A method for linking computed image features to histological semantics in neuropathology

B. Lessmann [a,b] , T.W. Nattkemper [b], V.H. Hans [c], A. Degenhard [a],

[a] *Theoretical Physics Department, University of Bielefeld, Germany*

[b] *Applied Neuroinformatics, University of Bielefeld, Germany*

[c] *Institute of Neuropathology, Evangelisches Krankenhaus Bielefeld, Germany*

Abstract

In medical image analysis, the image content is often represented by computed features that need to be interpreted at a clinical level of understanding to support the development of clinical diagnosis systems. Many features are of abstract nature, as for instance features derived from a wavelet transform. The interpretation and analysis of such features is difficult. This lack of coincidence between computed features and their meaning for a user in a given situation is commonly referred to as the *semantic gap*. In this work, we propose a method for feature analysis and interpretation based on the simultaneous visualization of feature and image domain. Histopathological images of meningioma WHO (World Health Organization) grade I are firstly color transformed and then characterized by features derived from the Discrete Wavelet Transform. The wavelet-based feature space is then visualized and explored using unsupervised machine learning methods. Our approach allows to analyze and select features regarding their relevance for the description of clinically relevant characteristics.

Key words: Self Organizing Map, Discrete Wavelet Transform, feature selection, Histopathology, semantic gap
PACS:

1 Introduction

In medical image analysis an increasing amount of processing tools is developed to assist clinical experts in diagnostic processes, including techniques for

Email address: lessmann@physik.uni-bielefeld.de (B. Lessmann).

automated classification or content based image retrieval. In such applications the image content is characterized by numerical image features. It is challenging to decide which features are the most appropriate ones for the specific purpose considered.

The most common method is to analyze the features regarding their discriminative power for the specific purpose. In case of automated classification one can e.g. perform classification on all possible subset of features and then select the feature set providing the highest classification score. However, this method of feature analysis strongly depends on the specific type of application, e.g. the applied classification methods, the size of the database and so on [10]. Furthermore it is still a black-box scenario, i.e. it does not provide an interpretation of the features in the image domain. While more simple and direct features such as color histograms are easy to interpret in the image domain, features based on transforms such as Fourier or Wavelet Transform are much more difficult to understand. In general this marks a crucial point of criticism by the physicians, who are interested in features or feature sets that can be interpreted to some degree at a clinical level of understanding. It is therefore desirable, to develop methods for linking numerical features and feature sets to clinical semantics.

The phenomenon of missing interpretability of numerical features is also called the *semantic gap* which is defined according to [22] in the following way: *The semantic gap is the lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data have for a user in a given situation.* Recently, the semantic gap has been discussed in the context of image retrieval. In [21] images are annotated using medical key words in order to improve retrieval performance. In [23] semantic metadata is extracted from visual descriptors by classifying images into semantic categories and organize a database based on concepts.

In this work we address the problem of the semantic gap regarding the interpretation of wavelet-based image features. To link local, morphological image characteristics to the space spanned by the wavelet-based features a Self Organizing Map (SOM) was employed. As a method of unsupervised learning the SOM is a powerful tool providing both the ability of clustering and data visualization.

SOMs have been considered to serve as interactive visualization tools e.g. regarding database visualization and browsing [6], content-based image retrieval [13] or building of a texture dictionary [17]. One application of SOMs on histological datasets is the investigation of human defined image characteristics as described in [9]. In that work the features used for clustering are based on human definition and rating, i.e. the occurrence of a specific histological characteristic (e.g. collagen fibers) has been graded on a scale of four

by a human observer.

Here, we demonstrate the ability of SOMs to serve as an interface for the analysis of numerical features. We demonstrate the method utilizing an example database of microscopic images of benign brain tumors. The database contains histopathological images of four subtypes of meningiomas. The subtypes are classified by a medical expert into four meningioma classes depending on textural characteristics at different scales. Therefore we have chosen from the variety of methods for texture characterization the Discrete Wavelet Transform as a method which allows to encode scale-dependent image information.

The SOM-based visualization of the feature space then enables us to establish a correlation between single features and histologically relevant image structures and thus the selection of a subset of clinically important features.

Preliminary experiments have already been described in [14] which suffered from a lack of interpretability. Utilizing a new color transform we can now show, how our approach can be used to bridge the gap between numerical features and histopathological terms and thus transfer clinical terms into the feature space.

2 Materials

In order to provide a test set of data for our system, we have chosen four histopathologically well defined subtypes of meningioma, a benign tumor of the coverings of the brain, i.e. the meninges [16]. The four subtypes are characterized by distinct features (Table 1) allowing a trained investigator to make an unequivocal diagnosis in most cases. Examples of subimages (256×256 pixel) are shown in Figure 1. However, intermediate features are especially common between fibroblastic, transitional and psammomatous subtypes.

Diagnostic tumor samples were derived from neurosurgical resections at the Bethel Department of Neurosurgery, Bielefeld, Germany for therapeutic purposes, routinely processed for formalin fixation and paraffin embedded. Four μm thick microtome sections were dewaxed on glass slides, stained with Mayer's haemalaun and eosin, dehydrated, and coverslipped with mouting medium (Eukitt®, O. Kindler GmbH, Freiburg, Germany). Archive cases from the years 2004 and 2005 were selected for representing typical features of each meningioma subtype. Slides were analyzed on a Zeiss Axioskop 2 plus microscope with a Zeiss Achroplan 40x/0,65 lens. After manually focusing and automated background correction, 1300×1030 pixels, 24 bit, true color RGB pictures were taken at standardized 3200 K light temperature in TIF format using Zeiss AxioVision 3.1 software and a Zeiss AxioCam HRc digital color

Subtype	Characteristics
Meningotheliomatous	Lobulated, cells forming a syncytium
Fibroblastic	Spindle-shaped cells, matrix abundant in collagen
Transitional	Whorls, few psammoma bodies, features of fibroblastic
Psammomatous	Transitional appearance, abundant psammoma bodies

Table 1
A summary of the histological features for the four tumor classes.

