16. T.H. Reiss, The Revised Fundamental Theorem of Moment Invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, no. 8, August (1991)

17. L.R. Schad, Correction of Spatial Distortion in Magnetic Resonance Agiography for Radiosurgical Treatment Planning of Cerebral Arteriovenous Malformations. *Magnetic Resonance Imaging* 10, 609–621, (1992)

18. U. Taneja, R.A.Robb, C.R. Jack jr., J.J. Camp, Tumor Volume Measurement with 3D Multispectral Analysis: Work in Progress. *Suppl. to Radiology* 197(P), 328, (1995)

19. M.W. Vannier et al., Multispectral Analysis of Magnetic Resonance Images. *Radiology* 154(1), 221–224, (1985)

20. *Wavelets for Texture Analysis*, http://www.ruca.ua.ac.be/~VisionLab/WTA.html, (1997)

21. D. Willett, C. Busch, F. Seibert, Fast Image Analysis using Kohonen Maps. In *Proceedings of the IEEE Workshop NNSP*, 461–470, (1994)

5. M. Gross, R. Koch, L. Lippert, A. Dreger, Multiscale Image Texture analysis in Wavelet Spaces. In *Proceeding of the IEEE – International Conference on Image Processing '94*, 412–416, (1994)

6. H. Handels, *Automatische Analyse mehrdimensionaler Bilddaten zur Diagnoseunterstützung in der MR–Tomographie*, Shaker, (1992)

7. H. Handels et al., *KAMEDIN: A telemedicine system for computer supported cooperative work and remote image analysis in radiology.* accepted for publication in: Computer Methods and Programs in Biomedicine, (1997)

8. R.M. Haralick and L.G. Shapiro, *Computer and Robot Vision*, Addison–Wesley (1992)

9. K.H. Höhne et al., 3D Visualization of Tomographic Volume Data using the Generalized Voxel Model. *The Visual Computer* 6, no. 1, 28–36, (1990)

10. M.K. Hu, Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, 179–187, (1962)

11. T. Kohonen, The Self–Organizing Map. *Proceedings of the IEEE* 78, no. 9, 1464–1480, (1990)

12. A. Krogh and, J. Vedelsby, Neural Network Ensembles, Cross Validation, and Active Learning. In *Advances in Neural Information Processing Systems,* G. Tesauro et al. (Eds.), MIT Press, 231–238, (1995)

13. K.I. Laws, Rapid texture identification. *Proc. SPIE Conference on Image Processing for Missile Guidance* 238, 376–380, (1980)

14. S. Livens et al., A Texture Analysis Approach to Corrosion Image Classification. *Microscopy, Microanalysis, Microstructures* 7(2), 1–10, (1996)

15. S.G. Mallat, A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11, 674–693, (1989)

designed to be capable of only two bands (mere double–echo). Additional studies on a subset of the test samples have proven that no significant decrease of classification rates must be expected on mere double–echo data.

This study shows that the presented pipeline is capable of inter–patient tissue segmentation. In addition investigation of inter–scanner generalization were made which resulted in lower classification rates. This will be the subject of further research.

Acknowledgments

References

1. C. Busch and M. Eberle, Morphological Operations for Color–Coded Images. In *Computer Graphics forum* 14, no. 3, C193–C204, (1995)

2. C. Busch and M. Gross, Interactive Neural Network Texture Analysis and Visualization for Surface Reconstruction in Medical Imaging. In *Computer Graphics forum* 12, no. 3, C49–C60, (1993)

3. T. Chang and C.C.J. Kuo, Texture Analysis and Classification with Tree–Structured Wavelet Transform. *IEEE Transactions on Image Processing* 2, 429–441, (1993)

4. C. K. Chui, *An Introduction to Wavelets*, Academic Press Inc., Boston San–Diego New York, (1992)

The semantically based postprocessing is performed using the operations from Section 5 and applying the rules shown in Figure 13. This leads to morphologically processed images as shown in Figure 12. The first step is an operation that removes isolated misclassified pixels in the scene as they might exist for grey matter, white matter, liquor, tumor, and fat tissue employing a relative transparent opening. The second step smoothes the contours and closes the regions. Then a sensitive closing is performed on each of the following anatomical classes: fat, white matter, grey matter, liquor. The set of protected color values is the set {tumor, fat, white matter, grey matter, liquor} without the color of the actual single sensitive closing step. The output of this smoothing procedure will contain a number of undefined regions. Thus the finishing operation is a limited filling of undefined regions. This can be achieved by alternating sensitive dilations on the anatomical classes.

7 Conclusion

This paper presents a fully automated segmentation pipeline for tomographic images. The main steps of this pipeline are texture classification, combination of classification results and knowledge based postprocessing using morphological operations for color–coded images.

The wavelet transform is investigated for texture analysis on a large dataset of approx. 70,000 test samples. The comparative evaluation of the classification capabilities leads to mean classification rates of up to 96% correct test samples. It can be shown that this texture discriminating method leads to higher classification rates than well known reference methods. It should be mentioned that these investigations did not consider performance aspects of different texture discriminating methods. Computing wavelet features requires more computational effort than traditional methods.

Despite the fact that the presented classification rates are based on multi–modal tomographic data with three bands (double–echo MRI with matched CT), the segmentation pipeline was
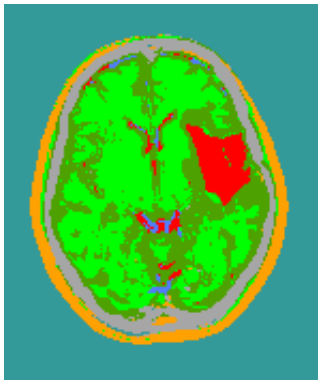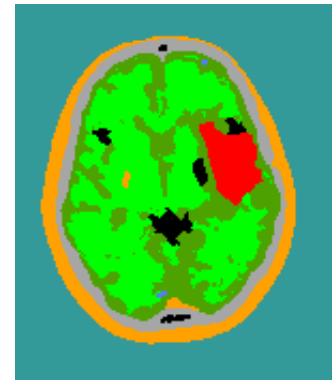
Figure 10. Step 1: classification of multi–band slice (see Fig. 8 to 9)
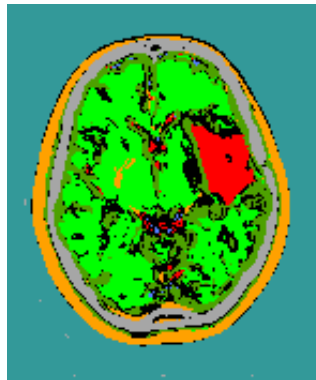
Figure 11. Step 2: combination of single results

Figure 12. Step 3: post–processing using anatomical knowledge

color–code	tissue type	semantically possible neighbors
(red)	tumor	{tumor, liquor, grey matter, white matter}
(blue)	liquor	{liquor, tumor, grey matter, (white matter), bone}
(bright green)	white matter	{white matter, tumor, (liquor), grey matter}
(dark green)	grey matter	{grey matter, tumor, liquor, white matter}
(orange)	fat tissue	{fat tissue, background, bone}
(teal)	background	{background, fat tissue}
(grey)	bone	{bone, liquor, fat tissue}
(black)	undefined	

Figure 13. Color–codes of classes used in classification results

A classification result for the selected slice of patient dataset 78 (see Figures 7 to 9) is shown in Figure 10 which was created using the implicit feature extraction method. The reduced classification rate of the patient data stems from misclassified *tumor* pixels inside the ventricular system. The color–codes of tissue types of the investigated brain section are given in Figure 13. Therein the color–code *black* is used for all pixels remaining in an undefined state, for example if none of the above classes were considered appropriate by a classifier.

