A Topologically Consistent Color Digital Image Representation by a Single Tree

Pablo Sánchez-Cuevas[1], Fernando Díaz-del-Río[1], Helena Molina-Abril[2(✉)], and Pedro Real[2]

[1] Department of Computer Architecture and Technology, University of Seville, Sevilla, Spain
[2] Department of Applied Mathematics I, University of Seville, Sevilla, Spain
habril@us.es

Abstract. A novel, flexible (non-unique) and topologically consistent representation called CRIT (Contour-Region incidence Tree) for a color 2D digital image I is defined here. The CRIT is a tree containing all the inter and intra connectivity information of the constant-color regions. Considering I as an abstract cell complex (ACC), its topological information can be packed as a smaller (in terms of cells) ACC, whose 2-cells are the different constant-color regions of I. This modus operandi overcomes the classical connectivity paradoxes of digital images by working with lower-dimensional cells such as 0-cells, 1-cells, and 2-cells. The CRIT structure allows to describe this smaller ACC in a non-redundant way. The proposed technique is based on the previous construction of the Homological Spanning Forest (HSF) structures for encoding homological information of the ACCs canonically associated to I, in terms of rooted trees connecting digital object elements without redundancy.

Keywords: Color 2D digital image · Abstract cell complex · Homological Spanning Forest · Contour-region incident tree

1 Introduction

Representing nD digital images simply using a rectangular array of pixel values, has several drawbacks. One of the most important is the "deficiencies" that this representation has with regards to getting consistent local and global spatial topological (region-contour, interior-boundary,...) information. Digital Topology has dealt with these problems of connectedness and continuity in this discrete context, proposing two kind of solutions in terms of enriched representations: (a) those based on neighborhood graphs [2,10]; (b) those based on abstract cell complexes (ACC) of dimension n [3]. The first ones propose for binary images, different neighborhood adjacency between pixels of the same color. They potentially suffer from connectivity paradoxes and incompleteness in adequately gathering all the information of topological interest. Nevertheless, there has been a lot of successful contributions of this type or image representation in the design of algorithms for modelling and analysing images [11]. ACC-based solutions work at inter-pixel level and with cellular structures in which connectivity is initially measured by the notion of incidence between cells of different dimension. Some important progress in this sense has also been published [4].

In this paper, working with a color 2D digital image I as an abstract cell complex(ACC) of dimension two, a new topological representation suitably containing all its intra and inter connectivity information of constant-color regions is defined by using a single tree. We call this representation the Contour Region incidence Tree (CRIT for short). CRITs promise to be simple, easy to manipulate and fast to compute in an almost fully parallel manner, due to the fact that the method is based on the HSF (Homological Spanning Forest) framework for topological parallel computing of 2D digital objects [7–9].

2 Related Works

Among all the topological data structures representing digital images that are available in the literature, there are numerous works dealing with the idea of representing images as neighborhood graphs whose nodes are the pixels of the image and whose edges are connecting two adjacent pixels (with regard, for example, 4-adjacency relationship, for square physical pixels), such that its topological information is condensed in a smaller graph, where 4-connected regions represent nodes in the graph and an edge between two nodes exists if the corresponding regions share a 4-connected frontier.

A classical example of such compact representations is the Region Adjacency Graph. RAGs are simple graphs (no parallel edges, no self-loops), which represent the neighboring relationships of whole regions. A region is a collection of connected pixels, which share some properties. RAGs effectively encode the image into a different representation, that stores the layout of the image, and describes which regions are next to each other and which are not. Unfortunately, some important topological information is missing in semantically correct RAGs of digital images, such as the number of frontiers between two regions (also called degree of contact) or, how these frontiers are connected to each other. RAGs are simple graphs, which implies a disadvantage if we want to encode more complex configurations of regions (i.e. region being included in another region). In this sense, we can affirm that RAG representations are not enough for segmentation and other image processing purposes. In [1], a consistent topological representation is the one that "shows the interactions between regions and then realize the topology of the image". Despite using this not very rigorous definition, we can venture to assert that RAGs are not topologically consistent representations, even in 2D. There are topologically different sets of inter-connections between regions in a continuous or digital image that are locally represented by the same graph pattern in its RAG. For example, Fig. 1 shows two topologically different color images and their RAG, from which they both can not be distinguished.

