

UNICAMP

UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

MATEUS SANGALLI

**Fuzzy Value-Based Supervised Ordering for
Multivalued Mathematical Morphology**

**Ordem Supervisionada Baseada em Valores
Fuzzy para Morfologia Matemática
Multivalorada**

Campinas

2019

Mateus Sangalli

Fuzzy Value-Based Supervised Ordering for Multivalued Mathematical Morphology

Ordem Supervisionada Baseada em Valores Fuzzy para Morfologia Matemática Multivalorada

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática Aplicada.

Dissertation presented to the Institute of Mathematics, Statistics and Scientific Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Applied Mathematics.

Supervisor: Marcos Eduardo Ribeiro do Valle Mesquita

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Mateus Sangalli e orientada pelo Prof. Dr. Marcos Eduardo Ribeiro do Valle Mesquita.

Campinas

2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

Sa58f Sangalli, Mateus, 1996-
Fuzzy value-based supervised ordering for multivalued mathematical morphology / Mateus Sangalli. – Campinas, SP : [s.n.], 2019.

Orientador: Marcos Eduardo Ribeiro do Valle Mesquita.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Morfologia matemática. 2. Imagem hiperespectral. 3. Aprendizagem supervisionada (Aprendizado do computador). 4. Lógica fuzzy. I. Mesquita, Marcos Eduardo Ribeiro do Valle, 1979-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Ordem supervisionada baseada em valores fuzzy para morfologia matemática multivalorada

Palavras-chave em inglês:

Mathematical morphology

Hyperspectral imaging

Supervised learning (Machine learning)

Fuzzy logic

Área de concentração: Matemática Aplicada

Titulação: Mestre em Matemática Aplicada

Banca examinadora:

Marcos Eduardo Ribeiro do Valle Mesquita [Orientador]

Peter Sussner

Roberto Hirata Junior

Data de defesa: 14-08-2019

Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: 0000-0003-0317-1932

- Currículo Lattes do autor: <http://lattes.cnpq.br/2403531762165180>

**Dissertação de Mestrado defendida em 14 de agosto de 2019 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). MARCOS EDUARDO RIBEIRO DO VALLE MESQUITA

Prof(a). Dr(a). PETER SUSSNER

Prof(a). Dr(a). ROBERTO HIRATA JUNIOR

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Acknowledgements

I would like to thank the University of Campinas and the Institute of Mathematics, Statistics and Scientific Computing for the opportunity to pursue my research interests. This work was done with the support of CNPq, National Council of Research and Scientific and Technological Development - Brazil, under the process number 131635/2018-2.

I would like to thank my friends and family for their support during my studies.

I want to express my gratitude towards Professor Marcos Valle, my supervisor, for his guidance during this research project. I also want to thank Professor Peter Sussner and Professor João Florindo for their helpful advice and constructive criticism regarding this project.

Resumo

Morfologia Matemática foi concebida como uma ferramenta para a análise e processamento de imagens binárias e foi subsequentemente generalizada para o uso em imagens em tons de cinza e imagens multivaloradas. Reticulados completos, que são conjuntos parcialmente ordenados em que todo subconjunto tem extremos bem definidos, servem como a base matemática para uma definição geral de morfologia matemática. Em contraste a imagens em tons de cinza, imagens multivaloradas não possuem uma ordem não-ambígua. Essa dissertação trata das chamadas ordens reduzidas para imagens multivaloradas. Ordens reduzidas são definidas por meio de uma relação binária que ordena os elementos de acordo com uma função h do conjunto de valores em um reticulado completo. Ordens reduzidas podem ser classificadas em ordens não-supervisionadas e ordens supervisionadas. Numa ordem supervisionada, o função de ordenação h depende de conjuntos de treinamento de valores de *foreground* e de *background*. Nesta dissertação, estudamos ordens supervisionadas da literatura. Também propomos uma ordem supervisionada baseada em valores *fuzzy*. Valores *fuzzy* generalizam cores *fuzzy* - conjuntos *fuzzy* que modelam o modo que humanos percebem as cores - para imagens multivaloradas. Em particular, revemos como construir o mapa de ordenação baseado em conjuntos *fuzzy* para o *foreground* e para o *background*. Também introduzimos uma função de pertinência baseada numa estrutura *neuro-fuzzy* e generalizamos a função de pertinência baseada no diagrama de Voronoi. Por fim, as ordens supervisionadas são avaliadas num experimento de segmentação de imagens hiperespectrais baseado num perfil morfológico modificado.

Palavras-chave: Morfologia Matemática, Conjuntos Fuzzy, Ordens Supervisionadas, Imagens Hiperespectrais

Abstract

Mathematical morphology has been conceived initially as a tool for the analysis and processing of binary images and has been later generalized to grayscale and multivalued images. Complete lattices, which are partially ordered sets in whose every subset has well defined extrema, serve as the mathematical background for a general definition of mathematical morphology. In contrast to gray-scale images, however, there is no unambiguous ordering for multivalued images. This dissertation addresses the so-called reduced orderings for multi-valued images. Reduced orderings are defined by means of a binary relation which ranks elements according to a mapping h from the value set into a complete lattice. Reduced orderings can be classified as unsupervised and supervised ordering. In a supervised ordering, the mapping h depends on training sets of foreground and background values. In this dissertation, we study some relevant supervised orderings from the literature. We also propose a supervised ordering based on fuzzy values. Fuzzy values are a generalization of fuzzy colors - fuzzy sets that model how humans perceive colors - to multivalued images other than color images. In particular, we review how to construct the fuzzy ordering mapping based on fuzzy sets that model the foreground and the background. Also, we introduce a membership function based on a neuro-fuzzy framework and generalize the membership function based on Voronoi diagrams. The supervised orderings are evaluated in an experiment of hyperspectral image segmentation based on a modified morphological profile.

Keywords: Mathematical Morphology, Fuzzy Sets, Supervised Ordering, Hyperspectral Images.

Contents

	Introduction	10
1	LATTICE THEORY AND REDUCED ORDERINGS	13
1.1	Lattices and Complete Lattices	13
1.2	Pre-Orderings and Reduced Orderings	14
2	MATHEMATICAL TOOLS	17
2.1	Support Vector Machines	17
2.2	Lattice Auto Associative Memories	19
2.3	Fuzzy Systems	20
2.3.1	Fuzzy Sets	20
2.3.2	Fuzzy Logic Connectives	21
2.3.3	Adaptive Neuro Fuzzy Inference System	25
3	MATHEMATICAL MORPHOLOGY	27
3.1	Erosion and Dilation	27
3.2	Opening and Closing	31
3.3	The Marginal and Lexicographical Approaches	33
4	MATHEMATICAL MORPHOLOGY ON REDUCED ORDERINGS	37
4.1	Mathematical Morphology on Reduced Orderings	37
4.2	Supervised Orderings	39
4.2.1	The Distance-Based Ordering	40
4.2.2	The SVM-Based Ordering	41
4.2.3	The LAAM-Based Ordering	42
4.2.4	Comparing Supervised Orderings	43
5	FUZZY VALUE-BASED MATHEMATICAL MORPHOLOGY	46
5.1	Ordering Mappings Based on Membership Functions	46
5.2	Neuro-Fuzzy Approach	50
5.3	Voronoi Diagram Fuzzy Partition	51
5.4	Comparing the Fuzzy Set-Based Orderings	54
6	CLASSIFICATION OF HYPERSPECTRAL IMAGES	57
6.1	Morphological Profile	57
6.2	Computational Experiment	58
6.2.1	Data	58

6.2.2	Experiment	59
6.2.3	Results	62
6.2.4	Discussion	63
7	CONCLUDING REMARKS	67
	Bibliography	69

Introduction

Mathematical morphology (MM) is a powerful non-linear image processing framework based on geometrical and topological concepts (Heijmans, 1995; Soille, 1999). Applications of MM include edge detection, segmentation, automatic image reconstruction, pattern recognition and image decomposition (Braga-Neto and Goutsias, 2004; Gonzalez-Hidalgo et al., 2015; Rittner et al., 2013; Serra, 2006).

The first morphological operators have been developed by Matheron and Serra in the 1960s for the analysis of binary images. Later, binary MM operators have been successfully generalized to deal with gray-scale images (Sternberg, 1986). Some gray-scale morphological operators were also developed using concepts from fuzzy logic and fuzzy set theory (Bloch, 2011; De Baets, 1997; Nachtgaeel and Kerre, 2001; Sussner and Valle, 2008).

Morphological operators are very well defined on complete lattices (Heijmans, 1995; Ronse, 1990). A complete lattice \mathbb{L} is a partially ordered non-empty set in which any subset admits both a supremum and an infimum (Birkhoff, 1993; Grätzer et al., 2003). Since the only requirement is a partial order with well-defined extrema operations, complete lattices allowed the development of morphological operators for multivalued data, including color and hyperspectral images (Aptoula and Lefèvre, 2007; Lézoray, 2016; Angulo, 2007). In contrast to gray-scale approaches, however, there is no natural ordering for vectors. Hence, most researches on multivalued MM consist on finding an appropriate ordering scheme for a given multivalued image processing task. The interested reader can find a detailed discussion on many approaches to multivalued MM, including color MM, on (Aptoula and Lefèvre, 2007; Angulo, 2007).

Among the many partial orderings used on color and hyperspectral MM, total orderings have been widely used because they avoid the appearance of *false values* (Aptoula and Lefèvre, 2008; Serra, 2009). For example, Hanbury and Serra (2002) introduced a conditional ordering on the CIELab space to color MM. Although total orderings avoid the appearance of *false values*, they are usually irregular in a metric space (Chevallier and Angulo, 2016). Specifically, Chevallier and Angulo showed that under mild conditions there always exist vectors $\mathbf{x}, \mathbf{y}, \mathbf{z}$ such that $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z}$ but $d(\mathbf{x}, \mathbf{z}) < d(\mathbf{x}, \mathbf{y})$, where d denotes a metric and “ \leq ” is a total order. In words, \mathbf{z} is more similar (or it is closer) to \mathbf{x} than \mathbf{y} in spite of the the inequalities $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z}$. Like false values, the irregularity issue may be a problem in some vector-valued image processing tasks.

Reduced orderings rank the values based on a surjective, and often real-valued, function h (Barnett, 1976; Goutsias et al., 1995). Reduced ordering can be divided into

unsupervised and supervised approaches. In the former, the mapping h depends only on the image being processed. For example, Louverdis and Andreadis (2004) proposed a reduced ordering in which colors are ranked using fuzzy IF-THEN rules. Also, Velasco-Forero and Angulo (2012) proposed a reduced ordering scheme using statistical depth functions. Another promising unsupervised ordering scheme, in which the surjective mapping is constructed from the values of an image, have been proposed by Lézoray (2016).

In contrast to the unsupervised reduced ordering schemes, supervised orderings are defined using a set of value references. For example, Sartor and Weeks (2001) proposed a reduced ordering scheme based on the distance to a reference color. Ordering schemes based on distances have also been investigated by many other researchers, including Al-Otum (2015); Angulo (2007); Comer and Delp (1999); Deborah et al. (2015); Ledoux et al. (2015); Valle and Valente (2016). The supervised ordering proposed by Velasco-Forero and Angulo (2011a), in which the surjective mapping is determined using support vector machines (SVMs) generalizes many of the distance-based approaches. Velasco-Forero and Angulo (2010) also proposed the ordering based on kriging, which coincides with the SVM-based ordering when there are only one foreground and one background color. Graña and Chyzhyk (2016) defined a supervised ordering based on lattice auto-associative memories, which computes the distance between the input and its recalled vector, that also generalizes the distance-based ordering.

Recently, we proposed a fuzzy-proposition-based reduced ordering (Sangalli and Valle, 2018) using fuzzy logic to model the subjective nature of colors. Precisely, the ordering mapping h_{FUZZY} is the membership function of a fuzzy set that is based on two other fuzzy sets; one corresponding to the foreground and the other representing the background and those fuzzy sets are modeled as a combination of fuzzy values. Fuzzy colors (Chamorro-Martínez et al., 2017), which attempt to solve the problem known as “semantic gap”, address the vagueness and subjectivity in the modeling of colors. Fuzzy values generalize the concept of fuzzy colors to other vector-valued spaces. The fuzzy-based reduced ordering also generalizes some distance-based approaches. In contrast to the SVM approach, however, the approach based on fuzzy colors does not involve the solution of an optimization problem. Furthermore, it can naturally take into account the imprecision used by humans to describe and perceive colors.

In this dissertation we will investigate supervised reduced orderings for multi-valued MM with a special focus in the fuzzy value-based approach. The dissertation is organized as follows: Chapter 1 review some basic concepts from lattice theory. Chapter 2 reviews some mathematical tools, including support vector machines, lattice associative memories and fuzzy sets. Chapter 3 reviews the theory of mathematical morphology based on complete lattices. Chapter 4 reviews mathematical morphology on multivalued images based on reduced orderings and we explore some examples of supervised orderings.

In Chapter 5 we introduce our approach based on fuzzy values. Chapter 6 contains an experiment of hyperspectral image segmentation using supervised orderings. Finally, the dissertation ends with some concluding remarks in Chapter 7.

1 Lattice Theory and Reduced Orderings

Lattices (Birkhoff, 1993) are a special type of partially ordered set where every finite subset has an infimum and a supremum. Lattices can also be defined by two binary operations referred to as join and meet. Complete lattices are a special type of lattice in which every subset, finite or infinite, has an infimum and a supremum. Complete lattices serve as the algebraic basis for mathematical morphology (Ronse, 1990).

Besides its role in mathematical morphology and fuzzy logic, lattices are used in the field of lattice computing (Graña, 2008; Kaburlasos et al., 2013), with the construction of lattice associative memories, morphological neural networks and fuzzy lattice neurocomputing.

1.1 Lattices and Complete Lattices

A pair (X, \leq) , where X is a set and \leq is a binary relation on X , is a partially ordered set, or poset, if, for all $x, y, z \in X$:

- $x \leq x$ (reflexivity)
- $y \leq x$ and $x \leq y \implies x = y$ (anti-symmetry)
- $x \leq y$ and $y \leq z \implies x \leq z$ (transitivity)

In that case, \leq is referred to as a partial ordering. An example of a poset is given by the real line with the usual ordering (\mathbb{R}, \leq) . Another example is the power set of a set X and the inclusion relation \subseteq , that is, $(\mathcal{P}(X), \subseteq)$.

A poset (X, \leq) is said to be a totally ordered set if, for all $x, y \in X$, $x \leq y$ or $y \leq x$ hold true. In this case, \leq is said to be a total ordering. The real line equipped with its usual ordering, (\mathbb{R}, \leq) , is a totally ordered set.

Let (X, \leq) be a poset. Given a subset $Y \subseteq X$ and $z \in X$ we say that z is the supremum of Y if $y \leq z \forall y \in Y$ and $z \leq w \forall w$ such that $y \leq w \forall y \in Y$. Similarly, we say that z is the infimum of Y if $z \leq y \forall y \in Y$ and $w \leq z \forall w$ such that $w \leq y \forall y \in Y$. If z is the supremum of Y we write $z = \sup Y$ and if z is the infimum of Y we write $z = \inf Y$. The supremum and infimum are unique when they exist.

Lattices are special types of partially ordered sets given by the following definition:

Definition 1. A partially ordered set (\mathcal{L}, \leq) is a lattice if every finite subset $X \subseteq \mathcal{L}$ admits a supremum and an infimum.

In a lattice (\mathcal{L}, \leq) the pairwise supremum and infimum, also referred to join and meet, are denoted, respectively, by \vee and \wedge . More specifically, for $a, b \in \mathcal{L}$

$$a \vee b := \sup\{a, b\} \quad \text{and} \quad a \wedge b := \inf\{a, b\}. \quad (1.1)$$

A complete lattice is a partially ordered set where every subset, finite or infinite, has a supremum and an infimum. Formally, complete lattices are defined as follows:

Definition 2. A partially ordered set (\mathcal{L}, \leq) is a complete lattice if every subset $X \subseteq \mathcal{L}$ admits a supremum and an infimum.

In a complete lattice, the supremum and infimum are denoted by the symbols \bigvee and \bigwedge , respectively. That is, given $X \subseteq \mathcal{L}$, we have

$$\bigvee X := \sup X \quad \text{and} \quad \bigwedge X := \inf X. \quad (1.2)$$

Given lattices $(\mathcal{L}, \leq_{\mathcal{L}})$ and $(\mathcal{M}, \leq_{\mathcal{M}})$, an operator $\psi : \mathcal{L} \rightarrow \mathcal{M}$ is a homomorphism (Birkhoff, 1993) if it is increasing with respect to the orderings $\leq_{\mathcal{L}}$ and $\leq_{\mathcal{M}}$, that is

$$a \leq_{\mathcal{L}} b \implies \psi(a) \leq_{\mathcal{M}} \psi(b) \quad \forall a, b \in \mathcal{L}. \quad (1.3)$$

Equivalently, a homomorphism preserves the join and meet operations, that is, ψ is a homomorphism if, and only if, for all $a, b \in \mathcal{L}$, we have

$$\psi(a \vee_{\mathcal{L}} b) = \psi(a) \vee_{\mathcal{M}} \psi(b) \quad \text{and} \quad \psi(a \wedge_{\mathcal{L}} b) = \psi(a) \wedge_{\mathcal{M}} \psi(b). \quad (1.4)$$

If, in addition to being a homomorphism, ψ is bijective, we say that ψ is an isomorphism. Finally, if there is an isomorphism $\psi : \mathcal{L} \rightarrow \mathcal{M}$, then $(\mathcal{L}, \leq_{\mathcal{L}})$ and $(\mathcal{M}, \leq_{\mathcal{M}})$ are said to be isomorphic.

Let (\mathbb{L}, \leq) be a (complete) lattice and \mathcal{D} be the point set. The partial ordering \leq induces an ordering on the set of functions from \mathcal{D} to \mathbb{L} , $\mathbb{L}^{\mathcal{D}}$, also referred as the set of \mathbb{L} -valued images on the point set \mathcal{D} , by pointwise application, that is

$$\mathbf{I} \leq \mathbf{J} \iff \mathbf{I}(p) \leq \mathbf{J}(p), \quad \forall p \in \mathcal{D}, \quad (1.5)$$

for all $\mathbf{I}, \mathbf{J} \in \mathbb{L}^{\mathcal{D}}$, such that $(\mathbb{L}^{\mathcal{D}}, \leq)$ is a (complete) lattice.

1.2 Pre-Orderings and Reduced Orderings

A preorder is a binary relation similar to an ordering, but with the anti-symmetry property removed.

Definition 3. A binary relation \leq on a set X is a preorder if it satisfies, for all $x, y, z \in X$

$$\bullet \ x \leq x \quad \text{(reflexivity)}$$

$$\bullet \ x \leq y \text{ and } y \leq z \implies x \leq z \quad \text{(transitivity)}$$

The absence of the anti-symmetry property results in the non-uniqueness of the extrema operators, that is, the supremum and the infimum, as defined for partial orderings, are not necessarily unique in a pre-ordered set.

Similarly to partial orderings, we say that a preorder relation \leq on X is a total preordering if, for all $x, y \in X$, $x \leq y$ or $y \leq x$.

There is a special kind of pre-ordering, called reduced ordering, that can be useful for dealing with vector-valued sets and serve as basis for the supervised orderings that will be studied later. Reduced orderings (Goutsias et al., 1995) rank the elements according to a surjective function h that maps to a complete lattice. More precisely, given a set \mathcal{V} , a complete lattice (\mathcal{L}, \leq) and a surjective function $h : \mathcal{V} \rightarrow \mathcal{L}$, the reduced ordering, or h -ordering, \leq_h is given by

$$\mathbf{x} \leq_h \mathbf{y} \iff h(\mathbf{x}) \leq h(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}. \quad (1.6)$$

Similarly, an equivalence relation $=_h$ is given by

$$\mathbf{x} =_h \mathbf{y} \iff h(\mathbf{x}) = h(\mathbf{y}), \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}. \quad (1.7)$$

Similarly to total preorderings and total orderings a reduced ordering \leq_h is said to be a total reduced ordering if, for all $\mathbf{x}, \mathbf{y} \in X$, $\mathbf{x} \leq_h \mathbf{y}$ or $\mathbf{y} \leq_h \mathbf{x}$, or equivalently, if the ordering \leq of the complete lattice \mathbb{L} is a total ordering.

An operator $\psi^h : \mathcal{V} \rightarrow \mathcal{V}$ is said to be h -increasing if

$$a \leq_h b \implies \psi^h(a) \leq_h \psi^h(b), \quad (1.8)$$

we have the following result regarding h -increasing operators:

Proposition 1 (Goutsias et al. (1995)). *An operator $\psi^h : \mathcal{V} \rightarrow \mathcal{V}$ is h -increasing if and only if there exists an increasing operator $\psi : \mathcal{L} \rightarrow \mathcal{L}$ such that*

$$h\psi^h = \psi h, \quad (1.9)$$

In this case, we write $\psi^h \xrightarrow{h} \psi$.

With this result, given an increasing operator $\psi : \mathcal{L} \rightarrow \mathcal{L}$, we can guarantee the existence of an h -increasing operator $\psi^h : \mathcal{V} \rightarrow \mathcal{V}$. Goutsias et al. (1995) used Proposition 1

to extend mathematical morphology based on complete lattices to reduced orderings, and Velasco-Forero and Angulo (2014) used it to develop morphological operators on multivalued images (see Algorithm 1 on Chapter 4). A flexibilized version of these approaches has been used by Sangalli and Valle (2018) to define reduced orderings under uncertainties.

2 Mathematical Tools

In this section we will review some tools and concepts that will be used throughout the dissertation. More specifically, we will look into support vector machines (SVM), lattice auto-associative memories (LAAM) and fuzzy systems

2.1 Support Vector Machines

A support vector machine (SVM) is a machine learning model that attempts to classify points into two classes by separating them by a hyperplane which maximizes the margin (i.e. the minimum distance from a training point to the hyperplane) (Cristianini et al., 2000). To start with, let us assume that we have a training set $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)$, consisting of points $\mathbf{x}_i \in \mathbb{R}^k$ and labels $y_i \in \{-1, 1\}$ for all $i \in \{1, \dots, N\}$, which is linearly separable, that is, there exists a hyperplane, represented by $\beta^T \mathbf{x} + \beta_0$, such that $\text{sign}(\beta^T \mathbf{x}_i + \beta_0) = y_i$, for all i . If a hyperplane maximizes the margin, the value of the margin is realized in two points, \mathbf{x}^+ and \mathbf{x}^- satisfying $\beta^T \mathbf{x}^+ + \beta_0 = -(\beta^T \mathbf{x}^- + \beta_0)$. Assuming, without loss of generality, that $\beta^T \mathbf{x}^+ + \beta_0 = -(\beta^T \mathbf{x}^- + \beta_0) = 1$ on this hyperplane, then the value of the margin is given by $\frac{1}{\|\beta\|^2}$ (Cristianini et al., 2000). In practice, the restriction of the points being linearly separable is relaxed, and the restriction $y_i(\beta^T \mathbf{x}_i + \beta_0) \geq 1 - \xi_i$, for positive variables ξ_i , is used. The hyperplane with maximum margin is given by the solution of the following quadratic optimization problem

$$\begin{aligned} & \underset{\beta, \beta_0}{\text{minimize}} \quad \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i, \\ & \text{subject to} \quad y_i(\beta^T \mathbf{x}_i + \beta_0) \geq 1 - \xi_i, \quad i = 1, \dots, N, \\ & \quad \quad \quad \xi_i \geq 0, \quad i = 1, \dots, N, \end{aligned} \tag{2.1}$$

where C is a positive constant and the term $C \sum_{i=1}^N \xi_i$ is added to the objective function to control the quantity of points not separated by the hyperplane.

