

TECHNICAL UNIVERSITY OF CRETE

DIPLOMA THESIS

Automated Classification of Pollen Grains from Microscope Images using Computer Vision and Semantic Web Technologies

Ioannis KOUTSOUKOS

INTELLIGENT SYSTEMS LABORATORY

School of Electronic and Computer Engineering

Diploma Thesis Committee

Professor Euripides G.M. Petrakis (supervisor).

Professor Michalis Zervakis.

Associate Professor Michail G. Lagoudakis.

November 2013

TECHNICAL UNIVERSITY OF CRETE

Abstract

Intelligent Systems Laboratory
School of Electronic and Computer Engineering

Undergraduate Student

Automated Classification of Pollen Grains from Microscope Images using Computer Vision and Semantic Web Technologies

by Ioannis KOUTSOUKOS

The work aims at the design and development of an image information system for the automatic recognition and classification of pollen grains. The collected samples will be photographed, magnified many times (sometimes thousands) and finally classified into genres: Once pollen has been identified, the species of the plant that produced it will be determined. This can be achieved with the aid of image analysis and pattern recognition techniques.

This is a rather complicated process because of the fact that the granules are three-dimensional, while the analysis has to be based on microscope images which are essentially two-dimensional. Multiple samples of each type of pollen taken under different orientations and viewpoints will be combined (the same granule might seem very different if seen from a different angle) to achieve a unique categorization of each pollen grain. The system was trained with images from the collection of reference samples of pollen grains of the Mediterranean Plant Conservation Unit (MPCU) of the Mediterranean Agronomic Institute of Chania (MAICh).

Digitization is realized using a light microscope. This collection consists of over 230 different species of flora, mainly originating from Crete. Machine vision techniques are applied to segment pollen images into regions (each corresponding to the region of grain) and subsequently acquire a set of features (mainly chromatic, texture and shape features) from each pollen grain.

The image database of pollen grains is organized as an ontology containing raw images and their feature sets organized in a categorization (by genre) of the pollen species, text descriptions and relationships between the genres of the species (e.g., similarities between them). Image recognition was implemented as a retrieval process by comparing vectors of low-level descriptors extracted from the input image and representative images of each species in the ontology respectively. A similarity measure was used to compute the similarity of an unknown (input or query image) and images in the ontology. The result list is ranked by decreasing similarity to the unknown image. Several ranking methods are used to estimate the species category where the image belonged to. The experimental results demonstrate that approximately 75-80% of the input images are correctly recognised. Our system works with an image dataset consisting of 501 images of pollen grains, belonging to 40 species, 35 genus and 19 families.

Acknowledgements

I would like to thank Prof. Euripides G.M. Petrakis for the cooperation and guidance throughout this work, Prof. Michalis Zervakis for his contribution and advices, Dr Sotirios Kampranis for his contribution on the initial idea for this work and his help on cooperating with MAICH, and finally Mrs Panagiota Gotsiou, Biologist B.Sc., M.Sc., Research Assistant of the MPCU of MAICH for her help and supervision on the sampling process of the pollen grains.

Contents

Abstract	ii
Acknowledgements	iv
List of Figures	vii
1 Introduction	1
2 Background and Related Work	3
2.1 Pollen	3
2.1.1 Angiosperms	3
2.1.2 Angiosperms' Life Cycle	4
2.1.3 Gamete Production in Angiosperms	4
2.1.4 Meiosis	6
2.1.5 Fertilization	6
2.1.6 Pollen Grain Structure	7
2.1.7 Pollen Grain Characteristics and Classification	9
2.2 Machine Vision and Image Processing	13
2.2.1 Color	13
2.2.2 Image Segmentation and Thresholding	17
2.2.3 Morphological Operators	20
2.2.4 Component Labeling	22
2.3 Ontologies	23
2.4 Related Work	23
2.5 Image Content Analysis	25
2.5.1 Color Descriptors	25
2.5.1.1 MPEG-7 Color Descriptors	25
2.5.1.2 Color (Histogram) Descriptors	31
2.5.2 Texture Descriptors	35
2.5.2.1 Edge Histogram Descriptor (EHD)	36
2.5.2.2 Tamura Descriptor (TD)	37
2.5.2.3 Gabor Descriptor (GD)	38
2.5.3 Hybrid Descriptors:	40
2.5.3.1 Color Edge Directivity Descriptor (CEDD)	40
2.5.3.2 Fuzzy Color Texture Histogram Descriptor (FCTHD)	41

2.5.3.3	Joint Composite Descriptor (JCD)	42
2.5.4	Geometrical Descriptors	43
3	Methodology	44
3.1	Sampling Method	44
3.1.1	Preparation of Pollen Reference Specimens	44
3.1.2	Laboratory Equipment	45
3.1.3	Image Capturing and Settings	47
3.2	Segmentation	48
3.2.1	Use of the HSI color model	48
3.2.2	Thresholding / Binarization	49
3.2.3	Morphological Operators	51
3.2.4	Component Labeling	53
3.2.5	Remarks	53
3.3	Ontology Construction	54
3.3.1	Pollen Ontology	54
3.4	Image Similarity	57
3.4.1	Feature Extraction	57
3.4.2	Gaussian Normalization	57
3.4.3	Image Similarity Measures	58
3.5	Pollen Species Estimation	59
3.5.1	Best Match	59
3.5.2	Max Occurrence	59
3.5.3	AVR	60
3.5.4	Summation of Instances	61
4	Experiments	62
4.1	Database Specifics	62
4.2	LIRE Similarity Measure Results	63
4.3	Comparison Of Category Estimation Methods	70
5	Conclusion	74
5.1	Future Work	74

List of Figures

2.1	Life cycle of a pea plant (genus <i>Pisum</i>).	5
2.2	Diagram illustrating the major organs (petals, sepals, stamen and pistil) of the typical angiosperm flower.	6
2.3	Photograph of the pollen sacs of an anthe. Four pollen sacs and a cross section of the fillament are visible(top). Diploid (2N chromosomes) microspore mother cells in the middle of the pollen sacs undergo meiosis to make microspores which later become pollen grains. The large green cells lining the pollen sacs provide nourishment for the developing pollen grains. Mature pollen grains (bottom). When the pollen sacs with their pollen grains mature, the sacs open and the numerous pollen grains within it will be released.	7
2.4	Formation of the male gametophyte plant. Special cells within the anther undergo meiosis (A-D), producing numerous haploid microspores, each of which is then surrounded by a tough, resistant cell wall. Each microspore divides once by mitosis, producing two haploid nuclei, and the structure develops into a pollen grain (E)-a male gametophyte- that is released from the anther when the pollen sacs mature. When the pollen grain is delivered to the stigma, a pollen tube grows out; its growth is governed by one of the two nuclei, the tube nucleus. The second nucleus divides once by mitosis to produce two sperm cells (F).	8
2.5	Fertilization of an angiosperm (A) Pollen grains land on the stigma and germinate forming pollen tubes that grow downward through the style and into the ovary. One of the pollen tubes shown here has reached the ovule within the ovary and discharged its sperm into the embryo sac. One sperm will fertilize the egg cell, and the other will inite with the tow polar nuclei to form a triploid nucleus, which will give rise to the endosperm. (B) Germinating pollen grain with pollen tube. The tough, resistant wall of the pollen grain protects the male gametophyte on its difficult journey from tha anther to the stigma. The pollen tube (below) will grow down through the tissues of the pistil to reach the ovule where it will discharge its sperm.	9
2.6	A pollen grain consists of a cell within a cell. The generative cell will undergo division to produce two sperm cells. One will fertilize the egg, and the other will join with the polar nuclei, yielding the endosperm. . . .	9
2.7	Sporoderm layers.	10
2.8	Apertures details: a) colpus; b) pore; c) colpus with margo; d) pore with anulus; e) colpus with pore; f) colpus with poroid; g) colpus and lolongate pore; h) colpus and lalongate pore; i) exopore and endopore; l) atrium; m) vestibulum; n) postvestibulum.	12

2.9	The visible color spectrum.	14
2.10	The C.I.E. Chromatic Diagram.	15
2.11	RGB Color Model.	16
2.12	Solid and Triangle of the HSI Color Model.	17
2.13	Example of an ontology	24
2.14	MPEG-7 Color Descriptors.	26
2.15	(a) Basic Unit of Haar transform (b) A schematic diagram of SCD generation.	27
2.16	Equivalent partitioning of HSV color space for different numbers of coefficients in the SCD.	27
2.17	Double cone representation of the HMMD color space.	28
2.18	HMMD color space quantization for CSD.	28
2.19	8x8 structuring element.	29
2.20	64 structuring points for 2 structuring element sizes.	29
2.21	The extraction process of the Color Layout Descriptor.	30
2.22	Stages of CLD computation.	31
2.23	Example of autocorellograms (b) for the images in (a)	35
2.24	Subimages and macroblocks decomposition.	36
2.25	Regularity of patterns (a)regular pattern, (b)irregular pattern.	39
2.26	A schematic diagram of CEDD.	41
2.27	A schematic diagram of FCTH.	42
2.28	Compact composite descriptors texture areas.	43
3.1	Nomarski Microscopy Arrangement.	45
3.2	ProgRes C12 Camera	47
3.3	ProgRes C12 Camera - Resolutions	47
3.4	Original capture from the microscope of the pollen species <i>Thymbra</i> of genus <i>Satureja</i>	48
3.5	Hue and Histogram of <i>Thymbra</i>	49
3.6	Saturation and Histogram of <i>Thymbra</i>	50
3.7	Otsu Thresholding on Hue.	50
3.8	Otsu Thresholding on Saturation.	51
3.9	Combination of the thresholded images on Hue and on Saturation.	51
3.10	Result after the use of morphological operators on Saturation.	52
3.11	Result after the use of Component Labeling algorithm. Each region is colored with a different shade of blue, that can't be distinguished with the eye.	54
3.12	Pollen Ontology. Note that not all classes, or individuals are shown.	56
4.1	The 40 query images used for evaluation.	63
4.2	Precision-Recall Diagram, based on species, for the 40 first answers using LIRE similarity measure.	64
4.3	Precision-Recall Diagram, based on genus, for the 40 first answers using LIRE similarity measure.	64
4.4	Retrieved list using LIRE similarity measure for Query 1.	65
4.5	Retrieved list using LIRE similarity measure for Query 2.	66
4.6	Retrieved list using LIRE similarity measure for Query 3.	67
4.7	Retrieved list using LIRE similarity measure for Query 4.	68

4.8 Retrieved list using LIRE similarity measure for Query 5.	69
4.9 Percentages of the first three answers for every category estimation method.	70
4.10 Results of example queries (1)	72
4.11 Results of example queries (2)	73

Dedicated to my family and friends.

Chapter 1

Introduction

The recognition of pollen grains leads to the identification of the originated flora (the produced the pollen grains) and has been in the center of research activities in many research fields, such as paleontology, medical science (allergies) and product certification (such as honey). The mass production of pollen by plants for reproduction, the durable nature of the grains for hundreds of years and their unique structure for each plant mark the importance for for this study.

Botanists use living pollen and spores (actuopalynology) in the study of plant relationships and evolution, while geologists (palynologists) may use fossil pollen and spores (paleopalynology) to study past environments, stratigraphy (the analysis of strata or layered rock) historical geology and paleontology. The correct identification of pollen is important in the study of vegetation and climatic history, in forensic palynology, in the analysis of origins of honey and beeswax, in the forecast of allergies caused of airborne pollen and in the alleviation of hay fever.

The common practice for pollen study is the observation of the grains through microscope, by qualified scientific personnel, and thus the matching with the known species of plants. This procedure is both time and resources consuming, as it demands the knowledge of a vast amount of different plants species and their pollen grains from the scientist, whereas each sample, containing tens of different types of pollen, must be examined separately.

Computer science and specifically machine vision can provide significant help to make pollen identification and classification much easier and much faster. With known techniques of image analysis, processing, segmentation, comparison and classification it is feasible to create a system that can provide fast and reliable automatic identification of pollen grains.

The most important obstacles that have to be resolved in order to build such a system are at first the collection of a large and representative pollen sample from various flora species, the discrimination of each pollen grain from images that contain many kinds of noise (background, flora particles, etc) and the training of the system to identify and classify the pollen with the least possible involvement of the qualified scientist in this process.

Related work, which will be discussed in detail in Chapter 2, used different methods of image segmentation based on color representation, shape, or contour form of the pollen grains, with similar results. Their bigger disadvantage was the limited access to pollen samples. The majority of these works used samples from, at most, 20 different flora species, a fact that constricts the utility of these systems. Furthermore, some related works were quite demanding in resources, as they needed supervision from a qualified scientist during the identification process.

In our work, thanks to our cooperation with the Mediterranean Plant Conservation Unit (MPCU) of the Mediterranean Agronomic Institute of Chania (MAICh), we had access to the collection of reference samples of pollen grains of the MPCU. This collection consists of over 230 different species of flora, mainly originated from Crete, and their pollen. All the different samples were digitized, using the equipment of MPCU (light microscope and camera) and had the images checked for accuracy and representativeness by Mrs Panagiota Gotsiou, Biologist B.Sc., M.Sc., Research Assistant of the MPCU. MPCU also provided the information of the flora of each sample (family / genus / species). Contrast to other works, the involvement of the qualified personnel ends at this stage of the process and does not have to interfere in the identification process. The retrieval and comparison of the images was implemented using the Lucene Image REtrival (LIRE), a library that provides a simple way to retrieve images based on their color and texture characteristics, also taking into consideration the size and circularity of the pollen grains. Our system uses 4 different methods in order to estimate the species of the given unknown image of pollen grain (best match, AVR, max occurrence and summation of instances).

Chapter 2

Background and Related Work

2.1 Pollen

For better understanding the utility and structure of pollen grains, it is necessary to analyze the life cycles and the fertilization process of plants and specifically of the plants that belong to the angiosperm class.

2.1.1 Angiosperms

Angiosperm plants are the first of the two classes of plants (the other class is gymnosperms) that belong to the division of Phanerogams or Spermatophytes. Phanerogams comprise those plants that produce seeds and are a subset of the land plants (embryophytes). Angiosperms are the flowering plants that produce seeds and pollen for their reproduction.

Angiosperms developed in the early to mid-Cretaceous period about 125 million years ago. Some analyses of morphological and molecular data support the view that the angiosperms are allied to the gnetopsids and to an extinct group known as the Bennettitales or cycadeoids. These groups have reproductive structures that can be interpreted as flowers. These early flower-bearing plants likely were derived from seed fern groups identified as caytonids or glossopterids. As the dominant plants on the earth's surface today, angiosperms number well over 200,000 species. They exhibit a wide array of habits, from perennial to annual and woody to herbaceous and occur in essentially all terrestrial habitats as well as in some aquatic ones.

The angiosperms are usually classified in two major groups, the dicotyledons and the monocotyledons. Cotyledon refers to a seed leaf, or the first leaf evident upon seed

germination. In dicotyledons, two seed leaves are formed, and in monocotyledons, one seed leaf is formed. The two groups differ in other important vegetative and reproductive characters. Several features of monocotyledons are considered to be derived or apomorphic. These include the solitary cotyledon, absence of a vascular cambium, dissected stele ('scattered' arrangement of vascular bundles in the stem), adventitious root system, sheathing leaf bases, and herbaceous habit. These unifying features of the monocotyledons serve to substantiate that the group is monophyletic [1].

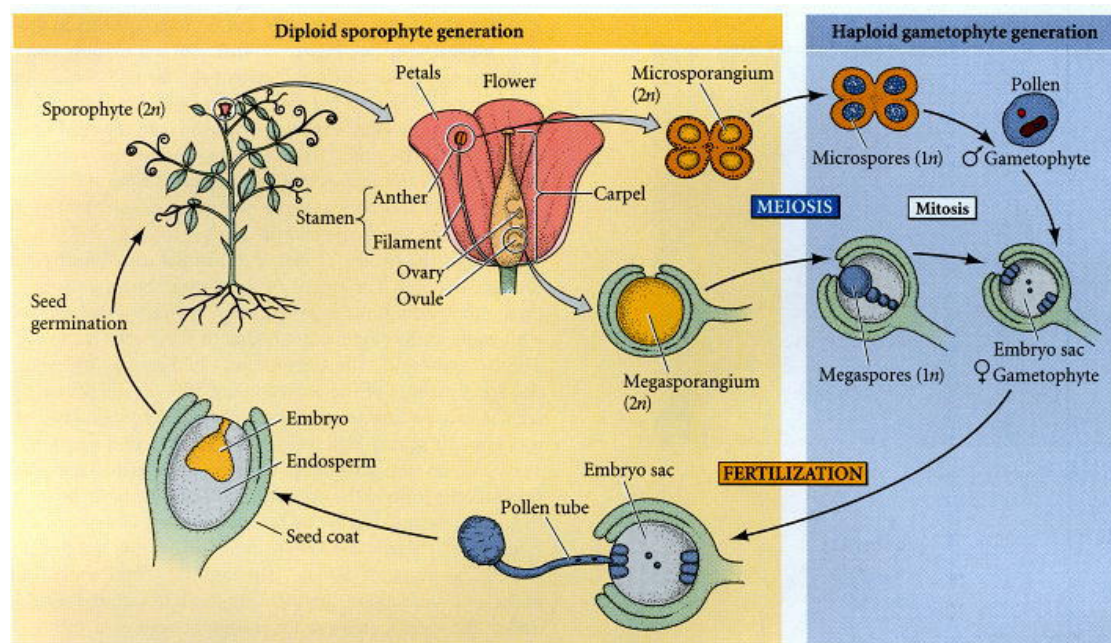
2.1.2 Angiosperms' Life Cycle

At first glance, angiosperms may appear to have a diplontic life cycle because the gametophyte generation has been reduced to just a few cells. However, mitotic division still follows meiosis in the sporophyte, resulting in a multicellular gametophyte, which produces eggs or sperm. All of this takes place in the organ that characterizes the angiosperms: the flower. Male and female gametophytes have distinct morphologies (i.e., angiosperms are heterosporous), but the gametes they produce no longer rely on water for fertilization. Rather, wind or members of the animal kingdom deliver the male gametophyte - pollen - to the female gametophyte. Another evolutionary innovation is the production of a seed coat, which adds an extra layer of protection around the embryo. The seed coat is also found in the gymnosperms. A further protective layer, the fruit, is unique to the angiosperms and aids in the dispersal of the enclosed embryos by wind or animals.

Life cycle of an angiosperm, represented in figure ?? . The sporophyte is the dominant generation, but multicellular male and female gametophytes are produced within the flowers of the sporophyte. Cells of the microsporangium within the anther undergo meiosis to produce microspores. Subsequent mitotic divisions are limited, but the end result is a multicellular pollen grain. The megasporangium is protected by two layers of integuments and the ovary wall. Within the megasporangium, meiosis yields four megaspores three small and one large. Only the large megaspore survives to produce the embryo sac. Fertilization occurs when the pollen germinates and the pollen tube grows toward the embryo sac. The sporophyte generation may be maintained in a dormant state, protected by the seed coat [2].

2.1.3 Gamete Production in Angiosperms

Angiosperm gametes are produced by the gametophyte generation. Angiosperm gametophytes are associated with flowers. The gametes they produce join to form the sporophyte. The study of embryonic development in plants is therefore the study of

FIGURE 2.1: Life cycle of a pea plant (genus *Pisum*).

early sporophyte development. In angiosperms, the sporophyte is what is commonly seen as the plant body. The shoot meristem of the sporophyte produces a series of vegetative structures. At a certain point in development, internal and external signals trigger a shift from vegetative to reproductive (flower-producing) development [3], [4]. Once the meristem becomes floral, it initiates the development of floral parts sequentially in whorls of organs modified from leaves (2.2). The first and second whorls become sepals and petals, respectively; these organs are sterile. The pollen-producing stamens are initiated in the third whorl of the flower. Each stamen consists of a stalk called a filament and a terminal pollen producing structure called a anther. The stamens contain four groups of cells, called the microsporangia (pollen sacs), within an anther. The microsporangia undergo meiosis to produce microspores. Unlike most ferns, angiosperms are heterosporous, so the prefix micro is used to identify the spores that mitotically yield the male gametophytes, pollen grains. The inner wall of the pollen sac, the tapetum, provides nourishment for the developing pollen. In the center of the flower is the female reproductive organ, the pistil (some species have more than one pistil per flower). Each pistil consists of an ovary at its base, a slender stalk called a style, which rises from the ovary and an enlarged top called a stigma. The pistil is derived from one or more modified spore-bearing leaves called carpels. Many flowers have all four kinds of floral organs-sepals, petals, stamens and pistils-but some flowers lack one or more of them.

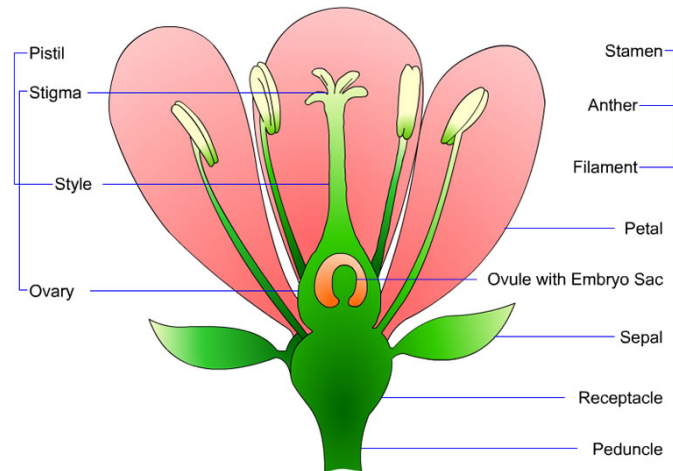


FIGURE 2.2: Diagram illustrating the major organs (petals, sepals, stamen and pistil) of the typical angiosperm flower.