More complex representations have been proposed to overcome these problems. Dual graphs [5], for instance, can be considered as an extension of the

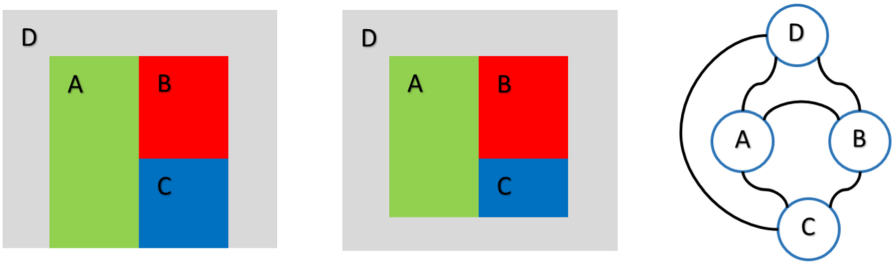


Fig. 1. Left: Two topologically different color images. Right: The RAG representation of these images (same for both)

RAG, were images are represented by a pair of graphs that are finite, planar, connected, not simple in general and dual to each other. These graphs have been used to build a hierarchy of image partitions called Dual Graph Pyramids. Although dual graphs are able to encode inclusion relationships and multiple borders, some ambiguities still remain unsolved, were for instance two images in which two regions are interchanged will have exactly the same dual graph representation. Furthermore, processing these two graphs is not time efficient, and its practical application to real images is quite limited.

Other structures such as combinatorial maps [6] overcome these ambiguities by representing the image as a set of half-edges called darts, and two permutation functions. Advancing in complexity, generalized maps have advantage over the classical combinatorial maps that they can also represent non-orientable objects. The problem with these more complicated structures (such as combinatorial and generalized maps) is their complexity and difficult manipulation. Summing up, from the viewpoint of considering a color 2D digital image I as a ACC object whose 2-dimensional cells are the physical pixels of I, RAG-based representations mainly represent the topology of the image using the connectivity graph involving all the incidence relations between 2-cells and 1-cells of I and combinatorial maps from the graph involving all the incidence relations between 0-cells and 1-cells of I. In 2D, using exclusively one of these graphs and topological duality properties, it is possible to extract topological information of I in terms of number of color-constant connected components and 1-dimensional homological holes, but it seems difficult to solve using these models, for instance, the problem of *homological hole's classification*: to determine the homological equivalence class to which a simple closed curve inside the object belongs to. This shortcoming and other more complex problems involving homotopy information (that is, in which the boundary information of any cell is treated as a loop structure) prevents these models from being fully "topologically consistent".

A CRIT representation of I is a subgraph of the connectivity graph of I, from which it is automatic to get a more compact and topologically informative ACC, whose 2-cells are the different constant-color regions of I. In this way, the previous difficulties encountered for preventing an enriched RAG description to be a topologically consistent representation are overcome, due to the fact that

CRIT description is based on the topological model of Homological Spanning Forest (HSF) of an ACC, which will be introduced in the following Section.

3 Building the CRIT

An abstract cell complex (or ACC, for short, [4]) can appropriately describe all the regions and their relations of a two dimensional image I. In this paper, pixels are considered to be cells of dimension 2. Interpixel cells of dimension 1 are the edges between pixels and the corners between pixels are defined as cells having dimension 0 (see Fig. 2). The keypoint in this ACC is that its cells can hold all the topological information of a 2-dimensional image composed of several constant-color regions, that is, a puzzle of objects (represented by two-dimensional cells) and their adjacency relations (represented by one and zero-dimensional cells).

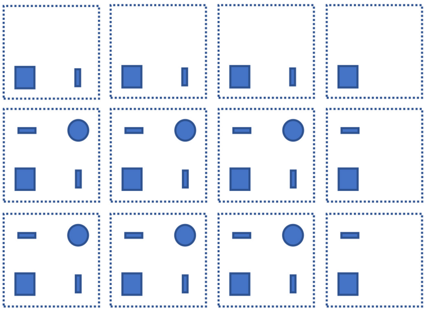


Fig. 2. ACC for a 2-dimensional image. Squares are 2-cells (image pixels), dotted lines are 1-cells (contours) and circles are 0-cells (crosses).

