



HAL
open science

Reasoning with spatial relations over high-content images

Nicolas Lomenie

► **To cite this version:**

Nicolas Lomenie. Reasoning with spatial relations over high-content images. IEEE World Congress on Computational Intelligence, WCCI'2010, hosting the International Joint Conference on Neural Networks, IJCNN'2010, Jul 2010, barcelone, Spain. hal-00497810

HAL Id: hal-00497810

<https://hal.science/hal-00497810>

Submitted on 6 Jul 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reasoning with spatial relations over high-content images

Nicolas Loménie

Abstract—Spatial relation and configuration modeling issues are gaining momentum in image analysis and pattern recognition fields in the perspective of mining high-content images or large scale image databases in a more expressive way than purely statistically. Continuing our previous efforts whereby we developed specific efficient morphological tools performing on mesh representation like Delaunay triangulations, we propose to formalize spatial relation modeling techniques dedicated to unorganized point sets. We provide an original mesh lattice framework more convenient for structural representations of large amount of image data by the means of interest points sets and their morphological analysis. The set of designed numerical operators is based on a specific dilation operator making it possible the representation of concepts like “between” or “left of” over sparse representations such as graphs. Then, for the sake of illustration and discussion, we apply these new tools to high-level queries in microscopic histo-pathological images and structural analysis of macroscopic images.

I. INTRODUCTION

Visual cognition or cognitive vision is a long-term goal aiming at putting together computer vision analysis tools and artificial intelligence reasoning methodologies in order to design robust systems for smart vision-based interfaces. As a first step towards this ambitious integrated research purpose, image understanding modules providing high-level image descriptions is a cornerstone, at least for the final user to seamlessly interact with visual data.

The major issue can be referred as the semantic gap filling that is the coincidence between low-level features given by the signal and the high-level concepts used to describe and handle images at a semantic level. Few works [1], [2] have dealt with such an issue in a sound theoretical way. For instance, in the field of spatial relation modeling, [7], [4] made a great extent of work for radiometric image modeling defined over regular grids, that is the mainstream way in which visual data are handled. In this work, we made an attempt to extend these seminal ideas to visual data represented over sparse representations such as interest point sets. In particular, Delaunay triangulation representations are favored for the neighborhood embedding of these visual data defined over irregular grids (see Fig. 1).

These mathematically well defined tools make it possible to consider enhanced visual systems with enabled reasoning capabilities. Not only they provide relevant algorithms to explore large images for instance in a more formal way as a robot would be expected to do by anchoring spatial relation concepts into the signal representation either as an image or a point set visual data. In addition, they open new opportunities

Nicolas Loménie is with the Department of Computer Sciences, University of Paris Descartes and with CNRS - IPAL, I2R, National University of Singapore; email: nicolas.lomenie@mi.parisdescartes.fr

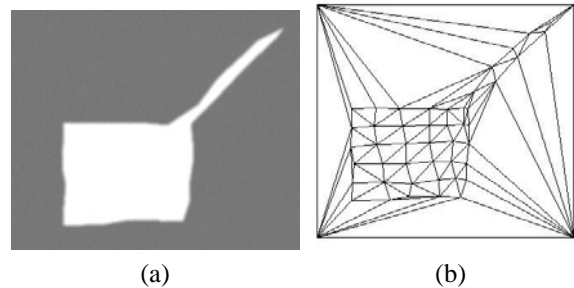


Fig. 1. Different representations of the same object: (a) gray intensity image I and (b) geometric mesh M

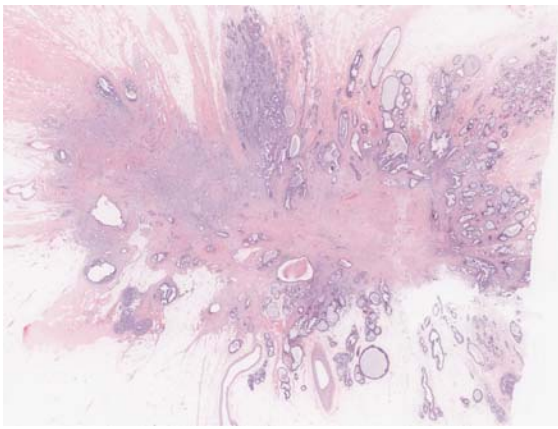
to perform high-level queries over large image databases featuring specific spatial organization requirements.