2.1.4 Meiosis

Meiosis within the Anther produces microspores which mature into the male gametophyte plant, the pollen grain. Each anther has four pollen sacs (fig. 2.3), in which special cells undergo meiosis to produce numerous haploid microspores (fig. 2.4), each of which is then surrounded by a tough, resistant cell wall. Each microspore divides once, producing two haploid nuclei, and the structure develops into a pollen grain, the male gametophyte plant (fig. 2.5). Pollen is released from the anther when the mature sacs split open (fig. 2.3). The next step is **pollination**, the transfer of pollen grains from the anther to the stigma of a carpel so that fertilization can occur. Plants depend on external agents such as wind, insects or birds to transfer the pollen grains from the anther to the stigma. More plants are cross-pollinated than self-pollinated i.e. the pollen of one plant is deposited on the stigma of another plant of the same species [2].

2.1.5 Fertilization

Angiosperms carry out double fertilization. The pollen grain germinates when it falls (or it is deposited) on the stigma of a pistil, which is usually tough and sticky. A pollen tube begins to grow, and the two nuclei of the pollen grain move into it. One nucleus, the tube nucleus divides, giving rise in two sperm cells, which are the male gametes (fig. 2.4 (F)) [2]. The pollen tube grows down through the tissue of the stigma and style and enters the ovary (fig. 2.5). When the tip of the pollen tube enters the ovule, it discharges the two sperm cells into the female gametophyte (embryo sac). Double fertilization then occurs: one sperm fertilizes the egg cell, and the zygote thus formed

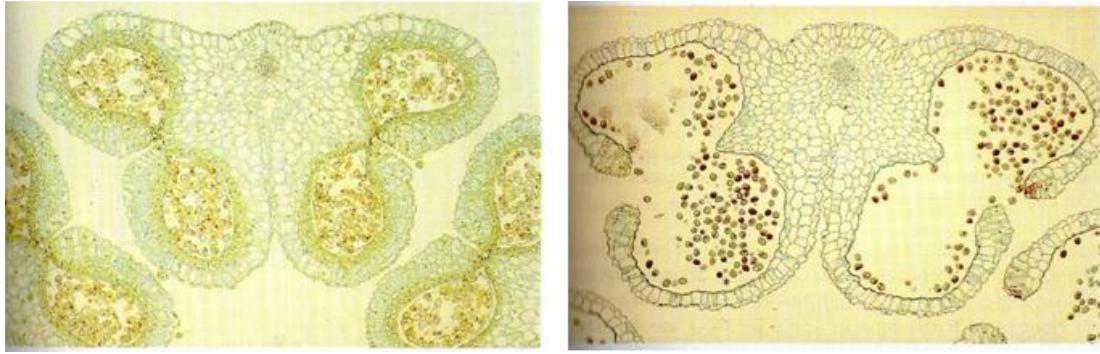


FIGURE 2.3: Photograph of the pollen sacs of an anthe. Four pollen sacs and a cross section of the filament are visible(top). Diploid ($2N$ chromosomes) microspore mother cells in the middle of the pollen sacs undergo meiosis to make microspores which later become pollen grains. The large green cells lining the pollen sacs provide nourishment for the developing pollen grains. Mature pollen grains (bottom). When the pollen sacs with their pollen grains mature, the sacs open and the numerous pollen grains within it will be released.

undergoes a series of mitotic divisions and develops into a tiny sporophyte embryo. The second sperm combines with the two polar nuclei to form a triploid nucleus. This nucleus undergoes a series of mitotic divisions, and a triploid tissue, called endosperm, is formed. The endosperm functions in the seed as a very rich source of stored food for the embryo. The embryo, together with the endosperm, becomes enclosed in a tough protective seed coat. The resulting composite structure, made up of embryo, endosperm, and seed coat, is called a seed. Seeds are evolutionary adaptations for dispersal and survival on land, allowing the embryo to wait in safety until conditions are favorable for growth.

Fertilization in angiosperms is accomplished without the use of water. The pollen is carried to the stigma by the wind or by an animal pollinator, and the sperm is delivered directly to the egg by the pollen tube. This method of fertilization does not require flagellated sperm to swim through the water to reach the egg, making this an important adaptation for life on land.

2.1.6 Pollen Grain Structure

The pollen grain is an extremely simple multicellular structure. The outer wall of the pollen grain, the exine, is composed of resistant material provided by both the tapetum (sporophyte generation) and the microspore (gametophyte generation). The exine consists of two layers: ectexina and endexina. The inner wall, the intine, is produced by the microspore. Intine is a flexible layer made of cellulose and pectina and extroverts from the apertures of the grain when the pollen germinates. Together, exine and intine are called sporoderm (fig. 2.7). Intine is the layer that separates the sporoderm from

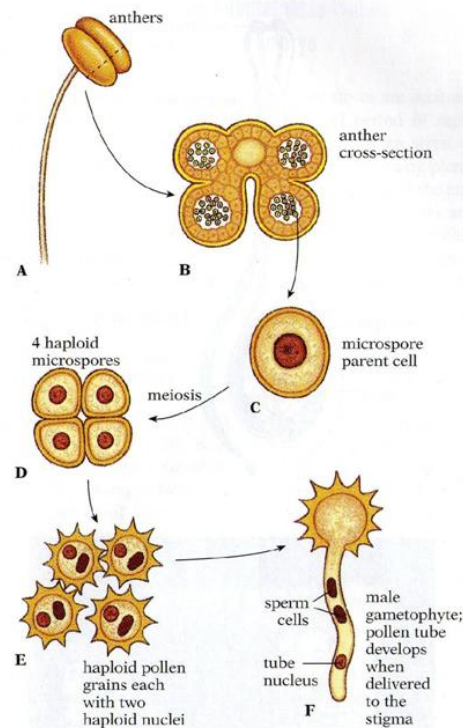


FIGURE 2.4: Formation of the male gametophyte plant. Special cells within the anther undergo meiosis (A-D), producing numerous haploid microspores, each of which is then surrounded by a tough, resistant cell wall. Each microspore divides once by mitosis, producing two haploid nuclei, and the structure develops into a pollen grain (E)-a male gametophyte- that is released from the anther when the pollen sacs mature. When the pollen grain is delivered to the stigma, a pollen tube grows out; its growth is governed by one of the two nuclei, the tube nucleus. The second nucleus divides once by mitosis to produce two sperm cells (F).

the cytoplasm. Cytoplasm is the cell mass inside the sporoderm which winds around the nucleus.

A mature pollen grain consists of two cells, one within the other (fig. 2.6). The tube cell contains a generative cell within it. The generative cell divides to produce two sperm. The tube cell nucleus guides pollen germination and the growth of the pollen tube after the pollen lands on the stigma of a female gametophyte. One of the two sperm will fuse with the egg cell to produce the next sporophyte generation. The second sperm will participate in the formation of the endosperm, a structure that provides nourishment for the embryo [5].

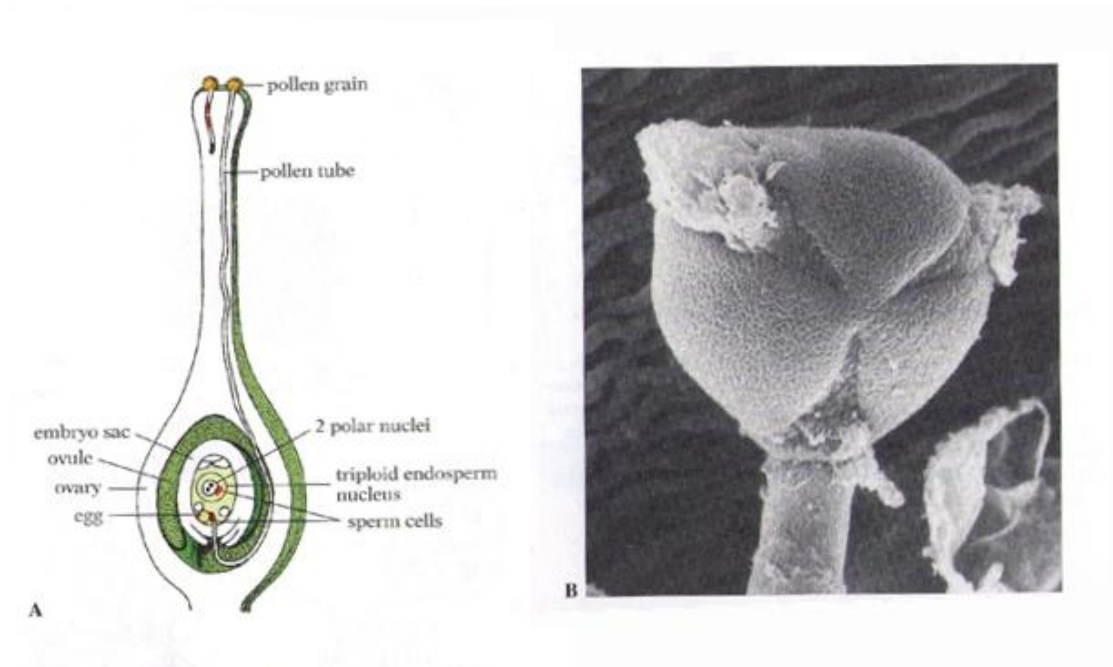


FIGURE 2.5: Fertilization of an angiosperm (A) Pollen grains land on the stigma and germinate forming pollen tubes that grow downward through the style and into the ovary. One of the pollen tubes shown here has reached the ovule within the ovary and discharged its sperm into the embryo sac. One sperm will fertilize the egg cell, and the other will unite with the two polar nuclei to form a triploid nucleus, which will give rise to the endosperm. (B) Germinating pollen grain with pollen tube. The tough, resistant wall of the pollen grain protects the male gametophyte on its difficult journey from the anther to the stigma. The pollen tube (below) will grow down through the tissues of the pistil to reach the ovule where it will discharge its sperm.

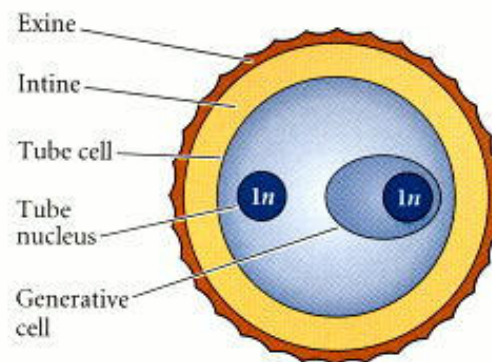


FIGURE 2.6: A pollen grain consists of a cell within a cell. The generative cell will undergo division to produce two sperm cells. One will fertilize the egg, and the other will join with the polar nuclei, yielding the endosperm.

2.1.7 Pollen Grain Characteristics and Classification

Pollen classification is based upon several characteristics of the pollen grains, such as symmetry, polarity, shape, size, structure, sculpture and apertures.

Polarity and symmetry

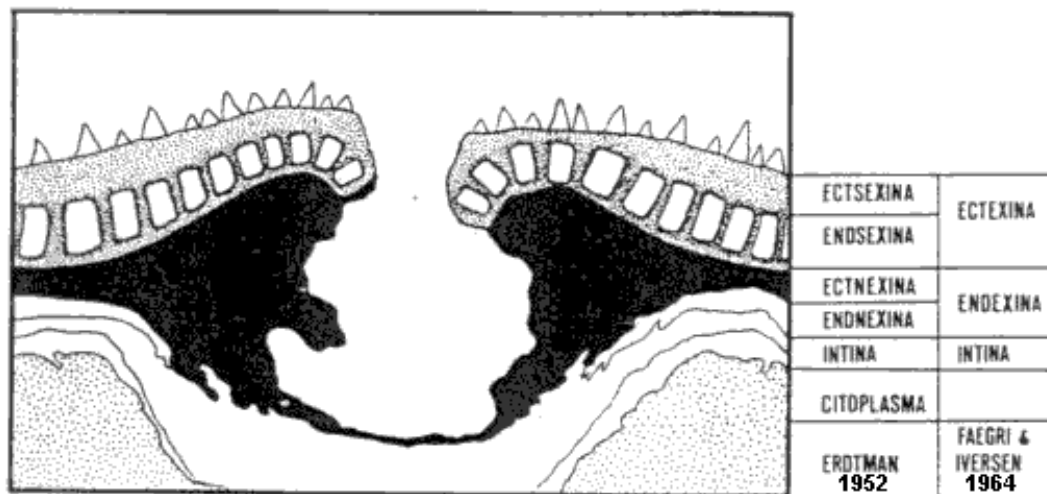


FIGURE 2.7: Sporoderm layers.

Pollen grains, resulting from the meiotic division of mother cells are gathered into groups and occasionally remain like this (Erdtman, 1952). They form diads, tetrads, or polyads. The tetrads (four granules) can be tetragonal, tetrahedral or rhomboidal. The polyads can be formed by a larger number of granules (8-10-12-16-24-32). Observing the position that the pollen grain first has in the tetrahedral agglomerate, 2 poles are noted: the proximal and the distal. If the poles are equal, the pollen is defined as isopolar and if they are unequal as anisopolar. Spherical pollens have no polarity (apolar).

Different types of pollen symmetry exist: there are isopolar granules in radial symmetry, i.e. granules with a horizontal symmetric plane and two or three vertical symmetric planes. Then there are anisopolar granules in radial symmetry, which have no horizontal symmetric plane. Isopolar granules in bilateral symmetry, with one horizontal symmetric plane and two vertical ones and anisopolar granules in bilateral symmetry, which have no horizontal symmetric plane. Lastly, there are asymmetrical granules which have no symmetric planes [6].

Indices

The description of a pollen is based on particular ratios and measures that best characterize that particular pollen. The morphology of a pollen grain is measured by the ratio of the length of the polar axis (an imaginary straight line connecting the two poles) to the equatorial diameter (P/E). The correlation between sporoderm and pollen size is measured by the ratio between sporoderm's thickness and the equatorial axis. In reality the indicator of the polar area is the ratio of this area to the equatorial diameter, but in practice only the distance between colpi (or pores) extremities near the poles is measured [7].

Shape and shape categories

The ratio between the polar axis and the equatorial diameter of a pollen gives us the shape of the pollen. Pollen shapes have been classified as follows [6]:

SHAPE	CATEGORY P/E
Perprolate	> 2
Prolate	2-1.33
Subprolate	1.33-1.14
Prolate-spheroidal	1.14- 1.00
Spheroidal	1
Oblate-spheroidal	1.00 - 0.88
Suboblate	0.88 - 0.75
Oblate	0.75 - 0.50
Peroblate	< 0.50

When observed in polar view, a pollen grain will have one of the following shapes: circular, semiangular, interexagonal, angular, interangular, semilobate, intersemilobate, lobate, interlobate, exagonal, intersemiangular, subangular and intersubangular [7], [8].

Sizes

These vary greatly, but are generally divided into six categories, although this number could be increased [7]:

< 10	very small pollens
10 - 25	small pollens
25 - 50	medium pollens
50 - 100	large pollens
100 - 200	very large pollens
> 200	giant pollens

Apertures

In some cases pollens do not have apertures (inaperturate), but more frequently they do. Traditional palynology usually characterizes apertures by their number (numerus), position (positio) and shape (caracter); the NPC parameter is fundamental in pollen description. The number can vary from N1 to N7; N1 is attributed to pollens with one aperture (e.g.: monocolpate); N7 is for pollens with several apertures (> 6). Inaperturate pollens are marked with NO. The position parameter in pollens has 7 variables -

P0 to P6; in P0 the location of the aperture is unknown; if the aperture is in a pole it is defined as proximal (P1) or distal (P3). P2 designates a pollen with two apertures, both at the poles. If the apertures are set in the equatorial zone then this is a zonal position (P4 and P5); pollens with apertures distributed over the whole granule surface (periaperturate with global position) are identified as P6. Apertures are also characterized by their shape (C = character). In C0 the aperture is indistinct, while in C1 it is represented by a more or less regular area (leptoma). In C2 the aperture is divided into three anastomosed arms (trichotomocolpate pollens) (some *Palmae*, many spores). C3 is the term used for sulcus-shaped apertures (colpus) and C4 for pore-shaped ones (porus) (fig. 2.8): We can refer to these as colpate and porate pollens. C5 is used for colporate pollens, i.e. pollens whose apertures have both a pore and a colpus. Lastly there are the C6 pororate pollens which have either a colpus or a pore externally and a circular oval part set in their centre [6].

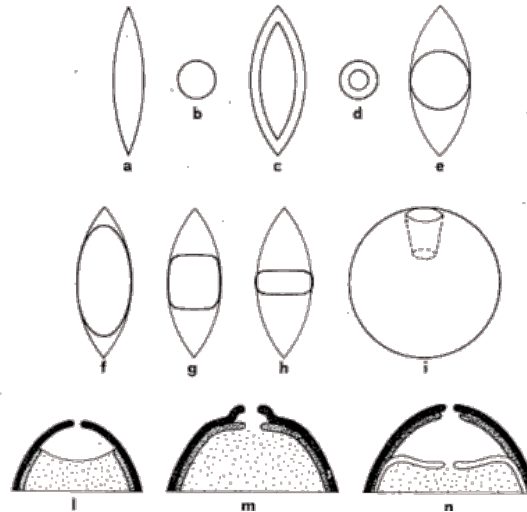


FIGURE 2.8: Apertures details: a) colpus; b) pore; c) colpus with margo; d) pore with anulus; e) colpus with pore; f) colpus with poroid; g) colpus and lalongate pore; h) colpus and lalongate pore; i) exopore and endopore; l) atrium; m) vestibulum; n) postvestibulum.

Sporoderm structure and sculpture

Generally, structure means the layers forming the sporoderm as a whole (clearly visible in section) and sculpture means the ornamental elements modelling the sporodermis surface (clearly visible in polar or equatorial view). However, in pollens with no tectum (external layer), the structural elements seen in section appear on the surface as ornaments (sculpture) in polar or equatorial view. In this case the two terms, structure and sculpture, coincide - only the view varies [9].

The classification based on sporoderm structure and sculpture is of great length and depth, which is not needed in this work. Thus, it is not presented here.

Color

Pollens treated with acids or alkali generally have a more or less bright brown coloration (they can be clarified by treatment with hydrochloric acid). Fresh compound pollen can present a distinctive coloration. As a rule, the color varies from a more or less intense yellow (presence of oils) to different shades of grey - green. If ether is used to decrease the pollen the yellow color fades. There are, however, some pollens with a typical coloration, for example violet or blue [5].

2.2 Machine Vision and Image Processing

2.2.1 Color

The use of color in image processing is important for two main reasons. Firstly, in automated image analysis, color is a very strong characteristic and in many cases simplifies the recognition of objects, on top of a scene. Secondly, the human eye can distinguish thousands color shades, in contrast with the, approximately, 24 shades of gray.

Color Fundamentals

Even if the perception of color from the human eye isn't fully deciphered, the nature of color can be described based on theoretical and practical results of research.

In 1666 Isaac Newton discovered that when a sunlight beam runs through a prism, it is analyzed to the visible color spectrum, with a fluctuation from violet to red. The spectrum can be divided into 6 large areas: violet, blue, green, yellow, orange and red. The transition from one color to another isn't abrupt, but smooth.

The perceived color, of an object, from the human eye is determined by the nature of the light that is reflected from the object. The visible light is structured by a rather limited range of frequencies, from the electromagnetic spectrum. An object that reflects with the same way all the wavelengths of the visible spectrum appears white. On the contrary, if an object "favors" a specific range of frequencies, it appears having a specific color shade (e.g. red). In case light is colorless the only characteristic it has is its intensity. Changing only the intensity of the light we can create different grey levels.

The visible color spectrum extends between 400 and 700 nm, on the electromagnetic spectrum (fig. 2.9). The three basic characteristics that are used to describe the quality of the light source are radiance, luminance and brightness. Radiance is the final amount of energy that the light source radiates and is often calculated in Watt. Luminance shows a measure of the amount of energy that an observer receives and is calculated

in Lumens. Finally, brightness is a subjective measure which practically cannot be calculated. Brightness incorporates the colorless meaning of intensity and is one of the basic characteristics that describe the feeling of the color [10].

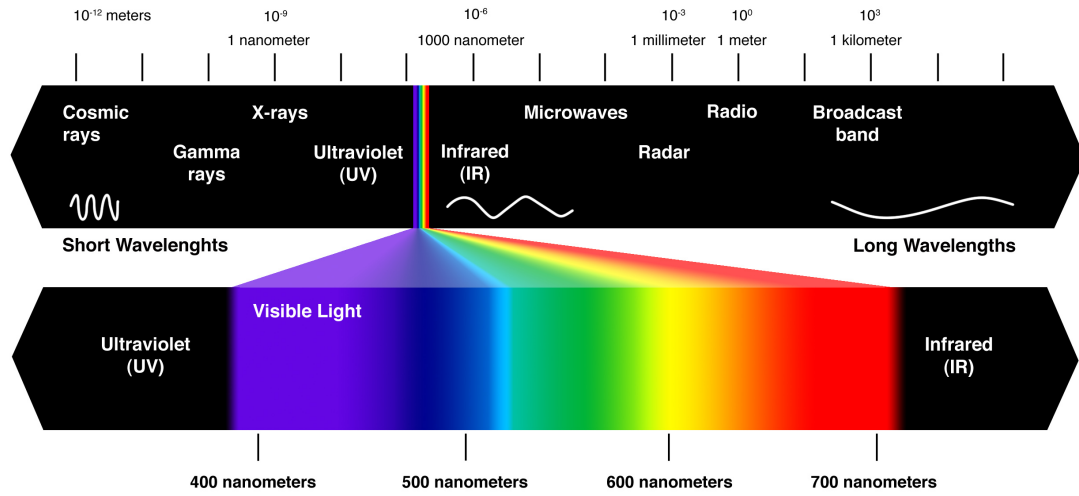


FIGURE 2.9: The visible color spectrum.

Based on the structure of the human eye all colors are considered combination of the three primary colors: (R)ed, (G)reen and (B)lue. For standardization reasons the international Commission on Illumination - Commission Internationale de l'Éclairage (CIE) set specific values for the three primary colors (red=700 nm, green=546.1 nm, blue=435.8 nm). Primary colors can be mixed in order to create the secondary colors: Magenta, Cyan and Yellow. The mix of the three primary colors or one secondary with the opposite primary, with certain intensity, produces the white color.

