AFRL-OSR-VA-TR-2015-0015

Metallic Material Image Segmentation by using 3D Grain Structure Consistency

Song Wang
SOUTH CAROLINA RESEARCH FOUNDATION

01/05/2015
Final Report

Air Force Research Laboratory
AF Office Of Scientific Research (AFOSR)/ RTC
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
28-12-2014	Final Performance Report	September 30, 2011-September 29, 201

4. TITLE AND SUBTITLE

Metallic Material Image Segmentation by using 3D Grain Structure Consistency and Intra/Inter-Grain Model Information

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-11-1-0327

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Wang, Song

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of South Carolina
901 Sumter St, STE 511
Columbia, SC 29208

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)
AFOSR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Distribution A: Approved for Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT

n this project, we conducted research on developing new methods and software tools to automatically segment various microscopic images of materials (especially metallic materials) to accurately extract their micro-structures, which determine mechanical and other important properties of the materials. The major accomplished work includes: 1) a general multi-label segmentation propagation framework to preserve the shape, appearance, and topology properties of the segments from slice to slice for 3D material image segmentation, 2) new algorithms for enforcing specified topology in image segmentation, 3) an interactive segmentation tool by allowing minimal and simplistic interactions for more accurate 3D material image segmentation, 4) a new clustering method to effectively and robustly segment the super alloy grains from 3D multichannel super alloy images, where each channel corresponds to a specific microscope setting, 5) faster clustering algorithms based on Edge-Weighted Centroid Voronoi Tessellation model by using propagation of the inter-slice consistency constraint for large-scale material image segmentation, and 6) a fully-automatic method to detect cracks from pavement images that can be used for pavement road maintenance.

15. SUBJECT TERMS

Material image segmentation, image processing, computer vision, topology-preserving segmentation, grain structures

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**	UU		**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U			

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

AFOSR Final Performance Report

Project Title:
Metallic Material Image Segmentation by using 3D Grain Structure Consistency and Intra/Inter-Grain Model Information

Award Number:
FA9550-11-1-0327

Dates Covered:
September 30, 2011 – September 29, 2014

Program Manager:
Dr. Arje Nachman
Air Force Office of Scientific Research (AFOSR)
875 North Randolph Street
Arlington, Virginia 22203
Tel (703) 696-8427
Fax (703) 696-8450
Email: SSNav@afosr.af.mil

Principal Investigator:
Song Wang
Department of Computer Science and Engineering
University of South Carolina
Columbia, SC 29208
Tel: (803) 777 2487
Fax: (803) 777 3767
Web: http://www.cse.sc.edu/~songwang
Email: songwang@cec.sc.edu

December 28, 2014

Summary

In this project, we conducted research on developing new methods and software tools to automatically segment various microscopic images of materials (especially metallic materials) to accurately extract their micro-structures, which determine mechanical and other important properties of the materials. The developed methods and tools can facilitate fast and accurate material structural and functional analysis, which can be used to accelerate the new material design and development. For this research, we have been working closely with material scientists from AFRL and other institutions. The major accomplished work includes:

1. Development of a general multi-label segmentation propagation framework to preserve the shape, appearance, and topology properties of the segments from slice to slice for 3D material image segmentation. In particular, we combine global labeling and local relabeling to count for both the global topology consistency between slices and possible local topology inconsistency that results from the disappeared or newly appeared segments in the propagation. The proposed framework can flexibly select and combine the preservation of different properties when segmenting different material images. A graph-cut algorithm is employed to guarantee the computational efficiency.

2. Development of a new strategy for enforcing specified topology in image segmentation. Most recent works on topology-constrained image segmentation focus on binary segmentation, where the topology is often described by the connectivity of both foreground and background. We developed a new multi-labeling method to enforce topology in multi-label image segmentation. In this case, we not only require each segment to be a connected region (*intra-segment topology*), but also require specific adjacency relations between each pair of segments (*inter-segment topology*). Both of them are important for metallic image segmentation, where microstructure consists of a large number of grains with complex adjacency relations.

3. Development of an interactive segmentation tool by allowing minimal and simplistic interaction from the user in an otherwise automatic algorithm. The developed interactive segmentation is able to simultaneously reduce the time taken to segment an image while achieving better segmentation results. More specifically, considering the specialized structure of materials images and level of segmentation quality required, we extended the multi-labeling approach such that it can not only handle a large number of grains but also quickly and conveniently allow manual addition and removal of segments in real time. In addition, multiple extensions were made to the interactive tools which increase the simplicity of the interaction, Finally, we developed a web interface for using the interactive tools in a client/server architecture.

4. Development of a *Multichannel Edge-Weighted Centroidal Voronoi Tessellation* (MCEWCVT) algorithm to effectively and robustly segment the superalloy grains from 3D multichannel superalloy images, where each channel corresponds to a specific microscope setting. MCEWCVT performs segmentation by minimizing an energy function which encodes both the multichannel voxel-intensity similarity within each cluster in the intensity domain and the boundary smoothness of segmentation in the 3D image domain. Based on MCEWCVT, we further developed a *Constrained Multichannel Edge-Weighted Centroidal Voronoi Tessellation* (CMEWCVT) algorithm which can take manual segmentation on a small number of selected 2D slices as constraints from the problem domain and incorporate them into the energy-minimization process to further improve the segmentation accuracy.

5. Development of a clustering algorithm based on Edge-Weighted Centroid Voronoi Tessellation (EWCVT) which uses propagation of the inter-slice consistency constraint. It can segment a 3D superalloy image, slice by slice, to obtain the underlying grain microstructures. With the propagation of the consistency constraint, the proposed method can automatically match grain segments between slices. On each of the 2D image slices, stable structures identified from the previous slice can be well-preserved, with further refinement by clustering the pixels in terms of both intensity and spatial information. We tested the developed algorithm on a 3D superalloy image and it outperforms the existing segmentation methods in terms of both segmentation accuracy and running time.

6. As an effort to extend our research results on material image segmentation to other application domains, we developed *CrackTree*, a fully-automatic method to detect cracks from pavement images, that can be used for pavement road maintenance. The developed method consists of three steps: 1) A geodesic shadow-removal algorithm to remove the pavement shadows while preserving the cracks; 2) building a crack probability map to enhance the connection of the crack fragments ; and 3) graph modeling of the fragments and a recursive tree-edge pruning algorithm to identify desirable cracks. Cracktree was evaluated on real pavement images and it achieves better performance than existing methods.

1 Multi-label Segmentation Propagation

We define *segmentation propagation* as the problem of transferring a segmentation from a segmented image U to an unsegmented image V, subject to predefined constraints. Specifically, given an image U, and a segmentation S^U of U such that S^U is a partition of the pixels in U into n segments $S^U = \{S_1^U, \ldots, S_n^U\}$ where

$$U = S_1^U \cup \ldots \cup S_n^U \text{ and } S_i^U \cap S_j^U = \emptyset, \ \forall i \neq j,$$

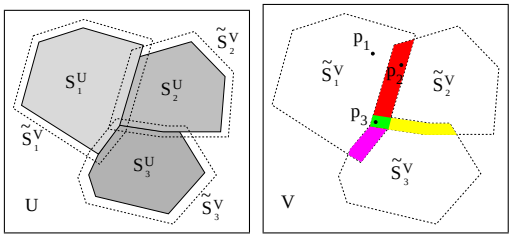
we wish to obtain a segmentation S^V of an image V, which contains the same objects as U, by *propagating* S^U to V, yielding S^V.

We define $\mathcal{A}$ to be the set of unordered pairs $\{S_i, S_j\}$, indicating segment S_i and S_j are neighbors. Let $\mathcal{P}$ be the set of pixel pairs that are neighbors. We formulate a solution to this problem as an MRF (Markov Random Field) energy minimization over the partitioning of pixels in V to find S^V, given in the following form:

$$E(S^V) = \sum_{p \in V} \Theta_p(S_i^V) + \sum_{\{p,q\} \in \mathcal{P}^V} \Phi_{pq}(S_i^V, S_j^V). \quad (1)$$

The unary term Θ_p describes a cost for assigning a particular pixel p to a segment S_i^V, and the binary term Φ_{pq} describes a cost for assigning two neighboring pixels p and q (i.e., $\{p, q\} \in \mathcal{P}^V$) to two segments S_i^V and S_j^V. Finding a global minimum of Eq. (1) is generally an NP-hard problem, but a locally-optimal solution can be efficiently found with a graph-cut algorithm [17, 2].

For the unary term, we make an assumption that all pairs S_i^U and S_i^V have good spatial overlap, varying most significantly around their boundaries. For every segment S_i^U, we construct a bounding region $\tilde{S}_i^V$, which contains all $p \in V$ that are within distance d of any pixel in S_i^U (d is called dilation size). Using the bounding region $\tilde{S}_i^V$, we set $\Theta_p(S_i^V) = 0$ for all $p \in \tilde{S}_i^V$ and $\Theta_p(S_i^V) = \infty$ for all $p \notin \tilde{S}_i^V$. An example of defining Θ_p is shown in Fig. 1, where the costs for p_1, p_2, and p_3 are shown for various assignments of S_1^V, S_2^V, and S_3^V.



