GETTING A MORPHOLOGICAL TREE OF SHAPES FOR MULTIVARIATE IMAGES:
PATHS, TRAPS, AND PITFALLS

Edwin Carlinet \textsuperscript{1,2}

\textsuperscript{1} EPITA Research and Development Lab. (LRDE)
14-16, rue Voltaire
FR-94276 Le Kremlin-Bicêtre, France
{firstname.lastname}@lrde.epita.fr

\textsuperscript{2} Université Paris-Est,
Laboratoire d’Informatique Gaspard-Monge (LIGM)
A3SI, ESIEE Paris, Cité Descartes, BP 99
FR-93162 Noisy-le-Grand, France

Thierry Géraud\textsuperscript{1,2}

\textbf{ABSTRACT}

The tree of shapes is a morphological tree that provides an
high-level hierarchical representation of the image suitable
for many image processing tasks. This structure has the
desirable properties to be self-dual and contrast-invariant
and describes the organization of the objects through level
lines inclusion. Yet it is defined on gray-level while many
images have multivariate data (color images, multispectral
images…) where information are split across channels. In
this paper, we propose some leads to extend the tree of shapes
on colors with classical approaches based on total orders,
more recent approaches based on graphs and also a new
distance-based method. Eventually, we compare these ap-
proaches through denoising to highlight their strengths and
weaknesses and show the strong potential of the new methods
compared to classical ones.

\textbf{Index Terms}— Mathematical Morphology; Tree of
Shapes; Color Image Processing; Filtering.

1. INTRODUCTION

Mathematical morphology (MM) offers a non-linear image
processing framework which is both simple and efficient. It
has been widely used in many image processing tasks as for
filtering, object detection, segmentation… Behind the scene,
MM operators rely on an ordering relation on values which
must form a lattice. Those operators usually apply to binary
and grayscale images, even if there are many attempts to deal
with color images \[\textsuperscript{11}\].

A particular class of MM operators, named connected op-
erators, have been widely investigated thanks to their contour-
preserving properties. Several authors \[\textsuperscript{2,3,4}\] propose to rep-
resent the inclusion of connected components in hierarchical
structures: the min-tree (resp. max-tree) states the inclusion
of the lower (resp. upper) level sets. Since then, many con-
ected operators have been developed on those trees. A self-
dual hierarchical representation of the image has been pro-
posed in \[\textsuperscript{5}\], which merges the min-tree and the max-tree into
a unique structure, called the tree of shapes. Despite its large
potential in many image processing and computer vision ap-
lications \[\textsuperscript{6,7,8,9}\], it is still under-exploited.

While most MM filters rely on a partial ordering relation
only, component-trees and trees of shapes need the ordering
to be total. The latter requirement hinders the extension of
these structures to color images. Some of the many suggested
workarounds are reviewed in Section \textsuperscript{2}. Basically, they con-
sist either in defining a total (pre)order on colors or dealing
directly with the partial order on a structure which is not a
tree anymore. The former approach is illustrated in Section \textsuperscript{3}
as a straightforward extension of the tree of shapes to colors.
The contributions of this paper are:

\begin{itemize}
  \item a natural extension of the algorithm \[\textsuperscript{10}\] from Géraud et al.
    using an alternative definition of shapes,
  \item a graph approach to organize the set of shapes and deal
directly with the partial order.
\end{itemize}

Both approaches are investigated in Sections \textsuperscript{4} and \textsuperscript{5}
respectively. Section \textsuperscript{6} eventually show a preliminary comparison
of these methods through denoising. Last we conclude and
give perspectives in Section \textsuperscript{7}.