Colors can be distinguished with the use of the characteristics of brightness, hue and saturation. Hue is highly dependent on the transcendent wavelength within a mix of light waves. Thus, hue represents the transcendent color which the observer perceives. Saturation refers to the relative purity or the quantity of white light when it is mixed with a shade of color. Saturation is inversely proportional to the quantity of white light.

Hue, along with saturation are called chromaticity. Consequently, a color can be fully described by its chromaticity and brightness. The values of red, green and blue that are needed to create any color are called tristimulus and are symbolized as x , y and z respectively.

A popular approach of the definition of colors is the C.I.E. chromatic diagram (fig. 2.10). The composition of each color is a function of red (x) and green (y). For each value of x and y the value of z results from the equation $x + y + z = 1$. Thus, we can find, for a given color, the percentage that each of the primary colors contributes [10].

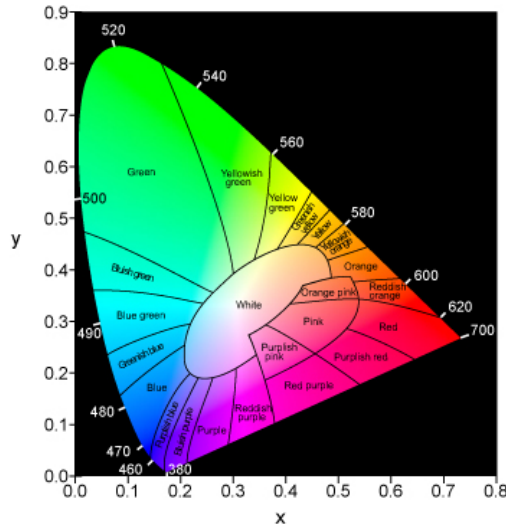


FIGURE 2.10: The C.I.E. Chromatic Diagram.

Color Models

The purpose of the use of color models is to facilitate the definition of colors for standardization. A color model is a three-dimensional coordinate system and a subsystem, in which each color is represented as a point. Most color models are adjusted to computer and peripheral hardware, such as monitors and printers, or to applications that need to manipulate color.

RGB Color Model

In this model every color is described by the three primary colors and is based on a Cartesian coordinate system. RGB color model can be represented by a cube, on the three corners - those on the axis - of which are the three primary colors (red, green and blue). The secondary colors (cyan, magenta and yellow) lie on the other three corners, whereas black is on the zero point of the axis and white on the direct opposite of black. Grey levels are represented on the line that connects the black and the white corner of the cube. Any other point within or on top of the cube represents a certain color. All values are normalized in the range $[0,1]$ (fig. 2.11).

Images that are based on this model consist of three independent levels, one for each primary color. When composed on a monitor, the three levels are combined and produce a compound colored image.

HSI Color Model

The HSI color model represents a color in terms of hue, saturation and intensity. It tries to imitate the way the human eye perceives color. Humans use to describe a scene, not in conditions of red, blue and green, but as hue, saturation and intensity. Hue represents the

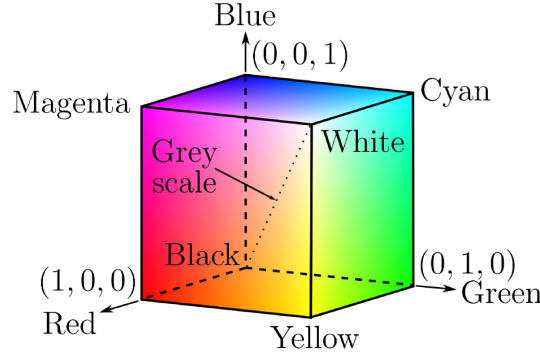


FIGURE 2.11: RGB Color Model.

color that is perceived as a result of its wavelength. Saturation represents how "clean" the image is. Intensity represents the brightness, the level of light energy. This model has two main advantages. Intensity is independent of color and Hue and Saturation are closely related to the way the human eye perceives color. These characteristics make the HSI color model an ideal tool for image processing.

The HSI color model can be described as a triangle (fig. 2.12) which represents the combinations of Hue and Saturation that can be represented by combinations of the RGB primary colors. The corners of the triangle correspond to the maximum amounts of each primary color available from the imaging system. Colorless pixels are shades of gray, corresponding to equal amounts of the primary colors and lie at the center of the triangle.

The HSI solid adds the dimension of image intensity, with black at the bottom of the solid and white at the top. Shades of gray run along the axis of the solid. Each cross section of the solid is an HSI triangle with the proportions of the primary colors constrained to produce a particular intensity value. The HSI solid narrows to a point at the top and bottom because white and black can only be represented by unique combinations of the RGB primaries.

The RGB components of an image can be converted to the HSI color representation, using the equations below, assuming that the RGB components have been normalized to the range [0,1]:

$$I = \frac{1}{3}(R + G + B)$$

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)]$$

$$H = \cos^{-1} \left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - R)}} \right)$$

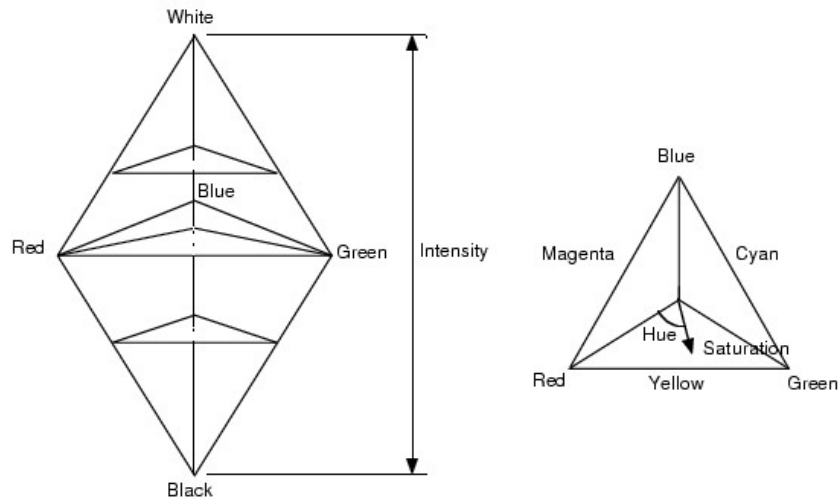


FIGURE 2.12: Solid and Triangle of the HSI Color Model.

The transformation from RGB to HSI is used to convert color images into a form appropriate for machine vision. Gray-level algorithms can be performed on the I component of the HSI representation. Segmentation can be performed on the H component to discriminate between objects with different hues. However, hue is not reliable for discrimination when the saturation is low, so the segmentation algorithms must be modified to leave pixels with low saturation unassigned.

Algorithms can divide the HSI solid into regions using thresholds on hue, saturation and intensity. These thresholds are easier to formulate and apply in the HSI model than in the RGB model provided by the imaging system.

2.2.2 Image Segmentation and Thresholding

Binary Images

An image contains a continuum of intensity values before it is quantized to obtain a digital image [11]. The information in an image is in these gray values. To interpret an image, the variations in the intensity values must be analyzed. A common practice in machine vision is to use binary vision systems. A binary image contains only two gray levels. This practice is very useful for many reasons. First of all, the algorithms for computing properties of binary images are well understood, from both machines and humans. They also tend to be less expensive and faster than vision systems that operate on gray level or color images. This is due to the significantly smaller memory and processing requirements of binary vision.

Thresholding

One of the most important problems in a vision system is to identify the subimages that represent objects. This operation, which is so natural and so easy for people, is surprisingly difficult for computers. The partitioning of an image into regions is called segmentation. Ideally, a partition represents an object. Formally, segmentation can be defined as a method to partition an image, $F[i, j]$, into subimages, called regions, P_1, \dots, P_k , such that each subimage is an object candidate. Each region P_i should satisfy a predicate; that is, all points of the partition have some common property. This predicate may be as simple as "*has uniform intensity*" but is more complex in most applications.

A binary image is obtained using an appropriate segmentation of a gray scale image. If the intensity values of an object are in an interval and the intensity values of the background pixels are outside this interval, a binary image can be obtained using a thresholding operation that sets the points that interval to 1 and points outside that range to 0. Thus, for binary vision, segmentation and thresholding are synonymous.

Thesholding is a method to convert a gray scale image into a binary image so that objects of interest are separated from the background. For thresholding to be effective in an object-background separation, it is necessary that the objects and background have sufficient contrast and that we know the intensity levels of either the objects or the background. In a fixed thresholding scheme, these intensity characteristics determine the value of a threshold.

The major problem with thresholding is that it considers only the intensity, not any relationships between the pixels. There is no guarantee that the pixels identified by the thresholding process are contiguous. Extraneous pixels that are not part of the desired region can easily be included, or isolated pixels within the region (especially near the boundaries of the region) can just as easily be missed. These effects get worse as the noise gets worse, simply because it is more likely that a pixels intensity does not represent the normal intensity in the region.

The common practice to set a global threshold or to adapt a local threshold to an area, is to examine the histogram of the image to find two or more distinct modes, one for the foreground (or the object) and one for the background. A histogram is a probability distribution:

$$P(g) = n_g/n$$

That is, the number of pixels n_g having greyscale intensity g as a fraction of the total number of pixels n [11].

Otsu Segmentation

A popular and reliable method for thresholding is Otsu's method [12]. Otsu tried to solve the problem of the overlapping ranges of intensities between the object and the background. This method tries to make each range as tight as possible in order to minimize their overlap. By adjusting the threshold one way, the spread of one is increased and the spread of the other is decreased. The goal then is to select the threshold that minimizes the combined spread.

Within-class variance is defined as the weighted sum of the variances of each cluster:

$$\sigma_{Within}^2(T) = n_B(T)\sigma_B^2(T) + n_O(T)\sigma_O^2(T)$$

where

$$n_B(T) = \sum_{i=0}^{T-1} p(i)$$

$$n_O(T) = \sum_{i=T}^{N-1} p(i)$$

$$\sigma_B^2(T) = \text{the variance of the pixels in the background (below threshold)}$$

$$\sigma_O^2(T) = \text{the variance of the pixels in the foreground (above threshold)}$$

and $[0, N - 1]$ is the range of intensity levels.

Subtracting the within-class variance from the total variance of the combined distribution, the between-class variance is computed:

$$\sigma_{Between}^2(T) = \sigma^2 - \sigma_{Within}^2(T) = n_B(T)[\mu_B(T) - \mu]^2 + n_O(T)[\mu_O(T) - \mu]^2$$

where σ^2 is the combined variance and μ is the combined mean.

Substituting $\mu = n_B(T)\mu_B(T) + n_O(T)\mu_O(T)$ and simplifying, we get:

$$\sigma_{Between}^2(T) = n_B(T)n_O(T)[\mu_B(T) - \mu_O(T)]^2$$

So, for each potential threshold T we

1. Separate the pixels into two clusters according to the threshold.
2. Find the mean of each cluster.
3. Square the difference between the means.

4. Multiply by the number of pixels in one cluster times the number in the other.

This depends only on the difference between the means of the two clusters, thus avoiding having to calculate differences between individual intensities and the cluster means. The optimal threshold is the one that maximizes the between-class variance (or, conversely, minimizes the within-class variance).

We can update $n_B(T)$, $n_O(T)$, and the respective cluster means $\mu_B(T)$ and $\mu_O(T)$ as pixels move from one cluster to the other as T increases. Using simple recurrence relations we can update the between-class variance as we successively test each threshold:

$$\begin{aligned} n_B(T+1) &= n_B(T) + n_T \\ n_O(T+1) &= n_O(T) - n_T \\ \mu_B(T+1) &= \frac{\mu_B(T)n_B(T) + n_T T}{n_B(T+1)} \\ \mu_O(T+1) &= \frac{\mu_O(T)n_O(T) - n_T T}{n_O(T+1)} \end{aligned}$$

2.2.3 Morphological Operators

Mathematical morphology gets its name from the study of shape. This approach exploits the fact that in many machine vision applications, it is natural and easy to think in terms of shapes when designing algorithms. A morphological approach facilitates shape-based, or iconic, thinking. The fundamental unit of pictorial information in the morphological approach is the binary image [11].

The *intersection* of any two binary images A and B , $A \cap B$, is the binary image which is 1 at all pixels p which are 1 in both A and B . Thus, $A \cap B = \{p | p \in A \text{ and } p \in B\}$.

The *union* of A and B , $A \cup B$, is the binary image which is 1 at all pixels p which are 1 in A or 1 in B (or 1 in both). Symbolically, $A \cup B = \{p | p \in A \text{ or } p \in B\}$.

Let Ω be a universal binary image (all 1) and A a binary image. The *complement* of A is the binary image which interchanges the 1s and 0s in A . Thus, $\bar{A} = \{p | p \in \Omega \text{ and } p \notin A\}$.

The *vector sum* of two pixels p and q with indices $[i, j]$ and $[k, l]$ is the pixel $p + q$ with indices $[i + k, j + l]$. The *vector difference* $p - q$ is the pixel with indices $[i - k, j - l]$.

If A is a binary image and p is a pixel, then the *translation* of A by p is an image given by $A_p = \{\alpha + p | \alpha \in A\}$.

Dilation

Translation of a binary image A by a pixel p shifts the origin of A to p . If $A_{b_1}, A_{b_2}, \dots, A_{b_n}$ are translations of the binary image A by the 1 pixels of the binary image $B = \{b_1, b_2, \dots, b_n\}$ then the union of the translations of A by the 1 pixels of B is called the *dilation* of A by B and is given by:

$$A \oplus B = \bigcup_{b_i \in B} A_{b_i}$$

Dilation has both associative and commutative properties. This, in a sequence of dilation steps the order of performing operations is not important. The fact allows breaking a complex shape into several simpler shapes which can be recombined as a sequence of dilations.

Erosion

The opposite of dilation is *erosion*. The erosion of a binary image A by a binary image B is 1 at a pixel p if and only if 1 pixel in the translation of B to p is also 1 in A . Erosion is given by:

$$A \ominus B = \{p | B_p \subseteq A\}$$

Often the binary image B is a regular shape which is used as a probe on image A and is referred to as a *structuring element*. Erosion plays a very important role in many applications. Erosion of an image by a structuring element results in an image that gives all locations where the structuring element is contained in the image.

Dilation and erosion exhibit a dual nature that is geometrical rather than logical and involves a geometric complement as well as a logical complement. The geometric complement of a binary image is called its *reflection*. The reflection of a binary image B is that binary image B' which is symmetric with B about the origin, that is:

$$B' = \{-p | p \in B\}$$

The geometric duality of dilation and erosion is expressed by the relationships:

$$\overline{A \oplus B} = \overline{A} \ominus B'$$

and

$$\overline{A \ominus B} = \overline{A} \oplus B'$$

Geometric duality contrasts with logical duality:

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

and

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

also called deMorgan's law.

Erosion and dilation are often used in filtering images. If the nature of noise is known, then a suitable structuring element can be used and a sequence of erosion and dilation operations can be applied for removing the noise. Such filters affect the shape of the objects in the image.

Opening - Closing

The basic operations of mathematical morphology can be combined into complex sequences. For example, an erosion followed by a dilation with the same structuring element (probe) will remove all of the pixels in regions which are small to contain the probe, and it will leave the rest. This sequence is called *opening*.

The opposite sequence, a dilation followed by an erosion, will fill in holes and concavities smaller than the structuring element. This is referred to as *closing*. Such filters can be used to suppress spatial features or discriminate against objects based upon size. The structuring element used does not have to be compact or regular, and can be any pattern of pixels. In this way, features made up of distributed pixels can be detected [11].

2.2.4 Component Labeling

One of the most common operations in machine vision is finding the connected components in an image [11]. The points in a connected component form a candidate region for representing an object. In computer vision most objects have surfaces. Points belonging to a surface project to spatially close points. The notion of "spatially close" is captured by connected components in digital images. It should be mentioned that connected component algorithms usually form a bottleneck in a binary vision system. The algorithm is sequential in nature, because the operation of finding connected components is a global operation. If there is only one object in an image, then there may not be a need for

finding the connected components; however, if there are many objects in an image and the object properties and locations need to be found, then the connected components must be determined.

A component labeling algorithm finds all connected components in an image and assigns a unique label to all points in the same component. In many applications, it is desirable to compute characteristics (such as size, position, orientation and bounding rectangle) of the components while labeling these components.

2.3 Ontologies

In computer and information science, an ontology is a formal representation of knowledge as a set of concepts within a domain, and the relationships between those concepts [13]. A popular definition for ontology states: An ontology is a formal, explicit specification of a shared conceptualization [14].

Ontologies are used in computer science to facilitate sharing and reuse. They are very useful because they give a shared and common understanding of a domain that can be communicated between people and application systems. The basic and most significant role of ontologies in knowledge and software engineering is to facilitate the construction of a domain model.

Ontologies have been proven to be useful for image annotation as they provide a way to enhance descriptions with conceptual meanings. In other words, using ontologies a computer is capable of generating meaningful descriptions rather than letting a human do it, as it would be time consuming and does not scale-up for large data sets [15]. Furthermore, the use of ontologies helps to recognize and use the possible semantic relations between the described concepts. The use of ontologies as a basis for defining visual vocabulary or as a framework for automatic image annotation increases the number of concepts an image annotation system can recognize and may as a means for improving the performance of image retrievals.

2.4 Related Work

In recent years, a number of notable work dealing with automated pollen recognition and classification from microscope images have been presented. Due to the complexity of the problem, there are many different approaches on every different part of the process. Variations can be noticed on the selection of the pollen databases used, image

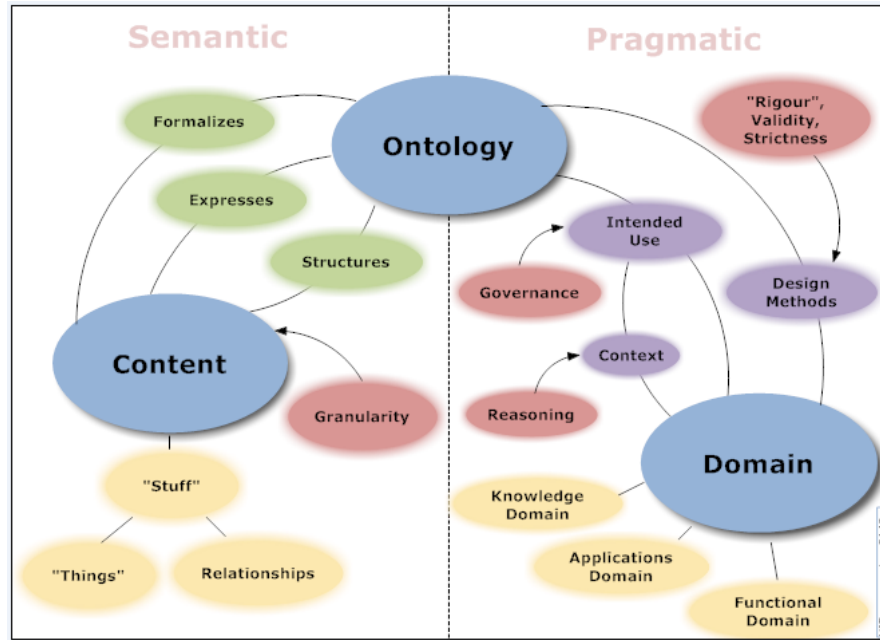


FIGURE 2.13: Example of an ontology

acquisition, the segmentation process applied on the images, the features of the pollen grains used for the recognition, or the methods of classification. A research, presented by G.P.Allen in [16], proposed the use of Sobel edge operator and morphological operators for the segmentation, in order to locate the pollen grains in an image. Texture features (a series of Wavelet transforms, Grey Level/Gradient Co-occurrence Matrix) and shape features (geometric, histogram and second moment) of the images were presented to a multi-layer perceptron for the classification of 3 different pollen species. Another work, [17], segmented the images by combining a simple semi-automatic algorithm to capture the pollen grains and a process using color contrast, decorrelation stretching, HSV color transformation and a thresholding algorithm. Features used in this work included geometrical features (area, convex hull and perimeter), Fourier descriptors and color features. A neural network classified 17 different pollen species from Costa Rica. A similar work, using the same pollen database, can be found in [18], tried to locate and classify the grains based on information extracted from their contour, using a Hidden Markov Model (HMM). A different approach can be found in the work [19], with the use of 2D and 3D image processing (multiple captures of the grain at different focal levels) and palynological information for the identification of pollen grains. Segmentation was achieved with split-and-merge segmentation with markovian relaxation, while detection and extraction of the pollen grains was done by chrominance and luminance analysis. Four pollen types of the Mediterranean area (along with 10 similar types) were selected for the image reference database of this study. A hierarchical approach, with the use of two neural networks for classification between pollen grains and debris, was presented

by France I. in [20]. The system was able to classify grains of 3 pollen species, from gray-scale images.

In our work we tried to combine several segmentation techniques that we came across, such as the use of the HSV/HSI color model, thresholding, morphological operators, according to the specifications of our own dataset. In terms of feature extraction and classification we based our work on an approach of an automatic image annotation system, that was developed in the Intelligent Systems Laboratory of ECE of TUC, by A.Dimas [21]. In this work, LIRE library was used for the extraction of the features of the images, while two similarity measures were used to compute similarity between unknown and classified images: LIRE similarity measure, which is a sum of the LIRE descriptors and a second similarity measure, which is a weighted similarity measure of LIRE features, with weights determined by a decision tree.

2.5 Image Content Analysis

In order to perform a successful image retrieval, image content descriptors were used in our work.

A descriptor is defined as a representation of a feature. A descriptor defines the syntax and semantics of the feature representation. Examples of low-level visual features include color, shape, motion, and texture. 4 types of descriptors are used. Color descriptors, texture descriptors, hybrid descriptors (a combination of color and texture) and geometrical descriptors.