(a) (b)

Θ	S_1^V	S_2^V	S_3^V
p_1	0	∞	∞
p_2	0	0	∞
p_3	0	0	0

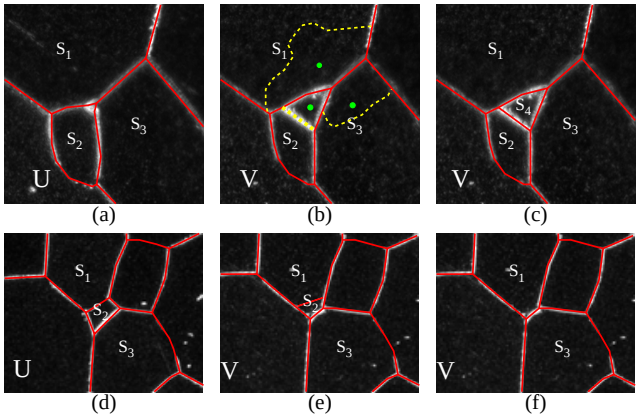
(c)

Figure 1: An illustration of defining the unary term in the proposed segmentation propagation. (a) Three adjacent S_i^U and associated $\tilde{S}_i^V$. (b) Three pixels which fall within various $\tilde{S}_i^V$. Specifically, p_1, p_2, and p_3, fall within one, two, and three bounding regions, respectively. (c) Unary term Θ defined for pixels p_1, p_2 and p_3 in (b).

To help preserve the topology in the propagation, we define the binary term as

$$\Phi_{pq}(S_i^V, S_j^V) = \begin{cases} 0, & i = j \\ \infty, & \{S_i^U, S_j^U\} \notin \mathcal{A}^U \\ g(p,q), & \{S_i^U, S_j^U\} \in \mathcal{A}^U \end{cases} \tag{2}$$

which introduces zero cost for pixels assigned to the same segment, an ∞ cost for pixels assigned to segments that are not adjacent in S^U, and a cost functional g for pixels that are assigned to segments that are adjacent in S^U. The ∞ cost in Eq. (2) enforces that two non-adjacent segments in U will not become adjacent after propagating to V. Function g represents a probability that pixels p and q are along segmentation boundaries and can be defined base don the intensities of p and q.



(a) (b) (c)

(d) (e) (f)

Figure 2: An illustration of local topology inconsistency by segment appearance (top row) and disappearance (bottom row). (a,d) Segmentation on U. (b,e) Segmentation with global labeling. Green dots and yellow dashed lines show how the local relabeling attempts to find newly appeared segment. (c,f) Segmentation after local relabeling.

For 3D metallic image, the grain adjacency relations are not always preserved between slices, since an existing grain may disappear or a new grain may appear when moving from one slice to its neighbors. Therefore, in segmenting 3D material images, we first use the above graph cut algorithm to segment the whole image V and then at each location of V, we relax topology constraints in Eq. (2) and perform local-relabeling to handle the possible appearance of new segments (top row of Fig. 2) and disappearance of existing segments (bottom row of Fig. 2). For handling the possible new

segment, we sample a set of seed points with new labels and then verify them using local relabeling by minimizing the same energy function, as illustrated in Fig. 2(b).

We tested the developed algorithm that combines global labeling and local re-labeling on a sequence of 11 microscopic Ti-21S titanium images. Each Ti-21S slice has a resolution of 750×525, and consists of ~ 120 β-Ti grains, which are the micro-structures of interest. These grains are all adjacent, meaning that there is no notion of a "background" in this material. We take the manually constructed ground truth segmentation on the first slice as an initialization and propagate it to all the other slices sequentially. Results on selected slices (cropped and zoomed view) are shown in Fig. 3, together with results from two comparison methods – watershed [14] and [16, 8]. Figure 4 shows the quantitative results of the proposed method and the comparison method, in terms of the F-measure against the manually constructed ground truth segmentation and the cardinality difference, which is the difference of the number of the segmented grains and the number of the ground-truth segments.

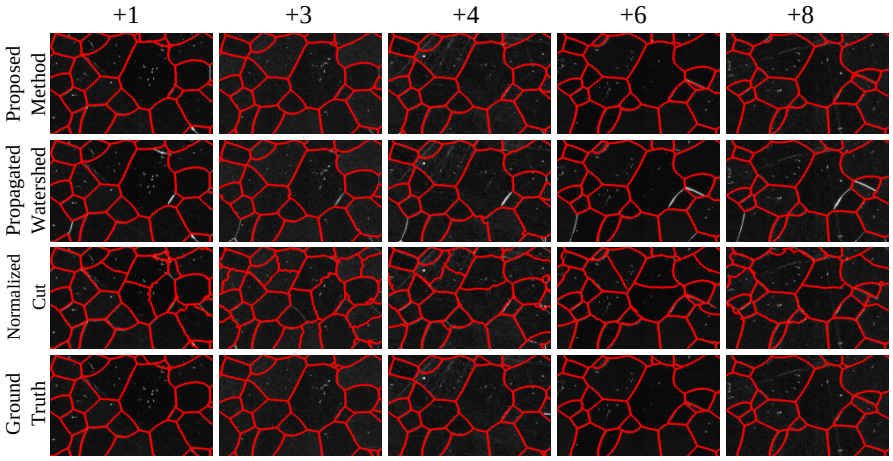
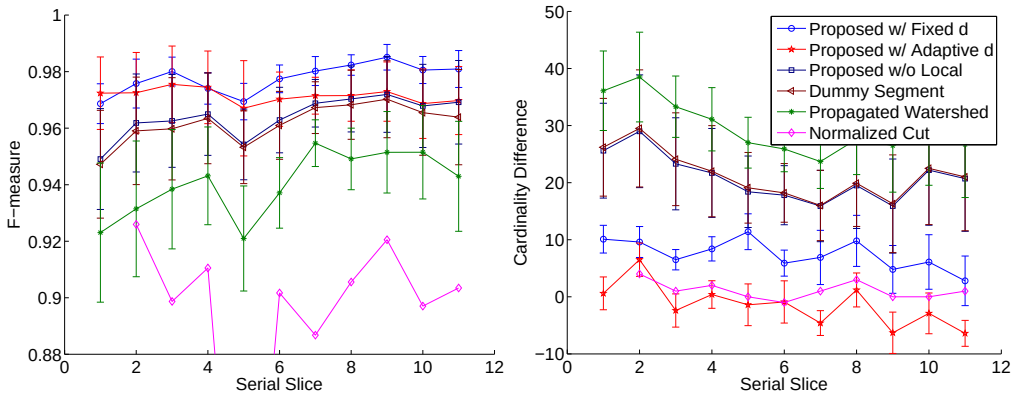


Figure 3: Zoomed view of segmentation results on selected slices using the proposed method, propagated watershed, and normalized cut, along with the ground truth. Each column shows a slice of different distance from the initial slice.

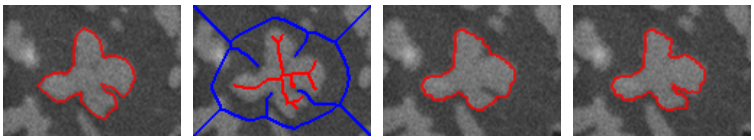


(a) F-measure (b) Cardinality Difference

Figure 4: (a) The segmentation F-measures for the proposed method, the watershed method, and normalized cut on the 11 Ti-21S slices. (b) Cardinality difference measure for all evaluated methods.

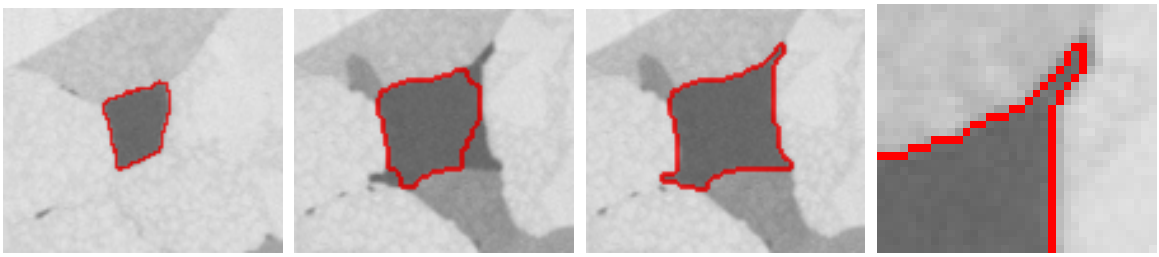
We also extended this method to preserve the shape and/or intensity of selected segments in the

segmentation propagation [18]. Examples are shown in Figs. 5 and 6, respectively.



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

Figure 5: Dendritic γ' precipitates in Rene88DT. (a) Segmentation of slice U, created manually. (b) Skeletonizaton of the segments in slice U, showing the shape of the foreground segment (red) and background (blue). (c) Segmentation results on slice V using the proposed method without shape preservation. (c) Segmentation result on slice V by incorporating the shape-preservation strategy, which keeps the label of the pixels along skeletons in propagation.



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

Figure 6: Grain structure of IN100 superalloy. (a) Segmentation of slice U, created manually. (b) Segmentation result on slice V using the proposed method without intensity preservation. (c) Segmentation result on slice V by preserving the intensity of the corresponding segments between U and V. (d) Zoomed view of upper-right corner of (c).

2 Topology-Preserving Image Segmentation

The graph-cut method introduced in Section 1 can not always preserve the topology in segmentation propagation. It can guarantee that non-adjacent segments do not become adjacent after propagation, but it cannot guarantee that adjacent segments stay adjacent after propagation. Based on the problem formulation and MRF energy function as defined in Section 1, as well as the special microstructures of metallic materials, we sequentially take each local ring structure for relabeling and proved that this can rigorously preserve the topology of segmentation.

