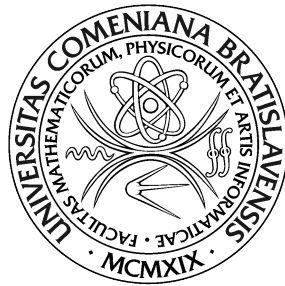


UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



UČENIE INVARIANTNÝCH SENZO-MOTORICKÝCH
REPREZENTÁCIÍ POHYBOV UCHOPOVANIA
POMOCOU NEURÓNOVÝCH SIETÍ

Diplomová práca

2018

Bc. Jakub Pospíchal

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



UČENIE INVARIANTNÝCH SENZO-MOTORICKÝCH REPREZENTÁCIÍ POHYBOV UCHOPOVANIA POMOCOU NEURÓNOVÝCH SIETÍ

Diplomová práca

Študijný program: Aplikovaná informatika
Študijný odbor: 2511 Aplikovaná informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Školiteľ: prof. Ing. Igor Farkaš, Dr.

Bratislava, 2018

Bc. Jakub Pospíchal



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

- Meno a priezvisko študenta:** Bc. Jakub Pospíchal
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: aplikovaná informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický
- Názov:** Učenie invariantných senzo-motorických reprezentácií pohybov uchopovania pomocou neurónových sietí
Learning invariant senso-motor representations of grasping movements using neural networks
- Anotácia:** Modelovanie vzniku invariantných reprezentácií je súčasťou procesu abstrakcie, ktorá vzniká počas ontogenézy. Teoretickým pozadím oblasti je teória zrkadliacich neurónov, ktorá ponúka vysvetlenie porozumenia pozorovaným akciám na základe mapovania vizuálnych vstupov na vlastný motorický repertoár.
- Cieľ:** 1. Navrhnete model neurónovej siete učiacej sa vizuálne a motorické reprezentácie pohybov robotického ramena, a ich prepojenie, s využitím teórie zrkadliacich neurónov.
 2. Navrhnete a otestujete neurálny mechanizmus na učenie invariantných reprezentácií pohybov ramena s využitím prepojenia oboch modulov, s cieľom implementovať hypotézu priameho mapovania.
- Literatúra:** Rebrová K., Pecháč M., Farkaš I.: Towards a robotic model of the mirror neuron system. In International Conference on Development and Learning and on Epigenetic Robotics, IEEE. 2013.
 Farkaš I., Rebrová K.: Bidirectional activation-based neural network learning algorithm. In Proceedings of the International Conference on Artificial Neural Networks (ICANN). Springer, 154-161, 2013.
- Vedúci:** prof. Ing. Igor Farkaš, Dr.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: prof. Ing. Igor Farkaš, Dr.
Dátum zadania: 14.12.2015
- Dátum schválenia:** 14.12.2015
- prof. RNDr. Roman Ďurikovič, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Čestne prehlasujem, že túto diplomovú prácu som
vypracoval samostatne len s použitím uvedenej literatúry
a za pomoci konzultácií u môjho školiteľa.

Bratislava, 2018

.....

Bc. Jakub Pospíchal

Pod'akovanie

Týmto by som sa chcel poďakovať môjmu školiteľovi prof. Ing. Igorovi Farkašovi, Dr. za vedenie, rady a pomoc pri tvorbe diplomovej práce. Ďalej by som sa chcel poďakovať mojej rodine a priateľom za podporu prejavenu počas celého štúdia.

Abstrakt

V tejto práci sa venujeme problematike zrkadliacich neurónov a jej aplikáciu v kognitívnej robotike. Navrhli sme model umelej rekurentnej samoorganizujúcej sa neurónovej siete využívajúcu metódu priameho párovania na prepojenie motorickej oblasti F5 a senzorickej oblasti STS. Vďaka tomuto prepojeniu sme vytvorili systém variantných a invariantných zrkadliacich neurónov, ktoré sme využili na klasifikáciu jednotlivých úchopov. Model je natrénovaný na dátach získaných pomocou simulátoru iCub. V práci analyzujeme výsledky simulácií a porovnávame ich s hodnotami nameraných počas experimentov.

Kľúčové slová: Zrkadliace neuróny, variantné a invariantné neuróny, uchopovanie objektov, iCub, neurónová sieť

Abstract

In this research we focused on the issue of mirror neuron and its application in cognitive robotics. We designed an artificial recurrent self-organizing neural network model which uses the method of direct matching for connecting motor area F5 with sensory area STS. Based on this connection we created a system of variant and invariant mirror neurons for classifying each particular grasps. The model is trained on data gained from iCub simulator. In the research we analyze the results of these simulations and compare them with values acquired from previous experiments.

Keywords: Mirror neurons, variant and invariant neurons, objects grasping, iCub, neural network

Obsah

| | | |
|----------|---|-----------|
| 1 | Úvod | 1 |
| 2 | Biologická inšpirácia | 3 |
| 2.1 | Prehľad problematiky | 3 |
| 2.1.1 | Zrkadliace neuróny | 3 |
| 2.1.2 | Variantné a invariantné neuróny | 8 |
| 3 | Použité modely | 10 |
| 3.1 | Prehľad modelov | 10 |
| 3.1.1 | RRNPB | 11 |
| 3.1.2 | MNS Mirror neuron system | 12 |
| 3.1.3 | Iné modely | 15 |
| 3.2 | Návrh modelu | 16 |
| 3.3 | MSOM | 19 |
| 3.3.1 | Rekurentné mapy | 19 |
| 3.3.2 | MSOM | 21 |
| 4 | Implementácia | 24 |
| 4.1 | iCub | 24 |
| 4.1.1 | Fyzický robot | 24 |
| 4.1.2 | Robotický simulátor | 26 |

| | |
|---|-----------|
| <i>OBSAH</i> | ix |
| 4.2 Databáza úchopov | 27 |
| 4.2.1 Vytvorenie úchopov | 27 |
| 4.2.2 Motorické a vizuálne reprezentácie | 30 |
| 4.3 Implementácia modelu | 34 |
| 4.3.1 F5 a STS | 34 |
| 4.3.2 Implementácia zrkadliacich neurónov | 35 |
| 5 Experimenty | 38 |
| 5.1 F5 a STS | 38 |
| 5.2 Zrkadliace neuróny | 45 |
| 5.3 Porovnanie máp s biologickými experimentami | 52 |
| 6 Záver | 62 |
| Zoznam literatúry | 64 |

Kapitola 1

Úvod

Fenomén zrkadliacich neurónov patrí medzi najväčšie objavy neurovedy, psychológie a kognitívnej vedy za posledné obdobie (Rizzolatti, 1988). Zrkadliace neuróny sú neuróny objavené v motorickej časti mozgu, ktoré reagujú aj na vizuálny vstup. Spolu s ich objavom sa vynorilo množstvo otázok o tomto systéme spojené najmä s ich štruktúrou a funkcionalitou. Ich existencia pomáha vysvetliť rôzne otázky spojené najmä s pochopením videnej udalosti, jej prenesením na svoj vlastný motorický aparát a na predikciu danej udalosti.

Pochopením tohto, do dnešných čias ešte stále nedostatočne popísaného systému, sa venuje značné množstvo prác, ktoré okrem iného skúmajú využitie týchto poznatkov v umelej inteligencii a prevažne v kognitívnej robotike. Vychádzajúc najmä z prác Caggiana a jeho kolegov (Caggian et al., 2009, Caggiano et al., 2011) sa zameriame na variantné a invariantné zrkadliace neuróny a využijeme ich vlastnosti na klasifikáciu úchopov pozorovaných z viacerých perspektív.

V našej práci predstavíme niektoré nami vybrané modely, ktorými sme sa nechali inšpirovať a navrhujeme si vlastný model, ktorý bude vychádzať

z biologicky ekvivalentných modulov. Daný systém budeme implementovať pomocou samoorganizovaných máp, konkrétne pomocou Merge SOM, ktorá pracuje s časovými sekvenciami.

Na zber dát potrebných na natréňovanie sietí použijeme simulátor iCub, ktorý je svojou detailnou konštrukciou robotického ruky vhodný na simuláciu viacerých typov úchopov. Ukážeme si výsledky analýz jednotlivých sietí a vplyv výberu parametrov na ich organizáciu. Experimenty na vytvorených mapách vyhodnotíme a porovnáme s biologickými experimentami na opiciach (makaky).

Kapitola 2

Biologická inšpirácia

2.1 Prehľad problematiky

Zrkadliace neuróny patria k najzaujímavejším a najvýraznejším objavom v neurovede z poslednej dekády minulého storočia. V tejto kapitole si opíšeme fenomén zrkadliacich neurónov a ich systému. Najskôr popíšeme ich biologický význam a štruktúru a načrtujeme s nimi spojenú problematiku.

2.1.1 Zrkadliace neuróny

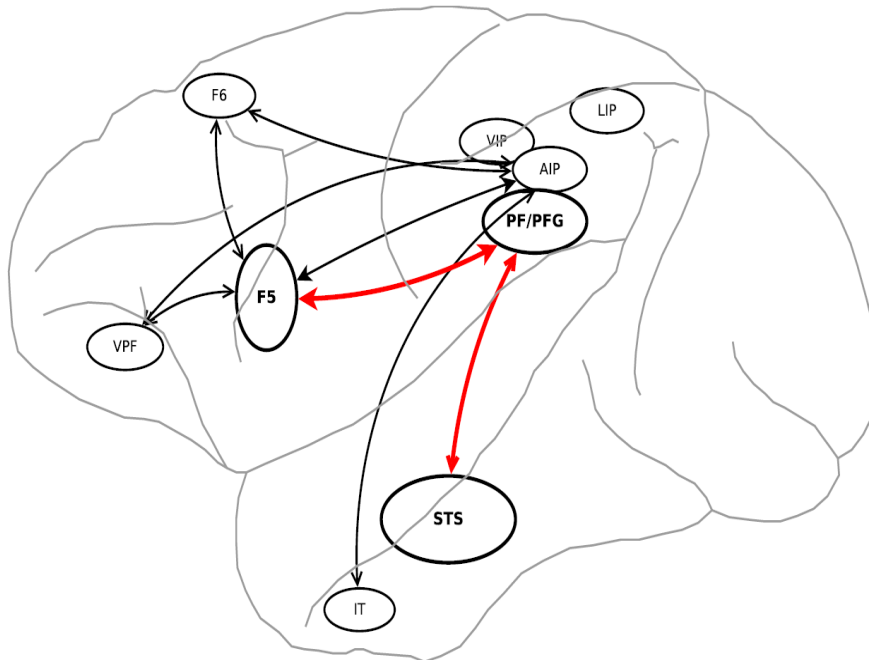
Od čias objavenia zrkadliacich neurónov na prelome 80. a 90. rokov (Pellegri-
rino et al., 1992, Rizzolatti, 1988) sa neustále vedú rôzne výskumy na túto
tému. Objavené boli v prefrontálnej kôre mozgu vo vyššej motorickej oblasti
F5 u opíc druhu Makaka sviňochvostého. Opiciam sa pripevnili elektródy
na hlavu a skúmali sa aktivity neurónov motorickej časti mozgu. Pri týchto
pokusoch sa objavili neuróny, ktoré boli aktívne nielen počas vykonávania
pohybov, ale zároveň pálili aj pri vizuálnom podnete. Opica najprv narábala
s rôznymi objektami a neskôr nehybne pozerala, ako experimentátor vyko-
náva rôzne pohyby či už so spomínanými objektami alebo bez nich. V tomto

experimente sa ukázalo, že tieto neuróny sú aktívne len v prípade, že subjekt pozoruje ako experimentátor narába s daným objektom. V prípade, že sleduje samotný objekt alebo pohyby experimentátora, ktoré sa neviazali na daný predmet zostali tieto neuróny neaktívne (Rizzolatti et al., 2001). Práve tieto a podobné zistenia vytvorili predpoklad, že tieto neuróny majú pomáhať lepšie definovať pohyb, akciu tým, že spoja vizuálny podnet s vlastným motorický aparátom. To znamená, že párovanie videného deja s motorickou skúsenosťou je nápomocné pri pochopení akcie a zároveň aj jej cieľa (Rizzolatti and Sinigaglia, 2010, Rizzolatti et al., 1996). Pri iných experimentoch sa ukázalo, že zrkadliace neuróny sú aktívne aj keď skúmaný subjekt nevidí celý objekt, ale má zakrytý výhľad (Rizzolatti and Sinigaglia, 2010). To naznačuje, že tento systém môže byť zodpovedný aj za predikcie jednotlivých akcií a spojených s daným predmetom. Podobný systém zrkadliacich neurónov sa objavil aj u človeka pomocou fMRI (Molenberghs et al., 2011). Vznik týchto neurónov je stále predmetom diskusie. S najväčšou pravdepodobnosťou sa však tieto neuróny vytvoria už v detstve spolu s motorickou časťou (Oztop and Arbib, 2002). Dôležitá skutočnosť je, že systém zrkadliacich neurónov nebol aktívny pri statickom obraze (Sebanz and Shiffrar, 2009). Pre aktiváciu bolo nutné subjektom buď ukazovať video nahranej akcie, alebo sa museli pozeráť na samotné predvádzanie pohybu.

Zrkadliace neuróny boli objavené aj v iných častiach mozgu. Spolu vytvárajú systém zrkadliacich neurónov (mirror neuron system, MNS), ktorý zastrešuje fronto-parietálny neurálny okruh, ktorý je zodpovedný za spomínané párovanie. Tento okruh u makakov pozostáva z troch častí:

- F5, oblasť v premotorickej kôre
- PF, oblasť v temennom laloku

- AIP, predná temenná oblasť

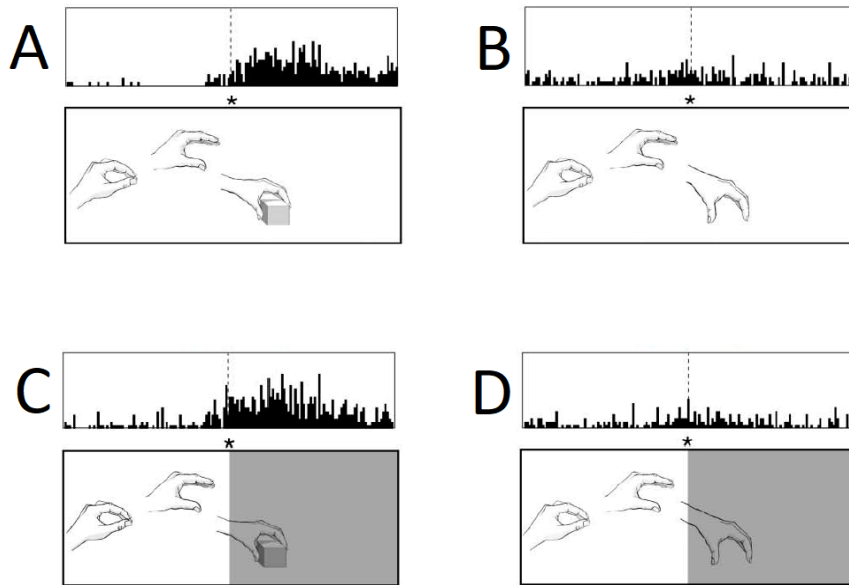


Obr. 2.1: Kľúčové oblasti fronto-parietálneho okruhu a ich vzájomné prepojenia. Prevzaté z (Rebrová and Farkaš, 2013a).

Okrem spomínaných oblastí sa zrkadliace neuróny našli aj v oblastiach F6, prednej medzitemennej oblasti (VIP) a bočnej medzitemennej oblasti (LIP). Veľmi dôležitú úlohu zohráva horná spánková brázda (superior temporal sulcus, STS), ktorá reaguje na vizuálne podnety. Informácie ďalej posiela do premotorickej kôry (F5) a temenného laloku (PF, AIP). Keďže reaguje len na jeden druh stimulu, na vizuálny, a nijak nereaguje na motorický aparát, nie je možné ju zaradiť do MNS ako skutočnú súčasť. Slúži však ako vstup pre ostatné časti systému.

Podľa rôznych štúdií dochádza k pochopeniu akcie práve pri interakciách medzi STS a F5 (Molenberghs et al., 2011). Nevyriešeným problémom zostáva určenie, ktorá časť má najväčší podiel na porozumení motorického deja.

Najrozšírenejšou teóriou je teória priameho párovania (direct-matching). V nej hrá dominantnú úlohu oblasť F5. Podľa nej sa vizuálne vnemy mapujú na vlastný motorický systém a STS v nej zohráva úlohu iba ako vstup. Naopak odporcovia tejto teórie tvrdia, že dominantnú časť tvorí STS a spracovanie pohybu prebieha najmä tu (Hickok and Hauser, 2010). Kompromisom je model, ktorý navrhol Tessitore s kolegami (Tessitore et al., 2010). Podľa tohto modelu obe časti spolupracujú. Tok informácií je obojsmerný a spracovanie deja zrkadliacimi neurónmi v oblasti F5 pomáha STS pri spracovaní vizuálnych vstupov. Zároveň tento model rieši problematiku variantných neurónov a navrhuje spôsob interpretácie pohybu ako skupinu parametrov určujúcich natočenie jednotlivých kĺbov. Dôležitá je úroveň porozumenia daného javu. V prípade hrubej abstrakcie pohybu nie je nutné mapovať daný pohyb na svoj vlastný pohybový aparát a spracovať ho na vizuálnej úrovni. Avšak pri komplikovanejšom pohybe, alebo ak vyžadujeme od subjektu zopakovať daný pohyb, je nutné lepšie porozumenie akcie, čo vyžaduje náročnejšie spracovanie.



Obr. 2.2: Na obrázku sú zachytené aktivácie neurónov v rôznych situáciach. (A) Neuróny sú aktívne pri akcii v ktorom figuruje objekt. (B) Neuróny nereagujú na pohyby, ktoré sa neviažu na žiaden objekt. (C) Neuróny sú aktívne aj v prípade, že daný objekt je schovaný za nepriehľadnou stenou, ale iba v prípade, že subjekt predtým videl, že za stenou sa daný objekt nachádza. (D) V prípade, že subjekt nevidel, že za stenou je položený objekt, neuróny zostávajú neaktívne. Takéto a im podobné experimenty ukazujú na fakt, že zrkadliace neuróny slúžia na predpovedanie cieľa akcie. Prevzaté z (Umiltà et al., 2001).

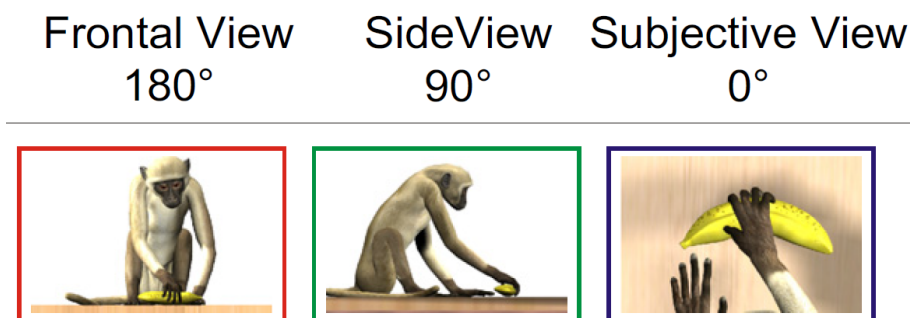
Caggiano s kolegami v štúdiách skúmali dôležitosť umiestnenia objektu a následnej vykonávanej akcii pre aktiváciu zrkadliacich neurónov. V prvej štúdií (Caggiano et al., 2009) sa zamerali na vzdialenosť vykonávanej akcie pred pozorovaným subjektom. Priestor rozdelili na dva úseky. Prvý bol malý okruh, v ktorom opica dosiahla na daný objekt a vedela s ním v tom danom okruhu manipulovať a na druhý, v ktorom už objekt nedosiahla. Experiment dokázal, že vzdialenosť nezohráva pre systém zrkadliacich neurónov veľkú

úlohu a mapovanie pozorovaného deja na motorickú časť prebieha aj pri väčšej vzdialenosti. V druhej štúdií sa zamerali najmä na pozorovanie akcie z rôznych perspektív, pričom zistili, že zrkadliace neuróny v F5 sa podobne ako v STS delia na variantné a invariantné (Caggiano et al., 2011).

2.1.2 Variantné a invariantné neuróny

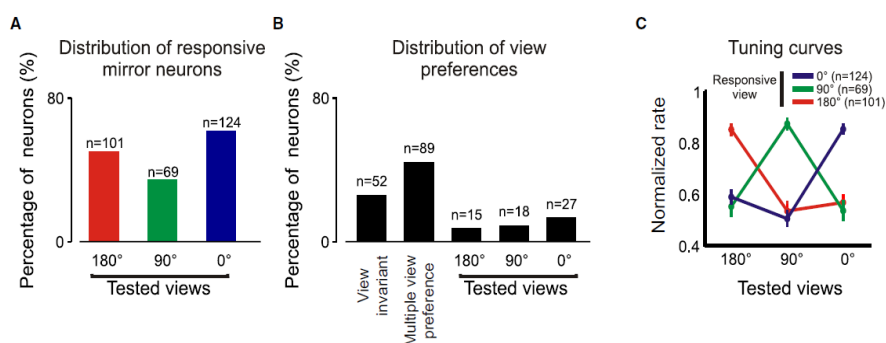
Zo štúdie Perreta a jeho kolegov (Harries et al., 1991, Perrett et al., 1989) vyplýva, že v STS sa nachádzajú neuróny, ktoré reagujú na pohyb rôzne. Väčšina z nich sú variantné, čo znamená, že reagujú na podnet len z určitej perspektívy, uhla. Obsahuje ale aj invariantné, ktoré pália nezávisle od uhla, pod akým pozorovateľ sleduje daný objekt a teda kódujú akciu všeobecným spôsobom. Invariantné neuróny sú hierarchicky nad variantnými, a prijímajú od nich aktivácie. Biologicky sú zároveň bližšie k frontálnej časti kôry, ktorá je zodpovedná za vytváranie konceptov. Existencia invariantných neurónov v motorickej oblasti F5 vedie k abstraktnému pochopeniu pozorovaného deja, čo umožňuje lepšie porozumenie akcií a jej cieľa.

