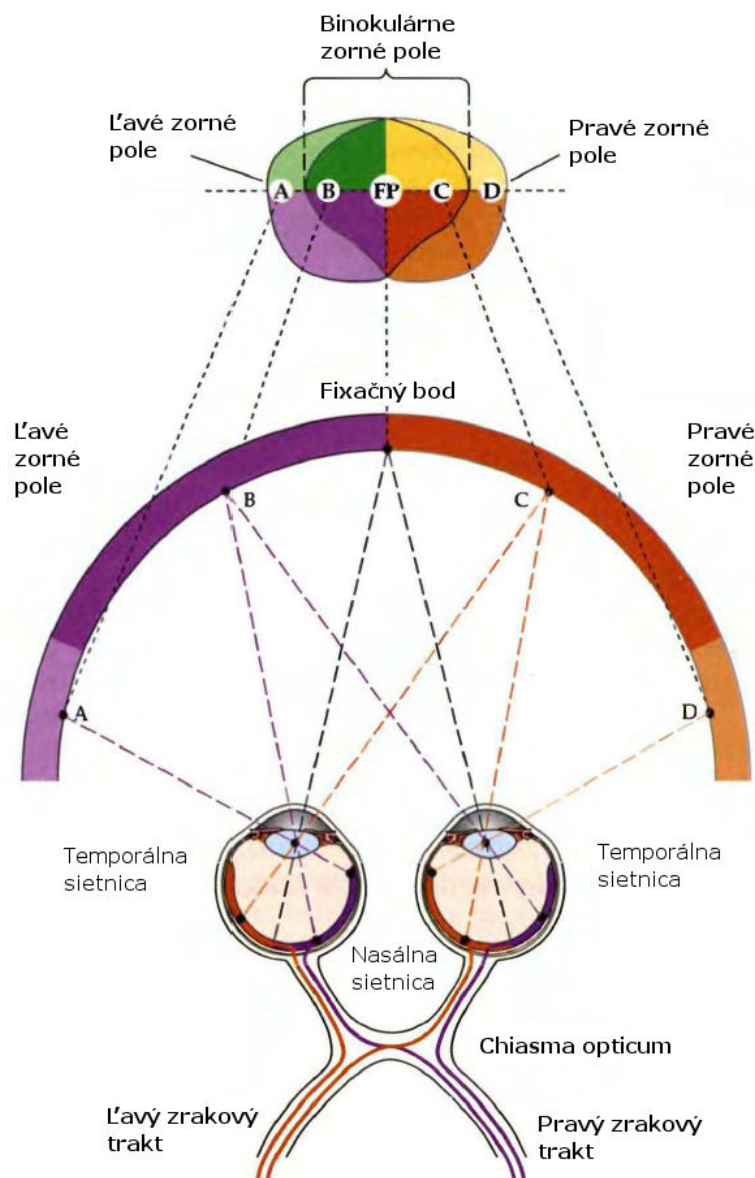
2.2 Retinotopická reprezentácia vizuálnej informácie

Priestorové usporiadanie gangliových buniek na sietnici je zachované aj vo väčšine cieľových štruktúr v medzmozgu. Väčšina týchto štruktúr získava informácie z oboch očí, je teda nutné, aby tieto informácie boli reprezentované ako koherentné mapy jednotlivých bodov prostredia. Vo všeobecnosti platí, že informácia z ľavej polovice videnej scény, či už z pravého alebo ľavého oka, je spracovaná v pravej polovici mozgu, a naopak.



Obr. 2 – 1: Projekcie gangliových buniek sietnice. Axóny gangliových buniek sú ukončené v štruktúrach LGN, superior colliculus, pretectum a v podlôžku. Pre väčšiu prehľadnosť sú zobrazené len križujúce axóny pravého oka. Obrázok prevzatý z (Purves, 2012).

Pre pochopenie priestorovej reprezentácie vizuálnej informácie z dvoch očí vo vyšších zrakových štruktúrach je nutné pochopiť ako sú zrkové podnety premietané na sieťnicu. Každé oko vníma časť videnej scény, ktorá definuje jeho zorné pole 2–2. Kvôli jednoduchšiemu popisu je každá sieťnica a jej korešpondujúce zorné pole rozdelené kvadranty. V tomto prípade je sieťnica rozdelená vertikálnou a horizontálnou čiarou, ktoré sa pretínajú v strede sieťnice 2–2. Vertikálna čiara rozdeľuje sieťnicu na nasálnu a temporálnu časť a horizontálna čiara na vyššiu (superior) a nižšiu časť (inferior). Korešpondujúce vertikálne a horizontálne čiary vo vizuálnej scéne sa pretínajú vo fixačnom bode a vytvárajú kvadranty vizuálnej scény. Kríženie svetelných lúčov vychádzajúcich z rôznych bodov videnej scény v zreničke spôsobuje, že obrázok objektu je na sieťnici invertovaný a obrátený z ľava do prava. Preto, objekty v temporálnej časti zorného poľa sú spracované nasálnou časťou sieťnice a objekty vo vyššej časti sú spracované nižšou časťou sieťnice.



Obr. 2–2: Projekcia binokulárnej časti zorného poľa. Body v binokulárnej časti ľavého zorného poľa (B) sú snímané nasálnou časťou sietnice ľavého oka a temporálnou časťou sietnice pravého oka. Body v binokulárnej časti pravého zorného poľa (C) sú snímané nasálnou časťou sietnice pravého oka a temporálnou časťou sietnice ľavého oka. Body, ktoré sa nachádzajú v monokulárnej časti pravého a ľavého zorného poľa (A a D) sú snímané nasálnymi časťami pravej a ľavej sietnice. Axóny gangliových buniek v nasálnej časti, na rozdiel od gangliových buniek v temporálnej časti, križujú zrakovú chiasmu. Obrázok prevzatý z (Purves, 2012).

Ak sú otvorené obidve oči, tak sú nasmerované na jeden cieľový bod v priestore, čo spôsobuje, že zorné polia obidvoch očí sa navzájom prekrývajú 2–2. Toto binokulárne pole sa skladá z dvoch navzájom symetrických častí. Ľavá časť binokulárneho poľa pokrýva nasálnu časť zorného poľa pravého oka a temporálnu časť zorného poľa ľavého oka. Pravá časť binokulárneho poľa pokrýva temporálnu časť zorného poľa pravého oka a nasálnu časť zorného poľa ľavého oka. Keďže temporálna časť zorného poľa je rozsiahlejšia ako nasálna časť, periférne časti videnej scény sú spracované monokulárne. Zvyšok môže byť videný obidvoma očami, teda jednotlivé body v priestore patria do nasálnej časti zorného poľa jedného oka a do temporálnej časti zorného poľa druhého oka.

Gangliové bunky, ktoré patria do nasálnej časti sietnice vytvárajú axóny, ktoré sú krížené v zrakovej chiasme. Bunky patriace do temporálnej časti sietnice vytvárajú axóny, ktoré smerujú do štruktúr na tej istej strane 2–2.

2.3 Priestorovo-časové vlastnosti neurónov v primárnej zrakovej kôre

Väčšina súčasných poznatkov o funkcionálnej organizácii zrakovej kôry pramení z pionierskych štúdií Davida Hubla a Torstena Wiesel (Hubel and Wiesel, 1959); (Hubel and Wiesel, 1962), ktorý mikroelektrodami u zvierat snímali reakcie neurónov v LGN a v zrakovej kôre na rôzne druhy stimulácie sietnice. Aktivácie neurónov v LGN boli v týchto testov veľmi podobné aktiváciám neurónov v sietnici, so selekciou pre neorientované svetelné vzory. Avšak, podnety, ktoré efektívne stimulovali neuróny v sietnici a v LGN boli z veľkej miery neúčinné pri neurónoch v zrakovej kôre (kortex). Namiesto toho bolo zistené, že väčšina kortikálnych neurónov reagovala len na obdĺžnikové vzory s určitou orientáciou. Receptívne polia kortikálnych neurónov sú teda nastavené na preferenciu hrán pod určitým uhlom (uhlová preferencia). Skúmaním aktivácií neurónov bolo zistené, že každá orientácia je zastúpená pri-

bližne rovnakým počtom neurónov v primárnej zrakovej kôre. V konečnom dôsledku teda každá orientácia vo vizuálnej scéne je reprezentovaná aktiváciou inej skupiny neurónov.

Pre lepšiu predstavu ako môžu byť vlastnosti obrazu vyjadrené skupinami neurónov s uhlovou preferenciou, môže byť obraz rozložený na frekvenčné komponenty a filtráciou rozdelený na množinu obrázkov, ktorých spektrálne rozloženie zodpovedá informácii, ktorá by bola vyjadrená neurónmi s určitou uhlovou preferenciou 2–3. Každý typ uhlovo-selektívneho neurónu sprostredkováva len zlomok informácie z videného obrazu. Dokopy však tieto neuróny poskytujú všetky informácie potrebné k rekonštrukcii snímaného obrazu.

Uhlová preferencia je len jedna z filtračných vlastností neurónov v primárnej zrakovej kôre. Podstatná časť neurónov je navyše selektívnych voči pohybu podnetu, napríklad reagujú omnoho silnejšie na pohyb z prava do ľava. Neuróny môžu byť charakterizované aj ich preferenciou určitej priestorovej frekvencie (veľkosť zmeny kontrastu v ich receptívnom poli).

2.4 Architektúra primárnej zrakovej kôry

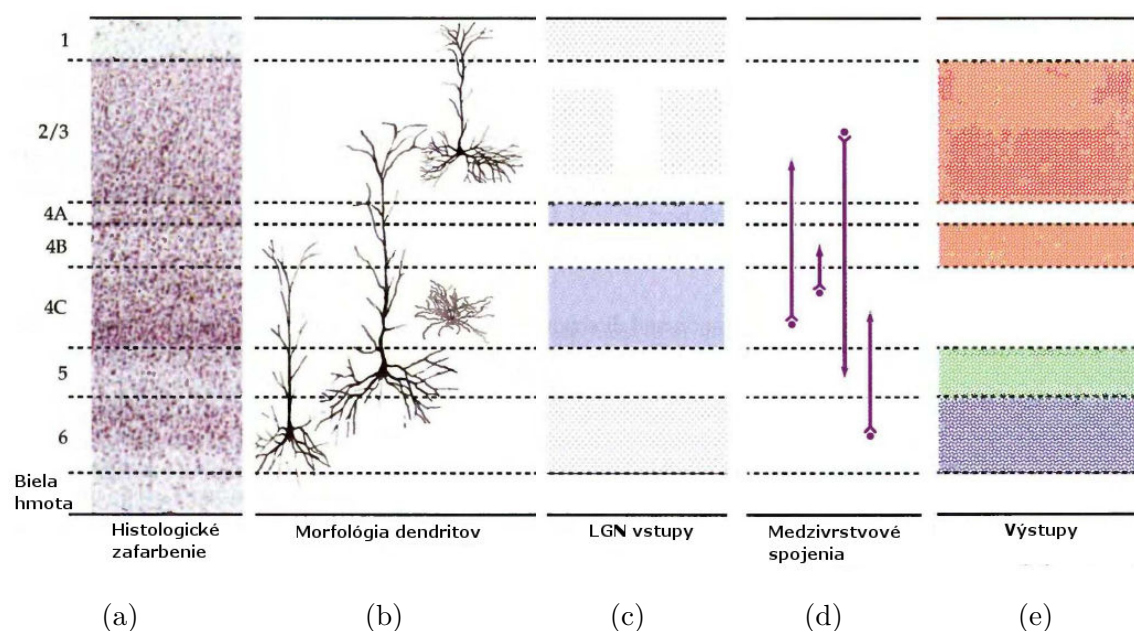
Zraková kôra je podľa zvyklosti delená na šesť bunkových vrstiev (vrstvy 1-6). Vzhľadom na komplexnosť zloženia V1, môžu byť vrstvy ďalej rozdelené na ďalšie vrstvy (napríklad vrstva 4C).

Axóny z LGN sú ukončené najmä vo vrstve 4C, tvorenej neurónmi, ktorých axóny sprostredkovávajú stimulácie z LGN do ďalších vrstiev. Pyramídové neuróny nachádzajúce sa vo vrchných vrstvách V1 sú zdrojom stimulácie pre oblasti sekundárnej zrakovej kôry (Van Essen et al., 1992). Neuróny v hlbších vrstvách smerujú svoje axóny do subkortikálnych oblastí vrátane LGN (Wandell, 1995). Vrstvová organizácia V1 teda slúži najmä na oddelenie skupín neurónov s rozdielnou konektivitou.