As illustrated in Fig. 7(a), using the segmentation S^U as the initial segmentation on V, we find a local ring structure that consists of one center segment and all segments adjacent to this center segment. The center segment is adjacent to every non-center segment in the ring, and is not adjacent to any segments outside of the ring. Additionally, in the general case, each non-center segment has a clockwise adjacent segment and a counterclockwise adjacent segment other than the center segment in the ring. We also require the existence of at least one pair of non-adjacent segments in a ring to activate the infinity penalty as defined in Eq. (2). This leads to the requirement that there be at least 4 non-center segments in a ring. If a ring contains only 2 or 3 non-center segments, as shown in Fig. 8(b,c), we can split one or two non-center segments along the radial direction, as shown in Fig. 8(d,e), to increase the number of non-center segments and introduce non-adjacency. This updates S^U which is then propagated to V, after which we merge such split segments together

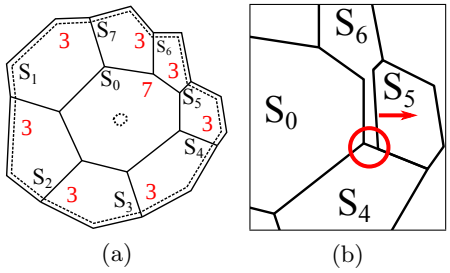


Figure 7: Local ring structure example. (a) Preserving inter-segment topology by fixing the label of pixels along the ring boundary and in the center segment (dashed lines). Red numbers indicate the numbers of segments adjacent to the indicated segment. (b) Cropped view of (a) illustrating the preservation of inter-segment topology while updating the ring.

to obtain the final segmentation S^V. Another degenerate case is when there is a single non-center segment in the ring, as shown in Fig. 8(a), which reduces to the binary segmentation problem, and the developed method can handle this degenerate case without splitting any segments.

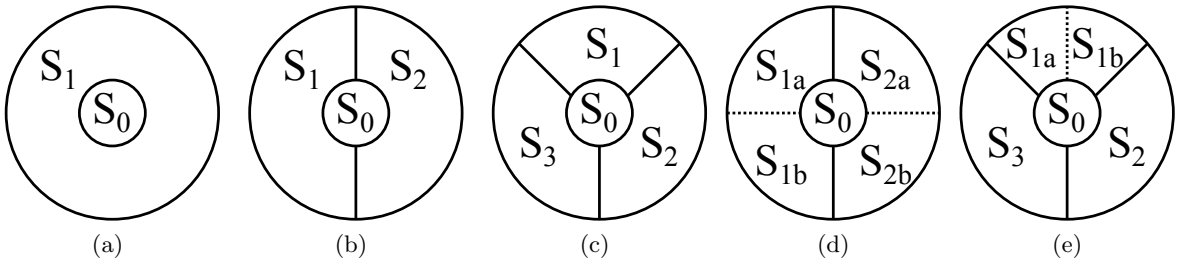


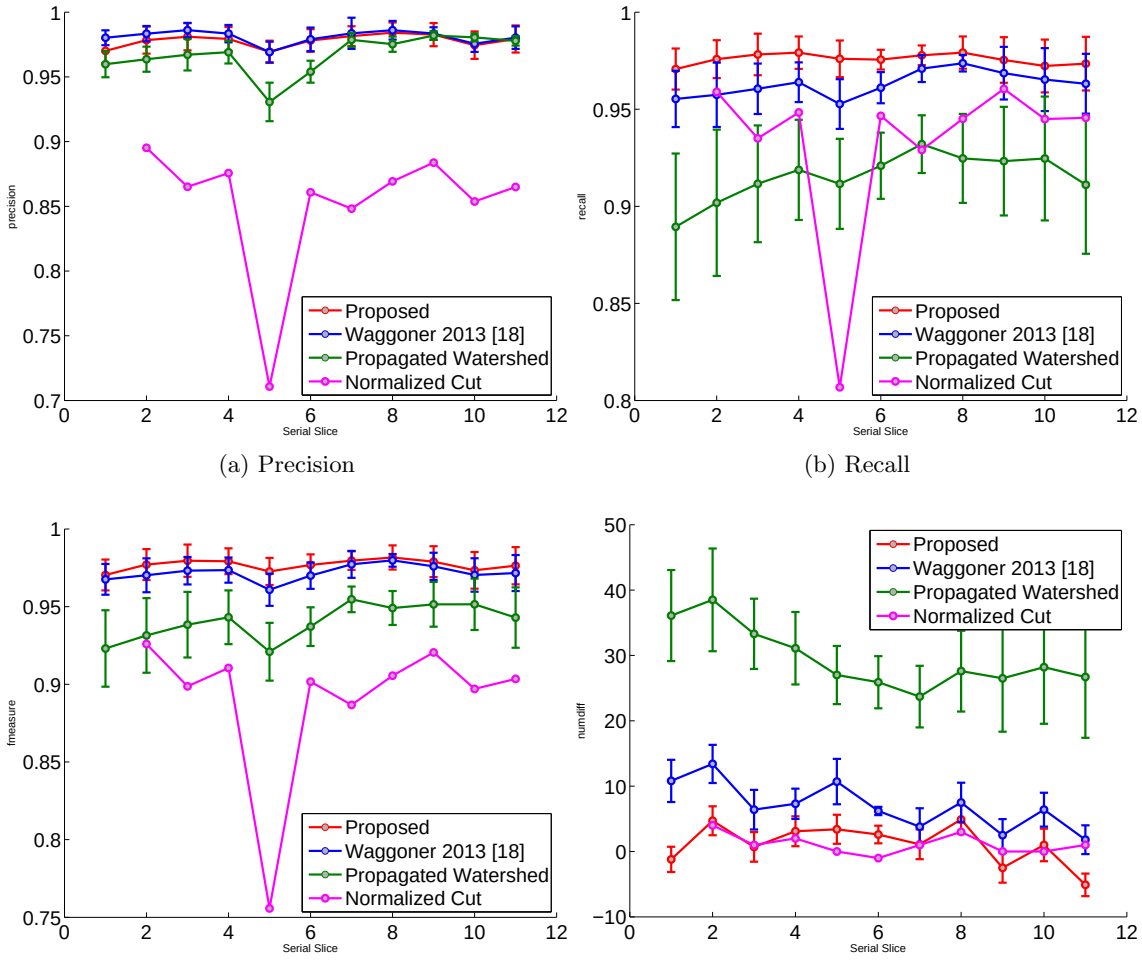
Figure 8: Illustration of the degenerate cases of ring structures. (a-c) Ring structures with 1, 2, and 3 non-center segments, respectively. (d-e) Non-center segment splitting to achieve four non-center segments for the rings in (b) and (c), respectively.

From such a local ring structure, together with the image V on which this ring is embedded, we can define Eq. (1) and use graph cut to update the segmentation in this ring. The primary issue is that we must preserve the topology of all the segments in S^U, not simply the topology inside this ring. Therefore, we fix the labels for all the pixels along the ring boundary, shown by the dashed contour in Fig. 7(a). This can be easily achieved by assigning unary-term values for such pixels to be zero if their labels are the same as before and infinity otherwise. This way, we insure that the adjacency relations between any segment in this ring and any segment outside this ring will remain unchanged after the labeling update in this ring. To avoid the disappearance of the center segment, we also select the centroid pixel of the center segment and require its label to be unchanged (dashed lines in the center of Fig. 7(a).

Based on this, we simply perform graph-cut based relabling with the binary term as defined in Eq. (2) within the image region defined by this ring to update its segmentation. As discussed above, this algorithm guarantees that non-adjacent segments remain non-adjacent, which, together with the constraints defined on the ring boundary, will also guarantee that adjacent segments in this ring remain adjacent. This is indeed the case because, 1) the adjacency between non-center segments has been preserved by the label constraints on the ring boundary, and 2) the center segment is still adjacent to every non-center segment. For 2), it can be proved by contradiction, as shown in Fig. 7(b), if the center segment S_0 becomes non-adjacent from a non-center segment, S_5, a pair of non-adjacent segments, S_4 and S_6 (be reminded that there are at least four non-center segments

in a ring), must become adjacent to separate S_0 and S_5. However, the proposed algorithm has an infinity penalty term in Eq. (2) specifically to prevent any non-adjacent segments from becoming adjacent. Based on this formulation and algorithm, we can also prove that the achieved segment is always a connected region. To update the segmentation of the whole image, we can sequentially and repeatedly perform such relabeling on each ring structure on V.

We tested the performance on the 11 Ti-21S titanium image slices (as described in Section 1) by propagating segmentation from the first slice to the last. Quantitative and qualitative results are shown in Figs. 9 and 10 respectively. Comparison methods are also included in these figures, where Waggoner 2013 indicates the graph-cut method introduced in Section 1.



(a) Precision (b) Recall

(c) F-measure (d) Cardinality Difference

Figure 9: Performance of the proposed ring-based relabeling method, the previous topology-preserving method [18], and two other comparison methods, on the Ti-21S dataset.

