IMAGE RECOGNITION WITH HARDWARE NEURAL NETWORKS

Karlis Berkolds
Riga Technical University
karlis.berkolds@rtu.lv

Abstract. One of the modern and potentially highly beneficial trends in precision agriculture is application of mobile robots. These robots provide capabilities like planting, watering, fertilizing, harvesting crops, thus saving time and resources. One of the major challenges to be addressed to the mentioned capabilities is the ability to recognize crops, weeds, soil etc. As distinguishing features of crops and their environment have wide variety, the recognition system should have the ability to adapt and learn thereby ensuring recognition of previously unseen scenes. The current state of the art of image recognition methodologies mostly focuses on various types of artificial neural networks that usually are implemented as software imitation. Benefits for using imitation are greater adaptability and easier modification, but the drawback is the required computing power. In mobile robots these resources usually are of great value due to the limited space and power source capacity. As an alternative to imitation we propose to use artificial hardware neural networks. The objective of the presented research is to determine artificial hardware neural network capabilities in image recognition and develop an image recognition system for mobile robots. In this paper existing solutions of artificial hardware neural network usage, solution structure, restrictions and possibilities are examined. Using the only hardware neuron controller available on the market CM1K, its supporting hardware and software are developed and presented for the image recognition sensor. Image recognition capabilities are tested by training neural networks using photos of objects of different colour and shape. In conclusion the acquired recognition results and CM1K capabilities of image recognition are analysed and discussed.

Keywords: hardware neural network, image recognition, CM1K neuron controller.

Introduction

Human vision is well developed, it performs massively parallel image processing, which can detect image meaning and context [1]. Closest equivalent to biological brain is assumed to be artificial neural network. Artificial neural network consists of individual units -neurons, which can process and save information and deliver it further to other neurons. The system of neurons has many properties that make it suitable for image recognition, like parallel architecture, ability to learn and adapt and damage tolerance.

Most current neural network solutions use software simulation on sequential von-Neuman architecture computers, thus loosing many positive neural network properties [2]. Specialized application hardware neural networks are cheaper and easier to use and they can take full advantage of neural network inherent parallelism. One of the possible solutions available on market is a chip CM1K.

In CM1K neurons are physically separate units. Information learning and recognition are processed simultaneously and neurons work independently. Parallel hardware accelerates the recognition process, simultaneously reducing computing and power consumption. These properties allow to design small, efficient, real-time image recognition sensors [3]. Existing alternatives to CM1K are designed for stationary image recognition and quality control in manufacturing. For example, CM1K predecessor zero instruction set computer ZISC is used in fishing industry, for fast and precise fish sorting equipment. ZISC was used because it could solve the nonlinear problem of fish recognition, provide reliable operation in minimal space and could recognize more than 360 fish per minute. The system can be trained depending on the fishing season conditions by fishermen themselves and achieve up to 98 % recognition precision [4]. For the same reasons CM1K is used in the vehicle plate recognition system. CM1K could recognize vehicle plate in 101µs, while for other technologies it takes milliseconds. It is also pointed out, that the CM1K system is more stable, than the software system, because hardware is used for most of the recognition and it has low power dissipation, which is important for portable systems [5]. There is also research, where hardware neural networks are suggested as solution for real-time robot weed control. Neural network classifies weeds and crops and makes a spray map for herbicides [6].

In this study, the only available neuron controller on market CM1K was used to examine its capabilities and limitations for real-time image recognition. Examining the existing solution structure and manufacturers' documentation, a simple image recognition sensor was developed. For the first sensor prototype the main aim was to learn about the technology and study its capabilities. Therefore, images of objects of different shapes and colours were used to train the sensor, to test the sensor learning and recognition abilities. Acquired results revealed that the sensor performed close to the expected results as it could recognize all objects and the recognition speed was near to the calculated. The results also indicated needed improvements for the current solution in order to use the sensor in mobile robots for real-time applications.

CM1K neural controller

CogniMem manufactured neuron controller CM1K is the only artificial hardware neural network available on the market at the moment. CM1K is completely parallel silicon neural network. It consists of 1024 identical neurons, which can store and process information simultaneously. Neurons cooperate through bi-directional and parallel neuron bus. Cooperation ensures that neuron will not learn unnecessary information, but could recognize novelty and conflicts. Parallel architecture's advantage is that learning and recognition time is independent from the neuron count in the network [7]. Neural network can be expanded by cascading chips via common communication bus and dedicated control signals. CM1K has a built in recognition engine, which can extract attributes from video or sensor data and create the feature vector. The vector is used for learning and recognizing patterns. Neural network uses two classifiers: Radial basis function (RBF) [7] and K-Nearest Neighbour (KNN) [7] classifier. CM1K low pin count, low power consumption and scalability make it suitable for sensors in mobile robots.

