UNIVERSITÀ DEGLI STUDI
DI MILANO

SCUOLA DI DOTTORATO IN INFORMATICA

DIPARTIMENTO DI SCIENZE DELL'INFORMAZIONE

CORSO DI DOTTORATO IN

INFORMATICA CICLO XXIII

TESI DI DOTTORATO DI RICERCA

Exploiting Higher Order Uncertainty in Image Analysis

INF/01

CANDIDATO
Alessio Ferone

TUTOR
Prof. Alfredo Petrosino

COORDINATORE DEL DOTTORATO
Prof. Ernesto Damiani

ANNO ACCADEMICO 2009/2010

Abstract

Soft computing is a group of methodologies that works synergistically to provide flexible information processing capability for handling real-life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions. Soft computing methodologies (involving fuzzy sets, neural networks, genetic algorithms, and rough sets) have been successfully employed in various image processing tasks including image segmentation, enhancement and classification, both individually or in combination with other soft computing techniques. The reason of such success has its motivation in the fact that soft computing techniques provide a powerful tools to describe uncertainty, naturally embedded in images, which can be exploited in various image processing tasks.

The main contribution of this thesis is to present tools for handling uncertainty by means of a rough-fuzzy framework for exploiting feature level uncertainty.

The first contribution is the definition of a general framework based on the hybridization of rough and fuzzy sets, along with a new operator called $\mathcal{RF}$-product, as an effective solution to some problems in image analysis. The second and third contributions are devoted to prove the effectiveness of the proposed framework, by presenting a compression method based on vector quantization and its compression capabilities and an HSV color image segmentation technique.

Acknowledgments

Thanks to Prof. Alfredo Petrosino and all the good people around me.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Soft Computing

The digital revolution, with the consequent development of computer hardware and software, has made available a huge amount of data. As a result, traditional statistical data summarization and database management techniques are just not adequate for handling data on this scale, and for extracting information or, rather, knowledge that may be useful for mining the domain in question and supporting the decision-making processes. The massive amount of data is generally characterized by the presence of not just numeric, but also textual, symbolic, pictorial and other type of data and may contain redundancy, errors, imprecision, and so on. In this scenario, soft computing provides a group of methodologies that works synergistically for handling real-life ambiguous situations. Its aim is to exploit imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions. The guiding principle is to devise methods of computation that lead to an acceptable solution at low cost by seeking for an approximate solution to an imprecisely/precisely formulated problem. Soft computing methodologies (involving fuzzy sets, neural networks, genetic algorithms, and rough sets) are most widely applied in the data mining process. Possibility Theory provides a natural framework for dealing with uncertainty. Neural networks and rough sets are widely used for classification and rule generation. Genetic algorithms (GAs) are involved in various optimization and search processes[68].

Each soft computing methodology has its own powerful properties and offer different advantages. For example, Fuzzy Logic is often used to model human reasoning and provide a natural mechanism for dealing with uncertainty. Neural networks are robust to noise and have a good ability to model highly non-linear relationships. Genetic algorithm is particularly useful for optimal search. Rough sets are very efficient in attribute reduction and rule extraction. On the other hand, these soft computing techniques also have some restrictions that do not allow their individual application in some cases. Fuzzy sets are dependent on expert knowledge. The training time of neural networks can be long when the input data are large and most neural network systems lack explanation facilities. The theoretical basis of genetic algorithm is weak, especially on algorithm convergence. Rough sets are sensitive to noise and present the NP problems on the choice of optimal attribute reduct and optimal rules. In order to cope with the drawbacks of individual approaches and leverage performance of data mining system, it is natural to develop hybrid systems by integrating two or more soft computing technologies. Each of them contributes a distinct methodology for addressing problems in its domain, in a cooperative, rather than a competitive, manner. The result is a more intelligent and robust system providing a human-interpretable, low cost, approximate solution, as compared to traditional techniques.

1.2 Neural Networks

Neural networks have proved to be a powerful tool for mining data, although they were earlier thought to be unsuitable for this task because of the lack of information suitable for verification or interpretation by humans. This has not prevented neural networks from being used, and sometimes abused, for a nearly every classification and regression tasks, both in supervised and unsupervised version. Recently there has been a growing interest aimed at filling this hole of knowledge, by extracting it from the trained networks in the form of symbolic rules [123]. In this way it is possible to identify the attributes that are the most significant determinants of the decision or classification. The main contribution of neural nets in the field of data mining is for rule extraction, classification and clustering. In general, the first step to extract knowledge from a connectionist model is to provide a representation of the trained neural network, in terms of its nodes and links. One or more hidden and output units are automatically selected to derive the rules, which can be combined to gain a more comprehensible rule set. Neural nets then provide high parallelism and optimization capability in the data domain. First a network is trained to achieve the required accuracy and then redundant connections pruned. Classification rules are generated by analyzing the network in terms of link weights and activation values of the hidden units[60].

1.3 Fuzzy Sets

The development of fuzzy logic has led to the rise of soft computing, becoming the earliest and most widely reported constituent of this field. Its aim is the modeling of imprecise and qualitative knowledge, as well the handling of uncertainty at various stages. Fuzzy logic is capable of encoding, to a certain extent, human reasoning in natural form. Despite a growing versatility of knowledge discovery systems, there is an important component of human interaction that is inherent to any process of knowledge representation, manipulation, and processing. Fuzzy sets are inherently inclined to cope with linguistic domain knowledge and produce more interpretable solutions. Fuzzy logic has been extensively used in many application fields to exploit its characteristic features, for instance, knowledge discovery in databases [24], clustering [126] [106] [95], web mining [78] and image retrieval [80] [28].

1.4 Rough Sets

Rough sets theory [92] has emerged as a major mathematical tool for handling uncertainty that arises from granularity in the domain of discourse, this is done by managing indiscernibility between objects in a set. It offers powerful tools to extract hidden patterns from data and therefore it is becoming very important in various application fields. A fundamental characteristic of a rough set-based learning system is to discover redundancies and dependencies between the given features of a problem to be classified. This is done by characterizing a given concept from below and from above, using lower and upper approximations. Recently rough sets theory has been extensively employed in various applications fields, although its use generally proceeds along two main directions:

1. Decision rule induction based on generation of discernibility matrices and reducts [74] [112].
2. Data filtration [102] by extracting elementary blocks from data based on equivalence relation.

1.5 Genetic Algorithm

Genetic Algorithms (GAs) are adaptive, robust, efficient, optimization methodologies based on principles of nature. They can also be viewed as searching algorithms, suitable in situations where the search space is large, because they explore a space using heuristics inspired by nature.

Any optimization problem has to be represented by using chromosomes, which are a codified representation of the real values of the variables in the problem. Then GAs optimize a fitness function to arrive at an optimal solution using certain genetic operators. Basically, a genetic algorithm uses a population of individuals, which are modified by using genetic operators in such a way as to eventually obtain the best individual.

GAs do not require or use derivative information and hence the most appropriate applications are problems where gradient information is unavailable or costly to obtain. Reinforcement learning is an example of such domain.In GAs, the only feedback used by the algorithm is information about the relative performance of different individuals.

1.6 Soft Computing in Image Analysis

Soft computing offers a novel approach to manage uncertainty in discovering data dependencies, relevance of features, mining of patterns, feature space dimensionality reduction, and classification of objects. Consequently, these techniques have been successfully employed for various image processing tasks including image segmentation, enhancement and classification, both individually or in combination with other soft computing techniques. Just to show some examples of such combinations, rough sets have been successfully combined with fuzzy sets, other than with neural networks, genetic algorithms, support vector machines for image segmentation, feature extraction, classification and enhancement, and many other tasks. Over the years the combination of two or more techniques has been proved to be effective in image processing, yielding algorithms which have overcome classical approaches. The reason of such success has its motivation in the fact that soft computing techniques provide a powerful tools to describe uncertainty, naturally embedded in images, which can be exploited in various image processing tasks.

1.7 Contributions

Many basic concepts of image analysis do not lend themselves well to precise definition. Uncertainties arise from deficiencies which can result from incomplete, imprecise and vague information in various stages of a image processing tasks. Classical image processing approaches do not consider image content as uncertain and hence tend to discard important, even though incomplete, parts of information. As stated above, soft computing techniques try to exploit uncertainty embedded in image, on the consideration that an incomplete or vague information can help the final task. Obviously different kind of uncertainties may be faced in image processing, and hence different techniques have to be employed to capture and elaborate them. In particular uncertainty can be exploited at lower level, i.e. at feature level, or at higher level, i.e. at semantic level. The aim of this thesis is to present three contributions for handling uncertainty by means of a rough-fuzzy framework for exploiting feature level uncertainty.

1.7.1 Rough Fuzzy Framework for Image Processing

The intrinsic presence of uncertainty when dealing with digital images processing and analysis, is the reason of the growing interest for the use of rough and fuzzy based techniques which have proved to be effective in this field. Moreover, the definition of combined frameworks of both theories has given more powerful tools to exploit their own characteristics.

The first contribution is the definition of a general framework based on the hybridization of rough and fuzzy sets as an effective solution to some problems in image analysis. In the contest of this framework, a new operator to compose rough fuzzy sets along with the proofs of its basics properties is presented. This new operator, called $\mathcal{RF}$-product, can be viewed as a multiresolution approach, i.e. as a sequence of composition of rough fuzzy sets.

1.7.2 Rough Fuzzy Vector Quantization

The second contribution is a compression method based on vector quantization. Feature extraction is based on the given definition of rough fuzzy sets and performed by partitioning each block in multiple rough fuzzy sets which are characterized by two approximation sets, containing inf and sup values over small portions within the block. Reconstruction of compressed images is performed exploiting $\mathcal{RF}$-product operator. The method is shown to efficiently encode images in terms of high peak signal to noise ratio (PSNR) values, while alleviating the blocking effect problem.

1.7.3 Rough Fuzzy Color Image Segmentation

The third contribution is a color image segmentation technique which exploits the given definition of rough fuzzy sets. The segmentation is performed by partitioning each block in multiple rough fuzzy sets that are used to build a lower and a upper histogram in the HSV color space. For each bin of the lower and upper histograms some measures are computed to find the best segmentation of the image. It is shown that the proposed method retains the structure of the color images leading to an effective segmentation.

1.8 How to read the thesis

The first part is devoted to the presentation of the rough-fuzzy framework for handling uncertainty at feature level. This part is composed by three chapters. Chapter two presents the mathematical foundations of rough sets, fuzzy sets and their hybridization. Also in this chapter a brief survey of applications of the hybridized rough and fuzzy sets is presented. The third chapter is a survey of techniques which employ rough sets and fuzzy sets theories in image analysis. The survey will be carried out following two points of view: the first one will illustrate methods where rough and fuzzy theories are employed separately; the second one will show techniques which exploit hybridization of rough and fuzzy theories. Chapter four presents the first contribution of this dissertation. First the rough fuzzy framework for the hybridization of rough sets and fuzzy sets theories is presented, followed by the definition of $\mathcal{RF}$-product along with the proof of its basic properties. Chapter five presents the second contribution, i.e. a compression algorithm which exploits the peculiarities of the proposed framework and the comparison with other compressions scheme. Chapter six presents the third contribution, that is the presentation of a segmentation algorithm based on the rough fuzzy framework exposed in Chapter three. The seventh chapter resumes the main results obtained in the thesis and proposes possible themes that could be further investigated in future.

2 Rough Sets and Fuzzy Sets

2.1 Introduction

An important issue, recently discussed with respect to the notion of a set, is vagueness. From a mathematical stand point, it is required that all the concepts must be exact. Although in computer science a growing interested is devoted to handle real life vague concepts just like humans do. Fuzzy set, proposed by Lofti Zadeh, is one of the first approach to handle vagueness by means of partial membership to a set. A different approach try to handle the vagueness introducing the concept of boundary of a set that represents the amount of knowledge about a set. An empty boundary region means that the set is crisp, otherwise the set is rough, i.e., the knowledge about the set is not sufficient to precisely define the set.

2.2 Fuzzy Sets theory

2.2.1 Fuzzy Logic

Since Zadeh firstly proposed the concept of "Fuzzy Sets" in 1965 [142], fuzzy set theory and its applications have been developed quickly and widely [61], [71]. Mainly in fields as control and artificial intelligence, it has been proved that fuzzy logic is a powerful mathematical tool for dealing with modeling and control aspects of complex processes, which transparently express the conflicting character of the precision of the model and the degree of its generality, i.e. the principle of incompatibility [94], [143].
Fuzzy sets was specifically designed to mathematically represent uncertainty and vagueness and to provide formalized tools for dealing

with the imprecision intrinsic to many problems. The notion of an infinite-valued logic was introduced in Zadeh seminal work “Fuzzy Set” where he described the mathematics foundation of fuzzy set theory, and by extension fuzzy logic. This theory considered the membership function to operate over the range of real numbers [0 , 1]. New operations for the calculus of logic were proposed, and showed to be in principle at least a generalization of classic logic. Fuzzy logic provides an inference morphology that enables approximate human reasoning capabilities to be applied to knowledge-based systems. The theory of fuzzy logic provides a mathematical tool to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning. The conventional approaches to knowledge representation lack the means for representing the meaning of fuzzy concepts. As a consequence, the approaches based on first order logic and classical probability theory do not provide an appropriate conceptual framework for dealing with the representation of commonsense knowledge, since such knowledge is by its nature both lexically imprecise and non-categorical. The development of fuzzy logic was motivated in large measure by the need for a conceptual framework which can address the issue of uncertainty and lexical imprecision. Some of the essential characteristics of fuzzy logic relate to the following sentences [144]:

- exact reasoning is viewed as a limiting case of approximate reasoning
- everything is a matter of degree
- knowledge is interpreted as fuzzy constraints on a collection of variables
- inference is viewed as a process of propagation of elastic constraints
- any logical system can be fuzzified

2.2.2 Fuzzy Sets

A classical "crisp" set is a collection of distinct objects. It is defined in such a way as to dichotomize the elements of a given universe of discourse into two groups: members and non-members. A crisp set can be defined by the so-called characteristic function. Let U be a universe of discourse, the characteristic function $\mu_A(x)$ of a crisp set A in U takes its values in $\{0,1\}$ and is defined such that

$$\mu_A(x) = \begin{cases} 1 & \text{iff } x \in A \\ 0 & \text{iff } x \notin A \end{cases} \tag{2.1}$$

We note that the boundary of a set A is rigid and sharp and performs a two-class dichotomization (i.e. $x \in A or x \notin A$) of the universe. A Fuzzy Set, on the other hand, introduces vagueness by eliminating the sharp boundary that divides members from non members in the group. Thus the transition between full membership and non-membership is gradual rather than crisp. Hence, fuzzy sets may be viewed as an extension and generalization of the basic concepts of crisp set [58]. A fuzzy set A in the universe of the discourse U can be defined as a set of ordered pairs

$$A = \{(x, \mu_A(x))|x \in U\} \tag{2.2}$$

where $\mu_A(\cdot)$ is called the membership function of A and $\mu_A(x)$ is the degree of membership of x in A.

Definition 1. *Fuzzy Set. Let U be a nonempty set. A Fuzzy Set A in U is characterized by its membership function*

$$\mu_A : U \to M \tag{2.3}$$

where $\mu_A(x)$ is interpreted as a the degree of membership of element x in fuzzy set A for each $x \in U$.

When $M = \{0, 1\}$ set A is nonfuzzy and $\mu_A(\cdot)$ is the characteristic function of the crisp set A. For fuzzy sets, the range of the membership function is a subset of the non-negative real numbers whose supremum is finite. In most general cases, M is a set on the unit interval [0 , 1]. We note that $\mu_A(x) \in [0, 1]$ indicates the membership grade of an element $x \in U$ in fuzzy set A and that it is not a probability because $\sum \mu_A(x) \neq 1$. Another way of representing a fuzzy set is through use of support of a fuzzy set. The support of a fuzzy set A is the crisp set of all $x \in U$ such that $\mu_A(x) > 0$. That is

$$Supp(A) = \{x \in U | \mu_A(x) > 0\} \tag{2.4}$$

A fuzzy set A whose support is a single point in U with $\mu_A(x) = 1$ is referred to as a fuzzy singleton. The height of a fuzzy set A is the supremum of $\mu_A(x)$ over U . That is

$$Height(A) = \sup_x \mu_A(x) \tag{2.5}$$

A fuzzy set is normalized when the height of the fuzzy set is unity (i.e. $Height(A) = 1$); otherwise it is subnormal. A nonempty fuzzy set A can always be normalized by dividing $\mu_A(x)$ by the height of A. Using the support of a fuzzy set A, we can simplify the representation of a fuzzy set A as

$$A = \sum_{i=1}^{n} \mu_i / x_i \tag{2.6}$$

where the summation indicates the union of the elements and μ_i is the grade of the membership of x_i.

Definition 2. *α-cut. An α-level set of a fuzzy set A of U is a nonfuzzy set denoted by A_α and is defined by*

$$A_\alpha = \begin{cases} \{x \in U | \mu_A(x) \geq \alpha\} & if\ \alpha \geq 0 \\ cl(Supp(A)) & if\ \alpha = 0 \end{cases} \tag{2.7}$$

where $cl(Supp(A))$ denotes the closure of the support of A.
Now we introduce the resolution principle which indicates that a fuzzy set A can be expanded in terms of its α-cuts.

Theorem 1. *Let A be a fuzzy set in the universe of the discourse U. Then the membership function of A can be expressed in terms of the characteristic functions of its α-cuts according to*

$$\mu_A(x) = \sup_{\alpha \in (0,1]} [\alpha \wedge \mu_{A_\alpha}(x)] \tag{2.8}$$

where $\wedge$ denotes the min *operation and μ_{A_α} is the characteristic function of the crisp set A_α,*

$$\mu_{A_\alpha} = \begin{cases} 1 & \textit{iff } x \in A_\alpha \\ 0 & \textit{otherwise} \end{cases} \tag{2.9}$$

Theorem 3 leads to the following representation of a fuzzy set A using the resolution principle. Let A be a fuzzy set in the universe of discourse U. Let αA_α denotes a fuzzy set with membership function

$$A = \bigcup_{\alpha \in \Lambda_A} \alpha A_\alpha \tag{2.10}$$

The resolution principle states that a fuzzy set A can be decomposed into αA_α, $\alpha \in (0,1]$. On the other hand, a fuzzy set A can be retrieved as union of its αA_α, which is called the *representation theorem*.
A fuzzy set A of U is called *convex* if A_α is a convex subset of U, $\forall \alpha \in (0,1]$. From this we define a *fuzzy number*

Definition 3. *A fuzzy number A is a fuzzy set of the real line with a normal, fuzzy convex and continuous membership function of bounded support. The family of fuzzy numbers will be denoted by $\Im$.*

A fuzzy set A is called triangular fuzzy number with center a, left width $\alpha > 0$ and right width $\beta > 0$ if its membership function has the following form

$$A(t) = \begin{cases} 1 - \frac{a-t}{\alpha} & \text{if } a - \alpha \leq t \leq a \\ 1 - \frac{t-a}{\beta} & \text{if } a \leq t \leq a + \beta \\ 0 \text{ otherwise} \end{cases} \tag{2.11}$$

We use the notation $A = (a, \alpha, \beta)$. The support of A is $(a-\alpha, b+\beta)$. A fuzzy set A is called *trapezoid fuzzy number* with tolerance interval $[a, b]$, left width β if its membership function has the following form

$$A(t) = \begin{cases} 1 - \frac{a-t}{\alpha} & \text{if } a - \alpha \leq t \leq a \\ 1 & \text{if } a \leq t \leq b \\ 1 - \frac{t-b}{\beta} & \text{if } a \leq t \leq a + \beta \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

We use the notation $A = (a, b, \alpha, \beta)$. The support of A is $(a-\alpha, b+\beta)$. A fuzzy set A is called *Gaussian fuzzy number* with tolerance interval $[a, b]$ if its membership function has the following form

$$G(t) = e^{-\frac{1}{2}(\frac{t-c}{\sigma})^2} \tag{2.13}$$

A fuzzy set A is called *Generalized-Bell fuzzy number* with tolerance interval $[a, b]$ if its membership function has the following form depending by the parameters a, b, c

$$GB(t) = \frac{1}{1 + |\frac{t-c}{a}|^{2b}} \tag{2.14}$$

Definition 4. *Any fuzzy number $A \in \Im$ can be described as*

$$A(t) = \begin{cases} L(\frac{a-t}{\alpha}) & \textit{if } t \in [a - \alpha, a] \\ 1 & \textit{if } t \in [a, b] \\ R(\frac{t-b}{\beta}) & \textit{if } t \in [b, b + \beta] \\ 0 & \textit{otherwise} \end{cases} \tag{2.15}$$

where $[a, b]$ is the core of A, and $R, L : [0, 1] \to [0, 1]$ are continuous and non-increasing shape functions with $L(0) = R(0) = 1$ and $L(1) =$

$R(1) = 0$. *We call this fuzzy interval of* $LR-$*type and refer to it by* $A = (a, b, \alpha, \beta)_{RL}$. *The supports of* A *is* $(a-\alpha, b+\beta)$. *Let* A *be a fuzzy number. If* $Supp(A) = \{x_0\}$ *then* A *is called a fuzzy point and we use the notation* $A = \overline{x}_0$.

2.2.3 t-norm, t-conorm and complement

A complement of a fuzzy set A, denoted as $\overline{A}$, is specified by a function $c : [0,1] \rightarrow [0,1]$ such that $\mu_{\overline{A}}(x) = c(\mu_A(x))$ where the function $c(\cdot)$ satisfies the following conditions:

1. Boundary conditions: $c(0) = 1$ and $c(1) = 0$
2. Monotonic properties: for any $x_1, x_2 \in U$, if $\mu_a(x_1) < \mu_a(x_2)$ then $c(\mu_a(x_1)) \geq c(\mu_a(x_2))$
3. Continuity: $c(\cdot)$ is a continuous function
4. Involution: $c(\cdot)$ is involutive, which means $c(c(\mu_A(x))) = \mu_A(x) \forall x \in U$

Next, let us discuss the intersection and union operations on fuzzy sets, which are often referred as *triangular norms (or t-norms)* and *triangular conorms (or t-conorms)*, respectively [50] [21] [22].
T-norms are two parameters functions of the form $t : [0,1] \times [0,1] \rightarrow [0,1]$ such that

$$\mu_{A \bigcap B}(x) = t[\mu_A(x), \mu_B(x)] \tag{2.16}$$

where the function $t(\cdot)$ satisfies the following conditions:

1. Boundary conditions: $t(0,0) = 0$, $t(\mu_A(x), 1) = t(1, \mu_A(x)) = \mu_A(x)$
2. Commutativity: $t(\mu_A(x), \mu_B(x)) = t(\mu_B(x), \mu_A(x))$
3. Monotonicity: if $\mu_A(x) \leq \mu_C(x)$ and $\mu_B(x) \leq \mu_D(x)$ then $t(\mu_A(x), \mu_B(x)) \leq t(\mu_C(x), \mu_D(x))$
4. Associativity: $t(\mu_A(x), t(\mu_B(x), \mu_C(x))) = t(t(\mu_A(x), \mu_B(x)), \mu_C(x))$

Typical t-norms are ($a \equiv \mu_A(x)$ and $b \equiv \mu_B(x)$)

1. Intersection: $a \wedge b = min(a, b)$
2. Algebraic product: $a \cdot b = ab$
3. Lukasiewicz: $a \odot b = max(0, a + b - 1)$
4. Drastic product:

$$a\hat{\cdot}b = \begin{cases} a & b = 1 \\ b & a = 1 \\ 0 & a, b < 1 \end{cases}$$

T-conorms (also called s-norms) are two parameters functions of the form $s : [0, 1] \times [0, 1] \rightarrow [0, 1]$ such that

$$\mu_{A \bigcup B}(x) = s[\mu_A(x), \mu_B(x)] \tag{2.17}$$

where the function $s(\cdot)$ satisfies the following conditions:

1. Boundary conditions: $s(1, 1) = 1$, $s(\mu_A(x), 0) = s(0, \mu_A(x)) = \mu_A(x)$
2. Commutativity: $s(\mu_A(x), \mu_B(x)) = s(\mu_B(x), \mu_A(x))$
3. Monotonicity: if $\mu_A(x) \leq \mu_C(x)$ and $\mu_B(x) \leq \mu_D(x)$ then $s(\mu_A(x), \mu_B(x)) \leq s(\mu_C(x), \mu_D(x))$
4. Associativity: $s(\mu_A(x), s(\mu_B(x), \mu_C(x))) = s(s(\mu_A(x), \mu_B(x)), \mu_C(x))$

Typical t-conorms are ($a \equiv \mu_A(x)$ and $b \equiv \mu_B(x)$)

1. Union: $a \vee b = max(a, b)$
2. Algebraic sum: $a \hat{+} b = a + b - ab$
3. Lukasiewicz: $a \oplus b = min(1, a + b)$
4. Disjoint sum: $a \bigtriangleup b = max\{min(a, 1 - b), min(1 - a, b)\}$

5. Drastic sum:

$$a\hat{\cdot}b = \begin{cases} a & b = 0 \\ b & a = 0 \\ 0 & a, b > 1 \end{cases}$$

The relations among various t-norms and t-conorms are described by the following Theorem

Theorem 2. *Let A and B to be fuzzy sets in the universe of discourse U. The t-norms are bounded by the inequality*

$$t_{dp}(a, b) = t_{min}(a, b) \leq t(a, b) \leq t_{max}(a, b) = min(a, b) \qquad (2.18)$$

where $t_{dp}(a, b)$ is the drastic product. Similarly, the t-conorms are bounded by the inequalities

$$max(a, b) = s_{min}(a, b) \leq s(a, b) \leq s_{max}(a, b) = s_{ds}(a, b) \qquad (2.19)$$

where $s_{ds}(a, b)$ is the drastic sum.

