Coverage Segmentation by Energy minimization

– A practical example of sub-pixel image processing using the Coverage model

Nataša Sladoje and Joakim Lindblad

sladoje@uns.ac.rs, joakim@cb.uu.se

Faculty of Technical Sciences
University of Novi Sad, Serbia

CCVW 2012, Zagreb, Croatia

Coverage segmentation

- *Coverage segmentation* is the task of extracting coverage information about objects in images.
  - Estimate a vector of coverage values for each image element.
- Acts as a separator between imaging and feature extraction.
  - Ensures that further processing can be made independent of imaging modalities and imaging conditions.

Coverage segmentation

Five methods which provide (approximate) coverage images:

1. Direct assignment of coverage values from a **continuous model**.
   - A. Tanács, C. Domokos, N. Sladoje, J. Lindblad, and Z. Kato. Recovering affine deformations of fuzzy shapes. SCIA 2009. LNCS-5575, pp. 735–744, 2009.
2. A method based on **mathematical morphology** and a double thresholding scheme.
   - N. Sladoje and J. Lindblad. High Precision Boundary Length Estimation by Utilizing Gray-Level Information. IEEE Trans. on PAMI, Vol. 31, No. 2, pp. 357–363, 2009.
3. A framework (and methods) for coverage segmentations of **graphs**.
   - F. Malmberg, J. Lindblad, I. Nyström. Sub-pixel segmentation with the image foresting transform. IWCIA 2009. LNCS-5852, pp. 201–211, 2009.
   - F. Malmberg, J. Lindblad, N. Sladoje, I. Nyström. A Graph-based Framework for Sub-pixel Image Segmentation. Theoretical Computer Science. Vol 412, No 15, pp. 1338-1349, 2011.
4. A method providing **local sub-pixel classification** extending any existing crisp segmentation.
   - N. Sladoje and J. Lindblad. Pixel coverage segmentation for improved feature estimation. ICIAP 2009. LNCS-5716, pp. 929-938, 2009.
5. An **energy based** method for regularized coverage segmentation.
   - J. Lindblad and N. Sladoje. Coverage Segmentation Based on Linear Unmixing and Minimization of Perimeter and Boundary Thickness. Pattern Recognition Letters. Vol 33, No. 6, pp. 728-738, 2012.

Problem formulation

- We observe a 2D digital image I with b spectral bands.
- Let N be the number of pixels of the image and let the image data be given as a matrix $I = [p_{i,k}]_{N\times b}$ such that a **row** contains intensities of **one pixel** in each of the observed bands, and a **column** represents the pixel intensities in **one band**, over the whole image.
- Our goal is to obtain a coverage segmentation of I corresponding to m classes (objects) existing in the image, i.e., each pixel is assigned a vector of length m whose components give the relative area of the pixel covered by each of the m classes.
- A coverage segmentation of the image I is a matrix $A = [\alpha_{i,j}]_{N\times m}$ where $\alpha_{i,j} \in [0, 1]$ is the coverage of the pixel with index i ($i = 1, 2 \ldots, N$) by a class (object) S_j. Assuming spatially non-overlapping classes S_j each row of A sums up to one.

Linear unmixing

Models based on linear unmixing of image intensities are common in the field of image processing, due to simplicity and wide applicability.

- We model the image intensities I as a **non-negative linear mixture** (a convex combination) **of pure class representatives** (a.k.a. end-members).[1]
- The pure class representatives can be written as a matrix $C = [c_{j,k}]_{m \times b}$, where $c_{j,k}$ is the (expected) image value of a class j in the band k.
- Using the introduced notation, we can, conveniently, express that I is approximately a linear mixture of the end-members as follows

$$I \approx A \cdot C \, .$$

Note: This notation suggests that the end-members $c_{j,k}$ are position invariant. This is not necessarily the case; we allow spatially varying class representatives $C = C(x)$. However, to not complicate notation, we write C as an $m \times b$ matrix, and not as an $N \times m \times b$ 3D tensor.

[1] Appropriate determination of end-members is a subject of many studies and outside the scope of this presentation.

