Decomposition Algorithms for the Dissection of Tissue Samples

Maike Buchin
maike.buchin@rub.de
Department of Computer Science,
Ruhr University Bochum
Germany

Leonie Selbach
leonie.selbach@rub.de
Department of Computer Science and
Center for Protein Diagnostics,
Ruhr University Bochum
Germany

KEYWORDS

shape decomposition, skeletonization, laser capture microdissection

1 INTRODUCTION

Laser capture microdissection (LCM) is an approach to extract specific regions of interest from heterogeneous tissue samples. A laser separates the boundary of a selected region from the surrounding tissue and transports the fragment into a collecting device. The analysis of disease-specific regions collected with LCM provides more accurate characterizations of the disease – for example on a genomic, transcriptomic or proteomic level. Thus, LCM is an established method in clinical studies regarding biomarker discovery and molecular characterization of diseases [1, 2]. Our research addresses one central problem of processing samples with LCM.

Biological problem. With LCM being more and more commonly used, there is a need to automate the procedures. This includes the automatic selection of regions of interest (ROI) that should be dissected. Digital pathology approaches provide the identification of disease-specific regions using microscopical data. However, these regions oftentimes cannot be successfully dissected without any further processing. Usually, this makes extensive user-interaction necessary. The main cause of failed dissections lies in the size and shape of the fragments. Our contribution includes a novel approach for an additional image-based processing of tissue samples to automatically obtain regions that can be successfully dissected. This is done by translating this biological and technical problem into a mathematical and computational one.

Mathematical problem. Tissue samples that are processed with LCM are thin sections and thus can be interpreted as 2D images. The identification of ROIs results in a binary mask. After a preprocessing step, the ROIs are given as a number of connected components without holes, which can be interpreted as simple polygons. Thus, the given problem can be modeled as a constrained *polygon decomposition problem*: Given a simple polygon P, find an optimal decomposition of P such that all subpolygons are feasible with respect to some additional constraint. For our application, the feasibility constraint and the optimization goal have to be chosen in such a way that the tissue fragments resulting from the decomposition have the correct morphology for LCM. Hence, it is necessary to find suitable mathematical definitions for both.

2 METHODS

The underlying biological problem inspired our research into different constrained decomposition problems. Our results include the development of (efficient) algorithms as well as proofs that some problems are NP-hard to solve. Here, we focus on one algorithmic approach [4] for the decomposition of tissue samples, which we implemented and evaluated in practice.

The proposed method computes an optimal decomposition of a polygon in two steps. The first step is to compute a skeleton of the polygon. Skeletons are frequently used in image processing applications [3]. They are given as a medial line that lies inside the shape and which represents the main morphological features. Because cancerous tissue regions often present themselves as highly complex and ramified shapes, we utilize the skeleton to take their morphology into account for the decomposition.

The second step consists of the skeleton-based decomposition algorithm. Here, only certain cut edges based on the skeleton are allowed. This together with the fact that subpolygons belonging to different skeleton branches are decomposed individually results in an efficient algorithm, which follows a dynamic programming approach. The basic idea is to check for each skeleton point whether a feasible decomposition of the corresponding subpolygon exists. If more than one possible decomposition exists, an optimal one is chosen with regard to a certain optimization goal.

This main algorithmic construct can be adapted for various applications by defining different feasibility criteria or optimization goals. For the application with LCM, the size and shape of the subpolygons has to be constrained. The size constraint is applied as a feasibility criterion. Namely, a polygon is *feasible* if its area is in between a lower and upper area bound. Regarding the shape constraint, it was observed that regular or round shapes (of the correct size) are more likely to be successfully dissected with LCM. On the other hand, a tissue fragment often tears if the shape is too irregular or presents very narrow regions. We considered different computational criteria to translate these observations into our algorithm. Eventually, we defined a criterion based on fatness of a polygon. A polygon is called α-fat if its aspect ratio is at most α. The aspect ratio (AR) of a polygon is defined as the ratio between the diameters of its minimum circumscribed circle and its maximum inscribed circle. This can be used as a roundness measure: The closer this value is to 1, the rounder the polygon is. We include the fatness into the algorithm the optimization goal. Thus, a decomposition is *optimal* if the fatness, i.e. the AR, of the subpolygons is maximized. We denote the proposed method by *MaxFat*.

womENcourage '21, 22-24 September, 2021, Prague, Czech Republic
2021.

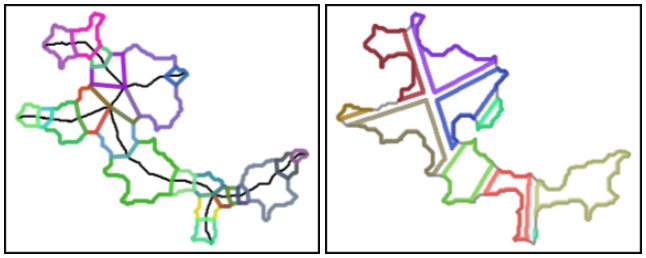

Figure 1: Comparison polygon decompositions computed with MaxFat (left) and BiSect (right)

3 EXPERIMENTAL RESULTS

The *MaxFat* method presented in Section 2 was implemented and evaluated in practice. So far there has been just limited research on shape decomposition methods for LCM. Therefore, we assessed our results in comparison to one heuristic decomposition method that was used in previous studies. This method, which we denote by *BiSect*, follows an iterative bisection approach and applies only a size constraint but no shape criterion or optimization goals.