Based on t-norms and t-conorms, we shall further introduce some operations of fuzzy sets that are central to fuzzy logic and fuzzy reasoning. Let A and B fuzzy sets in the universal sets U and V, respectively, we have

1. Fuzzy conjunction: the fuzzy conjunction of A and B is denoted as $A \wedge B$ and defined

$$\mu_{A \wedge B}(x, y) \triangleq t(\mu_A(x), \mu_B(y)) \qquad (2.20)$$

where t is a t-norm

2. Fuzzy disjunction: the fuzzy disjunction of A and B is denoted as $A \vee B$ and defined by

$$\mu_{A \vee B}(x, y) \triangleq s(\mu_A(x), \mu_B(y)) \qquad (2.21)$$

where s is a t-conorm

3. Fuzzy implication: the fuzzy implication of A and B is denoted as $A \rightarrow B$ and has five different definitions. In the following, t is a t-norm, s is a t-conorm, and the complement is the standard complement
   - Material implication: $A \rightarrow B = s(\overline{A}, B)$
   - Propositional calculus: $A \rightarrow B = s(\overline{A}, t(A, B))$
   - Extended propositional calculus: $A \rightarrow B = s(\overline{A} \times \overline{B}, B)$
   - Generalized of modus ponens: $A \rightarrow B = \sup\{k \in [0,1], t(A,k) \leq B\}$
   - Generalized of modus tollens: $A \rightarrow B = \inf\{k \in [0,1], s(B,k) \leq A\}$

2.2.4 Extension Principle

The *extension principle* introduced by Zadeh [145], is one of the most important tool of fuzzy sets theory. This principle allows the generalization of crisp mathematical concepts to the fuzzy set framework and extends point-to-point mappings to mappings for fuzzy sets. It provides a means for any function f that maps an n-tuple $(x_1, \ldots, x_n)$ in the crisp set U to a point in the crisp set V to be generalized to mapping n fuzzy subsets in U to a fuzzy subset in V. Hence, any mathematical relationship between non fuzzy elements can be extended to deal with fuzzy entities. Furthermore, the extension principle is very useful for dealing with set-theoretic operations for high-order fuzzy sets.
Given a function $F : U \rightarrow V$ and a fuzzy set A in U, where $A = \sum_{i=1}^{n} \mu_i / x_i$, the extension principle states that

$$f(A) = f(\sum_{i=1}^{n} \mu_i / x_i) = \sum_{i=1}^{n} \mu_i / f(x_i) \tag{2.22}$$

If more than one element of U is mapped to the same element of y in V by f, then the maximum among their membership grades is taken, that is,

$$\mu_{f(A)}(y) = \max_{x_i \in U : f(x_i) = y} [\mu_A(x_i)] \tag{2.23}$$

where x_i represents the elements that are mapped to the same y. Quite often the function f that is of interest maps n-tuple in U to a point V. The extension principle allows the function $f(x_1, \ldots, x_n)$ with $(x_1, \ldots, x_n)$ be an n-tuple to act on the n fuzzy subsets of U.

2.3 Rough Sets theory

2.3.1 Rough Sets

Rough set theory, introduced by Pawlak [90] in the early 1980s, is a mathematical approach that can be employed to handle imprecision, vagueness and uncertainty. Rough sets have many important advantages for data mining, such as providing efficient algorithms for finding hidden patterns in data, finding minimal sets of data, generating sets of decision rules from data, and offering straightforward interpretation of obtained results. In the last two decades, rough sets have widely been applied to data mining and rapidly established themselves in many real-life applications such as medical diagnosis, control algorithm acquisition and process control and image processing. The main advantage of rough set theory is that it needs no apriori knowledge or additional information about data, like, for instance, membership functions in fuzzy set theory.

The basic concept for data representation in the rough set framework is an information system. An information system I can be defined in terms of a pair (U, A)

$$I = (U, A) \tag{2.24}$$

where U is a non-empty finite set of objects and A is a non-empty finite set of attributes. Each attribute $a \in A$ can be viewed as a function that maps elements of U into a set V_a

$$a : U \rightarrow V_a \tag{2.25}$$

The set V_a is called the *value set* of attribute a. The value of attribute a for object x is said to be missing if $a(x)$ has not been observed. Values may be missing for a variety of reasons. How missing values should be interpreted and subsequently treated depends on the application domain. In the following, "⊤" will be used to denote a missing value, and is assumed to be a member of every value set. An information I defines a matrix M_{textitI}, of dimensions $|U| \times |U|$, called *discernibility matrix*. Each entry $M_{\text{textitI}}(x, y) \subseteq A$ consists of the set of attributes that can be used to discern between objects $x, y \in U$

$$M_{\text{textitI}}(x, y) = \{a \in A | \mathit{discerns}(a, x, y)\} \tag{2.26}$$

where $\mathit{discerns}(a, x, y)$ is defined as

$$\mathit{discerns}(a, x, y) \Leftrightarrow a(x) \neq a(y) \tag{2.27}$$

A discernibility matrix M_A defines a binary relation $R_A \subseteq U^2$, called *indiscernibility relation* with respect to A

$$xR_A y \Leftrightarrow M_A(a, y) = \emptyset \tag{2.28}$$

which expresses the pairs of objects that cannot be discerned between. A generalization of 2.28 would allow to include in the relation R_A all those objects that do not differ "enough". The properties of R_A vary according to how the discernibility function is defined. If it is defined as in 2.27, R_A is an equivalence relation, i.e. it shows three properties

1. *Reflexivity*: $xR_A x$, $\forall x \in U$
2. *Symmetry*: if $xR_A y$ then $yR_A x$, $\forall x, y \in U$
3. *Transitivity*: if $xR_A y$ and $yR_A z$ then $xR_A z$, $\forall x, y, z \in U$

Relations that are reflexive and symmetric but not transitive are sometimes referred to as *similarity relations* or *tolerance relations*. In standard rough set theory the indiscernibility relation is required to be an equivalence relation, but less restrictive extensions of rough

set theory do not require the transitivity condition to hold.
The *indiscernibility set* of an object $x \in U$ is denoted by $R_A(x)$, and consists of those objects that stand in relation to object x by R_A, that is

$$R_A(x) = \{y \in U | xR_| y\} \tag{2.29}$$

If R_A is an equivalence relation, then the indiscernibility sets are called *equivalence classes*. Equivalence relations induce a partition of the universe, meaning that all equivalence classes are disjoint and their union equals the full universe U. Vice versa, a partition also induces an equivalence relation. In the more general case of tolerance relations, the indiscernibility sets form a covering of U, meaning that the indiscernibility sets are allowed to overlap.
The basic idea behind rough sets is to construct approximations of sets using the binary relation R_A. The indiscernibility sets $R_A(x)$, also called *granules*, form basic building blocks from which subsets $X \subset U$ can be defined.
Let U be a finite set of objects and $R \subseteq U \times U$ be a binary relation. The sets U, R are the universe of discourse and an indiscernibility relation, respectively. The discernibility relation represents our lack of knowledge about elements of U. For simplicity, we assume that R is an equivalence relation. A pair (U, R) is called an approximation space, where U is the universe and R is an equivalence relation on U. Let X be a subset of U, i.e. $X \subseteq U$. Our goal is to characterize the set X with respect to R. Using only the indiscernibility relation, in general, we are not able to observe individual objects from U but only the accessible granules of knowledge described by this relation.

- The set of all objects which can be with certainty classified as members of X with respect to R is called the *R-lower approximation* of a set X with respect to R, and denoted by $\underline{R}(X)$, i.e.

$$\underline{R}(X) = \{x | R(x) \subseteq X\} \tag{2.30}$$

- The set of all objects which can be only classified as *possible* members of X with respect to R is called the *R-upper approximation* of a set X with respect to R, and denoted by $\overline{R}(X)$, i.e.

$$\overline{R}(X) = \{x|R(x) \cap X \neq \emptyset\} \tag{2.31}$$

- The set of all objects which can be decisively classified neither as members of X nor as members of $-X$ with respect to R is called the *boundary region* of a set X with respect to R, and denoted by $RN_R(X)$, i.e.

$$RN_R(X) = \overline{R}(X) - \underline{R}(X) \tag{2.32}$$

Definition 5. *A set X is called crisp (exact) with respect to R if and only if the boundary region of X is empty. A set X is called rough (inexact) with respect to R if and only if the boundary region of X is nonempty.*

The definitions of set approximations presented above can be expressed in terms of granules of knowledge in the following way. The lower approximation of a set is the union of all granules which are entirely included in the set; the upper approximation of a set is the union of all granules which have non-empty intersection with the set; the boundary region of a set is the difference between the upper and the lower approximation of the set. It is interesting to compare definitions of classical sets, fuzzy sets and rough sets. Classical set is a primitive notion and is defined intuitively or axiomatically. Fuzzy sets are defined by employing the fuzzy membership function, which involves advanced mathematical structures, numbers and functions. Rough sets are defined by approximations. Thus this definition also requires advanced mathematical concepts. The previous definitions of approximations respect the following properties:

1. $\underline{R}(X) \subseteq X \subseteq \overline{R}(X)$

2. $\underline{R}(\emptyset) = \overline{R}(\emptyset) = \emptyset; \underline{R}(U) = \overline{R}(U) = U$
3. $\overline{R}(X \cup Y) = \overline{R}(X) \cup \overline{R}(Y)$
4. $\underline{R}(X \cap Y) = \underline{R}(X) \cup \underline{R}(Y)$
5. $\underline{R}(X \cup Y) \supseteq \underline{R}(X) \cup \underline{R}(Y)$
6. $\overline{R}(X \cap Y) \subseteq \overline{R}(X) \cap \overline{R}(Y)$
7. $X \subseteq Y \rightarrow \underline{R}(X) \subseteq \underline{R}(Y)$ and $\overline{R}(X) \subseteq \overline{R}(Y)$
8. $\underline{R}(-X) = -\overline{R}(X)$
9. $\overline{R}(-X) = -\underline{R}(X)$
10. $\underline{R}(\underline{R}(X)) = \overline{R}(\underline{R}(X)) = \underline{R}(X)$
11. $\overline{R}(\overline{R}(X)) = \underline{R}(\overline{R}(X)) = \overline{R}(X)$

It is easily seen that the lower and the upper approximations of a set are, respectively the interior and closure of this set in the topology generated by the indiscernibility relation. One can define the following four basic classes of rough sets, i.e., four categories of vagueness:

1. A set X is roughly R-definable, iff $\underline{R}(X) \neq \emptyset$ and $\overline{R}(X) \neq U$
2. A set X is internally R-undefinable, iff $\underline{R}(X) = \emptyset$ and $\overline{R}(X) \neq U$
3. A set X is externally R-undefinable, iff $\underline{R}(X) \neq \emptyset$ and $\overline{R}(X) = U$
4. A set X is totally R-undefinable, iff $\underline{R}(X) = \emptyset$ and $\overline{R}(X) = U$

The intuitive meaning of this classification is the following. A roughly R-definable set X means that with respect to R we are able to decide for some elements of U that they belong to X and for some elements of U that they belong to $-X$. An internally R-undefinable set X means with respect to R we are able to decide for some elements of U that they belong to $-X$, but we are unable to decide for any element of U whether it belongs to X. An externally R-undefinable set X means that with respect to R we are able to decide for some elements of U that they belong to X, but we are unable to decide for any element of U whether it belongs to $-X$. A totally R-undefinable set X means that with respect to R we are unable to decide for any element of U whether it belongs to X or $-X$. In [90], Pawlack discusses two numerical characterization of imprecision of a subset X in the approximation space $\langle U, R\rangle$: accuracy and roughness. Accuracy of X, which is denoted by $\alpha_R(X)$, is the ratio of the number of objects on its lower approximation to that on its upper approximation, namely

$$\alpha_R(X) = \frac{|\underline{R}(X)|}{|\overline{R}(X)|} \tag{2.33}$$

The roughness of X, which is denoted by $\rho_R(X)$, is defined as $\rho_R(X) = 1 - \alpha_R(X)$. Note that the lower the roughness of a subset, the better is its approximation. Furthermore, the following conditions are noted

1. $0 \leq \rho_R(X) \leq 1$
2. $X = \emptyset \rightarrow \underline{R}(X) = \overline{R}(X) = \emptyset \rightarrow \rho_R(X) = 0$
3. $\rho_R(X) = 0$ iff X is definable in $\langle U, R\rangle$

2.3.2 Rough Membership Function

Rough sets can be also defined by using, instead of approximations, a rough membership function proposed in [91][132]. This view is called set-oriented as opposed to the above formulation called operator-oriented[136][137]. In classical set theory, either an element belongs to a set or it does not. The corresponding membership function is the characteristic function for the set, i.e. the function takes values 1 and 0, respectively. In the case of rough sets, the notion of membership is different. The rough membership function quantifies the degree of relative overlap between the set X and the equivalence class $R(x)$ to which x belongs. It is defined as follows:

$$\mu_X^R : U \rightarrow < 0, 1 > \tag{2.34}$$

where

$$\mu_X^R = \frac{|X \cap R(x)|}{|R(x)|} \tag{2.35}$$

and $|X|$ denotes the cardinality of X. The rough membership function expresses conditional probability that x belongs to X given R and can be interpreted as a degree that x belongs to X in view of information about x expressed by R. The rough membership function can be used to define approximations and the boundary region of a set as

$$\mu_X^R = 1 \text{ iff } x \in \underline{R}(X) \tag{2.36}$$

$$\mu_X^R = 0 \text{ iff } x \in U - \overline{R}(X) \tag{2.37}$$

$$0 < \mu_X^R(x) < 1 \text{ iff } x \in RN_R(X) \tag{2.38}$$

$$\mu_{U-X}^R = 1 - \mu_X^R, \forall x \in U \tag{2.39}$$

$$\mu_{X \cup Y}^R \geq \max\left(\mu_X^R(x), \mu_Y^R(x)\right), \forall x \in U \tag{2.40}$$

$$\mu_{X \cap Y}^R \leq \min\left(\mu_X^R(x), \mu_Y^R(x)\right), \forall x \in U \tag{2.41}$$

2.4 Rough and Fuzzy Hybridization

2.4.1 Rough Sets and Fuzzy Sets

The set-oriented view of rough sets is defined over a classical set algebra and associates a fuzzy set with each subset of the universe.

Vagueness arises in concept representation from the lack of information when defining a precise concept. In this view, rough membership functions can be thought as a special type of fuzzy membership functions, defined as probabilities simply derived by cardinality of sets. In fact one can use a probability function on the universe to define rough membership functions[152]. From a fuzzy logic point of view, lower and upper approximations can be defined with respect to a fuzzy set μ_A as

$$\underline{A} = \{x|\mu_A(x) = 1\} \tag{2.42}$$

$$\overline{A} = \{x|\mu_A(x) > 0\} \tag{2.43}$$

i.e., $\underline{A}$ and $\overline{A}$ are the core and the support of the fuzzy set μ_A, respectively. Although, in the theory of fuzzy sets, the membership value of an element does not depend on other elements, where in the theory of rough sets, with respect to an equivalence relation, the membership value of an element depends on other elements [12]. In the study of fuzzy sets, many types of fuzzy membership functions have been proposed so to implicitly specify the membership value of one element with respect to other elements [51].
It is clear that both theories provide means to handle vague concepts, even if from different points of view, and hence it is not surprising that many efforts have been performed to combine rough and fuzzy approach to obtain more general and powerful tools.

The two theories seem to complement each other and hence researchers have explored a variety of different ways in which these two theories interact with each other. The origins of both theories were essentially logical and hence, much of the hybridization between fuzzy and rough set theory is logically based. Moreover, rough set theory was proposed both for supervised and unsupervised learning.

Two combinations of rough set theory and fuzzy set theory lead to distinct generalization of classical set theory. By using an equivalence relation on the universe of discourse, one can introduce lower and upper approximations in fuzzy set theory to obtain an extended notion called rough fuzzy sets [23]. Alternatively, a fuzzy similarity relation can be used to replace an equivalence relation, that result in another notion called fuzzy rough sets [23].

The expressions for the lower and upper approximations of a set X depend on the type of relation R and whether X is a crisp or a fuzzy set. When X is a crisp or a fuzzy set and the relation R is a crisp or a fuzzy equivalence relation, the expressions for the lower and upper approximations of the set X are given by

$$\underline{R}X = \{(u, \underline{M}(u)) | u \in U\} \tag{2.44}$$

$$\overline{R}X = \{(u, \overline{M}(u)) | u \in U\} \tag{2.45}$$

where

$$\underline{M}(u) = \sum_{Y \in U/R} m_Y(u) \times \inf_{\varphi \in U} \max(1 - m_Y(\varphi), \mu_X(\varphi)) \tag{2.46}$$

$$\overline{M}(u) = \sum_{Y \in U/R} m_Y(u) \times \sup_{\varphi \in U} \min(m_Y(\varphi), \mu_X(\varphi)) \tag{2.47}$$

where the membership function m_Y is the membership degree of each element $u \in U$ to a granule $Y \in U/R$ and takes values in $[0, 1]$, and μ_X is the membership function associated with X and takes values in $[0, 1]$. When X is a crisp set, μ_X would take values only from the set $\{0, 1\}$. Similarly, when R is a crisp equivalence relation, m_Y would take values only from the set $\{0, 1\}$. Fuzzy union and intersection are chosen based on their suitability with respect to the underlying application of measuring ambiguity. The pair of sets $\langle \underline{R}X, \overline{R}X \rangle$ and the approximation space U/R are referred to differently, depending on whether X is a crisp or a fuzzy set and the relation R is a crisp or a fuzzy equivalence relation. The different combinations are listed in Table 2.1[109].

Table 2.1: Different combinations of rough and fuzzy sets.

X	R	$\langle \underline{R}X, \overline{R}X \rangle$	U/R
Crisp	crisp equivalence relation	rough set of X	crisp equivalence approximation space
Fuzzy	crisp equivalence relation	rough fuzzy set of X	crisp equivalence approximation space
Crisp	fuzzy equivalence relation	fuzzy rough set of X	fuzzy equivalence approximation space
Fuzzy	fuzzy equivalence relation	fuzzy rough fuzzy set of X	fuzzy equivalence approximation space

Hence the approximation of a crisp set in a fuzzy approximation space is called a fuzzy-rough set, and the approximation of a fuzzy set in a crisp approximation space is called a rough-fuzzy set, making the two models complementary [138]. In this framework, the approximation of a fuzzy set in a fuzzy approximation space is considered to be a more general model, unifying the two theories. In [104] a broad family of fuzzy-rough sets is constructed substituting min and max operators by different implicators and t-norms, and the properties of three well-known classes of implicators (S-, R- and QL-implicators) are investigated. Further research in the area of rough and fuzzy hybridization from different perspectives, can be found in [18], [122], [134], [140]. In [133] the properties of generalized fuzzy-rough sets are investigated, defining a pair of dual generalized fuzzy approximation operators based on arbitrary fuzzy relations, while in [67] a new approach introduces definitions for generalized fuzzy lower and upper approximation operators determined by a residual implication. Assumptions are found that allow a given fuzzy set-theoretic operator to represent a lower or upper approximation from a fuzzy relation. Different types of fuzzy relations produce different classes of fuzzy rough set algebras.

Other generalizations are possible in addition to the previous hybridization approaches. One of the first attempts at hybridizing the two theories is reported in [134], where the negative, boundary and positive regions of a rough set are expressed by means of a fuzzy membership function. All objects in the positive region have a membership of one, those belonging to the boundary region have a membership of 0.5, while those contained in the negative region have zero membership (i.e., they do not belong to the rough set). This construction leads to express a rough set as a fuzzy set, with suitable modifications to the rough union and intersection operators. Another approach that exploits the similarities between rough and fuzzy sets has been proposed in [96] where the author introduces the concept of shadowed set. The main idea comes from the consideration that a numeric fuzzy set representation may be too precise, and that is because a concept can be described only once its membership function has been defined. It is like requiring excessive precision in order to describe imprecise concepts. Shadowed set does not use exact membership values but adopts basic truth values and a zone of uncertainty (the unit interval), where elements may belong with certainty (membership of 1), possibility (unit interval) or not at all (membership of 0). This can be seen to be analogous to the rough set definitions for the positive, boundary and negative regions.

2.4.2 Applications of Hybridized Rough and Fuzzy Sets

Hybridized versions of rough and fuzzy sets have been employed in many applications fields to solve a huge variety of problems. The ratio of the use of these techniques is that of exploiting both uncertainty and indiscernibility deeply embedded in real life data.

In the field of supervised learning, many extensions to classical approaches have been propose. In [107] a fuzzy-rough nearest neighbor classification approach is presented, that employs a fuzzy-rough ownership function to incorporate fuzzy uncertainty, caused by overlapping classes, and rough uncertainty, through insufficient features. An extended version is presented in [129]. In [4] fuzzy upper and lower approximations are used to model rough uncertainty into the fuzzy K-NN classifier. In particular, the k nearest neighbors of the test pattern are used to compute its membership degree to the fuzzy lower and upper approximations for every class.

Another field in which theories of rough sets and fuzzy sets have found natural application is that of neural computation. Such hybridization has encountered increasing popularity because of the various aspects which could exploit its characteristics. Here we show a brief survey of different kind of hybridizations that have been proposed in literature. Lingras [55], [56] [57] showed how the use of upper and lower approximations of ranges of numbers may reduce training time and improve performance. Also he showed a way to combine rough neurons with fuzzy neurons. Different approach to the hybridization of rough and fuzzy in knowledge based networks are reported in [82]. In particular, Pal et al. [83] used an innovative combination of multiple technologies including rough sets, fuzzy sets, neurocomputing, and genetic algorithms that provided accelerated training and a compact network suitable for generating a minimum number of rules with high certainty values. The combination of rough and fuzzy neural computing is presented in [32] for classifying faults in high voltage power systems. They also showed how the combination of rough and fuzzy sets compares favorably with fuzzy neural computing. A different approach is employed by Zhang in [146], where the author combined the logical view of rough set theory with fuzzy neurocomputing. Rough set theory was used to eliminate inconsistent and redundant rules from the rule set of the fuzzy system, which is then used to create and train a simple fuzzy neural network. Other examples of rough fuzzy neurocomputing can be found in [25] and [15].

Combination of rough and fuzzy theories has made substantial progress also as extension of unsupervised learning techniques. Asharaf and Murty [2] describe a hybridized fuzzy-rough approach to clustering. Chimphlee, et al. [17] described how feature selection based on independent component analysis can be used for hybridized rough-fuzzy clustering of web user sessions. There are a number of approaches that combine the traditional supervised classification from rough set theory with unsupervised fuzzy clustering. For example, [150] uses a rough set-based fuzzy clustering algorithm in which the objects of fuzzy clustering are initial clusters obtained in terms of equivalence relations. The preponderance of many small classes is countered through secondary clustering on the basis of defining fuzzy similarity between two initial clusters. Wang, et al. [130] integrated fuzzy clustering and variable precision rough set theory. This approach was used to effectively discover association rules in process planning. Traditional hybridization of fuzzy and rough sets can also be seen in [98], where a texture segmentation algorithm is proposed to solve the problem of unsupervised boundary localization in textured images using rough fuzzy sets and hierarchical clustering. Pal [84] describes how rough-fuzzy initialization can be used for clustering with the example of multi-spectral image segmentation. Also he describes how rough set-based rule extraction can be combined with self-organizing maps. In [62] a generalized hybrid unsupervised learning algorithm which integrates both principles of rough and fuzzy sets is proposed. In [69] a novel clustering architecture is presented, by integrating the advantages of both fuzzy sets and rough sets, in which several subsets of patterns can be processed to find a common structure.

Rough and fuzzy hybridization has been also employed in the area of feature selection [41][121], one of the most successful applications of rough set theory. In [38][39][40], a method which employs fuzzy-rough sets to handle uncertainty in website classification is presented. This is a fuzzy extension of the rough set attribute reduction method (RSAR) [115] which only exploits indiscernibility relations by means of rough sets. The use of fuzzy equivalence classes allows the concept of crisp equivalence classes to be extended by the inclusion of a fuzzy similarity relation R on the universe, which determines the extent to which two elements are similar in R, resulting in fuzzy decision and conditional values. Another approach employs a greedy hill-climber to perform subset search, using a fuzzy dependency function both for subset evaluation and as a stopping criterion [42]. In [37] authors presented a new information measure for fuzzy equivalence relations which is used to redefine the dependency of a hybrid attribute set. This entropy based measure is useful to quantify the discernibility power of a fuzzy equivalence relation.

In [125], the concept of attributes reduction with fuzzy rough sets is proposed after a critical analysis of the algorithm in [38]. Authors pointed how the mathematical foundations of various aspects of the proposed approach are missing. For this reason a solid mathematical foundation is set up for attributes reduction with fuzzy rough sets, the structure of reduction is completely studied and an algorithm using discernibility matrix to compute all the attributes reductions is developed.

In [43] three new techniques for fuzzy rough features selection based on the use of fuzzy T-transitive similarity relations have been presented. In the first one, based on fuzzy lower approximations, similarity relations are used to construct approximations of decision concepts which are then evaluated through a new measure of dependency. The second one exploit information in the fuzzy boundary region to guide feature selection so to obtain a fuzzy-rough reduct. The last method represents a fuzzy extension of the discernibility matrix so that features have a certain degree of belongingness to each entry.

3 Rough and Fuzzy Sets in Image Analysis

This chapter presents an overview of rough and fuzzy set theories in the field of image processing [4]. Various methods have been proposed over the years as long as the theories developed and gained more solid theoretical foundation, with the aim of exploiting their fundamental characteristics: vagueness and indiscernibility handling. That is of particular interest if we consider the intrinsic presence of uncertainty when dealing with digital images processing and analysis.

Concepts represented in an image, e.g. a region, are not always crisply defined, hence uncertainty can arise within any processing phase and any decision made at a particular level will have an impact on all higher level activities. A recognition or vision system should have sufficient provision for representing and manipulating the uncertainties involved at every processing stage, so that the system can retain as much information content of the data as possible. The output of the system will then possess minimal uncertainty and, unlike conventional systems, will not be biased/affected much by lower level decisions. For instance, a gray tone image possesses ambiguity within pixels because of the possible multi-valued levels of brightness in the image. This indeterminacy, both in grayness and spatially, is due to inherent vagueness rather than randomness and hence many basic concepts of image analysis (e.g., edges, corners, boundary regions Fig. 3.1) do not lend themselves well to precise definition.

Over the years, many algorithms have been proposed to cope with intrinsic uncertainty in image analysis by exploiting either fuzzy or rough theory. Just to cite two examples, Bezdek [6] presented the famous Fuzzy C–Means often used in image segmentation, while in

Figure 3.1: Example of uncertainty in texture segmentation.

[76] authors presented an image segmentation technique based on rough set theory. Nevertheless, these kind of approaches only address vagueness or uncertainty present in an image. In the recent years a new trend emerged to try to exploit both theories at the same time. This new approaches evolved along two distinct research lines. Techniques belonging to the first one try to combine the two theories in different steps of the algorithm, thus exploiting fuzziness and roughness separately. The second one, which can be considered a more general approach, aims to hybridize both theories, thus exploiting at the same time fuzziness and roughness. As it will be clear at end of the chapter, rough and fuzzy based techniques have proved to be effective in the field of image analysis, as well as in other fields of pattern recognition¡.

