MULTI-SCALE COMPONENT-TREES FOR ENHANCED REPRESENTATION IN MULTIPLEX IMMUNOHISTOCHEMISTRY IMAGING

1238

Romain PERRIN [1], Aurélie LEBORGNE [1], Nicolas PASSAT [2], Benoît NAEGEL [1], Cédric WEMMERT [1]

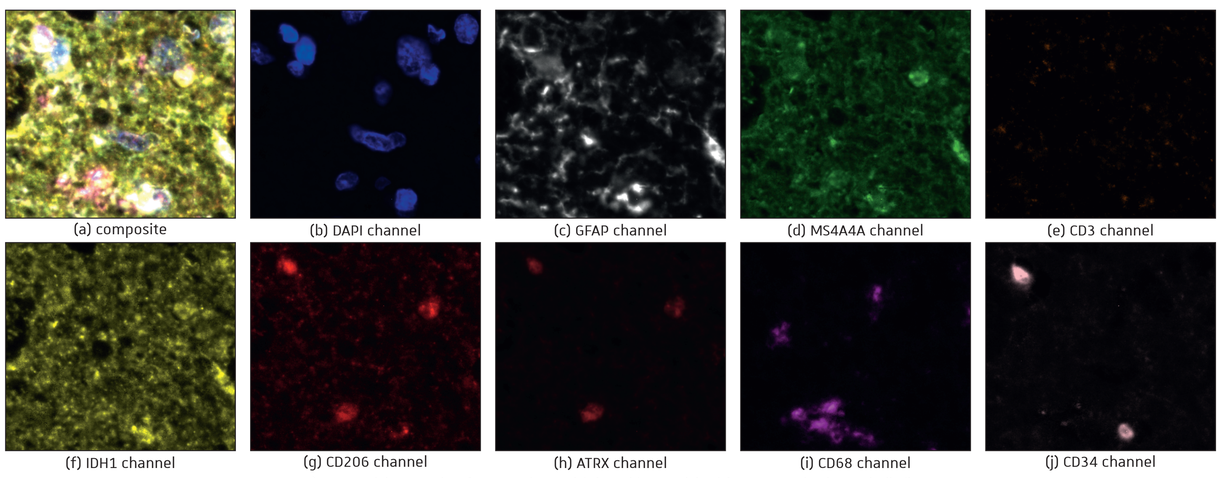
(1) ICube UMR 7357, University of Strasbourg, Strasbourg, France
(2) CReSTIC UR 3804, University of Reims Champagne-Ardenne, Reims, France

21st International Symposium on Biomedical Imaging (ISBI 2024)
Megaron Athens International Conference Center (MAICC)

Context

Multiplex immunohistochemistry imaging is a collection of innovative techniques allowing **simultanate staining** with multiple biomarkers on a same tissue. Analyzing these complex images remains a challenge even for machine learning methods. It could be beneficial to use an intermediary data structure to efficicntly represent these images. The Multi-Scale Component-Tree (MSCT) is a **hierarchical morphological data structure** allowing for a **efficient storage** of images across multiple scales. We illustrate the possibilities offered by this new structure on glioblastoma images, taking inspiration from the way pathologists manually process such data.

Multiplexed images

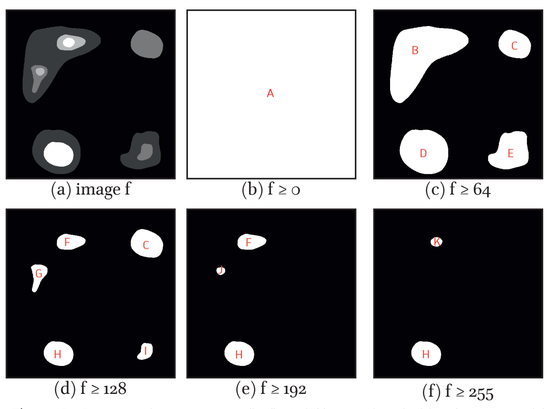


(a) composite (b) DAPI channel (c) GFAP channel (d) MS4A4A channel (e) CD3 channel
(f) IDH1 channel (g) CD206 channel (h) ATRX channel (i) CD68 channel (j) CD34 channel

Figure 1 : (a) Example of composite multiplexed image (a) with 9 semantic channels (b-j).

