9th International Conference Interdisciplinarity in Engineering, INTER-ENG 2015, 8-9 October 2015, Tirgu-Mures, Romania

Fast Probabilistic Pseudo-Morphology for Noise Reduction in Color Images

Radu-Mihai Coliban\textsuperscript{a,*}, Mihai Ivanovici\textsuperscript{a}, Iuliu Szekely\textsuperscript{b}

\textsuperscript{a}Transilvania University, B-dul Eroilor 29, Brasov, Romania
\textsuperscript{b}Sapientia University, Str. Matei Corvin 4, Cluj-Napoca, Romania

Abstract

Mathematical morphology is a popular framework for non-linear image processing, first introduced for binary and gray-level images, then extended to color and multivariate images. Various pseudo-morphological frameworks have been proposed as solutions to the problem of ordering multivariate data. We propose an improvement of the existing color probabilistic pseudo-morphology by computing the pseudo-extrema of a color set in a faster way, leading to a smaller execution time. We show the usefulness of the new construction in the context of noise reduction in color images using an OCCO filter, by comparing our approach with a series of color morphologies and pseudo-morphologies.

\textsuperscript{c}© 2016 The Authors. Published by Elsevier Ltd.

Published by Elsevier Ltd. Peer-review under responsibility of the “Petru Maior” University of Tirgu-Mures, Faculty of Engineering.

Keywords: Color image processing; Mathematical morphology; Probabilistic pseudo-morphology; Color noise reduction; OCCO filter

1. Introduction

Mathematical morphology (MM)[1][2] is a popular framework in non-linear image processing. Grayscale MM is based on lattice theory, which implies that a partial ordering is imposed on the pixel values in order to compute the infimum and supremum of the data set, which define the two basic morphological operators, erosion and dilation [3]. The extension of MM to the color domain is not trivial, since there are a multitude of ways in which vector values can be ordered [4]. The ordering schemes are generally classified in four groups [5]: marginal, reduced, conditional and partial, each having its advantages and disadvantages, depending on the application they are used in.

\textit{Pseudo-morphologies} focus on directly computing the extrema of a given data set, thus avoiding the ordering of vector values. While not respecting all the theoretical properties of their morphological counterparts, pseudo-morphological operators can be of interest in various applications, such as texture classification, noise reduction or multispectral data processing [6] [7]. We propose a new version of the Probabilistic Pseudo-Morphology [8] which improves the computation time and simplifies the parameter setup process, while generating mostly \textit{better} results (with respect to several criteria) in noise reduction tasks.

\textsuperscript{*}Corresponding author. Tel.: +40-268-478705.
\textit{E-mail address:} coliban.radu@unitbv.ro
2. Fast Probabilistic Pseudo-Morphology

The original definition of Probabilistic Pseudo-Morphology (denoted OrigPPM) uses the Chebyshev inequality [9] in order to compute pseudo-extrema of color vector values. Let $\xi$ be a random variable with a positive mean $\mu_\xi$ and positive standard deviation $\sigma_\xi$; the Chebyshev inequality states that: $P(|\xi - \mu_\xi| \geq k\sigma_\xi) \leq \frac{1}{k^2}$. Using the $k$ parameter, one may generate symmetrical intervals $[\mu_\xi - k\sigma_\xi, \mu_\xi + k\sigma_\xi]$ around the mean value $\mu_\xi$, with the bounds being more or less close to the actual maximum or minimum values of the distribution. In the case of OrigPPM, Principal Component Analysis (PCA) is applied locally, and the two vector pseudo-extrema of the data set are computed on the first principal component, based on the Chebyshev inequality:

$$
\begin{align*}
\tilde{E}^- &= (\mu_{\text{PCA}1} - k\sigma_{\text{PCA}1}, 0, 0) \\
\tilde{E}^+ &= (\mu_{\text{PCA}1} + k\sigma_{\text{PCA}1}, 0, 0)
\end{align*}
$$

