A Framework for Color-Texture Segmentation and Characterization of Natural Outdoor Images

A thesis submitted during 2015 to the University of Hyderabad in partial fulfillment of the award of a Ph.D. degree in School of Computer & Information Sciences

by

TILOTTAMA GOSWAMI

SCHOOL OF COMPUTER & INFORMATION SCIENCES

UNIVERSITY OF HYDERABAD

(P.O.) CENTRAL UNIVERSITY, GACHIBOWLI

HYDERABAD - 500 046,

INDIA

CERTIFICATE

This is to certify that the thesis entitled "**A Framework for Color-Texture Segmentation and Characterization of Natural Outdoor Images**" submitted by **Tilottama Goswami** bearing Regd. No. 10MCPC03 in partial fulfillment of the requirements for the award of **Doctor of Philosophy** in Computer Science is a bonafide work carried out by her under our supervision and guidance which is a plagiarism free thesis.

The Thesis has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Prof Arun Agarwal
Supervisor
School of Computer and
Information Sciences
University of Hyderabad

Prof C. R. Rao
Supervisor
School of Computer and
Information Sciences
University of Hyderabad

Dean
School of Computer and
Information Sciences
University of Hyderabad

DECLARATION

I, **Tilottama Goswami**, hereby declare that this Thesis entitled "**A Framework for Color-Texture Segmentation and Characterization of Natural Outdoor Images**" submitted by me under the guidance and supervision of **Professor Arun Agarwal and Professor C. R. Rao** is a bonafide research work which is also free from plagiarism. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma. I hereby agree that my thesis can be deposited in Shodganga/INFLIBNET.

A report on plagiarism statistics from the University Librarian is enclosed.

Date:

Signature of the Student

Name: **Tilottama Goswami**

Regd. No. **10MCPC03**

DEDICATION

'Start by doing what's necessary; then do what's possible and suddenly you are doing the impossible' - Francis of Assisi

I dedicate this thesis to my husband

Mr. Samir Goswami and my two lovely daughters

Prakriti and Sanskriti for their consistent support and encouragement.

ACKNOWLEDGEMENTS

I am thankful to God for giving me a beautiful lifetime opportunity - Ph.D. journey which is intellectually stimulating and at the same time emotionally exhausting. This journey has taught me a lot and enriched me with rich lifetime experiences. Someone has truly said,

> Success is a journey, not a destination. The doing is often more important than the outcome.

University of Hyderabad (UOH) has instilled this feeling in me. My supervisors Prof. Arun Agarwal and Prof. C. R. Rao, and Doctoral Research Committee members Prof. Atul Negi and Prof. Bapi Raju always encouraged me to attain high quality research work and guided me throughout the Ph.D. program. Prof. Arun Agarwal taught me to reinforce faith in myself while going through difficult times. Prof. Agarwal's domain expertise in image processing and computer vision helped me in expanding my knowledge in related areas of my research work. Prof. Rao introduced me to the world of statistics and its beauty in measuring the inherent image characteristics.

> Measure what is a measurable, and make measurable what is not so.. by Galileo Galilei.

I am thankful to the Dean, School of Computer & Information Sciences for providing me the required facilities during my work. I would like to thank office and technical staff of School of Computer & Information Sciences, Administration staff of University of Hyderabad for the support extended.

I am thankful to the Indian National Center for Ocean Information Services (INCOIS), Hyderabad for giving me the opportunity to work on Grid-computing and its architecture. I would like to thank UOH for granting me Basic Boarding and Lodging (BBL) scholarship and also, would like to acknowledge UGC for granting me Basic Scientific Research (UGC-BSR) fellowship during my Ph.D. program.

I would like to offer my special thanks to Ms. Usha Yadav for extending her help and taking care of my family, when I was fully engrossed with long hours of research work. I salute her professionalism mixed with such care and comfort, which I can never repay.

I am deeply indebted to my husband whose vision has made this journey of mine possible. I am indebted to my parents Late Sri T. K. Sen & Smt. Sabita Sen for implanting the seed for knowledge quest as a child. Later in my life, support and encouragement came from my parents Late Sri F. L. Goswami & Smt. Madhuri Goswami to pursue further in my quest. I seek their blessings upon successful completion of this journey. This thesis is dedicated to them.