Last, coming with the challenging great amount of visual data to process due to dramatic progress in both resolution and speed of the acquisition devices, satellite image as well as medical image end-users urge for more interactive, qualitative ways of interaction with the visual data they use to assess or diagnose daily critical situations. Even at the level of more long-term biological research, smart interactions with visual data such as microscopic information turn to become practical requirements due to the high-content - partly related to the size - of these images (see below Fig. 2 and Table I for an illustration of the amount of data for a single patient biopsy in histopathological imaging). Of course, satellite imaging with new resolution up to 0.7 cm a pixel over 24000x24000 pixel size images will have to face the same kind of issues in the near future.

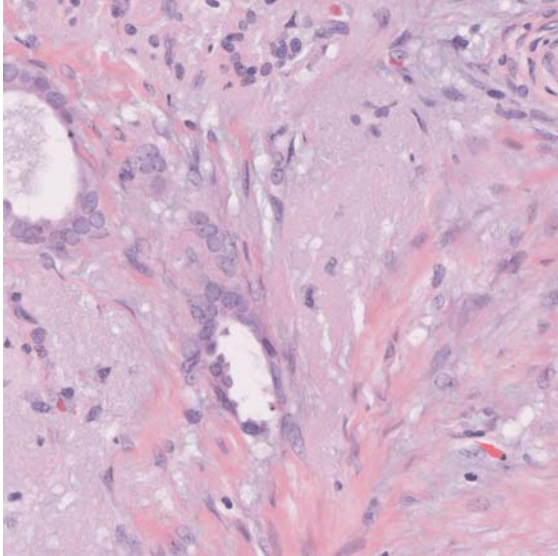
TABLE I
HISTOPATHOLOGICAL IMAGE DATA

	Size (pixels)	Size (bytes)
x1	1018x768	3.05 Mo
x10	3664x2763	39.54 Mo
x20	14657x11054	632.48 Mo
x40	58630x44216	7.77 Go

As image size scales up exponentially according to the progress in the field of optics, interactive time applications can not use mainstream image analysis algorithms for computational complexity issues. Working on sparse representations such as interest point sets can be a solution to this kind of real time issues by focusing on informative parts on the image. For instance, working on points of interest in the image of Fig. 2(a) scales down the size of visual data from 7.77 Go up to 1.5 Go. In that perspective, the adaptation of existing techniques to point set representations and so on to graph related representation should be a prospective field of research [5], [10]. Section II presents the theoretical



(a)



(b)

Fig. 2. (a) Whole Slide Image at resolution x1 (b) Sample image at resolution x40 out of over 2000 tiling the histopathological image

framework of the proposed operators including mathematical morphology and its adaptation to graph-based representations. Section III illustrates the relevance of these tools for image understanding and visual cognition modeling. Last, Section IV draws research perspectives both from practical and theoretical standpoints.

II. MORPHOLOGICAL TOOLBOX

Image analysis and computer vision in general are still considered as brittle technologies [6]. However, impressive achievements have been reached so far with a small amount of commercial product providing robust interfaces. Even the recently released Google Goggle tool performs quite well as long as the end-user does not go beyond strict limitations of usage such as natural object recognition. This is partly due to these tools relying on statistical representations and sub-optimal optimization processes. Yet in a non-closed world such as the Internet, such a framework needs endless and continuous learning procedures to proceed in a steady way. At the opposite, structural approaches are less dependent to

the closed-world assumption to provide stable and robust tools to interact with visual data. In particular, mathematical morphology is one of the only sound original theory that emerged from the image analysis community. As so, it constitutes a natural framework to put down the basis of a visual cognitive prospective theory.

A. Mathematical morphology

Mathematical morphology relies on the definition of relevant lattice structures with an order relation while statistical approaches mainly remain on vectorial space structures. By nature, its operators perform highly non linear type of filtering keen at delivering structural and morphological information and analysis about the visual entities in the images. Theoretically, the so-called structuring element is at the core of these flexible structural analysis capabilities even though in practice most of the processing lines use symmetric centered structuring elements such as open ball of a fixed radius.

Besides this, once a basic adjunction operator such as an erosion is defined over the specific lattice structure equipped with an involution operator, a raft of sophisticated mathematically sound operators can be easily built upon by composition within what is known as the pyramid of morphological operators. From that perspective, the usage of such a toolbox can be tailored to adapt to specific situations without a lot of additional theoretical developments. Once again, that makes mathematical morphology a suitable theory to lay down first fundamental steps towards a cognitive vision theory.

B. Adaptation to point sets

Let us consider now the case of an unorganized point set S in \mathbb{R}^2 . Let $del(S)$ be the Delaunay triangulation of this point set defining an embedded neighborhood structure.

