Alpha-Tree Segmentation of Human Anatomical Photographic Imagery

Wilbert Tabone
University of Malta
wilbert.tabone.15@um.edu.mt

Michael H.F. Wilkinson
University of Groningen
m.h.f.wilkinson@rug.nl

Anne E.J.V. Gaalen
University of Groningen
a.e.j.van.gaalen@umcg.nl

Janniko Georgiadis
University of Groningen
j.r.georgiadis@umcg.nl

George Azzopardi
University of Groningen
g.azzopardi@rug.nl

ABSTRACT

Segmentation of anatomical imagery is important in several areas, such as forensics, medical analysis and educational material. The manual segmentation of such images and the subsequent labelling of regions is a very laborious task. We propose an interactive segmentation scheme which we evaluate on a new data set of anatomical imagery. We use a morphological tree-based segmentation method, known as the alpha-tree, together with a Hu-moment thresholding mechanism in order to extract segments from a number of structures. Both qualitative and quantitative results in anatomical imagery of embalmed head, arm and leg specimens indicate that the proposed method can produce meaningful segmentation outputs, which could facilitate further refined labelling.

CCS CONCEPTS

• **Computing methodologies → Image segmentation**;

KEYWORDS

segmentation, anatomy, mathematical morphology, alpha tree, Hu moments.

1 INTRODUCTION

Segmentation of anatomical parts is not a trivial problem and both laborious and time consuming. A (semi) automatic segmentation and labelling process would alleviate these issues. The process presents a number of challenges such as the instances where neighbouring anatomical parts may have similar texture and colour, as

evidenced by Fig. 1. We propose an algorithm that is based on the alpha-tree combined with Hu moments for the segmentation of such images. In order to measure the performance of the proposed method, we employ an evaluation process consisting of both human-based qualitative and quantitative experiments. We perform experiments on a new proprietary photographic medical imagery set of various human anatomical specimens obtained from the University Medical Centre Groningen (UMCG) body donation programme.

The rest of the paper is organised as follows: Section 2 gives an account of related works, Section 3 describes the proposed methodology and Section 4 presents our results. Sections 5 and 6 provide a discussion and conclusions, respectively.

2 RELATED WORK

Automation of clinical segmentation using low-level model-free techniques such as region growing, edge detection and mathematical morphology are difficult due to the shape complexity and variability within and across the individual specimens [19]. The use of shape priors in the form of shape atlases and deformable model techniques have been applied to alleviate such problems and attempt to automate the process of manual segmentation [11].

The latter have been proven to be effective in matching, tracking and segmenting anatomical structures utilising the constraints derived from the image data and the a priori knowledge about the location, size and shape of the structures [8]. Examples of these are the *snake* and *scissors* approaches. The a priori knowledge allows the models to tolerate the variability and complexity of biological structures while aiding in overcoming the challenges faced by noise. The latter two issues are problems that may cause boundaries of structures to be indistinct and disconnecting but can be overcome through smoothing as a pre-processing step [19]. Furthermore, since deformable models are semi-automatic methods, interactive mechanisms are provided to medical experts which allow them to contribute their expertise to the segmentation task when needed [8].

To our knowledge there are no published projects which have segmented photographic human anatomical imagery utilising tree-based methods. We focus on these latter methods due to their flexible nature in effective and efficient re-computation with new parameters and overall computational feasibility. Tree-based methods such as the binary partition tree (BPT) [12] and the alpha tree [10], α-tree hereinafter, have also been applied in order to segment photographic images [2, 5, 6, 9]. It is noted that although the binary partition tree usually yields an effective segmentation result, its

computational complexity is very high: $O(n^2 \log n)$. This problem is caused at the lower level of the implementation as the neighbours of every region are reorganised and combined after each merge, and the link queue updated [10]. In contrast, the α-tree, with complexity of $O(n \log n)$, has a fixed merging order which is determined from the differences between neighbouring pixels. Moreover, the α-tree deals well with colour and can also work on vector images, unlike component trees such as min and max trees [13], which have found use in enhancing anatomical details in grey-scale medical imaging data interactively [18]. Similar interactive filtering and segmentation has been performed using α-trees [9]. We therefore choose to investigate the problem at hand with the α-tree algorithm [10], mainly based on its computational performance, interactivity and effectiveness in other applications [2, 3].