The dual formulation of (2.1) is given by (Cristianini et al., 2000)

$$\begin{aligned} & \text{maximize} \quad \sum_{i=1}^N \lambda_i - \sum_{i,j=1}^N y_i y_j \lambda_i \lambda_j \mathbf{x}_i^T \mathbf{x}_j, \\ & \text{subject to} \quad \sum_{i=1}^N y_i \lambda_i = 0 \\ & \quad \quad \quad 0 \leq \lambda_i \leq C, \quad i = 1, \dots, N, \end{aligned} \tag{2.2}$$

Gaussian RBF	$\mathbf{K}(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\ \mathbf{x} - \mathbf{y}\ ^2}{c}\right)$
Polynomial	$\mathbf{K}(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle + \theta)^d$
Sigmoidal	$\mathbf{K}(\mathbf{x}, \mathbf{y}) = \tanh(\kappa \langle \mathbf{x}, \mathbf{y} \rangle + \theta)$
Inv. multiquadratic	$\mathbf{K}(\mathbf{x}, \mathbf{y}) = \frac{1}{\sqrt{\ \mathbf{x} - \mathbf{y}\ ^2 + c^2}}$

Table 1 – Common kernels (Muller et al., 2001): Gaussian radial basis function, polynomial, sigmoidal and inverse multiquadratic kernels, with parameters $c \in \mathbb{R}_{>0}$, $d \in \mathbb{N}$, $\theta \in \mathbb{R}$ and $\kappa \in \mathbb{R}$.

where the λ_i are the lagrange multipliers of (2.1). The points \mathbf{x}_i such that the inequality $\lambda_i > 0$ holds true are called support vectors and the set of the indices of the support vectors is denoted \mathcal{V}_S . The parameters β and β_0 of the maximum margin hyperplane are given by

$$\beta = \sum_{i=1}^N y_i \lambda_i \mathbf{x}_i \quad \text{and} \quad \beta_0 = \frac{1}{|\mathcal{V}_S|} \sum_{i \in \mathcal{V}_S} (y_i - \beta^T \mathbf{x}_i). \quad (2.3)$$

The signed distance to the maximum margin hyperplane can be expressed as

$$f(\mathbf{x}) = \sum_{i=1}^N y_i \lambda_i \mathbf{x}_i^T \mathbf{x} + \beta_0. \quad (2.4)$$

A kernel \mathbf{K} is a real-valued function obtained by mapping two values from \mathbb{R}^n to a potentially much higher dimensional space \mathbf{F} , referred as feature space, by means of a function $\phi : \mathbb{V} \rightarrow \mathbf{F}$ and computing the dot product, that is, $\mathbf{K}(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$. Under some conditions, the kernel can be computed implicitly, that is, the function ϕ does not need to be computed, or even be known (Muller et al., 2001). Table 1 lists some common examples of kernels.

In (2.2) and in (2.4) the training points appear only through their dot product, allowing us to compute the SVM equations in a feature space with the use of a kernel, without the need for the training data to be linearly separable and allowing for a wider variety of decision boundaries. This method is referred to as the kernel trick (Cristianini et al., 2000). When applying the kernel trick to the SVM problem we obtain

$$\begin{aligned} & \text{maximize} \quad \sum_{i=1}^N \lambda_i - \sum_{i,j=1}^N y_i y_j \lambda_i \lambda_j \mathbf{K}(\mathbf{x}_i, \mathbf{x}_j), \\ & \text{subject to} \quad \sum_{i=1}^N y_i \lambda_i = 0 \\ & \quad \quad \quad 0 \leq \lambda_i \leq C, \quad i = 1, \dots, N, \end{aligned} \quad (2.5)$$

and the distance between the hyperplane and a point \mathbf{x} is given by

$$f(\mathbf{x}) = \sum_{i=1}^N y_i \lambda_i \mathbf{K}(\mathbf{x}_i, \mathbf{x}) + \beta_0, \quad (2.6)$$

where β_0 is given by

$$\beta_0 = \frac{1}{|\mathcal{V}_S|} \sum_{i \in \mathcal{V}_S} \left(y_i - \sum_{j=1}^N y_j \lambda_j K(\mathbf{x}_i, \mathbf{x}_j) \right). \quad (2.7)$$

2.2 Lattice Auto Associative Memories

A heteroassociative linear memory (Kohonen, 1977) can be built from input/output pairs $\{(\mathbf{x}^\xi, \mathbf{y}^\xi)\}_{\xi=1}^k$ as

$$W = \sum_{\xi=1}^k \mathbf{y}^\xi (\mathbf{x}^\xi)^T. \quad (2.8)$$

From an input \mathbf{x} , the heteroassociative linear memory yields

$$\mathbf{y} = W\mathbf{x} \quad (2.9)$$

as the recalled item. Similarly, lattice associative memories (LAM) are its morphological counterparts, proposed by Ritter et al. (1998). Erosive and dilative LAM are built as follows

$$W_{XY} = \bigwedge_{\xi=1}^k \mathbf{y}^\xi \times (-\mathbf{x}^\xi)^T, \quad (2.10)$$

$$M_{XY} = \bigvee_{\xi=1}^k \mathbf{y}^\xi \times (-\mathbf{x}^\xi)^T, \quad (2.11)$$

where \times is either \boxtimes or \boxdot , which are the max and min matrix products, given by

$$(A \boxtimes B)_{ij} = \bigvee_{k=1}^n \{A_{ik} + B_{kj}\}, \quad (2.12)$$

$$(A \boxdot B)_{ij} = \bigwedge_{k=1}^n \{A_{ik} + B_{kj}\}, \quad (2.13)$$

for real-valued matrices A and B of appropriate size. When $\mathbf{x}^\xi = \mathbf{y}^\xi$ for all $\xi \in \{1, \dots, k\}$, we say that the LAM is a lattice auto-associative memory (LAAM).

For a given input \mathbf{x} , the dilative memory recall is given by $W_{XX} \boxdot \mathbf{x}$ and the erosive memory recall is given by $M_{XX} \boxtimes \mathbf{x}$.

The LAAM presents some interesting properties (Sussner and Valle, 2006):

- $\forall \mathbf{x} \in X$, $W_{XX} \boxdot \mathbf{x} = \mathbf{x}$ and $M_{XX} \boxtimes \mathbf{x} = \mathbf{x}$;
- The sets of fixed points of M_{XX} and W_{XX} , denoted $\mathcal{F}(W_{XX})$ and $\mathcal{F}(M_{XX})$, are the same and satisfy:

$$\mathcal{F}(W_{XX}) = \mathcal{F}(M_{XX}) = \left\{ \bigvee_{j=1}^r \bigwedge_{\xi=1}^n (\mathbf{x}^\xi + c_j^\xi) \mid r \in \mathbb{N}, c_j^\xi \in \mathbb{R} \right\};$$

- $W_{XX} \boxtimes \mathbf{x} = \bigvee \{\mathbf{z} \in \mathcal{F}(W_{XX}) | \mathbf{x} \leq \mathbf{z}\};$
- $M_{XX} \boxtimes \mathbf{x} = \bigwedge \{\mathbf{z} \in \mathcal{F}(M_{XX}) | \mathbf{z} \leq \mathbf{x}\}.$

2.3 Fuzzy Systems

2.3.1 Fuzzy Sets

Fuzzy sets, introduced by Zadeh (1965), generalize the usual notion of sets by allowing for a continuous range of membership degrees. In classical set theory, an element either belongs to a set or it does not. Fuzzy sets relax this restriction by allowing for an element to have a degree of membership that can be represented by a number in the interval $[0, 1]$.

Fuzzy sets have also been used to generalize binary mathematical morphology to grayscale images (Nachtegaal and Kerre, 2001; Sussner and Valle, 2008; Bloch, 2011).

Mathematically, a fuzzy set over a set X is a set of pairs

$$A = \{(x, \mu_A(x)) | x \in X, \mu_A(x) \in [0, 1]\}, \quad (2.14)$$

where the so-called membership function $\mu_A : X \rightarrow [0, 1]$ represents the degrees which the elements of X belong to A . For a given $x \in X$, the equality $\mu_A(x) = 1$ means that x has full membership in A and $\mu_A(x) = 0$ means that x has no membership in A . Anything in between means that x has some degree of membership in A but has no full membership, with a greater value of $\mu_A(x)$ meaning a greater degree of membership of x in A . The set of all the fuzzy sets over a set X is denoted by $\mathcal{F}(X)$.

The inclusion relation of fuzzy sets is defined by the pointwise order of the membership functions, that is,

$$A \subseteq B \iff \mu_A \leq \mu_B \iff \mu_A(x) \leq \mu_B(x), \forall x \in X. \quad (2.15)$$

The pair $(\mathcal{F}(X), \subseteq)$ is a complete lattice. The set $[0, 1]$ is a complete lattice when equipped with its usual ordering, which implies that the set of membership functions is also a complete lattice with the pointwise ordering. Furthermore, the function $A \mapsto \mu_A$ that maps a fuzzy set to its membership function is an isomorphism between $(\mathcal{F}(X), \subseteq)$ and $([0, 1]^X, \leq)$. Because of this, a fuzzy set may be identified by its membership function.

The supremum and infimum of a family of fuzzy sets is defined in terms the supremum and infimum of the membership functions. Given a family of fuzzy sets

$\{A_i | i \in I\}$, the membership functions of the supremum $\bigvee_{i \in I} A_i$ and infimum $\bigwedge_{i \in I} A_i$ are

$$\mu_{\bigvee_{i \in I} A_i}(x) := \sup_{i \in I} \mu_{A_i}(x), \quad \forall x \in X \quad (2.16)$$

$$\mu_{\bigwedge_{i \in I} A_i}(x) := \inf_{i \in I} \mu_{A_i}(x), \quad \forall x \in X. \quad (2.17)$$

The set of crisp sets over the set X is given by $\mathcal{C}(X) = \{A \in \mathcal{F}(X) | (\forall x \in X) \mu_A(x) = 1 \text{ or } \mu_A(x) = 0\}$. The set $\mathcal{C}(X)$ is a complete lattice with the inclusion ordering induced by $\mathcal{F}(X)$ and is isomorphic to the power set of X with inclusion ordering under the mapping that assigns each set to its characteristic function, that is, for $A \in \mathcal{P}(X)$, we have

$$\mathcal{X}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases} \quad (2.18)$$

The set complement is also preserved by this isomorphism, that is, for $A \in \mathcal{P}(X)$, $\mathcal{X}_{A^c}(x) = 1 - \mathcal{X}_A(x)$.

The α -level of a fuzzy set A , for $\alpha \in (0, 1]$, is defined by

$$[A]^\alpha = \{x \in X | \mu_A(x) \geq \alpha\}. \quad (2.19)$$

A fuzzy set may be implicitly defined by its α -levels. A fuzzy set may be obtained from its α -levels by means of the following equation, for all $x \in X$

$$A(x) = \sup\{\alpha \in (0, 1] | x \in [A]^\alpha\}. \quad (2.20)$$

The support $\text{Supp}(A) \subseteq X$ of a fuzzy set A is the set of elements of X with non-zeros membership values in A , that is, $\text{Supp}(A) = \{x \in X | \mu_A(x) > 0\}$. Similarly, $\text{Core}(A) \subseteq X$ is the set of elements with full membership value in A , that is $\text{Core}(A) = \{x \in X | \mu_A(x) = 1\}$. A fuzzy set A is normal if $\text{Core}(A) \neq \emptyset$.

2.3.2 Fuzzy Logic Connectives

In classical logic, the value 1 represents “true” and the value 0 represents “false”. Also the “and” and “or” connectives are represented by the truth tables

and	0	1	or	0	1
0	0	0	0	1	1
1	0	1	1	1	0

and are equivalent to the operations \wedge and \vee on the $\{0, 1\}$ lattice with usual ordering.

The operations \wedge and \vee generalize the “and” and “or” connectives of classical logic. These operations belong to broader classes of operations that are able to generalize

the classical “or” and “and”. An increasing operation that generalizes the logical “or” is called a disjunction and an increasing operation that generalizes the logical “and” is called a conjunction. Notable special cases of conjunctions and disjunctions are the triangular-norms and triangular-conorms, also called t-norms and s-norms, which are defined below.

Definition 4 ((Nguyen and Walker, 2005)). Let $\triangle : [0, 1] \times [0, 1] \rightarrow [0, 1]$. If for all $x, y, z \in [0, 1]$

- $1 \triangle x = x;$ (neutral element)
- $x \triangle y = y \triangle x;$ (commutativity)
- $x \triangle (y \triangle z) = (x \triangle y) \triangle z;$ (associativity)
- if $x \leq y$ and $z \leq w$ then $x \triangle z \leq y \triangle w,$ (increasingness)

then \triangle is said to be a triangular-norm.

Definition 5 ((Nguyen and Walker, 2005)). Let $\nabla : [0, 1] \times [0, 1] \rightarrow [0, 1]$. If for all $x, y, z \in [0, 1]$

- $0 \nabla x = x;$ (neutral element)
- $x \nabla y = y \nabla x;$ (commutativity)
- $x \nabla (y \nabla z) = (x \nabla y) \nabla z;$ (associativity)
- if $x \leq y$ and $z \leq w$ then $x \nabla z \leq y \nabla w,$ (increasingness)

then ∇ is said to be a triangular-conorm.

The pairwise minimum \wedge is a triangular-norm. Likewise, the pairwise maximum \vee is a triangular-conorm. Some other notable examples of t-norms are:

- The product t-norm, given by $x \triangle_{\text{P}} y = xy;$
- The Lukasiewicz t-norm, given by $x \triangle_{\text{L}} y = 0 \vee (x + y - 1);$
- The drastic t-norm, given by

$$x \triangle_{\text{D}} y = \begin{cases} x, & \text{if } y = 1, \\ y, & \text{if } x = 1, \\ 0, & \text{if } x \neq 1, \text{ and } y \neq 1. \end{cases}$$

Some notable examples of s-norms are:

- The probabilistic sum s-norm, given by $x \nabla_P y = x + y - xy$;
- The Lukasiewicz s-norm, given by $x \nabla_L y = 1 \wedge (x + y)$;
- The drastic s-norm, given by

$$x \nabla_D y = \begin{cases} x, & \text{if } y = 0, \\ y, & \text{if } x = 0, \\ 1, & \text{if } x \neq 0, \text{ and } y \neq 0. \end{cases}$$

The operations \wedge and \vee are respectively the greatest t-norm and the least s-norm, that is, by considering the pointwise ordering on the functions from $[0, 1]^2$ to $[0, 1]$ we obtain, for any t-norm \triangle and s-norm ∇

$$\triangle \leq \wedge \leq \vee \leq \nabla. \quad (2.21)$$

Because of this, we can think of a t-norm, when applied pointwise to a pair of membership functions, as a stronger intersection. Dually, an s-norm can be viewed as a stronger union.

While (2.16) and (2.17) describe unions and intersections of arbitrary families of fuzzy sets, unions and intersections of finitely many fuzzy sets can be defined in terms of a t-norm and s-norm pair. For a given t-norm \triangle and s-norm ∇ , the union $A \cup B$ and the intersection $A \cap B$ of fuzzy sets A and B are given, in terms of their membership functions, by

$$\mu_{A \cup B}(x) = \mu_A(x) \triangle \mu_B(x), \quad \forall x \in X \quad (2.22)$$

$$\mu_{A \cap B}(x) = \mu_A(x) \nabla \mu_B(x), \quad \forall x \in X. \quad (2.23)$$

The logical negation is given by $\neg 1 = 0$ and $\neg 0 = 1$ is generalized by the strong fuzzy negation:

Definition 6 ((Nguyen and Walker, 2005)). *Let $\eta : [0, 1] \rightarrow [0, 1]$. If η satisfies*

- $\eta(0) = 1$ and $\eta(1) = 0$;
- η is nonincreasing;
- $\eta^2 = \text{id}$,

then η is said to be a strong fuzzy negation.

Given a fuzzy set A , its complement A^c is given by

$$(\forall x \in X) \quad \mu_{A^c}(x) = \eta(\mu_A(x)), \quad (2.24)$$

where η is a strong fuzzy negation.

A t-norm Δ and an s-norm ∇ are said to be dual with respect to a strong negation η if, for all $x, y \in [0, 1]$, $x \Delta y = \eta(\eta(x) \nabla \eta(y))$, or, equivalently, $x \nabla y = \eta(\eta(x) \Delta \eta(y))$. Let us denote the standard fuzzy negation by $\eta_{\mathbf{C}}(x) = 1 - x$. We have that

- \wedge and \vee are dual with respect to $\eta_{\mathbf{C}}$;
- $\Delta_{\mathbf{P}}$ and $\nabla_{\mathbf{P}}$ are dual with respect to $\eta_{\mathbf{C}}$;
- $\Delta_{\mathbf{L}}$ and $\nabla_{\mathbf{L}}$ are dual with respect to $\eta_{\mathbf{C}}$;
- $\Delta_{\mathbf{D}}$ and $\nabla_{\mathbf{D}}$ are dual with respect to $\eta_{\mathbf{C}}$.

The implication is a connective that satisfies the truth table

\rightarrow	0	1
0	1	1
1	0	1

A fuzzy implication is a mapping $I : [0, 1] \times [0, 1] \rightarrow [0, 1]$ such that, for all $y \in [0, 1]$, $I(\cdot, y)$ is increasing, $I(y, \cdot)$ is decreasing and satisfies the implication truth table.

An R-implication is an implication that can be obtained from a t-norm Δ by means of the equation

$$x \rightarrow y = \sup\{z \mid x \Delta z \leq y\}. \quad (2.25)$$

Some notable R-implications are

- The Gödel implication, which is the R-implication of \wedge , is given by

$$x \rightarrow_{\mathbf{G}} y = \begin{cases} 1, & \text{if } x \leq y, \\ y, & \text{if } x > y. \end{cases} \quad (2.26)$$

- The Goguen implication, which is the R-implication of $\Delta_{\mathbf{P}}$, is given by

$$x \rightarrow_{\mathbf{P}} y = \begin{cases} 1, & \text{if } x \leq y, \\ \frac{y}{x}, & \text{if } x > y. \end{cases} \quad (2.27)$$

- The Lukasiewicz implication, which is the R-implication of $\Delta_{\mathbf{L}}$, is given by

$$x \rightarrow_{\mathbf{L}} y = \begin{cases} 1, & \text{if } x \leq y, \\ 1 - x + y, & \text{if } x > y. \end{cases} \quad (2.28)$$

An S-implication is obtained from an s-norm ∇ and a fuzzy negation η from

$$x \rightarrow y = \eta(x) \nabla y, \quad (2.29)$$

and it is a more direct generalization of the classical implication.

The t-norms and s-norms are special cases of another operation, called uninorm, given by

Definition 7 ((Yager and Rybalov, 1996)). *Let $*$: $[0, 1] \times [0, 1] \rightarrow [0, 1]$. If there exists $e \in [0, 1]$ such that for all $x, y, z \in [0, 1]$:*

- $e * x = x$; (neutral element)
- $x * y = y * x$; (commutativity)
- $x * (y * z) = (x * y) * z$; (associativity)
- if $x \leq y$ and $z \leq w$ then $x * z \leq y * w$, (increasingness)

then $*$ is said to be a uninorm.

A simple example of a uninorm with a neutral element e is given by, for all $x, y \in [0, 1]$

$$x *_e y = \begin{cases} x \vee y, & \text{if } e \leq x \text{ and } e \leq y, \\ x \wedge y, & \text{if } x < e \text{ or } y < e. \end{cases} \quad (2.30)$$

An uninorm $*$ is self-dual with respect to a negation η if, for all $x, y \in [0, 1]$,

$$x * y = \eta(\eta(x) * \eta(y)). \quad (2.31)$$

An example of a self-dual uninorm with respect to the standard negation $\eta(x) = 1 - x$ is given by

$$x * y = \frac{xy}{xy + (1 - x)(1 - y)}, \quad \forall x, y \in [0, 1]. \quad (2.32)$$

2.3.3 Adaptive Neuro Fuzzy Inference System

The adaptative neuro fuzzy inference system (ANFIS)(Jang and Sun, 1997) is a neural network model equivalent to a Takagi-Sugeno-Kang type fuzzy inference system (FIS)(Sugeno and Kang, 1988). The Sugeno inference system is formulated by a set of fuzzy rules, as follows

$$\begin{array}{llllll} \text{If } x_1 \text{ is } A_{1,1} & \text{and} & \cdots & \text{and} & x_k \text{ is } A_{1,k} & \text{then } y = f_1(\mathbf{x}), \\ \text{If } x_1 \text{ is } A_{2,1} & \text{and} & \cdots & \text{and} & x_k \text{ is } A_{2,k} & \text{then } y = f_2(\mathbf{x}), \\ & & \vdots & & \vdots & \vdots \\ \text{If } x_1 \text{ is } A_{N,1} & \text{and} & \cdots & \text{and} & x_k \text{ is } A_{N,k} & \text{then } y = f_N(\mathbf{x}), \end{array}$$

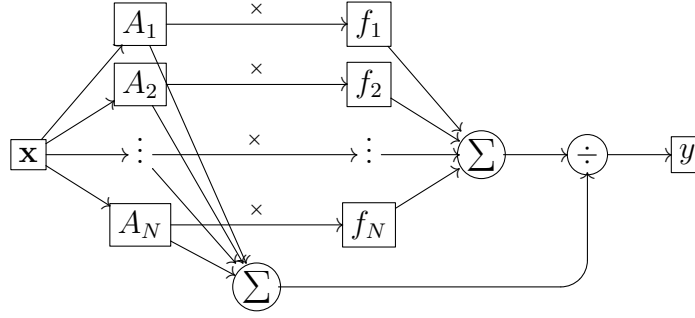


Figure 1 – Graph representation of the Sugeno inference system.

where $\mathbf{x} = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$ is a vector variable, $y \in \mathbb{R}$ is the output of the system, $A_{i,j} \in \mathcal{F}(\mathbb{R})$ for $i \in \{1, \dots, k\}$, $j \in \{1, \dots, N\}$ are the antecedent fuzzy sets and $f_j : \mathbb{R}^k \rightarrow \mathbb{R}$ are the consequent functions. The “and” connective is modeled after the product t-norm, and the output is given by the weighted mean of the consequent functions, with the weights given by the product of the antecedent fuzzy sets. Defining $\mu_{A_i}(\mathbf{x}) = \prod_{j=1}^k \mu_{A_{i,j}}(x_k)$, the output is obtained by

$$y = \frac{\sum_{i=1}^N \mu_{A_i}(\mathbf{x}) f_i(\mathbf{x})}{\sum_{i=1}^N \mu_{A_i}(\mathbf{x})}. \quad (2.33)$$

The inference system is represented as a graph in Figure 1. The ANFIS model uses parameterized membership functions for $A_{i,j}$ and attempts to obtain the optimal set of parameters for the regression of a set of input/output pairs of the form $\{(\mathbf{x}_i, y_i) | \mathbf{x}_i \in \mathbb{R}^k, y_i \in \mathbb{R}, 1 \leq i \leq M\}$, using a loss function such as the mean squared error and an optimization method like the gradient descent.

If all the $A_{i,j}$ are Gaussian functions, that is $A_{i,j}(x) = \exp\left(\frac{-(x - c_{i,j})^2}{2\sigma_{i,j}^2}\right)$, then the functions A_i will be Gaussian radial basis functions, that is, a function of the form

$$A(\mathbf{x}) = \exp\left(-\sum_{j=1}^k \frac{(x_j - c_j)^2}{\sigma_j^2}\right), \quad (2.34)$$

where $\mathbf{c} = (c_1, \dots, c_k) \in \mathbb{R}^k$ is the center and σ_j are the standard deviations for all j . The radial basis function network (RBFN)(Jang and Sun, 1997) uses a similar approach to the ANFIS, but the input is fed directly to the Gaussian radial basis functions.

3 Mathematical Morphology

Mathematical morphology was introduced by Matheron and Serra in the 1960s as a tool to analyze binary images and was later generalized to work on gray-scale images, including some approaches based on fuzzy set theory (Bloch, 2011; Sussner and Valle, 2008; Nachtgael and Kerre, 2001; Deng and Heijmans, 2002). Complete lattices constitute a general framework for mathematical morphology (Heijmans, 1995) and has allowed mathematical morphology to be used on multivalued images, including color images (Aptoula and Lefèvre, 2007; Angulo, 2007).

3.1 Erosion and Dilation

Mathematical morphology began as a tool to analyze binary images. A binary image is a set $\mathbf{I} \subseteq \mathcal{D}$, where \mathcal{D} is the point set, usually a subset of \mathbb{R}^2 or \mathbb{Z}^2 in the case of digital images. Assuming that $\mathcal{D} \subseteq E$ where $(E, +)$ is an abelian group, the dilation of an image $\mathbf{I} \subseteq \mathcal{D}$ by a pattern $S \subseteq E$, referred to as structuring element, is given by the Minkowski sum, that is

$$\delta_S(\mathbf{I}) := \mathbf{I} \oplus S = \{p \in \mathcal{D} \mid S_p^* \cap \mathbf{I} \neq \emptyset\}, \quad \forall \mathbf{I} \subseteq \mathcal{D}, \quad (3.1)$$

where $S_p^* = \{p - s \mid s \in S\}$. The dilation can be interpreted geometrically as the set of points such that the structuring element, translated by those points, “hits” the image. The visual effect of the dilation is to expand the image, as is seen on Figure 2(b).