CM1K neuron is cognitive and reactive memory, which can autonomously estimate L1 [7] and Lsup [7] distance between the incoming pattern and reference pattern in memory. Distance determines similarity with the reference pattern in neuron. Smaller distance means greater similarity. If the calculated pattern is within the neuron influence field, neuron returns affirmative response classification, which consists of the distance value and reference pattern category. After activation neuron can compare its result to other neuron results and refuse to participate in category determination, if its distance is larger than other neuron distances. If many neurons within the same category report small distance, it confirms the classification believability. Each neuron has its own distance calculation unit and all distance calculations for the incoming pattern happen in parallel. CM1K has two distance calculation functions L1 or Manhattan distance and L Sup. L1 emphasizes offset of the vector component sum between the incoming vector and the stored pattern. It is more used to classify patterns with different units of measurement. L Sup emphasizes the largest offset of the single vector component, between the incoming vector and the stored pattern [8].

CM1K neuron controller can be considered as a three-layer neural network: the input layer has 256 8bit inputs, Hidden layer consists of 1024 neurons, creating 262 144 8bit synapse connections and the output layer has 32 766 outputs [9]. All neurons are identical, they receive the same instructions and execute them simultaneously, and there is no need for external manager to organize interaction between them. In the moment of neural network initialization all neurons are idle, except the first one in the chain, which is ready to learn. During learning, neurons are used in increasing order to save reference models and categories and become committed. Each neuron can be used to recognize different contexts, for example, one part of neurons can be used for video recognition, while the other for sound recognition at the same time, thus creating sub networks for different kinds of information [10].

By learning, neurons create a decision space [8], which determines neural network recognition abilities. Decision space can be displayed as n dimensional graph, where n is the attribute or measurement count, which characterize the object. Attributes create the feature vector, which can be plotted in a graph. After plotting several vectors in the decision space, they start to form clusters of vectors with similar category. After these clusters, a new example category could be determined.

CM1K neural network can create nonlinear decision space, using RBF model generator and classify incoming vectors using RBF model or k-NN. Example of two dimensional decision spaces can be seen in Fig. 1. Generalizing the learnt knowledge, the neural network can recognize previously unseen patterns. Each neuron belongs to recognizable categories. Neuron has an influence field [8], which shape is determined by the distance calculation method.

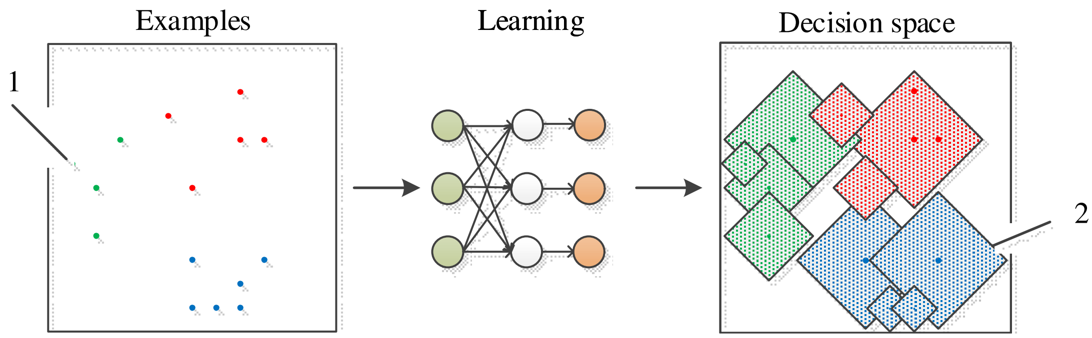

Fig. 1. **Example of two dimensional decision spaces [8]:**
1 – feature vector; 2 – neuron influence field

RBF classifier is capable of processing uncertainty. It is especially useful for anomaly and novelty detection. Classification determines, in which neuron influence fields the new vector is located. There are three possible outcomes [8].

- Unrecognized, the vector is outside the neuron influence fields. In this case new neuron is committed and the influence field is created around neuron, thus increasing the decision space.
- Recognized, the vector is inside one or more neuron influence fields within the same category. Decision space does not change.
- Uncertain, the vector is inside two or more neuron influence fields with different categories. Neurons, which are responsible for incorrect classification, shrink their influence fields till the vector is outside the field.