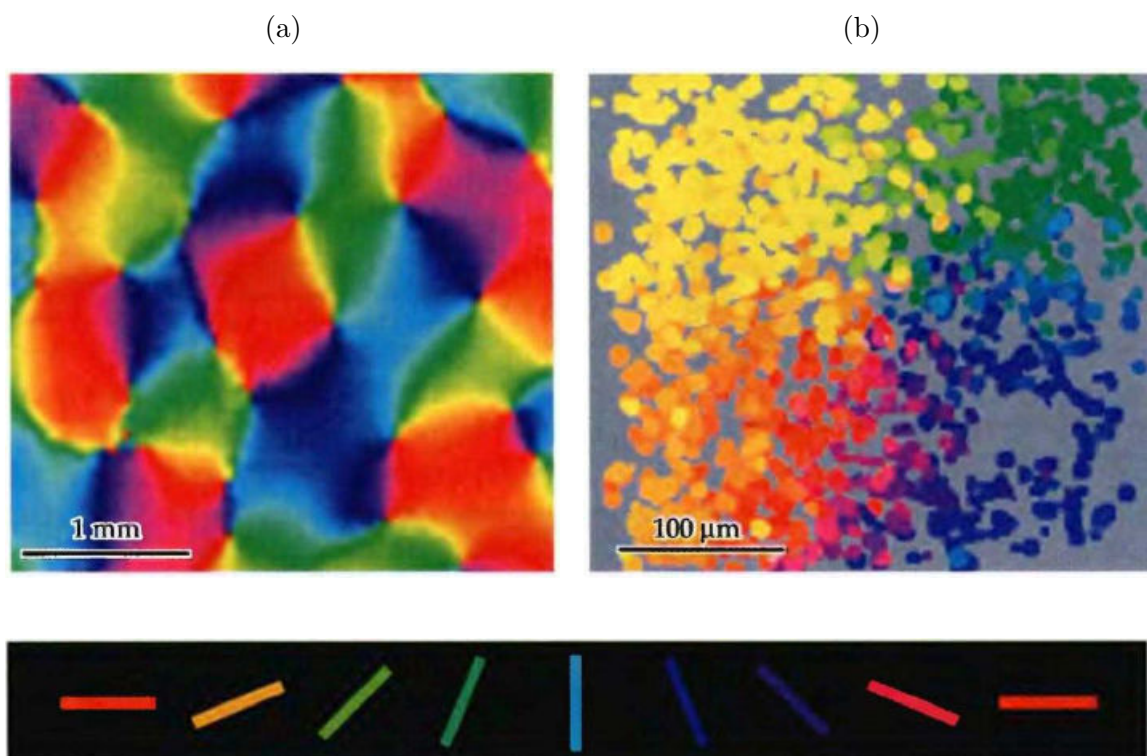
Obr. 2–3: Reprezentácia obrázku uhlovo selektívnymi neurónmi. Na ilustráciu vlastností obrázka, ktoré by boli reprezentované aktiváciami skupín neurónov s určitou uhlovou preferenciou, bola použitá 2D Fourierova transformácia. Jednotlivé obrázky okolo vstupného obrazu zobrazujú komponenty obrazu, ktoré by boli detekované neurónmi s horizontálnou, vertikálnou a diagonálnou uhlovou preferenciou. Obrázok prevzatý z (Van Hooser et al., 2013).

Neuróny v rámci vrstiev V1 formujú útvary známe ako stĺpce. V stĺpcoch sa nachádzajú neuróny, ktoré reagujú na tie isté podnety korešpondujúcej časti sietnice (Gilbert and Wiesel, 1989). Susedné stĺpce majú podobný, ale nie ten istý preferovaný typ podnetov (napríklad orientovaná čiara pohybujúca sa doprava). Dostupnosť zobrazovacích metód umožnila vizualizovať dvojrozmernú mapu uhlovej preferencie na povrchu zrakovej kôry 2–5. Väčšina mapy uhlovej preferencie vykazuje hladký priebeh zmeny preferencie. Tento priebeh je periodicky narušený bodmi nespojitosti, kde sa blízko seba nachádzajú neuróny s rôznymi preferenciami. Celý rozsah uhlových preferencií (0-180 stupňov) sa opakuje približne každý milimeter naprieč celou V1.



Obr. 2–4: Štruktúra primárnej zrakovej kôry. Primárna zrková kôra sa delí na šesť vrstiev, ktoré sa navzájom líšia hustotou, morfológiou a konektivitou neurónov. a Vizualizované telá neurónov v primárnej zrakovej kôre histologickým zafarbením. b Pyramídové bunky sú najpočetnejší typ neurónov v mozgovej kôre. Nachádzajú sa v každej vrstve s výnimkou 4C. Vrstva 4C je tvorená najmä hviezdovitými neurónmi. c Vrstvové rozdelenie vstupov z LGN. Axóny z LGN smerujú prevažne do vrstvy 4C a 4A, menšie časti aj do vrstiev 1, 2/3 a 6. d Vrstvová organizácia intrakortikálnych spojení. Neuróny vo vrstve 4C smerujú svoje axóny do vyšších vrstiev (4B a 2/3). Axóny neurónov z vrstvy 2/3 sú ukončené najmä vo vrstve 5. Axóny neurónov z vrstvy 6 smerujú do vrstvy 4C. e Vrstvové rozdelenie neurónov premietajúcich svoje aktivity do iných štruktúr. Spojenia so sekundárnou zrkovou kôrou sú tvorené neurónmi vo vrstvách 2/3 a 4B (červená). Spojenia s LGN sú tvorené neurónmi z vrstvy 6 (modrá). Spojenia so superior colliculus vznikajú vo vrstve 5. Obrázok prevzatý z (Solomon and Lennie, 2007).

Takéto opakovanie zabezpečuje, že v každej oblasti zorného poľa je k dispozícii plný rozsah uhlových preferencií.



Obr. 2 – 5: Kortikálne mapy primárnej zrakovej kôry. a Vizualizovaná mapa uhlovej preferencie neurónov zrakovej kôry. Farba indikuje prefereovanú orientáciu stĺpca neurónov; červená indikuje prítomnosť stĺpcov preferujúcich horizontálnu orientáciu, modrá prítomnosť stĺpcov preferujúcich vertikálnu orientáciu. Hladký priebeh zmeny preferencie narušajú body nespojitosti. b Pohľad na uhlové preferencie jednotlivých neurónov v bode nespojitosti. Z obrázku vyplýva, že susedné neuróny majú podobnú uhlovú preferenciu, s výnimkou stredu, kde susedné neuróny vykazujú takmer 90 stupňovú zmenu preferencie. Obrázok prevzatý z (Blasdel, 1992).

2.5 Spájanie vstupu z dvoch očí

Na rozdiel od neurónov z počiatočných štruktúr zrakovej dráhy, väčšina neurónov vo V1 je binokulárnych, reagujúcich na stimulácie z pravého aj ľavého oka. Vstup

z oboch očí je už síce prítomný aj na úrovni LGN, avšak kontralaterálne a ipsilaterálne axóny zo sietnice sú ukončené v rozdielnych vrstvách, takže neuróny v LGN sú monokulárne, reagujúce na vstup buď z ľavého, alebo z pravého oka (Erwin et al., 1995). Stimulácie z pravého a ľavého oka potom pokračujú oddelene až do vrstvy 4 V1, kde sú ukončené v neurónových stĺpcoch so striedajúcou sa okulárnou dominanciou. S postupom do iných vrstiev však dochádza ku konvergencii signálu z oboch očí. Väčšina neurónov mimo vrstvy 4 je binokulárna, avšak s určitou preferenciou jedného oka oproti druhému.

Spájanie vstupov z dvoch očí na úrovni V1 vytvára základ pre stereoskopické videnie, teda vnímanie hĺbky videných objektov. Keďže oči snímajú prostredie z iných uhlov, objekty, ktoré ležia pred, alebo za fixačným bodom sú premietané na nekorešpondujúce časti oboch sietnic. Rozdiel polohy týchto dvoch častí sa nazýva disparita.

Vlastnosti receptívnych polí niektorých binokulárnych neurónov vo V1 a v sekundárnej zrakovej kôre z nich robia dobrých kandidátov na extrakciu informácií o binokulárnej disparite. V takýchto neurónoch sú receptívne polia pre pravé a ľavé oko od seba trochu posunuté, teda nesnímajú presne rovnakú časť zorného poľa. Jeden typ týchto neurónov (far cells) reaguje na disparitu, ktorá vzniká ak sa snímaný objekt nachádza za fixačným bodom, zatiaľ čo iný typ (near cells) reaguje na disparitu, ktorá vzniká ak sa snímaný objekt nachádza pred fixačným bodom. Tretí typ neurónov reaguje výhradne na objekty ležiace v rovine fixačného bodu.

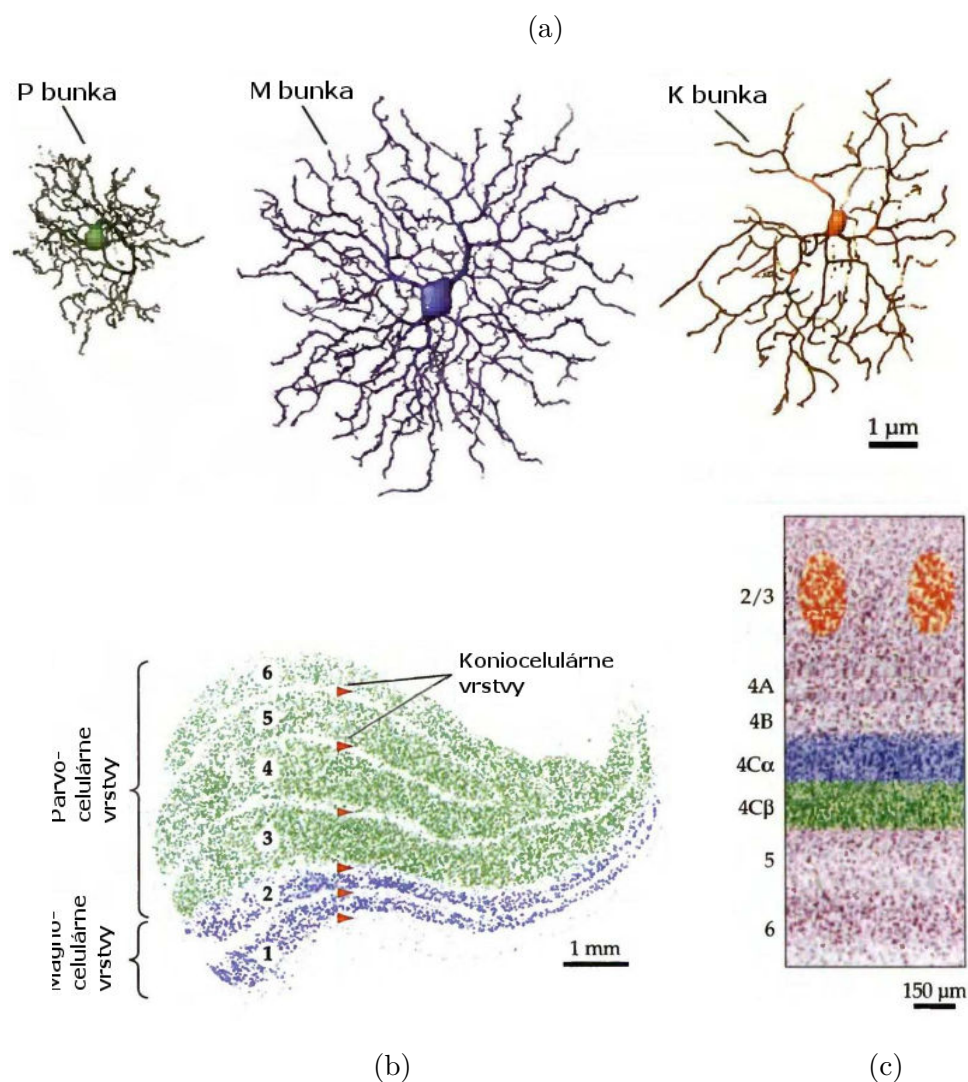
Prítomnosť binokulárnych aktivácií kortikálnych neurónov je podmienená rovnomernou aktivitou oboch očí počas ranného detstva (Callaway and Katz, 1991). Akýkoľvek faktor, ktorý vytvára nerovnosť v aktivite oboch očí, napríklad strabizmus, má za následok trvalé zníženie schopnosti jedného oka generovať aktivácie kortikálnych neurónov, čo má za následok zhoršenie schopnosti používať binokulárnu informáciu k odhadu hĺbky. Skorá detekcia a náprava zrakových problémov je teda nevyhnutná k správne fungovaniu zrakovej kôry.

2.6 Rozdelenie činností v rámci zrakovej dráhy

Okrem selektivity na podnety z jedného, alebo druhého oka sa vrstvy v LGN delia aj podľa veľkosti neurónov. Dve spodné vrstvy zložené z veľkých neurónov sa nazývajú magnocelulárne vrstvy, ostatné vrstvy zložené z malých neurónov sa nazývajú parvocelulárne vrstvy. Magnocelulárne a parvocelulárne vrstvy prijímajú vstupy z rôznych skupín gangliových buniek, ktoré majú takisto rozdielnu veľkosť. M gangliové bunky, ktorých axóny smerujú do magnocelulárnych vrstiev, majú väčšie telo, rozsiahlejší dendritický strom a axóny s väčším priemerom ako P gangliové bunky, ktorých axóny smerujú do parvocelulárnych vrstiev 2–6. Navyše axóny neurónov v jednotlivých vrstvách LGN sú ukončené v rôznych podvrstvách vrstvy 4C primárnej zrakovej kôry. Axóny z magnocelulárnych vrstiev sú ukončené vo vrchnej časti vrstvy 4C ($4C\alpha$) a axóny z parvocelulárnych vrstiev sú ukončené v spodnej časti vrstvy 4C ($4C\beta$) c.

Aktivačné vlastnosti M a P gangliových buniek objasňujú prínos magnocelulárnej a parvocelulárnej zrakovej dráhy k celkovému videniu. M gangliové bunky majú širšie receptívne pole a rýchlejšie axóny ako P bunky. M bunky reagujú na podnet prechodovo (aktivácie sa postupne strácajú), zatiaľ čo P bunky reagujú neprerušovane. P bunky, na rozdiel od M buniek, sú schopné sprostredkovať informáciu o farbe, pretože stred a okolie ich receptívneho poľa je stimulované rôznymi skupinami čapíkov (teda reagujú na rôznu kombináciu farieb). M gangliové bunky taktiež dostávajú vstupy z čapíkov, avšak ich receptívne pole je celé stimulované len jedným druhom čapíkov. To spôsobuje, že M bunky sú necitlivé na zmenu vlnovej dĺžky podnetu v ich receptívnom poli a teda, že nie sú schopné sprostredkovať informácie o farbe.