The erosion of an image \mathbf{I} by the structuring element S is defined by the Minkowski subtraction

$$\varepsilon_S(\mathbf{I}) := \mathbf{I} \ominus S = \{p \in \mathcal{D} \mid S_p \subseteq \mathbf{I}\}, \quad \forall \mathbf{I} \subseteq \mathcal{D}, \quad (3.2)$$

where $S_p = \{s + p \mid s \in S\}$. The erosion can be interpreted geometrically as the set of points such that the structuring element, translated by those points, “fits” the image. The visual effect of the erosion is to shrink the image, as is seen on Figure 2(c).

It turns out that $(\mathcal{P}(E), \subseteq)$ is a complete lattice with the supremum and the infimum given respectively by \bigcup and \bigcap . The erosion and dilation of binary images also commute with the infimum and supremum, respectively, that is, given a family of images $\{\mathbf{I}_j \subseteq \mathcal{D} \mid j \in A\}$, we have

$$\varepsilon_S \left(\bigcap_{j \in A} \mathbf{I}_j \right) = \bigcap_{j \in A} \varepsilon_S(\mathbf{I}_j), \quad \text{and} \quad \delta_S \left(\bigcup_{j \in A} \mathbf{I}_j \right) = \bigcup_{j \in A} \delta_S(\mathbf{I}_j). \quad (3.3)$$

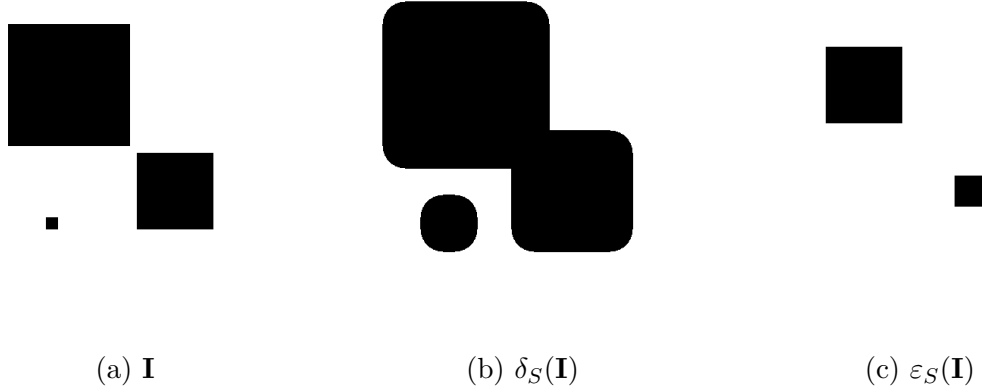


Figure 2 – A binary image \mathbf{I} , represented by the black part, on a 512×512 square on \mathbb{Z}^2 and its erosion and dilation by a disk structuring element of radius 30, given by $S = \{x \in \mathbb{Z}^2 \mid \|x\|_2 \leq 30\}$.

Serra (1988) used these properties to define an algebraic basis for mathematical morphology in a way that preserves the properties of the binary erosion and the binary dilation. Formally, dilations and erosions are algebraically defined as follows:

Definition 8 (Erosion). *Let (\mathcal{L}, \leq) be a complete lattice. We say that an operator $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}$ is an (algebraic) erosion if*

$$\bigwedge_{j \in A} \varepsilon(\mathbf{I}_j) = \varepsilon \left(\bigwedge_{j \in A} \mathbf{I}_j \right), \quad (3.4)$$

for every family $\{\mathbf{I}_j \mid j \in A\}$.

Definition 9 (Dilation). *Let (\mathcal{L}, \leq) be a complete lattice. We say that an operator $\delta : \mathcal{L} \rightarrow \mathcal{L}$ is an (algebraic) dilation if*

$$\bigvee_{j \in A} \delta(\mathbf{I}_j) = \delta \left(\bigvee_{j \in A} \mathbf{I}_j \right), \quad (3.5)$$

for every family $\{\mathbf{I}_j \mid j \in A\}$.

In words, erosions and dilations are defined as operators that commute, respectively, with the infimum and the supremum. By commuting with the infimum or the supremum, it is easy to see that erosions and dilations are always increasing operators.

The adjunction, a relation between operators in complete lattices, can also characterize erosions and dilations.

Definition 10 (Heijmans (1995)). *Let (\mathcal{L}, \leq) be a complete lattice and $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}$ and $\delta : \mathcal{L} \rightarrow \mathcal{L}$. We say that the pair (ε, δ) is an adjunction if*

$$(\forall \mathbf{I}, \mathbf{J} \in \mathcal{L}) \quad \delta(\mathbf{I}) \leq \mathbf{J} \iff \mathbf{I} \leq \varepsilon(\mathbf{J}). \quad (3.6)$$

The next proposition tells us some properties of an adjunction, in particular, it tells us that if (ε, δ) is an adjunction, then ε is an algebraic erosion and δ is an algebraic dilation.

Proposition 2 (Heijmans (1995)). *If (ε, δ) is an adjunction, then the following holds:*

$$\bigwedge_{j \in A} \varepsilon(\mathbf{I}_j) = \varepsilon\left(\bigwedge_{j \in A} \mathbf{I}_j\right), \quad (3.7)$$

$$\bigvee_{j \in A} \delta(\mathbf{I}_j) = \delta\left(\bigvee_{j \in A} \mathbf{I}_j\right), \quad (3.8)$$

$$\text{id} \leq \delta\varepsilon, \quad (3.9)$$

$$\varepsilon\delta \leq \text{id}, \quad (3.10)$$

$$\varepsilon\delta\varepsilon = \varepsilon, \quad (3.11)$$

$$\delta\varepsilon\delta = \delta, \quad (3.12)$$

$$\varepsilon(\mathbf{J}) = \bigvee \{\mathbf{I} \in \mathcal{L} \mid \delta(\mathbf{I}) \leq \mathbf{J}\}, \quad (3.13)$$

$$\delta(\mathbf{I}) = \bigwedge \{\mathbf{J} \in \mathcal{L} \mid \mathbf{I} \leq \varepsilon(\mathbf{J})\}. \quad (3.14)$$

Now we will look into some examples of erosions and dilations:

Definition 11 (Erosion and Dilation of a Lattice-Valued Image by a Flat Structuring Element). *Let (\mathbb{L}, \leq) be a complete lattice and $\mathcal{D} \subseteq E$, with $E = \mathbb{Z}^2$, a point set. The erosion and dilation of an image $\mathbf{I} \in \mathbb{L}^{\mathcal{D}}$ by a structuring element $S \subseteq E$ are given by*

$$\varepsilon_S(\mathbf{I})(p) = \bigwedge_{s \in S, p+s \in \mathcal{D}} \mathbf{I}(p+s), \quad \forall \mathbf{I} \in \mathbb{L}^{\mathcal{D}}, p \in \mathcal{D}, \quad (3.15)$$

$$\delta_S(\mathbf{I})(p) = \bigvee_{s \in S, p-s \in \mathcal{D}} \mathbf{I}(p-s), \quad \forall \mathbf{I} \in \mathbb{L}^{\mathcal{D}}, p \in \mathcal{D}. \quad (3.16)$$

Erosions and dilations can also be defined with non-flat structuring element, that is, structuring elements not defined by a set. Some notable examples of the use of non-flat structuring elements on the literature are the umbra approach (Serra, 1982) and fuzzy mathematical morphology (Bloch and Maitre, 1995; De Baets, 1997).

Example 1 (Gray-scale Erosion and Dilation). *Figure 3 illustrates the effect of erosion and the dilation of a gray-scale image, that is, when $\mathbb{L} \subseteq \bar{\mathbb{R}}$, with $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$, by a flat structuring element. We can notice that the erosion darkens the image, shrinking the lighter regions and expanding the darker regions, while the dilation lightens the image, shrinking the darker regions and expanding the lighter regions.*

Example 2 (Gray-scale Morphological Gradient). *An useful morphological operator for gray-scale images that can be obtained from the erosion and the dilation is the morphological gradient. In its simplest form, the morphological gradient ρ_S is obtained by*

$$\rho_S = \delta_{S^*} - \varepsilon_S, \quad (3.17)$$

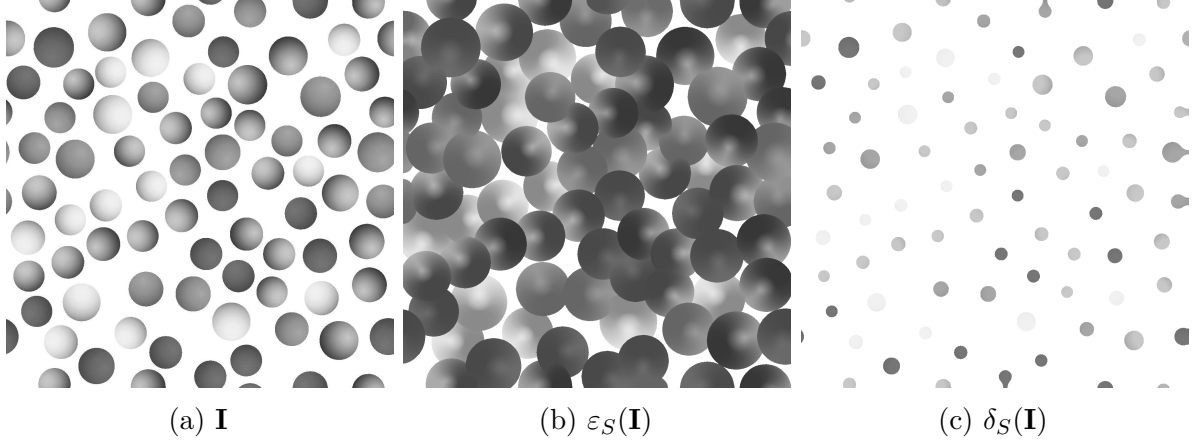


Figure 3 – A gray-scale image \mathbf{I} , where the pixel is greater the brighter they are, on a 1200×1200 square on \mathbb{Z}^2 (a) and its erosion and dilation by a disk of radius 30.

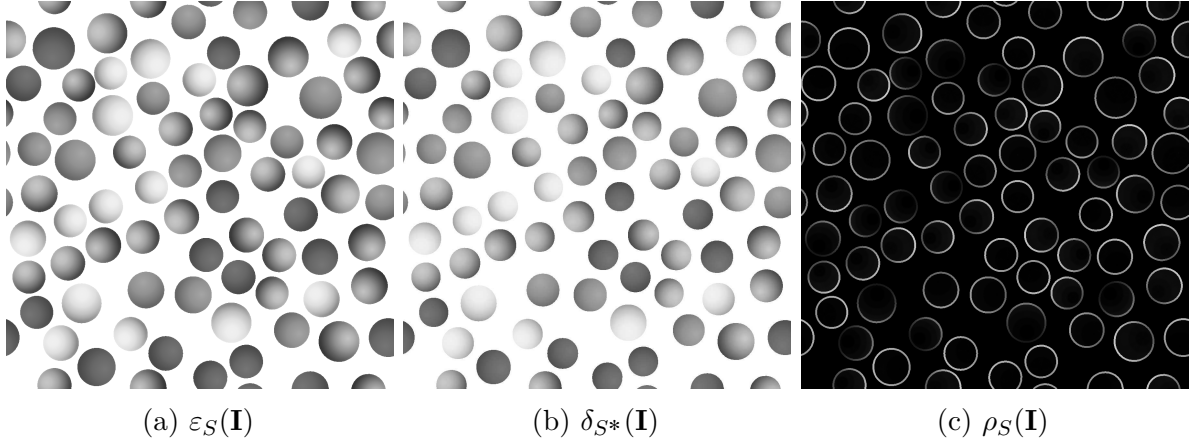


Figure 4 – Erosion, dilation and morphological gradient obtained from image \mathbf{I} from figure 3(a) using a disk of radius 3 as structuring element.

where $S^* = \{-s | s \in S\}$ is the reflected structuring element. Assuming that the origin O belongs to S , $\delta_{S^*} \geq \varepsilon_S$ and, therefore, ρ_S admits only positive values. The morphological gradient highlights the edges of objects in an image, setting its regions to high values, while simultaneously setting the flat regions to low values. Figure 4 exemplifies those properties using the image \mathbf{I} shown on figure 3(a). The morphological gradient is usually used as a preliminary operation on many image processing tasks, such as image segmentation.

Example 3 (Geodesic Erosions and Dilations). In the set of gray-scale images one can define geodesic erosions and geodesic dilations (Soille, 1999), which are also types of algebraic erosions and algebraic dilations. The geodesic dilation is computed for an image \mathbf{I} with a marker image \mathbf{J} , where $\mathbf{I} \leq \mathbf{J}$, from the following equations

$$\delta_{\mathbf{J}}^0(\mathbf{I}) = \mathbf{I}, \quad \delta_{\mathbf{J}}^i(\mathbf{I}) = \mathbf{J} \wedge \delta_N \delta_{\mathbf{J}}^{i-1}(\mathbf{I}), \quad (3.18)$$

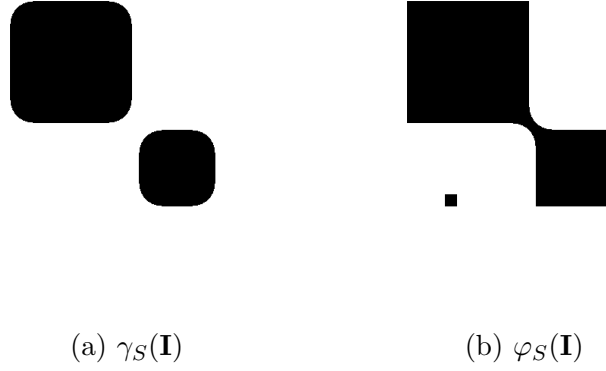


Figure 5 – The opening and closing by a disk of radius 30, of the image \mathbf{I} from figure 2.

and the geodesic erosion, for $\mathbf{I} \geq \mathbf{J}$,

$$\varepsilon_{\mathbf{J}}^0(\mathbf{I}) = \mathbf{I}, \quad \varepsilon_{\mathbf{J}}^i(\mathbf{I}) = \mathbf{J} \vee \varepsilon_N \varepsilon_{\mathbf{J}}^{i-1}(\mathbf{I}), \quad (3.19)$$

where the structuring element N is the Von Neumann neighborhood, given by $N = \{(0, 0), (0, 1), (1, 0), (-1, 0), (0, -1)\}$.

Example 4 (\boxminus and \boxplus). The operations \boxminus and \boxplus given by (2.12) and (2.13) also define algebraic erosions and dilations. For a fixed matrix A , the operator $A \boxminus \cdot$ is an algebraic erosion and the operator $A \boxplus \cdot$ is an algebraic dilation (Sussner and Esmi, 2011). Thus, given the dilative and erosive memories, W_{XY} and M_{XY} , respectively defined by (2.10) and (2.11), the dilative memory recall, given by $\delta(\mathbf{x}) = W_{XY} \boxplus \mathbf{x}$ is an algebraic dilation and the erosive memory recall, given by $\varepsilon(\mathbf{x}) = M_{XY} \boxminus \mathbf{x}$ is an algebraic erosion

3.2 Opening and Closing

Openings and closings are non-linear filters that removes small objects from an image. Originally, in binary mathematical morphology, the opening by a structuring element S is given by $\gamma_S = \delta_S \varepsilon_S$ and the closing by the structuring element S is given by $\varphi_S = \varepsilon_S \delta_S$. Given an image \mathbf{I} , the opening has the effect of removing details smaller than the structuring element from the foreground of the image while the closing has the effect of removing details smaller than the structuring element from the background. On Figures 5(a) and 5(b) we can see those effects on the image of Figure 2(a).

Such as dilations and erosions, opening and closings can be defined algebraically using complete lattices. To this end, let (\mathcal{L}, \leq) be a complete lattice and consider an operator $\psi : \mathcal{L} \rightarrow \mathcal{L}$. We say that ψ is:

- idempotent if $\psi^2 = \psi$;

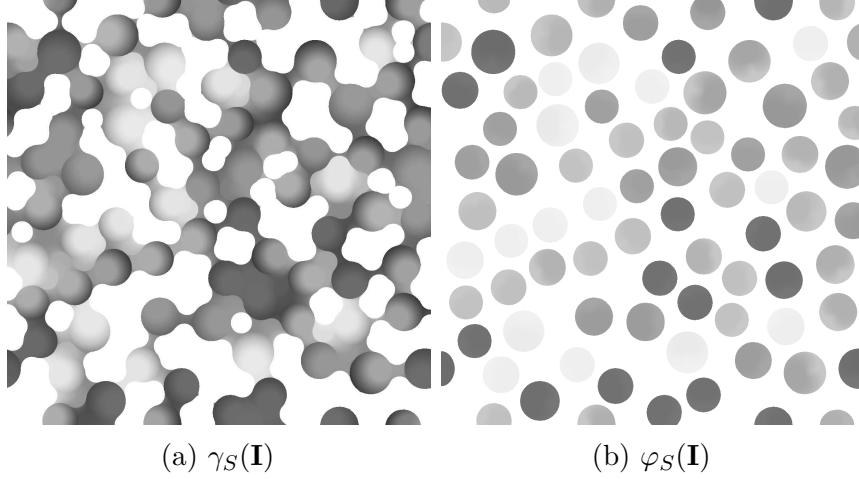


Figure 6 – The image from Figure 3(a) and its opening and closing by a disk of radius 30.

- extensive if $\text{id} \leq \psi$;
- anti-extensive if $\psi \leq \text{id}$.

A morphological filter is an operator that is increasing and idempotent. An algebraic opening is an anti-extensive morphological filter, and an algebraic closing is an extensive morphological filter. By (3.9), (3.10), (3.12) and (3.11), we conclude that, when (ε, δ) is an adjunction $\delta\varepsilon$ is an opening and $\varepsilon\delta$ is a closing. In particular, the operator $\gamma_S = \delta_S\varepsilon_S$ is the opening by the structuring element S and the operator $\varphi_S = \varepsilon_S\delta_S$ is the closing by the structuring element S .

Example 5 (Opening and Closing of a Gray-scale Image by a Structuring Element). *The compositions of the erosion and dilation by a structuring element is an opening in the set of images $\mathbb{L}^{\mathcal{D}}$. Figure 6 shows an example of an opening and a closing in the case of gray-scale images.*

Example 6 (Openings and Closings by Reconstruction). *The reconstruction $R_{\mathbf{J}}(\mathbf{I})$ of a marker image \mathbf{M} to an image \mathbf{I} can be built from geodesic dilations by $R_{\mathbf{J}}(\mathbf{I}) = \delta_{\mathbf{J}}^i(\mathbf{I})$ if $\delta_{\mathbf{J}}^i(\mathbf{I}) = \delta_{\mathbf{J}}^{i+1}(\mathbf{I})$. Similarly, the reconstruction by erosions is given by $R_{\mathbf{J}}^*(\mathbf{I}) = \varepsilon_{\mathbf{J}}^i(\mathbf{I})$ if $\varepsilon_{\mathbf{J}}^i(\mathbf{I}) = \varepsilon_{\mathbf{J}}^{i+1}(\mathbf{I})$. The opening by reconstruction of an image \mathbf{I} by a given structuring element S is given by*

$$\gamma_S^R(\mathbf{I}) = R_{\mathbf{I}}(\varepsilon_S(\mathbf{I})).$$

Similarly, the closing by reconstruction is defined by

$$\varphi_S^R(\mathbf{I}) = R_{\mathbf{I}}^*(\delta_S(\mathbf{I})).$$

Contrary to the opening by a structuring element, the opening by reconstruction preserves the shape of the objects that are not completely removed by initial erosion (Soille, 1999).

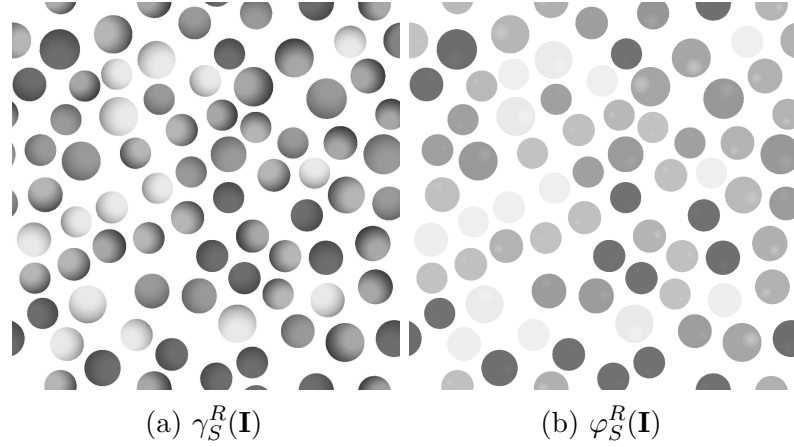


Figure 7 – The opening(a) and the closing(b) by reconstruction of the image from Figure 3(a) with a disk structuring element S of radius 30.

Similarly, contrary to the closing by a structuring element, the closing by reconstruction preserves the shape of objects in the background that are not completely removed by the initial dilation. Figure 7 illustrates these statements with examples of an opening by reconstruction and a closing by reconstruction. The structuring element and the image \mathbf{I} used were the same as the ones used in Figure 6.

3.3 The Marginal and Lexicographical Approaches

The marginal and lexicographical approaches are designed to deal with multivalued images. In this case, we will assume that the value set \mathbb{V} is a subset of $\bar{\mathbb{R}}^k$, for a positive integer k , where $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. The examples of this section considers RGB color images, that is, elements of the set $\mathbb{V}_{\text{RGB}}^{\mathcal{D}}$. An element $\mathbf{x} \in \mathbb{V}_{\text{RGB}}$ is of the form $\mathbf{x} = (x_{\text{R}}, x_{\text{G}}, x_{\text{B}})$, where $x_{\text{R}} \in [0, 1]$ corresponds to the level of red of the color represented by \mathbf{x} , $x_{\text{G}} \in [0, 1]$ corresponds to the level of green and $x_{\text{B}} \in [0, 1]$ corresponds to the level of blue.

A straightforward extension of the gray-scale MM to multivalued images, referred to as the marginal or component-wise approach, is obtained by processing separately each component (Aptoula and Lefèvre, 2007; Comer and Delp, 1999). In mathematical terms, the marginal approach is obtained by ordering the colors $\mathbf{x} = (x_1, \dots, x_k)$ and $\mathbf{y} = (y_1, \dots, y_k)$ as follows:

$$\mathbf{x} \leq_{\text{marg}} \mathbf{y} \iff (\forall i)(x_i \leq y_i) \quad (3.20)$$

where \leq denotes the usual ordering scheme of real numbers. One can easily check that \leq_{marg} is a partial ordering on \mathbb{V} . Also, it is not hard to show that $(\mathbb{V}, \leq_{\text{marg}})$ is a complete lattice. The elementary morphological operators of the marginal approach, given by (3.15) and (3.16) with the ordering defined by (3.20), are denoted respectively by ε_S^M and δ_S^M .

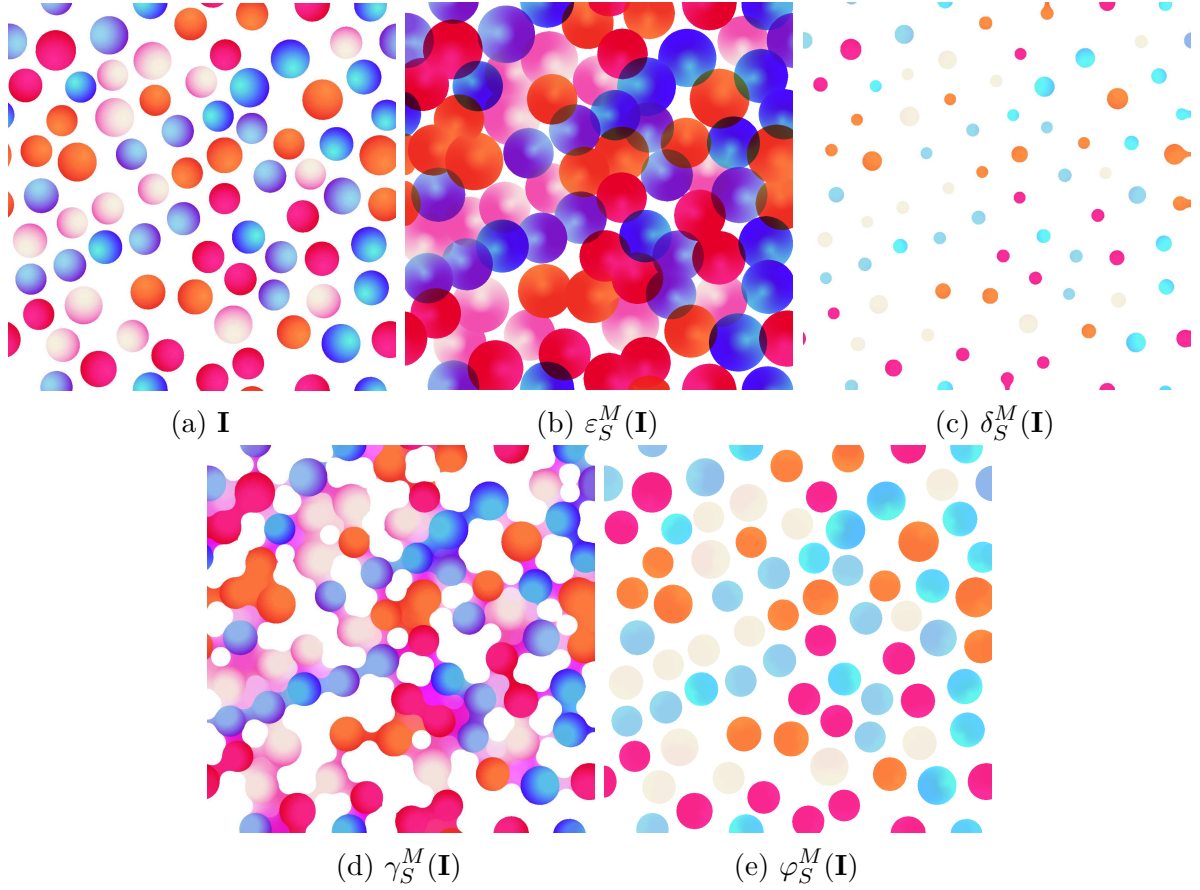


Figure 8 – A 1200×1200 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its erosion, dilation, opening, and closing by a disk of radius 30 with the marginal ordering. The color image \mathbf{I} was obtained from <https://www.freepik.com/>.