Figure 1: An embalmed specimen displaying uniform grey discolouration.

3 METHODOLOGY

3.1 Pre-processing

The photographed specimens used are embalmed and not fresh frozen specimens (FFS). The process of embalming for anatomy education typically utilises concentrated formaldehyde with a formalin concentration ranging from 37% to 40%. The solution is then usually injected into the artery under high pressure and allowed to be absorbed into the tissue for a number of hours until it is optionally drained out [7].

Such high concentration of formaldehyde mixed with the blood and a lack of red colouration agents (which are mostly used in non-medical embalming) lead to a uniform grey discolouration across the entire specimen (cadaver), an example of which can be seen in Fig. 1. This grey discolouration, known as "formaldehyde grey" or "embalmer's grey" [7] poses a challenge to the segmentation process since it would be more difficult to depend on colour in the region creating process when there is little or no distinction in colour between different regions. To alleviate this problem, we employ mean shift edge-preserving smoothing [1] in order to reduce high frequency noise while preserving the edges.

3.2 Segmentation: constrained connectivity, α-tree, and constrained α-tree

In constrained connectivity, image I is projected in a graph space where the vertices correspond to pixels and the edges to pairs of adjacent vertices. The main idea is that if the dissimilarity $d(x, y)$ between x and y, which are two adjacent elements, is less than or equal to the dissimilarity measure α, then the two are directly connected and hence an edge exists between the two elements; this makes them members of the same connected component denoted by α-CC [14, 15]. If $d(x, y) > \alpha$, then it suggests that there is no direct linkage between the elements, but are not excluded from belonging to the same α-CC.

In the simplest form of connected components, we use the Euclidean distance between the vector-valued pixels. We then obtain the segmentation for the given α by considering as connected all pixels which can be chained by a path of successive adjacent pixels whose dissimilarities do not exceed the threshold value α. The detail level of the segmentation can vary between fine to coarse by the adjustment of the α threshold from 0 to the maximum value. The 0-CC corresponds to the finest level whilst the maximum CC results in the coarsest level which matches the whole image definition domain [2]. Such a nested series of fine to coarse segmentation produces a hierarchy which can be viewed as a spatially rooted dendrogram [14] given the name of the α-tree [10].

We use an efficient union-find based algorithm to generate a tree representation of the totally ordered set of graph-space partitions. The processes on this tree can be launched interactively and in real-time from a separate module that allows the manipulation of the various attribute thresholds.

Given an image and set of all α values A, the set $\{\mathbf{P}\}^A$ of all the α-partitions of its definition domain E is re-arranged as a partition pyramid according to the dissimilarity range α, with the base of the pyramid corresponding to the finest partitions ($\alpha = 0$) of E and the tip to the coarse partitions ($\alpha_{max} - CC$).

Unfortunately there exists the problem of redundancy as some α-CCs, may persist in more than one level $\triangle_\alpha^A$ of the pyramid system. These α-CCs are replicas of previous connectivity constraints that appeared at a smaller α level and will persist until a merge occurs with another α-CC at a higher level. In order to solve this issue, [10] introduced an index mapping for these α-CCs which leads to a hierarchical partition structure constricted by rules of inclusion. These rules operate on the pyramid in order to generate the α-partition hierarchy which contains only the elements that appear for the first time at every pyramid level. The redundant data is not lost however, as the α-partition hierarchy is a lossless compression of the α-partition pyramid and hence every $\triangle_\alpha^A$ can be restored.

In order to solve the problem of leakage - where two regions with substantial colour differences could get inappropriately merged together, we adapt the global range parameter ω from [14] in order to constrain the segmentation.