Lattice structures.

$\mathcal{M}(S)$ is the set of meshes on the Delaunay triangulation of S $del(S)$, i.e., the set of mappings from the triangles T in $del(S)$ to ϕ_T values. T stands for any triangle in del . A mesh $M \in \mathcal{M}(S)$ is defined by $\{(T, \phi)\}_{T \in del(S)}$ or equivalently by a mapping $\phi : T \in del(S) \rightarrow [0, 1]$.

We define one complete lattice structure for a point set including an order relation, within the functional theory frame, called $\mathcal{L} = (\mathcal{M}(S), \leq)$, where the partial ordering \leq is defined by: $\forall M_1$ and $M_2 \in \mathcal{M}(S), M_1 \leq M_2 \iff \forall T \in del(S), \phi_T^1 \leq \phi_T^2$

Point set and mesh binarization. Let us also define the binarization operator bin_θ as follows :

$$\forall M \in \mathcal{M}(S),$$

$$bin_\theta(M) = \{T \in del(S), 1 \text{ if } \phi_T > \theta, 0 \text{ elsewhere}\} \quad (1)$$

Point set and mesh morphological operators.

The rank filtering operators need the computation of two additional values e_T and d_T for each triangle T , defined by:

$$\begin{aligned} e_T &= \min\{\phi_{T'} | T' \in \nu(T)\} \\ d_T &= \max\{\phi_{T'} | T' \in \nu(T)\} \end{aligned} \quad (2)$$

The neighborhood $\nu(T)$, playing the key role of a structuring entity, is the set of all triangles T' of $del(S)$ sharing at least one vertex with the triangle T , that is:

$$\nu(T) = \{T' \in Del | T' \cap T \neq \emptyset\} \quad (3)$$

Then, in the lattice framework, we define two operators $e(M)$ and $d(M)$ on the complete lattice \mathcal{L} by :

$$\forall M \in \mathcal{M}(S), \quad e(M) = \{T \in Del, e_T\} \text{ and } d(M) = \{T \in Del, d_T\} \quad (4)$$

with e_T and d_T defined in Eq. 2.

We proved in [3] that $e(M)$ and $d(M)$ are respectively erosion and dilation morphological operators on \mathcal{L} . We can define various operators now like opening $o(M)$ and closing $c(M)$:

$$\forall M \in \mathcal{M}(S), \quad o(M) = d \circ e(M) \text{ and } c(M) = e \circ d(M) \quad (5)$$

and benefits of all the inherited operators based on the erosion and the involution operators c :

$$\forall M \in \mathcal{M}(Del(S)), M^c = \{T \in Del, 1 - \phi_T\} \quad (6)$$

Thus, as expected,

$$\forall M \in \mathcal{M}(Del(S)), e(M) = d(M^c)^c \quad (7)$$

C. Modeling spatial relations

1) *The directional dilation*: The whole set of spatial relation concepts modeling is based on the directional dilation operator. Let α be the directional angle with regard to the horizontal axis of the representation plane. For any mesh $M \in \mathcal{L}$, the definition of the neighborhood $\nu(T)$ of a mesh triangle T acts as a directional structural element of direction α :

$$\nu_\alpha(T) = \{T' \in Del | T' \cap T \neq \emptyset \text{ and } \angle(T, T') < \alpha + \epsilon\} \quad (8)$$

with

$$\angle(T, T') = \angle(\overrightarrow{B_T B_{T'}}, (\overrightarrow{O \vec{x}})) \quad (9)$$

where B_T is the barycenter of the triangle T .

Eq. 8 defines an anisotropic neighborhood referred as ν_α , to be put in parallel with the isotropic neighborhood definition ν_{iso} of equation 3, where α stand for the angle with the horizontal axis $(\overrightarrow{O \vec{x}})$ and ϵ is the width or tolerance angle of the structuring element.

We can define the directional dilation as:

$$\forall M \in \mathcal{M}(S), d_{\alpha, \epsilon}(M) = d_{\nu_\alpha}(M) \quad (10)$$

where $d_{neighborhood}$ stands for the computation of d_T values over a specific system of *neighborhood*. When ϵ is not written, $\epsilon = \pi/8.0$ and $d_{\alpha, \epsilon} = d_\alpha$.