Figure 11 shows the result which has been achieved solely by combining classification data stemming from different classifiers (step 2 of the pipeline).

gray columns. The latter statistic represents the probability that any pixel in a dataset is correctly classified to one of the defined classes.

For comparison purposes concerning the investigated texture analysis methods, Table 6 shows class mean values for all methods used in step 1 of the segmentation pipeline.
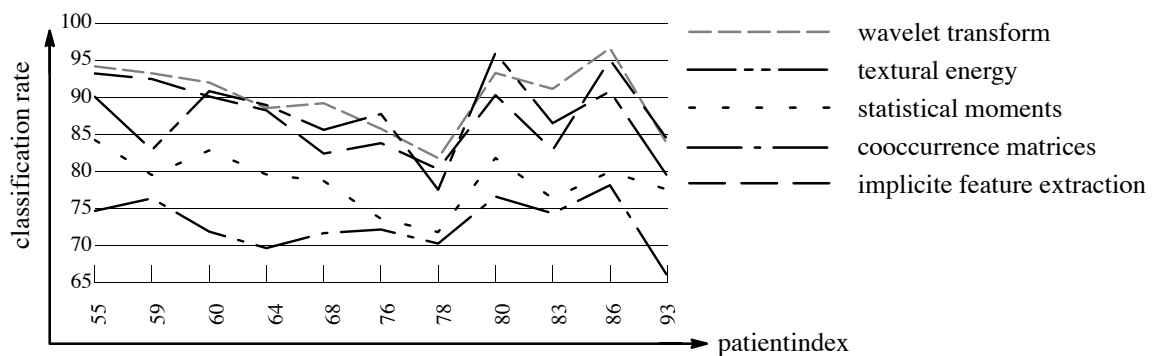
Tab. 6: Compared classification rates for the investigated texture analysis methods

The wavelet transform provides slightly higher classification rates than the reference methods. This supports the chosen multi resolution approach for texture discrimination. Furthermore the wavelet transform show the best minimum classification probability: the lowest result (referring to patientindex 78) is clearly higher than in all other methods. This can be interpreted as a better constancy with respect to the generalization capability.

It must be stated that there is no quantitative measure to judge the visual impression. The user definately expects homogenous closed regions with a low number of misclassifications inside the regions. This does not always correspond to the quantitative statistics in the tables. This is due to the fact that training areas are usually defined inside closed texture (tissue) regions. In order to judge the segmentation pipeline visually, the dataset stemming from the lowest result is used. Example slices from this dataset are shown in Figures 7 to 9 where the white polygon marks the defined isoline that was actually used for irradiation planing. Thus this polygon can be interpreted as the enlarged borders of the affected pathological areas.
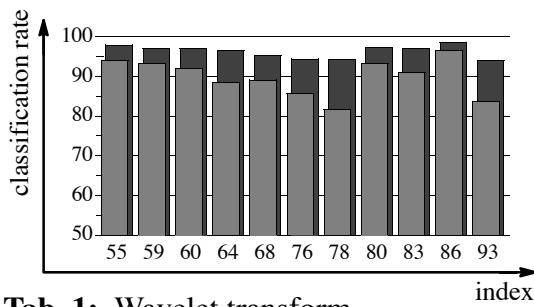
Tab. 1: Wavelet transform –
classification rates
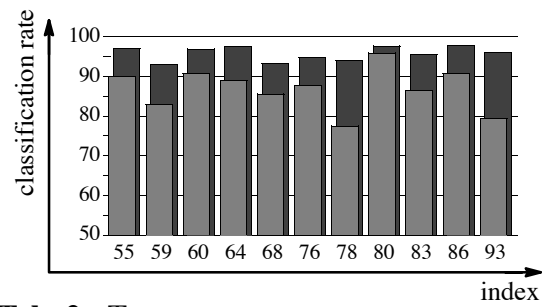
Tab. 2: Texture energy –
classification rates
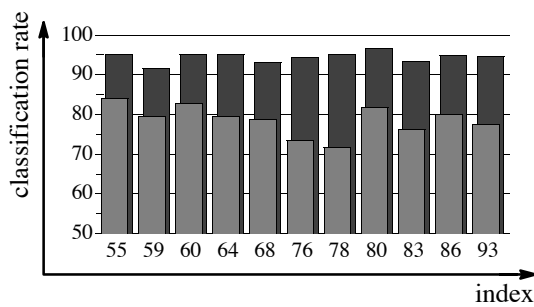
Tab. 3: Statistical moments –
classification rates
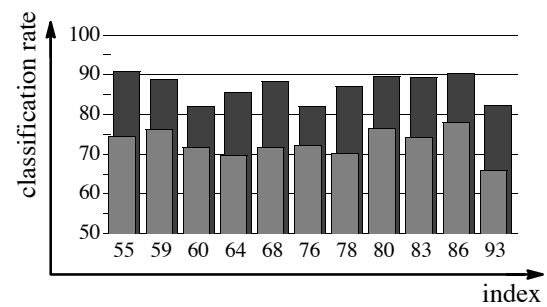
Tab. 4: Cooccurrence matricies –
classification rates
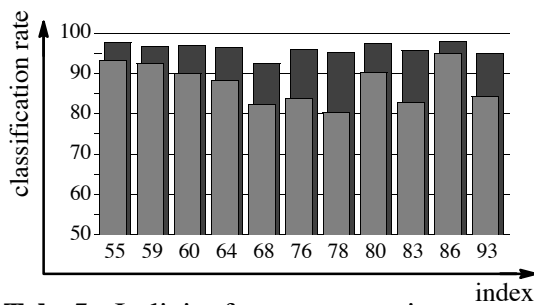
Tab. 5: Imlicite feature extraction –
classification rates

