

Koordinácia ruka–oko v simulátore humanoidného robota

Martin Kellner, Igor Farkaš

Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave
Mlynská dolina, 84248 Bratislava
farkas@fmph.uniba.sk

Abstrakt

Pri koordinácii ruky a oka, napríklad pri sledovaní objektu držaného v dlani, mozog okrem iného prepočítava súradnice medzi dvoma modalitami: propioceptívnou, určujúcou pozíciu ruky v priestore, a vizuálnou, určujúcou pozíciu časti ruky (napr. dlane) na sietnici v retinotopickom súradnicovom systéme, modulovaný pozíciou očí. Tento proces simulujeme pomocou nového modelu obojsmernej umelej neurónovej siete UBAL, ktorá sa učí požadovanú transformáciu na dátach zo simulátora humanoidného robota iCub. Experimenty ukazujú, že naučený model vykazuje pomerne dobrú presnosť a analýza potvrdzuje vznik rôznych typov receptívnych polí neurónov v skrytej vrstve, čo je známy biologický fenomén.

1 Úvod

Koordinácia ruky a očí je jednou z mnohých úloh, ktoré mozog musí vykonávať pri každodennej činnosti človeka, s čím súvisí prepočítavanie súradníc (Blohm, Khan a Crawford, 2009). Napríklad ak človek drží objekt v ruke a pozerá sa naň, objekt je v istej pozícii v (3D) priestore, no táto je súčasne reprezentovaná v rôznych súradnicových systémoch, tzv. referenčných rámcoch (Batista, 2002). Rôzne referenčné rámce týkajúce sa aj rôznych modalít sú navzájom prepojené, takže napríklad pri zavretých očiach a držaní objektu v nejakej polohe pred sebou vieme s istou presnosťou predpovedať, kde objekt je (a z pohľadu mozgu kde by dopadal obraz objektu v oku na sietnici). Platí to aj naopak, smerom od očí k ruke, pričom celá transformácia je modulovaná pozíciou očí, pretože človek môže vnímať objekt aj periférne, keď sa naň nepozerá.

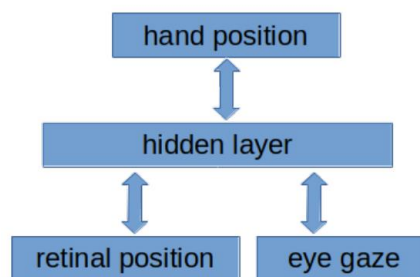
2 Model

Tento proces sme modelovali s využitím simulátora humanoidného robota iCub (Tikhanoff a spol., 2012), ktorý s výškou asi jeden meter odpovedá rozmerom asi tri a polročného dieťaťa. Robot iCub má namiesto očí dve kamery (s rozlíšením 640x480) a 53 stupňov

voľnosti, z toho 3 v očiach (horizontálny, vertikálny smer, a vergencia) a 7 v jednej ruke (a až 9 v dlani, ale tie sme nevyužívali). Simulátor nám poskytol dáta pre výpočtový model. Ukážka zo simulátora je na obr. 1 a znázornenie úlohy na obr. 2.



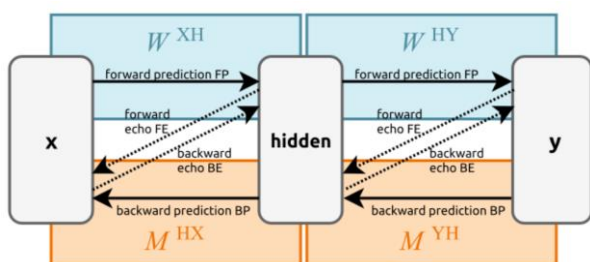
Obr. 1: Simulovaný robot iCub vníma svoju dľaň, aj keď sa na ňu nemusí pozeráť. Pre každú náhodne vygenerovanú pozíciu ruky danú propioceptívnymi súradnicami celej pravej ruky sa zaznamená pozícia očí ako aj obraz dlane na sietnici, čím vznikajú párované dáta odpovedajúce danej situácii.



Obr. 2: Schéma modelu transformácie pri koordinácii ruka–oko. Danej pozícii obrazu objektu (napr. dlane) na sietnici oka, modulovanej zameraním pohľadu (t.j. pozíciou očí v horizontálnom a vertikálnom smere) odpovedá pozícia ruky (a teda aj dlane) v priestore. Pre zjednodušenie bola pozícia hlavy fixná (iCub mal hlavu natočenú dopredu, hýbal však očami). Opačný smer transformácie nie je jednoznačný, no sieť sa naučí predikovať prototyp vizuálnej konfigurácie.