In Fig. 3 a 3×5 image with three regions is represented by its ACC. Not only regions are labelled with their colors A, B, C, but their contours and crosses are congruently labelled with those colors of their adjacent cells. Contours are a mix of two colors, and crosses are defined by the colors of their four surrounding pixels. Those 0 and 1 cells which are in the interior of a region, can be simply labelled by one color.

Once the different 0, 1, 2-dimensional objects are represented by a set of cells, their HSFs can be straightforwardly obtained for them. Roughly speaking, an HSF of a digital image I is a set of trees living at interpixel level within an abstract cell complex version of I (or ACC, for short, [4]) and appropriately connecting all cells without redundancy. This notion is in principle independent of the pixel's value within the image. Using this concept, any digital 2D image I can be seen as a cell complex formed exclusively by two trees (see [7,8]). One of them, is formed by all the cells of dimension 0 and some of dimension 1 and the other one, is formed by the rest of cells of dimension 1 and all the cells of dimension 2 of I. More concretely, crosses are only represented by a unique

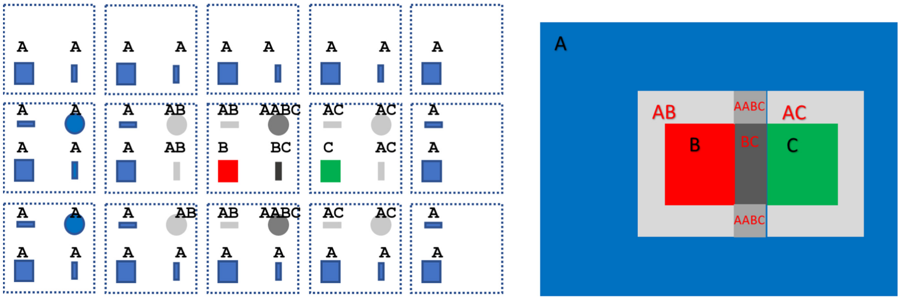
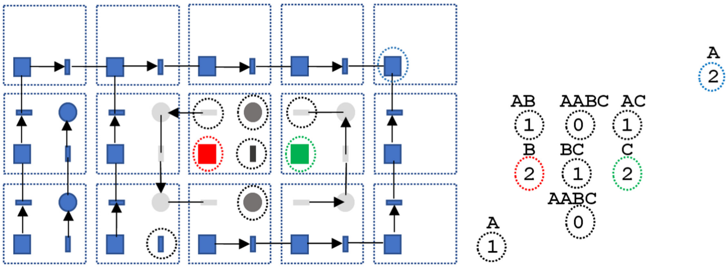


Fig. 3. Left: A 2-dimensional color image seen as an ACC. Regions are represented by 2-cells, contours by 1-cells and crosses by 0-cells. Each cell is labeled with its representative color. Right: The same image seen as a puzzle of 0, 1 and 2 dimensional objects.

tree (just one node, that is, a 0-cell). Contours are a set of 0 and 1-cells, whose "homology" information can be represented by one 1-cell and a set (which can be empty) of 0-cells representing the existence of holes. And finally, regions are a set of 0, 1 and 2-cells, whose homology information can be represented by one 2-cell and a set (which can be empty) of 1-cells representing the existence of holes. To fix ideas, the set of HSFs of Fig. 3 is depicted only by the cells representing their homology information in Fig. 4 (center), and a possible set of HSFs on the left of this image. The construction of HSFs is based on the concept of primal and dual vectors. Given two incidence cells c and c' of dimensions k and $k+1$ respectively, a primal (resp. dual) $(k, k+1)$ vector (c, c') (resp. (c', c)) connects the tail c (resp. c') with the head c' (resp. c). In this case, HSFs have been built following a boundary criterion for the primal vectors: from each k-cell one of its boundary $(k-1)$-cells (of the same object) has been chosen (represented by an arrow). Dual vectors are represented by a line, and they link each $(k-1)$-cell with all of its coboundary (k)-cells (of the same object).