Data fidelity term

Considering the task of finding a coverage segmentation A, which fulfils $I \approx A \cdot C$ as well as possible, we define the following data fidelity term $D(A)$, for a given image I and a given end-member matrix C

$$D(A) = \|I - AC\|^2 ,$$

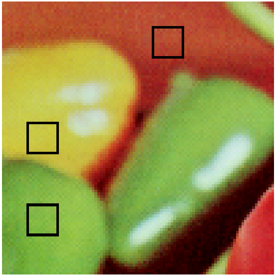
where $\|X\|$ is the Frobenius norm (Euclidean norm) of a matrix X.

Minimization of $D(A)$ (calculus of variations) constrained to $A \in \mathbb{A}_{N\times m}$ provides a linear unmixing segmentation.

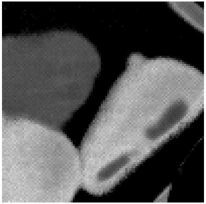
$$A^* = \arg\min_{A \in \mathbb{A}_{N\times m}} D(A)$$

Data fidelity term

- an illustrative example



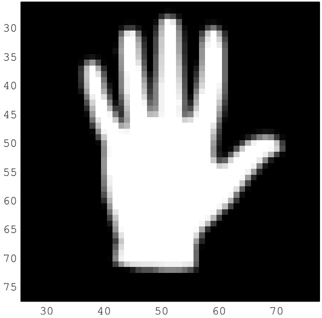
Example colour image with three training regions (defining the end-member matrix C) indicated.



$$A^* = \arg\min_{A \in \mathbb{A}_{N \times m}} D(A)$$

The lack of spatial information makes this type of coverage segmentation **noise sensitive**. Also, the resulting segmentation is generally **too fuzzy** (too many image pixels are classified as mixed).

Properties of coverage representations



- **homogeneous connected regions of “pure” pixels**
- **separated by thin layers of “mixed” pixels**

More energy terms

We add two more criteria to our (so far "too noisy" and "too fuzzy") segmentation model.

(i) we favour a smooth boundary of each object;

(ii) we favour objects with majority of pixels classified as pure, whereas mixed pixels appear only as thin boundaries between the objects.

Criterion (i) is implemented by inclusion of the (fuzzy) perimeter of the objects as a term in the energy function to minimize. Criterion (ii) is imposed by minimizing "thickness" of boundaries over the image, and also, to some extent, minimizing overall fuzziness of the image.

These requirements are combined into the following energy function:

$$J(A) = D(A) + \mu P(A) + \nu T(A) + \xi F(A) \ ,$$

where D, P, T, F are **data term**, overall **perimeter**, **boundary thickness**, and total **image fuzziness**, and $\mu, \nu, \xi \geq 0$ are weighting parameters.

Perimeter, thickness, and fuzziness

Perimeter $P(A)$ is the overall (fuzzy) perimeter of the m objects of a coverage segmentation A

$$P(A) = \frac{1}{2} \sum_{j=1}^{m} P(A_j) \ .$$

Thickness We define border thickness T of a coverage segmentation as

$$T(A) = \frac{1}{2} \sum_{j=1}^{m} T(A_j) \ ,$$

where the thickness of one component $T(A_j)$ is the sum of local thickness computed for all 2×2 tiles of the image:

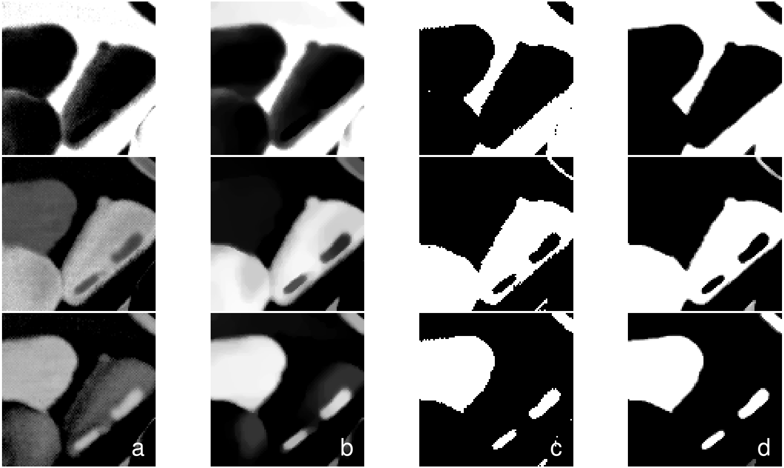
$$T(A_j) = \sum_{(\alpha_{1..4}) \in \tau_{2\times 2}(A_j)} \prod_{i=1}^{4} 4\alpha_i(1-\alpha_i) \ .$$

