Pattern Recognition 47 (2014) 721–735

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Probabilistic pseudo-morphology for grayscale and color images ☆

Alexandru Căliman [a,*], Mihai Ivanovici [a], Noël Richard [b]

[a] MIV² Laboratory, Department of Electronics and Computers, Transilvania University, Braşov 500024, România
[b] XLIM-SIC UMR CNRS 6172, Signals, Images and Communications Laboratory, University of Poitiers, France

ARTICLE INFO

Article history:
Received 1 March 2013
Received in revised form
17 July 2013
Accepted 23 August 2013
Available online 4 September 2013

Keywords:
Mathematical morphology
Fractal dimension
Texture analysis

ABSTRACT

Mathematical morphology offers popular image processing tools, successfully used for binary and grayscale images. Recently, its extension to color images has become of interest and several approaches were proposed. Due to various issues arising from the vectorial nature of the data, none of them imposed as a generally valid solution. We propose a probabilistic pseudo-morphological approach, by estimating two pseudo-extrema based on Chebyshev inequality. The framework embeds a parameter which allows controlling the linear versus non-linear behavior of the probabilistic pseudo-morphological operators. We compare our approach for grayscale images with the classical morphology and we emphasize the impact of this parameter on the results. Then, we extend the approach to color images, using principal component analysis. As validation criteria, we use the estimation of the color fractal dimension, color textured image segmentation and color texture classification. Furthermore, we compare our proposed method against two widely used approaches, one morphological and one pseudo-morphological.

1. Introduction

The mathematical morphology (MM) was founded by Matheron [1] and Serra [2] and became very popular in non-linear image processing. Initially, the MM had been introduced as a processing technique for binary images, which were regarded as sets; therefore, its elementary operations are based on the set theory [3]. However, the extension to sets of grayscale images, using the *umbra* concept [4,5], introduced a generalization of the basic morphological operations which were subsequently used in many image processing and analysis methods e.g. morphological filtering [6], watershed segmentation [7], etc. The grayscale morphology is based on the lattice theory, which implies a partial ordering of the data within the grayscale images. In this case, the lattice structure is not difficult to obtain, since the grayscale images are real functions and the set of real numbers implicitly possesses a lattice structure. However, while the extension from binary to grayscale images is a natural one, the extension to color or multi-spectral images is not straightforward, because of the vectorial nature of the data and the difficulty in finding a suitable ordering for it. Barnett introduced four types of vector orderings: marginal, reduced, conditional and partial [8]. When applied to color data, all these orderings have certain disadvantages, depending on application. For instance, the marginal ordering introduces false colors and the conditional ordering generates visual non-linearities from the human perception point of view [9]; the reduced and partial orderings are either relying on pre-orderings, thus lacking the anti-symmetry property, or behave like conditional orderings, generating perceptual non-linearities. In fact, the difficulty of extending MM to multivariate data does not consist in obtaining an ordering, but in obtaining a pertinent ordering from the human visual system point of view. There have been proposed a plethora of methods for color and multivariate MM, but few of them, only recently, referred to this linearity problem [10,11]. The paper written by Aptoula and Lefèvre in 2007 [12], which includes more than 70 references to different color MM methods, represents a relatively recent and comprehensive state-of-the-art in this field. Nevertheless, other approaches have been introduced recently. For instance, in [13] a method using the color data distribution in a partial ordering based on depth functions is presented, [14] proposes a graph-based approach using the Laplacian eigenmaps as a method for non-linear dimensionality reduction, thus resulting a reduced ordering, while[15] proposes a geometrical method based on the so-called Loewner order. Recently, there has also been a great interest in supervised methods for establishing orderings among vectorial data [16,17].

In addition to the proposed morphological approaches whose operations respect all the mathematical properties of the classical MM, there were also proposed pseudo-morphologies, which do not require an underlying order among the image data, focusing on computing directly the minimum and maximum of a given set [18,19]. This kind of approaches do not require a complete lattice

☆This paper is partially supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), ID76945 financed from the European Social Fund and by the Romanian Government.
* Corresponding author. Tel.: +40 268475751.
E-mail addresses: alexandru.caliman@gmail.com (A. Căliman), mihai.ivanovici@unitbv.ro (M. Ivanovici), noel.richard@univ-poitiers.fr (N. Richard).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.08.021

structure, consequently lacking a binary relation that is reflexive, anti-symmetric and transitive and thus, they cannot be theoretically considered *morphological*. However, they could be of practical interest in noise reduction, texture classification or multispectral remote sensing data processing [20,21].

In this paper we propose a probabilistic pseudo-morphology (PPM) approach based on the Chebyshev inequality, in which we estimate the two extrema (the *infimum* and the *supremum*) of a set instead of defining an underlying ordering for the entire data set. PPM is based on the choice of one parameter, which is capable of turning it either into a linear or non-linear approach (see details in Section 3). By using the statistical moments (both the mean and the variance) of the local data distribution, the proposed method is less influenced by the presence of noise. Then, we propose an extension to color or multivariate images. This extension has the advantages that it is full-vectorial and it generates adjustable operations i.e. linear versus non-linear vector filtering. Despite the fact that it introduces *false colors*, it is useful in various applications like morphological edge detection or texture description based on morphological operations. We demonstrate the usefulness of our approach in the context of color texture complexity estimation, textured image segmentation and color texture classification.

The next sections of the paper are organized as follows: Section 2 presents the general concepts and definitions for MM and Chebyshev inequality, on which our approach is based; Section 3 defines the probabilistic pseudo-morphology for grayscale images while in Section 4 the extension to multivariate images is presented; the paper is ended with a section of discussions and one of conclusions.