3 Interactive Material Image Segmentation

Based on the above-mentioned multi-label segmentation propagation framework, we developed several convenient interactive strategies to correct the possible errors in automatic segmentation and further improve the segmentation accuracy. The first strategy is to remove spurious segments. For

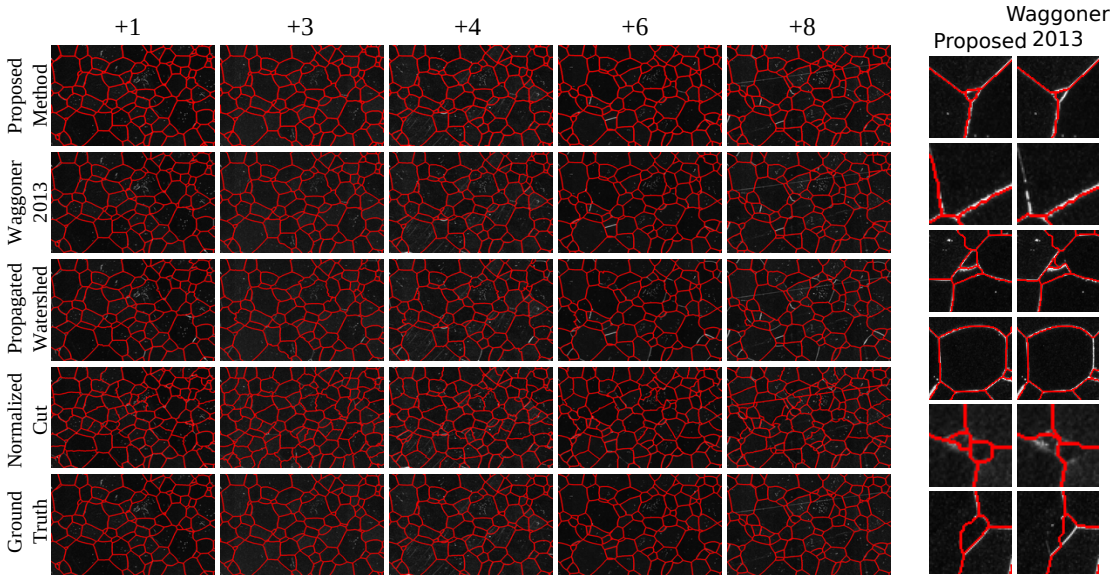


Figure 10: (a) Qualitative results for the Ti-21S dataset for the proposed method, the method in [18], the watershed method, and normalized cut. The distance from the initial template is shown by the numbers along the top. (b) The more subtle differences between the proposed method and the method in [18].

this interaction, we allow the user to select a spurious segment S_k^V for removal by clicking the mouse on this segment in a visualized segmentation of S^V. Instead of naively removing this segment by arbitrarily merging it into one of its neighbors, we use the same energy minimization discussed above to assign the individual pixels in S_k^V to potentially different neighboring segments. Like the local relabeling discussed in Section 1, we identify a local region in which we update the segmentation. Specifically, this local region consists of the specified S_k^V and its neighboring segments, e.g., S_1^V, S_2^V, S_3^V surrounding the spurious segment S_k^V in Fig. 11(a), and re-run the energy minimization within this local region after modifying the Θ term in a way that no pixel is allowed to be assigned to S_k^V (∞ cost), resulting in an updated segmentation in this local region, as shown by the example in Fig. 11(c).

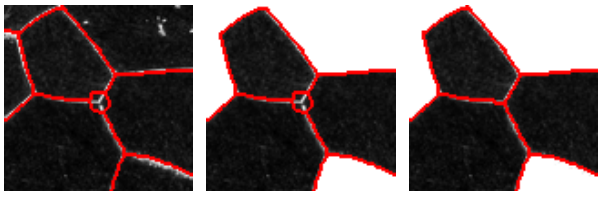
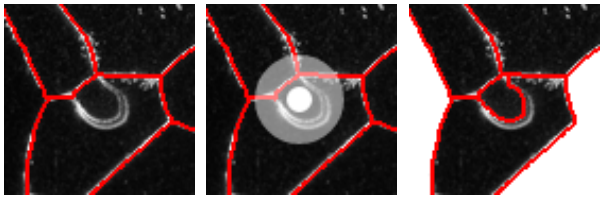


Figure 11: Example selection of a spurious segment S_k^V for removal. (a) Chosen S_k^V and surrounding segments. (b) Local region extracted around S_k^V. (c) The updated segmentation in the extracted local region.

The second strategy is to add a missing segment. Unlike removal, interactively annotating an additional structure cannot be solely formulated as a simple modification of the Θ term in the energy minimization formulation and for each missing segment, we must explicitly create a new segment at

the location interactively specified by the user. After interactively annotate a point (or a stroke) as seeds for a new segment, we explicitly enforce them to be part of the missing segment that is added, as shown by the inner circle in Fig.12(b), and their dilation pixels, excluding seed pixels, are *potentially* part of the missing segment that is added, as shown by the outer circle in Fig. 12(b). Similar to the removal interaction, we define a local region around the specified point to update the segmentation of S^V. Specifically, we define this region by taking all segments in S^V that contain one or more seed or dilation pixels. In this local region we modify the Θ term of the energy minimization in Eq. (1) to reflect the desirable labeling constraints on the seeds and dilated areas and then perform the relabeling to obtain an updated segmentation, as shown in Fig. 12c.



(a) (b) (c)

Figure 12: Annotating the addition of a missing segment. (a) Segmentation S^V with a missing segment near the center of the image. (b) Annotation of a center point c, along with a seed radius s and a dilation radius d, and the identified local region for updating the segmentation. (c) The updated segmentation of the local region shown in (b).

Because materials images can be very large and complex, it can take a significant amount of time for a human annotator to review the segmentation of such a large image to determine where it may require additional interaction. We developed a salient region detection approach that identifies candidate regions highlighting the areas most likely to need additional interaction. We focus on detecting the edges in the image that are not identified as segment boundaries, indicating a missing segment, for this salient region detection. As such, we identify prominent edges fragments that are not segmented during the propagation as candidate regions, and use a SVM classifier [7] to learn which candidates are truly salient regions, and which are noise that can be ignored by the human annotator. The extracted feature for SVM classifier reflects multiple shape and intensity properties for SVM classifier, including

- the total area of the region,
- the minor and major axis length of the ellipse fit to the region,
- the maximum intensity inside the region, and
- and the mean intensity inside the region.

These salient regions are later enclosed in a bounding box for easier visualization, as illustrated in Fig. 13.

Based on these research results, we built an interactive interface as a web application using the Django [9] web framework for the backend, and a custom single-page JavaScript client as the frontend. The software architecture is shown in Fig. 14 and the client interface is shown in Fig. 15.

4 Multichannel Clustering for Material Image Segmentation

Many superalloy images contains multichannel information – each channel corresponds to a specific microscopy setting. Multichannel imaging provides more information to identify the boundary

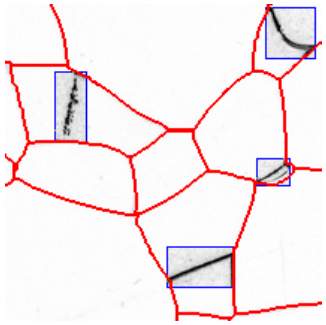


Figure 13: Sample results of salient region detection. Salient regions are surrounded a bounding box, highlighting locations indicating where segments may be missing.

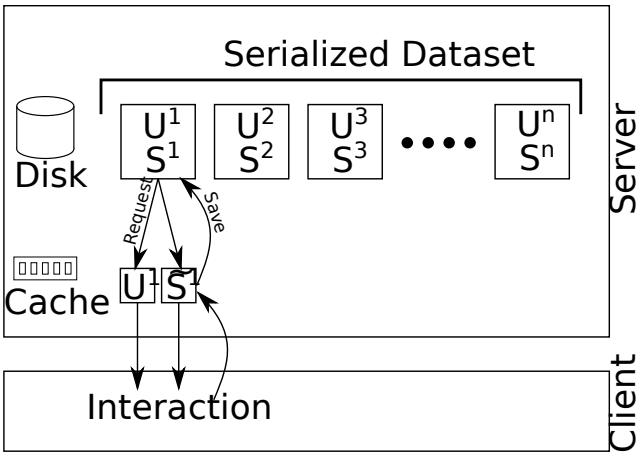


Figure 14: Overview of the client/server architecture used to implement the proposed approach. Large datasets are persisted on disk with both the underlying image (U^i) and segmentation from the automatic propagation approach (S^i) saved for retrieval. A cache allows multiple interactions that modify the segmentation S^i to be saved in memory, where the image and segmentation can be quickly retrieved and modified. The client may explicitly issue a "Save" request to persist changes made in the cache onto disk.

between adjacent grains since two adjacent grains may show similar intensity in one channel but different intensities in another channel. Figure 16 shows a 4-channel image of the same superalloy slice. We can see that adjacent grains g_1 and g_2 can be better separated in channels (a,b,d) than in channel (c). However, adjacent grains g_1 and g_3 can be better separated in (c) than in (a,b,d).