= correctly classified test samples
with respect to all pixels

= mean values over all classes
of correctly classified test samples

The dark gray columns in the above statistics refer to the evaluation of all pixels. Unfortunately the distribution of training samples among the classes is not homogenous, which results from their anatomical diversity. For example the definition of a large training area for the class *background* is much easier in compared to other tissue types, resulting in an overrepresentation of *background* pixels in the training set. This overrepresentation in combination with high classification rates leads to impressive statistics. In order to avoid this effect, the mean value of classification rates over all classes was calculated for each test, which is displayed by light
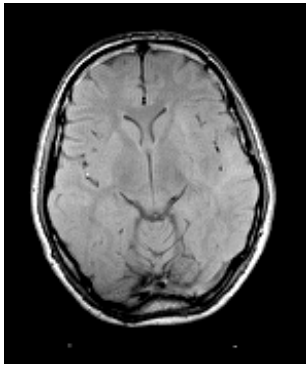
Figure 7. P_w–weighted MRI band
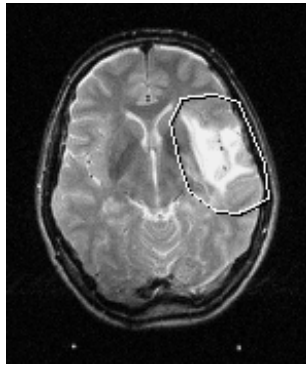
Figure 8. T_2–weighted MRI band
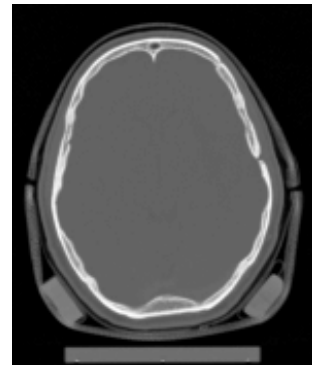
Figure 9. CT–band

The intention of this statistical evaluation is twofold: on the one hand the generalization capability of the segmentation process needs to be investigated. On the other hand a major aspect of this research is to find out whether the different texture analysis methods provide a better solution of this task. Thus the results presented in this chapter refer to classification rates that were achieved as classification results in step 1 of the segmentation pipeline.

The network design of the Kohonen Feature Map was optimized in previous work and has not been changed during the investigations. The output layer of the network was sized in three dimensions with 6x6x6 neurons while the input layer was adapted to the respective length of the feature extraction method. The training time was set to 40,000 cycles for the self–organization and 160,000 cycles for additional LVQ3 training.

In order to test the generalization capability the leaving–one–out method was applied. Thus for a single test the available ground truth data was split, while one of the *n* patient datasets was selected as test data and the training was performed on the remaining *n–1* datasets. Having trained the Kohonen Feature Map on 10 patients the 11th was used as reference test data to calculate the generalization capabilities. The results in relation to the involved texture analysis method are given in Table 1 to Table 5. For the investigations the Haar–wavelet was used for the wavelet transform.

tient analysis is possible at all an investigation was made on the base of datasets of 11 clinical studies. All of the tomographic datasets were acquired by a Siemens Magnetom with a double–echo sequence and physical acquisition parameters of TR=1800 ms for the repetition time and of TE=15/90 ms for the echo time. These parameters provide a proton weighted image (P_w–weighted) recorded at the first echo and a T_2–weighted image recorded at the second echo. The spatial resolution was approx. 1mm in all cases. In addition CT scans (Siemens Somatom) were recorded and matched to the MRI data using the stereotactical frame method [17]. All of the involved studies were investigated using histological probes resulting in a diagnosis for brain tumor of type astrozytoma and grade II respectively. Thus these datasets are providing comparable clinical cases and may be used for evaluation of the segmentation pipeline.

For the supervised classification ground truth data areas were set up for all of the specified classes. E.g. in each of the involved patient datasets suitable slices were selected and rectangular training areas were defined for the classes considered (*tumor, liquor, white matter, grey matter, bone, fat tissue* and *background*). Patterns from these training areas were tagged with class assignments and were stored in a database.

These samples contain a 3D environment for all bands (see examples in Figure 7 to 9) providing the texture analysis methods with the necessary texture window. Samples for the defined training areas resulted in a database of approx. 70,000 samples used for training and evaluation. The distribution of the samples over the patients is approx. homogenous.

element B that has the same color for all pixels. The closing behavior is then similar to a binary closing, as Figure 6 illustrates.
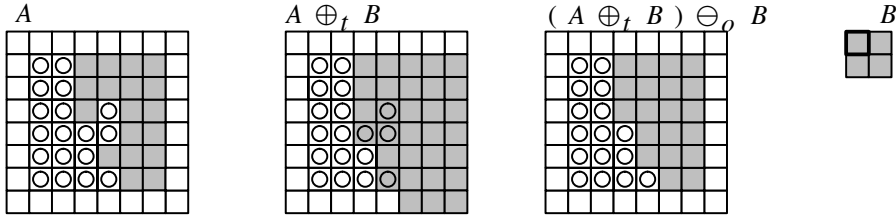
Figure 6. Superseding closing

Furthermore, the definition of a *sensitive closing* operation with

sensitive transparent closing: $A \bullet_{s(G),t} B = (A \oplus_t B) \ominus_{s(G),t} B$ (38)

sensitive superseding closing: $A \bullet_{s(G),o} B = (A \oplus_t B) \ominus_{s(G),o} B$ (39)

protects all pixels in A which have common color–values with protection set $G \subset F$. Note that in definitions 38 and 39 the erosion, not the dilation, was attributed with a sensitive operation mode. The usefulness of this choice is easy to prove. The definitions of the relative versions of 36 and 37 are straightforward

relative transparent closing: $A \bullet_{r(n),t} B = (A \oplus_t B) \ominus_{r(n),t} B$ (40)

relative superseding closing: $A \bullet_{r(n),o} B = (A \oplus_t B) \ominus_{r(n),o} B$ (41)

Again, the relativity which describes the matching percentage of the structuring element is assigned to the erosion.

6 Results

A major purpose of this work is to compare the suitability of different texture discrimination methods applied to MRI tomographic data. In order to answer the question whether inter–pa-
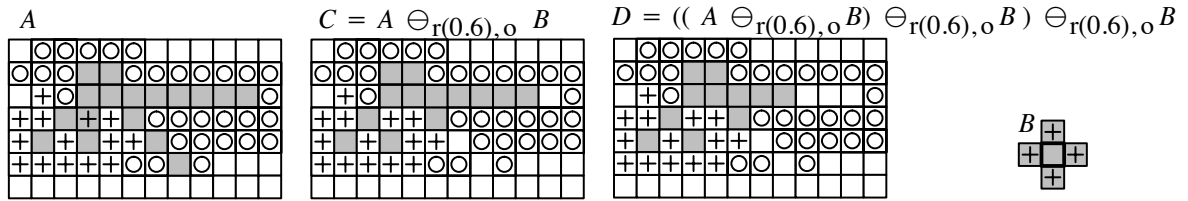
Figure 5. Relative erosion – a single and iterative application.

The matching condition must be fulfilled

for at least 60% of the structuring element B.