Caggiano sa vo svojom výskume (Caggiano et al., 2011) venoval existencii variantných a invariantných zrkadliacich neurónov v oblasti F5, ktoré reagovali na vizuálny vstup. V prvom experimente sa zamerali na rozdiel aktivácií neurónov, keď opica sledovala úchop naživo a porovnávali výsledky s aktiváciami, keď opica sledovala úchop z videa. Následne zistili, že aktivácie sú veľmi podobné a nenašiel sa žiaden veľký rozdiel medzi nimi. V druhej časti sa zamerali najmä na aktivácie 389 zrkadliacich neurónov z motorickej časti pri sledovaní akcie z troch rôznych perspektív, uhlov. Spredu (180°), z boku (90°) a z vlastnej perspektívy (0°), čiže opica pozerala na video, v ktorom sama uchopovala daný objekt.



Obr. 2.3: Tri rôzne perspektívy z ktorých opica sledovala daný dej. Prevzaté z (Caggiano et al., 2011).

Z pozorovaných 389 neurónov 201 (52%) výrazne reagovalo na vizuálny vnem. Z toho 149 (74%) neurónov reagovalo na vstup variantne, pričom každý významne preferoval aspoň jednu z daných perspektív, nie však všetky. Zvyšných 52 (26%) reagovalo na vstup invariantne, čiže vykazovali aktivitu zo všetkých troch perspektív. Z variantných neurónov až 89 (60%) pálilo na dve rôzne perspektívy a zvyšných 60 (30%) pálilo iba na jednu perspektívu. 27 (13%) na 0°, čiže zo svojej perspektívy, 15 (8%) na 180°, čiže spredu a 18 (9%) na bočnú perspektívu 90°.



Obr. 2.4: Populačná analýza odoziev neurónov na podnety z troch rôznych perspektív. Prevzaté z (Caggiano et al., 2011).

Kapitola 3

Použité modely

3.1 Prehľad modelov

Existuje niekoľko modelov napodobňujúcich funkciu zrkadliacich neurónov. Každý z nich ale pristupuje k problematike rôznym spôsobom. Vzhľadom na to, že ešte stále sme nenazbierali dostatočné množstvo poznatkov na to, aby sme s určitosťou objasnili funkcionality tohto komplexného systému nám dáva možnosť pristupovať k problematike alternatívne. Každý z modelov, ktorý si predstavíme, má špecifický prístup, čo nám ponúka množstvo rôznych inšpirácií a spôsobov ako chápať problematiku.

V mnohých modeloch sa však stretávame s tým, že sa neprihliada na rôznorodosť perspektív vizuálnych podnetov. Tento problém sa pokúsili riešiť najmä modely Tessitoreho a kolegov (Tessitore et al., 2010). Inšpiráciou sa nám stali aj modely Oztopa a Arbiba (Oztop and Arbib, 2002), ktoré sa zo všetkých spomínaných najviac približuje k biologickej predlohe. Po predstavení vybraných modelov si opíšeme nami navrhnutý model. V tejto kapitole si predstavíme aj model MSOM, ktorý využijeme pri implementácii modelu.

3.1.1 RRNPB

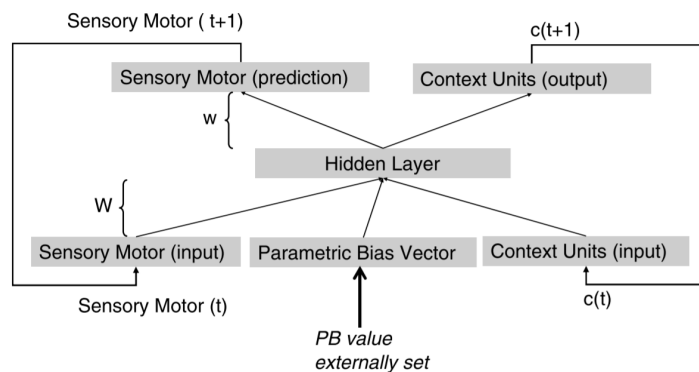
Prvý z modelov, ktorý si predstavíme je architektúra učenia RRNPB, alebo aj recurrent neural network with parametric biases (rekurentná neurónova sieť s parametrickými biasmi) (Ito and Tani, 2004). Tento model je zameraný na naučenie, napodobňovanie a samotné generovanie akcií. Vstupné informácie pre tento systém sú časovo-priestorové vzory, ktoré reprezentujú daný pohyb. Tieto vzory sú asociované s parametrickými bias vektormi. Po naučení bias vektory reprezentujú jednotlivé akcie a pomocou nich vie systém zopakovať, vykonať danú akciu. Model ma tri operačné módy:

- F5, Učiaci mód
- PF, Mód generovania akcií
- AIP, Rozpoznávací mód

Počas učiaceho módu sa siete predkladajú sensoricko-motorické vnemy, ktoré sú reprezentované ako dva vektory. Jeden reprezentuje vizuálny vnem ako pozíciu a pohyb ruky a druhý zase motorický vnem ako vektor uhlov jednotlivých kĺbov ruky. Sieť sa učí podobne ako klasická rekurentná sieť pomocou algoritmu spätného šírenia chyby. Jednotlivé váhy sa adaptujú tak, aby sieť čo najlepšie predpovedala dané vnemy. Okrem toho sa však vytvárajú aj spomínané parametrické bias vektory pre každú jednu akciu.

V druhom móde vieme generovať jednotlivé akcie. Po naučení sú spomínané parametrické bias vektory asociované s jednotlivými akciami a sú fixné. Vieme ich teda použiť ako vstup pre tento systém, ktorý následne vie vyprodukovať sensoricko-motorický reťazec, ktorý predstavuje danú akciu. Na vstup treba predkladať aj predikcie, ktoré pri tomto móde sieť vytvára.

V móde rozpoznávania sa vypočítavajú parametrické bias vektory len na základe pozorovaného deja. Na vstup prichádzajú len vektory reprezentujúce



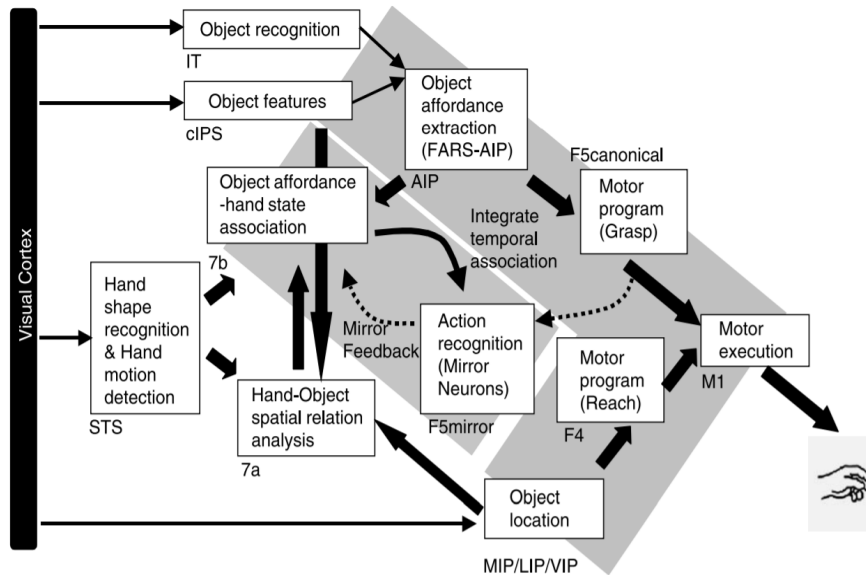
Obr. 3.1: Ukážka RNNPB systému v generovacom móde. Na vstup prichádza parametrický bias vektor a pomocou neho sa vypočítavajú sensorické a motorické reťazce. Tie už v ďalšom kroku slúžia aj ako vstup. Prevzaté z (Oztop et al., 2006).

vizuálny vnem a pomocou nich sa generujú predikcie. Vypočíta sa chybový signál a ten sa pomocou back-propagation aplikuje na parametrické bias vektory. Tie sa následne ďalej prepočítavajú a postupne konvergujú k jednotlivým akciám. Tento model má spojitosť so zrkadliacimi neurónmi práve vďaka schopnosti systému rozpoznávať ale aj generovať akcie. V tomto modeli fixné parametrické bias vektory vedú generovať akcie, ale zároveň vedú rozpoznávať akcie vďaka sensorickým vnemom postupným konvergovaním k danej akcii. Tieto vlastnosti sú veľmi podobné vlastnostiam zrkadliacich neurónov.

3.1.2 MNS Mirror neuron system

V tomto modeli si ako predlohu vývojári zobrali biologický model zrkadliacich neurónov u opíc (Oztop and Arbib, 2002). Na rozdiel od predchádzajúceho modelu, ktorý nezodpovedal skutočnému biologickému systému sa tento model snaží čo najviac priblížiť k skutočnému zloženiu systému zrkadliacich neurónov. Z viacerých štúdií vyplýva, že systém zrkadliacich neurónov nie je

E. Oztop et al. / Neural Networks 19 (2006) 254–271



Obr. 3.2: Schéma modelu MNS. Ako základné dva prvky modelu sú moduly 7b a F5. Tie predstavujú základné časti systému zrkadliacich neurónov a na obrázku sú zobrazené ako 2 šedé obdĺžniky. Na obrázku je vidno aj ich jednotlivé časti ako aj ich vstupy a vzájomné prepojenia. Prevzaté z (Oztop et al., 2006).

vrodený, ale postupne sa vytvára v rannom veku (Kohler et al., 2002). Tento model vychádza priamo z tejto hypotézy, ktorá vraví, že systém vzniká samo-skúmaním vlastných akcií, v našom prípade úchopov, čiže pri učení využíva vizuálne ale aj motorické informácie. I keď systém využíva do značnej miery samoorganizáciu, jednotlivé moduly sú presne definované a pospájané ako vidno na obrázku 3.2.

Základom modelu sú dve oblasti mozgu obsahujúce zrkadliace neuróny, 7b a F5. Súčasne do systému patria aj oblasti, ktoré neobsahujú zrkadliace neuróny, ale slúži ako vstup alebo na predspracovanie vnemov. Oblasť IT a cIPS na základe vizuálneho vstupu extrahujú základné vlastnosti objektov

ako je napríklad tvar, povrch, natočenie, pozícia a podobne. Modul AIP tak vie na základe ich výstupov vytvoriť takzvané affordances pre videný objekt. Afordancie si vieme predstaviť ako súbor akcií, ktoré vieme s daným objektom vykonať.

Na základe týchto informácií vie motoricky modul F5 vykonať danú akciu na blízko. Na dočiahnutie daného objektu je dôležitá spodná motorická kôra konkrétne F4. Tá dostáva informácie od MIP-LIP-VIP, ktorá určuje presnú pozíciu objektu. Takže informácie z F4, dôležité na dosiahnutie objektu a zároveň výstup z F5, ktorý určuje ako uchopiť objekt vstupujú do modulu M1, ktorý reprezentuje primárnu motorickú kôru, ktorá vykoná danú akciu.

Model využíva jednu z vlastností systému zrkadliacich neurónov a to, že neuróny nepália ak akcia nie je spojená s nejakým objektom. Na to model používa takzvaný „stav ruky“, ktorý reprezentuje aktuálny stav končatiny pomocou vizuálnej reprezentácie. Vizuálna informácia ako tvar a pohyb ruky pochádza z modulu STS, ktorý informácie získava z vizuálnej kôry a pomocou 7a modulu sa spracuje na jej priestorové vzťahy. Tento „stav ruky“ slúži ako vstup pre jadro systému zrkadliacich neurónov, F5 a 7b. Vizuálny vstup je asociovaný s afordanciami vytvorenými ako sme spomínali v predchádzajúcom kroku. Vďaka tomuto spôsobu asociácie systém zrkadliacich neurónov v F5 vie rozpoznávať akcie nezávisle od vykonávateľa. To znamená, že systém vie namapovať stav ruky na svoj vlastný motorický aparát v prípade, že pozoruje samého seba, ale aj v prípade, že pozoruje niekoho iného. To, že zrkadliace neuróny nedostávajú vstup priamo z vizuálnej kôry ale ako spomínaný „stav ruky“ spôsobuje to, že zostanú neaktívne ak akcia nie je zameraná na nejaký objekt, čiže neexistujú žiadne afordancie.

Okruh zrkadliacich neurónov bol tvorený doprednou sieťou a naučený pomocou algoritmu spätného šírenia chyby. V neskoršom modeli MNS2 bola

však nahradený rekurentnou sieťou, ktorá sa biologicky viac približovala k mozgu opíc. Tak isto sa podarili nasimulovať audio-vizuálne zrkadliace neuróny (Kohler et al., 2002) a vznikali aj odozvy v prípade, že objekt bol schovaný a akciu nebolo možné vidieť celú (Umiltà et al., 2001).

3.1.3 Iné modely

Za ďalší model, ktorý rieši problematiku zrkadliacich neurónov môžeme považovať MOSAIC (Modular selection and identification for control) (Haruno et al., 2001). I keď tento model bol vytvorený predovšetkým za účelom motorického ovládania, môžeme ho priradiť aj do skupiny modelov našej problematiky. Kľúčovými prvkami tohto modelu sú decentralizované moduly ovládačov, ktoré sú zodpovedné za vykonanie pohybov. Tieto moduly vytvárajú takzvané páry s prediktormi. Ovládače, ktoré čo najlepšie predpovedajú daný pohyb sa stávajú signifikantnejšie a preberajú väčšiu časť kontroly. Z toho pohľadu môžeme na tento model pozeráť aj ako na dobrý nástroj na rozpoznávanie akcií. Navyše tento model je špecifický tým, že namiesto jedného centralizovaného ovládača má viacero a práve modularita umožňuje modelu adaptovať sa na rôzne typy problémov.

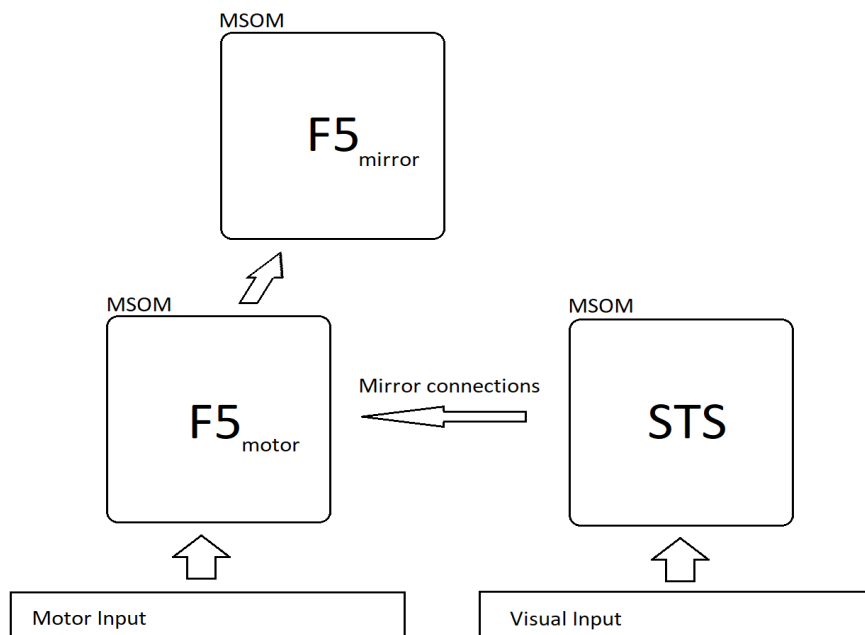
Vychádzajúc z modelu MOSAIC, páni Demiris a Hayes vytvorili takzvaný Modulárny model učenia a imitácie (Demiris and Hayes, 2002). Ten sa na rozdiel napríklad od modelu RRNPB nesnaží poňať celé správanie jednou sieťou, ale ku každému jednotlivému typu správania vytvorí zvláštny modul, PID kontrolér. Tieto moduly dostávajú na vstup informácie ako má daná akcia vyzeráť. Ich výstupom je reťazec motorických úkonov, ktoré sa majú vykonať a tie vstupujú do doprednej siete, ktorá predpovedá nasledujúci stav. Výstup jednotlivých modulov sa porovnáva so skutočným stavom a tak sa ohodnotí úspešnosť modulu. Ten modul, ktorý najlepšie predpovedá nasle-

dujúce stavy je potom považovaný za vhodného reprezentanta danej akcie, správania. V prípade, že žiaden z modulov dostatočne dobre neodhadol skutočné správanie, vytvorí sa nový modul, ktorý danému správaniu zodpovedá. Podobnosť so zrkadliacimi neurónmi je v tom, že sieť je aktívna, aj keď akcie vykonáva ale aj keď ich pozoruje.

Vyššie spomenuté modely sa stali inšpiráciou pri tvorbe nášho vlastného modelu. Za zmienku stoja aj iné modely ako napríklad model využívajúci evolučný prístup (Borenstein and Ruppin, 2005), model MSI (mental state inference) alebo model využívajúci obojsmerné prepojenie medzi F5 a STS (Rebrová and Farkaš, 2013b, Rebrová et al., 2013). Pre bližšie zoznámenie s jednotlivými modelmi môžete využiť prehľad modelov v článku Oztopa a Arbiba (Oztop et al., 2006).

3.2 Návrh modelu

Z predchádzajúcich modelov, ktoré sme si predstavili je vidno, že systém zrkadliacich neurónov je pomerne zložitý a obsiahnuť všetky jeho biologické aspekty je nateraz ťažko dosiahnuteľné. V našom modeli sme sa zamerali na spracovanie motorických a senzorických informácií, na komunikáciu medzi motorickou a vizuálnou oblasťou a na vznik variantných a invariantných zrkadliacich neurónov. V modeli nepredpokladáme nutnosť narábania s objektom na aktiváciu zrkadliacich neurónov. Cieľom nášho modelu je vytvorenie invariantných neurónov, ktoré budú vedieť rozoznať úchop nezávisle od perspektívy pozorovania. Tak isto sa v modeli zameriame na rozpoznávanie typu úchopu. V práci navrhne model, ktorý bude trénovaný na dátach získaných z robotického simulátoru iCub (Metta et al., 2008), ktorý si predstavíme neskôr. Náš model bude pozostávať z dvoch biologicky inšpirovaných



Obr. 3.3: Návrh modelu, inšpirovaný skutočnými biologickými oblasťami.

oblastí mozgu makaka a to z F5 a STS. F5 je motorická časť, ktorá okrem spracovania motorických vnemov bude obsahovať aj zrkadliace neuróny a bude zodpovedná aj za ich variantné a invariantné vlastnosti. STS bude vizuálna časť, ktorá bude zodpovedná za spracovanie vizuálnych vnemov.

Na vytvorenie oblastí F5 a STS použijeme samoorganizujúcu mapu MSOM (Merge SOM) alebo aj zlučovaciu samoorganizujúcu mapu (Strickert and Hammer, 2005), ktorú si predstavíme neskôr. Ako prvé v robotickom simulátore iCub vytvoríme databázu úchopov. Bude mať tri základne typy úchopov, ktoré sa budú líšiť najmä spôsobom, aký daný objekt uchopia, ale aj spôsobom, akým sa k nemu priblížia.

- Silový úchop (power grasp)
- Jemný úchop (precision grasp)

- Bočný úchop (side grasp)

Z nasimulovaných úchopov extrahujeme 16 obrazov (frames), ktoré budú reprezentovať daný uchop. Jeden obraz jedného úchopu bude reprezentovaný 2 reťazcami. Prvý, motorický bude poskytovať informácie o nastaveniach a hodnotách jednotlivých kĺbov robotického ramena. Tento reťazec bude slúžiť ako vstup pre motorickú oblasť F5. Druhý reťazec bude poskytovať vizuálnu reprezentáciu ramena. V našom prípade to budú predspracované vizuálne informácie o polohe robotickej ruky, tvorené súradnicami jednotlivých kĺbov. Na efektor sa budeme pozeráť zo 4 rôznych uhlov. Z vlastnej perspektívy (0°), sprava (90°), spredu (180°) a zľava (270°). To znamená, že jedna inštancia úchopu, bude mať 4 rôzne vizuálne reprezentácie vzhľadom k perspektíve.

Na základe týchto vstupov naučíme oblasti F5 a STS rozpoznávať jednotlivé typy úchopov a v rámci STS budeme rozlišovať aj perspektívy. Na vytvorenie zrkadliacich neurónov využijeme metódu jednosmerného prepojenia F5 a STS (direct matching). To znamená, že neuróny v F5 budú mať okrem základných vektorov motorických váh aj vektory váh, ktoré budú prepojené s STS. Výstupy z už natrénovaných oblastí F5 a STS pre jednotlivé úchopy budú slúžiť na natrénovanie zrkadliacich neurónov. Vhodným výberom parametrov chceme doceliť, aby sa tento systém zrkadliacich neurónov vyprofiloval tak, aby čo najlepšie spĺňal biologickú predlohu a vytvoril tak variantné, ale hlavne aj invariantné neuróny vzhľadom na perspektívu. To znamená, že chceme dosiahnuť, aby niektoré neuróny boli aktívne pri jednom type úchopu, ale zároveň pre viac ako jednu perspektívu. Výsledky budeme porovnávať s hodnotami nameranými pri pokusoch spomenutých v časti biologickej inšpirácie.

3.3 MSOM

V praxi sa často stretávame s dátami, ktoré sú reprezentované určitými postupnosťami, alebo sekvenciou jednotlivých vzoriek. V umelej inteligencii pomerne často pracujeme napríklad s DNA reťazcami alebo s rečou, čo sú typické príklady sekvenčných dát. Takéto typy dát zohrávajú veľkú úlohu aj pri spracovaní biologických signálov ako napríklad motorické alebo vizuálne stimuly. Práve preto modely, ktoré pracujú s takýmto typom dát sú dôležité nielen z praktického, ale aj z kognitívneho hľadiska. Pracovanie s takýmto typom dát vyžaduje zložitejšie modely, ktoré pracujú nielen s aktuálnym vstupom, ale berú do úvahy aj predchádzajúce vstupy. To spôsobuje, že modely bývajú spravidla výpočtovo a pamäťovo zložitejšie. Pre naše účely bolo vhodné použiť samoorganizujúce sa mapy s rekurentným kontextom (Kohonen, 1997).