Let N denote the number of channels, i.e., we have N images, $u^1, u^2, \cdots, u^N$, of the same superalloy sample. Denote

$$\mathbb{U} = \{\vec{u}(i,j,k) = (u^1, u^2, \cdots, u^N)^T \in \mathbb{R}^N\}_{(i,j,k)\in D}$$

be the images and

$$\mathcal{W} = \{\vec{w}_l = (w_l^1, w_l^2, \cdots, w_l^N)^T \in \mathbb{R}^N\}_{l=1}^L,$$

be the L typical intensity levels in each of these channels. We choose ∞-norm for defining the distance between $\vec{u}(i,j,k)$ and $\vec{w}_l$ to capture the grains with intensities that are distinct with its adjacent grains only in some (but not all) of the channels, i.e.,

$$\|\vec{x}\|_\infty = \max(|x^1|, |x^2|, \ldots, |x^N|)$$

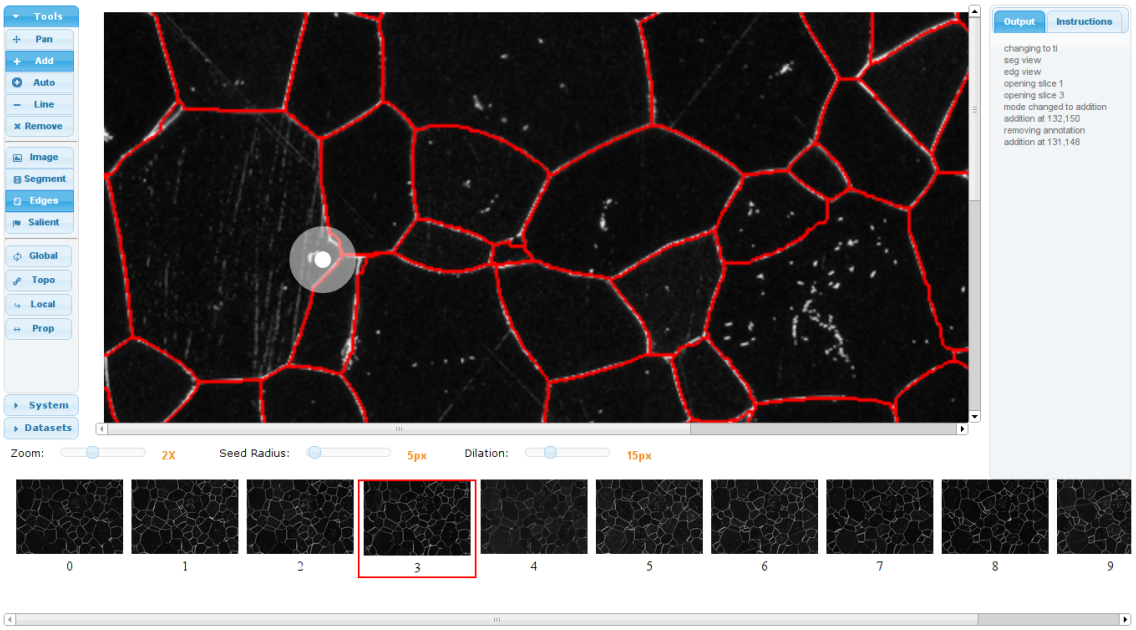


Figure 15: Client interface presented to the user for interaction.

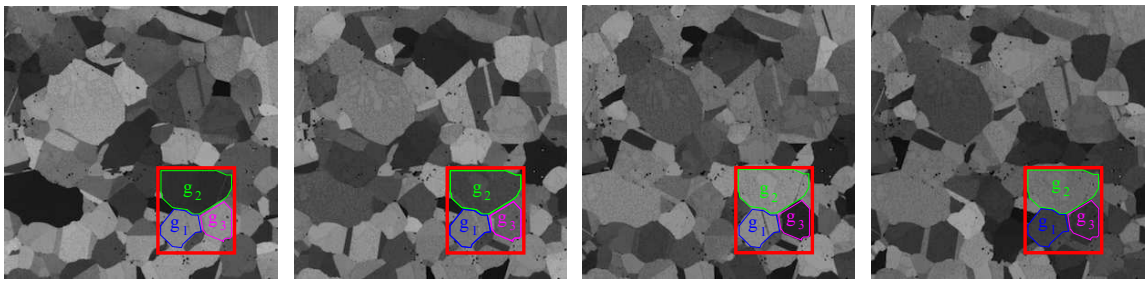


Figure 16: One slice of a superalloy sample with four image channels (4 different electronic microscope settings). (a) 4000_Series. (b) 5000_Series. (c) 6000_Series. (d) 7000_Series.

where $\vec{x} = (x^1, x^2, \ldots, x^N) \in \mathbb{R}^N$.

We then define the Multichannel Edge-Weighted Centroidal Voronoi Tessellation (MCEWCVT) energy function as

$$\begin{aligned} E(\mathcal{W};\mathcal{D}) &= E_C(\mathcal{W};\mathcal{D}) + \lambda E_L(\mathcal{D}) \\ &= \sum_{l=1}^{L} \sum_{(i,j,k)\in D_l} (1 + |\nabla \vec{u}(i,j,k)|) \|\vec{u}(i,j,k) - \vec{w}_l\|_\infty^2 \quad + \lambda \sum_{(i,j,k)\in D} \sum_{(i',j',k')\in \mathbb{N}_\omega(i,j,k)} \chi_{(i,j,k)}(i',j',k') \end{aligned} \tag{3}$$

where λ is a weighting function to control the balance between E_C and E_L.

In Eq. (3), E_L is an edge energy term, in which $\mathbb{N}_\omega(i,j,k)$ is a local neighborhood around voxel (i,j,k) and the characteristic function $\chi_{(i,j,k)} : \mathbb{N}_\omega(i,j,k) \to \{0,1\}$ is

$$\chi_{(i,j,k)}(i',j',k') = \begin{cases} 1 & \text{if } \pi_u(i',j',k') \neq \pi_u(i,j,k) \ , \\ 0 & \text{otherwise,} \end{cases}$$

where $\pi_u(i,j,k) : D \to \{1,\ldots,L\}$ tells which cluster that $\vec{u}(i,j,k)$ belongs to. The inclusion of the this edge energy term ensures the smoothness of the segmentation boundaries.

The optimization of the cluster energy function is similar to the traditional centroidal Voronoi tessellation algorithm or the K-means clustering algorithm, which alternately calculates the cluster mean and update the cluster assignment of each voxel. For the proposed energy function, the distance used for updating the cluster assignment can be computed by [4]

$$\begin{aligned} dist((i,j,k),\vec{w}_l) &= \sqrt{\rho(i,j,k)\|\vec{u}(i,j,k)-\vec{w}_l\|_\infty^2 + 2\lambda\tilde{n}_l(i,j,k)} \\ &= \sqrt{(1+|\nabla\vec{u}(i,j,k)|)\|\vec{u}(i,j,k)-\vec{w}_l\|_\infty^2 + 2\lambda\tilde{n}_l(i,j,k)} \end{aligned} \tag{4}$$

where $\tilde{n}_l(i,j,k) = |\mathbb{N}_w(i,j,k)| - n_l(i,j,k) - 1$ is the number of voxels in $\mathbb{N}_w(i,j,k)\backslash(D_l \bigcup (i,j,k))$ and the density function ρ is

$$\rho = 1 + |\nabla\vec{u}|.$$

We tested the proposed MCEWCVT algorithms on a Ni-based 3D superalloy image dataset [5, 4]. The dataset consists 4 channels of superalloy slice images taken under different electronic microscope parameters settings. Each slice was photographed as new facets appearing by keeping abrading the up-front facet of the superalloy sample. The size of each 2D slice is 671×671 and the number of slices in each channel is 170. The resolution within a slice is $0.2\mu m$ and resolution between slices is $1\mu m$. We linearly interpolate the 3D superalloy image with 4 additional slices between each pair of consecutive slices in the original data and this way, the interpolated data contain $169 \times 5 + 1 = 846$ slices. The testing dataset also comes with the ground truth segmentation created by manual segmentation on each 2D slice. Table 1 shows the performance of the developed MCEWCVT algorithm and other comparison methods.

Table 1: Segmentation performance of the proposed MCEWCVT algorithm and the comparison methods, using F-measures

Methods	MCEWCVT	random walks	power watersheds	mean shift	EM/MPM
F-measure	93.57	88.87%	88.27%	86.32%	80.82%
Methods	pbCanny	ucm	pbCGTG	pbBGTG	topological watersheds
F-measure	80.54%	79.52%	79.51%	78.8%	78.41%
Methods	srm	efficient-graph based	gpb	normalized cuts	3D watershed
F-measure	76.36%	75.30%	73.42%	71.19%	70.08%
Methods	2D watershed	2D level set	3D level set	CVT/K-means	
F-measure	68.94%	66.74%	65.06%	60.38%	

We also extended MCEWCVT to a *Constrained Multichannel Edge-Weighted Centroid Voronoi Tessellation* (CMEWCVT) algorithm that incorporates human annotated constraints [3]. For example, we can select a subset of metallic image slices and manually segment them. Then we run MCEWCVT over all the slices as described before, but subject to constraints that on the selected slices, the MCEWCVT segmentation results should be well aligned with the human annotated segmentation. This is achieved by examining the constraints after each iteration of MCEWCVT and making necessary changes if the constraints are not satisfied. Our experiments shows the F-measure-based segmentation performance on the Ni-based 3D superalloy image dataset can be further improved by introducing such constraints and using CMEWCVT [3].