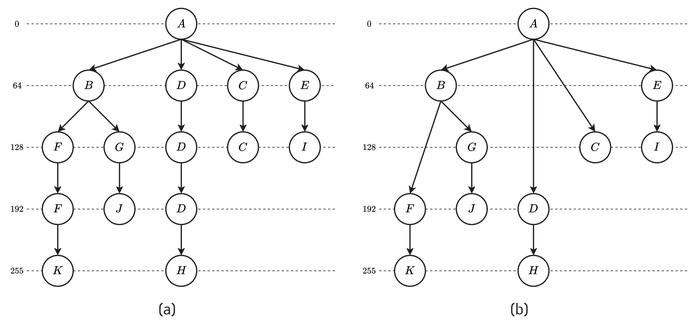
Images used in this paper come from a collaborative project around **glioblastoma**, a malignant cerebral tumor with poor prognosis, with neuropathologists of the Hannover Medical School (Germany). 62 images from 22 patients have been collected (3 µm tissue sections, x20 magnification 0.49 µm/pixel resolution). Regions ranging from the tumor center up to the sane tissue have been manually chosen and **multispectrally stained** with a resolution of 0.25 µm/pixel (Figure 1).

Component-tree



(a) image f (b) f ≥ 0 (c) f ≥ 64
(d) f ≥ 128 (e) f ≥ 192 (f) f ≥ 255

Figure 2 : Connected components (b-f) at different thresholds of an image (a).



(a) (b)

Figure 3 : Example of component-trees representing the image f (Figure 2.a).

The **component-tree** (or max-tree) [1] is a **hierarchical morphological graph-based model** offering an efficient encoding of images for a low algorithmic computation cost [2]. A component-tree encodes the **inclusion** relationships between **connected components** obtained from successive thresholdings (Figure 2.b-f) of a grayscale image. The tree root is the largest connected component containing the entire image (A) and its leaves are the connected components whose gray-level values correspond to the local maxima of the image (C, H, I, J et K). The resulting graph is a tree called the component-tree (Figure 2).

Multi-Scale Component-Tree (MSCT)

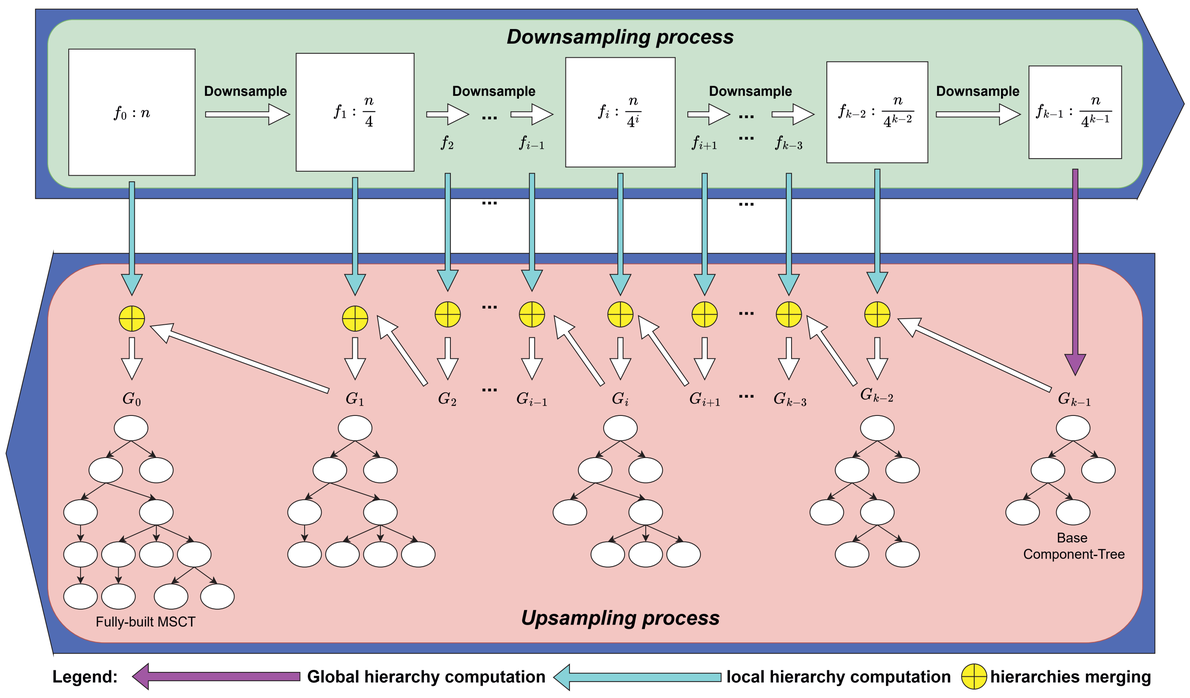
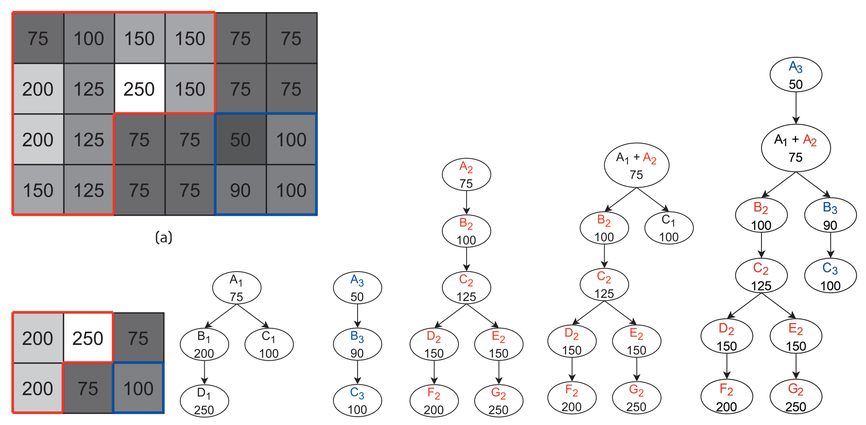
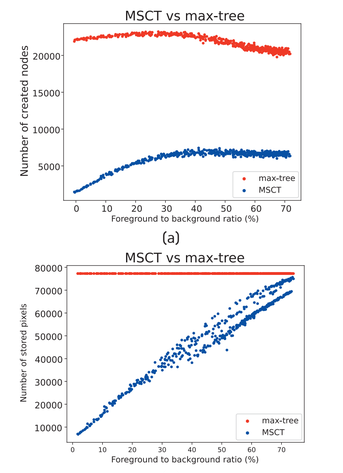