Prínos magnocelulárnej a parvocelulárnej dráhy k videniu bol experimentálne overený umelým prerušením jednotlivých dráh. Prerušenie magnocelulárnej dráhy malo síce malý efekt na ostrosť vnemu a vnímanie farieb, ale zato výrazne zredukovalo schopnosť zachytiť pohybujúci sa objekt. Naopak prerušenie parvocelulárnej dráhy



Obr. 2–6: Magno-, parvo- a koniocelulárne dráhy. a Morfológia M, P a K gangliových buniek sietnice. M bunky majú veľké telo a veľký dendritický strom. Sú napojené na magnocelulárne vrstvy LGN. P bunky majú menšie telo a dendritický strom. Sú napojené na parvocelulárne vrstvy LGN. K bunky majú malé telo a stredne veľký dendritický strom. Sú napojené na koniocelulárne vrstvy LGN. b Štruktúra ľudského LGN zobrazujúca magnocelulárne, parvocelulárne a koniocelulárne vrstvy. c Ukončenie axónov LGN v primárnej zrakovej kôre. Axóny z magnocelulárnych vrstiev sú ukončené vo vrstve 4Cα, z parvocelulárnych vrstiev vo vrstve 4Cβ a z koniocelulárnych vrstiev vo vrstvách 2 a 3. Obrázok prevzatý z (Purves, 2012).

nemalo žiadny efekt na vnímanie pohybu, avšak výrazne zredukovalo ostrosť vnemu a schopnosť vnímania farieb.

Okrem magnocelulárnej a parvocelulárnej dráhy existuje ešte aj koniocelulárna dráha. Neuróny patriace do koniocelulárnej dráhy sa nachádzajú medzi jednotlivými vrstvami LGN. Tieto neuróny sú stimulované prevažne čapíkmi, reagujúcimi na krátke vlnové dĺžky. Axóny týchto neurónov smerujú do vrstiev 2 a 3 V1. Presný prínos koniocelulárnej dráhy ešte nie je úplne objasnený, predpokladá sa však, že sa podieľa na vnímaní farieb. Je však preukázané, že koniocelulárna dráha je evolučne najstaršia dráha v zrakovom systéme.

2.7 Rozdelenie činností v sekundárnej zrakovej kôre

Anatomické štúdie mozgu opíc viedli k objavu viacerých oblastí v temennom, záhlavnom a spánkovom laloku, ktoré sa zapájajú do spracovania vizuálnej informácie. Každá z týchto oblastí je topograficky mapovaná a je do veľkej miery závislá na aktiváciách V1. Aktivačné vlastnosti neurónov v týchto oblastiach naznačujú, že každá oblasť je zodpovedná za spracovanie iných informácií z videnej scény. Napríklad stredná časť spánkového laloku obsahuje neuróny, ktoré sú selektívne voči smeru pohybu hrany nezávisle na jej farbe. Naopak, oblasť V4 obsahuje neuróny, ktoré sú selektívne voči farbe podnetu bez ohľadu na smer pohybu.

2.8 Laterálne prepojenia

Dlhú dobu sa laterálnym spojeniam v zrakovej kôre neprikladal význam a boli považované len za evolučný prebytok. Nedávne štúdie však naznačujú, že tieto spojenia majú výrazný vplyv na činnosť a vývoj zrakovej kôry (Ruthazer and Stryker, 1996); (Callaway and Katz, 1991). V týchto prácach bolo dokázané, že laterálne spojenia, tak ako aj aferentné spojenia, sa formujú na základe korelácií vo vstupných

podnetoch. V skutočnosti môže vývoj týchto dvoch typov spojení medzi sebou úzko súvisieť. Laterálne spojenia sa formujú približne v rovnakom čase keď aferentné spojenia vytvárajú kortikálne mapy (Dalva and Katz, 1994). Napriek tomu, že jednotlivé laterálne spojenia sú slabé, ich celkový efekt na aktivitu neurónov je zásadný a teda pravdepodobne do značnej miery ovplyvňujú vývoj zrkových centier.

Bolo navrhnutých niekoľko hypotéz popisujúcich funkciu laterálnych spojení. Pre stručnosť uvediem len niektoré z nich.

1. Laterálne prepojenia zohrávajú zásadnú úlohu pri formovaní kortikálnych máp (Bednar and Miikkulainen, 2003).
2. Môžu normalizovať aktivitu v jednotlivých zrkových centrách (Stemmler et al., 1995).
3. Môžu zabezpečovať invarianciu voči rotácii a zmene veľkosti (Edelman, 1996); (Wiskott and von der Malsburg, 1996).
4. Spôsobujú koncentráciu aktivity do niekoľkých bodov čím potláčajú redundantné aktivácie neurónov (Somers et al., 1996).

3 Základy výpočtovej neurovedy

Ako je vidieť z predchádzajúcej kapitoly, zrková kôra je vysoko komplexný dynamický systém a je obzvlášť náročné integrovať roztrúsené experimentálne pozorovania do konkrétneho, koherentného modelu popisujúceho stavbu a funkčnosť zrkovej kôry. Výpočtové modely predstavujú dôležitý nástroj v procese takejto integrácie, keďže sú to vlastne konkrétne implementácie danej teórie. Takýto model môže byť testovaný podobne ako zvieratá, alebo ľudia, buď na overenie danej teórie, alebo na vytvorenie základov pre ďalší výskum.

Táto práca popisuje výpočtový model inšpirovaný zrkovou kôrou, založený na zisťovaní v rámci výpočtovej neurovedy za posledných 30 rokov. Použité základné výpočtové modely sú popísané v tejto kapitole.

3.1 Výpočtové modely

Základnou charakteristikou výpočtových modelov je ich miera abstrakcie. Aj keď teoreticky je možné modelovať každý neurón na najnižšej možnej úrovni pri súčasnej technológii (teda na molekulárnej úrovni), v praxi je potrebné monitorovať len niekoľko vlastností neurónu a tak celý model by bol značne neefektívny. Prílišná detailnosť modelu taktiež sťažuje schopnosť pochopiť funkčnosť daného modelu a vytvárania predikcií na základe jeho správania. Našťastie, pre pochopenie celkového správania sa modelu je možné použiť modely s vyššou mierou abstrakcie a nie je potrebné používať takto detailne popísané simulácie.

V tejto časti sú popísané výpočtové modely neurónov a neurónových skupín s rôznou mierou abstrakcie 3–1, pričom pri každom modeli sú uvedené vlastnosti, ktoré s určitou mierou objasňujú funkčnosť zrkovej kôry.

3.1.1 Kompartmentálny model

Keďže sa predpokladá, že neuróny komunikujú cez akčné potenciály (AP), väčšina detailných modelov neurónov sa zameriava na to, ako sú AP generované a prenášané. Elektrické prúdy, ktoré vedú k vytvoreniu AP sú kontrolované rôznymi druhmi iontových kanálov v lipidovej membráne, ktorá obaľuje telo neurónu. Napätie na membráne (membránový potenciál) sa mení prechodom iónov cez tieto iontové kanály dnu, alebo von z neurónu a práve toto napätie určuje či neurón vygeneruje AP, alebo nie. Takéto zmeny stavu v rámci času zabezpečujú dynamické vlastnosti neurónu, ktoré môžu byť použité na kódovanie informácie.

Pre pochopenie správania sa neurónu je postačujúce modelovanie malých častí neurónovej membrány. Výpočtové modely membránových častí sú zvyčajne založené na Hodgkin-Huxleyho modeli excitačnej membrány (Hodgkin and Huxley, 1952). Skladá sa z dvojice diferenciálnych rovníc popisujúcich membránový potenciál V a počet c_i otvorených iontových kanálov z každého druhu iontových kanálov i (Gerstner and Kistler, 2002); (Rinzel and Ermentrout, 1989):

$$\begin{aligned} C \frac{dV}{dt} &= - \sum_i I_i(c_i, V) + I(t), \\ \frac{c_i}{dt} &= - \frac{c_i - c_{i,\infty}(V)}{\tau_i(V)}, \end{aligned} \tag{3.1}$$

kde C je membránová kapacita, $I_i(c_i, V)$ je prúd prechádzajúci cez iontový kanál i a $I(t)$ aplikovaný vstupný prúd. Pri konštantnom membránovom potenciáli V dosiahne c_i kludovú hodnotu $c_{i,\infty}$ za čas $\tau_i(V)$.

Hodgkin-Huxleyho rovnica popisuje len izolovanú časť lipidovej membrány (jednu zložku). Na modelovanie celého neurónu je ho nutné najprv rozdeliť na hlavné morfológické časti korešpondujúce axónom, dendritom a telu neurónu (Obrázok b). Každá z týchto častí je uvažovaná ako elektrický vodič, zvyčajne cylindrického tvaru (Rall, 1962);(Rall, 1977);(Rall and Agmon-Snir, 1998). Rovnice pre popis elektrických

vlastností vodiča môžu byť v jednoduchších prípadoch riešené analyticky, avšak v prípade detailného neurónového modelu musia byť riešené numericky. V takomto prípade sú morfológické časti neurónu rozložené na menšie zložky, z ktorých každá je popísaná membránovou rovnicou 3.1. Model celého neurónu potom pozostáva z množiny zložiek s vlastným membránovým potenciálom.

Zložkové modely umožňujú modelovať neurón na akejkoľvek úrovni abstrakcie. Môžu byť použité v prípadoch, kedy sú v rámci experimentálnych dát k dispozícii aj parametre modelu, ako šírka a tvar jednotlivých zložiek a rozloženie iontových kanálov na lipidovej membráne (Doya et al., n.d.). V takýchto prípadoch dokážu mapovať experimentálne dáta s podobnou presnosťou ako skutočné neuróny. Napríklad, pomocou zložkového modelu bol vytvorený detailný model pyramídových neurónov (Mainen and Sejnowski, 1998) a Purkyňových buniek nachádzajúcich sa v malom mozgu (Schutter and Bower, 1994).

Avšak veľké kortikálne štruktúry, ako napríklad mapa uhlovej preferencie, sú zložené z miliónov neurónov, z ktorých každý vytvára tisíce spojení (Wandell, 1995). Údaje o nastavení parametrov v takejto veľkej štruktúre sú dostupné len pre niekoľko neurónov a zvyšné miliardy parametrov pre zložkové modely by museli byť určené ručne. Veľké množstvo zložiek sťažuje pochopenie fungovania modelu, napríklad určenia, ktoré zložky sú zodpovedné za konkrétne výpočty. Navyše v súčasnosti je možné simulovať len niekoľko takto detailne popísaných neurónov vzhľadom na pamäťové a výkonnostné obmedzenia súčasných počítačov.

Model neurónovej siete použitý v tejto práci pozostáva z veľkého počtu neurónov, pri ktorých nie je nutné mať definované podrobné nastavenie neurónovej membrány. Preto je v tejto práci použitý neurónový model, ktorý poskytuje vyššiu mieru abstrakcie. Existujú tri hlavné druhy modelov neurónov s vyššou úrovňou abstrakcie: coupled oscillators, integrate-and-fire neuróny a firing-rate neuróny.

3.1.2 Coupled oscillators

Modely coupled oscillator sa zameriavajú modelovanie dynamických vlastností dvojice neurónov, prípadne skupín neurónov. Dynamika každého oscilátora je definovaná dvoma premennými x a y , ktoré reprezentujú stavy dvoch neurónov, z ktorých jeden je inhibičný a druhý excitačný (Obrázok c); (Horn and Opher, 1998); (Sabatini et al., 2004); (Terman and Wang, 1995); (von der Malsburg and Buhmann, 1992). Jednotlivé neuróny sú navzájom prepojené rekurzívnou slučkou, v ktorej excitačná jednotka aktivuje inhibičnú jednotku, ktorá následne inhibuje excitačnú jednotku. Aktivácie jednotiek môžu byť popísané sadou diferenciálnych rovníc (Wang, 1995):

$$\begin{aligned}\frac{dx}{dt} &= f(x) - y + z, \\ \frac{dy}{dt} &= \epsilon[g(x) - y],\end{aligned}\tag{3.2}$$

kde z je vstup, ϵ vyjadruje silu väzby medzi dvoma jednotkami, $f(x)$ a $g(x)$ sú oscilačné funkcie, zvolené tak, aby vznikala silná oscilácia medzi jednotkami. Napríklad, hyperbolická funkcia $f(x) = 3x - x^3 + 2$ a sigmoidálna funkcia $g(x) = a[1 + \tanh(x/b)]$, kde a je výška a b je sklon sigmoidy. Ak sa v takomto systéme zvýši x (na začiatku kvôli vstupu z), tak sa zvýšia aj hodnoty $f(x)$, $g(x)$ a y . Akonáhle začne $f(x)$ klesať, inhibícia z y vynuluje x . V dôsledku toho dôjde aj k vynulovaniu y a celý cyklus začne odznova.

Takýto oscilátor je možné vyjadriť aj ako jeden neurón, kde excitačná zložka bude zodpovedať membránovému potenciálu a inhibičná zložka zmene potenciálu, spôsobenej aktiváciou a deaktiváciou iontových kanálov (Wang, 2000). V pôvodnej forme sú jednotky oscilátora interpretované ako združená aktivita populácie neurónov rovnakého typu nachádzajúcich sa v rovnakom kortikálnom stĺpci (pyramídové pre excitačnú jednotku, hviezdicové pre inhibičnú). Takéto oscilátory môžu byť spojené do siete a na základe znamienka synapsie môžu byť ich fázy synchronizované, alebo desynchronizované. Podobné oscilačné neurónové siete už boli použité v segmentačných

úlohách. Napríklad letecké snímky, alebo snímky mozgu boli segmentované na homogénne oblasti (Liu and Wang, 1999); (von der Malsburg and Buhmann, 1992), prípadne reč bola segmentovaná od okolitého hluku (Wang and Brown, 1999).