2.5.1 Color Descriptors

Color is the most basic quality of visual content. One of the most recognizable elements of image content and is the most commonly used feature in image retrieval because of its invariance with respect to image scaling, translation and rotation. Color features are independent of image size and orientation and can be used for describing content in still images and video. A good reference on how color descriptors can be used in image retrieval is [22]

2.5.1.1 MPEG-7 Color Descriptors

MPEG-7, formally named "*Multimedia Content Description Interface*", is a standard for describing the multimedia content data that supports some degree of interpretation

of the information meaning, which can be passed onto, or accessed by, a device or a computer code. MPEG-7 visual descriptors are often used for image retrieval[23–25].

MPEG-7 supports 4 color descriptors (fig. 2.14). Dominant Color, Scalable Color, Color Structure and Color Layout. In this work, we do not use dominant color because of implementation issues.

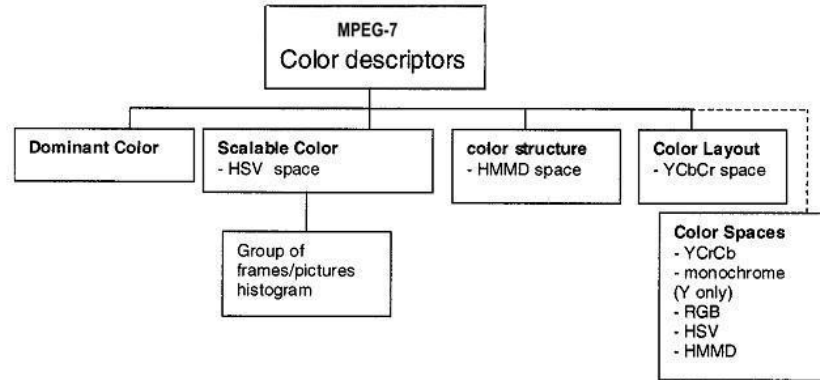


FIGURE 2.14: MPEG-7 Color Descriptors.

Scalable Color Descriptor (SCD)

The Scalable Color Descriptor is a Color Histogram in HSV Color Space, which is encoded by a Haar transform. Its binary representation is scalable in terms of bin numbers and bit representation accuracy over a broad range of data rates. The Scalable Color Descriptor is useful for image-to-image matching and retrieval based on color feature. Retrieval accuracy increases with the number of bits used in the representation.

The histogram values are extracted, normalized and non-linearly mapped into a 4-bit integer representation, giving higher significance to small values. The Haar transform is applied to the 4-bit integer values across the histogram bins. The basic unit of the transform consists of a sum operation and a difference operation (fig. 2.15), which relate to primitive low pass and high pass filters. Summing pairs of adjacent bins is equivalent to the calculation of a histogram with half number of bins. From the sums of every two adjacent Hue bin values out of the 256-bin histogram, we get a representation of a 128-bin histogram with 8 levels in H, 4 levels in S and 4 levels in V. If this process is repeated, the resulting 64, 32 or 16 sum coefficients from the Haar representation are equivalent to histograms with 64, 32 or 16 bins.

Figure 2.16 shows the equivalent partitioning of the HSV color space for different number of coefficients of the Haar transform. If an application does not require the full resolution,

limited number of Haar coefficients may simply be extracted from a 128, 64 or 32 bin histogram. This would still guarantee interoperability with another representation where all coefficients were extracted, but only to the precision of the coefficients that are available in both of the representations.

In order to use this descriptor to perform similarity retrieval, we use the matching function that is based on L1 metric. The matching function for Scalable Color is:

$$D_{SC} = \sum_{i=1}^N |H_A[i] - H_B[i]|$$

In principle, any other matching method suitable for histograms can be used, although it was found that L1 metric give very good retrieval performance in the MPEG-7 core experiments.

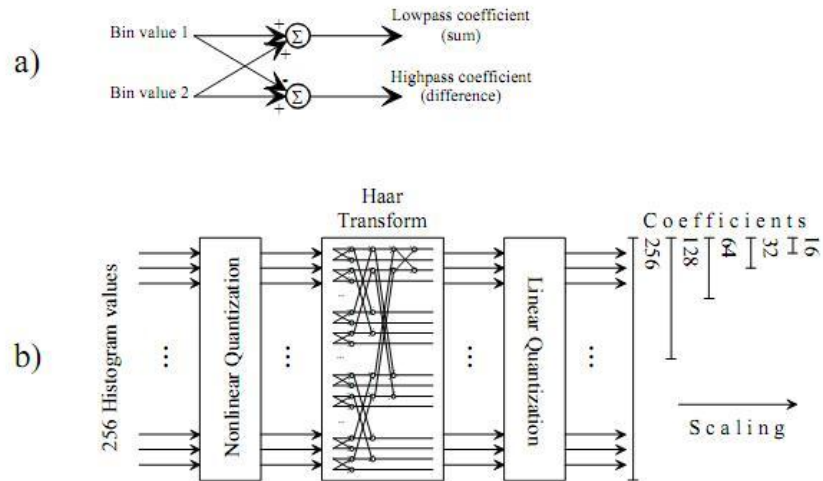


FIGURE 2.15: (a) Basic Unit of Haar transform (b) A schematic diagram of SCD generation.

# of coeff's	# of bins H	# of bins S	# of bins V
16	4	2	2
32	8	2	2
64	8	2	4
128	8	4	4
256	16	4	4

FIGURE 2.16: Equivalent partitioning of HSV color space for different numbers of coefficients in the SCD.

Color Structure Descriptor (CSD)

The Color Structure Descriptor is a color feature descriptor that captures both color content (similar to a color histogram) and information about the structure of this content (position of color).

The CSD works on a special version of the HMMD color space (fig. 2.15) dened by a non-uniform color space quantization. First, the HMMD color space is divided into five subspaces. This division is performed according to the Diff value where subspaces 0, 1, 2, 3, and 4 correspond respectively to Diff intervals $[0, 6]$, $[6, 20]$, $[20, 60]$, $[60, 110]$ and $[110, 255]$. Then a pixel is quantized along the Hue and Sum axes according to its subspace (fig. 2.18).

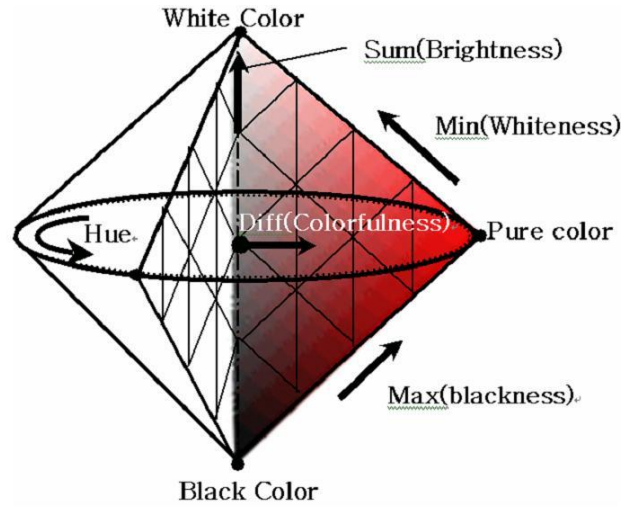


FIGURE 2.17: Double cone representation of the HMMD color space.

TABLE II
HMMD COLOR SPACE QUANTIZATION FOR CSD

Component	Subspace	Number of quantisation levels for different numbers of histogram bins			
		184	120	64	32
Hue	0	1	1	1	1
	1	8	4	4	4
	2	12	12	6	3
	3	12	12	4	2
	4	24			
Sum	0	8	8	8	8
	1	4	4	4	2
	2	4	4	4	4
	3	4	4	4	2
	4	2			

FIGURE 2.18: HMMD color space quantization for CSD.

Once the quantization step is done, the CSD is computed by visiting all locations in the image. At each location, the color C_m (with $m \in [0, M - 1]$) of all the pixels contained in the 8x8 structuring element overlaid are retrieved. The CSD bins are incremented according to these colors.

In other words, even if 14 of the 64 pixels are denoted by color C1 in the structuring element of figure 2.19, the bin C1 is only incremented of one.

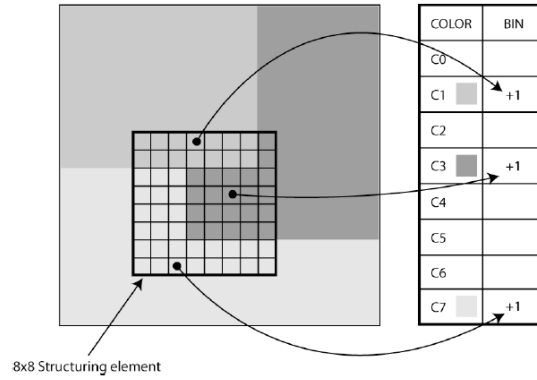


FIGURE 2.19: 8x8 structuring element.

This structuring element is exploited to avoid the lost of structure with typical histograms. The number of structuring points is always 64 and the distance between them increases with the image size (fig. 2.20).

Once the CSD histogram is computed, a non-linear quantization step is performed to obtain a 8-bits coding for each bins. Since the structure of the descriptor is the same, the same matching functions can be used. The default matching function is the L1 metric, as in the case of Scalable Color.

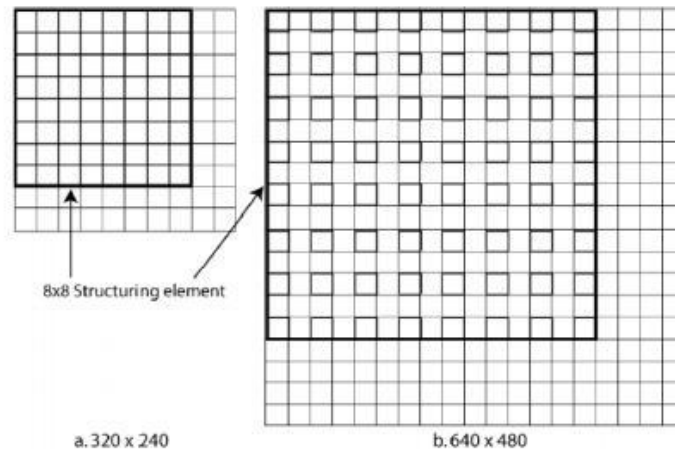


FIGURE 2.20: 64 structuring points for 2 structuring element sizes.

Color Layout Descriptor (CLD)

The CLD descriptor captures the spatial layout of the representative colors on a region or image. Representation is based on coefficients of the Discrete Cosine Transform. This is a very compact descriptor being highly efficient in fast browsing and search applications. It provides image-to-image matching as well as ultra high-speed sequence-to-sequence matching.

The Color Layout uses an array of representative colors for the image, expressed in the YCbCr color space, as the starting point for the descriptor definition. The size of the array is fixed to 8x8 elements to ensure scale invariance of the descriptor. The array obtained in this way is then transformed using the Discrete Cosine Transform (DCT), which is followed by zig-zag re-ordering (fig. 2.21).

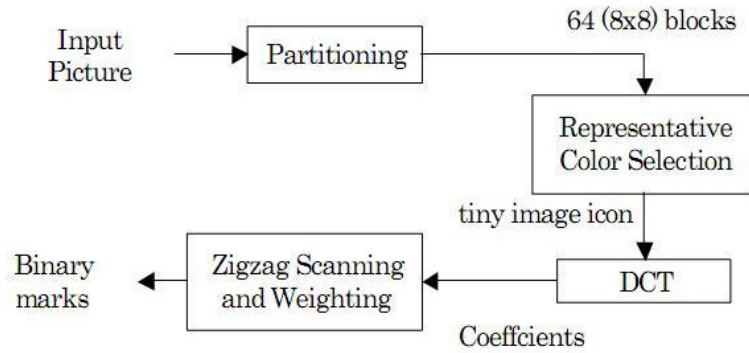


FIGURE 2.21: The extraction process of the Color Layout Descriptor.

A representative color was chosen for each block by averaging the values of all the pixels in each block. This results in three 8x8 arrays, one for each color component. This step is directly visualized in the first window of figure 2.22. Each 8x8 matrix was transformed to the YCbCr color space (second window of figure 2.22). Next each 8x8 matrix was transformed by 8x8 DCT to obtain 3 8x8 DCT matrices of coefficients, one for each YCbCr component (third window of figure 2.22). The CLD descriptor was formed by reading in zigzag order 6 coefficients from the Y-DCT-matrix and 3 coefficients from each DCT matrix of the two chrominance components. The descriptor is saved as an array of 12 values.

The default matching function is essentially a weighted sum of squared differences between the corresponding descriptor components.

$$D_{CL} = \sqrt{\sum_i w_i^y (Y_i - Y'_i)^2} + \sqrt{\sum_i w_i^b (C_{b_i} - C'_{b_i})^2} + \sqrt{\sum_i w_i^r (C_{r_i} - C'_{r_i})^2}$$

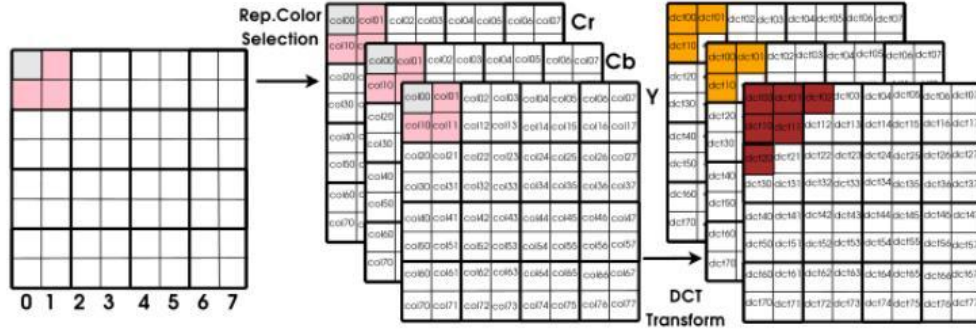


FIGURE 2.22: Stages of CLD computation.

where Y , Cb and Cr are the DCT coefficients of the respective color components, w_i^y , w_i^r , w_i^b are weights chosen to reflect the perceptual importance of the coefficients and the summation is over the number of coefficients.

CLD, SCD, CSD are described in detail in the following papers [25–27].

2.5.1.2 Color (Histogram) Descriptors

The color histogram Descriptors are statistics that can be viewed as an approximation of an underlying continuous distribution of colors values. They include simple color histogram, fuzzy color histogram and auto corellogram.

Simple Color Histogram Descriptor (SCHD)

The simple color histogram descriptor (SCHD) is a histogram descriptor that indicates the frequency of occurrence of every color in an image. The appealing aspect of the SCHD is its simplicity and ease of computation [28].

A Color histogram refers to the probability mass function of the image intensities. This is extended for color images to capture the joint probabilities of the intensities of the three color channels. More formally, the color histogram is defined by,

$$h_{A,B,C}(a,b,c) = N \text{Prob}(A = a, B = b, C = c)$$

where A , B and C represent the three color channels (R,G,B or H,S,V) and N is the number of pixels in the image.

Each pixel is associated to a specific histogram bin only on the basis of its own color, and color similarity across different bins color dissimilarity in the same bin are not

taken in account. Since any pixel in the image can be described three components in a certain colour space (for instance, red, green and blue components in RGB space or hue, saturation and value in HSV space) histogram, i.e., the distribution of the number of pixels for each quantized bin, can be defined for each component.

Computationally, the color histogram is formed by discretizing the colors within an image and counting the number of pixels of each color. Since the typical computer represents color images with up to 224 colors, this process generally requires substantial quantization of the color space. The main issues regarding the use of color histograms for indexing involve the choice of color space and quantization of the color space. When a perceptually uniform color space is chosen uniform quantization may be appropriate. If a non-uniform color space is chosen, then non-uniform quantization may be needed. The color histogram can be thought of as a set of vectors. For gray-scale images these are two dimensional vectors. One dimension gives the value of the gray-level and the other the count of pixels at the gray-level. For color images the color histograms are composed of 4-D vectors. This makes color histograms very difficult to visualize.

There are several distance formulas for measuring the similarity of color histograms using SCHD. In this work, we use the euclidean distance. Let h and g represent two color histograms. The euclidean distance between the color histograms h and g can be computed as:

$$d^2(h, g) = \sum_A \sum_B \sum_C (h(a, b, c) - g(a, b, c))^2$$

In this distance formula, there is only comparison between the identical bins in the respective histograms. Two different bins may represent perceptually similar colors but are not compared cross-wise. All bins contribute equally to the distance.

Fuzzy Color Histogram Descriptor (FCHD)

The fuzzy color histogram (FCHD), is a histogram descriptor that considers the color similarity of each pixel's color associated to all the histogram bins through fuzzy-set membership function.

Because each histogram bin represents a local color range in the given color space, color histogram represents the coarse distribution of the colors in an image. Two similar colors will be treated as identical provided that they are allocated into the same histogram bin. On the other hand, two colors will be considered totally different if they fall into two different bins even though they might be very similar to each other. This makes color histograms sensitive to noisy interference such as illumination changes and quantization

errors. Thus, fuzzy color histogram descriptor (FCHD) created to efficiently solve this aforementioned issue.

In contrast with Simple color histogram Descriptor (SCHD) which assigns each pixel into one of the bins only, FCHD considers the color similarity information by spreading each pixel's total membership value to all the histogram bins. FCHD is less sensitive to noisy interference such as lighting intensity changes and quantization errors than SCHD. Moreover, in contrast with quadratic histogram distance exploited for measuring the degree of similarity between SCHD's, simple Euclidean distance measurement over their FCHDs can yield similar retrieval results.

The color histogram is viewed as a color distribution from the probability viewpoint. Given a color space containing n color bins, the color histogram of image containing N pixels is represented as $H(I) = [h_1, h_2, h_3, \dots, h_n]$, where $h_i = N_i/N$ is the probability of a pixel in the image belonging to the i_{th} color bin, and N_i is the total number of pixels in the i_{th} color bin.

According to the total probability theory, h_i can be defined as follows:

$$h_i = \sum_{j=1}^N P_{i|j} P_j = \frac{1}{N} \sum_{j=1}^N P_{i|j}$$

where P_j is the probability of a pixel selected from image I being the j_{th} pixel, which is $\frac{1}{N}$, and $P_{i|j}$ is the conditional probability of the selected j_{th} pixel belonging to the i_{th} color bin. In the context of SCHD, $P_{i|j}$ is defined as:

$$P_{i|j} = \begin{cases} 1, & \text{if the } j_{th} \text{ pixel is quantized into the } i_{th} \text{ color bin} \\ 0, & \text{otherwise} \end{cases}$$

This definition leads to the boundary issue of SCHD such that the histogram may undergo abrupt changes even though color variations are actually small. This reveals the reason why the SCHD is sensitive to noisy interference such as illumination changes and quantization errors.

The FCHD essentially modifies probability $P_{i|j}$ as follows. Instead of using the probability $P_{i|j}$, we consider each of the N pixels in image I being related to all the n color bins via fuzzy-set membership function such that the degree of "belongingness" or "association" of the j_{th} pixel to the i_{th} color bin is determined by distributing the membership value of the j_{th} pixel, μ_{ij} , to the i_{th} color bin.

The fuzzy color histogram descriptor (FCHD) of image I can be expressed as:

$$F(I) = [f_1, f_2, f_3, \dots, f_n]$$

where

$$f_i = \sum_{j=1}^N \mu_{ij} P_j = \frac{1}{N} \sum_{j=1}^N \mu_{ij}$$

P_j has been defined in (1), and μ_{ij} is the membership value of the j_{th} pixel in the i_{th} color bin.

In contrast with SCHD, FCHD considers not only the similarity of different colors from different bins but also the dissimilarity of those colors assigned to the same bin. Therefore, FCHD effectively alleviates the sensitivity to the noisy interference. A useful reference for this descriptor is [29].

Auto - Corellogram Descriptor (ACD)

The Auto (color) Correlogram Descriptor (ACD) expresses how the spatial correlation of pairs of colors changes with stance. ACD includes the spatial correlation of colors, can be used to describe the global distribution of local spatial correlation of colors, it is easy to compute, and the size of the feature is fairly small.

Let I be an $n \times n$ image (for simplicity, we assume that the image is square). The colors in I are quantized into m colors $c_1, c_2, c_3, \dots, c_m$. (In practice m , is a constant) For a pixel $p = (x, y) \in I$, let $I(p)$ denote its color. Let $I_c = \{p | I(p) = c\}$. Thus, the notation $p \in I_c$ is synonymous with $p \in I, I(p) = c$. For convenience, we use the L_∞ -norm to measure the distance between pixels, i.e., for pixels $p_1 = (x_1, y_1), p_2 = (x_2, y_2)$, we define $|p_1 - p_2| = \max\{|x_1 - x_2|, |y_1 - y_2|\}$. We denote the set $\{1, 2, \dots, n\}$ by $[n]$. The histogram h of I is defined for $i \in [m]$ by:

$$h_{ci}(I) = n^2 Pr[p \in I_{ci}]$$

For any pixel in the image, $h_{ci}(I)/n^2$, gives the probability that the color of the pixel is c_i .

Let a distance $d \in [n]$ be fixed a priori. Then, the correlogram of I is defined for $i, j \in [m], k \in [d]$ as:

$$\gamma_{ci,cj}^{(k)}(I) = \sum_{p_1 \in I_{ci} p_2 \in I_{cj}} Pr[p_2 \in I_{cj} | |p_1 - p_2| = k]$$

Given any pixel of color c_i in the image, $\gamma_{c_i, c_j}^{(k)}$ gives the probability that a pixel at distance k away from the given pixel is of color c_j . Note that the size of the correlogram is $O(m^2d)$. The autocorrelogram of I captures spatial correlation between identical colors only and is defined by:

$$\alpha_c^{(k)}(I) = \gamma_{c,c}^{(k)}(I)$$

For example consider the simple case when $m = 2$ and $n = 8$. Two sample images are shown below in figure 2.23(a). The autocorrelograms corresponding to these two images are shown in figure 2.23(b). The change of autocorrelation of the foreground color with distance is perceptibly different for these images. Note that it is difficult to distinguish between these two images using histograms. More details about corellograms can be found here [30].