Danú schému sme implementovali pomocou nového modelu umelej neurónovej siete – UBAL (obr. 3), ktorá

je vhodná na učenie *obojsmerných* zobrazení medzi akýmkoľvek dvoma doménami (heteroasociácia), pričom využíva biologicky prijateľnejší algoritmus učenia (Malinovská a spol., 2018). Podobné úlohy v kognitívnej robotike týkajúce sa prepočítavania súradníc boli skôr riešené pomocou štandardného modelu neurónovej siete typu viacvrstvový perceptrón, napr. Zipser a Andersen (1988), a rozšírenie ich modelu s tréňovaním na dátach zo simulátora iCub (Švec a Farkaš, 2014), avšak tieto prístupy boli založené na algoritme spätného šírenia chyby, ktorý nie je v súlade s biologickými princípmi. Učenie v modeli UBAL využíva len šírenie aktivácií medzi neurónmi v oboch smeroch.



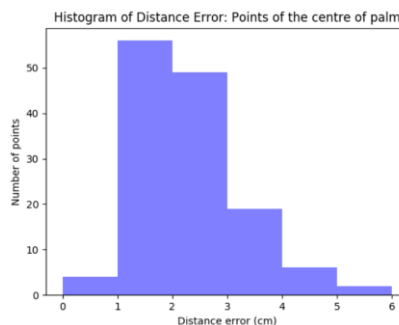
Obr. 3: Model obojsmernej neurónovej siete UBAL na prepojenie dvoch domén, vizuálnej (X) a proprioceptívnej (Y). Prevzaté z Malinovská a spol. (2018).

Doménu X predstavuje vizuálny systém, t.j. kombinácia obrazov na sietnici a pozície očí a doménu Y pozícia ruky (ktorej jednoznačne odpovedá pozícia dlane). Skrytá vrstva neurónov umožňuje premostenie oboch domén, pričom na nej vznikajú zaujímavé reprezentácie (receptívne polia neurónov). Ich vlastnosti závisia od kódovania informácie na vrstvách X a Y. Skúmali sme dva spôsoby: (a) lokalistické kódovanie, pri ktorom každý stupeň voľnosti je reprezentovaný reálnou hodnotou na jednom neuróne (s vhodným preškálovaním), t.j. 2 pre pozíciu očí, 2+2 pre pozíciu obrazu na sietniciach a 7 pre pozíciu ruky. (b) Pozičné kódovanie, pri ktorom je každý stupeň voľnosti reprezentovaný pozične na viacerých neurónoch (s gausovskými charakteristikami). Takto sme použili 13 neurónov pre pozíciu očí, 36 pre pozíciu ťažiska obrazu na sietnici a 33 na stupne voľnosti pravej ruky.

3 Výsledky

Model UBAL bol tréňovaný na vzorke 1590 a testovaný na vzorke 250 párov dát počas 200 epoch, pričom optimálne hyperparametre modelu boli nájdené experimentálne. Model sa naučil transformáciu celkom úspešne, nakoľko priemerná chyba (MSE) bola asi 2 cm v prípade lokalistických reprezentácií. Distribúcia predikcií z X do Y na testovacej množine je na obr. 4. Taktiež sme pozorovali, že pozičné kódovanie ako

biologicky inšpirovaná forma reprezentácie, umožnilo vznik rôznorodých receptívnych polí neurónov skrytej vrstvy pozorovaných v mozgu. Na druhej strane, v tejto verzii modelu bola predikcia o niečo horšia, s priemernou chybou asi 4 cm. Podarilo sa nám však úspešne overiť funkčnosť modelu UBAL v oblasti kognitívnej robotiky.



Obr. 4: Predikcie pozície dlane v modeli UBAL vyjadrením vzdialenosti od skutočnej pozície.

PodĎakovanie

Tento príspevok vznikol s podporou grantovej agentúry VEGA v rámci projektu 1/0796/18.

Literatúra

- Batista, A. (2002). Inner space: Reference frames, *Current Biology*, 12(11), 114–117.
- Blohm, G., Khan, A., Crawford, J. (2009). Spatial transformations for eye–hand coordination, *Encyclopedia of Neuroscience*, p. 203–211.
- Malinovská, K., Malinovský, L., Farkaš, I. (2018), Towards more biologically plausible error-driven learning for artificial neural networks. In *Artificial Neural Networks and Machine Learning – ICANN*, Lecture Notes in Computer Science, p. 228–231.
- Švec, M., Farkaš, I. (2014), Calculation of object position in various reference frames with a robotic simulator. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society, Quebec, Canada*.
- Tikhonoff, V., Fitzpatrick, P., Nori, F., Natale, L., Metta, G., Cangelosi, A. (2012), The iCub humanoid robot simulator. *IROS Workshop on Robot Simulators*.
- Zipser, D. and Andersen, R. A. (1988), A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331(6158), 679–684.