MEDICAL IMAGE SEGMENTATION ALGORITHMS USING A FUZZY EQUIVALENCE RELATION

PhD Dissertation

Martin I. Tabakow
Department of Computer Science
Wroclaw University of Technology
Poland

Supervisor: Prof. Dr hab. M. Kurzyński

Wrocław 2003

Martin I. Tabakow dedicates this monograph to his parents, his wife Marta and his children Patryk and Iwo.

Contents

Abstract

The subject of this dissertation is medical image segmentation by means of a fuzzy equivalence relation. The main areas considered are: membership functions and matrices, distance functions of the Minkowski class, distance functions of the Canberra class, linear convex combinations of such distance functions, fuzzy similarity relations, transitive max-min closures, fuzzy equivalence relations, α-cuts and fuzzy expected values, fuzzy c-means clustering, fuzzification, measures of fuzziness and linear convex combinations of such measures. The results obtained are used as a base for the development of a new method for medical image segmentation by introducing a properly defined fuzzy equivalence relation. A corresponding fuzzy equivalence relation-based image segmentation algorithm (in short: FERIS) is introduced and various experiments concerning the proposed approach are given. The major advantage of the FERIS algorithm lies in the ease of implementation. FERIS is a well-defined linear structure algorithm of polynomial complexity. The corresponding process of image segmentation is always realised in a unique way and the algorithm converges in a finite time. The obtained image segmentation is realised as an automated and resolution-independent computational process. The classical fuzzy c-means algorithm (in short: FCM) requires that the desired number of clusters be given in advance. This can be problematic when the clustering problem does not specify any desired number of clusters. The number of clusters should reflect the stricture of the given data. The above-proposed method is based on using a properly defined fuzzy equivalence relation. And so, it satisfies this need. Moreover, in contradistinction to the classical FCM algorithm, FERIS can be implemented for real-time applications. In experiments, this method demonstrated a high quality and promising performance for various classes of medical images. These experiments exhibited that the proposed algorithm is very robust to noise, spatial and temporal inhomogeneities. And finally, the presented method of image segmentation may be easily extended to other areas of application, e.g. such as: biology, geology, meteorology, urbanisation, chemistry, and so on, i.e. everywhere, where image processing is needed. In the last case, only some initial calibration of FERIS is necessary to be done.

The used designations and abbreviations

$\{x_1,x_2,...,x_n\}$	a finite set X of n different elements x_i (i=1,2,...,n) ;
$\{ x / \varphi(x) \}$	the set of all x such that $\varphi(x)$;
$f : X \rightarrow Y$	the function f is a mapping (into or onto) from X to Y;
$=_{df}$	equals by definition ;
$\sim,\wedge,\Rightarrow,\Leftrightarrow,\vee$	negation, conjunction, implication, equivalence and disjunction ;
$x \in X$ ($x \notin X$)	x is (is not) an element of X ;
$\forall , \exists , \exists!$	Universal ,existentional and uniqueness quantifiers ;
$\lvert X \rvert$	cardinality (or power) of a set X ;
$X \subseteq Y$ ($X \subset Y$)	X is a subset (a proper subset) of Y ;
$X \cup Y , X \cap Y , X - Y$	union ,intersection and difference of sets ;
$\bigcup_{i=1}^{m} X_i , \bigcap_{i=1}^{m} X_i$	generalised union and generalised intersection of n sets X_i ;
$\varnothing$	null (or empty) set ;
$X \times Y$	Cartesian product of sets ;
X^n	$X \times X \times ... \times X$ n times, $\underline{x} =_{df} (x_1,x_2,...,x_n) \in X^n$ denotes a vector (or also a transposed vector, i.e. $\underline{x}^T$: depending on the context) ;
$(\underline{x})_i$	the ith component of the vector $\underline{x} \in X^n$,i.e. x_i ($i \in \{1,...,n\}$) ;
X' , $\mathcal{P}(X)$	complement of a set (to the universal set) and the power set of X ;
dom(ρ), cod(ρ)	the domain and the codomain of a binary relation ρ ;
X/ρ	quotient set wrt (the binary relation) ρ ;
$[x]_\rho$	equivalence class under ρ (an equivalence relation) ;
$\rho \circ \sigma$	the composition of two binary relations ρ and σ ;
$\mathcal{N}$	the set of natural numbers ;
$\mathcal{R}$, $\mathcal{R}^*$, $\mathcal{R}_+$	the sets of all real numbers and all nonnegative real numbers, $\mathcal{R}_+ =_{df} \mathcal{R}^* - \{0\}$;
μ , μ(x)	a fuzzy subset of a set, say X and the degree of membership of x in μ ($x \in X$) ;
μ^*	the crisp set associated with a fuzzy set μ ;
d(μ, ν)	a distance between the fuzzy sets μ and ν ;
$[0,1] \subset \mathcal{R}_+$	the closed interval [0,1] ;

$\mathcal{P}(X)$ or $\mathcal{P}$ (if X is understood)	a family of fuzzy subsets of X ;
$\mathcal{FP}(X)$	the fuzzy power set of a set X ;
$\varphi(\mu)$	a measure φ of fuzziness of a fuzzy set μ ;
$H(\mu)$ entropy of fuzzy sets measure	entropy of fuzzy sets measure ;
M_ρ	the membership matrix of a fuzzy relation ρ ;
ρ_α	the α-cut wrt the fuzzy relation ρ ;
$\pi_\alpha =_{df} \pi(\rho_\alpha)$	the partition generated by ρ_α ;
G	is a graph ;
$\subseteq$	a partition refinement relation ;
supp(μ), sup $\{\cdot\}$, inf $\{\cdot\}$	the support of a fuzzy set μ , supremum, infimum ;
c (or C)	the number of fuzzy classes in a fuzzy partition ;
(n) modulo m	$= b_2$, if $n = b_1 \cdot m + b_2$ (n,m,b_1,b_2 – integers, $0 \leq b_2 < m$) ;
$\mathbf{x},\mathbf{v},\mathbf{x}_k, \ldots, \|\cdot\|, \lvert\cdot\rvert$	some vectors , a norm in $\mathcal{R}^p$, the absolute value ;
$U = [u_{ik}]_{c \times n}$	a fuzzy pseudopartition matrix ;
$\mathcal{I}$	an image (a finite set of pixels $p \in \mathcal{I}$) ;
FCM	Fuzzy C – Means clustering ;
PI_m	the Performance Index associated with a given pseudopartition $\mathcal{P}$;
$A(\cdot)$	one argument predicate ;
$\lambda_{(\rho,\sigma)}$	a linear convex combination of the distance functions ρ and σ ;
μ_{dark} , μ_{grey1} , μ_{grey2} , μ_{grey3} , μ_{bright} , $\mathcal{M}$	the dark class, grey class and bright class membership functions , $\mathcal{M} =_{df} \{\mu_{dark}, \mu_{grey1}, \mu_{grey2}, \mu_{grey3}, \mu_{bright}\}$;
IMT, USG	intima-media thickness, ultrasonography ;
MRI (or MRI imaging), CT	magnetic resonance imaging, computed tomography ;
d, g, b, $H(\mathcal{I})$	dark, grey, bright, the histogram of $\mathcal{I}$;
RGB, YUV, HSB	red, green, blue; luminance, saturation, value; hue, saturation, brightness;
FEV (WFEV), FERIS	the (weighted) fuzzy expected value , fuzzy equivalence relation-based image segmentation ;
L , T	the total number of grey levels, a threshold value ;
τ	the total computational time of FERIS ;
SISD (SIMD)	single instruction, single (multiple) data ;
CNS	central nervous system ;
max, min	the logical operations maximum and minimum ;
Iff, wrt	if and only if , with respect to.

Introduction

Since time immemorial, vision in general and images in particular have played an important an essential role in human life. Nowadays, the field of image processing also has numerous scientific, commercial, industrial and military applications. All these applications result from the interaction between fundamental scientific research on the one hand, and the development of new and high-standard technology on the other hand. Regarding the scientific component, quite recently the scientific community became familiar with 'fuzzy techniques' in image processing, which make use of the framework of fuzzy sets and related theories. The theory of fuzzy sets was initiated by Zadeh (1965), and is one of the most developed models to treat imprecision and uncertainty. Instead of the classical approach that an object belongs or does not belong to a set, the concept of a fuzzy set allows a gradual transition from membership to nonmembership, providing partial degrees of a membership. Fuzzy techniques are often complementary to existing techniques and can contribute to the development of better and more robust methods, as has already been illustrated in numerous scientific branches (Kerre and Nachtegael 2000).

The digital revolution and the computer's processing power in combination with imaging modalities such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography, ultrasound, etc., have helped humans to better understand the complex human anatomy and its behaviour to a certain extent. Computer power and medical scanner data alone are not enough; we need the art to extract the necessary boundaries, surfaces, and segmented volumes these organs in the spatial and temporal domains. This art of organ extraction is segmentation.

Image segmentation is essentially a process of pixel classification, wherein the image pixels are segmented into subsets by assigning the individual pixels to classes. These segmented organs and their boundaries are very critical in the quantification process for physicians and medical surgeons, in any branch of medicine, which deals with imaging (Suri, Setarehdan and Singh 2002). Moreover the field of medicine has become a very attractive area for the application of fuzzy set theory. This is due to the large role imprecision and uncertainty play in the field (Mordeson, Malik and Cheng 2000). The first step of any fuzzy applications is the so-called 'fuzzification' of the considered input space. In general, any such fuzzification can be realised by means of various techniques.

There have been many different families of segmentation algorithms proposed in past years. These algorithms can be categorised into edge-based (Marchisio, Koperski and Sanella 2000), clustering-based (Noordam,Van Den Broek and Buydens 2000, Pham, Wagner and Clark 2001), region-based (Fradkin, Roux, Maitre and Leloglu 1999), and split/merge (Tyagi and Bayoumi, 1989, Tu, Zhu and Shum 2001) algorithms. Some advanced algorithmic approaches to medical image segmentation ((Suri, Setarehdan and Singh 2002) with applications in cardiology, neurology, mammography and pathology were also presented.

Among the above used algorithms or approaches, clustering-based algorithms are ones of those proved to be suited for medical or in general for remotely sensed imagery (Mordeson, Malik and Cheng 2000, Li, Dong and Gao 2002).

Hence *fuzzy clustering* plays a fundamental role in the image segmentation process. Thus it can be observed large research intensivity in the field, e.g. (Tizhoosh 1998). Two main clustering segmentation approaches based on crisp and fuzzy methods have been developed. The crisp clustering segmentation algorithms generate clusters such that each pixel in an image is assigned to exactly one cluster. However, fuzzy segmentation algorithms try to cope with each cluster as a fuzzy set, and each pixel in a image has a membership value (ranging between 0 and 1) associated to each cluster, measuring how much the pixel belong to that particular cluster. In the fuzzy segmentation algorithms, the most popular is the *fuzzy c-means algorithm* (Bezdek 1981) and many research works have been proposed to steep up this classical approach, e.g. multi – modality fuzzy c – means (Bloch I. 1994), using a performance index of a fuzzy pseudopartition (Klir and Yuan 1995), a modified fuzzy c – means algorithm by using of extended Xie-Beni index (Xie and Beni 1991, Pal and Bezdek 1995), conditional fuzzy c – means algorithm, i.e. using some conditional variables (Pedrycz 1996), pixel resistance and neighbourhood oriented approach (Mohamed, Ahmed and Farag 1998), density-weighted fuzzy c-means (Chen and Wang 1999), improved robust fuzzy clustering (Melek, Emami and Goldenberg 1999), modified fuzzy c-means (Thitimajshima 2000), fuzzy clustering with ε -insensitive loss function (Łęski 2001), adaptive fuzzy c – means (e.g. see: Suri, Setarehdan and Singh 2002), etc.

Although the fuzzy c-means algorithm is powerful in image segmentation, there is still a drawback encountered, namely the desired number of clusters should be specified in advance. This is a disadvantage whenever the clustering problem cannot specify any desired number of clusters. The situations are often for remotely sensed image segmentation, because the ground truth is always not available for these images (Mordeson, Malik and Cheng 2000, Li, Dong and Gao 2002).

The number of clusters should reflect the structure of the given data. Methods based on *fuzzy equivalence relations* satisfy this need (Mordeson, Malik and Cheng 2000). In comparison with the classical approach, the use of some well-defined fuzzy equivalence relations seems to be a new and more attractive stage in the fuzzy clustering investigations. An introductory research (Helgason, Jobe, Malik and Mordeson 1999, Helgason, Jobe, Mordeson, Malik and Cheng 1999) was related to medical diagnosis of patients (an analysis of data from 30 patients for measuring the degree of casual efficacy or conditions reflecting the abnormalities of blood flow, coagulation, and vascular walls damage). The next investigations were related to the possibility of using fuzzy equivalence relations in medical image segmentation (Tabakow 2001), image segmentation of an urban environment (Li, Dong and Gao 2002), etc.

The subject of this dissertation is medical image segmentation by means of a fuzzy equivalence relation. The main areas considered are: fuzzy sets and relations, distance functions of the Minkowski class, transitive max-min closure algorithms, fuzzy c-means clustering algorithms, fuzzification, fuzzy equivalence relations and measures of fuzziness. The results obtained are used as a base for the development of a new method for medical image segmentation by introducing a properly defined fuzzy equivalence relation. A corresponding algorithm is introduced and various experiments concerning the proposed approach are given.

The more detailed structure of the work is as follows. Chapter 1 introduces some basic definitions for this monograph. Here are given such notions as: fuzzy sets, fuzzy relations, fuzzy equivalence relations, distance functions, measures of fuzziness, etc. Also a mathematical background is given in the Appendix. A short introduction to the classical fuzzy c-means clustering is given in the next chapter. Some next modifications are also considered. Chapter 3 presents the main results obtained in this work. Here a method for medical image segmentation by using a properly defined fuzzy equivalence relation is introduced. A corresponding algorithm is specified in the next Chapter 4.. An example implementation of this algorithm and various experiments concerning this approach are also given. In experiments, the proposed method demonstrated promising performance for various classes of medical images.

I should like to thank Prof. Marek Kurzyński of Wroclaw University of Technology for his helpful suggestions during this work. Also I should like to thank my brother Pawel for his competent and professional advice concerning the medical experiments included in the last section of this dissertation.

1. Basic notions

Chapter 1 introduces some basic definitions concerning the fuzzy mathematics. Here are given such notions as: fuzzy sets, fuzzy relations, fuzzy equivalence relations, distance functions, measures of fuzziness, etc. The most of the considered material is mainly under Mordeson, Malik and Cheng (2000). Also a mathematical background is given in the Appendix.

1.1 Fuzzy sets

Let X be a set and $A \subseteq X$. The *characteristic function* χ of A is defined as follows.
$\chi : X \to \{0,1\}$ such that: $\chi(x) =_{df}$ if $x \in A$ then 1 else 0 (for any $x \in X$).

Definition 1.1
A *fuzzy subset* μ of a set X is a function of X into the closed interval $[0,1]$. By $\mu(x) \in [0,1]$ we shall denote the *degree of membership* of $x \in X$ in μ. The set of all fuzzy subsets of X, i.e. the *fuzzy power set* of X is denoted by $\mathcal{FP}(X)$.

Let $X = \{x_1, \ldots, x_n\}$ is finite and μ is a fuzzy subset on X. Provided there is no ambiguity the following notion will be used sometimes: $\mu =_{df} \mu(x_1) / x_1 + \ldots + \mu(x_n) / x_n$. Also, the corresponding vector form will be used, i.e. $\mu =_{df} (\mu(x_1), \ldots, \mu(x_n))$ (if X is understood).

Definition 1.2
Let μ be a fuzzy subset of X and $\alpha \in [0,1]$. The *α-cut* wrt μ is the crisp set $\mu_\alpha =_{df} \{x \in X / \mu(x) \geq \alpha\}$. The *support* of μ is defined by the crisp set $\text{supp}(\mu) =_{df} \{x \in X / \mu(x) > 0\}$.

Definition 1.3
Let μ and ν be fuzzy subsets of a set X and $x \in X$. The *union* and the *intersection* of μ and ν are defined as follows: $(\mu \cup \nu)(x) =_{df} \max\{\mu(x), \nu(x)\}$, $(\mu \cap \nu)(x) =_{df} \min\{\mu(x), \nu(x)\}$. The *complement* of fuzzy subset μ is defined as: $\mu'(x) =_{df} 1 - \mu(x)$.

Any fuzzy set μ can be decomposed wrt the corresponding levels $\alpha \in [0,1]$, i.e. the following theorem is satisfied.

Theorem 1.1

$$\mu = \bigcup_{\alpha \in [0,1]} \alpha \cdot \mu_\alpha \ . \ \square$$

Next by μ^* we shall denote *the crisp set associated with* μ, i.e. $\mu^*(x) =_{df}$ if $\mu(x) \geq 0.5$ then 1 else 0 (for any $x \in X$).

1.2 Fuzzy relations and fuzzy equivalence relations

One can think of a *crisp relation* as representing the presence or absence of association or interaction between the elements of two or more sets. A generalisation can be obtained by using the notion of fuzzy subsets to allow for various degrees or strengths of dependency between elements. Let X and Y be sets. A *fuzzy relation* or *binary fuzzy relation* $\rho(X,Y)$ (or in short ρ), is the fuzzy subset of the product $X \times Y$, i.e. a function of $X \times Y$ into the closed interval $[0,1]$. Below we shall assume X and Y are finite. A convenient representation of any binary fuzzy relation $\rho(X,Y)$ is the corresponding *membership matrix* $M_\rho =_{df} [m_{ij}]_{|X| \times |Y|}$, where $m_{ij} =_{df} \rho(x_i,y_j)$ (for any pair $(x_i,y_j) \in X \times Y$).

Let $\mu(X,Y)$ and $\nu(Y,Z)$ be two fuzzy relations. The *max-min composition*[*] of μ and ν, i.e. $\mu \circ \nu$ produces a new fuzzy relation $\rho(X,Z)$ defined as follows: $\rho(x,z) =_{df} [\mu \circ \nu](x,z) = \max\{\min\{\mu(x,y),\nu(y,z)\} / y \in Y\}$. It can be shown that any such composition is associative but not commutative. So it can be generalised for more than two (but a finite set) of fuzzy relations. Compositions of binary fuzzy relations can be performed conveniently in terms of the well-known membership matrices of the relations. Next we shall say a given fuzzy relation ρ is *defined in* X if $\rho(X,X)$.

Definition 1.4

Let ρ be a fuzzy relation defined in X and $x,y,z \in X$. Then:

(i) ρ is *reflexive* iff $\rho(x,x) = 1$

(ii) ρ is *symmetric* iff $\rho(x,y) = \rho(y,x)$

(iii) ρ is *transitive (or max-min transitive)* iff $\rho(x,z) \geq \max\{\min\{\rho(x,y),\rho(y,z)\} / y \in X\}$.

---

[*] In general the following three different compositions can be used: $\sup\{\min\{\mu(x,y),\nu(y,z)\}/ y \in Y\}$, $\inf\{\max\{\mu(x,y),\nu(y,z)\}/ y \in Y\}$ and $\sup\{\{\mu(x,y) \cdot \nu(y,z)\}/ y \in Y\}$. Here '$\cdot$' denotes the usual arithmetic product. According to Mordeson, Malik and Cheng (2000) the max-min composition is assumed in this research (since the considered sets are finite, the supremum is interpreted as the logical operation maximum).

According to Definition 1.4 (iii) it follows that ρ is transitive iff $\rho(x,z) \geq \min\{\rho(x,y),\rho(y,z)\}$ (for any $x,y,z \in X$).

The *transitive (max-min) closure* ρ^+ of a given fuzzy relation ρ can be obtained by using the following simple algorithm.

Algorithm 1.1

(1) Let $\hat{\rho} =_{df} \rho \cup (\rho \circ \rho)$;

(2) If $\hat{\rho} \neq \rho$, set $\rho =_{df} \hat{\rho}$. Go to step (1);

(3) Stop: $\rho^+ =_{df} \hat{\rho}$. □

Example 1.1

Let $\rho =_{df} .6 / (a,b) + .4 / (a,c) + .5 / (b,c)$ is a fuzzy relation on $X =_{df} \{a,b,c\}$. According to Algorithm 1.1 we can obtain (the corresponding steps are identified):

(1) $\hat{\rho} = .6 / (a,b) + .5 / (a,c) + .5 / (b,c) \neq \rho$;

(2) Since $\hat{\rho} \neq \rho$ then $\rho =_{df} .6 / (a,b) + .5 / (a,c) + .5 / (b,c)$;

(1) $\hat{\rho} = .6 / (a,b) + .5 / (a,c) + .5 / (b,c)$;

(2) $\hat{\rho} = \rho$;

(3) $\rho^+ =_{df} .6 / (a,b) + .5 / (a,c) + .5 / (b,c)$. □

Any fuzzy relation ρ on X can be presented in a unique way by a labelled directed 1-graph $G =_{df} (X,U)$ called a *fuzzy graph* of ρ, where any edge $(x,y) \in U$ is labelled by corresponding $\rho(x,y) \in [0,1]$.

A fuzzy relation ρ defined in X is said to be a *similarity relation* iff ρ is reflexive and symmetric. Any similarity relation ρ is a *fuzzy equivalence relation* iff ρ is transitive.

Theorem 1.2

If ρ is a fuzzy similarity relation on X, where $|X| = n \in \mathcal{N}$, then $\rho^{(n-1)}$ is the max-min transitive closure of ρ (see below). □

Algorithm 1.2

The transitive closure, $\rho^+ = \rho^{(n-1)}$ can be determined by calculating the following relations iteratively.

Let $\rho^{(1)} =_{df} \rho$, $\rho^{(2)} =_{df} \rho \circ \rho$, etc. For any $k = 1,2, \ldots$ compute $\rho^{(2^k)} = \rho^{(2^{k-1})} \circ \rho^{(2^{k-1})}$, until no new relation is produced or until $2^k \geq n-1$. □

Consider a given fuzzy equivalence relation ρ. According to Definition 1.2 any α-cut ρ_α is a crisp equivalence relation on X. Let $\pi_\alpha =_{df} \pi(\rho_\alpha)$ be the *partition generated by* ρ_α. Clearly, two elements x and y belong to the same equivalence class under ρ_α iff $\rho(x,y) \geq \alpha$.

1.3 Distances and fuzziness measures

Let μ and ν be fuzzy subsets of a set X. The determination of similarity between μ and ν is a result of measuring a *distance* $d(\mu,\nu)$ wrt some space X. This distance is a *metric* iff it is nonnegative, reflexive, symmetric, and follows the triangle inequality. Any such distance may be measured in a variety of ways (Diday and Simon 1980, Mordeson, Malik and Cheng 2000, Wolfram S. et al. 2003). Some example distances $d(\mu,\nu)$ are illustrated below.

Hamming ("city block" or Manhattan) distance
$$d_1(\mu,\nu) =_{df} \sum_{x \in X} |\mu(x) - \nu(x)|$$

Euclidean distance
$$d_2(\mu,\nu) =_{df} \sqrt{\sum_{x \in X} (\mu(x) - \nu(x))^2}$$

Minkowski distance
$$d_3(\mu,\nu) =_{df} \left(\sum_{x \in X} |\mu(x) - \nu(x)|^q \right)^{\frac{1}{q}}, \; q \in \mathcal{R}_+$$

Canberra distance
$$d_4(\mu,\nu) =_{df} \sum_{x \in X} \frac{|\mu(x) - \nu(x)|}{\mu(x) + \nu(x)}$$

Chebyshev distance
$$d_5(\mu,\nu) =_{df} \max\{|\mu(x) - \nu(x)| / x \in X\}$$

The Hamming and Euclidean distances are particular cases of the Minkowski distance (for q = 1 and 2, respectively).

Let $\mathcal{R}_\bullet$ denote the set of all nonnegative real numbers, X be a set and $\mathcal{FP}(X)$ the fuzzy power set of X.

Definition 1.5

A *measure of fuzziness* is a mapping $\varphi : \mathcal{FP}(X) \rightarrow \mathcal{R}_\bullet$ such that $\varphi(\mu) \in \mathcal{R}_\bullet$ (for any fuzzy subset $\mu \in \mathcal{FP}(X)$).

A measure of fuzziness corresponds to some type of uncertainty, namely fuzziness (or vagueness). Some example measures of fuzziness $\varphi(\mu)$ are illustrated below.

Hamming ("city block" or Manhattan) measure of fuzziness
$$\varphi_1(\mu) =_{df} \frac{2d_1(\mu,\mu^*)}{|X|}$$

Euclidean measure of fuzziness
$$\varphi_2(\mu) =_{df} \frac{2d_2(\mu,\mu^*)}{\sqrt{|X|}}$$

Minkowski measure of fuzziness
$$\varphi_3(\mu) =_{df} \frac{2d_3(\mu,\mu^*)}{\sqrt[q]{|X|}}, \quad q \in \mathcal{R}_+$$

Canberra measure of fuzziness
$$\varphi_4(\mu) =_{df} \frac{2d_4(\mu,\mu^*)}{|X|}$$

Chebyshev measure of fuzziness
$$\varphi_5(\mu) =_{df} 2d_5(\mu,\mu^*)$$

Complement indistinguishability of fuzzy sets measure
(the use of Hamming distance for a measure of the fuzziness of a fuzzy subset by the lack of its distinction from its complement)
$$\varphi_6(\mu) =_{df} \sqrt[n]{|X|} - \left(\sum_{x\in X} |2\mu(x)-1|^n\right)^{\frac{1}{n}}, \quad n \in \mathcal{N}$$

Entropy of fuzzy sets measure
$$\varphi_7(\mu) =_{df} -\sum_{x\in X}\left(\mu(x)\lg_2 \mu(x) + \mu'(x)\lg_2 \mu'(x)\right),$$
$$H(\mu) =_{df} \varphi_7(\mu), \text{ where } 0 < \mu(x) < 1 \text{ (for any } x \in X)$$

The Hamming and Euclidean measures are particular cases of the Minkowski measure of fuzziness (for $q = 1$ and 2, respectively).