5.4. Opening

With the definitions in 5.2. and 5.3. one can combine erosion and dilation for a linked color–coded morphological operation. Corresponding to the binary standard the aim is to smooth regions. The operation is defined using a transparent erosion and transparent dilation and requires a final intersection with the input image A.

$$\text{transparent opening:} \quad A \circ_t B = ((A \ominus_t B) \oplus_t B) \cap A \qquad (34)$$

An interesting variation to Equation 34 can be derived if the involved erosion that detects the region assigned to the reference pixel's color is encoded in a relative way.

$$\text{relative transparent opening:} \quad A \circ_{r(n),t} B = ((A \ominus_{r(n),t} B) \oplus_t B) \cap A \quad (35)$$

5.5. Closing

The counterpart in combining dilation and erosion defines the closing of a color–coded image in its transparent and superseding versions

$$\text{transparent closing:} \quad A \bullet_t B = (A \oplus_t B) \ominus_t B \qquad (36)$$

$$\text{superseding closing:} \quad A \bullet_o B = (A \oplus_t B) \ominus_o B \qquad (37)$$

Note that in Equations 36 and 37 one must restrict the definition area from structuring element B to achieve a proper working operation. Thus, one should require a *monochrome* structuring

5.3. Erosion

Although erosion is considered the morphological opposite to dilation, this is not true in all cases. The transparent version is defined as

transparent erosion:

C = A $\ominus_t$ B: (32)
C = A
for all a $\in$ A and A(a) $\cap$ B(0) $\neq$ {} **do**   /* reference pixel matches a */
    if (for all b $\in$ B : A(a+b) $\cap$ B(b) $\neq$ {} **)** /* structuring element matches */
            C(a) = C(a) $\cup$ B(0)
    else
            C(a) = C(a) \ B(0)

Using the transparent erosion the image position that is covered by the reference pixel of the structuring element will still keep its previous color.

If one works with color–coded operations, the shape that is encoded by an erosion operator will soon be considered to be too strong or too precise. The matching condition of a structuring element might be fulfilled only in a limited number of cases. Negative impacts on the desired processing behavior can be reduced with a modified *relative* definition of the erosion.

relative transparent erosion:

C = A $\ominus_{r(n),t}$ B; with n $\in$]0;1]**:** (33)
C = A
for all a $\in$ A and A(a) $\cap$ B(0) $\neq$ {} **do**   /* reference pixel matches */
    if (for n $_*$ card B **or more** b $\in$ B : A(a+b) $\cap$ B(b) $\neq$ {} **)**   /* a defined
                        percentage of pixels matches at covered pixelposition? */
            C(a) = C(a) $\cup$ B(0)
    else
            C(a) = C(a) \ B(0)

Relative transparent or superseding erosions enable us to achieve a more fuzzy description of the matching condition, as for example: "The structuring element *B* matches at position *a*, if a percentage *n* of the neighbored pixels match the corresponding pixels". The percentage *n* is given with respect to the cardinality of the structuring element B. If *n* is set to 1, the relative erosion is equivalent to the definition of erosion in Equation 32 .

must require that any operation can only add or supersede color values to the color of a pixel *a* if the actual color of that pixel *A(a)* does not include any write–protected color value.

sensitive transparent dilation:

```
C = A ⊕s(G),t B;  with G ⊂ F:
C = A
for all a ∈ A and A(a) ∩ B(0) ≠ {} do        /* reference pixel matches a */
    for all b ∈  B do                          /* for all points of B */
            if ( A(a+b) ∩ G = {} )             /* if the covered pixel is not protected */
                C(a+b) = C(a+b)  ∪ B(b)
```

$$(30)$$

The write–protected color values are elements of a color subspace $G \subset F$. One can define sensitive instances of the dilation versions and denote them with an index *s(G)* that expresses the sensitivity to a protected color space. The corresponding *sensitive superseding dilation* is defined straightforwardly by replacing the union `C(a+b)=C(a+b)∪B(b)` with an explicit assignment `C(a+b)=B(b)` with the color of the covering element. Figure 4 illustrates an example for a sensitive superseding dilation on an image that contains one protected color.
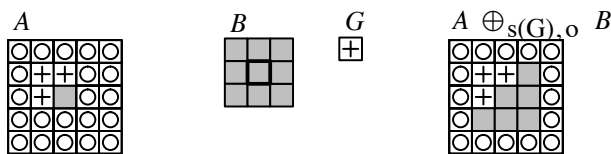
Figure 4. Sensitive superseding dilation of image *A*

In some applications it might be desirable to include a certain condition, for example, the correlation with a second image. In this case, both versions of dilation are extended in a way that the condition is defined with image *C*.

conditional transparent dilation:

```
D = A ⊕c,t B | C:
D = A
for all a ∈ A and A(a) ∩ B(0) ≠ {} do        /* reference pixel matches a */
    for all b ∈  B do                          /* for all points of B */
            if ( B(b) ∩  C(a+b) ≠ {} )         /* if b matches c=a+b */
                D(a+b) = D(a+b) ∪ B(b)
```

$$(31)$$

25 is fulfilled for all elements *b* of *B* and their corresponding *covered* pixels in *A,* then it is a *match* between object *B* and object *A* :

object B matches at pixel a to object A: $B(b) \cap A(a + b) \neq \{\} \; \forall \; b \in B$ (26)

Operating with structuring element *B* on image *A* one can differ between two morphological operation modes, either *superseding* (subindex o) or *transparent* (subindex t). The characteristics of *superseding* operations are similar to those of binary operations. A pixel *b* with an assigned color *B(b)* will supersede the previous color *A(a),* if *B* covers *a.*. The contrary *transparent* operation will preserve the information that was set in image *A.*

transparent operation: $C(a) = A(a) \cup B(b)$ (27)

5.2. Dilation

Similar to binary operations, the dilation is one of the two basic operations in color–coded morphology. Contrary to binary operations the definitions for color–coded operations are generally defined with pseudo code.

For the superseding version of dilation, the color of structuring element *B* will replace a color that was originally set at a covered position in *A*

superseding dilation:

```
C = A ⊕o B:                                                    ( 28 )
C = A
for all a ∈ A and A(a) ∩ B(0) ≠ {} do      /* reference pixel of B matches a */
    for all b ∈  B do                      /* for all points of  B */
         C(a+b) = B(b)
```

with associated definition area for structuring element B:

$\{B \mid B$ is object, $B(i) = B(j) \; \forall \; i,j \in B\}$ (29)

The definition of color–coded operations defines a powerful instrument for realizing *semantic sensitive processing* if the operations themselves are sensitive. Following this concept, one

5.1. Definition of Color–Coded Images

First of all an application–dependent color space *F* as the set containing all *k* possible color values has to be defined.

$$\textit{color space:} \quad F = \{f_1, \ f_2, \ \ldots, f_k\} \tag{21}$$

Furthermore a color set $\mathcal{F}$ of all possible subsets of *F* shall be defined as

$$\textit{color set:} \quad \mathcal{F} = \{\{ \ \}, \{f_1\}, \ldots, \{f_k\}, \ \{f_1, \ f_2\}, \ldots \ldots, \{f_1, \ f_2, \ \ldots, f_k\}\} \tag{22}$$

Note that the empty set itself is an element of $\mathcal{F}$. One can now define a color–coded image as an object A, which is a subset $A \subseteq Z^N$ and an associated function $\mathcal{A}$, which defines a projection from A to the set $\mathcal{F}$:

$$\textit{color–coded image A / object A:} \quad A \subseteq Z^N, \ \mathcal{A} : A \to \mathcal{F} \tag{23}$$

According to this definition, function $\mathcal{A}(a)$ assigns a color–code to each pixel *a*. The color–code – simply named *color* – is one element from the color set $\mathcal{F}$. The color itself might consist of an arbitrary number of color values $f_1, f_2,\ldots$ from the color space.