In this case, the homology of region A can be condensed to one 2-cell and one 1-cell, whereas regions B and C can be described only by one 2-cell each, thus conforming the table on the right of the Figure. The three contours have no holes, that is, each of them is represented by one 1-cell, called AB, AC and BC. Each one of the two crosses are given by one 0-cell (named $AABC$). Finally, note that this process is fully generic for any 2-dimensional image with the 4-adjacency criterion

To this point, a color image can be condensed into a set of representative homology cells, each one being the root of a certain HSF tree in the case of two-dimensional objects. Additionally, these cells can be labelled with the colors that embody the relations among objects (see Fig. 4, right). The next step is taken here for efficiency purposes. Instead of managing relations among these cells or linking them in a graph (like in the case of RAG), it is preferable to condense all the homology information into a unique tree, called CRIT (Contour Region incidence Tree). CRIT combines the information of the flat regions, contours and



#	dim	color
1	0	AABC
2	0	AABC
3	1	AB
4	1	AC
5	1	BC
6	1	A
7	2	A
8	2	B
9	2	C

Fig. 4. Left: A possible set of HSFs for the image on Fig. 3. Cells representing homology information are surrounded by a dashed circle. Center: The set of homology information of the image indicating their dimension and representative color. Right: A table with the homology information of the image.

crosses of the original color image, plus the relations between them. According to Fig. 4, this supposes the conversion of the table of Fig. 4 (right) into a unique tree. To do this, some linking rules must be defined so as to prevent graphs. This can be achieved for instance by:

– For those cells of the same object, the same boundary criterion of the HSF building can be obeyed, that is, linking each representative (k)-cell with only one ($k-1$)-cell, until a 0-cell is reached. This comprises a new set of primal vectors.
– Inserting additional primal vectors by linking the rest of k-cells with one ($k-1$)-cell (of other object). Although, this usually supposes that 0-cells would remain disconnected to the rest of cells, for a simple image, the CRIT building may finish here. This is the case of Fig. 5, which contains a simple image of 4 pixels of different colors, so that a cross of 4 regions exists in the middle.
– Introducing additional links from the 0-cells to only one of their coboundary 1-cells, that is, a new set of dual vectors. This step must forbid forming graphs, for instance, by defining a direction for the links and preventing that the dual vectors go backwards along the selected direction.

In order to systematically build a CRIT in the most convenient way, a color incidence criterion can be followed. One possible set of rules is the following:

a) Firstly, "closing" each object by linking each k-cell with one of its boundary ($k-1$)-cells (of the same object, that is, the same color). These are represented by continuous arrows in this paper. For instance, in Fig. 6, only one primal vector exists (region A).
b) Secondly, inserting the rest of additional primal vectors by linking the rest of k-cells with one ($k-1$)-cell (of other object) having the most similar color to the k-cell. These are represented by a dashed arrows in Fig. 6. No matter that two or more ($k-1$)-cells match in the similar color criterion. For example,

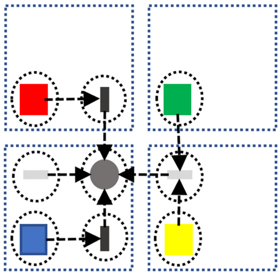


Fig. 5. A simple image of 4 pixels having different colors, so that a cross of 4 regions exist in the middle. In addition, this image can built a CRIT just by inserting additional primal vectors from the homology representative k-cells to one $(k-1)$-cell (of other object).

contour BC can be linked to the upper or the inferior cross. Similarly region B may link with contour BC or with the AB, and so on. For this image, five of these additional primal vectors must be inserted.

c) Finally, dual vectors are represented by double arrows in the same Figure. Note that the criterion has been "going outwards", that is, from the center to the image border. Thus, each 0-cell goes to the 1-cell (hole of region A) that surrounds the composite region B, C.

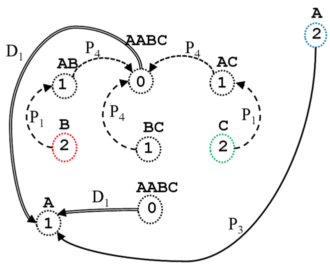


Fig. 6. Building the CRIT by linking the homology group cells of the image in Fig. 3

This criterion is followed in all the figures of this paper. Previous rules can be straightforwardly determined for bidimensional color images. More concretely, objects can be classified into five cases (Fig. 7, top), and therefore, only seven primal vectors and two dual vectors can exist (Fig. 7, bottom). Besides the link cases that appear in Fig. 6, the rest of possible links are drawn for a different color image in Fig. 8.