The experiment was conducted on ten FFPE (formalin-fixed paraffin-embedded) lung tissue samples from patients with non-small-cell lung carcinoma. Using pixel spectra from infrared microscopic images of the samples, the different tissue types were identified and the tumor class was chosen as the region of interest (ROI). After a preprocessing procedure, the input consisted of 441 individual ROI polygons, which were decomposed with both methods. The resulting fragments were extracted using LCM and the number of successfully dissected fragments was documented.

As depicted in Figure 1, MaxFat predominantly produces rounder shapes, whereas fragments computed with BiSect are often irregular and have narrow regions. The experiments confirm that the number of successfully collected fragments is higher with MaxFat than with BiSect. On average, for MaxFat 95.44 % of fragments were labeled successful and only 80.98 % for BiSect. However, the critical factor to determine the quality of the decomposition methods is the tissue yield, i.e. the amount of collected tissue area.

For both methods, the loss of tissue material appears at two places. On the one hand, the total area of the fragments does not equal the total area of input polygons. This is due to the nature of the algorithms. For example, a feasible decomposition of a polygon might not exist with the constraints given by the MaxFat approach. On the other hand, tissue loss occurs whenever a fragment contained in the computed decompositions could not be successfully collected with LCM. We define the overall success rate of the methods as the percentage of successfully collected tissue area in relation to the original input area. For example, if the computed decomposition had a area loss of 8.58 % and of those remaining 91.42 % then 10 % could not be collected with LCM, the resulting overall success rate amounts to 82.28 %. Figure 2 shows the amount of tissue loss and overall success rates of both methods. Over all ten samples, the tissue yield of MaxFat was at least 10 % and up to 30 % higher. Thus, the proposed method considerably increases the amount of collected protein and DNA material for further analysis.

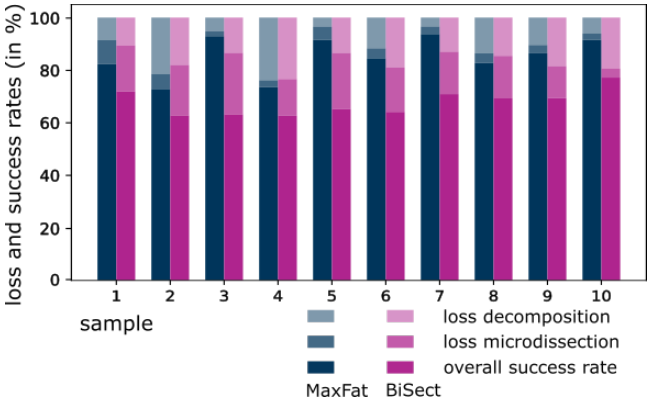

Figure 2: Comparison of the amount of tissue loss and the overall success rates for MaxFat and BiSect. On average, the results for the decomposition loss (LD), the microdissection loss (LM) and the success rate (SR) are as follows. MaxFat: 10.77% (LD), 4.54% (LM) and 85.21% (SR). BiSect: 16.35% (LD), 18.38% (LM) and 68.21% (SR).

4 CONCLUSION AND FUTURE WORK

We developed a polygon decomposition algorithm as a novel approach to preprocess tissue samples for LCM. We translated technical constraints into image processing and mathematical computations to improve tissue yield. As far as we know, this is the first method for this application, which was built on theoretical grounds and validated in practice. With this, our work contributes to further optimization and automation of LCM and thus enhances its suitability for larger scale clinical studies. Note that the quality of the decompositions obtained by the presented method is influenced by the skeleton. Hence, recent improvements for skeletonization might further improve our results.

Moreover, our method not only applies to LCM, but can be adjusted for other applications by including different constraints into the framework. Note that the runtime of the algorithm depends on the choice of criteria. Therefore, our research includes the analysis of other algorithmic approaches.

REFERENCES

[1] Soma Datta, Lavina Malhotra, Ryan Dickerson, Scott Chaffee, Chandan K Sen, and Sashwati Roy. 2015. Laser capture microdissection: Big data from small samples. *Histology and Histopathology* 30, 11 (2015), 1255.

[2] Frederik Großerueschkamp, Thilo Bracht, Hanna C Diehl, Claus Kuepper, Maike Ahrens, Angela Kallenbach-Thieltges, Axel Mosig, Martin Eisenacher, Katrin Marcus, Thomas Behrens, et al. 2017. Spatial and molecular resolution of diffuse malignant mesothelioma heterogeneity by integrating label-free FTIR imaging, laser capture microdissection and proteomics. *Scientific Reports* 7, 1 (2017), 1–12.

[3] Punam K Saha, Gunilla Borgefors, and Gabriella Sanniti di Baja. 2016. A survey on skeletonization algorithms and their applications. *Pattern Recognition Letters* 76 (2016), 3–12.

[4] Leonie Selbach, Tobias Kowalski, Klaus Gerwert, Maike Buchin, and Axel Mosig. 2020. Shape Decomposition Algorithms for Laser Capture Microdissection. In *20th International Workshop on Algorithms in Bioinformatics (WABI 2020) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 172).* 13:1–13:17.