ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

Hardware accelerated real time classification of hyperspectral imaging data for coffee sorting

Andreas Backhaus[1], Jan Lachmair[2], Ulrich Rückert[2], Udo Seiffert[1]

1- Biosystems Engineering, Fraunhofer Institute for Factory Operation and Automation (IFF), Magdeburg - Germany

2- Cognitronics & Sensor Systems, University Bielefeld, Bielefeld - Germany

Abstract. Hyperspectral imaging has been proven to be a viable tool for automated food inspection that is non-invasive and on-line capable. In this contribution a hardware implemented Self-Organizing Feature Map with Conscience (CSOM) is presented that is capable of on-line adaptation and recall in order to learn to classify green coffee varieties as well as coffee of different roast stages. The CSOM showed favourable results in some datasets compared to a number of classical supervised neural network classifiers. The massive parallel neural hardware architecture allows for constant processing times at different map sizes.

1 Introduction

Quality control of coffee products, from basic green coffee to the finished roasted coffee by hyperspectral imaging has been shown to offer the means for a non-invasive, on-line and automated screening method to control large product quantities [1, 2]. For example green coffee has to be sorted for various defects to the bean or the stability of the finished product has to be evaluated. In order to characterize these materials beyond their colour and shape, their narrowly sampled spectrum in the short wave infra-red (SWIR) range is collected using hyperspectral imaging. Each pixel of an acquired image contains a full spectrum which is characteristic for the material's chemical composition and forms a high-dimensional data vector or pattern. Due to the spatial resolution, loose materials like coffee can be inspected spatially for sorting purposes.

For classification an Artificial Neural Network is learnt from exemplary data. In order to be able to teach the system as well as recall from the model in real-time an implementation in integrated hardware is desirable. This opens the possibility of integrating the data analysis within the camera hardware for a smart camera set-up. In this contribution we present a hardware accelerated implementation of the Self-Organizing Feature Map with Conscience [3, 4, 5, 6]. A post-labeled CSOM was used to classify green coffee beans as well as beans of varying roasting degrees. The classification accuracy is benchmarked to other prototype based Neural Networks like Radial Basis Function (RBF) Network, Generalized Relevance Learning Vector Quantization (GRLVQ) [7] and Supervised Relevance Neural Gas (SRNG) [8]. Additionally, a Multilayer Perceptron (MLP) Network and a Support Vector Machine (SVM) were trained as comparison.

2 Data Acquisition

Coffee beans of each class were recorded separately. Beans and a standard optical PTFE (polytetrafluoroethylene) calibration pad were positioned on a translation

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

table. Hyperspectral images were recorded using a HySpex SWIR-320m-e line camera (Norsk Elektro Optikk A/S). Spectra are from the short-wave infrared range (SWIR) of 970 nm to 2,500 nm at 6 nm resolution yielding a 256 dimensional spectral vector per pixel. The camera line has a spatial resolution of 320px and can be recorded with a maximum frame rate of 100fps. Radiometric calibration was performed using the vendors software package. Coffee beans were segmented from background via Neural Gas clustering. Dataset A comprised of four different green coffee varieties, two varieties of Arabica and two varieties of Robusta. Dataset B and C comprised of Robusta and Arabica coffee at five different roasting stages. Recordings of a rubber conveyor belt served as the background class. All spectra were normalized to unit length.

3 Self-Organizing Feature Map with Conscience

The CSOM algorithm extends the distance calculation of classical Self-Organizing Feature Maps [9] by a neuron specific offset value g. This bias value g_i (the conscience of neuron i) is permanently adapted and holds the winning neuron off to win the competition for being the best match in a short time again [3]. As a result a magnification factor of $\rho = 1$ can also be reached for higher dimensional data [6]. The magnification factor ρ describes the relation between the point density $D()$ of the neuron prototypes (SOM weights ($\mathbf{w}$), receptive field centres) in the input space and the probability density function $P()$ of the stimuli data as $D(w) \propto P(w)^\rho$ [5, 6]. A magnification factor of $\rho = 1$ leads to a maximization of the implied information in the trained lattice and makes the CSOM well suited for handling sparsely presented clusters.