Fuzziness The inclusion of an overall fuzziness term allows better control of the fuzziness in the resulting segmentation.

$$F(A) = \sum_{i=1}^{N} \sum_{j=1}^{m} 4\alpha_{i,j}(1-\alpha_{i,j}) \ .$$

Different terms

- an illustrative example



(a) Minimization of **D**ata term alone (linear unmixing). (b) Minimization of **D**ata and **P**erimeter terms. (c) Minimization of **D**ata and **F**uzziness terms. (d) Minimization of **all** the suggested energy terms.

Minimization

The sought coverage segmentation A^* is obtained by minimizing the complete energy functional J over the set of valid coverage segmentations:

$$A^* = \arg\min_{A \in \mathbb{A}_{N\times m}} J(A).$$

A convex constrained large scale non-convex optimization problem.

Encouraged by good results obtained when addressing problems of similar structure and dimensionality we decided to use the **Spectral Projected Gradient** (SPG) method.

The SPG method requires **differentiating the energy function $J(A)$**. The partial derivative of $J(A)$ w.r.t. an individual coverage value $\alpha_{i,j}$ is

$$\frac{\partial(J(A))}{\partial\alpha_{i,j}} = \frac{\partial(D(A))}{\partial\alpha_{i,j}} + \mu\frac{\partial(P(A))}{\partial\alpha_{i,j}} + \nu\frac{\partial(T(A))}{\partial\alpha_{i,j}} + \xi\frac{\partial(F(A))}{\partial\alpha_{i,j}}.$$

Minimization

All the included terms are either pixel-wise (data and fuzziness), or utilize only a 2×2 neighbourhood (perimeter and thickness terms). Therefore only 9 pixel values affect $\frac{\partial(J(A))}{\partial \alpha_{i,j}}$, making differentiation quite “manageable”.

The energy function J is, unfortunately, highly **non-convex**, and minimization of J is far from trivial. Care has to be taken to not end up in a sub-optimal local minimum of the energy function.

To reach as good as possible result, solutions of numerically easier problems are used as starting guesses when addressing more difficult ones. We initiate the process with a unmixing based on the data term alone. This is followed by introduction of the perimeter term and an iterative part where the weights of the two fuzziness regulating terms ν and ξ are **gradually increased**.

The iteration continues until the Fuzziness of the solution is lower than twice the Perimeter. This **stopping criterion** utilizes the fact that a correct coverage representation typically fulfils this relation.

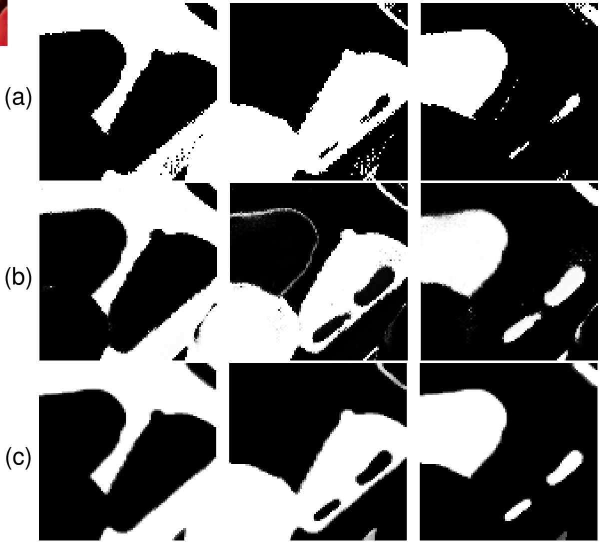
Method 5: Algorithm

Alg. 1. Coverage segmentation

Parameters: $\mu, \nu_0, \xi_0, \rho \geq 0$.

