

# A Fully Neural Approach to Image Segmentation and Image Coding

## Abstract

A fully neural approach to image segmentation, image coding and feature analysis of facial images is presented. An algorithmic quadtree image decomposition method is used to generate training data for a group of multi-layer-perceptrons (MLP), which are trained with a standard error-backpropagation rule. The generated segmented image is equivalent to its quadtree representation and is used as input for following facial feature extraction and variable-block-size vector quantization (VBSVQ) steps. The codebook generation for the VBSVQ image coder and the image coding itself is done by a Kohonen feature map. The extracted facial image features, e.g. eyes, nose and mouth, may serve as input data for new coding techniques like semantic coding or techniques based on motion estimation.

## 1 Introduction

A very important field of digital image transmission is face-to-face image transmission using (digital) telephony lines. Due to the low bitrate capacity of these lines (ISDN 64 kbit/s) sophisticated coding methods are required. The purpose of this paper is to present an image coder based on image segmentation combined with vector quantization in a fully neural environment. Such an approach offers significant advantages both in performance and implementation complexity. An additional application based on the segmentation step is the extraction of important facial features which is done by neural networks, too. This facial feature recognition is done by the investigation of a binary representation of the source image different to the approach of Vincent [4], which uses fine resolution feature locators to investigate the original grayscale picture. The cascable VLSI-system BACCHUS developed at Darmstadt University of Technology (DUT) seems to be a suitable hardware implementation for the facial feature recognition [1,2,3]. This neuron like binary associative memory serves as hardware implementation in the fields of industrial quality control, handwritten letter recognition and image compression. In prior work an image coding with the procedure of vector quantization has been implemented successfully [1,2]. The resulted feature representation of the facial image is the starting point for the recognition of the whole face, which is necessary for new 3rd generation coding schemes like semantic or 3-D model based coding [9].

## 2 Quadtree image segmentation

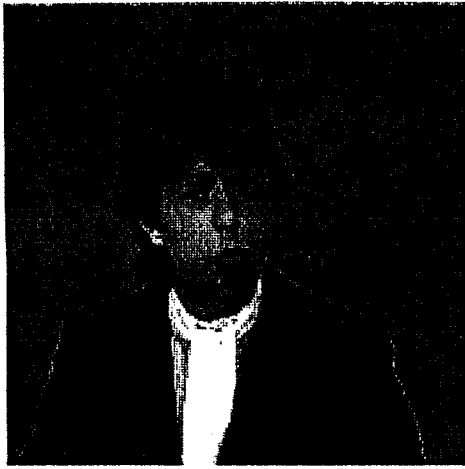
A quadtree represents a digital image of a square region by hierarchically partitioning the region into quadrants and subquadrants until all subquadrants are uniform with respect to image value. This approach to image representation was first proposed by Klinger [5] and is very popular in the field of pattern recognition, image analysis and image coding [6,7,8]. As partitioning condition the tolerated variation of brightness within a single region is very common. A very important property of the quadtree image segmentation algorithm is the preservation of edges and detailed structures on the one hand and the merger of homogeneous areas to larger regions on the other hand. Especially this adaptive property predestines the quadtree image segmentation as good preprocessing candidate for the VBSVQ. The sizes of the finally created square regions are dependent on their positions in the quadtree. A region at tree-level  $l$  has the size of  $2^l * 2^l$  pels. The quadtree segmentation algorithm is implemented in the programming language C to create training data for the image segmentation neural networks.

## 4 Simulation results

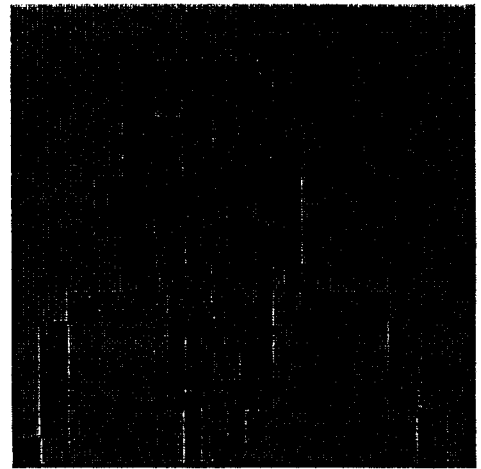
Figure 2 shows simulation results of the segmentation and VBSVQ processes. Starting with the 256\*256\*8 bits picture CLAIRES.LUM (a) Figure 2(b) is the result of a quadtree segmentation process done by four individually trained MLPs. Each MLP is responsible for the generation of 16\*16, 8\*8, 4\*4 and 2\*2 pels image blocks. Figure 2 (c-h) is a step-by-step demonstration of the VBSVQ. The segmented image (d) describes the location of the different sized image blocks. Each block is assigned to the best matching codebook-block. The generation of the different sized codebook-blocks and the assignment of the best matching blocks is done by a Kohonen Feature Map. The compression ratio of the shown 128\*128\*8 bits Picture is 14:1, the SNR is about 31 dB. The VBSVQ applied to the 256\*256\*8 bits Picture CLAIRES.LUM realizes a compression ratio of about 36:1 preserving the same good visual image quality.