Except the *conscience* g of the neurons, identifying the winning neuron as minimum of the p-norm distance between the input stimulus and the receptive field center is equal to classical Self-Organizing Feature Maps (Eq. 1). The calculation specification for the individual bias value can be seen in Eq. 2. The parameter γ is one of three user parameters to influence the training data flow. All user controlled parameters (α, β, γ) have to decrease during the training in order to strengthen already learned informations.

$$\|\mathbf{x}(t)-\mathbf{w}_c(t)\|_p - g_c(t) < \|\mathbf{x}(t) - \mathbf{w}_i(t)\|_p - g_i(t) \ \forall i \neq c \tag{1}$$

$$g_i(t) = \gamma(t)\left(\frac{1}{l} - F_i(t)\right) \tag{2}$$

$$F_i(t) = F_i(t-1) + \beta(t-1)(y_i - F_i(t-1)) \tag{3}$$

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) + h_{c,i}(t)[\mathbf{x} - \mathbf{w}_i]; \quad h_{c,i}(t) = \begin{cases} \alpha(t) & \text{if } \|\mathbf{r}_c - \mathbf{r}_i\|_p \leq 1 \\ 0 & \text{if } \|\mathbf{r}_c - \mathbf{r}_i\|_p > 1 \end{cases} \tag{4}$$

The value l represents the number of neurons in the lattice and is constant during runtime. The other inner parameter next to g_i is the winning frequency F_i (Eq. 3) depending on β and the value y_i. The value y_i is 1 for the winning neuron w_c and 0 for all other neurons. As can be seen in equation Eq. 4, within a learning step the neighborhood function of the CSOM uses a constant α value which leads to a good-natured self organizing feature map for hardware acceleration.

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

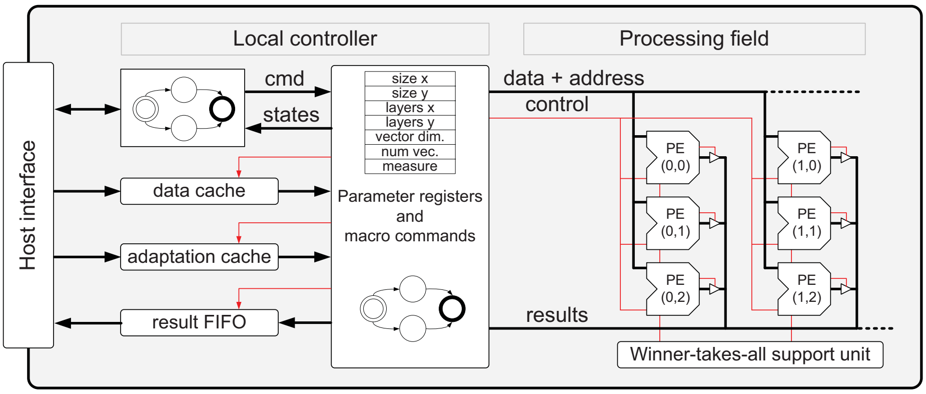

Fig. 1: Principle of the hardware accelerator consisting of a number of processing elements (PEs) to simulate the Self-Organizing Feature Map. The embedded local controller translates macro commands to control signals for the processing elements. An additional cache structure is used to store the incoming spectra from the camera and to provide the learning parameter.

4 Hardware Realisation

The realized hardware accelerator is capable for conscience SOM calculation as well as classical Self-Organizing Feature Map calculation. The hardware accelerator depends on the highly flexible gNBX core [10] extended by control logic for conscience and classical SOM, cache structures for data and adaptation value management and an additional communication interface to connect the accelerator to the host system (see Fig. 1). The gNBX core itself is a generic designed artificial neural net core, which is fully written in VHDL. Inside the core is a $n \times m$ lattice of massive parallel processing elements each capable to simulate one or more artificial neurons through resource sharing operations. The main features of the processing elements are the adjustable calculation accuracy and the embedded hardware multipliers for the $p1$ and $p2$ distance metric as well as the local memories for the synaptic neuron circuits and the pipelined data path structure.

Due to the high similarity between the inner parameter calculation shown in Eq. 2 and Eq. 3 and the neurons weights update calculation for classical Self-Organizing Maps (identical to Eq. 4), the conscience algorithm has been implemented by adding additional registers and by ensure correct signed calculations. The additional registers are necessary to store the user parameters β and γ and the neurons *conscience* g which causes the need of signed calculations. The permanently updated winning frequency is stored inside the local memory so that the needed memory space for one neuron increases to $dim_r + dim_w + 1$ instead of $dim_r + dim_w$ for classical SOM. Whereas dim_r is the dimension of the neuron lattice and dim_w is the number of synaptic circuits.

