A NEW HYBRID ALGORITHM BASED ON WATERSHED METHOD, CONFIDENCE CONNECTED THRESHOLDING AND REGION MERGING AS PREPROCESSING FOR STATISTICAL CLASSIFICATION OF GENERAL MEDICAL IMAGES

Gerald Zwettler[a], Werner Backfrieder[a,b]

[a]Bio- and Medical Informatics, Research and Development Department,
Upper Austria University of Applied Sciences, Austria
[b]School of Informatics, Communication and Media, Upper Austria University of Applied Sciences, Austria

[a]gerald.zwettler@fh-hagenberg.at, [b]werner.backfrieder@fh-hagenberg.at

ABSTRACT

Segmentation of morphology in medical image data is a highly context specific and differs from various imaging modalities, necessitating the use of sophisticated mathematical models and algorithms to achieve good results. In this work an algorithm is presented for pre-segmentation of general medical input data, based on a watershed-segmentation strategy utilizing both, original intensities and derived gradient magnitudes for region growing. The number of resulting pre-classified regions is iteratively reduced to a user-defined threshold using merge metrics, accounting for the similarity of intensity profiles of two neighboring regions to merge, as well as the height of the gradient barriers to overcome and geometric aspects like sphericity and size of the border area with respect to the total region size. Based on such a context-independent pre-segmentation, the resulting manageable number of regions can be further merged and classified, utilizing texture features and a priori statistical models. Results are presented from brainweb database.

Keywords: watershed-segmentation, statistical image classification, texture features

1. INTRODUCTION

Accurate and automated segmentation of medical image data is of high importance in a very broad range of medical applications. The evaluation of geometric properties like position, size and extent of anatomical structures necessitates previously performed image segmentation and is of high relevance in the domain of surgery planning as well as disease progression. Based on such an anatomical classification, e.g. in case of liver tumors, the position of the lesion with respect to the supplying vessel systems as well as the parenchyma size and shape can be assessed prior to liver lobe resection (Zwetter, Backfrieder, Swoboda, and Pfeifer 2009). When combining data from different imaging modalities like CT/MRI for high anatomical resolution and SPECT/PET from the functional imaging domain, metabolic activity can be quantified utilizing segmentation masks of the corresponding anatomical region (Beyer, Schwenzer, Bisdas, Claussen, and Pichler 2010). Evaluation of therapy success and disease progression gets feasible in a quantitative way if morphological segmentations of the same patient at different points in time are available (Kuhnigk, Dicken, Bornemann, Bakai, Wormanns, Krass, and Peitgen 2006). Furthermore, in virtual reality scenarios, patient-specific segmentation models based 3D tracking and navigation, true 3D vision and rapid prototyping of haptic patient models (Wulf, Vitt, Gehl, and Busch 2001; Torres, Staskiewicz, Sniezynski, Drop, and Maciejewski 2001) are highly applicable for the task of surgical training and planning (Stone 2011).

To achieve accurate segmentation of particular organs and anatomical structures, semi-automated concepts like region growing (Gonzalez and Wintz 1987) or live-wire contour definition (Barrett and Mortensen 1997) can be applied. Due to the high demand for user-interaction these concepts are improper for clinical use. Although requiring rarely no a priori knowledge and thus favoring generic segmentation, utilizing these strategies only rather awkwardly shaped structures with homogenous intensities can be segmented. Although not directly applicable for medical applications, these concepts are of high relevance preparing manual reference segmentations for model-driven segmentation approaches.

Utilizing deformable models (McInerney and Terzopoulos 1996) and incorporating a priori knowledge, particular structures with low anatomical variability can be precisely segmented. The generic use of deformable models for segmentation of arbitrary anatomical structures is not practicable, because the model parameters need proper adjustment to the target morphology and fine structures showing inhomogenous intensities remain hard to segment. Statistical Shape Models (Cootes, Taylor, Cooper, and Graham 1992) can be automatically derived from a set of reference segmentations, allowing for the use of geometric features of the target structures to segment. If trying to model several structures utilizing statistical shape models, the problem of shape overlapping as well as topological changes remain unconsidered. Furthermore,

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2012
978-88-97999-13-3; Backfrieder, Bruzzone, Longo, Novak, Rosen, Eds.

73

modeling very thin and sparsely connected structures with respect to low imaging resolution or partial volume effects cannot be achieved with geometric shape modeling. Active appearance models (Cootes, Edwards, and Taylor 1998) introduce statistical properties of the targets structure expected intensity profile besides geometric features, thus being more robust in case of anatomical variability. Segmentation of small, loosely connected structures and changes in topology remain still unaddressed. Utilizing level sets (Osher and Sethian 1988), topological changes and anatomical variability can be handled, but adapting the steering parameters curvature, propagation and advection, only rather compact structures can be modeled. Furthermore, the level set parameters must be adjusted for each structure to segment, rather than deriving them from a set of reference segmentations.

To facilitate generic segmentation of arbitrary anatomical structures, a context and modality invariant pre-segmentation is required as fundament for texture based merging for classification. Classifiers like fuzzy connectedness (Gammage and Chaudhary 2006) and confidence connected thresholding (Ibanez, Schroeder, Ng, and Cates 2003) necessitate initial seed points for separating clusters of homogenous intensity distributions. For application of k-means clustering (Kanungo, Mount, Netanyahu, Piatko, Silverman, and Wu 2002), no seeds are required but regions of the expected k clusters are defined according to expected constant intensity profiles within the entire input image volume.