where $\mu_{\text{PCA}1}$ and $\sigma_{\text{PCA}1}$ represent the mean and the standard deviation of the first principal component of the color data. This approach embeds both a linear behavior, given by the mean estimation, and a non-linear behavior, given by the parameter $k$, which controls the error in extrema estimation. The probabilistic pseudo-extrema are obtained by representing $\tilde{E}^-$ and $\tilde{E}^+$ in the initial coordinate system, thus resulting $E^-$ and $E^+$. However, the sign associated with the pseudo-extrema in the PCA basis cannot reflect their order, because the PCA is performed through data rotations. In order to establish which of the two pseudo-extrema is associated with pseudo-erosion or pseudo-dilation, they are ordered using a dot product operator, according to three pairs of global reference colors: $\{R_0^-, R_0^+, R_1^-, R_1^+, R_2^-, R_2^+\}$. The references are positioned on perpendicular lines, thus forming an orthogonal coordinate system, and they may be set manually or automatically.

Therefore, given a color image $f$, with the support $D_f$, and a flat structuring element (SE) $g$, with the support $D_g$, the pseudo-erosion $\varepsilon_\xi(f)$ and pseudo-dilation $\delta_\xi(f)$ are defined for $x \in D_f$ as:

$$
[\varepsilon_\xi(f)](x) = \bigvee_{z \in D_g} f(x + z) \overset{\Delta}{=} \begin{cases} \\
\arg \min_i [\overline{R_0^- R_0^+} \cdot \overline{R_0^+ g}] & \text{if } \overline{R_0^- R_0^+} \cdot \overline{E^- E^+} = 0 \\
\arg \min_i [\overline{R_1^- R_1^+} \cdot \overline{R_1^+ g}] & \text{if } \overline{R_0^- R_0^+} \cdot \overline{E^- E^+} = 0
\end{cases}
$$

$$
[\delta_\xi(f)](x) = \bigvee_{z \in D_g} f(x - z) \overset{\Delta}{=} \begin{cases} \\
\arg \max_i [\overline{R_0^- R_0^+} \cdot \overline{R_0^+ g}] & \text{if } \overline{R_0^- R_0^+} \cdot \overline{E^- E^+} = 0 \\
\arg \max_i [\overline{R_2^- R_2^+} \cdot \overline{R_2^+ g}] & \text{if } \overline{R_0^- R_0^+} \cdot \overline{E^- E^+} = 0
\end{cases}
$$

where $i \in \{E^-, E^+\}$ and $E^-$ and $E^+$ are the pseudo-extrema of the local data given by $D_f \cap D_g$. The need for three pairs of reference coordinates in the construction is justified by the fact that orthogonality can occur between the local pseudo-extrema vector $E^- E^+$ and global reference vectors $\overline{R_0^- R_0^+}$ and $\overline{R_1^- R_1^+}$, making impossible the labeling of the two pseudo-extrema using the dot product.

We propose a new construction (denoted FastPPM), which uses a single pair of reference color vectors $\{R^-, R^+\}$, which determine the reference axis:

$$
R_{\text{axis}} = \frac{R^+ - R^-}{||R^+ - R^-||}
$$

The colors in the data set given by $D_f \cap D_g$ are projected on this axis, resulting in a set of scalar projection values $p = \{p_1, p_2, \ldots\}$; the projection of a local color $C_i$ is the dot product $p_i = C_i \cdot R_{\text{axis}}$.

The two pseudo-extrema are computed using Chebyshev’s inequality using the mean color value in RGB and the standard deviation of the projection values:

$$
\begin{align*}
E^- &= \mu_{\text{RGB}} - k \cdot \sigma_p \cdot R_{\text{axis}} \\
E^+ &= \mu_{\text{RGB}} + k \cdot \sigma_p \cdot R_{\text{axis}}
\end{align*}
$$
No further ordering of the pseudo-extrema is required, since by construction $E^-$ is always closer to $R^-$, while $E^+$ is closer to $R^+$; thus, pseudo-erosion and pseudo-dilation are defined as:

$$[\varepsilon_g(f)](x) = \bigwedge_{z \in D_y} f(x + z) \triangleq E^- \quad (6)$$

$$[\delta_g(f)](x) = \bigvee_{z \in D_y} f(x - z) \triangleq E^+ \quad (7)$$

Fig. 1(a)-(c) presents the test image Lena and the results for pseudo-dilation using OrigPPM and FastPPM, respectively. In both cases a $5 \times 5$ square-shaped SE was used, with the parameter $k = 1$, while the reference colors ($R^-_g$ and $R^+_g$ for OrigPPM; $R^-_g$ and $R^+_g$ for FastPPM) were chosen to be black ([0,0,0] in the RGB space) and white ([255,255,255]). It can be noticed that both results are smoother than the original image, highlighting the linear behavior embedded in the pseudo-morphological operators. Moreover, the two results appear to be very similar; the $35 \times 35$ crops in Fig. 1(d)-(e) show that there are differences in color in Lena’s hair - the FastPPM dilation is more purplish, compared to OrigPPM. Generally, the color differences are more pronounced on the edges of objects. These observations underline the fact that FastPPM is an approximation of OrigPPM; for the same values of the input parameters (i.e. reference colors, $k$, SE size), the results are visually similar but they are not identical.

From the complexity point of view, the PCA algorithm runs in $O(mn^2 + n^3)$ [10], while the dot product that FastPPM is based on runs in $O(mn)$, where $m$ is the number of vectors and $n$ the dimension of the vectors. The graph in Fig. 2 presents a comparison of the running time for a pseudo-dilation between our Matlab implementations of OrigPPM and FastPPM on the 256 $\times$ 256 version of Lena, as a function of the size of the structuring element; since we are dealing with color data, in this case $n = 3$ (the number of image channels), while $m$ is the square of the structuring element size (the number of pixels used for the computation of an output value). The theoretical speedup achieved by replacing PCA with the dot product operator is $3 + \frac{9}{m}$. From the graph, it can be noticed that FastPPM improves the computational speed by roughly a factor of 3, irrespective of the structuring element size. The execution speed is not affected by the actual content of the image. The tests were run on a Core2Duo E7500 machine.
3. Noise reduction

3.1. Methodology

In order to show the usefulness of the new construction in the context of an application, we use an evaluation methodology presented in [4], by which we make a comparison between morphological and pseudo-morphological frameworks using a noise reduction task with an open-close-close-open (OCCO) filter [11], defined as the pixelwise average between open-close and close-open. Morphological *opening* is defined as the dilation of the previously eroded image, $\gamma_g(f) = \delta_g(\varepsilon_g(f))$, while its dual is called *closing*: $\phi_g(f) = \varepsilon_g(\delta_g(f))$. Thus, the definition of the OCCO filter is:

$$OCCO_g(f) = \frac{1}{2} \phi_g(\gamma_g(f)) + \frac{1}{2} \gamma_g(\phi_g(f))$$  \hspace{1cm} (8)

We use images corrupted with zero-mean additive Gaussian noise, with $\sigma = 32$ and inter-channel correlation factor $\rho = 0$ (uncorrelated noise) and $\rho = 0.9$ (correlated noise). Figures 3 and 4 depict $100 \times 100$ crops of the test images (the classical test image Lena and a medical image consisting of cervical cells - Pap Smear), along with their versions corrupted with Gaussian noise.

We first compare the new approach with OrigPPM in order to highlight the behavior of the new operators as a function of $k$, then we proceed to a more thorough study using various unanimously-accepted color morphological and pseudo-morphological frameworks. For both probabilistic approaches, the reference colors ($R^-_0$ and $R^+_0$ for OrigPPM; $R^-_o$ and $R^+_o$ for FastPPM) were chosen to be black and white, based on the observation that when correlated noise is added to images, the resulting color distribution will be oriented on the black-white axis, while uncorrelated noise has no definite direction. The SE used was square-shaped, of size $3 \times 3$.