Figure 8 shows some of the morphological operators of the marginal ordering in the \mathbb{V}_{RGB} color space.

Although the marginal approach yielded excellent results in computational experiments concerning the removal of Gaussian noise (Aptoula and Lefèvre, 2007), it does not take into account the correlations between the value coordinates. In fact, certain features can be removed or enhanced in one of the coordinates but not in the others. As a consequence, there is the possibility of introducing *false values*, changing the value balance, or altering the edges of objects (Comer and Delp, 1999). Figure 9 illustrates this effect. Notice that there are some new colors in the erosion and dilation obtained using the marginal ordering (3.20). This effect is also observed on Figure 8(b): it is possible to see a darker region of colors that were not present on the original image where some of the circles overlap. This is because the correlation between the color components was not taken into account. These undesired effects can be avoided by endowing the color space with a total ordering instead of a partial ordering.

In contrast to the marginal approach, values are ranked sequentially in the lexicographical approach. Formally, the lexicographical ordering, denoted by \leq_{lex} , is defined

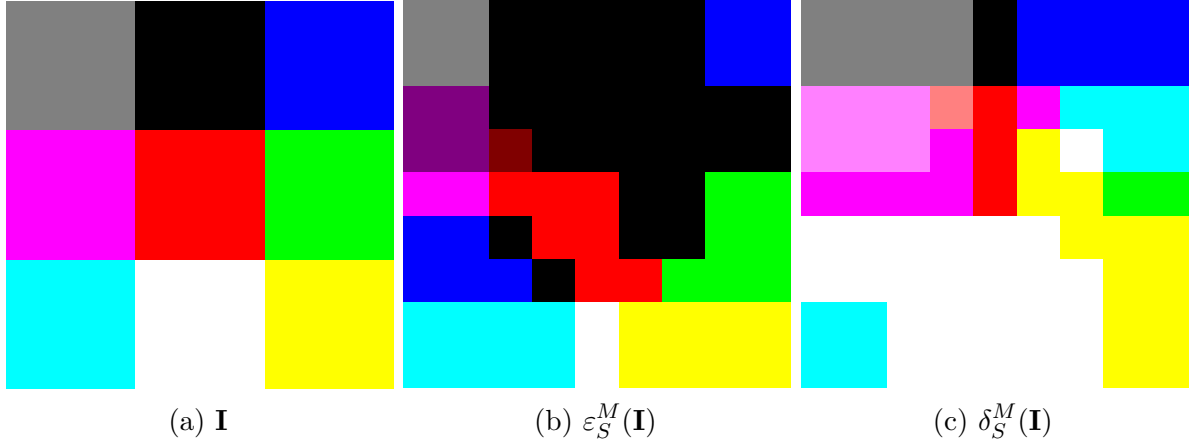


Figure 9 – An example of the problem of *false values* in the marginal ordering. A 9×9 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its erosion and dilation by the marginal ordering approach with a cross-shaped structuring element given by $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

by means of the following equation for $\mathbf{x}, \mathbf{y} \in \mathbb{V}$:

$$\mathbf{x} \leq_{\text{lex}} \mathbf{y} \iff \begin{cases} \mathbf{x} = \mathbf{y} \text{ or,} \\ \exists i \in \{1, \dots, k\} : \forall j < i, x_j = y_j \text{ and } x_i < y_i. \end{cases} \quad (3.21)$$

One can easily show that \leq_{lex} is a total ordering and that $(\mathbb{V}, \leq_{\text{lex}})$ is a complete lattice. The lexicographical erosion and the lexicographical dilation of a multivalued image by a structuring element S , denoted by ε_S^L and δ_S^L , are given respectively by (3.15) and (3.16) with the ordering defined by (3.21). Figure 10 illustrates some of the morphological operators of the lexicographical ordering on the \mathbb{V}_{RGB} color space. In contrast to the marginal ordering operators, there is no overlap between the different objects.

The lexicographical approach has been widely used in multivalued MM partially because it prevents the presence of *false values*. It turns out, however, that this ordering scheme prioritizes excessively the first condition in the lexicographical cascade (Aptoula and Lefèvre, 2007). Furthermore, the lexicographical orderings and total orderings in general are affected by the issue of irregularity (Chevallier and Angulo, 2016), which may distort the output of some morphological operators. Briefly, the irregularity issue states that, if (\mathbb{V}, d) is a metric space and \leq a total ordering on \mathbb{V} , under some circumstances, which usually holds in the case of multivalued sets, there are points $a, b, c \in \mathbb{V}$ and real numbers $r, R \in (0, +\infty)$ such that $R > r$ and

$$\begin{aligned} a &\leq b \leq c, \\ d(a, b) &\geq R, \\ d(a, c) &\leq r. \end{aligned}$$

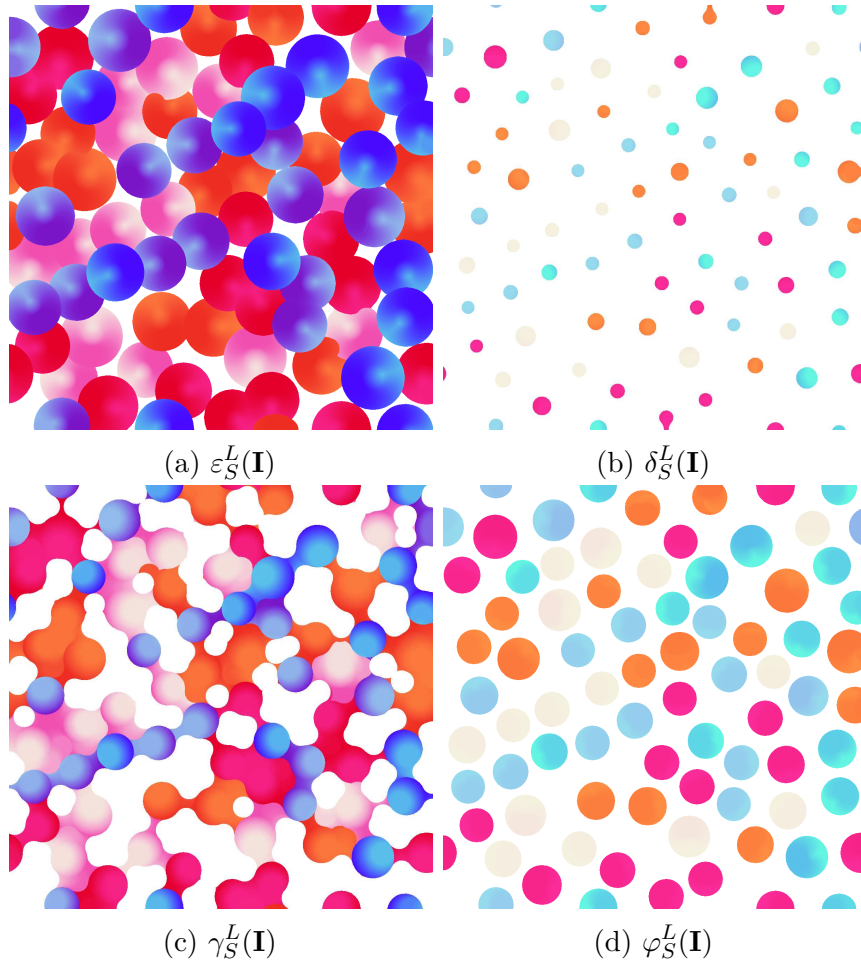


Figure 10 – The erosion, dilation, opening, and closing of the image \mathbf{I} from figure 8(a) by a disk of radius 30 with the lexicographical ordering.

This states that the topology produced by the total ordering cannot reproduce the natural topology of a vector-valued set. A consequence of this lemma is that, for any given total ordering on a vector-valued set, it is possible to find an image where the erosion and the dilation introduce important irregularities.

4 Mathematical Morphology on Reduced Orderings

In a reduced ordering, elements are ranked according to the value of a mapping h . Reduced orderings were interpreted in the context of multivalued mathematical morphology by Goutsias et al. (1995). Reduced orderings have been built using a variety of different approaches, for example, Louverdis et al. (2002) proposed a reduced ordering based on fuzzy IF-THEN rules, Chanussot and Lambert (1998) proposed a reduced ordering based on space-filling curves, and Velasco-Forero and Angulo (2012) proposed an ordering based on anomalies in the images. There are some approaches that take sets of colors, or values, as input to compute the mapping h . For example, Comer and Delp (1999), Goutsias et al. (1995) and Velasco-Forero and Angulo (2014) proposed orderings based on the Euclidean, Mahalanobis and kernel-induced distances to a reference value, respectively. It turns out that these distance-based orderings can be seen as a special case of the SVM supervised ordering (Velasco-Forero and Angulo, 2011a).

In this chapter, we will study the fundamental definitions and propositions of reduced orderings and look at some of the approaches available in the literature, with focus on the supervised orderings.

4.1 Mathematical Morphology on Reduced Orderings

In the case of a multivalued value set \mathbb{V} , when there is no natural ordering, one of the possible approaches is defining a reduced ordering. The advantage reduced orderings is that they can reduce a vector-valued image to a gray-scale image and still retain the useful information, depending on the ordering mapping h that was used. Indeed, by using a total reduced ordering on a finite value set, which is the case with digital multivalued images, we can compute morphological operators while avoiding false values, and without having to resort to prioritizing one dimension, as in the lexicographical ordering. Nonetheless, in practice, ordering mappings are bijective when restricted to the range of a digital image, so total reduced orderings still suffer from the issue of irregularity (Chevallier and Angulo, 2016).

Since our object of interest are images, we will assume that the sets \mathcal{V} and \mathcal{L} are obtained from $\mathcal{V} = \mathbb{V}^{\mathcal{D}}$ and $\mathcal{L} = \mathbb{L}^{\mathcal{D}}$, where \mathbb{L} is a complete lattice. In the case that $h : \mathbb{V} \rightarrow \mathbb{L}$ is given, we obtain $h : \mathcal{V} \rightarrow \mathcal{L}$ by defining for all $x \in \mathcal{D}$, $\mathbf{I} \in \mathcal{V}$, $h(\mathbf{I})(x) = h(\mathbf{I}(x))$.

Erosions and dilations are then obtained by a generalization of the adjunction

(Goutsias et al., 1995).

Definition 12. Let \mathcal{L} be a complete lattice, $h : \mathcal{V} \rightarrow \mathcal{L}$ a surjective function, $\varepsilon^h : \mathcal{V} \rightarrow \mathcal{V}$ and $\delta^h : \mathcal{V} \rightarrow \mathcal{V}$. We say that $(\varepsilon^h, \delta^h)$ is an h -adjunction if

$$(\forall \mathbf{I}, \mathbf{J} \in \mathcal{V}) \quad \delta^h(\mathbf{I}) \leq_h \mathbf{J} \iff \mathbf{I} \leq_h \varepsilon^h(\mathbf{J}). \quad (4.1)$$

In this case we say that ε^h is an h -erosion and that δ^h is an h -dilation.

Because the infimum and the supremum are not well defined in a reduced ordering, we can not use (3.15) and (3.16) to obtain h -erosions and h -dilations. However there is a way to obtain h -erosion and h -dilations on \mathcal{V} from erosions and dilations on \mathcal{L} . Indeed, if $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}$ is an erosion and $\varepsilon^h : \mathcal{V} \rightarrow \mathcal{V}$ is such that $\varepsilon^h \xrightarrow{h} \varepsilon$, as defined in Proposition 1.9, then ε^h is an h -erosion (Goutsias et al., 1995). An analogous result applies for dilations.

An efficient way to compute an h -increasing operator from a corresponding gray-scale increasing operator is by using a lookup table. Suppose that \mathcal{D} is finite and that we have a increasing operator ψ on \mathcal{L} that does not introduce new values, that is for all $\mathbf{I} \in \mathcal{L}$, $\text{range}(\psi(\mathbf{I})) \subseteq \text{range}(\mathbf{I})$. Algorithm 1, written in Matlab notation, proposed by Velasco-Forero and Angulo (2014), yields from ψ an operator ψ^h such that $\psi^h \xrightarrow{h} \psi$. The inputs to the algorithm are an $n_1 \times n_2 \times k$ -dimensional array `im` that represents a multivalued image \mathbf{I} on a $n_1 \times n_2$ grid, an $n_1 \times n_2$ -dimensional array `h_function`, that represents the gray-scale image $h(\mathbf{I})$, and a function `psi` that computes a flat gray-scale operator ψ . The algorithm works by creating an integer-valued image `im_latt` such that the value i is assigned to the position with the i -th smallest value in `h_function`. Ties are resolved by the spatial position, if two pixels have the same value, the one that appears first when scanning the image will be the one with the least value in `im_latt`. The increasing gray-scale operator is computed on `im_latt`, resulting in `im_psi`. The output multivalued `im_out` is obtained by applying the inverse operation used to obtain `im_latt`. We have that `im_out` is the digital representation of the image $\psi^h(\mathbf{I})$ such that $\psi^h \xrightarrow{h} \psi$.

Similarly to operators on complete lattices, we say that an operator $\psi^h : \mathcal{V} \rightarrow \mathcal{V}$ is:

- h -idempotent if $(\psi^h)^2 =_h \psi^h$;
- h -antiextensive if $\psi^h \leq_h \text{id}$;
- h -extensive if $\text{id} \leq_h \psi^h$.

An operator γ^h is an h -opening if it is h -increasing, h -idempotent and h -antiextensive. Similarly, φ^h is a h -closing if φ^h is h -increasing, h -idempotent and h -extensive. Analogous

Algorithm 1 – Computation of an h -increasing operator from a corresponding increasing gray-scale operator. In this algorithm, the sort function returns a sorted vector and the index permutation. The reshape function reshapes an array to a specific dimension.

Input: Multivariate $n_1 \times n_2 \times k$ image **im**, flat increasing gray-scale operator **psi** and the preorder function, represented by a vector **h_function**, with size $n_1 \times n_2$.

```

im = reshape(im,  $n_1 n_2$ ,  $k$ ) ;
h_function = reshape(h_function,  $n_1 n_2$ ) ;
[b, idx] = sort(h_function) ;
im_latt(b) = 1 : ( $n_1 n_2$ ) ;
im_latt = reshape(im_latt,  $n_1$ ,  $n_2$ ) ;
im_psi = psi(im_latt) ;
im_out = reshape(im(b(im_psi(:)), :,  $n_1$ ,  $n_2$ ,  $k$ ) ;

```

to erosions and dilations, if $\gamma : \mathcal{L} \rightarrow \mathcal{L}$ is an opening and $\gamma^h \xrightarrow{h} \gamma$ then γ^h is an h -opening. Likewise, if $\varphi : \mathcal{L} \rightarrow \mathcal{L}$ is a closing and $\varphi^h \xrightarrow{h} \varphi$ then φ^h is an h -closing. We can then use Algorithm 1 with a gray-scale opening (or closing) of $h(\mathbf{I})$ given as input to an h -opening (or h -closing) of \mathbf{I} . If $(\varepsilon^h, \delta^h)$ is an h -adjunction, then an h -opening and an h -closings are obtained from $\varepsilon^h \delta^h$ and $\delta^h \varepsilon^h$, respectively.

Here we will make a distinction between some types of reduced orderings that are seen in the literature, namely, supervised and unsupervised orderings. In supervised orderings the function h depends on sets of values that represent what is meant to be the foreground and the background pixels of an image. In an unsupervised ordering, the function h depends only on the input image \mathbf{I} .

In the next section we will look at the literature for some examples of supervised orderings which are the focus of this work.

4.2 Supervised Orderings

When some information about the ordering is provided, such as values that are foreground or background, we may use a supervised ordering (Velasco-Forero and Angulo, 2011a). Precisely, the input to a supervised ordering is a set of foreground values $F = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ and a set of background values $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. It is desirable that the values $h(\mathbf{f}_i)$ are the highest possible value of the lattice and the values of $h(\mathbf{b}_j)$ the lowest possible values, that is

$$h(\mathbf{f}_i) = \top, \quad \forall i = 1, \dots, m, \quad \text{and} \quad h(\mathbf{b}_j) = \perp, \quad \forall j = 1, \dots, n, \quad (4.2)$$

where $\top = \bigvee h(\mathbb{V})$ and $\perp = \bigwedge h(\mathbb{V})$ denote respectively the largest and the least values of the image of the mapping h .

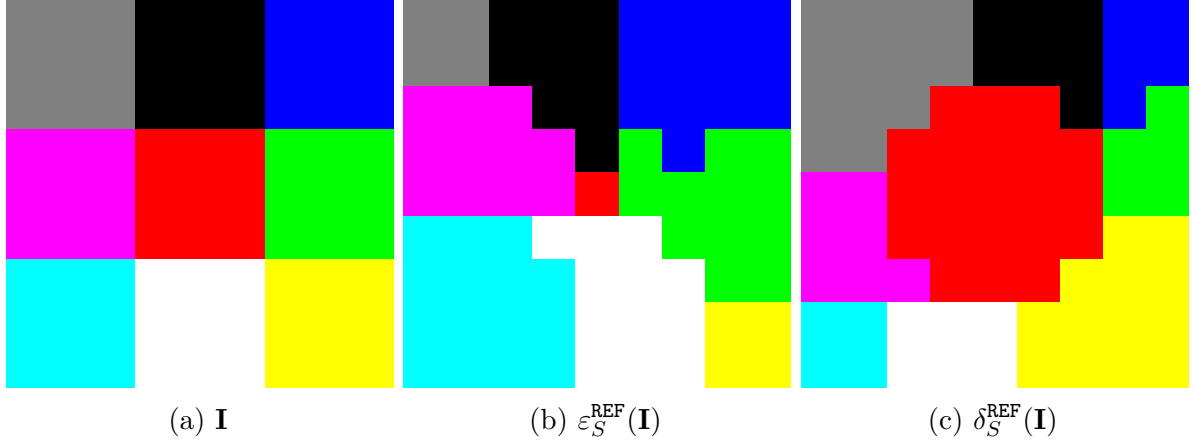


Figure 11 – An illustration of the effect of the distance based ordering. A 9×9 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its h_{REF} -erosion and h_{REF} -dilation were computed, using the red of the original image as the reference color, the euclidean distance and a cross-shaped structuring element given by $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

4.2.1 The Distance-Based Ordering

The distance-based approach ranks the elements based on the distance to a fixed reference value \mathbf{r} (Angulo, 2007). Most generally, the mapping h_{REF} is obtained from the distance from a reference color \mathbf{r} obtained from a metric d , e.g. the euclidean distance (Comer and Delp, 1999) or the Mahalanobis distance (Goutsias et al., 1995),

$$h_{\text{REF}}(\mathbf{x}) = -d(\mathbf{x}, \mathbf{r}). \quad (4.3)$$

In the ordering given by (4.3), the greatest element in the value set is the reference value, and the further a value is from the reference value, the lower it is. This is exemplified on Figure 11, where the h_{REF} -erosion and h_{REF} -dilation were computed with red as the reference value, using Algorithm 1. We can notice how the red part of the image has grown in the dilation and shrunk in the erosion. The greatest value can be controlled by the parameter \mathbf{r} - the maximum will always be \mathbf{r} - but the least value, although being the farthest value to the reference, has no simple interpretation (Velasco-Forero and Angulo, 2014).

We will regard d as a kernel induced distance (Velasco-Forero and Angulo, 2014). For a given kernel $\mathbf{K}(\mathbf{x}, \mathbf{y}) = \langle \phi(\mathbf{x}), \phi(\mathbf{y}) \rangle$ its induced distance is given by $d(\mathbf{x}, \mathbf{y}) = \|\phi(\mathbf{x}) - \phi(\mathbf{y})\|_2$, which is equivalent to

$$d(\mathbf{x}, \mathbf{y})^2 = \mathbf{K}(\mathbf{x}, \mathbf{x}) + \mathbf{K}(\mathbf{y}, \mathbf{y}) - 2\mathbf{K}(\mathbf{x}, \mathbf{y}). \quad (4.4)$$

This ordering scheme is able to be computed for a training set of values $T = \{\mathbf{t}_1, \dots, \mathbf{t}_n\}$. Velasco-Forero and Angulo (2014) used local linear combinations of the

type

$$h_{\text{REF}}(\mathbf{x}) = \sum_{i=1}^n \lambda_i(\mathbf{x}) d(\mathbf{x}, \mathbf{t}_i), \quad (4.5)$$

where, $\lambda_i(\mathbf{x}) \neq 0$ only if $i = \operatorname{argmin}_j d(\mathbf{x}, \mathbf{t}_j)$. We will use a similar approach, where

$$h_{\text{REF}}(\mathbf{x}) = \bigwedge_{i=1}^n d(\mathbf{x}, \mathbf{t}_i), \quad (4.6)$$

which is the distance from the set T to the vector \mathbf{x} .

In this approach, we have no direct control over the position of the small values in the value set, the background set is always empty. By considering $F = T$ and $B = \emptyset$, the conditions from (4.2) are always satisfied for h_{REF} determined by (4.6).

4.2.2 The SVM-Based Ordering

The SVM-supervised ordering (Velasco-Forero and Angulo, 2011a, 2014) assigns the foreground set $F = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ with positive values and the background set $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ with negative labels and makes use of the SVM to compute an ordering mapping h_{SVM} . By using the SVM formulation in this data, the resulting problem becomes

$$\begin{aligned} & \text{maximize} \sum_{i=1}^m \lambda_{\mathbf{f}_i} + \sum_{j=1}^n \lambda_{\mathbf{b}_j} - \frac{1}{2} \sum_{i,l=1}^m \lambda_{\mathbf{f}_i} \lambda_{\mathbf{f}_l} \mathbf{K}(\mathbf{f}_i, \mathbf{f}_l) - \frac{1}{2} \sum_{j,l=1}^n \lambda_{\mathbf{b}_j} \lambda_{\mathbf{b}_l} \mathbf{K}(\mathbf{b}_j, \mathbf{b}_l) \\ & \quad + \frac{1}{2} \sum_{k=1}^m \sum_{i=1}^n \lambda_{\mathbf{f}_k} \lambda_{\mathbf{b}_i} \mathbf{K}(\mathbf{f}_k, \mathbf{b}_i) \\ & \text{subject to} \sum_{i=1}^m \lambda_{\mathbf{f}_i} - \sum_{j=1}^n \lambda_{\mathbf{b}_j} = 0 \quad \text{and} \quad 0 \leq \lambda_{\mathbf{f}_i}, \lambda_{\mathbf{b}_j} \leq C, \end{aligned} \quad (4.7)$$

The ordering mapping h_{SVM} , derived from (2.6), is given by

$$h_{\text{SVM}}(\mathbf{x}) = \sum_{i=1}^m \lambda_{\mathbf{f}_i} \mathbf{K}(\mathbf{x}, \mathbf{f}_i) - \sum_{j=1}^n \lambda_{\mathbf{b}_j} \mathbf{K}(\mathbf{x}, \mathbf{b}_j) + \beta_0. \quad (4.8)$$

In the simplest case, where there is only one foreground value \mathbf{f} and one background value \mathbf{b} , the ordering mapping h_{SVM} is given by

$$h_{\text{SVM}}(\mathbf{x}) = \frac{\mathbf{K}(\mathbf{f}, \mathbf{x}) - \mathbf{K}(\mathbf{b}, \mathbf{x})}{\mathbf{K}(\mathbf{x}, \mathbf{x}) - \mathbf{K}(\mathbf{f}, \mathbf{b})}. \quad (4.9)$$

The mapping h_{SVM} does not necessarily satisfy (4.2). We can see this by using $\mathbb{V} = \mathbb{R}^2$, a Gaussian RBF kernel with $\sigma = 1$, $\mathbf{f} = (0.5, 0.5)$ and $\mathbf{b} = (1, 1)$. With this, and using (4.9), we have $h_{\text{SVM}}(0, 0) \approx 8.84 > 1 = h_{\text{SVM}}(\mathbf{f})$.