In this work we present a generic concept for pre-segmentation on arbitrary medical image input data. Starting at local minima positions in a gradient magnitude representation of the image volume, regions are grown similar to the watershed algorithm (Vincent and Soille 1991; Beare and Lehmann 2006). To overcome limitations of the watershed algorithm along homogenously and slightly increasing areas, the input image is incorporated during the region growing process. The initially classified regions (catchment basins) are not merged solely due to watershed level tolerance, but a metric, also considering region neighborhood, geometric properties and similarity of intensity profiles. Thus, an arbitrary input image can be pre-classified independent of context and imaging modality at a user-specified number of target regions, defining granularity of the pre-processing step. Based on this pre-segmentation, the final classification can be performed, utilizing statistical feature values automatically extracted from the set of reference segmentations.

2. DATA
For testing of generic pre-segmentation, a number of $n=20$ T1-weighted MRI datasets from the simulated *brainweb* database (Cocosco, Kollokian, Kwan, and Evans 1997; Kwan, Evans, and Pike 1999) and associated reference segmentations are used.

Further test runs and validations are performed utilizing $n=12$ anonymous multi-modal patient studies, comprising morphologic image acquisitions (T1, T2, PD, …) as well as functional imaging (SPECT, PET). For patient data, the reference segmentations are achieved in a semi-automated way by applying image processing pipelines modeled and evaluated using the MeVisLab image processing platform (Ritter 2007; MeVis 2012). The image segmentation pipeline comprises filtering, region growing, mathematical morphology, image arithmetics and live-wire contour definition, see chapter 3.6 for more details.

3. METHODOLOGY
For the pre-segmentation strategy, in a first step input image data is filtered to smooth the intensity topography and suppress noise and artifacts. Later, the gradient magnitude is extracted as derivation of the original intensities for balancing differences in intensity level and reducing region definitions to their boundaries. Based on the gradient magnitude image, local minima regions are detected and enlarged utilizing region growing. Local minima forming autonomous regions are iteratively dilated with respect to a stepwise increased water surface. Finally for the grown regions image statistics are calculated, essential for pair-wise region merging until the number of pre-segmented regions falls below the target region count. For illustration of the processing chain, see Fig. 1.
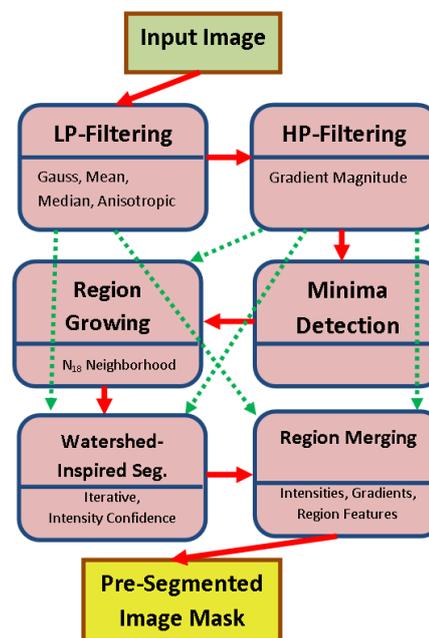

Figure 1: The sequential process chain (solid red arrows) comprises low-pass filtering, high-pass filtering, minima detection region growing, the watershed-type segmentation procedure and finally region merging until the convergence criterion is reached. Filtered input image and the gradient magnitude representation are required as input for several particular process steps (dashed green arrows).

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2012
978-88-97999-13-3; Backfrieder, Bruzzone, Longo, Novak, Rosen, Eds.

74

3.1. Filtering

As all local minima of input image derived gradient magnitude representation result in an additional initial region to grow, their number has to be limited by applying some lowpass filtering for smoothing the intensities. Limiting the number of autonomous regions prior to the final merge procedure is essential with respect to runtime and memory resources.

For lowpass filtering classic filter methods like mean, Gaussian and the median rank filter are evaluated. As the image intensity smoothing should not be performed at the expense of weakened edges and reduced inter object contrast, diffusion filtering has to be applied. The applied filters are taken from the ITK image processing library (Ibanez, Schroeder, Ng, and Cates 2003).

The filter configuration discussed in the following section refers to an input image with iso-voxel spacing of *1mm* in each dimension. In case of data with different spacing configuration, additional noise or artifacts to handle, an individual additional pre-processing is required to prepare the data at sufficient quality and with properties needed.

The major smoothing and fine tuning is accomplished applying several diffusion filter runs. Modules `itk::GradientAnisotropicDiffusion` and `itk::CurvatureAnisotropicDiffusion` are sequentially applied with parameterization *(6;0.125;10.0)* and *(3;0.125;10.0)* for *iteration*, *time-step* and *conductance* value. Utilizing both, gradient and curvature preservation aspects in diffusion equation, leads to balanced results, compared to an increased number of iterations for only one diffusion filter type in the image processing chain.

Besides the diffusion filtering, median filter module `itk::MedianImageFilter` is utilized with *radius 1* at the end of the filtering pipeline to minimize local outliers, preserved during diffusion filtering.

3.2. Gradient Magnitude Calculation

Besides the input image intensities, a gradient representation indicating closed borders between neighboring separated regions is required.

For choice of the gradient calculation strategy, impermeability between two neighboring regions to separate is the most important criterion. Impermeability is achieved, if the gradient values of an enclosed region don't show any kind of low gradient value by-pass connections to neighboring regions. In case of smooth intensity changes, only gradients of low magnitude remain to separate the regions, see Fig. 2. The weakest gradient borders delimit applicability of gradient-based segmentation strategies. To check the level of impermeability, the possible flood-fill level for well defined regions is evaluated. For details please refer to the results section.