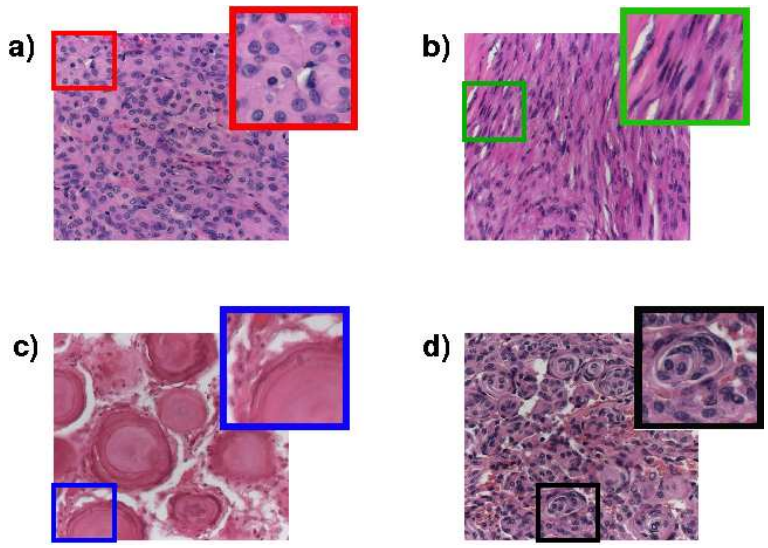


Fig. 1. Example subimages for the different subtypes of meningiomas: a) - meningotheliomatous, b) - fibroblastic, c) - psammomatous, d) - transitional.

camera (Carl Zeiss AG, Oberkochen, Germany). Five cases were selected for each diagnostic group and four different photomicrographs were taken of each case, resulting in a set of 80 pictures. Each original picture was truncated to 1024×1024 pixels and then subdivided in a 4×4 subset of 256×256 pixel pictures. This resulted in a database of 1280 subimages for further analysis.

3 Methods

3.1 Color channels

In many applications RGB (Red, Green, Blue) images are transformed into a color space more suitable for human perception, i.e. the HSV (Hue, Saturation, Value) color space [11] or the L*u*v color space [5]. Since pathology images are limited regarding their colors we transform the RGB values in order to enhance special image structures. In the following, we have computed two transformed images from the RGB values (R(x,y), G(x,y), B(x,y)). Firstly, an intensity value is computed by averaging the three color channels for each pixel and location (x, y) according to

$$h_1(x, y) = \frac{1}{3} * [\mathrm{R(x, y)} + \mathrm{G(x, y)} + \mathrm{B(x, y)}]. \tag{1}$$

The images computed according to equation (1) will further be denoted as *intensity images*. Secondly, a transform in color space is designed to extract the cell nuclei from the image, since they appear to exhibit significant characteristics for distinguishing the tumor classes. However, the image color of a single stained tumor section inevitably depends to some extent on changes during the preparation. To reduce dependence on these color changes, we apply a mean shift to all images and color channels. The average color, in the following indexed by av, is supposed to be very close to the color of those structures providing the largest areas in the image, in this case the cytoplasm or the psammoma bodies.

$$\mathrm{R_{shift}(x, y) = R(x, y) - R_{av}} \tag{2}$$
$$\mathrm{G_{shift}(x, y) = G(x, y) - G_{av}} \tag{3}$$
$$\mathrm{B_{shift}(x, y) = B(x, y) - B_{av}} \tag{4}$$

The preparation procedure described above using the routine H&E stain leads to a blue coloration of the cell nuclei in contrast to the surrounding cytoplasm usually showing a pink color. We point out that this holds for the dataset analyzed since it contains only WHO grade I meningioma. Some other types of meningioma, e.g. the clear cell meningioma (WHO grade II) provide a coloring different from the one described.
After applying the mean shift to the color channels, those image structures, which are "bluer" than the surrounding tissue, are characterized by $\mathrm{B_{shift}(x, y)} > \mathrm{R_{shift}(x, y)}$.

By computing

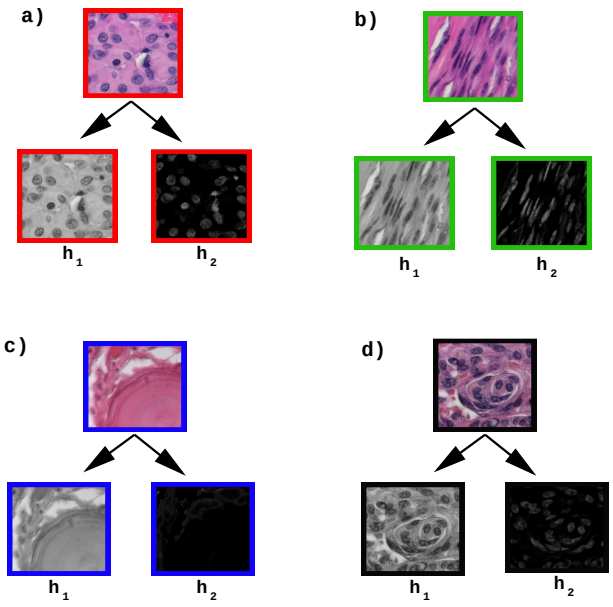


Fig. 2. Example subimages for the transform in color space. Each of the subimages already shown above is displayed as intensity image h_1 and with enhanced cell nuclei h_2: a) - meningotheliomatous, b) - fibroblastic, c) - psammomatous, d) - transitional.