Qualitative concepts like spatial relations depend both on the context and on the reference object. A fuzzy version of for all the morphological operators acting on meshes following the algorithmic procedure in Alg. 1 can deliver more flexible results for reasoning purposes. From this fundamental spatial relation, by composition, we can define a bunch of qualitative spatial relation in a quantitative way such as “between” as briefly illustrated in the following section.

Algorithm 1 Fuzzy Dilation d^f

INPUT: a mesh M defined over the lattice \mathcal{L}

for all $i = 0$ to $N + 1$ **do**

for all $T \in del(S)$ **do**

$d_T = 0$;

$d_T = \max\{d_T, \max_{T' \in \nu(T)} \{\phi_T, \phi_{T'} + (1 - i/N) - 1\}\}$;

end for

for all $T \in del(S)$ **do**

$\phi_T = d_T$;

end for

end for

OUTPUT a resulting mesh $d^f(M)$

2) *The relation “South East”*: Directional relations can now be defined in a crisp or a fuzzy mode over any point set. For instance, the absolute directional spatial relation like “south east of” is defined as follows:

$$\begin{aligned} SouthEast_{dil}(M_1) &= d_{7\pi/4}^n(M_1) \\ \text{with } n/d_{7\pi/4}^n(M_1) &= d_{7\pi/4}^{n+1}(M_1) \end{aligned} \quad (11)$$

where d^n stands for $d \circ d \circ \dots \circ d$ n times. The *SouthEast*^f fuzzy operator is directly obtained by combining the definitions 10 and 11 associated with the directional neighborhood ν_α into the algorithmic procedure used to compute d^f . In the following Sec. III, we extent both the scope of spatial relations and of use cases for microscopic as well for macroscopic images.

III. RESULTS AND APPLICATIONS

In this section, we illustrate the efficiency of these tools either to perform automatic analysis or to interact in a semi-automatic way with the visual data in the perspective of a visual cognition paradigm. We achieved structural analysis results over a variety of images (macroscopic/microscopic, natural/document) to assess the effectiveness of the proposed toolbox.

A. Structural analysis

The Lukas-Kanade tracking algorithm of objects over video sequence is based on multi-scale robust point set extractors. We illustrate the structural analysis that can be performed onto this specific sparse representation of a macroscopic image in Fig. 3 and Fig. 4. For the extraction

of the starting sub-mesh of interest, we use the concept of $\alpha - objects$ [9], [8], [3]. The mapping ϕ associates to each triangle T a ϕ_T value corresponding to intrinsic properties of the triangles. For Fig. 3(b), ϕ_T corresponds to a measure of the size of the triangle T defined by its three vertices A,B,C: $\phi_T = \max(AB, AC, BC)$. For Fig. 4(a), ϕ_T corresponds to a measure of the shape of the triangle: $\phi_T = 1/\rho_T$, with ρ_T the radius of the circumscribed sphere to T . The starting sub-mesh corresponds to the simplicial complex $C(S) = T \in Del(S), \phi_T$. Then, the θ threshold is chosen as a median value: $\theta_{opt} = 0.5 * \text{median}_{T \in Del}(\phi_T)$. The starting sub-mesh of interest is $M = \text{bin}_{\theta}(C(S))$. Then, various morphological operators based on the basic erosion (and dilation) operator can derive relevant structural outcomes, in particular outlining face features such as the eyes and the mouth regions (see Fig. 3(d) and (f)). Note that the operators work as well on triangles as on edges in a contour mode [3]. The mouth region and eyes region are obtained in Fig. 3(f) as open-like mesh filtered:

$$\text{MouthAndEyes}(M) = d \circ e \circ e(C) \quad (12)$$

Fig. 4 illustrates another result if we choose the shape criterion for ϕ_T .

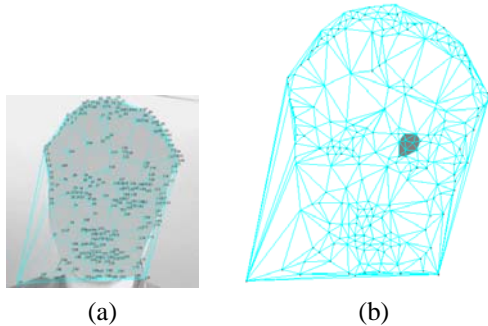


Fig. 4. (a) Thresholding with the Shape Criteria M_{shape} (b) Order 5 opening-like filtering of M_{shape} : $d \circ e^5(C(S))$

B. Document analysis

In [14], the logical structure of a document image with irregular layout - like the location of title and author regions - is determined by the way of a Delaunay triangulation-based method. This method is very robust and can also establish the major threshold values, such as the line space and the inter-character space, automatically, from a document image itself. Some features are measured from different classified Delaunay triangle groups in the Delaunay triangulation model. These feature measures can be improved by the use of morphological mesh operators to clean out the shape of the various regions in the document.