The `itkGradientMagnitudeImageFilter` filter algorithm is applied for deriving the edge representation from input image. Gradient magnitude values are calculated at floating-point precision. As discrete values are required for further merging neighboring pixels of same value into larger regions, the gradient magnitude is normalized to gradient range *[0;maxVal]*, with *maxVal=200* and rounded to integer type. This normalization ensures a constant level of segmentation granularity, independent from pixel type of input image data.
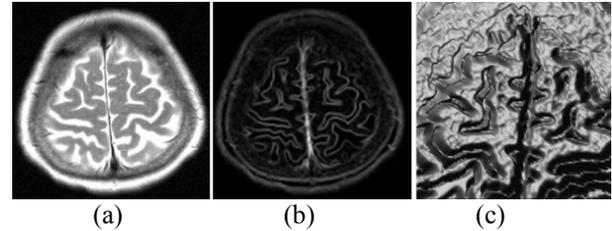

(a)   (b)   (c)

Figure 2: Transversal brain slice (a) and calculated gradient magnitude (b). Although brain windings are mostly enclosed by high gradient magnitude values, the smooth intensity decline in ventral direction leads to smaller gradients along the borders, as illustrated utilizing a depth profile (c).

3.3. Local Minima Detection

Each local gradient minimum in N_{26} is a possible starting position of a new region to be grown in the course of watershed-type pre-segmentation. For local minimum detection, all pixels in the gradient magnitude representation of the input image are evaluated with respect to their direct neighbors. If all neighbors show equal or larger values, the current position is marked as local minimum, see Equ. 1.

In areas of homogeny gradient values, an over-classification of local minima is inevitable. Consequently, the marked positions are only region start positions for following initial region growing and candidates to form up a new region. The number of regions to grow is therefore significantly smaller compared to the total number of detected minima.

$$LM(p) := \forall q \in N_{26}(p) \mid I(q) \geq I(p) \qquad (1)$$

3.4. Initial Region Growing

The detected local minima positions are seed point candidates for growing regions of identical gradient magnitude utilizing N_{18} neighborhood.

Besides gradient magnitude as region growing criterion, similarity in filtered intensities is incorporated as second aspect for an adjacent voxel to be added to a particular region. The intensity topography must be accounted for, as regions of same gradient magnitude can show huge differences in intensity values due to marginal but steady changes.

A threshold *T* is defined for the acceptable difference in intensity of a region to be grown with respect to matching gradient magnitude values. A voxel *v* can be added to an adjoined region *R*, if intensity limits after union are still within the defined tolerance interval *T*.

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2012
978-88-97999-13-3; Backfrieder, Bruzzone, Longo, Novak, Rosen, Eds.

75

$$\text{grow}(v, R) := \begin{array}{l} N_{18}(v, R) \wedge \text{grad}(v) == \text{grad}(R) \wedge \\ \left(\left| \text{maxI}(v \cup R) - \text{minI}(v \cup R) \right| \leq T \right) \end{array} \quad (2)$$

Due to incorporation of the input image intensities, large differences in intensities of a grown region can be prevented.

In arbitrary medical image data, the variance of local intensities is expected to be quite inhomogeneous, e.g. lower variance σ for background intensity distribution compared to objects and structures belonging to the patient's anatomy. Standard deviation of local intensities can be calculated for all voxel positions, thus the generation of a standard deviation map, indicating image areas with slightly homogenous voxel values and areas with high variance, is performed. To balance these local differences in the intensity profile of the entire image I, an adaptive region growing algorithm will segment the initial regions based on a tolerance interval T_R, individually calculated and updated for each region R.

$$\text{aGrow}(v, R) := \begin{array}{l} N_{18}(v, R) \wedge \text{grad}(v) == \text{grad}(R) \wedge \\ \left(\left| \text{maxI}(v \cup R) - \text{minI}(v \cup R) \right| \leq T(r) \right) \end{array} \quad (3)$$
$$T(R) := \quad T \cdot \left(\frac{1}{2} + \frac{\text{meanStdDev}(R)}{2 \cdot \text{meanStdDev}(I)} \right).$$

The standard deviation value for each pixel is calculated as mask operation with radius *r=5*, incorporating all neighboring voxels at a maximum Euclidean distance of *r*. Results are smoothed applying an *11x11x11* average kernel. With this adaptive region growing strategy, similar granularity of pre-segmented regions can be accomplished, balancing local variation in intensity standard deviation.

3.5. Hybrid Watershed-Inspired Pre-Segmentation

Based on the regions resulting from initial region growing, the tolerance level for gradients is iteratively increased, thus processing the border elements of the regions and growing neighbors with a matching gradient value. Only voxels with a gradient value of exactly the particular region start gradient value (from the local minima) plus the current gradient increment are candidates for growing. The gradient tolerance level is increased when none of the regions can grow at the current gradient tolerance level. The described algorithm refers to classic watershed segmentation and is labeled as GRG_WS.

To stronger incorporate the intensity features, the grow conditions described in Equ. 2 and Equ.3 are also applied for growing the regions, resulting in a static intensity interval segmentation strategy (SII_WS) and an adaptive intensity interval segmentation strategy (AII_WS), see Code Listing 1.