Besides the visual assessment of results, we employ three widely-used error measures in order to evaluate the performances of the different frameworks [4][12]: the Mean Average Error (MAE), the Normalized Mean Square Error (NMSE) and the perceptual $\Delta_Lab$ error.

The definitions for MAE and NMSE, measuring image quality preservation and noise suppression, are as follows:

$$MAE = \frac{1}{3MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \sum_{i=1}^{3} |f_i(x,y) - f'_i(x,y)|$$  \hspace{1cm} (9)

$$NMSE = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} ||f(x,y) - f'(x,y)||^2}{\sum_{x=1}^{M} \sum_{y=1}^{N} ||f(x,y)||^2}$$  \hspace{1cm} (10)
where \( f(x, y) = [f_1(x, y), f_2(x, y), f_3(x, y)] \) and \( f'(x, y) = [f'_1(x, y), f'_2(x, y), f'_3(x, y)] \) are the original and the processed pixel values in the images of size \( M \times N \), represented in the RGB color space.

The \( \Delta_{Lab} \) error is a measure of the perceptual closeness between the original and the filtered image, defined as:

\[
\Delta_{Lab} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} ||\bar{f}(x, y) - \bar{f}'(x, y)||
\]

(11)

where \( \bar{f} \) and \( \bar{f}' \) are the CIELab representations of the original and OCCO processed image, respectively.

### 3.2. Comparison against OrigPPM

Figures 5 - 8 present a comparison between filtering results obtained by OrigPPM and FastPPM for three values of the \( k \) parameter: 0.5, 1, and 1.5, in order to highlight the transition between the linear and non-linear behavior of the pseudo-morphological operators. In the case of uncorrelated noise (Figures 5 and 6), for \( k = 0.5 \) the results are similar between the two approaches and a pronounced blurring effect is noticeable. For \( k = 1 \), the blurring effect is less evident and some small distortions appear in the images filtered with OrigPPM, which become extremely noticeable in the case of \( k = 1.5 \), to the point of rendering the result unusable. The results obtained with FastPPM do not show such distortions, thus obtaining good filtering results from a human perception point of view, with a low blurring effect.

The explanation for the distortions in the OrigPPM results lies in the fact that in the original construction the pseudo-extrema are always chosen on the PCA axis with the largest variance; thus, for large values of \( k \), they are far from the local mean. There are situations when the pseudo-extrema may be situated outside the RGB color space, leading to saturation on one or more channels, which distorts the results in the iterating filtering process. In the case of FastPPM, the pseudo-extrema are not tied to the direction of maximum variance, thus being generally closer to the mean value of the local color distribution. Consequently, the phenomenon of computing pseudo-extrema outside the RGB cube is less frequent for large values of \( k \), leading to less distortions.

Fig. 3. Lena, 100 \times 100 crops - original and corrupted with Gaussian noise.

Fig. 4. Pap Smear, 100 \times 100 crops - original and corrupted with Gaussian noise.
Fig. 5. Filtering results for Lena with uncorrelated noise, using OrigPPM (top row) and FastPPM (bottom row).

Fig. 6. Filtering results for Pap Smear with uncorrelated noise, using OrigPPM (top row) and FastPPM (bottom row).

Fig. 7. Filtering results for Lena with correlated noise, using OrigPPM (top row) and FastPPM (bottom row).
For correlated noise (Figures 7 and 8), the results are relatively similar between the two frameworks. This happens because, after the addition of correlated noise, the color distribution of the resulting image tends to be orientated on the black-white direction (which is also the chosen reference color axis for both frameworks).

3.3. Comparison against other morphological and pseudo-morphological frameworks

In this section we present a comparison between our proposed approach (FastPPM), the original probabilistic approach (OrigPPM), two classic color morphology approaches (marginal and lexicographic [4], implemented in both the RGB and CIELab color spaces - denoted MargRGB, MargLab, LexRGB and LexLab) and two more recent approaches, a group-invariant morphological approach based on marginal ordering (GroupInv) [13] and a pseudo-morphological approach based on lexicographic ordering ($\alpha$-trimmed) [6]. In order to illustrate the performances of the methods under comparison using the three objective measures presented in section 3.1, the error values obtained after filtering the two test images are presented in Tables 1 - 4. The error values for the noisy images are also given as reference - denoted as Noisy.