3.3.1 Rekurentné mapy

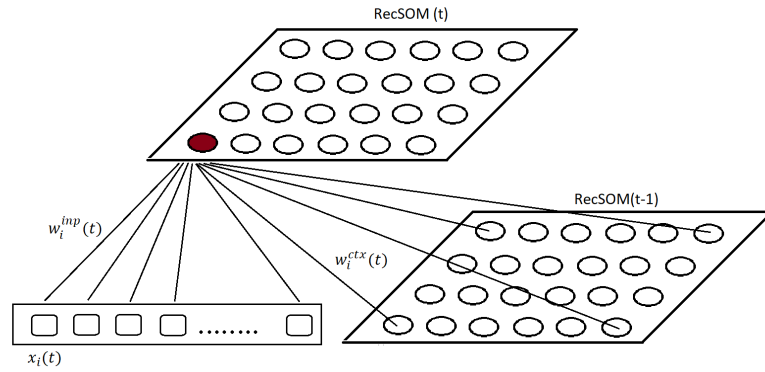
Všetky modely spomenuté v tejto práci vychádzajú zo základného modelu samoorganizujúcej sa mapy, ktorej autorom je Kohonen (Kohonen, 1997). Tento typ sietí je typický príklad modelu reprezentujúci typ učenia bez učiteľa. To znamená, že algoritmus nedostáva informácie o požadovanom výsledku počas tréningu. Váhy sa adaptujú tak, aby odzrkadľovali štatistickú povahu vstupných dát, ktoré sú reprezentované vektormi. Typickou vlastnosťou samoorganizujúcich sa sietí je, že extrahujú charakteristické črty vstupných dát, pričom zobrazenie zachová topológiu. Mapa je tvorená spravidla mriežkou $\mathcal{N} \times \mathcal{N}$ neurónov, ktorý každý má vektor váh $\mathbf{w}_i \in R^n$, ktorý sa postupne

adaptuje podľa pravidla:

$$\Delta \mathbf{w}_i = \gamma \cdot h_\sigma(d(i^*, i)) \cdot (\mathbf{x} - \mathbf{w}_i) \quad (3.1)$$

kde vektor \mathbf{x} je vstupný vektor reprezentujúci jednu vzorku dát a \mathbf{w} je vektor váh i -teho neurónu, ktorý práve adaptujeme. i^* je pozícia víťazného neurónu takého, ktorého vzdialenosť od aktuálneho vstupu \mathbf{x} bola najmenšia. d je funkcia, ktorá vracia vzdialenosť i -teho neurónu, ktorý práve upravujeme od víťaza v danej mriežke a h_σ je klesajúca funkcia okolia, ktorá definuje rozsah kooperácie medzi neurónmi a spravidla má tvar Gaussovej krivky.

Tento klasický typ máp však nezohľadňuje postupnosť vzorov prichádzajúcich na vstup. Pre takýto typ dát je treba základný model rozšíriť o rekurentné spojenia. Existuje viacero modelov rekurzívnych samoorganizujúcich máp (SOM). Do tejto kategórie modelov môžeme zaradiť napríklad TKM (Temporal Kohonen Map) (Chappell and Taylor, 1993), RSOM (rekurentná SOM) (Hagenbuchner et al., 2003), RecSOM (rekurzívna SOM) (Varsta et al., 1997), SOMSD (SOM for Structural Data) (Voegtlin, 2002) a nami zvolená MSOM resp. Merge SOM (Strickert and Hammer, 2005). V tejto práci sa nebudeme podrobne venovať popisom ostatných sietí, iba popíšeme fungovanie MSOM. Všetky hore spomenuté modely používajú klasický spôsob zmeny váh. Každá z nich však funkciu d počíta inak, pričom každá berie rôznym spôsobom do úvahy časový kontext. Typický príklad je RecSOM, ktorá má okrem klasického váhového vektoru aj kontextový váhový vektor asociovaný s kontextovým vektorom. Kontextový vektor reprezentuje aktiváciu danej mapy v predchádzajúcom kroku, viď obrázok 3.4. To zabezpečuje, že výstup nezávisí iba od aktuálneho vstupu, ale aj od predchádzajúcich.



Obr. 3.4: Model RecSOM. Okrem klasického vstupu má aj kontextový v podobe výstupu mapy z predchádzajúceho časového kroku.

3.3.2 MSOM

Model Merge SOM alebo aj zlučovacia SOM, je tvorený mriežkou $\mathcal{N} \times \mathcal{N}$ neurónov, ktoré majú okrem základných váh $\mathbf{w}_i^{\text{inp}} \in R^n$ aj kontextové váhy, $\mathbf{w}_i^{\text{ctx}} \in R^n$, ktoré sú tak isto ako základné váhy o veľkosti vstupu vid' obrázok 3.5. Vzdialenosť neurónu od vstupu sa počíta nasledovne:

$$d_i(t) = (1 - \alpha) \cdot \|\mathbf{x}(t) - \mathbf{w}_i^{\text{inp}}(t)\|^2 + \alpha \cdot \|\mathbf{q}(t) - \mathbf{w}_i^{\text{ctx}}(t)\|^2 \quad (3.2)$$

Vektor $\mathbf{x}(t)$ je vstupný vektor v čase t a $\mathbf{w}_i^{\text{inp}}$ je vektor vstupných váh v čase t . Parameter α , pre ktorý platí $0 < \alpha < 1$ je konštanta, ktorá určuje pomer, do akej miery má dôležitosť Euklidovská vzdialenosť označená ako $\|\cdot\|$, vstupných váh od aktuálneho vstupu a kontext. $\mathbf{w}_i^{\text{ctx}}$ je vektor kontextových váh, pomocou ktorých počítame vzdialenosť neurónu od \mathbf{q} , čo je kontextový deskriptor. Kontextový deskriptor je lineárnou kombináciou váh vífazného neurónu v prechádzajúcom časovom kroku a vypočítame ho ako:

$$\mathbf{q}(t) = (1 - \beta) \cdot \mathbf{w}_{b(t-1)}^{\text{inp}}(t) + \beta \cdot \mathbf{w}_{b(t-1)}^{\text{ctx}}(t) \quad (3.3)$$

$b(t - 1)$ je označenie pre víťaza v predošlom kroku, čiže $b(t - 1) = \arg \min_i \{d_i(t - 1)\}$. Parameter β , $0 < \beta < 1$, podobne ako α určuje pomer, tentokrát váh a kontextových váh predchádzajúceho víťaza, ktorý má vplyv na kontextový deskriptor. Zmena váh je podľa klasického Hebbovho pravidla:

$$\Delta \mathbf{w}_i^{\text{inp}}(t) = \gamma_1 \cdot h_{ib} \cdot (\mathbf{x}(t) - \mathbf{w}_i^{\text{inp}}(t)) \quad (3.4)$$

$$\Delta \mathbf{w}_i^{\text{ctx}}(t) = \gamma_2 \cdot h_{ib} \cdot (\mathbf{q}(t) - \mathbf{w}_i^{\text{ctx}}(t)) \quad (3.5)$$

Parameter γ_1 a γ_2 predstavujú rýchlosť učenia, $0 < \gamma < 1$, h_{ib} je klesajúca funkcia okolia, v našom prípade je reprezentovaná Gaussovou krivkou

$$h_{ib} = \exp\left(\frac{-d(i, b)^2}{\sigma^2}\right) \quad (3.6)$$

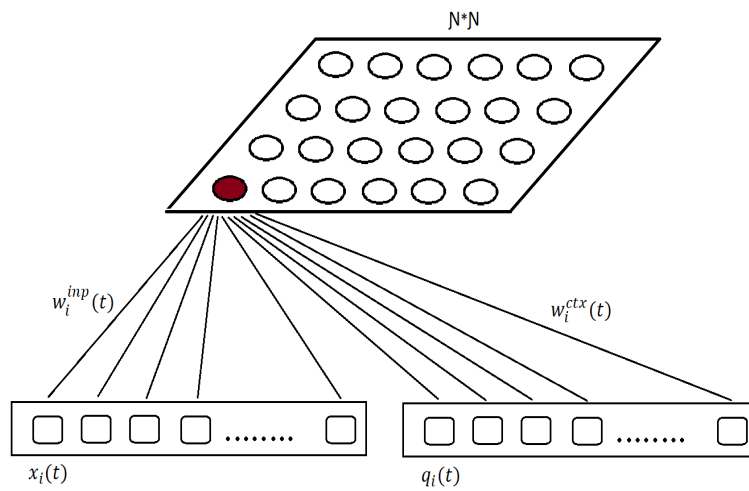
Funkcia d vracia vzdialenosť i -teho neurónu od víťaza b . Parameter σ , $1 < \sigma$, podobne ako parameter alpha postupne klesá, čo nám zaručuje čoraz menšie okolie kooperácie medzi neurónmi počas tréovania. Tréovanie prebieha tak, že jednotlivé vektory danej sekvencie predkladáme sieti v chronologickom poradí. Pre prvý vektor \mathbf{x} v danej sekvencii nemáme ako vypočítať kontextový deskriptor, takže preň platí:

$$\mathbf{q}(t) - \mathbf{w}^{\text{ctx}} = 0 \quad (3.7)$$

Výstup neurónu i pre daný vstup v čase t vypočítame:

$$y_i(t) = \exp(-d_i(t)) \quad (3.8)$$

Výhodou Merge SOM na rozdiel napríklad od známejšej RecSOM je, že



Obr. 3.5: Model MSOM. Na rozdiel od RecSOM a iných jej podobných máp má MSOM kontextový vstup len veľkosť vstupného vektoru.

má veľkosť kontextu len o veľkosti vstupu a nie o veľkosti celej mapy. V prípade, že nám postačuje výpočtovo jednoduchší kontext, môže byť Merge SOM efektívnou náhradou RecSOM (Farkaš and Vančo, 2007).

Kapitola 4

Implementácia

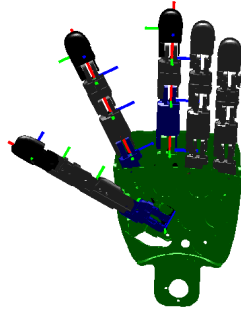
4.1 iCub

4.1.1 Fyzický robot

Robot iCub je 1 meter vysoký humanoidný robot navrhnutý tak, aby výzorom pripomínal 2,5 ročné dieťa (Metta et al., 2008). Návrh na vytvorenie takéhoto robota pochádza od Robotického konzorcia európskych univerzít a skonštruovaný bol v Italian Institute of Technology v Janove. Vyvinutý bol predovšetkým za účelom praktického overenia rôznych hypotéz z oblasti umelej inteligencie a kognitívnej robotiky. Robot iCub je úmyselne postavený ako malé dieťa z dôvodu otestovania hypotéz zaoberajúcimi sa kognitívnymi schopnosťami malých detí. Keďže dieťa skúma svet pomocou interakcie s prostredím, musel aj robot mať podobné fyzické parametre ako dieťa.

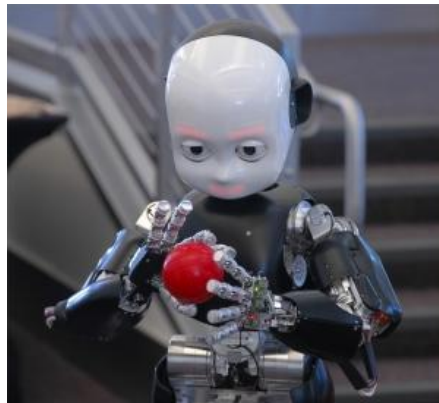
Robot iCub je za týmito účelmi používaný vo viac ako 20 laboratóriách po celom svete, ale využívaný je najmä na univerzitách v západnej Európe. Robot má až 53 stupňov voľnosti, 7 v oboch ramenách, 9 v oboch rukách (3 palec, 2 ukazovák, 2 prostredník, 1 pre malíček a prostredník, 1 na rozťaho-

vane prstov do strán), 6 v hlave (3 krk a 3 oči), 3 v páse/torze a 6 v oboch nohách.



Obr. 4.1: Stupne voľnosti na ľavej ruke.

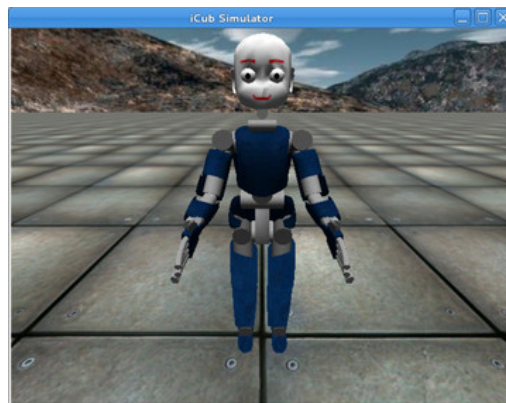
Veľký počet stupňov voľnosti v rukách je veľmi vhodný práve pre experimenty s rôznymi spôsobmi uchopovania objektov a preto sme si v tejto práci vybrali práve tohto robota. Má na tvári zabudované dve kamery v oblastiach očí a dva mikrofóny z boku hlavy. Takisto má LED senzory, ktoré imitujú obočie a ústa, čím môže robot simulovať mimiku tváre. Robot prešiel viacerými zmenami od prvotnej podoby a to predovšetkým čo sa týka končatín. Ruky sú teraz ohybnejšie a vedia lepšie uchopiť jednotlivé objekty a zmenou prešli aj dolné končatiny, ktoré teraz umožňujú prirodzenejšiu chôdzu. Musí byť však pripojený na externý zdroj energie, čo jeho pohyb, hlavne na väčšie vzdialenosti značne limituje. V novej verzii je robot obalený na viacerých miestach umelou kožou, ktorá obsahuje väčšie množstvo senzorov citlivých na tlak.



Obr. 4.2: Humanoidný robot iCub.

4.1.2 Robotický simulátor

Robot iCub je pomerne drahý a preto sa Robotické konzorcium rozhodlo vytvoriť simulátor, ktorý bude čo najhodnovernejšie napodobňovať skutočného robota (Tikhanoff et al., 2008). Tikhanoff a Fitzpatrick takýto simulátor vytvorili, je voľne šíriteľný a pre viacero laboratórií a univerzít je dostatočnou náhradou skutočného robota.



Obr. 4.3: Robotický simulátor iCub.

Robot je ovládaný pomocou YARP (Yet Another Robot Platform), čo je multiplatformový voľne šíriteľný framework, ktorý riadi komunikáciu medzi

jednotlivými časťami robota metódou peer-to-peer. Simulátor je navrhnutý tak, aby programovanie bolo skoro totožné pre reálneho robota ako aj pre simulátor. Je napísaný v jazyku C++, ale umožňuje programovanie robota aj v iných jazykoch, ako napríklad java, python alebo matlab. Tak isto obsahuje rôzne programy, ktoré uľahčujú prácu s robotom. Napríklad motorgui pomocou grafického rozhrania umožňuje vidieť všetky hodnoty kĺbov, tým sa myslí uhol daných kĺbov a zároveň môže užívateľ tieto hodnoty ručne meniť, čo sa hneď odrazí na robotovi v simulátore.

4.2 Databáza úchopov

Ako prvé sme vytvorili základnú databázu úchopov pomocou robotického simulátoru iCub. Rozhodli sme sa vytvoriť po 10 inštancií úchopov na každý z troch rôznych typov úchopov. Databáza bude teda obsahovať ich motorické a vizuálne reprezentácie. Jeden úchop bude mať iba jednu motorickú reprezentáciu, ale až 4 vizuálne, teda po jednej z každej perspektívy.

4.2.1 Vytvorenie úchopov

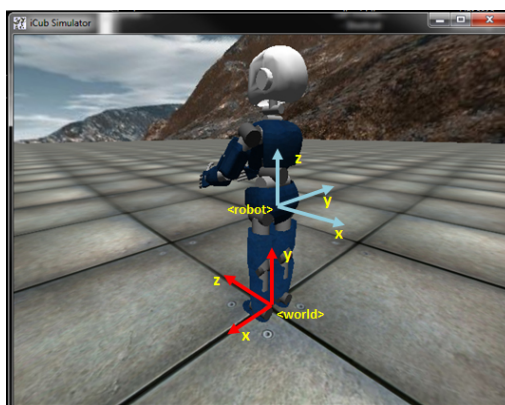
Ako sme spomínali v návrhu, cieľom bolo vytvoriť tri typy úchopov. Silový (power), jemný (precision) a bočný (side). Vychádzali sme z prirodzeného spôsobu ako uchopovať rôzne objekty. Rozdiel medzi nimi nemal byť iba v spôsobe, akým treba objekt uchopiť, ale aj spôsob, akým sa k nim robotická ruka priblížila. Aby sme vytvorili dostatočne hodnoverné úchopy najprv sme vytvorili tri typy objektov s rôznymi parametrami tak, aby sa hodili pre tri rôzne úchopy. Pre silový úchop bola vytvorená pomerne veľká kocka, ktorú nebolo možné uchopiť zvyšnými dvoma spôsobmi. Pre jemný úchop sme vytvorili maličkú guľičku, ktorú bolo možné uchopiť len troma prstami a bočný

úchop sme zase využili pri uchopovaní objektov valcového tvaru. iCub simulátor umožňuje vytváranie objektov pomocou „world“ portu. Tri základné objekty hranol, guľa, valec sa dajú vytvoriť pomocou príkazov:

- `world mk box` (tri parametre veľkosti) (tri parametre pozície) (tri parametre farby)
- `world mk sph` (parameter veľkosti) (tri parametre pozície) (tri parametre farby)
- `world mk cyl` (dva parametre veľkosti) (tri parametre pozície) (tri parametre farby)

Keď už sme mali vytvorené objekty mohli sme pristúpiť k vytváraniu samotných úchopov. V našom modeli sme sa rozhodli jeden úchop rozdeliť na dve fázy. Fázu približovanie (reaching) a na fázu úchopu (grasping). Ako prvé sme sa snažili vyriešiť približovanie pomocou knižnice iKin. Táto knižnica je určená na doprednú a inverznú kinematiku. Pomocou inverznej kinematiky vieme nastaviť parametre všetkých kĺbov od ramena až po dlaň tak, aby stred dlane bol presne na mieste, ktoré definujeme pomocou súradníc. Tento spôsob riešenia sa zdal byť na prvý pohľad najjednoduchší a časovo nenáročný. Avšak prvý problém, ktorým sme sa museli zaoberať, bol rôzny súradnicový systém pre „svet“ (myslíme tým prostredie simulátoru) a samotného robota, vid' obrazok 4.4.

Po vyriešení tohto problému sme už mohli použiť inverznú kinematiku na priblíženie stredu dlane k objektu. Už po pár pokusoch sme zistili, že robotické rameno pôsobí neprirodzene. Navyše toto riešenie spôsobovalo, že niektoré kĺby sa ohýbali veľa a iné zase málo. Najväčší problém spôsobovala dlaň. Tá sa ohýbala prílišne a pôsobila veľmi neprirodzene. Potrebovali sme, aby os dlane pretínala stred objektu. Keďže ku každému objektu, vzhľadom



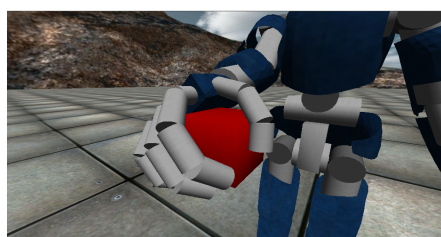
Obr. 4.4: 2 rôzne súradnicové systémy v robotickom simulátore iCub.

na typ úchopu, sa približovalo iným spôsobom, nebolo možné dlaň zafixovať. Pri jemnom úchope sa rameno malo približovať z vrchu, pri silovom šikmo a pri bočnom z boku. Aj keď sme dlaň dostali na požadované miesto nebola orientovaná správne. Preto sme sa rozhodli pre iný spôsob riešenia.

Nakoniec sme sa rozhodli využiť nástroj motorgui. V motorgui vieme nastavovať jednotlivé hodnoty kĺbov pre celú končatinu. Celá končatina má 16 stupňov voľnosti. My sme ich rozdelili na dve skupiny podľa toho, v ktorej fáze úchopu s nimi narábame. Prvých 8 reprezentuje kĺby od ramena až po dlaň vrátane. Pre už vytvorené objekty sme teda našli správne hodnoty kĺbov tak, aby dlaň bola dostatočne blízko a vedeli sme uchopiť objekt. Zvyšných 8 stupňov voľnosti máme v prstoch a tie sme využili pri samotnom uchopovaní.

V simulátore sa tvorcovia zamerali predovšetkým na pohyblivosť palca, ukazováka a prostredníka. Tie majú spolu 7 stupňov voľnosti a zvyšné dva prsty iba jeden. Keď už sme mali dlaň na správnom mieste, mohli sme objekt začať uchopovať. Tak isto ako predtým, aj teraz sme menili hodnoty kĺbov prstov pomocou motorgui tak, aby úchop bol čo najprirodzenejší. Teraz sme už mali všetky nastavenia kĺbov pre dané typy úchopu, viď obrázok 4.5.

V našom navrhnutom modeli nepracujeme s objektmi. To znamená, že



Power



Precision



Side

Obr. 4.5: Tri rôzne typy objektov uchopované troma rôznymi typmi úchopov.