Introducing the new intensity conditions for the growing process, it is not assured that all voxels can be classified. Remaining voxels are iteratively assigned to the neighboring region with lowest difference in

intensity. The intensity tolerance limit is set to 2 and is temporarily increased, whenever in the course of one full iteration, not a single voxel can be assigned to one of the existing regions, see Code Listing 2.

```
regions = initialRegionGrowing;
int gradientTolerance = 1;
while(gradientTolerance < maxGradVal)
  while(stillChanges)
    for(region r : regions)
      for(voxel u : r.border)
        for(voxel v : u.N18)
          if(unclassified(v) &&
              grad(v) == (r.initGrad
                + gradientTolerance) &&
              aGrow(v, R))
            r.addToRegion(v);
            v.classVal = r.regID;
            r.addToBorderVect(v);
            stillChanges = true;
  gradientTolerance++;
  for(region r : regions)
    r.updateBorderVect(gradientTolerance);
```

Code Listing 1: Illustration of the watershed-inspired pre-segmentation algorithm based on adaptive intensity intervals (AII_WS).

```
int intensityTolerance = 2;
while(unclassifiedVoxel.size > 0)
  bool changes = false;
  for(voxel u : unclassifiedVoxel)
    for(voxel v : u.N18)
      region r = region(v);
      if(r != NULL &&
          insideTolerance(u.intVal, r,
            intensityTolerance))
        r.addToRegion(u);
        u.classVal =r.regID;
        r.updateIntToleranceInterval();
        changes = true;
  if(changes)
    intensityTolerance = 2;
  else
    intensityTolerance += 2;
```

Code Listing 2: Algorithm for iteratively assigning the voxels remaining unclassified during AII_WS to a neighboring region.

3.6. Calculation of Region Features for Merging

Based on the results of the pre-segmentation, the number of regions has to be significantly reduced applying region merging strategies. For classic watershed algorithm, a tolerance value for differences in the starting gradient value of neighboring regions is the key criterion for merging two separated regions *u* and *v* into one. This concept lacks steering of the resulting number of regions to remain.

For the presented intensity interval and adaptive intensity interval pre-segmentation strategy, besides the gradient value *gradW(u,v)*, also differences in the regions intensities as *intW(u,v)* and length of the borders as geometric property *borderW(u,v)* and indicator for sphericity are taken into account, with *surf(u,v)* as voxel count on the border between regions *u* and *v* and *surf(u)* as the total count of neighboring region voxels of a region *u*.

$$intW(u,v) := \frac{(meanInt(u) - meanInt(v))^2}{maxIntTolerance},$$ (4)

$$maxIntTolerance = \frac{(maxInt(I) - minInt(I))}{10}$$

$$gradW(u,v) := \frac{\max\begin{pmatrix} meanG(u,v) - \min G(u), \\ meanG(u,v) - \min G(v) \end{pmatrix}}{\max GradTolerance},$$ (5)

$$maxGradTolerance = \frac{(maxG(I) - minG(I))}{10}$$

$$borderW(u,v) := \min\begin{pmatrix} \dfrac{surf(u) - surf(u,v)}{surf(u)}, \\ \dfrac{surf(v) - surf(u,v)}{surf(v)} \end{pmatrix}$$ (6)

Thus, for every pair of neighboring regions *u* and *v* connected at a border *b*, the discussed features can be evaluated and a numerical merge feature value *mergeW(u,v)* can be evaluated, thus indicating similarity and weight for merging regions *u* and *v*. The weighting factors $w_g=2$, $w_i=7$ and $w_b=1$ lead to good results as combining the different aspects and outperforming each single criterion.

$$mergeW(u,v) :=$$
$$w_g \cdot gradW(u,v) + w_i \cdot intW(u,v) + w_b \cdot borderW(u,v).$$ (7)

Based on a sorted and steadily updated list of merge values, the total region count can be reduced by particularly merging the region pair showing the lowest merge value, until the target number of *R* remaining regions is reached.

3.7. Semi-Automated Reference Segmentation

For semi-automated generation of reference segmentations, MeVisLab (MeVis 2012) rapid prototyping platform is utilized. Particular anatomical structures are segmented using the `LiveWire` module (Handels 2009; Barrett and Mortensen 1997) if the shape is well defined and `RegionGrowing` module is applied for segmentation of smaller structures lacking insularity. Besides the segmentation modules, anisotropic filtering for initial smoothing and morphological post-processing (opening/closing utilizing `Morphology`) in case of artifact occurrence are additional process steps in the image processing pipeline.

In contrast to brainweb reference segmentations, not the entire dataset will be semi-automatically segmented. For each imaging modality and test case, only particular anatomical structures, like gray matter, white matter and ventricle for brain MRI data, are chosen for algorithmic validation.

4. RESULTS

The presented watershed-inspired pre-segmentation strategy is tested utilizing the first *n=6* samples and associated reference segmentations of the brainweb database. The image data is scaled to 8bit and iso-voxel spacing at size *256x307x256* using *sinc*-window-based *Lanczos 3* interpolator (Burger and Burge 2008).

4.1. Evaluation of the Filter Chain

At first the filter chain is tuned with respect to the filter types and parameterization. As one quality criterion, the image mean intensity level should be kept constant. The smoothing effect can be estimated, evaluating changes on the voxel intensity's standard deviation and the average voxel change. Besides noise reduction, the key goal of this filtering process is to balance inhomogeneity in local voxel neighborhood and thus, significantly reducing the number of initial regions to be detected during the first region growing run. Finally, the classification result quality after executing the filter chain is evaluated utilizing the reference segmentations. A comparison of different filter chain settings is given in Table 1.

Based on the low number of initial regions, a small difference in mean intensity and preferably high precision, the chains Ga10 (0.01;145,622;0.98) and Ga6Ca3M1 (0.16;119,805;0.966) are best balancing the criterions, as charted in Fig. 3.

Table 1: Evaluation of different filter chains, utilizing average (A), median (M), gradient anisotropic diffusion (Ga), curvature anisotropic diffusion (Ca) and Gaussian filtering (G) at different kernel sizes.