The authors in [14] define (α, ω)-connectivity as:

$$\alpha, \omega - CC = \max(\alpha' - CC(p) \parallel \alpha' \le \alpha \wedge R(\alpha' - CC(p)) \le \omega) \quad (1)$$

where R denotes the range of differences between a pixel p_i and its neighbour p_{i+1}. This (α, ω)-connectivity based on the above definition employs two predicates, with the first returning true if the maximum indexed component α_i in the tree is less than or equal to α, whilst the second would restrict the total element colour variation to be less than or equal to ω. Therefore, now a region not only has to conform to CC-threshold α but also to a new constraint

ω, which globally optimises the merges based on its value. Hence, the merge is forbidden if the resulting range of colours exceed ω.

3.3 Post-processing: vector attributes

In order to enhance the localisation of the different elongated structures irrespective of their position, scale and rotation, we adapt Hu's geometric moments [4]. This entails the assignment of vector attributes [17] to every node of the α-tree which are calculated incrementally. When a new region is created through the merger of the other regions, the attributes of the new region are derivable from the existing attributes which together form an attribute vector. A reference vector is created for every object of interest, against which the attribute vector of every node must be compared to with some distance measure, such as the Euclidean distance, in order to determine if the object should be accepted if its distance is below a defined threshold.

There exist a variety of types of attributes to use, such as area and colour. In order to detect certain shapes, attributes which describe these shapes are required and it is desirable that such attributes are invariant to location, scale or rotation. For the context of our system, the use of moments is required in order to focus the segmentation on interesting structures such as those having a degree of elongation. This category includes ligaments, bones and veins which could be highlighted through the use of the first Hu moment.

Hu [4] defines these image moments as:

$$M_{pg} = \int \int_{\mathbb{R}^2} x^p y^p f(x, y) \, dx \, dy \tag{2}$$

From these moments, the central moments can then be calculated. Their main property is that they are translation invariant and hence if the pixels of a shape are shifted by the same amount, the yielded result will be the same. These central moments are given by:

$$\mu_{pq} = \int \int_{\mathbb{R}^2} (x - \overline{x})^p (y - \overline{y})^q f(x, y) \, dx \, dy \tag{3}$$

$$\text{where,} \quad \overline{x} = \frac{m_{10}}{m_{00}}, \overline{y} = \frac{m_{01}}{m_{00}} \tag{4}$$

The normalised moments can then be calculated from the central moments. Besides being translation invariant, normalised moments are also scale invariant. These are given by:

$$\eta_{pg} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{5}$$

$$\text{where,} \quad \gamma = \frac{p + q}{2} + 1 \tag{6}$$

Further to the above moments, new moments which are also invariant to rotation can then be derived. These are the Hu's invariant moments and the first four are given below:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{7}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{8}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \tag{9}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \tag{10}$$

3.4 Interactive α-tree

In order to allow further dynamic experimentation on images using different parameters which can be dynamically adjusted, we developed a GUI application using Matlab. The tool offers the possibility of α-tree parameter and Hu moment threshold manipulation. Furthermore, we use colour coding to render the resulting map of the proposed segmentation method. This enhances the Gestalt psychological grouping effect for the user and facilitates the evaluation of the results obtained by the combination of parameters. This feature produces the final labelled result which could be one of three main output types. These are the segmentation results utilising only the λ threshold and the optional constraining ω factor, the results utilising the guide colour system and the ones using highlighted elongated structures through the utilisation of the moments. The latter output structures can also be further highlighted by toggling the guide colour system. A number of outputs using a combination of thresholds and the guide colour system are given in Fig. 2(d-f, j-l)[1].

4 EVALUATION PROCEDURE AND RESULTS

4.1 Data Set

The photographs used for our research are of specimens from the body donation programme at the University Medical Center Groningen (UMCG). Each specimen has been embalmed in a 4% concentrate formaldehyde solution. A 50 Megapixel Canon EOS DSLR was utilised in order to capture the 5760x3840 photographs with 240 pixels/inch resolution.