For the classification of a spectrum a calculation accuracy of 16 Bit is used with the ability of up to seven neurons per processing element (2048 local addresses) mapped on a single Xilinx Virtex4FX100 FPGA. Through this configuration up to 100 processing elements fit to the Virtex4 each using about $\approx$ 310 Slices, 1 embedded multiplier and 2 block ram memories. The work-

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

Table 1: Mean quantisation error and topographic error for the hard- and software implementation (training set); averaged across cross-validations; standard deviations in brackets.

	mean quantisation error			
Method	Map Size	Dataset A	Dataset B	Dataset C
HW CSOM	15x1	0.192 (0.003)	0.240 (0.007)	0.243 (0.008)
SW CSOM	15x1	0.184 (0.001)	0.225 (0.002)	0.226 (0.002)
HW CSOM	50x1	0.166 (0.001)	0.187 (0.001)	0.189 (0.002)
SW CSOM	50x1	0.167 (0.000)	0.183 (0.001)	0.184 (0.000)
	topographic error			
Method	Map Size	Dataset A	Dataset B	Dataset C
HW CSOM	15x1	0.010 (0.005)	0.039 (0.022)	0.036 (0.016)
SW CSOM	15x1	0.000 (0.000)	0.026 (0.007)	0.027 (0.015)
HW CSOM	50x1	0.065 (0.012)	0.077 (0.010)	0.076 (0.010)
SW CSOM	50x1	0.051 (0.007)	0.051 (0.007)	0.072 (0.018)

ing frequency of the hardware accelerator is 50 MHz using the RAPTOR-X64 Prototyping System [11] as test environment.

5 Machine Learning

For the CSOM, a post labelling approach yielded a class label per neural weight vector. The most frequent class label of all best matching data vectors per weight vector was kept. The CSOM was trained with a 1D map of 15 neurons as well as 50 neurons. A 2D map was trained as well but rejected due to its high topographic error. As comparison, a CSOM implementation with and without OpenMP optimization was run on an Intel Core i7 950 Quad Core processor at 3.07 GHz. For evaluation the topographic and quantisation error were calculated. Topographic error is the proportion of all data vectors for which first and second best matching units are not adjacent units.

The MLP and RBF were set with 20 neurons in the hidden layer and trained with conjugate gradient with momentum term and 1ofN coding of labels at their output. Standard GRLVQ and SRNG were trained with stochastic gradient and 4 neurons per class. Learn rate decreased over time. Finally a C-SVM with linear kernel was used as well. Other kernels were considered but yielded poorer accuracy.

Dataset A was a 5 class problem, datasets B and C were 6 class problems. For each class, 2000 samples where randomly chosen from the original data. Each dataset was divided according to a 10-fold cross-validation and 10 separate classification models were learned and their training and test accuracy averaged.

6 Results

In Table 1 the mean quantisation error and topographic error for the hardware (HW) and software (SW) CSOM are shown. Topographic error is small and quantisation error is comparable between map sizes. These results show that the hardware implementation with fix point arithmetic has created a comparable model in terms of quality to a software implementation.

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

Table 2: Test accuracy of classification; averaged across cross validations; standard deviation in brackets.

Method	Dataset A	Dataset B	Dataset C
HW CSOM 1x15	0.509 (0.014)	0.813 (0.019)	0.803 (0.026)
SW CSOM 1x15	0.506 (0.013)	0.826 (0.018)	0.803 (0.018)
HW CSOM 1x50	0.559 (0.019)	0.893 (0.015)	0.885 (0.015)
SW CSOM 1x50	0.555 (0.013)	0.904 (0.008)	0.891 (0.013)
RBF	0.962 (0.008)	0.946 (0.010)	0.917 (0.011)
MLP	0.435 (0.041)	0.722 (0.045)	0.661 (0.043)
GRLVQ	0.548 (0.022)	0.813 (0.025)	0.725 (0.016)
SRNG	0.684 (0.010)	0.812 (0.012)	0.873 (0.007)
SVM	0.969 (0.007)	0.966 (0.005)	0.944 (0.009)

Table 2 shows the classification results for all used classifiers. Average test accuracy and their standard deviation across the 10-fold cross-validation are shown. The RBF and SVM classifier show a robust performance across datasets. SRNG and GRLVQ show good results on Dataset B and C but poor results on Dataset A. The MLP shows average results across all Datasets. It is apparent that the CSOM performs equally in hardware and software at the level of the supervised approaches of the SRNG and GRLVQ at comparable map size of 1×15 and can increase performance with the 1×50 map in Dataset B and C.