For both probabilistic frameworks there are three different optimal values of the $k$ parameter, each one minimizing one of the errors under consideration, determined by the image content and type of noise added. Depending on the application, the value of the parameter can be chosen in order to either minimize one of the errors (MAE for image quality preservation, NMSE for noise suppression or $\Delta_{Lab}$ for perceptual closeness to the original) or a trade-off between the three. For the comparison presented in this section, we have chosen to use the median of the three values, which is presented in parentheses in the tables, ensuring that at least one error will be minimized, while generating acceptable results for the other criteria.

A general observation to be made is that the approaches based on marginal ordering (MargRGB, MargLab, GroupInv) generally perform better than their lexicographic-based counterparts (LexRGB, LexLab, $\alpha$-trimmed), and their performances are similar regardless of the type of noise they are used against. In all situations, the more recently developed approaches (GroupInv and $\alpha$-trimmed) obtain lower errors in comparison to the classical approaches they are based on (marginal and lexicographic, respectively).

The error tables show that both PPM approaches prove their filtering superiority for the images under consideration with respect to all three error criteria, regardless of the type of noise used. In the case of uncorrelated noise (Tables 1 and 2), when the behavior is closer to that of a non-linear filter (higher optimal $k$ value), FastPPM shows improvements with respect to all error criteria over OrigPPM. In the case of correlated noise (Tables 3 and 4), OrigPPM obtains slightly better performances in the case of the Lena Image (MAE, $\Delta_{Lab}$), while for the Pap Smear the differences in performance are rather small. Thus, the proposed framework proves its usefulness in the context of OCCO filtering by obtaining consistently good results against both types of noise.
Table 1. Errors for Lena against uncorrelated noise.

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>NMSE ×10^2</th>
<th>Δ_ΔLab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>24.76</td>
<td>4.81</td>
<td>24.78</td>
</tr>
<tr>
<td>MargRGB</td>
<td>9.70</td>
<td>0.89</td>
<td>8.45</td>
</tr>
<tr>
<td>MargLab</td>
<td>9.98</td>
<td>0.93</td>
<td>8.20</td>
</tr>
<tr>
<td>LexRGB</td>
<td>18.14</td>
<td>3.90</td>
<td>17.59</td>
</tr>
<tr>
<td>LexLab</td>
<td>18.09</td>
<td>3.68</td>
<td>17.27</td>
</tr>
<tr>
<td>GroupInv</td>
<td>9.07</td>
<td>0.81</td>
<td>7.67</td>
</tr>
<tr>
<td>α-trimmed</td>
<td>15.68</td>
<td>2.14</td>
<td>16.23</td>
</tr>
<tr>
<td>OrigPPM (k=0.4)</td>
<td>8.95</td>
<td>0.86</td>
<td>6.69</td>
</tr>
<tr>
<td>FastPPM (k=1.3)</td>
<td>7.91</td>
<td>0.64</td>
<td>6.28</td>
</tr>
</tbody>
</table>

Table 2. Errors for Pap Smear against uncorrelated noise.

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>NMSE ×10^2</th>
<th>Δ_ΔLab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>22.18</td>
<td>2.16</td>
<td>21.17</td>
</tr>
<tr>
<td>MargRGB</td>
<td>8.75</td>
<td>0.29</td>
<td>6.94</td>
</tr>
<tr>
<td>MargLab</td>
<td>9.18</td>
<td>0.33</td>
<td>6.79</td>
</tr>
<tr>
<td>LexRGB</td>
<td>17.54</td>
<td>2.94</td>
<td>15.68</td>
</tr>
<tr>
<td>LexLab</td>
<td>17.60</td>
<td>2.91</td>
<td>14.38</td>
</tr>
<tr>
<td>GroupInv</td>
<td>8.23</td>
<td>0.27</td>
<td>6.24</td>
</tr>
<tr>
<td>α-trimmed</td>
<td>14.17</td>
<td>0.92</td>
<td>14.29</td>
</tr>
<tr>
<td>OrigPPM (k=0.1)</td>
<td>7.57</td>
<td>0.22</td>
<td>5.13</td>
</tr>
<tr>
<td>FastPPM (k=1.2)</td>
<td>7.15</td>
<td>0.19</td>
<td>5.02</td>
</tr>
</tbody>
</table>

Table 3. Errors for Lena against correlated noise.