An illustrative example of the h_{SVM} -erosion and the h_{SVM} -dilation is shown on Figure 12. In this example we can see that the model behaved as expected, that is, in the erosion, the foreground color, red, was shrunk while the background color, green, was expanded. Conversely, on the dilation, red was expanded and green was shrunk.

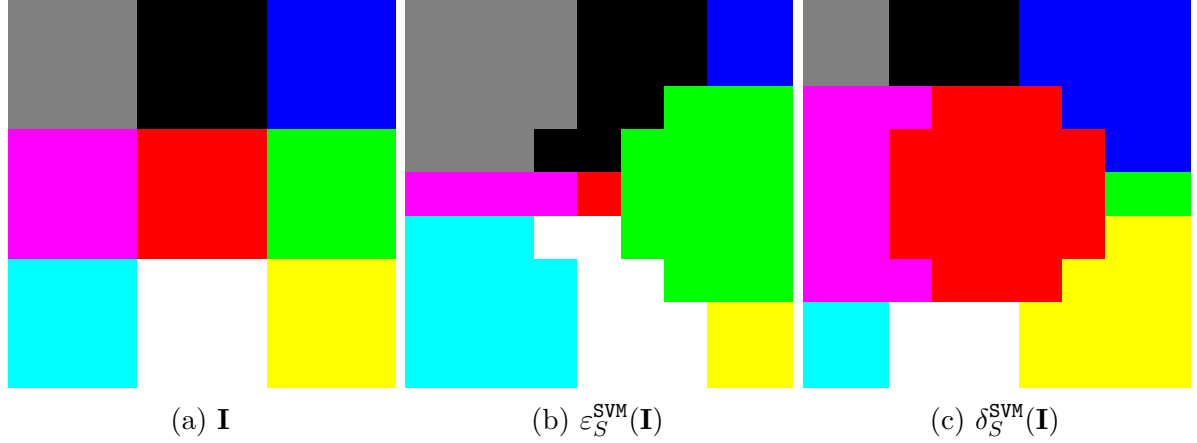


Figure 12 – An illustration of the effect of the SVM-based ordering. A 9×9 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its h_{SVM} -erosion and h_{SVM} -dilation were computed, using the red of the original image as the foreground color and the green as the background color, a Gaussian radial basis kernel and a cross-shaped structuring element given by $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

4.2.3 The LAAM-Based Ordering

Graña and Chyzhyk (2016) proposed an h -ordering based on the LAAM and the idea of a classifier that uses the Chebyshev distance between the input vector and the recalled vector (Sussner and Valle, 2006). Given a set of foreground values $F = \{\mathbf{f}^\xi\}_{\xi=1}^N$, the mapping h_F proposed by Graña and Chyzhyk (2016) at a value \mathbf{x} is given by the Chebyshev distance from the vector recalled by either the dilative LAAM or by the erosive LAAM and the input. Formally,

$$h_X(\mathbf{x}) = -d_{\text{C}}(\mathbf{x}^\#, \mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{V}. \quad (4.10)$$

where $\mathbf{x}^\# = M_{FF} \boxtimes \mathbf{x}$ or $\mathbf{x}^\# = W_{FF} \boxtimes \mathbf{x}$.

Although the mapping h_F only takes into account foreground values, it can be easily adapted to take into account both foreground and background values. Indeed, given a set F of foreground values and a set B of background values (Graña and Chyzhyk, 2016), the mapping h_{LAAM} is given by

$$h_{\text{LAAM}}(\mathbf{x}) = h_F(\mathbf{x}) - h_B(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbb{V}. \quad (4.11)$$

This mapping also does not always satisfy (4.2). As a counter-example, consider the foreground value $\mathbf{f} = (0.9, 0.1)$ and the background value $\mathbf{b} = (0.1, 0.9)$ and the input point $\mathbf{x} = (1, 0)$. The value $h_{\text{LAAM}}(\mathbf{x}) = 0.8$ is greater than $h_{\text{LAAM}}(\mathbf{f}) = 0.7$.

An illustrative example of the lattice auto-associative memory based ordering is shown on Figure 13. As expected, the erosion shrunk the foreground color while expanding the background color, while the dilation expanded the foreground color while shrinking the background color.

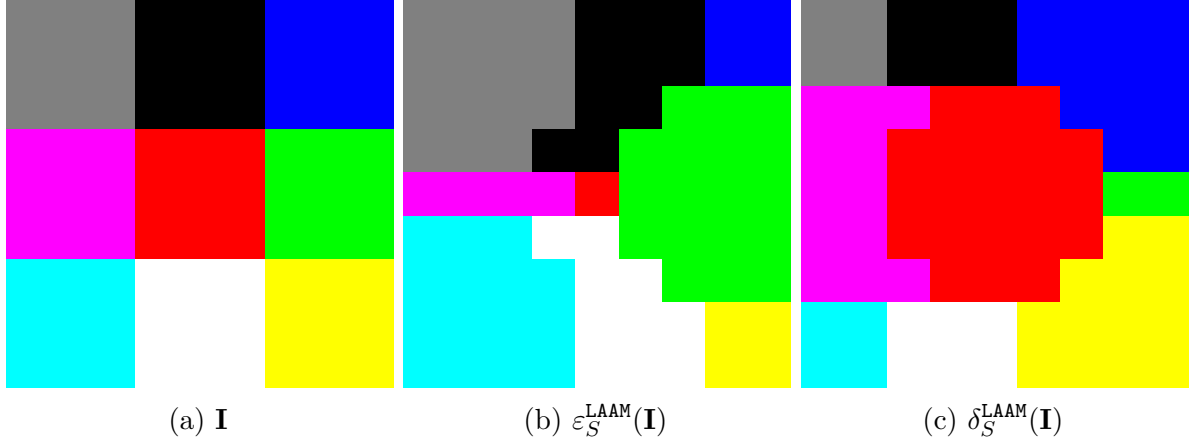


Figure 13 – An illustration of the effect of the LAAM-based ordering. A 9×9 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its h_{LAAM} -erosion and h_{LAAM} -dilation were computed, using the red of the original image as the foreground color and the green as the background color and a cross-shaped structuring element given by $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

4.2.4 Comparing Supervised Orderings

Let us provide an illustrative comparison between the supervised orderings using the natural color image, shown in Figure 14(a). Consider the foreground and background colors given, respectively, by the color values inside the green and red bullets in Figure 14(b).

The segmentation is computed as follows: Given an image \mathbf{I} and sets of foreground and background colors, denoted by F and B , and a supervised ordering mapping h that is trained on these sets, we compute a filtered version of h , denoted $\psi(h(\mathbf{I}))$. In this case the filter is given by the composition of an opening by reconstruction and a closing by reconstruction, both with a disk structuring element of radius 5. The image is submitted to a Otsu threshold (Otsu, 1979), yielding a binary image $B = \mathcal{X}_{>t}(\psi(h(\mathbf{I})))$, where t is the threshold obtained by the Otsu's method and $\mathcal{X}_{>t}$ is the threshold function by this value. Finally the edges are obtained by an inner gradient on the binary image B , given by $E = \rho'_S(B) = B - \varepsilon_S(B)$, where $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$. The results are evaluated by Pratt's Figure of Merit (FoM) (Abdou and Pratt, 1979) with respect to the ground truth of the original image. Recall that for an image of edges E and a ground truth GT , the FoM can be computed by

$$\text{FoM}(E, \text{GT}) = \frac{1}{\max\{\text{card}(E), \text{card}(\text{GT})\}} \sum_{x \in E} \frac{1}{1 + \alpha d(x, \text{GT})^2}, \quad (4.12)$$

where α is a scaling constant and was set as $\alpha = 1$ and d is the euclidean distance between x and the ground truth set GT , that is $d(x, \text{GT}) = \inf\{d(x, y) | y \in \text{GT}\}$.

In the Berkeley Segmentation Dataset, the ground truth images corresponds to the boundaries detected by humans. There are more than one human-segmented ground



Figure 14 – (a)An image from the Berkeley Segmentation Dataset (Martin et al., 2001) and (b)the image plus the labeled foreground(green) and background(red) colors.

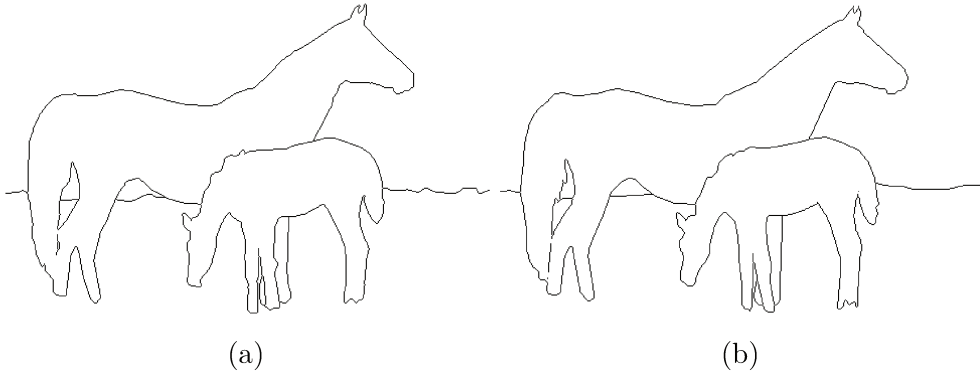


Figure 15 – Two different ground truths for the image from Figure 14(a).

truth images for each natural image in the dataset. Figure 15 shows two different ground truths for the image shown in Figure 14(a). To evaluate the results, the average over the FoM values of the detected edges and each of the ground truths of the image shown in Figure 14(a) is used.

The aim of this experiment is to measure how well these supervised orderings can separate the foreground from the background on the example image. As we can see, this method does not work well on separating foreground objects from other foreground objects, as well as separating background objects from other background objects.

The results of the segmentation for the h_{LAAM} and h_{SVM} orderings are shown in Figure 16. The h_{SVM} used a Gaussian RBF kernel with $\gamma = \frac{1}{3\text{Var}(X)}$, where X is given by the concatenation of $(\mathbf{I}(x_1)_1, \dots, \mathbf{I}(x_N)_1, \mathbf{I}(x_1)_2, \dots, \mathbf{I}(x_N)_3)$ and $\{x_i | i \in \{1, \dots, N\}\} = \mathcal{D}$. In words, X is the flattened image \mathbf{I} .

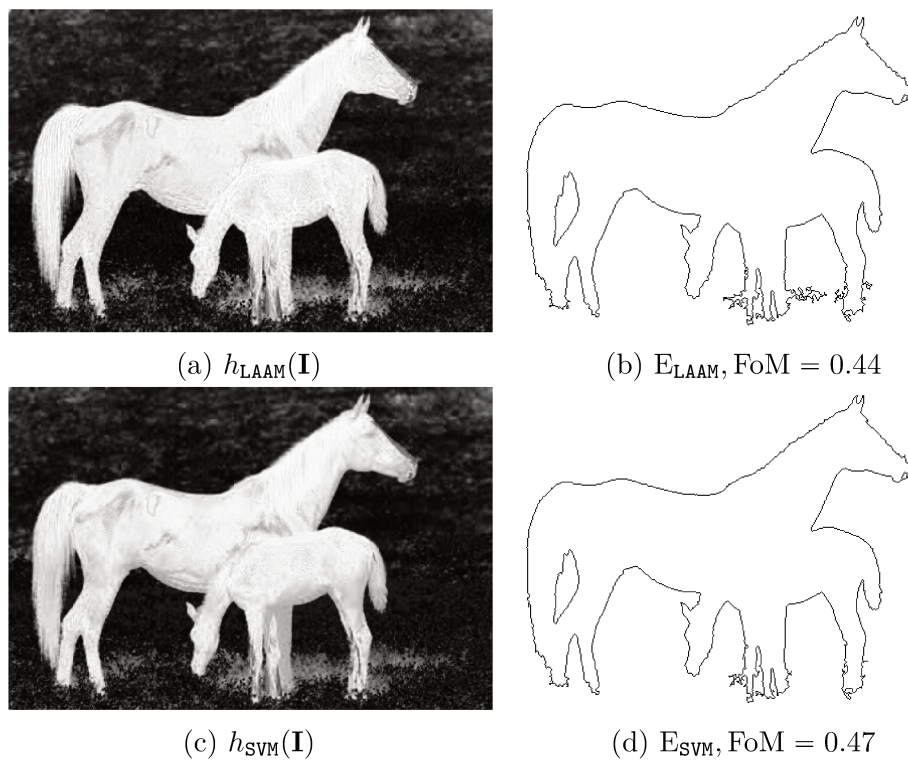


Figure 16 – Supervised orderings and segmentations of Figure 14(a) by different approaches. The FoM values are shown below each of the edges images.

5 Fuzzy Value-Based Mathematical Morphology

Fuzzy colors (Chamorro-Martínez et al., 2017) use fuzzy sets in color spaces to find a suitable mapping between the digital representation of colors and the color terms used by humans, they also solve the problem known as the semantic gap, a problem that involves issues like defining color similarity and modeling color properties.

In this chapter we will review and explore the reduced ordering based on fuzzy sets introduced by Sangalli and Valle (2018). More specifically, this ordering is given by the membership function of a fuzzy set obtained by the combination of several fuzzy colors, or fuzzy values, in the case of other types of multivalued images, by using fuzzy connectives such as t-norms, s-norms and uninorms. The intended result is a reduced ordering that can be defined by a human user in an interactive way and that can be easily interpreted. Furthermore, we apply to our reduced ordering the fuzzy partition based on Voronoi diagrams proposed by Chamorro-Martínez et al. (2017) and also extend our approach to a neuro-fuzzy framework, which aims to obtain the reference values by minimizing a loss function, instead of using the user's input.

5.1 Ordering Mappings Based on Membership Functions

Like the h_{SVM} -ordering proposed by Velasco-Forero and Angulo, in this section we generalize the distance-based approach by considering sets B and F of background and foreground values. However, instead of defining the mapping h in terms of the solution of a quadratic problem, we use concepts from fuzzy set theory, namely, the concept of fuzzy color:

Definition 13 (Chamorro-Martínez et al. (2017)). *A fuzzy color on a color space \mathbb{V} is a linguistic label whose semantics are given by a normal fuzzy subset of \mathbb{V} .*

Fuzzy colors attempt to model colors by considering the imprecision, subjectivity, and context dependency widely used by humans to describe them. Since this approach can be applied to other types of multivalued images other than color images, we will use the terms fuzzy value to refer to a normal fuzzy set on \mathbb{V} .

By using a fuzzy set H to characterize the values of an image, we are able to use its membership function $h_{\text{FUZZY}} := \mu_H$ as an ordering mapping, and thus rank the values according to some concept, that may be supplied by an user or learned by a machine. One straightforward approach to this is to assume that foreground and background fuzzy

sets are available, analogous to a supervised ordering. In other words, assume that the foreground and background fuzzy sets \tilde{F} and \tilde{B} are known. To obtain the value $h_{\text{FUZZY}}(\mathbf{x})$ for $\mathbf{x} \in \mathbb{V}$ it makes sense to consider the fuzzy proposition “the color \mathbf{x} is a foreground color and it is not a background color”. Thus, the function h_{FUZZY} can be determined by

$$h_{\text{FUZZY}}(\mathbf{x}) := \mu_{\tilde{F}}(\mathbf{x}) \triangle \eta(\mu_{\tilde{B}}(\mathbf{x})), \quad \forall \mathbf{x} \in \mathbb{V}, \quad (5.1)$$

where \triangle and η are a triangular norm and a strong fuzzy negation, respectively. Alternatively, the dual proposition is achieved by the negation of (5.1), but with the functions $\mu_{\tilde{F}}$ and $\mu_{\tilde{B}}$ also negated. The dual proposition is given by the following identity, for all $\mathbf{x} \in \mathbb{V}$

$$\begin{aligned} h_{\text{FUZZY}}(\mathbf{x}) &:= \eta(\eta(\mu_{\tilde{F}}(\mathbf{x})) \triangle \eta(\eta(\mu_{\tilde{B}}(\mathbf{x})))) \\ &= \eta(\eta(\mu_{\tilde{F}}(\mathbf{x})) \triangle \mu_{\tilde{B}}(\mathbf{x})) \\ &= \mu_{\tilde{F}}(\mathbf{x}) \nabla \eta(\mu_{\tilde{B}}(\mathbf{x})), \end{aligned} \quad (5.2)$$

where ∇ is a s-norm dual to \triangle . This definition also has a simple interpretation, namely “the color \mathbf{x} is a foreground color or it is not a background color”. In contrast to (5.1), the membership function given by (5.2) assigns high values to colors that are neither foreground nor background. It is also equivalent to $\mu_{\tilde{B}}(\mathbf{x}) \rightarrow \mu_{\tilde{F}}(\mathbf{x})$, where \rightarrow is a fuzzy s-implication. A more general approach is obtained by substituting the conjunction or disjunction of the fuzzy proposition in (5.1) and (5.2) by a generic connective that can be represented by a uninorm $*$

$$h_{\text{FUZZY}}(\mathbf{x}) := \mu_{\tilde{F}}(\mathbf{x}) * \eta(\mu_{\tilde{B}}(\mathbf{x})), \quad \forall \mathbf{x} \in \mathbb{V}, \quad (5.3)$$

The approaches of (5.1) and (5.2) treat the background and foreground differently. On one hand, (5.1) is more restrictive on what is considered foreground, or alternatively, biased toward the background because values with high membership in both foreground and background are effectively treated as background. On the other hand, (5.2) is the exact opposite: it is restrictive on what is considered background. In contrast, choosing a self-dual uninorm in (5.3) would result in a function that treats background and foreground equally, and could assign a intermediate degree of membership to values that have a high membership in both foreground and background, meaning uncertainty in its membership to both of them.

Another interesting property of (5.3) when $*$ is self-dual is that its negation, $\eta(h_{\text{FUZZY}})$ can be obtained by exchanging the roles of the background and the foreground sets, that is $\eta(\mu_{\tilde{F}}(\mathbf{x}) * \eta(\mu_{\tilde{B}}(\mathbf{x}))) = \mu_{\tilde{B}}(\mathbf{x}) * \eta(\mu_{\tilde{F}}(\mathbf{x}))$.

It turns out that the fuzzy sets \tilde{F} and \tilde{B} can be derived from crisp sets of foreground and background colors by obtaining a fuzzy set that is a fuzzyfication of each of the elements of these crisp sets. In other words, given a value $\mathbf{r} \in \mathbb{V}$, we can obtain a fuzzy set R from \mathbf{r} . One straightforward way to do this is to use a decreasing one-dimensional

Gaussian	$G(\mathbf{x}; \mathbf{r}, \gamma) = \exp(-\gamma d(\mathbf{x}, \mathbf{r})^2)$
Bell-shaped	$B(\mathbf{x}; \mathbf{r}, a, b) = \frac{1}{1 + \left(\frac{d(\mathbf{x}, \mathbf{r})}{a}\right)^{2b}}$
Conical	$C(\mathbf{x}; \mathbf{r}, c) = 0 \vee \left(1 - \frac{d(\mathbf{x}, \mathbf{r})}{c}\right)$

Table 2 – Some membership functions for fuzzy values that can be obtained by using a unidimensional membership function on the distance to a reference \mathbf{r} . The parameters are $\gamma \in \mathbb{R}_{>0}$ $a, b, c \in \mathbb{R}$.

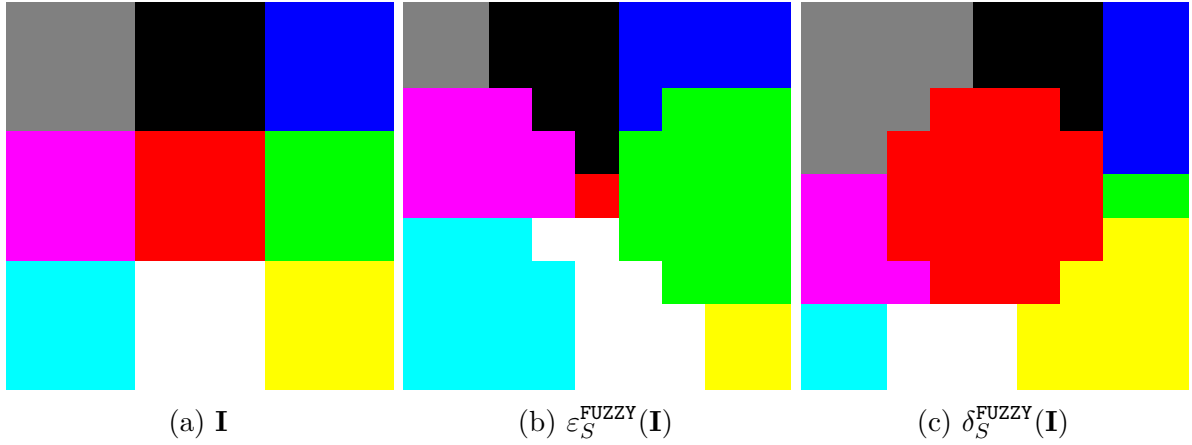


Figure 17 – An illustration of the effect of the fuzzy proposition based ordering. A 9×9 image \mathbf{I} in the \mathbb{V}_{RGB} color space and its h_{FUZZY} -erosion and h_{FUZZY} -dilation were computed, using a gaussian membership function with the red of the original image as the foreground set and with the green as the background set, the negation, uninorm and s-norm used were $1 - x$, \wedge and \vee , and a cross-shaped structuring element given by $S = \{(0, 0), (1, 0), (0, 1), (-1, 0), (0, -1)\}$.

membership function of the distance to the reference color, that is, given a decreasing function $f: \mathbb{R}^{\geq 0} \rightarrow [0, 1]$, the membership function with respect to a reference color \mathbf{r} is given by $\mu_R(\mathbf{x}) = f(d(\mathbf{x}, \mathbf{r}))$. Examples of such functions are given in Table 2.

Figure 17 shows the erosion and dilation with this ordering scheme. Using colors from the image \mathbf{I} from figure 17a) and a Gaussian membership function with $\sigma = 1$ to construct the foreground and background sets.