Finally Table 3 compares the processing speed of the hardware CSOM with an efficient software implementation (with OpenMP) on general purpose hardware with map sizes of 1×15 and 1×50. Processing times cover only the calculation of the best matching unit from one data vector and the adaptation of all weight vectors. Data transfer time as well as the time for data pre- and post-processing were not considered and would depend on the final system realisation. In this paper we focus on the ability of the system to label the hyperspectral image pixelwise in tune with the maximum framerate possible. The HySpex SWIR-320m-e line camera produces 100×320 spectra per second. In order to label the image on-line, the theoretical processing time per spectra is $31.25\mu s$. The processing times in Table 3 show that a hardware CSOM and an efficiently implemented software CSOM, if it is run on a top of the line desktop processor, can be used to label as well as learn the Self-Organizing Feature Map on-line. The massive parallel CSOM hardware has the advantage that processing times did not grow with larger map size as they were for the software implementation.

7 Conclusion

We have shown that a CSOM implemented on a hardware accelerator, which can be integrated efficiently into a hyperspectral camera system, is suitable to classify spectral data in this case for coffee sorting purposes but still lacks a level of robustness to changing sorting tasks. Furthermore the on-line adaptation speed opens the possibility to teach the system on-line in order to build a model from the incoming data and label them to their identity. We compared the classification performance to a number of other prototype based neural network architectures and found that a Radial Basis Function Network approach showed the most robust results in the classification task. In future work, the hardware accelerator will be adopted to supervised approaches. The integration

ESANN 2012 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 25-27 April 2012, i6doc.com publ., ISBN 978-2-87419-049-0. Available from http://www.i6doc.com/en/livre/?GCOI=28001100967420.

Table 3: Processing time measurements for the Hardware CSOM and an efficient software implementation on a Intel Core i7 950 Quad Core processor at 3.07GHz; Massive parallel neural hardware shows its advantage at increasing map sizes; Processing time for one spectral vector; averaged across a 1000 samples.

Method	Map Dimension	best match (μs)	best match & weight update (μs)
HW CSOM	15x1	6.62	12.08
HW CSOM	50x1	6.62	12.08
SW CSOM	15x1	29.75	35.67
SW CSOM (OpenMP)	15x1	4.36	7.16
SW CSOM (OpenMP)	50x1	12.08	14.66

of efficient neural network processing hardware that is capable of dealing with the complexity of the high-dimensional spectral pattern adaptively to the task at hand will be important for the creation of smart camera systems. These can be used as stationary or mobile systems with applications in food inspection, medical diagnostics or smart farming.

Acknowledgements

We would like to thank Röstfein Magdeburg as well as Thomas Villmann and Marika Kästner for providing the recorded coffee samples.

References

[1] A.G. Fiore, R. Romaniello, G. Peri, and C. Severini. Quality assessment of roasted coffee blends by hyperspectral image analysis. In *In Proc. 22nd International Conference on Coffee Science*, Campinas, Brazil, 2008.

[2] A. Backhaus, F. Bollenbeck, and U. Seiffert. High-throughput quality control of coffee varieties and blends by artificial neural networks from hyperspectral imaging. In *Proceedings of the 1st International Congress on Cocoa, Coffee and Tea (CoCoTea)*, volume 1, pages 88–92, Novara, Italy, 2011.

[3] D. DeSieno. Adding a conscience to competitive learning. In *Proc. IEEE Int Neural Networks Conf*, pages 117–124, 1988.

[4] D.R. Dersch and P. Tavan. Asymptotic level density in topological feature maps. *IEEE Transactions on Neural Networks*, 6(1):230–236, Jan 1995.

[5] H.-U. Bauer, R. Der, and M. Herrmann. Controlling the magnification factor of self-organizing feature maps. *Neural Computation*, 8(4):757–771, 1996.

[6] E. Merenyi, A. Jain, and T. Villmann. Explicit magnification control of self-organizing maps for "forbidden" data. 18(3):786–797, 2007.

[7] B. Hammer and T. Villmann. Generalized relevance learning vector quantization. *Neural Networks*, 15:1059–1068, 2002.

[8] B. Hammer, M. Strickert, and T. Villmann. Supervised neural gas with general similarity measure. *Neural Processing Letters*, 21:21–44, 2005.

[9] T. Kohonen. *Self-Organizing Maps.* New York : Springer-Verlag, 1997.

[10] C. Pohl, M. Franzmeier, M. Porrmann, and U. Ruckert. gNBX - reconfigurable hardware acceleration of self-organizing maps. In *Proc. IEEE Int Field-Programmable Technology Conf*, pages 97–104, 2004.

[11] M. Porrmann, J. Hagemeyer, C. Pohl, J. Romoth, and M. Strugholtz. RAPTOR a scalable platform for rapid prototyping and FPGA-based cluster computing. In *Proceedings of the International Conference on Parallel Computing, ParCo2009, Symposium on Parallel Computing with FPGAs*, Lyon, France, 1 - 4 September 2009.