Three main specimens consisting of an arm, head and leg were chosen in order to demonstrate the effectiveness of our method on different kinds of structures. Each specimen was photographed hundreds of times from varying angles producing a collection of thousands to proes, of which two distinct angles per specimen where selecteddof imaguce a set of six main test images, Fig. 2(a-c, g-i).

4.2 Evaluation

We split the evaluation routine into three main categories. First, the qualitative evaluation of a subset of segmentation outputs from the developed system by experts in the anatomical field. Then, we compared the segmentation results of specimens photographed from different angles, and finally we made an overall examination of the segmentation results of the entire collection of tests in order to gather qualitative data from the experts.

We used a subset of segmentation results from the photographic data set. We first segmented every image with only the λ value set and then by a repetition of the previous test together with alternating values of the moments threshold (0.0, 0.5, 1.0). This was followed by a final test which included alternating values of ω (42, 46, 49). The constant λ value of 7 was chosen after being commended by the medical experts as producing results that retained the most structure details, as can also be said with respect to the ω factor values of 42, 46 and 49. Furthermore, the values of 0.5 and 1.0 were

---

[1]Each resultant image from Fig. 2 corresponds to the following parameter group numbers (Img.No.) from Table 2: (c) - 2, (e) - 9, (f) - 1 (j) - 2, (k) - 2, (l) - 9.
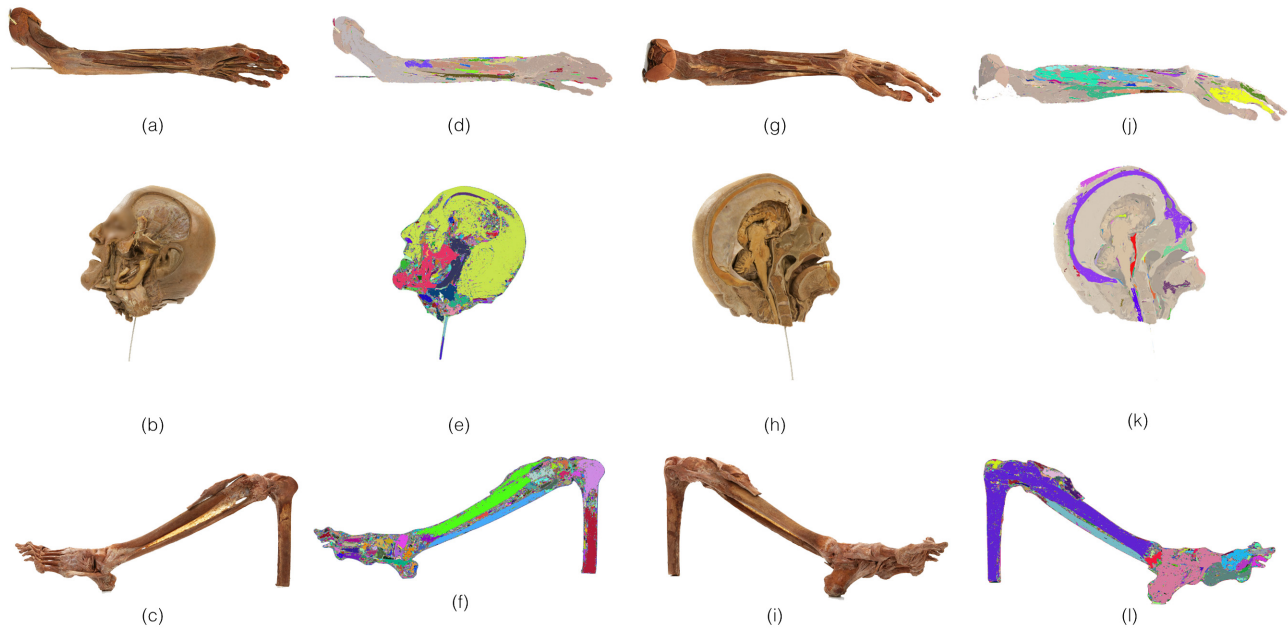
Figure 2: (a-c, g-i): Images chosen to be segmented in order to create the evaluation image set. These were assigned the following labels, *(a)*: ARM1, *(g)*: ARM2, *(b)*: HEAD1, *(h)*: HEAD2, *(c)*: LEG1, *(i)*; LEG2. (d-f, j-l): Sample of segmented and colour labeled results using a combination of α-tree thresholding, Hu moments thresholding and filtering.