Example 1.2

Let $\mu =_{df} .6 / a + .4 / b + .5 / c$ be a fuzzy set on $X =_{df} \{a,b,c\}$. The following measures of fuzziness are obtained: $\varphi_1(\mu) = .866_{(6)} \approx .867$, $\varphi_2(\mu) = .871779... \approx .872$, $\varphi_3(\mu) / _{q=1.5} = .869162... \approx .869$, $\varphi_4(\mu) = 1.055_{(5)} \approx 1.056$, $\varphi_5(\mu) = 1.000$, $\varphi_6(\mu) / _{n=2} = 1.44920... \approx 1.449$, and $\varphi_7(\mu) = H(\mu) = 2.941901... \approx 2.942$. □

A short introduction to the classical fuzzy c-means clustering is given in the next chapter.

2. Fuzzy c-means clustering

We first consider the classical fuzzy c-means clustering method. Some next modifications (extensions or improvings) are presented in the second section. The most of the considered material is mainly under Mordeson, Malik and Cheng (2000) and Suri, Setarehdan and Singh (2002).

2.1 The classical fuzzy c-means clustering method

Clustering plays a fundamental role in searching for structures in data. The problem of clustering in a finite set of data X is to find several cluster centers that correctly characterise relevant classes of X. These classes must form a partition of X in such a way that the degree of association is strong for data within members of the partition and weak for data in different members. This requirement is often too strong in practical applications. Consequently, it is desirable to replace it with a weaker requirement. A weaker requirement which can be used is that of a fuzzy partition or a fuzzy pseudopartition on X. The problem is then referred to as *fuzzy clustering*. Fuzzy pseudopartitions are often called fuzzy c-partitions, where c designates the number of fuzzy classes in the partition. A basic method of fuzzy clustering is based on fuzzy c-partitions and is called a *fuzzy c-means clustering method* (Bezdek 1981). The following is a brief review of this method.

Definition 2.1

Let $X =_{df} \{x_1, x_2, \ldots, x_n\}$ be a set of given data. A *fuzzy pseudopartition* or *fuzzy c-partition* of X is a family of fuzzy subsets of X, denoted by $\mathcal{P} =_{df} \{\mu_1, \mu_2, \ldots, \mu_c\}$ such that:

$$\sum_{i=1}^{c} \mu_i(x_k) = 1 \text{ (for all } k = 1,2, \ldots ,n) \text{ and } 0 < \sum_{k=1}^{n} \mu_i(x_k) < n \text{ (for all } i = 1,2, \ldots ,c),$$

where c is a positive integer.

Example 2.1

Let $\mu_1 =_{df} .6 / x_1 + .3 / x_2 + .8 / x_3$ and $\mu_2 =_{df} .4 / x_1 + .7 / x_2 + .2 / x_3$ are two fuzzy subsets on $X =_{df} \{x_1, x_2, x_3\}$. Then $\mathcal{P} = \{\mu_1, \mu_2\}$ is a fuzzy pseudopartition or fuzzy 2-partition of X, where $n = 3$ and $c = 2$. This pseudopartition corresponds toa partition in the usual sense by assuming $\mu_i =_{df} \mu_i^*$ $(i = 1,2)$. □

Assume now that X, in general, is a set of vectors, i.e. $X =_{df} \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, where $\mathbf{x}_k =_{df} (x_{k1}, x_{k2}, \ldots, x_{kp}) \in \mathcal{R}^p$ (for all $k = 1,2, \ldots, n$). The problem of fuzzy clustering is to find pseudopartition $\mathcal{P}$ and the associated cluster centers by which the structure of the data is represented as good as possible. This requires some criterion expressing the general idea that associations be strong within clusters and weak between them. The last criterion can be introduced in terms of a performance index of $\mathcal{P}$, based upon cluster centers. The c cluster centers, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_c$ associated with a given pseudopartition $\mathcal{P} =_{df} \{\mu_1, \mu_2, \ldots, \mu_c\}$ are calculated as follows:

$$\mathbf{v}_i =_{df} \frac{\sum_{k=1}^{n} (\mu_i(\mathbf{x}_k))^m \mathbf{x}_k}{\sum_{k=1}^{n} (\mu_i(\mathbf{x}_k))^m} \quad \text{(for all } i = 1,2, \ldots, c),$$

where $m > 1$ is a real number that governs the influence of membership grades. The *performance index* PI_m of $\mathcal{P}$ is then defined as follows:

$$PI_m =_{df} \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_i(\mathbf{x}_k))^m \left\| \mathbf{x}_k - \mathbf{v}_i \right\|^2,$$

where $\| \cdot \|$ is some inner product-induced norm metric in $\mathcal{R}^p$ and $\left\| \mathbf{x}_k - \mathbf{v}_i \right\|^2$ the distance between $\mathbf{x}_k$ and $\mathbf{v}_i$. Hence, the performance index PI_m measures the weighted sum of distances between cluster centers and elements in the corresponding fuzzy clusters. The smaller value of PI_m will implicate a better fuzzy pseudopartition. Consequently, the clustering problem is an optimisation problem. This problem can be specified as follows: find $\mathcal{P}^*$ such that $PI_m(\mathcal{P}^*) =_{df} \min\{ PI_m(\mathcal{P}) \mid \mathcal{P} \subseteq \mathcal{FP}(X) \}$.

For simplicity, next the following designations will be used: $u_{ik} =_{df} \mu_i(\mathbf{x}_k)$ and $d_{ik} =_{df} \left\| \mathbf{x}_k - \mathbf{v}_i \right\|$ (for any $i = 1, 2, \ldots, c$ and $k = 1, 2, \ldots, n$). Hence, the i^{th} fuzzy subset μ_i has the form $U_i =_{df} \{u_{i1}, u_{i2}, \ldots, u_{in}\}$. The c fuzzy sets form the rows of a *fuzzy pseudopartition matrix* $U = [u_{ik}]_{c \times n}$.

The fuzzy c-means clustering algorithm (in short: FCM algorithm) iteratively optimises the objective function $PI_m =_{df} \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m (d_{ik})^2$ to approximate its minimums. This leads to the clustering of the given data set into c fuzzy subsets. A version of the classical FCM algorithm is illustrated below (Clark 1994).

Algorithm 2.1 (FCM algorithm)

(1) Let t = 0. Select an initial fuzzy pseudopartition matrix $U^{(0)}$;

(2) Calculate the c cluster centers $\mathbf{v}_i^{(t)}$ for $U^{(t)}$ and the chosen value of m:

$$v_{ij}^{(t)} =_{df} \frac{\sum_{k=1}^{n} (u_{ik}^{(t)})^m x_{kj}}{\sum_{k=1}^{n} (u_{ik}^{(t)})^m}, \quad j = 1, 2, \ldots, p$$

(3) Update $U^{(t+1)}$ by the following procedure. For each $\mathbf{x}_k \in X$, if $d_{ik}^{(t)} \neq 0$ for all $i = 1, 2, \ldots, c$ then define

$$u_{ik}^{(t+1)} =_{df} \frac{1}{\sum_{s=1}^{c} \left(\frac{d_{ik}^{(t)}}{d_{sk}^{(t)}} \right)^{\frac{2}{m-1}}}$$

else, define $u_{ik}^{(t+1)}$ for $i \in I_k^{(t)} =_{df} \{s \in \{1, 2, \ldots, c\} / d_{sk}^{(t)} = 0\}$ by any nonnegative real numbers satisfying $\sum_{i \in I_k^{(t)}} u_{ik}^{(t+1)} = 1$, and define $u_{ik}^{(t+1)} = 0$ for $i \in \{1, 2, \ldots, c\} - I_k^{(t)}$;

(4) Compare $U^{(t)}$ and $U^{(t+1)}$, where

$$\Delta U =_{df} \sum_{i=1}^{c} \sum_{k=1}^{n} \left\| u_{ik}^{(t+1)} - u_{ik}^{(t)} \right\|.$$

If $\Delta U \leq \varepsilon$, stop; otherwise, set $t =_{df} t + 1$. Go to (2). □

According to step (3) of the above algorithm any distance $d_{rk}^{(t)} =_{df} \left\| \mathbf{x}_k - \mathbf{v}_r^{(t)} \right\|$ (for any r and k). The stop condition for ΔU in step (4) can be modified by using such matrix norms as: the Frobeius F-norm $\left(\sum_{i=1}^{c} \sum_{k=1}^{n} \left| u_{ik}^{(t+1)} - u_{ik}^{(t)} \right| \right)^{\frac{1}{2}}$ or also the Chebyshev norm $\max\{ \left| u_{ik}^{(t+1)} - u_{ik}^{(t)} \right| / i = 1, \ldots, c ; k = 1, \ldots, n\}$, etc. (e.g. see: Sheen 2001).

Example 2.2

Consider the fuzzy subsets $\mu_1 =_{df} .6 / \mathbf{x}_1 + .3 / \mathbf{x}_2 + .8 / \mathbf{x}_3$ and $\mu_2 =_{df} .4 / \mathbf{x}_1 + .7 / \mathbf{x}_2 + .2 / \mathbf{x}_3$ on X, where it is assumed that $\mathbf{x}_k \in \mathcal{R}^2$ are vectors (for k = 1,2,3) and p = 2. According to Example 2.1 $\{\mu_1, \mu_2\}$ is a fuzzy pseudopartition of $X =_{df} \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$, where

$n = 3$ and $c = 2$. Also assume that $m = 2$, $\varepsilon = .35$, and $\|\cdot\|$ is the Euclidean distance.

The data set X consisting of 3 points (i.e. vectors) is as follows:

$$X =_{df} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 1 & 3 \\ 4 & 2 \end{bmatrix}.$$

According to Algorithm 2.1 we can obtain (the corresponding steps are identified):

(1) The initial fuzzy pseudopartition matrix

$$U^{(0)} = \begin{matrix} \mu_1 \\ \mu_2 \end{matrix} \begin{bmatrix} .60 & .30 & .80 \\ .40 & .70 & .20 \end{bmatrix};$$

(2) The obtained cluster centers matrix

$$V^{(0)} =_{df} \begin{matrix} \mathbf{v}_1^{(0)} \\ \mathbf{v}_2^{(0)} \end{matrix} \begin{bmatrix} 2.43 & 2.08 \\ .94 & 2.71 \end{bmatrix};$$

(3) Since the distance matrix

$$D^{(0)} =_{df} \begin{bmatrix} 2.43 & 1.79 & 2.43 \\ 1.17 & 1.71 & 3.14 \end{bmatrix}$$

then each $d_{ik}^{(0)} \neq 0$ and the updated pseudopartition matrix

$$U^{(1)} = \begin{bmatrix} .19 & .48 & .63 \\ .81 & .52 & .37 \end{bmatrix};$$

(4) We have: $\Delta U = 1.52^{*} > .1$ and $t = 1$. Assuming the F-norm we can obtain $\sqrt{\Delta U} = 1.23$. For $\Delta U =_{df} \max\{ \left| u_{ik}^{(t+1)} - u_{ik}^{(t)} \right| / i = 1, 2 ; k = 1, 2, 3\}$ we have $\Delta U = .41$ (this norm is used below) ;

(2) $$V^{(1)} = \begin{matrix} \mathbf{v}_1^{(1)} \\ \mathbf{v}_2^{(1)} \end{matrix} \begin{bmatrix} 2.74 & 2.35 \\ .77 & 2.25 \end{bmatrix};$$

(3) $$D^{(1)} = \begin{bmatrix} 2.76 & 1.86 & 1.30 \\ .81 & .78 & 3.24 \end{bmatrix}$$ and

$$U^{(2)} = \begin{bmatrix} .08 & .15 & .86 \\ .92 & .85 & .14 \end{bmatrix};$$

(4) $\Delta U = .33 < .35$.

---

* Any scalar can be considered as one-component vector. Then the notions of vector norm and its absolute value coincide.

The algorithm stops for t = 1. The obtained fuzzy pseudopartition is given by $U^{(2)}$. We have the following two cluster centers: $\mathbf{v}_1$ = (2.74, 2.35) and $\mathbf{v}_2$ = (.77, 2.25).

Example medical image segmentation under the above FCM algorithm is illustrated in Appendix B. □

The algorithm is based on the assumption that the desired number of clusters c is given and, in addition, a particular distance, a real number m ∈ (1,∞), and a small positive real ε, serving as a stopping criterion, are chosen. In general, the number of iterations of the FCM algorithm will increase for decreasing values of ε and also will depend on the used initial fuzzy pseudopartition matrix $U^{(0)}$ as well as on the used definition for ΔU in step (4). Moreover, the computational complexity of this algorithm is depending on the used values for c and n. For example, the initial fuzzy pseudopartition matrix $U^{(0)}$ can be initialised as follows (Clark 1994): $u_{ik}^{(0)} =_{df}$ if (k – 2i) modulo (2c) = 0 then 1 else 0 (for any i = 1, ... ,c and k = 1, ... ,n). Some next modifications of the FCM algorithm are considered below.

2.2 Modified approaches

Many research works have been proposed to steep up this classical approach, e.g. multi – modality fuzzy c – means (Bloch I. 1994), using a performance index of a fuzzy pseudopartition (Klir and Yuan 1995), a modified fuzzy c – means algorithm by using of extended Xie-Beni index (Xie and Beni 1991, Pal and Bezdek 1995), conditional fuzzy c – means algorithm, i.e. using some conditional variables (Pedrycz 1996), pixel resistance and neighbourhood oriented approach (Mohamed, Ahmed and Farag 1998), density-weighted fuzzy c-means (Chen and Wang 1999), improved robust fuzzy clustering (Melek, Emami and Goldenberg 1999), modified fuzzy c-means (Thitimajshima 2000), fuzzy clustering with ε - insensitive loss function (Łęski 2001),adaptive fuzzy c – means (e.g. see: Suri, Setarehdan and Singh 2002), etc. Most of the methods are based on the fuzziness of the partition, e.g. $u_{ik} =_{df} \frac{1}{1+e^{md_{ik}-s}}$, where s is a threshold value at the level .5 of μ (Bloch I. 1994).

The *extended Xie-Beni index* (Xie and Beni 1991, Pal and Bezdek 1995) is given by

$$PI_{X-B,m} =_{df} \frac{\sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_i(\mathbf{x}_k))^m \left\|\mathbf{x}_k - \mathbf{v}_i\right\|}{n\cdot[\min\{\left\|\mathbf{v}_j - \mathbf{v}_i\right\|^2 / i \neq j; i,j = 1,...,c\}]}.$$

The Xie – Beni and the extended Xie-Beni indices are the most reliable methods for discovering the number of clusters. Other methods, which may be used in the future, are measures of cluster typicality (Mordeson, Malik and Cheng 2000).

The *conditional FCM algorithm* (Pedrycz 1996) is based on a conditional variable b_k associated with the k^{th} data point $\mathbf{x}_k$, where $\sum_{i=1}^{c} u_{ik} = b_k$ (k = 1, ... ,n). The cluster membership constraint is now conditioned on the membership values of the variables b_k. The partition update equation (i.e. step (3) of the FCM algorithm) is modified as follows:

$$u_{ik}^{(t+1)} =_{df} \frac{b_k}{\sum_{s=1}^{c}\left(\frac{d_{ik}^{(t)}}{d_{sk}^{(t)}}\right)^{\frac{2}{m-1}}} .$$

This algorithm helps to reveal structure in data based on rule of the form: "If the patient is deficient in folate, then cluster the data".

Consider a given image $\mathcal{I}$. Let $\mathbf{x}_k \in \mathcal{I}$ is a pixel. In the FCM algorithm the clustering of $\mathbf{x}_k$ with class i depends on the membership value u_{ik} .If we consider a noisy image, the FCM does not have a method to overcome this problem. The *modified FCM algorithm* (Mohamed, Ahmed and Farag 1998) is based on an interpretation of any distance $d_{ik} =_{df} \| \mathbf{x}_k - \mathbf{v}_i \|$ as the *resistance of pixel* $\mathbf{x}_k$ to be clustered with class i. This resistance can be tolerated by the neighbouring pixels $\mathbf{x}_j$. The neighbouring pixels work to decrease the pixel's resistance by a fraction that depends on the membership value of $\mathbf{x}_j$ with cluster i, u_{ij} . Let J_k be the set of all neighbours associated with $\mathbf{x}_k$ and $p_{kj} =_{df} \| k - j \|$ be the *relative location between the two pixels* $\mathbf{x}_k$ and $\mathbf{x}_j$. So, any p_{kj} measures the proximity of pixel $\mathbf{x}_k$ to its neighbour pixel $\mathbf{x}_j$ ($\mathbf{x}_k , \mathbf{x}_j \in \mathcal{I}$). Hence, the modified value of any distance d_{ik} is defined as follows:

$$\hat{d}_{ik} =_{df} d_{ik} \cdot \left(1 - \alpha \cdot \frac{\sum_{j \in J} u_{ij} \cdot p_{kj}}{\sum_{j \in J} p_{kj}}\right),$$

where $\alpha \in [0,1]$ is a constant (a convenient result can be obtained assuming $\alpha = .5$). For $\alpha = 0$ we have the original FCM algorithm (without considering the neighbours).

The total effect of the neighbouring pixels tries to pull its neighbour toward its class without neglecting the effect of the pixel itself. This works as an adaptive filter applied during the segmentation. Each iteration of the modified FCM enhances the filtering effect while emphasising the edges.

In the *adaptive FCM algorithm*, the objective function PI_m is decomposed into three components using some finite difference operators (similar to derivatives in the continuous domain). As some regularisation terms two additional coefficients are used (to penalise a large amount of variation and also the corresponding multiplier fields having discontinuities).These two terms are set on the basis of image inhomogeneity. The multiplier field $m = m(\mathbf{x}_k)$ is a slowly varying function ($\mathbf{x}_k \in \mathcal{J}$).The updated adaptive FCM preserves the advantages of FCM while being applicable to images with intensity homogeneities (Pham and Prince 1999).

The major advantage of the fuzzy c-means clustering technique lies in the ease of implementation. The major weaknesses of this method are: (1) The algorithm is not fast enough to be implemented for real-time applications. (2) The performance of the algorithm depends upon a large number of user parameters, such as error threshold and number of iterations. In particular (as it was shown in Example 2.2) instead of the Euclidean norm, some other stop conditions could be used, e.g. such as the Frobeius F-norm, the Chebyshev norm, etc. Since any pseudopartition should correspond to some partition, the correct realisation of any such correspondence may be a non-trivial problem (e.g. if $u_{ik} = .5$, for some i and k). (3) The choice of the initial cluster is important and needs to be carefully selected. (4) The algorithm is not very robust to noise, spatial and temporal inhomogeneities (Suri, Setarehdan and Singh 2002).

The FCM sensitivity to initialisation has been also criticised by Frigui and Krishnapuram (1999). Although the fuzzy c-means algorithm is powerful in image segmentation, there is still a drawback encountered, namely the desired number of clusters should be specified in advance. This is a disadvantage whenever the clustering problem cannot specify any desired number of clusters. The situations are often for remotely sensed image segmentation, because the ground truth is always not available for these images (Mordeson, Malik and Cheng 2000, Li, Dong and Gao 2002).

3. Medical image segmentation using a fuzzy equivalence relation

The fuzzy c-means method requires that the desired number of clusters be given in advance. This can be problematic when the clustering problem does not specify any desired number of clusters. The number of clusters should reflect the structure of the given data. Methods based on fuzzy equivalence relations satisfy this need. This chapter presents the main results obtained in this work. Here a new method for medical image segmentation by using a properly defined fuzzy equivalence relation is introduced. This research is a continuation of the introductory work (Tabakow 2001).

3.1 The image segmentation model

Image segmentation is essentially a process of pixel classification, wherein the image pixels are segmented into subsets by assigning the individual pixels to classes. These segmented organs and their boundaries are very critical in the quantification process for physicians and medical surgeons, in any branch of medicine, which deals with imaging. Hence, any such segmentation is a process of partitioning an image into some regions (or classes) such that each region is homogeneous and none of the union of two adjacent regions is homogeneous. A formal definition of the image segmentation notion is given below.

Definition 3.1

Let $\mathcal{J}$ be a given finite set of all image pixels, $I =_{df} \{I_1, I_2, \ldots, I_k\}$ be a finite family of connected nonempty subsets, i.e. clusters of $\mathcal{J}$ and $A(\cdot)$ be one argument homogeneity predicate defined over clusters of connected pixels. Then *image segmentation* is a partition of the set $\mathcal{J}$ into a finite family of connected nonempty subsets of $\mathcal{J}$ (i.e. clusters) $\{I_1, I_2, \ldots, I_k\}$, such that:

(i) $I_r \cap I_s = \varnothing$, for $r \neq s$ ($r,s \in \{1,2, \ldots ,k\}$),

(ii) $\bigcup_{s=1}^{k} I_s = \mathcal{J}$,

(iii) $A(I_s) = true$ (for any $I_s \subseteq \mathcal{J}$), and

(iv) $A(I_r \cup I_s) = false$ (for any two different neighbours I_r and I_s).

The segmentation process described below assumes the using of a properly defined fuzzy equivalence relation. The specification of such a relation is given in the next section.

3.2 Specification of a fuzzy equivalence relation

A fuzzy equivalence relation induces a crisp equivalence and hence a crisp partition of each of its α-cuts. Thus the fuzzy clustering problem can be thought as the problem of identifying an appropriate fuzzy equivalence relation on the given data. This cannot usually be done directly, but a fuzzy similarity relation can be determined in terms of an appropriate distance function applied to the given data. Then the transitive closure of this fuzzy similarity relation is the appropriate fuzzy equivalence relation.

Let $X =_{df} \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\} \subset \mathcal{R}^p$ be a set of data. The fuzzy relation ρ on X can be defined in terms of an appropriate distance function of the Minkowski class as follows.

Definition 3.2

Let $\rho : X \times X \rightarrow [0,1]$ such that $\rho(\mathbf{x}_i,\mathbf{x}_k) =_{df} 1 - \delta \cdot \left(\sum_{j=1}^{p} \left| x_{ij} - x_{kj} \right|^q \right)^{\frac{1}{q}}$ (for any $\mathbf{x}_i,\mathbf{x}_k \in X$),

where $q \in \mathcal{R}_+$ and δ is the reciprocal of the largest distance in X, i.e. $\delta =_{df} \dfrac{1}{\max\left\{ \left(\sum_{j=1}^{p} \left| x_{ij} - x_{kj} \right|^q \right)^{\frac{1}{q}} / \mathbf{x}_i,\mathbf{x}_k \in X; \mathbf{x}_i \neq \mathbf{x}_k \right\}}$. Then ρ is said to be a *distance function of the Minkowski class*.

According to Definition 3.2 for any two vectors $\mathbf{x},\mathbf{y} \in X$ we have: $\rho(\mathbf{x},\mathbf{x}) = 1$ and $\rho(\mathbf{x},\mathbf{y}) = \rho(\mathbf{y},\mathbf{x})$. Hence, the following proposition is satisfied.

Proposition 3.1

Let ρ be a distance function of the Minkowski class defined in X. Then ρ is a fuzzy similarity relation. □

Distance functions of the Hamming and Euclidean classes are particular cases of the Minkowski class (for $q = 1$ and 2, respectively). Some other classes of such functions can be also obtained, e.g. a *distance function of the Canberra class* can be introduced as follows:

$\sigma(\mathbf{x}_i, \mathbf{x}_k) =_{df} 1 - \gamma \cdot \sum_{j=1}^{p} \left| \frac{x_{ij} - x_{kj}}{x_{ij} + x_{kj}} \right|$, where γ is the reciprocal of the largest distance in X, i.e.

$$\gamma =_{df} \frac{1}{\max\{\sum_{j=1}^{p} \left| \frac{x_{ij} - x_{kj}}{x_{ij} + x_{kj}} \right| \; / \; \mathbf{x}_i, \mathbf{x}_k \in X; \mathbf{x}_i \neq \mathbf{x}_k\}} .$$

Proposition 3.2

Let σ be a distance function of the Canberra class defined in X. Then σ is a fuzzy similarity relation. □

Definition 3.3

Let ρ and σ be two distance functions of the Minkowski and Canberra classes, respectively. The *linear convex combination* of ρ and σ, denoted by $\lambda_{(\rho,\sigma)}$ or λ (if ρ and σ are understood), is a new distance function defined as follows: $M_\lambda =_{df} \alpha \cdot M_\rho + \beta \cdot M_\sigma$, where $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. We shall say that $\lambda_{(\rho,\sigma)}$ is a *compound distance function* wrt ρ and σ.

The similarity property for $\lambda_{(\rho,\sigma)}$ is preserved, i.e. the following proposition is satisfied.

Proposition 3.3

$\lambda_{(\rho,\sigma)}$ is a fuzzy similarity relation.

Proof:

Let $\lambda_{(\rho,\sigma)}$ be given. According to Definition 3.3 we have $M_\lambda =_{df} \alpha \cdot M_\rho + \beta \cdot M_\sigma$, where $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. According to Propositions 3.1 and 3.2 it follows that the corresponding membership matrices M_ρ and M_σ are symmetric. Assume that $M_\rho =_{df} [a_{ij}]_{n \times n}$ and $M_\sigma =_{df} [b_{ij}]_{n \times n}$, where $n =_{df} |X|$. Also let $M_\lambda =_{df} [c_{ij}]_{n \times n}$. We have: $c_{ij} = \alpha \cdot a_{ij} + \beta \cdot b_{ij} = \alpha \cdot a_{ji} + \beta \cdot b_{ji} = c_{ji}$ and $c_{ii} = \alpha \cdot a_{ii} + \beta \cdot b_{ii} = \alpha \cdot 1 + \beta \cdot 1 = \alpha + \beta = 1$ (for any i,j = 1,2, ... ,n). □

Obviously, the previous two classes can be obtained for $(\alpha,\beta) = (1,0)$ and $(\alpha,\beta) = (0,1)$, respectively. In a natural way the last proposition can be generalised for more than two distance functions. In general, for more than two, e.g. $k \geq 3$, distance functions, we can

obtain: $M_\lambda =_{df} \sum_{i=1}^{k} \alpha_i \cdot M_{\rho_i}$, where $\alpha_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i = 1$ (a more formal treatment is omitted). Some example fuzzy similarity relations are given below.