For a clear description of the impact of a structuring element *B* on image *A* one can define:

$$\textit{object B at pixel p covers pixel q in image A:} \quad \exists b \in B \ \textit{with } p + b = q \tag{24}$$

The definition of a color–code for a pixel requires a specific understanding of matching. Therefore, one can define a *match* between two object elements as

$$\textit{pixel a matches pixel b:} \quad A(a) \cap B(b) \ \neq \ \{\} \tag{25}$$

In this case, the elements *a* and *b* have at least one common color value from the color space in their colors. If object *B* is moved with its reference pixel *B(0)* to position *a*, and condition

In Equation 19 $\delta_{i,j}$ is the Kronecker delta and cls_r^k $[x, y]$ defines the class assignment of classi-fier k. The index r refers to a dataset of step 1 of the pipeline. The class assignment cls in the united dataset (index v) can then be calculated by evaluating the maximum of the single votes

$$w_{cls}[x, y] \;=\; \max_c \; \{w_c[x, y]\} \qquad\qquad (\; 20 \;)$$

The class assignment from Equation 20 is the output of step 2 of the segmentation pipeline and is based on the analysis of all contributing classifiers. This result can be considered to be much more reliable than choosing one and disregarding the rest. Visual results of this step are given in Section 6. They show that some misclassified pixels are removed after this step. Neverthe-less there some still remain. Thus the next section will review a postprocessing method as the final step of the pipeline.

5 Knowledge Based Post Processing

Classification results and combined results, as they are computed from step 1 and step 2 of the segmentation pipeline provide, color–coded images that can be seen as a *theme map* of the in-vestigated brain section. Since the processing of step 1 and 2 is pixel based, misclassifications do occur and should be detected and eliminated on the base of the knowledge about a *normal* brain anatomy.

It is essential to keep the importance of *semantic sensitive processing* in mind. If one takes a tomographic image from the brain and submit it to a postprocessing analysis, one needs anoth-er treatment of a *liquor* pixel adjacent to *white–* or *grey matter* than for a *liquor* pixel adjacent to a *background* pixel. The latter case does not satisfy the simple semantic rule that all anatom-ically possible neighbors of a *background* pixel must be either *background* or *fat tissue*.

In order to consider this situation the segmentation pipline incorporates specific morphologi-cal operations for color–coded images that were presented in [1] and are briefly reviewed here.

In this case the input of all K classified datasets is considered and pixels in the final vote are only set, if all contributing classifiers vote for the same class assigment. This method provides a highly reliable decision but obviously leaves a fair chance that manny pixels will be set to an undefined state.

Widely spread and more promising than boolean operations is the technique of combining several classifiers considering statistical weights to an collection [12]. This leads to the problem of deciding how to determine the contributing weights. Some solutions employ multilayer perceptrons which propagate acivities of ouput layer neurons to the collection. This technique can not be applied to Kohonen Feature Maps, since no equivalent information is recorded. Thus one can involve class dependent bias rates that can be calculated from statistical evaluation of a classifier on ground truth data. Equation 18 gives a bias value p_{cls}^{k} which defines the reliability of classifier k with respect to the class cls

$$p_{cls}^{k} = \frac{RP}{(RP + FN)} \quad \frac{RP}{(RP + FP)} \qquad (18)$$

where RP, FN, FP can be calculated as percentage values from the confusion matricies of classifier k:

- **RP (right–positive):**   classifier k recognized RP pixles of class cls correct.

- **FN (false–negativ):**   classifier k falsely rejected FN pixels of class cls

- **FP (false–positive):**   classifier k falsely assigned FP pixels
  of other classes to class cls

The bias values from Equation 18 are used as weights in a modified majority voting as follows, where a single vote w_c for class c is given by:

$$w_c[x, y] = \sum_{k=1}^{K} \delta_{c, cls_r^k [x,y]} \quad p_{cls}^{k} \qquad \forall \ c \qquad (19)$$

- Cooccurrence Matrices   Introduced by Haralick [8]

  describing second order statistics

- Textural Energy   According to Laws [13]

  Convolution of the texture window

  filter kernels

- Implicit Feature Extraction   Used in combination with KFM in [2]

4 Combination of Classification Results

The second step of the proposed segmentation pipeline in Figure 1 combines multiple classifiers that are trained on the same task but use different methods. This technique is often referred to as a network ensemble [12] even though the subjects which are to be combined are classification results rather than the classifiers themselves. Furhtermore if we think of the objects we realize that this step is independent of the classification method used (statistical classifier, neural network, ...) and can easily be applied to non–neural classifiers or mixed ensembles. Nevertheless most techniques require information about the statistical reliability of the involved classifier. In order to solve the task, boolean operators, statistically weighted combinations and fuzzy set techniques have been considered in this work. The latter hardly justify the computational overhead with respect to the improvement of classification result and is not covered in this paper.

A rather simple method to combine the classification results is to apply boolean operators. Equation 17 gives a combination method involving the AND operator:

$$class_v[x, y] = \begin{cases} class_r^1[x, y] & \text{if } class_r^1[x, y] = ... = class_r^K[x, y] \\ undefined & \text{else} \end{cases} \qquad (17)$$

three spatial orientations. According to [15] it can be shown that these orientations correspond to the maximum sensitivity of the human vision system.

Aside from the selectivity aspect, the suitability of the transform for feature extraction is funded in its hierarchical representation. For derivation of features we decompose a texture window W with iteration depth m. The features F_{mi} for a central pixel [x,y] are then taken from the detail coefficients of the wavelet representation:

$$ F_{mi} = d^i_{m,n,j} \qquad | \ n = \frac{x}{m} \wedge j = \frac{y}{m} \qquad\qquad (16) $$

If different frequency bands are biased with different weights one can dynamically control the receptive behavior of the texture window at distinct resolutions. Due to the localization of the wavelet transform in time and frequency space the derived features describe the texture at a localized position in the signal. Features for a 3D texture window are calculated by accumulating detail coefficients over adjacent slices in a feature vector. Further accumulation steps include associated bands of the multi–modal image sequence.