5 Propagation based Edge-Weighted Centroid Voronoi Tessellation (EWCVT)

While MCEWCVT and CMEWCVT can achieve the state-of-the-art performance on material image segmentation, it takes intensive computation time, because it directly operates on a large number of voxels. To address this issue, we applied EWCVT slice by slice, while a propagation of the structural consistency is considered when moving from one slice to another [22]. Given two consequent image slices I^i and I^{i+1}, their segmentation results can be defined as $S^i = \{s_1^i, \ldots, s_{m_i}^i\}$ and $S^{i+1} = \{s_1^{i+1}, \ldots, s_{m_{i+1}}^{i+1}\}$ where m_i and m_{i+1} are the number of segments (grains) in I^i and I^{i+1} respectively. The segment structure of the segmentation S^i on the image slice I^i can be represented by a graph of segments in S^i, denoted as $G^i\left(\mathcal{V}^i, \mathcal{E}^i\right)$, where each vertex in $\mathcal{V}^i$ is a segment and the edge weights in $\mathcal{E}^i$ measure the strength of the adjacency of two neighbor segments (directly connected). Typically, given two segments, we use the number of pixels located on the boundary shared by them as their edge weight.

The stable segment structure of S^i on I^i can be defined as a connected subgraph $G_*^i\left(\mathcal{V}_*^i, \mathcal{E}_*^i\right)$ of G^i. Specifically, it holds that

$$\mathcal{V}_*^i = \left\{s_p^i \in \mathcal{V}^i \mid \left|s_p^i\right| \geq \alpha\right\} \tag{5}$$

and

$$\mathcal{E}_*^i = \left\{\mathcal{E}_{(p,q)}^i \in \mathcal{E}^i \mid \mathcal{E}_{(p,q)}^i \geq \beta, \quad S_p^i, S_q^i \in \mathcal{V}_*^i\right\}, \tag{6}$$

where the parameter $\alpha > 0$ is the minimal size of segments that are defined as stable ones, and the parameter $\beta > 0$ is the minimal length of boundaries that are stable. In the developed algorithm, we preserve the stable grain structure of S^i on I^i when propagated to achieve segmentation S^{i+1} on I^{i+1}. Unstable grains and their adjacency, caused by the difference of two consequent image slices, are determined by the image information on I^{i+1}, using the standard EWCVT clustering.

On the 170-slice Ni-based 3D superalloy image dataset as described in Section 4, we found that this new method achieve the comparable accuracy to MCEWCVT, as shown in Table 2. However, compared with the MCEWCVT algorithm, this new method achieves a 5x speed up.

Table 2: Quantitative comparison of segmentation on the 170-slice Ni-based 3D superalloy image dataset.

Methods	Precision	Recall	F-score
EWCVT [19]	0.838385	0.962131	0.896005
MeanShift [6]	0.911927	0.844106	0.876707
GraphBased [10]	0.704163	0.928424	0.800891
SRM [15]	0.81018	0.800006	0.805061
gPb [1]	0.828988	0.866076	0.847126
NormalizedCuts [16]	0.736609	0.691646	0.71342
3D Levelset [20]	0.739025	0.581001	0.650554
3D Watershed [14]	0.864594	0.589135	0.700767
StreamGBH [21]	0.454185	0.792653	0.577479
Proposed	**0.957377**	**0.896125**	**0.925739**

6 CrackTree: Automatic Crack Detection from Pavement Images

As an extension to other application domains, we developed a graph-based algorithm to detect cracks from road pavement images [23]. The diagram of the proposed method is illustrated in Fig. 17. We first developed a new geodesic shadow-removal algorithm to remove the pavement shadows. Compared to many classical shadow-removal algorithms, the geodesic shadow-removal algorithm can automatically identify and more accurately model the large penumbra areas with strong particle textures. After removing the shadows, we construct a crack probability map using tensor voting [13] on detected noisy crack pixels and crack fragments. We then construct a graph model by sampling crack seeds from the crack probability map, construct the minimum spanning tree (MST) of the graph, and conduct recursive edge pruning in the MST to identify the final crack curves. In practice, different cracks or crack fragments may show different widths. In this work, we focus on detecting the location and shape of the crack curves, but not the crack width.

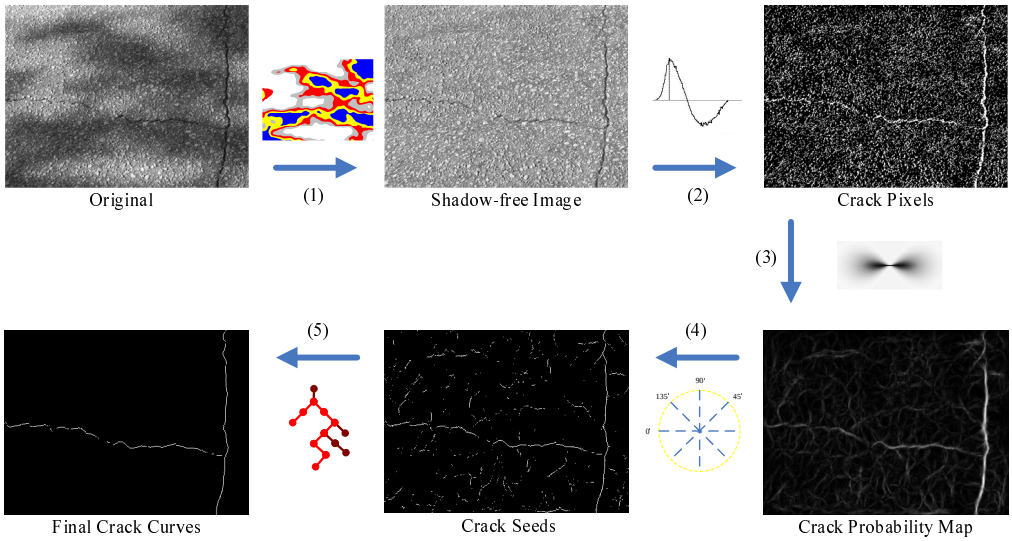


Figure 17: Flow chart of the developed CrackTree method. (1) Geodesic shadow removal, (2) local intensity-difference analysis, (3) tensor voting, (4) crack seed sampling, and (5) minimum spanning tree construction and edge pruning.

We conducted experiments on 34 real pavement images with cracks. The quantitative performance is shown in Table 3. Detection results on five pavement images are shown in Fig. 18, with comparison to four comparison methods [11, 12].

Table 3: Crack detection performance on 34 real images, with comparison to other existing methods.

Method	pbCGTG		gpb		pbCanny		Seg-ext		CrackTree	
With Shadow Removal?	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Precision	0.32	**0.34**	0.34	**0.36**	0.30	**0.30**	0.35	**0.57**	0.60	**0.79**
Recall	0.36	**0.36**	0.34	**0.49**	0.19	**0.21**	0.45	**0.63**	0.59	**0.92**
F-measure	0.34	**0.35**	0.34	**0.41**	0.23	**0.25**	0.39	**0.59**	0.59	**0.85**

Figure 18: Crack detection on five images (column 1 through 5). Row 1: original images. Row 2: shadow-removal results. Row 3: cracks detected by the proposed CrackTree. Row 4: cracks detected by the Seg-ext method. Row 5: cracks detected by pbCanny (with best F-measure). Row 6: cracks detected by gpb (with best F-measure). Row 7: cracks detected by pbCGTG (with best F-measure). Row 8: ground-truth cracks.

References

[1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *TPAMI*, 33:898–916, 2011.

[2] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[3] Y. Cao, L. Ju, and S. Wang. Grain segmentation of 3D superalloy images using multichannel EWCVT under human annotation constraints. In *European Conference on Computer Vision*, pages 244–257, 2012.

[4] Y. Cao, L. Ju, Y. Zhou, and S. Wang. 3D superalloy grain segmentation using a multichannel edge-weighted centroidal voronoi tessellation algorithm. *IEEE Trans. Image Process.*, 22:4123–4135, 2013.

[5] Y. Cao, L. Ju, Q. Zou, C. Qu, and S. Wang. A multichannel edge-weighted centroid Voronoi tessellation algorithm for 3D superalloy image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 17–24, 2011.

[6] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

[7] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[8] T. Cour, F. Benezit, and J. Shi. Spectral segmentation with multiscale graph decomposition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1124–1131, 2005.

[9] Django Software Foundation. Django (version 1.5).

[10] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59:167–181, 2004.

[11] F. Liu, G. Xu, Y. Yang, X. Niu, and Y. Pan. Novel approach to pavement cracking automatic detection based on segment extending. In *International Symposium Knowledge Acquisition and Modeling*, pages 610–614, 2008.

[12] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *International Conference on Computer Vision*, pages 416–423, 2001.

[13] G. Medioni, M. Lee, and C. Tang. *A computational framework for segmentation and grouping*. Elsevier Sicence, 2000.

[14] F. Meyer. Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125, 1994.

[15] R. Nock and F. Nielsen. Statistical region merging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1452–1458, 2004.

[16] J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[17] O. Veksler. *Efficient graph-based energy minimization methods in computer vision*. PhD thesis, Cornell University, Ithaca, NY, USA, 1999.

[18] J. Waggoner, Y. Zhou, J. Simmons, M. D. Graef, and S. Wang. 3D materials image segmentation by 2D propagation: A graph-cut approach considering homomorphism. *IEEE Transactions on Image Processing*, 22:5282–5293, 2013.

[19] J. Wang, L. Ju, and X. Wang. An edge-weighted centroidal voronoi tessellation model for image segmentation. *IEEE Transactions on Image Processing,*, 18:1844 –1858, 2009.

[20] R. T. Whitaker. A level-set approach to 3d reconstruction from range data. *International Journal of Computer Vision*, 29:203–231, 1998.