3.1 Combined use of rough and fuzzy sets

In this section an overview of methods that combine rough and fuzzy theories is presented, with regard to different tasks in image processing. As stated above, these techniques try to exploit both theories separately but in a coordinated way.

3.1.1 Image Segmentation

In [72][73], the fusion of rough set theory and fuzzy c-means is used for color image segmentation. The technique aims to segment natural images characterized by regions with gradual color variations. Core centers are evaluated through approximations obtained by rough set theoretic, so to reduce the computational complexity required by standard fuzzy c-means (FCM). FCM based segmentation strategies requires a–priori information about the number of clusters and their means as initialization points. The proposed technique extracts color information from the image employing rough set approximations on the segments and presents it as input to FCM for the soft evaluation of the segments. The advantage of the proposed technique is to analyze colors utilizing all the 3-dimension (RGB) as one entity, where many algorithms works on single bands. However, employing FCM requires the definition of a distance between colors that, due to non perceptive uniformity of the RGB color space, can lead to inconsistencies.

In [5] authors propose a color fuzzy decision algorithm to face segmentation in a color image. Main characteristic of the proposed algorithm is the use of fuzzy decision marking to segment image without user interaction, while rough sets are adopted to merge segments and choose the face region in each image. Use of different color quantization in YCbCr color space and fuzzy decision algorithm allows rough sets to correctly merge face skin regions, but only partially addressing the problem of detection in bigger images.

3.1.2 Features Selection

A method to select an optimal group of bands in hyperspectral images based on rough sets and Fuzzy C-Means clustering is proposed in [116]. First Fuzzy C-Means clustering algorithm is used to classify the original bands into equivalent band groups since adjacent bands in hyperspectral image always show strong correlation. The concept of attribute dependency in Rough Sets is used to define the distance between a group and the cluster center. Then data are reduced by selecting only the band with maximum grade of fuzzy membership from each of the groups. In this way the number of bands is decreased while preserving most useful information. Last step consists in either selecting only one from each of the groups or composing the images in each group linearly. The proposed technique exploits one of the most usefull characteristic of rough sets, i.e. attribute dependency, but employs a clustering algorithm which needs the number of clusters to be a–priori known.

Hassanien [34] introduced a hybrid scheme that combines the advantages of fuzzy sets and rough sets in conjunction with statistical feature extraction techniques. First step consists in a fuzzy image processing as pre-processing technique to enhance the contrast of the whole image, to extracts the region of interest and then to enhance the edges surrounding the region of interest. Next, features from the segmented regions of the interested regions are extracted using the grey-level co-occurrence matrix. Rough set is used for generation of all reducts, that contains minimal number of features, and hence rules. Although rough set rules generation allow to identify significant attributes very accurately, the major drawback of this technique is the number of clusters to segment the image which can vary depending on the image.

3.1.3 Image Evaluation

Due to the complexity of fused image quality evaluation, in [135] a hybrid model of knowledge reduction is constructed by means of rough set theory and Fuzzy Support Vector Machine (FSVM). The proposed model combines the reduction ability of rough sets with the classification ability of the FSVM. A reduced information table is obtained by reducing the number of evaluation criteria, without information loss, through rough set method. The reduced information is used to develop classification rules and train fuzzy support vector machine. The proposed approach needs a sufficient number of training samples to overcome FSVM and SVM due to the reduced number of attributes considered during the training phase.

3.1.4 Detection

Gao et al. [29] present a feature reduction method based on rough set theory and fuzzy c-means, to extract rules for shot boundary detection. Based on the characteristics of differences from the classification capability of various features to different shot transition, the correlation between features can be defined using the classification ability of attributes (or dependence between attributes) in rough set theory. Then, by employing fuzzy c-means algorithm, the optimal feature reduction can be obtained. The first step consists in extracting conditional attributes from video sequences. Then, by calculating their correlation, the importance of conditional attributes can be computed. Selected features are obtained by clustering feature attributes with fuzzy c-means. For each class, the fuzzy if-then rule is generated for decision with fuzzy inference. Also this technique presents the same limits of the other approaches based on Fuzzy C-Means, that is the number of clusters.

In [52] two combined classifiers have been discussed in the field of landmines detection. In the first classifier Hebb Net learning is used with rough set theory and in the second one fuzzy filter neural network is used with the rough set theory. Rough sets have been applied to classify the landmine data because in this theory no prior knowledge of rules are needed, hence these rules are automatically discovered from the database. The rough logic classifier uses lower and upper approximations for determining the class of the objects. The neural network is for training the data, and has been used especially to avoid the boundary rules given by the rough sets that do not classify the data with cent percentage probability. Although the combined use of rough set and fuzzy filter classifier gives good results, it can partially reduce the problem of ambiguous patterns belonging to the boundary region.

In [3] a method for object labeling, based on the uncertainty measurement of a fuzzy similarity is presented. The labeling is performed on objects detected in a scene, based on information provided by a set of different sensors. First the fuzzy similarity is computed between the detected object and a rough set of possible prototypes, followed by a measurement of the uncertainty induced by the observation. For all results obtained from each sensor, the global uncertainty, corresponding to the most likely label, is computed. The proposed technique aims to improve the labeling process by suppressing the inconsistent observations and making new labeling determinations. Different prototypes are used in this process, corresponding to different observation distances and positions. Also, from this observations, uncertainty variation can be analyzed, as determined by the switch from one prototype to another. Problems may arise in complex scenes where inconsistencies can be faced at different resolutions, resulting in erroneous labels assignments.

3.2 Hybridization of rough and fuzzy sets

In this section techniques that exploit a different approach to combination of rough and fuzzy theories are summarized. These methods mainly employ the concept of rough-fuzzy sets and fuzzy-rough sets as generalization of their constituent theories. The aim of rough and fuzzy hybridization is to exploit, at the same time, uncertainty and vagueness as a whole, thus leading to better results.

3.2.1 Image Segmentation

Pal [84] describes how rough-fuzzy initialization can be used for clustering with the example of multi-spectral image segmentation.

Mitra et al. [69] introduced a hybrid clustering architecture, in which several subsets of patterns can be processed together with an objective of finding a common structure. A detailed clustering algorithm is developed by integrating the advantages of both fuzzy sets and rough sets, and a measure of quantitative analysis of the experimental results is provided for synthetic and real-world data. Rough sets are used to model clusters in terms of upper and lower approximations, which are weighted by a pair of parameters while computing cluster prototypes. The use of rough sets help to control uncertainty among patterns in the boundary region, during collaboration between the modules. Memberships are used to enhance the robustness of clustering as well as collaboration. The main limitation of the proposed rough fuzzy c-means relies on the optimal selection of the parameters which can vary among different datasets.

In [62] the development of a generalized methodology, which integrates c-means algorithm, rough sets, and probabilistic and possibilistic memberships of fuzzy sets is presented. This formulation is geared toward maximizing the utility of both rough and fuzzy sets with respect to knowledge-discovery tasks. Several measures are defined based on rough sets to evaluate the performance of rough-fuzzy clustering algorithms. The effectiveness of the proposed algorithm is demonstrated, along with a comparison with other related algorithms, in the task of image segmentation. Also in this case, as in the previous technique, the main drawback is the optimal selection of the parameters.

In [110] the combined use of rough and fuzzy set theory to measure the ambiguities in images is proposed. Rough set theory is used to capture the indiscernibility among nearby gray values, whereas fuzzy set theory is used to capture the vagueness in the boundaries of the various regions. A measure called rough-fuzzy entropy of sets is proposed to quantify image ambiguity. By using this measure, a characteristic measure of an image called the average image ambiguity (AIA) is presented. The rough-fuzzy entropy measure is used to perform various image processing tasks such as object / background separation, multiple region segmentation and edge extraction. The performance are compared to those obtained using existing fuzzy and rough set theory based image ambiguity measures. Although the results are very promising, finding the optimal values of the input parameters can be a tricky task.

In [70] an application of rough-fuzzy clustering is presented for synthetic as well as CT scan images of the brain. The algorithm generates good prototypes even in the presence of outliers. The rough-fuzzy clustering simultaneously handles overlap of clusters and uncertainty involved in class boundary, hence yielding the best approximation of a given structure in unlabeled data. The number of clusters is automatically optimized in terms of various validity indexes. Comparison with other partitive algorithms is also presented. Experimental results demonstrate the effectiveness of the proposed method in CT scan images, and is validated by medical experts. The hybrid approach proposed in this paper aims to maximize the utility of both fuzzy and rough sets so to improve the performance of fuzzy c–means and rough c–means. Nevertheless the computation complexity is increased due to the simultaneous us of the two models.

In [63] a comprehensive investigation into rough set entropy based thresholding image segmentation techniques has been performed. Simultaneous combining entropy based thresholding with rough sets results in rough entropy thresholding algorithm. Standard RECA (Rough Entropy Clustering Algorithm) and Fuzzy RECA combined with rough entropy based partitioning routines have been proposed. Rough entropy clustering incorporates the notion of rough entropy into clustering model taking advantage of uncertainty in analyzed data. Based on the test reported in the article, Standard and Fuzzy RECA seem to have similar performance. This result could be deeper investigated because the fuzzy version, at least in principle, should better capture uncertainty in data and hence lead to a better segmentation.

In [148] an improved hybrid algorithm called rough-enhanced fuzzy c-means (REnFCM) algorithm is presented for segmentation of brain

MR images. The enhanced fuzzy c-means algorithm can speed up the segmentation process for gray-level image, especially for MR image segmentation. Experimental results indicate that the proposed algorithm is more robust to the noises and faster than many other segmentation algorithms, although at the cost of higher computation complexity.

Jiangping et al. [44] propose a fuzzy-rough approximation method for image segmentation. Based on the graph theory combined with the shortest path algorithm of watershed transformation, the paper presents a shortest path segmentation algorithm based on rough fuzzy grid, where to each fuzzy rough grid of the digital image is assigned a shortest path. The proposed method was applied in Traditional Chinese Medicine (TCM) tongue image segmentation experiment, where the algorithm has proved to avoid oversegmentation of the image.

In [45] a method to segment tongue image based on the theory of fuzzy rough sets is presented. The proposed method called Fuzzy Rough Clustering Based on Grid, extracts condensation points by means of fuzzy rough sets, and quarters the data space layer by layer. The algorithm has been used in tongue image segmentation of Traditional Chinese Medicine (TCM). Results indicate that the algorithm avoids oversegmentation of the image.

A multithresholding algorithm for color image segmentation is presented [75] using the concept of A-IFS histon obtained from Atanassov's Intuitionistic Fuzzy Set (A-IFS) representation of the image. A-IFS histon, an encrustation of the histogram, consists of the pixels that belong to the set of similar color pixels. In a rough set theoretic sense, A-IFS histon and the histogram can be correlated to upper and lower approximations, respectively. A multithresholding algorithm, using roughness index, is then employed to get optimum threshold values for color image segmentation. The qualitative and quantitative comparison of the proposed method against the histogram based and the conventional histon-based segmentations proves its superiority.

Performance of the proposed algorithm also proves that exploiting uncertainty and vagueness in color images can lead to good results when dealing with difficult concept like colors.

3.2.2 Edge Detection

Petrosino et al. [99] presented a multi-scale method based on the hybrid notion of rough fuzzy sets, coming from the combination of two models of uncertainty rough sets and fuzzy sets. Marrying both notions lead to consider, as instance, approximation of sets by means of similarity relations or fuzzy partitions. The most important features are extracted from the scale spaces by unsupervised cluster analysis, to successfully tackle image processing tasks. [100] describes a feed-forward layered ANN, whose operations are based on those of C-calculus [23], able to operate on a single image at a time. Within this framework C-calculus born as a method of representing fuzzy image subsets [7]. Its applications to shrinking, expanding and filtering [8] naturally lead to use it as a mathematical framework for designing a hierarchical neural network to deal with image analysis. The employed ANN is provided with a feedback mechanism and is structured in a hierarchical architecture. Application of the proposed network to edge detection is reported. The advantage of the proposed framework relies on the possibility of building a hierarchy of rough fuzzy sets, i.e., the possibility of exploiting uncertainty and vagueness at different resolutions.

In [131], image processing based on rough sets theory is discussed in detail. The paper presents a binary fuzzy rough set model based on triangle modulus, which describes binary relationship by upper approximation and lower approximation. Given an image described by binary relationship, the upper approximation and lower approximation can be used to represent the image. An edge detection algorithm by the upper approximation and the lower approximation of image is presented, and image denoising also is discussed. The proposed model is well fit for processing image that have gentle gray change.

3.2.3 Texture Segmentation

Traditional hybridization of fuzzy and rough sets can be found in [98], where a texture segmentation algorithm is proposed to solve the problem of unsupervised boundary localization in textured images using rough fuzzy sets and hierarchical clustering.

In [149], rough set theory is applied to multiple scale texture-shape recognition. Multiple-scale texture-shape recognition approach tries to cluster textures and shapes. In a multi resolution approach, texture and shape should be analyzed at different level. It becomes evident that an exact representation is not a feasible option (both practically and conceptually) therefore one needs to look at some viable approximation. This is realized by means of rough sets to construct the generalized approximate space by employing its fuzzy function and rough inclusion function. In this paper, according to the data set extracted from images, the fuzzy function and rough inclusion function of generalized approximate space is defined. Also statistical measures to denote the threshold of the fuzzy function and the importance degree of each extracted feature is used. A rough set based image texture recognition algorithm is proposed and compared with many other methods.

In [20], authors propose a rough content-based image quality measure. The image is partitioned into three parts: edges, textures and flat regions according to their gradient. In each part, the rough fuzzy integral is applied as the fuzzy measure of the similarity. The overall image quality metric is calculated based on the different importance of each part.

Based on FRM model, in [147] a rough neural network suitable for decision system modeling is proposed. It can implement smooth fuzzy partition of universe space by adaptive G-K (Gaustafason-Kessel) clustering algorithm, which overcomes the defect of traditional reduct calculation based on rough data analysis method. By making advantage of characteristics of fuzzy clusters obtained by adaptive G-K clustering algorithm, significant generalization ability enhancement is achieved. By making full use of learning ability of neural network, FRM_RNN_M improve the adaptability and achieve comprehensive soft decision-making ability. The experiment results of classifying Brodatz texture image indicate that FRM_RNN_M is superior to traditional Bayesian and learning vector quantization (LVQ) methods.

3.2.4 Image Classification

Mao et al. [64] proposed a fuzzy Hopfield net model based on rough-set reasoning for the classification of multispectral images. The main purpose is to embed a rough-set learning scheme into the fuzzy Hopfield network to construct a classification system called a rough-fuzzy Hopfield net (RFHN). The classification system is a paradigm for the implementation of fuzzy logic and rough systems in neural network architecture. Instead of all the information in the image being fed into the neural network, the upper- and lower-bound grey levels, captured from a training vector in a multispectral image, are fed into a rough-fuzzy neuron in the RFHN. Therefore, only 2/N pixels are selected as the training samples if an N-dimensional multispectral image was used.

Wang et al. [129] proposed a nearest neighbor classification algorithm based on fuzzy-rough set theory (FRNNC). First, they make every training sample fuzzy rough and use nearest neighbor algorithm to remove training sample points in class boundary or overlapping regions. Then mountain clustering method is used to select representative cluster center points, and finally Fuzzy-Rough Nearest neighbor algorithm (FRNN) is applied to classify the test data. The proposed method is applied to hand gesture image recognition and the results show that it is more effective and performs better than other nearest neighbor methods.

In [113] presents a combined approach of neural network classification systems with a fuzzy-rough set-based feature reduction method is presented. Unlike transformation-based dimensionality reduction techniques, this approach retains the underlying semantics of the selected feature subset. This is very important to help ensure that classification results are understandable by the user. Following this approach, the conventional multi-layer feedforward networks, which are sensitive to the dimensionality of feature patterns, can be expected to become effective on classification of images whose pattern representation may otherwise involve a large number of features. The proposed scheme has been applied to the real problem of normal and abnormal blood vessel image classification involving different cell types.

Authors in [13] present a study of the classification of large-scale Mars McMurdo panorama image. Three dimensionality reduction techniques, based on fuzzy-rough sets, information gain ranking, and principal component analysis respectively, are each applied to this complicated image data set to support learning effective classifiers. The work allows the induction of low-dimensional feature subsets from feature patterns of a much higher dimensionality. To facilitate comparative investigations, two types of image classifier are employed, namely multi-layer perceptrons and K-nearest neighbors. Experimental results demonstrate that feature selection helps to increase the classification efficiency by requiring considerably less features, while improving the classification accuracy by minimizing redundant and noisy features.

The focus of the study in [46] is an analysis of the effect of the granularity on indiscernability relation of objects. In this study, authors have applied the Rough Set Theory, to handle the imprecision due to granularity of the structure of the satellite image.Rough set and rough-fuzzy theory offer a better and transparent choice to have faster, comparable and effective results.

3.2.5 Detection

Han et al. [33] present a feature reduction method based on Rough-Fuzzy Set, by which the dissimilarity function for shot boundary detection is obtained. By calculating the correlation between conditional attributes, the importance of conditional attributes in the rough set can be obtained. Due to the ambiguity of the set of feature, the class precision of rough-fuzzy is given. Then, the importance of conditional attributes is defined as the rough-fuzzy operator by the product of the importance of conditional attributes in the rough set and the class precision of rough-fuzzy. According to the proportion of each feature, the top k features can be obtained. The dissimilarity function is generated by weighting these important features.

Human face detection plays an important role in application such as video surveillance, human computer interface, face recognition, and face image database management. In [151] an attribute reduction method based on fuzzy rough set is applied for face recognition. This paper mainly quotes attribute reduction of fuzzy rough sets to deal with the face data, while the recognition process uses neural network ensemble. The method avoids losing of information caused by rough set attribute reduction. A fuzzy similarity relation is used to replace an equivalence relation so that the dispersing of data is canceled. As a result, the recognition accuracy is improved.

In [101] authors presents a scheme for human faces detection in color images under unconstrained scene conditions, such as the presence of a complex background and uncontrolled illumination. The proposed method adopts a specialized unsupervised neural network, to extract skin colour regions in the Lab colour space, obtained from the integration of the rough fuzzy set based scale space trans- form and neural clustering. A correlation-based method is then applied for the detection of ellipse regions. Experiments on three benchmark face databases demonstrate the ability of the proposed algorithm in detecting faces also in difficult conditions.

4 Rough Fuzzy Product

Granular computing is based on the concept of information granule, that is a collection of similar objects which can be considered as indistinguishable. Partition of an universe into granules offers a coarse view of the universe where concepts, represented as subsets, can be approximated by means of granules. In this framework, rough set theory can be regarded to as a family of methodologies and techniques that make use of granules [90]. The focus of rough set theory is on the ambiguity caused by limited discernibility of objects in the domain of discourse. Granules are formed as objects and are drawn together by the limited discernibility among them. Granulation is of particular interest when a problem involves incomplete, uncertain or vague information. In such cases, precise solutions can be difficult to obtain and hence the use of techniques based on granules can lead to a simplification of the problem at hand.

At the same time, multivalued logic can be applied to handle uncertainty and vagueness present in information system, the most visible of which is the theory of fuzzy sets [142]. In this framework, uncertainty is modelled by means of functions that define the degree of belongingness of an object to a given concept. Hence membership functions of fuzzy sets enable efficient handling of overlapping classes.

The hybrid notion of rough fuzzy sets comes from the combination of these two models of uncertainty to exploit, at the same time, properties like coarseness, by handling rough sets [93], and vagueness, by handling fuzzy sets. In this combined framework, rough sets embody the idea of indiscernibility between objects in a set, while fuzzy sets model the ill-definition of the boundary of a subclass of this set.

Combining both notions leads to consider, as instance, approximation of sets by means of similarity relations or fuzzy partitions. The rough fuzzy synergy is hence adopted to better represent the uncertainty in granular computation.

Nevertheless, some considerations are in order. Classical rough set theory is defined over a given partition, although several equivalence relations, and hence partitions, can be defined over the universe of discourse. Different partitions correspond to a coarser or finer view of the universe, because of different information granules, thus leading to coarser or finer definition of the concept to be provided. Then a substantial interest arises about the possibility of exploiting different partitions and, possibly, rough sets of higher order. Some approaches have been presented to exploit hierarchical granulation [139] where various approximations are obtained with respect to different levels of granulation. Considered as a nested sequence of granulations by a nested sequence of equivalence relations, this procedure leads to a nested sequence of rough set approximations and to a more general approximation structure. Hierarchical representation of the knowledge is also used in [141] to build a sequence of finer reducts so to obtain multiple granularities at multiple layers. The hierarchical reduction can handle problem with coarser granularity at lower level so to avoid incompleteness of data present in finer granularity at deeper layer. A different approach is presented in [103] where authors report a Multi-Granulation model of Rough-Set (MGRS) as an extension of Pawlak's rough set model. Moreover, this new model is used to define the concept of approximation reduct as the smallest attribute subset that preserves the lower approximation and upper approximation of all decision classes in MGRS.

The hybridization of rough and fuzzy sets reported here has been observed to possess a viable and effective solution to some of the most difficult problems in image analysis. The model exhibits a certain advantage of having a new operator to compose rough fuzzy sets, called $\mathcal{RF}$-product, able to produce a sequence of composition of rough fuzzy sets in a hierarchical manner. Theoretical foundations and properties, together with an example of application for image compression are described in the following sections.

4.1 Rough fuzzy sets: a background

Let us start from the definition of a rough fuzzy set given by Dubois and Prade [23]. Let U be the universe of discourse, X a fuzzy subset of U, such that $\mu_X(u)$ represents the fuzzy membership function of X over U, and R an equivalence relation that induces the partition $U/R = \{Y_1, \dots, Y_p\}$ (from now on denoted as $\mathcal{Y}$) over U in p disjoint sets, i.e. $Y_i \bigcap Y_j = \emptyset\ \forall i, j = 1, \dots, p$ and $\bigcup_{i=1}^{p} Y_i = U$. The lower and upper approximation of X by R, i.e. $\underline{R}(X)$ and $\overline{R}(X)$ respectively, are fuzzy sets defined as

$$\mu_{\underline{R}(X)}(Y_i) = \inf\{\mu_X(u)|Y_i = [u]_R\} \tag{4.1}$$
$$\mu_{\overline{R}(X)}(Y_i) = \sup\{\mu_X(u)|Y_i = [u]_R\} \tag{4.2}$$

i.e. $[u]_R$ is a set such that (4.1) and (4.2) represent the degrees of membership of Y_i in $\underline{R}(X)$ and $\overline{R}(X)$, respectively. The couple of sets $< \underline{R}(X), \overline{R}(X) >$ is called *rough-fuzzy set* denoting a fuzzy concept (X) defined in a crisp approximation space (U/R) by means of two fuzzy sets ($\underline{R}(X)$ and $\overline{R}(X)$). Specifically, identifying $\pi_i(u)$ as the function that returns 1 if $u \in Y_i$ and 0 if $u \notin Y_i$, and considering $Y_i = [u]_R$ and $\pi_i(u) = 1$, the following relationships hold:

$$\mu_{\underline{R}(X)}(Y_i) = \inf_u max\{1 - \pi_i(u), \mu_X(u)\} \tag{4.3}$$
$$\mu_{\overline{R}(X)}(Y_i) = \sup_u min\{\pi_i(u), \mu_X(u)\} \tag{4.4}$$

To emphasize that the lower and upper approximations of the fuzzy subset X are, respectively, the infimum and the supremum of the membership functions of the elements of a class Y_i to the fuzzy set X, we can define a rough-fuzzy set as a triple

$$RF_X = (\mathcal{Y}, \mathcal{I}, \mathcal{S}) \tag{4.5}$$

where $\mathcal{Y} = \{Y_1, \ldots, Y_p\}$ is a partition of U in p disjoint subsets $Y_1, \ldots, Y_p$, and $\mathcal{I}, \mathcal{S}$ are mappings of kind $U \to [0,1]$ such that $\forall u \in U$,

$$\mathcal{I}(u) = \sum_{i=1}^{p} \underline{\nu_i} \times \mu_{Y_i}(u) \tag{4.6}$$

$$\mathcal{S}(u) = \sum_{i=1}^{p} \overline{\nu_i} \times \mu_{Y_i}(u) \tag{4.7}$$

where

$$\underline{\nu}_i = \inf\{\mu_X(u) | u \in Y_i\} \tag{4.8}$$

$$\overline{\nu}_i = \sup\{\mu_X(u) | u \in Y_i\} \tag{4.9}$$

for the given subsets $\mathcal{Y} = \{Y_1, \ldots, Y_p\}$ and for every choice of function $\mu : U \to [0,1]$. $\mathcal{Y}$ and μ uniquely define a rough-fuzzy set as stated below

Definition 6. *Given a subset $X \subseteq U$, if μ is the membership function μ_X defined on X and the partition $\mathcal{Y}$ is made with respect to an equivalence relation $\mathcal{R}$, i.e. $\mathcal{Y} = U/\mathcal{R}$, then X is a fuzzy set with two approximations $\overline{R}(X)$ and $\underline{R}(X)$, which are again fuzzy sets with membership functions defined as (4.8) and (4.9), i.e. $\underline{\nu}_i = \mu_{\underline{R}(X)}$ and $\overline{\nu}_i = \mu_{\overline{R}(X)}$. The pair of sets $< \overline{R}(X), \underline{R}(X) >$ is then a rough fuzzy set.*

Let us recall the generalized definition of rough set given in [109]. Expressions for the lower and upper approximations of a given set X are

$$\underline{R}(X) = \{(u, \mathcal{I}(u)) | u \in U\} \tag{4.10}$$

$$\overline{R}(X) = \{(u, \mathcal{S}(u)) | u \in U\} \tag{4.11}$$

$\mathcal{I}$ and $\mathcal{S}$ are defined as:

$$\mathcal{I}(u) = \sum_{i=1}^{p} \mu_{Y_i}(u) \times \inf_{\varphi \in U} \max(1 - \mu_{Y_i}(\varphi), \mu_X(\varphi)) \tag{4.12}$$

$$\mathcal{S}(u) = \sum_{i=1}^{p} \mu_{Y_i}(u) \times \sup_{\varphi \in U} \min(\mu_{Y_i}(\varphi), \mu_X(\varphi)) \tag{4.13}$$

where μ_{Y_i} is the membership degree of each element $u \in U$ to a granule $Y_i \in U/R$ and μ_X is the membership function associated with X.