To sum up, note that a CRIT representation is consistent, so that any topological representation can be extracted from it. For instance, RAGs can be straightforwardly calculated by simply selecting the 1-dimensional objects (they

#	Object description	Set of homology groups (representative cells)
O_2	A region without holes	2
O_{2_1}	A region with one or more holes	2, 1, …, 1
O_1	A contour without holes	1
O_{1_0}	A contour with a hole	1, 0
O_0	A cross (of 3 or 4 regions)	0

#	Additional Primal Vectors (description)	Example
P_1	2 → 1: from object O_2 to O_1	Fig. 6 from B to AB, from C to AC
P_2	2 → 1: from object O_2 to O_{1_0}	Fig. 8 from 2-cell of E to 1-cell of ED
P_3	2 → 1: from object O_{2_1} to O_{2_1}	Fig. 6 from 2-cell of A to 1-cell of A
P_4	1 → 0: from object O_1 to O_0	Fig. 6 from AB to AABC, from AC to AABC, from BC to AABC
P_5	1 → 0: from object O_{1_0} to O_{1_0}	Fig. 8 from 1-cell of ED to 0-cell of ED
P_6	1 → 0: from object O_{2_1} to O_{1_0}	Fig. 8 from 1-cell of E to 0-cell of F
P_7	1 → 0: from object O_{2_1} to O_0	Fig. 8 from 1-cell of F to FGH

#	Additional Dual Vectors (description)	Example
D_1	0 → 1: from object O_0 to O_{2_1}	Fig. 6 from AABC to 1-cell of A (two vectors)
D_2	0 → 1: from object O_{1_0} to O_{2_1}	Fig. 8 from 0-cell of ED to 1-cell of E

Fig. 7. Top: The five possible objects for bidimensional color images from a topological point of view. Bottom: Possible edges among topological objects for bidimensional color images.

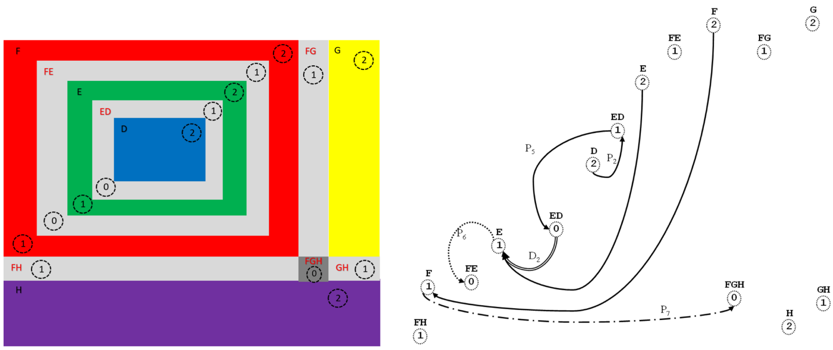


Fig. 8. Left: A 2-dimensional color image that contains additional links that are not present in Fig. 6. Each homology representative cell is drawn with a dashed circle and its dimension. Right: Homology representative cells and the additional links (P_2, P_6, P_7, D_2). Each cell is labeled with its representative color and its dimension. The rest of additional links are not drawn for clarity purposes.

are the edges of the RAG) and the 2-dimensional ones (they compose the RAG nodes). Each RAG edge directly associates the two nodes of its representative color. However, an interesting circumstance is when two regions share two (or more) contours (Fig. 9). In this case, regions K and L share two common contours, namely $KL1$, $KL2$. Usual RAG representation is shown in Fig. 9, right, where the edge between K and L is sometimes weighted by a factor of 2 to indicate this double contact. Nonetheless, common RAGs do not represent the hole of M, nor the fact that K and L are contained in it.