Výhoda oscilačných modelov spočíva v tom, že celý model je definovaný len dvoma premennými, čo umožňuje jednoduchšiu analýzu v porovnaní s kompartmentálnym modelom. Aktivity jednotiek môžu byť vyjadrené ako dvojrozmerný fázový obraz, čo uľahčuje analýzu rôznych druhov správania sa modelu, napríklad identifikáciu uzavretého oscilačného cyklu.

Vo všeobecnosti poskytujú oscilačné modely popis neurónu na vyššej úrovni abstrakcie ako kompartmentálne modely, čo umožňuje jednoduchšiu analýzu a aplikáciu daného modelu. Avšak existujú modely s ešte vyššou mierou abstrakcie, kde model je tvorený len jednou premennou definujúcou membránový potenciál.

3.1.3 Integrate-and-Fire neuróny

Na popísanie stavu integrate-and-fire neurónov je použitá len jedna premenná definujúca membránový potenciál neurónu (Obrázok d). Tento typ neurónov akumuluje membránový potenciál zo vstupného signálu, ak prekročí prah vygeneruje AP a po každom AP vynuluje potenciál. Toto správanie je vyjadrené rovnicou:

$$C \frac{dV}{dt} = I(t) - \frac{V}{R}, \quad (3.3)$$

kde V je membránový potenciál, C je kapacitancia, R je odpor a $I(t)$ je vstupný prúd. Dôsledkom pôsobenia vstupného prúdu $I(t)$ dochádza k postupnému rastu membránového potenciálu. $-V/R$, takzvaný "leak" člen, spomaľuje rast potenciálu a bez vstupného prúdu vracia potenciál späť do kľudovej hladiny. Ak membránový potenciál stúpne až nad prahovú úroveň, neurón generuje AP a potenciál je následne nastavený späť na kľudovú hodnotu.

Bolo navrhnutých niekoľko variácií modelov integrate-and-fire neurónov. Jedna z najefektívnejších variácií je model s dynamickým prahom (Eckhorn et al., 1990). Prah sa v tomto modeli zvýši ihneď po vygenerovaní AP a potom postupne klesá, simulujúc tak refraktórnu periódu skutočných neurónov. "Leaky" synapsie ako aj dynamický prah môžu byť popísané spoločnou rovnicou:

$$x(t) = X(t) * K(t), \quad (3.4)$$

kde $*$ je operátor konvolúcie, $x(t)$ je membránový potenciál, alebo prah v čase t a $X(t)$ vstupný signál reprezentujúci prijatý, alebo vygenerovaný AP. Konvolučné jadro $K(t)$ je dané ako:

$$K(t) = \begin{cases} \exp(-\lambda t) & t \geq 0 \\ 0 & \text{inak} \end{cases} \quad (3.5)$$

kde λ určuje rýchlosť zostupu. Vygenerovanie AP spôsobí exponenciálne klesanie membránového potenciálu v čase.

Konvolučná rovnica môže byť vyjadrená pomocou rovnice (Eckhorn et al., 1990):

$$x(t) = X(t) + x(t-1) \exp(-\lambda), \quad (3.6)$$

kde t predstavuje diskretný časový krok (iteráciu). Akýkoľvek vstup z $X(t)$ spôsobí skok premennej $x(t)$, ktorá potom postupne klesá o hodnotu $\exp(-\lambda)$. Touto jednoduchou rekurzívnou rovnicou je možné vyjadriť komplexnú dynamickú povahu neuróna. Celá časová dynamika neurónu je zahrnutá v jednej premennej, čo robí tento model vhodný pre veľkorozmerné simulácie.

Integrate-and-Fire model neurónu je jednoduchý a teoreticky dobre popísaný. Bol niekoľkokrát použitý v praktických aplikáciách, vrátane segmentácie statických a

pohybujúcich sa objektov, zvukovej analýzy, regulácie aktuátorov, simulácií sekvencnej pamäti a rozpoznávania vzorov (Campbell et al., 1999); (Eckhorn et al., 1990); (Hugh and Henriquez, 2002); (Kuhlmann et al., 2002); (Rehn and Lansner, 2004); (Reitboeck et al., 1993).

3.1.4 Firing-rate neróny

Doposiaľ popísané neurónové modely môžu byť použité na modelovanie správania sa jednotlivých neurónov a temporálneho kódovania, ktoré sa dá použiť v segmentačných úlohách. Avšak na simulovanie správania sa zrakovej kôry na vyššej úrovni nie je nutné používať takto detailne definované modely. Celková aktivita je často dôležitejšia ako simulovanie temporálnej dynamiky neurónov. Jednotlivé udalosti generovania AP môžu byť zlúčené do všeobecnej aktivácie vyjadrujúcej rýchlosť aktivácie neurónu (firing-rate). Takisto aktivácie jednotlivých neurónov môžu byť agregované do jednej výpočtovej jednotky. Napríklad, rýchlosť aktivácií neurónov zrakovej kôry kóduje orientáciu a polohu podnetov (Hubel and Wiesel, 1968). Zameranie sa len na rýchlosť aktivácie umožňuje zjednodušiť výpočtový model neurónu.

Rýchlosť aktivácie je znovu závislá na membránovom potenciáli neurónu. Potenciál s je vypočítaný ako suma aktivácií η_k pre všetky neuróny k , ktoré svoje aktivácie posielajú do neurónu, vynásobená váhou synapsie w_k :

$$s = \sum_k \eta_k w_k, \quad (3.7)$$

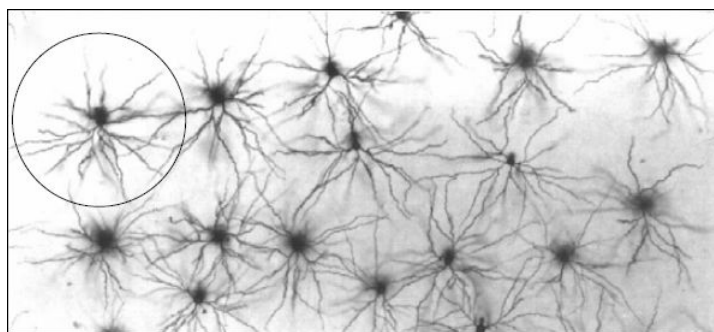
Rýchlosť aktivácie η je následne určená aplikáciou sigmoidálnej funkcie σ na membránový potenciál:

$$\eta = \sigma(s) = \frac{1}{1 + \exp(-s)}, \quad (3.8)$$

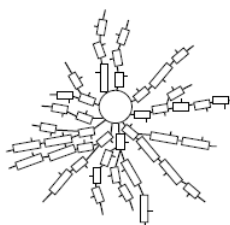
Hodnota aktivácie neurónu, teda jeho aktivačná rýchlosť, je definovaná v intervale

$< 0, 1 >$, približne modelujúc aktivačné funkcie skutočných neurónov. V mnohých prípadoch, vrátane tejto práce, môže byť ako σ použitá aj po častiach lineárna sigmoidálna funkcia, ktorá sa rýchlejšie vypočíta a dostatočne aproximuje hodnoty pôvodnej sigmoidálnej funkcie.

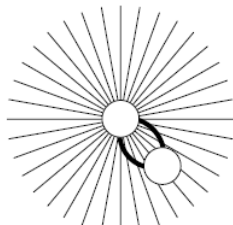
Firing-rate neuróny môžu byť použité na simuláciu sietí s veľkým počtom neurónov. Väčšina neurónových sietí, hlavne pri spracovaní jazyka, simulácii uvažovania a pamäte, rozpoznávaní reči a rozpoznávaní obrazových vzorov sú založené na firing-rate neurónových modeloch.



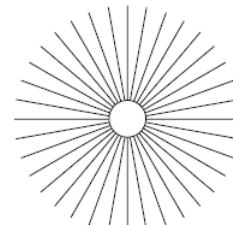
(a)



(b)



(c)



(d)

Obr. 3 – 1: Výpočtové modely neurónov a neurónových sietí. Biologické neuróny môžu byť modelované na rôznych úrovniach abstrakcie. a Mikroskopický obrázok pyramídových buniek v $1.4mm \times 0.7mm$ veľkej oblasti 3. vrstvy spánkovo-záhlavnej oblasti (TEO) makaka. Aj keď sa jedná len o zobrazenie zlomku neurónov v horizontálnom reze, dokazuje to komplexnosť štruktúry jednotlivých neurónov a ich spojení. b Podrobný compartmental model ľavého horného neurónu (zakrúžkovaný neurón). Každá zložka reprezentuje malú časť dendritu, spojenia medzi zložkami sú zobrazené čiarami. c Coupled oscillator model neurónu, pozostávajúci z excitačnej a inhibičnej jednotky s rekurentnou väzbou a s váhovaným spojením s iným neurónom v sieti. d Model, kde je aktivácia neurónu popísaná jedinou premennou, buď membránovým potenciálom (integrate-and-fire model) alebo priemerným počtom AP za jednotku času (firing-rate model). Obrázok prevzatý z (Miiikkulainen et al., 2005).

4 Použitie pulzných neurónových sietí pri spracovaní obrazu

V tejto kapitole sú popísané niektoré modely biologicky inšpirovaných pulzných neurónových sietí a ich vzťah k modelu použitému v tejto práci.

4.1 LISSOM

LISSOM (laterally interconnected synergetically self-organizing map) je biologicky inšpirovaný model ľudskej zrakovej kôry, zameriavajúci sa hlavne na konštrukciu kortikálnych zrakových oblastí (Mäkeläinen, 1991). Gro modelu tvorí hierarchické usporiadanie dvojrozmerných vrstiev neurónov reprezentujúcich rozličné oblasti mozgu: vstupná vrstva (fotoreceptory sietnice), dve vrstvy LGN neurónov (ON-center a OFF-center) vrstva kortikálnych neurónov modelujúca primárnu zrakovú kôru (V1). Neuróny v LGN aj v kortikálnej vrstve vytvárajú aferentné spojenia s neurónmi predchádzajúcej vrstvy nachádzajúcimi sa v receptívnom poli neurónu. Neuróny v kortikálnej vrstve navyše vytvárajú rekurentné laterálne prepojenia medzi neurónmi jednej vrstvy. Laterálne prepojenia sa delia na excitačné medzi susednými neurónmi a inhibičné medzi vzdialenejšími neurónmi. Váhy aferentných a laterálnych spojení sú určené pomocou normalizovaného Hebbovho učenia na základe externých vstupov.

Po natrénovaní siete predstavuje každá vrstva inú úroveň abstrakcie. LGN sa chová ako filter, ktorý extrahuje zo vstupného obrazu hrany a čiary ignorujúc nezaujímavé oblasti. Neuróny vo V1 potom reagujú len na výskyt orientovaných čiar vo výstupoch z LGN. LISSOM sieť je preto zameraná najmä na simuláciu formovania kortikálnej mapy orientácie v rámci vrstvy V1 počas perinatálneho vývoja mozgu.

4.2 PCNN

PCNN (Pulse Coupled Neural Network) je biologicky inšpirovaný model neurónovej siete na základe pozorovaní synchronných aktivít medzi neurónmi v zrakovej kôre cicavcov (Eckhorn et al., 1990).

Štandardná PCNN je jednovrstvová, dvojrozmerná neurónová sieť, ktorej štruktúra kopíruje štruktúru vstupného obrazu. To znamená, že každý neurón je asociovaný práve s jedným obrazovým prvkom obrazovej matice.

Pulzovo viazaný neurón je špeciálny typ takzvaného “Leaky integrator” neurónu s laterálnymi excitačnými prepojeniami so susednými neurónmi. Vstup do neurónu sa skladá z dvoch častí: aferentný vstup a laterálny vstup. Pri počítaní každej zložky sa do výpočtu zahrňa aj časť danej hodnoty z predchádzajúcej iterácie, čím sa modeluje efekt krátkodobej pamäte. Interný stav neurónu je vypočítaný ako súčet aferentného a laterálneho vstupu. Interný stav neurónu je potom porovnaný s hodnotou dynamického prahu a ak je väčší, tak dochádza k excitácii neurónu.

Akonáhle neurón excituje, hodnota prahu je podstatne zvýšená a následne postupom času sa jeho hodnota znižuje až pokiaľ nedosiahne kritickú hodnotu a neurón znova neexcituje. Takáto správanie imituje refraktórnu periódu biologických neurónov. Výstupom PCNN je séria binárnych obrazov, ktoré sú reprezentované maticou v jednotlivých iteračných krokoch, pričom všetky vygenerované binárne obrazy majú rovnaký rozmer ako vstupná obrazová matica. Využíva sa preto najmä pri segmentácii obrazu.