Example 3.1

(1) Let ρ be a distance function of the Euclidean class (i.e. q = 2) and X be the data set considered in Example 2.2, where p = 2 (*see Section 2.1*). Since $\delta = \frac{1}{\max\{\sqrt{2},\ 4,\ \sqrt{10}\}}$ $= \frac{1}{4}$, we can obtain: $\rho(\mathbf{x}_1,\mathbf{x}_2) = 1 - \frac{\sqrt{2}}{4} = .64644... \approx .65$, $\rho(\mathbf{x}_1,\mathbf{x}_3) = 1 - \frac{4}{4} = 0$ and $\rho(\mathbf{x}_2,\mathbf{x}_3) = 1 - \frac{\sqrt{10}}{4} = .20943... \approx .21$. The *membership matrix* of the obtained fuzzy similarity relation ρ, denoted by M_ρ, is as follows: $M_\rho = \begin{bmatrix} 1 & .65 & 0 \\ .65 & 1 & .21 \\ 0 & .21 & 1 \end{bmatrix}$.

(2) Assume now that σ is a distance function of the Canberra class. For the same X and p, we have: $\gamma = \frac{1}{\max\{\frac{6}{5},\ 1,\ \frac{4}{5}\}} = \frac{5}{6}$. Hence: $\sigma(\mathbf{x}_1,\mathbf{x}_2) = 1 - \frac{5}{6}\cdot\frac{6}{5} = 0$, $\sigma(\mathbf{x}_1,\mathbf{x}_3) = 1 - \frac{5}{6}\cdot 1 = \frac{1}{6} = .166_{(6)} \approx .17$ and $\sigma(\mathbf{x}_2,\mathbf{x}_3) = 1 - \frac{4}{5}\cdot\frac{5}{6} = \frac{2}{6} = \frac{1}{3} =$ $.333_{(3)} \approx .33$. The following membership matrix is obtained: $M_\sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & .33 \\ .17 & .33 & 1 \end{bmatrix}$.

(3) Let now $\alpha = \beta = .5$. The following distance function can be introduced (after rounding off):

$$M_\lambda = \begin{bmatrix} 1 & .33 & .09 \\ .33 & 1 & .27 \\ .09 & .27 & 1 \end{bmatrix}.$$

(4) Let $M_\lambda =_{df} \alpha_1 \cdot M_{\rho_1} + \alpha_2 \cdot M_{\rho_2} + \alpha_3 \cdot M_\sigma$, where $M_{\rho_1} =_{df} M_\rho$ for q = i $\in \{1,2\}$. Here M_{ρ_2} and M_σ are the above two fuzzy relations, as in (1) and (2), and M_{ρ_1} is the

corresponding distance function of the Hamming class: $M_{\rho_1} = \begin{bmatrix} 1 & .5 & 0 \\ .5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Let $\alpha_1 = \alpha_2 = \frac{1}{4}$ and $\alpha_3 = \frac{1}{2}$. Then we can obtain: $M_\lambda = \begin{bmatrix} 1 & .29 & .09 \\ .29 & 1 & .22 \\ .09 & .22 & 1 \end{bmatrix}$. □

The above defined binary relations ρ, σ and λ are fuzzy similarity relations, but not necessary a fuzzy equivalence relations, e.g. according to the last example we have: $M_{\hat{\rho}} = \begin{bmatrix} 1 & .65 & .21 \\ .65 & 1 & .21 \\ .21 & .21 & 1 \end{bmatrix} \neq M_\rho$ and ρ is not transitive. Similarly, σ and λ (considered in (3) and (4)) are not transitive. Hence, the transitive closure must often be determined. This can be done by Algorithm 1.1 or Algorithm 1.2 (*see Section 1.2*). This latter algorithm is more efficient than the former since ρ, σ and λ are some similarity relations. Without loss of generality, only distance functions of the Minkowski class are assumed below.

3.3 Fuzzification of the input space

The first step of any fuzzy applications is the so-called 'fuzzification' of the considered input space. Fuzzification does mean that we assign the image (its grey levels, features, segments, etc.) with one or more membership values regarding to the interesting properties (e.g. brightness, homogeneity, noisiness, edginess etc.). After transformation of image into the so-called membership plane (*fuzzification*), a suitable fuzzy approach aggregates and/or modifies the membership values. To achieve new results (modified grey-levels, segmented image regions, classified objects, etc.), the output of membership plane should be decoded (*defuzzification*). It means that the membership values are retransform into the grey level plane. Generally, there are three ways for image fuzzification: *histogram - based grey level fuzzification*, *local fuzzification*, and *feature fuzzification* (Tizhoosh 1998).

Any such fuzzification can be realised by means of various techniques. Next we shall concentrate our attention to the histogram based grey level fuzzification approach. Below a histogram based fuzzy equivalence relation segmentation process is described. Also, without loss of generality, only monochromatic (i.e. grey scale) images are analysed.

Consider a given image $\mathcal{I}_0$ in terms of its grey levels. Any such image can be presented as a composition of the following three basic classes of colours: *dark*, *grey* and *bright* (in short: *d - class*, *g – class* and *b – class*, respectively). For each such class, an appropriate membership function can be introduced wrt the grey level scale gradation of $\mathcal{I}_0$. A formal definition of the corresponding membership functions is given below. Also the above g – class is decomposed into three different subclasses, denoted as g_1 *– class*, g_2 *– class* and g_3 *– class* (below the corresponding membership functions are denoted by μ_{grey1}, μ_{grey2} and μ_{grey3}, respectively). Hence, the whole grey level scale gradation of $\mathcal{I}_0$ is assumed to be a composition of the corresponding five fuzzy sets (an extension of this analysis is Section 4.2 of Chapter 4).

Definition 3.4

The *dark class*, *grey class* and *bright class membership functions*, denoted by μ_{dark}, μ_{grey1}, μ_{grey2}, μ_{grey3} and μ_{bright} respectively, are defined as follows (a graphical representation of the last membership functions is shown in Figure 3.1 below):

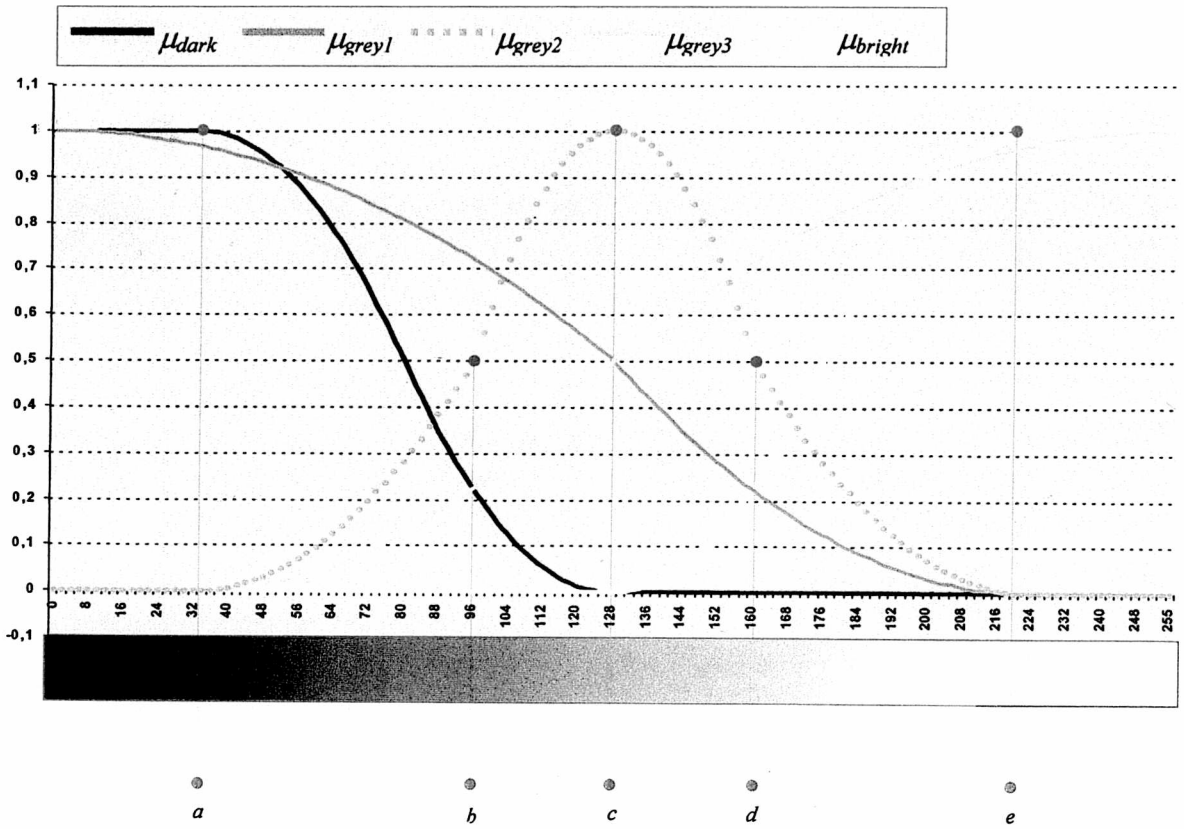


Figure 3.1 A graphical representation of μ_{dark}, μ_{grey1}, μ_{grey2}, μ_{grey3} and μ_{bright}

where a, b, c, d, e $\in$ [0,255] are some expert-estimated parameters, also μ : [0,255] $\rightarrow$ [0,1] (for any $\mu \in \{\mu_{dark}, \mu_{grey1}, \mu_{grey2}, \mu_{grey3}, \mu_{bright}\}$). The determination of the last parameters can be also realised in an automated way (a fuzzy expected value is used: see Definition 4.1 of Section 4.1).

$$(1) \qquad \mu_{dark}(x) =_{df} \begin{cases} 1 & \text{for } 0 \le x \le a, \\ 1 - 2 \cdot \left(\frac{x-a}{c-a}\right)^2 & \text{for } a < x \le \frac{a+c}{2}, \\ 2 \cdot \left(\frac{x-c}{c-a}\right)^2 & \text{for } \frac{a+c}{2} < x < c, \\ 0 & \text{for } c \le x \le 255. \end{cases}$$

$$(2) \qquad \mu_{grey1}(x) =_{df} \begin{cases} 1 - \frac{1}{2}\left(\frac{x}{c}\right)^2 & \text{for } 0 \le x \le c, \\ \frac{1}{2} \cdot \left(\frac{x-e}{e-c}\right)^2 & \text{for } c < x < e, \\ 0 & \text{for } e \le x \le 255. \end{cases}$$

$$(3) \qquad \mu_{grey2}(x) =_{df} \begin{cases} 0 & \text{for } 0 \le x \le a \ \vee \ e \le x \le 255, \\ \frac{1}{2} \cdot \left(\frac{x-a}{b-a}\right)^2 & \text{for } a < x < b, \\ 1 - \frac{1}{2} \cdot \left(\frac{x-c}{c-b}\right)^2 & \text{for } b \le x \le c, \\ 1 - \frac{1}{2} \cdot \left(\frac{x-c}{d-c}\right)^2 & \text{for } c < x \le d, \\ \frac{1}{2}\left(\frac{x-e}{e-d}\right)^2 & \text{for } d < x < e. \end{cases}$$

$$(4) \qquad \mu_{grey3}(x) =_{df} \begin{cases} 0 & \text{for } 0 \le x \le a, \\ \frac{1}{2} \cdot \left(\frac{x-a}{c-a}\right)^2 & \text{for } a < x < c, \\ 1 - \frac{1}{2} \cdot \left(\frac{x-255}{255-c}\right)^2 & \text{for } c \le x \le 255. \end{cases}$$

$$(5)\qquad \mu_{bright}(x) =_{df} \begin{cases} 0 & for\ \ 0 \le x \le c, \\ 2\cdot\left(\dfrac{x-c}{e-c}\right)^2 & for\ \ c < x \le \dfrac{c+e}{2}, \\ 1-2\cdot\left(\dfrac{x-e}{e-c}\right)^2 & for\ \ \dfrac{c+e}{2} < x < e, \\ 1 & for\ \ e \le x \le 255. \end{cases}$$

3.4 Specification of the initial membership matrix

Let $\mathcal{I}_0$ be a given image and $\mathcal{M} =_{df} \{\mu_{dark}, \mu_{grey1}, \mu_{grey2}, \mu_{grey3}, \mu_{bright}\}$ be the corresponding set of membership functions obtained under Definition 3.4 . A properly defined image segmentation process can be obtained by considering in a separated way any membership function $\mu \in \mathcal{M}$. Hence, the segmented image can be interpreted as a composition wrt the above five basic functions.

Let $\mu \in \mathcal{M}$, $i \in \{0, 1, \ldots, 255\} \subset [0,255]$ be a given colour and $\mu_i =_{df} \mu(i) \in [0,1]$ be the value of the considered membership function for i. Also let $x_{i+1} =_{df} \mu_i$. The following data set can be introduced: $X =_{df} \{x \in [0,1] / x =_{df} x_{i+1}$ for $i = 0, 1, \ldots, 255\} \subset [0,1]$, where $0 < |X| \le 256$.

Definition 3.5

Let $\rho : X \times X \to [0,1]$ such that $\rho(x_t, x_k) =_{df} 1 - \delta_\varphi \cdot \left|\varphi(\mu_{i-1}) - \varphi(\mu_{k-1})\right|$ (for any $x_i, x_k \in X$), where $\varphi(\mu_{s-1})$ is a measure of fuzziness wrt μ_{s-1} ($s \in \{i,k\}$) and $\delta_\varphi =_{df} \dfrac{1}{\max\{\left|\varphi(\mu_{i-1}) - \varphi(\mu_{k-1})\right| / x_t, x_k \in X; x_t \neq x_k\}}$. Then ρ is said to be a *modified distance function of the Minkowski class.*

According to Definition 3.5 we have: $\rho(x,x) = 1$ and $\rho(x,y) = \rho(y,x)$ (for any $x,y \in X$). Then the following proposition is satisfied.

Proposition 3.4

Let ρ be a modified distance function of the Minkowski class defined in X. Then ρ is a fuzzy similarity relation. □

In the case of colour image segmentation, instead of one-element sets, ordered triples should be associated with the corresponding membership functions or using some transformation from the classical colour model RGB to other colour model, e.g. YUV, HSB, etc. Without loss of generality of this approach, the image segmentation process can be properly defined considering only one component (e.g. the luminance component Y for YUV, the brightness component B for HSB, etc.: this is omitted). Moreover, instead of one measure of fuzziness, a *linear convex combination of a subset of measures of fuzziness* can be used, i.e. $\varphi(\mu) =_{df} \sum_{i=1}^{k} \alpha_i \cdot \varphi_i(\mu)$, where any $\alpha_i \geq 0$ and $\sum_{i=1}^{k} \alpha_i = 1$. The next proposition follows directly by Definition 1.5 .

Proposition 3.5

Any linear convex combination of a set of measures of fuzziness is also a measure of fuzziness. □

Proposition 3.6

Let $\mu =_{df} (\mu(x_1), ..., \mu(x_n))$ and $\nu =_{df} (\nu(x_1), ..., \nu(x_n))$ be two fuzzy subsets on X and ν be the complement of μ, i.e. $\nu = \mu' = (1 - \mu(x_1), ..., 1 - \mu(x_n))$. Assume that $\mu(x_k) \neq .5$ (for any $k = 1, ..., n$). Then, the crisp set associated with ν, i.e. $\nu^* = (\mu^*)'$.

Proof:

(1) Let $\mu(x_k) > .5$. Then $\mu^*(x_k) = 1$ and $1 - \mu(x_k) = \nu(x_k) < .5$. Hence $\nu^*(x_k) = 0$.

(2) In a similar way, for $\mu(x_k) < .5$ we can obtain: $\mu^*(x_k) = 0$, $1 - \mu(x_k) = \nu(x_k) > .5$ and $\nu^*(x_k) = 1$. □

Proposition 3.7

Let $\mu =_{df} (\mu(x_1), ..., \mu(x_n))$ and $\nu =_{df} (\nu(x_1), ..., \nu(x_n))$ be two fuzzy subsets on X such that $\nu = \mu'$. Then the corresponding Minkowski measures of fuzziness are equal, i.e. $\varphi_3(\nu) = \varphi_3(\mu)$.

Proof:

Assume that $\mu(x_k) \neq .5$ (for any $k = 1, ..., n$) and $\nu = \mu'$. Since $\varphi_3(\nu) =_{df} \frac{2 \cdot \left(\sum_{k=1}^{n} |\nu(x_k) - \nu^*(x_k)|^q \right)^{\frac{1}{q}}}{\sqrt[q]{|X|}}$ according to Proposition 3.6 we have:

$$|\nu(x_k) - \nu^*(x_k)| = |1 - \mu(x_k) - \nu^*(x_k)|$$
$$= |\mu(x_k) + \nu^*(x_k) - 1|$$
$$= |\mu(x_k) + (1 - \mu^*(x_k)) - 1|$$
$$= |\mu(x_k) - \mu^*(x_k)|.$$

Hence $\varphi_3(\nu) = \varphi_3(\mu)$. Let now $\mu(x_k) = .5$ (for some k). Then we can obtain: $|\mu(x_k) - \mu^*(x_k)| = |.5 - 1| = .5$. On the other hand: $|\nu(x_k) - \nu^*(x_k)| = |1 - \mu(x_k) - \nu^*(x_k)| = |.5 - 1| = .5$. □

In a similar way as in Definition 3.5 the following definition can be introduced.

Definition 3.6

Let $\rho : X \times X \rightarrow [0,1]$ such that $\sigma(x_i, x_k) =_{df} 1 - \gamma_\varphi \cdot \left| \frac{\varphi(\mu_{i-1}) - \varphi(\mu_{k-1})}{\varphi(\mu_{i-1}) + \varphi(\mu_{k-1})} \right|$ (for any $x_i, x_k \in X$), where $\varphi(\mu_{s-1})$ is a measure of fuzziness wrt μ_{s-1} ($s \in \{i,k\}$) and

$\gamma_\varphi =_{df} \dfrac{1}{\max\left\{ \left| \dfrac{\varphi(\mu_{i-1}) - \varphi(\mu_{k-1})}{\varphi(\mu_{i-1}) + \varphi(\mu_{k-1})} \right| \; / \; x_i, x_k \in X; x_i \neq x \right\}}$. Then σ is said to be a *modified distance function of the Canberra class.*

Proposition 3.8

Let σ be a modified distance function of the Canberra class defined in X. Then σ is a fuzzy similarity relation. □

The above data set X is a natural ordered set. Hence, the following definition can be introduced.

Definition 3.7

Let ρ be a fuzzy relation on X and $Y \subseteq X$ be an ordered subset of X ($|Y| \geq 2$). By ρ/Y we shall say that the fuzzy relation ρ is *restricted* wrt the subset Y iff $M_{\rho/Y}$ is a submatrix of M_ρ, where $dom(\rho/Y) =_{df} dom(\rho) \cap (Y \times Y)$.

The restricted ρ/Y will correspond to some fuzzy subgraph wrt the fuzzy graph of ρ.

Proposition 3.9

Let ρ be a fuzzy equivalence relation on X and ρ/Y be a restricted fuzzy relation wrt $Y \subseteq X$ ($|Y| \geq 2$). Then ρ/Y is a fuzzy equivalence relation on Y.

Proof:

Let ρ be a fuzzy equivalence relation on X and ρ/Y be a restricted fuzzy relation wrt $Y \subseteq X$ ($|Y| \geq 2$). Hence, from Definition 1.4 it follows that ρ is reflexive, symmetric and transitive, i.e. $\rho(x,x) = 1$, $p(x,y) = \rho(y,x)$, and $\rho(x,z) \geq \max\{\min\{\rho(x,y),\rho(y,z)\} / y \in X\}$ (for any $x,y,z \in X$: without loss of generality, the max-min transitivity is assumed).

Let $a,b,c \in Y$. Since $Y \subseteq X$ then $a,b,c \in X$. In particular, in accordance with the assumption, the above three properties of ρ are also satisfied for $x =_{df} a$, $y =_{df} b$, and $z =_{df} c$, i.e. for any $a,b,c \in Y$. Hence, the restricted ρ/Y is a fuzzy equivalence relation on Y. □

And so, any quadrate submatrix of M_ρ corresponds to a fuzzy equivalence relation, if ρ is a fuzzy equivalence relation. Moreover, according to Propositions 3.3 and 3.9, the above property is preserved for any linear convex combination, i.e. the following proposition is satisfied.

Proposition 3.10

Let $\{\rho_i / i = 1, \ldots, k\}$ be a set of fuzzy similarity relations and λ be the corresponding linear convex combination wrt the above relations. If λ is transitive on X and $Y \subseteq X$ ($|Y| \geq 2$) then λ/Y is a fuzzy equivalence relation on Y. □

The last two propositions give us some possibility of using fuzzy equivalence relations as a hierarchical clustering process. In fact, after using the transitive closure algorithm for the original image, in the next processing of a given subimage (i.e. subset of pixels), any transitive closure is not necessary to be done.This is not true in the classical FCM because a subset of a pseudopartition is not a pseudopartition.

Some example modified distance functions are given below.

Example 3.2

Let $\mu =_{df} 1 / a + .86 / b + .45 / c + 0 / d$ be a fuzzy set on $X =_{df} \{a,b,c,d\}$ (a hypothetical set of "colours").

(1) According to Proposition 3.9, the corresponding matrix M_ρ is given below (after rounding off):

$$M_\rho = \begin{bmatrix} 1 & .69 & 0 & 1 \\ .69 & 1 & .31 & .69 \\ 0 & .31 & 1 & 0 \\ 1 & .69 & 0 & 1 \end{bmatrix}, \text{ where } \delta_\varphi = \frac{10}{9}.$$

(2) Assuming the same set X, the matrix M_σ is as follows:

$$M_{\sigma/Y} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & .47 & 0 \\ 0 & .47 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \text{ where } \gamma_\varphi = 1.$$

(3) Assuming $\alpha = \beta = .5$, the following linear convex combination λ of ρ and σ can be obtained:

$$M_\lambda = \begin{bmatrix} 1 & .35 & 0 & 1 \\ .35 & 1 & .39 & .35 \\ 0 & .39 & 1 & 0 \\ 1 & .35 & 0 & 1 \end{bmatrix}.$$

An analysis of the corresponding transitivity properties is omitted here. □

A specification of the corresponding equivalence classes is described in the next section.

3.5 Selection of the equivalence classes

Any fuzzy relation ρ on X (as a fuzzy set) can be uniquely represented in terms of its α-cuts. And so, according to Theorem 1.1 we have: $\rho = \bigcup_{\alpha \in [0,1]} \alpha \cdot \rho_\alpha$.

Definition 3.8

Let ρ be a fuzzy relation on X. Assume that π_α and π_β are the corresponding partitions generated by the crisp equivalence relations ρ_α and ρ_β (for $\alpha, \beta \in (0,1]$). We shall say that π_α is a *refinement* of π_β, i.e. $\pi_\alpha \sqsubseteq \pi_\beta \Leftrightarrow \forall A \in \pi_\alpha \exists B \in \pi_\beta (A \subseteq B)$.

Let $\Pi(\rho) =_{df} \{ \pi_\alpha / \alpha \in (0,1] \}$. The following proposition is satisfied (Mordeson, Malik and Cheng 2000).

Proposition 3.11

$\forall \pi_\alpha, \pi_\beta \in \Pi(\rho) (\pi_\alpha \sqsubseteq \pi_\beta \Leftrightarrow \alpha \geq \beta)$. □

Proposition 3.12

$(\Pi(\rho),\sqsubseteq)$ is a partial ordered set wrt $\sqsubseteq$. □

The equivalence classes formed by the levels of refinement of an fuzzy equivalence relation can be interpreted as grouping elements that are equivalent, i.e. similar to each other wrt the corresponding degree of association not less than α .

Example 3.3

Consider the matrix M_ρ of the previous Example 3.2 . The transitivity property of ρ can be verified by using Algorithm 1.1 . Since ρ is not transitive, after the 2^{th} iteration, a corresponding transitive (max-min) closure is obtained. So we have:

$$M_{\rho^+} = \begin{bmatrix} 1 & .69 & .31 & 1 \\ .69 & 1 & .31 & .69 \\ .31 & .31 & 1 & .31 \\ 1 & .69 & .31 & 1 \end{bmatrix}.$$

The codomain of ρ^+, i.e. $\text{cod}(\rho^+) = \{.31, .69, 1\}$. Hence ρ^+ is associated with a sequence of three *nested partitions*, for $\alpha \in \{.31, .69, 1\}$ and $\alpha > 0$. For example, assuming $\alpha = .69$ the following crisp equivalence relation $\rho^+_{.69}$ can be obtained: $M_{\rho^+.69} = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$.

The hypothetical set of colours X is partitioned into the following two subsets of colours: $I_1 = \{a, b, d\}$ and $I_2 = \{c\}$ (according to the identical columns of $M_{\rho^+.69}$). The corresponding graph of $\rho^+_{.69}$, i.e. $G_{\rho^+.69}$ is shown in Figure 3.3 below. We have two components, i.e. subgraphs generated by the subsets I_1 and I_2 . □

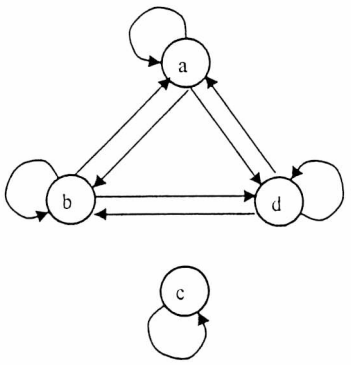


Figure 3.2 The graph $G_{\rho^+.69}$

Next we shall assume some automated image segmentation process. Hence, there exists a necessity of introducing an appropriate α-cut and so a value for $\alpha \in (0,1]$ (another possibility is presented in Sections 4.1 and 4.2: using a fuzzy expected value).