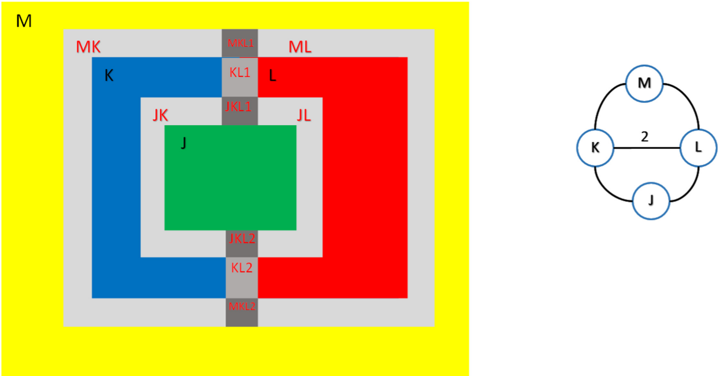


Fig. 9. Left: An image with two regions K and L having two common boundaries, namely $KL1$, $KL2$. Because of external M and internal J regions, there are also 4 crosses that can be represented by the same colors in two couples ($MKL1, MKL2, JKL1, JKL1$). Right: Classical RAG representation for this image.

4 Conclusions and Future Work

A CRIT model of a color 2D digital image I (seen as an ACC) is a new topologically consistent tree-representation having as nodes the different 0-cells, 1-cells and 2-cells of I and having as edges some of the incidence relationships between 0-cells and 1-cells or 1-cells and 2-cells. The set of the rest of incidence relationships for any node in a CRIT that are not expressed as edges of it is suitably stored as node information. From a CRIT, it is automatic to construct the smaller ACC that encodes the homological information of I, having as 2-cells the different constant-color connected regions of I, as 1-cells, the different connected cracks between two regions and as 0-cells the cross points of I. Moreover, although this model is not unique for one image and its construction depends on some fusion criterion for the regions, it is also straightforward to change one CRIT into another one. In a near future, we plan to advance in the following issues:

- to design and implement an almost-full parallel algorithm for computing a CRIT of a digital image of dimension $n \times m$ of complexity near $\log(n+m)$.
- to design fast algorithms for transforming one CRIT into another one that meets specific incidence requirements;
- to define new geometric and topological features or characteristics associated to a CRIT.

Acknowledgments. This work has been supported by the Ministerio de Ciencia e Innovacion of Spain and the AEI/FEDER (EU) through the research project PID2019-110455GB-I00 (Par-HoT) and the IV-PP of the University of Seville.

References

1. Fiorio, C.: A topologically consistent representation for image analysis: the frontiers topological graph. In: Miguet, S., Montanvert, A., Ubéda, S. (eds.) DGCI 1996. LNCS, vol. 1176, pp. 151–162. Springer, Heidelberg (1996). https://doi.org/10.1007/3-540-62005-2_13
2. Klette, R.: Cell complexes through time. Communication and Information Technology Research Technical report 60 (2000)
3. Kovalevsky, V.: Algorithms and data structures for computer topology. In: Bertrand, G., Imiya, A., Klette, R. (eds.) Digital and Image Geometry. LNCS, vol. 2243, pp. 38–58. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45576-0_3
4. Kovalevsky, V.: Geometry of Locally Finite Spaces. Baerbel Kovalevski, Berlin (2008)
5. Kropatsch, W.: Building irregular pyramids by dual-graph contraction. In: IEEE Proceedings-Vision, Image and Signal Processing, pp. 366–374 (1995)
6. Lienhardt, P.: Topological models for boundary representation: a comparison with n-dimensional generalized maps. Comput.-Aided Des. **23**(1), 59–82 (1991)
7. Molina-Abril, H., Real, P.: Homological spanning forest framework for 2D image analysis. Ann. Math. Artif. Intell. **64**(4), 385–409 (2012)
8. Diaz-del Rio, F., Real, P., Onchis, D.: A parallel homological spanning forest framework for 2d topological image analysis. Pattern Recognit. Lett. **83**, 49–58 (2016)
9. Diaz-del Rio, F., Sanchez-Cuevas, P., Molina-Abril, H., Real, P.: Parallel connected-component-labeling based on homotopy trees. Pattern Recognit. Lett. **131**, 71–78 (2020)
10. Rosenfeld, A.: Adjacency in digital pictures. Inf. Control **26**, 24–33 (1974)
11. Saha, P., Strand, R., Borgefors, G.: Digital topology and geometry in medical imaging: a survey. IEEE. Trans. Med. Imaging **34**(9), 1940–1964 (2015)