2. HIERARCHICAL STRUCTURES ON
MULTIVARIATE DATA

2.1. Processing Colors with Trees

Among hierarchical representations of images, component
trees rely on a data (pixel values) ordering that needs to be
total. When dealing with multivariate data, authors generally
have two approaches. They either define a new total ordering
relation on their data, or adapt their structures to deal with the
natural partial order. Many authors have attempted to define
orders on colors (see for example \[\textsuperscript{11,12}\] or Aptoula and
Lefèvre \[\textsuperscript{13}\], Velasco-Forero and Angulo \[\textsuperscript{14}\] for a rather
complete survey of color orders for MM processing) but few of them [15, 16, 17] have been involved in studying their influences on component trees (and even fewer on the tree of shapes [13]). While it may seem straightforward to construct a component tree (or a tree of shapes) once a total ordering relation is chosen, there are still some issues about the reconstruction process in case of preorders [15, 17]. Few people have been involved in the second approach that considers the structure implied by a partial order. However, some recent works [19, 20, 21] have proposed to generalize the component trees to component graphs, and that idea can be transposed to the tree of shapes.

2.2. Reminders about the Tree of Shapes

Let $u$ be an image defined over a domain $\Omega$ (typically $\mathbb{Z}^2$ for 2D images), with values on a set $V$ equipped with an ordering relation. We note $\leq$ whenever the considered relation is total and $\preceq$ whenever it is partial.

The lower level set at level $\lambda \in V$ is defined by $[u \leq \lambda] = \{ p \in \Omega, u(p) \leq \lambda \}$: the upper level set $[u \geq \lambda]$ is defined similarly. Level sets will also be called cuts (as in [10]) in the rest of the paper. Let $\mathcal{C}(X) \subset \mathcal{P}(\Omega)$ denote the connected components of the set $X \subset \Omega$, $\mathcal{L}(\lambda) = \mathcal{C}([u \leq \lambda])$ the lower components at level $\lambda$, and $\mathcal{L} = \cup_{\lambda \in V} \mathcal{L}(\lambda)$ the set of lower connected components. The set $\mathcal{L}$ endowed with the inclusion relationship is a hierarchical structure named the min-tree. Its dual $\mathcal{U}$, based on upper level sets, forms the max-tree.

Let $X \subset \Omega$ and $cc(X, \partial \Omega)$ be the connected component of $X$’s complement which connects with the domain’s border. Then, let $\text{Sed}$ be the hole filling operator defined as $\text{Sed}(X) = \Omega \setminus cc(X, \partial \Omega)$. Then, a shape is a connected components without holes from $\mathcal{L}$ and $\mathcal{U}$. We call a lower shape of level $\lambda$, an element of $\mathcal{S}_\lambda^- = \{ \text{Sed}(X), X \in \mathcal{L}(\lambda) \}$, the whole set of lower shapes is noted $\mathcal{S}^-$. $\mathcal{S}_\lambda^+$ and $\mathcal{S}^+$ are defined in the same way from $\mathcal{U}(\lambda)$. Any two elements from the set of shapes $\mathcal{S} = \mathcal{S}^+ \cup \mathcal{S}^-$ being either disjoint or nested, $(\mathcal{S}, \subseteq)$ also forms a tree called the tree of shapes.

2.3. Grain Filters and Reconstruction

Given a total order $(V, \leq)$, an image can be reconstructed from the set of shapes by:

$$
\hat{u}(x) = \max \{ \lambda \mid x \in A \in \mathcal{S}_\lambda^-, x \not\in B \in \mathcal{S}_{\lambda'}^+, B \subset A \} \\
\quad = \min \{ \lambda \mid x \in A \in \mathcal{S}_\lambda^+, x \not\in B \in \mathcal{S}_{\lambda'}^-, B \subset A \} \quad (1)
$$

The grain filter consists in removing shapes having a size below a given threshold and reconstructing from a subset of $\mathcal{S}$. Algorithmically speaking, filtering means pruning the tree of shapes and merging the deleted nodes with the first parent node verifying the criterion. Reconstructing consists in assigning to each pixel, the level of the node it belongs to.