If we rewrite (4.8) and (4.9) as

$$\underline{\nu}_i = \mu_{\underline{R}(X)}(Y_i) = \inf_{\varphi \in U} \max(1 - \mu_{Y_i}(\varphi), \mu_X(\varphi)) \tag{4.14}$$

$$\overline{\nu}_i = \mu_{\overline{R}(X)}(Y_i) = \sup_{\varphi \in U} \min(\mu_{Y_i}(\varphi), \mu_X(\varphi)) \tag{4.15}$$

and considering a Boolean equivalence relation R, we arrive at the same definition of rough fuzzy set as given in (4.3) and (4.4). Indeed, considering (4.14) and the equivalence relation R, $\mu_Y(\varphi)$ takes values in $\{0, 1\}$ hence the expression $1 - \mu_Y(\varphi)$ equals 0 if $\varphi \in Y$ or 1 if $\varphi \notin Y$. Furthermore the max operation returns 1 or $\mu_X(\varphi)$ depending on the fact that $\varphi \in Y$ or $\varphi \notin Y$. The operation inf then returns the infimum of such values, that is the minimum value of $\mu_X(\varphi)$. The same applies to (4.15).

4.2 Hierarchical refinement of Rough-fuzzy sets

Rough set theory allows to partition the given data into equivalence classes. Nevertheless, given a set U, it is possible to employ different equivalence relations and hence produce different data partitions. This leads to a choice of the partition that represents the data in the best manner. For example, let us consider N-dimensional patterns, with $N = 4$ as in Table 4.1

Table 4.1: Example of data

	A_1	A_2	A_3	A_4
u_1	a	a	b	c
u_2	b	c	c	c
u_3	a	b	c	a
u_4	c	b	a	b

Let $\mathcal{Y}_i$ be the partition obtained applying the equivalence relation R_{A_i} on the attribute A_i. We may get from Table 4.1 the following four partitions

$$\begin{aligned}
\mathcal{Y}^1 &= \{\{u_1, u_3\}, \{u_2\}, \{u_4\}\} \\
\mathcal{Y}^2 &= \{\{u_1\}, \{u_2\}, \{u_3, u_4\}\} \\
\mathcal{Y}^3 &= \{\{u_1\}, \{u_2, u_3\}, \{u_4\}\} \\
\mathcal{Y}^4 &= \{\{u_1, u_2\}, \{u_3\}, \{u_4\}\}
\end{aligned} \tag{4.16}$$

that, without any apriori knowledge, have potentially the same representation data power. To exploit all the possible partitions by means of simple operations, we propose to refine them in a hierarchical manner, so that partitions at each level of the hierarchy retain all the important informations contained into the partitions of the lower levels. The operation employed to perform the hierarchical refinement is called *Rough–Fuzzy product* ($\mathcal{RF}$-product) and is defined by:

Definition 7. *Let $RF^i = (\mathcal{Y}^i, \mathcal{I}^i, \mathcal{S}^i)$ and $RF^j = (\mathcal{Y}^j, \mathcal{I}^j, \mathcal{S}^j)$ be two rough fuzzy sets defined, respectively, over partitions $\mathcal{Y}^i = (Y_1^i, \ldots, Y_p^i)$ and $\mathcal{Y}^j = (Y_1^j, \ldots, Y_p^j)$ with $\mathcal{I}^i$ (resp. $\mathcal{I}^j$) and $\mathcal{S}^i$ (resp. $\mathcal{S}^j$) indicating the measures expressed in Eqs. (4.6) and (4.7). The $\mathcal{RF}$–product between two rough-fuzzy sets, denoted by $\otimes$, is defined as a new rough fuzzy set*

$$RF^{i,j} = RF^i \otimes RF^j = (\mathcal{Y}^{i,j}, \mathcal{I}^{i,j}, \mathcal{S}^{i,j})$$

where $\mathcal{Y}^{i,j} = (Y_1^{i,j}, \ldots, Y_{2p-1}^{i,j})$ *is a new partition whose equivalence classes are*

$$Y_k^{ij} = \begin{cases} \bigcup_{\substack{s=1\\q=h}}^{\substack{s=h\\q=1}} Y_q^i \cap Y_s^j & h = k, \quad k \leq p \\ \\ \bigcup_{\substack{s=h\\q=p}}^{\substack{s=p\\q=h}} Y_q^i \cap Y_s^j & h = k - p + 1, \quad k > p \end{cases} \tag{4.17}$$

and $\mathcal{I}^{i,j}$ *and* $\mathcal{S}^{i,j}$ *are*

$$\mathcal{I}^{i,j}(u) = \sum_{k=1}^{2p-1} \underline{\nu}_k^{i,j} \times \mu_k^{i,j}(u) \tag{4.18}$$

$$\mathcal{S}^{i,j}(u) = \sum_{k=1}^{2p-1} \overline{\nu}_k^{i,j} \times \mu_k^{i,j}(u) \tag{4.19}$$

and

$$\underline{\nu}_k^{ij} = \begin{cases} \sup\limits_{\substack{s=1,\dots,h\\ q=h,\dots,1}} \{\underline{\nu}_q^i, \underline{\nu}_s^i\} & h = k, \qquad k \le p \\ \\ \sup\limits_{\substack{s=h,\dots,p\\ q=p,\dots,h}} \{\underline{\nu}_q^i, \underline{\nu}_s^i\} & h = k - p + 1, \qquad k > p \end{cases} \tag{4.20}$$

$$\overline{\nu}_k^{ij} = \begin{cases} \inf\limits_{\substack{s=1,\dots,h\\ q=h,\dots,1}} \{\overline{\nu}_q^i, \overline{\nu}_s^i\} & h = k, \qquad k \le p \\ \\ \inf\limits_{\substack{s=h,\dots,p\\ q=p,\dots,h}} \{\overline{\nu}_q^i, \overline{\nu}_s^i\} & h = k - p + 1, \qquad k > p \end{cases} \tag{4.21}$$

Let us pick up the example shown in Table 4.1, and consider partitions $\mathcal{Y}^1$ and $\mathcal{Y}^2$ obtained from equivalence relations R_{A_1} and R_{A_2} defined on U by attributes A_1 and A_2, respectively. In terms of rough–fuzzy sets they are $RF^1 = (\mathcal{Y}^1, \mathcal{I}^1, \mathcal{S}^1)$ and $RF^2 = (\mathcal{Y}^2, \mathcal{I}^2, \mathcal{S}^2)$. Partitions $\mathcal{Y}^1$ and $\mathcal{Y}^2$ are defined as follows

$$\{u_4\} = Y_1^1 \qquad \{u_3, u_4\} = Y_1^2$$
$$\{u_2\} = Y_2^1 \qquad \{u_2\} = Y_2^2$$
$$\{u_1, u_3\} = Y_3^1 \qquad \{u_1\} = Y_3^2$$

The refined partition $\mathcal{Y}^{1,2}$ defined on U by both attributes, corresponds to the partition obtained by $\mathcal{RF}$-producting RF^1 and RF^2. The new partition $\mathcal{Y}^{1,2}$ is obtained by (in matrix notation)

–	–	$(Y_3^1 \cap Y_1^2)$	$(Y_2^1 \cap Y_1^2)$	$(Y_1^1 \cap Y_1^2)$
–	$(Y_3^1 \cap Y_2^2)$	$(Y_2^1 \cap Y_2^2)$	$(Y_1^1 \cap Y_2^2)$	–
$(Y_3^1 \cap Y_3^2)$	$(Y_2^1 \cap Y_3^2)$	$(Y_1^1 \cap Y_3^2)$	–	–

The final partition is obtained by joining sets by column as explained in Eq. 4.17

$$
\begin{aligned}
Y_1^{1,2} &= \{(Y_1^1 \cap Y_1^2)\} \\
Y_2^{1,2} &= \{(Y_2^1 \cap Y_1^2) \cup (Y_1^1 \cap Y_2^2)\} \\
Y_3^{1,2} &= \{(Y_3^1 \cap Y_1^2) \cup (Y_2^1 \cap Y_2^2) \cup (Y_1^1 \cap Y_3^2)\} \\
Y_4^{1,2} &= \{(Y_3^1 \cap Y_2^2) \cup (Y_2^1 \cap Y_3^2)\} \\
Y_5^{1,2} &= \{(Y_3^1 \cap Y_3^2)\}
\end{aligned}
$$

Hence

$$
\mathcal{Y}^{1,2} = \{Y_1^{1,2}, Y_2^{1,2}, Y_3^{1,2}, Y_4^{1,2}, Y_5^{1,2}\}
$$

and $\mathcal{I}$ and $\mathcal{S}$ of the new rough–fuzzy set, computed as in (4.18) and (4.19), are

$$
\begin{aligned}
\underline{\nu}_1^{1,2} &= \sup\{\inf\{\underline{\nu}_1^1, \underline{\nu}_1^2\}\} \\
\overline{\nu}_1^{1,2} &= \inf\{\sup\{\overline{\nu}_1^1, \overline{\nu}_1^2\}\} \\
\underline{\nu}_2^{1,2} &= \sup\{\inf\{\underline{\nu}_2^1, \underline{\nu}_1^2\}, \inf\{\underline{\nu}_1^1, \underline{\nu}_2^2\}\} \\
\overline{\nu}_2^{1,2} &= \inf\{\sup\{\overline{\nu}_2^1, \overline{\nu}_1^2\}, \sup\{\overline{\nu}_1^1, \overline{\nu}_2^2\}\} \\
\underline{\nu}_3^{1,2} &= \sup\{\inf\{\underline{\nu}_3^1, \underline{\nu}_1^2\}, \inf\{\underline{\nu}_2^1, \underline{\nu}_2^2\}, \inf\{\underline{\nu}_1^1, \underline{\nu}_3^2\}\} \\
\overline{\nu}_3^{1,2} &= \inf\{\sup\{\overline{\nu}_3^1, \overline{\nu}_1^2\}, \sup\{\overline{\nu}_2^1, \overline{\nu}_2^2\}, \sup\{\overline{\nu}_1^1, \overline{\nu}_3^2\}\} \\
\underline{\nu}_4^{1,2} &= \sup\{\inf\{\underline{\nu}_3^1, \underline{\nu}_2^2\}, \inf\{\underline{\nu}_2^1, \underline{\nu}_3^2\}\} \\
\overline{\nu}_4^{1,2} &= \inf\{\sup\{\overline{\nu}_3^1, \overline{\nu}_2^2\}, \sup\{\overline{\nu}_2^1, \overline{\nu}_3^2\}\} \\
\underline{\nu}_5^{1,2} &= \sup\{\inf\{\underline{\nu}_1^3, \underline{\nu}_2^3\}\} \\
\overline{\nu}_5^{1,2} &= \inf\{\sup\{\overline{\nu}_1^3, \overline{\nu}_2^3\}\}
\end{aligned}
\tag{4.22}
$$

The rough–fuzzy set obtained by $RF^1 \otimes RF^2$ is thus defined by

$$
RF^{1,2} = (\mathcal{Y}^{1,2}, \mathcal{I}^{1,2}, \mathcal{S}^{1,2})
$$

where

$$\mathcal{I}^{1,2}(u) = \sum_{i=1}^{5} \underline{\nu}_i^{1,2} \times \mu_{Y_i^{1,2}}(u)$$
$$\mathcal{S}^{1,2}(u) = \sum_{i=1}^{5} \overline{\nu}_i^{1,2} \times \mu_{Y_i^{1,2}}(u) \tag{4.23}$$

In case of partitions of different cardinalities, it is sufficient to add empty sets to have partitions of the same size.

4.3 Characterization of $\mathcal{RF}$–product

Let us recall that a partition $\mathcal{Y}$ of a finite set U is a collection $\{Y_1, Y_2, \ldots, Y_p\}$ of nonempty subsets (equivalence classes) such that

$$Y_i \cap Y_j = \emptyset \qquad \forall i, j = 1, \ldots, p \tag{4.24}$$
$$\bigcup_i Y_i = U \tag{4.25}$$

Hence, each partition defines an equivalence relation and, conversely, an equivalence relation defines a partition, such that the classes of the partition correspond to the equivalence classes of the relation.
Partitions are partially ordered by reverse refinement $\mathcal{Y}^i \subseteq \mathcal{Y}^j$. We say that $\mathcal{Y}^i$ is finer than $\mathcal{Y}^j$ if every equivalence class of $\mathcal{Y}^i$ is contained in some equivalence class of $\mathcal{Y}^j$, that is, for each equivalence class Y_h^j of $\mathcal{Y}^j$, there are equivalence classes $Y_1^i, \ldots, Y_p^i$ of $\mathcal{Y}^i$ such that $Y_h^j = Y_1^i, \ldots, Y_p^i$. If $E(\mathcal{Y}^i)$ is the equivalence relation defined by the partition $\mathcal{Y}^i$, then $\mathcal{Y}^i \subseteq \mathcal{Y}^j$ iff $\forall u, u' \in U, (u, u') \in E(\mathcal{Y}^i) \implies (u, u') \in E(\mathcal{Y}^j)$, that is, $E(\mathcal{Y}^i) \subseteq E(\mathcal{Y}^j)$.

The set $\Pi(U)$ of partitions of a set U forms a lattice under the partial order of reverse refinement. The minimum is the partition where an equivalence relation is a singleton, while the maximum is the partition composed by one single equivalence relation. The meet $\mathcal{Y}^i \wedge \mathcal{Y}^j \in \Pi(U)$ is the partition whose equivalence classes are given by $Y_k^i \cap Y_h^j \neq \emptyset$, where Y_k^i and Y_h^j are equivalence classes of $\mathcal{Y}^i$ and $\mathcal{Y}^j$, respectively. In terms of equivalence relations

$$R_{\mathcal{Y}^i \wedge \mathcal{Y}^j} = R_{\mathcal{Y}^i} \cap R_{\mathcal{Y}^j} \tag{4.26}$$

is the largest equivalence relation contained in both $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$. The join $\mathcal{Y}^i \vee \mathcal{Y}^j$ is a partition composed by the equivalence classes of the transitive closure of the union of the equivalence relations defined by $\mathcal{Y}^i$ and $\mathcal{Y}^j$. In terms of equivalence relations

$$\begin{aligned} R_{\mathcal{Y}^i \vee \mathcal{Y}^j} =& R_{\mathcal{Y}^i} \cup R_{\mathcal{Y}^i} \circ R_{\mathcal{Y}^j} \cup R_{\mathcal{Y}^i} \circ R_{\mathcal{Y}^j} \circ R_{\mathcal{Y}^i} \cup \ldots \\ & \cup R_{\mathcal{Y}^j} \cup R_{\mathcal{Y}^j} \circ R_{\mathcal{Y}^i} \cup R_{\mathcal{Y}^j} \circ R_{\mathcal{Y}^i} \circ R_{\mathcal{Y}^j} \cup \ldots \end{aligned} \tag{4.27}$$

where $R_x \circ R_y$ denotes the composition of the equivalence relations R_x and R_y and is the smallest equivalence relation containing both $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$. $\mathcal{I}$ and $\mathcal{S}$ are defined in (4.6) and (4.7). Firstly, we prove that the $\mathcal{RF}$-product yields an equivalence relation

Theorem 3. *Let $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$ be equivalence relations on a set U. Then $E = R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ is an equivalence relation on U.*
Proof. *E is an equivalence relation iff (1) $\forall E_i, E_j \in E, E_i \cap E_j = \emptyset$ and (2) $\cup E = U$.*

1. *Given that $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$ are equivalence relations, $\mathcal{Y}^i$ and $\mathcal{Y}^j \in \Pi(U)$. $\forall u \in U, \exists R \in \mathcal{Y}^i$ and $\exists T \in \mathcal{Y}^j$ such that $u \in R$ and $u \in T$. Then $u \in R \cap T$. If $u \in R \cap T$ then $u \notin P \cap Q, \forall P \in \mathcal{Y}^i (P \neq R)$ and $\forall Q \in \mathcal{Y}^i (Q \neq T)$. The union of the intersections in the $\mathcal{RF}$-product ensure that u belongs to a single equivalence class $E_i \in E$ and hence $\forall E_i E_j \in E, E_i \cap E_j = \emptyset$.*

2. *Given that $\forall u \in U, \exists E_i \in E$ such that $u \in E_i$. Then $\cup E_i = U$.*

The operation $\mathcal{RF}$-product is commutative:

Theorem 4. *Let $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$ be equivalence relations on a set U. Then $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j} = R_{\mathcal{Y}^j} \otimes R_{\mathcal{Y}^i}$.*
Proof. *The property can be easily proved by first noting that intersection is a commutative operation. The matrix representing the $\mathcal{RF}$–product is built row-wise in $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ (that is each row is the refinement of an equivalence class of $R_{\mathcal{Y}^j}$ by all equivalence classes of $R_{\mathcal{Y}^i}$), while it is built column-wise in $R_{\mathcal{Y}^j} \otimes R_{\mathcal{Y}^i}$ (that is each column is the refinement of an equivalence class of $R_{\mathcal{Y}^j}$ by all equivalence classes of $R_{\mathcal{Y}^i}$). In both cases the positions considered at the union step are the same, thus yielding the same result.*

Next we prove two theorems which bound the level of refinement of the partitions induced by the $\mathcal{RF}$-product.

Theorem 5. *Let $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$ be equivalence relations on a set U. It holds that $R_{\mathcal{Y}^i} \cap R_{\mathcal{Y}^j} \subseteq R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$.*
Proof. *From Eq. 4.17 it can be easily seen that each equivalence class of $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ is the union of some equivalence classes of $R_{\mathcal{Y}^i} \cap R_{\mathcal{Y}^j}$. Then each equivalence class of $R_{\mathcal{Y}^i} \cap R_{\mathcal{Y}^j}$ is contained in an equivalence class of $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$. Hence $R_{\mathcal{Y}^i} \cap R_{\mathcal{Y}^j} \subseteq R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$.*

Theorem 6. *Let $R_{\mathcal{Y}^i}$ and $R_{\mathcal{Y}^j}$ be equivalence relations on a set U. It holds that $(R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) \otimes (R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) = R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$.*
Proof. *From Theorem 3 $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ is an equivalence relation. Then $(R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) \cap (R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) = R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ and from Eq. 4.17 it derives that each equivalence relation of $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ is equal to only one equivalence relation of $(R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) \cap (R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j})$.*

Another interesting property of the $\mathcal{RF}$-product is that partition $E = R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ can be seen as the coarsest partition with respect to the sequence of operations

$$E = R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$$
$$E' = E \otimes R_{\mathcal{Y}^j} = (R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) \otimes R_{\mathcal{Y}^j} \subseteq E$$
$$E'' = E \otimes R_{\mathcal{Y}^i} = (R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}) \otimes R_{\mathcal{Y}^i} \subseteq E$$

In other words, at each iteration, the $\mathcal{RF}$-product produces a finer partition with respect to the initial partition. It is worth noting that, at each iteration

$$E = E' \otimes E'' \tag{4.28}$$

Viewed from another perspective, the $\mathcal{RF}$-product can be seen as a rule generation mechanism. Suppose that it is possible to assign a label to each equivalence class of a partition. Then $R_{\mathcal{Y}^i} \otimes R_{\mathcal{Y}^j}$ represents a partition whose equivalence classes are consistent with the labels of the operands. Consider the following partitions on a set U

$$\begin{aligned} \mathcal{Y}^1 &= \{Y^1_{low}, Y^1_{medium}, Y^1_{high}\} \\ \mathcal{Y}^2 &= \{Y^2_{low}, Y^2_{medium}, Y^2_{high}\} \end{aligned} \tag{4.29}$$

where $low = 1$ $medium = 2$ $high = 3$ and

$$\begin{aligned} \{u_4\} &= Y^1_{low} & \{u_3, u_4\} &= Y^2_{low} \\ \{u_2\} &= Y^1_{medium} & \{u_2\} &= Y^2_{medium} \\ \{u_1, u_3\} &= Y^1_{high} & \{u_1\} &= Y^2_{high} \end{aligned}$$

Applying $\mathcal{RF}$-product we get

$$\mathcal{Y}^{1,2} = \{Y^{1,2}_{low}, Y^{1,2}_{medium/low}, Y^{1,2}_{medium}, Y^{1,2}_{medium/high}, Y^{1,2}_{high}\} \tag{4.30}$$

where

$$
\begin{aligned}
&Y^{1,2}_{low} = \{(Y^1_{low} \cap Y^2_{low})\}\\
&Y^{1,2}_{medium/low} = \{(Y^1_{medium} \cap Y^2_{low}) \cup (Y^1_{low} \cap Y^2_{medium})\}\\
&Y^{1,2}_{medium} = \{(Y^1_{high} \cap Y^2_{low}) \cup (Y^1_{medium} \cap Y^2_{medium}) \cup (Y^1_{low} \cap Y^2_{high})\}\\
&Y^{1,2}_{medium/high} = \{(Y^1_{high} \cap Y^2_{medium}) \cup (Y^1_{medium} \cap Y^2_{high})\}\\
&Y^{1,2}_{high} = \{(Y^1_{high} \cap Y^2_{high})\}
\end{aligned}
\tag{4.31}
$$

Analyzing the new partition we note how the equivalence classes are consistent with the composition of the original ones, i.e.:

a) $u \in U$ belongs to “low” class in $Y^{1,2}$ if it belongs to “low” class in Y^1 and “low” class in Y^2;

b) $u \in U$ belongs to “medium/low” class in $Y^{1,2}$ if it belongs to “low” class in Y^1 and “medium” class in Y^2 or to “medium” class in Y^1 and “low” class in Y^2;

c) $u \in U$ belongs to “medium” class in $\mathcal{Y}^{1,2}$ if it belongs to “medium” class in Y^1 and Y^2 or to “high” class in Y^1 and to “low” class in Y^2 or to “high” class in Y^2 and to “low” class in Y^1;

d) $u \in U$ belongs to “medium/high” class in $Y^{1,2}$ if it belongs to “high” class in Y^1 and “medium” class in Y^2 or to “medium” class in Y^1 and “low” class in Y^2;

e) $u \in U$ belongs to “high” class in $Y^{1,2}$ if it belongs to “high” class in Y^1 and Y^2.

5 Rough Fuzzy Vector Quantization

Vector quantization has been extensively investigated to reduce transmission bit rate or storage space for speech and image signals and for image coding [30], [77]. In image coding, block coding methods involve the partition of the image into small rectangular blocks and the extraction of a set of features from an image block to be arranged in a single vector. Each vector is then compared to a set of standard vector prototypes in a codebook and the codeword index, identifying the best match to the input vector, is transmitted as the corresponding reproduction vector. The receiver reconstructs the image using the codewords corresponding to the indexes sent in place of the original vectors. The same is done if the aim of the vector quantization is to save storage space. Due to the independent processing of each block, coding loss can produce discontinuities in the image, since the reconstructed pixels of one block will, most likely, not match with the pixels of the next one; this problem is commonly named *blocking effect*. This phenomenon is especially apparent when very low numbers of coefficients are retained in the coding, i.e. in low bit rate systems. In some applications, like medical and remote sensing imaging, the visual annoyance due to the blocking effect becomes critical, since fundamental information may be lost.

There are two approaches in literature to tackle this kind of problem: (i) the design of block coding methods that capture information between blocks and(ii) the application of post-processing methods to reduce the introduced noise between blocks.We propose a method that belongs to the first class of approaches,taking inspiration from the consideration that border pixels, responsible for the blocking effect, could be seen as pixels with a certain degree of uncertainty to belong to a block of an other adjacent one. Within this framework, two problems are to be faced: the codebook design, based on a particular measure, and the choice of vectors to be quantized in order to retain as much as possible information, saving space but also image quality over the block borders, where uncertainty usually appears. The rise of several major seminal theories including fuzzy logic, rough sets, neural networks and their combination (the soft-computing paradigm in brief) allows to incorporate imprecision and incomplete information, and to model very complex systems, making them a useful tool in many scientific areas. In particular, the notions of rough fuzzy sets and learning may offer viable and effective solutions to some of the most difficult problems in image analysis [99]. The hybrid notion of rough fuzzy sets comes from the combination of two models of uncertainty like coarseness by handling rough sets [?] and vagueness by handling fuzzy sets [142]. Rough sets embody the idea of indiscernibility between objects in a set, while fuzzy sets model the ill-definition of the boundary of a sub class of this set. Marrying both notions leads to consider, as instance, approximation of sets by means of similarity relations or fuzzy partitions. The rough fuzzy synergy is hence adopted to better represent the uncertainty in block coding methods. The fuzzy logic is not new to image coding problems: some approaches based on fuzzy relation equation and fuzzy transforms have been recently reported in literature [35][97]. The synergy of fuzzy and rough sets we propose should, in principle, better tackle the problem. Specifically, in this section we discuss an original method [3] that exploits the uncertainty of data structure to deal with the above mentioned problems [47]. Feature extraction is based upon rough fuzzy sets and performed by partitioning each block in multiple rough fuzzy sets which are characterized by two approximation sets, containing inf and sup values over small portions within the block. The method is shown to efficiently encode images in terms of high peak signal to noise ratio(PSNR)values, while alleviating the blocking problem. A comparison with other fuzzy-based coding/decoding methods and with the DCT and JPEG methods is performed to show the performance of the proposed method with respect to the PSNR values. By using approximately the same compression rates, the proposed method performs very well for low compression rates and, in some cases, the achieved PSNR values result to be quite similar to those obtained by JPEG compression.

5.1 Feature Discovery

Let us consider an image I defined over a set $U = [0, ..., H-1] \times [0, ...W-1]$ of picture elements, i.e. $I : u \in U \to [0,1]$. Let us also consider a grid, superimposed on the image, whose cells Y_i are of dimension $w \times w$, such that all Y_i constitute a partition over I, i.e. eqs (4.24) and (4.25) hold and each Y_i^1, for $i = 1 \ldots p$, has dimension $w \times w$ and $p = H/w + W/w$. The size w of each equivalence class will be referred to as *scale*.

Each cell of the grid can be seen as an equivalence class induced by an equivalence relation $\mathcal{R}$ that assigns each pixel of the image to a single cell. Given a pixel u, whose coordinates are u_x and u_y, and a cell Y_i of the grid, whose coordinates of its upper left point are $x(Y_i)$ and $y(Y_i)$, u belongs to Y_i if $x(Y_i) \leq u_x \leq x(Y_i) + w - 1$ and $y(Y_i) \leq u_y \leq y(Y_i) + w - 1$. In other words, we are defining a partition U/R of the image induced by the relation $\mathcal{R}$, in which each cell represents an equivalence class $[u]_{\mathcal{R}}$. Also suppose that equivalence classes can be ordered in some way, for instance, from left to right.

Moreover, given a subset X of the image, not necessarily included or equal to any $[u]_{\mathcal{R}}$, we define the membership degree $\mu_X(u)$ of a pixel u to X as the normalized gray level value of the pixel.