Table 1: Selected subset of filtering parameter combinations.

Number of iterations (i)	Kernel width (h)
2	40
2	80
3	20
3	40

Table 2: A list of parameters used for evaluation. i⇒filter iteration count, h⇒ filter kernel width.

Img. No.	Pre-proc.	i	h	λ	ω	Mom. Thres.
1	N	0	0	7	0	0.0
2	N	0	0	7	0	0.5
3	N	0	0	7	0	1.0
4	Y	0	0	7	42	0.0
5	Y	0	0	7	46	0.0
6	Y	0	0	7	49	0.0
7	Y	2	40	7	0	0.0
8	Y	2	80	7	0	0.0
9	Y	3	20	7	0	0.0
10	Y	3	40	7	0	0.0

selected as the standard values for the moment threshold parameter during the experiment.

In total, we therefore performed 60 segmentation procedures for each test image, resulting in a collection of 360 resultant images. In order to conduct the expert evaluation, we filtered down the collection into a smaller subset of 10 segmentation results per image for a total of 60 images. Each set of 10 images had the properties listed in Table 2. The filtering parameter combinations selected are given in Table 1. These values were selected for as they produced promising results. High *i* and *h* values produced unstable results due to the high level of smoothing. This led to the entire structure body in the image being grouped into the same α-CC during segmentation producing a singular connected structure.

4.3 Human evaluation

For each of the specimens in Fig. 2(a-c, g-i), we asked two medical experts to list all the observable structures which we subsequently sorted into categories such as muscles, tendons, and bones, among others, and compiled them into lists to be used in the generation of six questionnaires corresponding to each specimen. We placed two Likert scales for *distinctiveness* and *completeness* underneath each

structure label. These terms respectively correspond to *precision* and *recall*, and are defined as follows.

Distinctiveness: how well a structure can be distinguished from other structures.
Completeness: how much of the structure is observable.

We weighted the scales from 1 to 10, with 1 being the lowest weighting. For each test image (10 for each specimen), the experts examined the structures on a separate viewport and rated the distinctiveness and completeness of its labelling. Once this process was concluded, we calculated and subsequently normalised the averages of distinctiveness and completeness. Furthermore, we performed the *F*-measure calculation for every tested image from each subset.

Table 3: F-measurement results.

| Test Image No. | F-Measure | | | |
	Arm 1	Arm 2	Head 1	Head 2
1	**0.500**	**0.494**	0.584	**0.405**
2	0.343	0.299	0.216	0.153
3	0.193	0.115	0.150	0.100
4	0.484	0.429	0.578	0.379
5	0.499	0.423	0.586	0.383
6	0.488	0.446	0.578	0.390
7	0.341	0.485	0.528	0.251
8	0.272	0.314	0.454	0.224
9	0.383	0.485	**0.639**	0.397
10	0.185	0.337	0.329	0.202

Table 4: Difference between the F-measure values between Arm 1 - Arm 2 and Head 1 - Head 2.

| Image Number | F-Measure Difference | |
	Arm	Head
1	0.006	0.179
2	0.044	0.063
3	0.078	0.050
4	0.055	0.199
5	0.076	0.203
6	0.042	0.188
7	0.144	0.277
8	0.042	0.230
9	0.102	0.242
10	0.152	0.127
Average	**0.074**	**0.176**