5 Výpočtový model

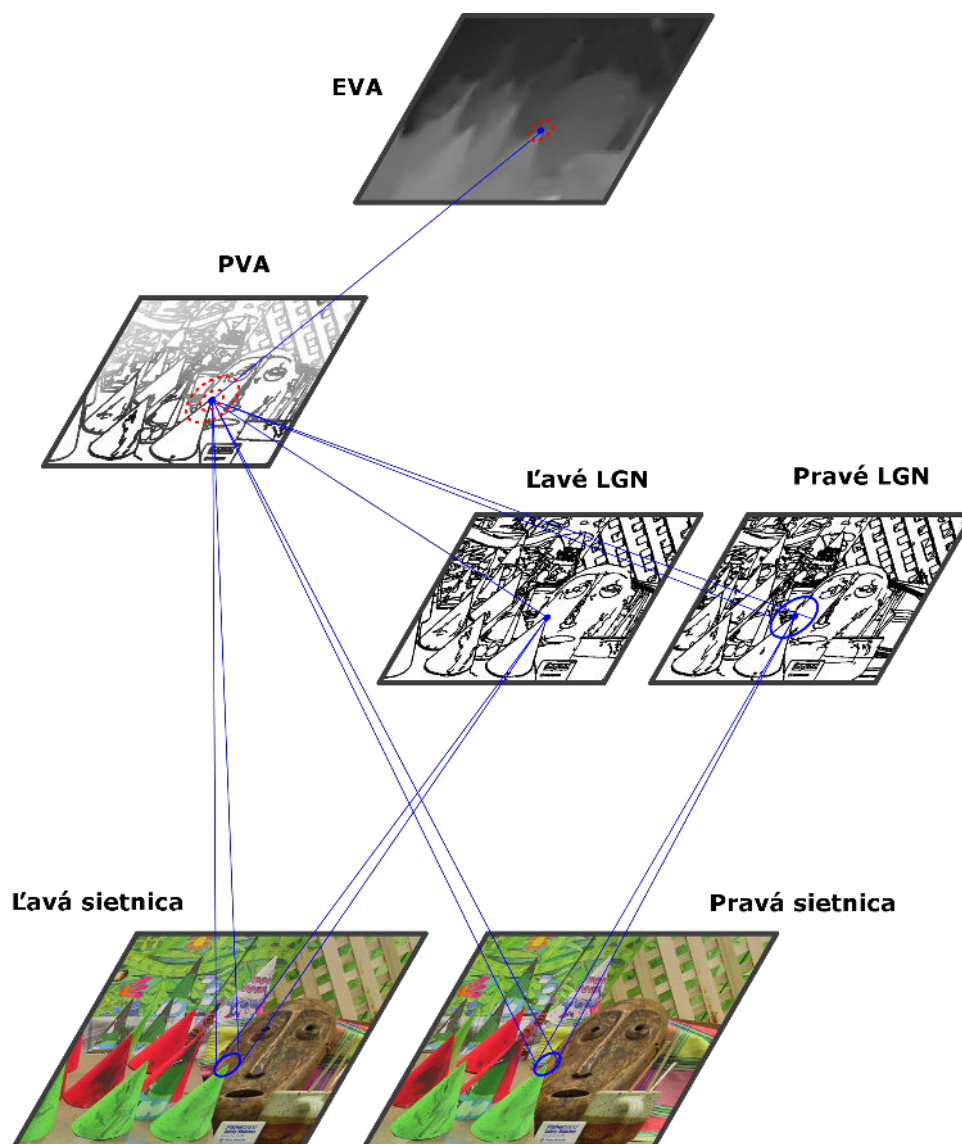
Táto práca sa zaoberá spracovaním stereovízneho obrazu s použitím neurónovej siete založenej na modeli pulzovo viazanej neurónovej siete (PCNN). PCNN je biologicky inšpirovaný model neurónovej siete, vytvorený na základe pozorovaní synchronných aktivít medzi neurónmi v zrakovej kôre cicavcov (Lindblad and Kinser, 1998). Táto kapitola detailne popisuje základnú architektúru modelu.

5.1 Architektúra

BCNN pozostáva z niekoľkých oblastí zrakovej kôry, ktoré sú nevyhnutné pre spracovanie stereovízneho obrazu. Základná architektúra je zobrazená na Obrázku 5–1. Gro modelu tvorí hierarchické usporiadanie dvojrozmerných vrstiev neurónov reprezentujúcich rozličné oblasti mozgu: dve vstupné vrstvy (dve sietnice), dve vrstvy LGN neurónov (ON-center a OFF-center) a dve vrstvy kortikálnych neurónov: primárna zraková oblasť (Primary Visual Area - PVA) a sekundárna zraková oblasť generujúca hĺbkovú mapu vstupného obrazu (Extrastriate Visual Area - EVA).

Neuróny v LGN aj v kortikálnych vrstvách vytvárajú aferentné spojenia s neurónmi predchádzajúcej vrstvy nachádzajúcimi sa v receptívnom poli neurónu. Neuróny v jednotlivých vrstvách sú navyše vzájomne prepojené, vytvárajúc laterálne interakcie medzi neurónmi jednej vrstvy. Váhy aferentných a laterálnych spojení sú nemenné a určené počas inicializácie.

Zrakový signál sa šíri postupne do vyšších vrstiev, z ktorých každá transformuje signál do relevantnejšej formy. ON a OFF vrstvy LGN slúžia na filtrovanie uniformných segmentov obrazu, ponechávajú len hrany a čiary. Vo vrstve V1 dochádza k určeniu disparity aktivovaných neurónov ľavej časti LGN na základe aktivácií neurónov v pravej časti LGN a obrazov na sietnici. Vo vrstve EVA je potom aproximovaná disparita pre všetky časti vstupného obrazu, ktoré boli odfiltrované v



Obr. 5 – 1: Architektúra modelu BSNN. Obrázok zachytáva receptívne polia jedného neurónu a vzorovú aktiváciu každej vrstvy po prezentácii jednoduchého obrazového podnetu na pravú a ľavú sietnicu. Po vykreslení podnetov na sietnice sa generovaná aktivita šíri prostredníctvom aferentných spojení až do vyšších vrstiev. LGN sa chová ako filter, ktorý extrahuje zo vstupných obrazov hrany a čiary. Následne sa v LGN pôsobením laterálnych spojení sústreďuje aktivita do okrajových častí detegovaných hrán. Vo vrstve PVA sa určí disparita ľavého hranového obrazu na základe pravého hranového obrazu a obrazov na sietnici. Vo vrstve EVA je potom pomocou laterálnych spojení aproximovaná disparita v chýbajúcich častiach.

LGN.

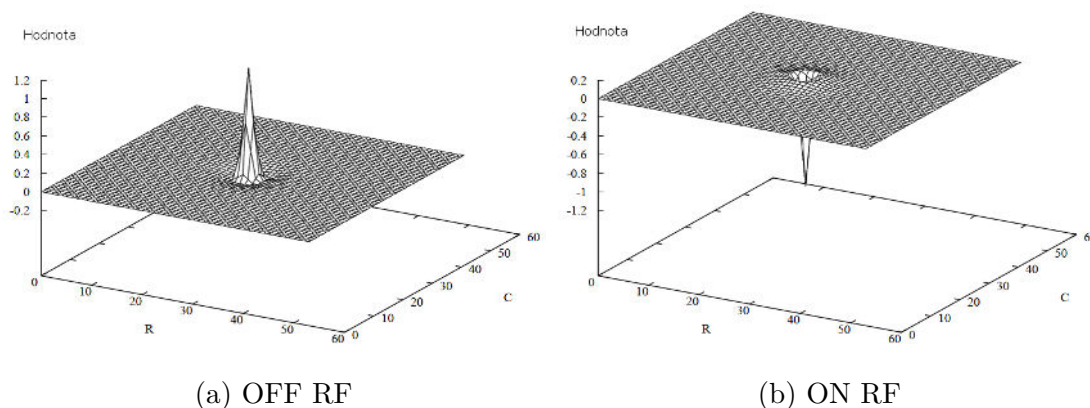
Vrstva PVA je inšpirovaná primárnou zrakovou kôrou, ktorej väčšina neurónov je binokulárnych, teda reagujúcich na vstupy z pravého aj z ľavého oka. Keďže neuróny v primárnej zrakovkej kôre dostávajú vstup z gangliových buniek sietnice a LGN, ktoré slúžia na extrakciu hrán, predpokladá sa, že vo V1 dochádza k výpočtu disparity len pre neuróny, ktoré boli aktivované prítomnosťou hrán v ich receptívnom poli. Vo vrstve EVA, ktorá je inšpirovaná sekundárnou zrakovou kôrou, dochádza následne k vytvoreniu úplnej hĺbkovej mapy. Keďže presné miesto určovania disparity v primárnej a sekundárnej zrakovkej kôre nie je známe, zvolili sa pre kortikálne vrstvy neurónovej siete všeobecné názvy ako FSA a EVA s cieľom zdôrazniť, že biologický ekvivalent týchto vrstiev sa neviaže na presnú pozíciu v zrakovom systéme.

5.1.1 Konektivita v LGN vrstve

Váhy spojení v LGN boli určené rozdielom Gaussiánov (DoG) tak, aby aproximovali receptívne polia skutočných neurónov v LGN u dospelého človeka. Najprv je určené ťažisko receptívneho poľa na základe pozície neurónu vo vstupnej vrstve korešpondujúcemu neurónu, ktorého receptívne pole určujeme. Takéto mapovanie zaisťuje, že LGN bude mať také isté dvojrozmerné topografické usporiadanie ako sietnica. Na základe pozície každého ťažiska sú následne určené jednotlivé váhy pomocou rozdielu dvoch normalizovaných Gaussiánov. Konkrétne, pre ON-center neurón na pozícii (i, j) s ťažiskom (x_c, y_c) je váha $L_{xy,ij}$ spojenia pre každý fotoreceptor sietnice (x, y) v receptívnom poli daná nasledovne:

$$L_{xy,ij} = \frac{\exp\left(-\frac{(x-x_c)^2+(y-y_c)^2}{\sigma_c^2}\right)}{\sum_{uv} \exp\left(-\frac{(u-x_c)^2+(v-y_c)^2}{\sigma_c^2}\right)} - \frac{\exp\left(-\frac{(x-x_c)^2+(y-y_c)^2}{\sigma_s^2}\right)}{\sum_{uv} \exp\left(-\frac{(u-x_c)^2+(v-y_c)^2}{\sigma_s^2}\right)}, \quad (5.1)$$

kde σ_c určuje šírku centrálného Gaussiánu a σ_s šírku okolitého Gaussiánu. Váhy OFF-center neurónov sú opačné k váham ON-center neurónov, teda sú vypočítané



Obr. 5 – 2: Receptívne polia ON a OFF neurónov. ON a OFF neuróny sú v sieti HLISSOM modelované ako rozdiel dvoch Gaussiánov (centrálneho a okolitý). Gaussiány sú síce normalizované, takže ich sumy sú rovnaké, ale centrálny Gaussián je koncentrovaný na omnoho menšej ploche. ON neuróny majú excitačný stred a inhibičné okolie a, a OFF neuróny majú inhibičný stred a excitačné okolie b. Takto formované receptívne polia fungujú ako filter, zvýrazňujúci hrany a čiary a ignoruje uniformné segmenty vstupného obrazu. Obrázok prevzatý z (Bednar, 2002).

ako rozdiel okolitého a centrálného Gaussiánu. Obrázok 5 – 2 zobrazuje typické ON a OFF receptívne polia.

Aj keď majú OFF neuróny také isté váhy ako ON neuróny, líšiace sa len znamienkom, ich činnosť nie je redundantná. Každý neurón je prahovaný tak, aby nadobúdala len kladné aktivácie, takže ON a OFF neuróny nebudú nikdy súčasne aktívne na tej istej pozícii. Toto prahovanie je inšpirované skutočnosťou, že v biologických systémoch neexistuje negatívna aktivácia neurónov. Výsledkom je, že ON a OFF vrstvy poskytujú komplementárnu informáciu, ako v modeli, tak aj v biologických systémoch.

Neuróny v LGN sú navyše medzi sebou spojené laterálnymi prepojeniami. Tie sa delia na excitačné medzi susednými neurónmi a inhibičné medzi vzdialenejšími neurónmi. Keďže je jedna hrana detegovaná viacerými susednými neurónmi, laterálne spojenia spôsobujú, že sa aktivácie neurónov po niekoľkých iteráciách skonvergujú do okrajových oblastí detegovaných hrán. Bez použitia laterálnych spojení by bola počítaná

disparita pre všetky hranové neuróny, čo by viedlo k zvýšeniu lokálnych nepresností disparitnej mapy a zvýšeniu výpočtovej náročnosti. Váhy laterálnych spojení sú určené ako normalizovaný Gaussián:

$$W_{xy,ij} = \frac{\exp(-\frac{(x-x_c)^2+(y-y_c)^2}{\sigma^2})}{\sum_{uv} \exp(-\frac{(u-x_c)^2+(v-y_c)^2}{\sigma^2})}, \quad (5.2)$$

kde $W_{xy,ij}$ je laterálna váha (E, alebo I) z presynaptického neurónu (x, y) do neurónu (i, j) a σ určuje šírku Gaussiánu.

5.1.2 Konektivita v kortikálnych vrstvách

Kortikálne vrstvy neurónovej siete sú topograficky mapované. Zo znamená, že majú rovnakú veľkosť ako predchádzajúca vrstva a každý neurón v danej vrstve zodpovedá neurónu na korešpondujúcej pozícii v predchádzajúcej vrstve. Každý neurón v kortikálnej vrstve je aferentne spojený práve s týmto jedným korešpondujúcim neurónom. Váha aferentného spojenia je nemenná a rovná 1.

Vrstva EVA navyše implementuje laterálne excitačné spojenia, ktoré spájajú neurón so sebou samým a s jeho najbližšími susedmi v kruhovom okolí. Laterálne váhy sú taktiež nemenné a určené ako normalizovaný Gaussián rovnicou 5.2.

5.2 Generovanie aktivácií neurónov

Pred každým novým spracovaním obrazového podnetu sú aktivácie každého neurónu v sieti nastavené na nulu. Systém následne obdrží vstupný podnet prostredníctvom aktivácie fotoreceptorov na sietniciach. Signál zo sietnic postupuje cez ON a OFF vrstvy LGN, kde je pomocou laterálnych spojení sústredený na okraje detegovaných hrán, až do kortikálnych oblastí, kde dôjde k určeniu disparity hrán a vytvoreniu plnej disparitnej mapy.