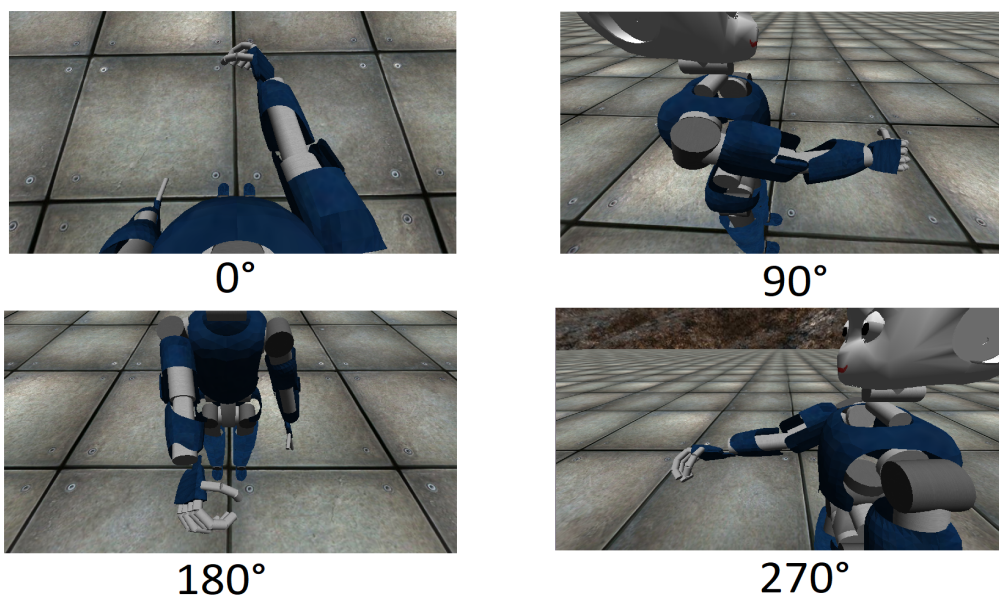
keď už sme mali hodnoty typické pre dané typy úchopov, mohli sme generovať úchopy bez potreby vytvárať objekty. Rozhodli sme sa vytvoriť 10 rôznych inštancií pre každý typ úchopu. Každý zo 16 stupňov voľnosti sme rozumne zašumili, aby sme dostali rôzne inštancie úchopov. Zašumenie prebiehalo tak, že sme pre každý kĺb zvolili interval, z ktorého môže byť vybraná hodnota pre daný kĺb. Tieto intervaly sme určili z pozorovania. Programy naprogramované v jazyku C++ na vytvorenie týchto úchopov možno nájsť v priloženom súbore zdrojových kódov v priečinku TeachersGrasp.

4.2.2 Motorické a vizuálne reprezentácie

Keď už sme mali vytvorené jednotlivé úchopy, bolo treba z nich extrahovať motorické a vizuálne informácie, ktoré budú slúžiť ako vstup pre naše neuronové siete. Podľa modelu, ktorý sme navrhli sme jeden úchop reprezentovali pomocou 16 obrazov (frames). Keďže rýchlosť, akou sa kĺby pohybujú sa dá

prispôbiť, mohli sme celý úchop vytvoriť tak, aby trval 4 sekundy. Motorické dáta sme preto mohli získať pomerne jednoducho a to tak, že sme si každých 0,25 sekundy vypýtali informácie o aktuálnom uhle kĺbu. Toto vytváranie motorických sekvencií je naprogramované spolu s vytváraním jednotlivých úchopov z predchádzajúcej časti.

Získanie vizuálnej informácie bolo však náročnejšie a vyžadovala si viacero fáz spracovania. V simulátore sa vieme pohybovať kamerou, čo nám umožňuje nasnímať jednotlivé úchopy zo 4 rôznych perspektív ako sme uviedli v návrhu, viď obrázok 4.6.



Obr. 4.6: Štyri rôzne perspektívy, z ktorých sledujeme úchopy.

V simulátore neexistuje zabudovaný nástroj, ktorý by vedel pretransformovať 3D súradnice jednotlivých kĺbov na 2D súradnice vzhľadom ku kamere. Preto sme museli vymyslieť spôsob, ako tieto súradnice kĺbov získať. Na základe ohnutia jednotlivých kĺbov sme zistili súradnice kĺbov vzhľadom k robotovi a pretransformovali ich na súradnicový systém „sveta“.

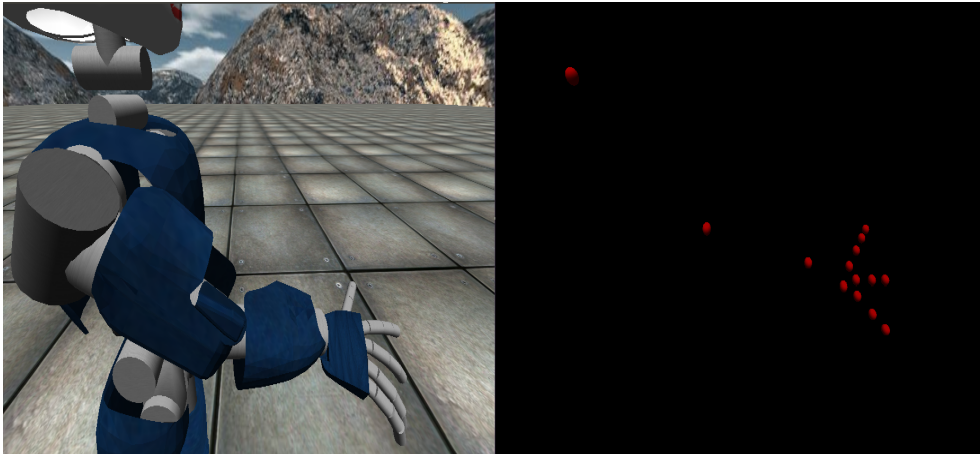
Na tento účel sme opäť využili knižnicu iKin, ktorá nám na základe nastavenia kĺbov vedela poskytnúť súradnice kĺbov. Pôvodne vedela knižnica pracovať len s takzvaným iCubArm. Ten vedel reprezentovať kĺby iba od ramena po dlaň. Bolo preto nutné využiť rozšírenie iCubFinger, ktoré predstavuje jeden prst. Avšak existovali len tri možné inštancie prstu a to thumb (palec), index (ukazovák) a middle (prostredník) a pre zvyšné dva prsty nebolo možné iCubFinger použiť. Preto sme sa rozhodli, že vizuálna reprezentácia nebude obsahovať informácie o zvyšných dvoch prstoch. Keďže v motorickej sekvencii sú reprezentované iba jednou hodnotou a pri žiadnom úchope nezohrávajú dôležitú úlohu, môžeme v našom modeli pracovať iba s tromi prstami bez toho, aby to narušilo cieľ práce.

Na to, aby sme ku každému obrazu (frame) získali adekvátne súradnice kĺbov sme použili už vytvorené motorické sekvencie. Využili sme pritom takzvané „chains“ obsiahnuté v iKin. Tie vytvárajú reťazce kĺbov tak, ako idú za sebou v robotickom ramene. Pre rameno a prsty sú „chains“ rôzne a aby správne fungovali bolo potrebné ich správne pospájať. Potom sme im postupne nastavovali hodnoty motorických sekvencií a získali sme 3D súradnice kĺbov vzhľadom k robotovi. Podobne ako pri vytváraní úchopov, aj teraz sme ich pretransformovali na súradnice „sveta“.

Keď už sme mali tieto súradnice, mohli sme na ich miestach vytvárať statické objekty. Keďže simulátor neumožňuje vytvoriť objekt na mieste, kde sa nachádza nejaká časť robota bez toho, aby ním pohl, vytvárali sme tieto objekty pred ním. Zároveň sme vytvorili čierne pozadie, aby sa nám ľahšie extrahovali potrebné súradnice.

V tomto kroku sme teda mali červené guľičky na miestach, kde boli kĺby (len o pol metra posunuté) na čiernom pozadí. Následne sme spravili screenshot obrazu a vytvorili sme program, ktorý z nich vedel extrahovať potrebné

súradnice. Porovnanie reálneho úchopu a reprezentácie jednotlivých kĺbov môžete vidieť na obrázku 4.7.



Obr. 4.7: Vizuálna reprezentácia kĺbov ramena.

V praxi celé vytváranie vizuálnej sekvencie vyzeralo nasledovne.

- Pre daný frame sme získali motorické nastavenie kĺbov.
- Pomocou iKin sme získali 3D súradnice kĺbov vzhľadom k robotovi.
- Súradnice sme pretransformovali na súradnice sveta.
- Pol metra pred ním sme vytvorili čierne pozadie.
- Na súradniciach, kde sa vyskytovali kĺby, sme vytvorili červené guľičky.
- Spravili sme screenshot a následne sme z obrazu vyextrahovali súradnice stredov guľičiek.

Toto sme zopakovali pre všetky 4 perspektívy. Keďže sme potrebovali 16 obrazov pre každý z 30 úchopov zo 4 perspektív, tak sme vytvorili scripty, ktoré tento proces automatizujú. Nájsť ich možno v priečinku TeachersGrasp

priloženom k zdrojovým kódom. Na získanie súradníc sme vytvorili program v C#, ktorý z vygenerovaných guľičiek extrahoval súradnice ich stredov.

4.3 Implementácia modelu

Po vytvorení databázy úchopov sme sa mohli pustiť do implementácie modelu. V tejto podkapitole si prejdeme finálnu verziu programu. Program je napísaný v jazyku Java a jeho základnými časťami sú moduly F5 a STS. Hlavná trieda je trieda Som, ktorá vytvára triedy F5.java a STS.java a spúšťa tréning, testovanie, vyhodnocovanie a vykresľovanie máp.

4.3.1 F5 a STS

Ako sme spomenuli, vyššie oblasti F5 a STS implementujeme pomocou MSOM. F5 aj STS majú rovnaké konštruktory pozostávajúce z 9 parametrov:

$$F_5(\alpha, \beta, \gamma_{\text{Start}}, \gamma_{\text{End}}, \lambda_{\text{Start}}, \lambda_{\text{End}}, \mathcal{N}_x, \mathcal{N}_y, Epochs) \quad (4.1)$$

Parametre α a β sú analogické k parametrom MSOM, ktorej fungovanie sme si popísali vyššie. γ_{Start} je rýchlosť učenia, ktorú má sieť na začiatku tréningu, pričom monotónne klesá až na γ_{End} . Veľkosť okolia λ_{Start} , podobne ako predchádzajúci parameter, postupne klesá na λ_{End} . Ich aktuálne hodnoty sa vypočítavajú pri tréningu v závislosti od aktuálnej epochy. \mathcal{N}_x a \mathcal{N}_y sú veľkosti máp a $Epochs$ je celkový počet epoch.

Po zavolaní konštruktora sa ako prvé vytvorí sieť s požadovanými parametrami. Neurón je reprezentovaný triedou Neuron.java pričom má dva vektory váh, vstupný a kontextový. Následne sa načíta databáza. Jeden obraz je reprezentovaný dvoma vektormi. Prvý je vstup pre F5, čo sú moto-

rické hodnoty, teda uhly kĺbov reprezentované 16 hodnotami a druhý pre STS, to sú zase 2D súradnice 14 bodov, čiže 28 súradníc. Jednotlivé hodnoty sú najprv normalizované na interval $[-1;1]$. Po predspracovaní dát nasleduje učenie siete. Implementovali sme všetky potrebné vzorce z kapitoly 3, pričom rovnicu 3.2, ktorú používame na zistenie vzdialeností vstupu od neurónov sme implementovali bez obmeny nasledovne:

$$d_i(t) = (1 - \alpha) \cdot \|\mathbf{x}(t) - \mathbf{w}_i^{\text{inp}}(t)\|^2 + \alpha \cdot \|\mathbf{q}(t) - \mathbf{w}_i^{\text{ctx}}(t)\|^2 \quad (4.2)$$

Po naučení metóda `activations()` nám vytvorí výstup siete tak, že celú databázu ešte raz vloží na vstup siete, pričom si ukladá jednotlivé aktívacie neurónov. Následne nám trieda `Som.java` pomocou triedy `GraphicPanel.java` zobrazí grafický výstup siete. Motorickú sieť vizualizujeme pomocou jedného zobrazovacieho panelu, na ktorom vidíme zobrazenú selektivitu neurónov vzhľadom k typu úchopu po natrénovaní. Senzorickú sieť zase vizualizujeme pomocou dvoch panelov, na jednom vzhľadom k typu úchopu a druhým vzhľadom k perspektíve. Bližšie si ukážeme v kapitole 5. V triede `F5.java` a `STS.java` je viacero pomocných metód, napríklad na ukladanie a zobrazovanie rôznych medzivýsledkov, ale tie tu nebudeme opisovať.

4.3.2 Implementácia zrkadliacich neurónov

Po naučení F5 a STS môžeme pristúpiť k učeniu zrkadliacich prepojení medzi nimi. Ako prvé sa vytvorí nové prepojenia medzi F5 a STS tak, že každému neurónu z F5 vytvoríme dva nové vektory váh, ktoré budú reprezentovať tieto spojenia. Toto prepojenie bude analogické ako u MSOM tak, ako v prvých dvoch prípadoch. Prvý, novovytvorený vektor bude mať veľkosť STS mapy, ktorá mu bude vytvárať vstup. Rovnako veľký bude aj kontextový

vektor, ktorý bude fungovať rovnako ako v predchádzajúcich prípadoch na ukladanie kontextu. To znamená, že v tomto bode majú neuróny v F5 štyri váhové vektory. Vstupný váhový vektor, ktorý porovnávame s motorickým vstupným vektorom a k nemu prislúchajúci kontextový vektor a zároveň zrkadliaci váhový vektor, ktorý porovnávame s výstupom STS a k nemu prislúchajúci kontextový vektor.

Zrkadliace prepojenia sa učia podobne ako doteraz. Rozdiel je v tom, že teraz nemáme len jeden vstupný vektor, ale dva. Jeden motorický a druhý senzorický. Pri tréningu zrkadliaceho systému už neprekladáme motorické vstupy, ale len vizuálne. Namiesto toho využívame už natrénovanú F5 a jej aktivácie si pre každý jeden úchop zapamätávame. Motorický vektor je teda vektor vzdialeností neurónov z F5 (d_i) od daného vstupu, ktorý sme si zapamätali pri zisťovaní aktivácií. Už natrénovanej STS predkladáme jednotlivé vizuálne vstupné dáta a jej výstup požívame ako vstup pre neuróny v F5. Takže rovnica, ktorá určuje vzdialenosť neurónov od takéhoto vstupu vyzerá nasledovne:

$$d_{\text{vis}}(t) = (1 - \alpha) \cdot \|\mathbf{s}(t) - \mathbf{w}_i^{\text{inp}}(t)\|^2 + \alpha \cdot \|\mathbf{q}(t) - \mathbf{w}_i^{\text{ctx}}(t)\|^2 \quad (4.3)$$

$$d_i(t) = \sigma \cdot d_{\text{mot}} + (1 - \sigma) \cdot d_{\text{vis}} \quad (4.4)$$

d_{mot} je zapamätaná vzdialenosť daného neurónu od motorického vstupu vypočítaný rovnicou 4.2 v predchádzajúcej časti a d_{vis} je vzdialenosť zrkadliacich vektorov \mathbf{w}_i od výstupu \mathbf{s} z STS. Treba doplniť, že jeden typ úchopu má len jedno motorické vyjadrenie, ale má až 4 rôzne vizuálne vzhľadom na rôzne perspektívy. Parameter σ určuje pomer, do akej miery má dôležitosť

zapamätaná motorická informácia a do akej vizuálna. Tento nami vytvorený parameter značne ovplyvňuje tréovanie a správanie siete, ako si povieme v kapitole 5. Týmto prístupom, akým tréujeme sieť, spájame vizuálne informácie s motorickými. Keďže spájame výstupy zo senzorickej časti so zapamätanou motorickou informáciou, ako keby sme mapovali videný obraz na náš motorický aparát, čo je zmyslom zrkadliacich neurónov.

Podobne ako v predchádzajúcej časti, aj tu po naučení získame výstup siete pomocou metódy `activationsMirror()`. Na rozdiel od učenia, kde bolo potrebné použiť aj motorickú informáciu, pri zisťovaní aktivácií zrkadliacich neurónov používame iba vizuálny vstup, čo znamená, že porovnáваме len výstup STS s vektorom zrkadliacich váh. Po natréovaní a zobrazení zrkadliacich neurónov ešte trieda `Som.java` vypočíta rôzne štatistické údaje o sieti, ktoré nám ju pomôžu ohodnotiť.

Kapitola 5

Experimenty

V tejto časti vyhodnotíme experimenty a porovnáme nami dosiahnuté výsledky s reálne nameranými hodnotami spomenutými v kapitole 2. Úspešnosť sietí vyjadríme pomocou niekoľkých, nami zvolených mier a ukážeme, ako voľba jednotlivých parametrov ovplyvní sieť. Ako prvé sa budeme zaoberať oblasťami F5 a STS a následne vytvorením variantných a invariantných zrkadliacich neurónov.

5.1 F5 a STS

Ako prvé sme natrénovali oblasti F5 a STS pomocou motorických a senzoryckých vstupov tak, ako sme spomenuli v implementácií. Zaujímali nás predovšetkým parametre α a β a ich vplyv na organizáciu siete. Parametre, ako rýchlosť učenia γ , veľkosť okolia λ alebo veľkosť máp, sme určili na začiatku a po pár pokusoch sa ukázali ako vhodné a ich obmeny nemali veľký vplyv na organizáciu siete. Preto počas experimentov používame tie isté hodnoty pre parametre:

- F5: $\gamma = [1;0.001]$, $\lambda = [3;1]$, veľkosť mapy 12×12

- STS: $\gamma = [1;0.001]$, $\lambda = [4;1]$, veľkosť mapy 16×16

Následne sme opakovali tréning map pomocou rôznych α a β parametrov. Úspešnosť map sme vyhodnocovali pomocou troch mier:

- Kvantizačná chyba, označujeme ju ako q_{err} . Vypočítavame ju ako $q_{err} = \sum d(t)_{b(t)}$. Táto miera vyjadruje súčet vzdialeností váhových vektorov vífazov od vstupných vektorov. Táto miera nám hovorí ako „presná“ bola sieť pri vyhodnocovaní.
- Distribúcia vífazov, $Dist = \frac{|b_i|}{N \times N}$. Táto miera vyjadruje počet neurónov, ktoré sa stali aspoň raz vífazmi, predelený celkovým počtom neurónov v sieti.
- Entropia, $Entropy = \sum_{i=1}^n \frac{w_i}{|X|} \cdot \log(\frac{w_i}{|X|})$. n je počet neurónov v mape a w_i je počet vífazstiev neurónu i na množine X . Entropia je miera neusporiadanosti v mape.

Po natréningu sietí s rôznymi hodnotami parametrov α a β sme vyrobili tabuľku hodnôt našich troch mier:

| α | β | q_{err} | Dist. | Entropy |
|----------|---------|-----------|-------|---------|
| 0.1 | 0.1 | 432.16 | 0.7 | 3.2 |
| 0.1 | 0.2 | 436.82 | 0.67 | 3.17 |
| 0.1 | 0.3 | 433.91 | 0.67 | 3.11 |
| 0.1 | 0.4 | 414.7 | 0.66 | 3.16 |
| 0.1 | 0.5 | 387.52 | 0.69 | 3.1 |
| 0.1 | 0.6 | 361.54 | 0.63 | 3.06 |
| 0.1 | 0.7 | 329.42 | 0.58 | 3.02 |
| 0.1 | 0.8 | 280.83 | 0.33 | 2.48 |
| 0.1 | 0.9 | 217.02 | 0.3 | 2.41 |
| 0.2 | 0.1 | 427.51 | 0.69 | 3.19 |
| 0.2 | 0.2 | 438.56 | 0.64 | 3.1 |
| 0.2 | 0.3 | 421.56 | 0.67 | 3.15 |
| 0.2 | 0.4 | 402.66 | 0.71 | 3.22 |
| 0.2 | 0.5 | 401.2 | 0.69 | 3.19 |
| 0.2 | 0.6 | 368.55 | 0.65 | 3.07 |
| 0.2 | 0.7 | 366.68 | 0.51 | 2.9 |
| 0.2 | 0.8 | 312.16 | 0.56 | 2.97 |
| 0.2 | 0.9 | 217.09 | 0.39 | 2.61 |
| 0.3 | 0.1 | 453.39 | 0.65 | 3.18 |
| 0.3 | 0.2 | 444.1 | 0.65 | 3.14 |
| 0.3 | 0.3 | 421.4 | 0.63 | 3.11 |
| 0.3 | 0.4 | 412.78 | 0.69 | 3.16 |
| 0.3 | 0.5 | 399.86 | 0.65 | 3.14 |
| 0.3 | 0.6 | 363.95 | 0.69 | 3.19 |
| 0.3 | 0.7 | 355.77 | 0.55 | 2.93 |
| 0.3 | 0.8 | 273.19 | 0.58 | 3.01 |
| 0.3 | 0.9 | 235.45 | 0.17 | 1.76 |
| 0.4 | 0.1 | 436.93 | 0.67 | 3.19 |
| 0.4 | 0.2 | 444.36 | 0.66 | 3.14 |
| 0.4 | 0.3 | 421.69 | 0.67 | 3.16 |
| 0.4 | 0.4 | 410.09 | 0.68 | 3.13 |
| 0.4 | 0.5 | 390.71 | 0.73 | 3.22 |
| 0.4 | 0.6 | 361.57 | 0.72 | 3.19 |
| 0.4 | 0.7 | 339.48 | 0.66 | 3.13 |
| 0.4 | 0.8 | 277.38 | 0.56 | 2.94 |
| 0.4 | 0.9 | 196.3 | 0.33 | 2.45 |
| 0.5 | 0.1 | 443.77 | 0.62 | 3.07 |
| 0.5 | 0.2 | 429.55 | 0.69 | 3.18 |
| 0.5 | 0.3 | 424.34 | 0.75 | 3.25 |
| 0.5 | 0.4 | 401.27 | 0.69 | 3.21 |
| 0.5 | 0.5 | 379.58 | 0.74 | 3.23 |

| α | β | q_{err} | Dist. | Entropy |
|----------|---------|-----------|-------|---------|
| 0.5 | 0.6 | 345.36 | 0.74 | 3.25 |
| 0.5 | 0.7 | 321.84 | 0.63 | 3.14 |
| 0.5 | 0.8 | 264.46 | 0.63 | 3.11 |
| 0.5 | 0.9 | 182.33 | 0.44 | 2.84 |
| 0.6 | 0.1 | 437.7 | 0.65 | 3.14 |
| 0.6 | 0.2 | 435.22 | 0.74 | 3.23 |
| 0.6 | 0.3 | 413.54 | 0.68 | 3.18 |
| 0.6 | 0.4 | 411.39 | 0.71 | 3.18 |
| 0.6 | 0.5 | 376.34 | 0.74 | 3.23 |
| 0.6 | 0.6 | 344.06 | 0.76 | 3.25 |
| 0.6 | 0.7 | 296.75 | 0.75 | 3.24 |
| 0.6 | 0.8 | 285.8 | 0.64 | 3.08 |
| 0.6 | 0.9 | 199.12 | 0.48 | 2.89 |
| 0.7 | 0.1 | 459.78 | 0.63 | 3.1 |
| 0.7 | 0.2 | 434.17 | 0.67 | 3.17 |
| 0.7 | 0.3 | 423.74 | 0.7 | 3.22 |
| 0.7 | 0.4 | 400.31 | 0.7 | 3.22 |
| 0.7 | 0.5 | 374.32 | 0.78 | 3.27 |
| 0.7 | 0.6 | 334.33 | 0.83 | 3.36 |
| 0.7 | 0.7 | 305.21 | 0.83 | 3.35 |
| 0.7 | 0.8 | 267.58 | 0.65 | 3.15 |
| 0.7 | 0.9 | 195.88 | 0.59 | 3.05 |
| 0.8 | 0.1 | 433.98 | 0.67 | 3.19 |
| 0.8 | 0.2 | 414.31 | 0.66 | 3.15 |
| 0.8 | 0.3 | 415.35 | 0.69 | 3.19 |
| 0.8 | 0.4 | 383.32 | 0.7 | 3.23 |
| 0.8 | 0.5 | 363.82 | 0.74 | 3.24 |
| 0.8 | 0.6 | 321.73 | 0.81 | 3.35 |
| 0.8 | 0.7 | 297.46 | 0.79 | 3.26 |
| 0.8 | 0.8 | 228.16 | 0.69 | 3.2 |
| 0.8 | 0.9 | 160.33 | 0.47 | 2.86 |
| 0.9 | 0.1 | 420.26 | 0.63 | 3.13 |
| 0.9 | 0.2 | 403.39 | 0.65 | 3.14 |
| 0.9 | 0.3 | 363.79 | 0.7 | 3.22 |
| 0.9 | 0.4 | 355.32 | 0.69 | 3.19 |
| 0.9 | 0.5 | 322.34 | 0.72 | 3.26 |
| 0.9 | 0.6 | 296.84 | 0.72 | 3.2 |
| 0.9 | 0.7 | 246.47 | 0.78 | 3.34 |
| 0.9 | 0.8 | 192.78 | 0.82 | 3.39 |
| 0.9 | 0.9 | 140.78 | 0.72 | 3.26 |