$$h_2(x,y) = (max_{\mathrm{RB_{shift}}}(x,y) - \mathrm{R_{shift}(x,y)}) * \mathrm{S} \tag{5}$$

$$\text{with} \quad max_{\mathrm{RB_{shift}}}(x,y) = max\{\mathrm{R_{shift}(x,y)}, \mathrm{B_{shift}(x,y)}\} \tag{6}$$

all image structures with $\mathrm{B_{shift}(x,y)} < \mathrm{R_{shift}(x,y)}$ are set to zero. The remaining structures, which are supposed to be mainly cell nuclei, are retained.

The factor S is the saturation as defined in the HSV color space [7]

$$\mathrm{S} = \frac{max_{\mathrm{RGB}}(x,y) - min_{\mathrm{RGB}}(x,y)}{max_{\mathrm{RGB}}(x,y)} \tag{7}$$

$$\text{with} \quad max_{\mathrm{RGB}}(x,y) = max\{\mathrm{R(x,y)}, \mathrm{G(x,y)}, \mathrm{B(x,y)}\}. \tag{8}$$

This factor suppresses artefacts occurring from the white areas in the images. Especially in psammomatous meningiomas the white areas are due to artificial cracks during tissue processing and do not represent relevant staining properties of the tumors. The resulting images of the transform h_2 will be denoted as *cell nuclei images* in the further description. Figure 2 shows examples of the transform.

3.2 Discrete Wavelet Transform

In this work, wavelet based features are used for tissue characterization. Wavelet based multiresolution analysis enables to assess the scale-dependent informa-

tion in signals and images [19]. A signal f is decomposed using the *Discrete (Dyadic) Wavelet Transform* into a basis of shifted and dilated versions of a basic wavelet or *mother wavelet* ψ [4]

$$f(x) = \sum_{(j,k)} d_{j,k}\psi_{j,k}(x) \ , \tag{9}$$

$$\text{with} \qquad \psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k) \ . \tag{10}$$

Here the index j indicates the dilation or *scaling step* while k refers to translation or shifting. The wavelet coefficients $d_j(k)$ are given by the scalar product $d_j(k) = \langle f(x), \psi_{j,k}(x)\rangle$ or $d_j(k) = \langle f(x), \tilde{\psi}_{j,k}(x)\rangle$ in case of *biorthogonal wavelets* with the dual wavelet $\tilde{\psi}$ [4]. An efficient calculation of these coefficients is accomplished by the *Fast Wavelet Transform* (FWT), an algorithm allowing the coefficients to be calculated in a stepwise manner. On the first scale the signal is decomposed into its *details* and the remaining signal, i.e. the *approximation*, reflecting the particular scale. The details are described by the wavelet coefficients of this scale while the approximation is represented by the scaling function coefficients. The procedure can be iterated by a further decomposition of the approximation into details and approximation of the next coarser scale [18].

In two dimensions, the wavelets are product functions of the one-dimensional wavelet- and scaling-function. We have chosen to use the decomposition procedure, which is called *Nonstandard-Decomposition* in [24]. The corresponding coefficients $d_{j,o}(k_x, k_y)$ therefore do not only reflect a particular scale j, but also a particular orientation o in the images. The two-dimensional wavelet functions to be considered are $\psi(x)\phi(y)$, $\phi(x)\psi(y)$, $\psi(x)\psi(y)$ and $\phi(x)\phi(y)$. The coefficients corresponding to the latter one are the two-dimensional scaling function coefficients, further decomposed in the next step, while the coefficients corresponding to the first functions provide the details in horizontal (x) direction, vertical (y) direction and the diagonal details.

3.2.1 Wavelet-based image features

Texture is probably the most important descriptor for tissue characterization in medical textbooks. Several different texture features have been evaluated in literature, including e.g. co-occurrence matrices, line-angle-ratio statistics [1,8], Gabor filters [8] or wavelet based features [26]. A comprehensive overview of texture features can be found in [20]. In this work the texture characterization does not involve a prior segmentation as e.g. in [5] but is performed on the entire subimages. To access the scale-dependent information we decided to use multiresolution wavelet features, as calculated by the discrete wavelet transform. We use a pair of symmetric, biorthogonal wavelets, developed by

Cohen, Daubechies and Feauveau. These pairs of scaling and wavelet-functions are indicated as $\phi_2, \tilde{\phi}_{2,2}, \psi_{2,2}, \tilde{\psi}_{2,2}$ in [3]. As described in [2] and [25] the l_1 or l_2-norm of the detail coefficients corresponding to one scale and orientation can be used as powerful texture features. In this work the l_1-norm, i.e. the mean absolute coefficient (MAC) of each scale and orientation has been computed according to $l_1(d_{j,o}) = \sum_{k_x,k_y} |d_{j,o}(k_x, k_y)|$. The final feature vectors are based on these MACs. As shown in [14] two types of features based on these MACs were found to be appropriate to characterize the four types of tissue. Firstly, a mean absolute coefficient for each scale j is computed according to

$$f_1(j) = \sum_o \mathrm{MAC}(o, j), \quad j = 1..8, \quad o = o_1, o_2, o_3. \tag{11}$$

Again j is the scale index, while index o indicates the orientation in the image. The indices o_1 and o_2 indicate coefficients in vertical or horizontal direction, while o_3 indicates the diagonal details.