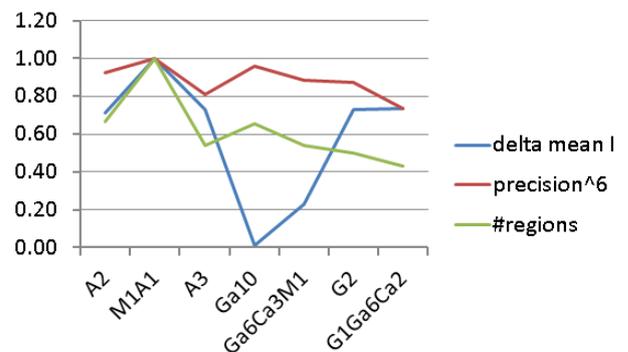| filter chain	ΔI	σI	ΔI_{mean}	#reg	prec
unfiltered	0.00	0.00	0.00	899,386	1.000
A2	6.05	8.28	0.50	148,393	0.974
M1A1	4.39	5.93	0.70	222,375	0.987
A3	7.66	10.76	0.51	119,767	0.952
Ga10	2.46	2.23	0.01	145,622	0.980
Ca10	8.95	34.07	5.09	898,315	0.982
Ga6Ca3M1	3.35	3.95	0.16	119,805	0.966
G2	6.97	9.71	0.51	110,991	0.964
G1Ga6Gc2	5.59	7.39	0.52	96,122	0.937

Figure 3: Charting of the three key features ΔI_{mean}, *#regions* and *precision* as ratio with respect to filter chain M1A1. Despite precision, a low feature value is desired.

With the discussed filter chain parameterization Ga6Ca3M1 smoothing of local intensity homogeneity can be achieved but borders and gradients are preserved, see Fig. 4.
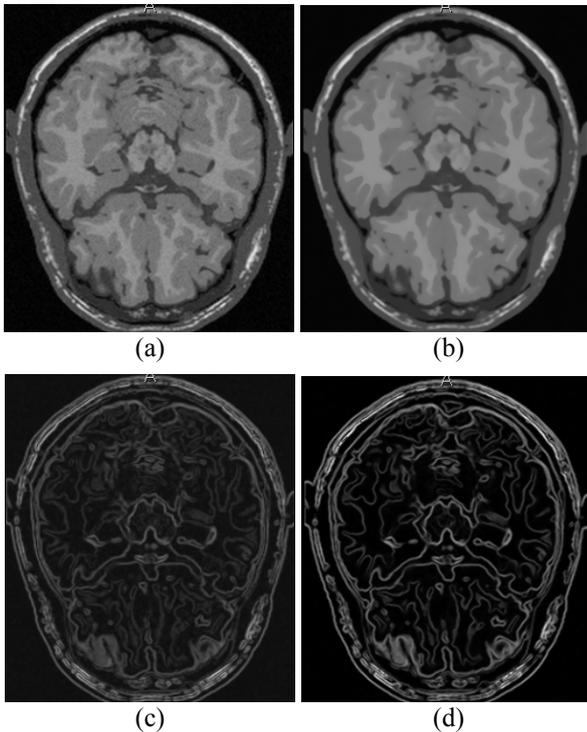

(a)   (b)

(c)   (d)

Figure 4: Input image intensities (a), *slice #76* of *subject04*, are significantly smoothed applying the filter chain (b). Furthermore, results of gradient magnitude calculation after smoothing (d) show less variance due to suppressed noise compared to gradient magnitude calculation on the original image data (c).

4.2. Tests on Initial Region Growing

Region growing is performed based on the gradient magnitude representation after processing the filter chain. For common watershed segmentation, neighbors in N_{18} with identical gradient value are grouped together. N_{18} adjacency demands two neighboring voxels to share at least one edge and thus excludes the eight corner diagonal members of N_{26}. Focusing on the gradient values leads to high variability in intensities, see Table 2. With our presented strategies for initial region growing, intra-region intensity homogeneity can be enforced.

For calculation of the voxel standard deviation map to be used for adaptive region growing (AII), a radius of r=5 with subsequent 11x11x11 average kernel application is sufficient for handling local intensity aspects, see Fig. 6.

As presented in Table 2, the restrictions on the intensity range for initial region growing significantly improve quality of the results. Despite a slightly increased number of regions, configuration AII4 best balances the different criterions, see Fig. 5. Enforcing regions to show identical gradient and intensity values (SII0) does not lead to improved results but an

increased number of regions to handle in the next processing steps.

Table 2: Comparison of gradient-based region growing (GRG) with static intensity interval restriction (SII) and adaptive intensity interval restriction (AII) at different interval sizes. Maximum inner region intensity difference (max ΔI) and mean region intensity difference (mean ΔI) indicate homogeneity of intensities. The grow ratio reflects, how many voxels are processed during this first classification run.

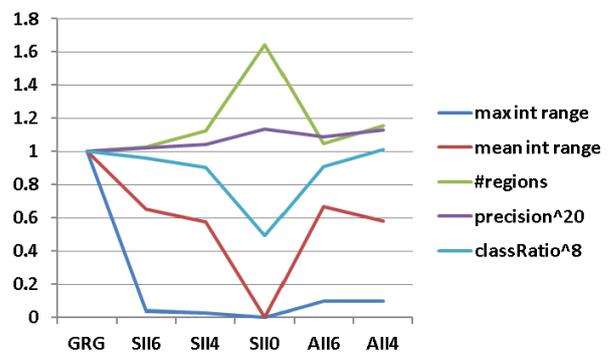| RG strategy	max ΔI	mean ΔI	#reg	grow ratio	prec
GRG	151	0.703	119,805	0.592	0.966
SII6	6	0.458	123,056	0.589	0.967
SII4	4	0.405	134,569	0.584	0.968
SII0	0	0.000	196,984	0.542	0.972
AII6	15	0.469	125,495	0.585	0.970
AII4	10	0.401	138,323	0.592	0.972