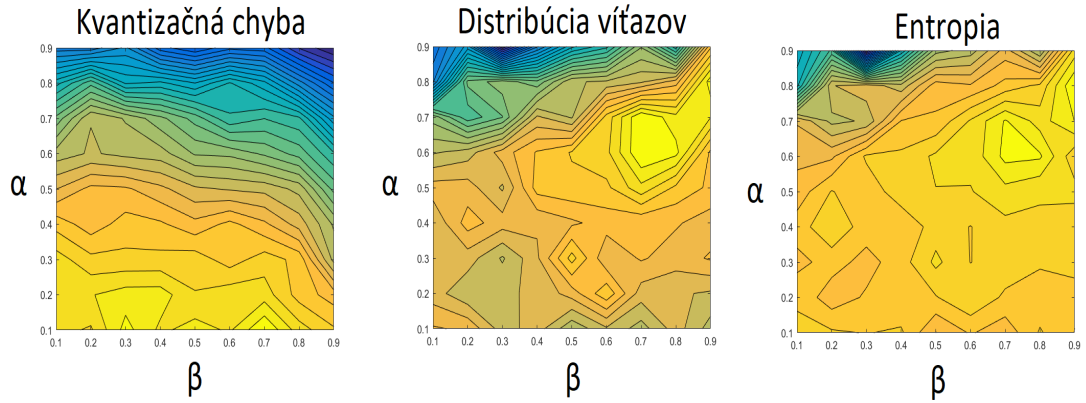
Obr. 5.1: Tabuľka zobrazujúca hodnoty jednotlivých hodnotiacich mier siete F5 vzhľadom k výberu parametrov α a β .

| α | β | q_{err} | Dist. | Entropy |
|----------|---------|-----------|-------|---------|
| 0.1 | 0.1 | 781.8 | 0.76 | 3.78 |
| 0.1 | 0.2 | 781.01 | 0.77 | 3.8 |
| 0.1 | 0.3 | 754.53 | 0.74 | 3.76 |
| 0.1 | 0.4 | 727.3 | 0.76 | 3.75 |
| 0.1 | 0.5 | 704.18 | 0.73 | 3.73 |
| 0.1 | 0.6 | 678.96 | 0.68 | 3.63 |
| 0.1 | 0.7 | 631.48 | 0.61 | 3.37 |
| 0.1 | 0.8 | 566.56 | 0.51 | 2.98 |
| 0.1 | 0.9 | 440.57 | 0.29 | 2.44 |
| 0.2 | 0.1 | 784.31 | 0.79 | 3.81 |
| 0.2 | 0.2 | 770.48 | 0.79 | 3.8 |
| 0.2 | 0.3 | 743.67 | 0.76 | 3.77 |
| 0.2 | 0.4 | 729.75 | 0.77 | 3.79 |
| 0.2 | 0.5 | 708.04 | 0.73 | 3.72 |
| 0.2 | 0.6 | 677.86 | 0.67 | 3.56 |
| 0.2 | 0.7 | 629.67 | 0.62 | 3.43 |
| 0.2 | 0.8 | 566.47 | 0.43 | 2.89 |
| 0.2 | 0.9 | 397.66 | 0.35 | 2.64 |
| 0.3 | 0.1 | 781.18 | 0.76 | 3.77 |
| 0.3 | 0.2 | 773.29 | 0.76 | 3.78 |
| 0.3 | 0.3 | 736.47 | 0.78 | 3.8 |
| 0.3 | 0.4 | 714.72 | 0.75 | 3.79 |
| 0.3 | 0.5 | 698.71 | 0.71 | 3.71 |
| 0.3 | 0.6 | 664.75 | 0.68 | 3.62 |
| 0.3 | 0.7 | 603.77 | 0.64 | 3.5 |
| 0.3 | 0.8 | 536.45 | 0.45 | 3.07 |
| 0.3 | 0.9 | 423.16 | 0.38 | 2.76 |
| 0.4 | 0.1 | 780.97 | 0.74 | 3.75 |
| 0.4 | 0.2 | 764.85 | 0.76 | 3.8 |
| 0.4 | 0.3 | 724.23 | 0.77 | 3.79 |
| 0.4 | 0.4 | 698.9 | 0.76 | 3.79 |
| 0.4 | 0.5 | 677.64 | 0.73 | 3.76 |
| 0.4 | 0.6 | 676.5 | 0.7 | 3.66 |
| 0.4 | 0.7 | 604.33 | 0.62 | 3.53 |
| 0.4 | 0.8 | 520.3 | 0.52 | 3.17 |
| 0.4 | 0.9 | 446.39 | 0.38 | 2.72 |
| 0.5 | 0.1 | 780.03 | 0.75 | 3.79 |
| 0.5 | 0.2 | 758.08 | 0.75 | 3.78 |
| 0.5 | 0.3 | 726.03 | 0.77 | 3.79 |
| 0.5 | 0.4 | 698.41 | 0.79 | 3.8 |
| 0.5 | 0.5 | 669.52 | 0.71 | 3.72 |

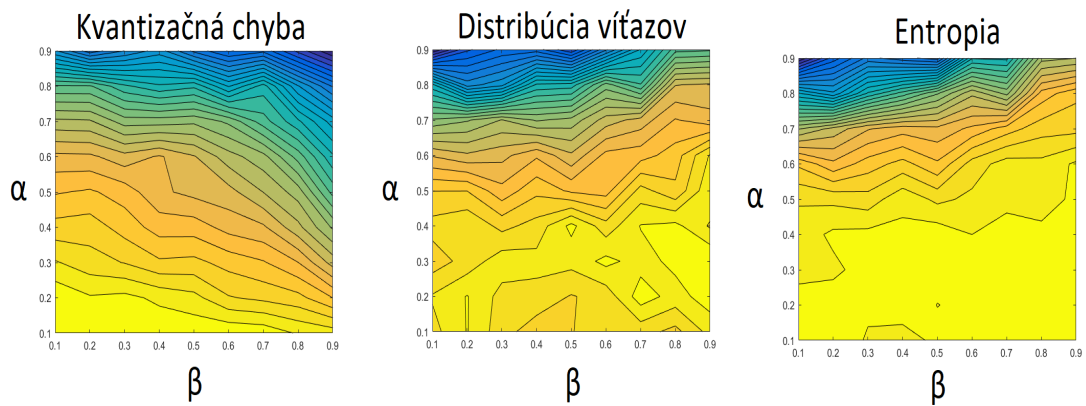
| α | β | q_{err} | Dist. | Entropy |
|----------|---------|-----------|-------|---------|
| 0.5 | 0.6 | 660.2 | 0.67 | 3.6 |
| 0.5 | 0.7 | 607.83 | 0.63 | 3.52 |
| 0.5 | 0.8 | 529.38 | 0.53 | 3.28 |
| 0.5 | 0.9 | 402.84 | 0.34 | 2.68 |
| 0.6 | 0.1 | 781.65 | 0.77 | 3.79 |
| 0.6 | 0.2 | 738.57 | 0.76 | 3.79 |
| 0.6 | 0.3 | 715.57 | 0.79 | 3.8 |
| 0.6 | 0.4 | 684.02 | 0.76 | 3.78 |
| 0.6 | 0.5 | 664.61 | 0.71 | 3.74 |
| 0.6 | 0.6 | 633.94 | 0.71 | 3.71 |
| 0.6 | 0.7 | 592.31 | 0.66 | 3.56 |
| 0.6 | 0.8 | 484.24 | 0.59 | 3.41 |
| 0.6 | 0.9 | 374.01 | 0.42 | 3.06 |
| 0.7 | 0.1 | 779.5 | 0.76 | 3.78 |
| 0.7 | 0.2 | 733.17 | 0.79 | 3.81 |
| 0.7 | 0.3 | 707.36 | 0.78 | 3.8 |
| 0.7 | 0.4 | 677.09 | 0.79 | 3.8 |
| 0.7 | 0.5 | 642.54 | 0.76 | 3.76 |
| 0.7 | 0.6 | 617.24 | 0.71 | 3.75 |
| 0.7 | 0.7 | 555.79 | 0.68 | 3.62 |
| 0.7 | 0.8 | 524.54 | 0.55 | 3.32 |
| 0.7 | 0.9 | 386.75 | 0.49 | 3.05 |
| 0.8 | 0.1 | 766.79 | 0.75 | 3.79 |
| 0.8 | 0.2 | 726.86 | 0.78 | 3.8 |
| 0.8 | 0.3 | 690.37 | 0.79 | 3.81 |
| 0.8 | 0.4 | 647.61 | 0.79 | 3.8 |
| 0.8 | 0.5 | 619.81 | 0.74 | 3.74 |
| 0.8 | 0.6 | 573.49 | 0.73 | 3.74 |
| 0.8 | 0.7 | 520.65 | 0.7 | 3.67 |
| 0.8 | 0.8 | 455.98 | 0.67 | 3.57 |
| 0.8 | 0.9 | 339.05 | 0.56 | 3.35 |
| 0.9 | 0.1 | 761.29 | 0.77 | 3.79 |
| 0.9 | 0.2 | 705.65 | 0.79 | 3.82 |
| 0.9 | 0.3 | 652 | 0.8 | 3.82 |
| 0.9 | 0.4 | 608.98 | 0.77 | 3.8 |
| 0.9 | 0.5 | 558.53 | 0.78 | 3.79 |
| 0.9 | 0.6 | 520.17 | 0.77 | 3.78 |
| 0.9 | 0.7 | 478.27 | 0.69 | 3.67 |
| 0.9 | 0.8 | 405.29 | 0.67 | 3.63 |
| 0.9 | 0.9 | 297.32 | 0.57 | 3.37 |

Obr. 5.2: Tabuľka zobrazujúca jednotlivé hodnotiace miery siete STS vzhľadom k výberu parametrov α a β .

Tabuľky sme sa rozhodli reprezentovať pomocou kontúrových grafov:



Obr. 5.3: Kontúrový graf zobrazujúci hodnoty jednotlivých mier siete F5.



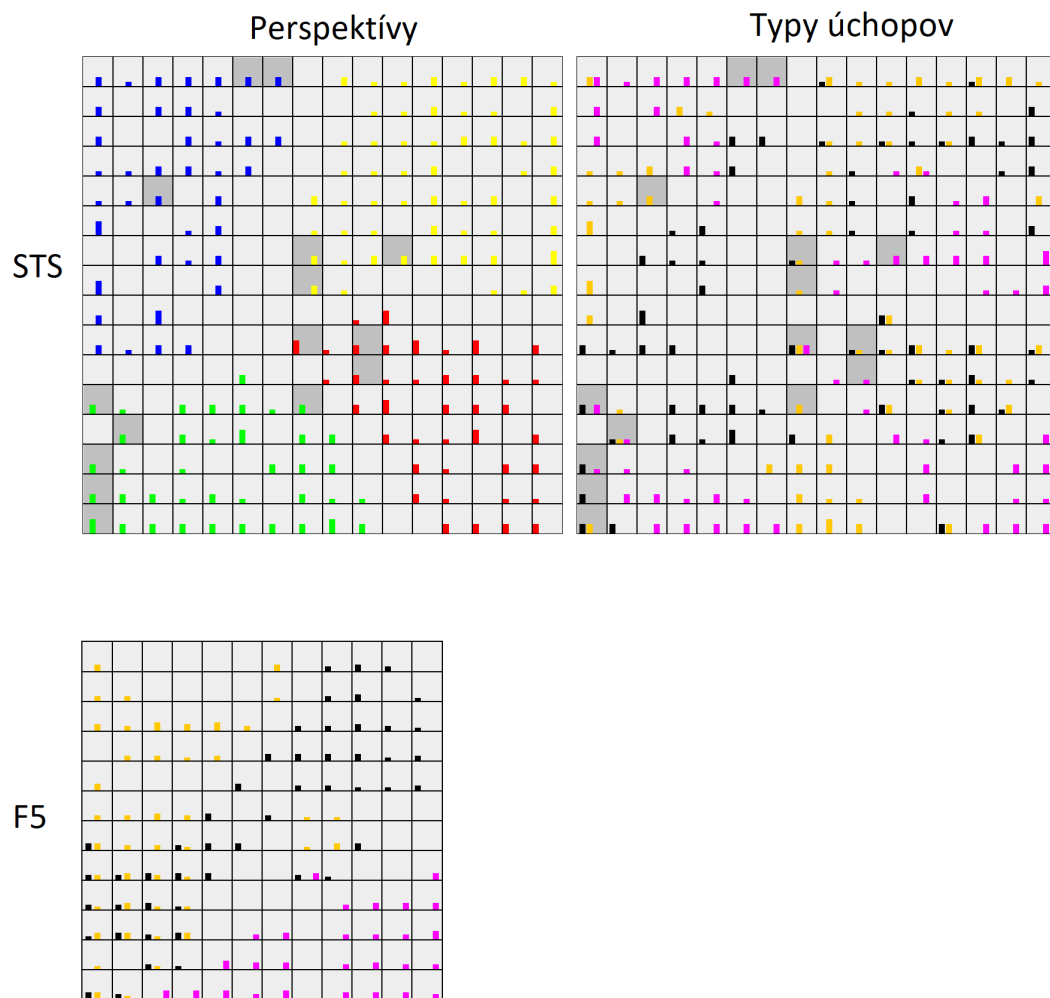
Obr. 5.4: Kontúrový graf zobrazujúci hodnoty jednotlivých mier siete STS.

V prípade kvantizačnej chyby považujeme mapy, ktoré mali túto mieru menšiu, za lepšie natrénované. Z kontúrových grafov vidíme, že kvantizačná chyba je závislá predovšetkým od parametru α . Čím väčšia α , tým sa nám chyba znižuje. To znamená, že čím väčší dôraz kladieme na kontextovú

zložku, tým lepšie sa sieť natrénuje. Naopak, pri distribúcií víťazov a entropií uprednostňujeme mapy, ktoré mali tieto dve miery väčšie. Ako môžeme vidieť na kontúrových grafoch, tieto dve miery úzko súvisia a majú podobný profil. Pre sieť F5 sú tieto miery najväčšie pre vysoké hodnoty α aj β . V prípade STS sme najlepšie výsledky pre dané dve miery dosiahli v prípade, že α nepresiahla hodnoty 0,4 a β bola vysoká, okolo 0,8. Na základe týchto výsledkov sme vybrali dve siete, ktoré obstáli najlepšie vzhľadom k všetkým trom parametrom. Pre F5 to bola sieť s $\alpha = 0,8$ a $\beta = 0,9$ a STS s $\alpha = 0,4$ a $\beta = 0,8$.

Tieto siete sme si následne nechali zobrazit', vid' obrázok 5.5. Každý neurón siete je reprezentovaný jedným štvorčekom. Jednotlivé stĺpce zobrazujú počet, koľko krát sa neurón stal víťazom pre daný typ alebo perspektívu úchopu. STS môže mať dve reprezentácie, zobrazenie aktivácií neurónov vzhľadom na typ úchopu a vzhľadom na perspektívu. Pre lepšie zobrazenie sme použili logaritmus, čo nám lepšie vykreslilo neuróny, ktoré sa stali víťazmi len párkrát. Víťazov sme hľadali po každom jednom frame. Farebná konvencia zobrazovania stĺpcov je nasledovná:

- čierna – Silový úchop (power grasp)
- oranžová – Jemný úchop (precision grasp)
- fialová – Bočný úchop (side grasp)
- červená – 0° (vlastná perspektíva)
- zelená – 90° (pohľad sprava)
- modrá – 180° (pohľad oproti)
- žltá – 270° (pohľad zľava)



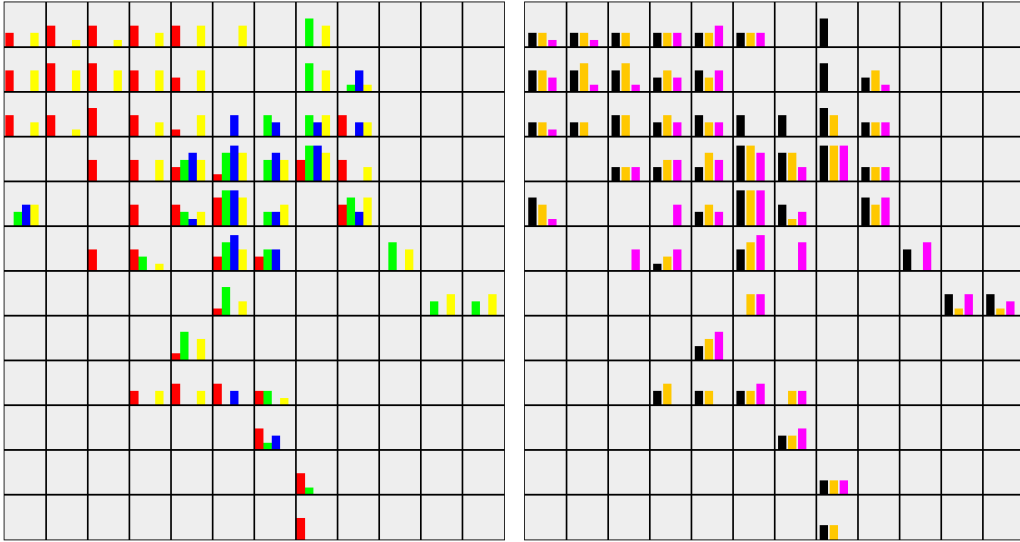
Obr. 5.5: Zobrazenie sietí F5 a STS. STS vľavo zobrazuje výstup vzhľadom na perspektívy a vpravo vzhľadom na typ úchopu. Motorický výstup zobrazujeme len vzhľadom na typu úchopov. Šedou sú zobrazené neuróny aktívne najmä na prvé 4 framy.

Ako môžeme vidieť, neuróny v F5, ktoré reagujú na bočný úchop, sú pekne oddelené od zvyšku a takmer vôbec sa neprekrývajú so zvyšnými dvoma typmi úchopov. Na rozdiel od toho máme v mape viacero neurónov, ktoré reagujú aj na jemný aj na silový úchop. Pri bližšom skúmaní sme zistili, že

Ľavá dolná štvrt' mapy reaguje predovšetkým na prvé framy úchopov. Keďže jemný a silový úchop majú veľmi podobné približovanie sa k objektu, sieť nebola schopná dostatočne dobre sa naučiť rozlišovať tieto prvé framy. Avšak pri uchopovaní je sieť už dostatočne dobre natrénovaná, aby ich vedela rozlíšiť. V STS sa nám pekne oddelili úchopy vzhľadom na perspektívy. Ani jeden neurón nereaguje na viac ako jednu perspektívu. V rámci perspektívy sa vyčlenili jednotlivé typy úchopov. Tu už môžeme vidieť, že sieť mala väčšie problémy rozlíšiť jednotlivé typy úchopov v rámci jednej perspektívy. Najhoršie na tom sú, podobne ako v F5, jemný a silový úchop. Táto organizácia vychádza z toho, že vizuálne dáta je pomerne ľahké rozdeliť vzhľadom na perspektívu, ale ťažké vzhľadom na typ úchopov. Najhoršie je spravidla rozlíšiť silový a jemný úchop z vlastnej perspektívy. Tento problém rozlíšiť tieto dva úchopy vychádza aj z toho, že sme museli v našej implementácii odignorovať malíček a prstenník, ktoré mohli poskytnúť viac informácií ako tieto dva úchopy rozlíšiť. V rámci STS sa nám už nevytvorila oblasť, ktorá podobne ako v F5 bola aktívna predovšetkým na fázu približovania. Na obrázku sme nechali na šedo vykresliť neuróny, ktorý boli aktívne najmä na prvé štyri framy a teda najmä na približovanie.