3. TOTAL (PRE-)ORDERS APPROACH

A classification of ordering relations can rely on their algebraic properties (totality, anti-symmetry...) or with respect to the way these relations are built. Barnett [22] proposed to classify them into four groups: marginal (M-ordering), conditional (C-ordering), partial (P-ordering) and reduced (R-ordering). The former is a component-wise ordering that deals separately with each channel, i.e., a partial order, and it will be discussed in Section 5. The three others provide total (pre)orders and are discussed below.

With C-ordering, vectors are ordered by mean of one, several, or all of their marginal components. The most well-known C-ordering is the lexicographical ordering, that is a total order. If only several components participate in the comparison, it yields a total-preorder. For example let $v, w \in \mathbb{R}^n$, the lexicographical ordering $\leq_L$ using only the two first components $(v \leq_L w$ iff $v_1 < w_1 \lor v_1 = w_1 \land v_2 \leq w_2)$ is a total-preorder. Colors $(1, 1, 2)$ and $(1, 1, 3)$ are considered as equivalent by the above relation. The main pitfall of C-orderings lies in the importance given to the first components. Considering the RGB space for example, it implies that the red component is more relevant than the others. Workarounds like the sub-quantization of first components (also known as $\alpha$-lexicographical ordering [23, 24]) enables to lower the importance given to the first dimension.

With R-ordering, vectors are reduced to scalar values using a mapping $h : \mathbb{R}^n \rightarrow \mathbb{R}$ s.t. $v < w \iff h(v) < h(w)$. If $h$ is injective then each index is mapped to a unique color and the relation is a total order, otherwise it is a total preorder. Typical examples of R-ordering are distance-based ordering. It consists in choosing a reference vector (or reference set of vectors) $v_{\text{ref}}$ and the order relation is built upon a distance to $v_{\text{ref}}$ ($v \leq_R w$ iff $d(v, v_{\text{ref}}) \leq d(w, v_{\text{ref}})$). The main drawback of distance-based orders lies in the choice of reference vectors. The method in Section 4 can be seen as a local R-ordering where the reference vector is the parent shape level.

Image Restitution. While some applications (e.g., object detection, segmentation) are not concerned with the restitution of the image, filtering requires reconstructing from the pruned tree. If the order used to construct the tree is a not anti-symmetric (preorder), then a single node is associated with several levels and reconstruction issues arise. In [15], two strategies are proposed: $P_{\text{mean}}$ and $P_{\text{median}}$. During restitution, the former assigns the mean value of the node’s pixels whereas the latter assigns the median value (based on the total lexicographical order). Tushabe and Wilkinson [17] extend those ideas with: (1) reconstructing the removed shapes with the mean color of the parent shape and leave others pixels intact (Mean Parent (MP)), (2) reconstructing the removed shapes with the closest color in the parent shape (Nearest Color (NC)) and (3) computing the geodesic influence of the parent shape pixels in the removed shape and assigning the closest pixel color (Nearest Neighbor (NN)). The last strategy is not a connected filter, thus it is not considered in the comparison in Section 6.

4. A DISTANCE-BASED APPROACH

While the initial definition of the tree of shapes [25, 26] relies on the upper and lower level sets only, it encodes two pieces of information: the level sets and the level lines. In Morse theory, level lines are either defined by iso-contours
\[ u = \lambda \] or in terms of contours of lower or upper level sets. In discrete topology, the second definition is more convenient because it does not require \( u \) to be continuous. Nevertheless, a recent tree of shapes computation algorithm uses a continuous interpretation scheme of the image using an interval-valued map on a Khalimsky grid where isolines are well-defined. The algorithm works as follows: it starts a propagation from the border of the image and floods every level set in a continuous way, i.e., \( \lambda \) is the level currently flooded, the next flooded level is either \( \lambda + 1 \) or \( \lambda - 1 \). Once the image has been totally flooded, a union-find step processes the pixels in the reverse order of the propagation and builds the tree of shapes. The algorithm thus provides a recursive definition of shapes in gray levels: given a shape \( s_\lambda \) at level \( \lambda \), one can generate a set \( S' \) of \( s_\lambda \) sub-shapes with:
\[
S' = \{ s_\lambda = \text{Sed}(X), \ X \in \mathcal{C}(s_\lambda \backslash \{u = \lambda\}) \}
\]
with \( \lambda' = \lambda + 1 \) or \( \lambda' = \lambda - 1 \).