[21] C. Xu, C. Xiong, and J. J. Corso. Streaming hierarchical video segmentation. In *European Conference on Computer Vision*, 2012.

[22] Y. Zhou, L. Ju, Y. Cao, J. Waggoner, Y. Lin, J. Simmons, and S. Wang. Edge-weighted centroid voronoi tessellation with propagation of consistency constraint for 3d grain segmentation in microscopic superalloy images. In *CVPR Workshop on Perception Beyond the Visible Spectrum (PBVS)*, 2014.

[23] Q. Zou, Y. Cao, Q. Li, Q. Mao, and S. Wang. Cracktree: Automatic crack detection from pavement images. *Pattern Recognition Letters*, 33:227–238, 2012.

Personnel:

Faculty:

Song Wang*

Graduate Students:

Dhaval Salvi* (Ph.D. in 2014)
Yu Cao* (Ph.D. in 2013)
Jarrell Waggoner* (Ph.D. in 2013)
Zhiqi Zhang (Ph.D. in 2012)
Andrew Temlyakov (Ph.D. in 2012)
Jun Zhou (Ph.D. student, started from 2010)
Dazhou Guo (Ph.D. student, started from 2010)
Yuewei Lin* (Ph.D. student, started from 2011)
Youjie Zhou* (Ph.D. student, started from 2012)
Kang Zheng* (Ph.D. student, started from 2012)
Xiaochuan Fan* (Ph.D. student, started from 2012)
Yang Mi* (Ph.D. student, started from 2013)
Hongkai Yu (Ph.D. student, started from 2014)
Kareem Ezz El-Deen (Ph.D. student, started from 2014)

* Partially supported by this AFOSR grant.

Publications:

1. R. Meng, S. Nelakuditi, S. Wang, R. Choudhury. OmniView: A Mobile Collaborative System for Assisting Drivers with a Map of Surrounding Traffic, to appear in *International Conference on Computing, Networking and Communications (ICNC)*, Anaheim, CA, 2015

2. J. Waggoner, Y. Zhou, J. Simmons, M. De Graef, S. Wang. Topology-Preserving Multi-Label Image Segmentation, to appear in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa Beach, HI, 2015

3. D. Salvi, K. Zheng, Y. Zhou, S. Wang. Distance Transform Based Active Contour Approach for Document Image Rectification, to appear in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa Beach, HI, 2015

4. Y. Zhou, L. Ju, S. Wang. Multiscale Superpixels and Supervoxels Based on Hierarchical Edge-Weighted Centroidal Voronoi Tessellation, to appear in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa Beach, HI, 2015

5. K. Zheng, Y. Lin, Y. Zhou, D. Salvi, X. Fan, D. Guo, Z. Meng, S. Wang. Video-based Action Detection using Multiple Wearable Cameras, *ECCV ChaLearn Looking at People Workshop*, Zurich, Switzerland, 2014

6. Q. Zou, Y. Cao, Q. Li, C. Huang, S. Wang. Chronological Classification of Ancient Paintings using Appearance and Shape Features, *Pattern Recognition Letters*, 49:146-154, 2014

7. X. Fan, K. Zheng, Y. Zhou, S. Wang. Pose Locality Constrained Representation for 3D Human Pose Reconstruction, *European Conference on Computer Vision (ECCV)*, Zurich, Switzerland, 2014

8. Y. Zhou, L. Ju, Y. Cao, J. Waggoner, Y. Lin, J. Simmons, S. Wang. Edge-Weighted Centroid Voronoi Tessellation with Propagation of Consistency Constraint for 3D Grain Segmentation in Microscopic Superalloy Images, *CVPR Workshop on Perception Beyond the Visible Spectrum (PBVS)*, Columbus, OH, 2014

9. Q. Zou, Y. Cao, Q. Li, Q. Mao, S. Wang. Automatic Inpainting by Removing Fence-Like Structures in RGBD Images, *Machine Vision and Applications*, 25(7):1841-1858, 2014

10. J. Waggoner, Y. Zhou, J. Simmons, M. De Graef, S. Wang. Graph-Cut Based Interactive Segmentation of 3D Materials-Science Images, *Machine Vision and Applications*, 25(6):1615-1629, 2014 (PDF)

11. S. Singh, R. Meng, S. Nelakuditi, Y. Tong, S. Wang. SideEye: Mobile Assistant for Blind Spot Monitoring, *International Conference on Computing, Networking and Communications (ICNC)*, Honolulu, Hawaii, 403-407, 2014

12. J. Waggoner, Y. Zhou, J. Simmons, M. De Graef, S. Wang. 3D Materials Image Segmentation by 2D Propagation: A Graph-Cut Approach Considering Homomorphism, *IEEE Transactions on Image Processing*, 22(12):5282-5293, 2013

13. Y. Cao, L. Ju, Y. Zhou, S. Wang. 3D Superalloy Grain Segmentation Using a Multichannel Edge-Weighted Centroidal Voronoi Tessellation Algorithm, *IEEE Transactions on Image Processing*, 22(10):4123-4135, 2013

14. J. Dong, F. Chen, Q. Cheng, S. Wang. A New Criterion for Choosing Planar Subproblems in MAP-MRF Inference, *Neurocomputing*, 120:453-460, 2013

15. Y. Cao, D. Barrett, A. Barbu, S. Narayanaswamy, H. Yu, A. Michaux, Y. Lin, S. Dickinson, J. Siskind, S. Wang. Recognizing Human Activities from Partially Observed Videos, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2658-2665, Portland, OR, 2013

16. J. Waggoner, Y. Zhou, J. Simmons, A. Salem, M. De Graef, S. Wang. Interactive Grain Image Segmentation using Graph Cut Algorithms, *Proceedings of SPIE Volume 8657 (Computational Imaging XI)*, Burlingame, CA, 2013

17. A. Temlyakov, P. Dalal, J. Waggoner, D. Salvi, S. Wang. Shape and Image Retrieval by Organizing Instances Using Population Cues, *IEEE Workshop on the Applications of Computer Vision (WACV)*, 303-308, Clearwater, FL, 2013

18. D. Salvi, J. Waggoner, A. Temlyakov, S. Wang. A Graph-Based Algorithm for Multi-Target Tracking with Occlusion, *IEEE Workshop on the Applications of Computer Vision (WACV)*, 489-496, Clearwater, FL, 2013

19. D. Salvi, J. Zhou, J. Waggoner, S. Wang. Handwritten Text Segmentation using Average Longest Path Algorithm, *IEEE Workshop on the Applications of Computer Vision (WACV)*, 505-512, Clearwater, FL, 2013

20. J. Tian, C. Qu, W. Xu, S. Wang. KinWrite: Handwriting-Based Authentication Using Kinect, *Annual Network & Distributed System Security Symposium (NDSS)*, San Diego, CA, 2013

21. Y. Lin, Y. Y. Tang, B. Fang, Z. Shang, Y. Huang, S. Wang. A Visual-Attention Model Using Earth Mover's Distance based Saliency Measurement and Nonlinear Feature Combination, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):314-328, 2013

22. F. Chen, Q. Cheng, H. Liu, W. Xu, S. Wang. A New Inference Framework for Dependency Networks, *Communications in Statistics Theory and Methods*, 42(1):56-75, 2013

23. B. C. Munsell, A. Temlyakov, C. Qu, S. Wang. Person Identification Using Full-Body Motion and Anthropometric Biometrics from Kinect Videos, *3rd Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams*, Firenze, Italy, 2012

24. A. Barbu, A. Bridge, Z. Burchill, D. Coroian, S. Dickinson, S. Fidler, A. Michaux, S. Mussman, S. Narayanaswamy, D. Salvi, L. Schmidt, J. Shangguan, J. M. Siskind, J. Waggoner, S. Wang, Y. Yin, Z. Zhang. Video In Sentences Out, *Conference on Uncertainty in Artificial Intelligence (UAI)*, 102-112, Catalina Island, CA, 2012

25. Y. Cao, L. Ju, S. Wang. Grain Segmentation of 3D Superalloy Images Using MultiChannel EWCVT under Human Annotation Constraints, *European Conference on Computer Vision (ECCV)*, Firenze, Italy, 2012

26. P. Dalal, S. Wang. Landmark Sliding for 3D Shape Correspondence, *Intelligent Data Analysis for Real-Life Applications: Theory and Practice*, IGI Global, 57-71, 2012

27. J. Waggoner, J. Simmons, M. De Graef, S. Wang. Graph Cut Approaches for Materials Segmentation Preserving Shape, Appearance, and Topology, *International Conference on 3D Materials Science*, 147-152, Seven Springs, PA, 2012

28. Z. Zhang, S. Fidler, J. Waggoner, Y. Cao, S. Dickinson, J. M. Siskind, S. Wang. Superedge Grouping for Object Localization by Combining Appearance and Shape Information, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3266-3273, Providence, RI, 2012

29. X. Zhou, S. Wang, H. Chen, T. Hara, R. Yokoyama, M. Kanematsu, H. Fujita. Automatic Localization of Solid Organs on 3D CT Images by a Collaborative Majority Voting Decision based on Ensemble Learning, *Computerized Medical Imaging and Graphics*, 36(4):304-313, 2012

30. Q. Cheng, F. Chen, W. Xu, S. Wang. Recursive SumProduct Algorithm for Generalized Outer-Planar Graphs, *Information Processing Letters*, 112(11):449-456, 2012