5.2 Zrkadliace neuróny

Keď už sme mali natrénované moduly F5 a STS, mohli sme pristúpiť k natrénovaniu zrkadliacich prepojení a k vzniku variantných a invariantných neurónov. Ešte pred zavedeným parametra σ sme sa pomocou rovnice 4.2 pokúšali natrénovať tieto prepojenia pomocou postupu, ako sme opísali v implementácii. Pre nami zvolené parametre $\alpha = 0,3$ a $\beta = 0,7$ sa natrénovali jednotlivé zrkadliace prepojenia ako môžeme vidieť na obrázku 5.6.

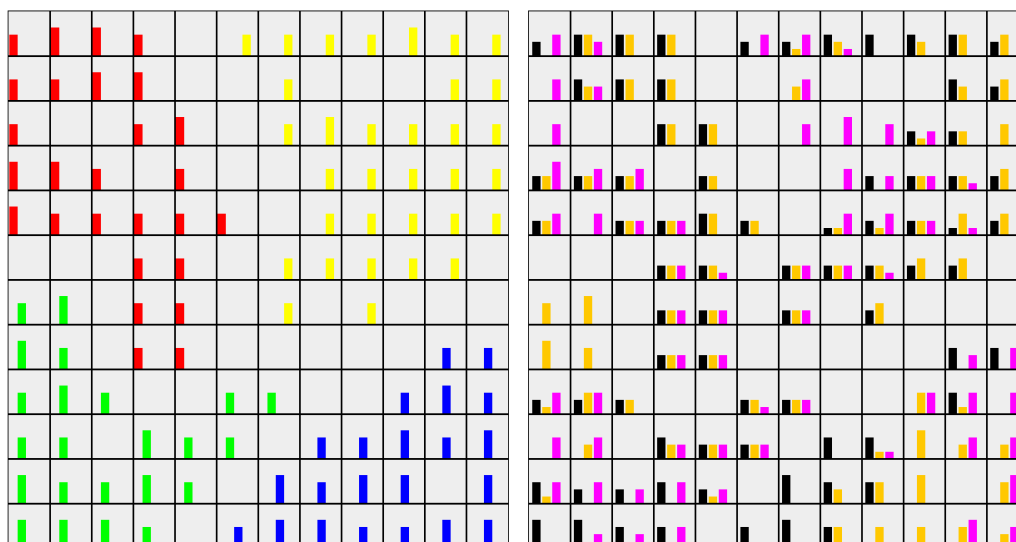


Obr. 5.6: Zobrazenie mapy F5 a jej zrkadliacich neurónov, vľavo vzhľadom na perspektívy, vpravo vzhľadom na typ úchopov.

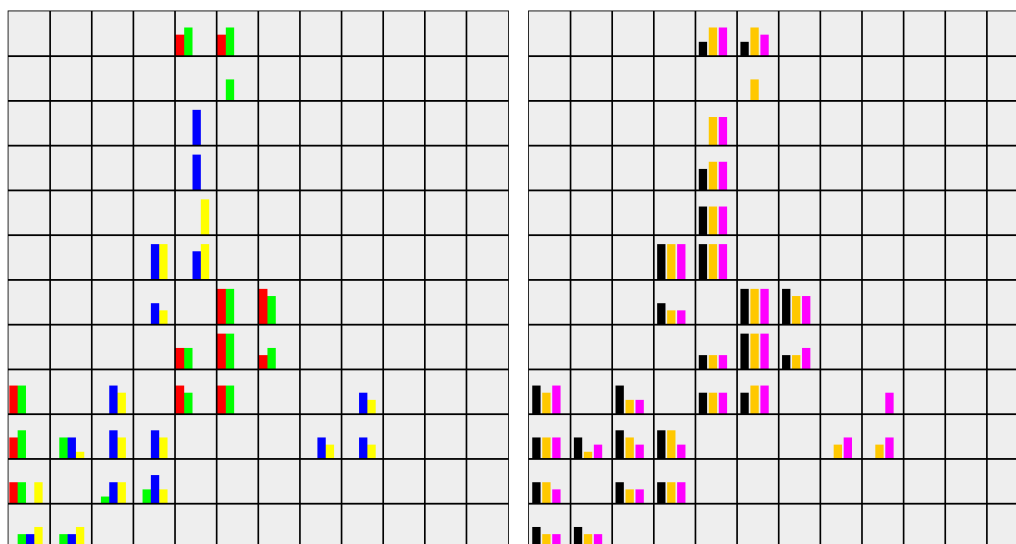
I keď sa nám podarilo vytvoriť niekoľko perspektívne invariantných neurónov, organizácia mapy nebola dobrá. Problém bol v tom, že okrem perspektív boli neuróny invariantné aj vzhľadom na úchop. To znamená, že sme mali neuróny, ktoré pálili takmer na všetko. Tie ale pre nás nemajú veľkú výpovednú hodnotu a nemohli sme pomocou nich rozlíšiť dané úchopy.

Tento problém nám čiastočne pomohol vyriešiť parameter σ , ktorý, ako bolo spomenuté v implementácií, určuje pomer dôležitosti motorických a vizuálnych informácií. V prípade, že sme určili $\sigma = 0.1$, tak sieť mala tendenciu rozlišovať úchopy najmä na základe vizuálnej informácie, kde prevládala najmä perspektíva. Ako môžeme vidieť na obrázku 5.7, zrkadliace neuróny sa vyprofilovali podobne ako tie v STS. Ak sme určili $\sigma = 0.9$, sieť sa viac snažila vyprofilovať ako F5, ale nedosahovala očakávanú organizáciu. Opäť nám vznikli neuróny čo boli aktívne skoro vždy a nevedeli sme rozlíšiť jednotlivé

úchopy, viď obrázok 5.8.

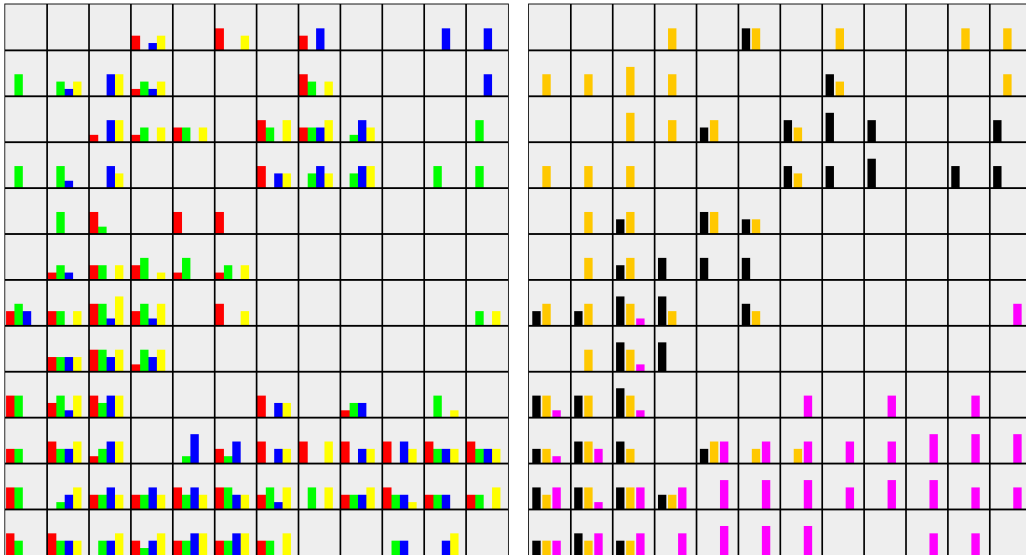


Obr. 5.7: Organizácia mapy F5 v prípade zvýšeného dôrazu na perspektívy.



Obr. 5.8: Organizácia mapy F5 v prípade zvýšeného dôrazu na typ úchopy.

Ako problém sa ukázal predovšetkým výstup z STS, ktorý sme počítali pomocou rovnice 3.8. Jednotlivé d_i boli až príliš podobné a ani pomocou spomínanej funkcie sme nevedeli docieľiť, aby sieť vedela lepšie tieto dáta klasifikovať. Preto sme pristúpili na riešenie pomocou One-hot. To znamená, že pre daný vizuálny vstup sme namiesto celej matice, ktorá reprezentovala d_i pre každý neurón mapy STS, posielali na vstup zrkadliacim neurónom maticu, kde na mieste víťaza pre daný vstup bola 1 a zvyšku matice sme dali 0. Pomocou tohto postupu a parametru $\sigma = 0.9$ sme dostali sieť zobrazenú na obrázku 5.9.



Obr. 5.9: Aktivácia zrkadliacich neurónov po zmene výstupu z STS pomocou metódy One-hot.

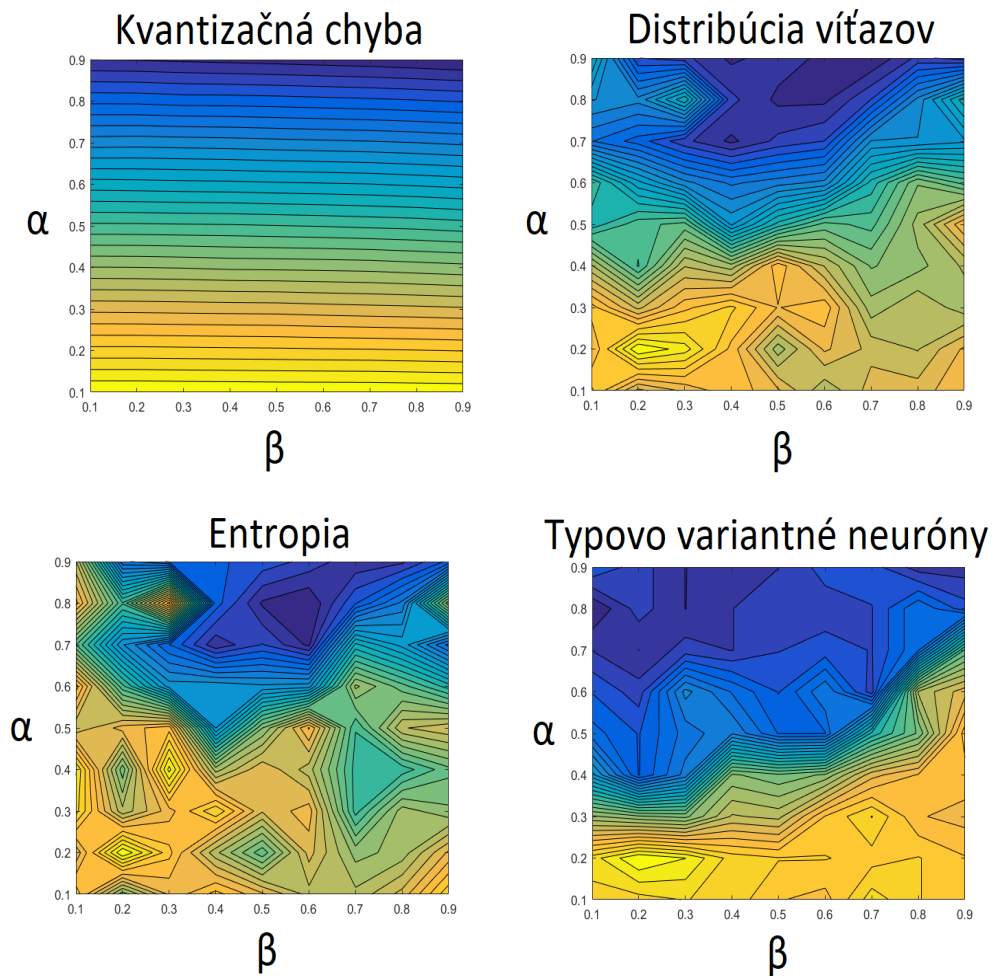
Mapa má už požadovanú organizáciu. Stále máme neuróny, ktoré sú in-variantné aj na typ aj na perspektívu úchopu, ale máme aj neuróny, ktoré sú selektívne pre jeden typ úchopu a viacero perspektív. Pomocou týchto neurónov už vieme rozlíšiť jeden typ úchopu pozorovaný z viacero perspektív.

Metódou One-hot sme dosiahli, že informácie z STS lepšie klasifikujú jednotlivé úchopy a pomocou parametra σ sme zase dosiahli, že sieť uprednostňuje motorické informácie, ktoré už má naučené.

Po tom, čo sme vytvorili vhodnú implementáciu nášho modelu, sme mohli pristúpiť podobne ako pri F5 a STS k hľadaniu vhodných parametrov α a β a k vyhodnocovanie mier jednotlivých máp. Pre vyhodnocovanie systému zrkadliacich neurónov sme si zaviedli ďalšiu mieru, označenú ako Type-var, ktorá nám určuje percento typovo variantných zrkadliacich neurónov v F5. To znamená počet neurónov, ktoré pália iba na jeden typ úchopu. Výsledky podobne ako predtým sme si nechali zobrazit' pomocou tabuľky a kontúrového grafu.

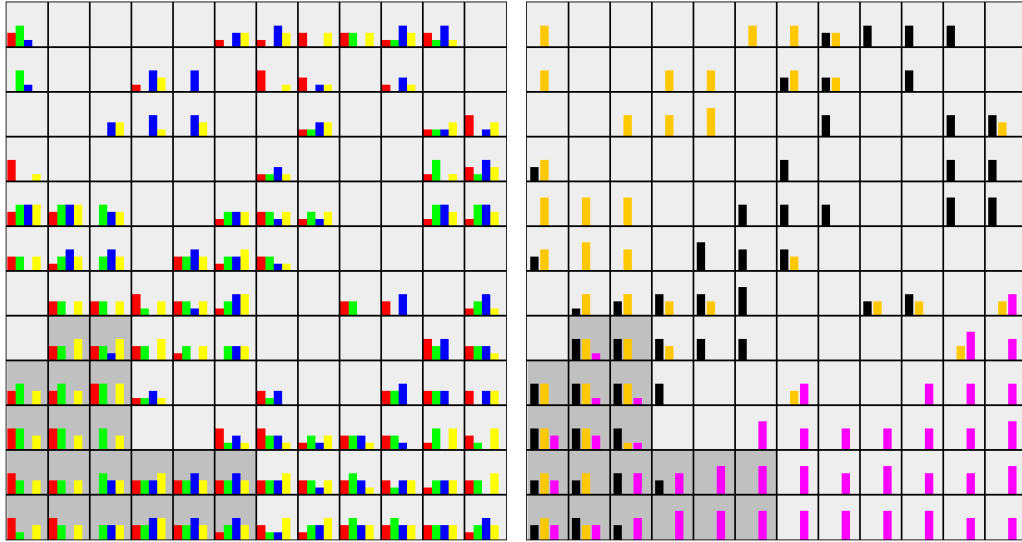
| α | β | q_{err} | Dist. | Entropy | Type-Var |
|----------|---------|-----------|-------|---------|----------|
| 0.1 | 0.1 | 1.68 | 0.63 | 1.69 | 39.58 |
| 0.1 | 0.2 | 1.5 | 0.61 | 1.66 | 40.97 |
| 0.1 | 0.3 | 1.33 | 0.63 | 1.71 | 29.17 |
| 0.1 | 0.4 | 1.16 | 0.58 | 1.72 | 15.28 |
| 0.1 | 0.5 | 0.97 | 0.51 | 1.65 | 8.33 |
| 0.1 | 0.6 | 0.8 | 0.53 | 1.69 | 6.25 |
| 0.1 | 0.7 | 0.61 | 0.44 | 1.6 | 3.47 |
| 0.1 | 0.8 | 0.43 | 0.47 | 1.68 | 0.69 |
| 0.1 | 0.9 | 0.25 | 0.49 | 1.65 | 6.25 |
| 0.2 | 0.1 | 1.68 | 0.59 | 1.62 | 39.58 |
| 0.2 | 0.2 | 1.5 | 0.67 | 1.74 | 44.44 |
| 0.2 | 0.3 | 1.33 | 0.59 | 1.64 | 27.78 |
| 0.2 | 0.4 | 1.15 | 0.53 | 1.61 | 6.25 |
| 0.2 | 0.5 | 0.97 | 0.53 | 1.67 | 6.25 |
| 0.2 | 0.6 | 0.79 | 0.49 | 1.59 | 4.17 |
| 0.2 | 0.7 | 0.61 | 0.42 | 1.52 | 2.08 |
| 0.2 | 0.8 | 0.43 | 0.46 | 1.59 | 4.17 |
| 0.2 | 0.9 | 0.25 | 0.41 | 1.49 | 4.17 |
| 0.3 | 0.1 | 1.67 | 0.6 | 1.67 | 40.28 |
| 0.3 | 0.2 | 1.5 | 0.65 | 1.69 | 43.06 |
| 0.3 | 0.3 | 1.33 | 0.62 | 1.67 | 27.08 |
| 0.3 | 0.4 | 1.15 | 0.58 | 1.73 | 11.11 |
| 0.3 | 0.5 | 0.97 | 0.54 | 1.68 | 9.72 |
| 0.3 | 0.6 | 0.79 | 0.46 | 1.53 | 12.5 |
| 0.3 | 0.7 | 0.61 | 0.4 | 1.48 | 4.17 |
| 0.3 | 0.8 | 0.42 | 0.5 | 1.69 | 2.08 |
| 0.3 | 0.9 | 0.24 | 0.4 | 1.48 | 2.08 |
| 0.4 | 0.1 | 1.67 | 0.62 | 1.69 | 39.58 |
| 0.4 | 0.2 | 1.5 | 0.62 | 1.64 | 39.58 |
| 0.4 | 0.3 | 1.32 | 0.63 | 1.71 | 31.94 |
| 0.4 | 0.4 | 1.14 | 0.57 | 1.63 | 27.78 |
| 0.4 | 0.5 | 0.97 | 0.47 | 1.53 | 11.81 |
| 0.4 | 0.6 | 0.79 | 0.44 | 1.53 | 9.03 |
| 0.4 | 0.7 | 0.6 | 0.39 | 1.44 | 3.47 |
| 0.4 | 0.8 | 0.42 | 0.4 | 1.48 | 3.47 |
| 0.4 | 0.9 | 0.24 | 0.39 | 1.47 | 2.78 |
| 0.5 | 0.1 | 1.67 | 0.6 | 1.67 | 37.5 |
| 0.5 | 0.2 | 1.5 | 0.57 | 1.6 | 38.89 |
| 0.5 | 0.3 | 1.32 | 0.61 | 1.66 | 31.25 |
| 0.5 | 0.4 | 1.14 | 0.62 | 1.67 | 26.39 |
| 0.5 | 0.5 | 0.97 | 0.51 | 1.62 | 7.64 |
| 0.5 | 0.6 | 0.78 | 0.45 | 1.52 | 6.25 |
| 0.5 | 0.7 | 0.6 | 0.4 | 1.46 | 4.86 |
| 0.5 | 0.8 | 0.42 | 0.39 | 1.44 | 4.86 |
| 0.5 | 0.9 | 0.23 | 0.4 | 1.47 | 3.47 |
| 0.6 | 0.1 | 1.67 | 0.58 | 1.64 | 36.81 |
| 0.6 | 0.2 | 1.49 | 0.6 | 1.66 | 38.89 |
| 0.6 | 0.3 | 1.32 | 0.62 | 1.67 | 37.5 |
| 0.6 | 0.4 | 1.14 | 0.58 | 1.64 | 27.78 |
| 0.6 | 0.5 | 0.96 | 0.54 | 1.68 | 7.64 |
| 0.6 | 0.6 | 0.78 | 0.46 | 1.53 | 10.42 |
| 0.6 | 0.7 | 0.59 | 0.41 | 1.44 | 5.56 |
| 0.6 | 0.8 | 0.41 | 0.39 | 1.43 | 4.17 |
| 0.6 | 0.9 | 0.23 | 0.39 | 1.46 | 5.56 |
| 0.7 | 0.1 | 1.67 | 0.6 | 1.64 | 40.97 |
| 0.7 | 0.2 | 1.49 | 0.59 | 1.62 | 38.19 |
| 0.7 | 0.3 | 1.31 | 0.57 | 1.58 | 40.28 |
| 0.7 | 0.4 | 1.14 | 0.56 | 1.58 | 34.03 |
| 0.7 | 0.5 | 0.95 | 0.53 | 1.59 | 20.83 |
| 0.7 | 0.6 | 0.77 | 0.51 | 1.63 | 4.17 |
| 0.7 | 0.7 | 0.59 | 0.47 | 1.56 | 4.86 |
| 0.7 | 0.8 | 0.41 | 0.41 | 1.47 | 4.86 |
| 0.7 | 0.9 | 0.22 | 0.38 | 1.44 | 5.56 |
| 0.8 | 0.1 | 1.67 | 0.6 | 1.66 | 38.89 |
| 0.8 | 0.2 | 1.49 | 0.58 | 1.63 | 38.89 |
| 0.8 | 0.3 | 1.31 | 0.58 | 1.63 | 36.11 |
| 0.8 | 0.4 | 1.13 | 0.56 | 1.58 | 36.11 |
| 0.8 | 0.5 | 0.95 | 0.57 | 1.65 | 26.39 |
| 0.8 | 0.6 | 0.77 | 0.56 | 1.6 | 30.56 |
| 0.8 | 0.7 | 0.58 | 0.47 | 1.54 | 9.03 |
| 0.8 | 0.8 | 0.4 | 0.46 | 1.5 | 9.03 |
| 0.8 | 0.9 | 0.21 | 0.4 | 1.45 | 2.78 |
| 0.9 | 0.1 | 1.67 | 0.61 | 1.66 | 37.5 |
| 0.9 | 0.2 | 1.48 | 0.6 | 1.63 | 37.5 |
| 0.9 | 0.3 | 1.3 | 0.57 | 1.63 | 35.42 |
| 0.9 | 0.4 | 1.12 | 0.58 | 1.58 | 36.81 |
| 0.9 | 0.5 | 0.94 | 0.61 | 1.65 | 38.19 |
| 0.9 | 0.6 | 0.76 | 0.53 | 1.6 | 34.03 |
| 0.9 | 0.7 | 0.57 | 0.47 | 1.54 | 25 |
| 0.9 | 0.8 | 0.39 | 0.51 | 1.5 | 5.56 |
| 0.9 | 0.9 | 0.2 | 0.4 | 1.45 | 2.78 |

Obr. 5.10: Tabuľka zobrazujúca jednotlivé hodnotiace miery siete vzhľadom k výberu parametrov α a β . Pre stĺpec kvantizačnej chyby platí, že hodnoty sme predelili 1000.