Definition given by Equation (2) and the algorithm ensure that shapes will form a tree. However, it requires to choose a next flooding level \( \lambda' \) in order to continue the propagation in a hole. In grayscale, we choose either \( \lambda + 1 \) or \( \lambda - 1 \) which is the closest upper or lower level available in the hole’s border. To mimic this continuous behavior, we propose to choose also the closest level in terms of a distance which is equivalent to impose locally a total (pre)order based on a distance to a reference color. Equation (2) becomes:
\[
S' = \{ s_\lambda = \text{Sed}(X), \ X \in \mathcal{C}(s_\lambda \backslash \{u = \lambda\}) \}
\]
with \( \lambda' = \min_{v \in u(d_{\text{Sed}}(X))} \| \lambda - v \| \)

Note that Equation (3) is equivalent to the original definition in grayscale levels. Also, since \( \lambda' \) may not be unique, one can combine the distance with a lexicographical cascade to avoid the pitfall of multiple minima.

5. A GRAPH APPROACH

Since a natural total order is not obvious, some approaches try to deal with the natural partial order that does exist. A set of shapes \( S \) is endowed with the inclusion relationship, the Haas diagram of \((S, \subseteq)\) is a directed acyclic graph (DAG) which is our processing structure. We consider two sets for \( S \):
1) \( \tilde{S} = \{ \text{Sed}(X), \ X \in \mathcal{L}_x \cup \mathcal{U}_x \} \)
2) \( \tilde{S} = \{ \text{Sed}(X), \ X \in \bigcup_i \mathcal{C}(\{u_i \leq \lambda_i\}) \cup \mathcal{C}(\{u_i \geq \lambda_i\}) \} \)

In other words, \( \tilde{S} \) is the set of all shapes extracted from every cuts \( \{u \leq \lambda\} \) and \( \{u \geq \lambda\} \), while \( \tilde{S} \) is the set of marginal shapes. In Figure 2 where \( \tilde{G} \) denotes the shape-graph associated with the cover relation of \((\tilde{S}, \subseteq)\) and \( G \) the one of \((S, \subseteq)\), we highlight the difference between the two sets of shapes and shows that \( \tilde{G} \) is finer that \( G \). Actually, any lower marginal cut can be obtained from \( \{u \leq \{T, \ldots, T, \lambda_i, T, \ldots, T\}\} \) and upper marginal cuts as well), thus \( \tilde{S} \subset \tilde{S} \). However, \( \tilde{S} \) and \( \tilde{G} \) have an interest from a computational point of view because, contrary to \( G \), \( \tilde{G} \) can be computed efficiently by merging the marginal trees of shapes built on individual channels.