4.4 Quantitative results

From the results that we present in Table 3, it is evident that the experts penalised experiments which used the constraining parameter ω and the Hu moment thresholding system since *Test Image 1* had the highest *F*-measure in most cases. *Test Image 9* has the same properties bar being filtered through the mean-shift with $i = 3$ and $h = 20$. From this we learned that the anatomists preferred a segmented image with lots of preserved detail rather than a distilled version which filters out most of the noise and retains only thresholded structures. Such is the case with the Hu moment 0, which preserves structures with a degree of elongation such as nerves, veins, bones and tendons. The F-measure scores of 0.534 and 0.561 for images 1 and 9, respectively, of the LEG subset re-affirmed the high confidence which the experts have on these parameter combinations.

We also calculated the difference between the values of the respective specimen columns in the results table in order to assess if the segmentation process produced similar results for the viewpoint pairs of the same specimens. The total average *F*-Measure variation between each experiment of the two arm structures was 0.074, whilst for the head structures was 0.176. As can be observed in Table 4, the process performed more uniformly on the arm specimen when compared to the head specimen since the difference in viewpoint for the head was far greater than for the arm, Fig. 2. The insignificance in value variation suggests that the segmentation results for slightly different viewpoints produce similar labelling. Hence, we inferred that the calculated moments did indeed aid in allowing the segmentation process to be invariant to translation, scale and rotation.
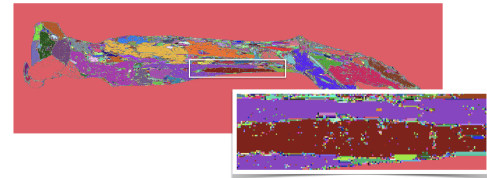
4.5 Qualitative results

The qualitative report commended the quality of the labelled results pertaining to the thin tendons running at oblique directions and the main nerve of the two arm views (volar and dorsal). Furthermore, the algorithm performed well on the nails, tendons and muscles that had an oblique fibre direction. This suggests that the segmentation process, with Hu moments, labelled elongated structures well whilst the thresholding system may have aided in enhancing the contrast of structures such as the nails, leading to a positive outcome.

5 DISCUSSION

In this work we introduced an unsupervised image processing approach which attempts to segment and label photographic imagery of our new human specimen data set in an automatic fashion. There are various directions in which this work can be extended. One direction is to investigate a pixel marker system which would allow the user to select a pixel at any particular location of the image using a pointing device and subsequently a segmentation would occur for only the ROI which bounds the selected pixel. This would be similar to the way snakes and scissors from the deformable models category, work. Another direction would be to investigate learning algorithms in order to automatically determine the best set of parameter values. A key area of work is optimal computation of dissimilarities between neighbouring pixels, as these dictate merge order, and therefore the α-tree topology. The use of other edge preserving pre-filters such as bilateral filters [16] should also be investigated. A more thorough validation on many more samples must also be carried out. A final improvement to the output images would be to seal their pixel 'holes' or 'islands'. An example of an image containing such 'holes' is depicted in the detailed inset of Figure 3.

Figure 3: One of the images from the arm subsets with an inset detailing the pixel 'holes'.

Such structures could be sealed using techniques from digital image inpainting which includes techniques such as: PDE based image inpainting, exemplar based image inpainting, hybrid image inpainting and appearance-based techniques. Other approaches are dilation and morphological closing. Furthermore, we believe that

our application can be utilised for educational purposes whereby the images can be used in the generation of virtual atlases and could also contribute to clinical applications that may require surgery guidance. This work opens a new technical window whereby further research could lead to the creation of tools which can contribute to the aforementioned areas.

6 CONCLUSIONS

The results indicate that the proposed approach is promising in achieving an acceptable result which can be used as an initialisation for medical experts in producing a textually labelled result following slight corrections and fine-tuning. Such a process still makes our system far more efficient than traditional anatomical labelling. We have also shown that the established α-tree segmentation algorithm performed well on a new challenging application of distinguishing different regions in photographic imagery of biological human specimens.