Obr. 5.11: Kontúrové grafy hodnotiacich mier.

Podobne, ako v predchádzajúcej časti, aj teraz sa kvantizačná chyba ukázala závislá najmä od parametra α a pomerne nezávislá od β . Zvyšné tri miery mali najlepšie výsledky pre parametre $\alpha = 0.2$ a $\beta = 0.2$ a preto sme túto sieť vybrali ako najlepšiu a nechali zobrazit', viď obrázok 5.12.



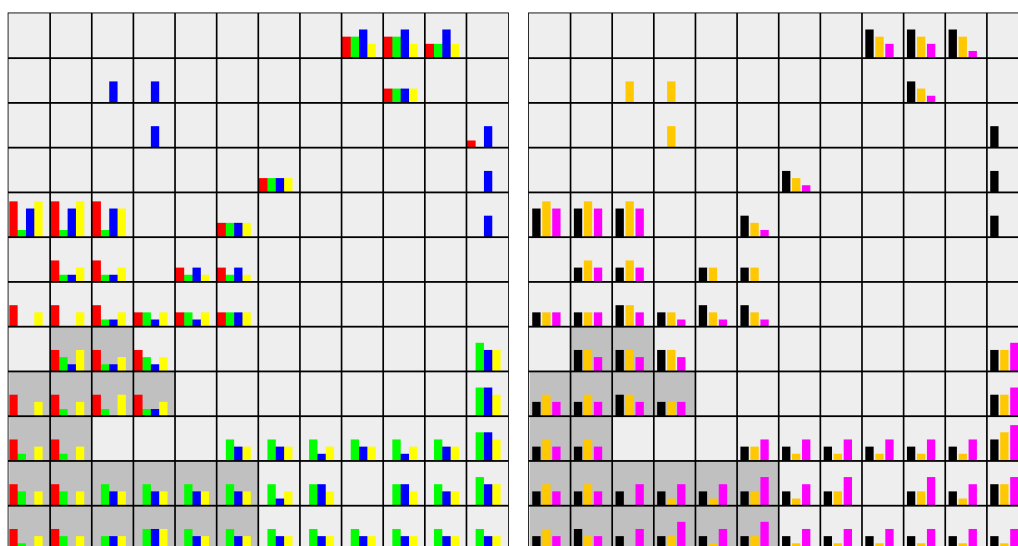
Obr. 5.12: Sieť zrkadliacich neurónov, ktorá dosiahla najlepšie výsledky.

Ako môžeme vidieť, sieť sa vyprofilovala podobne ako predchádzajúca, ale lepšie rozložila svoje zdroje a vieme pomocou nej lepšie klasifikovať jednotlivé úchopy. Keďže sieť má profil podobný ako F5 aj tu môžeme vidieť, že neuróny aktívne predovšetkým na približovanie, sa nachádzajú v ľavom dolnom rohu. Tieto neuróny sú invariantné aj na typ úchopu. To znamená, že sieť nevie na základe prvých frame-ov určiť o aký úchop ide. To je však očakávaná vlastnosť siete.

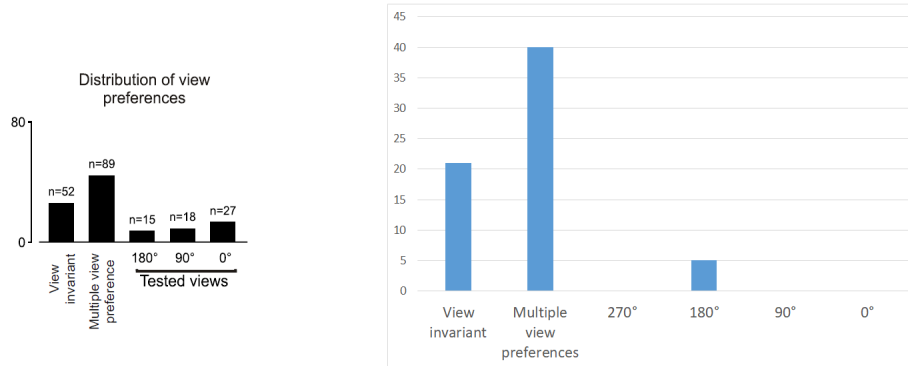
5.3 Porovnanie máp s biologickými experimentami

V našej práci sme sa nechali inšpirovať najmä prácou Caggiana a kolegov (Caggiano et al., 2011) a preto porovnáme nami dosiahnuté výsledky s ich nameranými. V grafe na obrázku 2.4 v kapitole 2 môžeme vidieť jednotlivé

počty pre invariantné a variantné neuróny. V spomínanej práci pracovali s tromi perspektívami, my sme však pridali aj pohľad zľava (270°). Mapa, ktorú sme vyhodnotili ako najlepšiu, sa však nepribližovala k výsledkom nameranými biológmi. Preto sme sa rozhodli z natrénovaných máp nájsť takú, čo najlepšie odzrkadľuje reálne hodnoty. Mapu s parametrami $\alpha = 0.8$ a $\beta = 0.2$ sme ohodnotili pomocou našich mier ako nie veľmi dobrú. Ako vidíme na obrázku 5.13, mapa má prevažne typovo invariantné neuróny avšak perspektívne sa najviac podobá nameraným výsledkom, vid' porovnanie na obrázku 5.14.

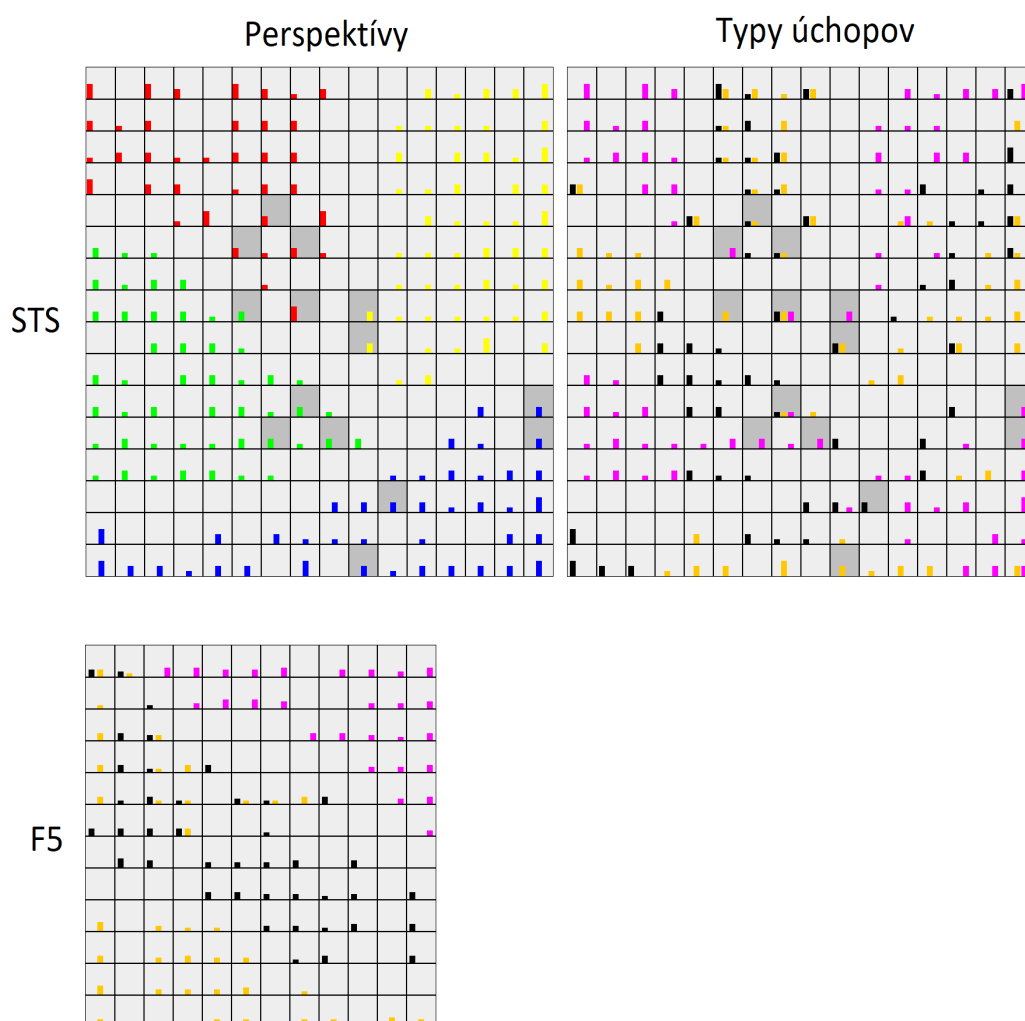


Obr. 5.13: Sieť, ktorá najlepšie napodobňuje reálne výsledky, avšak nevhodná na klasifikáciu typov úchopu.



Obr. 5.14: Porovnanie počtu variantných a invariantných neurónov s prácou Caggiano. Ako invariantné berieme všetky neuróny, ktoré pálili na všetky 4 perspektívy. Pod „Mutliple View Preferences“ rozumieme všetky neuróny, čo pálili na dve alebo tri rôzne perspektívy. Následne máme počty selektívnych neurónov pre jednotlivé perspektívy.

V ďalších experimentoch sme sa rozhodli preskúmať, aký vplyv má výber siete F5 a STS na tvorbu zrkadliacich neurónov. V predchádzajúcich experimentoch sme pracovali len s tými, čo uspeli v našom hodnotení najlepšie. Ale ako ukázalo porovnávanie výsledkov, nie vždy voľba tých najlepších parametrov odzrkadľuje biologicky namerané hodnoty. Preto sme sa rozhodli vybrať aj siete, ktoré nedopadli najlepšie a pomocou nich vybudovať systém zrkadliacich neurónov. Keďže tréning takýchto sietí je časovo náročné, nemohli sme si dovoliť nakombinovať všetky možné hodnoty parametrov a nájsť požadovanú sieť. Namiesto toho sme zvolili pár sietí, ktoré sme považovali za vhodné a tie skúmali. Najlepšie z toho vyšla sieť, ktorá mala takéto kombinácie parametrov, F5 $\alpha = 0.3$, $\beta = 0.7$ a pre STS $\alpha = 0.3$ a $\beta = 0.7$. Ich zobrazenie môžeme vidieť na obrázku 5.15.



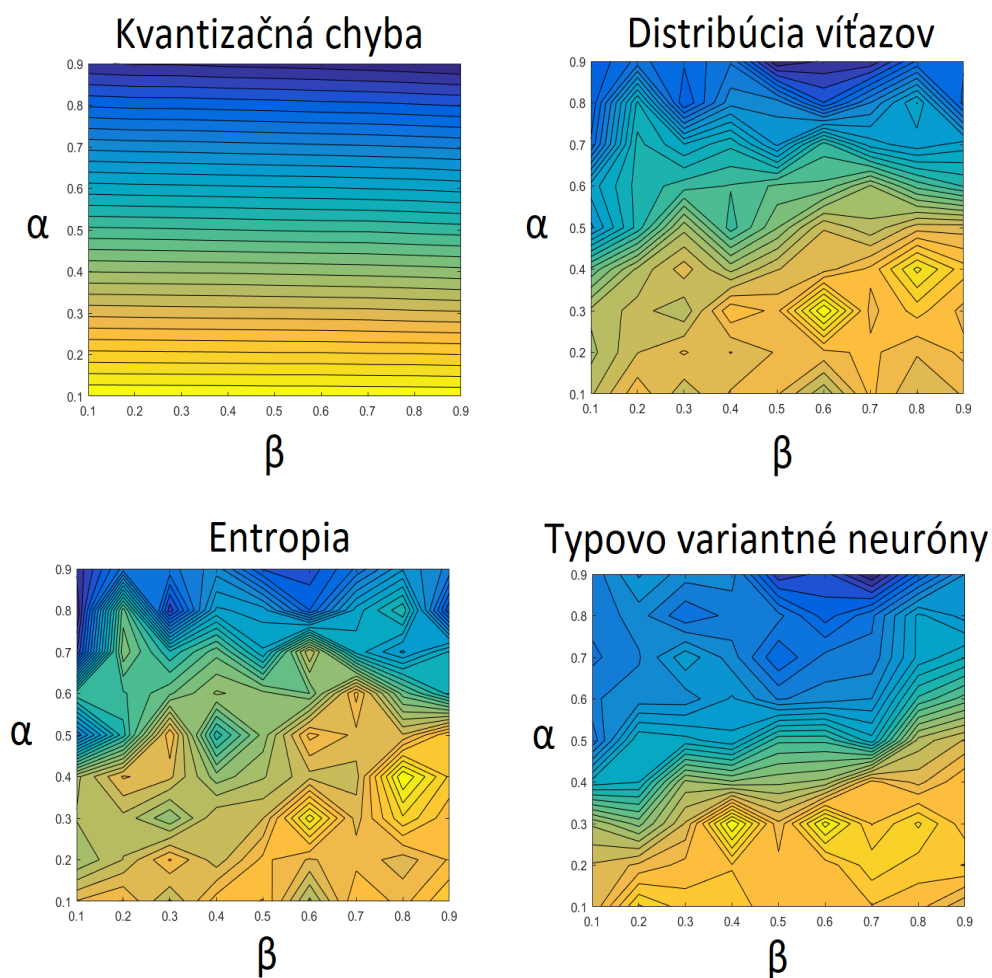
Obr. 5.15: Vybrané mapy F5 a STS, ktoré majú horšie výsledky, ale sú vhodnejšie na tvorbu zrkadliacich neurónov.

Pre tieto dve vybrané siete sme mohli spraviť rovnakú analýzu ako pri predchádzajúcom experimente. Ako dôležité sa ukázalo to, že v F5 je málo neurónov aktívnych pri jemnom a zároveň silovom úchope. Aj tentokrát sa vytvorila oblasť (vľavo hore), kde sa združovali neuróny aktívne najmä pri prvých frame-och. Opäť sme skúmali vplyv parametrov α a β na mapy a ich výsledky sme vyjadrili mierami a následne kontúrovými grafmi.

| α | β | q_{err} | Dist. | Entropy | Type-Var |
|----------|---------|-----------|-------|---------|----------|
| 0.1 | 0.1 | 1.68 | 0.6 | 1.66 | 42.36 |
| 0.1 | 0.2 | 1.5 | 0.58 | 1.6 | 42.36 |
| 0.1 | 0.3 | 1.33 | 0.59 | 1.62 | 39.58 |
| 0.1 | 0.4 | 1.15 | 0.54 | 1.61 | 29.17 |
| 0.1 | 0.5 | 0.98 | 0.45 | 1.43 | 21.53 |
| 0.1 | 0.6 | 0.8 | 0.47 | 1.54 | 23.61 |
| 0.1 | 0.7 | 0.62 | 0.4 | 1.35 | 22.22 |
| 0.1 | 0.8 | 0.43 | 0.38 | 1.33 | 22.92 |
| 0.1 | 0.9 | 0.26 | 0.38 | 1.34 | 22.92 |
| 0.2 | 0.1 | 1.68 | 0.63 | 1.67 | 48.61 |
| 0.2 | 0.2 | 1.5 | 0.61 | 1.63 | 43.75 |
| 0.2 | 0.3 | 1.33 | 0.6 | 1.63 | 36.81 |
| 0.2 | 0.4 | 1.15 | 0.59 | 1.66 | 29.86 |
| 0.2 | 0.5 | 0.97 | 0.51 | 1.53 | 29.86 |
| 0.2 | 0.6 | 0.79 | 0.51 | 1.54 | 24.31 |
| 0.2 | 0.7 | 0.61 | 0.51 | 1.58 | 23.61 |
| 0.2 | 0.8 | 0.43 | 0.49 | 1.57 | 24.31 |
| 0.2 | 0.9 | 0.25 | 0.41 | 1.41 | 26.39 |
| 0.3 | 0.1 | 1.67 | 0.6 | 1.63 | 46.53 |
| 0.3 | 0.2 | 1.5 | 0.63 | 1.67 | 45.14 |
| 0.3 | 0.3 | 1.32 | 0.59 | 1.59 | 42.36 |
| 0.3 | 0.4 | 1.15 | 0.62 | 1.65 | 38.89 |
| 0.3 | 0.5 | 0.97 | 0.58 | 1.67 | 29.17 |
| 0.3 | 0.6 | 0.79 | 0.52 | 1.57 | 24.31 |
| 0.3 | 0.7 | 0.61 | 0.48 | 1.52 | 27.08 |
| 0.3 | 0.8 | 0.43 | 0.38 | 1.36 | 21.53 |
| 0.3 | 0.9 | 0.25 | 0.4 | 1.42 | 24.31 |
| 0.4 | 0.1 | 1.67 | 0.63 | 1.68 | 46.53 |
| 0.4 | 0.2 | 1.5 | 0.61 | 1.64 | 45.83 |
| 0.4 | 0.3 | 1.32 | 0.65 | 1.64 | 52.08 |
| 0.4 | 0.4 | 1.15 | 0.58 | 1.59 | 37.5 |
| 0.4 | 0.5 | 0.96 | 0.52 | 1.52 | 27.78 |
| 0.4 | 0.6 | 0.78 | 0.53 | 1.6 | 26.39 |
| 0.4 | 0.7 | 0.6 | 0.5 | 1.57 | 24.31 |
| 0.4 | 0.8 | 0.42 | 0.44 | 1.48 | 22.92 |
| 0.4 | 0.9 | 0.24 | 0.42 | 1.44 | 24.31 |
| 0.5 | 0.1 | 1.67 | 0.63 | 1.69 | 45.14 |
| 0.5 | 0.2 | 1.5 | 0.63 | 1.67 | 45.14 |
| 0.5 | 0.3 | 1.32 | 0.63 | 1.64 | 43.75 |
| 0.5 | 0.4 | 1.14 | 0.61 | 1.61 | 38.19 |
| 0.5 | 0.5 | 0.96 | 0.56 | 1.59 | 33.33 |

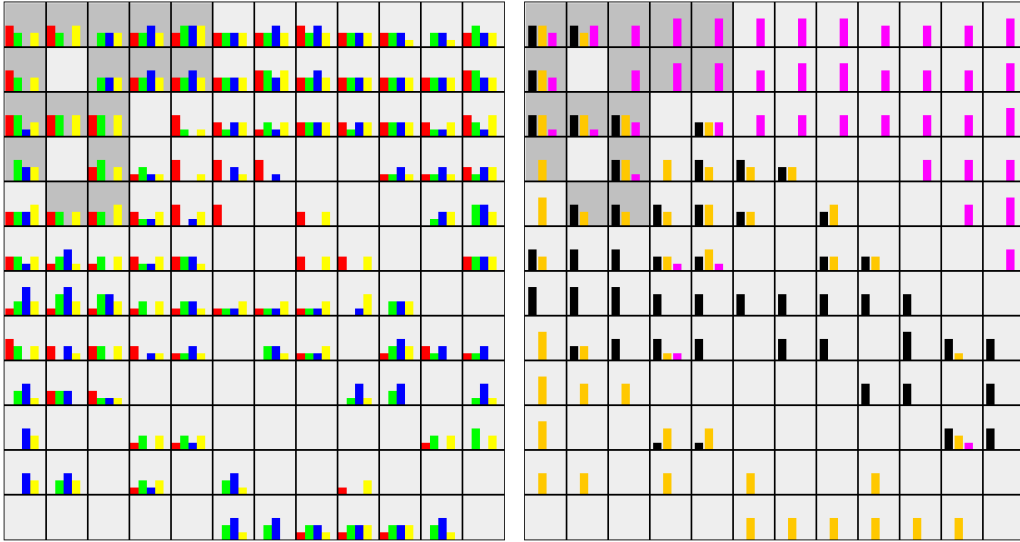
| α | β | q_{err} | Dist. | Entropy | Type-Var |
|----------|---------|-----------|-------|---------|----------|
| 0.5 | 0.6 | 0.78 | 0.54 | 1.59 | 23.61 |
| 0.5 | 0.7 | 0.6 | 0.45 | 1.49 | 20.14 |
| 0.5 | 0.8 | 0.42 | 0.44 | 1.47 | 22.92 |
| 0.5 | 0.9 | 0.23 | 0.33 | 1.38 | 16.67 |
| 0.6 | 0.1 | 1.67 | 0.59 | 1.61 | 45.14 |
| 0.6 | 0.2 | 1.5 | 0.63 | 1.66 | 45.14 |
| 0.6 | 0.3 | 1.32 | 0.71 | 1.73 | 51.39 |
| 0.6 | 0.4 | 1.14 | 0.62 | 1.64 | 38.89 |
| 0.6 | 0.5 | 0.96 | 0.61 | 1.68 | 33.33 |
| 0.6 | 0.6 | 0.78 | 0.54 | 1.57 | 24.31 |
| 0.6 | 0.7 | 0.59 | 0.52 | 1.63 | 22.92 |
| 0.6 | 0.8 | 0.41 | 0.38 | 1.4 | 19.44 |
| 0.6 | 0.9 | 0.23 | 0.35 | 1.37 | 18.06 |
| 0.7 | 0.1 | 1.67 | 0.64 | 1.68 | 45.83 |
| 0.7 | 0.2 | 1.49 | 0.64 | 1.67 | 47.92 |
| 0.7 | 0.3 | 1.31 | 0.63 | 1.65 | 45.83 |
| 0.7 | 0.4 | 1.14 | 0.64 | 1.64 | 45.14 |
| 0.7 | 0.5 | 0.96 | 0.59 | 1.65 | 29.17 |
| 0.7 | 0.6 | 0.77 | 0.58 | 1.68 | 25 |
| 0.7 | 0.7 | 0.59 | 0.47 | 1.5 | 22.92 |
| 0.7 | 0.8 | 0.41 | 0.43 | 1.48 | 22.22 |
| 0.7 | 0.9 | 0.22 | 0.35 | 1.41 | 13.89 |
| 0.8 | 0.1 | 1.67 | 0.61 | 1.67 | 45.14 |
| 0.8 | 0.2 | 1.49 | 0.63 | 1.65 | 46.53 |
| 0.8 | 0.3 | 1.31 | 0.65 | 1.71 | 48.61 |
| 0.8 | 0.4 | 1.13 | 0.69 | 1.75 | 43.06 |
| 0.8 | 0.5 | 0.95 | 0.62 | 1.64 | 40.28 |
| 0.8 | 0.6 | 0.77 | 0.53 | 1.57 | 32.64 |
| 0.8 | 0.7 | 0.58 | 0.46 | 1.46 | 27.78 |
| 0.8 | 0.8 | 0.4 | 0.48 | 1.53 | 27.78 |
| 0.8 | 0.9 | 0.21 | 0.4 | 1.46 | 20.14 |
| 0.9 | 0.1 | 1.67 | 0.63 | 1.7 | 42.36 |
| 0.9 | 0.2 | 1.49 | 0.64 | 1.67 | 47.22 |
| 0.9 | 0.3 | 1.3 | 0.63 | 1.67 | 45.14 |
| 0.9 | 0.4 | 1.12 | 0.65 | 1.7 | 45.83 |
| 0.9 | 0.5 | 0.94 | 0.62 | 1.7 | 42.36 |
| 0.9 | 0.6 | 0.76 | 0.51 | 1.51 | 35.42 |
| 0.9 | 0.7 | 0.57 | 0.47 | 1.51 | 29.86 |
| 0.9 | 0.8 | 0.39 | 0.39 | 1.37 | 25.69 |
| 0.9 | 0.9 | 0.2 | 0.4 | 1.43 | 25 |

Obr. 5.16: Tabuľka zobrazujúca jednotlivé hodnotiace miery siete vzhľadom k výberu parametrov α a β . Pre stĺpec kvantizačnej chyby platí, že hodnoty sme predelili 1000.