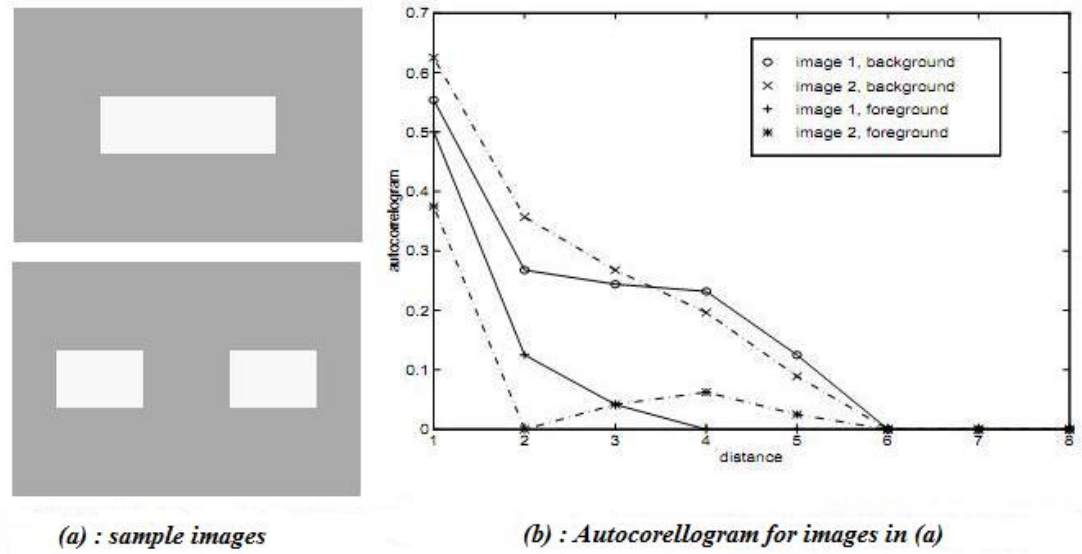


FIGURE 2.23: Example of autocorellograms (b) for the images in (a)

2.5.2 Texture Descriptors

Image texture has emerged as an important visual primitive to search and browse through large collections of similar looking patterns. An image can be considered as a mosaic of textures and texture features associated with the regions can be used to index the image data. Color descriptors has shown good performance for discriminating images based on color. However, in many cases, color is usually insufficient for discriminating between images with the same color but different texture. Texture features are

capable of recognizing repeated patterns in an image, analyzing the energy distribution in the frequency domain.

2.5.2.1 Edge Histogram Descriptor (EHD)

The edge histogram descriptor (EHD) captures the spatial distribution of edges, somewhat in the same spirit as the CLD. The edge histogram descriptor represents the spatial distribution of five types of edges, namely four directional edges and one non-directional edge.

Since edges play an important role for image perception, it can retrieve images with similar semantic meaning. Thus, it primarily targets image-to-image matching (by example or by sketch), especially for natural images with non-uniform edge distribution. In this context, the image retrieval performance can be significantly improved if the edge histogram descriptor is combined with other Descriptors such as the color histogram descriptor. Besides, the best retrieval performances considering this descriptor alone are obtained by using the semi-global and the global histograms generated directly from the edge histogram descriptor as well as the local ones for the matching process.

This descriptor is implemented as follows: Firstly, a gray-intensity image is divided in 4×4 sub-images. Each sub-image has its own local histogram with 5 bins. These 5 bins correspond to the 5 edge types: vertical, horizontal, 45 diagonal, 135 diagonal, and isotropic.

In order to fill the local histograms each sub-image is divided in macroblock. A macroblock is composed by 22 macropixels and is associated to an edge type. Figure 2.24 gives details on how the macroblock are built.

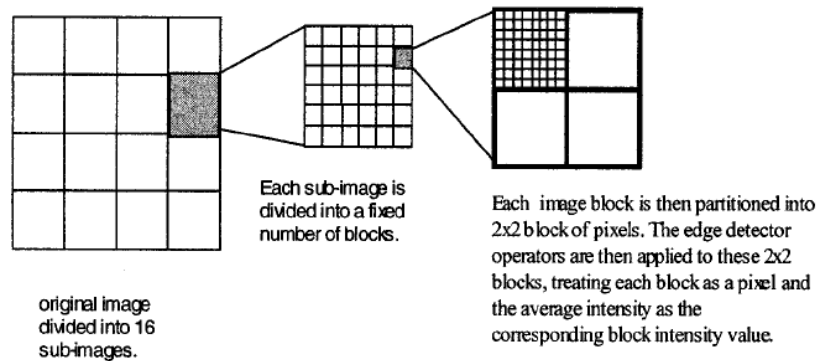


FIGURE 2.24: Subimages and macroblocks decomposition.

To associate a macroblock with an edge type, a convolution with 5 simple edge detectors is performed and the one with the strongest reply is linked to the macroblock:

$$\begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix} \begin{vmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{vmatrix} \begin{vmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{vmatrix} \begin{vmatrix} 2 & -2 \\ -2 & 2 \end{vmatrix}$$

The local histogram is therefore built counting the result of each macroblock. Finally the global histogram is quantized in 3 bits per bin (EHD(i) 2 [0 7]).

Note that there are a total of 80 bins, 3 bits/bin, in the edge histogram. One can use the 3-bit number as an integer value directly and compute the L1 distance between two edge histograms. More Information about EHD can be found in [31].

2.5.2.2 Tamura Descriptor (TD)

Tamura et. al.[32] proposed texture features that correspond to human visual perception. They defined six textural features (coarseness, contrast, directionality, line-likeness, regularity and roughness) and compared them with psychological measurements for human subjects. The first three features described below attained very successful results and are used in our evaluation, both separately and as joint values.

Coarseness has a direct relationship to scale and repetition rates and was seen by Tamura as the most fundamental texture feature. An image will contain textures at several scales coarseness aims to identify the largest size at which a texture exists, even where a smaller micro texture exists. Computationally one first takes averages at every point over neighborhoods the linear size of which are powers of 2. The average over the neighborhood of size $2^k \times 2^k$ at the point (x, y) is:

$$A_k(x, y) = \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} f(i, j) / 2^{2k}$$

Then at each point one takes differences between pairs of averages corresponding to non-overlapping neighborhoods on opposite sides of the point in both horizontal and vertical orientations. In the horizontal case this is:

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)|$$

At each point, one then picks the best size which gives the highest output value, where k maximizes E in either direction. The coarseness measure is then the average of $S_{opt}(x, y) = 2^{k_{opt}}$.

Contrast aims to capture the dynamic range of grey levels in an image, together with the polarization of the distribution of black and white. The first is measured using the standard deviation of grey levels and the second the kurtosis α_4 . The contrast measure is therefore dened as:

$$F_{con} = \sigma / (\alpha_4)^n$$

where,

$$\alpha_4 = \mu_4 / \sigma^4$$

Experimentally, Tamura found $n = 1/4$ to give the closest agreement to human measurements. This is the value we used in our experiments.

Directionality is a global property over a region. The feature described does not aim to differentiate between different orientations or patterns, but measures the total degree of directionality. Two simple masks are used to detect edges in the image. At each pixel the angle and magnitude are calculated. A histogram, Hd , of edge probabilities is then built up by counting all points with magnitude greater than a threshold and quantizing by the edge angle. The histogram will reflect the degree of directionality. To extract a measure from Hd the sharpness of the peaks are computed from their second moments. Finally distances between images vectors were calculated upon feature vectors using the Manhattan metric. An interesting reference to Tamura descriptor and its image retrieval use is [33]

2.5.2.3 Gabor Descriptor (GD)

Gabor descriptor is a texture descriptor based on a multiresolution decomposition using gabor wavelets and is based on gabor filtering [34].

The descriptor has two parts: The first part relates to a perceptual characterization of texture in terms of structuredness, directionality and coarseness (scale). This part is called the perceptual browsing component (PBC). The second part provides a quantitative description that can be used for accurate search and retrieval and is referred to as the similarity retrieval component (SRC). Both of the components are derived from a multi resolution Gabor filtering.

From the multi resolution decomposition, a image is decomposed into a set of filtered images. Each of these images represents the image information at a certain scale and at a certain orientation. The PBC captures the regularity (or the lack it) in the texture pattern.

$$PCB = [u_1, u_2, u_3, u_4, u_5]$$

Regularity (u_1): u_1 represents the degree of regularity or structuredness of the texture. A larger value of u_1 indicates a more regular pattern. Consider the two patterns in the figure 2.25. Pattern (a) is intuitively more regular than (b), and hence should have a larger u_1 compared to (b).

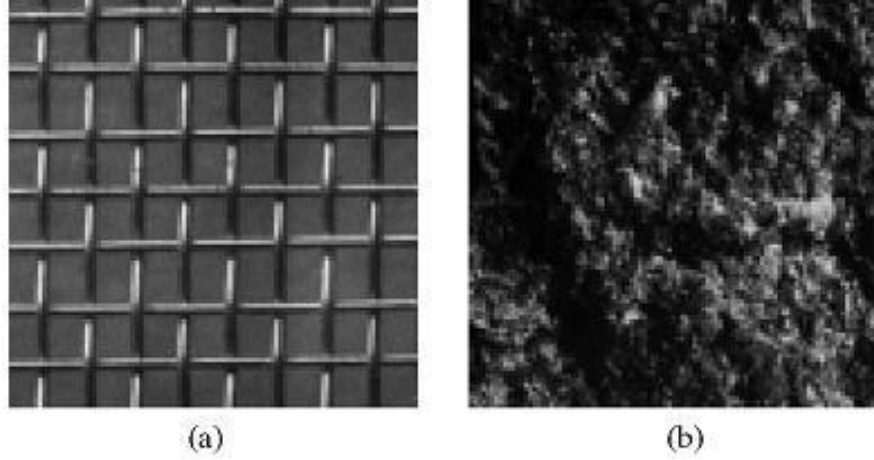


FIGURE 2.25: Regularity of patterns (a)regular pattern, (b)irregular pattern.

Directionality (u_2, u_3): These represent the two dominant orientations of the texture. The accuracy of computing these two components often depends on the level of regularity of the texture pattern.

Scale (u_4, u_5): These represent two dominant scales of the texture. Similar to directionality, the more structured the texture, the more robust the computation of these two components.

Finally to Compute the similarity retrieval component (SRC). The mean μ_{mn} and the standard deviation σ_{mn} of the magnitude of the transform coefficients are used to form the SRC:

$$\mu_{mn} = \int \int |W_{mn}(x, y)| dx dy$$

and

$$\sigma_{mn} = \sqrt{\int \int (|W_{mn}(x, y)| - \mu_{mn})^2 dx dy}$$

The similarity retrieval component (SRC) vector is now constructed using μ_{mn} and σ_{mn} . For S scales and K orientations, this results in a vector:

$$SRC = [\mu_{11}\sigma_{11} \dots \mu_{SK}\sigma_{SK}]$$

In order to use this descriptor to perform similarity retrieval, we use a distance measure on the feature vector of GD. Consider two image patterns i and j . Then the distance between the two patterns is:

$$d(i, j) = \sum_m \sum_n d_{mn}(i, j)$$

where

$$d_{mn}(i, j) = \left| \frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})} \right|$$

$\alpha(\mu_{mn})$ and $\alpha(\sigma_{mn})$ are the standard deviations of the respective features over the entire database, and are used to normalize the individual feature components.

2.5.3 Hybrid Descriptors:

Hybrid descriptors can be formulated by incorporating color and texture to a new descriptor. We will describe three low-level descriptors, used in this work, which contain both color and texture information.

2.5.3.1 Color Edge Directivity Descriptor (CEDD)

The Color edge directivity descriptor (CEDD) is a descriptor that combines, in one histogram, color and texture information [35]. CEDD size is limited to 54 bytes per image, rendering this descriptor suitable for use in large image databases.

First, the image is divided in a fixed number of blocks (e.g. 3x3). In order to extract the color information, a set of fuzzy rules undertake the extraction of a Fuzzy-Linking histogram. This histogram stems from the HSV color space. Twenty rules are applied to a three-input fuzzy system (one for each HSV) in order to generate eventually a 10-bin quantized histogram. Each bin corresponds to a preset color. The number of blocks assigned to each bin is stored in a feature vector. Then, 4 extra rules are applied to a two input fuzzy system, in order to change the 10-bins histogram into 24-bins histogram, importing thus information related to the hue of each color that is presented.

Next, the 5 digital filters that were proposed in the MPEG-7 Edge Histogram Descriptor are also used for exporting the information which is related to the texture of the image, classifying each image block in one or more of the 6 texture regions that has been fixed, shaping thus the 144 bins histogram. With the use of the Gustafson Kessel fuzzy classifier 8 regions are shaped, which are then used in order to quantize the values of the 144 CEDD factors in the interval 0-7, limiting thus the length of the descriptor in 432 bits. Schematic diagram of CEDD is presented in figure 2.26.

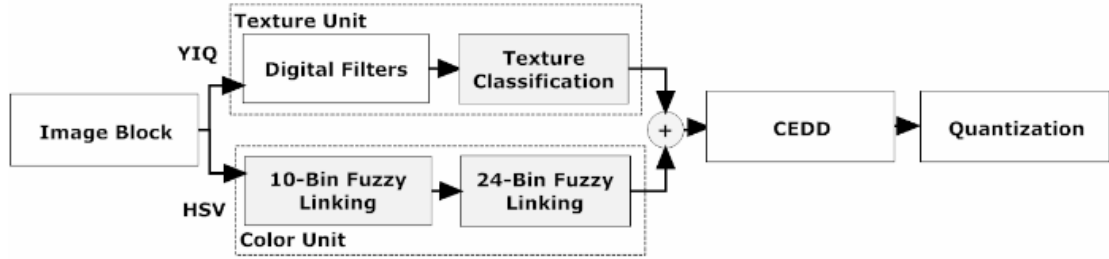


FIGURE 2.26: A schematic diagram of CEDD.

For the measurement of the distance of CEDD between images, Tanimoto coefficient is used:

$$T_{ij} = t(x_i, x_j) = \frac{x_i^T x_j}{x_i^T x_i + x_j^T x_j - x_i^T x_j}$$

Where x^T is the transpose vector of x .

2.5.3.2 Fuzzy Color Texture Histogram Descriptor (FCTHD)

The Fuzzy Color Texture Histogram Descriptor (FCTHD) results from the combination of 3 fuzzy systems. FCTHD size is limited to 72 bytes per image. FCTHD works exactly the same with CEDD with a little difference in texture information extraction.

Initially the image is segmented in a present number of blocks. Next it extracts the same color information as CEDD. For the extraction of texture information each image block is transformed with Haar Wavelet transform and a set of texture elements are exported. These elements are used as inputs in a third fuzzy system which converts the 24-bins histogram in a 192-bins histogram, importing texture information in the proposed feature. Eight rules are applied in a three-input fuzzy system. For the quantization process Gustafson Kessel fuzzy classifiers are used. You can see the schematic diagram of FCTH in figure 2.27.

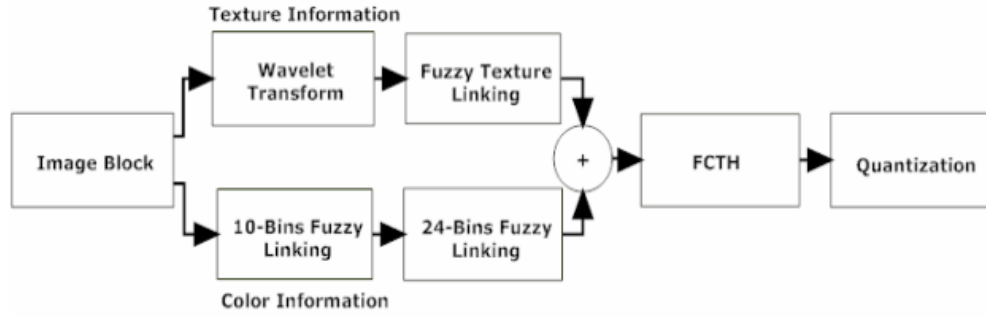


FIGURE 2.27: A schematic diagram of FCTH.

For the measurement of the distance of FCTH between the images, Tanimoto coefficient is used. Additional information about the descriptor can be found at [36].

2.5.3.3 Joint Composite Descriptor (JCD)

Joint Composite Descriptor (JCD) is a combined compact vector that contains color and texture information at the same time. JCD successfully combines CEDD and FCTH[37].

As mentioned above, the structure of CEDD and FCTH descriptors consists of n texture areas. Each texture area is separated into 24 sub regions, with each sub region describing a color. Both CEDD and FCTH use the same color information, as it results from 2 fuzzy systems that map the colors of the image in a 24-color custom palette. To extract texture information, CEDD uses a fuzzy version of the five digital filters proposed by the EHD forming 6 texture areas.

In contrast, FCTH uses the high frequency bands of the Haar wavelet Transform in a fuzzy system, to form 8 texture areas. The types of texture areas adopted by each descriptor are illustrated below in figure 2.28.

This new descriptor is made up of 7 texture areas, with each area made up of 24 sub regions that correspond to color areas. The colors that represent these 24 sub regions are: (0) White, (1) Grey, (2) Black, (3) Light Red, (4) Red, (5) Dark Red, (6) Light Orange, (7) Orange, (8) Dark Orange, (9) Light Yellow, (10) Yellow, (11) Dark Yellow, (12) Light Green, (13) Green, (14) Dark Green, (15) Light Cyan, (16) Cyan, (17) Dark Cyan, (18) Light Blue, (19) Blue, (20) Dark Blue, (21) Light Magenta, (22) Magenta, (23) Dark Magenta. The texture areas are as follows: JCD(0) Linear Area, JCD(1) Horizontal Activation, JCD(2) 45 Degrees Activation, JCD(3) Vertical Activation, JCD(4) 135 Degrees Activation, JCD(5) Horizontal and Vertical Activation and JCD(6) Non directional Activation.

	0	1	2	3	4	5	6	7
CEDD	Linear	Non Directional	Horizontal Activation	Vertical Activation	45 Degree Diagonal	135 Degree Diagonal	-	-
FCTH	Linear Low Energy	Horizontal Low Energy	Vertical Low Energy	Both Directions Low Energy	Linear High Energy	Horizontal High Energy	Vertical High Energy	Both Directions High Energy

FIGURE 2.28: Compact composite descriptors texture areas.

2.5.4 Geometrical Descriptors

Two simpler, but equally important measures are the area and the circularity. The area of an object constitute all the pixels of an image that belong to the certain object. Area represents the size of the object, when compared to other objects, on the same scale.

Circularity is measured by the isoperimetric inequality: $\frac{4A\pi}{P^2} \leq 1$, where A is the Area of the object and P its perimeter. According to this measure, a circle is the most circular object (i.e., has the largest circular value). For other objects the ratio is smaller. As a circle is tilted and assumes the shape of an ellipse, the area decreases but the perimeter does not decrease as rapidly, so the circularity decreases. At the extreme angle of tilt, the ellipse is squeezed into a line and the circularity goes to zero. For digital images, circularity provides a good measure of how circular is an object in relation to a circle.

Chapter 3

Methodology

In this section the methodology used for our application is presented. This includes the process of capturing pollen images from the samples of MAICH, the segmentation of the images in order to create our pollen database, the feature extraction for each image, the construction of the pollen ontology and finally the image authentication and classification.

3.1 Sampling Method

3.1.1 Preparation of Pollen Reference Specimens

Pollen can be recovered from fresh flowers, or from pressed or dried specimens. For reference material to be useful to others accurate identification of the plant is essential and this is done by specialized botanists consulting the appropriate Floras, taxonomic keys and herbarium specimens. The identification of the plants used for this work has been done by botanists of the Herbarium of MAICH (voucher specimens of the plants are kept in the herbarium MAICH).

There are two methods of sample preparation for pollen reference specimens: the acetolytic one (first of all described by Erdtman, 1952), and the one prescribed by the International Commission for Plant-Bee Relationship (Louveaux et al. 1978). With acetolysis, concentrated sulphuric acid is used during preparation and the cytoplasm and the intina of the granules are destroyed. The pollen resembles the results of the second method in fresh honey analysis, while in a pollen that is several years old (other cases of palynology, such as for e.g. in soil samples) and no longer has its cytoplasm, the pollen resembles an acetolytic compound.

In the present work, the second method has been used with some minor modifications as following: Flowers were collected fresh. A drop of a 20% glucose solution is placed on a microscope slide. The anthers of the flowers are kept with tweezers and shaken above the microscope slide which then is placed on a heated plate (round 40 C) to dry; it is then covered with a cover glass that has a drop of Entellan (Merck1.07961.0500). In some cases, staining with fuchsin (0,1% in ethanol, Merck 1.00983.2511) is performed in order to be able to examine some pollen characteristics such as pores, colps, ornamentation, etc.

3.1.2 Laboratory Equipment

Microscope

For the examination of each specimen slide we used a Leica DM L B optical Microscope. The microscope consists of several basic stands and a range of modular components allowing an almost unlimited variety of individual outfits. It uses a 12 V 30 W halogen lamp. Photos have been taken with a magnification of 400x with normal lens and with the use of Nomarski devices (prisms) [38].

Nomarski (DIC) Microscopy

Nomarski or differential interference contrast (DIC) microscopy (invented by Georges Nomarski in the mid-1950s) is used to image, living or stained specimens, which contain little or no optical contrast when viewed using brightfield illumination [39].