Filtering and Reconstruction. One can reconstruct \( u \) from \( \tilde{G} \) using the same principle of Equation (1) where the max and min are respectively replaced by the infimum and the supremum. Algorithmically speaking, a node of the DAG \( \tilde{G} \) is filtered with the value of the least common ancestor (LCA) of its parents (LCA strategy). If the LCA is both a lower and an upper shape with different level, the mean of the levels is chosen as the restitution value. This restitution rule tends to create color artifacts since the LCA may be distant in the DAG hierarchy. Thus, we propose another strategy derived from Equation (1). A removed node is filtered with the color of the closest (in terms of the color distance) parent node (NC strategy). In Figure 3a, the least common ancestor of the shapes to remove is the root, thus the LCA strategy reconstructs the shapes with value \((1,1,1)\). On the contrary, with the NC strategy, the \( e, f \) shape can be reconstructed with value \((1,1,0)\) or \((0,1,1)\) depending on the chosen distance. Those values are closer to the subject one \((1,0,0)\). Reconstructing \( u \) from \( \tilde{G} \) is straightforward and results in a component-wise restitution, i.e., \( \tilde{u}(x) = (\tilde{u}_1(x), \ldots, \tilde{u}_n(x)) \) where \( \tilde{u}_i(x) \) is given by Equation (1) based on shapes \( \delta_i \) and \( \tilde{S}_{i} \).
We compare the different approaches for denoising. Even if this application does not take advantage of high-level abstraction provided by the Tree/Graph of Shapes, it is still a good assessment of the meaningfulness and the coherence of the representation. The dataset is composed by 65 color images with a strong chromatic component used for image compression and color constancy test \cite{28, 29}. A Gaussian noise ($\mu = 0$, $\sigma = 30$) is first applied independently on each channel and we then compare the three approaches: total (pre)order-based, distance-based and graph-based. For the first category, we evaluate the RGB lexicographical order (namely Lex) and preorders based on lightness (in La*b* and HLS spaces), brightness (as the average of the three RGB components) and chrominance (in La*b* space, $C = \sqrt{a^2 + b^2}$). They are referred to $L_{\text{La}^*b^*}$, $L_{\text{HLS}}$, Brightness and $C_{\text{La}^*b^*}$. Values are quantized to 256 levels and each method was tested with the MP, NC and $P_{\text{mean}}$ strategies. The second and third classes of methods are those described in Sections \ref{sec:methods} and \ref{sec:distance-based}. We use the Peak Signal to Noise Ratio (PSNR) to assess the denoising quality. We apply grain filters of size 2 to 500, and we retain the highest PSNR score for each pair (method, image)\footnote{An in-depth description of the evaluation and complete results are available on \url{http://lrde.epita.fr/wiki/Publications/carlinet.14.lcip}}.

Table 1 shows the min, max and average PSNR score for each method, as well as the average grain size $\lambda$ for which the maximum score has been obtained. The shape-graph approach outperforms the others, then come preorders based on $P_{\text{mean}}$ and MP strategies and then, the distance-based approach. The Nearest Color strategy performs the worst in our benchmark. It is not surprising that the best methods are those that tend to blur the image by averaging color when reconstructing (MP, and $P_{\text{mean}}$) or by reconstructing marginally (Shape-Graph $\hat{G}$). Even if they create false colors (colors not in the original image), the new colors may (or may not) make sense in the original image. Indeed, in Figure \ref{fig:original}, the false colors generated by marginal reconstruction are close to the original ones. In Figures \ref{fig:method} and \ref{fig:methods}, false colors create visible artifacts (in the background and the flower) due to averaging pixels with a similar brightness but a different chrominance. Figure \ref{fig:methods} has even more visible artifacts for different reasons: no false colors are created (since shapes are assigned with closest colors from the image) but it may reconstruct with a noise pixel value. Also, (pre)order based methods require a high filtering grain size to perform a good reconstruction (see the first column of Table 1), meaning that they cannot retain details in presence of noise. On the contrary, the distance-based approach preserves the details while preventing most color artifacts (Figure \ref{fig:methods}). Thus it is well adapted for fine objects retrieval.