Secondly, we use a feature set f_2 describing whether the image structures are anisotropic, i.e. orientated in a preferred orientation.

$$f_2(j) = |\mathrm{MAC}(o_1, j) - \mathrm{MAC}(o_2, j)| + c\,\mathrm{MAC}(o_3, j), \quad j = 1..8 \tag{12}$$

In case of image structures mainly orientated in horizontal or vertical direction, either the MAC for o_1 or o_2 should be significantly increased, while the other one is correspondingly decreased. Therefore $|MAC(o_1, j) - MAC(o_2, j)|$ reaches a high value. Regarding image structures orientated in a diagonal direction the first part of equation (12) vanishes, but, at the same time, the second part increases. The normalization factor c assures that the added MACs have equalized variances such that the influences of both parts of the sum are comparable. In contrast to [14] each component of the second feature set is not normalized to the corresponding component of the first set. We point out, that a similar approach to characterize images by scale- and orientation-related texture measures using wavelets has already used for the characterization of corrosion images [15]. However, there the computed features used differ from the ones we computed in this work.

3.3 Self Organizing Map

The Self Organizing Map (SOM) is a clustering approach from the field of artificial neural networks [12]. A set of reference or prototype vectors $\{\mathbf{u_j}\}, \mathbf{u_j} \in \mathbb{R}^n$, is trained according to a given data set of feature vectors $\{\mathbf{x_i}\}, \mathbf{x_i} \in \mathbb{R}^n$. Here, n is the number of features. The SOM is represented by a two-

dimensional square grid, which consists of nodes each associated to one of the reference vectors. During each learning step t a vector from the input space $\mathbf{x}_k$ is selected and the nearest reference vector $\mathbf{u}_m$ (winner node) is identified by $|\mathbf{u}_m - \mathbf{x}_k| = \min_j |\mathbf{u}_j - \mathbf{x}_k|$. The nodes are updated according to $\mathbf{u}_j(t+1) = \mathbf{u}_j(t) + h_{mj}(t)[\mathbf{x}_k - \mathbf{u}_j(t)]$. The function h_{mj} is defined as $h_{mj} = \alpha(t) \exp -\frac{|r_m - r_j|^2}{2\sigma^2(t)}$, where $\alpha(t)$ is the learning rate factor. The vectors $\mathbf{r}_m$ and $\mathbf{r}_j$ are the coordinates of the nodes in the SOM grid, with the winner node m. The function $\sigma(t)$ also decreases with time in order to reduce the neighborhood included in the training procedure with time. The decreasing of both functions freezes the SOM after a given number of iterations [12]. After a successful training procedure the reference vectors depict the data distribution in the data set preserving the topology.

3.3.1 SOM-based exploration procedure

The SOM is applied to visualize and explore a database content based on specific image features. Since the images have a size of 256×256 pixels, features of eight scales can be computed. The features of the coarsest two scales are neglected, since the associated coefficients encode details corresponding to the total or a quarter image, which does not seem to be senseful. Considering the two sets of features (f_1, f_2), the two color channels (h_1, h_2) described above and the remaining six scales a total number of 24 possible features has to be taken into account.

The training procedure has been accomplished two times, firstly using the first set of features f_1 (12 feature vector components), secondly using the second set of features f_2 (12 feature vector components). The results are shown in the Figures 3 and 4 respectively. An 8×8 SOM grid has been utilized for the exploration procedure. In each figure three types of SOM visualizations are presented. At the top the clustering result is shown. For each feature vector, i.e. each subimage, the nearest reference vector of the trained SOM can be computed. In this way, each subimage is mapped to one node in the SOM grid. In the visualization at the top each subimage is symbolized by a point of the color corresponding to the class of tissue. This point is displayed at the specific node, thus each node is visualized by the corresponding cluster of feature vectors mapped to this node. In the middle the same SOM grid is visualized in another mode. Here each node is marked with one of the subimages of the respective cluster. To be explicit, the subimage associated with the nearest feature vector has been chosen from the dominating tumor class at the specific node. In this way, the distribution of histological features can be explored. At the bottom of the figure the Component Plane Maps are shown, a special visualization of the reference vector components. Each little square is a visualization of the SOM grid and one component of the reference vectors. The gray value shows the distribution of this component on

the SOM grid. Since the reference vectors have twelve components in both cases, twelve Component Plane Maps are shown, one map for each component of the reference vector. Black color symbolizes low values of the component, white color indicates high values.

3.3.1.1 Feature set f_1 Starting with the analysis of Figure 3, the following conclusions can be derived. The clustering results visualized at the top shows that this set of features is appropriate to discriminate psammomatous (blue) and transitional (black) images. However, the meningotheliomatous (red) and the fibroblastic (green) classes are still strongly mixed. This is especially due to the feature vector components $1-6$, which are the features of the intensity image h_1. In the Component Plane Maps these components vary strongly from the top to the bottom of the SOM grid. By exploring the corresponding subimages it becomes clear, that these components are linked to tissue inhomogeneities in the extracellular matrix, which occur in both classes. Furthermore the high scale component (component six) additionally introduces an inner-class separation in the psammomatous group of tissue (from top right to bottom right). The visualization of histological characteristics reveals that this component is especially high in those images showing large cracks in the tissue (bottom right of the SOM grid). Since the amount of cracks in the tissue varies strongly in psammomatous tissue it is not considered to be of diagnostic relevance. We will therefore neglect the components $1-6$ of feature set f_1. The visualization further reveals some redundancy in the components $7-12$ of this feature set. These components show a very similar distribution. Some of them can therefore also be neglected. From the feature set f_1 only the components $9-10$ are chosen for tissue characterization. These components correspond to the scales 3 and 4 of the cell nuclei image h_2.