Based on the mass centers of the connected components, Fig. 5 presents such ideas. The gray area corresponds to the dilated mesh for an θ threshold of $\theta_{opt} = 0.75 * \text{median}_{T \in Del}(\phi_T^{size})$ and extracts the specific text area in the document not related to the title areas.

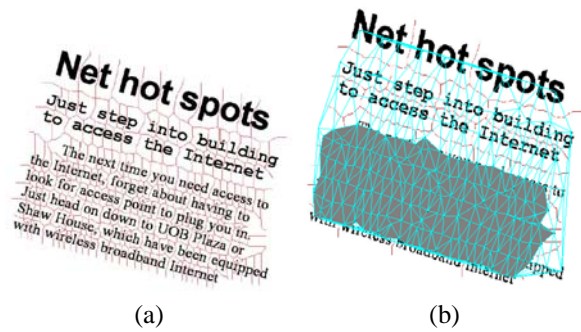


Fig. 5. (a) Mass centers of the connected components, (b) the gray area corresponds to the dilated of the mesh $\text{bin}_{\theta}(C(S))$

C. High-level spatial relation queries

In histopathological images, the need to explore high-content, large images to draw a diagnosis makes it relevant the development of smart interfaces to perform high-level semantic queries involving spatial relations [12]. The diagnosis of breast cancer for instance is based on the analysis of spatial distributions of cells and lumina, which are the mammal ducts to be short. The answer to the query “Show me the cells around the lumina” or the action “Digitalize the cell images between the lumina1 and the lumina2” in the case of a motorized microscope are of major importance to enhance the diagnosis process and to make it possible the digitalized exploration of such data. As for now, the $\phi_T = 1$ values will be displayed white triangles and the $\phi_T = 0$ values will be displayed as transparent triangles. Fig. 6 illustrates two such biological lumina under the form of two meshes M_1 and M_2 (Fig. 6 (b)).

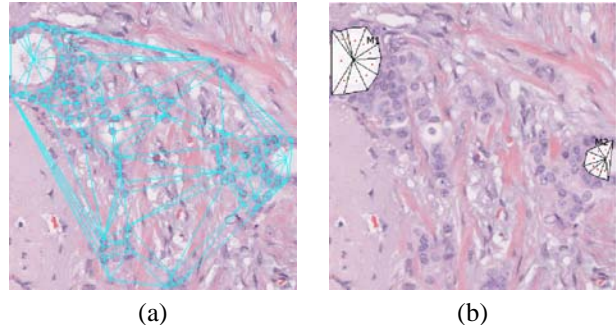


Fig. 6. (a) A small 1024x1024 pixel size images out of a Whole Slide Image (b) Two sub-meshes of interest M_1 and M_2 corresponding to two biological lumina objects.

Fig. 7 illustrates the successive directional dilations of the relation $\text{SouthEast}(M_1)$ until stabilization yielding to the south-east crisp region of the lumina M_1 . In Fig. 7(h) a filtered version by an final isotropic opening is provided: $o_{\text{viso}}(d_{7\pi/4}^n(M_1))$.

Fig. 8 illustrates the *between* relation between M_1 and M_2 in a crisp definition (Fig. 8(e)) and in a fuzzy definition (Fig. 8(f)). Let α be the angle between the x axis and the line joining the centers of gravity of M_1 and M_2 . Fig. 8(a) and (b) illustrate the $d_{\alpha}^n(M_1)$ and $d_{2\pi-\alpha}^n(M_2)$ sub-meshed