The fuzzy sets \tilde{F} and \tilde{B} can be obtained by a set of foreground and background values. Let $\mathbf{f}_1, \dots, \mathbf{f}_m$ be the foreground values and $\mathbf{b}_1, \dots, \mathbf{b}_n$ the background values and let F_1, \dots, F_m and B_1, \dots, B_m be fuzzy values obtained from the foreground and background in a manner such as the ones from Table 2. It makes sense to define the fuzzy set \tilde{F} as the union of the foreground fuzzy values and the fuzzy set \tilde{B} as the union of the background

values. Given a triangular-conorm ∇ , \tilde{F} and \tilde{B} can be obtained as follows, for all $\mathbf{x} \in \mathbb{V}$

$$\mu_{\tilde{F}}(\mathbf{x}) = \mu_{F_1}(\mathbf{x}) \nabla \cdots \nabla \mu_{F_m}(\mathbf{x}), \quad (5.4)$$

$$\mu_{\tilde{B}}(\mathbf{x}) = \mu_{B_1}(\mathbf{x}) \nabla \cdots \nabla \mu_{B_n}(\mathbf{x}). \quad (5.5)$$

As shown by the next theorem is possible to guarantee, under some circumstances, that the the ordering mapping will satisfy (4.2), that is

$$h(\mathbf{f}_i) = \top, \quad \forall i = 1, \dots, m, \quad \text{and} \quad h(\mathbf{b}_j) = \perp, \quad \forall j = 1, \dots, n.$$

Theorem 1. *Let $\mathbf{f}_1, \dots, \mathbf{f}_m$ be foreground values, $\mathbf{b}_1, \dots, \mathbf{b}_n$ be background values and let \tilde{F} and \tilde{B} be the foreground and background fuzzy sets. If, for all $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$, we have $\mu_{\tilde{F}}(\mathbf{f}_i) = 1$, $\mu_{\tilde{B}}(\mathbf{b}_j) = 1$, then the following affirmations hold:*

1. *If, for all $i \in \{1, \dots, m\}$, $\mu_{\tilde{B}}(\mathbf{f}_i) = 0$ then the mapping h_{FUZZY} , defined by (5.1), satisfies (4.2);*
2. *If, for all $j \in \{1, \dots, n\}$, $\mu_{\tilde{F}}(\mathbf{b}_j) = 0$ then the mapping h_{FUZZY} , defined by (5.2), satisfies (4.2);*
3. *If, for all $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$, $\mu_{\tilde{B}}(\mathbf{f}_i) = 0$ and $\mu_{\tilde{F}}(\mathbf{b}_j) = 0$ then the mapping h_{FUZZY} , defined by (5.3), satisfies (4.2);*

Proof. Let us consider each case separately

1. For $i \in \{1, \dots, m\}$ assume $\mu_{\tilde{B}}(\mathbf{f}_i) = 0$. We have

$$\begin{aligned} h_{\text{FUZZY}}(\mathbf{f}_i) &= \mu_{\tilde{F}}(\mathbf{f}_i) \triangle \eta(\mu_{\tilde{B}}(\mathbf{f}_i)) \\ &= 1 \triangle \eta(0) \\ &= 1 \triangle 1 \\ &= 1. \end{aligned}$$

Now for $j \in \{1, \dots, n\}$ we have

$$\begin{aligned} h_{\text{FUZZY}}(\mathbf{b}_j) &= \mu_{\tilde{F}}(\mathbf{b}_j) \triangle \eta(\mu_{\tilde{B}}(\mathbf{b}_j)) \\ &= \mu_{\tilde{F}}(\mathbf{b}_j) \triangle \eta(1) \\ &= \mu_{\tilde{F}}(\mathbf{b}_j) \triangle 0 \\ &= 0. \end{aligned}$$

2. The proof of the second affirmation comes from the dual of the first affirmation.

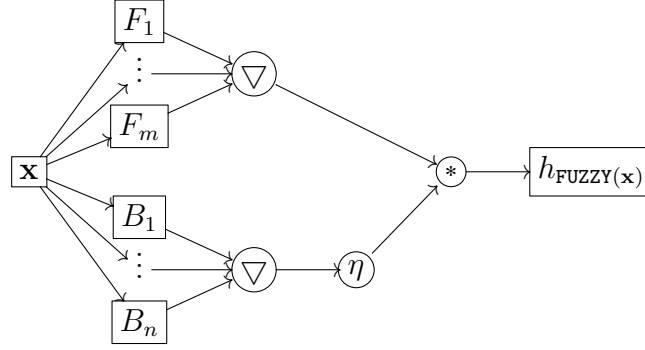


Figure 18 – Graph representation of the proposed supervised ordering.

3. For all $i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$ assume $\mu_{\tilde{B}}(\mathbf{f}_i) = 0$ and $\mu_{\tilde{F}}(\mathbf{b}_j) = 0$. We have, for $i \in \{1, \dots, m\}$

$$\begin{aligned}
 h_{\text{FUZZY}}(\mathbf{f}_i) &= \mu_{\tilde{F}}(\mathbf{f}_i) * \eta(\mu_{\tilde{B}}(\mathbf{f}_i)) \\
 &= 1 * \eta(0) \\
 &= 1 * 1 \\
 &= 1.
 \end{aligned}$$

Now for $j \in \{1, \dots, n\}$ we have

$$\begin{aligned}
 h_{\text{FUZZY}}(\mathbf{f}_i) &= \mu_{\tilde{F}}(\mathbf{b}_i) * \eta(\mu_{\tilde{B}}(\mathbf{b}_i)) \\
 &= 0 * \eta(1) \\
 &= 0 * 0 \\
 &= 0.
 \end{aligned}$$

□

5.2 Neuro-Fuzzy Approach

The combination of (5.4), (5.5) and (5.3) can be represented in the network of Figure 18. By using Gaussian membership functions (i.e. (2.34)) or another type of parameterizable membership function, one can obtain this supervised ordering from sets of foreground and background values similar to the ANFIS and the RBF networks, instead of directly using them as the Gaussian function centers. The advantage of this approach is that it does not need as many fuzzy sets as the number of foreground and background values. Additionally, the uninorm, s-norm and fuzzy negation may be parameterizable. In this approach, input and output pairs are of the form $(\mathbf{f}_i, 1)$ for foreground values and $(\mathbf{b}_j, 0)$ for background values. The parameters are obtained by minimizing a loss function.

The centers of the Gaussian functions obtained through this method are not necessarily close to the foreground and background values given as input. As an example, consider membership function $\mu_R(\mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{r}\|^2)$, where $\|\cdot\|$ is the L_2 norm, the \vee s-norm, \wedge uninorm and $1 - x$ negation, and the foreground and background values given by $\mathbf{f} = (1, 0, 0)$ and $\mathbf{b} = (0, 1, 0)$ respectfully. Using these values as centers we obtain $h_{\text{FUZZY}}(\mathbf{f}) \approx 0.865$ and $h_{\text{FUZZY}}(\mathbf{b}) = 0$. When obtaining parameters through optimization with respect to the mean squared error, and with the Adam optimizer (Kingma and Ba, 2014), from the tensorflow implementation¹, the centers obtained were $\mathbf{f}' \approx (0.733, 0.053, 0)$ and $\mathbf{b}' \approx (-0.148, 1.114, 0)$ and the h values $h_{\text{FUZZY}}(\mathbf{f}) \approx 0.928$ and $h_{\text{FUZZY}}(\mathbf{b}) = 0.042$.

The sets \tilde{F} and \tilde{B} obtained from (5.4) and (5.5) may fail to satisfy the hypothesis of Theorem 1, for instance, when using Gaussian and bell-shaped membership functions. Choosing $c \leq \min_{\mathbf{f} \in \tilde{F}, \mathbf{b} \in \tilde{B}} d(\mathbf{f}, \mathbf{b})$ guarantees that the conical membership function will satisfy Theorem 1. In the next section we will study a fuzzy partition of the value space introduced by Chamorro-Martínez et al. (2017) that can be built from a training set and satisfies the conditions of Theorem 1.

5.3 Voronoi Diagram Fuzzy Partition

Chamorro-Martínez et al. (2017) used the Voronoi diagram of a set of prototype colors to obtain a fuzzy color space - a set of fuzzy colors to use as linguistic terms in a color space. In this section we will make use of this approach to obtain foreground and background fuzzy sets from crisp sets of background and foreground values.

We first wish to obtain a fuzzy set R from a so-called positive prototype $\mathbf{r} \in F \cup B$ and $S = \{\mathbf{s}_1, \dots, \mathbf{s}_l, \mathbf{r}\} = F \cup B$ such that its α -levels are rescalings of the Voronoi cell corresponding to \mathbf{r} . For a given metric d , the Voronoi cell corresponding to \mathbf{r} in the Voronoi diagram of S is given by

$$C_{\mathbf{r}} = \{\mathbf{y} \in \mathbb{V} | d(\mathbf{y}, \mathbf{r}) \leq d(\mathbf{y}, \mathbf{s}_i), i = 1, \dots, l\}, \quad (5.6)$$

and let us denote the scaling of the cell $C_{\mathbf{r}}$ by a factor $\lambda \in \mathbb{R}^{\geq 0}$ centered at $\mathbf{r} \in \mathbb{R}^k$ by

$$\lambda \cdot C_{\mathbf{r}} = \{\mathbf{r} + \lambda(p - \mathbf{r}) | p \in C_{\mathbf{r}}\}. \quad (5.7)$$

In practice, we are concerned with the case where the $\lambda \in [0, 2]$. When $\lambda = 0$ the rescaling is simply \mathbf{r} and when $\lambda = 2$, the rescaling intersects with points from S other than \mathbf{r} (namely, with the points that define of its neighboring cells).

The Voronoi diagram can only create a crisp partition of the value set. To obtain the fuzzy partition we will require a continuous, surjective and decreasing interpolation

¹ <https://www.tensorflow.org/>

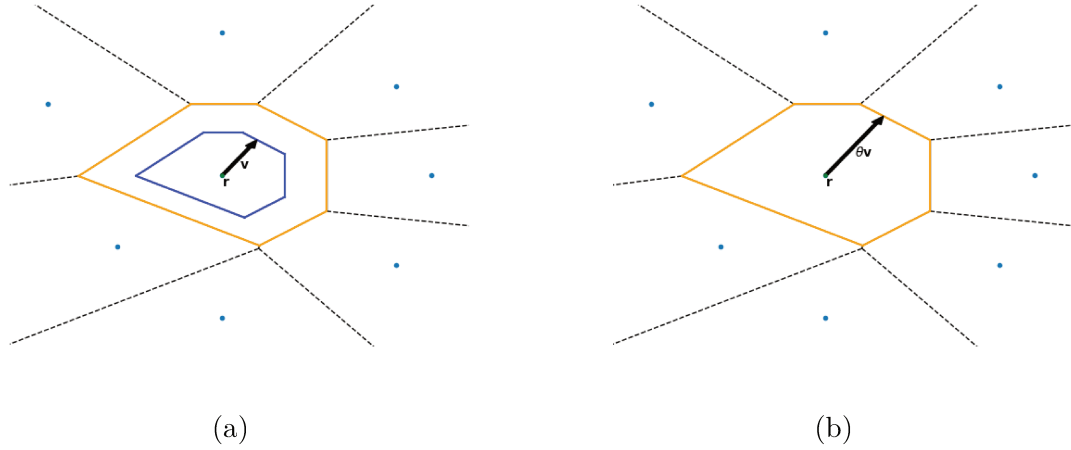


Figure 19 – A Voronoi diagram with the Voronoi cell of $C_{\mathbf{r}}$ (in orange) and the rescaled Voronoi cell $\lambda \cdot C_{\mathbf{r}}$ (in blue) with a vector \mathbf{v} such that $\mathbf{r} + \mathbf{v}$ is at the boundary of $\lambda \cdot C_{\mathbf{r}}$ (a) and the same diagram with $\mathbf{r} + \theta\mathbf{v}$, where $\theta = \frac{1}{\lambda}$, falling in the boundary of $C_{\mathbf{r}}$ (b).

function $f : \mathbb{R}^{\geq 0} \rightarrow [0, 1]$. Chamorro-Martínez et al. (2017) used a piecewise linear interpolation function. Let $f^{\leftarrow}(\alpha) = \sup\{\lambda \in \mathbb{R}^{\geq 0} \mid f(\lambda) = \alpha\}$. The membership function of R is defined in terms of its α -levels:

$$[R]^{\alpha} = f^{\leftarrow}(\alpha) \cdot C_{\mathbf{r}}, \quad \forall \alpha \in [0, 1]. \quad (5.8)$$

In words, a point that is at the border of the scaling of $C_{\mathbf{r}}$ by a factor λ has membership $f(\lambda)$ in R .

To explicitly compute the membership function μ_R , at a point \mathbf{x} we must find the greatest α such that $\mathbf{x} \in [\mu_R]^{\alpha}$. With this we have,

$$\begin{aligned} \mu_R(\mathbf{x}) &= \sup\{\alpha \in (0, 1] \mid \mathbf{x} \in f^{\leftarrow}(\alpha) \cdot C_{\mathbf{r}}\} \\ &= \sup\{f(\lambda) \mid \lambda \in \mathbb{R}^{>0}, \mathbf{x} \in \lambda \cdot C_{\mathbf{r}}\} \\ &= f(\inf\{\lambda \mid \lambda \in \mathbb{R}^{>0}, \mathbf{x} \in \lambda \cdot C_{\mathbf{r}}\}). \end{aligned} \quad (5.9)$$

Thus, we need to find the lowest λ such that $\mathbf{x} \in \lambda \cdot C_{\mathbf{r}}$.

Let us assume, for the moment, that the distance d is the Euclidean distance. In this case, consider $\mathbf{v} = \mathbf{x} - \mathbf{r}$, that is $\mathbf{x} = \mathbf{r} + \mathbf{v}$. We have that $\mathbf{r} + \frac{\mathbf{v}}{\lambda} \in C_{\mathbf{r}} \iff \mathbf{x} \in \lambda \cdot C_{\mathbf{r}}$. Find the least λ such that $\mathbf{x} \in \lambda \cdot C_{\mathbf{r}}$ is equivalent to finding the greatest θ such that $\mathbf{r} + \theta\mathbf{v} \in C_{\mathbf{r}}$, that is, the greatest θ such that

$$\|\mathbf{r} + \theta\mathbf{v} - \mathbf{r}\|^2 = \|\theta\mathbf{v}\|^2 \leq \|\mathbf{r} + \theta\mathbf{v} - \mathbf{s}\|^2$$

for all $\mathbf{s} \in S$. Figure 19 shows an example of such a \mathbf{v} with $\mathbf{r} + \mathbf{v}$ falling at the boundary of $\lambda \cdot C_{\mathbf{r}}$ and $\mathbf{r} + \theta\mathbf{v}$ at the boundary of $C_{\mathbf{r}}$. Writing this in the form of an optimization

problem gives us

$$\begin{aligned} & \text{maximize } \theta \\ & \text{subject to } \|\theta \mathbf{v}\|^2 \leq \|\mathbf{r} + \theta \mathbf{v} - \mathbf{s}_i\|^2, \quad i = 1, \dots, l \\ & \theta \geq 0. \end{aligned} \tag{5.10}$$

The right-hand side of the first restriction in (5.10) can be modified as follows

$$\begin{aligned} \|\mathbf{r} + \theta \mathbf{v} - \mathbf{s}_i\|^2 &= \langle \mathbf{r} + \theta \mathbf{v} - \mathbf{s}_i, \mathbf{r} + \theta \mathbf{v} - \mathbf{s}_i \rangle \\ &= \langle \mathbf{r} - \mathbf{s}_i, \mathbf{r} - \mathbf{s}_i \rangle + 2\langle \theta \mathbf{v}, \mathbf{r} - \mathbf{s}_i \rangle + \langle \theta \mathbf{v}, \theta \mathbf{v} \rangle \\ &= \|\mathbf{r} - \mathbf{s}_i\|^2 + 2\langle \theta \mathbf{v}, \mathbf{r} - \mathbf{s}_i \rangle + \|\theta \mathbf{v}\|^2, \end{aligned}$$

and that restriction can be modified to

$$\begin{aligned} \|\theta \mathbf{v}\|^2 &\leq \|\mathbf{r} + \theta \mathbf{v} - \mathbf{s}_i\|^2 \\ \iff \|\theta \mathbf{v}\|^2 &\leq \|\mathbf{r} - \mathbf{s}_i\|^2 + 2\langle \theta \mathbf{v}, \mathbf{r} - \mathbf{s}_i \rangle + \|\theta \mathbf{v}\|^2 \\ \iff 2|\theta| \langle -\mathbf{v}, \mathbf{r} - \mathbf{s}_i \rangle &\leq \|\mathbf{r} - \mathbf{s}_i\|^2 \\ \iff 2|\theta| \langle \mathbf{r} - \mathbf{x}, \mathbf{r} - \mathbf{s}_i \rangle &\leq \|\mathbf{r} - \mathbf{s}_i\|^2 \\ \iff |\theta| &\leq \frac{\|\mathbf{r} - \mathbf{s}_i\|^2}{2\langle \mathbf{r} - \mathbf{x}, \mathbf{r} - \mathbf{s}_i \rangle}. \end{aligned}$$

Therefore, the solution to (5.10), when it exists, is given by

$$\theta_{\mathbf{r}, S, \mathbf{x}}^* = \min_{\langle \mathbf{r} - \mathbf{x}, \mathbf{r} - \mathbf{s} \rangle \geq 0, \mathbf{s} \in S} \frac{\|\mathbf{r} - \mathbf{s}\|^2}{2\langle \mathbf{r} - \mathbf{x}, \mathbf{r} - \mathbf{s} \rangle}, \tag{5.11}$$

In practice, the minimum of (5.11) needs to be computed only over the neighboring cells of \mathbf{r} . The value of the membership function is given by

$$\mu_R(\mathbf{x}) = f\left(\frac{1}{\theta_{\mathbf{r}, S, \mathbf{x}}^*}\right), \tag{5.12}$$

and, when the problem (5.10) is unfeasible, for the sake of continuity, we define $\mu_R(\mathbf{x}) = 1$.

Notice that (5.11) can be computed only from inner products. We propose to take advantage of this to compute it in a feature space with the help of a kernel \mathbf{K} . This allows for a greater flexibility to the final shape of the membership functions. We can compute (5.11) in a feature space by

$$\theta_{\mathbf{r}, S, \mathbf{x}}^* = \min_{\mathbf{K}(\mathbf{r}, \mathbf{r}) - \mathbf{K}(\mathbf{r}, \mathbf{s}) - \mathbf{K}(\mathbf{x}, \mathbf{r}) + \mathbf{K}(\mathbf{x}, \mathbf{s}) \geq 0, \mathbf{s} \in S} \frac{1}{2} \frac{\mathbf{K}(\mathbf{r}, \mathbf{r}) + \mathbf{K}(\mathbf{x}, \mathbf{x}) - 2\mathbf{K}(\mathbf{x}, \mathbf{r})}{\mathbf{K}(\mathbf{r}, \mathbf{r}) - \mathbf{K}(\mathbf{r}, \mathbf{s}) - \mathbf{K}(\mathbf{x}, \mathbf{r}) + \mathbf{K}(\mathbf{x}, \mathbf{s})}. \tag{5.13}$$

The membership function corresponding to a class may be obtained from the union, using an s-norm, of the membership functions corresponding to the centers obtained from that class. Doing this, each class will have a corresponding fuzzy set

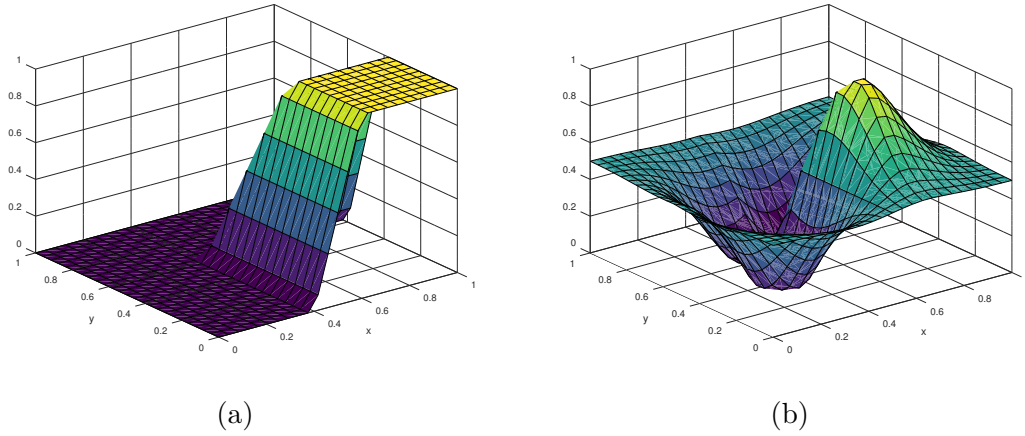


Figure 20 – The Voronoi membership function of a prototype $\mathbf{r} = (0.6, 0.4)$ and a set $S = \{(0.6, 0.6), (0.4, 0.4), (0.4, 0.6), (0.6, 0.4)\}$ using linear interpolating functions in $[0, 2]$ and (a) the inner product kernel function and (b) the Gaussian RBF kernel function with $c = 0.05$.

and a corresponding h -ordering. A foreground/background ordering may be obtained by classifying the data into those two classes, obtaining the fuzzy sets \tilde{F} and \tilde{B} and then computing (5.2).

An advantage of this approach is that if the interpolation function is 0 outside $[0, 2]$, that is $f([0, 2]^c) = \{0\}$, and $f(0) = 1$, it satisfies the conditions of Theorem 1, guaranteeing fulfillment of (4.2).

Figure 20 shows examples of this membership function obtained from the Voronoi diagram fuzzy partition using the inner product kernel and the Gaussian RBF. Notice how the membership function that uses the Gaussian kernel is much more able to account for uncertainty in the points that are neither foreground nor background than the inner product one.

Examples of supervised erosion and dilation with these approaches can be seen on Figure 21. In this case, the erosion obtained from the Gaussian Kernel, with $\gamma = 0.5$, achieved results identical to the ones obtained from the erosion obtained from the inner product. Similarly, the dilation obtained from the Gaussian kernel was identical to the dilation obtained from the inner product. We can see that the erosion and dilation behave as expected, that is, the erosion shrunk the foreground and expanded the background and the dilation expanded the foreground and shrunk the background.

5.4 Comparing the Fuzzy Set-Based Orderings

Using the approach described in the previous chapter to compute the segmentation of an image into foreground and background, we segmented the image shown in

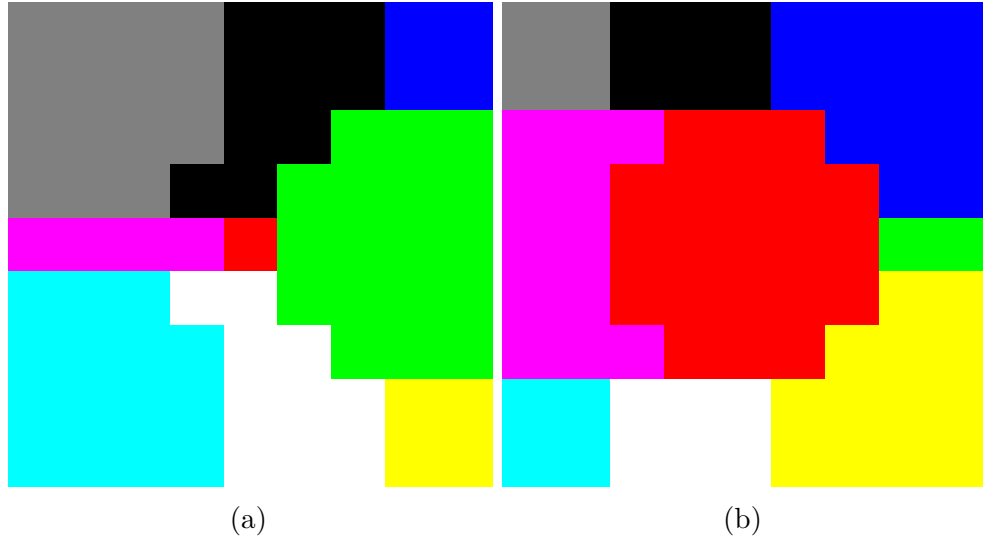


Figure 21 – Example of (a)erosion and (b)dilation with the inner product Voronoi membership function with a cross structuring element S using image from Figure 11(a) using red as foreground and green as background. The foreground and background were combined using the \wedge t-norm and the standard fuzzy negation.

Figure 14 for a variety of parameters. The mean FoM value between the segmentations and the ground truth images can be found in Table 3. As we can see, the fuzzy ordering with the gaussian membership function and the \vee uninorm and \vee s-norm were able to surpass the h_{svm} . In this section, $*$ is the self-dual uninorm given by (2.32) and the negation is given by $\eta(a) = 1 - a$.

The parameters for the Gaussian membership function, the Gaussian RBF and the bell-shaped function are all given by $a = \gamma = \frac{1}{3\text{Var}(X)}$ and $b = 1$ where X is given by the flattened image \mathbf{I} being processed.

Visual interpretation of the gray-scale images $h(\mathbf{I})$ and their corresponding segmentations are shown in Figure 22. Figures 22(a), (c) and (e) are computed using the same type of membership function for individual foreground and background fuzzy values, and the same s-norm to compute their union, the only difference between them is how to combine the foreground and background fuzzy sets. In the case of Figure 22(a), the foreground and background were combined using the s-norm \vee , and as a result, the resulting gray-scale image is biased towards foreground values. On Figure 3(c), the foreground and background were combined using the t-norm \wedge , resulting in a bias towards background values. On Figure 3(e) the self-dual uninorm $*$ was used, and as a consequence, the resulting image does not appear to be biased to neither foreground nor background.

	Gaussian MF	Bell MF	Voronoi w/ $\langle \cdot, \cdot \rangle$	Voronoi w/ Gaussian RBF
(\wedge, \vee)	0.386847	0.361106	0.455740	0.452803
$(*, \vee)$	0.465290	0.463997	0.464667	0.464588
(\vee, \vee)	0.484880	0.379021	0.444111	0.452130
(\triangle_P, ∇_P)	0.373185	0.329323	0.406735	0.402238
$(*, \nabla_P)$	0.428827	0.432984	0.463899	0.464588
(∇_P, ∇_P)	0.352563	0.164731	0.411878	0.430259

Table 3 – The mean FoM values for the segmentation of image Figure 14(a) with different sets of parameters (where the rows are a uninorm, s-norm pair and the collums are a type of membership function) of the fuzzy approach.