2. Theoretical notions

2.1. General aspects about MM

In this section we briefly present the general concepts and definitions of MM's basic operations, using the notations from [22]. Table 1 presents all the notations we use and propose within this article.

The erosion and the dilation, the fundamental operations of MM, are defined within a complete lattice as the operations which distribute over the infima and the suprema [3]. However, these definitions are not suitable for a practical implementation. The more popular definitions, which are often used in implementations, are based on the concept of *structuring element* (SE), which is a relatively small set used for probing the image f, which is analyzed. An origin is associated with any SE g, within its definition domain $\mathcal{D}_g$. This origin helps positioning the SE at every coordinate within the initial image definition domain $\mathcal{D}_f$ i.e. at every pixel coordinate x. In order to avoid mixing the spatial units of $\mathcal{D}_f$ with the pixel values $C_x \in \mathcal{S}_{\mathcal{D}f}$ ($\mathcal{S}_{\mathcal{D}f} \subset \mathbb{R}$ for grayscale images or $\mathcal{S}_{\mathcal{D}f} \subset \mathbb{R}^n$ for multivariate images) and because of the fact that there is no pertinent meaning of adding two multivariate image data, it is very common to use *flat* SEs, which are defined only through their origin and their shape, given by $\mathcal{D}_g$. Using these notations, the erosion and the dilation of an image *f*, using a *flat* SE *g*, are defined as follows [23]:

Table 1
Notations.

f, $\mathcal{D}_f$	Image function and its support
$\mathcal{S}_{\mathcal{D}f}$	The f function's range of values
$\widetilde{\mathcal{S}_{\mathcal{D}f}}$	The f function's codomain expressed in PCA basis
$x=(i,j)$	Spatial coordinates for the pixel at line i and column j
$C_x, \widetilde{C_x}$	Grayscale, color or multivariate coordinates of the x pixel, expressed in the initial or PCA space
g, $\mathcal{D}_g$	Structuring element and its spatial domain of definition
$[\delta_g(f)](x)$, $[\varepsilon_g(f)](x)$	Dilation and erosion of image f, using the structuring element g, computed for pixel x
ξ, μ_ξ, σ_ξ	A random variable, its mean value and standard deviation
$\mathfrak{E}_\alpha$, $\mathfrak{E}_\beta$	Probabilistic pseudo-extrema
$(\mathfrak{R}_i^+, \mathfrak{R}_i^-)$	The ith pair of global references

$$[\varepsilon_g(f)](x) = \bigwedge_{z \in \mathcal{D}_g} f(x+z), \quad \forall x \in \mathcal{D}_f \tag{1}$$

$$[\delta_g(f)](x) = \bigvee_{z \in \mathcal{D}_g} f(x-z), \quad \forall x \in \mathcal{D}_f \tag{2}$$

where $\bigwedge$ and $\bigvee$ are the *infimum* and *supremum* operators. It can be noticed that the basic morphological operations involve finding an infimum and a supremum for the points within a local region, given by the SE positioning. The extension of these operations to multivariate images is not straightforward, due to the difficulty of defining such extrema for vectorial data.

2.2. The Chebyshev inequality

Our approach is based on the Chebyshev inequality (3), which expresses the upper bound of the probability that any random variable ξ takes values farther from its average value, outside of a specified interval [24]. The inequality stands for any distribution, as long as the mean and the standard deviation are finite [25]. Let ξ be a random variable with the mean μ_ξ and the standard deviation σ_ξ; the Chebyshev inequality states that

$$P\{|\xi-\mu_\xi| \geq k\sigma_\xi\} \leq \frac{1}{k^2}. \tag{3}$$

Using the k parameter, one may generate symmetrical intervals around the mean value, delimited by bounds which are more or less closed to the real maximum or minimum values of the distribution. The bounds of this confidence interval are given by (4). We define these bounds as the probabilistic pseudo-extrema, $\mathcal{E}^+$ and $\mathcal{E}^-$, in the sense of the Chebyshev inequality:

$$\begin{cases} \mathcal{E}^+ & \triangleq \mu_\xi + k\sigma_\xi \\ \mathcal{E}^- & \triangleq \mu_\xi - k\sigma_\xi \end{cases} \tag{4}$$

Using an appropriate k value, the probabilistic extrema and the real maximum and minimum values may be more or less closed to each other, but only for symmetrical distributions there is a unique k for which they coincide.

3. A direct application for grayscale images

Our first proposal aims at the direct application of the previously described notions to grayscale images. Thus, we consider an image $f: \mathcal{D}_f \to \mathcal{S}_{\mathcal{D}f}$, with $\mathcal{S}_{\mathcal{D}f} \subset \mathbb{R}$. The histogram of any local neighborhood of the image is an estimate of the probability density function associated with the pixel data. In this case, the Chebyshev inequality defines an interval depending on the k parameter and the standard deviation of the pixel values. The error between the real maximum and minimum of the local data set and the estimated pseudo-extrema based on the Chebyshev's inequality is a function of k. We define the grayscale pseudo-dilation and pseudo-erosion operations as

$$[\varepsilon_g(f)](x) = \bigwedge_{z \in \mathcal{D}_g} f(x+z) \triangleq \mu_\xi - k\sigma_\xi, \quad \forall x \in \mathcal{D}_f \tag{5}$$

$$[\delta_g(f)](x) = \bigvee_{z \in \mathcal{D}_g} f(x-z) \triangleq \mu_\xi + k\sigma_\xi, \quad \forall x \in \mathcal{D}_f \tag{6}$$