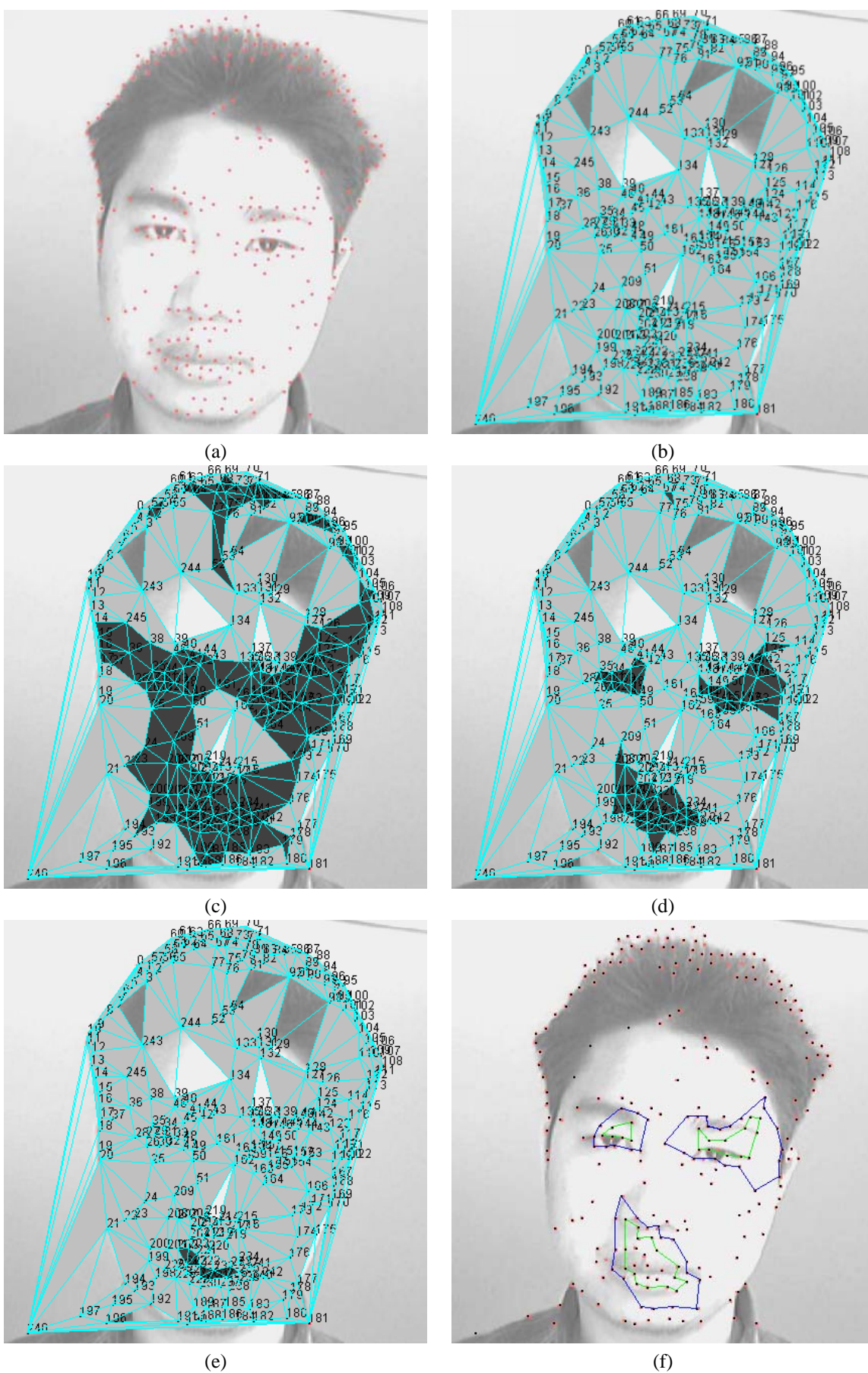


Fig. 3. (a) Kanade-Lukas point set extraction (b) Thresholding with the Size Criteria M_{size} . In black, (c) order 1 erosion-like filtering of M_{size} , (d) order 2 erosion-like filtering of M_{size} , (e) order 3 erosion-like filtering of M_{size} , (f) order 2 opening-like filtering of M_{size}