REFERENCES

[1] Dorin Comaniciu and Peter Meer. 2002. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 5 (2002), 603–619. https://doi.org/10.1109/34.1000236

[2] D. Ehrlich, T. Kemper, X. Blaes, and P. Soille. 2013. Extracting building stock information from optical satellite imagery for mapping earthquake exposure and its vulnerability. *Natural Hazards* 68, 1 (2013), 79–95. https://doi.org/10.1007/s11069-012-0482-0

[3] Lionel Gueguen, GK Ouzounis, M Pesaresi, and P Soille. 2012. Tree based representations for fast information mining from VHR images. In *Proceedings of the ESA-EUSC-JRC Eight Conference on Image Information Mining, Prof. Mihai Datcu, Ed.*

[4] Ming-Kuei Hu. 1962. Visual pattern recognition by moment invariants. *IRE transactions on information theory* 8, 2 (1962), 179–187.

[5] Zhi Liu, Liquan Shen, and Zhaoyang Zhang. 2011. Unsupervised image segmentation based on analysis of binary partition tree for salient object extraction. *Signal Processing* 91, 2 (2011), 290–299.

[6] Zhi Liu, Jie Yang, and Ning Song Peng. 2005. An efficient face segmentation algorithm based on binary partition tree. *Signal Processing: Image Communication* 20, 4 (2005), 295–314.

[7] Robert G Mayer. 1990. *Embalming: history, theory, and practice.* Appleton & Lange.

[8] T. McInerney and D. Terzopoulos. 1996. Deformable models in medical image analysis. In *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis.* IEEE, 171–180. https://doi.org/10.1109/MMBIA.1996.534069

[9] G Ouzounis, V Syrris, L Gueguen, and P Soille. 2012. The switchboard platform for interactive image information mining. In *Proceedings of 8th conference on image information mining, ESA-EUSC-JRC, Joint Research Centre of the European Commission.*

[10] Georgios K Ouzounis and Pierre Soille. 2012. The Alpha-Tree Algorithm. *JRC Scientific and Policy Report* (2012).

[11] Kilian M Pohl, John Fisher, Ron Kikinis, W Eric L Grimson, and William M Wells. 2005. Shape Based Segmentation of Anatomical Structures in Magnetic Resonance Images. In *Computer Vision for Biomedical Image Applications.* Springer, 489–498.

[12] Philippe Salembier and Luis Garrido. 2000. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE transactions on Image Processing* 9, 4 (2000), 561–576.

[13] P. Salembier, A. Oliveras, and L. Garrido. 1998. Anti-extensive connected operators for image and sequence processing. *IEEE Trans. Image Proc.* 7 (1998), 555–570.

[14] Pierre Soille. 2008. Constrained connectivity for hierarchical image decomposition and simplification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30, 7 (2008), 1132–1145. https://doi.org/10.1109/TPAMI.2007.70817

[15] Pierre Soille. 2011. Preventing chaining through transitions while favouring it within homogeneous regions. In *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing.* Springer, 96–107.

[16] C. Tomasi and R. Manduchi. 1998. Bilateral Filtering for Gray and Color Images. *International Conference on Computer Vision* (1998), 839–846. https://doi.org/10.1109/ICCV.1998.710815

[17] E. R. Urbach, N. J. Boersma, and M. H. F. Wilkinson. 2005. Vector-Attribute Filters. In *Mathematical Morphology: 40 Years On, Proc. Int. Symp. Math. Morphology (ISMM) 2005.* Paris, 95–104.

[18] Michel A. Westenberg, Jos B. T. M. Roerdink, and Michael H. F. Wilkinson. 2007. Volumetric attribute filtering and interactive visualization using the max-tree representation. *IEEE Transactions on Image Processing* 16, 12 (2007), 2943–2952.

[19] Stefan Zachow, Michael Zilske, and Hans-Christian Hege. 2007. *3D reconstruction of individual anatomy from medical image data: Segmentation and geometry processing.* ZIB.