Table 1. Min, max, and average PSNR scores of the methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Best Grain Size</th>
<th>Min PSNR</th>
<th>Max PSNR</th>
<th>Avg. PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preorder $L_{\text{La}^<em>b^</em>}$ (NC)</td>
<td>2</td>
<td>36.21</td>
<td>39.60</td>
<td>37.43</td>
</tr>
<tr>
<td>Preorder $L_{\text{La}^<em>b^</em>}$ (MP)</td>
<td>5</td>
<td>36.32</td>
<td>38.43</td>
<td>37.43</td>
</tr>
<tr>
<td>Preorder $L_{\text{HLS}}$ (NC)</td>
<td>7</td>
<td>36.34</td>
<td>39.36</td>
<td>37.57</td>
</tr>
<tr>
<td>Preorder Brightness (NC)</td>
<td>8.5</td>
<td>36.39</td>
<td>39.60</td>
<td>37.67</td>
</tr>
<tr>
<td>Order Lex</td>
<td>100</td>
<td>36.42</td>
<td>47.78</td>
<td>38.10</td>
</tr>
<tr>
<td>Distance-Based</td>
<td>15</td>
<td>37.02</td>
<td>41.14</td>
<td>38.76</td>
</tr>
<tr>
<td>Preorder $L_{\text{La}^<em>b^</em>}$ (MP)</td>
<td>15</td>
<td>36.31</td>
<td>41.17</td>
<td>38.92</td>
</tr>
<tr>
<td>Preorder $L_{\text{La}^<em>b^</em>}$ ($P_{\text{mean}}$)</td>
<td>200</td>
<td>36.20</td>
<td>41.45</td>
<td>39.03</td>
</tr>
<tr>
<td>Preorder $L_{\text{La}^<em>b^</em>}$ (MP)</td>
<td>300</td>
<td>36.77</td>
<td>41.38</td>
<td>39.34</td>
</tr>
<tr>
<td>Preorder $L_{\text{HLS}}$ ($P_{\text{mean}}$)</td>
<td>300</td>
<td>36.65</td>
<td>41.49</td>
<td>39.59</td>
</tr>
<tr>
<td>Preorder Brightness (MP)</td>
<td>7</td>
<td>36.34</td>
<td>39.36</td>
<td>37.57</td>
</tr>
<tr>
<td>Preorder Brightness ($P_{\text{mean}}$)</td>
<td>5</td>
<td>36.32</td>
<td>38.43</td>
<td>37.43</td>
</tr>
<tr>
<td>Preorder Brightness (MP)</td>
<td>100</td>
<td>36.60</td>
<td>42.69</td>
<td>39.94</td>
</tr>
<tr>
<td>Preorder Brightness ($P_{\text{mean}}$)</td>
<td>200</td>
<td>36.41</td>
<td>43.19</td>
<td>40.40</td>
</tr>
<tr>
<td>Shape-Graph $\hat{G}$ (marginal)</td>
<td>50</td>
<td>37.59</td>
<td>50.21</td>
<td>41.60</td>
</tr>
</tbody>
</table>

6. EXPERIMENTS

We have presented three classes of methods to extend the tree of shapes to multivariate images that have been evaluated through a denoising contest. We have shown that classical approaches consisting in equipping the data with a total order or a total preorder perform well but highly depend on the restitution rules and suffer from color artifacts. We have also presented two new approaches: the first is an algorithmic re-definition of shapes based on distances, and the second is the component-graph extended to the tree of shapes. They both produced promising results in our evaluation that highlights their strong potential for more complex tasks in computer vision. In an incoming work, we will further investigate the capacities of these methods for object retrieval and simplification of images where chrominance matters by adapting tree of shapes processing algorithms to shape-graphs.

Fig. 4. Top 5 methods performing the best denoising on the flower image.

7. CONCLUSION

We have presented three classes of methods to extend the tree of shapes to multivariate images that have been evaluated through a denoising contest. We have shown that classical approaches consisting in equipping the data with a total order or a total preorder perform well but highly depend on the restitution rules and suffer from color artifacts. We have also presented two new approaches: the first is an algorithmic re-definition of shapes based on distances, and the second is the component-graph extended to the tree of shapes. They both produced promising results in our evaluation that highlights their strong potential for more complex tasks in computer vision. In an incoming work, we will further investigate the capacities of these methods for object retrieval and simplification of images where chrominance matters by adapting tree of shapes processing algorithms to shape-graphs.
References