Definition 3.9

Let ρ be a fuzzy equivalence relation on X and $M_\rho =_{df} [m_{ij}]_{|n\times n}$ be the membership matrix of ρ, where $n =_{df} |X|$. Then the α-cut ρ_α is specified in a unique way as follows:

$$\overline{\alpha} =_{df} \frac{2\cdot\sum_{i=2}^{n}\left(\sum_{s=1}^{i-1} m_{is}\right)}{n\cdot(n-1)} .$$

Example 3.4

According to Definition 3.9, in the case of Example 3.3 we can obtain $\overline{\alpha} = .55166... \approx .55$. Since $\{(x,y) \in X \times X / .55 \le \rho^+(x,y) < .69\} = \varnothing$ we have: $\rho^+_{.55} = \rho^+_{.69}$, i.e. the same partition of the set X can be obtained. Moreover, it can be observed that instead using the mean arithmetic value $\overline{\alpha}$ the mean geometric or also the mean quadratic values can be used. The last two values are as follows: $\sqrt[6]{.69^2\cdot.31^3} = .49199... \approx .49$ and $\sqrt{\frac{2\cdot.69^2+3\cdot.31^2+1^2}{6}} = .61107... \approx .61$, respectively. It can be observed that: $\frac{.49+.61}{2} = .55 = \overline{\alpha}$. Here the mean harmonic value $\frac{1}{\frac{2}{.69}+\frac{3}{.31}+\frac{1}{1}} = .07365... \approx .07$ is omitted. □

3.6 Segmentation of the input image

Image segmentation is essentially a process of pixel classification, wherein the image pixels are segmented into subsets by assigning the individual pixels to classes. Consider a given image $\mathcal{I}_0$. The obtained pixel classification for $\mathcal{I}_0$ should be realised in an unique way. This process should be in accordance with the assumed basic classes of colours (see Definition 3.4). Since the considered data set of all possible colours (in grey level scale gradation) $X =_{df} \{0, 1, ..., 255\}$ is a natural ordered set, it is assumed below that this order is preserved for any subset of X. Hence, according to Definition 3.4 (see Figure 3.1) the following five cluster selection rules are introduced (see Example 4.2).

Definition 3.10

Let I_s be a cluster obtained in accordance with Definition 3.4 .The following five *cluster selection rules* are introduced:

$d-class$ rule: a cluster $I_s \in d-class$ iff I_s is selected as the first class from left to right.

$g_1-class$ rule: a cluster $I_s \in g_1-class$ iff $I_s \in (\{d-class\} \cup \{b-class\} \cup \{g_2-class\} \cup \{g_3-class\})'$.

$g_2-class$ rule: a cluster $I_s \in g_2-class$ iff I_s is selected as the cluster to which belongs FEV under Definition 4.1 or I_s is the first cluster from left or from right to FEV.

$g_3-class$ rule: a cluster $I_s \in g_3-class$ iff $I_s \in (\{d-class\} \cup \{b-class\} \cup \{g_1-class\} \cup \{g_2-class\})'$.

$b-class$ rule: a cluster $I_s \in b-class$ iff I_s is selected as the first class from right to left.

The following definition of $A(\cdot)$ is introduced: $A(I_s) = true$ iff I_s satisfies the above five cluster selection rules. Hence, for any two different clusters belonging to two different classes I_r and I_s we have: $A(I_r) = true$ and $A(I_s) = true$. Consider now the subset of colours $I^* =_{df} I_r \cup I_s \subseteq X$. Assume that $A(I^*) = true$. According to Definition 3.10 we have a partition of the set of all possible colours X into five classes. The set I^* should satisfy the cluster selection rules. Hence I^* should belong to exactly one of the above defined classes. However there is not a colour that belongs to more than one class. We have a contradiction. Hence $A(I^*) = false$ and the used interpretation is unique. In accordance with Definition 3.1, the obtained partition is image segmentation. Hence, the following proposition is satisfied.

Proposition 3.13

The above described image segmentation is a well-defined process. □

A corresponding algorithm and various experiments concerning the proposed approach are given in the next chapter.

4. The algorithm, implementation and experiments

The main objective of image enhancement is to generate a new image such that the new image is more suitable for an application than the original image. One of the methods for image enhancement is histogram equalisation (Schneider and Craig 1992). In this chapter, the fuzzy expected value is used to improve the image quality wrt the distance of all grey levels from this value (Section 4.1). A fuzzy equivalence relation-based image segmentation algorithm is presented in the next section. Here, only the main steps of this algorithm are discussed. An example implementation and various experiments concerning this approach are given in Section 4.3. In experiments, the proposed method demonstrated promising performance for various classes of medical images.

4.1 Equalisation using fuzzy expected value

The notion of a fuzzy expected value is introduced below (Schneider and Craig 1992, Kerre and Nachtegael 2000). The determination of the introduced in Definition 3.4 parameters a, b, c, d_1, d_2, e, and f realised in an automated way by means of this expected value. Provided there is no ambiguity, in the next considerations we shall assume that the data set (i.e. universe) X is an image $\mathcal{I}$ of *size* $m \times n$, i.e. $|\mathcal{I}| = m \cdot n$ pixels, with L *grey levels* g = 0, 1, ... , L – 1.

Definition 4.1

Let μ be a fuzzy subset on X (an universe) and $\mu_T \subseteq X$ be the α-cut wrt μ for $\alpha =_{df} T \in [0,1]$, where T is a threshold value. Assume that $\varphi(\mu_T)$ is a measure of fuzziness defined as follows: $\varphi(\mu_T) =_{df} \frac{|\mu_T|}{|X|}$. Let $F =_{df} \sup\{\min\{T, \varphi(\mu_T)\} / T \in [0,1]\}$. Let i be the colour index corresponding to F. Then the *fuzzy expected value* (in short: FEV) of μ is defined as the colour associated with i.

Algorithm 4.1 (FEV-equalisation)

(1) Calculation of the image histogram ;

(2) Calculation of the most typical value $K =_{df}$ FEV ;

(3) Calculation of the distance measure D of each grey level $g \in \{0, 1, \ldots, L-1\}$ from the most typical value K:

$$D =_{df} \sqrt{|K^2 - g^2|}$$

(4) Generation of the new grey levels $\hat{g}$:

$$\hat{g} =_{df} \begin{cases} \max\{0, K-D\} & \text{if } g < K \\ \min\{L-1, K+D\} & \text{if } g > K \\ K & \text{otherwise} \end{cases}$$

End. □

In a similar way the *weighted fuzzy expected value* (in short: WFEV, denoted below by s) can be calculated by $\hat{s} =_{df} \frac{\sum_{i=0}^{L-1} \mu_i \cdot e^{-\beta|\mu_i - s|} \cdot n_i^{\lambda}}{\sum_{i=0}^{L-1} e^{-\beta|\mu_i - s|} \cdot n_i^{\lambda}}$, where: β and λ are constants (e.g. β = .1, λ = 2), n_i is the frequency of i^{th} grey level (histogram), and $\mu_i = \mu(i)$ is the membership value of the i^{th} grey level (e.g. normalised in [0,1], here $e^{-\beta|\mu_i - s|}$ is a distance function). In general, the last equation is of the form: $s_{k+1} = F(s_k)$ and can be solved iteratively assuming $s_0 =_{df}$ FEV. Then the obtained solution is called the WFEV. Hence, the step (2) of Algorithm 4.1 can be modified as follows: $K \in$ {FEV, WFEV} (this is omitted).

Example 4.1

(1) Let X be a hypothetical example image $\mathcal{I}$ of size 10 × 10, i.e. having 100 pixels, with L = 8 grey levels $g =_{df} i$ (i = 0, 1, ... ,7). Also, let $n_i =_{df} |\{p \in \mathcal{I} / \text{colour of } p = i\}|$ be the number of pixels having colour i (for any i). Obviously $\sum_{i=0}^{7} n_i = 100$. Then, according to Definition 4.1 we can obtain: $F = \sup\{\min\{T, \varphi(\mu_T)\} / T \in [0,1]\} = \max\{\min\{T, \varphi(\mu_T)\} / T \in \{\mu_i / i = 0, 1, \ldots, 7\}\} = .57$. This corresponds to the colour index 4. Hence FEV = 145. The corresponding computation of FEV (after rounding off) is given in the table below. An illustration of the used pallet of colours is shown in Figure 4.1 .

Colour value	Colour index i	$\mu_i =_{df} \mu(i)$	n_i	$\lvert\mu_T\rvert$	$\varphi(\mu_T)$	$\min\{\mu_i, \varphi(\mu_T)\}$
2	0	0	11	100	1	0
15	1	.14	5	89	.89	.14
71	2	.29	24	84	.84	.29

128	3	.43	2	60	.6	.43
145	4	.57	14	58	.58	.57
174	5	.71	31	44	.44	.44
198	6	.86	6	13	.13	.13
237	7	1	7	7	.07	.07

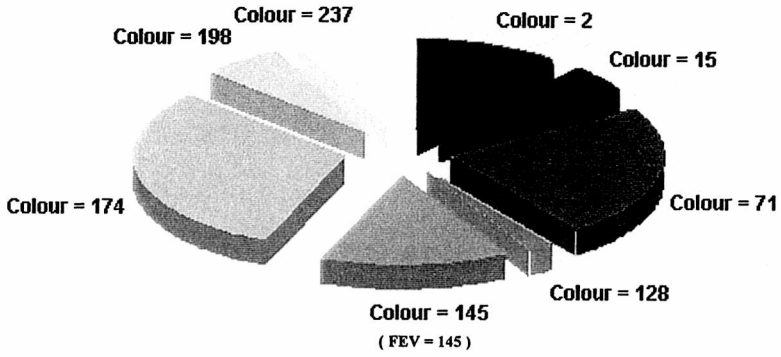
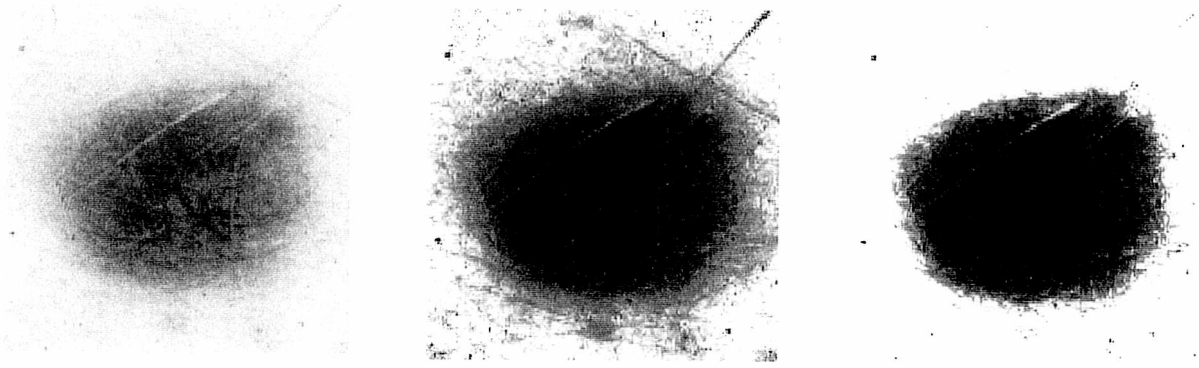


Figure 4.1 The used pallet of colours in Example 4.1(1)

(2) A comparative example for the classical and fuzzy histogram equalisation approaches is given in Figure 4.2 below. The obtained fuzzy equalisation is under Algorithm 4.1 . Hence, after using a corresponding FEV- preprocessing, the obtained image is more suitable for an application than the image under the classical histogram equalisation approach. □



Skin cancer | Classical histogram equalisation | FEV histogram equalisation

Figure 4.2 A comparison between classical and fuzzy histogram equalisation

An illustration of FEV-computation for α (a specification of the used α-cut) is given in the next example.

Example 4.2

Let M_{ρ^+} be the matrix considered in Example 3.3 . Since M_{ρ^+} is a symmetric matrix, it is sufficient to consider the elements that are below the main diagonal. Any such element is interpreted as a *hypothetical grey level* (the corresponding number of all such levels is denoted also by L). Also any μ_i is interpreted as a "colour". .Since L = 3 we can obtain:

Colour index i	$\mu_i =_{df} \mu(i)$	n_i	$\lvert\mu_T\rvert$	$\varphi(\mu_T)$	$\min\{\mu_i,\varphi(\mu_T)\}$
0	.31	3	6	1	.31
1	.69	2	3	3/6	.50
2	1	1	1	1/6	$.16_{(6)}$

Hence α_{FEV} = .69 . □

In general, assuming that M_{ρ^+} is n × n square matrix, the number of the above hypothetical grey levels $L = \frac{n\cdot(n-1)}{2} + r - \sum_{i=1}^{r} n_i$, where r is the number of the subsets of identical elements (below the main diagonal) and n_i is the cardinality of the i^{th} subset (i = 1, 2, ... ,r).

4.2 Histogram based ε-fuzzification

A more exact process of fuzzification can be obtained by analysing some properties of the image colours directly following from the obtained histogram. Below two kinds of distances are considered: colour distances and pixel number distances. Hence, we have a possibility of identification of the corresponding colour concentration points wrt the used measure of fuzziness φ . In consequence, a better quality of the whole segmentation process can be obtained. Moreover, this process can be simplified by considering only the corresponding colour concentration points. The following algorithm describes the obtained preprocessing.

Algorithm 4.2

(1) Input the histogram $H(\mathcal{J}_0)$ associated with the original image $\mathcal{J}_0$;

(2) Let i_{min} and i_{max} be the first and the last colours in $H(\mathcal{I}_o)$, respectively. Also let h_{min} and h_{max} be the minimal and the maximal pixel numbers associated with $H(\mathcal{I}_o)$;

(3) For any different two colours i and j define: $\alpha_{ij} =_{df} \frac{|i - j|}{i_{max} - i_{min}}$ and $\beta_{ij} =_{df} \frac{|h(i) - h(j)|}{h_{max} - h_{min}}$ (i,j = 0, 1, ... ,255), where h(k) is the number of all pixels having colour $k \in \{i,j\}$. Let $\mu =_{df} \alpha_{ij} / x_1 + \beta_{ij} / x_2$ be a fuzzy set ;

(4) Let φ be a measure of fuzziness. Compute $\psi(\mu) =_{df} \frac{\varphi(\mu)}{\varphi_{max}(\mu)}$;

(5) If $\psi(\mu) \leq \varepsilon$ then colours i and j belong to the environment of the same colour concentration point, where $\varepsilon \in [0,1]$ is a priori given value ;

(6) Let $\pi =_{df} \pi(\{0, 1, ... ,255\})$ be the obtained partition under (5). For any subset of colours $A \subseteq \pi$ select the *colour concentration point* p_A such that $h(p_A) =_{df} \max\{ h(i) / i \in A \}$

(7) Define the reduct $\hat{H}(\mathcal{I}_o)$ by considering only the set of colour concentration points. □

An illustration of Algorithm 4.2 is given in Example 4.3 below.

4.3 The algorithm specification

A fuzzy equivalence relation-based image segmentation algorithm (in short: FERIS) is presented below. Provided there is no ambiguity and for convenience, here only the main steps of this algorithm are included. It is shown that FERIS is a well-defined algorithm, i.e. for any *original image* $\mathcal{I}_o$ there exists exactly one *segmented image* $\mathcal{I}_s$ obtained from $\mathcal{I}_o$ under FERIS. And so, the corresponding process of image segmentation is always realised in a unique way and the algorithm converges in a finite time. The computational complexity of FERIS is also considered. A comparative analysis wrt the classical FCM (see Algorithm 2.1) is also given.

Algorithm 4.3 (FERIS algorithm)

(1) Input of the original image $\mathcal{I}_0$;

(2) Using of an appropriate image preprocessing and FEV-computation of the parameters a, b, c, d, e $\in$ [0,255] and identification of the corresponding colour concentration points under Algorithm 4.2 ;

(3) Fuzzification of the input space and specification of the membership functions μ_{dark} ,μ_{grey1} ,μ_{grey2} , μ_{grey3} and μ_{bright} ;

(4) Selection of a measure of fuzziness φ and an appropriate distance function ρ . Specification of the corresponding fuzzy similarity relation ;

(5) Determination of the initial membership matrix M_ρ ;

(6) Calculation of the transitive closure of ρ and definition of the corresponding fuzzy equivalence relation ρ^+ ;

(7) Selection of $\alpha \in (0,1]$ and specification of the used α-cut: Define $\alpha =_{df} \alpha_{FEV}$ (i.e. FEV for α) and $\hat{\rho}_\alpha =_{df} (\rho^+)_\alpha$;

(8) Selection of the equivalence classes belonging to the obtained partition $\pi(\hat{\rho}_\alpha)$;

(9) Specification of the obtained clusters by using the cluster selection rules ;

(10) Final segmentation of the original image and generation of the segmented image $\mathcal{I}_S$. End of FERIS. □

In accordance with the above Algorithm 4.3, after the input of $\mathcal{I}_0$ (1), first an appropriate image preprocessing is realised and the corresponding parameters a, b, c, d, and e are computed by using FEV. The process of identification of the corresponding colour concentration points under Algorithm 4.2 is realised in a unique way (2). In the next step (3), a fuzzification of the input space and specification of the membership functions μ_{dark} ,μ_{grey1} ,μ_{grey2} , μ_{grey3} and μ_{bright} is realised. This is in accordance with Definition 3.4 . The process of a measure of fuzziness selection (4) and an appropriate distance function is realised in an automated way. Hence, it is assumed an Euclidean measure of fuzziness under Definition 1.5, i.e. $\varphi =_{df} \varphi_2(\mu)$. However, there may exist two different colours i and k (i,k $\in$ {0, 1, ... ,255}) such that the complement $\mu'(i) = \mu(k)$ ($\mu \in \mathcal{M}$). According to Proposition 3.7 the

corresponding Minkowski measures of fuzziness are equal. Since $q = 2$ we have: $\varphi_2(\mu_i) = \varphi_2(\mu_k)$. Then the uniqueness condition is preserved by assuming $\varphi_2(\mu_k) =_{df} 1 - \varphi_2(\mu_i)$. This coincides with the practical case. The used fuzzy similarity relation ρ is the (modified) distance function of the Minkowski class under Definition 3.5 . After the determination of the initial membership matrix M_ρ (5), the max-min transitivity property is verified. (6) . Without loss of generality, the corresponding fuzzy equivalence relation ρ^+ is generated under Algorithm 1.1 . A very good approximation for α can be obtained using $\alpha =_{df} \alpha_{FEV}$ (i.e. FEV for α) under Definition 4.1 or α_{WFEV} (see: Section 4.1). Another possible value for α may be the corresponding mean arithmetic value $\overline{\alpha}$ under Definition 3.9 . In this version of FERIS the fuzzy expected value α_{FEV} is assumed (7). And so, the α-cut $\hat{\rho}_\alpha =_{df} (\rho^+)_\alpha$ is defined in an unique way. The process of selection of the equivalence classes belonging to the family $\pi(\hat{\rho}_\alpha)$ under step (8) is realised in an automated way by analysis of the subsets of identical rows (an equivalent analysis of the subsets of identical columns can be also accepted). The specification of the obtained clusters (9) is realised under d – , g_1 –, g_2 –, g_3 – and b – class cluster selection rules given in Definition 3.10 . According to the obtained partition of the set of all clusters, the final segmentation of the original image is realised and the segmented image $\mathcal{J}_s$ is generated (10).

In accordance with the above given description of FERIS, in any step (k) of this algorithm (k = 1, 2, ... ,10) the associated with this step formula (-e) is (are) strictly defined. Since the size of $\mathcal{J}_o$ is finite then any such computation is realised in a finite time. The used interpretation of Definition 3.1 is unique. This is in accordance with the introduced cluster selection rules. Hence, according to Proposition 3.12 the above described image segmentation is a well-defined process. The compound function realised by FERIS can be interpreted as a composition of the sequence of the functions F_i realised in each step (i) of this algorithm. In fact, the condition $dom(F_{i+1}) \subseteq cod(F_i)$ is satisfied for any i. Provided there is no ambiguity, the compound function is also denoted by FERIS. We have: $\mathcal{J}_s = FERIS(\mathcal{J}_o)$. The corresponding process of image segmentation is always realised in a unique way and the algorithm converges in a finite time. The obtained image segmentation is realised as an automated computational process. The following proposition is satisfied.

Proposition 4.1

FERIS is a well-defined algorithm, i.e. $\forall \mathcal{I}_0 \; \exists! \; \mathcal{I}_s \; (\mathcal{I}_s = \text{FERIS}(\mathcal{I}_0))$. □

In contradistinction to the FCM algorithm (see Algorithm 2.1, also commentary at the end of Chapter 2), FERIS is a linear structure algorithm of polynomial complexity. In consequence the total computational time τ of FERIS will be less than this one for FCM. An approximative calculation of the computational time complexity of FERIS can be obtained assuming the worst case, i.e. considering only the max-min transitive closure calculation of ρ according to step (6). For a given class of images some single algorithm calibration is possible. Hence, some steps, e.g. such as (4) and (5), will be executed only one time. Next we shall assume that $\tau < 6 \cdot \tau_6$, where τ_6 is the time consumed in step (6). According to Theorem 1.2, the max-min transitive closure, $\rho^+ = \rho^{(n-1)}$ can be obtained by using no more than $(n - 1)$ times the operation '$\circ$'. Hence, assuming SISD architecture (i.e. single instruction, single data) the computational time complexity of $\rho \circ \rho$ is of size n^3, where $n =_{df} |X|$. Hence $\tau_6 = k \cdot n^3(n-1)$, where k is a constant and $\tau < 6 \cdot k \cdot n^3(n-1)$. For $n = 256$, assuming 2GHz Pentium IV processor and also 4 cycles for one elementary operation $\min\{a,b\}$ $(a,b \in [0,1]$, e.g. using the compare instruction CMP) we can obtain: $k = \frac{4}{2 \cdot 10^9}$ [sec] and $6 \cdot n^3(n-1) = 6 \cdot 256^3 \cdot 255 = 25669140480 \approx 26 \cdot 10^9$. Hence, the total computational time $\tau < 52$ [sec]. Assuming 3GHz we have: $\tau < 35$ [sec]. For a 2 GHz SIMD architecture (i.e. single instruction, multiple data) the corresponding time $\tau <$ .20054... $\approx$ .2 [sec] (here n^3 is reduced to n^2). A high speed computation can be obtained using a systolic architecture. In the last case n^3 is reduced to $(3n - 1)$. And so, we have: $\tau < \frac{24 \cdot (3n-1)(n-1)}{2 \cdot 10^9} = .00234702$ [sec] $\approx$ 2 [msec]. Some experimental results for τ_6 are given in the next section. An illustration of Algorithm 4.3 is given below.

Example 4.3

Multiple studies have shown that the carotid artery intima-media thickness (IMT), as measured non-invasively by ultrasonography (USG), is directly associated with an increased risk of cardiovascular disease. Because it has been shown to be an independent predictor of cardiovascular disease after adjustment for traditional risk factors, it is the only non-invasive imaging test currently recommended by the American Heart Association for inclusion in the evaluation of risk. However, it remains unclear how much additional information beyond that

afforded by traditional risk factors is gained by inclusion of IMT in risk profiles. Change in IMT is increasingly being used as the end point in interventional trials. Meaningful differences in progression rates have been shown in progression rates in trials of either lipid-lowering drugs or calcium channel blockers involving several hundred subjects over a period of several years. Acceptance of a standardised protocol for measuring IMT change would facilitate comparison of results from the many trials using this technique. However, uncertainty about which measure of IMT offers the best end point has inhibited methodological standardisation.

An example USG image segmentation using FERIS is given below. The original image $\mathcal{I}_o$ is shown in Figure 4.3(a). The obtained d – class clustering is shown in Figure 4.3(b). The g_1 – class clustering is given in Figures 4.3(c,d), where $|\{ g_1 - \text{class}\}| = 2$ clusters.

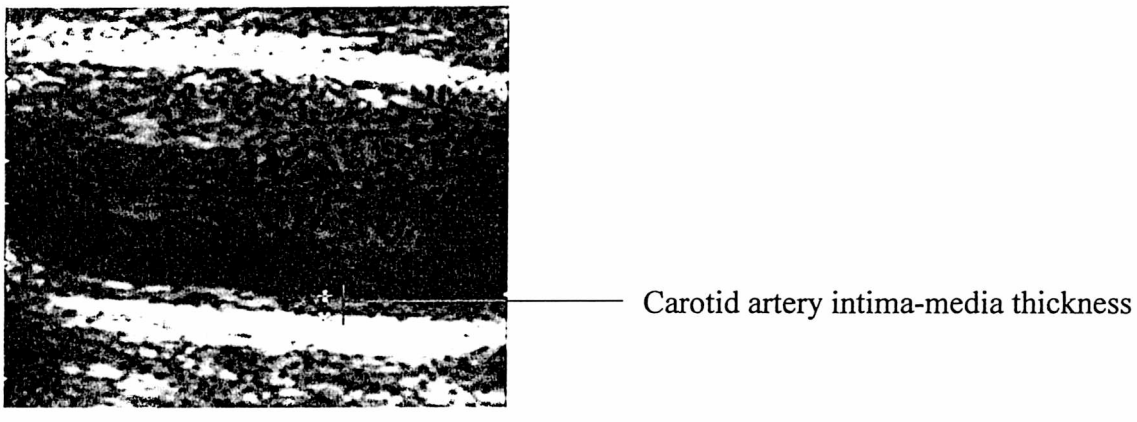


Figure 4.3(a) An example USG image $\mathcal{I}_o$

In a similar way the obtained g_2 – class clustering is given in Figure 4.3(e). The obtained g_3 – class clustering corresponds to Figure 4.3(f). The b – class clustering is shown in Figure 4.3(g) below. In accordance with the corresponding medical requirements the obtained segmented image $\mathcal{I}_s$ is defined as the whole g – class shown in Figure 4.4(a) below. This is equivalent of using the complement $(\{\text{d} - \text{class}\} \cup \{ \text{b} - \text{class} \})'$. The obtained four clusters, belonging to the whole g – class, are identified by 4 different colours as it is shown in Figure 4.4(b). This is important wrt the observed disease characteristics, e.g. such as: lipid-lowering, calcium channel blockers, etc.