If we consider different scales, the partitioning scheme yields many partitions of the same image and hence various approximations $\overline{R}(X)$ and $\underline{R}(X)$ of the subset X. For instance, other partitions can be obtained by a rigid translation of $\mathcal{Y}^1$ in the directions of 0^o, 45^o and

90^o of $w-1$ pixels, so that for each partition a pixel belongs to a shifted version of the same equivalence class Y_j^i.

If we consider four equivalence classes, Y_j^1 Y_j^2 Y_j^3 Y_j^4 as belonging to these four different partitions, then there exists a pixel u with coordinates u_x, u_y such that u belongs to the intersection of Y_j^1 Y_j^2 Y_j^3 Y_j^4. Hence each pixel can be seen as belonging to the equivalence class

$$Y_j^{1,2,3,4} = Y_j^1 \cap Y_j^2 \cap Y_j^3 \cap Y_j^4 \tag{5.1}$$

of the partition obtained by $\mathcal{RF}$-producting the four rough fuzzy set to which Y_j^i, with $i = 1, \ldots, 4$, belongs, i.e.

$$RF_X^{1,2,3,4} = RF_X^1 \otimes RF_X^2 \otimes RF_X^3 \otimes RF_X^4 \tag{5.2}$$

The $\mathcal{RF}$-product behaves as a filtering process according to which the image is filtered by a minimum operator over a window $w \times w$ producing $\mathcal{I}$ and by a maximum operator producing $\mathcal{S}$. Iterative application of this procedure consists in applying the same operator to both results $\mathcal{I}$ and $\mathcal{S}$ obtained at the previous iteration.

As instance, X defines the contour or uniform regions in the image. On the contrary, regions appear rather like fuzzy sets of grey levels and their comparison or combination generates more or less uniform partitions of the image. Rough fuzzy sets, as defined in (4.5), seem to capture these aspects together, trying to extract different kinds of knowledge in data.

This procedure can be efficiently applied to image coding/decoding, getting rise to the method *rough fuzzy vector quantization* (RFVQ)[**?**]. The image is firstly partitioned in non–overlapping k blocks X_h of dimension $m \times m$, such that $m \geq w$, that is $X = \{X_1, \ldots, X_k\}$ and $k = H/m + K/m$.

Considering each image block X_h, a pixel in the block can be characterized by two values that are the membership degrees to the lower and upper approximation of the set X_h. Hence, the feature extraction process provides two approximations $\underline{R}(X_h)$ and $\overline{R}(X_h)$ characterized by $\mathcal{I}$ and $\mathcal{S}$ as defined in (4.6) and (4.7) where

$$\begin{aligned} \underline{\nu}_i &= \mu_{\underline{R}(X_h)}(Y_i) = \inf\{\mu_{X_h}(u)|Y_i = [u]_R\} \\ \overline{\nu}_i &= \mu_{\overline{R}(X_h)}(Y_i) = \sup\{\mu_{X_h}(u)|Y_i = [u]_R\} \end{aligned} \tag{5.3}$$

and $[u]_R$ is the granule that defines the resolution at which we are observing the block X_h. For a generic pixel $u = (u_x, u_y)$ we can compute the coordinates of the upper left pixel of the four equivalence classes containing u, as shown in Fig. 5.2:

$$\begin{aligned} &u_x = x_1 + w - 1 \Rightarrow x_1 = u_x - w + 1 \\ &u_y = y_1 + w - 1 \Rightarrow y_1 = u_y - w + 1 \\ &u_x = x_2 \Rightarrow x_2 = u_x \\ &u_y = y_2 + w - 1 \Rightarrow y_2 = u_y - w + 1 \\ &u_x = x_3 + w - 1 \Rightarrow x_3 = u_x - w + 1 \\ &u_y = y_3 \Rightarrow y_3 = u_y \\ &u_x = x_4 \Rightarrow x_4 = u_x \\ &u_y = y_4 \Rightarrow y_4 = u_y \end{aligned}$$

where the four equivalence classes for pixel u are

$$\begin{aligned} &Y_j^1 = (x_1, y_1, \underline{\nu}_j^1, \overline{\nu}_j^1) \\ &Y_j^2 = (x_2, y_2, \underline{\nu}_j^2, \overline{\nu}_j^2) \\ &Y_j^3 = (x_3, y_3, \underline{\nu}_j^3, \overline{\nu}_j^3) \\ &Y_j^4 = (x_4, y_4, \underline{\nu}_j^4, \overline{\nu}_j^4) \end{aligned}$$

For instance, if we choose a granule of dimension $w = 2$ for a generic $j - th$ granule of the $i - th$ partition, equations in (6.3) become:

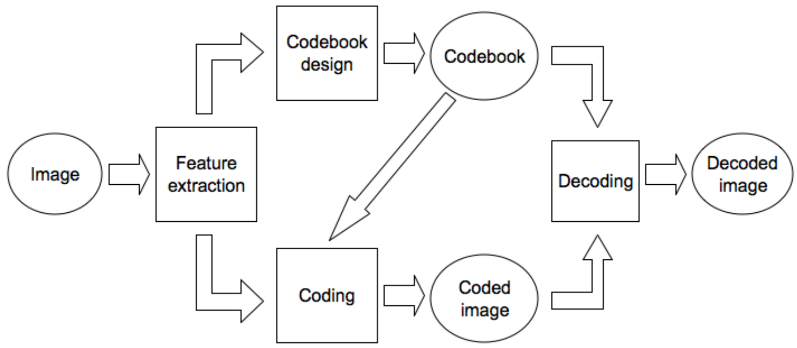


Figure 5.1: Schematic of the proposed method

$$\begin{aligned} \underline{\nu}_j^i &= \inf\{(u_x + a, u_y + b) | a, b = 0, 1\} \\ \overline{\nu}_j^i &= \sup\{(u_x + a, u_y + b) | a, b = 0, 1\} \end{aligned}$$

The compression method performed on each block X_h is composed of three phases: *codebook design*, *coding* and *decoding*. The entire process can be schematized as shown in Figure 5.1.

A vector is constructed by retaining the values $\underline{\nu}_j^i$ and $\overline{\nu}_j^i$ at positions u and $u + (w-1, w-1)$ in a generic block X_h, or equivalently $\underline{\nu}_j^1, \overline{\nu}_j^1, \underline{\nu}_j^3, \overline{\nu}_j^3$. The vector has hence dimension m^2 consisting of $m^2/2$ inf values and $m^2/2$ sup values. The vectors so constructed and extracted from a training image set are then fed to a quantizer in order to construct the codebook. The aim of vector quantization is the representation of a set of vectors $u \in X \subseteq R^{m^2}$ by a set of C prototypes (or codevectors) $V = \{v_1, v_2, ..., v_C\} \subseteq R^{m^2}$. Thus, vector quantization can also be seen as a mapping from an m^2-dimensional Euclidean space into the finite set $V \subseteq R^{m^2}$, also referred to as the codebook. Codebook design can be performed by clustering algorithms, but it is worth noting that the proposed method relies on the representation capabilities of the vector to be quantized and not on the quantization algorithm, to determine optimal codevectors, i.e. Fuzzy C-Means, Generalized Fuzzy C-Means or any analogous clustering algorithms can be adopted. Pseudocode of the codebook design procedure follows.

Algorithm 1 CODEBOOK DESIGN

Given N IMAGES of dimension $H \times W$
for $i = 1$ to n **do**
  for $h = 0$ to H -1 **do**
    for $k = 0$ to W - 1 **do**
      save INF and SUP pixel values from a block of dimension $H_B \times W_B$ located at coordinates h,k of IMAGES(i), by sliding a window of dimension 2×2.
      VECTOR $\leftarrow$ arrange a one-dimensional array using INF and SUP values $k \leftarrow k + H_B$
    end for$h \leftarrow h + W_B$
  end for
  save VECTOR
  $i \leftarrow i + 1$
end for
apply a quantization algorithm to the data read from IMAGES
save codebook for compression/decompression

5.1.1 Coding

The process of coding a new image proceeds as follows. For each block X_h the features extracted are arranged in a vector, following the same scheme used for designing the codebook, and compared with the codewords in the codebook to find the best match, i.e. the closest codeword to the block. In particular, for each block, inf and sup values are extracted from a window of size 2×2 shifted by one pixel into the block. All the extracted values are arranged in a one-dimensional array, i.e. for block dimension $m \times m$ and a window dimension 2×2 the array is represented by m^2 elements consisting of $m^2/2$ inf values and $m^2/2$ sup values. Doing so, the identificative number (out of C) of the winning codeword, i.e. the best match to the coded block, is saved in place of the generic block X_h. Pseudocode of the coding procedure follows.

Algorithm 2 CODING

```
Given the IMAGE of dimension H × W to be compressed
for h = 0 to H - 1 do
  for k = 0 to W - 1 do
    INF, SUP ← extract inf and sup values of pixels by sliding a window of dimen-
    sion 2 × 2 inside a block of dimension H_B × W_B
    VECTOR ← arrange a one-dimensional array using INF and SUP values
    CODEWORD_NUMBER ← find the closest codeword to VECTOR
    BLOCK(h,k) ← CODEWORD_NUMBER
  end for
end for
```

5.1.2 Decoding

Given a coded image, the decoding step firstly consists in the substitution of the identificative codeword number with the codeword itself, as reported in the codebook. The codeword consists of $m^2/2$ inf and $m^2/2$ sup values, instead of the original m^2 values of the block. To proceed to the original block construction, we apply the theory as follows. As stated above, each pixel can be seen as belonging to the block of the partition obtained by $\mathcal{RF}$-producting the four equivalence classes (6.1) and (6.2). Specifically, the blocks contained into the codeword are not the original ones, but those chosen to represent the block, i.e.

$$Y_j^{1,2,3,4} = Y_j^1 \cap Y_j^2 \cap Y_j^3 \cap Y_j^4 \tag{5.4}$$

where Y_j^i is a set of the partition of the rough-fuzzy set intersecting the generic block of the image X_h. The result of the $\mathcal{RF}$-product operation, with respect to a single block, is represented by another rough fuzzy set, characterized by lower and upper approximations. These values are used to fill the missing values into the decoded block.

In detail, being q_r the codeword corresponding to a generic block, the decoded block $X_{decoded}$ is constructed by filling the missing values, i.e. the original $\underline{\nu}_j^2, \overline{\nu}_j^2, \underline{\nu}_j^4, \overline{\nu}_j^4$ as combination of them, like average, median, etc., yielding $\tilde{\underline{\nu}}_j^2, \tilde{\overline{\nu}}_j^2, \tilde{\underline{\nu}}_j^4, \tilde{\overline{\nu}}_j^4$

The reconstructed block $X_{decoded}$, again a rough fuzzy set, is obtained by $\mathcal{RF}$-producting the four equivalence classes Y_j^1 Y_j^2 Y_j^3 Y_j^4, yielding the following

$$
\begin{aligned}
&Y_j^{1,2,3,4} = Y_j^1 \cap Y_j^2 \cap Y_j^3 \cap Y_j^4 \\
&\tilde{\mathcal{I}}^{1,2,3,4}(u) = \sum_j \tilde{\underline{\nu}}_j^{1,2,3,4} \times \mu_{Y_j}^{1,2,3,4}(u) \\
&\tilde{\mathcal{S}}^{1,2,3,4}(u) = \sum_j \tilde{\overline{\nu}}_j^{1,2,3,4} \times \mu_{Y_j}^{1,2,3,4}(u)
\end{aligned}
$$

where

$$
\tilde{\underline{\nu}}_j^{1,2,3,4} = \sup\{\underline{\nu}_j^1, \tilde{\underline{\nu}}^2{}_j, \underline{\nu}_j^3, \tilde{\underline{\nu}}_j^4\} \qquad \tilde{\overline{\nu}}_j^{1,2,3,4} = \inf\{\overline{\nu}_j^1, \tilde{\overline{\nu}}_j^2, \overline{\nu}_j^3, \tilde{\overline{\nu}}_j^4\} \tag{5.5}
$$

Lastly, under the assumption of local smoothness an estimate of the original grey values at the generic position u can be computed composing $\tilde{\underline{\nu}}_j^{1,2,3,4}$ and $\tilde{\overline{\nu}}_j^{1,2,3,4}$, as instance averaging or simply using only one of them. Pseudocode of the decoding procedure follows.

5.2 Related works

The section describes related works with the aim to highlight the differences with the proposed method and to give evidence of their functionalities with regards to the performance evaluation made and

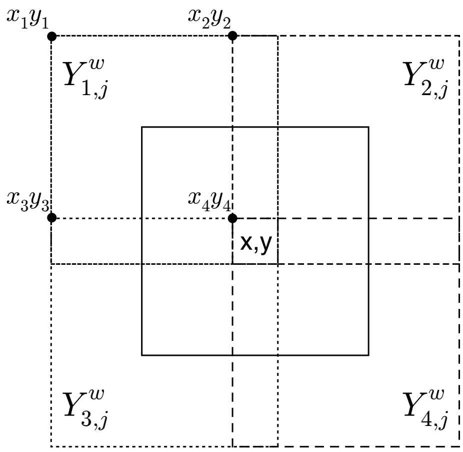


Figure 5.2: Equivalence classes coordinates.

Algorithm 3 DECODING

```
Given COMPRESSED_IMAGE and the CODEBOOK
Output decompressed IMAGE
for h = 0 to H - 1 do
   for k = 0 to W - 1 do
      CODE ← read from COMPRESSED_IMAGE the number of the codeword cor-
      responding to the block at position (h,k)
      INF, SUP ← CODEBOOK(CODE)
      fill the missing pixels by averaging INF and SUP values
   end for
end for
IMAGE ← apply the product operation
```

described in Section6. Specifically, we shall describe the approaches based on fuzzy relational equations [35] and fuzzy transforms [97].

5.2.1 Fuzzy relation equation

The main idea of the approach reported in [35] is based on the consideration that an $H \times W$ gray scale image, whose values have been normalized, could be seen as a fuzzy relation R over discrete interval $[0, ..., H-1]$ and $[0, ..., W-1]$:

$$R : (i, j) \in [0, ..., H-1] \times [0, ..., W-1] \rightarrow [0, 1]$$

such that $R(i, j) = P(i, j)/255$ where $P(i, j)$ is the gray level of pixel (i,j).
Once matrix R is created, it is divided in blocks of dimension $h_B \times w_B$ that are compressed in blocks of dimension $k_B \times l_B$, where $k_B \leq h_B$ and $l_B \leq w_B$. Compression is performed using a *fuzzy relation equations* system of the *max-t* type (Lukasiewicz t-norm). Recomposing compressed blocks leads to another fuzzy relation $G : (p, q) \rightarrow [0, 1]$ that represents the *compression* of the fuzzy relation R.
Decompression of each block to the original size is performed using a *fuzzy relation equations* system of the type $min - \rightarrow_t$, where $\rightarrow_t$ represents the Lukasiewicz "residuum" operator. Recomposing decompressed blocks leads to another fuzzy relation $D : (i, j) \in [0, ..., H-1] \times [0, ..., W-1] \rightarrow [0, 1]$ whose values can be shown to be very similar to the original ones.
The key elements for solving fuzzy equations system are the codebooks employed in compression and decompression. Codebooks are two matrices of dimension $h \times H$ and $w \times W$, respectively, in which each row is a fuzzy set characterized by Gaussian membership function:

$$A_{pi} = exp\left[-\alpha\left(p\frac{m}{k} - i\right)^2\right]$$

$$B_{qj} = exp\left[-\alpha\left(q\frac{n}{h} - j\right)^2\right]$$

Parameter α takes values in $[0.1, 0.2, ..., 1.0]$ and it is optimized to minimize RMSE for each decompressed block.
Let $R : (i, j) \in [0, ..., H-1] \times [0, ..., W-1] \rightarrow [0, 1]$ be a fuzzy relation, $A_1...A_h : [0, ..., H-1] \rightarrow [0, 1]$ and $B_1...B_w : [0, ..., W-1] \rightarrow [0, 1]$ fuzzy sets and $G : (p.q) \in [0, ..., H-1] \times [0, ..., W-1] \rightarrow [0, 1]$ a fuzzy relation, then

$$G_{pq} = \bigcup_{i=0}^{H-1} \bigcup_{j=0}^{W-1} [(A_{pi} t B_{qj}) t R_{ij}]$$

can be seen as a system of $h \times w$ fuzzy equations, where G is said compression of R with respect to codebooks A e B.
The solution of the same system in the unknown R, given by:

$$D_{ij} = \bigcap_{p=0}^{h-1} \bigcap_{q=0}^{w-1} [(A_{pi} t B_{qj}) \rightarrow {}_t G_{pq}]$$

is said decompression of R with respect to codebooks A e B. It can be shown that the fuzzy relation D is a solution of the previous system, such that $D \geq R$ for each solution R, i.e. $D_{ij} \geq R_{ij} \forall (i, j) \in [0, ..., H-1] \times [0, ..., W-1]$.
As stated above, this technique is applied to blocks R_B (submatrices) of dimension $H_B \times W_B$ that are compressed to blocks of dimension $h_B \times w_B$, where $h_B \leq H_B$ and $W_B \leq W_B$, so:

$$G_B(p, q) = \bigcup_{i=0}^{H_B-1} \bigcup_{j=0}^{W_B-1} [(A_{pi} t B_{qj}) t R_B(i, j)] : p = 0..h_B - 1, q = 0..w_B - 1$$

$$D_B(i, j) = \bigcap_{k=0}^{h_B-1} \bigcap_{h=0}^{w_B-1} [(A_{pi} t B_{qj}) \rightarrow {}_t G_B(p, q)] : i = 1..H_B - 1, j = 0..W_B$$

Recomposing single blocks leads to the fuzzy relation D.

5.2.2 Fuzzy transform

A fuzzy transform (F-transform) [97]is a function in one variable that associates a suitable n-dimensional vector to a continuous function

f over the interval $[a, b]$ by using the fuzzy sets $A_1, ..., A_n$ forming a fuzzy partition of [a,b]. An inverse F-transform is used to convert the n-dimensional output vector to a continuous function, which equals f within an arbitrary ϵ. The same process can be applied to functions in two variables continuous in th Cartesian product $[a, b] \times [c, d]$. An image, whose values has been normalized with respect to the length of the gray scale, can be considered as fuzzy matrix R. It is divided in blocks that are compressed using the discrete F-transform of the membership function f of R. These block are decompressed with the inverse discrete F-transform and recomposed to give a new image. The idea is to consider an $H \times W$ gray scale image, whose values has been normalized, as a fuzzy relation R over discrete intervals $[0, ..., H - 1]$ and $[0, ..., W - 1]$:

$$R : (i, j) \in [0, ..., H - 1] \times [0, ..., W - 1] \to [0, 1]$$

such that $R(i, j) = P(i, j)/255$ where $P(i, j)$ is the gray level of pixel (i,j).

Let the fuzzy sets $A_1, ..., A_h$ and $B_1, ..., B_w$, with $h < H$ and $w < W$, form a fuzzy partition of the real intervals $[0, H - 1]$ and [0,W-1]. Once a matrix R is created, it is divided in blocks (R_B) of dimension $H_B \times W_B$ that are compressed to blocks of dimension $h_B \times w_B$, where $h_B < H_B$ and $w_B < W_b$, using the discrete fuzzy transform

$$F_{kl}^B = \frac{\sum_{j=0}^{W_B-1} \sum_{i=0}^{H_B-1} R_b(i,j)A_k(i)B_l(j)}{\sum_{j=0}^{W_B-1} \sum_{i=0}^{H_B-1} A_k(i)B_l(j)}, \quad \forall k = 0...h_B - 1, l = 0...w_B.$$

Each compressed block is decompressed with the discrete inverse F-transform

$$R_{h_B w_B}^F(i, j) = \sum_{k=0}^{h_B-1} \sum_{l=0}^{w_B-1} F_{kl}^B A_k(i)B_l(j), \quad \text{defined in } \{0, ..., H_B - 1\} \times \{0, ..., W_B - 1\}.$$

5.3 Examples and Performance Analysis

To test the performances of the proposed coding scheme, the method has been applied to a set of test images of size 256×256 with 8

Compression	Uncompressed block	Compressed block
0.03	16 X 16	3 X 3
0.06	8 X 8	2 X 2
0.14	8 X 8	3 X 3
0.25	8 X 8	4 X 4

Table 5.1: Compression rates for fuzzy relation equations and fuzzy transforms methods.

Compression	Number of clusters	Block dimension
0.03	16	4 X 4
0.06	256	4 X 4
0.14	32	2 X 2
0.25	256	2 X 2
0.44	16384	2 X 2

Table 5.2: Compression rates for Composite Fuzzy Vector Quantization

bits/pixel and, among others, results for four test images: Bridge, Camera, Lena and House (Figs. 9,11,13,15) in terms of compression rates, MSE and PSNR values are reported. Tables 5.1 and 5.2 summarize compression rates adopted during the test.

The ISODATA quatizer has been trained on a set of images of size 256×256 with 8 bits/pixel: Baboon, Bird, Bridge, Building, Camera, City, Hat, House, Lena, Mona, Salesman (Fig. 8). The adopted sets were composed of 10 images, that is the image to be compressed was never considered in the construction of the codebook.

The results include also those achieved by the methods based on fuzzy relation equations and fuzzy transforms, as long as the DCT and JPEG methods [19].

For each images, PSNR values of the proposed method is compared with PSNR values achieved by other methods. In the following tables, FTR stands for fuzzy transforms and FEQ stands for fuzzy relation equations. As can be seen in Tables 5.3 - 5.5, RFVQ outperforms FTR, FEQ and DCT methods, while RFVQ's PSNR is lower than JPEG's one. For image "House" (see Table 5.6), a slightly

Compression	RFVQ	FTR	FEQ	DCT	JPEG
0.03	21.8205	20.7262	11.0283	18.6115	22.6985
0.06	23.7830	21.4833	14.2812	19.4849	24.7253
0.14	25.0959	23.2101	16.4632	20.8430	28.1149
0.25	26.2261	24.6975	19.7759	22.5470	31.2148
0.44	26.8410	27.0960	23.7349	26.1490	37.2367

Table 5.3: PSNR values for image Bridge

Compression	RFVQ	FTR	FEQ	DCT	JPEG
0.03	21.0179	20.6304	11.8273	18.1489	25.5207
0.06	23.8881	21.5427	15.4535	19.4447	28.4293
0.14	24.4189	23.5428	17.4869	22.1506	33.4379
0.25	26.0532	25.0676	20.5530	24.0288	38.8007
0.44	26.7560	27.4264	23.7706	25.5431	45.5878

Table 5.4: PSNR values for image Camera

lower PSNR than the one obtained by FTR may be observed. For image "Bridge" (see Table 5.3) the proposed method gains results quite close to JPEG when considering lower compression rate. Analyzing results shown in Tables refBridge-5.6, we can observe that the proposed method performs well for higher compression rates while it looses efficiency for lower compression rates. The reason for that resides in the quantization algorithm. Indeed, in order to obtain a compression rate of 0.44 a large number of clusters has to be computed (precisely 16384), but in this situation the quantity of centroids does not assure that the optimal choice will be done when selecting the most approximating codeword.

Compression	RFVQ	FTR	FEQ	DCT	JPEG
0.03	23.6305	23.5685	12.5959	18.1489	29.8727
0.06	26.0631	24.5514	17.1275	23.0445	32.4369
0.14	27.0874	26.8100	19.7528	24.8803	35.7345
0.25	28.5235	28.4310	23.2983	27.4874	37.5461
0.44	29.1465	30.8003	26.9285	29.7911	38.4881

Table 5.5: PSNR values for image Lena

Compression	RFVQ	FTR	FEQ	DCT	JPEG
0.03	22.3483	22.9525	11.8965	20.2155	30.0249
0.06	24.9479	23.8517	16.5426	21.2327	32.0180
0.14	25.9605	26.4038	19.9876	23.2612	34.2460
0.25	27.0839	28.1763	23.8031	26.5368	35.1001
0.44	27.6749	31.5114	28.7464	30.7693	35.7719

Table 5.6: PSNR values for image House

From the visual quality standpoint, we may observe a higher quality of RFVQ with respect to the images obtained by FEQ and FTR, while JPEG quality still remains not comparable. In particular FEQ produces images with a marked blocking effect (see Figures 5.4(d), 5.6(d), 5.8(d) and 5.10(d)) and, while FTR eliminates blocking effect, images appear too blurred with a consequent loss of details (see Figures 5.4(c), 5.6(c), 5.8(c) and 5.10(c)). The results of RFVQ are remarkable: it does not suffer a lot of the blocking effect, while loosing only a small amount of details (see Figures 5.4(b), 5.6(b), 5.8(b) and 5.10(b)).

To furthermore confirm these conclusion, Figs. 5.12 - 5.18 show images obtained applying RFVQ with various compression rates.