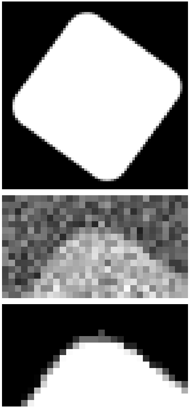
$A_0 = \left[\frac{1}{m}\right]_{N \times m}; \nu = \nu_0; \xi = \xi_0;$
$A = \arg\min D(A_0)$ by SPG;
repeat
    $A \leftarrow \arg\min J(A; I, C, \mu, \nu, \xi)$ by SPG;
    $f = F(A)/(2P(A));$
    $\nu \leftarrow \nu(1 + \rho \cdot f);$
    $\xi \leftarrow \xi(1 + \rho \cdot f);$
until $f \leq 1$

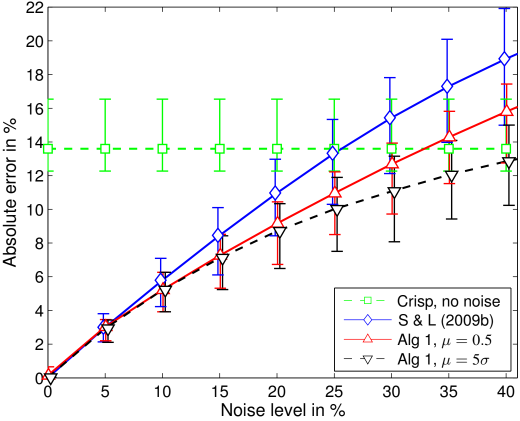
Qualitative evaluation



Segmentation result obtained by: (a) linear discriminant analysis, (b) fuzzy c-means clustering, (c) the proposed method.

Quantitative evaluation
- noise sensitivity

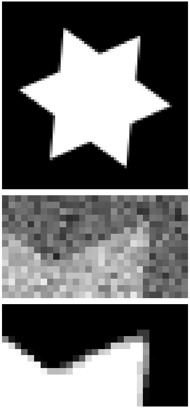


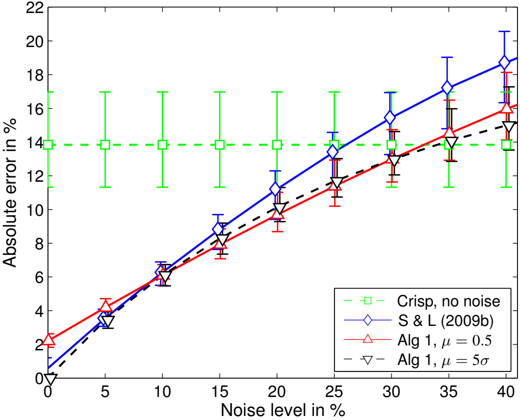


Left: (top) Synthetic test objects. (middle) Part of object with 30% noise added. (bottom) Coverage segmentation result for 30% noise. **Right:** Average absolute error of coverage values of object *border pixels* for different noise levels. Lines show averages for 50 observations and bars indicate max and min errors.

Quantitative evaluation

- noise sensitivity





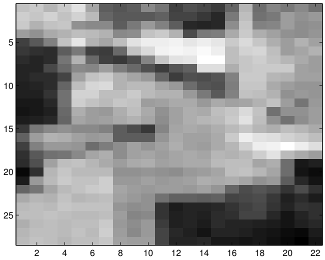
Left: (top) Synthetic test objects. (middle) Part of object with 30% noise added. (bottom) Coverage segmentation result for 30% noise. **Right:** Average absolute error of coverage values of object *border pixels* for different noise levels. Lines show averages for 50 observations and bars indicate max and min errors.