3.2. Reference Methods

The previous section reviewed the wavelet transform and showed how texture features can be derived from the wavelet representation. In order to investigate the efficiency of the wavelet transform several traditional texture discriminating methods are also implemented in the presented pipeline. The reference methods are briefly shown here and can be found in the referred literature.

- Statistical Moments   Initially proposed by Hu [10] and
                             lately corrected by Reiss [16]
                             describing first order statistics

$$\hat{\psi}(0) \;=\; \int\limits_{-\infty}^{\infty} \psi(x)\ dx \;=\; 0 \qquad\qquad (\,10\,)$$

must hold for the basis wavelet and thus wavelet $\psi_m\ (x)$ works like a band–pass filter. The corresponding scaling function $\phi_m\ (x)$ might be considered as a low–pass filter.

In order to apply the WT to multidimensional signals, tensor product extensions $V_m^2 = V_m \otimes V_m$ are involved. The 2D operation is then performed on function $f(x,y) \in L^2(\mathbb{R}^2)$ which bases on the 2D scaling function

$$\phi^2(x,y) \;=\; \phi(x)\ \phi(y) \qquad\qquad (\,11\,)$$

as the product of two 1D scaling functions. The decomposition of a 2D space into subspaces is realized by using

$$V_{m-1}^2 \;=\; V_m^2 \oplus W_m^{2,1} \oplus W_m^{2,2} \oplus W_m^{2,3} \qquad\qquad (\,12\,)$$

where V_m^2 is given by the tensor product of two identical subspaces. The spaces $W_m^{2,1}, W_m^{2,2}, W_m^{2,3}$ are defined by the following 2D wavelet basis functions

$$\psi_{mnj}^{2,1}\ (x,y) \;=\; 2^{-m}\ \phi(2^{-m}\ x - n)\ \psi(2^{-m}\ y - j) \qquad\qquad (\,13\,)$$

$$\psi_{mnj}^{2,2}\ (x,y) \;=\; 2^{-m}\ \psi(2^{-m}\ x - n)\ \phi(2^{-m}\ y - j) \qquad\qquad (\,14\,)$$

$$\psi_{mnj}^{2,3}\ (x,y) \;=\; 2^{-m}\ \psi(2^{-m}\ x - n)\ \psi(2^{-m}\ y - j) \qquad\qquad (\,15\,)$$

which discriminate vertical, horizontal and diagonal points of variation in the signal. Due to the orthonormal design of the above basis functions a property is implied which is of high importance for texture analysis: the orthogonal wavelet representation of *f(x,y)* is a decomposition of the signal into independent frequency channels, which allows a coarse selectivity of

$$\int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \qquad\qquad (5)$$

where $\hat{\psi}(\omega)$ is the fourier transform of the wavelet function.

The wavelet transform is in close relation to Mallats multi resolution analysis [15]: if a discrete function $\tilde{f}(x) \in L^2(\mathbb{R})$ is considered as an element of a vector space V_0 and furthermore this space V_0 is decomposed into hierarchical subspaces V_m then $\tilde{f}(x)$ can be described by the contribution from the single subspaces m with $m = 1, 2, ...$. The decomposition of space V_0 is computed iteratively, where

$$V_{m-1} = V_m \oplus W_m \qquad\qquad (6)$$

defines the splitting of a space V_{m-1} into the complementary subspaces V_m and W_m. The space W_m is spanned using the set of orthonormal functions $\{\psi_{m,n} \mid n \in \mathbf{Z}\}$. According to Equation 6 the space V_m is orthogonal to W_m and is spanned itself by a set of orthonormal *scaling functions* $\{\phi_{m,n} \mid n \in \mathbf{Z}\}$ which are derived in analogy to Equation 4 from a mother function:

$$\phi_{m,n}(x) = 2^{-\frac{m}{2}} \phi(2^{-m} x - n) \qquad\qquad (7)$$

Employing Equations 4 and 7 , the coefficients $c_{m,n}$ and $d_{m,n}$ of the wavelet representation can be calculated as follows:

$$c_{m,n} = \langle f(x), \phi_{m,n} \rangle \qquad\qquad (8)$$
$$d_{m,n} = \langle f(x), \psi_{m,n} \rangle \qquad\qquad (9)$$

The detail coefficient describes the signal in different frequency bands. This stems from the fact, that according to Equation 5

from compression, motion analysis and volume rendering to texture analysis [5] and [14]. A good overview of the current activities can be found in [20].

One decided upon the WT for application on texture analysis since it provides unique characteristics. In contrast to the fourier transform the WT is localized in the time domain as well as in the frequency domain. Thus it allows to derive localized contributions of energy to the textured signal in well separated frequency channels. This section will briefly investigate the WT and its suitability for the segmentation approach. For a deeper understanding of the theory the reader is referred to [4] and [15].

The wavelet transform $L^2(\mathbb{R})$ is an integral transform of the function $f(x)$

$$WT\,(a,b)\;=\;\int_{-\infty}^{\infty} f(x)\;\psi_{a,b}\;dx \qquad\qquad (\,1\,)$$

where the orthogonal *wavelet* $\psi_{a,b}$ is employed as a basis function of the operation. It is derived by a dilation a and a translation b from an unique mother wavelet. The discrete and dyadic transform WT^D is realized using discrete factors

$$a\;=\;2^m \qquad\quad m \in \mathbf{Z} \qquad\qquad (\,2\,)$$

and

$$b\;=\;n2^m \qquad\quad n,m \in \mathbf{Z} \qquad\qquad (\,3\,)$$

where m represents the depth of iteration. With

$$\psi_{m,n}\,(x)\;=\;2^{-\frac{m}{2}}\;\psi\;(2^{-m}\,x\,-\,n) \qquad\qquad (\,4\,)$$

one can define a family of wavelet basis functions. The following equation must hold for the corresponding mother wavelet

pipeline considers the coherence of neighbored slices in tomographic data. Thus where the feature extraction method is not extended for a 3D window, features are accumulated in the feature vector. Usually the data is analyzed as a multi–modal image and texture features from windows on different bands of the dataset are contributing to a single feature vector, describing the voxel position of the 3D dataset (see Figure 3).
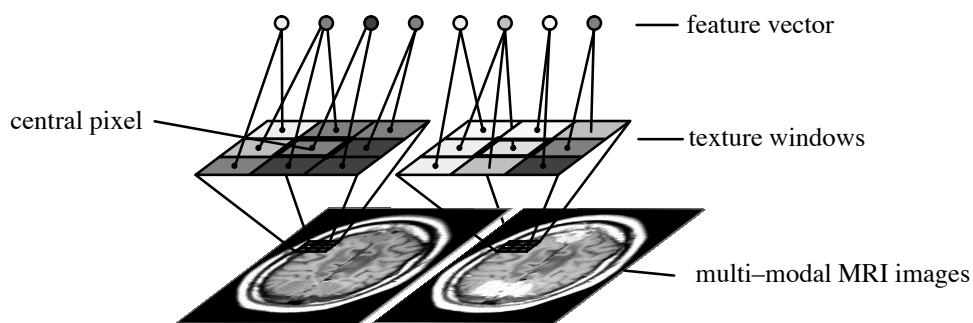
Figure 3. Texture windows on multi–modal images

Traditional work on texture description and analysis is driven by a statistical understanding of texture [8]. Early approaches employed first and second order statistics. Others investigated spectral properties or textural energy [13]. More recently texture analysis research has been influenced by the research done on wavelet transforms [15]. The common interest is due to the excellent properties of the transformation such as the localization of the representation and the multi resolution analysis capabilities. So far promising results have been achieved, e.g. the work of Chang [3]. In contrast to the work mentioned the focus of this paper is, to apply the wavelet transformation as a feature extraction method on small sized texture windows. The following sections will briefly review the wavelet transform.