Figure 5.3: Bridge original

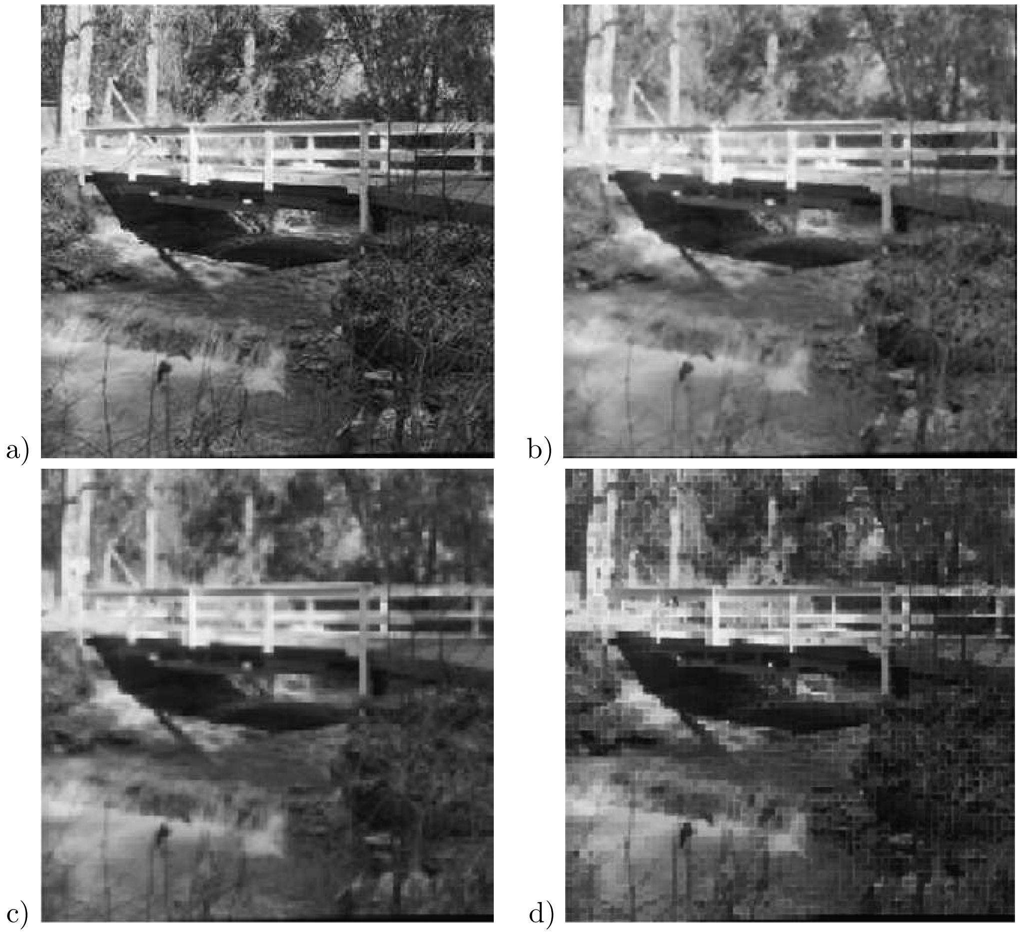


Figure 5.4: Bridge compression rate 0.25. a) JPEG, b) RFVQ, c) FTR, d) FEQ

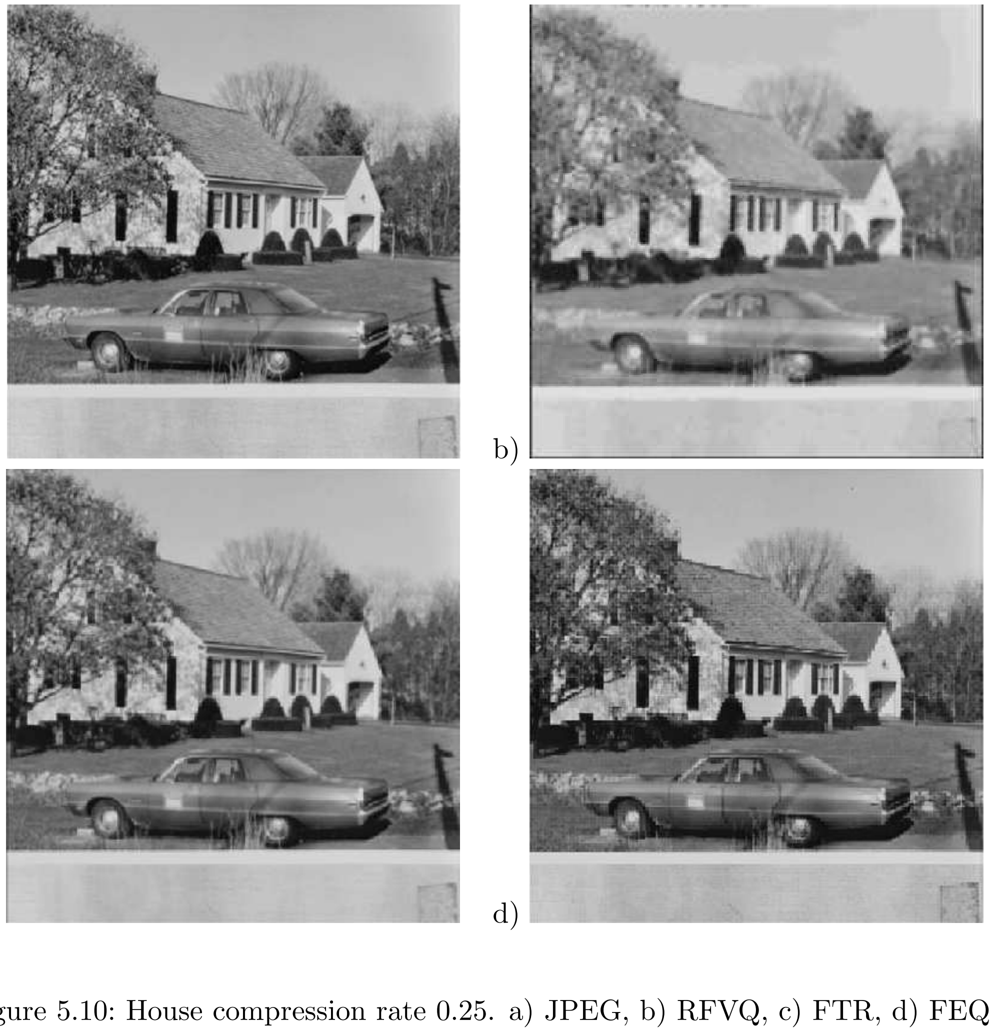
Figure 5.5: Camera original

Figure 5.6: Camera compression rate 0.25. a) JPEG, b) RFVQ, c) FTR, d) FEQ

Figure 5.7: Lena original

Figure 5.8: Lena compression rate 0.25. a) JPEG, b) RFVQ, c) FTR, d) FEQ

Figure 5.9: House original



a) b) c) d)

Figure 5.10: House compression rate 0.25. a) JPEG, b) RFVQ, c) FTR, d) FEQ

Figure 5.11: Bridge original

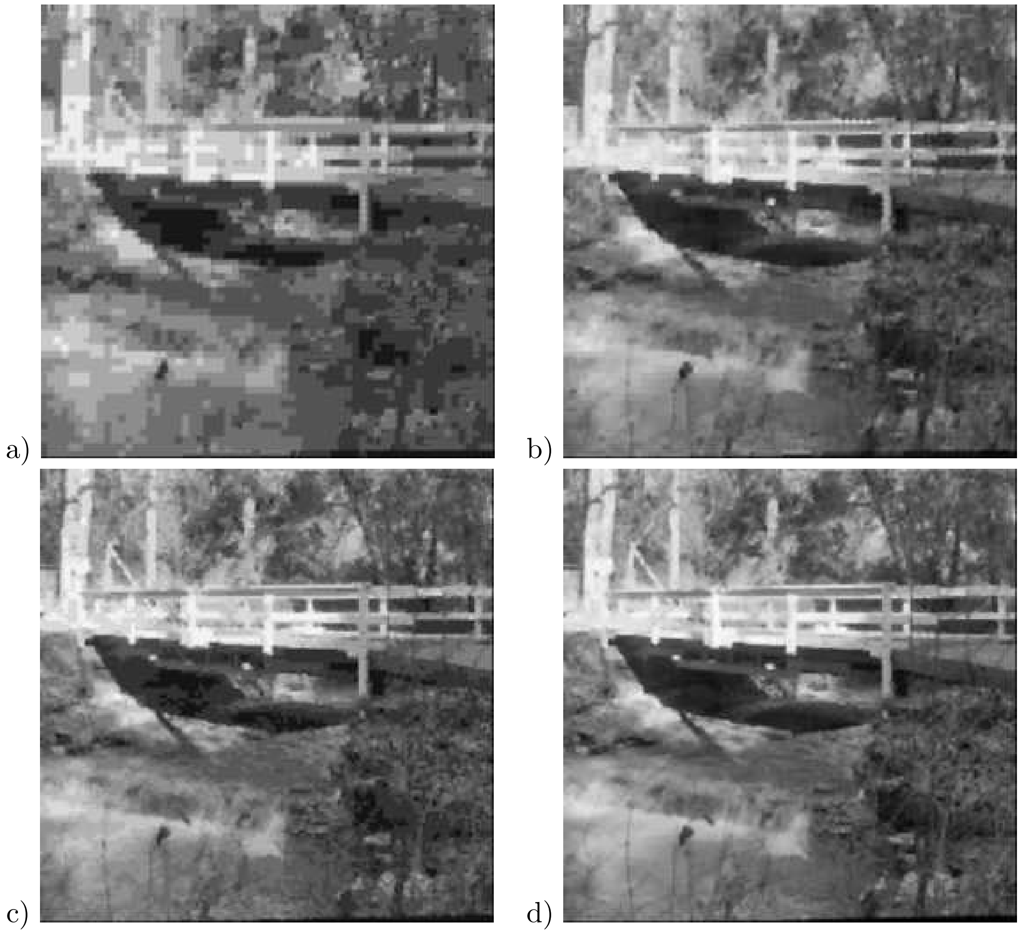
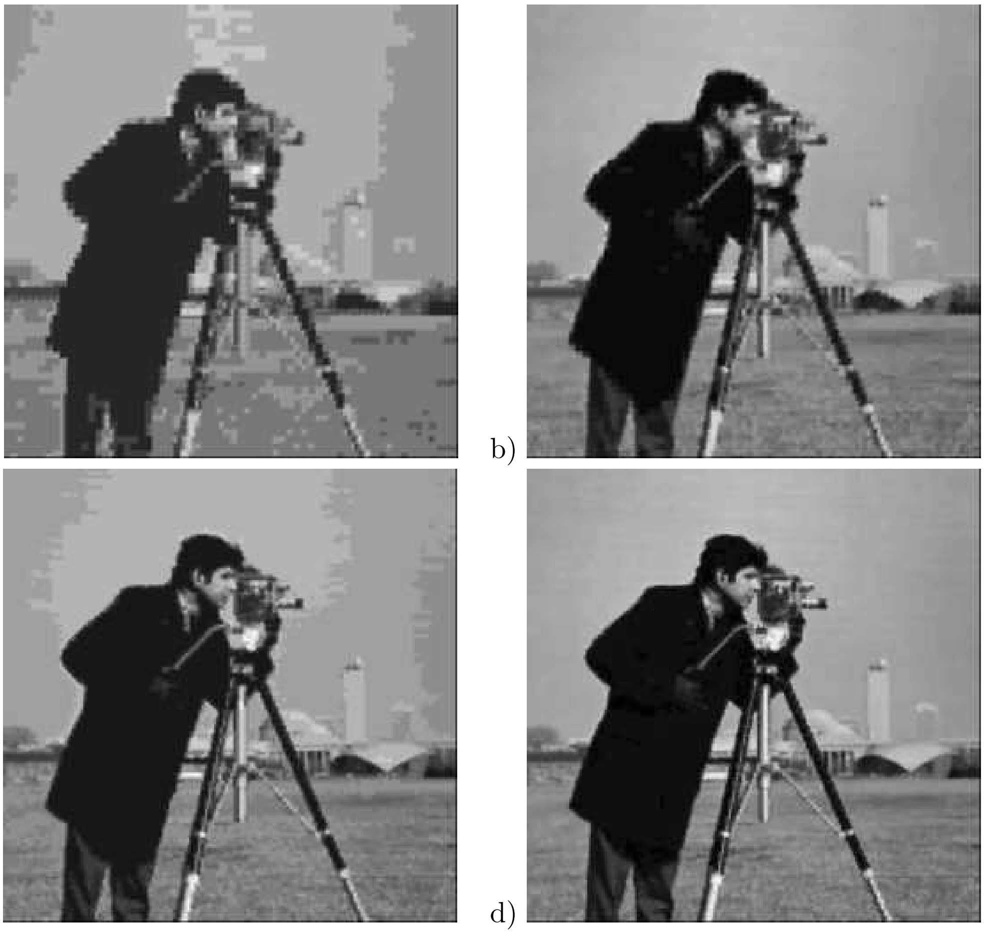


Figure 5.12: Bridge compression rate a) 0.03, b) 0.06, c) 0.14, d) 0.25

Figure 5.13: Camera original



a) b) c) d)

Figure 5.14: Camera compression rate a) 0.03, b) 0.06, c) 0.14, d) 0.25

Figure 5.15: Lena original

Figure 5.16: Lena compression rate a) 0.03, b) 0.06, c) 0.14, d) 0.25

Figure 5.17: House original

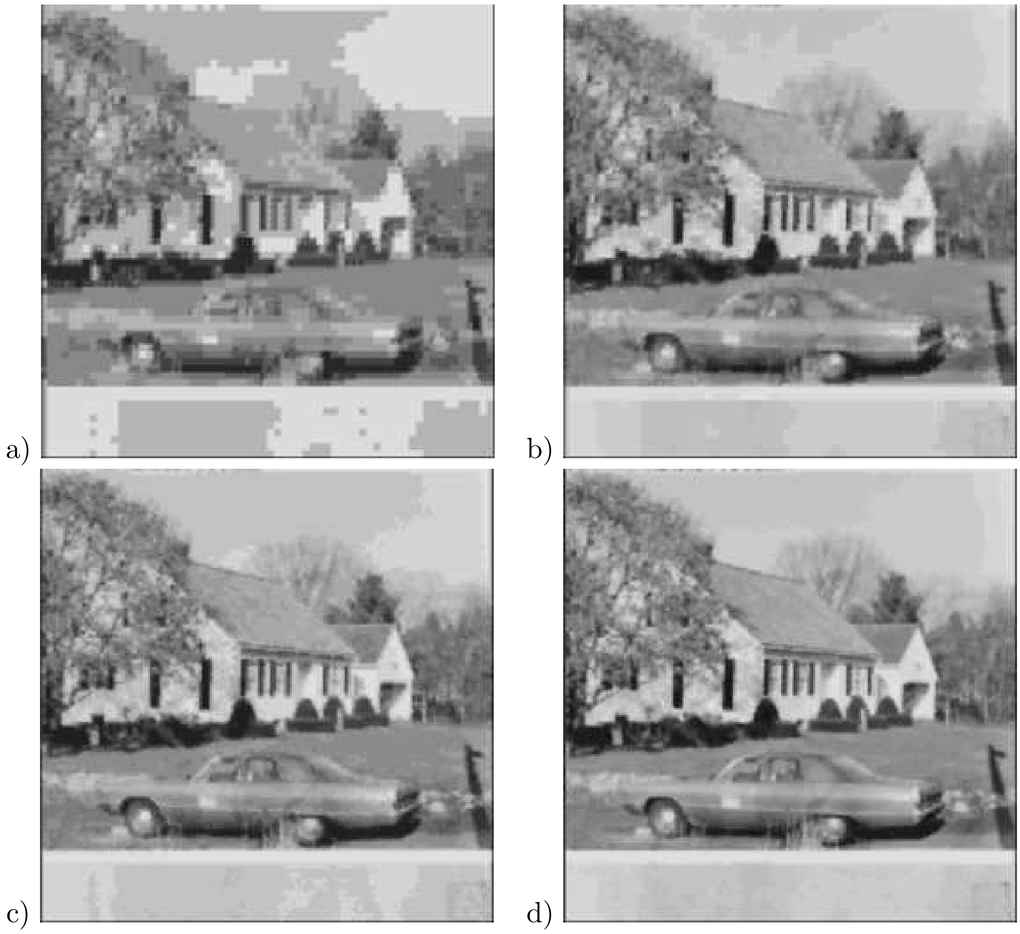


Figure 5.18: House compression rate a) 0.03, b) 0.06, c) 0.14, d) 0.25

6 Rough Fuzzy Color Image Segmentation

Image segmentation is one of the most challenging tasks in image processing, being the basic pre-processing step of many computer vision and pattern recognition problems [88]. Image segmentation consists in partitioning an image into different non-overlapping regions, where each region is characterized by homogeneity of gray levels, colors, texture or other criteria. From the computation point of view, color image segmentation is of particular interest because the huge amount of information held by colors can make the task very difficult to perform; although it can give fundamental information about the image to be analyzed. Different techniques have been reported over the years which can be classified [118] in: a) pixel based, which comprise histogram based algorithms [111] and clustering algorithms [127]; b) region based, which comprise region growing algorithms [124] and split and merge algorithms [89]; c) edge based, which comprise local [14] and global [53] algorithms to find region contours. Among them, one of the most used is represented by histogram based techniques because it needs no a–priori information about the image. The task consists in finding clusters corresponding to regions of uniform colors, identified by peaks in the histogram. Color images, being characterized by three dimensional scatterograms, make more difficult the search for peaks, either in the whole histogram or in each color channel independently. The major drawback of these methods is that they do not take into account the spatial correlation between adjacent pixels, while images usually show this property. Cheng et al. [16] employed a fuzzy homogeneity approach to extract homogeneous regions in a color image. The proposed method introduces the concept of homogram build considering intensity variation in pixels neighborhood. In [72] the concept of encrustation of the histogram (histon), which is a contour plotted on the top of each primary color histogram, is presented. In a rough-set theoretic sense, the histon represents the upper approximation of the color regions, that is a collection of pixels possibly belonging to the same region, while the histogram represents the lower approximation. An histogram-based technique is employed on the histon to obtain the final segmentation. Mushrif and Ray [76] presented a segmentation scheme, based on the concept of histon, which employs the roughness index. Roughness is large when the boundary contains a large number of elements, hence it will be small in the boundary between two objects and it will be large in region with uniform color. In this chapter a histogram based technique which exploits a generalized definition of rough–fuzzy sets and a particular operation called rough–fuzzy product in the HSV color space is presented [1].

6.1 HSV color space

Color models can be classified in two groups:

- Hardware-oriented which are designed considering device properties. Examples are RGB, CMY, YUV.
- User-oriented color models which are based on human perception. In these models, colors are represented by hue, saturation and value, where hue indicates the wavelength of the color, saturation measures the amount of white in a color, and value measures the color intensity. Examples are HLS, HSV and HSB.

The first user-oriented color model was proposed by Munsell, specifically designed for artists. Several approximations of the Munsell color model have been developed which separate luminance from hue and saturation. Among these models, one of the most popular is HSV (hue, saturation, value). The HSV model can be represented by a cone, where the cone axis represents the gray values while hue and saturation are represented, respectively, by the angle around the vertical axis and the distance from the central axis. Here we propose to employ the HSV color space because it allows to specify colors in a way similar to the human perception of colors. Moreover, the light intensity is explicit and separated from chromaticity, and this allows change detection invariant to modifications of illumination strength.

6.2 Rough Fuzzy Color Histogram

In this section we introduce the Rough Fuzzy Color Histogram that will be used to segment color images. Although the procedure is very similar to that employed in the RFVQ (presented in the previous chapter), here we repeat the the main ideas to highlight the necesary modifications needed to build the histogram.

Let us consider an image I defined over a set $U = [0, ..., H-1] \times [0, ...W-1]$ of picture elements, i.e. $I : u = (u_x, u_y) \in U \rightarrow [h(u), s(u), v(u)]$. Let us also consider a grid, superimposed on the image, whose cells Y_i are of dimension $w \times w$, such that all Y_i constitute a partition over I, i.e. eqs (4.24) and (4.25) hold and each Y_i^1, for $i = 1 \ldots p$, has dimension $w \times w$ and $p = H/w + W/w$. The size w of each equivalence class will be referred to as *scale*.

Each cell of the grid can be seen as an equivalence class induced by an equivalence relation $\mathcal{R}$ that assigns each pixel of the image to a single cell. Given a pixel u, whose coordinates are u_x and u_y, and a cell Y_i of the grid, whose coordinates of its upper left point are $x(Y_i)$ and $y(Y_i)$, u belongs to Y_i if $x(Y_i) \leq u_x \leq x(Y_i) + w - 1$ and $y(Y_i) \leq u_y \leq y(Y_i) + w - 1$. In other words, we are defining a partition U/R of the image induced by the relation $\mathcal{R}$, in which each cell represents an equivalence class $[u]_{\mathcal{R}}$. Also suppose that equivalence classes can be ordered in some way, for instance, from left to right.

Moreover, given a subset X of the image, not necessarily included or equal to any $[u]_{\mathcal{R}}$, we define the membership degree $\mu_X(u)$ of a pixel u to X as the normalized hue component of the pixel.

If we consider different scales, the partitioning scheme yields many partitions $\mathcal{Y}$ of the same image and hence various approximations $\overline{R}(X)$ and $\underline{R}(X)$ of the subset X. For instance, given a partition $\mathcal{Y}^i$, other partitions can be obtained by a rigid translation in the directions of 0^o, 45^o and 90^o of $w-1$ pixels, so that for each partition a pixel belongs to a shifted version of the same equivalence class Y^i_j.

If we consider four equivalence classes, Y^1_j, Y^2_j, Y^3_j and Y^4_j belonging to four partitions $\mathcal{Y}^1$ $\mathcal{Y}^2$ $\mathcal{Y}^3$ $\mathcal{Y}^4$, then there exists a pixel u with coordinates (u_x, u_y) such that u belongs to the intersection of Y^1_j, Y^2_j, Y^3_j and Y^4_j. Hence each pixel can be seen as belonging to the equivalence class

$$Y_j^{1,2,3,4} = Y_j^1 \cap Y_j^2 \cap Y_j^3 \cap Y_j^4 \tag{6.1}$$

of the partition obtained by $\mathcal{RF}$-producting the four rough fuzzy set to which Y^i_j, with $i = 1, \ldots, 4$, belongs, i.e.

$$RF_X^{1,2,3,4} = RF_X^1 \otimes RF_X^2 \otimes RF_X^3 \otimes RF_X^4 \tag{6.2}$$

The $\mathcal{RF}$-product behaves as a filtering process according to which the image is filtered by a minimum operator over a window $w \times w$ producing $\mathcal{I}$ (4.18) and by a maximum operator producing $\mathcal{S}$ (4.19). Iterative application of this procedure consists in applying the same operator to both results $\mathcal{I}$ and $\mathcal{S}$ obtained at the previous iteration.

As instance, X defines the contour or uniform regions in the image. On the contrary, regions appear rather like fuzzy sets of hue values and their comparison or combination generates more or less uniform partitions of the image. Rough fuzzy sets, as defined in (4.5), seem to capture these aspects together, trying to extract different kinds of knowledge in data.

The image is firstly partitioned in non–overlapping k blocks X_h of dimension $m \times m$, such that $m \geq w$, that is $X = \{X_1, \ldots, X_k\}$ and

$k = H/m + K/m$. Considering each image block X_h, a pixel in the block can be characterized by two values that are the membership degrees to the lower and upper approximation of the set X_h. Hence, the feature extraction process provides two approximations $\underline{R}(X_h)$ and $\overline{R}(X_h)$ characterized by $\mathcal{I}$ and $\mathcal{S}$ as defined in (4.6) and (4.7) where

$$\underline{\nu}_i = \mu_{\underline{R}(X_h)}(Y_i) = \inf\{\mu_{X_h}(u)|Y_i = [u]_R\}$$
$$\overline{\nu}_i = \mu_{\overline{R}(X_h)}(Y_i) = \sup\{\mu_{X_h}(u)|Y_i = [u]_R\} \quad (6.3)$$

and $[u]_R$ is the granule that defines the resolution at which we are observing the block X_h. For a generic pixel $u = (u_x, u_y)$ we can compute the coordinates of the upper left pixel of the four equivalence classes containing u, as shown in Fig. 5.2:

$$\begin{aligned}
u_x &= x_1 + w - 1 \Rightarrow x_1 = u_x - w + 1\\
u_y &= y_1 + w - 1 \Rightarrow y_1 = u_y - w + 1\\
u_x &= x_2 \Rightarrow x_2 = u_x\\
u_y &= y_2 + w - 1 \Rightarrow y_2 = u_y - w + 1\\
u_x &= x_3 + w - 1 \Rightarrow x_3 = u_x - w + 1\\
u_y &= y_3 \Rightarrow y_3 = u_y\\
u_x &= x_4 \Rightarrow x_4 = u_x\\
u_y &= y_4 \Rightarrow y_4 = u_y
\end{aligned}$$

where the four equivalence classes for pixel u are

$$\begin{aligned}
Y_j^1 &= (x_1, y_1, \underline{\nu}_j^1, \overline{\nu}_j^1)\\
Y_j^2 &= (x_2, y_2, \underline{\nu}_j^2, \overline{\nu}_j^2)\\
Y_j^3 &= (x_3, y_3, \underline{\nu}_j^3, \overline{\nu}_j^3)\\
Y_j^4 &= (x_4, y_4, \underline{\nu}_j^4, \overline{\nu}_j^4)
\end{aligned}$$

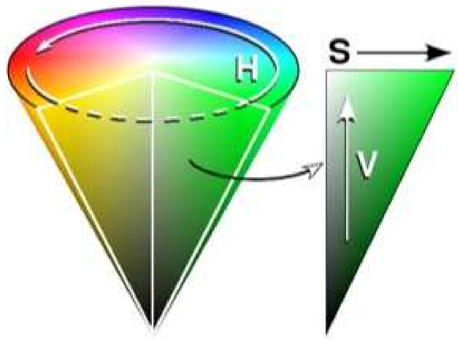


Figure 6.1: A wedge of the HSV color space.

For instance, if we choose a granule of dimension $w = 2$ for a generic $j - th$ granule of the $i - th$ partition, equations in (6.3) become:

$$\begin{aligned} \underline{\nu}_j^i &= \inf\{(u_x + a, u_y + b) | a, b = 0, 1\} \\ \overline{\nu}_j^i &= \sup\{(u_x + a, u_y + b) | a, b = 0, 1\} \end{aligned}$$

Let us now consider the HSV color space represented by a cone and a segment $[q, q + qt - 1]$ on the maximum circumference, where $0 \leq q \leq 359$ and $[qt_{\min} \leq qt \leq qt_{\max}]$ is the wedge dimension. This interval contains a certain amount of colors. In particular, if we imagine to cut the HSV cone in wedges, each one contains all the possible combination of saturation and value given a portion of hue (fig. 6.1).

Our goal is to describe each wedge using the blocks of the image, under the assumption that blocks with similar colors will fall in the same wedge. Each block, of dimension $w \times w$, characterized by a minimum ($h_m = \mathcal{I}$) and a maximun ($h_M = \mathcal{S}$) hue value, can be (fig. 6.2) : i) totally contained into a wedge of dimension qt
(i.e. $q \leq h_m \leq h_M < qt + q$), ii) partially contained into a wedge
(i.e. $q \leq h_m$ or $h_M < qt + q$), iii) not contained at all.

Hence, we can consider the set of blocks whose h_m and h_M are

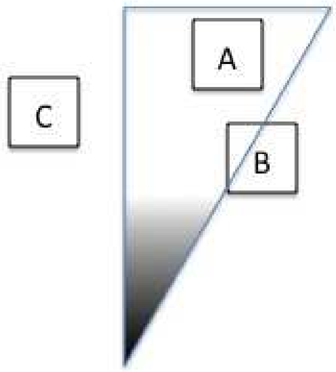


Figure 6.2: Block contained into the wedge (A), block partially contained into the wedge (B), block not contained into the wedge (C).