3.3.1.2 Feature set f_2 As described above this set of features is constructed to describe preferred orientations in the image. As expected the separation of the fibroblastic and the meningotheliomatous class as visualized at the top of the figure is significantly increased compared to feature set f_1. However, the Component Plane Maps reveal that the components 6 and 12 associated to very coarse scale details do not show a distribution corresponding to some histological interpretation and are therefore neglected. The remaining components show strong redundancy, so in a next step we only choose the components $3, 4, 9, 10$ for tissue characterization. These components correspond to the scales three and four of both color channels h_1 and h_2. The six selected features are summarized in table 2.

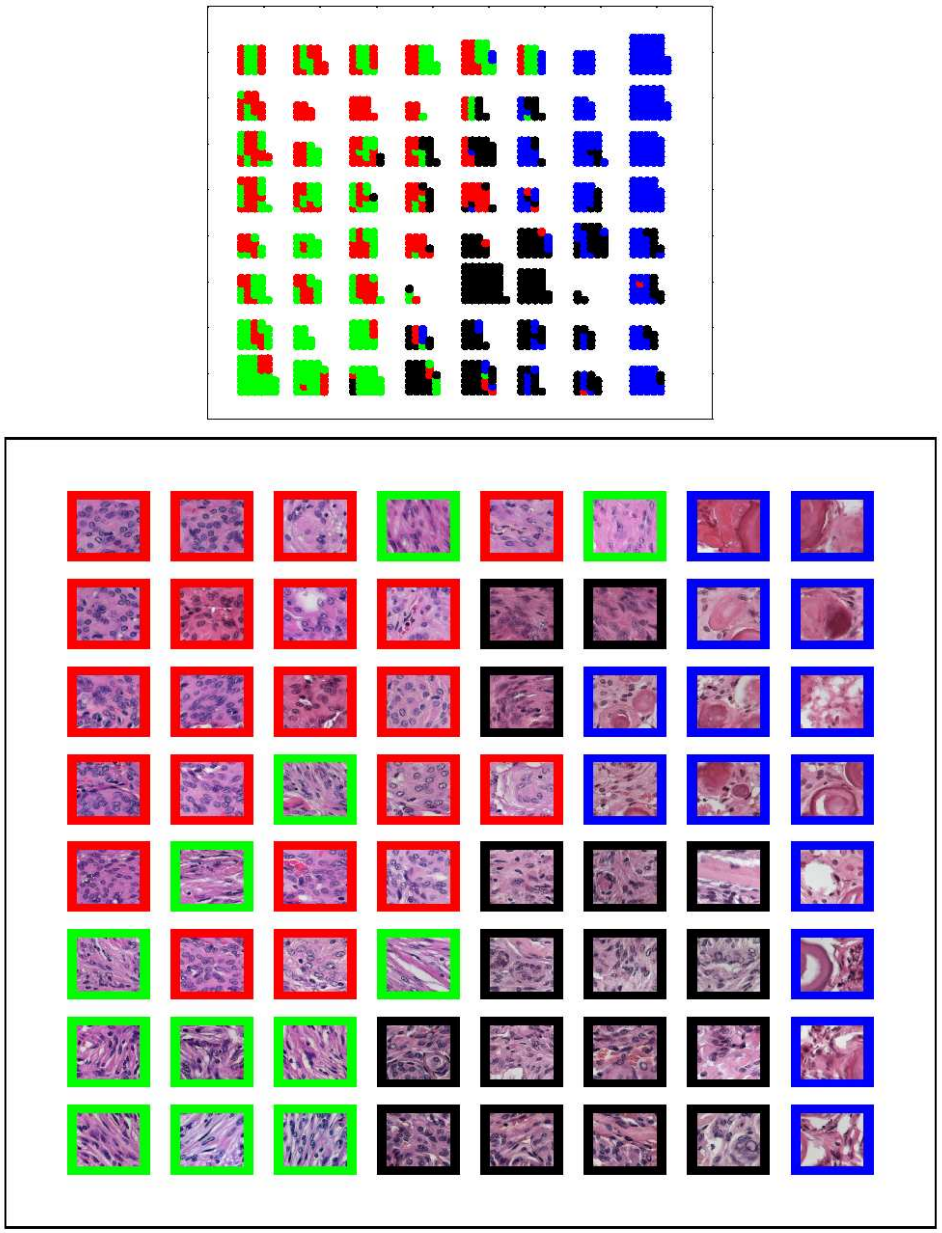


Fig. 3. Visualization of the SOM training result based on the feature set f_1. Top: Visualization of cluster structure. Center: Visualization of histological characteristics. Bottom: Component Plane Maps.