Figure 5: Charting of the features from Table 2 as ratio with respect to region growing strategy GRG. Despite precision and the classification ratio, a low feature value is desired.
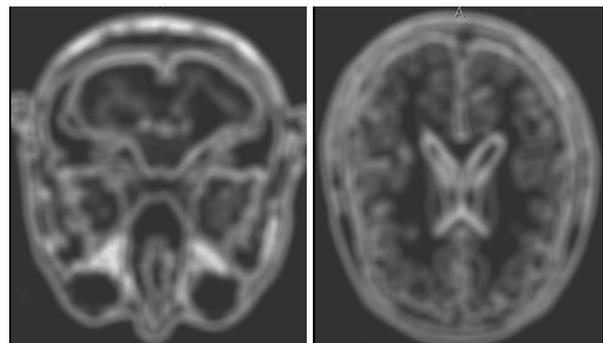

Figure 6: Two slices of the calculated standard deviation map at Euclidean radius r=5 with subsequent 11x11x11 average filter application. Results are smooth enough to adequately reflect local intensity variation.

4.3. Testing Watershed-Inspired Pre-Segmentation

Based on the results of initial region growing for GRG, SII6 and AII4, the results after classification of all voxels are analyzed, see Table 3. Both, static intensity interval and adaptive intensity interval significantly improve results compared to common gradient-based watershed segmentation. Furthermore, the mean inter-

region intensity range (mean ΔI) is reduced by a factor of 2.0, relevant for subsequent neighbor region merge process. The key criterions precision and mean inter-region intensity range can be both improved compared to GRG_WS, see Fig. 7.

Table 3: Comparison of common gradient-based watershed segmentation (GRG_WS) with the newly developed concepts for static intensity interval restriction (SII6_WS) and adaptive intensity interval restriction (AII4_WS). Maximum inner region intensity difference (max ΔI) and mean region intensity difference (mean ΔI) indicate homogeneity of intensities. The grow ratio reflects, how many voxels are processed during this first classification run.

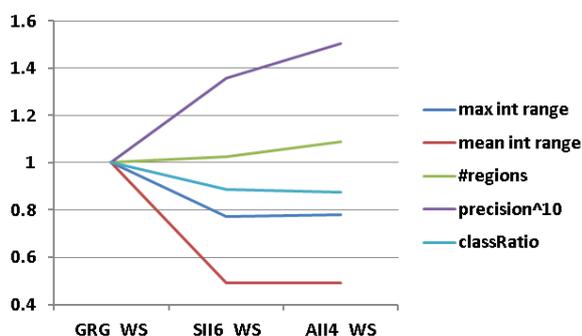| RG strategy	max ΔI	mean ΔI	#reg	grow ratio	prec
GRG_WS	215.5	21.974	119,955	1.00	0.883
SII6_WS	166.2	10.818	123,056	0.89	0.910
AII4_WS	168.2	10.768	130,466	0.87	0.920

Figure 7: Charting of the features from Table 3 as ratio with respect to GRG_WS watershed segmentation. High precision and a low value for both, mean and max intensity range are highly recommended.

Evaluating results in detail, GRG_WS is outperformed with respect to precision for each particular dataset. Furthermore, SII6_WS is outperformed by the adaptive AII4_WS for each particular brainweb sample, see Table 4 for the first *n=6* samples.

Table 4: Precision and mean ΔI for the first *n=6* brainweb datasets. Precision of SII6_WS higher compared to GRG_WS and precision of AII4_WS higher compared to SII6_WS. Same for decrease in mean ΔI, exempt from sample [5]sj20.

data	GRG_WS		SII6_WS		AII4_WS	
	mean ΔI	prec	mean ΔI	prec	mean ΔI	prec
[1]sj04	22.1	0.881	11.0	0.910	10.9	0.915
[2]sj05	20.6	0.888	10.4	0.915	10.3	0.924
[3]sj06	21.3	0.849	11.0	0.883	10.6	0.908
[4]sj18	24.9	0.885	11.7	0.913	11.7	0.920
[5]sj20	21.9	0.904	10.3	0.928	10.5	0.931
[6]sj38	21.1	0.889	10.6	0.913	10.6	0.921

Results after AII4_WS are presented in Fig. 8. Regions showing rather related region labels are visualized in similar colors. Due to a still large number of regions (>100,000), the coloring reflects a trend of growing IDs from left to right according to the processing order.
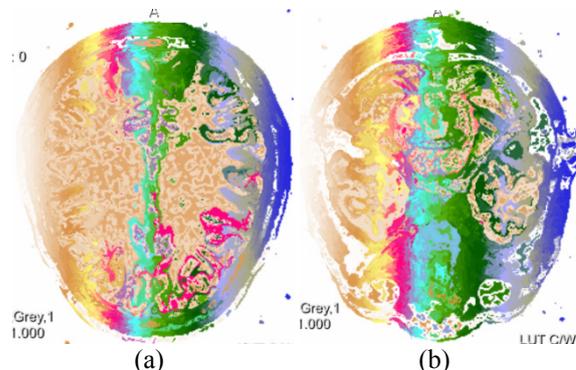

(a)   (b)

Figure 8: Color representation of regions resulting from AII4_WS. The MRI morphology is cognizable for slices #146 (a) and #55 (b) of dataset [1]sj04.