Obr. 5.17: Kontúrové grafy hodnotiacich mier.

Na základe týchto výsledkov sme sa rozhodli vybrať za najúspešnejšiu mapu s hodnotami $\alpha = 0.3$, $\beta = 0.6$. Mapa, viď obrázok 5.18, sa vyprofilovala dobre a mala pomerne málo typovo invariantných neurónov, pričom mala veľký počet perspektívne invariantných.

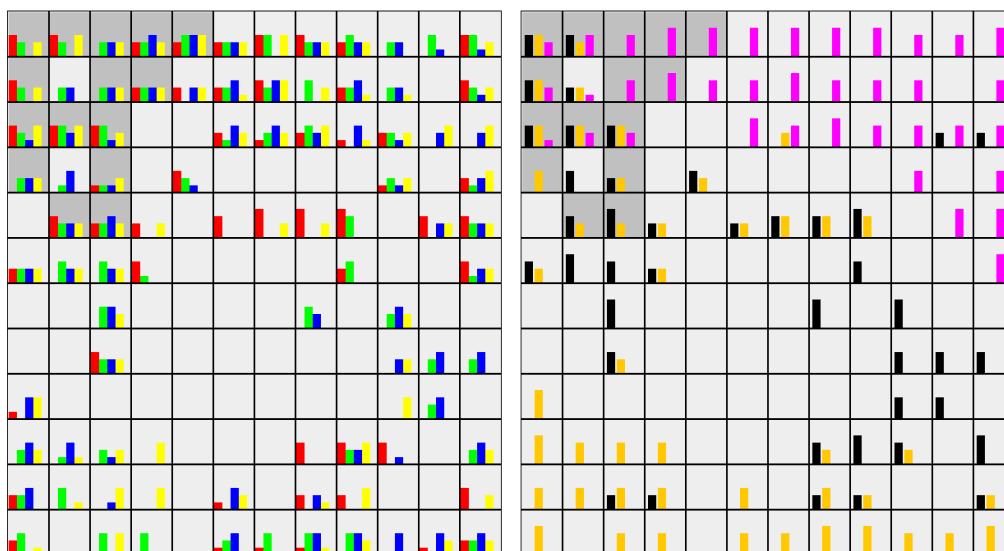


Obr. 5.18: Sieť zrkadliacich neurónov natrénovaná pomocou parametrov $\alpha = 0.3$ a $\beta = 0.6$, ktorá dosiahla najlepšie výsledky.

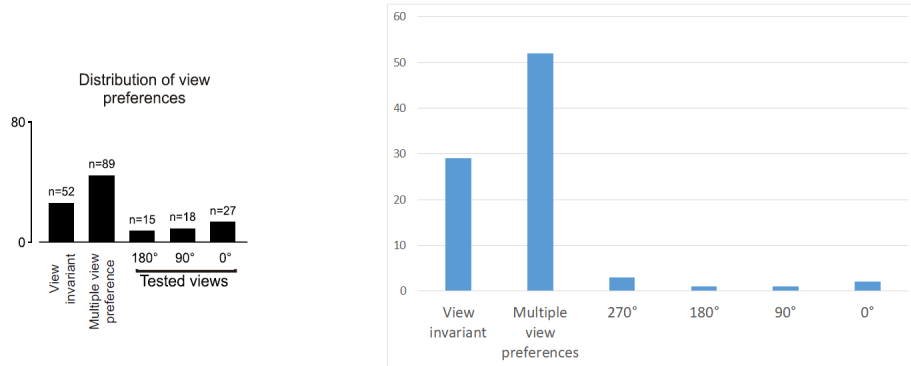
Opäť sme porovnali tieto výsledky s biologickými výsledkami a podobne, ako v predchádzajúcich experimentoch sme došli k záveru, že sieť, čo má najlepšie výsledky, nezodpovedá reálnym modelom. Vychádza to z toho, že my sme sieť hodnotili pomocou mier ako napríklad entropia a distribúcia víťazov. Tieto miery ohodnocujú lepšie tie siete, ktoré čo najrovnomernejšie rozložia svoje prostriedky. V našom prípade to znamená, že vzhľadom na to, že parametrom σ uprednostňujeme motorické informácie, tak sieť, ktorá čo najrovnomernejšie rozloží perspektívy, má najlepšie ohodnotenie. To spôsobuje, že máme veľa invariantných neurónov. Toto je pre nás požadovaná vlastnosť, avšak nekoreluje s biologickými hodnotami.

Podobne, ako v predchádzajúcom experimente, aj teraz sme našli mapu, ktorá nemá najlepšie hodnotenie z pohľadu našich mier, avšak už pomerne dobre koreluje s nameranými výsledkami. Mapu sme si nechali zobrazit', vid'

obrázok 5.19. Sieť má postačujúce variantné a invariantné vlastnosti a zároveň počty variantných a invariantných neurónov korelujú s biológmi nameranými hodnotami, viď porovnanie na obrázku 5.20.



Obr. 5.19: Sieť dosahujúca výsledky najviac podobné s experimentálnymi hodnotami.

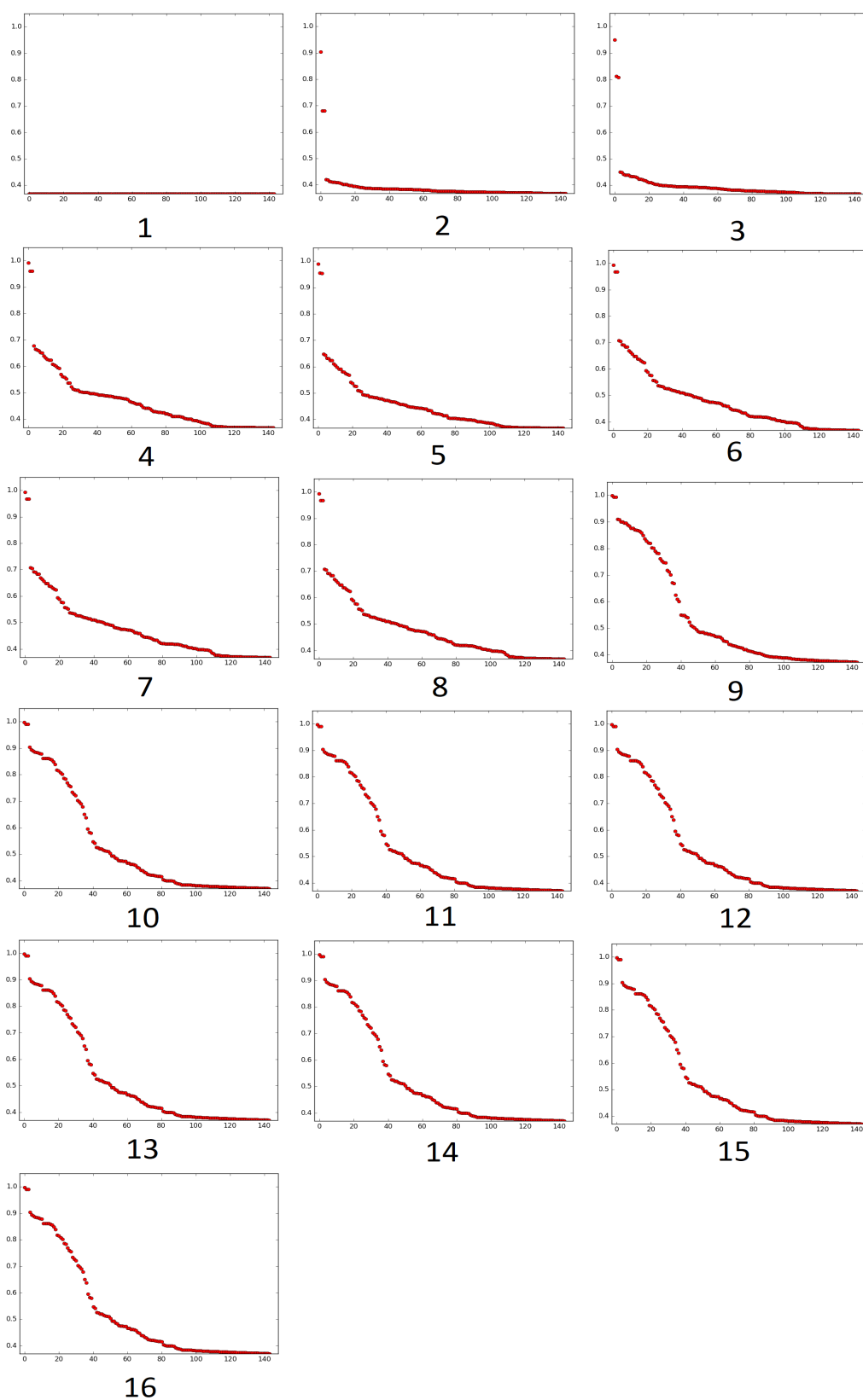


Obr. 5.20: Porovnanie našich výsledkov s experimentálnymi z pohľadu počtov invariantných a variantných neurónov.

Ako posledné sme chceli preskúmať profil aktivácií neurónov počas jednej sekvencie. Podobne ako u STS aj pri zrkadliacich neurónoch boli hodnoty vektora d_i pre jeden frame veľmi podobné a výpočet výstupu siete pomocou rovnice 3.2 bol nedostatočný. Takisto nadobúdali rôzne hodnoty, tak sme namiesto spomínanej funkcie zaviedli:

$$y_x = \frac{1}{1 + e^{k(x - \langle d_j(x) \rangle)}} \quad (5.1)$$

Konštanta k určuje strmosť sigmoidy, v našom prípade sme zvolili $k=10$. x je aktuálna vzdialenosť d_i neurónu, ktorého výstup počítame a $\langle d_j(x) \rangle$ je stredná hodnota všetkých d_i pre aktuálny vstup. Pomocou tejto rovnice vieme zobrazit profil aktivácií pre jednu sekvenciu tak, ako to vidíme na obrázku 5.21. Pre prvé framy sa postupne zvyšuje aktivácia pre najlepšie neuróny, pričom sa vždy spravidla tri pomerne rýchlo priblížia k 1. V druhej polovici sekvencie sa už aktivácie nemenia. To je spôsobené tým, že výstup z STS predkladáme sieti pomocou One-hot a pri uchopovaní (posledných 8 frame-ov) sieť už identifikovala úchop a výstup z STS je teda rovnaký.



Obr. 5.21: Zobrazenie typického profilu aktivácií neurónov počas jednej sekvencie.

Kapitola 6

Záver

V našej práci sme si predstavili fenomén zrkadliacich neurónov a problematiku s nimi spojenú. Vychádzajúc z viacerých prác, zaoberajúcich sa týmto problémom, sme ukázali jeho potenciál na riešenie viacerých problémov spojených s prepojením vizuálnej a motorickej informácie. V práci sme sa zamerali najmä na modelovanie ich variantných a invariantných vlastností, čo sa v ostatných modeloch nerobilo.

Predstavili sme si viacero modelov, ktoré sa zaoberali podobnou problematikou a sami sme následne navrhli model, ktorý sme implementovali pomocou Merge SOM. V experimentoch sa ukázal výber tohto typu samoorganizujúcej sa mapy ako dostatočný a vhodne klasifikoval jednotlivé úchopy pri menšej časovej a pamäťovej náročnosti. Nami zvolený prístup prepojenia F5 a STS pomocou priameho párovania úspešne vytvorili systém zrkadliacich neurónov, ktoré sme následne využili na perspektívne invariantnú klasifikáciu úchopov. Na základe viacerých experimentov a analýz jednotlivých máp sme docielili vznik máp s pomerne dobrou organizáciou zrkadliacich neurónov. Pri experimentoch sme došli aj k záveru, že mapy vyhodnotené nami zvolenými mierami ako najlepšie, svojimi parametrami nezodpovedajú expe-

rimentálnym hodnotám. Pre dosiahnutie takýchto máp sme museli pristúpiť k mapám, ktoré obstáli horšie, avšak stále mali dostatočne dobré výsledky. Ako nemenej dôležité, ako výber správnych parametrov, bol výber máp F5 a STS, pomocou ktorých sme sa systém zrkadliacich neurónov pokúsili vytvoriť. Aj tu sa ukázalo, že nie vždy dosiahneme najlepšie výsledky ak pracujeme s najlepšie vyhodnotenými mapami. Nakoniec sa nám však podarilo vytvoriť organizáciu, ktorá spĺňala naše predstavy o perspektívne invariantnej klasifikácii úchopov a zároveň korelovala so skutočnými biologickými hodnotami.

V tejto práci sme pracovali so zjednodušeným modelom a zámerne sme obchádzali problematiku spojenú s afordanciami. Do budúcnosti je možné tento model rozšíriť o tento fenomén a priblížiť sa tak k skutočnému biologickému systému. Tak isto by bolo vhodné navrhnúť a použiť iné miery, ktoré by siete podobné skutočným nameraným hodnotám vyhodnotil lepšie. Veríme, že nami navrhnutý systém zrkadliacich neurónov by sa dal využiť v kognitívnej robotike, napríklad na natréňovanie robota iCub-a zopakovať videné úchopy pomocou doprednej siete.

Literatúra

- E. Borenstein and E. Ruppin. The evolution of imitation and mirror neurons in adaptive agents. *cognitive systems research*, 6(3). 2005.
- V. Caggian, L. Fogassi, G. Rizzolatti, P. Thier, and A. Casile. Mirror neurons differentially encode the peripersonal and extrapersonal space of monkeys. *science*, 324(5925):403–406. 2009.
- V. Caggiano, L. Fogassi, G. Rizzolatti, J. K. Pomper, P. Thier, M. A. Giese, and A. Casile. View-based encoding of actions in mirror neurons of area f5 in macaque premotor cortex. *current biology*, vol. 21, no. 2, pp. 144–148. 2011.
- G. Chappell and J. Taylor. The temporal kohonen map. *neural networks*, 6:441–445. 1993.
- Y. Demiris and G. Hayes. Imitation as a dual-route process featuring predictive and learning components: A biologically-plausible computational model. in k. dautenhahn, and c. nehaniv (eds.), *imitation in animals and artifacts*. cambridge, ma: Mit press. 2002.
- I. Farkaš and P. Vančo. Spracovanie postupností symbolov pomocou rekurzívnych neurónových máp. in *kognícia a umelý život vii*, slezská univerzita, opava. 99-106. 2007.

- M. Hagenbuchner, A. Sperduti, and A. Tsoi. A self-organizing map for adaptive processing of structured data. *ieee transactions on neural networks*, 14(3):491–505. 2003.
- M. Harries, R. Bevan, J. Hietanen, P. Benson, D. Perrett, M. Oram, and S. Thomas. Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *experimental brain research*, 86(1):159-173. 1991.
- M. Haruno, D. M. Wolpert, and M. Kawato. Mosaic model for sensorimotor learning and control. *neural computation*, 13(10), 2201–2220. 2001.
- G. Hickok and M. Hauser. (mis)understanding mirror neurons. *current biology*, 20(14):593–594. 2010.
- M. Ito and J. Tani. On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system. *adaptive behavior*, 12(2), 93–115. 2004.
- E. Kohler, C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, and G. Rizzolatti. Hearing sounds, understanding actions: Action representation in mirror neurons. *science*, 297(5582), 846–848. 2002.
- T. Kohonen. *Self-organizing maps*. springer. 1997.
- G. Metta, G. Sandini, D. Vernon, L. Natale, and F. Nori. The icub humanoid robot: an open platform for research in embodied cognition. in *proceedings of the 8th workshop on performance metrics for intelligent systems*, pages 50–56. acm. 2008.
- P. Molenberghs, R. Cunnington, and J.B. Mattingley. Brain regions with mirror properties: A meta-analysis of 125 human fmri studies. *neuroscience and biobehavioral reviews*. 2011.

- E. Oztop and M. Arbib. Schema design and implementation of the grasp-related mirror neuron system. *biological cybernetics*, 87:116–140. 2002.
- E. Oztop, D. Wolpert, and M. Kawato. Mental state inference using visual control parameters. *brain research cognitive brain research*, 22(2), 129–151. 2005.
- E. Oztop, M. Kawato, and M.A. Arbib. Mirror neurons and imitation: A computationally guided review. *neural networks* 19, 254–271. 2006.
- G. Pellegrino, L. Fadiga, L. Fogassi, V. Gallese, and G. Rizzolatti. Understanding motor events: a neurophysiological study. *experimental brain research*, 91(1):176–180. 1992.
- D. Perrett, MH Harries, R. Bevan, S. Thomas, PJ Benson, AJ Mistlin, AJ Chitty, JK Hietanen, and JE Ortega. Frameworks of analysis for the neural representation of animate objects and actions. *journal of experimental biology*, 146(1):87–113. 1989.
- K. Rebrová and I. Farkaš. Robotický model systému zrkadliacich neurónov: experimentálna analýza, kognícia a umelý život. 2013a.
- K. Rebrová and I. Farkaš. Bidirectional activation-based neural network learning algorithm. in *proceedings of the international conference on artificial neural networks (icann)*, springer, 154-161. 2013b.
- K. Rebrová, M. Pecháč, and I. Farkaš. Towards a robotic model of the mirror neuron system. in *international conference on development and learning and on epigenetic robotics*, iee. 2013.
- G. Rizzolatti. Functional organization of inferior area 6 in the macaque

- monkey: li. area f5 and the control of distal movements. *exp. brain res.* 71, 491–507. 1988.
- G. Rizzolatti and C. Sinigaglia. The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *nature reviews neuroscience*, 11(4):264–74. 2010.
- G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *cognitive brain research*, 3(2):131–141. 1996.
- G. Rizzolatti, L. Fogassi, and V. Gallese. Neurophysiological mechanisms underlying the understanding and imitation of action. *nature rev. neuroscience*, 2:661–670. 2001.
- N. Sebanz and M. Shiffrar. Detecting deception in a bluffing body: The role of expertise. *psychonomic bulletin and review*, 16(1):170–175. 2009.
- M. Strickert and B. Hammer. Merge som for temporal data, *neurocomputing*, vol. 64, pp. 39–71. 2005.
- G. Tessitore, R. Prevede, E. Catanzariti, and G. Tamburrini. From motor to sensory processing in mirror neuron computational modelling. *biological cybernetics*, 103(6):471–485. 2010.
- V. Tikhanoff, A. Cangelosi, P. Fitzpatrick, G. Metta and L. Natale, and a F. Nori. An open-source simulator for cognitive robotics research: The prototype of the icub humanoid robot simulator. in *proc. of 8th workshop on perf. metrics for intel. systems*, 57–61. 2008.
- M. A. Umiltà, E. Kohler, V. Gallese, L. Fogassi, L. Fadiga, G. Keysers, and G. Rizzolatti. I know what you are doing. a neurophysiological study. *neuron*. 31:155–165. 2001.

M. Varsta, J. del R. Milan, and J. Heikkonen. A recurrent self-organizing map for temporal sequence processing. in proc. icann'97, 7th international conference on artificial neural networks, volume 1327 of lecture notes in computer science, pages 421–426. springer, berlin. 1997.

T. T. Voegtlin. Recursive self-organizing maps. neural networks, 15(8–9):979–991. 2002.

Príloha

Ku práci je priložené CD, ktoré obsahuje:

- Zdrojový kód jednotlivých sietí napísaný v jazyku Java v priečinku Som.
- Zdrojový kód generovania úchopov pomocou iCub napísaný v jazyku C++ v priečinku TeacherGrasp.
- Grafické výstupy zobrazené v práci.
- Samotnú prácu vo formáte pdf.