5.2.1 Aktivácie v LGN

Neuróny v ON a OFF vrstvách LGN počítajú svoje aktivácie ako súčet aferentnej a laterálnej stimulácie. Najprv je vypočítaná aferentná stimulácia a_{ij} LGN neurónu (i, j) ako normalizovaná suma váhovaných aktivácií fotoreceptorov v jeho receptívnom poli:

$$a_{ij} = \gamma_L \sum_{kl} L_{kl,ij} \chi_{kl}, \quad (5.3)$$

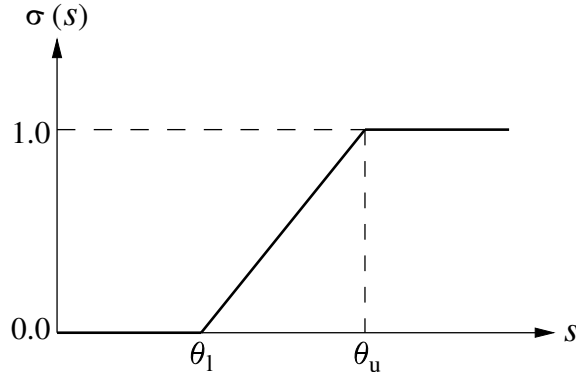
kde $L_{kl,ij}$ je aferentná váha z neurónu (i, j) do fotoreceptoru sietnice (k, l) a χ_{kl} je hodnota aktivácie receptoru (k, l) v receptívnom poli daného neurónu. γ_L je škálovacia konštanta určujúca veľkosť vplyvu aferentných stimulácií a zabraňuje tak saturácii neurónu. Počiatočná aktivácia LGN neurónov je potom vypočítaná aplikovaním aktivačnej funkcie na aferentnú stimuláciu každého neurónu:

$$\xi_{ij}(0) = \sigma(a_{ij}), \quad (5.4)$$

kde σ predstavuje po častiach lineárnu aproximáciu sigmoidálnej aktivačnej funkcie, zvolenú hlavne kvôli menšej výpočtovej náročnosti ako je u klasickej sigmoidálnej funkcie (Obrázok 5–3):

$$\sigma(s) = \begin{cases} 0 & s \leq \theta_l \\ (s - \theta_l)/(\theta_u - \theta_l) & \theta_l < s < \theta_u \\ 1 & s \geq \theta_u \end{cases} \quad (5.5)$$

V každom nasledujúcom časovom kroku je aktivácia LGN neurónov počítaná ako súčet aferentnej a laterálnej stimulácie. Laterálna stimulácia je určená ako rozdiel excitačnej a inhibičnej stimulácie neurónov v rámci LGN. Konkrétne, aktivácia ξ_{ij} v čase t LGN neurónu (i, j) je počítaná ako



Obr. 5 – 3: Aktivačná funkcia modelu BSNN. Pre aktiváciu je potrebné aby vstup do neurónu bol väčší ako prah θ_l , pri prekročení hornej hranice θ_u už dochádza k saturácii neurónu. Výstupné hodnoty sú ohraničené intervalom $[0, 1]$.

$$\xi_{ij}(t) = \sigma \left(a_{ij} + \gamma_E \sum_{kl} E_{kl,ij} \xi_{kl}(t-1) - \gamma_I \sum_{kl} I_{kl,ij} \xi_{kl}(t-1) \right), \quad (5.6)$$

kde $E_{kl,ij}$ je laterálna excitačná váha z neurónu (k, l) do neurónu (i, j) a $I_{kl,ij}$ je laterálna inhibičná váha. $\xi_{kl}(t-1)$ je aktivita neurónu (k, l) v predchádzajúcom kroku. γ_E a γ_I sú škálovacie konštanty určujúce veľkosť vplyvu laterálnych excitačných a laterálnych inhibičných interakcií.

5.2.2 Aktivácie v PVA

Aktivácie neurónov v PVA určujú disparitu aktivovaných neurónov v ľavej časti LGN na základe aktivácií neurónov v pravej časti LGN a aktivácií fotoreceptorov pravej a ľavej sietnice. Aktivácia neurónu (i, j) v PVA je určená ako rozdiel pozície (i, j) neurónu v ľavej časti LGN a takého neurónu v pravej časti LGN, pri ktorom sa aktivácie receptorov v jeho receptívnom poli na pravej sietnici najviac podobajú aktiváciám receptorov v receptívnom poli (i, j) neurónu nachádzajúcom sa na ľavej sietnici. Keďže sa predpokladá, že sú vstupné obrazy zarovnané a teda platí epi-polárne obmedzenie, skúmajú sa len tie neuróny pravej časti LGN, ktoré ležia na

rovnakej horizontálnej línii ako neurón z ľavej LGN. Konkrétne je aktivácia PVA neurónov definovaná ako

$$s_{ij} = \begin{cases} i - u_{ij} & \xi_{\lambda ij} > 0 \\ 0 & \text{inak} \end{cases} \quad (5.7)$$

kde $\xi_{\lambda ij}$ je aktivácia (i, j) neurónu v ľavej časti LGN a u_{ij} je pozícia korešpondujúceho neurónu z pravej časti LGN, určenej ako

$$u_{ij} = \left\{ x \mid \xi_{\rho xj} > 0 \wedge k_{(i,j;x,j)} = \min_{d \in (0 \dots D)} k_{(i,j;x+d,j)} \right\}_0 \quad (5.8)$$

kde $\xi_{\rho xj}$ je aktivácia (x, j) neurónu v pravej časti LGN. Keďže sa disparita počíta len pre tie neuróny z ľavej LGN, ktoré boli aktivované prítomnosťou hrán v ich receptívnom poli, predpokladá sa, že korešpondujúci neurón z pravej LGN bude taktiež aktivovaný a teda nie je nutné prehľadávať receptívne polia všetkých neurónov pravej časti LGN. D vyjadruje šírku prehľadávania v rámci pravej časti LGN. $k_{(i,j;x,j)}$ je podobnosť aktivácií receptorov ľavej sietnice nachádzajúcich sa v receptívnom poli (i, j) neurónu ľavej LGN a aktivácií receptorov pravej sietnice v receptívnom poli (x, j) neurónu pravej LGN. Konkrétne je táto podobnosť určená rovnicou:

$$k_{(i,j;x,j)} = \sum_{\Delta r = -\phi}^{\phi} \sum_{\Delta c = -\phi}^{\phi} |\chi_{\lambda(i+\Delta r, j+\Delta c)} - \chi_{\rho(x+\Delta r, j+\Delta c)}|, \quad (5.9)$$

kde $\chi_{\lambda(i,j)}$ vyjadruje aktiváciu (i, j) receptora ľavej sietnice a $\chi_{\rho(x,j)}$ vyjadruje aktiváciu (x, j) receptora pravej sietnice. ϕ určuje šírku prehľadávania. Výstup PVA neurónu je potom určený ako

$$\eta_{ij}(t) = \sigma(s_{ij}), \quad (5.10)$$

kde σ predstavuje aktivačnú funkciu definovanú rovnicou 5.5.

5.2.3 Aktivácie v EVA

V poslednej kortikálnej vrstve dochádza k aproximovaniu disparity chýbajúcich segmentov a vytvoreniu úplnej kortikálnej mapy. Aktivácia každého neurónu je vypočítaná z aferentného a laterálneho príspevku. Najprv je vypočítaná aferentná stimulácia v_{ij} kortikálneho neurónu (i, j) ako váhovaná aktivácia korešpondujúceho neurónu z ľavej LGN:

$$v_{ij} = \gamma_A \eta_{ij}, \quad (5.11)$$

kde η_{ij} je aktivácia korešpondujúceho neurónu z vrstvy PVA a γ_A je škálovacia konštanta.

Počiatočná aktivácia neurónov je potom vypočítaná aplikovaním aktivačnej funkcie na aferentnú stimuláciu každého neurónu:

$$\vartheta_{ij}(0) = \sigma(v_{ij}), \quad (5.12)$$

kde σ predstavuje aktivačnú funkciu definovanú rovnicou 5.5. Po počiatočnej aktivácii dôjde pôsobením laterálnych excitačných spjení k postupnému usádzaniu kortikálnej aktivity až do stabilného stavu. V každom časovom kroku v tomto diskretnom časovom intervale, neuróny kombinujú aferentné a laterálne excitačné stimulácie:

$$\vartheta_{ij}(t) = \sigma \left(v_{ij} + \gamma_M \sum_{kl} M_{kl,ij} \vartheta_{kl}(t-1) \right), \quad (5.13)$$

kde $M_{kl,ij}$ je laterálna excitačná váha z neurónu (k, l) do neurónu (i, j) a $\vartheta_{kl}(t-1)$ je aktivita neurónu (k, l) v predchádzajúcom kroku. γ_M určuje veľkosť vplyvu laterálnych excitačných interakcií.

Počas usádzania kortikálnej aktivity vo vrstve EVA je aferentná stimulácia nemenná. Pôsobením laterálnych spojení dôjde počas niekoľkých iterácií k rozptýleniu aktivácií neurónov v rámci celej EVA a k výpočtu úplnej disparitnej mapy.

6 Experimenty

Táto kapitola sa zaoberá aplikáciou navrhnutého modelu pulznej neurónovej siete do reálneho prostredia. V kapitole budú popísané vlastnosti použitého modelu ako aj jeho optimálne nastavenia. Takto definovaný model bude následne použitý v experimentoch zameraných na spracovanie stereovízneho obrazu.

6.1 Použitý model

Keďže vrstvy BSNN modelu sú mapované rovnako ako sietnice, teda jeden neurón vrstvy zodpovedá jednému neurónu z predchádzajúcej vrstvy, majú vrstvy rovnakú veľkosť ako sietnice. V experimentoch boli použité obrázky s rozlíšením 640x480 pixelov, preto každá vrstva modelu obsahovala práve 640x480 neurónov.

Po prezentácii vstupných obrazov na sietnice došlo počas 3 iterácií k dostatočnému skonvergovaniu aktivácií v LGN vrstvách. Následne bola vo vrstve PVA vypočítaná disparita detegovaných hrán a vo vrstve EVA počas 100 iterácií vytvorená kompletná disparitná mapa.

Veľkosť aferentného receptívneho poľa LGN neurónov bola určená definovaním šírky centrálnych a okolitých Gaussiánov na $\sigma_c = 0.3$ a $\sigma_s = 3$. Veľkosť laterálnych receptívnych polí bola určená taktiež definovaním šírky Gaussiánov na $\sigma_e = 0.5$ a $\sigma_i = 1$. Škálovacie konštanty LGN neurónov boli nastavené ako $\gamma_L = 0.0039$, $\gamma_E = 1.0$ a $\gamma_I = 1.0$.

Šírka prehľadávania v rámci neurónov pravej časti LGN, vyjadrujúca maximálnu možnú disparitu, bola nastavená ako $D = 100$. Veľkosť prehľadávacieho okna pri prehľadávaní aktivácií fotoreceptorov sietnic bola nastavená na $\phi = 2$.

Keďže aktivácie PVA neurónov sú normalizované v intervale $< 0, 1 >$, veľkosť aferentnej škálovacej konštanty EVA neurónov bola nastavená na $\gamma_A = 1.0$. Veľkosť laterálneho excitačného receptívneho poľa EVA neurónov bola určená nastavením

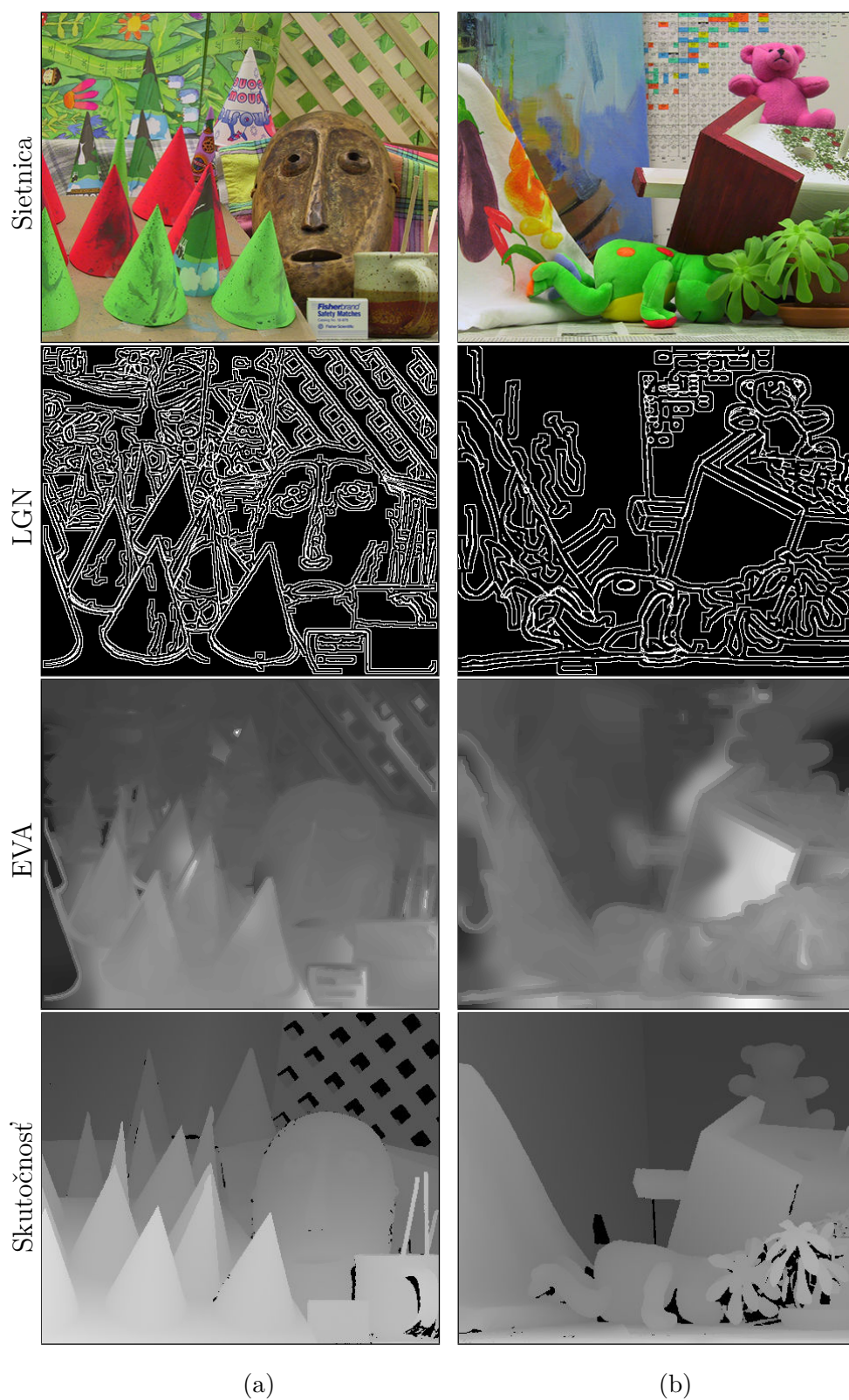
šírky Gaussiánu na $\sigma_m = 0.5$. Škálovacia konštanta γ_M bola nastavená na hodnotu 0.8.