Segmentation errors produced by GRG_WS and AII4_WS are compared in Fig. 9. Most of the wrongly classified voxels result from filtering effects and deficits of the reference segmentation, like differentiating between air outside and inside the head, see Fig. 10.
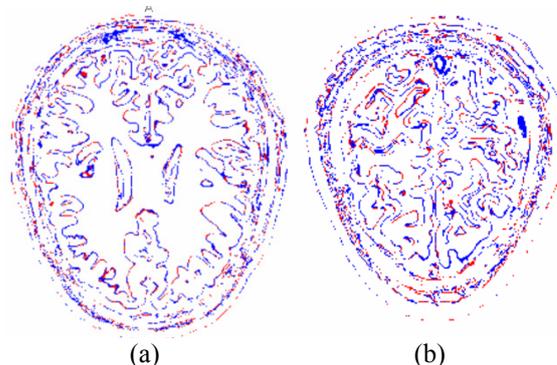

(a)   (b)

Figure 9: Color coded differences of errors produced from GRG_WS and AII4_WS for slices #140 and #182 of dataset [1]sj04. Discrepancies only appearing for GRG_WS shown in blue and those only resulting from AII4_WS in red.

4.4. Testing Calculation of Merge Features
Tests show that in contrast to classic watershed segmentation approach, the target number of remaining regions can be precisely pre-defined by the user.

Due to combination of different types of features, the achievable precision can be significantly increased, see Table 5 and Fig. 11. For common watershed segmentation only the tolerance in gradient level at the region borders is utilized as merge criterion. Results in Table 5 show that the watershed region merge metric is outperformed by incorporating intensity level and border ratio.
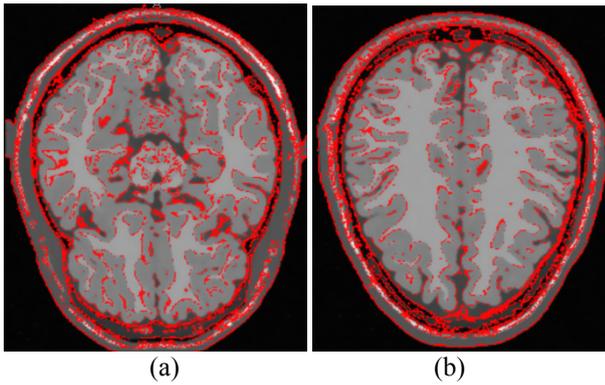
(a)   (b)

Figure 10: Error (red) of AII4_WS with respect to the original image data at slices #70 and #155 of dataset [1]sj04.

Table 5: Comparison of result quality utilizing different types of merge features. Analysis runs are performed based on AII4_WS pre-segmentation results. For the cumulated weight defined in Equation 7, the weights for intensity level similarity (squared diff in intensities), difference in gradient level and the border ratio are combined, thus leading to increased precision. The difference in mean region gradient level (squared diff in gradients) showed to be no proper metric for region merging.

merge strategy	#reg= 2500	#reg= 5000	reg= 7500
random (region IDs)	0.474	0.458	0.477
squared diff in intensities	0.809	0.803	0.836
squared diff in gradients	0.514	0.497	0.517
difference in gradient level	0.735	0.736	0.766
border ratio	0.622	0.603	0.629
cumulated weight	0.862	0.831	0.866

5. DISCUSSION

The incorporation of original image intensities in the process of region growing and gradient-based watershed-like segmentation showed to improve results significantly.

Due to the fact that the number of local gradient minima is low in anatomically relevant areas of high pixel intensity variance, further region seed candidates will be evaluated in future. Besides the gradient values, local minima will also be searched for, utilizing the input image intensity voxel mask. It is thereby expected to increase the number of region seed candidates by around 20-30%, mainly in the sparse area of high variance.

Furthermore the algorithm for iteratively assigning the voxels remaining unclassified during AII4_WS algorithm execution will be improved. If unclassified voxels show higher similarity to their neighbors being unclassified too, additional regions should be formed to keep the intra-region intensity range low.
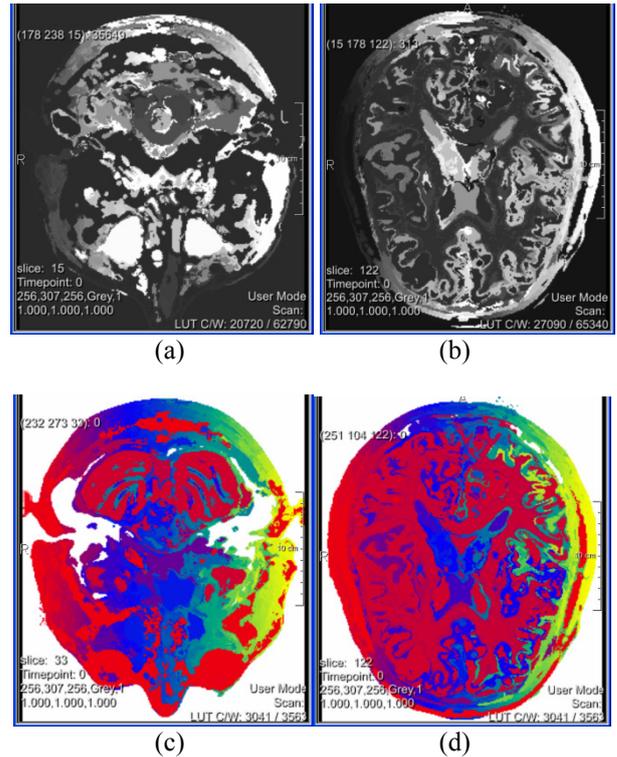

(a)   (b)

(c)   (d)

Figure 11: Final merged regions of dataset [1]sj04 after AII4_WS pre-segmentation and region merging utilizing cumulated metrics. Visualization with gray value colormap (a-b) and in color (c-d). Compared to Fig. 8, the color trend from left to right is replaced by solid and well defined regions representing anatomical structures.