The following steps are identified by FERIS (for simplicity, steps (1), (5), (6), and the specification of the obtained α-cuts $\hat{\rho}_\alpha$ in (7) are omitted here).

(2),(3): Obtained FEV-values: a = 45, b = 70, c = 83, d = 105, e = 148 ;
Here c = 83 is the global FEV for [0,255], a = 45 is the corresponding local FEV for [0,c) and e = 148 is the local FEV for (c,255]. The remaining values are computed wrt the intersection of the curves μ_{dark} and μ_{grey3} (for b = 70) and μ_{grey1} and μ_{bright} (for d = 105). The process of identification of the colour concentration points under Algorithm 4.2 is shown in Figure 4.5(a,b) below (an Euclidean measure of fuzziness is used). The accepted values for ε are as follows: $\varepsilon = 0$ (dark, bright) and $\varepsilon = .6$ ($grey_1$, $grey_2$, $grey_3$) ;

(4): It is assumed an Euclidean measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ ;

(5),(6): The corresponding determination of M_ρ and the calculation of ρ^+ are omitted in this example ;

(7): $\alpha_{FEV\,/\,dark}$ = .632232462877624, $\alpha_{FEV\,/\,grey1}$ = .521403367206105, $\alpha_{FEV\,/\,grey2}$ = .617647058823529, $\alpha_{FEV\,/\,grey3}$ = .522058823529412, $\alpha_{FEV\,/\,bright}$ = .50968595323434 (the specification of the obtained α-cuts $\hat{\rho}_\alpha$ is omitted here) ;

(8),(9): $I_s \in$ d – class: {{16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47}},
$I_s \in g_1$ – class: {{48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68},{69,70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81}},
$I_s \in g_2$ – class: {{82, 83, 84, 85, 86, 87, 88, 89, 90, 96, 97, 98, 99, 100}},
$I_s \in g_3$ – class: {{101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140}},
$I_s \in$ b – class: {{141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 230, 231, 232, 235}} ;

(10): The segmented image $\mathcal{I}_s =_{df} \{g_1 - \text{class}\} \cup \{g_2 - \text{class}\} \cup \{g_3 - \text{class}\} = \{g - \text{class}\}$. □

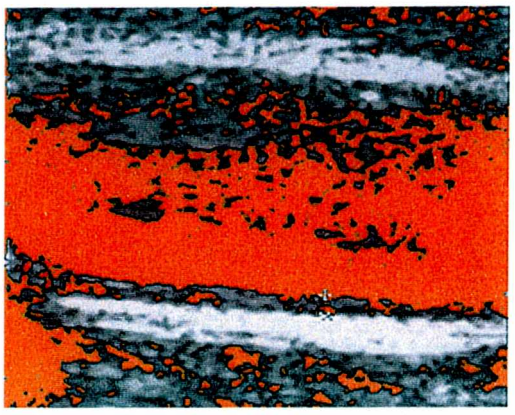


Figure 4.3(b) d – class clustering

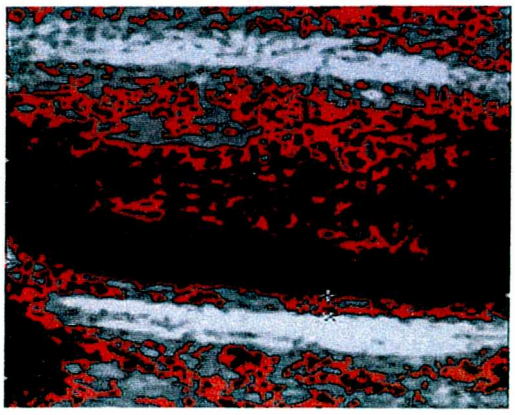


Figure 4.3(c) g_1 – class clustering

Figure 4.3(d) g_1 – class clustering

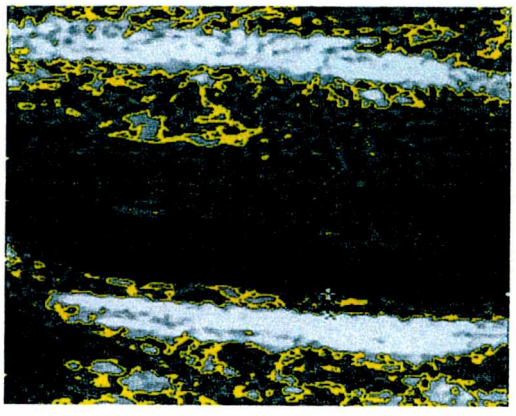


Figure 4.3(e) g_2 – class clustering

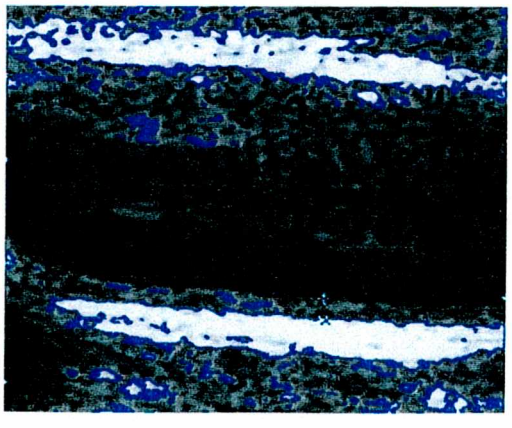


Figure 4.3(f) g_3 – class clustering

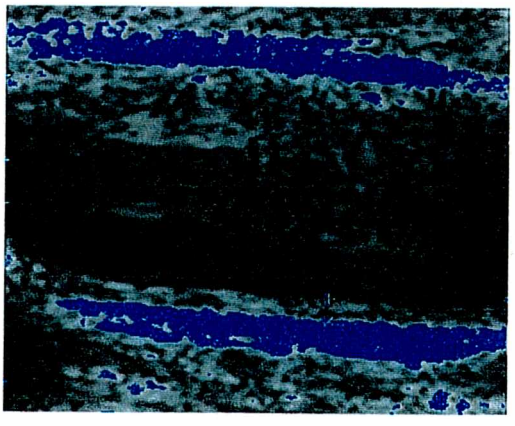


Figure 4.3(g) b – class clustering

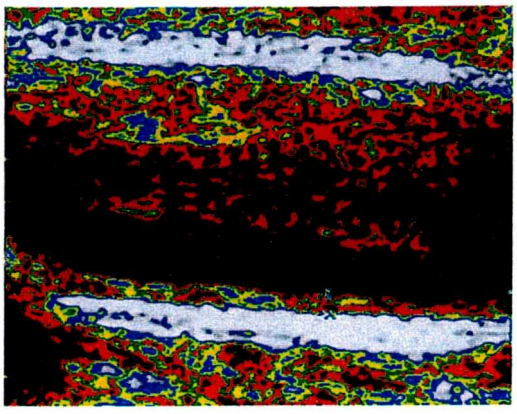


Figure 4.4 (a) The whole grey class

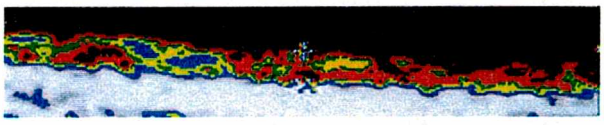


Figure 4.4 (b) The intima-media thickness

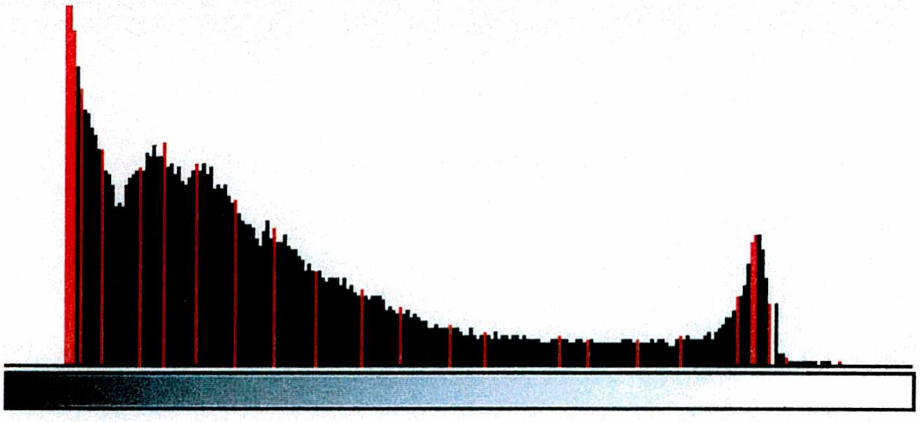


Figure 4.5(a) The histogram of the IMT with marked concentration points. The used parameters are: $\varepsilon = .1$ and Chebyshev measure of fuzziness

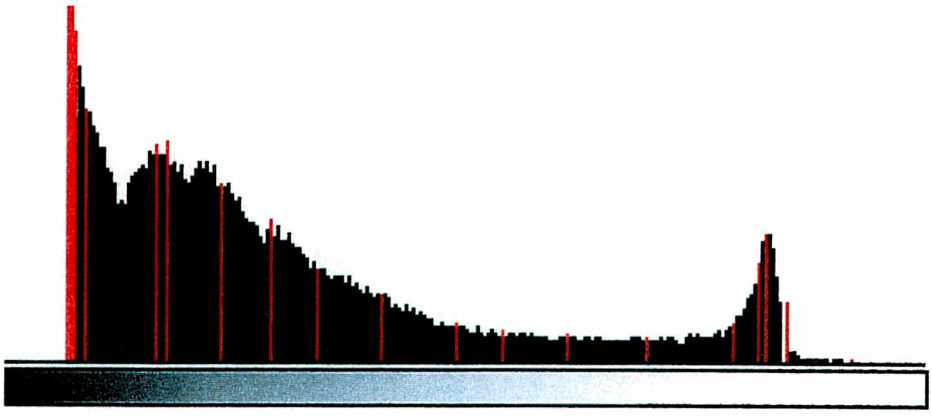


Figure 4.5(b) The histogram of the IMT with marked concentration points. The used parameters are: $\varepsilon = .1$ and Euclidean measure of fuzziness

An example implementation of this algorithm and various experiments concerning this approach are given below.

4.4 Implementation and experiments

The proposed fuzzy equivalence relation-based image segmentation algorithm has been implemented using Delphi 5.0. Some experiments using FERIS are given below. At the beginning an illustration concerning computed tomography (CT imaging) is presented.

Example 4.4 (CT imaging)
Unlike other medical imaging techniques, such as conventional x-ray imaging (radiography, CT enables direct imaging and differentiation of soft tissue structures, such as liver, lung tissue, and fat. CT is especially useful in searching for large space occupying lesions, tumours and metastasis and can not only reveal their presence, but also the size, spatial location and extent of a tumour. CT imaging of the head and brain can detect tumours, show blood clots and blood vessel defects, show enlarged ventricles (caused by a build up of cerebrospinal fluid) and image other abnormalities such as those of the nerves or muscles of the eye. Due to the short scan times of 500 milliseconds to a few seconds, CT can be used for all anatomic regions, including those susceptible to patient motion and breathing. For example, in the thorax CT can be used for visualisation of nodular structures, infiltrations of fluid, fibrosis (for example from asbestos fibers), and effusions (filling of an air space with fluid).

By using image segmentation we have a possibility of improving the accuracy of the corresponding radiological CT image interpretation as follows.

- Consider the original image $\mathcal{I}_{01}$ of Figure 4.5(a), where a large human brain tumour of the left frontal lobe is shown. A wide area of brain oedema surrounds this tumour. The oedema has an irregular shape. The border between the oedema and the surrounding tissue is not enough sharp for an exact radiological analysis. The oedema occupies a considerable part of the left brain hemisphere. It has a tendency for diffusion into the opposite hemisphere. The main task of the segmentation process is an isolation of the oedema from the remaining pathological changes in the brain (before and after the surgery). The magnitude of the area of oedema is an important indirect indicator of the neoplasmatic process expansiveness.

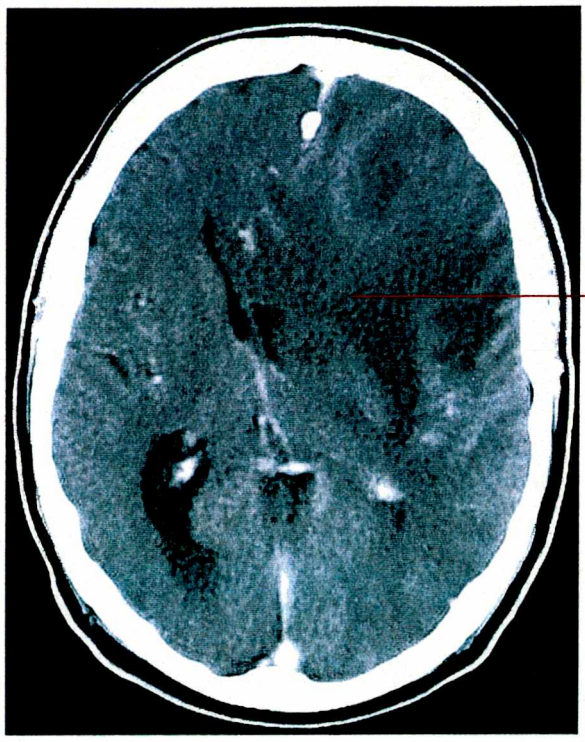
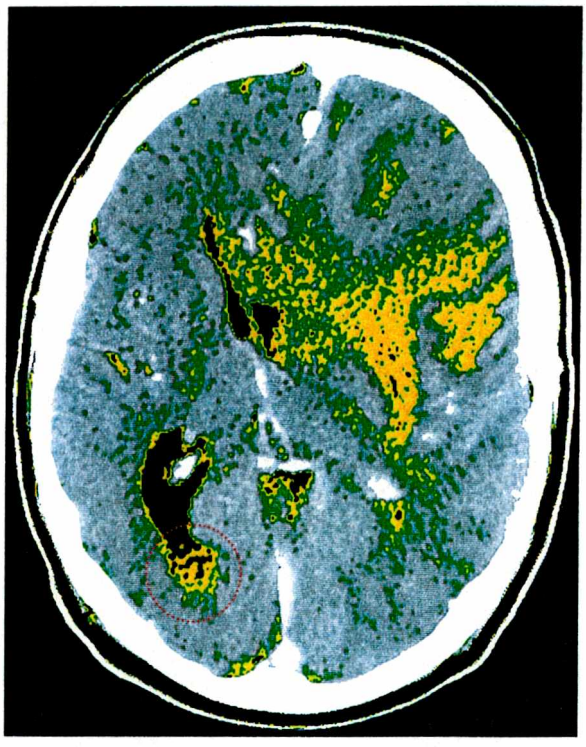


Figure 4.5(a) The original image $\mathcal{I}_{o1}$



■ With black colour is segmented the cerebro-spinal fluid

■ With yellow and green colours is segmented the area of oedema

□ With white colour are segmented the plexi choroidei and the brain falx

Figure 4.5(b) The segmented image $\mathcal{I}_{s1}$

The following parameters where used: a = 82, c = 117, e = 188, ε = .3 (dark, $grey_1$, bright). The clusters associated with $grey_2$ and $grey_3$ are not important for this case. So they are omitted. It is assumed a Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ. Hence, the following classes are obtained: {d – class} = {{3, 6, 9, 12, 15, 19, 22, 25, 28, 31, 35, 38, 41, 44, 47, 51, 54, 57, 60}}, {g_1 – class} = {{63, 66, 70, 73, 76, 79, 82}, {86, 89, 92, 95, 98, 102}} and {b – class} = {{178, 181, 184, 188, 191, 194, 197, 200, 204, 207, 210, 213, 216, 219, 223, 226, 229, 232, 235, 239, 242, 245, 248, 251}}. The isolation of the oedematous area from the remaining changes in the brain is realised in a unique way by comparing $\mathcal{I}_{o1}$ with $\mathcal{I}_{s1}$. Here it is assumed that the user is familiar with the anatomic structure of the brain. During the segmentation process another pathology has been also recognised as it is shown in Figure 4.5(b), i.e. a diffusion of the cerebro-spinal fluid from the enlarged occipital horn of the right lateral ventricle. A process of active hydrocephalus resulting from compression of the frontal ventricular horns by the pathologic zone in the left frontal lobe might cause this fluid transudation. This suspicion was confirmed by a 'FLAIR-projection'. Next are discussed two original CT images $\mathcal{I}_{oi}$ obtained at the same time (for two different planes of scan: see Figure 4.7 at the end of Example 4.4) after the surgery and the corresponding segmented images $\mathcal{I}_{si}$ under FERIS (i = 1,2).

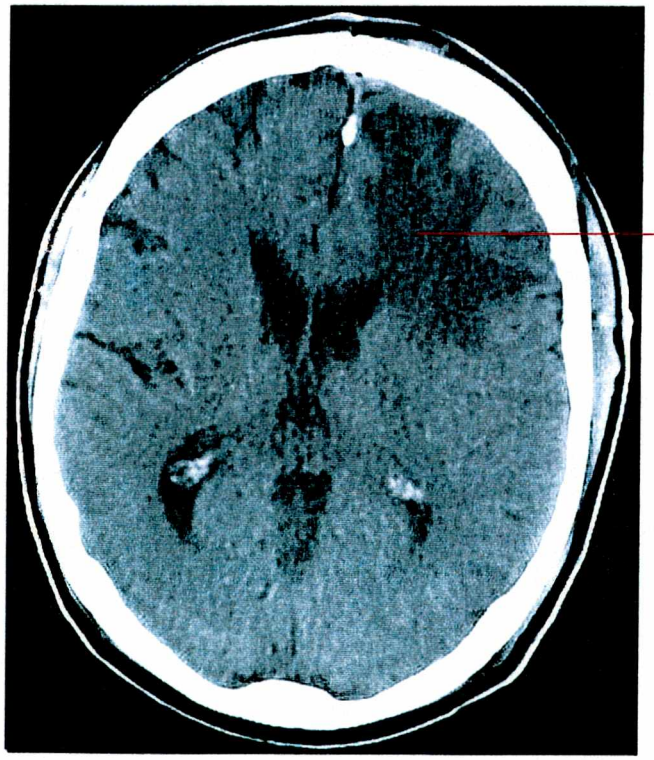
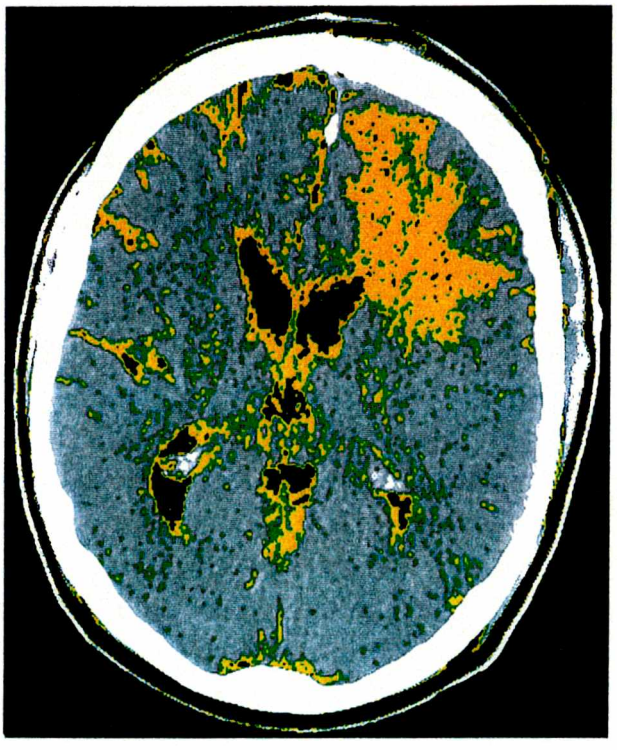


Figure 4.5(c) The original image $\mathcal{I}_{o2}$

Let first consider the original image $\mathcal{I}_{o2}$ after the surgery shown in the above Figure 4.5(c) The corresponding segmented image $\mathcal{I}_{s2}$ under FERIS is shown in Figure 4.5(d) below.



■ With black colour is segmented the cerebro-spinal fluid

■ ■ With yellow and green colours is segmented the oedema area

□ With white colour are segmented the plexus choroideus and the brain falx

Figure 4.5(d) The segmented image $\mathcal{I}_{s2}$

The following parameters have been used: a = 51, c = 111, e = 188, all other parameters have been used with the same values as in the case of $\mathcal{I}_{s1}$ shown in Figure 4.5(b). Hence, the following classes have been obtained: {d – class} = {{3, 6, 9, 12, 15, 19, 22, 25, 28, 31, 35, 38, 41, 44, 47, 51}}, {g_1 – class} = {{54, 57, 60, 63, 66, 70, 73, 76, 79, 82}, {86, 89, 92, 95}} and {b – class} = {{178, 181, 184, 188, 191, 194, 197, 200, 204, 207, 210, 213, 216, 219, 223, 226, 229, 232, 235, 239, 242, 245, 248, 251}}. It can be observed that a reduction process of the oedema area is started. This means that the considered tumour might be removed successfully.

Let now consider the original image $\mathcal{I}_{o3}$, where the tumour location after the surgery is shown in Figure 4.5(e). The corresponding segmented images $\mathcal{I}_{s3}$ is given in Figure 4.5(f) below. The following parameters have been used: a = 51, c = 114, e = 191, ε = .3 (dark, $grey_1$, $grey_3$, bright), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ. Then the next classes were obtained: {d – class} =

{{3, 6, 9, 12, 15, 19, 22, 25, 28, 31, 35, 38, 41, 44, 47, 51, 54, 57, 60}}, {g_1 – class} = {{63, 66, 70, 73, 76, 79}, {82, 86, 89, 92, 95}}, {g_3 – class} = {{149, 153, 156, 159, 162, 165, 168}} and {b – class} = {{172, 175, 178, 181, 184, 188, 191, 194, 197, 200, 204, 207, 210, 213, 216, 219, 223, 226, 229, 232, 235, 239, 242, 245, 248, 251}}.

The segmentation and characterisation of the tumour location is of much importance for the radiological analysis. Turning to the above discussed case, in the first thirty days after the operation of tumour resection there is no any possibility to determine whether the site of tumour removal contains blood or remains of the tumour (the both have the same brightness). The only way is to segment and characterise this area in a temporally way. A tendency for regression of the area of brain oedema (in comparison with the state before operation) will correspond to the assumption of blood existence. Otherwise, a more likely would be the assumption of tumour remains existence.

The segmented image $\mathcal{I}_{s3}$ under FERIS is shown in Figure 4.5(f) below.

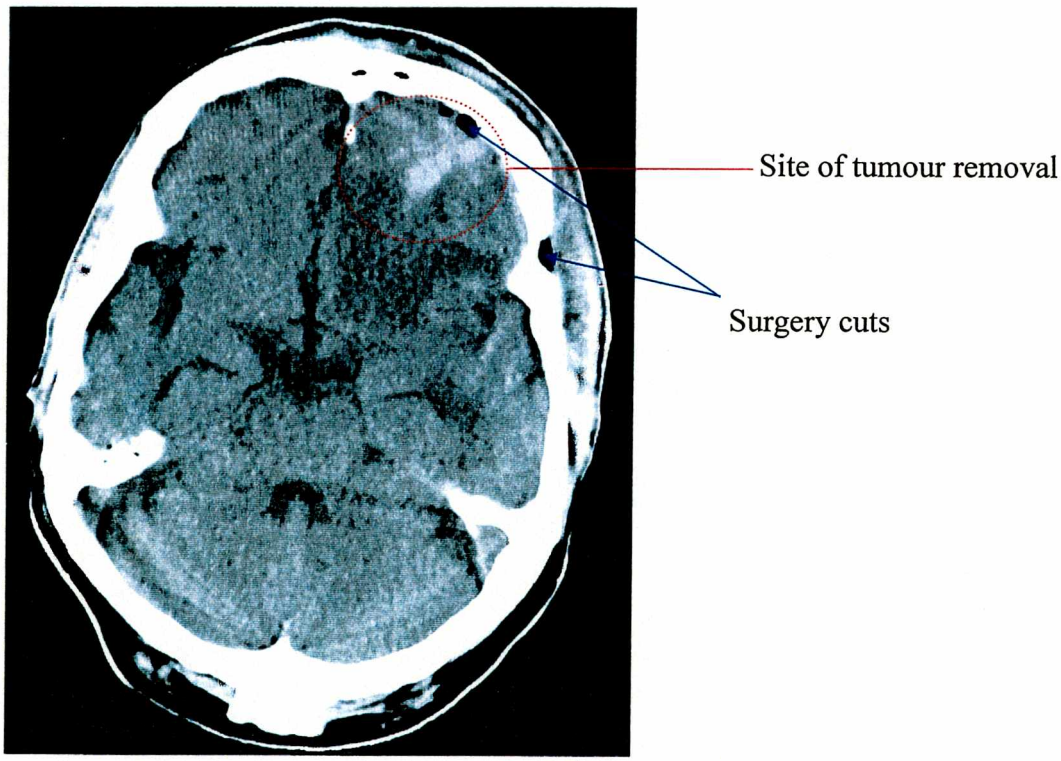
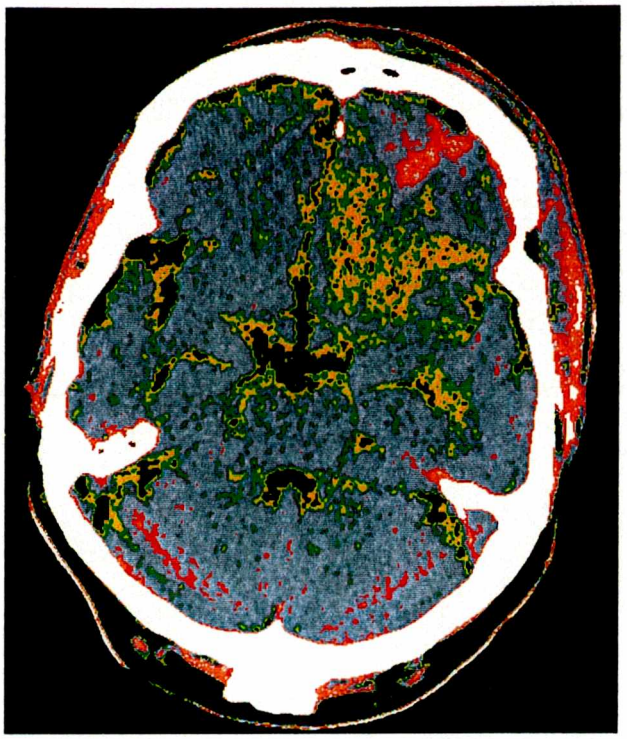


Figure 4.5 (e) The original image $\mathcal{I}_{o3}$



■ With black colour is segmented the cerebro-spinal fluid

■ ■ With yellow and green colours is segmented the oedema area

■ ■ With red and orange colours is segmented the area of the tumour location after the surgery

Figure 4.5(f) The segmented image $\mathcal{I}_{s3}$

- The next considered case ($\mathcal{I}_{o4}$ of Figure 4.6 (a)) is a suspicion of a process of destruction of the myelin sheet enwrapping neurones in the central area of the brain. It is very important to segment the focus of demyelination (i.e. the process of destruction of myelin) in purpose of its characterisation, see $\mathcal{I}_{s4}$ of Figure 4.6 (b) below.