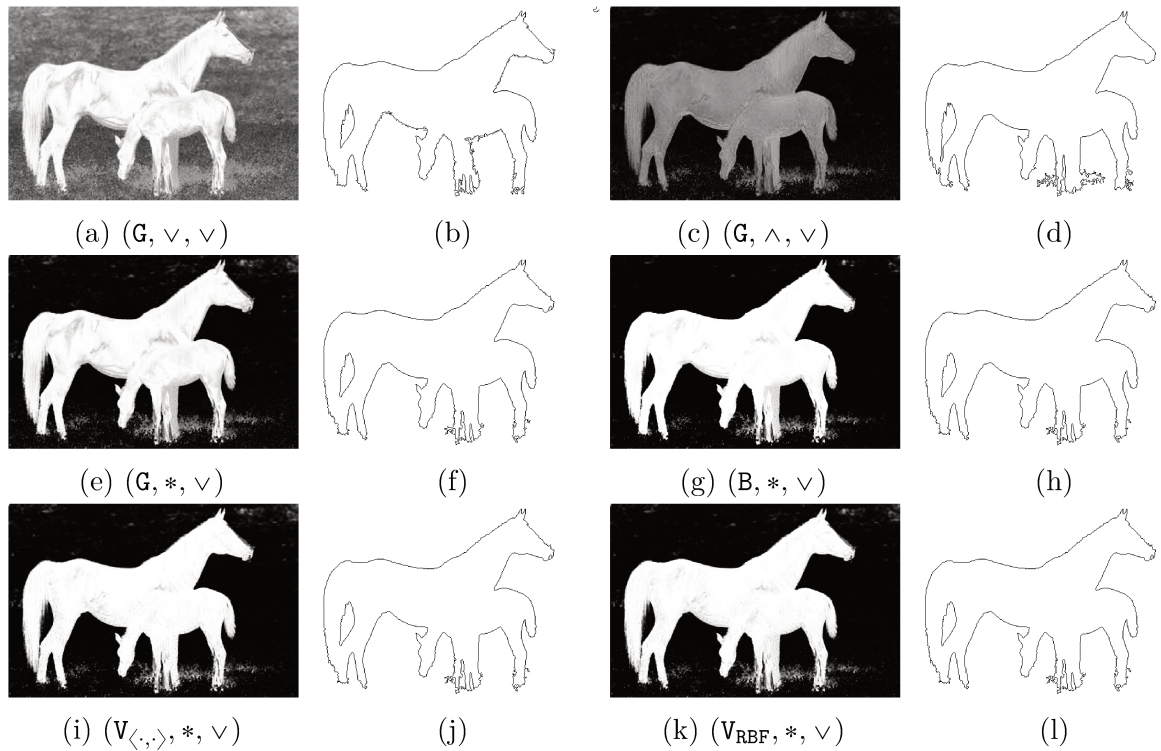


Figure 22 – Examples of ordering mappings, followed by their edge images, using the fuzzy ordering approach with varying uninorm, s-norm and membership functions. The captions denote the membership function used, the uninorm and the s-norm, respectively, of the ordering mapping and its edges. G denotes the Gaussian membership function, B denotes the bell-shaped membership function, $V_{\langle \cdot, \cdot \rangle}$ denotes the Voronoi membership function with the inner product kernel and V_{RBF} denotes the Voronoi membership function with the Gaussian RBF kernel.

6 Classification of Hyperspectral Images

Hyperspectral images possess a great resolution in the spectral domain. Their value set \mathbb{V} is high-dimensional. Hyperspectral images are useful for the classification of remote sensed urban areas. Initial attempts at pixel-wise classification of hyperspectral images used only the pixel value (Landgrebe, 2002). Using morphological transformations, Pesaresi and Benediktsson (2001) built a morphological profile (MP), making use of both the spatial and spectral data. The MP was later extended to the so-called extended morphological profile (EMP), with the help of principal component analysis (Benediktsson et al., 2005). Tarabalka et al. (2010) used the watershed transform to perform a spatial correction of a baseline classification, which improved the results.

Within the framework of supervised orderings, Velasco-Forero and Angulo (2011b) used the SVM-supervised ordering and the morphological leveling to build a classifier based on spatial and spectral features. More recently Graña and Chyzhyk (2016) extended the approach of Tarabalka et al. (2010) to work with supervised orderings. Velasco-Forero and Angulo (2011b) also suggested the combination of morphological profiles and supervised orderings.

In this chapter, we are going to explore the morphological profile and extend it to the context of supervised orderings. This approach is evaluated by computational experiments.

6.1 Morphological Profile

Morphological profiles are built from families of morphological operators, more specifically from granulometries and anti-granulometries. Granulometries can be viewed as families of openings with increasing size. Dually, anti-granulometries can be viewed as families of closings with increasing size. Here we will build the morphological profile as concatenation of the opening and closing profiles, according to Fauvel et al. (2008).

The opening profile of an image \mathbf{I} , denoted by $\mathbf{OP}(\mathbf{I})$, is a multivalued image defined as follows:

$$\mathbf{OP}_i(\mathbf{I})(p) = \gamma^{(i)}(\mathbf{I})(p), \quad \forall p \in \mathcal{D}, i \in \{0, 1, \dots, l\}, \quad (6.1)$$

where $\{\gamma^{(i)}\}_{i=0}^l$ is a family of openings by reconstruction with structuring elements of increasing size with respect to i . Dually, the closing profile of an image \mathbf{I} , denoted by $\mathbf{CP}(\mathbf{I})$, is defined as follows:

$$\mathbf{CP}_i(\mathbf{I})(p) = \varphi^{(i)}(\mathbf{I})(p), \quad \forall p \in \mathcal{D}, i \in \{0, 1, \dots, l\}, \quad (6.2)$$

where $\{\varphi^{(i)}\}_{i=0}^l$ is a family of closings by reconstruction with structuring elements of increasing size with respect to i . For convenience, the size of the structuring element when $i = 0$ is always 0, meaning $\gamma^{(0)} = \varphi^{(0)} = \text{id}$. The MP is then defined as the concatenations of the opening and closing profiles, more specifically,

$$\mathbf{MP}(\mathbf{I}) = \{\mathbf{CP}_l(\mathbf{I}), \dots, \mathbf{CP}_1(\mathbf{I}), \mathbf{I}, \mathbf{OP}_1(\mathbf{I}), \dots, \mathbf{OP}_l(\mathbf{I})\}. \quad (6.3)$$

The EMP (Fauvel et al., 2008) is defined as the concatenation of the morphological profiles of the principal components, that is

$$\mathbf{EMP}(\mathbf{I}) = \{\mathbf{MP}(\mathbf{I}_{\text{PC}_1}), \dots, \mathbf{MP}(\mathbf{I}_{\text{PC}_m})\}, \quad (6.4)$$

where \mathbf{I}_{PC_i} is image obtained from the i -th principal component of the pixel values and m is the number of chosen principal components.

Graña and Chyzhyk (2016) used a set of supervised ordering mappings $\{h_\omega : \omega \in \Omega\}$ where $\Omega = \{\omega_1, \dots, \omega_n\}$ is the set of classes and h_ω is trained to discriminate between class ω and the rest, and used these mapping to perform a spatial correction on the classification based on spectral data computed by a SVM classifier.

Notice that the images \mathbf{I}_{PC_i} can be obtained from a linear unsupervised h -mapping, that is, there exists a mapping h_{PC_i} such that $\mathbf{I}_{\text{PC}_i} = h_{\text{PC}_i}(\mathbf{I})$.

Based on the EMP and on the supervised mappings $\{h_\omega : \omega \in \Omega\}$ we propose to define the supervised morphological profile (SMP) as the concatenation of the morphological profiles of the mappings h_ω , $\omega \in \Omega$. The SMP is then computed as follows

$$\mathbf{SMP}(\mathbf{I}) = \{\mathbf{MP}(h_{\omega_1}(\mathbf{I})), \dots, \mathbf{MP}(h_{\omega_n}(\mathbf{I}))\}. \quad (6.5)$$

The SMP and the EMP can be viewed in a more general sense as the concatenation of the morphological profiles of several ordering mappings. The SMP is expected to select features that are more relevant to the process of classification, since the supervised orderings already use the information of the training data to compute a preliminary classification.

6.2 Computational Experiment

6.2.1 Data

The Pavia University hyperspectral image, provided by Prof. Paolo Gamba from the Pavia University, at northern Italy, was detected by a ROSIS sensor during a flight campaign. After some samples and bands were discarded, the image has 610×340 pixels with 103 bands and the labeled pixels are divided in 9 classes as described in Table 4. In Figure 23 a sample band of the Pavia University image is shown, together with the ground truth for the same image, showing the labeled pixels.

Index	Class	Number of Pixels
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare soil	5029
7	Bitumen	1330
8	Self-blocking bricks	3682
9	Shadows	947

Table 4 – Index, class name and number of pixels per class for each label of the Pavia University hyperspectral image.

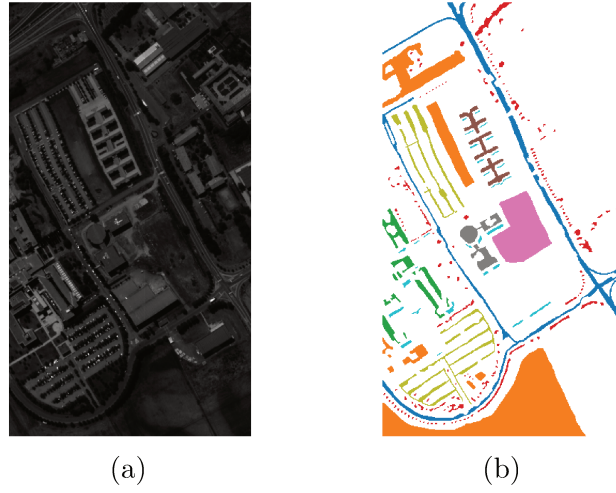


Figure 23 – A sample band of the Pavia University image(a) and the ground truth for the Pavia University image labeled pixels(b), where the white pixels are the unlabeled pixels.

6.2.2 Experiment

For each of the classes $\{\omega_1, \dots, \omega_9\}$ the h mappings $h_{\text{LAAM}}^\Omega = (h_{\text{LAAM}}^{\omega_i})_{i=1}^9$, $h_{\text{SVM}}^\Omega = (h_{\text{SVM}}^{\omega_i})_{i=1}^9$, $h_{\text{VORONOI}}^\Omega = (h_{\text{VORONOI}}^{\omega_i})_{i=1}^9$ and $h_{\text{FUZZY}}^\Omega = (h_{\text{FUZZY}}^{\omega_i})_{i=1}^9$ were trained, based on a set of training data, to discriminate between class ω_i and the others. More specifically, the ordering mappings were computed as follows:

- The $h_{\text{LAAM}}^{\omega_i}$ mapping was obtained based on the foreground only version of h_{LAAM} , obtained by setting $B = \emptyset$ and using (4.10), for each class i .
- The $h_{\text{SVM}}^{\omega_i}$ was computed in a one-versus-the-rest approach, that is trained to discriminate between class i and the others using a RBF kernel function with parameter $\gamma = \frac{1}{k \text{Var}(\bar{X})}$, where k is the spectral dimension size and \bar{X} is all of the scalar values

of the training set and $C = 10^3$. The SVM was trained using python3.6 sklearn library¹, in which the implementation of the SVM classifier is based on libsvm².

- Since the dimension of the data is high for the fuzzy method with the voronoi membership functions, the centers of the Voronoi diagram were obtained from the medoids obtained from the k-medoids method (Kaufmann and Gupta, 1988). To obtain the centers, the k-medoids of the training data are computed and then the Voronoi membership functions are computed, using the medoids and their classes. When there is no representative for one of the classes in the k-medoids, the medoid of the points in that class is used. The $h_{\text{VORONOI}}^{\omega_i}$ mappings were computed using various numbers of clusters obtained from the k-medoids, using GitHub user letiantian's implementation³, the number of clusters which yielded the highest scores on a validation set was used on the data. The dot product kernel function and a piecewise linear interpolation function $f : \mathbb{R}_{>0} \rightarrow [0, 1]$, which interpolates linearly between the points $(0, 1)$ and $(2, 0)$ and is 0 outside $[0, 2]$, that is,

$$f(x) = 0 \vee \left(1 - \frac{x}{2}\right), \quad (6.6)$$

were used.

- The $h_{\text{FUZZY}}^{\omega_i}$ mappings were computed using the optimization approach with the Adam Optimizer and a fixed learning rate given by 10^{-4} . We used the standard fuzzy negation $\eta_S(x) = 1 - x$, the s-norm \vee and the uninorm given, for some parameter e that coincides with its neutral element, by

$$a *_e b = \begin{cases} a \vee b, & \text{if } e \leq a \text{ and } e \leq b, \\ a \wedge b, & \text{if } a < e \text{ or } b < e, \end{cases}$$

which can be rewritten as

$$a *_e b = a \wedge b + H(a \wedge b - e)(a \vee b - a \wedge b), \quad (6.7)$$

where H is a step function given by, for all $x \in \mathbb{R}$,

$$H(x) = \begin{cases} 1, & \text{if } x \leq 0, \\ 0, & \text{if } x > 0. \end{cases} \quad (6.8)$$

Since (6.7) is non-differentiable with respect to e and the derivative is 0 everywhere it exists, it is not adequate for optimization using the gradient, so during the training stage we use the approximation given by

$$a \tilde{*}_e b = a \wedge b + \sigma(\alpha(a \wedge b - e))(a \vee b - a \wedge b), \quad (6.9)$$

¹ <https://scikit-learn.org/stable/>

² <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

³ <https://github.com/letiantian/kmedoid>

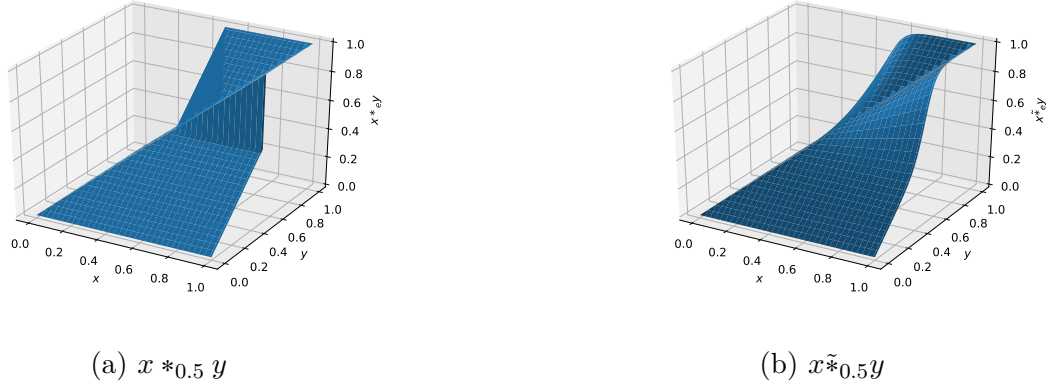


Figure 24 – Comparison between the surfaces of (6.7) and (6.9) with $e = 0.5$ and $\alpha = 20$.

where σ is the sigmoid function and α is some arbitrary scaling constant, which was set to $\alpha = 20$. Figure 24 shows a comparison between the surfaces of (6.7) and (6.9). The surfaces are very similar except close to the lines given by $x = e$ and $y = e$. Outside of training, when computing the output from a set of inputs we use (6.7). The loss function used was the Huber loss (Huber, 1964). The number of background and foreground sets was chosen based on a validation set. Gaussian membership functions were used, and the initialization was done as follows: the initial centers of the background and foreground sets were obtained from the fuzzy c-means algorithm, the γ parameters of the gaussian functions were initialized as 1 and the e parameter of the uninorm was initialized as 0.5. Random initialization was also tested on the validation sets.

The h -mappings were also computed on an image obtained by the first 20 PCs of the hyperspectral image. The results were compared with an EMP classification, using 9 PCs, so it has the same dimension as the SMP, and a purely spectral classification.

The Pavia University image values as well as the EMP's and the SMPs' were normalized to fit the interval $[-1, 1]$.

The training/validation set consists of 1% randomly chosen labeled pixels of each class, the test set consists of the remaining labeled pixels. The training set is chosen as half of the training/validation set and the other half is the validation set. The experiments were conducted multiple times, with different training/validation splits. The SMP was computed using these h -mappings, with disk structuring elements of sizes 3, 6, 9 and 12. Using the SMP as the features, an SVM was trained in a one-vs-the-rest approach on the training set with the parameter γ of the RBF as $\frac{1}{l\text{Var}(Z)}$ where l is the dimension of the inputs to the classifier and Z the flattened inputs and with the penalty parameter $C = 10^3$. The classifier was trained using python3.6 sklearn library. The trained model

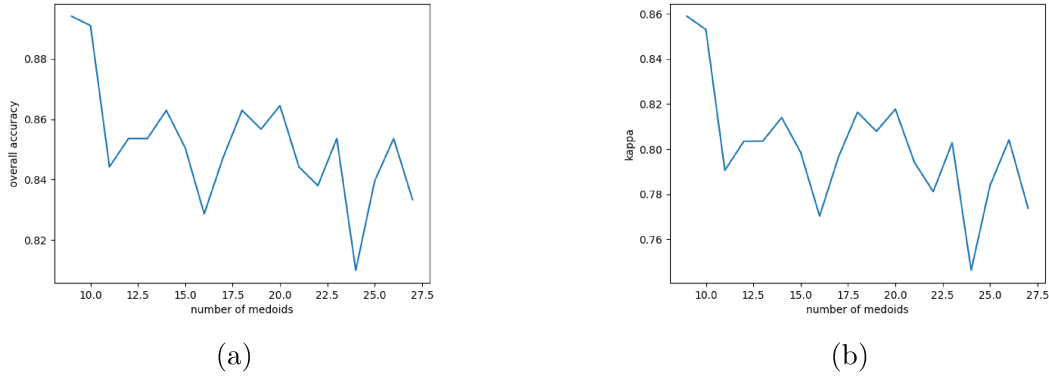


Figure 25 – Graphs of the values of overall accuracy(a) and kappa coefficient(b) for the quantities of clusters between 9 and 27.

was then evaluated on the test set.

The performance is measured by the overall accuracy, average accuracy of each class and the κ coefficient (Graña and Chyzhyk, 2016; Richards John and Xiuping, 1999). The κ coefficient measures accuracy considering the possibility that it is by random chance.

6.2.3 Results

Using a number of clusters varying from 9 to 27, the h_{VORONOI} was computed using 3 different splits of the training and validation sets. The graphs showing the mean overall accuracies and kappa coefficients of these measurements are shown in Figure 25(a) and Figure 25(b). The chosen number of medoids was 10.

As for the h_{FUZZY} obtained through optimization, the experiment was done 3 times on the validation set. The mean values of the accuracy of the non-random initialization approach are shown on Table 5 and the mean κ coefficients are shown on Table 6. The mean values of the accuracy of the random initialization approach are shown on Table 7 and the mean κ coefficients are shown on Table 8. The performance of the non-random initialization approach was significantly better than the performance of the random initialization approach, so the experiment was done with non-random initialization. The number of foreground sets n_{fg} and the number of background sets n_{bg} were chosen based on these tables, more specifically, the maximum κ value of Table 6, where $n_{\text{fg}} = 1$ and $n_{\text{bg}} = 3$ was chosen and the second to maximum κ value, where $n_{\text{fg}} = 2$ and $n_{\text{bg}} = 3$ was also chosen. In this chapter, these approaches are denoted Fuzzy (1,3) and Fuzzy (2,3), respectively.

The results of the hyperspectral image classification computational experiment are shown in Table 9. The results are a mean of 30 repetitions of the experiment. Examples of the results are shown in Figure 28. The boxplot of the κ coefficients obtained from all

$n_{fg} \backslash n_{bg}$	2	3	4	5	6
1	0.9502	0.9595	0.9533	0.9533	0.9502
2		0.9548	0.9470	0.9470	0.9548
3			0.9548	0.9533	0.9470

Table 5 – Values of mean accuracy of the SMP classification on the validation set using the h_{FUZZY} approach. Using non-random initialization.

$n_{fg} \backslash n_{bg}$	2	3	4	5	6
1	0.9310	0.9441	0.9354	0.9353	0.9315
2		0.9379	0.9264	0.9268	0.9378
3			0.9373	0.9353	0.9268

Table 6 – Values of the mean κ coefficient of the SMP classification on the validation set using the h_{FUZZY} approach. Using non-random initialization.

$n_{fg} \backslash n_{bg}$	2	3	4	5	6
1	0.7975	0.8629	0.7664	0.8115	0.7991
2		0.8380	0.7352	0.8100	0.8380
3			0.8551	0.8614	0.8785

Table 7 – Values of mean accuracy of the SMP classification on the validation set using the h_{FUZZY} approach. Using random initialization.

of the approaches is shown in Figure 26. The Hasse Diagram of paired Student's t-test is shown in Figure 27.

6.2.4 Discussion

As we can see, the h_{LAAM} classification achieved the greatest score, but the other supervised ordering approaches, in particular the h_{SVM} and the h_{FUZZY} approaches achieved similar performances. Moreover they all surpassed the EMP classification. This is further confirmed by Figure 27. As expected, the approach based only on spectral data achieved the worst performance. The $h_{VORONOI}$ classification also did not achieve a relatively good performance, obtaining worse results than the EMP classification, possibly because it is not well suited to large quantities of data.

These results suggest that the SMP yielded competitive results and is able to surpass the EMP, depending on the ordering being used. Furthermore, the h_{FUZZY} ordering was able to achieve competitive results.

$n_{\text{fg}} \backslash n_{\text{bg}}$	2	3	4	5	6
1	0.7268	0.8120	0.6680	0.7430	0.7302
2		0.7787	0.6253	0.7403	0.7800
3			0.8044	0.8125	0.8333

Table 8 – Values of the mean κ coefficient of the SMP classification on the validation set using the h_{FUZZY} approach. Using random initialization.

Method	Overall accuracy	κ
Spectral only	0.8187	0.7573
EMP	0.9361	0.9149
h_{SVM}	0.9561	0.9416
h_{LAAM}	0.9584	0.9445
h_{VORONOI}	0.8781	0.8367
h_{FUZZY} with $(n_{\text{fg}}, n_{\text{bg}}) = (1, 3)$	0.9547	0.9398
h_{FUZZY} with $(n_{\text{fg}}, n_{\text{bg}}) = w(2, 3)$	0.9542	0.9391

Table 9 – Mean accuracies and mean κ for the hyperspectral image experiment.