Segmentation of hyperspectral data

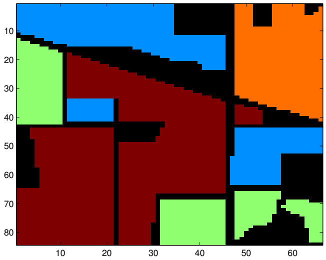
- Test on a publicly available[1] 220 band **hyperspectral data set** from an Airborne Visible/Infrared Imaging Spectrometer (AVIRIS).
- The same data is used in Villa et al.[2] allowing direct performance comparison.
- Available ground truth classification is crisp. Approximate coverage values are created by binning 3×3 pixels into a lower resolved image.
- The 220 bands are highly correlated, making the Euclidean distance (in the Data term) unsuitable as a distance measure. We therefore decorrelate the data initially by a **whitening** transformation.
- For each class, 20 **non-mixed pixels** from the low resolution image are randomly selected as **training data**. From these pixels the matrix C is computed.

[1] `https://engineering.purdue.edu/~biehl/MultiSpec/`

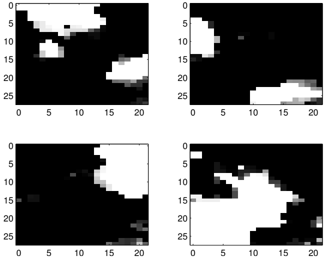
[2] A. Villa et al. "Spectral unmixing for the classification of Hyperspectral images at a finer spatial resolution." IEEE J. Selected Topics Signal Proc. 5 (3), 512-533. 2011.



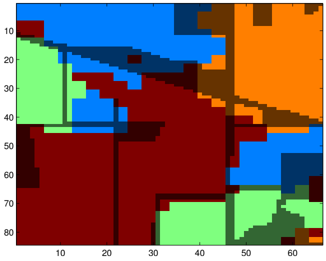
(a) One band (30 out of 220) of a low resolution image obtained by averaging of 3×3 blocks in the original image



(b) Ground truth for the high resolution image, with unclassified pixels presented in black



(c) A coverage segmentation (into four classes) of (a)



(d) Crisp segmentation derived from (c) at the same spatial resolution as (b)

Quantitative evaluation of results

- The method of Villa et al. (2011) performs sub-pixel classification. (SVM-based coverage segmentation is followed by spatial high resolution assignment by means of simulated annealing optimization.)

- To compare our results, we generate two high resolution distributions of coverage:
  1. "Stupid" method: Perform crisp classification and scale up by a factor 3
  2. Optimal method: Distribute the coverage to best match the ground truth

  This provides **lower and upper bounds** of accuracy for a possible sub-pixel assignment of the coverage values.

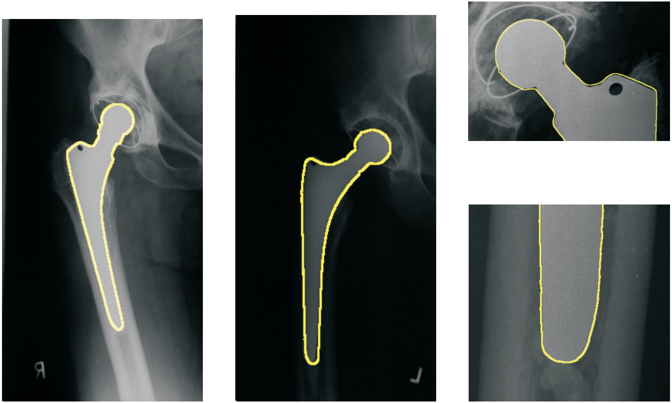
	Accuracy [%]	CPU time [s]
Villa et al.,2011	90.65	58 (88 incl. SA)
Proposed	[92.59,94.74]	4.5