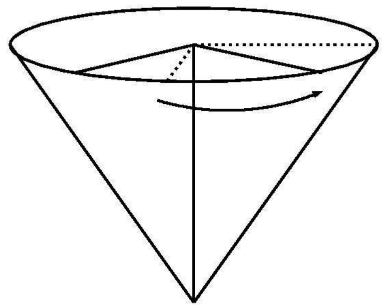
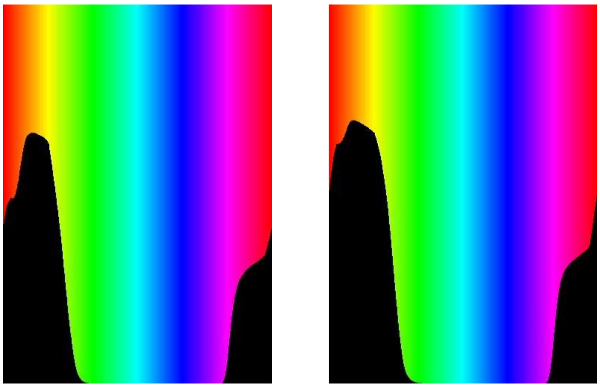


Figure 6.3: Shifting wedge over the hue circle.

contained into the wedge as the lower approximation of the object represented by the wedge itself, while the set of blocks whose h_m and h_M are partially contained into the wedge as its upper approximation. Clearly, it is very unlikely that all the colors into the wedge are represented into the image, i.e., the lower and upper approximation are the rough–fuzzy representation, induced by the partition of the image and hence by the equivalence relation, of an implicit object contained into the wedge.

Now consider a wedge of dimension $[q_i, q_i + qt - 1]$, $i = 0, \ldots, 359$ moving on the hue circle towards increasing hue values, starting from $q_1 = 0$. At each step the wedge is shifted of an offset x, i.e. $q_{i+1} = q_i +$ x, and the cardinality of the lower and upper approximation of the wedge is computed (with respect to the number of blocks). This pro-

Figure 6.4: Example image.



(a) (b)

Figure 6.5: a) Lower histogram and b) Upper histogram.

cedure yields the Rough Fuzzy Color Histogram, composed by the *Lower Histogram* ($\underline{H}$) and the *Upper Histogram* ($\overline{H}$) of the image. Repeating the same procedure for each wedge dimension $qt_{\min} \leq qt \leq qt_{\max}$, many Rough Fuzzy Color Histrograms are produced according to the possible values of qt. Figure 6.4 and 6.5 depict the lower and upper histograms of Figure 6.4.

It should be remainded that, if for a given pixel the saturation equals 0, the hue component is undefined and the pixel is characterized only by the value component, i.e. only by its gray level intensity. To overcome this problem, it is possible to exclude all the pixels with a saturation value lower than a given threshold ϵ and segment them

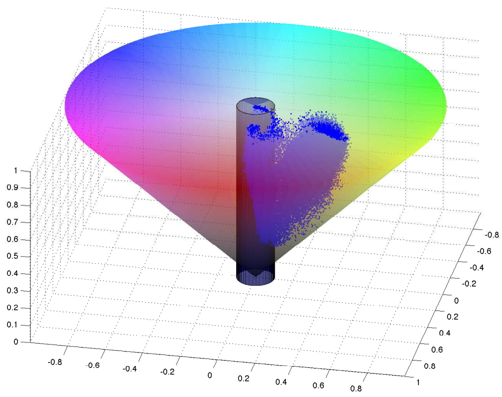


Figure 6.6: Distribution of the pixels in the HSV cone and saturation threshold $\epsilon = 0.2$.

separately (for instance employing a segmentation algorithm for gray scale images). Figure 6.6 shows the distribution of the pixels and a cylinder centered on the axis of the HSV cone, used to cut off the pixels characterized by saturation values ≤ 0.2.

6.3 Image segmentation by Rough Fuzzy Color Histogram

The goal of the proposed method is to find regions characterized by uniform colors. The segmentation of a color image is performed in the HSV color space by choosing the wedges that are better represented in a rough–fuzzy sense. The choice is guided by the rough accuracy of the wedge, i.e. each wedge s_i has an accuracy computed by means of the corresponding bin in the lower and upper histogram

$$\alpha_i = \frac{\underline{H}(i)}{\overline{H}(i)} \tag{6.4}$$

The wedge with highest accuracy is the wedge that is better represented with respect to the number of blocks belonging to lower and upper approximations of the wedge. Clearly, this can not be the only discriminant index to obtain a good segmentation. First of all because the accuracy, as computed in eq. 6.4, does not take into account the number of blocks, and hence the number of pixels contained into the wedge, but only their ratio. For instance, consider two wedges s_i and s_j and their lower and upper bins $\underline{H}(i) = 10$, $\overline{H}(i) = 20$, $\underline{H}(j) = 20$ and $\overline{H}(i) = 40$, respectively. The accuracies are

$$\alpha_i = \frac{\underline{H}(i)}{\overline{H}(i)} = 0.5$$
$$\alpha_j = \frac{\underline{H}(j)}{\overline{H}(j)} = 0.5 \tag{6.5}$$

Hence it is not possible to discriminate between wedges s_i and s_j. Moreover using only the accuracy does not take into account saturation and value of each pixel. To overcome both problems, it is also necessary to consider the number of pixels belonging to the wedge and their saturations and values. The first problem is tackled by weighting the accuracy of each wedge by the fraction of pixels whose hue value belongs to the wedge, i.e.

$$\tilde{\gamma} = \frac{N_{wedge}(s_i)}{N_{tot}(I)} \tag{6.6}$$

where $N_{wedge}(s_i)$ represents the number of pixels whose hue value belongs to the wedge and $N_{tot}(I)$ represents the number of pixels of the image I. More precisely, we use the following index

$$\gamma = 1 - \tilde{\gamma} \tag{6.7}$$

to limit the uncontrolled growth of the wedge, which otherwise would tend to include many pixels with very different hue values. Provided that regions of uniform colors are searched into the image, we need an index to measure the color uniformity of the pixels belonging to the the wedge and then use this index to weight the accuracy. To this aim we propose to employ a measure of the dispersion of the pixels falling into a wedge with respect to saturation and value. The main idea is that for a given hue q different shades are determined by saturation and value of each pixel. Hence a region characterized by uniform color will present a narrow scatter, while a region characterized by non uniform colors will have a sparse scatter. To compute the compactness of saturation and value into a wedge s_i, we propose the following index

$$\tilde{\delta} = \frac{1}{N_{wedge}(s_i)} \times \sqrt{\sum_{x \in s_i} (x - \mu_{s_i})^T (x - \mu_{s_i})} \tag{6.8}$$

where $x = [x_{saturation}, x_{value}]$. This index can be considered as the weighted squared root of the track of the covariance matrix. Also in this case we use the form

$$\delta = 1 - \tilde{\delta} \tag{6.9}$$

which yields higher values whenever the dispersion index is low (i.e. whenever the color is uniform). The final index, τ, is computed by composing α, γ and δ indeces (eqs. 6.4, 6.7 and 6.9)

$$\tau = \alpha \times (w_1 \times \gamma + w_2 \times \delta) \tag{6.10}$$

where w_1 and w_2, with $w_1 + w_2 = 1$, are parameters used to weight the fraction of pixels falling into a wedge and the saturation–value dispersion, respectively. A higher value for w_1 will lead to wedges comprising few pixels characterized by a low saturation–value dispersion, whilst a higher value for w_2 will produce wider wedges, with a larger number of pixels presenting a lower saturation–value dispersion. The index τ, computed for all the wedges, is used to segment the image. Firstly, the wedge with the highest τ value is selected as the region which is better represented into the image. Next all the wedges that intersect the first one are removed to avoid overlapping regions. For instance, consider s_i the wedge with the highest τ value corresponding to the hue segment $q_{s_i}, q_{s_i} + qt - 1$, then all the wedges s_j such that $q_{s_i} \leq q_{s_j} + \tilde{qt} - 1 < q_{s_i} + qt - 1$, with $\tilde{qt}$ varying in $[qt_{\min}, qt_{\max}]$, are removed. Next the wedge with the highest τ value, among those not removed in the previous step, is selected, and so on until no more wedges are left.

6.4 Experimental results

Quantifying the performance of a segmentation algorithm is a challenging task. Since image segmentation is an ill-defined problem, no single ground truth segmentation is available against which the output of an algorithm may be compared. Rather the comparison is to be made against the set of all possible perceptually consistent interpretations of the image, of which only a small fraction is usually available. Over the years different approaches have been proposed to evaluate the segmentation quality, as instance:

1. The Variation of Information (VoI) metric [66] defines the distance between two segmentations as the average conditional entropy of a segmentation with respect to the other, and thus roughly measures the amount of randomness in one segmentation which cannot be explained by the other.

2. The Global Consistency Error (GCE) [65] measures the extent to which one segmentation can be viewed as a refinement of the other. Segmentations related in this manner are considered to be consistent, since they could represent the same natural image segmented at different scales.

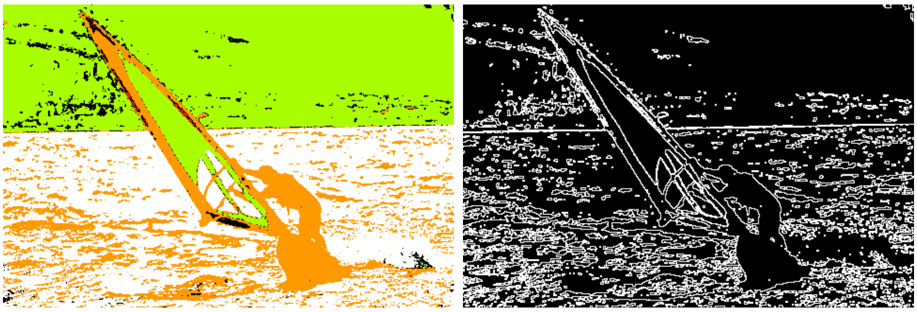
3. The Boundary Displacement Error (BDE) [27] measures the average displacement error of boundary pixels between two segmented images. Particularly, it defines the error of one boundary pixel as the distance between the pixel and the closest pixel in the other boundary image.

Here we employ the Probabilistic Rand Index (PRI) [128] that counts the fraction of pairs of pixels whose labellings are consistent between the computed segmentation and the ground truth, averaging across multiple ground truth segmentations to account for scale variation in human perception. For each image, the quality of the segmentation is evaluated by comparing it with all the available segmentations of the same image.

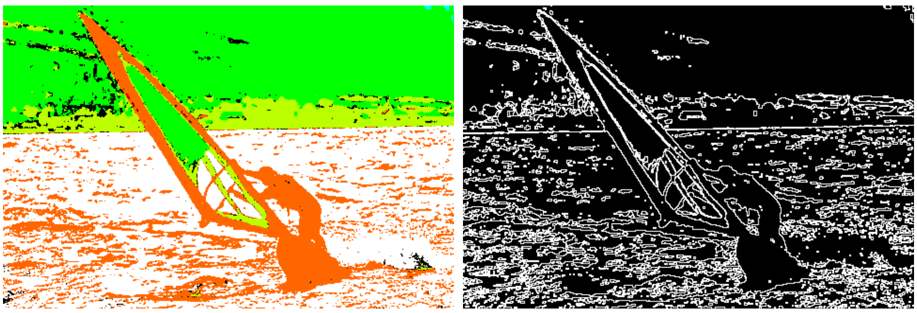
The performance of the proposed algorithm were tested using "The Berkeley Segmentation Dataset" [65], a dataset composed of 12,000 hand-labeled segmentations, from 30 human subjects, of 300 gray scale and color images of dimension 481×321 and 321×481. The tests were performed on the 100 color test images, out of the 300. Threshold has been fixed to $\epsilon = 0.2$; all the pixels presenting a saturation value lower than ϵ have been segmented by employing another threshold $\eta = 0.5$, i.e., pixels are labelled as "white" if their value component is greater than η, as "black" otherwise. For each image, the original image, the segmented image and the edges of the regions are showed. The first test aims to show how the dimension of the granules can affect the segmentation process. Analyzing Figures 6.7 and 6.9 it is possible to note how a larger granule dimension allows to produce wedges able to enclose more similar hues (see the green segment in the background of Fig. 6.7 (b) and 6.7 (c)) so to suppress small hue variations, while smaller granule dimension tends to beeter differentiate between similar hues. A larger granule size can be useful to segment images that show larger hue variance and hence obtain better PRI (Fig. 6.9 (b) and 6.9 (c)).

The aim of the second test is to show how the parameters w_1 and w_2 can be used to obtain different kinds of segmentation by weighting the importance of the number of pixels into the wedge with respect to the saturation–value dispersion. Higher values of w_1 mean that wedges enclosing few pixels are privileged, while higher value of w_2 privilege wedges characterized by higher saturation–value dispersion. Figure 6.11 shows results for image 6.11 using weights (a) $w_1 = 0.8$ and $w_2 = 0.2$ (b) $w_1 = 0.6$ and $w_2 = 0.4$ (c) $w_1 = 0.5$ and $w_2 = 0.5$ and granule dimension $w = 2$. Using parameters (a), the algorithm is able to discriminate between the mountain and the sky (Fig. 6.11), while with parameters (b) and (c) the algorithm yields a segment composed by both the mountain and sky thus producing a lower PRI (Fig. 6.11 (b) and 6.11 (c)).

Figure 6.7: Image 62096.



(a)



(b)

Figure 6.8: Segmented image a) $w_1 = 0.8$ $w_2 = 0.2$ $w = 4$ $PRI = 0.829390$; b) $w_1 = 0.8$ $w_2 = 0.2$ $w = 2$ $PRI = 0.804270$.

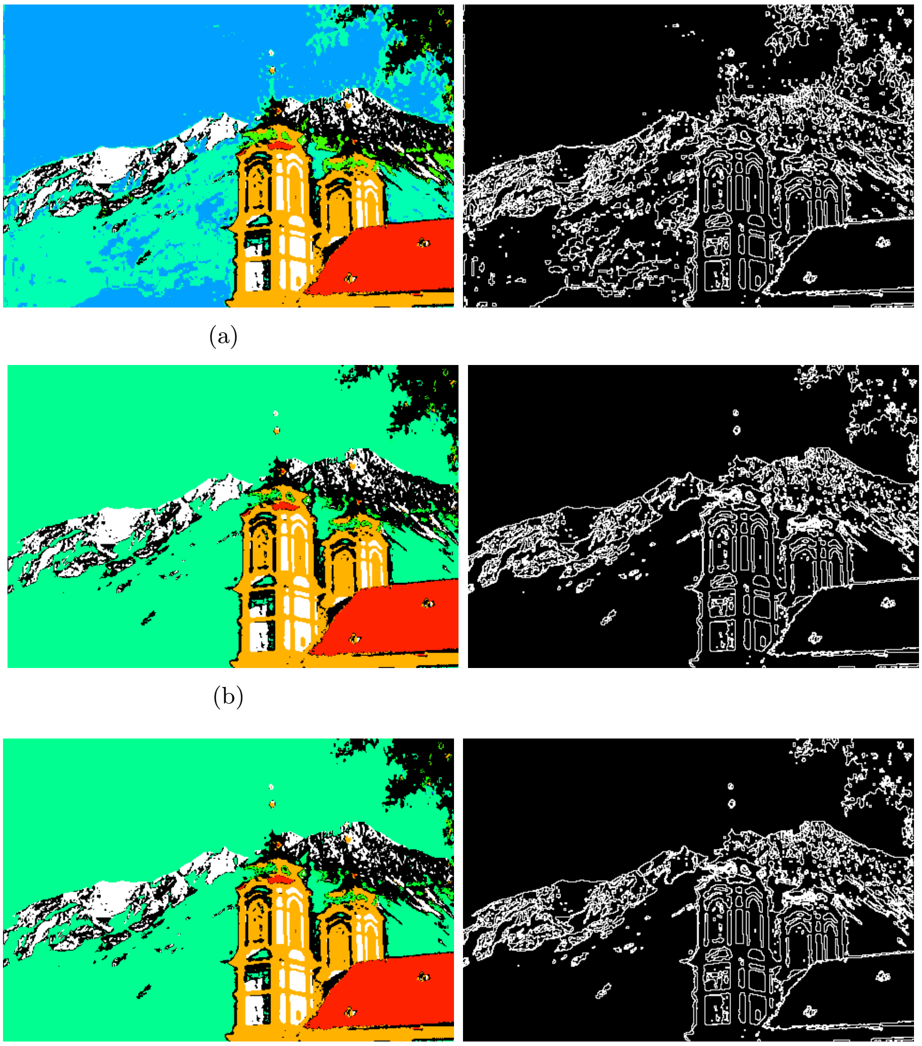
Figure 6.9: Image 189080.

(a)

(b)

Figure 6.10: Segmented image a) $w_1 = 0.8$ $w_2 = 0.2$ $w = 4$ $PRI = 0.818182$; b) $w_1 = 0.8$ $w_2 = 0.2$ $w = 4$ $PRI = 0.847405$.

Figure 6.11: Image 126007.



(a)

(b)

(c)

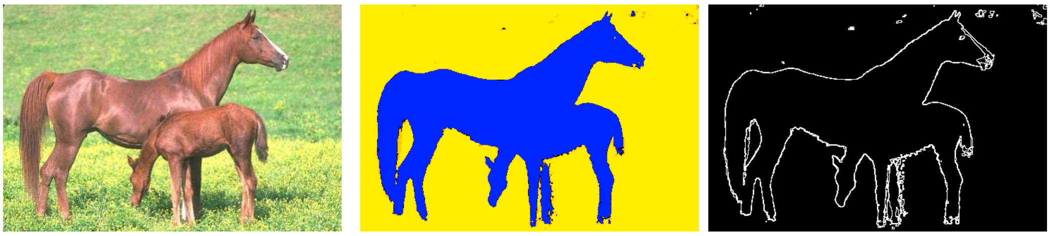
Figure 6.12: Segmented image a) $w_1 = 0.8$ $w_2 = 0.2$ $w = 2$ $PRI = 0.846230$; b) $w_1 = 0.6$ $w_2 = 0.4$ $w = 2$ $PRI = 0.719006$; c) $w_1 = 0.5$ $w_2 = 0.5$ $w = 2$ $PRI = 0.719006$.

Granule dimension w	$w_1 = 0.8$, $w_2 = 0.2$	$w_1 = 0.6$, $w_2 = 0.4$	$w_1 = 0.5$, $w_2 = 0.5$
$w = 2$	0.678028	0.663410	0.654179
$w = 4$	0.661959	0.636016	0.624948
$w = 8$	0.640885	0.621997	0.619233
$w = 16$	0.623986	0.613345	0.609314
$w = 32$	0.618413	0.601521	0.590546

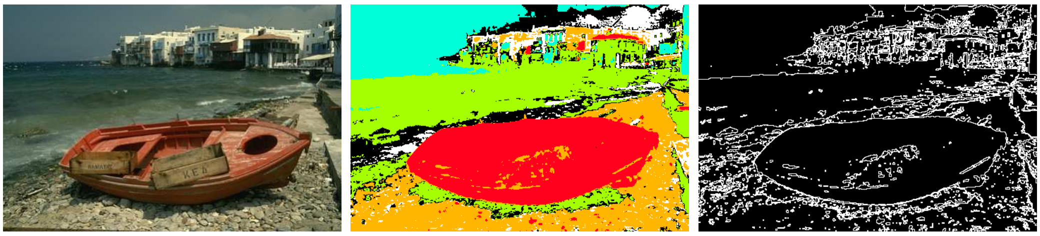
Table 6.1: PRI values for the 100 test images of the BSD.

Table 6.1 summarizes results obtained with different parameter configurations in terms of mean PRI computed over the 100 color images used for testing the algorithm. As can be seen, best results are obtained using small granule dimension and giving importance to the number of pixels over the saturation–value dispersion. Here we want to point out that, altough this configuration gives the best results on average, this does not imply that good results could not be obtained for single images employing different values (as reported in the previous examples).

Figure 6.13 show results for images 113044, 118020, 118035 and 361010 of the BSD.



(a)



(b)

(c)

(d)

Figure 6.13: Segmented image a) Image 113044 $w_1 = 0.6$ $w_2 = 0.4$ $w = 2$ $PRI = 0.774117$, b) Image 118020 $w_1 = 0.6$ $w_2 = 0.4$ $w = 2$ $PRI = 0.826157$, c) Image 118035 $w_1 = 0.7$ $w_2 = 0.3$ $w = 2$ $PRI = 0.870635$, d) Image 361010 $w_1 = 0.8$ $w_2 = 0.2$ $w = 2$ $PRI = 0.86228$.

7 Conclusions

Soft computing methodologies have been successfully employed in various image processing tasks including image segmentation, enhancement and classification, both individually or in combination with other soft computing techniques. In this thesis we have presented a model to manage uncertainty by means of a rough-fuzzy framework for exploiting feature level uncertainty.

7.1 Rough Fuzzy Product in Image Processing

The hybrid notion of rough fuzzy sets comes from the combination of rough and fuzzy models of uncertainty to exploit, at the same time, properties like coarseness, by handling rough sets , and vagueness, by handling fuzzy sets. In this combined framework, rough sets embody the idea of indiscernibility between objects in a set, while fuzzy sets model the ill-definition of the boundary of a sub class of this set. Marrying both notions leads to consider, as instance, approximation of sets by means of similarity relations or fuzzy partitions. The rough fuzzy synergy is hence adopted to better represent the uncertainty in granular computation. We have described a general framework based on the hybridization of rough and fuzzy sets and propose it as a viable and effective solution to some of the most difficult problems in image analysis. Also a new operator to compose rough fuzzy sets along with the proofs of its basics properties has been presented. This new operator, called $\mathcal{RF}$-product, can be viewed as a multiresolution approach, i.e. as a sequence of composition of rough fuzzy sets. We have presented a compression method, based on vector quantization, which employs the proposed rough-fuzzy framework. Feature extraction is based on the given definition of rough fuzzy sets while reconstruction of compressed images is performed exploiting $\mathcal{RF}$-product operator. Coding and decoding images by means of the proposed method gives good results when compared with other fuzzy-based methods, not only in terms of lower PSNR values but also, from a visual quality standpoint, remarkably reducing the blocking effect typical of block-based schemes. A color image segmentation technique, which exploits the given definition of rough fuzzy sets, has been also presented. The segmentation is performed by employing the definitions of lower and a upper histogram in the HSV color space build upon blocks of the image defined as rough fuzzy sets. Each bin of the lower and upper histograms presents some characteristic measures used to find the best segmentation of the image. It is shown that the proposed method is able to retain the structure of the color images leading to an effective segmentation.

7.2 Future Works

In this section we present possible evolutions of this thesis. With respect to the proposed compression algorithm, it can be noted that for lower compression rates performances are comparable with those of JPEG, although for higher compression rates it doesn't seem to perform at its best. Although the cause of this behaviour has been found in the quantization procedure, studies are necessary to further investigate other quantization algorithms. Further studies will be also directed to exploit the proposed method for coding and decoding color images. Concerning the presented segmentation technique, ongoing work is devoted to automatically select the best parameters. Also the algorithm relies only on color distribution without considering any spatial information. This can be a limitation for cluttered scenes where the color based segmentation could lead to oversegmentation. Another line of research is devoted to embed spatial relationships between image blocks to prevent oversgmentation and obtain more stable results.

Bibliography

[1] A. Ferone, A. Petrosino, S.K. Pal, *Rough-fuzzy based HSV Color Histogram for Image Segmetation*, Technical Report, RT-DSA-UNIPARTHENOPE-01-11, 2011

[2] A. Petrosino, A. Ferone, *Feature Discovery through Hierarchies of Rough Fuzzy Sets*, Granular Computing and Intelligent Systems: Design with Information Granules of Higher Order and Higher Type, Ed. S.M. Chen, W. Pedrycz, 2011 (to appear)

[3] A. Petrosino, A. Ferone, Rough fuzzy set-based image compression, Fuzzy Sets and Systems, Vol. 160, Issue 10, pp. 1485-1506, 2009

[4] A. Petrosino, S. K. Pal, A. Ferone, *Soft Computing Techniques in Image Analysis*, Handbook of Soft Computing in Video Surveillance, Ed. A. Petrosino S. K. Pal, L. Maddalena, 2011 (to appear)

Bibliography

[1] C. Aggarwal, N. Ta, J. Feng, J. Wang, M. J. Zaki. XProj, *A Framework for Projected Structural Clustering of XML Documents*, KDD Conference, 2007

[2] S. Asharaf, M.N. Murty, *An adaptive rough fuzzy single pass algorithm for clustering large data sets*, Pattern Recognition, vol. 36, no. 12, pp. 3015-3018, 2004.