Figure 4 : Construction principle of a Multi-Scale Component-Tree involving two steps (downsampling then upsampling) [5].



(a) (b) (c) (d) (e) (f) (g)

Figure 5 : Example of Multi-Scale Component-Tree construction (g) on a simple image using two scales (a-b). (g) is the result of mergingthe MSCT (c) with partial max-trees at higher scale (d-e).



(a) (b)

Figure 6 : Storage performances.

The **Multi-Scale Component-Tree** (MSCT) [5] is a multi-scale extension of the concept of component-tree [1] where nodes may contain flat zones composed of pixels at multiple scales. The MSCT is built from a set of **downsampled images**. A first component-tree is computed on the smallest downsampled image. Iteratively, nodes of the tree are selected according to a criterion based on the **Maximally Stable Extremal Regions** (MSER) [3]. Partial component-trees are then computed on the regions of these selected nodes **projected on a higher scale image**, then merged on the MSCT, replacing the old nodes (Figure 5). This process is repeated k-1 times until the original image scale is reached (Figure 4). Thanks to its multi-scale capabilities, the MSCT is able to generate a significantly lower amount of nodes compared to a regular component-tree (Figure 6.a) and produces an **efficient image representation and storage** of its pixels (Figure 6.b).

Feature vectors

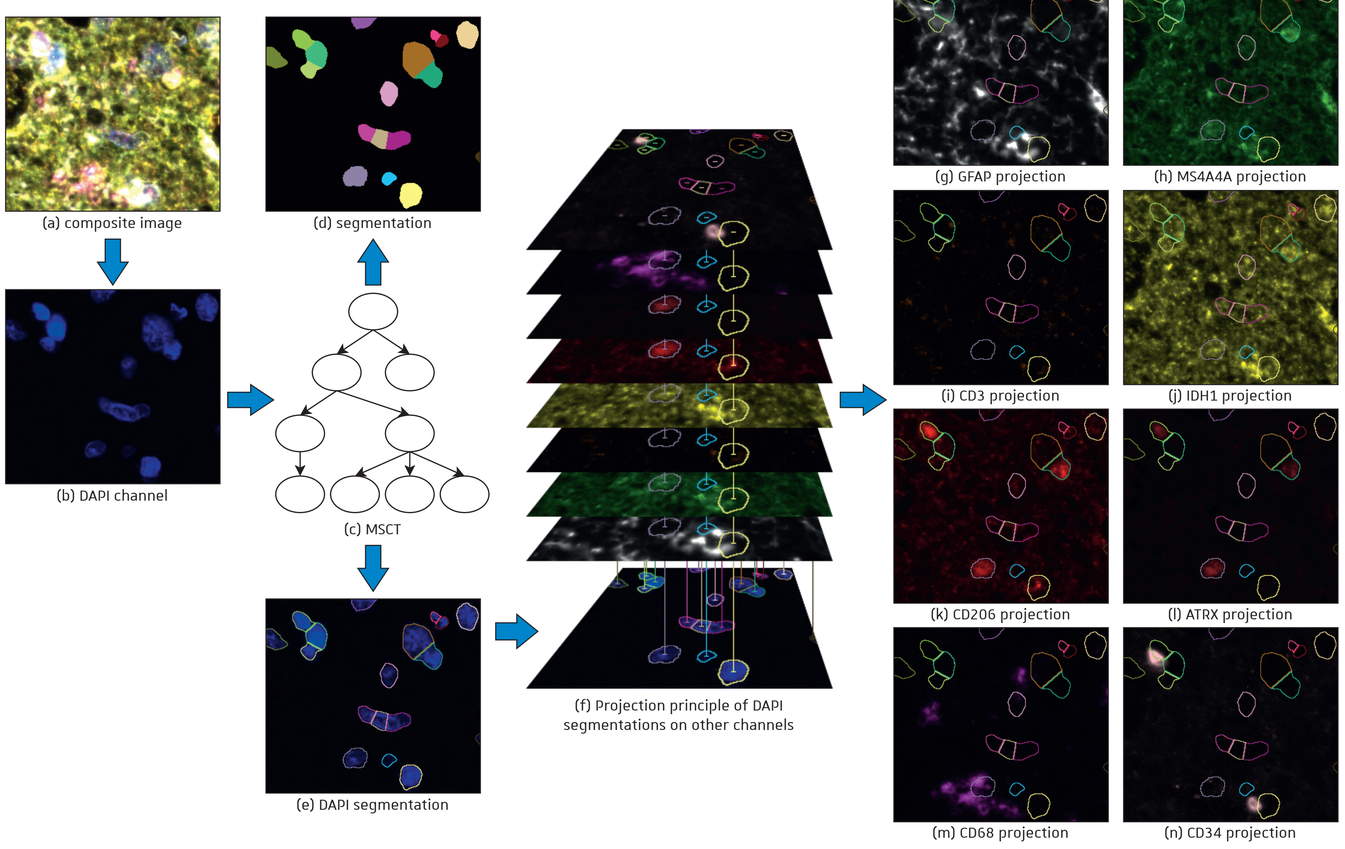
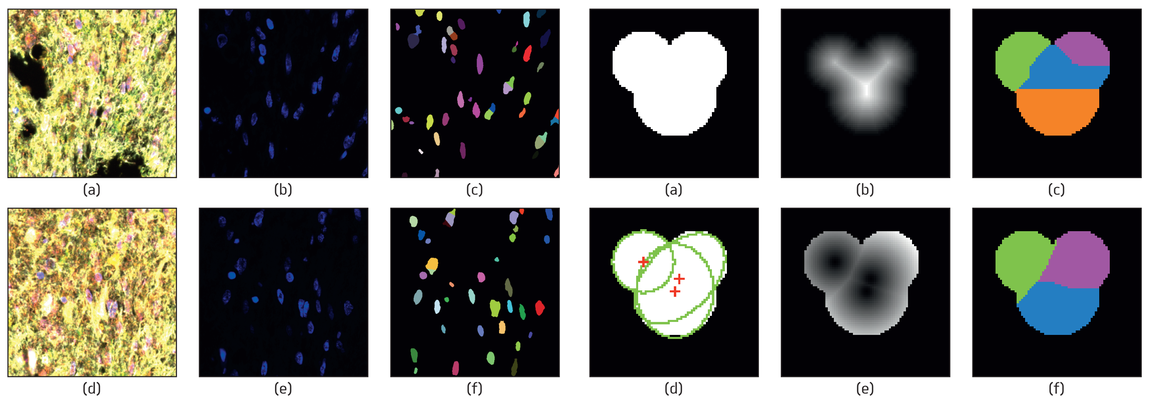


Figure 7 : Construction principle of a feature vector from a multiplexed image (a) by extracting cell segmentations (d) and projecting them on other channel (e-n) aggregating pixel values.