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>NMSE ×10^2</th>
<th>Δ_ΔLab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>24.68</td>
<td>4.78</td>
<td>12.77</td>
</tr>
<tr>
<td>MargRGB</td>
<td>9.53</td>
<td>0.68</td>
<td>6.11</td>
</tr>
<tr>
<td>MargLab</td>
<td>9.65</td>
<td>0.90</td>
<td>5.32</td>
</tr>
<tr>
<td>LexRGB</td>
<td>13.32</td>
<td>2.69</td>
<td>9.06</td>
</tr>
<tr>
<td>LexLab</td>
<td>12.73</td>
<td>2.55</td>
<td>8.71</td>
</tr>
<tr>
<td>GroupInv</td>
<td>9.42</td>
<td>0.86</td>
<td>5.88</td>
</tr>
<tr>
<td>α-trimmed</td>
<td>11.82</td>
<td>1.32</td>
<td>7.56</td>
</tr>
<tr>
<td>OrigPPM (k=1.0)</td>
<td>8.31</td>
<td>0.72</td>
<td>4.70</td>
</tr>
<tr>
<td>FastPPM (k=1.1)</td>
<td>8.38</td>
<td>0.72</td>
<td>4.94</td>
</tr>
</tbody>
</table>

Table 4. Errors for Pap Smear against correlated noise.

<table>
<thead>
<tr>
<th>Method</th>
<th>MAE</th>
<th>NMSE ×10^2</th>
<th>Δ_ΔLab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noisy</td>
<td>22.19</td>
<td>2.16</td>
<td>10.85</td>
</tr>
<tr>
<td>MargRGB</td>
<td>8.72</td>
<td>0.29</td>
<td>5.23</td>
</tr>
<tr>
<td>MargLab</td>
<td>8.88</td>
<td>0.30</td>
<td>4.46</td>
</tr>
<tr>
<td>LexRGB</td>
<td>13.69</td>
<td>2.47</td>
<td>8.34</td>
</tr>
<tr>
<td>LexLab</td>
<td>13.33</td>
<td>2.44</td>
<td>8.06</td>
</tr>
<tr>
<td>GroupInv</td>
<td>8.61</td>
<td>0.29</td>
<td>5.01</td>
</tr>
<tr>
<td>α-trimmed</td>
<td>10.31</td>
<td>0.78</td>
<td>6.45</td>
</tr>
<tr>
<td>OrigPPM (k=0.8)</td>
<td>7.12</td>
<td>0.19</td>
<td>3.81</td>
</tr>
<tr>
<td>FastPPM (k=1.0)</td>
<td>6.99</td>
<td>0.18</td>
<td>3.82</td>
</tr>
</tbody>
</table>

4. Conclusions

In this paper we propose a new definition for Probabilistic Pseudo-Morphology. The new construction improves the computational speed with respect to the original framework, by computing the local pseudo-extrema using the dot product instead of using Principal Component Analysis. The new framework also simplifies the parameter specification process, by using a single reference axis instead of an orthogonal axis system. In the context of color noise reduction, the new construction achieves less distortion in the case of non-linear behavior with respect to the original approach, and also proves to have superior filtering capabilities when compared to a series of relevant morphological and pseudo-morphological approaches.

Acknowledgement

The research was partially funded by the PN-II-PT-PCCA-2013-4-0202 project - Intelligent System for Automatic Assistance of Cervical Cancer Diagnosis, contract no. 7/2014, financed by UEFISCDI, 2014-2016.

References