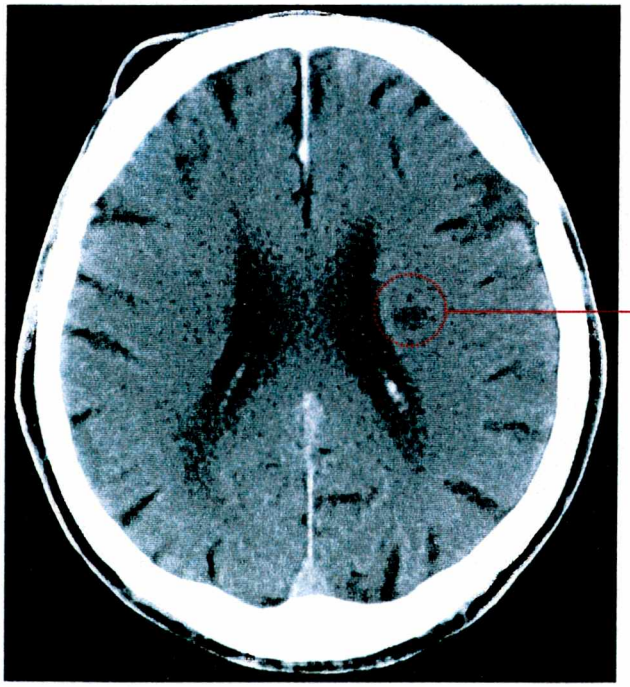
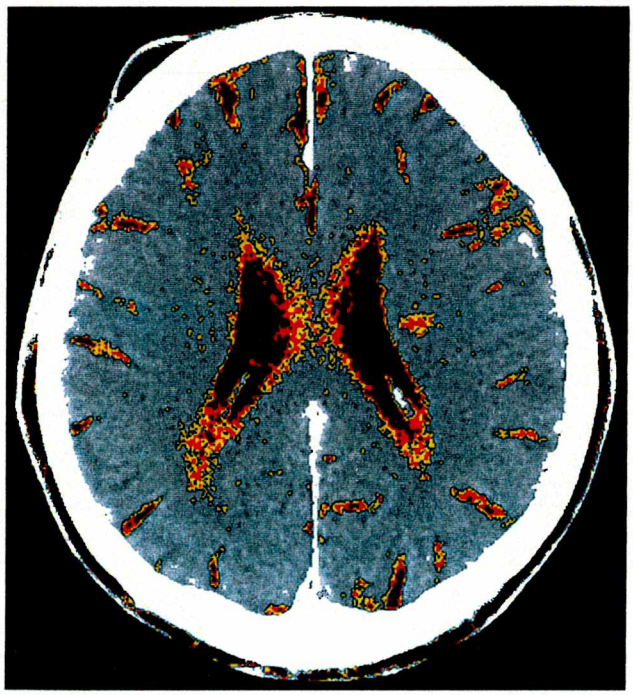


Figure 4.6(a) The original image $\mathcal{I}_{o4}$

The following parameters have been used: a = 76, c = 108, e = 184, ε = .1 (dark, $grey_1$, bright), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ.



■ With black colour is segmented the cerebro-spinal fluid

■ ■ With red and yellow colours is segmented the demyelination area

□ With white colour are segmented the plexus choroideus and the brain falx

Figure 4.6(b) The segmented image $\mathcal{I}_{s4}$

Below an image magnification of the focus of central demyelination is shown (Figure 4.6(c): before and after the segmentation process). Here a zooming factor 400% has been used.

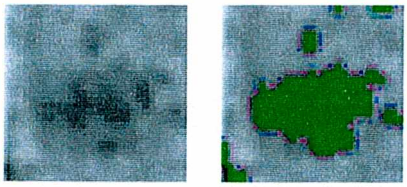


Figure 4.6(c) The magnified region

According to the above-considered examples, the manner of obtaining CT scans in different planes (below denoted by θ) is illustrated in Figure 4.7. □

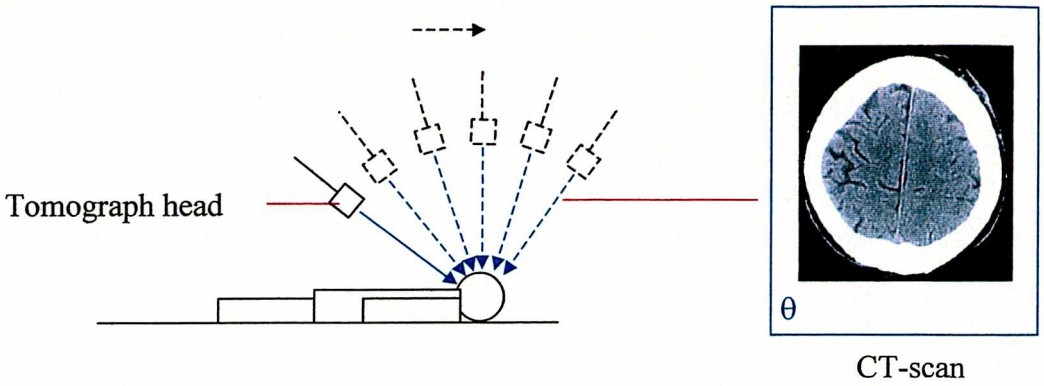


Figure 4.7 CT scanning

Example 4.5 (*Histology imaging: neurosurgery*)

The central nervous system (CNS) consists of neurones and glial cells (Figure 4.8). Neurones constitute about half the volume of the CNS and glial cells make up the rest. Glial cells provide support and protection for neurones. They are thus known as the 'supporting cells' of the nervous system. The four main functions of glial cells are: to surround neurones and hold them in place, to supply nutrients and oxygen to neurones, to insulate one neurone from another, and to destroy and remove the carcasses of dead neurones (clean up). The three types of CNS supporting cells are astrocytes, oligodendrocytes, and microglia.

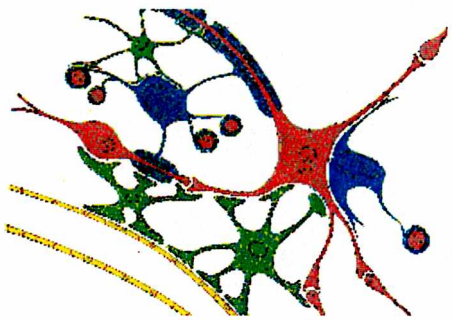


Figure 4.8 A diagram of the component cells of the Central Nervous System

- The subject of histology is study of tissues and cells under a microscope. In this example we illustrate the application of FERIS in image processing of microscopy samples. The original microphotographs of a culture of olfactory glial cells obtained from the olfactory bulb are shown in Figure 4.9(a, c). The identification and the characterisation of the olfactory glial cells are indispensable for their use in neurosurgery wrt the spinal cord injury treatment. This is because of their unique property for stimulating the process of neuroregeneration. Some results of the segmentation process of these cells are sown in Figures 4.9(b, d), respectively.

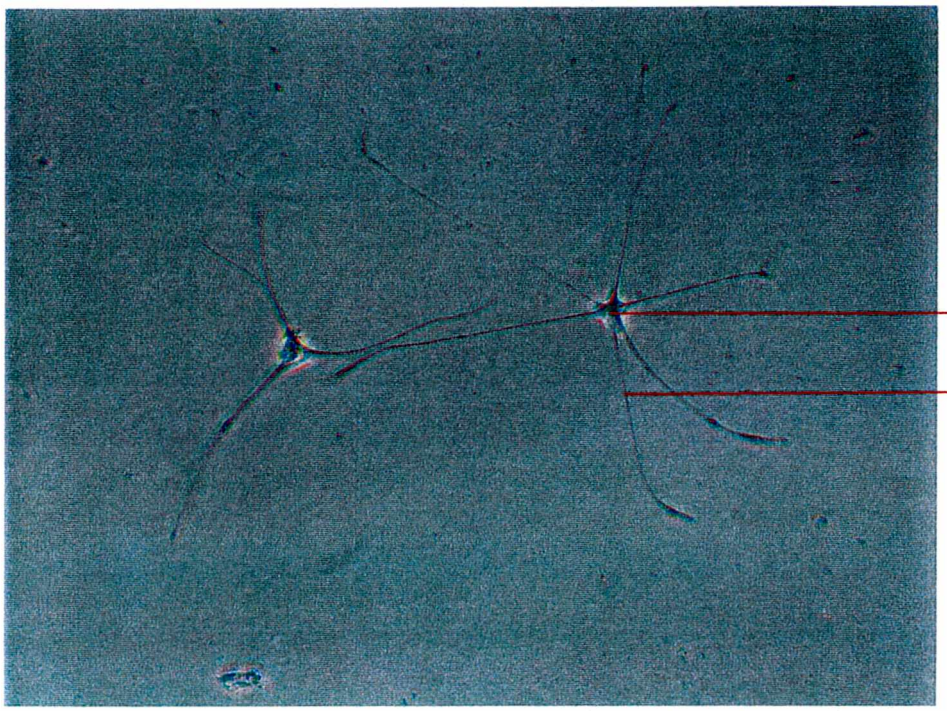
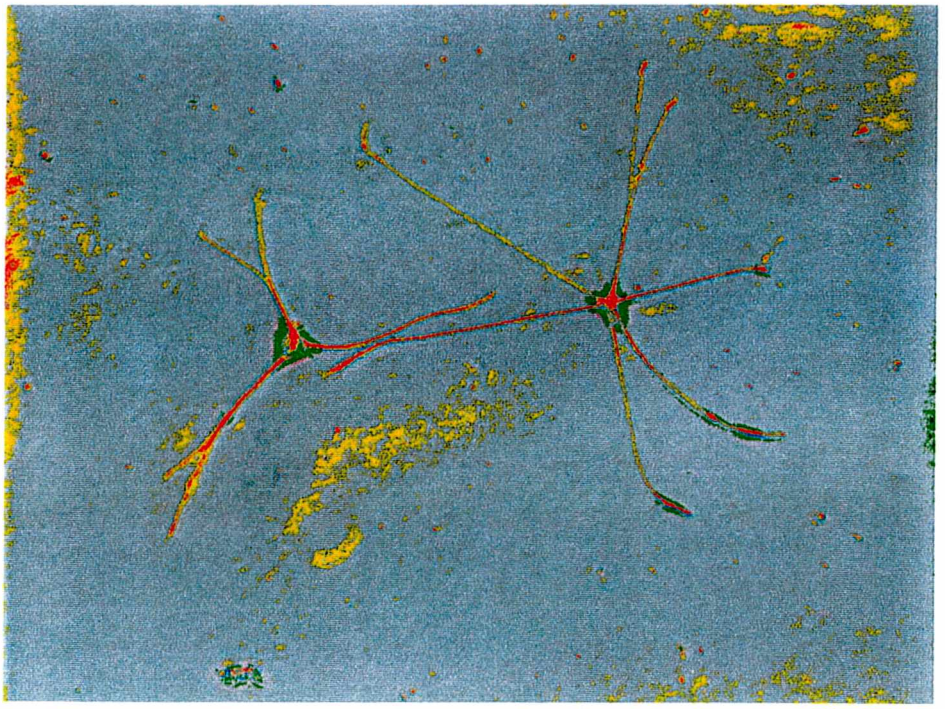


Figure 4.9(a) Microscopic image of olfactory glial cells



With red, orange and yellow colours are segmented the podia of the olfactory glial cells

With green colour are segmented the surroundings of the olfactory glial cells cytoplasm

Figure 4.9(b) The segmented image

The olfactory glial cells cytoplasm is identified in a unique way by their surroundings because of their 'star' shape. Here the following parameters have been used: a = 103, c = 123, e = 142, ε = .1 (dark, $grey_1$, bright), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ. Then the following classes have been obtained: {d – class} = {{35, 37, 38, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103}}, {g_1 – class} = {{104, 105, 106}, {107, 108, 109, 110, 111, 112, 113}} and {b – class} = {{142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 204, 205, 206, 207, 208, 209, 210, 211, 212, 220, 222}}.

- The next example segmentation is given in Figure 4.9(c, d), where the following parameters have been used: a = 93, c = 122, e = 159, ε = .1 (dark, $grey_1$, bright), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ. □

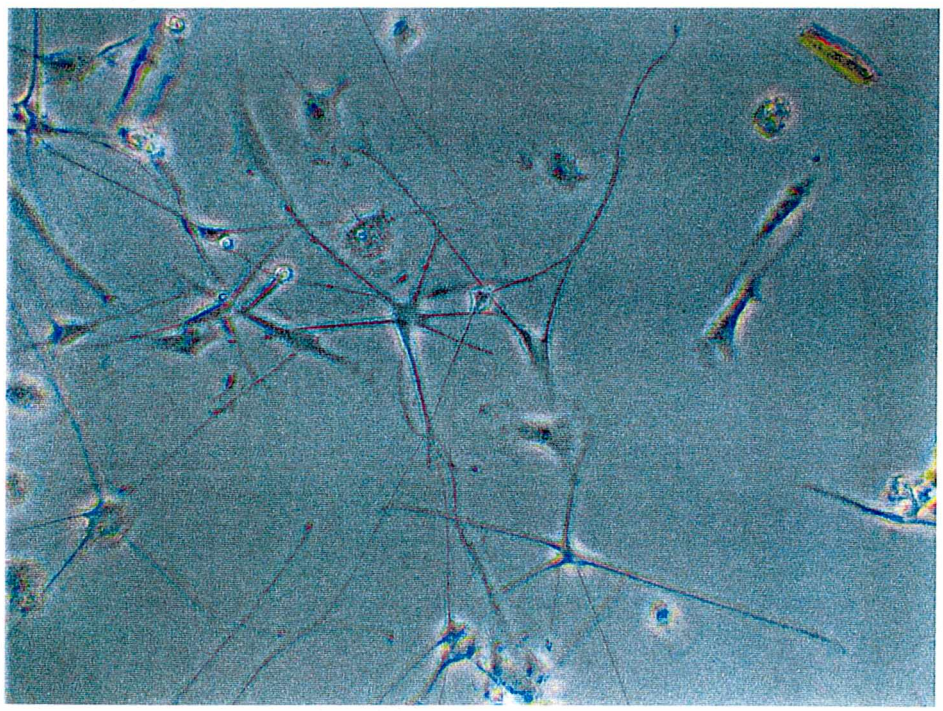
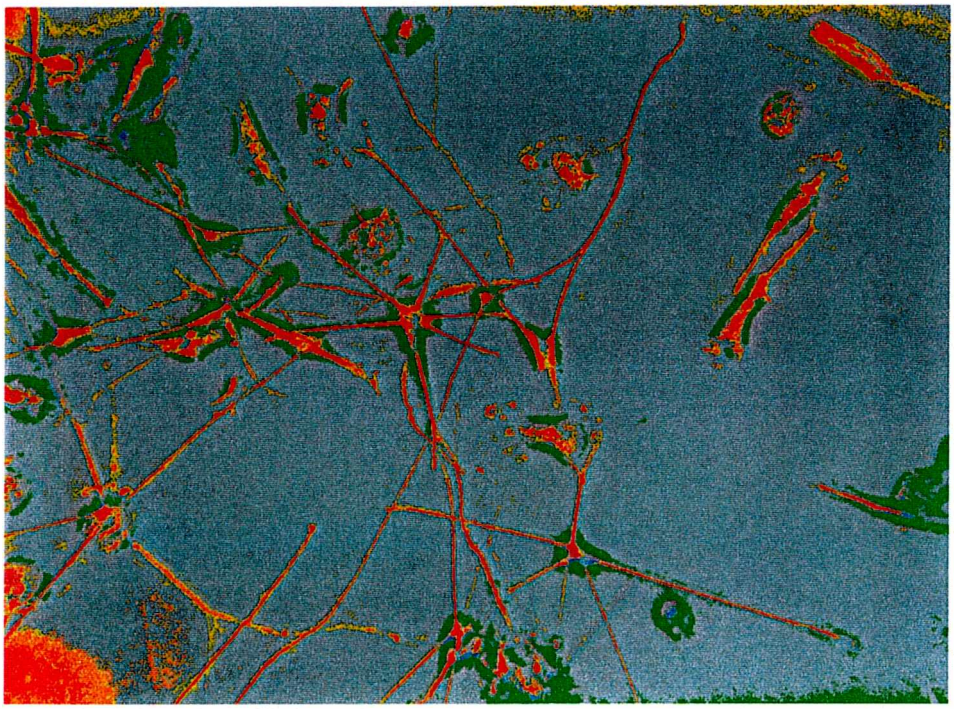


Figure 4.9(c) Microscopic image of mixed culture of glial cells



With red, orange and yellow colours are segmented the podia of the olfactory glial cells

With green colour are segmented the surroundings of the olfactory glial cells cytoplasm

Figure 4.9(d) The segmented image

The presented method of image segmentation may be easily extended to other areas of application. Below an illustration is given in the area of biology.

Example 4.6 (*Histology imaging: plant cells*)

- Let consider the plant cells shown in Figure 4.10(a). Any such cell is highly organized with many functional units or organelles. One or more membranes limit most of these units. Here the main segmentation task is a distinction of the corresponding cell membranes, wrt their degree of tearing (for a given fluid concentration) in different phases of analysis.

The following parameters have been used: a = 146, c = 174, e = 191, ε = .6 (dark, grey_1, grey_2), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ.

The following classes have been obtained: {d – class} = {{7, 9, 10, 16, 19, 20, 23, 26, 27, 28, 29, 31, 33, 34, 35, 36, 38, 39, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129}}, {g_1 – class} = {{130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154}, {155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170}} and {g_2 – class} = {{171, 172, 173, 174, 175, 176, 177, 178, 179, 180}}.

The segmented image is shown in Figure 4.10(b) below.

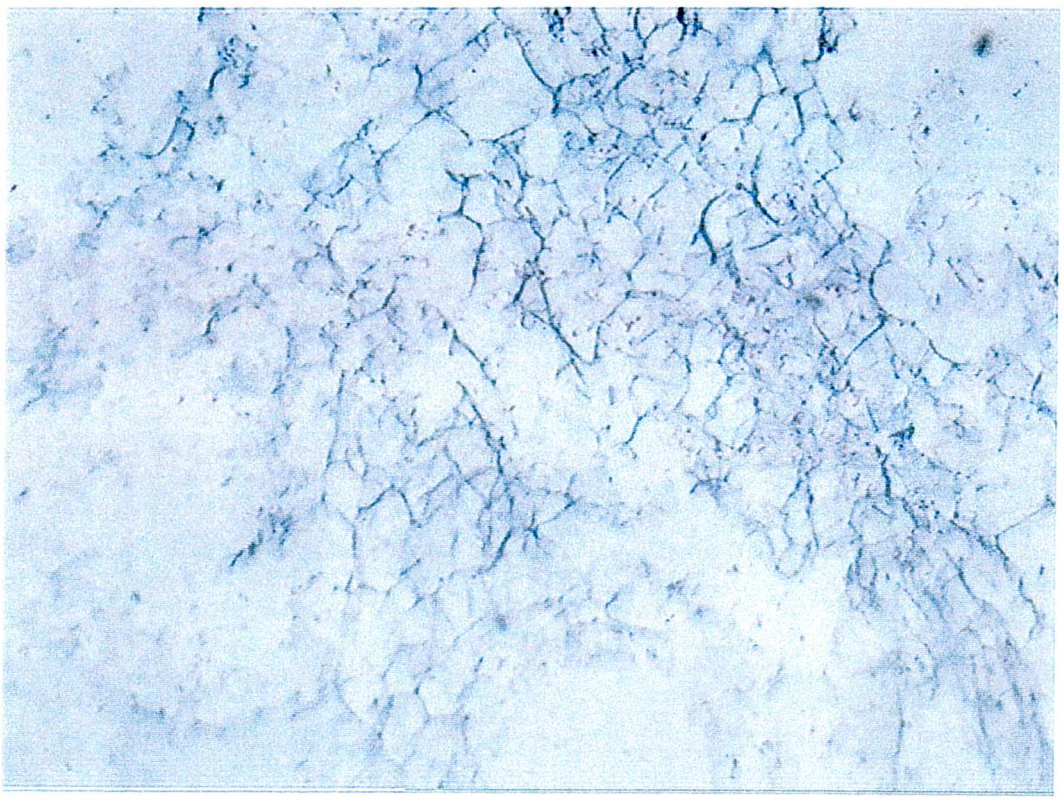


Figure 4.10(a) The microscopic image of plant cells

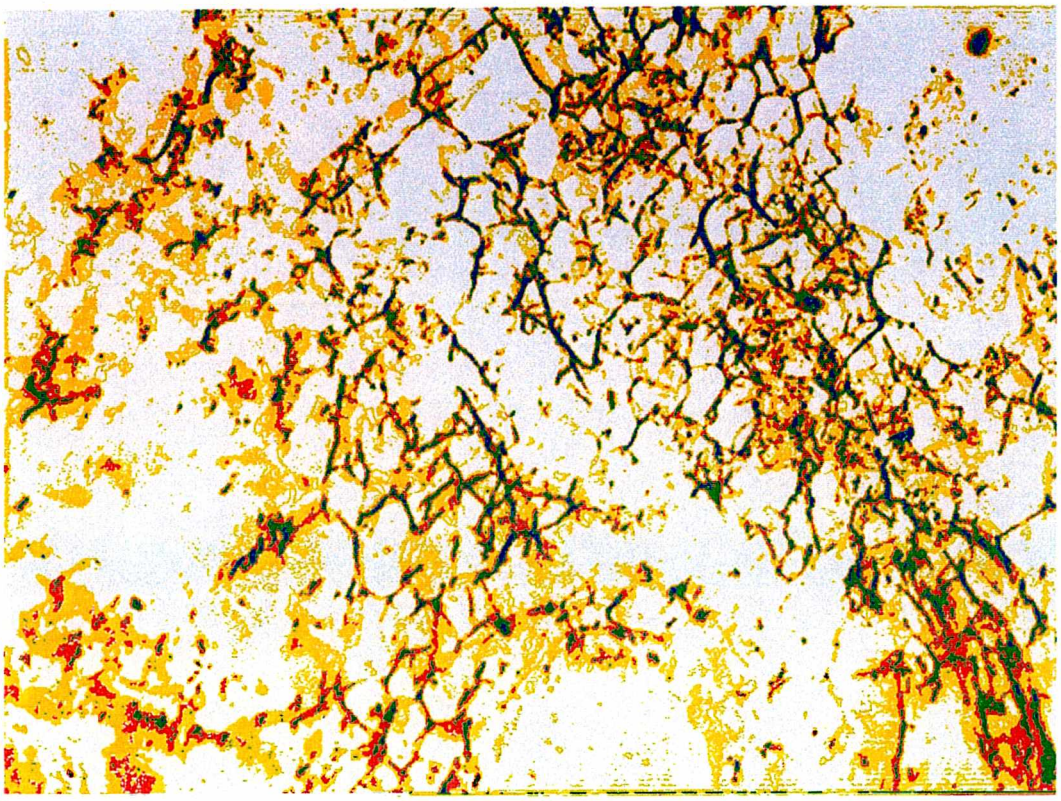


Figure 4.10(b) The segmented image

■ ■ ■ ■

With blue, green, red and yellow colours are segmented the cell membranes considering the increasing degree of tearing.

□

With grey colour is segmented the fluid between the cells.

- To illustrate the obtained quality of the segmentation process another type of cells have been considered as it is shown in Figure 4.10(c). The corresponding segmentation is illustrated in Figure 4.10(d).

The following parameters have been used: a = 109 c = 144, e = 193, ε = .6 (dark, $grey_1$, $grey_2$), ε = .45 ($grey_3$), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ.