REFERENCES

Barrett, W.A., Mortensen, E.N., 1997. Interactive Live-Wire Boundary Extraction. *Medical Image Analysis* 1:331-341.

Beare, R., and Lehmann, G., 2006. The watershed transform in ITK – discussion and new developments. *Insight Journal*, Jannuary – June 2006.

Beyer, T., Schwenzer, N., Bisdas, S., Claussen C.D., and Pichler, B.J., 2010. *MR/PET – Hybrid Imaging for the Next Decade. MAGNETOM Flash* 3/2010.

Burger, W., and Burge, M.J., 2008. *Digital Image Processing – An Algorithmic Introduction Using Java*. New York: Springer, 402pp.

Gonzalez, R.C., Wintz, P., 1987. *Digital Image Processing*. London: Addison Wesley.

McInerney, T., Terzopoulos, D., 1996. Deformable Models in Medical Image Analysis : A Survey. *Medical Image Analysis* 1 (2):91-108.

Cocosco, C.A., Kollokian, V., Kwan, R.K.-S., Evans, A.C., 1997. BrainWeb: Online Interface to a 3D MRI Simulated Brain Database. *NeuroImage* 5(4):425.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1992. Training Models of Shape from Sets of Examples. *Proceedings of the British Machine Vision Conference*, 9-18. September 22-24, Leeds (United Kingdom).

Cootes, T.F., Edwards, G.J., Taylor, C.J., 1998. Active Appearance Models. *Proceedings of the 5th European Conference on Computer Vision*, 484-498. June 2-6, Freiburg (Germany).

Gammage, C., and Chaudhary, V., 2006. On Optimization and Parallelization of Fuzzy Connected Segmentation for Medical Imaging. *Proc. of the 20th International Conf. on Advanced Information Networking and Applications*, 623-627. April 18-20, Vienna (Austria).

Handels, H., 2009. Medizinische Bildverarbeitung. Wiesbaden (Germany) : 116-124.

Ibanez, L., Schroeder, W., Ng, L., and Cates, J., 2003. *The ITK Software Guide: The Insight Segmentation and Registration Toolkit*. New York (United States) : Kitware Inc.

Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., and Wu, A.Y., 2002. An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 24:881-892.

Kuhnigk, J.-M., Dicken, V., Bornemann, L., Bakai, A., Wormanns, D., Krass, S., and Peitgen, H.-O., 2006. Morphological Segmentation and Partial Volume Analysis for Volumetry of Solid Pulmonary Lesions in Thoracic CT Scans. *IEEE Transactions on Medical Imaging* 25(4):417-434.

Kwan, R.K.-S., Evans, A.C., Pike, G.B., 1999. MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging* 18(11):1085-1097.

MeVis, 2012. *MeVisLab : medical image processing and visualization*, MeVis Medical Solutions, Bremen (Germany). Available from: http://www.mevislab.de/developer/documentation/ [04/2012].

Osher, S., Sethian, J.A., 1988. Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations. *Journal of Computational Physics* 79:12-49.

Ritter, F., 2007. Visual Programming for Prototyping of Medical Applications. *IEEE Visualization* 2007 workshop.

Stone, R.J., 2011. The (human) science of medical virtual learning environments. *Philosophical Transactions of the Royal Society* 366:276-285.

Torres, K., Staskiewicz, G., Sniezynski, M., Drop, A., and Maciejewski, R., 2011. Application of rapid prototyping techniques for modeling of anatomical structures in medical training and education. *Folia morphologica* 70(1):1-4.

Vincent, L., and Soille, P., 1991. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(6):583-598.

Wulf, J., Vitt, K.D., Gehl, H.-B., and Busch, L.C., 2001. Anatomic Accuracy in Medical 3D Modeling. *Phidias Rapid Prototyping in Medicine* Newsletter No. 7, December 2011.

Zwettler, G., Backfrieder, W., Swoboda, R., Pfeifer, F., 2009. Fast Fully-automated Model-driven Liver Segmentation Utilizing Slicewise Applied Levelsets on Large CT Data. *Proceedings of the 21st European Modeling and Simulation Symposium*, 161-166. September 23-25, Tenerife (Canary Islands, Spain).

AUTHORS BIOGRAPHY

Gerald A. Zwettler was born in Wels, Austria and attended the Upper Austrian University of Applied Sciences, Campus Hagenberg where he studied software engineering for medicine and graduated Dipl.-Ing.(FH) in 2005 and the follow up master studies in software engineering in 2009. In 2010 he has started his PhD studies at the University of Vienna at the Institute of Scientific Computing. Since 2005 he is working as research and teaching assistant at the Upper Austrian University of Applied Sciences at the school of informatics, communications and media at the Campus Hagenberg in the field of medical image analysis and software engineering with focus on computer-based diagnostics support and medical applications. His e-mail address is gerald.zwettler@fh-hagenberg.at and the research web page of the Research & Development department at campus Hagenberg, he is employed at, can be found under the link http://www.fh-ooe.at/fe/forschung.

Werner Backfrieder received his degree in technical physics at the Vienna University of Technology in 1992. Then he was with the Department of Biomedical Engineering and Physics of the Medical University of Vienna, where he reached a tenure position in 2002. Since 2002 he is with the University of Applied Sciences Upper Austria at the division of Biomedical Informatics. His research focus is on Medical Physics and Medical Image Processing in Nuclear Medicine and Radiology with emphasis to high performance computing. Recently research efforts were laid on virtual reality techniques in the context of surgical planning and navigation.

Proceedings of the International Workshop on Innovative Simulation for Health Care, 2012
978-88-97999-13-3; Backfrieder, Bruzzone, Longo, Novak, Rosen, Eds.

81