Some application examples

1. Affine registration of digital X-ray and CT images utilizing improved moments estimation
   - A. Tanács, C. Domokos, N. Sladoje, J. Lindblad, and Z. Kato. Recovering affine deformations of fuzzy shapes. SCIA 2009. LNCS-5575, pp. 735–744, 2009.
   - A. Tanács, J. Lindblad, N. Sladoje, and Z. Kato. Estimation of linear deformations of 3D objects. ICIP 2010, IEEE, pp. 153-156, Hong Kong, 2010.
2. Histomorphometrical study from microscopy images, using coverage representation and feature estimates.
   - N. Sladoje, J. Lindblad. Pixel coverage segmentation for improved feature estimation. ICIAP 2009. LNCS-5716, pp. 929-938 Vietri sul Mare, Italy, 2009.
3. Coverage segmentation of a CT image, followed by precise feature estimates
   - F. Malmberg, J. Lindblad, I. Nyström. Sub-pixel segmentation with the image foresting transform, IWCIA 2009, LNCS- 5852, pp. 201-211, 2009.
   - F. Malmberg, J. Lindblad, N. Sladoje, and I. Nyström. A Graph-based Framework for Sub-pixel Image Segmentation. Theoretical Computer Science, Vol. 412, No. 15, pp. 1338-1349, 2011

Application 1 – Registration from moments

Affine registration of digital X-ray images of hip-prosthesis implants for follow up examinations

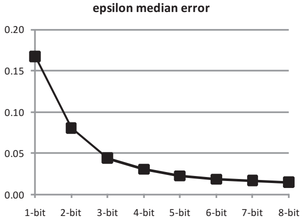


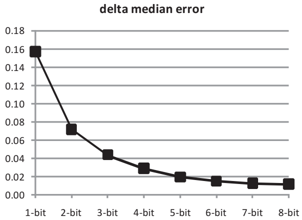
Real X-ray registration results. (a) and (b) show full X-ray observation images and the outlines of the registered template shapes. (c) shows a close up view of a third study around the top and bottom part of the implant.

Application 1 – Registration from moments

Affine registration of digital X-ray images of hip-prosthesis implants for follow up examinations

Coverage values used for improved moments' estimation in a registration process.



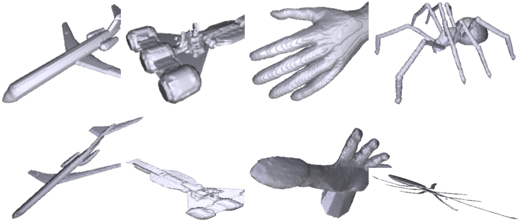


Registration results of 2000 synthetic images using different quantization levels of the coverage representation.

$$\epsilon = \frac{1}{m} \sum_{p \in T} \left\| (\mathbf{A} - \widehat{\mathbf{A}})\boldsymbol{p} \right\|, \quad \text{and} \quad \delta = \frac{|R \triangle O|}{|R| + |O|},$$

Application 1 – Registration from moments

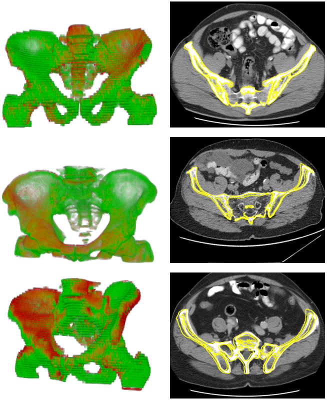
Same thing in 3D



Examples from the image database: template objects (top) and their affine deformed observations (bottom).

Table: Median error values for different supersampling levels n.

n	ϵ	δ	Time (sec)
1	0.0361	0.1555	1.54
2	0.0108	0.0627	1.56
4	0.0069	0.0470	1.54
8	0.0065	0.0402	1.52



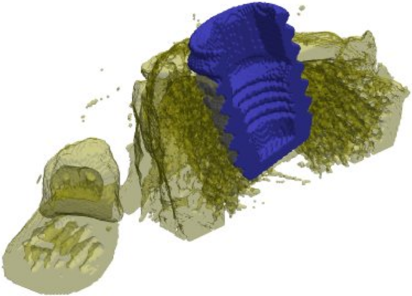
Registration of pelvic CT data

Application 2 – Contact length estimation

Histomorphometrical study from microscopy images

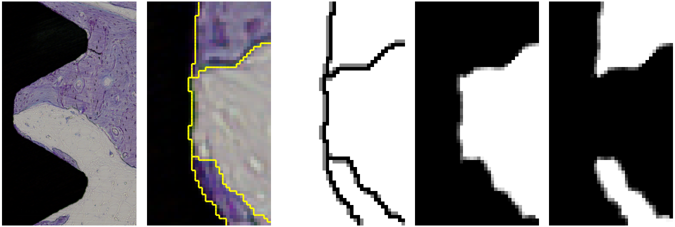
Measure bone implant integration for the purpose of evaluating new surface coatings which are stimulating bone regrowth around the implant.