After few learning iterations decision space could recognize most of population with high precision, but it has problems with the rest of the population. If recognition gives an uncertain result, then it can be resolved the following operations.

- Choosing the category after neuron with the smallest distance
- Selecting *N* neurons with the smallest distance. The best match is determined using probabilities.
- Writing a correct answer and using other neural network with different context to get a better result without uncertainty.

Influence field has two threshold values, the maximum influence field (Maxif) and minimum influence field (Minif) value. Maxif is set when new neuron is created and determines the initial size of the influence field. Decision space mapping can be changed from moderate to careful by changing the Maxif value. Minif is the value below which the influence field cannot shrink. Uncertainty area can be influenced by changing the Minif value [8].

Decision space is created during learning and its shape is determined by the order of examples. To create precise decision space, examples must be taught repeatedly, until stable enough recognition quality is achieved. Decision space can be considered as stable if after two learning iterations no new neurons are committed. If the example is recognized, then no changes to the decision space are applied. In the case of incorrect classification, no new neurons are committed - neurons only shrink their influence fields. Increasing the number of new examples, the decision space continues changing and it is possible, that the initial examples are no longer recognizable. By repeatedly learning examples, neural network commits new neurons. This is the reason why learning order is important. To analyse the learning results, a learning curve can be used, it shows committed neuron count against new examples. If the examples are significant neurons can generalize knowledge and the committed neuron count reaches stable value. If the learning curve is close to straight line it means that the learning vector has insufficient information for the learning example and generalizing knowledge. By examining learning curves for different categories, recognition difficulties can be determined [8].

Sensor structure

Studying existing solutions structures [4; 5; 11], which use CM1K neuron controller or its predecessor and examining manufacturers' suggested architectures [12], one can see that the sensor structures are very similar and consist of three main parts: video sensor, neuron controller and field-programmable gate array (FPGA).

FPGA is suggested as the main control device for the whole system, as it can be configured to work in a parallel mode without losing performance. As the research purpose is to determine CM1K capabilities in image recognition, a microcontroller [13] is used instead of FPGA, Fig. . Microcontroller is easier to program and it has more built in interfaces for debugging and connection to other devices. The main control device is used for image pre-processing and feature extraction. But as CM1K has built-in vector extraction from the video sensor it receives raw image data from camera. Microcontroller is used for image forwarding, CM1K control and debugging.

CMOS video camera is used as a video sensor. Its main features are: low resolution, small size, low power consumption. In the current solution the image data are sent through the microcontroller, so the image data can be collected for analysis and debugging. To increase the system speed, the video sensor should be connected directly to the neuron controller trough parallel port.

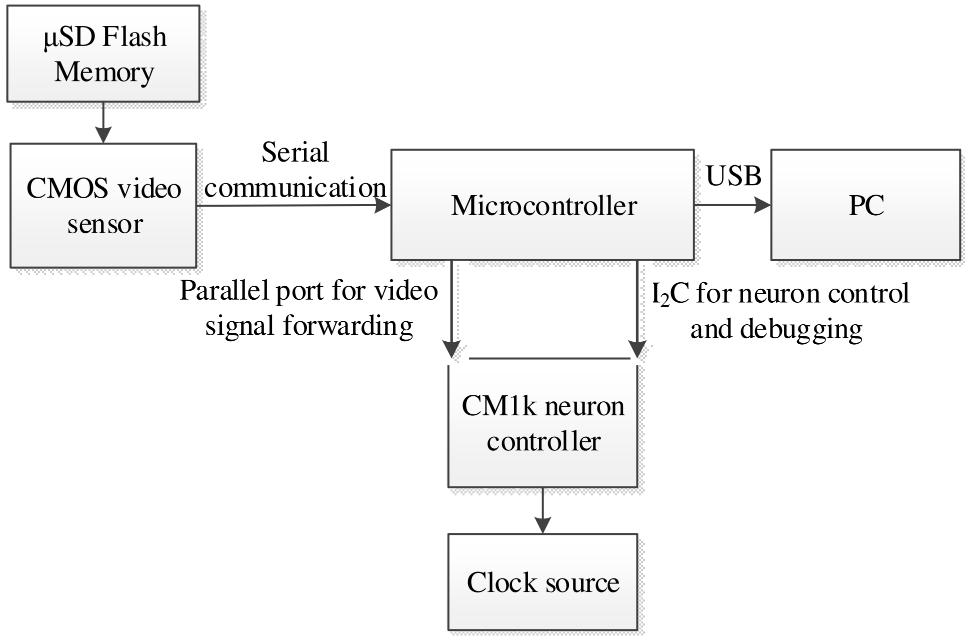

Fig. 2. **Sensor structure**

Control program provides device communication and neural network learning and recognition functions. Commands are sent from the computer to the microcontroller trough serial port. CM1K control interface consists of two main commands: learning and recognizing pattern. All operations are executed by reading and writing 15 CM1K register trough I2C slave controller [10].