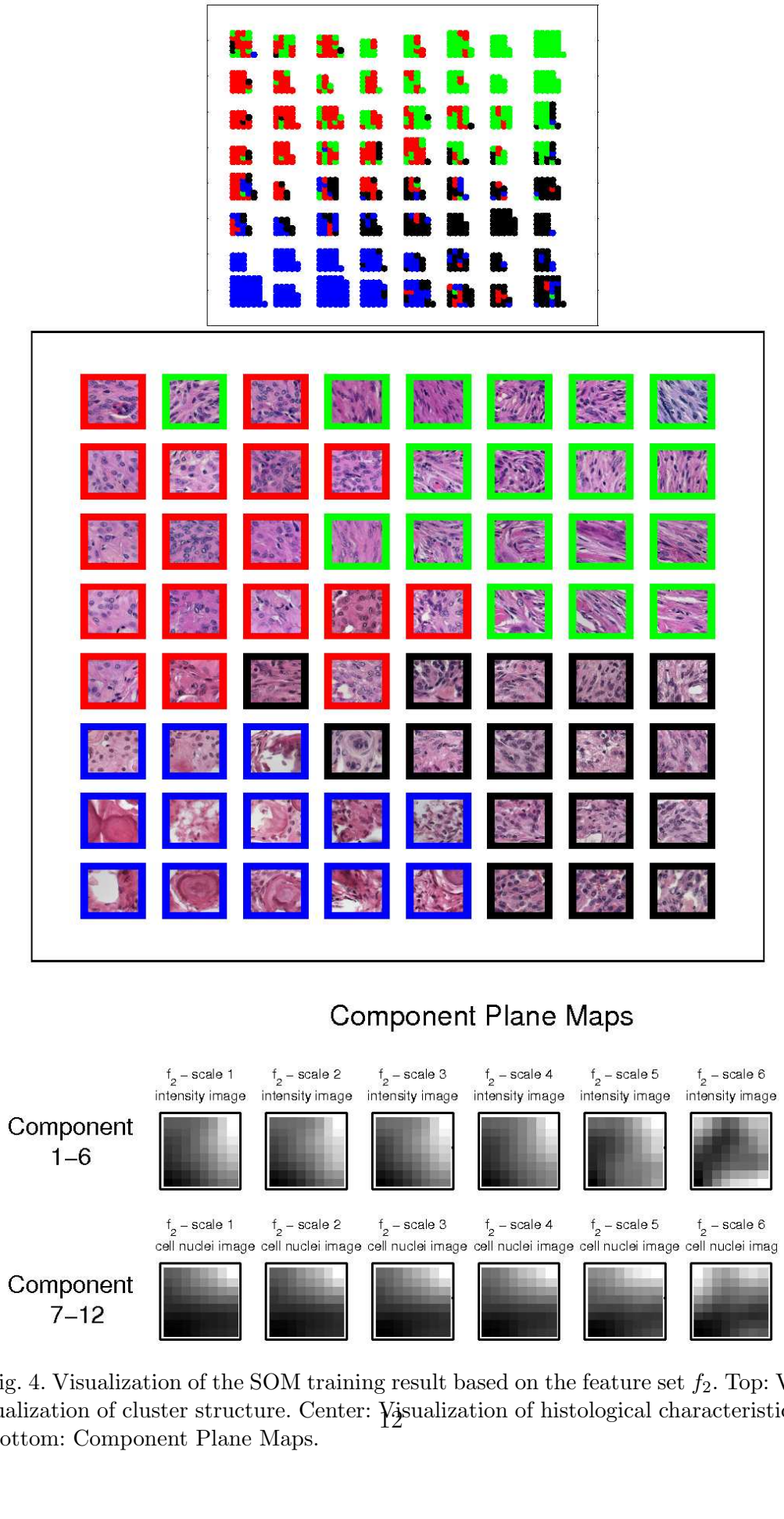


Fig. 4. Visualization of the SOM training result based on the feature set f_2. Top: Visualization of cluster structure. Center: Visualization of histological characteristics. Bottom: Component Plane Maps.

set of features	h_1	h_2
f_1		scale 3 & 4
f_2	scale 3 & 4	scale 3 & 4

Table 2
Final set of features employed for tissue characterization.

4 Results and Discussion

In the foregoing section a subset of six features is selected from a total set of 24 features. Here, the SOM training based on these selected features is repeated. The result is shown in Figure 5. Obviously, although only six features have been selected from a group of 24 features, these features are clearly appropriate for a good separation of the four classes of tissue. Due to a feature selection based on histological interpretation, the components of the reference vectors can be clearly linked to particular image structures at a clinical level of understanding. This is shown in Figure 6. This Figure shows the reference vectors associated to each node as a bar plot. Due to the visualization in the image domain in Figure 5, we are able to link the characteristics of the reference vectors to the clinically relevant image structures, which is shown in the table beneath. For example low values of all components of the feature vector are characteristic for psammoma bodies while high values in all components of the feature vector are typical for fibroblastic tissue, representing orientated image structures and a significant amount of cell nuclei.

To prove that the subset of features selected above is more appropriate for class separation than the combination of all possible features, a measure rating the class separation in the feature space is computed. Firstly, the center of each class c in the feature space $\bar{\mathbf{x}}_c$ and the average distance of each class center from the centers of the remaining three classes $dist(c)$ is determined. Secondly,the inner-class variance of each class σ_c is determined. All values are normalized to the length of the feature vectors. This normalization assures, that the results of feature vectors providing different dimensions are comparable. A good separation of one class from the others is achieved, if the distance of the class center to the other class centers is high and the inner-class variance is small. As a measure of the separation we therefore compute the ratio $r_c = \frac{dist(c)}{\sqrt{\sigma_c}}$. The higher this ratio the better the class separation. In Table 3 the ratios r_c for the four classes based on different subsets of features are shown. The selected subset of six features clearly outperforms the other combinations of features.