3.1. Wavelet Transform

Recently respect for the wavelet transform (WT) has been growing in the computer graphics community. Due to its excellent properties it has been involved in a variety of applications

or statistically weighted combinations with respect to class–specific reliability of single classifiers are appropriate for this task. This subject will be covered in Section 4.

While the second step might be optional, the third step incorporates an important postprocessing of the classified result. Its objective is to detect and eliminate single misclassifications as isolated pixels or failures stemming from mixel effects at texture borders. This postprocessing requires a minimum size for a region to survive and thus provides a segmentation result with homogenous and smoothed areas. From these areas contour lines can be taken for further processing, e.g. generating object surfaces for 3D visualization. This postprocess filtering is performed with morphological operators, which are extended for color–coded images in order to perform a class–overlapping filtering and to incorporate basic anatomical knowledge. The mathematical foundations are therefore reviewed in Section 5 and applied to the segmentation problem of brain tissues.

3 Feature Extraction

The segmentation pipeline that was proposed in the previous section is based on texture oriented analysis of tomographic images. As outlined in Section 1 the focus of this work is to perform the desired segmentation on a very limited number of bands from a multi–modal tomographic dataset. Thus the approach balances the reduced amount of information with information extracted from the data itself. When less information is contributed from a number of bands, more information must be derived by the feature extraction method by taking the neighborhood of the respective pixel position into account. This corresponds to the common understanding of texture as an attribute representing the spatial arrangement of the gray levels of pixels in a region [20]. The region may be investigated by different texture analysis techniques which will be the subject of this section. The investigated region is called *texture window* and will be configured in its size by the different techniques. In general the segmentation

tion reliability it can be shown that the class assignment step performed by the Kohonen Feature Map can be optimized in two efficient ways: first, due to the homogeneity of the data vector to be classified there is a high probability that two subsequent vectors will result in the same class decision. Taking geometrical relations in the feature space into account one can exclude most of the feature vectors from consideration. These geometrical aspects are realized by precalculated hyperspheres around the cluster centroids in the feature space and by effective use of the triangle relation. Furthermore minimal distances can be approximated for proper initial values of the search algorithm. These aspects which are implemented in the presented segmentation pipeline are described in detail in [21]. It was shown that these modifications of the search algorithm of the Kohonen Feature Map reduce the required computation time to about 5%. Second, the calculation time can be further reduced by implementing the Kohonen Feature Map on an appropriate supercomputer. Vector processors are adequate to handle the kernel operations of the KFM since they require a small number of SIMD steps (Single Instruction Multiple Data) for the processing of a feature vector. The implementation of the time consuming texture classification step of the designed pipeline on an adequate supercomputer (Siemens–Fujitsu S400/40) with two vector processors, as compared to a medium sized workstation (HP715/50), reduces the required processing time down to 2% with respect to the kernel operations of the KFM.

The second step of the segmentation pipeline combines classification results which were calculated for the same patient dataset to a single united one. The idea is, to apply an ensemble of networks [12] to the problem. This approach offers two advantages: on the one hand the discriminating capabilities of different feature extraction methods are combined into a single result. On the other hand the reliability of the result is improved since a collection of classifiers will reduce the risk of failures stemming from a specific feature extraction technique. Combining the different results R^k to a united result V requires a combination rule. Boolean operations

using reliable training data [2]. Therefore large numbers of samples are required which are defined by experts and stored in a database. This database then is used in the training phase for the organization of the network. Figure 2 shows the two phases of the classification process: In Phase 1 the KFM is organized on the data stored in the database. In Phase 2 the KFM is applied to comparable image data performing a class assignment for each pixel. For a good generalization capability of the network one must provide a large number of training samples for the supervised organization. Thus a collection of image patterns that are attributed with class assignments is stored in a database. Since the pipeline shall rely on a very limited number of input bands, feature extraction methods are incorporated in the classification step in order to enhance the discrimination of patterns. The pipeline includes the wavelet transform for feature extraction methods (see Section 3) as well as some classical methods. The suitability within a tomographic application is investigated in Section 6 of this paper.
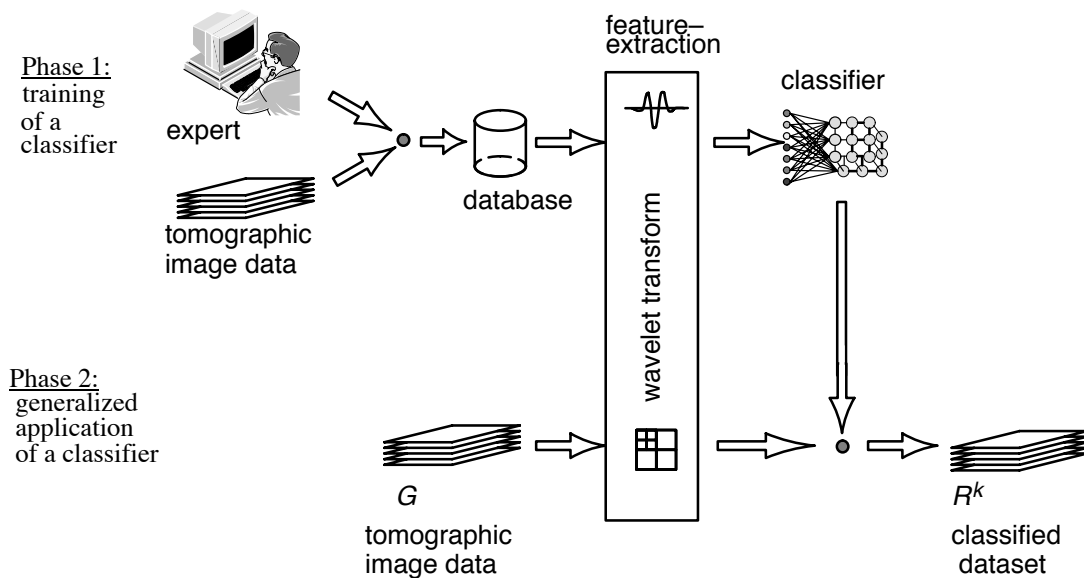

Figure 2. Texture classification according to the first step of the segmentation pipeline

A major issue in the segmentation pipeline is the system performance, i. e. the time required for feature extraction and the class assignment of the respective feature vector in the application phase. While computational expense of feature extraction is often balanced by classifica-

interaction in the analysis process. This excludes semi–automatic techniques such as region–growing.