In transmitted light DIC, light from the lamp is passed through a polarizer located beneath the condenser, in a manner similar to polarised light microscopy [40]. The polarized light then passes through a modified Wollaston prism (below the condenser) that splits the entering beam of light into two beams travelling in slightly different directions but vibrating perpendicular to each other, and therefore unable to recombine to cause interference. The distance between the two beams is called the "shear" distance, and is always less than the resolving ability of the objective, to prevent the appearance of double images. The split beams enter and pass through the specimen where their paths are altered by the specimen's varying thickness, slopes, and refractive indices. When the parallel beams enter the objective, they are focused above the rear focal plane where they enter a

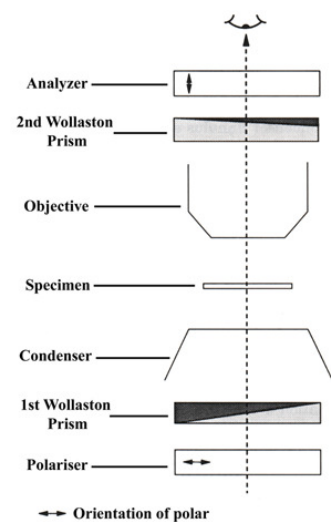


FIGURE 3.1: Nomarski Microscopy Arrangement.

second modified Wollaston prism that recombines the two beams at a defined distance outside of the prism. This removes the shear and the original path difference between the beam pairs. However, the parallel beams are no longer the same length because of path changes caused by the specimen. In order for the parallel beams to interfere with each other, the vibrations of the beams of different path length must be brought into the same plane and axis. This is accomplished by placing a second polarizer (analyzer) above the upper Wollaston beam-combining prism. The light then proceeds toward the eyepiece where it can be observed as differences in intensity and color. DIC microscopy causes one side of an object to appear bright (or colored) while the other side appears darker (or a different colour). This shadow effect gives a pseudo three-dimensional appearance to the specimen, but is not a true representation of the geometry of the specimen, because it is based on optical thickness. The pseudo 3D appearance of the specimen can also be profoundly affected by its orientation i.e. rotation of the specimen by 180 degrees can change a hill into a valley or vice versa. Therefore, DIC microscopy is not suitable for accurate measurement of actual heights and depths.

Camera and Software

The pictures of the pollen specimens, examined on the microscope, were captured with the use of ProgRes C12 Plus camera and ProgRes Capture Pro 2.1 software. ProgRes is an optical, fine mechanical precision device. It has been built and tested according to the regulations specified in EN60950 / VDE0805. ProgRes digital cameras are designed and built for use in clean and dry locations for digital image acquisition [41].

The technical characteristics of the camera are shown in figure 3.2 and the supported resolutions in figure 3.3. The supported image formats are the following:

- BMP - 8 bits per color channel
- TIFF 8/16 bits per color channel
- JPEG high/medium/low quality

The ProgRes camera is connected to the computer with an IEEE 1394 (FireWire) interface. Operating the ProgRes camera is done via the ProgRes Capture Pro 2.1 application software. The software provides several tools for controlling the camera, with more important, for our work, the control of light exposure of the lens. This feature was very helpful on acquiring clean and light balanced images, something that had an impact on the segmentation process. The software, also, provides the option to insert a scale bar on the images. This way it allows keeping all the images at the same scale, as there are many variations on the size of pollen grains.

Technical Data ProgRes® C12^{plus}

Image Sensor	SONY ICX282AQ Super HAD
Type	2/3" Interline Transfer Color (Type: Bayer Mosaic) CCD
Active Pixels	2580 x 1944
Active Sensor Area	8.8 mm x 6.6 mm (corresponding aspect ratio 4:3)
Pixel Pitch	3.4 μ m x 3.4 μ m
IR Cut Off Filter	Hoya C-500S
Sensor Cooling	None
Camera Electronics	JENOPTIK LOS – LC5-LowNoise Electronics
AD Conversion	12 bit
Dynamic Range	Typ. 60 dB (according to raw image)
Power Supply	Via IEEE1394a Firewire
Rated Voltage	9 V – 33 V
Power Consumption	6 W
Data Interface	IEEE1394a FireWire
Optical Connector	C-Mount, 1" x 1/32" Warning: max. penetration depth for objectives: 7 mm
Indicator	LED (status indicator red / green)
Tripod Thread	3/8" and 1/4" reduction insert
Dimensions	145 mm x 93 mm x 123 mm (Height x Width x Depth)
Weight	780 g

FIGURE 3.2: ProgRes C12 Camera

Please see the following tables for the resolutions available for ProgRes® C12^{plus}.

Image Capture	
Resolution	Description
2580 x 1944	full sensor resolution
2580 x 1944 F	full sensor resolution without anti-moiré filter
1290 x 972 VGA	for capturing moving objects
644 x 490 VGA	for capturing moving objects
1290 x 972 2x bin	2 x 2 binning
860 x 648 3x bin	3 x 3 binning
645 x 486 4x bin	4 x 4 binning
516 x 389 5x bin	5 x 5 binning

FIGURE 3.3: ProgRes C12 Camera - Resolutions

3.1.3 Image Capturing and Settings

For capturing the images from the specimens we followed the following process:

At first, we examined each specimen with the microscope and tried to find certain areas of the specimen that contained distinctive pollen grains of the species. Then, each area was captured at two different focal points. One that focused on the sporoderm of the grains and another that focused on the cytoplasm, in order to capture all the distinctive characteristics of the grain anatomy. Light exposure had to be adjusted according to every different specimen, as there were many variations between the different tiles.

This process was repeated, after the application of the Nomarski filter on the microscope. Because of the double polarization that the Nomarski filter applies on the light of the microscope the first images were too dark. In order to acquire brighter images, light exposure was adjusted accordingly along with the application of more light by the microscope lamp. The result images were better, compared to the dark ones, but still the quality was worse than the original light microscope images. The images were too bright for the 3D effect, that the Nomarski filter applied, to be clearly and distinguishably visible.

With a view to acquire the best images possible, we selected the best settings that the equipment could provide. Specifically, we captured the images with a resolution of 2580x1944 pixels and saved them using JPEG high quality file format. All images were captured, through the microscope, using 400x optical zoom, by using the 40x objective and the 10x eyepieces.

3.2 Segmentation

After acquiring the images of the pollen grains, the next step was to isolate each grain into a new image. Examining the specimen image, the first problem that had to be resolved was the elimination of the background. Background is the main element of noise in our images, as it covers the biggest part of the images. Background had to be extracted in a way that the pollen grains wouldnt be distorted.



FIGURE 3.4: Original capture from the microscope of the pollen species Thymbra of genus Satureja.

3.2.1 Use of the HSI color model

The choice of working on a different color model than the RGB, for the background subtraction, was based on the fact that in the majority of the samples the main factor that differentiates the background from the pollen grains is the saturation and intensity, rather than the color. The HSI color model allows better processing of these measures,

in contrast to the RGB model, which limits the choices of processing to just the manipulation of the three primary colors.

Hints, occurred by the observation of the images, that there was a distinctive difference between the saturation of the background and the grains were confirmed by examining the histograms of both Hue and Saturation vectors of the HSI color model. The histograms noted that of the two curves, with the first representing the pixels that compose the background and the second the pixels that belong to the grains. An example can be seen in figure 3.5 for Hue and in figure 3.6 for Saturation.

It was noticeable that the curves were more distinct on the Saturation histogram. This observation led to the assumption that with a thresholding technique we could remove most of the background and keep the grains.

The conversion from RGB to HSI color model was done with the use of the equations presented in Chapter 2.

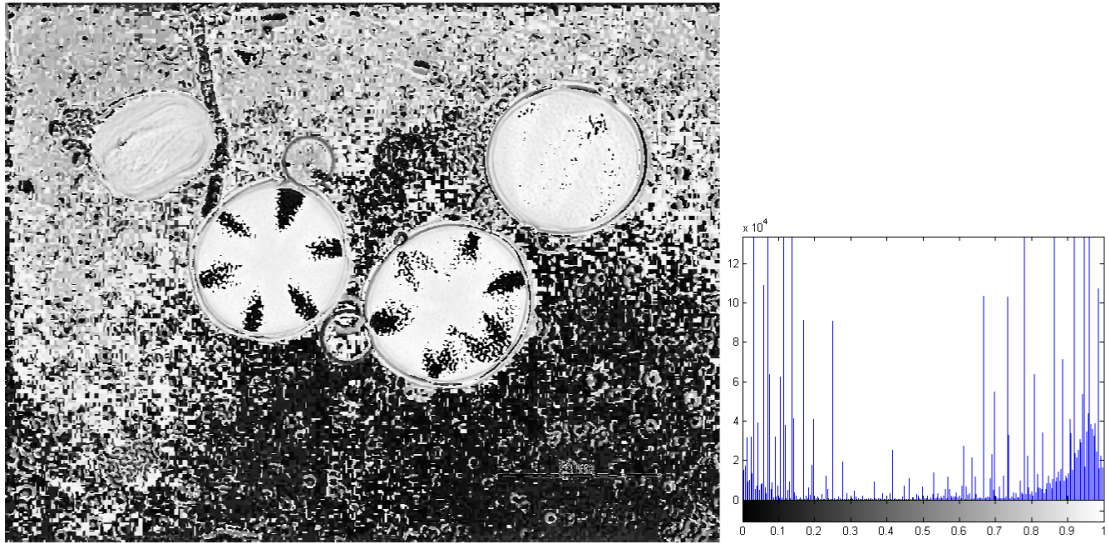


FIGURE 3.5: Hue and Histogram of Thymbra.

3.2.2 Thresholding / Binarization

Thresholding allowed the subtraction of the background of the images and retain the pollen grains. Otsu's method, presented in Chapter 2, was used.

In order to obtain the best result, we had to experiment with thresholding on both Hue and Saturation attributes of the HSI color model.

Due to the nature of the Hue attribute, as it contains the information about the color of the image, thresholding led to results that had either less or more than the information

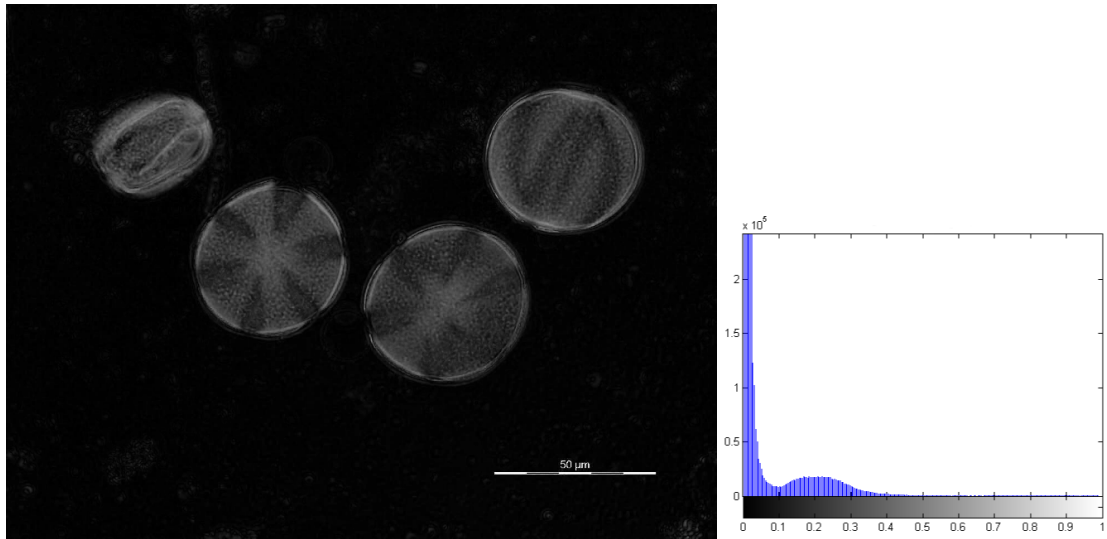


FIGURE 3.6: Saturation and Histogram of Thymbra.

needed. In some cases, many pixels of the background were included in the output image and in other cases many pixels of the grains were excluded (fig. 3.7).

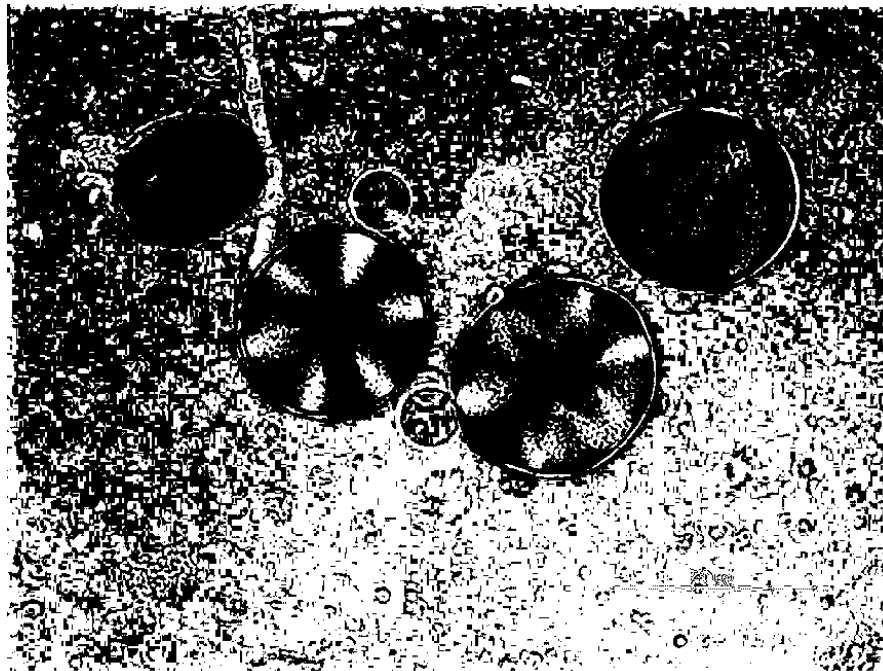


FIGURE 3.7: Otsu Thresholding on Hue.

Thresholding applied on the Saturation attribute provided the best results, without nevertheless avoiding to lose information of the pollen grains (fig. 3.8).

We, also, experimented on a combination of the output images of thresholding Hue and Saturation, combining the pixels that survived from each procedure. The goal was to take advantage of some regions, mainly on the sporoderm of the grains, that survived with thresholding on Hue and not on Saturation attribute. The result, still, maintained

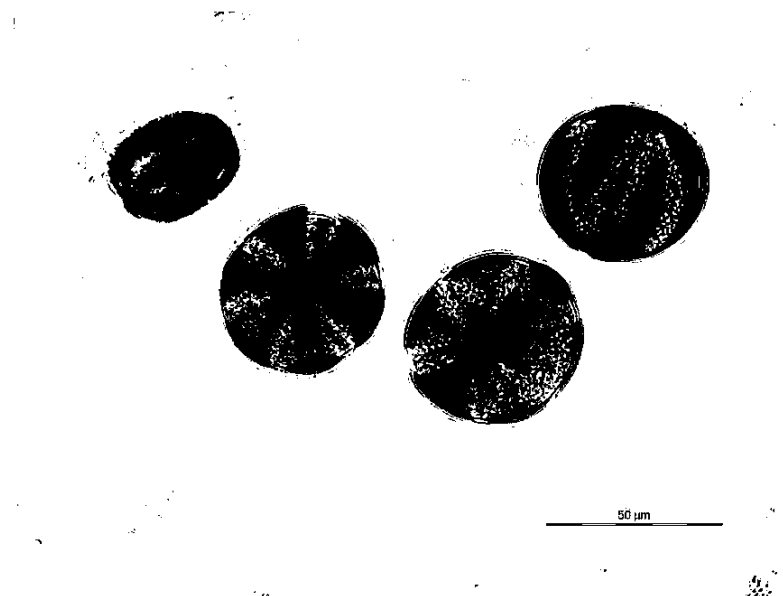


FIGURE 3.8: Otsu Thresholding on Saturation.

all the noise that Hue applied on the output, a fact that made this try inefficient (fig. 3.9).



FIGURE 3.9: Combination of the thresholded images on Hue and on Saturation.

3.2.3 Morphological Operators

The result images after thresholding contained most of the vital information (pixels) of the pollen grains, but also contained noise regions and gaps in the grains. Noise regions

survived the thresholding process mainly due to defects on the specimens(trapped air, bubbles of the preservative substance, dirt on the objective lenses of the microscope, etc). In order to expunge the noise regions and fill the gaps of the grains we resorted to the use of morphological operators, presented in Chapter 2.

First, we used the opening operator (erosion followed by dilation) to remove the small noise regions. This process also removed some of the pixels of the grains, which were, still, recovered in the next step. Then, we used the closing operator (dilation followed by erosion), on the modified image, to close the gaps in the grains and form compact regions, each one of which represents one pollen grain, containing as much detailed information as possible.

As structuring element we ended up, after tests, using a circle. Structuring elements that resemble to the objects being processed tend to provide better results. Pollen grains, for the most part, are circular or ellipsoid. The size of the structuring element depended on each grain, due to the size differences, but for the majority of the species we used a circle with a radius of 20 pixels.

After applying the morphological operators, the results were binary images where the regions detected as pollen grains were marked as connected components, with black colored pixels (fig. 3.10). Some noise regions survived the process, as they were large enough to contain the structuring element.

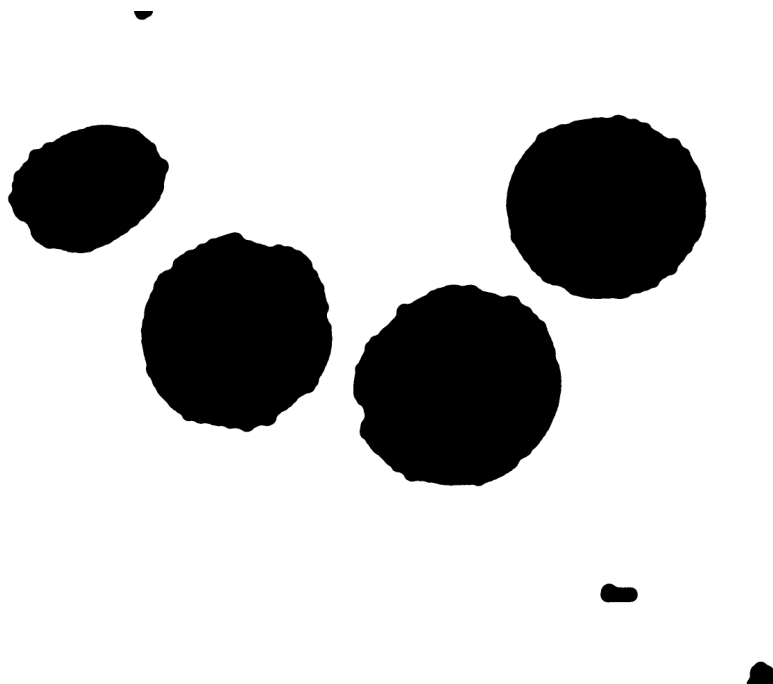


FIGURE 3.10: Result after the use of morphological operators on Saturation.

3.2.4 Component Labeling

In order to isolate each pollen grain and create a new image for each one of them, it was necessary to use a component labeling algorithm.

The algorithm used in this work was developed in the Institute of Information Science, Taiwan, in 2003. The main step of this algorithm is to use a contour tracing technique to detect the external contour and possible internal contours of each component, and also to identify and label the interior area of each component. Labeling is done in a single pass over the image, while no re-labeling is required. An example of the result is presented in figure 3.11.

For each component detected a new image was created and stored, using the corresponding pixels of the original image of each specimen. In this stage of our methodology, small regions of noise, detected as components, were not counted as pollen grains and consequently not stored. It was safe to exclude these noise regions as they were too small to be compared to the size of the grains of any of the pollen species used in this work.

Still, the specimens contained parts of the plants or not fully developed pollen grains that survived the segmentation process, mainly due to their size. These parts were also detected by the component labeling process, as connected components. Exclusion with a simple automated process, as above, was not an option as it might also remove pollen grains. These images were removed manually after the segmentation process.

Note here, that this procedure didn't demand the involvement of a qualified scientist, as the plant parts could be easily identified by observation and differentiated from the pollen grains.

3.2.5 Remarks

An important problem that this method came across was the fact that in some specimen tiles the pollen grain were contiguous. That meant that no background, to be detected by the thresholding process, was between the grains and the results were big areas that contained more than one grain. Morphological operators didn't manage to separate the grains as there were no holes between the grains for the structure elements to eliminate. In some cases, the grains were not contiguous, but too close to each other. There, the structure elements tended to close those small gaps that they didn't manage to fill in. Consequently, component labeling detected these areas as one and labeled them accordingly. The final results were images that contained more than one, contiguous, pollen grains. Still, the background was subtracted and it was possible to edit the images and manually separate the grains.

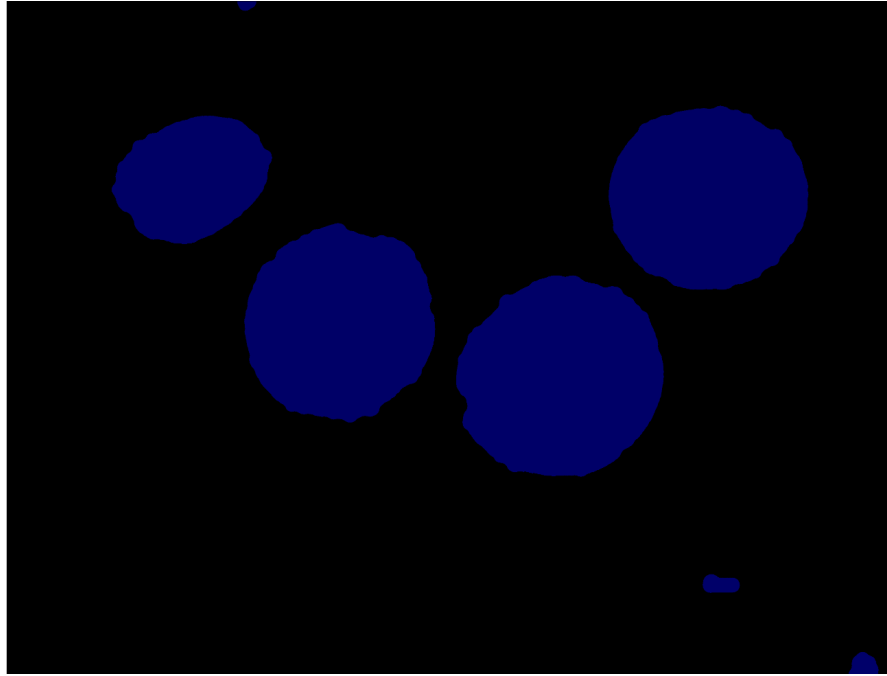


FIGURE 3.11: Result after the use of Component Labeling algorithm. Each region is colored with a different shade of blue, that can't be distinguished with the eye.