Pattern learning starts by turning on recognition module and sending image trough parallel video port. Recognition module extracts the feature vector and broadcasts it to neurons. Then the category of the pattern is sent to the neuron controller. If the sent pattern is novelty new neuron is added to the network, if else, the network stays the same [14].

Pattern recognition is similar to learning. Feature vector is broadcasted to neurons and recognized. Response from "firing" neurons consists of the distance between the pattern and pattern stored in neuron and the category of neuron. If the classification is incorrect, neurons can be taught by changing the category register [14].

Image recognition experiments

Image recognition capabilities are tested by training neural networks using photos of objects of different colours and shapes. The objects are placed on a platform in constant distance, which is used for learning and recognition.

Neural network is trained by a set of 42 different shape and colour objects. The objects are divided in 10 groups, 3 groups with similar shape but different colour and 7 groups with different shape but similar colour. Each object is freely placed on a platform with white background. Fig. 3 shows examples of a group with black objects.

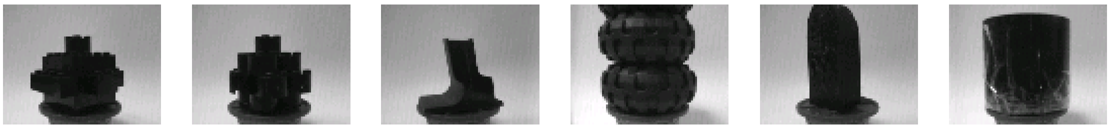

Fig. 3. **Example of black objects with different shape**

Neural network is trained using supervised learning. Objects are shown in constant order to the neural network. If the object is recognized correctly, the next object is shown. If the object category is wrong or it is not recognized, the correct category is written and the step is repeated till successful recognition. Object group is taught till 5 times in row there are no wrong classifications and no new neuron is committed, the decision space is considered stable.

Image size is 60x80 pixels. Feature vector is acquired by dividing the image in 192 5x5 pixel pieces and calculating average value of each piece. Image recognition time is 186μs using 27 MHz clock source. It takes 4800 clock cycles for sending the image, 192 to broadcast the vector to the neural network and 37 for recognition. Recognition time remains constant regardless of information in the picture or neural network size.

Results and discussion

Recognition results show that to recognize objects with different colour and similar shape more committed neurons are needed than for objects with different shape and similar colour, Fig. 4.

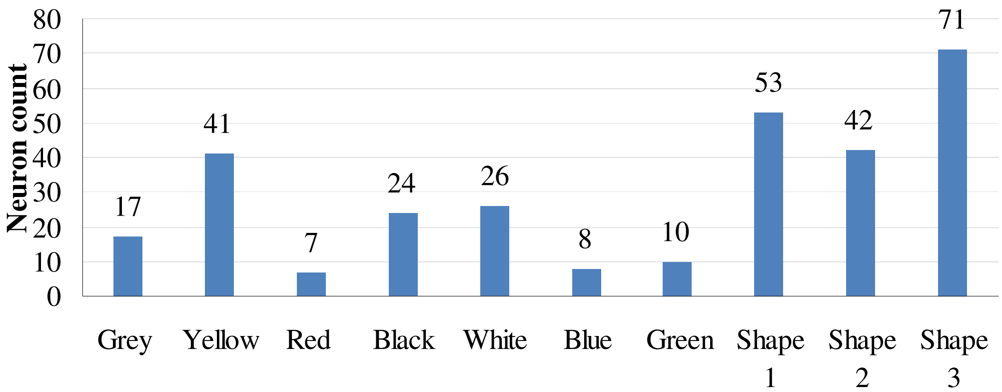

Fig. 4. **Committed neuron count for different object groups**

To recognize complicated objects more learning iterations, Fig. 5, and committed neurons were needed.