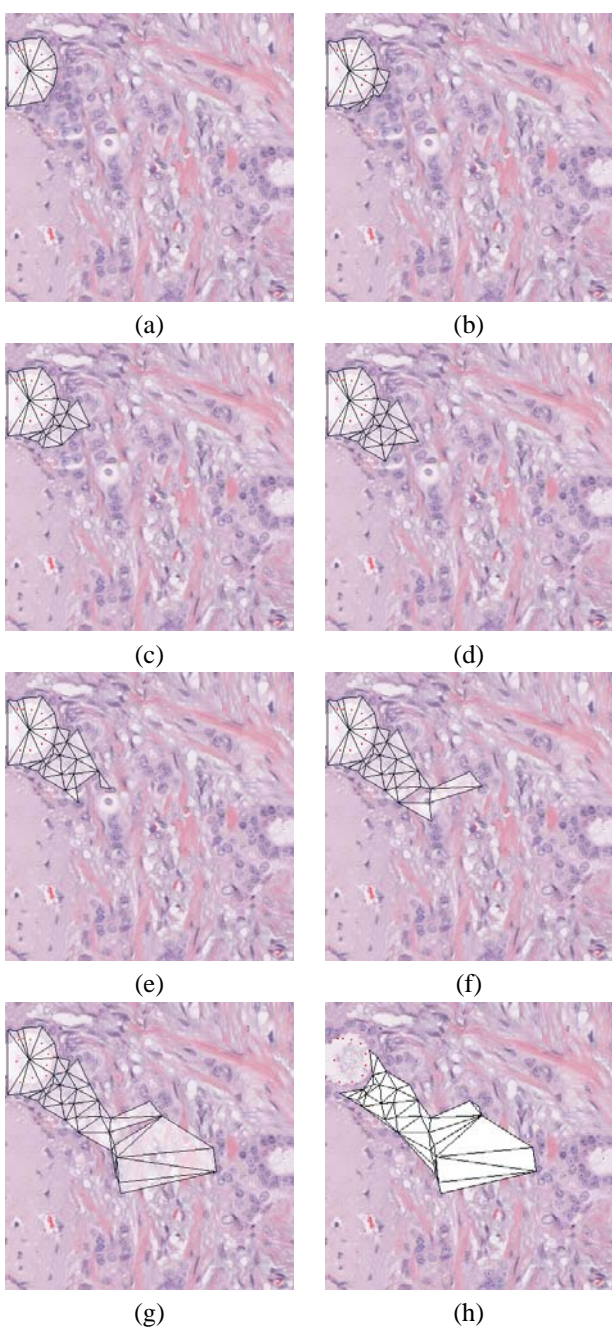


Fig. 7. (a) The sub-meshed region M_1 of the luminal. $M_1 \cup$ successive directional dilations of order (b) 1 (c) 2 (d) 3 (e) 4 (f) 5 (g) the region corresponding to the extent $south - east(M_1) = d_{7\pi/4}^n(M_1)$ (h) after an isotropic opening $o(d_{7\pi/4}^n(M_1))$

regions from which the *between* region is computed as their intersection. Fig. 8(c) and (d) illustrate the $d_\alpha^f(M_1)$ and $d_{2\pi-\alpha}^f(M_2)$ sub-meshed regions from which the *between* region is computed as a $t - norms$ outcome given by $t(a, b) = \min(a, b)$ adapted to the mesh lattice structure. In Fig. 8(f), the outcome shows clearly that the cells right in the middle of M_1 and M_2 achieve a higher fuzzy membership to the fuzzy set *between* than the cells at the frontiers in the darker triangles corresponding to lower fuzzy memberships

to the fuzzy set giving a more realistic and contextual answer to the spatial query than the crisp outcome.