3.3 Ontology Construction

After acquiring all the different images of the pollen grains, a structure was needed to store the images, the content of the descriptors and the information of the different species, as well as the relations connecting each species with its semantic information. We chose to construct a pollen ontology considering the ease of use, the expansion potentiality, the retrieval efficiency, the ability to develop an annotation system and the capability of recognising and using semantic relations between the species and the information that describe them (i.e. size, shape, polarity, etc.). For example, comparing the polarity of *Salvia officinalis*, which is isopolar and *Echium vulgare*, which is anisopolar, the ontology helps the computer understand that those two attributes are disjoint. Using ontologies as a basis for defining visual vocabulary or as a framework for automatic image annotation increases the number of concepts an image annotation system can recognize and may as a means for improving the performance of image retrievals.

3.3.1 Pollen Ontology

The backbone of the pollen ontology consists of a simple IS-A hierarchy of plants. Our ontology is a specialization of a certain division of plants classification, phanerogams, also known as spermatophytes. Phanerogams comprise those plants that produce seeds, they are a subset of the land plants (embryophytes) and they are divided into two classes:

angiosperms and gymnosperms. Angiosperms are the flowering plants that produce seeds and pollen for their reproduction.

In our ontology, 40 different species of the angiosperms class are included. They belong to two subclasses (monocotyledonae and dycotyledonae), 19 different families and 35 different genera.

With this approach, a specie belonging to a class in the pollen ontology will inherit properties belonging to upper classes.

The pollen ontology consists of two main class hierarchies. The pollen hierarchy, which is the basic class categorization of pollen and the pollen traits hierarchy, which is the class categorization of the high and low level pollen traits (descriptors). These two hierarchies are connected through relations between the instances of corresponding classes.

Pollen Hierarchy

The pollen class hierarchy was based on the categorization of each pollen specie used in this work, provided by the MPCU of MAICH. The 40 leaf classes in this hierarchy represent the different pollen species. For each species, an individual - belonging to the corresponding class, was created, containing the name of the specie as a data property (hasName) and a number of object properties (hasLowLevelInstance), connecting the individual with each one of grain images assigned to the specific specie. Finally, an object property (hasHighLevel) was assigned to each individual for the future purpose of holding information that characterize each species. All relations will be discussed later in detail. A part of the pollen hierarchy can be seen in figure 3.12.

Pollen Traits Hierarchy

Pollen traits class hierarchy, is an hierarchy of pollen traits that can characterize each specie and each grain image. Pollen traits hierarchy is divided into two main components. Low level and high level (fig. 3.12).

Low level traits are all the corresponding low level features used in image content analysis to recognize an image of its content. In this ontology, the low level traits are used for the training of the system. Each grain image is represented by an individual (LowLevelInstance#), which belongs to the LowLevel type of the hierarchy. Individuals contain, as data properties, each descriptor presented in image content analysis, as well as the path where the image is saved in the disk. Each individual is connected with the corresponding specie through the object property hasLowLevelInstance.

High level traits are a proposal of characteristics that can describe, to a large extent, a pollen species, such as size, polarity, symmetry, shape, number of apertures and other

palynological features. High level traits could be very useful in an annotation process of the pollen species. The HighLevel class, in our ontology, is connected with the corresponding specie through the object property hasHighLevel.

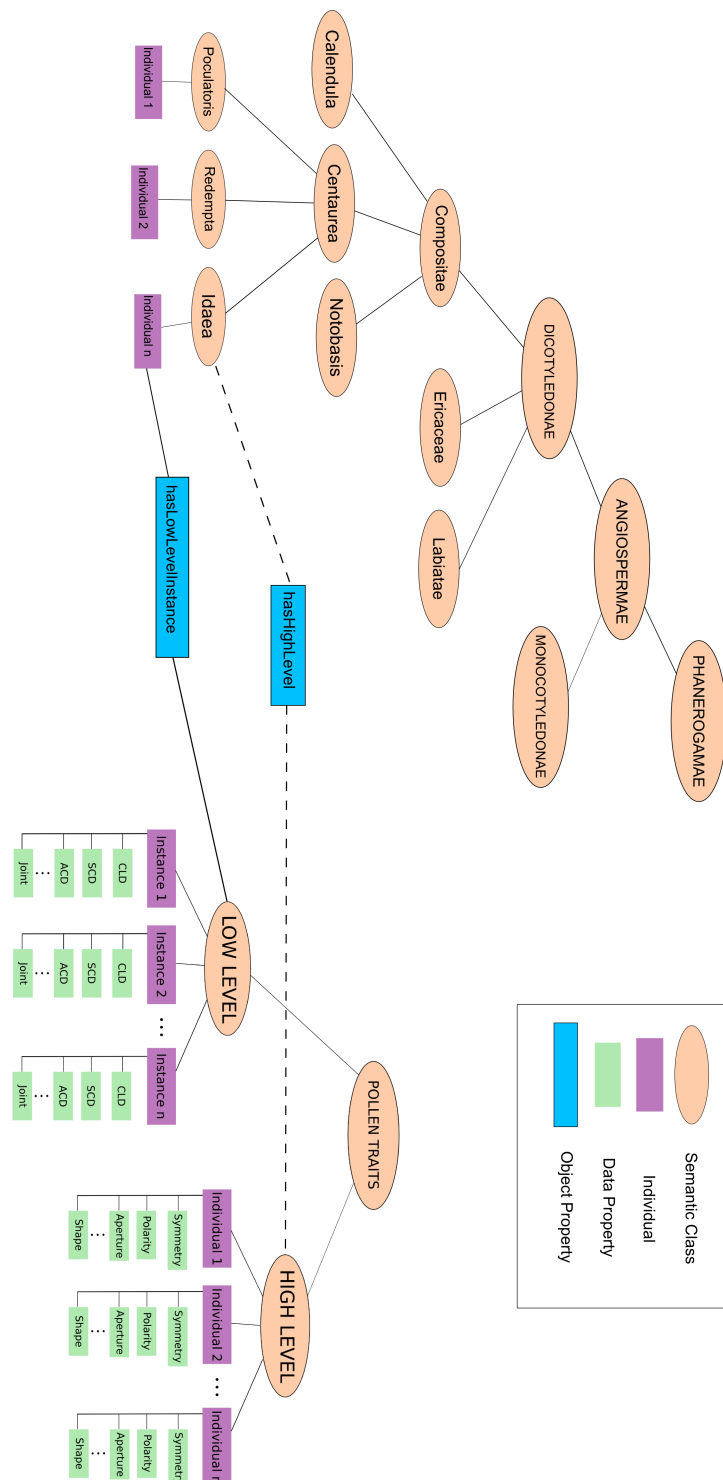


FIGURE 3.12: Pollen Ontology. Note that not all classes, or individuals are shown.

3.4 Image Similarity

In order to estimate the similarity between a pair of images LIRE similarity measure was used. LIRE similarity measure is the summary of the normalized (in range 0-1) low-level descriptors.

3.4.1 Feature Extraction

Our system starts with an unknown image of a single pollen grain as input. Afterwards, the image features (image content) are extracted. Image features are represented as visual descriptors, which are number vectors. For example, Color Layout Descriptor (CLD) has a vector representation of this form: [50 13 22 25 12 27z22,...] while Scalable Color Descriptor (SCD) has this form: [scalablecolor; 0; 64; -173 48 -3 57 4 9 22 30 -24...]. With this vector representation, visual descriptors of a pair of images cannot be directly compared i.e., we cannot compare CLD1 descriptor of image1 with CLD2 descriptor of image2. Instead, the distance of these descriptors can be used. The distance between two descriptors of a pair of images is a number from 0 to +infinity (Notice that comparison can be performed only between visual descriptors of the same type, i.e. CLD1 with CLD2 ,SCD1 with SCD2 etc). For example, the distance between a pair of CLD descriptors can be a number between 0-470 while SCD can be a number between 0-550. The lower the distance ,the higher the similarity of the pair of descriptors compared. When the distance zero, the visual descriptors compared are the same.

3.4.2 Gaussian Normalization

We compute the similarity of a pair of images as a function of feature distances extracted from the compared images. However, the different range of distances each pair of descriptors have (i.e., range of CLD distance is 0-470 but range of SCD distance is 0-550 etc.) yield the problem of not giving equal emphasis to each descriptor. We can use normalization techniques to solve this problem.

In order to achieve normalization, we must have a considerable number of distances for each descriptor. In our work, we use nearly 5000 distances for each descriptor based on the comparison between 501 database images of 40 different pollen species.

The most simple normalization technique is the L1 metric: for each element i in a ranked list of k elements having distance d , the D normalized distance is:

$$D = \frac{d_i}{d_{max}}$$

where d_i is the i_{th} distance and d_{max} is the maximum distance.

A second popular normalization metric, is linear scaling to unit range, L2 metric:

$$D = \frac{d_i - d_{min}}{d_{max} - d_{min} + 1}$$

where d_{max} and d_{min} are the maximum and minimum values of d .

The above normalization techniques are not useful when a limited database of images is available. This is because the distances of each descriptor usually fall into a small subrange of the entire possible range. As a result, the linear normalization will possibly compact the distances into a very narrow and indiscriminate range within $[0, 1]$ and distances of descriptors from unknown images (not in our database) will be mapped to different subranges within $[0, 1]$ which makes these distances incomparable with the ones of our database.

Gaussian normalization can solve this problem as it is capable of normalizing distances in a range within $[0, 1]$ that follow a gaussian distribution with standard deviation equal to 1. The Gaussian normalization is defined:

$$D = \frac{1}{2} \left(1 + \frac{d_i - \mu}{3 * \sigma} \right)$$

where μ is the mean value and σ is the standard deviation of the distances computed from our database images.

Thus, the presence of a few abnormally large or small values does not bias the importance of a feature measurement in computing the similarity between two images. In our work, we use the Gaussian normalization to normalize distances in the range within $[0, 1]$.

3.4.3 Image Similarity Measures

Given an unknown image, the database images of the ontology are searched using a similarity measure and the most similar to the unknown image are retrieved. In this work, LIRE similarity measure was used.

LIRE Similarity Measure

LIRE similarity measure is a function of LIRE feature distances computed from the compared images. LIRE similarity measure is defined as:

$$S_{LIRE}(A, B) = \sum_i (1 - d_i)$$

where A, B is the pair of images and d_i is the Gaussian normalized distance of the i_{th} descriptor. Note that we used $1 - d_i$ and not d_i distance. This was done because the distances range within $[0, 1]$ was needed and not in $[1, 0]$, as similarity - not distance - was computed. In other words, the similarity increased when the distance d_i decreased.

LIRE similarity measure can compare pairs of images good enough especially when these images differ a lot. For example, if we had to compare an image containing a building with an image containing a pollen grain, LIRE similarity measure would give a really low similarity denoting that these images do not match. However, in most cases, images compared are identical and therefore cannot be compared by using only LIRE similarity measure. An overall measure that could give relative importance in each descriptor was needed.

3.5 Pollen Species Estimation

Our system uses five methods to estimate the species of an unknown pollen grain image. These methods use the retrieval results list which is a ranked list containing the database images of the pollen ontology retrieved in a decreasing similarity order.

3.5.1 Best Match

Best match selects the species of the most similar image to the unknown image i.e. the class belonging to the retrieved database image with the highest similarity score.

3.5.2 Max Occurrence

Max Occurrence selects the species that has the maximum number of instances in the first 15 answers. If multiple classes have the same number of instances, best match between competing classes is employed.

3.5.3 AVR

For a query q with a ground-truth size of $NG(q)$, we define $rank(k)$ as the rank of the k_{th} ground-truth image on the top- N result list. The average retrieval rank is then computed as follows:

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{rank(k)}{NG(q)}$$

AVR is used as the standard metric of specie estimation of an unknown image (correctly classified image), based on the retrieval results. A ranked list of instances is obtained answering the image query. The result list includes all the classes of our ontology (species). For each specie class AVR is computed. From all AVR computed, the class having the best AVR (the least) would be the class the query image belongs to. If multiple classes have the same AVR, best match is employed between competing classes.

The algorithm of semantic category estimation, using AVR, is shown below:

$rankspeciesimage(q_{i1})$ = best first rank in ranked list for picture from species i for the query q.

...

$rankspeciesimage(q_{in})$ = n best rank in ranked list for picture from species i for the query q.

query q ;

List $list$ = similar retrieved images in decreasing order of similarity.

$list[0]$ = best image

for $eachspecies == i$ do

$AVR[i] = (rankspeciesimage(q_{i1}) + ... + rankspeciesimage(q_{in}))/n$

end for

$bestavr = bubblesort(AVR, species)$

$bestavr[0] = bestspecies$

3.5.4 Summation of Instances

Summation of instances selects the pollen species belonging to the instances that achieve the highest summation of similarities in the retrieved list.

Chapter 4

Experiments

In this chapter the experiments are presented. The evaluation of our system was done using 40 query single pollen grain images. Each query image is compared with all the images in our database, according to its visual content. In the next sections the retrieved results for each query, using LIRE similarity measure, are presented and discussed.

The basic IR evaluation tools used were precision and recall:

$$precision = \frac{|\{Relevant\ Images\} \cap \{Retrieved\ Images\}|}{|\{Retrieved\ Images\}|}$$

$$recall = \frac{|\{Relevant\ Images\} \cap \{Retrieved\ Images\}|}{|\{Relevant\ Images\}|}$$

Precision and recall helped measure how well our system performed. In our work we used average retrieval-recall as a quality measure:

$$P(r) = \sum_{i=1}^{N_q} \frac{P_i(r)}{N_q}$$

where N_q number of queries and $P_i(r)$ is precision at recall level r for i_{th} query.

4.1 Database Specifics

The database consists of 501 pollen grain images of 40 different pollen species, belonging to 35 different genus and 19 families. Different families and species were selected in order to have a variety of pollen shape, size and ornamentation. An important factor for the

selection of the species, that were included in the database, was the quality of the acquired images during the sampling process (number of unique grains, quality of the tiles), as well as the results of the segmentation process. The selected species had an adequate number of unique pollen grains that could be extracted using the segmentation techniques noted before. Species whose images contained only contiguous grains could not be included in our database.

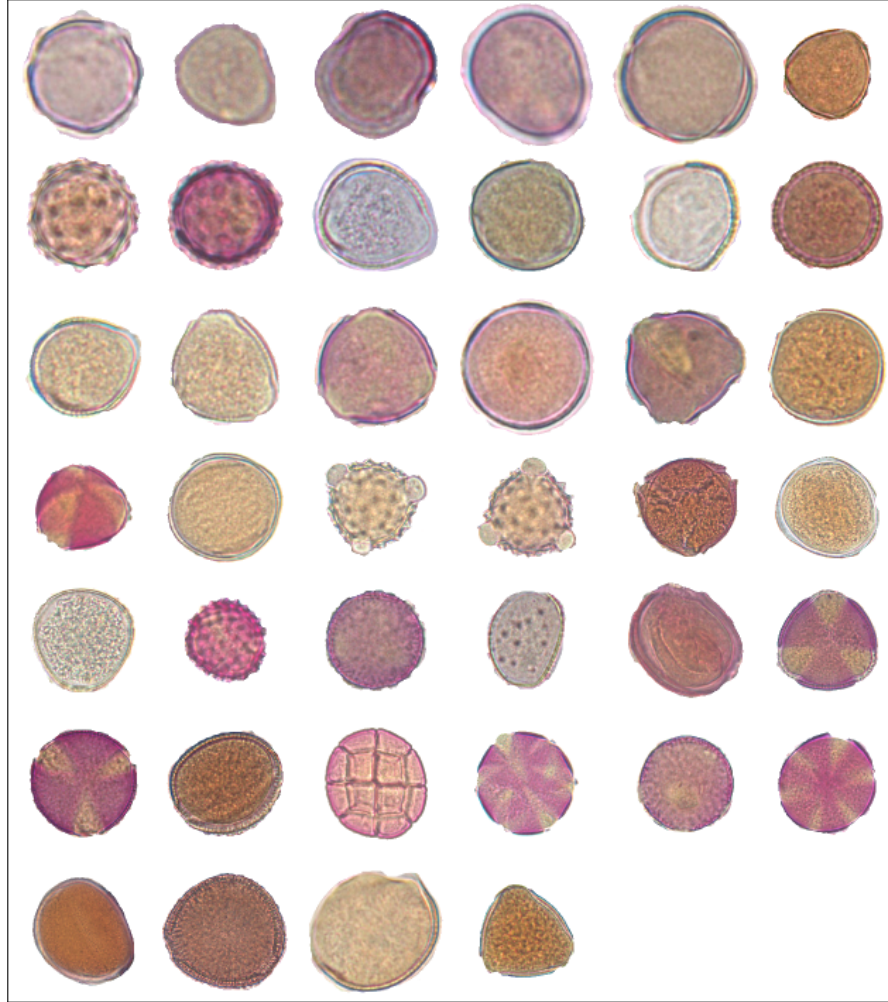


FIGURE 4.1: The 40 query images used for evaluation.

4.2 LIRE Similarity Measure Results

We used 40 queries (fig. 4.1) and computed the precision recall for the first 40 answers. For each query tested, all the instances of the species the query actually belongs to, are noted as the relevant images.

The precision of the query set was tested requesting the species, as well as the genus of the pollen grain. The two diagrams are presented in figures 4.2 and 4.3. It can be

noted that the diagram when the genus of the pollen grains, and not the species, are requested, is slightly corrected as expected. Specifically, the retrieved results for the first few answers seem to match successfully the species and the genus of the query image.

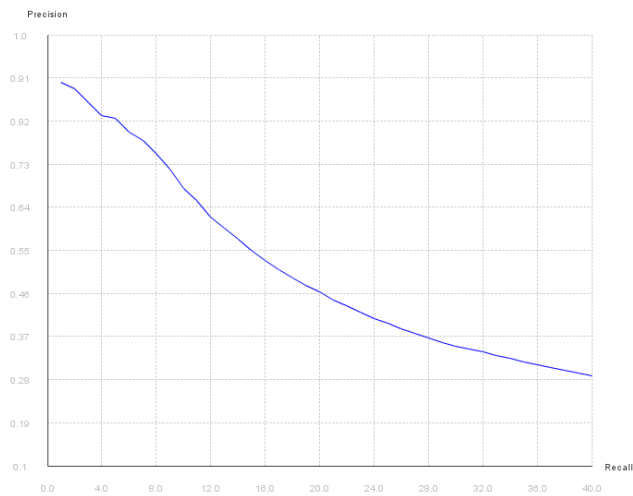


FIGURE 4.2: Precision-Recall Diagram, based on species, for the 40 first answers using LIRE similarity measure.

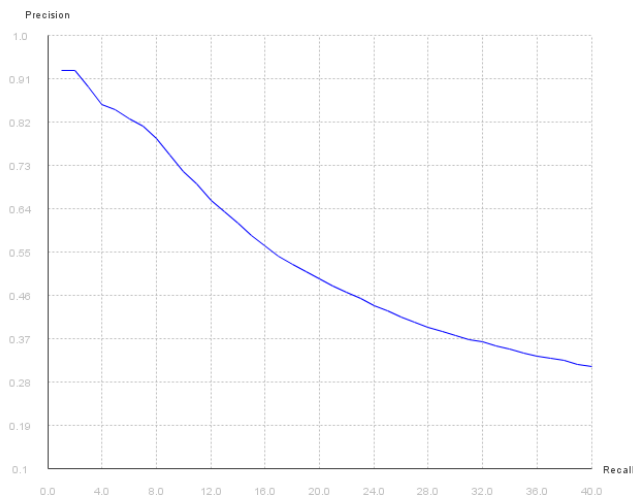


FIGURE 4.3: Precision-Recall Diagram, based on genus, for the 40 first answers using LIRE similarity measure.

Below, five example queries and their corresponding result list are presented. Note that LIRE similarity measure is normalized in range between $[0, 1]$.

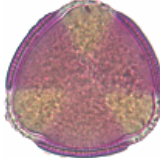
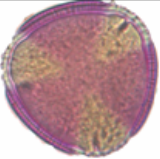
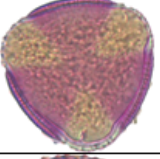
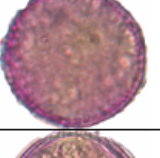
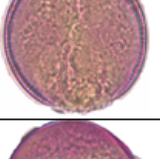
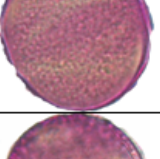
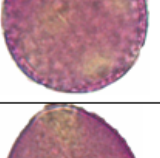
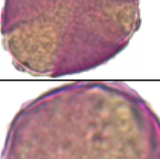
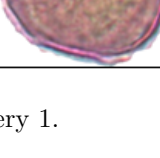

QUERY 1 Convolvulus Cneorum 		Species Genus LIRE Similarity Measure Image	
Cneorum	Convolvulus	0.9391717581289709	
Cneorum	Convolvulus	0.9378824154143918	
Juniperinus	Dianthus	0.9300733272337747	
Cneorum	Convolvulus	0.9284623557243472	
Thymbra	Satureja	0.9261318165371578	
Xylorrhizus	Dianthus	0.9250306215847577	
pes-caprae	Oxalis	0.9231372755449738	
Bellardia	Trixago	0.9229492591081391	

FIGURE 4.4: Retrieved list using LIRE similarity measure for Query 1.