One MSCT is built for each multiplexed image (Figure 7.a-c). The connected components containing **nuclei** are extracted from the tree by **filtering** its nodes (Figure 7.d-e). The resulting segmented objects form **masks** that are **projected** on the remaining semantic channels (Figure 7.f). For each connected component n_i, on each channel f^j other than the DAPI channel (f^0), the values of pixels $f^j(p)$ are summed. The final **feature vector** c_i of an underlying object n_i is composed ot the k-1 sums of **physical measures** associated with biomarkers from channels 1 to k-1 inside its segmented mask (Figure 7.g-n). Formally, $c_i = \left[\sum_p f^j(p) \mid p \in n_i\right]_{j=1}^{k}$

Results



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

Figure 8 : MSCT segmentation results (c,f,i,l) on DAPI channels (b,e,h,k) of the composite multiplexed images (a,d,g,j).

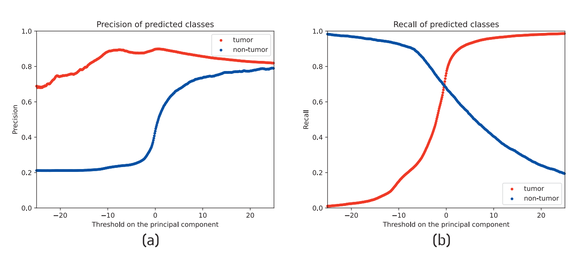
(a) (b) (c) (d) (e) (f)

Figure 9 : Segmentation + watershed (c) and DTECMA + watershed (f) of (a).

The connected components extracted from the MSCT may contain multiple **overlapping objects** (Figure 8). In order to identify single objects (nuclei), an **ellipse fitting model DTECMA** [4] is applied. Using the computed ellipse centers, a **Euclidean distance transform** (distance to the closest center) is computed and used to improve the quality of the final segmentation obtained though **watershed** (Figure 9).
62 multiplexed images of 1396 x 1860 x 9 pixels have been processed. 24 296 **expert-labeled dot annotations** have been used to evaluate the quality of the segmentations obtained from the MSCT.
On these 62 images, 18 494 nuclei were **true positives**, 6 457 **false positives** et 5 802 **false negatives**, for a total **precision** of **p** = 0.741 and a **recall** of **r** = 0.761 [6].

Perspectives

The feature vectors computed from the MSCT can be used for **unsupervised classification tasks**.
A multiplexed image with k channels produces a feature vector of k-1 dimensions. A **PCA reduction** step may be applied to project these vectors in 2D. A threshold may then be chosen to create a **binary classifier** with two classes : immune cells and tumor cells (Figure 10). The expressive power of these specific vectors was insufficient to create a complete classifier with all sub-types of cells (lymphocytes, macrophages...). Complementary **attributes** may be introduced to augment the expressive power of these vectors (compacity, circularity, optical density...).



(a) (b)

Figure 10 : Precision (a) and recall (b) depending on the chosen PCA threshold.

References

[1] P. Salembier, A. Oliveras, L. Garrido, Anti-extensive connected operators for image and sequence processing, IEEE Transactions on Image Processing, vol. 7, pp. 555-570, 1998.
[2] E. Carlinet, T. Géraud, A comparative review of component tree computation algorithms, IEEE Transactions on Image Processing, vol. 23, pp. 3885-3895, 2014.
[3] J. Matas, O. Chum, M. Urban, T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, Image and Vision Computing, vol. 22, pp. 761-767, 2004.
[4] T. Zou, T. Pan, M. Taylor, H. Stern, Recognition of overlapping elliptical objects in a binary image, Pattern Analysis and Applications, vol. 24, pp. 1193-1206, 2021.
[5] R. Perrin, A. Leborgne, N. Passat, B. Naegel, C. Wemmert, Multi-Scale Component-Tree: An Hierarchical Representation of Sparse Objects, IAPR Third International Conference on Discrete Geometry and Mathematical Morphology (DGMM), Florence, Italy, april 2024.
[6] R. Perrin, A. Leborgne, N. Passat, B. Naegel, C. Wemmert, Multi-Scale Component Trees for Enhanced Representation in Multiplex Immunohistochemistry Imaging, IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece, may 2024.