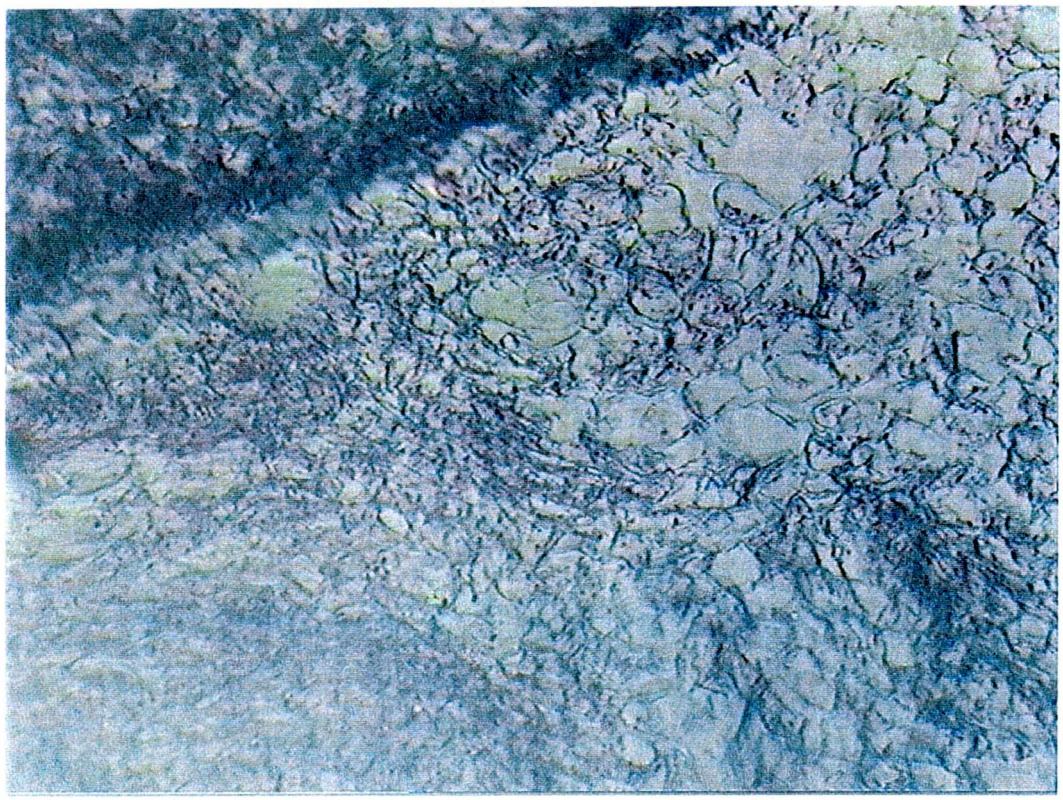


Figure 4.10(c) Microscopic image of plant cells

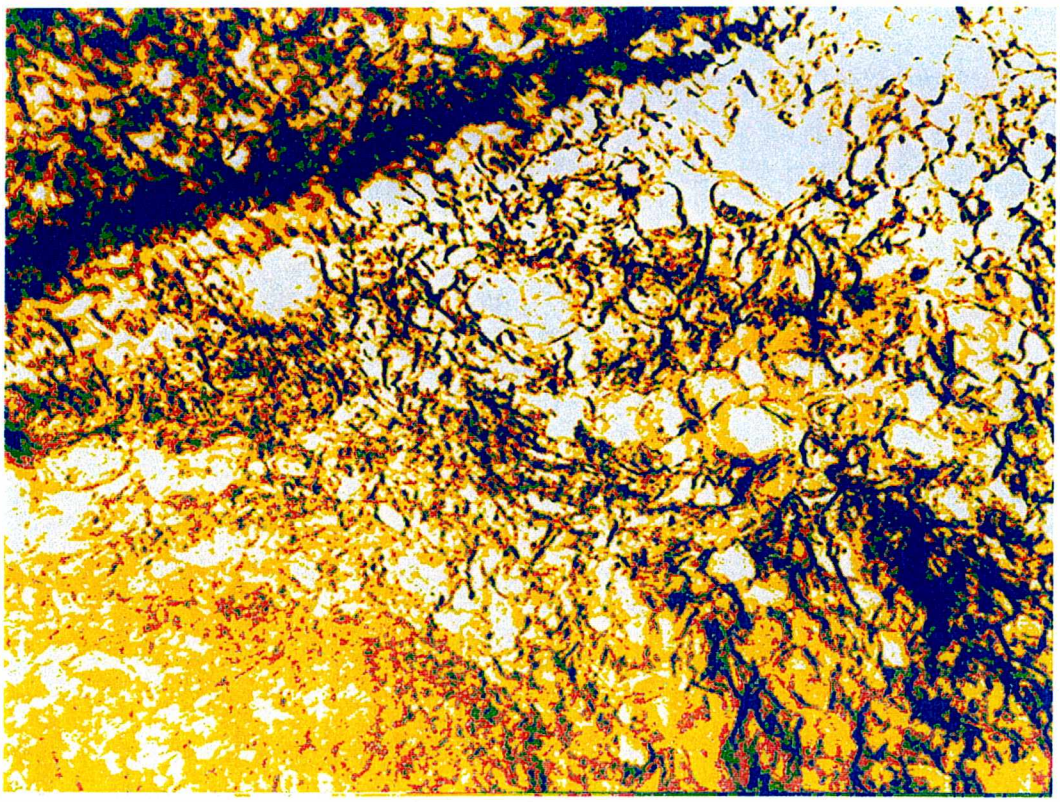


Figure 4.10(d) The segmented image

■ ■ ■ ■

With blue, green, red and yellow colours are segmented the cell membranes considering the increasing degree of tearing.

□

With grey colour is segmented the fluid between the cells. □

There exists some similarity between the above-analysed microscopic images of plant cells, shown in Figures 4.10(a, c), and also some types of histopathological images related to: emphysema, chronic bronchitis, asthma, etc. This similarity is illustrated in Figure 4.11 below.

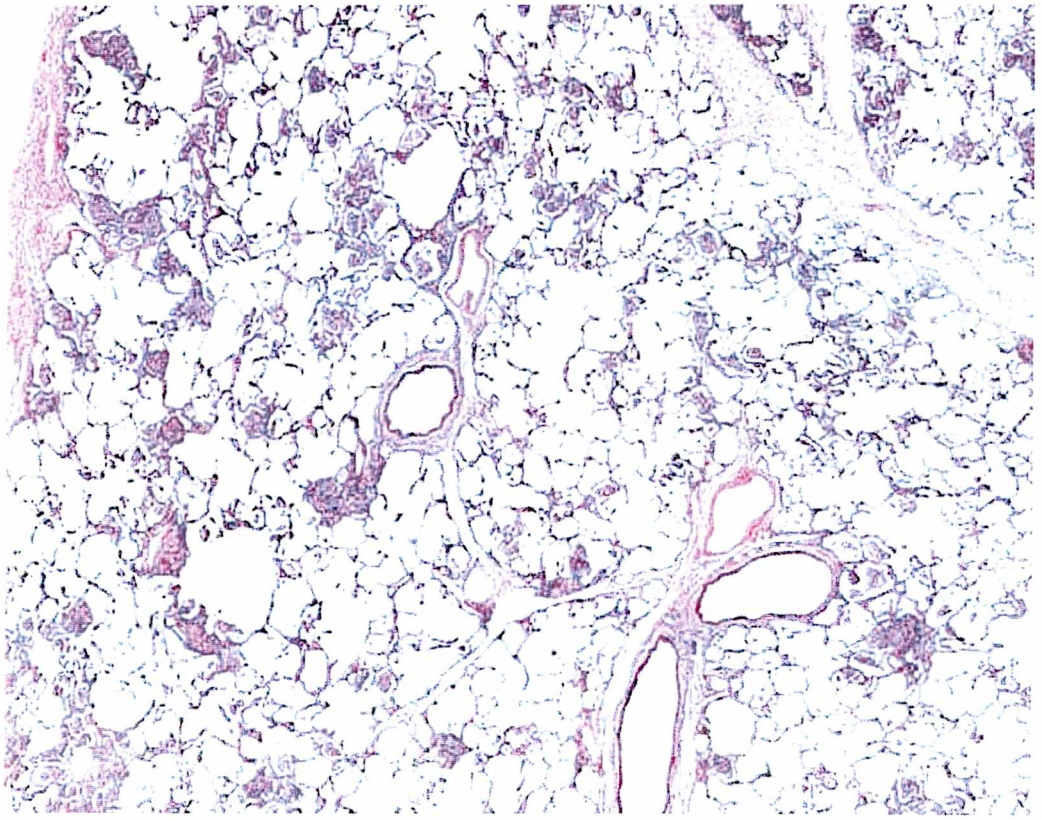


Figure 4.11 Example histopathological image

Here, two acini (terminal bronchioles and their alveoli) show interstitial infiltrates around bronchioles and a patchy interstitial and alveolar infiltrate.

Example 4.7 (Fluid impurity classification)[*]

This example is an illustration of a fully automated system for recognition of a fluid impurity, for use in pharmacy and medicine. The system configuration is shown in Figure 4.12 below. First, a sample of the considered fluid is applied on the input of the system. Then a microscopy imaging system, connected with a computer, scans the sample and partitions it into a set of images. Each image is analysed as far as the fluid impurity elements are segmented, recognised and classified. Some example images and the corresponding segmentation results are shown in Figures 4.13 (a - f). Three different original images are considered below (each of them having a different degree of impurity, shown in Figures 4.13(a, c, e)). The segmented images are given in Figures 4.13(b, d, f), respectively.

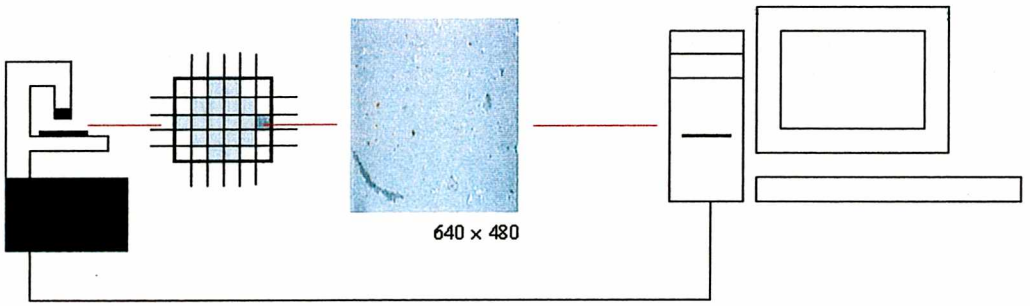


Figure 4.12 Fully automated system for fluid impurity analysis

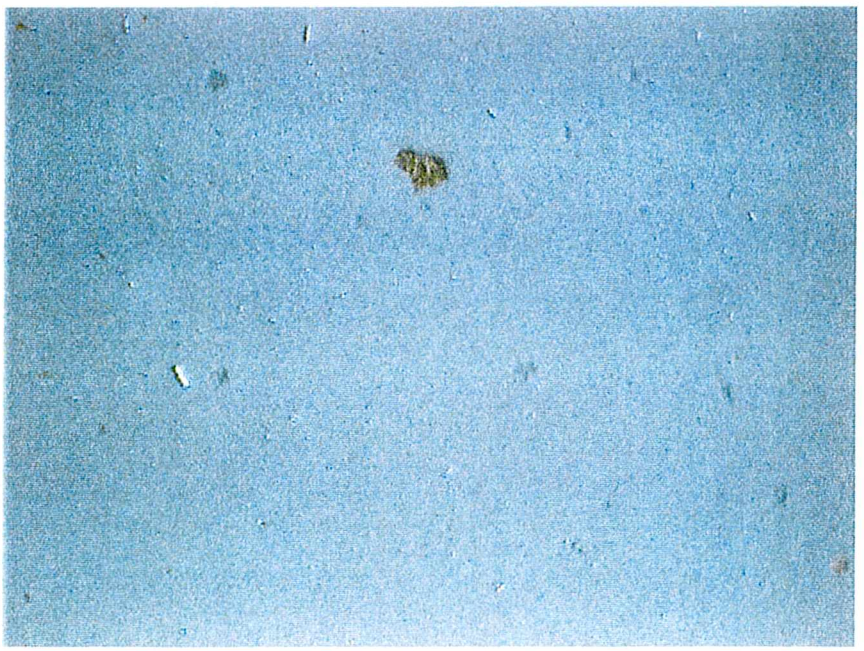
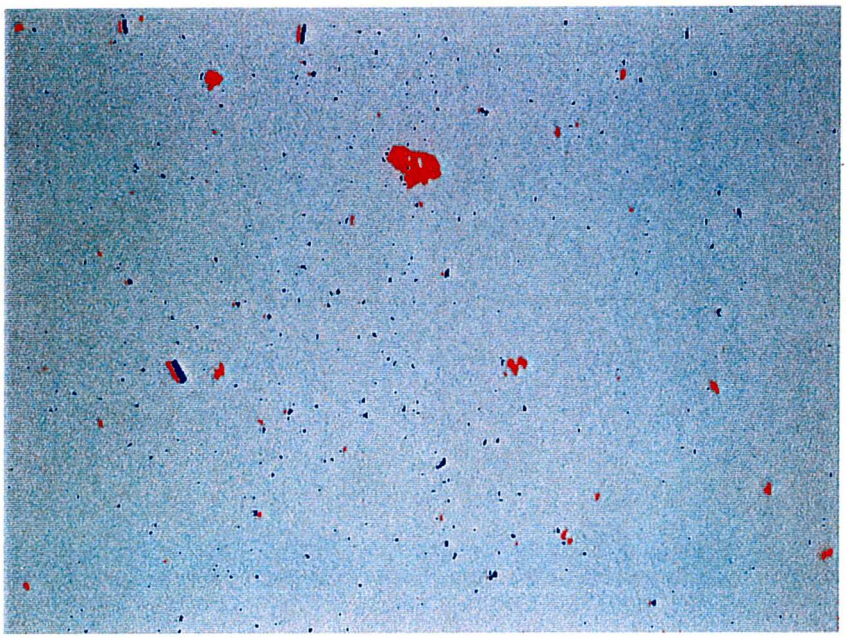


Figure 4.13(a) The first image of the fluid impurity

[*] A FERIS implementation for Stomil Sanok S.A., production section for pharmacy (ul. Reymonta 19, 38-500 Sanok, Poland).



With red and blue colours are segmented the fluid impurity elements.

Figure 4.13(b) The first segmented image

The following parameters have been used: a = 98, c = 132, e = 148, ε = .1 (dark, bright), Chebyshev measure of fuzziness for φ and a (modified) distance function of the Minkowski class for ρ. The corresponding parameters in the next two cases have been obtained in a similar way (this is omitted). □

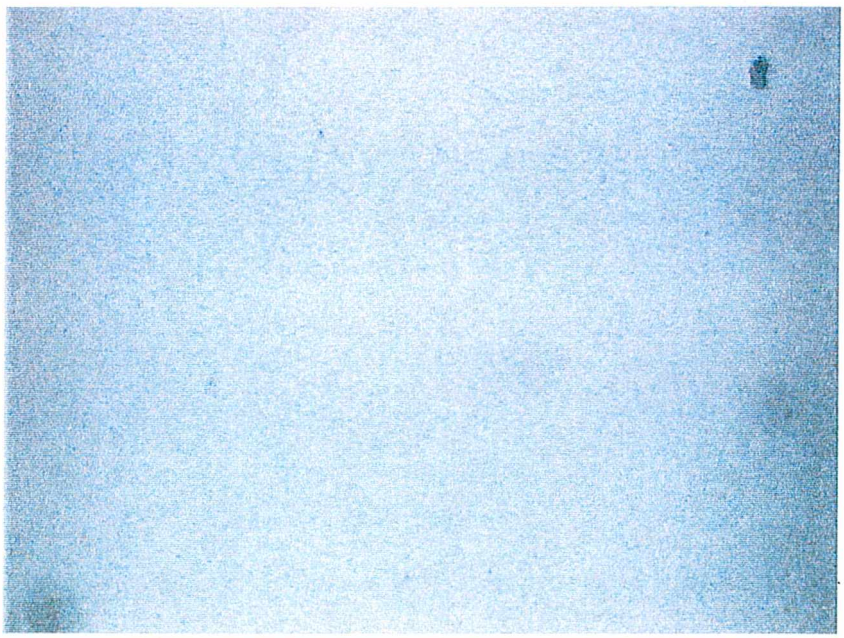
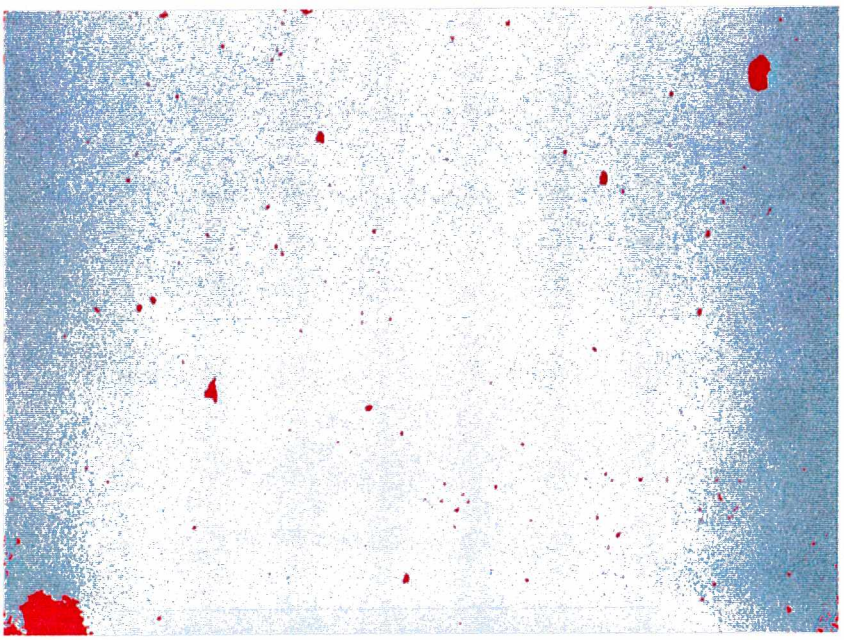


Figure 4.13(c) The second image of the fluid impurity



■

With red colour are segmented the fluid impurity elements.

Figure 4.13(d) The second segmented image

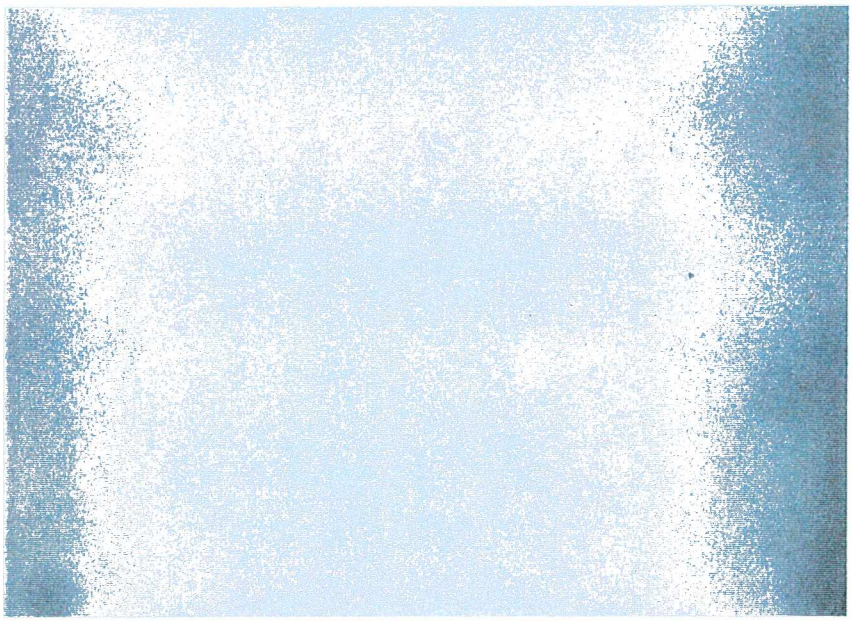
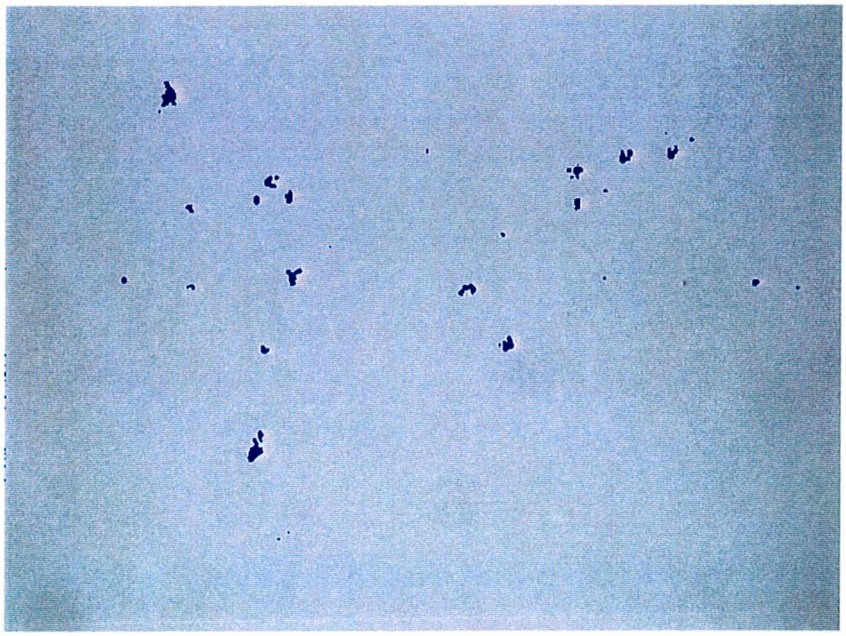


Figure 4.13(e) The third image of the fluid impurity



■ With blue colour are segmented the fluid impurity elements.

Figure 4.13(f) The third segmented image

The following observations may be useful in summarizing the presented experiments in this section. So, after the whole segmentation process some post processing operations, e.g. noise filtration and other morphological operations (such as erosion and dilatation) may be useful to eliminate elements (i.e. subsets of pixels corresponding to some regions of the segmented image, e.g. see Example 4.7) under a priori given requirements. Moreover, in accordance with the considerations given in the previous section, for a given class of images some single algorithm calibration is necessary to be done. This may implicate a restricted version of the introduced in Definition 3.10 cluster selection rules. For example, in addition to the global FEV, some local FEV-values may be very useful. In general, any such calibration will depend of the considered area of application of FERIS, i.e. the class of images under segmentation. But this is a problem in common for any such algorithm. And so, it is beyond this work to take on a complete analysis of this problem. Finally, in experiments the proposed method demonstrated a high quality and promising performance for various classes of medical images.

Conclusions

This chapter summarises the research contributions of this dissertation and discusses several areas for future work following from it.

This dissertation considers properties of fuzzy sets and relations and uses them as a foundation for the development of new algorithms for medical image segmentation.

The main areas from the theory of fuzzy sets and relations considered are: membership functions and matrices, distance functions of the Minkowski class, distance functions of the Canberra class, linear convex combinations of such distance functions, fuzzy similarity relations, transitive max-min closures, fuzzy equivalence relations, α-cuts and fuzzy expected values, fuzzy c-means clustering, fuzzification, measures of fuzziness and linear convex combinations of such measures. The primary contributions of the dissertation are the following:

- an exact version of the used image segmentation model;
- an introduction of new modified distance functions of the Minkowski and the Canberra classes in the process of image segmentation, and also some linear convex combinations of such functions;
- a histogram-based grey level fuzzification using corresponding classes of colours and a new, formally defined appropriate set of membership functions $\mathcal{M} =_{df} \{\mu_{dark}, \mu_{grey1}, \mu_{grey2}, \mu_{grey3}, \mu_{bright}\}$;
- a new histogram based ε-fuzzification preprocessing algorithm ;
- the proposed estimation of the used parameters a, b, c, d, e $\in$ [0,255] for any $\mu \in \mathcal{M}$ in an automated way by means of a fuzzy expected value or also a weighted fuzzy expected value;
- a formal definition of a linear convex combination of fuzzy measures;
- a formal definition of a restricted fuzzy similarity relation and the possibility of using fuzzy equivalence relations as a hierarchical clustering process (i.e. top-down image segmentation);
- a formal definition of the α-cut ρ_α using a (weighted) fuzzy expected value;
- a formal definition of the used set of cluster selection rules;

- a new fuzzy equivalence relation-based image segmentation linear structure algorithm of polynomial complexity (FERIS algorithm);
- a new possibility of using FERIS in the case of colour image segmentation by using the corresponding colour model component;
- a new possibility of obtaining a high speed segmentation by FERIS according to the classical systolic architecture for square matrices multiplication by using the logical operations minimum and maximum (instead of the arithmetical operations of multiplication and addition, respectively);
- some definitions and proofs concerning the above basic properties of fuzzy relations, distance functions and linear convex combinations of such functions (including 10 definitions and 12 propositions).

The major advantage of the FERIS algorithm lies in the ease of implementation. FERIS is a well-defined linear structure algorithm of polynomial complexity. The corresponding process of image segmentation is always realised in a unique way and the algorithm converges in a finite time. The obtained image segmentation is realised as an automated and resolution-independent computational process. The classical FCM algorithm requires that the desired number of clusters be given in advance. This can be problematic when the clustering problem does not specify any desired number of clusters. The number of clusters should reflect the stricture of the given data. The above-proposed method is based on using a properly defined fuzzy equivalence relation. And so, it satisfies this need. Moreover, in contradistinction to the classical FCM algorithm, FERIS can be implemented for real-time applications. In experiments, this method demonstrated a high quality and promising performance for various classes of medical images. These experiments exhibited that the proposed algorithm is very robust to noise, spatial and temporal inhomogeneities. And finally, the presented method of image segmentation may be easily extended to other areas of application, e.g. such as: biology, geology, meteorology, urbanisation, chemistry, and so on, i.e. everywhere, where image processing is needed. In the last case, only some initial calibration of FERIS is necessary to be done.

References

[1] Bezdek J.C., *Pattern recognition with fuzzy objective function algorithms*. Plenum Press, New York (1981) 256pp.