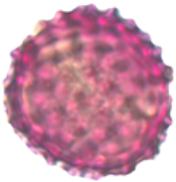
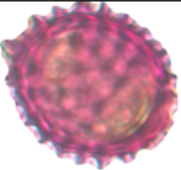
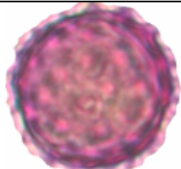
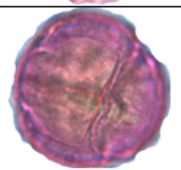
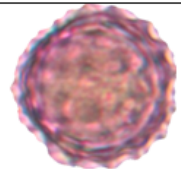
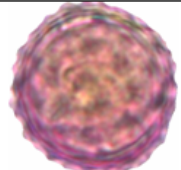
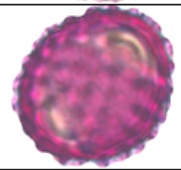
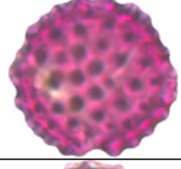
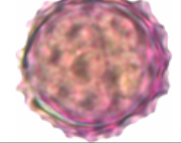
QUERY 2 Calendula Arvensis 	Species	Genus	LIRE Similarity Measure	Image
	Arvensis	Calendula	0.9337220804501069	
	cf.maritimum	Senecio	0.9330804995894122	
	Unedo	Arbutus	0.9325021402483983	
	cf.maritimum	Senecio	0.9314073181508008	
	cf.maritimum	Senecio	0.9312885118012539	
	Arvensis	Calendula	0.930717860645484	
	Arvensis	Calendula	0.9306509493172417	
	cf.maritimum	Senecio	0.9303443522115431	

FIGURE 4.5: Retrieved list using LIRE similarity measure for Query 2.

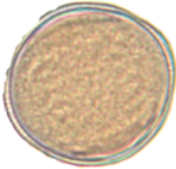

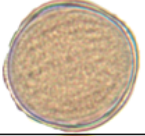

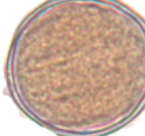

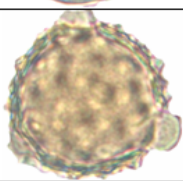

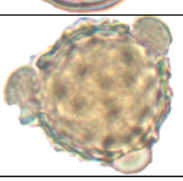
<div><div>QUERY 3</div><div>Borago Officinalis</div><div></div></div>	Species	Genus	LIRE Similarity Measure	Image
	Officinalis	Borago	0.9395553263880254	
	Officinalis	Borago	0.93843470404078	
	Officinalis	Borago	0.9369513780741219	
	Officinalis	Borago	0.9274210027127164	
	Officinalis	Borago	0.9256471551048636	
	Glauc	Nicotiana	0.924895677051597	
	Officinalis	Borago	0.9244924689472098	
	Officinalis	Borago	0.9235454809719519	

FIGURE 4.6: Retrieved list using LIRE similarity measure for Query 3.


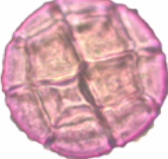
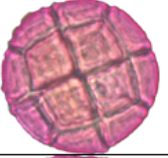
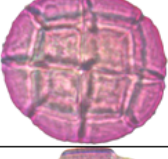
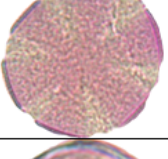
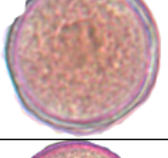
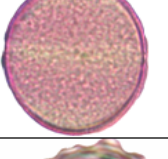
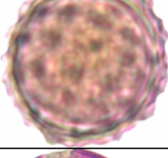
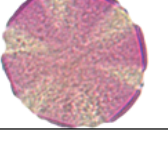
QUERY 4 Acacia Cyanophylla 	Species	Genus	LIRE Similarity Measure	Image
	Cyanophylla	Acacia	0.9391717581289709	
	Cyanophylla	Acacia	0.9378824154143918	
	Cyanophylla	Acacia	0.9300733272337747	
	Thymbra	Satureja	0.9284623557243472	
	Asiaticus	Ranunculus	0.9261318165371578	
	Thymbra	Satureja	0.9250306215847577	
	cf.maritimum	Senecio	0.9231372755449738	
	Thymbra	Satureja	0.9229492591081391	

FIGURE 4.7: Retrieved list using LIRE similarity measure for Query 4.






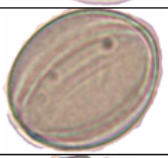
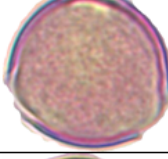


<p>QUERY 5</p> <p>Scilla autumnalis</p> 	Species	Genus	LIRE Similarity Measure	Image
	Autumnalis	Scilla	0.9330269139465427	
	Autumnalis	Scilla	0.92793013794107	
	Pseudodictamnus	Ballota	0.9258710843768715	
	Pseudodictamnus	Ballota	0.9255284649022555	
	Pseudodictamnus	Ballota	0.9248953595814621	
	Glaucal	Nicotiana	0.9241949354646024	
	Pseudodictamnus	Ballota	0.9239140556116929	
	Officinalis	Borago	0.9230310216735357	

FIGURE 4.8: Retrieved list using LIRE similarity measure for Query 5.

4.3 Comparison Of Category Estimation Methods

The species of the images are estimated based on further analysis of the retrieval results. For each unknown image, a ranked list of images is produced in decreasing order of similarity. As previously mentioned, four methods have been tested for the species estimation based on retrieval results: Best match, AVR, max occurrence and summation of instances. In table 4.9 the results for each method are presented. Success rates were calculated for both the results based on species and those based on genus. These results were taken using 40 queries.

Category Estimation Based on Species

Category Estimation Method	1 st Answer	2 nd Answer	3 rd Answer	Overall
<i>Best Match</i>	85	7.5	0	92.5
<i>AVR</i>	77.5	10	5	92.5
<i>Max Occurrence</i>	67.5	17.5	2.5	87.5
<i>Sum of Instances</i>	67.5	10	5	82.5

Category Estimation Based on Genus

Category Estimation Method	1 st Answer	2 nd Answer	3 rd Answer	Overall
<i>Best Match</i>	87.5	7.5	0	95
<i>AVR</i>	80	7.5	5	92.5
<i>Max Occurrence</i>	70	15	2.5	87.5
<i>Sum of Instances</i>	72.5	7.5	5	85

FIGURE 4.9: Percentages of the first three answers for every category estimation method.

With LIRE similarity measure the species estimation methods performed as follows:

Best Match 85% of the queries belong to the same species as the first ranked image, 7.5% of the queries belong to the same species as the second best image in the ranked list, and 0% of queries belong to the same species with the third best image in the ranked list. There is a slight improvement in the percentages of the first answer (87.5%), when the genus is requested. If the query image is not identified correctly at the first two results it is unlikely that it would be identified of later answers, with this method.

AVR 77.5% of the queries belong to the species with the best AVR, 10% of the queries belong to the species with the second best AVR, and 5% of the queries belong to the species with the third best AVR. When the genus is requested a slight "relocation" of a 2.5% from the second to the first answer can be noted.

Max occurrence 67.5% of the queries belong to the species that has the maximum number of instances in first 15 answers, 17.5% of the queries belong to the species that has the second maximum number of instances in first 15 answers, and 2.5% of the queries belong to the class that has the third maximum number of instances in first 15 answers. With this method too, when the genus is requested, there is a slight "relocation" of a 2.5% from the second to the first answer.

Summation of instances 67.5% of the queries belong to the species with the highest summation, 10% of the queries belong to the species with the second highest summation, and 5% of the queries belong to the class with the third highest summation. There is a slight improvement in the percentages of the first answer (72.5%), when the genus is requested.

Best match and AVR proved to be the most reliable methods, while summation of instances method is the most unreliable, having though high success rates. Max occurrence is slightly worse than the two best. A factor that influenced the results of all the methods, but mainly max occurrence and summation of instances, was the different number of images for each species, that were included in the database. Some species that had many images, when ranked high in the result list, would push in lower ranking images of the same species with the query image, because of a slight difference in the result. Another notable issue was the differences that exist between the images of the same species. Images that capture the two poles of a grain, or different focal points, can produce distant scores in the retrieved list. That was the reason, that best match method could not identify images after the first two answers.

Below, the results of the five example queries, for each category estimation method, are presented.

Query 1: *Convolvulus cneorum*

Species	Genus	AVR Score	Image
Juniperinus	Dianthus	0.9269470297018704	
Cneorum	Convolvulus	0.9264235336046436	
Thymbra	Satureja	0.9222347486790784	
pes-caprae	Oxalis	0.9212357138898704	
Xylorrhizus	Dianthus	0.9195872771190764	

Species	Genus	Max Occurrence (#)	Image
Bellardia	Trixago	5	
Cneorum	Convolvulus	4	
Thymbra	Satureja	1	
Xylorrhizus	Dianthus	1	
pes-caprae	Oxalis	1	

Species	Genus	Summation of Instances Score	Image
Cneorum	Convolvulus	0.921617706204185	
Xylorrhizus	Dianthus	0.911417428945228	
pes-caprae	Oxalis	0.9094373649708707	
Thymbra	Satureja	0.9074232474580792	
Cretica	Tulipa	0.9073235499648867	

Query 2: *Calendula arvensis*

Species	Genus	AVR Score	Image
Unedo	Arbutus	0.9285791747055754	
Arvensis	Calendula	0.9264758822271975	
cf.maritimum	Senecio	0.9261218038103173	
cyanophylla	Acacia	0.92492016560455	
Cretica	Tulipa	0.9233765778365993	

Species	Genus	Max Occurrence (#)	Image
cf.maritimum	Senecio	10	
Arvensis	Calendula	3	
Unedo	Arbutus	1	
Juniperinus	Dianthus	1	

Species	Genus	Summation of Instances Score	Image
cf.maritimum	Senecio	0.9235883974383702	
Arvensis	Calendula	0.9223780804555368	
xylorrhizus	Dianthus	0.9138088258187123	
Juniperinus	Dianthus	0.9136943689451007	
thymbra	Satureja	0.9119688428053896	

Query 3: *Borago officinalis*

Species	Genus	AVR Score	Image
Officinalis	Borago	0.9334516176534482	
Autumnalis	Scilla	0.922373801517239	
Tomentosa	Galactites	0.9215432532419976	
Repens	Trifolium	0.920288232026004	
Italicum	Echium	0.9186313025886769	

Species	Genus	Max Occurrence (#)	Image
Officinalis	Borago	6	
Tomentosa	Galactites	6	
pseudodictamnus	Ballota	2	
Autumnalis	Scilla	1	

Species	Genus	Summation of Instances Score	Image
Officinalis	Borago	0.9201871365797621	
Tomentosa	Galactites	0.9160547496621342	
Pseudodictamnus	Ballota	0.9127962821364406	
Asiaticus	Ranunculus	0.9111296956121754	
Glauca	Nicotiana	0.9093710666675054	

FIGURE 4.10: Results of example queries (1)

Query 4: *Acacia cyanophylla*

Species	Genus	AVR Score	Image
Cyanophylla	Acacia	0.9134385321435736	
Thymbra	Satureja	0.9122143603293225	
cf.maritimum	Senecio	0.9102276146295948	
autumnalis	Scilla	0.9100758100132053	
Tomentosa	Galactites	0.9096470915826191	

Species	Genus	Max Occurence (#)	Image
Thymbra	Satureja	4	
Cyanophylla	Acacia	3	
cf.maritimum	Senecio	3	
Asiaticus	Ranunculus	1	
autumnalis	Scilla	1	

Species	Genus	Summation of Instances Score	Image
Cyanophylla	Acacia	0.9101323309419151	
Thymbra	Satureja	0.9064605062659647	
Asiaticus	Ranunculus	0.9053292709708586	
Tomentosa	Galactites	0.9041725501945893	
autumnalis	Scilla	0.9014514363266221	

Query 5: *Scilla autumnalis*

Species	Genus	AVR Score	Image
Autumnalis	Scilla	0.927788848875288	
Pseudodictamnus	Ballota	0.9234725386426144	
Pseudodictamnus	Ballota	0.9231717264900163	
Glaucia	Nicotiana	0.9196146591476481	
arcturus	Verbascum	0.9194971361222927	

Species	Genus	Max Occurence (#)	Image
Pseudodictamnus	Ballota	6	
Autumnalis	Scilla	3	
Glaucia	Nicotiana	2	
Officinalis	Borago	2	
arcturus	Verbascum	2	

Species	Genus	Summation of Instances Score	Image
Pseudodictamnus	Ballota	0.9186557020855527	
Glaucia	Nicotiana	0.9157273814558605	
Arcturus	Verbascum	0.913341005097909	
Autumnalis	Scilla	0.9132054829243492	
Tomentosa	Galactites	0.91106538442008	

FIGURE 4.11: Results of example queries (2)

Chapter 5

Conclusion

The system presented in this work is capable of estimating the species of pollen images, using image content and a pollen ontology. It can be part of a fully automated system of pollen identification, where pollen images, captured from pollen traps or honey samples, would be presented and identified to species. An important part of our work is the information gained by the process of capturing the pollen images from a light microscope and applying several image segmentation techniques in order to obtain single pollen grain images. Image background can be subtracted successfully using thresholding on the HSI color model, while pollen grains can be discriminated from the image using morphological operators and component labeling.

The process of estimating the species of an unknown image is implemented in steps. A query image is provided by the user to obtain similar images from the ontology. Image matching is implemented using image content descriptions. 14 visual descriptors are used. The LIRE similarity measure between images is proposed as similarity measure. The species of an unknown image is computed based on AVR (Average Retrieval Rank), summation of instances, max occurrence and best match.

Our experiments show that automated image recognition using ontologies and image content analysis can be successful in percentages near 75% for the first answer, given a good and variant training set for the classification stage and a descriptive ontology.

5.1 Future Work

Possible future work could expand our system with the addition of an annotation system that would characterize thoroughly each pollen grain. The High Level class, the object and data properties that are already part of our ontology could accept information

about the size, polarity, symmetry, shape, apertures and other palynological features of each species. These features could be cross-checked and enriched with information from palynological sites such as [\[42\]](#).

In a machine vision view, the solution of the contiguous pollen grains issue would help include more species in the classification process. This might be achieved with the use of edge detection and edge following techniques on the images, after the background subtraction. Another domain of palynological interest, for future work, is counting the pollen grains in images with several different species. The measure of the content rate of each species would help certificate the floral origin of honeys, or forecast which allergen pollen is more dangerous each day.

Bibliography

- [1] University of Washington. Angiosperms: The flowering plants. URL <http://courses.washington.edu/bot113/spring/WebReadings/PdfReadings/ANGIOSPERMS.pdf>.
- [2] Scott Gilbert. *Developmental biology*. Sinauer Associates, Sunderland, Mass, 2000. ISBN 0-87893-243-7.
- [3] Smith SME McDaniel CN, Singer SR. Developmental states associated with the floral transition. pages 153:59–69, 1992. URL <http://www.ncbi.nlm.nih.gov/pubmed/1516752>.
- [4] Dean C. Levy Y Y. The transition to flowering. 1998. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC526001/>.
- [5] Giancarlo Albore. *Textbook of melissopalynology*. Apimondia Pub. House, Bucharest, 1997. ISBN 973-605-023-8.
- [6] G Erdtman. *Pollen morphology and plant taxonomy*. E.J. Brill, Leiden, 1952. ISBN 9004081224.
- [7] G.O.W Kremp. *Morphologic encyclopedia of palynology*. Tuscon, University of Arizona Press, 1965.
- [8] Iversen J. Faegri K. *Textbook of Pollen Analysis*. Blackwell Scientific Publs, 1964.
- [9] Groupe de travail morphologie sporo-pollinique et numericulture Association des palynologues de langue francaise. C Caratini, C Tissot. *Structure et terminologie de la paroi sporo-pollinique: Documents pour la reunion du Groupe de Travail Morphologie sporopollinique et numriculture*. Association des palynologues de langue francaise, 1975.
- [10] Papamarkos. *Digital image processing and analysis*. Gkiourdas, 2005. ISBN 960-387-352-7.
- [11] Ramesh Jain. *Machine vision*. McGraw-Hill, New York, 1995. ISBN 0-07-032018-7.

- [12] A Threshold Selection Method from Gray-Level Histograms. *Systems, Man and Cybernetics, IEEE Transactions*, 9:62–66, 1979. ISSN 0018-9472. URL <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4310076>.
- [13] Thomas C. Jepsen. Just what is an ontology, anyway? *IT Professional*, 11(5):22–27, 2009. URL <http://doi.ieeeecomputersociety.org/10.1109/MITP.2009.105>.
- [14] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition - Special issue: Current issues in knowledge modeling*, 5(2): 199–200, 1993.
- [15] Liang-Tien Chia Huan Wang, Song Liu. Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. *MULTIMEDIA '06 Proceedings of the 14th annual ACM international conference on Multimedia*, pages 109–112, 2006. URL <http://dl.acm.org/citation.cfm?doid=1180639.1180672>.
- [16] Marsland Stephen Arnold Greg Flemmer Rory Flenley John Fountain David Allen Gary, Hodgson Bob. Automatic recognition of light-microscope pollen images. 2006. URL <http://mro.massey.ac.nz/handle/10179/155>.
- [17] Carlos M. Travieso Jorge Arroyo-Hernández Santiago T. Pérez Jesús B. Alonso Federico Mora-Mora Jaime Roberto Ticay-Rivas, Marcos del Pozo-Baños. Pollen classification based on geometrical, descriptors and colour features using decorrelation stretching method. In *Engineering Applications of Neural Networks / Artificial Intelligence Applications and Innovations (EANN/AIAI (2))*, pages 342–349, 2011.
- [18] Juan Carlos Briceño Carlos M. Travieso Norma Monzón García, Víctor Alfonso Elizondo Chaves. Pollen grains contour analysis on verification approach. In *HAIS (1)*, pages 521–532, 2012.
- [19] Monique Thonnat Jordina Belmonte Carmen Galan Pierre Bonton-Rgis Tomczak Alain Boucher, Pablo J. Hidalgo. Development of a semi-automatic system for pollen recognition. *Aerobiologia*, 29(3-4):195–201, 2002. ISSN 0393-5965. URL <http://link.springer.com/article/10.1023%2FA%3A1021322813565>.
- [20] H.F.Lamb G.A.T. Duller I. France, A.W.G. Duller. A comparative study of approaches to automatic pollen identification. In *British Machine Vision Conference*, 1997.
- [21] Euripides G.M. Petrakis Antonios Dimas, Pyrros Koletsis. Dogi : An annotation system for images of dog breeds. In *International Conference on Image Analysis and Recognition (ICIAR)*, pages 410–417, 2012.

- [22] Vinay Modi. Color descriptors from compressed images. signal processing.
- [23] Franca Debole Fabrizio Falchi Claudio Gennaro Fausto Rabitti Giuseppe Amato, Paolo Bolettieri. Using mpeg-7 for automatic annotation of audiovisual content in elearning digital libraries. In *Italian Research Conference on Digital Libraries (IRCDL)*, pages 1–12, 2008.
- [24] Theofilos P. Mailis Eddie Cooke Yannis S. Avrithis-Noel E. O’Connor Evaggelos Spyrou, Hervé Le Borgne. Fusing mpeg-7 visual descriptors for image classification. In *Internet Corporation for Assigned Names and Numbers — ICANN (2)*, pages 847–852, 2005.
- [25] Athena Vakali, Mohand-Said Hacid, and Ahmed K. Elmagarmid. Mpeg-7 based description schemes for multi-level video content classification. *Image Vision Comput.*, 22(5):367–378, 2004.
- [26] Dr V. Kannan R. Balasubramani. Efficient use of mpeg-7 color layout and edge histogram descriptors in cbir systems. *Global Journal of Computer Science and Technology*, 9(4), 2009.
- [27] José María Martínez Sanchez, Rob Koenen, and Fernando Pereira. Mpeg-7: The generic multimedia content description standard, part 1. *IEEE MultiMedia*, 9(2): 78–87, 2002.
- [28] Dr. I.V. MURALI KRISHNA P.S.SUHASINI, Dr. K.SRI RAMA KRISHNA. Cbir using color histogram processing. *Journal of Theoretical and Applied Information Technology*, 2009.
- [29] Ju Han and Kai-Kuang Ma. Fuzzy color histogram and its use in color image retrieval. *IEEE Transactions on Image Processing*, 11(8):944–952, 2002.
- [30] Jing Huang, Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image indexing using color correlograms. In *Computer Vision and Pattern Recognition (CVPR)*, pages 762–768, 1997.
- [31] Danco Davcev Katarina Trojacanec, Georgina Mirceva. Application of edge histogram descriptor and region shape descriptor to mris.
- [32] S. Mori H. Tamura and T. Yamawaki. Texture features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-8(6):460–473, 1978.
- [33] Peng Wu, B. S. Manjunath, Shawn Newsam, and H. D. Shin. A texture descriptor for browsing and similarity retrieval. *Sig. Proc.: Image Comm.*, 16(1-2):33–43, 2000.

- [34] Haixia Jiang Xuemin Qian Zhi Li, Guizhong Liu. Image copy detection using a robust gabor texture descriptor. In *LS-MMRM '09 Proceedings of the First ACM workshop on Large-scale multimedia retrieval and mining*, pages 65–72, 2009.
- [35] Savvas A. Chatzichristofis and Yiannis S. Boutalis. Cedc: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In *International Colour Vision Society (ICVS)*, pages 312–322, 2008.
- [36] Savvas A. Chatzichristofis and Yiannis S. Boutalis. Fcth: Fuzzy color and texture histogram - a low level feature for accurate image retrieval. In *International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, pages 191–196, 2008.
- [37] Nikos Papamarkos Yiannis S. Boutalis Konstantinos Zagoris, Savvas A. Chatzichristofis. Automatic image annotation and retrieval using the joint composite descriptor.
- [38] Leica Microsystems Wetzlar Gmbh. Leica dm l instructions, 1998.
- [39] LOCI biophotonics research laboratory of University of Wisconsin-Madison. Differential interference contrast (nomarski) imaging. URL <http://loci.wisc.edu/optical-sectioning/differential-interference-contrast-nomarski-imaging>.
- [40] John Innes Centre. Nomarski (dic) microscopy. URL http://www.jic.ac.uk/microscopy/more/T5_5.htm.
- [41] JENOPTIK. Progres capture pro 2.1 operation manual.
- [42] Ricciardelli D'Albore M. Ricciardelli D'Albore G., Galarini R. Mediterranean melissopalynology. URL <http://www.izsum.it/Melissopalynology/>.