VISION/GESTURE RECOGNITION USING SPIKING NEURONS

October 6, 2014

By

Qian Liu

School of Computer Science

Contents

List of Tables

List of Figures

Abstract

The extraction of the report not the research.

The aim, objectives and motivations are...

What have been done...

The results of experiment...

One sentence about the conclusions and future plan. About half page long.

Chapter 1

Introduction

1.1 Aim

Objectives and Motivations behind it.

1.2 Why is it important

Why is it important to research on vision process in the brain.

Chapter 2

Background

Literature Review

2.1 Posture/Gesture Recognition

All the literature review I can borrow from the previous report.

A pattern or an object in a two-dimensional image can be described with four properties [1]: position, geometry (size, area and shape), colour and texture, and trajectory. Appearance-based methods are the most direct approaches to perform pattern recognition. The test image is compared with all the templates to find the best match on one particular or a combination of properties. In terms of classification algorithms, distance measure methods (nearest neighbour, k-means clustering), support vector machine (SVM), multi-layer perceptron (MLP) neural networks and statistical methods, e.g. Gaussian mixture model (GMM) have been applied successfully in visual recognition. Since the 2D projection of an object changes under various illuminations, viewing angles, relative positions and distances (size), it is impossible to represent all appearances of an object in different conditions. Robust matching methods are employed, such as edge matching [2], the divide-and-conquer approach [3], gradient matching [4], etc. Moreover, feature based methods are used to improve reliability, robustness and classification efficiency. Among various feature extraction methods, the scale-invariant feature transform (SIFT)[5] and the sped-up robust features (SURF)[6] methods are well-accepted recently in the field. However, to find a proper feature for a specific object still remains an open question and there is not any process as accurate, general and effective as the brain.

Turning to biology for answers is always the way to explore the field of visual pattern recognition. Riesenhuber and Poggio [7] presented a biologically-inspired model following the organization of the visual cortex which has the ability to represent relative position- and scale-invariant features. Integrating a rich set of visual features became available using a feed-forward hierarchical pathway.

More and more attention has been drawn into the investigation of spiking neural networks for vision processing. Pattern information can be encoded in the delays between the pre- and post-synaptic spikes since the spiking neurons are capable of computing radial basis functions (RBFs)[8]. A further study [9] has stated that spatio-temporal information can be also stored in the exact firing time instead of the relative delay. Maass [10] has proved mathematically that:

1) networks of spiking neurons are computationally more powerful than the first and second generation of neural network models;

2) a concrete biologically relevant function can be computed by a single spiking neuron, replacing hundreds of hidden units in a sigmoidal neural net;

3) any function that can be computed by a small sigmoidal neural net can also be computed by a small network of spiking neurons.

Applications of SNN-based vision processing have been successfully carried out. A two-layered SNN has been trained using spike time dependent plasticity (STDP) and employed for a character recognition task [11]. Lee and co-authors [12] have implemented the direction selective filters in real time using spiking neurons. The direction selective filters here are considered as a layer of convolution module in the model of so called convolution neural network [13]. Different features, such as Gabor filter features (scale, orientation and frequency) and shape can be modeled as layers of feature maps. Rank order coding, as an alternative to conventional rate-based coding, treats the first spike the most important and has well applied to an orientation detection training process [14]. Nengo [15] is a graphical and scripting based software package for simulating large-scale neural systems and has been used to build the world's largest functional brain model, Spaun [16]. An FPGA implementation of a Nengo model for digit recognition has been reported [17]. Deep Belief Networks (DBNs), the 4th generation of artificial neural network, has shown a strong ability in solving classification problems. A recent study [18] has resoundingly mapped an offline-trained DBN onto an efficient event-driven spiking neural network for a digit recognition task.

So far to our knowledge, no complete hand posture/gesture recognition application study has been published. The main reasons lie in the complex shape of the hands.

Hand segmentation and feature extraction usually takes the colour into account and involves wavelets, e.g. using a Kalman filter. Thus, shape-only recognition of the hand posture will be a challenge. In the terms of gesture recognition, the hidden Markov model (HMM) has shown its ability to recognize dynamic gestures [19]. However, with their instinctive temporal processing, SNNs have the potential to deliver dynamic gesture recognition.

2.2   Biology Aspect

A lot to work on this part. May or may not include neuron models.

2.3   Platforms

2.3.1   Vision Processing Front-ends

2.3.2   SNNs Back-ends

2.3.3   SpiNNaker distinguishing features

Should include detailed description and more focus on FPGA connecting to sensors.

Chapter 3

CNNs Models

Models description.

3.1 Model Description

Layer by layer

Biology background on V1, especially for Gabor Filters.

3.2 Training and test

3.3 Experiments

Chapter 4

Recognition on SpiNNaker

4.1 Moving to spiking neurons

Mathematical approach.

Need to work hard on.

4.2 Recorded data test

New experiment to work on.

New tool need to be fixed (Spike Source Array), and maybe new ways of reloading.

4.3 Real-time test

Videos will help here.

Chapter 5

Research Plan

5.1 Biology Aspect

5.2 Tracking

5.3 HMM for gesture recognition

Chapter 6

Conclusion

Just a conclusion

Bibliography

[1] S. G. Wysoski, L. Benuskova, and N. Kasabov, "Fast and adaptive network of spiking neurons for multi-view visual pattern recognition," *Neurocomputing*, vol. 71, no. 13, pp. 2563–2575, 2008.