

Predikcia dynamických systémov rekurentnými neurónovými sieťami

Študijný odbor: Umelá inteligencia

Autor: Ing. Peter Trebatický

Vedúci dizertačnej práce: Prof. RNDr. Jiří Pospíchal, DrSc.

Máj 2009

Rekurentné neurónové siete pri predikcii časových radov vo všeobecnosti dosahujú lepšie výsledky ako dopredné siete. Alternatívou k nim sa ukazujú byť neurónové siete s echo stavmi. Na úlohe korekcie poškodeného textu som ukázal, že dosahujú mierne lepšie výsledky ako už známa metóda s použitím Markovovského modelu. Väčšia časť práce je zameraná na alternatívy tréningu rekurentných neurónových sietí založené na modifikácii Kalmanovej filtrácie. Podrobne popisujem tréning pomocou filtrov: rozšírený Kalmanov filter, „Unscented“ Kalmanov filter (UKF), nprKF filter a ich zdieľané verzie UKFj a nprKFj. Filter UKFj bol v kontexte rekurentných neurónových sietí popísaný pravdepodobne prvýkrát v mojej práci. Prínosom tejto práce je uvedenie jednoduchších rovníc jednotlivých filtrov, nakoľko sú upravené čisto pre účely tréningu rekurentných sietí. Porovnávam jednotlivé filtre navzájom a s gradientovou metódou Skrátené spätné šírenie chýb v čase (BPTT(h)). Ukazujem, že výsledky rekurentných neurónových sietí tréningovaných týmito pokročilými metódami sú konzistentne lepšie v porovnaní s BPTT(h). Podobne rozšírený Kalmanov filter dosahuje horšie výsledky oproti ostatným spomínaným filtrom, ktoré zase dosahujú navzájom porovnateľné výsledky. V práci popisujem, akým spôsobom možno urýchliť ich výpočet prostredníctvom grafickej karty. Moja práca je jedna z prvých (ak nie prvá) zaoberajúca sa tréningom rekurentných neurónových sietí s využitím procesora na grafických kartách.

Prediction of Dynamical Systems by Recurrent Neural Networks

Degree Course: Artificial intelligence

Author: Ing. Peter Trebatický

Supervisor: Prof. RNDr. Jiří Pospíchal, DrSc.

May 2009

Recurrent neural networks in general achieve better results in prediction of time series than feedforward networks. Echo state neural networks seem to be one alternative to them. I have shown on the task of text correction, that they achieve slightly better results compared to already known method based on Markovov model. The major part of this work is focused on alternatives to recurrent neural networks training that are based on Kalman filtration modifications. I describe in detail the training by filters: Extended Kalman Filter, Unscented Kalman Filter (UKF), nprKF Filter and their joint versions UKFj and nprKFj. Filter UKFj in context of recurrent neural networks was probably firstly described in my work. Contribution of this work is presentation of simpler equations for individual filters, because they are modified specifically for recurrent neural network training. I compare individual filters with each other and also with gradient descent method Truncated Backpropagation Through Time (BPTT(h)). I show the results are consistently better when comparing recurrent neural networks trained by these advanced methods with BPTT(h). In the like manner, Extended Kalman Filter achieves worse results compared to the other filters, which on the other hand achieve comparable results with each other. I describe how to speed up their computation by utilizing the graphics card. My work is one of the first (if not the first) that focuses on recurrent neural network training utilizing the processor on graphics card.