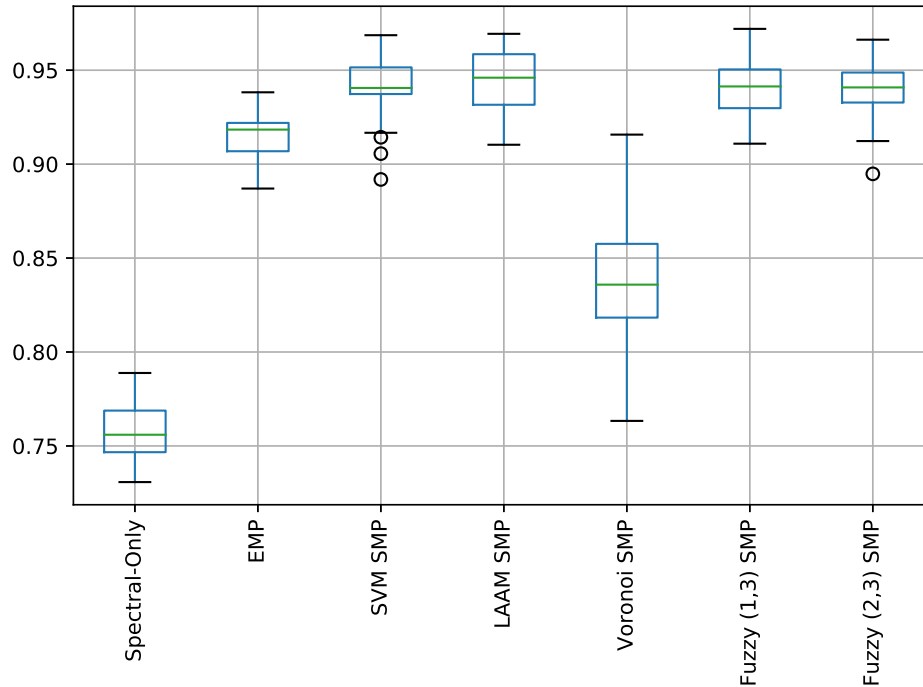
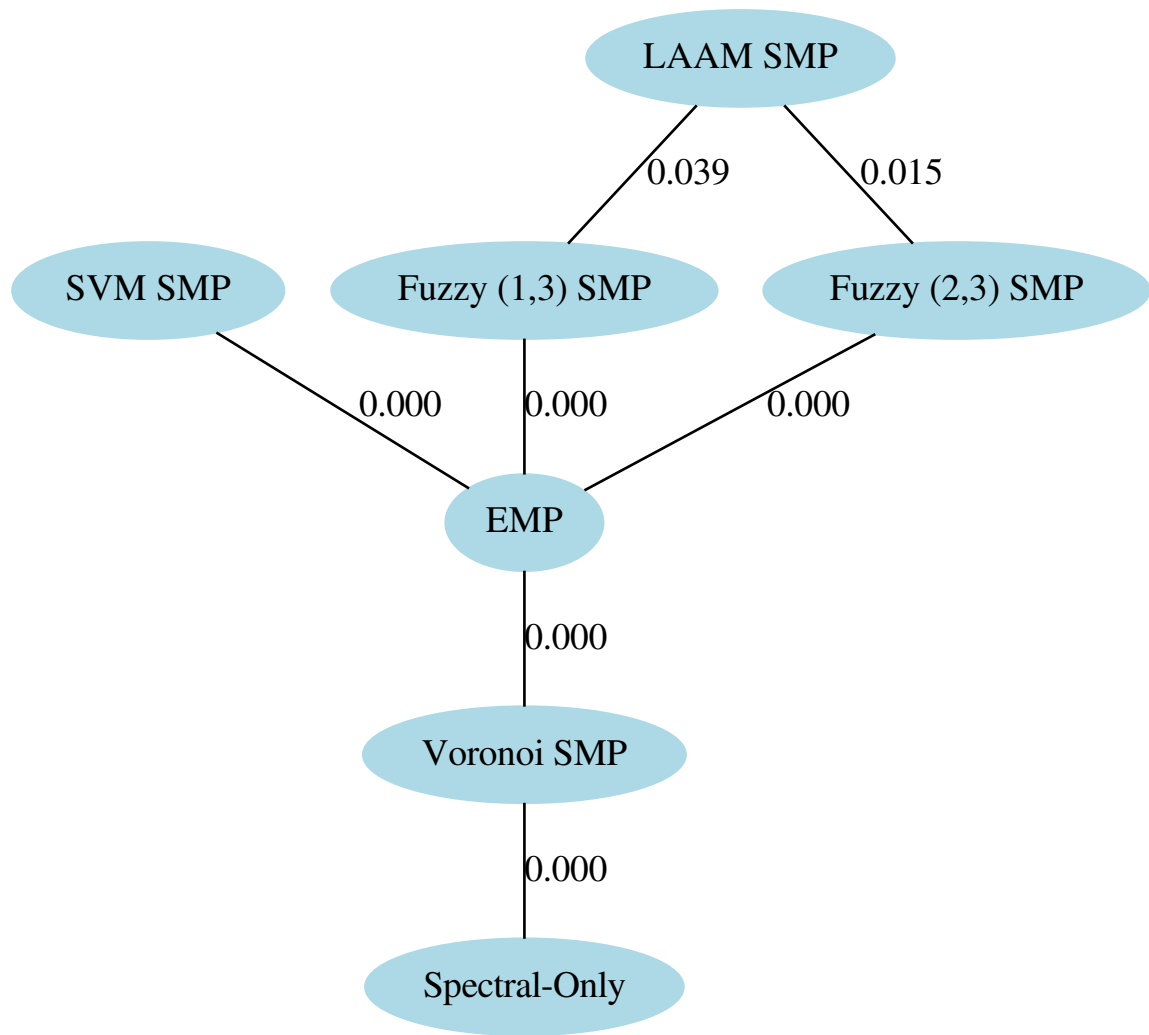


Figure 26 – Boxplot of the κ coefficients obtained from the 30 repetitions of the experiment using the various classification approaches.



Hasse diagram of paired Student's t-test
(confidence level at 95.0%)

Figure 27 – Hasse diagram of paired Student's t-test of the κ coefficients obtained from the 30 repetitions of the experiment using the various classification approaches. The p-values are shown beside the edges with 3 decimal places.

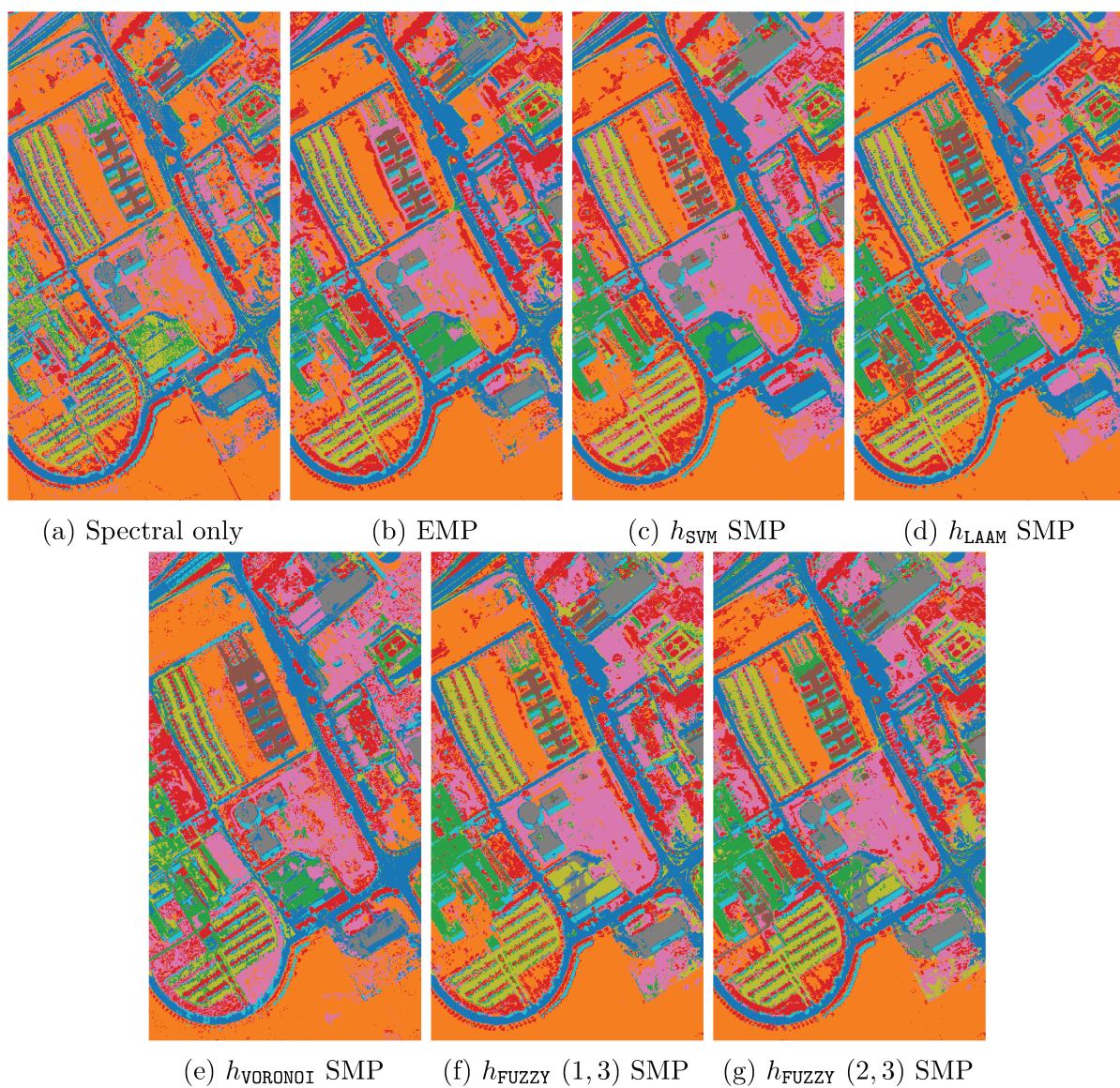


Figure 28 – Example classifications of the Pavia University image using various approaches.

7 Concluding Remarks

The main focus of this dissertation was supervised orderings, in particular, the supervised ordering given by fuzzy propositions defined at Chapter 5.

Concepts of the theory of lattices were reviewed on Chapter 1. Some mathematical tools, including fuzzy logic and the ANFIS, were reviewed on Chapter 2. Mathematical morphology was reviewed on Chapter 3. Mathematical morphology based on reduced orderings was reviewed on Chapter 4, and the concept of supervised ordering was revised. A supervised ordering is a reduced ordering that is defined through a set of reference values, usually divided into foreground and background sets. The supervised orderings studied at that chapter were the distance based ordering, the support vector machine based ordering and the lattice auto-associative memory based ordering. An illustrative example was used to compare some of those approaches qualitatively.

Chapter 5 introduced the ordering based on fuzzy sets of values. Precisely, the ordering mapping of the supervised ordering was given by the membership function h_{FUZZY} of a fuzzy set. To obtain this fuzzy set, first, fuzzy sets for foreground values and fuzzy sets for background values are obtained. After obtaining the foreground and background fuzzy sets, the membership function h_{FUZZY} is defined by a fuzzy proposition based on the foreground and background fuzzy sets. Two particular cases of the propositions that were used can be written for a value \mathbf{x} as “ \mathbf{x} is a foreground value and it is not a background value” and “ \mathbf{x} is a foreground value or it is not a background value”, and in the more general case, the “and” and the “or” are traded for a connective expressed by a uninorm $*$. These models were also shown, under some hypothesis, to satisfy the desired conditions of supervised orderings, namely that the foreground values are the greatest elements and that the background values are the least elements in the reduced ordering.

The main difference on the approaches of Chapter 5 is how to obtain the primary foreground and background fuzzy sets. On the first approach the membership function is given by, for a reference value \mathbf{r} that is either foreground or background, $\mu_R(\mathbf{x}) = f(d(\mathbf{x}, \mathbf{r}))$, where $f : [0, +\infty) \rightarrow [0, 1]$ is a decreasing function and d is a metric. On the second approach, the model is formulated as a neuro-fuzzy model and the values are obtained through minimization of the loss function. On the last approach, the fuzzy sets are obtained via interpolation in the Voronoi cells of the sets of the foreground and background colors. While the first and last approach are more suited to smaller data and with human supervision, the neuro-fuzzy approach was shown to achieve good results in a hyperspectral image classification in Chapter 6. The neuro-fuzzy model also possess more flexibility than the other two, admitting the possibility of using it in conjunction

with deep learning architectures. An interesting possibility that can be explored is the combination of the Voronoi cells approach and the neuro-fuzzy approach.

On Chapter 6 the problem of pixelwise classification of hyperspectral images was studied. Some approaches from the literature, such as the morphological profile and the EMP were studied and a new approach, based on supervised orderings and referred to as SMP, was proposed. The idea of the SMP is to use supervised ordering mappings to build a classification of the spectral information and then use a morphological profile to obtain features that carry both spectral and spatial information. The supervised ordering approaches were tested on the Pavia University image and compared with a spectral-only SVM classification and the EMP classification. The supervised orderings achieved the best results, with the exception of the fuzzy ordering based on the Voronoi diagram, which is more suited to smaller data. The h_{LAAM} achieved the best results, but both the h_{SVM} and the neuro-fuzzy model achieved competitive results. The effect of the initialization on the neuro-fuzzy approach was also tested and it was verified that initializing it with the centers of a fuzzy clustering of the foreground and background training sets managed to improve the performance.

In conclusion, supervised orderings were shown to be flexible and to have potential in applications such as hyperspectral image segmentation. A possible next step in this topic might be to study their application in other fields of computer vision, such as medical image segmentation, and study the interaction of these models with other models of computer vision, for example trying to use the neuro-fuzzy model in conjunction of deep learning models such as convolutional neural networks.

Bibliography

- I. E. Abdou and W. K. Pratt. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of the IEEE*, 67(5):753–763, 1979. pages 43
- H. M. Al-Otum. A novel set of image morphological operators using a modified vector distance measure with color pixel classification. *J. Visual Communication and Image Representation*, 30:46–63, 2015. URL <http://dblp.uni-trier.de/db/journals/jvcir/jvcir30.html#Al-Otum15>; <http://dx.doi.org/10.1016/j.jvcir.2015.02.010>. pages 11
- J. Angulo. Morphological colour operators in totally ordered lattices based on distances: Application to image filtering, enhancement and analysis. *Computer Vision and Image Understanding*, 107(1–2):56–73, July-August 2007. Special issue on color image processing. pages 10, 11, 27, 40
- E. Aptoula and S. Lefèvre. A Comparative Study on Multivariate Mathematical Morphology. *Pattern Recognition*, 40(11):2914–2929, November 2007. pages 10, 27, 33, 34, 35
- E. Aptoula and S. Lefèvre. On Lexicographical Ordering in Multivariate Mathematical Morphology. *Pattern Recognition Letters*, 29(2):109–118, January 2008. pages 10
- V. Barnett. The ordering of multivariate data. *Journal of Royal Statistical Society A*, 3: 318–355, 1976. pages 10
- J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):480–491, 2005. pages 57
- G. Birkhoff. *Lattice Theory*. American Mathematical Society, Providence, 3 edition, 1993. pages 10, 13, 14
- I. Bloch. Lattices of fuzzy sets and bipolar fuzzy sets, and mathematical morphology. *Information Sciences*, 181(10):2002–2015, May 2011. pages 10, 20, 27
- I. Bloch and H. Maitre. Fuzzy mathematical morphologies: a comparative study. *Pattern Recognition*, 28(9):1341–1387, 1995. pages 29
- U. Braga-Neto and J. Goutsias. Supremal multiscale signal analysis. *SIAM Journal of Mathematical Analysis*, 36(1):94–120, 2004. pages 10

- J. Chamorro-Martínez, J. M. Soto-Hidalgo, P. M. Martínez-Jiménez, and D. Sánchez. Fuzzy color spaces: A conceptual approach to color vision. *IEEE Transactions on Fuzzy Systems*, 25(5):1264–1280, 2017. pages 11, 46, 51, 52
- J. Chanussot and P. Lambert. Total ordering based on space filling curves for multivalued morphology. In *Proceedings of the Fourth International Symposium on Mathematical Morphology and Its Applications to Image and Signal Processing*, ISMM '98, pages 51–58, Norwell, MA, USA, 1998. Kluwer Academic Publishers. pages 37
- E. Chevallier and J. Angulo. The Irregularity Issue of Total Orders on Metric Spaces and Its Consequences for Mathematical Morphology. *Journal of Mathematical Imaging and Vision*, 54(3):344–357, Mar 2016. doi: 10.1007/s10851-015-0607-7. pages 10, 35, 37
- M. L. Comer and E. J. Delp. Morphological operations for color image processing. *Journal of Electronic Imaging*, 8(3):279–289, 1999. pages 11, 33, 34, 37, 40
- N. Cristianini, J. Shawe-Taylor, et al. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000. pages 17, 18
- B. De Baets. Fuzzy Morphology: A Logical Approach. In B. M. Ayyub and M. M. Gupta, editors, *Uncertainty Analysis in Engineering and Science: Fuzzy Logic, Statistics, and Neural Network Approach*, pages 53–67. Kluwer Academic Publishers, Norwell, 1997. pages 10, 29
- H. Deborah, N. Richard, and J. Hardeberg. Spectral Ordering Assessment Using Spectral Median Filters. In J. A. Benediktsson, J. Chanussot, L. Najman, and H. Talbot, editors, *Mathematical Morphology and Its Applications to Signal and Image Processing*, volume 9082 of *Lecture Notes in Computer Science*, pages 387–397. Springer International Publishing, 2015. doi: 10.1007/978-3-319-18720-4_33. pages 11
- T. Deng and H. Heijmans. Grey-scale morphology based on fuzzy logic. *Journal of Mathematical Imaging and Vision*, 16(2):155–171, 2002. pages 27
- M. Fauvel, J. Benediktsson, J. Chanussot, and J. Sveinsson. Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(11):3804–3814, Nov 2008. doi: 10.1109/TGRS.2008.922034. pages 57, 58
- M. Gonzalez-Hidalgo, S. Massanet, A. Mir, and D. Ruiz-Aguilera. On the Choice of the Pair Conjunction-Implication Into the Fuzzy Morphological Edge Detector. *IEEE Transactions on Fuzzy Systems*, 23(4):872–884, Aug 2015. doi: 10.1109/TFUZZ.2014.2333060. pages 10

- J. Goutsias, H. J. A. M. Heijmans, and K. Sivakumar. Morphological Operators for Image Sequences. *Computer vision and image understanding*, 62:326–346, 1995. pages 10, 15, 37, 38, 40
- M. Graña. A brief review of lattice computing. In *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2008)*, pages 1777–1781, Hong Kong, China, June 2008. pages 13
- M. Graña and D. Chyzyk. Image understanding applications of lattice autoassociative memories. *IEEE transactions on neural networks and learning systems*, 27(9):1920–1932, 2016. pages 11, 42, 57, 58, 62
- G. Grätzer et al. *General Lattice Theory*. Birkhäuser Verlag, Basel, Switzerland, 2nd edition edition, 2003. pages 10
- A. Hanbury and J. Serra. Mathematical Morphology in the CIELAB Space. *Image Analysis and Stereology*, 21:201–206, 2002. pages 10
- H. J. A. M. Heijmans. Mathematical Morphology: A Modern Approach in Image Processing Based on Algebra and Geometry. *SIAM Review*, 37(1):1–36, 1995. pages 10, 27, 28, 29
- P. Huber. Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1):73–101, Mar. 1964. pages 61
- J.-S. R. Jang and C.-T. Sun. *Neuro-fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1997. ISBN 0-13-261066-3. pages 25, 26
- V. G. Kaburlasos, S. E. Papadakis, and G. A. Papakostas. Lattice computing extension of the fam neural classifier for human facial expression recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 24(10):1526–1538, 2013. pages 13
- A. Kaufmann and M. Gupta. *Fuzzy Mathematical Models in Engineering and Management Science*. North-Holland, Amsterdam, 1988. pages 60
- D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. pages 51
- T. Kohonen. *Associative Memory – A System Theoric Approach*. Berlin: Springer-Verlag, 1977. pages 19
- D. Landgrebe. Hyperspectral image data analysis. *IEEE Signal processing magazine*, 19(1):17–28, 2002. pages 57

- A. Ledoux, N. Richard, A.-S. Capelle-Laizé, and C. Fernandez-Maloigne. Perceptual color hit-or-miss transform: application to dermatological image processing. *Signal, Image and Video Processing*, 9(5):1081–1091, 2015. doi: 10.1007/s11760-013-0537-z. pages 11
- O. Lézoray. Complete lattice learning for multivariate mathematical morphology. *Journal of Visual Communication and Image Representation*, 35:220–235, Feb 2016. doi: 10.1016/j.jvcir.2015.12.017. pages 10, 11
- G. Louverdis and I. Andreadis. Soft Morphological Filtering Using a Fuzzy Model and its Application to Colour Image Processing. *Formal Pattern Analysis and Applications*, 6(4):275–268, Feb. 2004. doi: 10.1007/s10044-003-0193-y. pages 11
- G. Louverdis, I. Andreadis, and P. Tsalides. New fuzzy model for morphological colour image processing. *IEE Proceedings-Vision, Image and Signal Processing*, 149(3):129–139, 2002. pages 37
- D. Martin, C. Fowlkes, D. Tal, and J. Malik. A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In *Proceedings of the 8th International Conference on Computer Vision*, volume 2, pages 416–423, July 2001. pages 44
- K.-R. Muller, S. Mika, G. Ratsch, K. Tsuda, and B. Scholkopf. An introduction to kernel-based learning algorithms. *IEEE transactions on neural networks*, 12(2):181–201, 2001. pages 18
- M. Nachtegaele and E. E. Kerre. Connections between binary, gray-scale and fuzzy mathematical morphologies. *Fuzzy Sets and Systems*, 124(1):73–85, 2001. pages 10, 20, 27
- H. T. Nguyen and E. A. Walker. *First course in fuzzy logic*. CRC Press, 2005. pages 22, 23
- N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-9(1):62–66, Jan. 1979. pages 43
- M. Pesaresi and J. A. Benediktsson. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE transactions on Geoscience and Remote Sensing*, 39(2):309–320, 2001. pages 57
- A. Richards John and J. Xiuping. Remote sensing digital image analysis: an introduction, 1999. pages 62
- G. X. Ritter, P. Sussner, and J. L. D. de Leon. Morphological Associative Memories. *IEEE Transactions on Neural Networks*, 9(2):281–293, 1998. pages 19

- L. Rittner, J. Campbell, P. Freitas, S. Appenzeller, G. B. Pike, and R. Lotufo. Analysis of Scalar Maps for the Segmentation of the Corpus Callosum in Diffusion Tensor Fields. *Journal of Mathematical Imaging and Vision*, 45:214–226, 2013. pages 10
- C. Ronse. Why Mathematical Morphology Needs Complete Lattices. *Signal Processing*, 21(2):129–154, 1990. pages 10, 13
- M. Sangalli and M. E. Valle. Color mathematical morphology using a fuzzy color-based supervised ordering. In *North American Fuzzy Information Processing Society Annual Conference*, pages 278–289. Springer, 2018. pages 11, 16, 46
- L. J. Sartor and A. R. Weeks. Morphological operations on color images. *Journal of Electronic Imaging*, 10(2):548–559, 2001. doi: 10.1117/1.1353199. pages 11
- J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982. pages 29
- J. Serra. *Image Analysis and Mathematical Morphology, Volume 2: Theoretical Advances*. Academic Press, New York, 1988. pages 28
- J. Serra. A Lattice Approach to Image Segmentation. *Journal of Mathematical Imaging and Vision*, 24:83–130, 2006. pages 10
- J. Serra. The “False Colour” Problem. In M. H. Wilkinson and J. B. Roerdink, editors, *Mathematical Morphology and Its Application to Signal and Image Processing*, volume 5720 of *Lecture Notes in Computer Science*, pages 13–23. Springer Berlin Heidelberg, 2009. pages 10
- P. Soille. *Morphological Image Analysis*. Springer Verlag, Berlin, 1999. pages 10, 30, 32
- S. Sternberg. Grayscale Morphology. *Computer Vision, Graphics and Image Processing*, 35:333–355, 1986. pages 10
- M. Sugeno and G. Kang. Structure identification of fuzzy model. *Fuzzy sets and systems*, 28(1):15–33, 1988. pages 25
- P. Sussner and E. L. Esmi. Morphological Perceptrons with Competitive Learning: Lattice-Theoretical Framework and Constructive Learning Algorithm. *Information Sciences*, 181(10):1929–1950, May 2011. doi: 10.1016/j.ins.2010.03.016. pages 31
- P. Sussner and M. E. Valle. Grayscale Morphological Associative Memories. *IEEE Transactions on Neural Networks*, 17(3):559–570, May 2006. pages 19, 42
- P. Sussner and M. E. Valle. Classification of Fuzzy Mathematical Morphologies Based on Concepts of Inclusion Measure and Duality. *Journal of Mathematical Imaging and Vision*, 32(2):139–159, Oct. 2008. pages 10, 20, 27

- Y. Tarabalka, J. Chanussot, and J. A. Benediktsson. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition*, 43(7): 2367–2379, 2010. pages 57
- M. E. Valle and R. A. Valente. Mathematical Morphology on the Spherical CIELab Quantale with an Application in Color Image Boundary Detection. *Journal of Mathematical Imaging and Vision*, 2016. doi: 10.1007/s10851-016-0674-4. Accepted for publication. pages 11
- S. Velasco-Forero and J. Angulo. Morphological processing of hyperspectral images using kriging-based supervised ordering. In *2010 IEEE International Conference on Image Processing*, pages 1409–1412. IEEE, 2010. pages 11
- S. Velasco-Forero and J. Angulo. Supervised ordering in \mathbb{R}^p : Application to morphological processing of hyperspectral images. *IEEE Transactions on Image Processing*, 20(11): 3301–3308, 2011a. pages 11, 37, 39, 41
- S. Velasco-Forero and J. Angulo. Multiclass ordering for filtering and classification of hyperspectral images. In *2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pages 1–4. IEEE, 2011b. pages 57
- S. Velasco-Forero and J. Angulo. Random projection depth for multivariate mathematical morphology. *IEEE Journal of Selected Topics in Signal Processing*, 6(7):753–763, 2012. pages 11, 37
- S. Velasco-Forero and J. Angulo. Vector ordering and multispectral morphological image processing. In *Advances in Low-Level Color Image Processing*, pages 223–239. Springer, 2014. pages 16, 37, 38, 40, 41
- R. R. Yager and A. Rybalov. Uninorm aggregation operators. *Fuzzy Sets and Systems*, 80(1):111–120, 1996. pages 25
- L. A. Zadeh. Fuzzy Sets. *Information and Control*, 8(3):338–353, 1965. pages 20