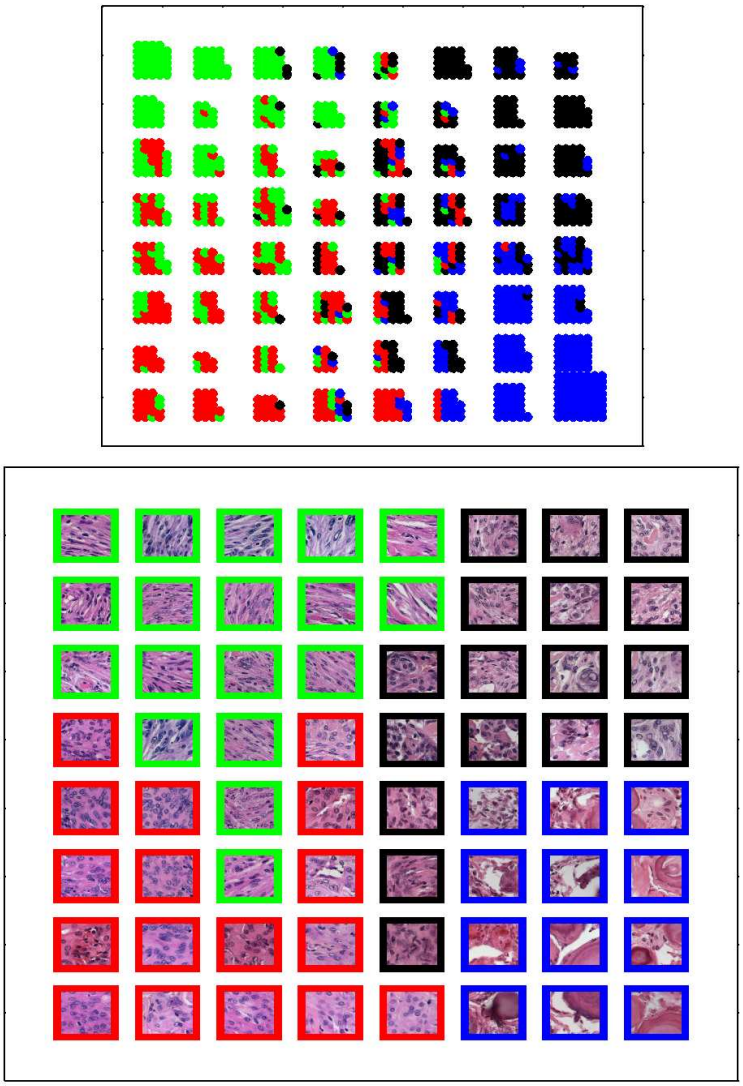
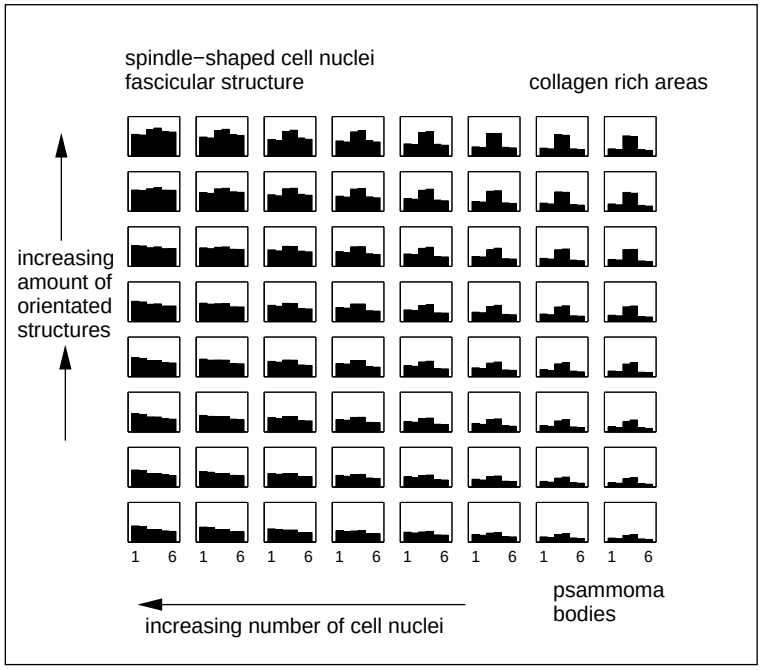


Fig. 5. Result of the SOM training based on the selected features. Top: Visualization of cluster structure. Bottom: Visualization of histological characteristics.

5 Conclusions

We presented a framework for the semantic exploration of a wavelet-based feature space. The introduction of a pre-processing step to enhance particular image structures and the simultaneous visualization of the feature and image space provides a method for the exploration of correlations of clustering results in the feature space to histological image characteristics.

The visualization methods presented allows the inclusion of medical expert knowledge in the process of feature selection and thus the selection of a subset of features with clinical relevance.



Component	histological interpretation
1+2	number of cell nuclei
	especially low values in psammoma bodies and areas of collagen
3+4	Amount of anisotropic tissue inhomogeneities
	especially high values in case of parallel aligned cracks in fibroblastic tissue
5+6	Anisotropic shaped cell nuclei
	especially high values in fibroblastic tissue

Fig. 6. Top: A histological feature map derived from the visualization procedure. The Map allows to clearly link numerical features to histological semantics (Bottom).

Due to selection of a small subset of features computational costs are reduced and difficulties associated with large feature vectors, sometimes called the *curse of dimensionality* [10] can be avoided. Additionally, by utilizing a class separation measure it was proven that the selected subset of features leads to a separation of the four classes in the feature space with highest accuracy.

Features	r_c of tumor classes			
	meningo-theliomatous	fibro-blastic	psammo-matous	transi-tional
set f_1	1.295	1.334	1.923	1.601
set f_2	1.403	1.466	2.101	1.487
all features	1.369	1.386	1.968	1.518
selected features	**1.735**	**1.741**	**2.992**	**1.954**

Table 3
Separation of tumor classes according to different feature combinations.