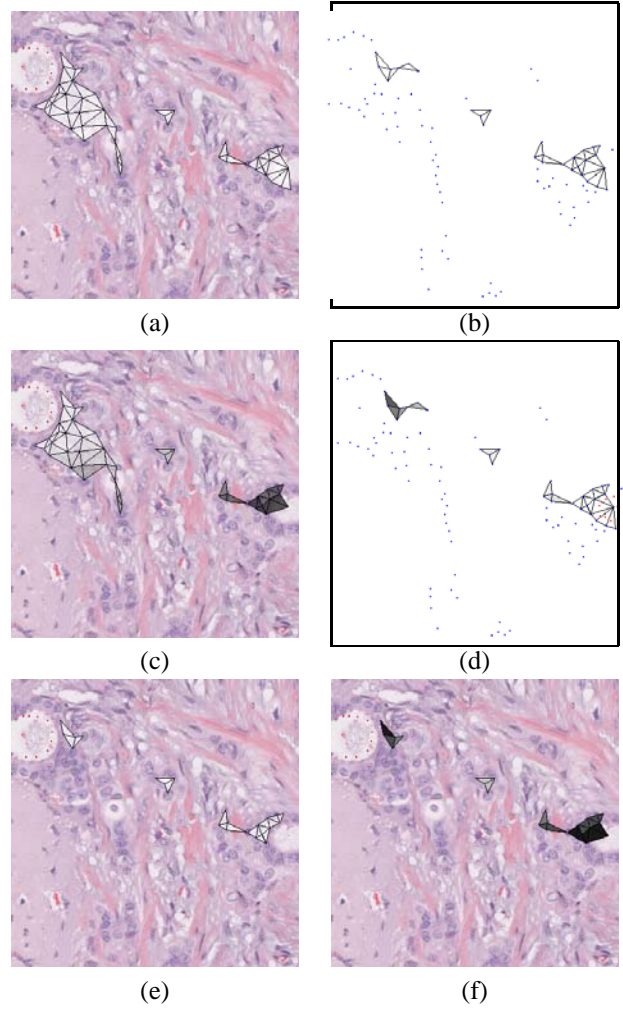


Fig. 8. (a) $d_\alpha(M_1)$ (b) $d_{2\pi-\alpha}(M_2)$ (c) $d_\alpha^f(M_1)$ (d) $d_{2\pi-\alpha}^f(M_2)$ (e) $between_{crisp}(M_1, M_2)$ (f) $between_{fuzzy}(M_1, M_2)$

IV. CONCLUSION

Reasoning at a high-level of conceptualization within digital images is of high interest with the advent of new available high-content large images. We put down the basic principles of a framework to reason about spatial configurations and spatial relations over unorganized point sets. The framework of mathematical morphology has proved to provide sound operators to be combined in order to anchor high-level semantic concepts into quantitative outcomes for efficient interaction and reasoning purposes with visual data. From a theoretical point of view, two open issues remain:

- How to shape specific neighborhood system as flexible structuring entity but in a practical way [11]?
- How to take into account the topological and geometrical features involving the edges and the points in addition to the triangles [13]?

From the cognitive vision perspective, psycho-visual studies or at least ergonomic studies must prove the relevance of such

a theoretical toolbox to enhance the cognitive experience with handling visual data and leverage interactive queries to improve practical application outcomes.

ACKNOWLEDGMENT

The authors acknowledge the support of Ludovic Roux - IPAL for the cognitive virtual microscope project and Jacques Klossa - TRIBVN and National University Hospital of Singapore for providing histopathological images and inspiring fruitful discussions about the new challenges in microscopic imaging systems.

REFERENCES

- [1] E. Clementini, P. Di Felice, Approximate Topological Relations, *International Journal of Approximate Reasoning* (16) (1997) 173-204.
- [2] P. Matsakis, L. Wendling, A new way to represent the relative position between real objects, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (7) (1999) 634-643.
- [3] N. Lomenie, G. Stamon, Morphological Mesh filtering and alpha-objects, *Pattern Recognition Letters* 29 (10) (2008), 1571-1579.
- [4] O. Colliot, O. Camara, I. Bloch, Integration of Fuzzy Spatial Relations in Deformable Models - Application to Brain MRI Segmentation, *Pattern Recognition* (8) (2006) 1401-1414.
- [5] V.T. Ta, O. Lezoray, A. Elmoataz, S. Schupp, Graph-based Tools for Microscopic Cellular Image Segmentation, *Pattern Recognition, Special Issue on Digital Image Processing and Pattern Recognition Techniques for the Detection of Cancer* 42 (6) (2009) 1113-1125.
- [6] ECVision: The European Research Network for Cognitive Computer Vision Systems, "A Research Roadmap of Cognitive Vision", IST Project IST-2001-35454, August 2005.
- [7] I. Bloch, O. Colliot, R.M. Cesar, On the ternary spatial relation "between", *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics* 36 (2) (2006) 312-327.
- [8] H. Edelsbrunner, D.G. Kirkpatrick, On the shape of set of points in the plane, *IEEE Trans. on Information Theory* (29) (1983) 551-559.
- [9] H. Edelsbrunner, E.P. Mucke, E.P., Three-dimensionnal alpha-shapes, *ACM Transactions on Graphics* 13 (1) (1994) 43-72.
- [10] H. Heijmans, P. Nacken, A. Toet, L. Vincent, Graph Morphology, *Journal of Visual Communication and Image Representation* 3 (1) (1992) 24-38.
- [11] L. Vincent, Graphs and Mathematical Morphology, *Signal Processing* 16 (4) (1989) 365-388.
- [12] S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated Grading of Breast Cancer Histopathology Using Spectral Clustering with Textural and Architectural Image Features, in: *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI*, (29) (2008) 496-499.
- [13] J. Cousty, L. Najman, J. Serra, Some morphological operators in graph spaces, in: *Computer Sciences Springer, Lecture Notes* (ed), *Mathematical Morphology and Its Application to Signal and Image Processing*, vol. 5720 (2009) 149-160.
- [14] Y. Xiao, H. Yan, Location of title and author regions in document images based on the Delaunay triangulation, *Image and Vision Computing*, vol. 22 (2004) 319-329.