[3] C. Barna, *Object labeling method using uncertainty measurement*, Intl. Workshop on Soft Computing Applications SOFA 09, pp. 225-228, 2009

[4] H. Bian and L. Mazlack, *Fuzzy-Rough Nearest-Neighbor Classification Approach*, Proceeding of the 22nd International Conference of the North American Fuzzy Information Processing Society (NAFIPS), pp. 500-505, 2003

[5] O. Byun, I. Park, D. Baek, S. Moon, *Video object segmentation using color fuzzy determination algorithm*, The 12th IEEE International Conference on Fuzzy Systems, FUZZ 03., vol.2, pp. 1305-1310 2003

[6] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, NewYork, 1981

[7] E. R. Caianiello, *A calculus of hierarchical systems, Proc. 1st Inf. Congress on Pattern Recognition*, Washington D. C., 1973

[8] A. Apostolico, E. R. Caiainiello, E. Fischetti, S. Vitulano, *C-calculus: an elementary approach to some problems in pattern recognition*, Pattern Recognition, 10, pp. 375-387

[9] E.R. Caianiello, A. Petrosino, *Neural networks, fuzziness and image processing*, V. Cantoni (Ed.), Machine and Human Perception: Analogies and Divergencies, Plenum Press, NewYork, pp. 355-370, 1994

[10] E.R. Caianiello, A.Ventre, *A model for C-calculus*, International Journal of General Systems, Vol. 11, 153-161, 1985

[11] R. Challoo, D.A. Clark, R.A. McLauchlan, S.I. Omar, *A fuzzy neural hybrid system*, IEEE World Congress on Computational Intelligence, 3, pp. 1654 - 1657, 1994

[12] S. Chanas, D. Kuchta, *Further remarks on the relation between rough and fuzzy sets, Fuzzy Sets and Systems*, Vol. 47, Issue 3, pp. 391-394, 1992

[13] S. Changjing, D. Barnes, S. Qiang, *Effective Feature Selection for Mars McMurdo Terrain Image Classification*, Intl. Conf. on Intelligent Systems Design and Applications ISDA 09, pp.1419-1424, 2009

[14] M. Chapron, *A new chromatic edge detector used for color image segmentation*, In Proc. 11th Int. Conf. on Pattern Recognition, vol. 3, pp. 311-314, 1992

[15] W.C. Chen, N.B. Chang, J.C. Chen, *Rough set-based hybrid fuzzy-neural controller design for industrial wastewater treatment*, Water Research, vol. 37, no. 1, pp. 95-107, 2003

[16] H.D. Cheng, X.H. Jiang, J. Wang, *Color image segmentation based on homogram thresholding and region merging*, Pattern Recognition 35, pp. 373-393, 2002

[17] S. Chimphlee, N. Salim, M.S.B. Ngadiman, W. Chimphlee, S. Srinoy, *Independent Component Analysis and Rough Fuzzy based Approach to Web Usage Mining*, Proc. Artificial Intelligence and Applications, 2006

[18] M. De Cock, C. Cornelis, E.E. Kerre, *Fuzzy rough sets: beyond the obvious*, IEEE International Conference on Fuzzy Systems, vol. 1, pp. 103-108, 2004

[19] F. DiMartino, V. Loia, I. Perfilieva, S. Sessa, *An image coding/decoding method based on direct and inverse fuzzy transforms*, International Journal of Approximate Reasoning, Vol. 48, pp. 110-131, 2008

[20] W. Dong, Q. Yu, C.N. Zhang, H. Li, *Image Quality Assessment Using Rough Fuzzy Integrals*, 27th International Conference on Distributed Computing Systems Workshops ICDCSW 2007, pp. 1-5, 2007

[21] D. Dubois, H. Prade, *Fuzzy Sets and Systems: Theory and Applications*, New York:Academic Press, 1985

[22] D. Dubois, H. Prade, *A Review of Fuzzy Sets Aggregation Connectives*, Inf. Science, 36:85-121, 1985

[23] D. Dubois, H. Prade, *Rough Fuzzy Sets and Fuzzy Rough Sets*, International Journal of General Systems, Vol. 17, Issue 2-3, pp. 191 - 209, 1990

[24] U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy, Eds., *Advances in Knowledge Discovery and Data Mining*, Menlo Park, CA: AAAI/MIT Press, 1996

[25] L. Feng, J-J. Qiu, Y.J. Cao, *Performance of the novel rough fuzzy neural network on short-term load forecasting*, Proc. Power Systems Conference and Exposition, pp. 543-547, 2004

[26] G. Flake, R. Tarjan, M. Tsioutsiouliklis, *Graph Clustering and Minimum Cut Trees*, Internet Mathematics, 1(4), 385-408, 2003

[27] J. Freixenet, X. Munoz, D. Raba, J. Marti, and X. Cuff, *Yet Another Survey on Image Segmentation: Region and Bound-*

ary Information Integration, Proc. European Conf. Computer Vision, pp. 408-422, 2002

[28] H. Frigui, *Adaptive image retrieval using the fuzzy integral*, in Proc. NAFIPS 99, New York, pp. 575-579, 1999

[29] X. Gao, B. Han, H. Ji, *A Shot Boundary Detection Method for News Video Based on Rough Sets and Fuzzy Clustering*, Image Analysis and Recognition, LNCS, Vol. 3656/2005, pp. 231-238, 2005

[30] A. Gersho, R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, Boston, MA, 1992

[31] D. Gibson, R. Kumar, A. Tomkins, *Discovering Large Dense Subgraphs in Massive Graphs*, VLDB Conference, 2005

[32] L. Han, J. F. Peters, S. Ramanna, R. Zhai, *Classifying Faults in High Voltage Power Systems: A Rough-Fuzzy Neural Computational Approach*, Proc. 7th International Workshop on Rough Sets, Data Mining, and Granular-Soft Computing, pp. 47-54, 1999

[33] B. Han, X. Gao, H. Ji, *A Shot Boundary Detection Method for News Video Based on Rough-Fuzzy Sets*, International Journal of Information Technology, Vol. 11, No. 7, 2005

[34] A.E. Hassanien, *Fuzzy-rough hybrid scheme for breast cancer detection*, Image and Computer Vision Journal, Elsevier, vol. 25, Issue 2, pp. 172-183, 2007

[35] K. Hirota, W. Pedrycz, *Fuzzy relational compression*, IEEE Trans. on Systems, Man, and Cybernetics 29 (3), pp. 407-415, 1999

[36] X. Hu, N. Cercone, *Mining knowledge rules from databases: A rough set approach*,in Proc. 12th Int. Conf. Data Eng. Washington, DC, pp. 96-105, 1996

[37] Q. Hu, D. Yu, Z. Xie, *Information-preserving hybrid data reduction based on fuzzy-rough techniques*, Pattern Recognition Letters, 27, pp. 414-423, 2006

[38] R. Jensen, Q. Shen, *Fuzzy-rough attribute reduction with application to web categorization*, Fuzzy Sets and Systems, vol. 141, pp. 469-485, 2004

[39] R. Jensen, Q. Shen, *A rough set-aided system for sorting WWW bookmarks*, in: N. Zhong et al. (Eds.), Web Intelligence: Research and Development, pp. 95-105, 2001

[40] R. Jensen, Q. Shen, *Fuzzy-rough sets for descriptive dimensionality reduction*, Proc. 11th Internat. Conf. on Fuzzy Systems, pp. 29-34, 2002

[41] R. Jensen, Q. Shen, *Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough Based Approaches*, IEEE Trans. Knowl. Data Eng., vol. 16, no. 12, pp. 1457-1471, 2004

[42] R. Jensen, Q. Shen, *Fuzzy-Rough Sets Assisted Attribute Selection*, IEEE Trans. on Fuzzy Systems, vol. 15, no. 1, pp. 73-89, 2007

[43] R. Jensen, Q. Shen, *New Approaches to Fuzzy-Rough Feature Selection*, IEEE Tans. on Fuzzy Systems, vol. 17, no. 4, 2009

[44] L. Jiangping, W. Yuke, *A Shortest Path Algorithm of Image Segmentation Based on Fuzzy-Rough Grid*, Intl. Conf. on Computational Intelligence and Software Engineering CiSE 09, pp.1-4, 2009

[45] L. Jiangping, P. Baochang, W. Yuke, *Tongue Image Segmentation Based on Fuzzy Rough Sets*, Intl. Conf. on Environmental Science and Information Application Technology ESIAT 09, vol.3, pp.367-369, 2009

[46] M. Juneja, E. Walia, P.S. Sandhu, R. Mohana, *Implementation and comparative analysis of rough set, Artificial Neural Network (ANN) and Fuzzy-Rough classifiers for satellite image classification*, Intl. Conf. on Intelligent Agent & Multi-Agent Systems IAMA 09, pp. 1-6, 2009

[47] N.B. Karayiannis, *A methodology for constructing fuzzy algorithms for learning vector quantization*, IEEE Trans. on Neural Networks, Vol. 8, pp. 505-518, 1997

[48] N.B. Karayiannis, *Generalized fuzzy k-means algorithms and their application in image compression*, in: Proc. SPIE: Applications of Fuzzy Logic Technology II, Vol. 2493, pp.206-217, 1995

[49] N.B. Karayiannis, *An axiomatic approach to soft learning vector quantization and clustering*, IEEE Trans. on Neural Networks, Vol. 10, pp. 1153-1165, 1999

[50] E.P. Klement, R. Mesiar, E. Pap, *Triangular Norms*, Kluwer Academic Publishers,1995

[51] G. J. Klir, B. Juan, *Fuzzy Sets and Fuzzy logic, Theory and Applications*, Prentice-Hall, New Jersey, 1995

[52] S. Kumar, S. Atri, L. Mandoria, *A Combined Classifier to Detect Landmines Using Rough Set Theory and Hebb Net Learning & Fuzzy Filter as Neural Networks*, Intl. Conf. on Signal Processing Systems, pp.423-427, 2009

[53] S.Z. Li, *Markov Random Field Modeling in Computer Vision*, Ed. Tosiyasu L. Kunii, Springer-Verlag, Berlin, 1995

[54] Y. Linde, A. Buzo, R.M. Gray, *An algorithm for vector quantizer design*, IEEE Trans. on Communications, Vol. 28, pp. 84-95, 1990

[55] P. Lingras, *Rough Neural Networks*, Proc. International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'96), pp. 1445-1450, 1996

[56] P. Lingras, *Unsupervised Learning Using Rough Kohonen Neural Network Classifiers*, Proc. Symposium on Modelling, Analysis and Simulation, CESA'96 IMACS Multiconference, pp. 753-757, 1996

[57] P. Lingras, *Fuzzy-Rough and Rough-Fuzzy Serial Combinations in Neurocomputing*, Neurocomputing Journal, vol. 36, pp. 29-44, 2001

[58] C.T. Lin, C.S.G. Lee, *Neural Fuzzy Systems: a Neuro-fuzzy Synergism to Intelligent Systems*, Upper Saddle River-NJ: Prentice Hall, 1996

[59] C. T. Lin, C. S. G. Lee, *Real-time supervised structure/parameter learning for fuzzy neural networks*, Proc. IEEE International Conference of Fuzzy Systems, 2(1), pp. 1283 - 1291, 1992

[60] H. J. Lu, R. Setiono, and H. Liu, *Effective data mining using neural networks*, IEEE Trans. Knowledge Data Eng., vol. 8, pp. 957-961, 1996

[61] J. L.Ma, *A Bibliography of Fuzzy Systems* , Xueyuan Press, Beijing, 1989

[62] P. Maji, S.K. Pal, *Rough Set Based Generalized Fuzzy C-Means Algorithm and Quantitative Indices*, IEEE Trans. on Systems Man and Cybernetics, vol.37, no.6, pp.1529-1540, 2007

[63] D. Malyszko, J. Stepaniuk, *Standard and Fuzzy Rough Entropy Clustering Algorithms in Image Segmentation*, Rough Sets and Current Trends in Computing, LNCS, Vol. 5306/2008, pp. 409-418, 2008

[64] C-W. Mao, S-H. Liu, J-S. Lin, *Classification of multispectral images through a rough-fuzzy neural network*, Optical engineering, vol. 43, pp. 103-112, 2004

[65] D. Martin, C. Fowlkes, D. Tal, J. Malik, *A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics*, in: Proc. 8th Int'l Conf. Computer Vision, vol. 2, pp. 416-423, 2001

[66] M. Meila, *Comparing Clusterings: An Axiomatic View*, Proc. 22nd IntÕl Conf. Machine Learning, pp. 577-584, 2005

[67] J.S. Mi and W.X. Zhang, *An axiomatic characterization of a fuzzy generalization of rough sets*, Information Sciences, vol. 160, no. 1-4, pp. 235-249, 2004

[68] S. Mitra, S.K. Pal, P. Mitra, *Data mining in soft computing framework: a survey*, IEEE Trans. on Neural Networks, vol.13, no.1, pp.3-14, 2002

[69] S. Mitra, H. Banka, W. Pedrycz, *Rough-Fuzzy Collaborative Clustering*, IEEE Trans. on Systems Man and Cybernetics, vol.36, no.4, pp.795-805, 2006

[70] S. Mitra, B. Barman, *Rough-Fuzzy Clustering: An Application to Medical Imagery*, Rough Sets and Knowledge Technology, LNCS, Vol. 5009/2008, pp. 300-307, 2008

[71] M. Mizumoto, *Fuzzy Sets and Their Operators*, Information and Control, 50, pp. 166-174, 1981

[72] A. Mohabey, A.K. Ray, *Rough set theory based segmentation of color images*, Fuzzy Information Processing Society, 2000. NAFIPS. 19th International Conference of the North American, pp.338-342, 2000

[73] A. Mohabey, A.K. Ray, *Fusion of rough set theoretic approximations and FCM for color image segmentation*, IEEE Intl. Conf. on Systems Man and Cybernetics, vol.2, pp.1529-1534, 2000

[74] T. Mollestad and A. Skowron, *A rough set framework for data mining of propositional default rules*, in Lecture Notes Comput. Sci., vol. 1079, pp. 448-457, 1996

[75] M.M. Mushrif, A.K. Ray, *A-IFS Histon Based Multithresholding Algorithm for Color Image Segmentation*, IEEE Signal Processing Letters, vol.16, no.3, pp.168-171, 2009

[76] M.M. Mushrif, A.K. Ray, *Color image segmentation: Rough-set theoretic approach*, Pattern Recognition Letters, Vol. 29, Issue 4, pp. 483-493, 2008

[77] N.M. Nasrabadi, R.A. King, *Image coding using vector quantization: a review*, IEEE Trans. on Communications, Vol. 36, pp. 956-971, 1988

[78] O. Nasraoui, R. Krishnapuram, and A. Joshi, *Relational clustering based on a new robust estimator with application to web mining*, in Proc. NAFIPS 99, New York, pp. 705-709, 1999

[79] S. J. Nowlan, *Gain variation in recurrent error propagation networks*, Complex Systems 2, pp. 305 - 320, 1988

[80] S. K. Pal, A. Ghosh, and M. K. Kundu, Eds., *Soft Computing for Image Processing*. Heidelberg, Germany: Physica-Verlag, 2000

[81] S.K. Pal, A. Skowron (Eds.), *Rough-Fuzzy Hybridization: A New Trend in Decision Making*, Springer Verlag, Singapore, 1999

[82] S.K. Pal, S. Mitra, *Neuro-Fuzzy Pattern Recognition : Methods in Soft Computing*, New York: John Wiley, 1999

[83] S.K. Pal, S. Mitra, P. Mitra, *Rough-Fuzzy MLP: Modular Evolution, Rule Generation, and Evaluation*, IEEE Trans. on Knowledge and Data Engineering, vol. 15, no. 1, pp. 14-25, 2003

[84] S.K. Pal, *Pattern Recognition Algorithms for Data Mining*, Chapman and Hall, 2004

[85] S.K. Pal, *Multispectral Image Segmentation Using the Rough-Set-Initialized EM Algorithm*, IEEE Trans. On Geoscience and Sensing, Vol. 40, No. 11, 2002

[86] S.K. Pal, *Rough-Fuzzy Granulation, Rough Entropy and Image Segmentation*, First Asia International Conference on Modelling & Simulation AMS 07, pp. 3-6, 2007

[87] S. K. Pal, S. Mitra, *Multilayer perceptron, fuzzy sets, and classification*, IEEE Trans. Neural Networks, 3(5), pp. 683 - 697, 1992

[88] N.R. Pal, S.K. Pal, *A review on image segmentation techniques*, Pattern Recognition 26 (9), pp. 1277-1294, 1993

[89] D.K. Panjwani, G. Healey, *Markov random field models for unsupervised segmentation of textured color images*, IEEE Trans. Pattern Anal. Mach. Intell. 17 (10), pp. 939-954, 1995

[90] Z. Pawlak, *Rough sets*, Int. J. of Information and Computer Sciences, 11, 5, 341-356, 1982.

[91] Z. Pawlak, A. Skowron, *Rough membership function*, in: R. E Yeager, M. Fedrizzi and J. Kacprzyk (eds.), Advaces in the Dempster-Schafer of Evidence, Wiley, New York, 1994, 251-271.

[92] Z. Pawlak, *Rough Sets, Theoretical Aspects of Reasoning about Data*, Dordrecht, Kluwer, 1991

[93] Z. Pawlak, *Granularity of knowledge, indiscernibility and rough sets*, Proceedings of IEEE International Conference on Fuzzy Systems, 106-110, 1998

[94] W. Pedrycz, *Fuzzy Control and Fuzzy Systems*, New York: John Wiley, 1989

[95] W. Pedrycz, *Conditional fuzzy c-means*, Pattern Recognition Lett., vol. 17, pp. 625-632, 1996

[96] W. Pedrycz, *Shadowed sets: bridging fuzzy and rough sets*, in Rough-Fuzzy Hybridization: A New Trend in Decision Making, Springer Verlag, Singapore, pp. 179-199, 1999

[97] I. Perfilieva, *Fuzzy transforms and their applications to data compression*, IEEE Internat. Conf. on Fuzzy Systems, pp.294-299, 2005

[98] A. Petrosino, M. Ceccarelli, *Unsupervised Texture Discrimination Based on Rough Fuzzy Sets and Parallel Hierarchical Clustering*, Proc. IEEE International Conference on Pattern Recognition, pp. 1100-1103, 2000

[99] A. Petrosino, G. Salvi, *Rough fuzzy set based scale space transforms and their use in image analysis*, Int. J. of Approximate Reasoning, vol. 41, Issue 2, pp. 212-228, 2006

[100] A. Petrosino, P. Feng, *A fuzzy hierarchical neural network for image analysis*, Systems, Man and Cybernetics, 1993. Intl. Conf. Systems Engineering in the Service of Humans, vol. 2, pp.657-662, 1993

[101] A. Petrosino, G. Salvi, *A Rough Fuzzy Neural Based Approach to Face Detection*, Intl. Conf. on Image Processing, Computer Vision and Pattern Recognition IPCV10, 2010

[102] L. Polkowski, A. Skowron, *Rough Sets in Knowledge Discovery 1 and 2*, Heidelberg, Germany: Physica-Verlag, 1998

[103] Y. Qian, J. Liang, Y. Yao, C. Dang, *MGRS: A multi-granulation rough set*, Information Sciences, Vol. 180, Issue 6, pp 949-970, 2010

[104] A.M. Radzikowska, E.E. Kerre, *A comparative study of fuzzy rough sets*, Fuzzy Sets and Systems, vol. 126, no. 2, pp. 137-155, 2002

[105] M. Rattigan, M. Maier, D. Jensen, *Graph Clustering with Network Sructure Indices*, ICML, 2007

[106] S. Russell, W. Lodwick, *Fuzzy clustering in data mining for telco database marketing campaigns*, in Proc. NAFIPS 99, New York, NY, pp. 720-726, 1999

[107] M. Sarkar, *Fuzzy-Rough nearest neighbors algorithm*, Proc. IEEE International Conference on Systems, Man and Cybernetics, pp. 3556-3561, 2000

[108] D. Sen, S. K. Pal, *Histogram Thresholding using Beam Theory and Ambiguity Measures*, Fundamenta Informaticae, Vol. 75, No. 1-4/2007, pp. 483-504, 2007

[109] D. Sen, S.K. Pal, *Generalized Rough Sets, Entropy, and Image Ambiguity Measures*, IEEE Trans. on Systems Man and Cybernetics, vol.39, no.1, pp.117-128, 2009

[110] D. Sen, S.K. Pal, *Measuring Ambiguities In Images Using Rough And Fuzzy Set Theory*, Signal Processing, Intl. Conf. Communications and Networking ICSCN 08, pp.333-338, 2008

[111] L. Shafarenko, M. Petrou, J.V. Kittler, *Histogram based segmentation in a perceptually uniform color space*, IEEE Trans. Image Process. 7 (9), pp. 1354-1358, 1998

[112] N. Shan, W. Ziarko, *Data-based acquisition and incremental modification of classification rules*, Comput. Intell., vol. 11, pp. 357-370, 1995

[113] C. Shang, Q. Shen, *Aiding neural network based image classification with fuzzy-rough feature selection*, IEEE Intl. Conf. on Fuzzy Systems FUZZ-IEEE 2008 IEEE World Congress on Computational Intelligence, pp.976-982, 2008

[114] J. Shawe-Taylor, N. Cristianini, *Kernels Methods for Pattern Analysis*, Cambridge University Press, 2004

[115] Q. Shen, A. Chouchoulas, *A modular approach to generating fuzzy rules with reduced attributes for the monitoring of complex systems*, Eng. Appl. Artif. Intell. 13 (3) (2000) 263-278

[116] H. Shi, Y. Shen, Z. Liu, *Hyperspectral bands reduction based on rough sets and fuzzy C-means clustering*, Proceedings of the 20th IEEE IMTC 03, vol.2, pp. 1053-1056, 2003

[117] P. Smets, *Implications and Modus Ponens in Fuzzy Logic*, Conditional Logic in Expert Systems, Elsevier Science Publisher, pp. 235 - 268, 1991

[118] W. Skarbek, A. Koschan, *Colour image segmentation: A survey*, Technical Report 94-32, Computer Science Department, Berlin Technical University, 1994

[119] A. Skowron, *Extracting laws from decision tables-A rough set approach*, Comput. Intell., vol. 11, pp. 371-388, 1995

[120] A. Sperduti and A. Starita, *Supervised neural networks for the classification of structures*, IEEE Trans. Neural Network, vol. 8, no. 2, pp. 429-459, 1997

[121] K. Thangavel, A. Pethalakshmi, *Dimensionality reduction based on rough set theory: A review*, Applied Soft Computing 9, pp. 1-12, 2009

[122] H. Thiele, *Fuzzy rough sets versus rough fuzzy sets - an interpretation and a comparative study using concepts of modal*

logics, Technical report no. CI-30/98, University of Dortmund, 1998.

[123] A. B. Tickle, R. Andrews, M. Golea, and J. Diederich, *The truth will come to light: Directions and challenges in extracting the knowledge embedded within trained artificial neural networks*, IEEE Trans. Neural Networks, vol. 9, pp. 1057-1068, 1998

[124] A. Trémeau, P. Colantoni, *Regions adjacency graph applied to color image segmentation*, IEEE Trans. Image Process. 9 (4), pp. 735-744, 2000

[125] E. C. C. Tsang, D. Chen, D. S. Yeung, X. Wang, J. W. T. Lee, *Attributes Reduction Using Fuzzy Rough Sets*, IEEE Trans. on Fuzzy Systems, vol. 16, No. 5, pp. 1130-1141, 2008

[126] I. B. Turksen, *Fuzzy data mining and expert system development*, Proc. IEEE Int. Conf. Syst., Man, Cybern., pp. 2057-2061, 1998

[127] T. Uchiyama, M.A. Arbib, *Color image segmentation using competitive learning*, IEEE Trans. Pattern Anal. Mach. Intell. 16 (12), pp. 1197-1206, 1994

[128] R. Unnikrishnan, C. Pantofaru, and M. Hebert, *A Measure for Objective Evaluation of Image Segmentation Algorithms*, Proc. CVPR Workshop Empirical Evaluation Methods in Computer Vision, 2005

[129] X. Wang, J. Yang, X. Teng, N. Peng, *Fuzzy-Rough Set Based Nearest Neighbor Clustering Classification Algorithm*, Lecture Notes in Computer Science, vol. 3613/2005, pp. 370-373, 2005

[130] Z. Wang, X. Shao, G. Zhang, H. Zhu, *Integration of Variable Precision Rough Set and Fuzzy Clustering: An Application to Knowledge Acquisition for Manufacturing Process Planning,*

Proc. 10th International Conference, RSFDGrC 2005, pp. 585-593, 2005

[131] D. Wang, M. Wu, *Binary fuzzy rough set model based on triangle modulus and its application to image processing*, IEEE Intl. Conf. on Cognitive Informatics ICCI 09, pp. 249-255, 2009

[132] S. K. M. Wong, W. Ziarko, *Comparison of the probabilistic approximate classification and the fuzzy set model*, Fuzzy Sets and Systems 21, pp. 357-362, 1987

[133] W.Z. Wu, J.S. Mi, and W.X. Zhang, *Generalized fuzzy rough sets, Information Sciences*, vol. 151, pp. 263-282, 2003.

[134] M. Wygralak, *Rough sets and fuzzy sets - some remarks on interrelations*, Fuzzy Sets and Systems, vol. 29, no. 2, pp. 241-243, 1989

[135] G. Xian, B. Zeng, *A novel hybrid model for information processing basing on rough sets and fuzzy SVM*, Intl. Conf. on Multimedia and Ubiquitous Engineering MUE08, pp.320-323, 2008

[136] Y. Y. Yao, *Two views of the theory of rough sets in finite universes*, International Journal of Approximate Reasoning 15, pp. 291-317, 1996

[137] Y. Y. Yao, *A comparative study of fuzzy sets and rough sets*, Journal of Information Sciences, vol. 109, Issues 1-4, pp. 227-242, 1998

[138] Y.Y. Yao, *Combination of rough and fuzzy sets based on a-level sets*, T.Y. Lin, N. Cereone (Eds.), Rough Sets and Data Mining: Analysis of Imprecise Data, Kluwer Academic Publishers, pp. 301-321, 1997

[139] Y.Y. Yao, *Information granulation and rough set approximation*, International Journal of Intelligent Systems, Vol. 16. No. 1, 87-104, 2001

[140] D.S. Yeung, D. Chen, E.C.C. Tsang, J.W.T. Lee, and W. Xizhao, *On the Generalization of Fuzzy Rough Sets*, IEEE Trans. on Fuzzy Systems, vol. 13, no. 3, pp. 343-361, 2005.

[141] L. Yurong, Q. Bin, *Hierarchical Reduction Algorithm of Rough Sets*, Sixth International Conference on Intelligent Systems Design and Applications, vol. 1, pp.497-502, 2006

[142] L.A. Zadeh, *Fuzzy Sets, Information and Control*, vol. 8, pp. 338-353, 1965

[143] L.A. Zadeh , *The Concept of a Linguistic Variable and Application to Approximate Reasoning*, Part I, Information Sciences, 8, pp. 199-249, 1975

[144] L.A. Zadeh , *Knowledge Representation in Fuzzy Logic*, In An Introduction to Fuzzy Logic Applications in Intelligent Systems R. R. Yager and L. A. Zadeh eds., pp. 1-25, Boston: Kluwer Academic, 1992

[145] L.A. Zadeh , *Fuzzy Sets as a Basis for a Theory of Possibility*, Fuzzy Sets and Systems, pp. 3-28, 1978

[146] J. Zhang, *Fuzzy neural networks based on rough sets for process modeling*, Proc. Fifth International Symposium on Instrumentation and Control Technology, pp. 844-847, 2003.

[147] D. Zhang, Y. Wang, H. Huang, *Rough neural network modeling based on fuzzy rough model and its application to texture classification*, Neurocomputing, Vol. 72, pp. 2433-2443, 2009

[148] W. Zhang, C. Li, Y. Zhang, *A new hybrid algorithm for image segmentation based on rough sets and enhanced fuzzy c-means*

clustering, IEEE Intl. Conf. on Automation and Logistics ICAL 09, pp.1212-1216, 2009

[149] Z. Zheng, H. Hu, Z. Shi, *Rough Set Based Image Texture Recognition Algorithm*, Knowledge-Based Intelligent Information and Engineering Systems, LNCS, Vol. 3213/2004, pp. 772-778, 2004

[150] Y. Zhao, X. Zhou, G. Tang, *A rough set-based fuzzy clustering*, Proc. Second Asia information retrieval symposium, pp. 401-409, 2005

[151] L. Zhou, W. Li, Y. Wu, *Face Recognition Based on Fuzzy Rough Set Reduction*, International Conference on Hybrid Information Technology ICHIT'06, Vol. 1, pp. 642-646, 2006

[152] W. Ziarko, *Probabilistic Rough Sets, Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, LNCS, Vol. 3641, 283-293, 2005