6.2 Výsledky experimentov

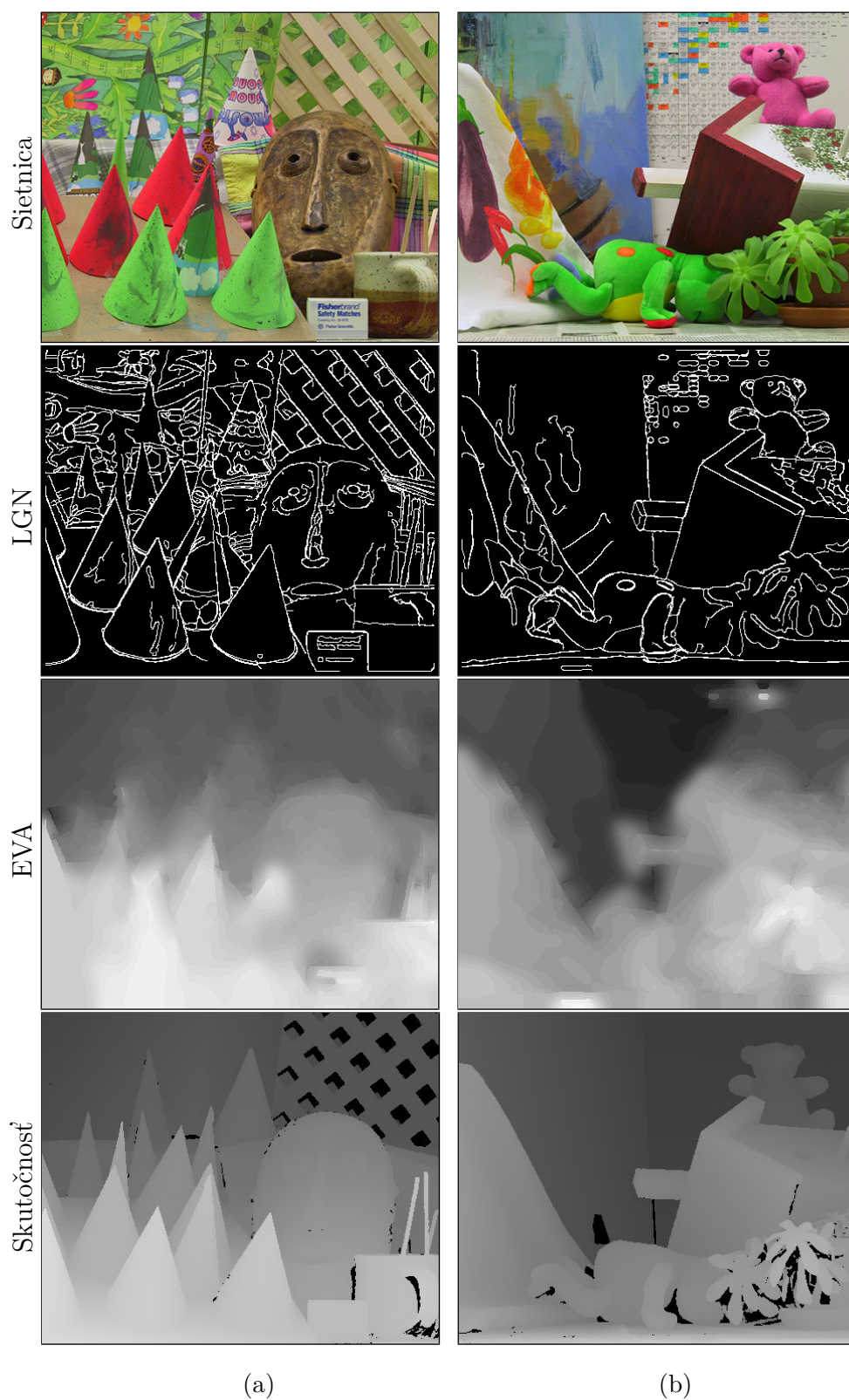
Celkovo boli realizované štyri experimenty s cieľom overiť funkčnosť modelu. Každý experiment pozostával zo spracovania dvoch dvojíc (pravý a ľavý) obrázkov a vytvorenia korešpondujúcej disparitnej mapy.

Prvý experiment bol realizovaný na obrázkoch z reálneho prostredia (Obrázok 6–1). Druhý experiment zachytáva správanie sa modelu v prípade absencie laterálnych spojení v LGN (Obrázok 6–2). V takomto prípade nedochádza k potlačeniu redundantných aktivácií susedných neurónov reagujúcich na výskyt tej istej hrany. Tretí experiment zobrazuje správanie sa modelu v prípade absencie celej vrstvy LGN, pričom vrstva PVA je spojená priamo so sietnicami (Obrázok 6–3). V takomto prípade sa vypočítava disparita každého bodu ľavej sietnice hľadaním korešpondujúceho bodu v pravej sietnici na celej epipolárnej línii. Vrstva EVA už len následne skopíruje úplnú disparitnú mapu vytvorenú vrstvou PVA. Takéto usporiadanie modelu imituje klasický "brute-force" prístup k počítaniu disparitnej mapy. V štvrtom experimente je LGN vrstva nahradená Canny hranovým detektorom, ktorý implementuje podobný mechanizmus potláčania redundantných aktivácií ako BSNN (Obrázok 6–4). Avšak na rozdiel od použitého modelu, Canny detektor potláča okrajové okrajové aktivácie zhluky reagujúceho na tú istú hranu a ponecháva tak len jednopixelovú líniu nachádzajúcu sa v strede daného zhluku.

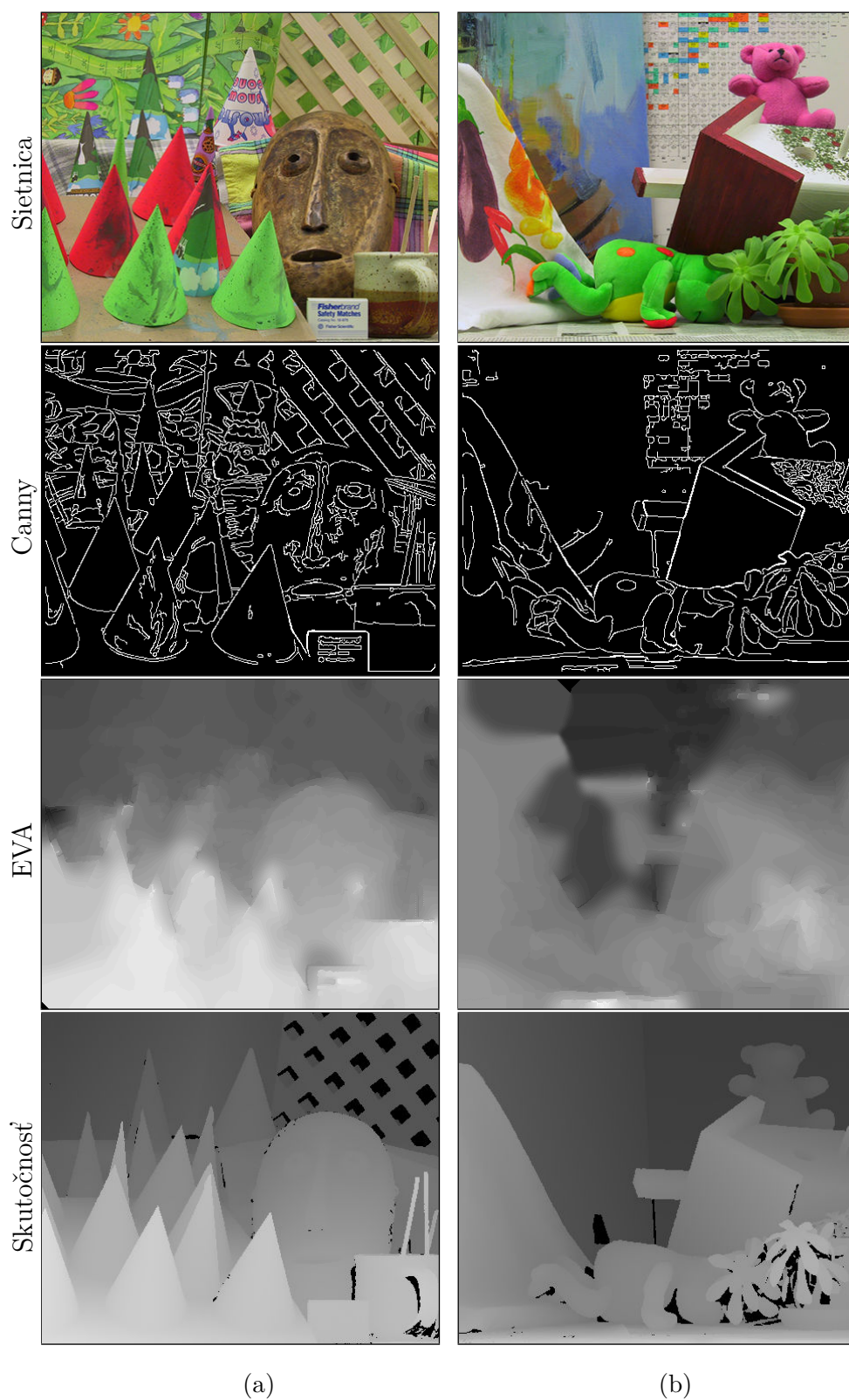
Prvý riadok z vykonaných experimentov zobrazuje obrazový podnet vykreslený na ľavú sietnicu. Druhý riadok zobrazuje korešpondujúce aktivácie neurónov v LGN. V prípade 3. experimentu druhý riadok zobrazuje extrahované hrany z obrazových podnetov pomocou Canny hranového detektora. V prípade 4. experimentu, neuróny v LGN vrstve len skopírovali aktivácie fotoreceptorov sietnice a teda nedošlo k ex-



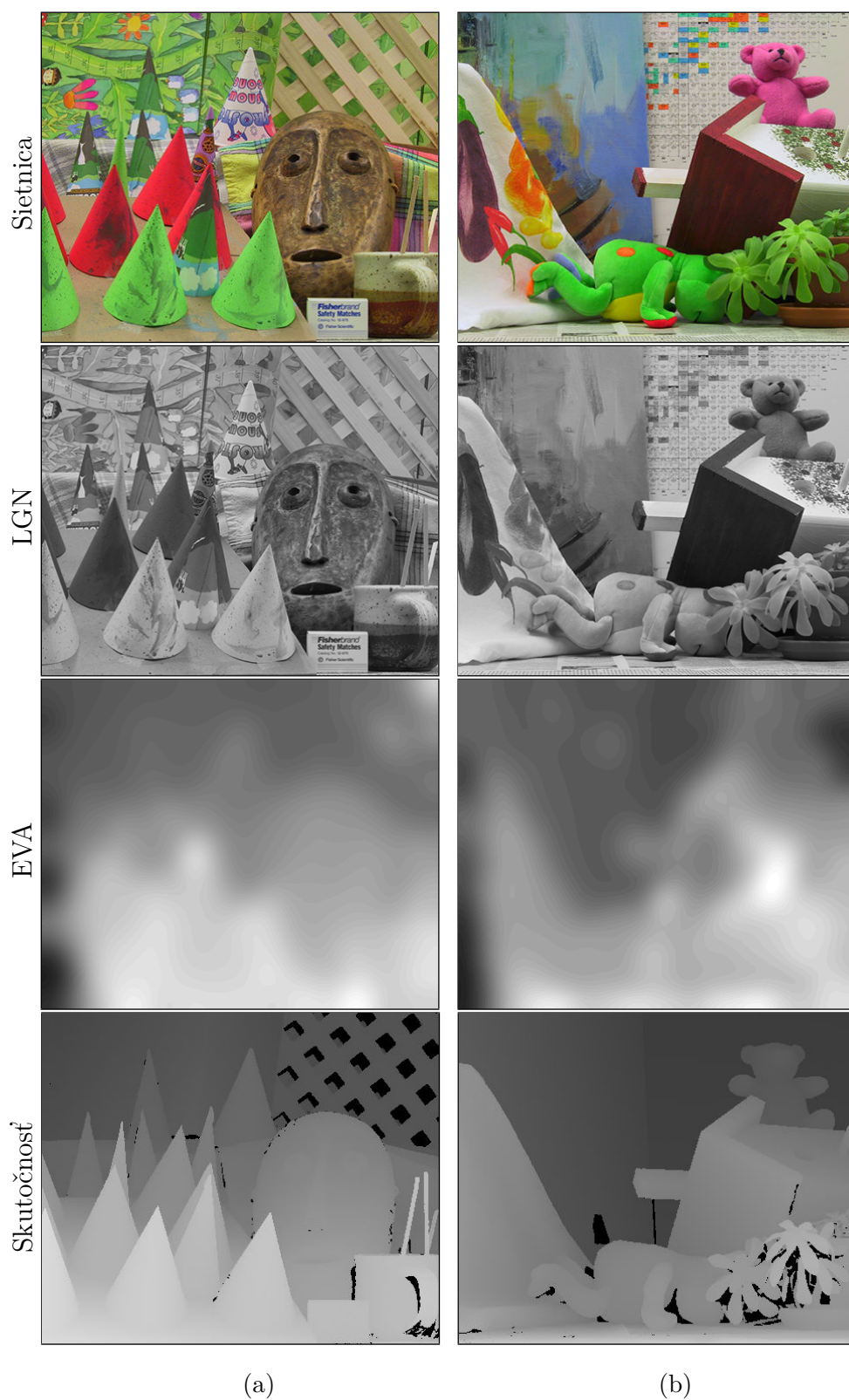
Obr. 6 – 1: Disparitná mapa obrázkov z reálneho prostredia.