Due to this constraint the segmentation is based on a fully automatic three–step pixelbased approach, where no user interaction is necessary. The main steps of this process are shown in Figure 1.
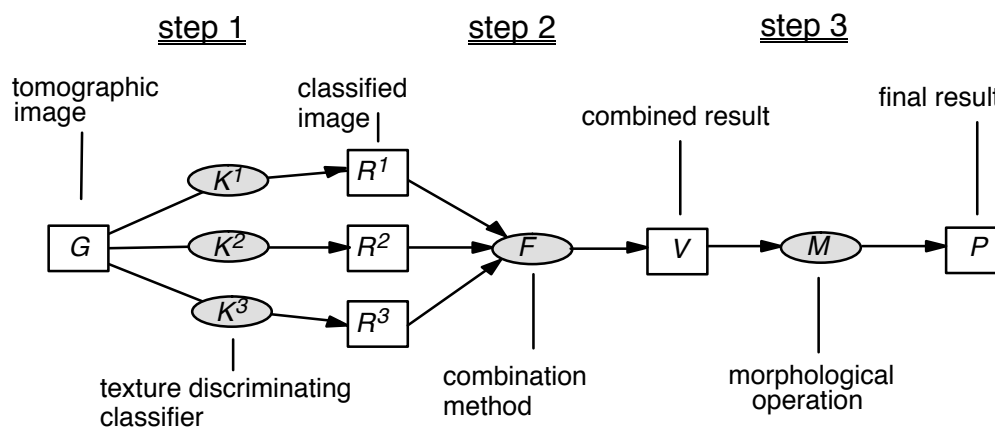
Figure 1. Three–step segmentation pipeline

The different steps of the pipeline can be described as follows:

- Step 1 integrates texture analysis techniques and an effective classifier scheme.

- Step 2 applies an ensemble of classifiers to a given dataset, where each classifier incorporates a single feature extraction method.

- Step 3 performs a postprocessing of the combined result while controlling the pixel–oriented approaches based on anatomical knowledge.

The first step of the pipeline analyzes the presented data and assigns of each pixel to a member of the predefined class set. This step uses the Kohonen Feature Map (KFM) which has proven its suitability to medical image analysis in a variety of studies [11]. Even though most applications use the KFM for cluster analysis of subspace mapping it can be shown that after self–organization of the network an additional learning vector quantization step [11] provides a supervised classifier. This step performs a class assignment to each neuron which is realized

spin density (pw weighted image band) characteristics of the tissue. In order to compensate for the lack of information compared to multiple input bands, the acquired images are analyzed using texture analysis techniques for proper discrimination of tissue types.

Section 2 presents a segmentation pipeline which separates an MRI dataset into seven different tissue types. While a Kohonen Feature Map is used as a classification scheme throughout this work the wavelet transform as a feature extraction method is investigated in Section 3. Section 4 will consider ensembles of classifiers and Section 5 reviews a postprocessing method for color–coded images. Section 6 will show the generalizing classification capabilities of the wavelet transform as compared to traditional texture describing methods on a test set of 11 comparable clinical studies.

2 Segmentation Pipeline

This paper focuses on the automatic analysis of tomographic data acquired by MRI scanners. The intention is to support the diagnosis by computer based segmentation of different tissue types, such as the classes *tumor, liquor* (cerebospinal fluid), *white matter, grey matter, bone, fat tissue* and *background*. For this purpose a pipeline was created that localizes pathological tissue (*tumor*) in a given dataset. As a result the process provides a color–coded dataset in which each pixel is classified as one of the classes defined previously.

An important aspect for the design of this pipeline is the integration of the segmentation process into a teleconsultation system [7]. In that case the segmentation process provides remote image analysis within a network of user groups linked by ISDN lines to a central super computer. In this scenario the site which acquires the image data is locally separated from the site where the data is analyzed. This implies a specific prerequisite for the design of the image processing pipeline: Due to offline execution at a locally separated supercomputer there is no user

1 Introduction

Segmentation techniques of multi–modal tomographic data is an increasingly important step in medical imaging. On the one hand magnetic resonance imaging (MRI) supplies high quality tomographic data characterizing specific tissue types with varying physical acquisition parameters. On the other hand computer tomography data (CT) can be matched with MRI data using algorithm such as [17] in order to provide multi–modal image information. The classification of tissue types in these datasets contributes to the diagnosis process in the field of volumetric rendering as well as surface reconstruction; both rely on a robust presegmentation of the initial data. This is required in order to compute a high quality 3D visualization of pathological tissue that might be used for irradiation or surgery planning. Furthermore, volume measurements are essential in order to control, for example, brain tumor regression in correlation with irradiation treatment [18].

Applying image analysis techniques to multi–modal images often involves interactive or semiautomatic techniques [9]. Nevertheless there are a number of automatic segmentation approaches which are applying multi–dimensional image analysis techniques especially to MRI data [19], [6]. Common drawbacks are either the requirement for a large number of variations as input data for the classification method or overlapping clusters in the feature space. Although MRI could provide a high number of bands of multivariate nature these recording sequences are not suitable for clinical use since the necessarily long acquisition times are too stressful for a patient.

The overall goal of this work is to reduce the amount of data that needs to be acquired. In other words the presented segmentation pipeline aims at a reduction of the necessary input information to those datasets used in everyday clinical studies, where double–echo sequences of MRI scans record images that are describing spin–spin relaxation (T_2 weighted image band) and

Wavelet Based Texture Segmentation
of Multi–Modal Tomographic Images

Christoph Busch

Computer Graphics Center

Wilhelminenstraße 7

D–64283 Darmstadt

Tel: +49 6151 155 236

Fax: +49 6151 155 480

E–mail: busch@igd.fhg.de

Abstract

This paper presents a segmentation pipeline for computer–based automatic analysis of multi–modal tomographic images. It is a computer based support for the localization of pathological tissues such as brain tumors. The segmentation pipeline of the presented approach includes texture analysis, classification with a modified Kohonen Feature Map, a collection of classifiers and knowledge based morphological postprocessing. Furthermore this paper presents a statistical investigation that compares the wavelet transform to classical texture analysis methods. Patient data which was acquired using magnetic resonance imaging (MRI) and computer tomography (CT) is used for this investigation.

Key Words: texture analysis, wavelet transform, kohonen feature map, magnetic resonance imaging, color–coded images, morphological operations