## 5 Conclusion and Future Aspects

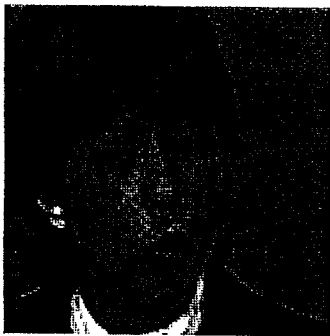
A neural image segmentation system has been presented. The segmented image is equivalent to its quadtree representation and serves as starting point for image coding and image analysis. The adaptive VBSVQ procedure offers increased compression rates *and* decreased image errors in comparison with the fixed block size VQ. It has to be mentioned that this system realizes *intraframe* coding, but an implementation of the presented system in a hybrid coder similar to the MPEG or CCITT H.261 coder is possible. Future work will concentrate on this extension to interframe coding, coding schemes based on face detection and the application of the available binary associative memory BACCHUS in the field of face detection.



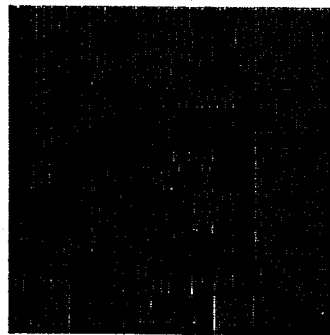
(a)



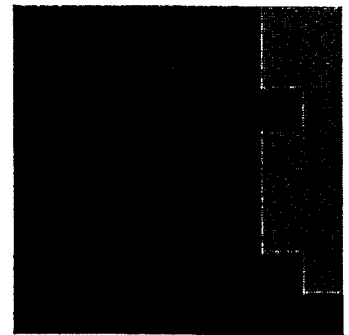
(b)



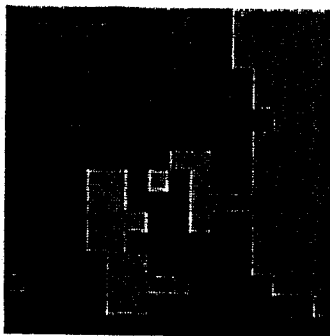
(c)



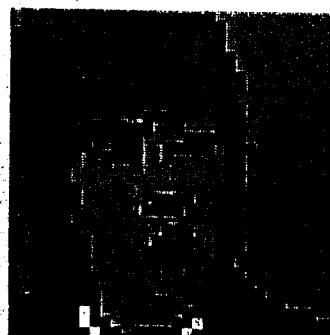
(d)



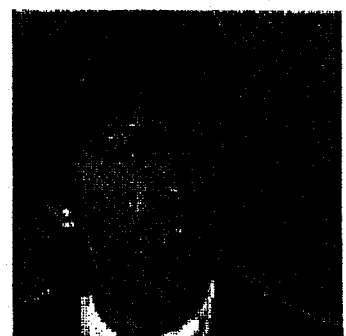
(e)



(f)



(g)



(h)

Figure 2: Simulation results

## References

- [1] A. Koenig, M. Glesner,  
"An Approach to the Application of Dedicated Neural Network Hardware for Real Time Image Compression",  
ICANN 91, Helsinki, 24-29 June 1991
- [2] A. Koenig, W. Poechmueller, M. Glesner,  
"A Flexible Neural Network implemented as a Neural Coprocessor to a Von Neumann Architecture", *Proc. of the 2nd Int. Conf. on MICROELECTRONICS FOR NEURAL NETWORKS*, Oct. 16-18, Munich, 1991
- [3] W. Poechmueller, M. Glesner,  
"A Cascadable VLSI Architecture for the Realization of Large Binary Associative Networks", *International Workshop on VLSI for Artificial Intelligence and Neural Networks*, Oxford, 1990
- [4] J.M. Vincent,  
"Facial Feature Location in Coarse Resolution Images by Multi-Layered Perceptrons", ICANN 91, Helsinki, 24-29 June 1991
- [5] A. Klinger and M.L. Rhodes,  
"Organization and access of image data by areas",  
*IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-1, pp. 50—60, Jan. 1979
- [6] S.L. Horowitz and T. Pavlidis,  
"Picture Segmentation by a Tree Traversal Algorithm", *Journal of the Association for Computing Machinery*, Vol. 23, pp. 368—388, Apr. 1976
- [7] H. Samet,  
"Computing Geometric Properties of Images Represented by Linear Quadrees", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-7, Feb. 1985
- [8] Y. Cohen, M. Landy and M. Pavel,  
"Hierarchical Coding of Binary Images", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-7, May 1985
- [9] R. Forchheimer and T. Kronander,  
"Image Coding—From Waveforms to Animation",  
*IEEE Trans. Acoust. Sp. Signal Proc.*, Vol. 37, pp. 2008—2021, Dec. 1989
- [10] T. Kohonen,  
*Self-Organization and Associative Memory*, Third Edition, Springer 1989
- [11] D.E. Rumelhart, G.E. Hinton and R.J. Williams,  
"Learning internal representations by error propagation", in *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Vol.1, MIT Press, pp. 318—362