Obr. 6 – 2: Disparitná mapa v prípade absencie laterálnych spojení v LGN.



Obr. 6–3: Disparitná mapa v prípade použitia Canny detektoru.



Obr. 6 – 4: Disparitná mapa v prípade absencie LGN neurónov.

trahovaniu hrán zo vstupného obrazu. V treťom riadku je zobrazená vygenerovaná úplná disparitná mapa vrstvou EVA. Štvrtý riadok zobrazuje skutočnú disparitnú mapu daného podnetu.

Z experimentov vyplýva, že BSNN môže slúžiť ako efektívna metóda pri spracovaní stereovízneho obrazu a vytváraní disparitnej mapy, pričom k vytvoreniu úplnej disparitnej mapy postačuje určiť disparity jednotlivých hrán. Plná disparitná mapa je následne vygenerované pôsobením laterálnych interakcií vo vrstve EVA. Z experimentov taktiež vyplýva, že prítomnosť laterálnych spojení už vo vrstve LGN je esenciálna pre správne vytvorenie disparitnej mapy. Pôsobením excitačných a inhibičných laterálnych spojení dôjde totiž k vytlačeniu aktivácií v LGN na okraj detegovaných hrán a potlačeniu redundantných aktivácií ostatných neurónov.

7 Záver

7.1 Splnenie pokynov na vypracovanie

1. **Vypracovať prehľad problematiky biologicky inšpirovaných metód pre spracovanie obrazu so zameraním na pulzné neurónové siete (spiking neural networks)**

Táto úloha bola splnená v 4. kapitole vytvorením stručného popisu danej problematiky a používaných metód.

2. **Návrh a implementácia systému využívajúceho pulzné neurónové siete na spracovanie stereovízneho obrazu s cieľom vytvorenia 3D modelu priestoru**

V 5. kapitole je teoreticky popísaná základná architektúra použitého modelu.

3. **Navrhnuť a realizovať experimentálne overenie funkčnosti systému.**

Realizáciou experimentov sa zaoberá 6. kapitola, kde časť 5.1 popisuje nastavenia modelu použitého v experimentoch a v časti 5.2 sú uvedené a popísané výsledky experimentov.

4. **Vypracovať dokumentáciu podľa pokynov vedúceho diplomovej práce**

Táto úloha bola splnená vypracovaním diplomovej práce, používateľskej a systémovej príručky.

7.2 Zhodnotenie, výstupy a ďalší výskum

V práci bol na spracovanie stereovízneho obrazu použitý biologicky inšpirovaný model pulznej neurónovej siete, detailne popísanej v 4. kapitole. Konkrétne bol model použitý na vytvorenie úplnej disparitnej mapy videnej scény.

Ďalší výskum na základe tejto práce by sa mohol sústrediť na zlepšenie detekcie hrán použitím viackanálového LGN, podobne ako v biologických systémoch. V takomto prípade by bolo LGN rozšírené o vrstvy, ktoré by filtrovali hrany v jednotlivých farebných kanáloch vstupného obrazu. Následne by boli výstupy jednotlivých LGN vrstiev integrované do jednotného hranového obrazu. Taktiež by mohlo byť zlepšené počítanie disparity vo vrstve PVA implementovaním laterálnych spojení, ktoré by dokázali vyhladiť vypočítanú disparitu jednotlivých hrán a odstrániť tak lokálne nepresnosti.

Literatúra

Bednar, J. A. (2002). *Learning to See: Genetic and Environmental Influences on Visual Development*, PhD thesis, Department of Computer Sciences, The University of Texas at Austin. Also Technical Report AI-TR-02-294.

URL: <http://nn.cs.utexas.edu/?bednar:phd02>

Bednar, J. A. and Miikkulainen, R. (2003). Self-organization of spatiotemporal receptive fields and laterally connected direction and orientation maps, *Neurocomputing* **52–54**: 473–480.

Blasdel, G. G. (1992). Differential imaging of ocular dominance columns and orientation selectivity in monkey striate cortex, *The Journal of Neuroscience* **12**: 3115–3138.

Callaway, E. M. and Katz, L. C. (1991). Effects of binocular deprivation on the development of clustered horizontal connections in cat striate cortex, *Proceedings of the National Academy of Sciences of the United States of America* **88**: 745–749.

Campbell, S. R., Wang, D. L. and Jayaprakash, C. (1999). Synchrony and desynchrony in integrate-and-fire oscillators., *Neural Computation* **11**(7): 1595–1619.

Dalva, M. B. and Katz, L. C. (1994). Rearrangements of synaptic connections in visual cortex revealed by laser photostimulation, *Science* **265**: 255–258.

Doya, K., Selverston, A. I. and Rowat, P. F. (n.d.). A hodgkin-huxley type neuron model that learns slow non-spike oscillation, *In Cowan et*, pp. 15–5.

Eckhorn, R., Reitboeck, H. J., Arndt, M. and Dicke, P. (1990). Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex, *Neural Computation* **2**: 293–307.

Edelman, S. (1996). Why have lateral connections in the visual cortex?, *in* J. Sirosh, R. Miikkulainen and Y. Choe (eds), *Lateral Interactions in the Cortex: Structure and Function*, The UTCS Neural Networks Research Group, Austin, TX.

-
- Erwin, E., Obermayer, K. and Schulten, K. (1995). Models of orientation and ocular dominance columns in the visual cortex: A critical comparison, *Neural Computation* **7**: 425–468.
- Gerstner, W. and Kistler, W. M. (2002). *Spiking Neuron Models: Single Neurons, Populations, Plasticity*, Cambridge University Press.
- Gilbert, C. D. and Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex, *The Journal of Neuroscience* **9**: 2432–2442.
- Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve., *The Journal of physiology* **117**(4): 500–544.
- Horn, D. and Opher, I. (1998). Collective excitation phenomena and their applications.
- Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurons in the cat’s striate cortex, *Journal of Physiology* **148**: 574–591.
- Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat’s visual cortex, *Journal of Physiology (London)* **160**: 106–154.
- Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex, *Journal of Physiology (London)* **195**: 215–243.
- Hugh, N. and Henriquez (2002). A simulator for the analysis of neuronal ensemble activity: application to reaching tasks, *Neurocomputing* **44?46**: 847?854.
- Kuhlmann, L., Burkitt, A. N., Paolini, A. and Clark, G. M. (2002). Summation of spatiotemporal input patterns in leaky integrate-and-fire neurons: Application to neurons in the cochlear nucleus receiving converging auditory nerve fiber input., *Journal of Computational Neuroscience* **12**(1): 55–73.
-

-
- Lindblad, T. and Kinser, J. M. (1998). *Image processing using pulse-coupled neural networks.*, Perspectives in neural computing, Springer.
- Liu, X. and Wang, D. L. (1999). Range image segmentation using a relaxation oscillator network., *IEEE Transactions on Neural Networks* **10**(3): 564–573.
- Mainen, Z. F. and Sejnowski, T. J. (1998). Modeling Active Dendritic Processes in Pyramidal Neurons, in C. Koch and I. Segev (eds), *Methods in Neuronal Modeling*, second edn, MIT Press, Cambridge, MA, pp. 170–209.
- Miikkulainen, R. (1991). Self-organizing process based on lateral inhibition and synaptic resource redistribution, *In Proceedings of the International Conference on Artificial Neural Networks*, North-Holland, pp. 415–420.
- Miikkulainen, R., Bednar, J. A., Choe, Y. and Sirosh, J. (2005). *Computational Maps in the Visual Cortex.*, Springer.
- Purves, D. (2012). *Neuroscience*, Sinauer Associates, Incorporated.
URL: <http://books.google.sk/books?id=B5YXRAAACAAJ>
- Rall, W. (1962). Theory of physiological properties of dendrites, *Annals of the New York Academy of Science* **96**: 1071–1092.
- Rall, W. (1977). Core conductor theory and cable properties of neurons, in E. R. Kandel (ed.), *Handbook of Physiology: Volume I*, American Physiology Society, Bethesda, MD, pp. 39–97.
- Rall, W. and Agmon-Snir, H. (1998). Cable theory for dendritic neurons, in C. Koch and I. Segev (eds), *Methods in Neuronal Modeling: From Ions to Networks*, 2 edn, MIT Press, chapter 2, pp. 27–92.
- Rehn, M. and Lansner, A. (2004). Sequence memory with dynamical synapses., *Neurocomputing* **58-60**: 271–278.
- Reitboeck, H., Stoecker, M. and Hahn, C. (1993). Object separation in dynamic
-

- neural networks, *Proceedings of the IEEE International Conference on Neural Networks* (San Francisco, CA), Vol. 2, Piscataway, NJ: IEEE, pp. 638–641.
- Rinzel, J. and Ermentrout, G. B. (1989). Methods in neuronal modeling, MIT Press, Cambridge, MA, USA, chapter Analysis of Neural Excitability and Oscillations, pp. 135–169.
- URL:** <http://dl.acm.org/citation.cfm?id=94605.94613>
- Ruthazer, E. S. and Stryker, M. P. (1996). The role of activity in the development of long-range horizontal connections in area of the ferret, *The Journal of Neuroscience* **16**: 7253–7269.
- Sabatini, S. P., Solari, F. and Secchi, L. (2004). A continuum-field model of visual cortex stimulus-driven behaviour: emergent oscillations and coherence fields., *Neurocomputing* **57**: 411–433.
- Schutter, E. and Bower, J. M. (1994). An active membrane model of the cerebellar Purkinje cell. II, *Journal of Neurophysiology* **71**: 401–419.
- Solomon, S. G. and Lennie, P. (2007). The machinery of colour vision, *Nature Reviews Neuroscience* **8**: 276–286.
- Somers, D. C., Toth, L. J. and Todorov, E. (1996). Variable gain control in local cortical circuitry supports context-dependent modulation by long-range connections, in J. Sirosh, R. Miikkulainen and Y. Choe (eds), *Lateral Interactions in the Cortex: Structure and Function*, The UTCS Neural Networks Research Group, Austin, TX.
- Stemmler, M., Usher, M. and Niebur, E. (1995). Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics., *Science* **269**: 1877–1880.
- Terman, D. and Wang, D. (1995). Global competition and local cooperation in a network of neural oscillators, *Physica D* **81**: 148–176.

-
- Van Essen, D., Anderson, C. and Felleman, D. (1992). Information processing in the primate visual system: an integrated systems perspective, *Science* **255**(5043): 419–423.
- Van Hooser, S. D., Roy, A., Rhodes, H. J., Culp, J. H. and Fitzpatrick, D. (2013). Transformation of receptive field properties from lateral geniculate nucleus to superficial v1 in the tree shrew., *J Neurosci* **33**(28): 11494–505.
URL: <http://www.biomedsearch.com/nih/Transformation-receptive-field-properties-from/23843520.html>
- von der Malsburg, C. and Buhmann, J. (1992). Sensory segmentation with coupled neural oscillators, *Biological Cybernetics* **67**: 233–242.
- Wallace, M. T., Mchaffie, J. G. and Stein, B. E. (1997). Visual response properties and visuotopic representation in the newborn monkey superior colliculus, *Journal of Neurophysiology* **78**(5): 2732–2741.
- Wandell, B. (1995). *Foundations of vision*, Sinauer Associates.
URL: <http://books.google.sk/books?id=dVRRAAAAMAAJ>
- Wang, D. (1995). Emergent synchrony in locally coupled neural oscillators., *IEEE Transactions on Neural Networks* **6**(4): 941–948.
- Wang, D. (2000). On connectedness: A solution based on oscillatory correlation, *Neural Computation* **12**:131–139.
- Wang, D. L. and Brown, G. J. (1999). Separation of speech from interfering sounds based on oscillatory correlation., *IEEE Transactions on Neural Networks* **10**(3): 684–697.
- Wiskott, L. and von der Malsburg, C. (1996). Face recognition by dynamic link matching, in J. Sirosh, R. Miikkulainen and Y. Choe (eds), *Lateral Interactions in the Cortex: Structure and Function*, The UTCS Neural Networks Research Group, Austin, TX.
-

Zoznam príloh

Príloha A DVD médium – diplomová práca s prílohami v elektronickej podobe, zdrojové texty.

Príloha B Používateľská príručka

Príloha C Systémová príručka