6 Acknowledgments

The authors would like to thank Axel Saalbach, Technical Faculty, University of Bielefeld and Oliver Schulz, Max-Planck-Institute Dortmund for the provided software resources. Birgit Lessmann would like to thank the University of Bielefeld and the federal state Nordrhein-Westfalen for financial support.

References

[1] S. Aksoy and R.H. Haralick. Textural features for image database retrieval. In *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries (CVPR'98)*, pages 45–49, Santa Barbara, California, 1998.

[2] T. Chang and C.-C.J. Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Imag. Process.*, 2(4):429–441, 1993.

[3] A. Cohen, I. Daubechies, and J. Feauveau. Biorthogonal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, 45:485–559, 1992.

[4] I. Daubechies. *Ten Lectures on Wavelets*. CBMS-NFS Series Appl. Math., SIAM, 1991.

[5] D.J. Foran, D. Comaniciu, P. Meer, and L.A. Goodell. Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories and telemicroscopy. *IEEE Trans. Inf. Techn. Biomed.*, 4(4):265–273, 2000.

[6] K.A. Han and S.H. Myaeng. Image organization and retrieval with automatically constructed feature vectors. In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 157–165, 1996.

[7] D. Hearn and M.P. Baker. *Computer Graphics*. Prentice-Hall Intternational Inc., 1997.

[8] P. Howarth, A. Yavlinsky, D. Heesch, and S. Rueger. Medical image retrieval using texture, locality and colour. In C. Peters, P. Clough, and J. Gonzalo, editors, *5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004, Bath, UK*, Lecture Notes in Computer Science, pages 740–749. Springer, 2004.

[9] J.R. Iglesias-Rozas and N. Hopf. Histological heterogeneity of human glioblastomas investigated with an unsupervised neural network (SOM). *Histology and Histopathology*, 20:351–356, 2005.

[10] A.K. Jain, R.P.W. Duin, and J. Mao. Statistical pattern recognition: A review. *IEEE Trans. Patt. Anal. Machine Intell.*, 22(1):4–37, 2000.

[11] J. Jelonek, K. Krawiec, R. Slowinski, and J. Szymas. Intelligent decision support in pathomorphology. *Pol. J. Pathol.*, 50(2):115–118, 1999.

[12] T. Kohonen. *Self Organizing Maps*. Springer, Berlin, Heidelberg, 1995.

[13] J. Laaksonen, M. Koskela, Sami Laakso, and Erkki Oja. Picsom - content-based image retrieval with self-organizing maps. *Pattern Recognition Letters*, 22:1199–1207, 2000.

[14] B. Lessmann, V. Hans, A. Degenhard, and T.W. Nattkemper. Feature-space exploration of pathology images using content-based database visualization. In J.M. Reinhardt and J.P.W. Pluim, editors, *Proceedings of SPIE Medical Imaging*, volume 6144, 61445Z, San Diego, California, USA, 2006.

[15] S. Livens, P. Scheunders, G. van de Wouwer, D. van Dyck, H. Smets, J. Winkelmans, and W. Bogaerts. A texture analysis approach to corrosion image classification. *Microscopy, Microanalysis, Microstructures*, 7(2):1–10, 1996.

[16] DN. Louis, B.W. Scheithauer, H. Budka, A. von Deimling, and J.J. Kepes. Meningiomas. In P. Kleihues and W.K.Cavenee, editors, *World Health Organization Classification of Tumours. Pathology & Genetics. Tumours of the Nervous System.*, pages 176–184. IARC Press, Lyon, 2000.

[17] W. Y. Ma and B. S. Manjunath. Image indexing using a texture dictionary. In *SPIE conference on Image Storage and Archiving System*, volume 2606, pages 288–298, 1995.

[18] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, San Diego, London, 1999.

[19] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Trans. Patt. Anal. Machine Intell.*, 14:710–723, 1992.

[20] B.S. Manjunath and W.Y. Ma. Texture features for image retrieval. In V. Castelli and L.D. Bergman, editors, *Image Databases - Search and Retrieval of Digital Imagery*. John Wiley and Sons, Inc, New York, 2002.

[21] H. Mueller, P. Ruch, and A. Geissbuhler. Enriching content-based image retrieval with automatically extracted mesh terms. In *GMDS conference*, pages 266–269, Innsbruck, Austria, 2004.

[22] A.W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval systems at the end of the early years. *IEEE Trans. Patt. Anal. Machine Intell.*, 22(12):1349–1380, 2000.

[23] J. Stauder, J. Sirot, H. Le Borgne, E. Cooke, and N. E. O'Connor. Relating visual and semantic image descriptors. In *Proc. European Workshop for the Integration of Knowledge, Semantics and Digital Media Technology (EWIMT)*, London, 2004.

[24] E.J. Stollnitz, T.D. Derose, and D.H. Salesin. *Wavelets for Computer Graphics*. Morgan Kaufmann Publishers, San Francisco, 1996.

[25] G. van de Wouwer, P. Scheunders, and D. van Dyck. Statistical texture characterization from discrete wavelet representation. *IEEE Trans. Imag. Proc.*, 8(4):592–598, 1999.

[26] G. van de Wouwer, B. Weyn, P. Scheunders, W. Jacob, E. van Marck, and D. van Dyck. Wavelets as chromatin texture descriptors for the automated identification of neoplastic nuclei. *J. Microscopy*, 197:25–35, 1999.