Local unmixing segmentation followed by area and boundary estimates.



Application 2 – Contact length estimation

Histomorphometrical study from microscopy images



(a) (b) (c) (d) (e)

(a): The screw-shaped implant (black), bone (purple) and soft tissue (light grey). (b) Part of a crisp (manual) segmentation of (a). (c) The set of re-evaluated pixels. (d) and (e) Pixel coverage segmentations of the soft tissue and the bone region, respectively.

Result:

Approximately a **30% reduction of errors** on average, as compared to when using estimates from the crisp starting segmentation.

Application 3 – Precise volume estimation

Coverage segmentation of CT data, followed by feature estimates

User assisted graph based segmentation of the spleen, for medical diagnosis based on accurate feature estimates.

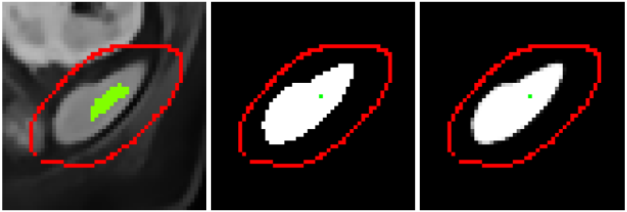


Fig. 3. Segmentation of the spleen in a slice from a CT volume. (Left) Seed-point regions used in the experiment. The green pixels define all object seeds, while the red pixels define background seeds. Single pixels from the green region were used to define object seeds. (Middle) Example result of crisp IFT. (Right) Example result of the proposed sub-pixel IFT.

Table 1. Statistics on the measured area for the 41 segmentations in the experiment. (Areas are given in number of pixels.)

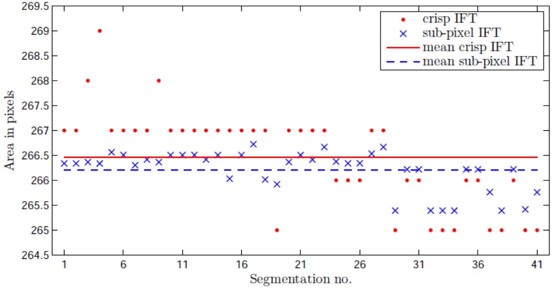
Method	Mean area	Min area	Max area	σ
Crisp IFT	266.5	265	269	0.98
Sub-Pixel IFT	266.2	265.4	266.7	0.40

Result: 50% reduction of standard deviation of estimates, as compared to when using estimates from the crisp starting segmentation.

Application 3 – Precise volume estimation

Coverage segmentation of CT data, followed by feature estimates

User assisted graph based segmentation of the spleen, for medical diagnosis based on accurate feature estimates.



Result: Assuming that the mean result is correct, **more than 3 times reduction of the maximal error**, as compared to when using estimates from the crisp starting segmentation.

Visualization improvement

Lateral ventricles

Conclusions

- A number of coverage segmentation methods have been developed and are shown to work well for estimating coverage values in noisy conditions and for real imaging situation.
- The presented energy based coverage segmentaion method delivers fast and accurate segmentation results under varying imaging conditions.
- The flexible energy minimization approach allows easy introduction of additional terms and/or manipulation of existing ones, thereby simplifying further development and adjustment to specific conditions.

The coverage model, including coverage segmentation, coverage representation, and feature extraction provides:

- separation of imaging and analysis by the coverage segmentation step,
- preservation of object information in the coverage representation,
- high precision feature extraction and even super resolution reconstruction, from the information rich coverage representation,
- reduction of discretization effects, more stable results with improved rotation and translation invariance,

without making high level assumptions about topological properties, often required in continuous models, and without the potentially difficult interpretation of results from more unrestricted fuzzy approaches.