A causal extraction scheme in top-down pyramids for large images segmentation

Romain Goffe[1], Guillaume Damiand[2] and Luc Brun[3]

[1] SIC-XLIM, Université de Poitiers,
CNRS, UMR6172, F-86962, Futuroscope Chasseneuil, France
goffe@sic.univ-poitiers.fr

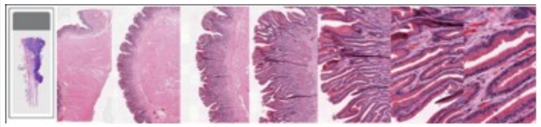
[2] LIRIS, Université de Lyon,
CNRS, UMR5205, F-69622, Villeurbanne, France
guillaume.damiand@liris.cnrs.fr

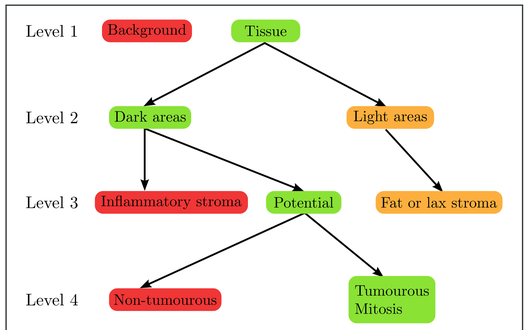
[3] GREYC, ENSICAEN,
CNRS, UMR6072, F-14050, Caen, France
luc.brun@greyc.ensicaen.fr

Acknowledgments

This research is part of the FoGrImMi project, supported by the ANR foundation under grant ANR-06-MDCA-008-01/FOGRIMMI.

Breast Cancer Histological Context





Simplified scheme of pathologists' analysis

Main identified requirements

Topological criteria

The evaluation is based on adjacency and inclusion criteria between the objects identified in the image. For example, a tumourous area is a region that includes strictly a high number of mitosis.

Large images: up to $32\,000^2$ pixels

The size of the images produced by Aperio scanners appears to be the main technical constraint. Existing models are not designed to handle such an amount of data.

Hierarchical analysis

Images are produced at different resolutions that fit the pathologists' analysis. For each resolution, interesting areas are sorted out and then further explored at higher resolution. This analysis involves a top-down construction scheme.

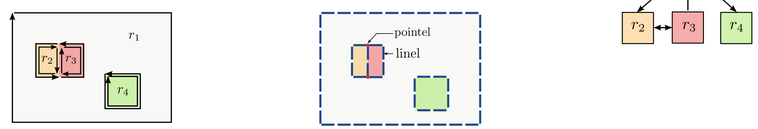
Our objectives

Propose a workflow for automatic analysis:

- definition of a model
- top-down construction scheme
- application to segmentation

Step 1: Model Definition

Topological Maps



Combinatorial maps:

- Finite set of darts (half-edges)
- β_1 permutation: turn around a face
- β_2 involution: get opposite face

⇒ Describe multi-adjacency relationships

Geometrical matrix:

- based on interpixel elements (pointels, linels, pixels)
- represent regions boundaries
- association with darts (embedding)

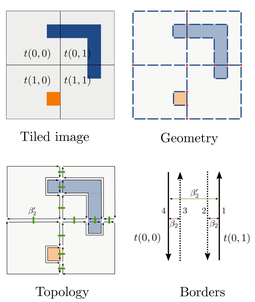
⇒ Describe geometrical information

Tree of regions:

- parent/child relationships
- union-find structure

⇒ Describe inclusion relationships

Tiled maps



Topological tiles:

- geometrical subdivision
- new β'_2 operator on borders

entails →

Fictive borders:

- mark on linels that belong to tiles' borders
- use segmentation criterion to determine real/fictive linels

solution ↓

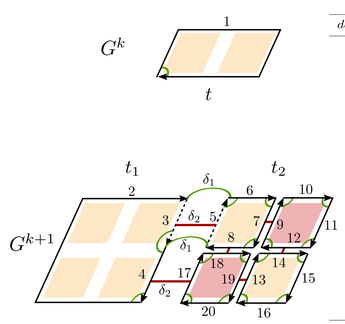
Redefine β_i to abstract the subdivision:

- δ_2 involution: if β'_2 is defined (tiles' borders) then $\delta_2 = \beta'_2$, otherwise $\delta_2 = \beta_2$
- δ_1 permutation: for each dart d; $\delta_1(d) = \beta_1((\beta'_2 \circ \beta_1)^n(d))$ with $n = min\{p \in N \mid ligne(\beta_1((\beta'_2 \circ \beta_1)^p(d))$ is real$\}$

Tiled top-down pyramids

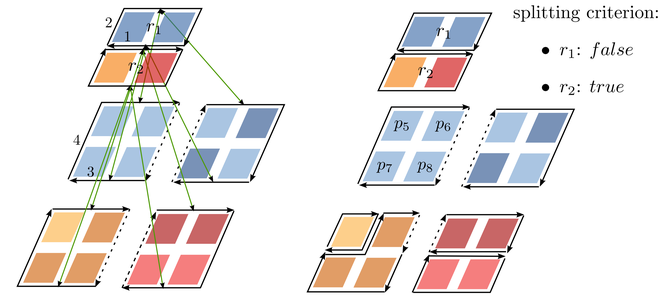
Features:

- stack of tiled maps
- up/down relationships between darts, regions and tiles
- splitting/merging operations on regions
- retrieval operations for edges and faces between different levels
- swap/load operations on tiles



dart	δ_1	δ_2	up	down
1	1	∞	-	2
2	6	∞	1	-
3	4	5	-	-
4	2	17	-	-
5	6	3	-	-
6	7	∞	1	-
7	8	9	-	-
8	4	18	-	-
9	10	7	-	-
10	11	∞	1	-
11	12	∞	1	-
12	9	14	-	-
13	14	19	-	-
14	15	12	-	-
15	16	∞	1	-
16	13	∞	1	-
17	18	4	-	-
18	19	8	-	-
19	20	13	-	-
20	17	∞	1	-

Step 2: Construction from large images



splitting criterion:

- r_1: *false*
- r_2: *true*

Constraints:

- Causality: ensure up/down relationships
- Top-down: only refine areas of interest

Structures:

- *Embedding structure*: determines if a boundary exists in upper levels
- *Regions structure*: gives segmentation criteria of corresponding up region when processing a pixel

Global algorithm:

- traverse pixels from top/left to bottom/right
- make a new region for current pixel
- compare with left and top neighbors:
  - merge with neighbors
  - start a new region

Regions comparison:

- *top-down*: merge automatically if splitting criterion is false
- *causality*: forbid merging if parent regions are different
- *segmentation*: involve criteria for remaining cases

Causal structure

- topological representation
- consistency of hierarchy (no loss of information)
- adjacency and inclusion relationships

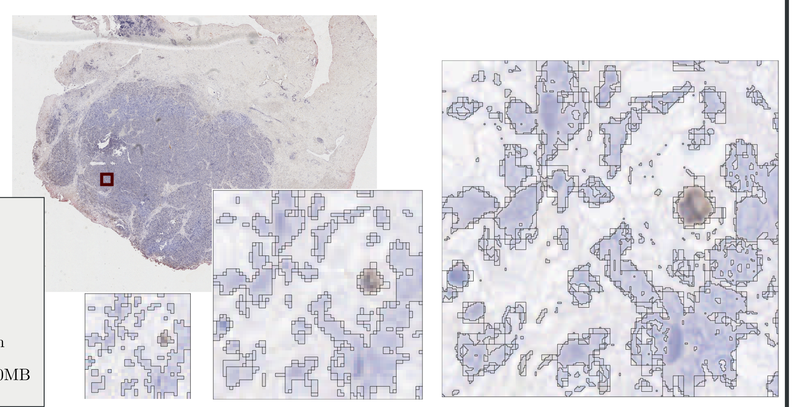
Constant tile size

- user-defined rectangular tiles
- swap/load operations for memory bounding
- ensures memory feasiblity

Top-down scheme

- focus of attention over selected regions
- reduces memory usage
- generic model for various segmentation criteria

Step 3: Application to Segmentation

Features:

- preliminary results
- combination of criteria
- main segmentation with k-means algorithm
- required memory < 150MB