[2] Bloch I., *Fuzzy classification for multi-modality image fusion*. Int. Conf. on Image Processing ICIP 1994,Austin, Texas,USA, Nov.13 – 16,vol.1 (1994)628 – 632.

[3] Chen, J.L. and Wang J.H., *A new robust clustering algorithm-density-weighted fuzzy c-means*. Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, Tokyo, Japan (1999) 90 - 94.

[4] Clark M.C., *Segmenting MRI volumes of the brain with knowledge-based clustering*. MS.Thesis, University of South Florida USA (1994) 76pp.

[5] Diday E. and Simon, J.C., *Clustering analysis*. Digital Pattern Recognition, Fu K.S. ed. , Springer-Verlag, Berlin (1980) 47 - 94.

[6] Fradkin, M., Roux M., Maitre H. and Leloglu U.M., Surface reconstruction from multiple aerial images in dense urban areas. Proceedings of IEEE Computer Society Conference on CVPR, Fort Collins, Colorado,USA (1999) 264 - 267.

[7] Frigui H. and Krishnapuram R., A robust competitive clustering algorithm with applications in computer vision. IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.21,no.5 (1999) 450 – 465.

[8] Helgason C.M., Jobe T.H.,Malik D.S. and Mordeson J.N., *Analysis of stroke pathogenesis using hierarchical fuzzy clustering techniques*. 5th International Conference on Information Systems, Analysis and synthesis: ISAS'99 Orlando, Florida USA (1999) 477 –482.

[9] Helgason C.M., Jobe T.H., Mordeson J.N., Malik D.S. and Cheng S.C., *Discarnation and interactive variables as conditions for disease*. The 18th International Conference of the North American Fuzzy Information Society NAFIPS'99, New York City USA (1999) 298 – 303.

[10] Kerre E.E. and Nachtegael M. eds., *Fuzzy Techniques in image processing*. Physica Verlag, Heidelberg,New York (2000) 413pp.

[11] Klir G.J. and Yuan B., *Fuzzy sets and fuzzy logic. Theory and application.* Prentice Hall, Upper Saddle NJ (1995) 574pp.

[12] Kwiatkowski J., Kwiatkowska W., Kawa K. and Tabakow M., *The computerised system for measurement of the intima-media thickness.* Proc. of the Conf. on Medical Informatics and Technologies. MIT 2000.,Ustroń, Poland (2000) 31 – 39.

[13] Li Y.L., Dong H. and Gao X., *A fuzzy relation based algorithm for segmenting color aerial images of urban environment.* Integrated System for Spatial Data Production, Custodian and Decision Support Symposium ,Xi'an, P.R.China (2002) 271 – 274.

[14] Łęski J., Fuzzy clustering with ε -insensitive loss function. Conference on Computer Recognition Systems KOSYR'01, Milków, in Polish (2001) 81 – 86.

[15] Marchisio G.B., Koperski K. and Sanella M., *Querying remote sensing and GIS repositories with spatial association rules.* International Geoscience and Remote Sensing Symposium IGARSS\'00, Honolulu, Hawaii, USA (2000) 3054 – 3056.

[16] Melek W. W., Emami M.R. and Goldenberg A.A., 1999. *An improved robust fuzzy clustering algorithm.* Proceedings of IEEE International Conference on Fuzzy Systems, Seoul, South Korea (1999) 1261 - 1265.

[17] Mohamed N.A., Ahmed M.N. and Farag A., *Modified fuzzy c-mean in medical image segmentation.* Alliant Health System and the NSF grants ESC-9505674 USA (1998) 4pp.

[18] Mordeson J.N., Malik D.S. and Cheng S.C., *Fuzzy mathematics in medicine.* Physica Verlag, Heidelberg,New York (2000) 256pp.

[19] Noordam, J. C., Van Den Broek W.H.A.M. and L .M. C. Buydens L.M.C., *Geometrically guided fuzzy c-means clustering for multivariate image segmentation.* Proceedings of 15th International Conference on Pattern Recognition, Barcelona, Spain (2000) 462 -465.

[20] Pal N.R. and Bezdek J.C., *On cluster validity for the fuzzy c-means model.* IEEE Trans. on fuzzy systems 3 (1995) 370 – 379.

[21] Pedrycz W., *Conditional fuzzy c-means.* Pattern Recognition Letters 17 (1996) 625 – 631.

[22] Pham D.L. and Prince J.L., *Adaptive fuzzy segmentation of magnetic resonance images.* IEEE Trans. on Med. Imag., vol.18, no.9, (1999) 737 – 752.

[23] Pham T., Wagner M. and Clark D., *Applications of genetic algorithms, geostatistics and fuzzy c-means clustering to image segmentation.* Proceedings of the Congress on Evolutionary Computation Seoul, South Korea (2001) 741 – 746.

[24] Schneider M. and Craig M., *On the use of fuzzy sets in histogram equalization.* Fuzzy Sets and Systems 45, North-Holland (1992) 271 – 278.

[25] Sheen D., *Lecture notes on numerical linear algebra.* South Korea, Seoul National University (2001).

[26] Suri J.S., Setarehdan S.K. and Singh S. eds., *Advanced algorithmic approaches to medical image segmentation*. Springer-Verlag, London Ltd. (2002) 660pp.

[27] Tabakow M., *Using fuzzy set theory in medical image information processing: basic notions and definitions.* Reports of the Department of Computer Science of Wroclaw University of Technology ser. PRE nr 7/01, Wroclaw, in Polish (2001) 35pp.

[28] Tabakow M., *Using fuzzy sets and box – counting dimension in medical image processing.* Conference on Computer Recognition Systems KOSYR'01, Milków, in Polish (2001) 185 – 189.

[29] Tabakow M., *Medical image segmentation algorithms using a fuzzy equivalence relation.* Reports of the Department of Computer Science of Wroclaw University of Technology ser. PRE nr 11/01, Wroclaw, in Polish (2001) 14pp.

[30] Tabakow M., *Medical image segmentation algorithms using a fuzzy equivalence relation.* Conference on Computer Recognition Systems KOSYR'03, Milków, in Polish (2003) 239 – 243.

[31] Thitimajshima, P., 2000. *A new modified fuzzy c-means algorithm for multispectral satellite images segmentation.* International Geoscience and Remote Sensing Symposium IGARSS\'00, Honolulu, Hawaii, USA (2000) 1684 - 1686.

[32] Tizhoosh, H.R. *Fuzzy Image Processing: Potentials and State of the Art.* IIZUKA'98, 5th International Conference on Soft Computing, Iizuka, Japan, October 16-20,vol.1 (1998) 321 – 324.

[33] Tu, Z., Zhu S.C. and Shum H.Y., 2001. *Image segmentation by data driven Markov chain Monte Carlo.* Proceedings of 8th IEEE International Conference on Computer Vision, Vancouver, British Columbia, Ca (2001) 131 - 138.

[34] Tyagi, A. and Bayoumi M., 1989. *A systolic array for image segmentation using split and merge procedure.* Proceedings of the 32nd Midwest Symposium on Circuits and Systems, Champaign, Illinois, USA (1989) 345 - 348.

[35] Wolfram S. et al., *Digital image processing. user's guide.* Chapter 7., Wolfram Research, Inc.USA (2003).

[36] Xie X.L. and Beni G., *A validity measure for fuzzy clustering.* IEEE Transactions on Pattern Analysis and Machine Intelligence.8 (1991) 841—847.

[37] Zadeh L., *Fuzzy sets*, Information and Control 8, (1965) 338-353.

Appendix A: Mathematical notions

A mathematical background used in this work is given below. Only some introductory notions concerning the classical set theory are presented. Any other notions, e.g. such as functions and operations, algebraic systems, vectors and matrices, etc. are omitted. A more complete description of the considered material is contained in any good treatment of discrete mathematics (in particular in any literature related to discrete mathematical structures).

A1. Sets

As usual we use, for sets X, the notation $x \in X$ and $Y \subseteq X$ to denote that x is an *element* of X and Y is a *subset* of X, i.e. $Y \subseteq X \Leftrightarrow_{df} \forall x\,(x \in Y \Rightarrow x \in X)$. We shall say that $X = Y \Leftrightarrow_{df} \forall x\,(x \in X \Leftrightarrow x \in Y)$, i.e. iff $X \subseteq Y$ and $Y \subseteq X$. Also $Y \subset X \Leftrightarrow_{df} Y \subseteq X \wedge Y \neq X$. $\mathcal{P}(X) =_{df} \{Y / Y \subseteq X\}$ denotes the *power set* of X, e.g. $\mathcal{P}(\{a,b,c\}) = \{\varnothing,\{a\},\{b\},\{c\},\{a,b\},\{a,c\},\{b,c\},\{a,b,c\}\}$.

Let X and Y be sets. As usual, $X \cup Y$, $X \cap Y$ and $X - Y =_{df} \{x \in X / x \notin Y\}$ denote the *union* of X and Y, the *intersection* of X and Y and the *set difference*, i.e. the complement of Y in X. In general, the *complement* of a set X to the universal set (i.e. the *universe*) $\mathcal{U}$ is denoted by $X' =_{df} \mathcal{U} - X$. Since any $x \in \mathcal{U}$: $x \in X' \Leftrightarrow x \notin X$. Also we have: $x \in X \cup Y \Leftrightarrow_{df} x \in X \vee x \in Y$ and $x \in X \cap Y \Leftrightarrow_{df} x \in X \wedge x \in Y$. The *generalised union and intersection* of an arbitrary (finite or not) sequence of sets are defined as follows: $x \in \bigcup_{i \in I} X_i \Leftrightarrow_{df} \exists i \in I\,(x \in X_i)$ and $x \in \bigcap_{i \in I} X_i \Leftrightarrow_{df} \forall i \in I\,(x \in X_i)$, where $I \subseteq \mathcal{N}$. *De Morgan's laws* are given by: $(\bigcup_{i \in I} X_i)' = \bigcap_{i \in I} X_i'$ and $(\bigcap_{i \in I} X_i)' = \bigcup_{i \in I} X_i'$.

A2. Relations

The *(ordered) pair* of elements x,y is denoted by (x,y), in general $(x,y) \neq (y,x)$. The *Cartesian product* of X and Y is denoted by $X \times Y =_{df} \{(x,y) / x \in X \wedge y \in Y\}$. Let

$X_1,...,X_n$ be sets. So, the generalised form of the last product $X_1 \times ... \times X_n =_{df} \{(x_1,...,x_n) / x_1 \in X_1 \wedge ... \wedge x_n \in X_n\}$, where the n-tuple $(x_1,...,x_n)$ is denoted by $\underline{x} =_{df} (x_1,...,x_n)$. The Cartesian product of k copies of X is denoted by X^k, $X^1 = X$. Any subset ρ of the Cartesian product $X \times Y$ is said to be a *binary relation* between the elements of the sets X and Y. Hence $\rho \subseteq X \times Y$ and we write $x\rho y$ for $(x,y) \in \rho$. The *opposite relation* to ρ is denoted by $\rho^{-1} =_{df} \{(x,y) / (y,x) \in \rho\} \subseteq Y \times X$. The *complement* of ρ is denoted by ρ', i.e. $(x,y) \in \rho'$ iff $(x,y) \notin \rho$. The *domain* of ρ is denoted by $dom(\rho) =_{df} \{x \in X / \exists y \in Y (x\rho y)\}$. The *codomain* of ρ (called also image or range of ρ), i.e. $cod(\rho) =_{df} dom(\rho^{-1})$.

Next we shall concentrate our attention mainly to the notions of order and equivalence relations (some other notions, e.g. such as irreflexivity, asymmetry, connectivity, similarity, transitive closure, etc. are omitted below).

Let ρ be a *binary relation over* X, i.e. $\rho \subseteq X \times X$. For any $x,y,z \in X$, we shall say ρ is: (i) *reflexive* iff $x\rho x$, (ii) *symmetric* iff $x\rho y \Rightarrow y\rho x$, (iii) *antisymmetric* (weak antisymmetry) iff $x\rho y \wedge y\rho x \Rightarrow x = y$, (iv) *transitive* iff $x\rho y \wedge y\rho z \Rightarrow x\rho z$. We shall say ρ an *order relation* (also called a *partial order*) iff (i),(iii) and (iv) are satisfied. We shall say X is a *partial ordered set* wrt ρ (in short: a *poset*). As usual the partial order relation ρ is denoted by $\geq$. The poset X wrt $\geq$ is denoted by $(X,\geq)$. The *strong order* is defined as: $x > y$ iff $x \geq y \wedge x \neq y$. Any finite poset can be presented as a graph such that an edge (x,y) is drawn iff $x > y \wedge \sim \exists z (x > z > y)$.

Also ρ is an *equivalence relation* iff (i),(ii) and (iv) are satisfied. Any equivalence relation generates some partition over X (and vice versa). Any element $[x]_\rho \subseteq X$ of this partition is said to be an *equivalence class* under ρ. So $y \in [x]_\rho \Leftrightarrow_{df} y\rho x$. The *quotient set* wrt ρ is denoted by $X/\rho =_{df} \{ [x]_\rho / x \in X \}$.

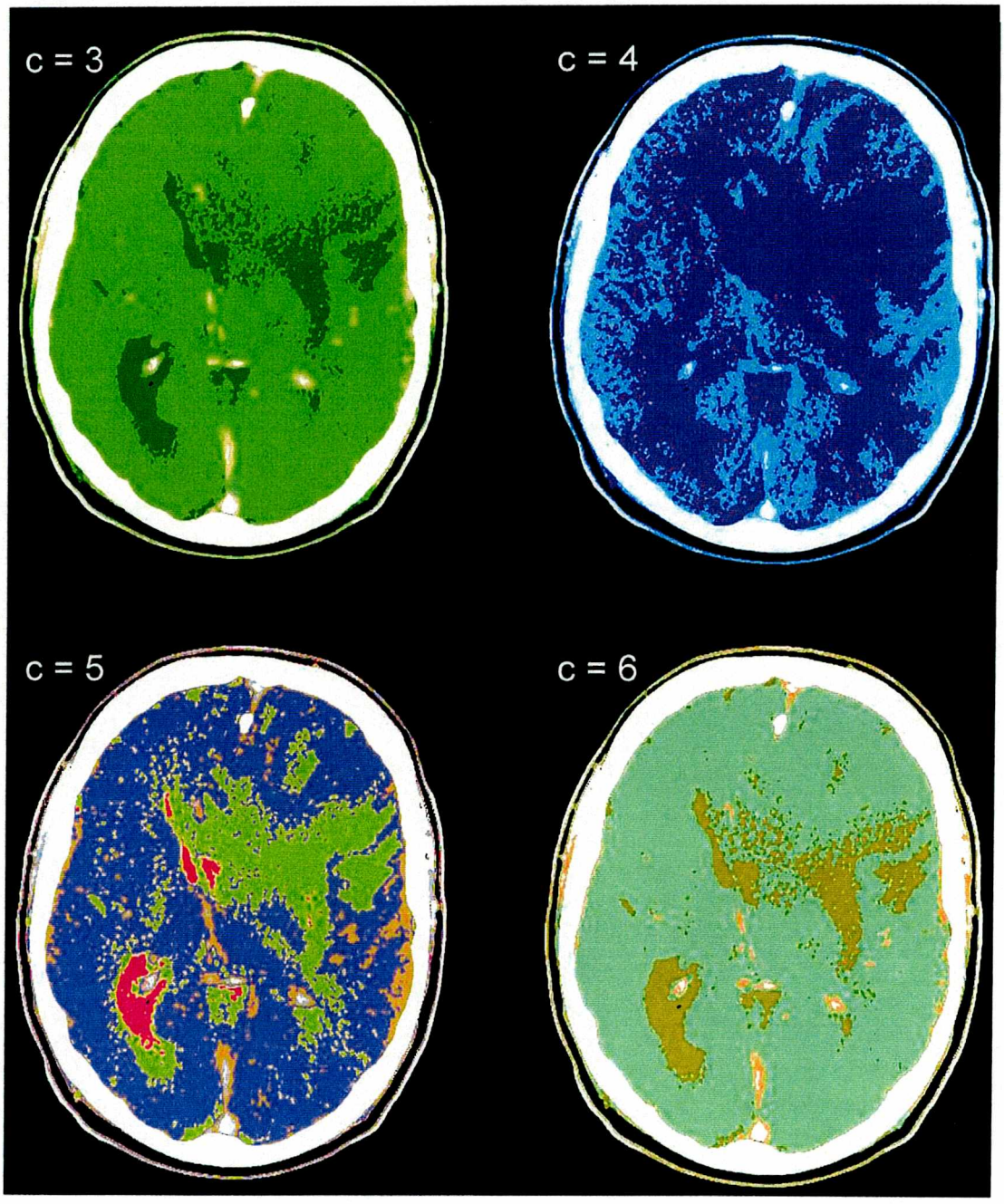
Let $\rho \subseteq X \times Y$ and $\sigma \subseteq Y \times Z$ be two binary relations. The *composition* $\rho \circ \sigma \subseteq X \times Z$, of ρ and σ is defined as follows: $x (\rho \circ \sigma) z$ iff $\exists y \in Y (x\rho y \wedge y\sigma z)$, for any $x \in X$ and $z \in Z$.

Let A be a set and $\mathcal{P}$ be a collection of nonempty subsets of A. Then $\mathcal{P}$ is a *partition* of A iff (1) for all $B, C \in \mathcal{P}$, either $B = C$ or $B \cap C = \varnothing$ and (2) $A = \bigcup B / B \in \mathcal{P}$.

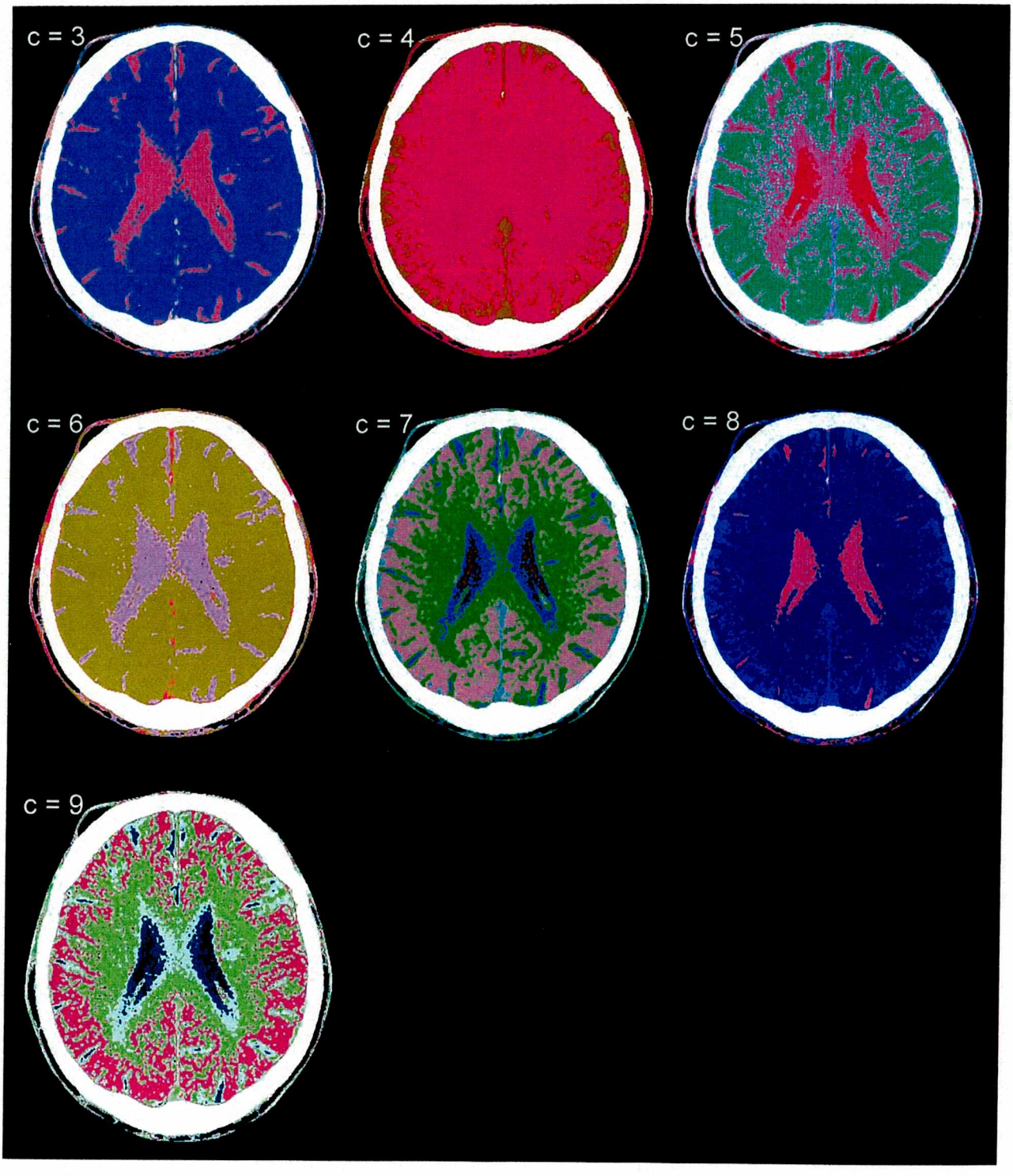
It can be shown that any equivalence relation generates some partition and vice versa.

Appendix B: Example fuzzy c-means image segmentation

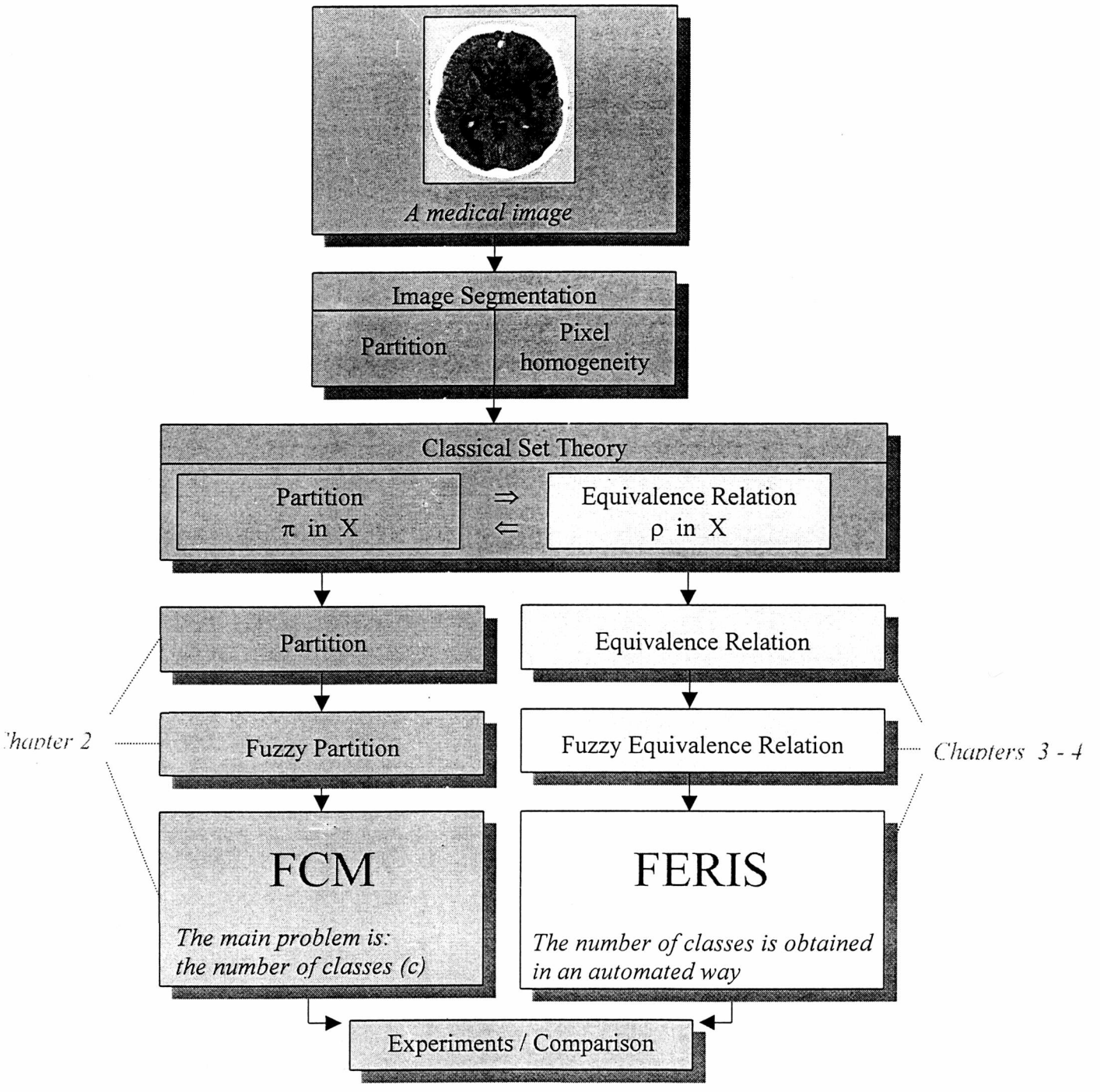
B1. This is an illustration of using FCM for $\mathcal{I}_{o1}$ of Figure 4.5(a). As it is shown below, a similar segmentation to $\mathcal{I}_{s1}$ of Figure 4.5(b) (under FERIS, see Example 4.4: *CT imaging*) can be obtained assuming the cardinality of the corresponding initial pseudopartition $c = 5$.



B2. This is an illustration of using FCM for $\mathcal{I}_{o4}$ of Figure 4.6(a). As it is shown below, a similar segmentation to $\mathcal{I}_{s4}$ of Figure 4.6(b) (under FERIS, see Example 4.4: *CT imaging*) can be obtained assuming the cardinality of the corresponding initial pseudopartition c = 9.



Appendix C: Main idea and structure of this dissertation



Index

Appendix: mathematical notions