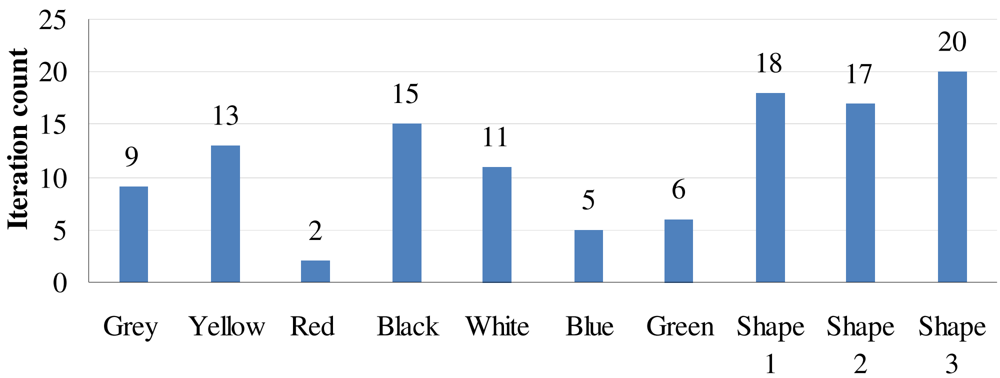

Fig. 5. **Iteration count to reach stable decision space for different object groups**

For example, in the previously seen black object group, the first and second objects are very similar, that is why more neurons and iterations were needed to create stable decision space. The results differ significantly between different groups, because the object shapes and colours are similar, but not identical.

Because of hardware restrictions the used images are monochrome and it complicates colour recognition. Hardware improvements, which would allow use of colour images, increase the power consumption, reduce recognition speed or reduce the image quality.

Conclusions

CM1K neural network usage after creation of hardware and software is very simple and consists of two functions for learning and recognitions. Using only basic functionality of the neural network, high speed and precision can be achieved. Neural network could classify correctly all object images. Depending on the image complexity, the committed neuron and learning iterations count, required for stable decision space, changed. Because of hardware restrictions, different shapes are easier recognized than different colours. Neural network size does not influence the recognition time and for the current solution it is about 186 μs.

Future work includes hardware and software improvements to increase the recognition speed and functionality. Some of the possible functionality includes image search and real-time object tracking. Image search would allow determining the learnt object location in the image and its classification. Picture of a field could be analysed and give information about the cultivated plant and weed locations for watering or weed control. As CMIK has low power consumption it can be used in small cheap image recognition sensors for precision agriculture.

Acknowledgement

This research has been supported by the Latvian state research program 2014-2017 project “The Next Generation of Information and Communication Technologies (NexIT)”.

References

1. Mitchell T.M. Machine Learning. MCGraw-Hill, 1997. 432 p.
2. Liao Y. Neural Networks in Hardware: A Survey. University of California, 2003. 18 p.
3. CogniMem Tehnologies, Inc. CM1K Datasheet, 2012. 2 p.
4. Menendez A., Paillet G., Fish Inspection System Using a Parallel Neural Network Chip and the Image Knowlege Builder Application. AI Magazine, vol. 29, 2008, pp. 21-28.
5. Liu Y., Wei D., Zhang N., Zhao M., Vehicle-License-Plate Recognition Based on Neural Networks. Proceeding of the IEEE international conference “Information and Automation”, June, 2011, Shenzen, China, pp. 363-366.
6. Chris G., David S., Crop verses Weed Recognition Artificial neural networks: Neural Network Plant Recognition for Vision Based Robotics Weed Control. Proceeding of the ASAE conference “Real-Time Image Applications”, 2001, Sacramento.
7. CogniMem Tehnologies, Inc., CogniMem Tehnology Reference Guide. 2012, 30 p.
8. CogniMem Tehnologies, Inc., CogniMem Neural Network Tehnology Decision space mapping, 2011, 18 p.
9. LeivaL., Acosta N., Hardware Radial Basis Function Neural Network Automatic Generation. JCS&T , vol. 11, 2011, pp. 15-21.
10. CogniMem Tehnologies, Inc., CM1K Hardware manual. 2012, 59 p.
11. Malasne N., Yang F., Paindavoine., Real time implementation of face tracking. University of Burgundy, France, 2003, 4 p.
12. CogniMem Tehnologies, Inc., CM1K Arhitectures, 2011, 6 p.
13. MBED LPC1768 “ARMmbed”. [online] [30.03.2016]. Available at: https://developer.mbed.org/platforms/mbed-LPC1768/
14. CogniMem Tehnologies, Inc., CogniMem Tehnical brief. 2012, 2 p.