31. J. Waggoner, J. Simmons, S. Wang. Combining Global Labeling and Local Relabeling for Metallic Image Segmentation, *Proceedings of SPIE Volume 8296 (Computational Imaging X)*, Burlingame, CA, 2012

32. B. C. Munsell, A. Temlyakov, M. Styner, S. Wang. Pre-organizing Shape Instances for Landmark-Based Shape Correspondence, *International Journal of Computer Vision*, 97(2):210-228, 2012

33. Q. Zou, Y. Cao, Q. Li, Q. Mao, S. Wang. CrackTree: Automatic Crack Detection from Pavement Images, *Pattern Recognition Letters*, 33(3):227-238, 2012

34. F. Chen, Q. Wang, S. Wang, W. Zhang, W. Xu. Object Tracking via Appearance Modeling and Sparse Representation, *Image and Vision Computing*, 29(11):787-796, 2011

35. S. Wang, J. Waggoner, J. Simmons. Graph-cut Methods for Grain Boundary Segmentation, *JOM Journal of the Minerals, Metals and Materials Society*, 63(7):49-51, 2011

36. Y. Cao, L. Ju, Q. Zou, C. Qu, S. Wang. A Multichannel Edge-Weighted Centroidal Voronoi Tessellation Algorithm for 3D Superalloy Image Segmentation, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 17-24, Colorado Springs, CO, 2011

Dissertations:

Four Ph.D. dissertations are produced under the support of this AFOSR grant. These dissertations are publicly available at the University of South Carolina Library and ProQuest Dissertations & Theses Database.

1. Document Image Analysis Techniques for Handwritten Text Segmentation, Document Image Rectification and Digital Collation, Dhaval Salvi, Ph.D., 2014

   Abstract: Document image analysis comprises all the algorithms and techniques that are utilized to convert an image of a document to a computer readable description. In this work we focus on three such techniques, namely (1) Handwritten text segmentation (2) Document image rectification and (3) Digital Collation.

   Offline handwritten text recognition is a very challenging problem. Aside from the large variation of different handwriting styles, neighboring characters within a word are usually connected, and we may need to segment a word into individual characters for accurate character recognition. Many existing methods achieve text segmentation by evaluating the local stroke geometry and imposing constraints on the size of each resulting character, such as the character width, height and aspect ratio. These constraints are well suited for printed texts, but may not hold for handwritten texts. Other methods apply holistic approach by using a set of lexicons to guide and correct the segmentation and recognition. This approach may fail when the lexicon domain is insufficient. In the first part of this work, we present a new global non-holistic method for handwritten text segmentation, which does not make any limiting assumptions on the character size and the number of characters in a word.

   Digitization of document images using OCR based systems is adversely affected if the image of the document contains distortion (warping). Often, costly and precisely calibrated special hardware such as stereo cameras, laser scanners, etc. are used to infer the 3D model of the distorted image which is used to remove the distortion. Recent methods focus on creating a 3D shape model based on 2D distortion information obtained from the document image. The performance of these methods is highly dependent on estimating an accurate 2D distortion grid. These methods often affix the 2D distortion grid lines to the text line, and as such, may suffer in the presence of unreliable textual cues due to preprocessing steps such as binarization. In the domain of printed document images, the white space between the text lines carries as much information about the 2D distortion as the text lines themselves. Based on this intuitive idea, in the second part of our work we build a 2D distortion grid from white space lines, which can be used to rectify a printed document image by a dewarping algorithm.

   Collation of texts and images is an indispensable but labor-intensive step in the study of print materials. It is an often used methodology by textual scholars when the underlying manuscript

of the text is nonexistent. Although various methods and machines have been designed to assist in this labor, it still remains an expensive and time-consuming process, often requiring travel to distant repositories for the painstaking visual examination of multiple original copies. Efforts to digitize collation have so far depended on first transcribing the texts to be compared, thus introducing into the process more labor and expense, and also more potential error. Digital collation will instead automate the first stages of collation directly from the document images of the original texts, thereby speeding the process of comparison. We describe such a novel framework for digital collation in the third part of this work.

2. 3D Grain Segmentation in Superalloy Images using Multichannel Edge-weighted Centroidal Voronoi Tessellation based Methods, Yu Cao, Ph.D., 2013

   Abstract: Accurate grain segmentation on 3D superalloy images is very important in materials science and engineering. From grain segmentation, we can derive the underlying superalloy grains' micro-structures, based on which many important physical, mechanical and chemical properties of the superalloy samples can be evaluated. However, grain segmentation is usually a very challenging problem since: 1) even a small 3D superalloy sample may contain hundreds of grains; 2) carbides and noises may degrade the imaging quality; and 3) the intensity within a grain may not be homogeneous. In addition, the same grain may present different appearances, i.e. intensities, under different microscope settings. In practice, a 3D superalloy image may contain multichannel information where each channel corresponds to a specific microscope setting. In this work, we develop a Multichannel Edge-Weighted Centroidal Voronoi Tessellation (MCEWCVT) algorithm to effectively and robustly segment the superalloy grains in 3D multichannel superalloy images. MCEWCVT performs segmentation by minimizing an energy function which encodes both the multichannel voxel-intensity similarity within each cluster in the intensity domain and the smoothness of segmentation in the 3D image domain. Based on MCEWCVT, we further develop a Constrained Multichannel Edge-Weighted Centroidal Voronoi Tessellation (CMEWCVT) algorithm which can take manual segmentation on a small number of selected 2D slices as constraints from the problem domain, and incorporate them into the energy minimization process to further improve the segmentation accuracy. We quantitatively evaluate the MCEWCVT and the CMEWCVT algorithms on an authentic Ni-based dataset and two synthesized datasets against ground-truth segmentation. The qualitative and quantitative comparisons among the MCEWCVT, the CMEWCVT and 18 existing image segmentation algorithms on the authentic dataset demonstrate the effectiveness and robustness of the MCEWCVT and the CMEWCVT algorithms. In addition, the experiments on two synthesized datasets indicate that the optimal algorithm parameters found in the testing on the authentic dataset can be used on other superalloy datasets which have similar size and number of grains. Major results of this dissertation are summarized in Section 4 of this report.

3. Multi-Label Segmentation Propagation for Materials Science Images Incorporating Topology and Interactivity, Jarrell Waggoner, Ph.D., 2013

   Abstract: Segmentation propagation is the problem of transferring the segmentation of an image to a neighboring image in a sequence. This problem is of particular importance to materials science, where the accurate segmentation of a series of 2D serial-sectioned images of multiple, contiguous 3D structures has important applications. Such structures may have prior-known shape, appearance, and/or topology among the underlying structures which can be considered to improve segmentation accuracy. For example, some materials images may have structures with a specific shape or appearance in each serial section slice, which only changes minimally from slice to slice; and some materials may exhibit specific topology which

constrains their structure or neighboring relations. In this work, we develop a framework for materials image segmentation that segments a variety of materials image types by repeatedly propagating a 2D segmentation from one slice to another, and we formulate each step of this propagation as an optimal labeling problem that can be efficiently solved using the graph-cut algorithm. During this propagation, we propose novel strategies to enforce the shape, appearance, and topology of the segmented structures, as well as handling local topology inconsistency. Most recent works on topology-constrained image segmentation focus on binary segmentation, where the topology is often described by the connectivity of both foreground and background. We develop a new multi-labeling approach to enforce topology in multiple-label image segmentation. In this case, we not only require each segment to be a connected region (intra-segment topology), but also require specific adjacency relations between each pair of segments (inter-segment topology). Finally, we integrate an interactive approach into the proposed framework that improves the segmentation by allowing minimal and simplistic human annotations. We justify the effectiveness of the proposed framework by testing it on various 3D materials images, and we compare its performance against several existing image segmentation methods. Major results of this dissertation are summarized in Sections 1,2 and 3 of this report.

4. Object Localization by Combining Shape and Appearance Features, Zhiqi Zhang, Ph.D., 2012

Abstract: Object localization is an important task in computer vision, which is usually handled by searching for an optimal subwindow that tightly covers the object of interest. Both boundary-based shape and region-based appearance features are important to accurate object localization. For some objects, shape feature might be more important and for some objects, appearance feature might be more important. However, current state-of-the-art object localization methods either focus on shape feature or focus on appearance feature, and efficiently combining shape and appearance features to achieve object localization is a very challenging research topic in computer vision. In addition, the subwindows considered in previous work are usually limited to rectangles or other specified, simple shapes. With such specified shapes, there may not exist a subwindow that can cover the object of interest tightly. As a result, the desired subwindow around the object of interest may not be optimal in terms of the localization objective function and cannot be detected by the subwindow search algorithm. In my dissertation, to address the above problems we proposed new approaches to combine shape and appearance features for object localization, in a globally optimal fashion, using graph-theoretic models and algorithms. We first develop an edge grouping based free-shape subwindow search algorithm for object localization, where no specific shape features of individual object classes are considered. We just generally require the bounding contour (free-shape subwindow) to be aligned with detected edges and cover the desired object appearance features that are learned from a training set. This requirement is quantified and integrated into the localization objective function based on the widely-used bag of visual words technique. We then extend the edge grouping based freeshape subwindow search method to super-edge grouping method, where both the shape and appearance features of specific object classes are learned and then integrated to the object localization algorithm. Experiments show that our proposed method, by integrating both boundary-based shape feature and region-based appearance feature, can produce better localization performance than the previous state-of-the-art subwindow search methods.