A special gratitude I give to my research scholar friends, Sandhya, Akhter, Vijaya Lakshmi, Supriya, Vani, Lakshmi, Deepthi and Tarun. Heartfelt thanks to my lovely daughters Prakriti and Sanskriti for extending their support in every possible way and filling my world with immense joy and hope.

ABSTRACT

An abundance of various automated and semi-automated segmentation techniques can be found in literature, that cater to wide range of image analysis and understanding applications. To cater such a wide coverage of algorithms, appropriate taxonomy is required, which can cater to all the aspects of the segmentation. We have proposed Multi-Faceted Hierarchical Image Segmentation Taxonomy (MFHIST), where every facet is considered as a design task. The collection of six facets are presented in a hierarchical manner, which are namely - scope, requirement, control, feature, image representation and approach specifications. Universal algorithm for automating image segmentation without having any prior knowledge about the application is a big challenge and poses an open problem in the research area even today. Segmentation problem is subjective in nature, as the perceptual grouping plays a crucial part in human visual system, and varies from one person to another. The domain of natural image segmentation is found to be a natural choice for the vision researchers as it involves a substantial level of subjectivity. Recent trend shows the motivation of computer vision researchers in integration of the two most fundamental descriptors - color and texture to simulate the humans intuitive psychophysical approach, in order to find a more reliable segmentation. Segmentation of color-texture natural images has its own inherent complexity consisting of random variations in texture-tonal mix. We propose a computationally inexpensive generic hybrid color-texture feature integration methodology applied to natural color image segmentation. The unsupervised image segmentation is based on new Hybrid Color Space and Orthogonal Polynomial operators which reasonably segments image having a good mix of texture and tonal variations present in natural images. The algorithm uses the edge based texture extraction using a set of difference operators based on orthogonal polynomials on a set of hybrid color planes. We have proposed adaptive feature vector representation where number of features vary from image to image. Cluster

and region metadata is generated in the process. Berkeley Segmentation Dataset has been used for our experiments and quantitative evaluation.

For automatic texture characterization, the intrinsic property of the object surface has to be modeled or mathematically defined. Texture is a measure of variability aspect, which is sometimes very difficult to model due to its subjective and irregular nature. Natural images have multi-textural components which in most of the cases have non-regular and non-periodic variations in surfaces. We have proposed to capture the local variability in visual appearance in terms of orientation specific based on Analysis Of Variance, which is an inferential statistical technique. Progressive sampling is devised to achieve the appropriate sampling density. The collection of the features obtained from image samples is rich in capturing both luminance and chrominance information. The characterization is represented by string code representation based on ANOVA grading. The semantic meaning can be extracted from the string code using our proposed polling method. Experiments have also been conducted for Vistex natural textures, Prague mosaic textures. Based on the survey, we set to choose our path for unsupervised segmentation method for any image without any prior knowledge about its content. From the survey, it is evident qualitatively and quantitatively, a segmentation method cannot segment all images to a satisfactory level. The importance of appropriate segmentation and informative representation aid the image understanding, which inspired the present study and hence we have come up with a framework for providing meaningful representation of the given image objectively. Our proposed framework initially can segregate the input images into four categories - 1) Object-based image, 2) Heterogeneous color based texture image, 3) Homogeneous color based texture image, 4) Pure toned image. We have developed a conceptual framework as a part of low and mid-level image analysis component having two components:

1 Segmentation approach: Color-texture image segmentation and segment characterization

2 Block approach: Block generation and blocks characterization

In special cases of segmentation, where not a single distinct object can be extracted from an image because of its highly textured nature, the framework directly processes the texture characterization to know the basic orientation embedded in the image using the block approach. Thus, when any image is made input to the framework, it outputs metadata from segmentation approach(if applicable) and from block approach. More focused and detailed algorithms can then be applied for higher level analysis in future.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

RGB	Red Green Blue color space
HSV	Hue Saturation Value color space
CMYK	Cyan Magenta Yellow Key(black) color space
L*a*b*	L* for Luminance; a* and b* for chrominance
HVS	Human Visual System
PCA	Principal Component Analysis
HCS	Hybrid Color Space
OP	Orthogonal Polynomials
HVS	Human Visual System
ANOVA	Analysis of Variance
BSD	Berkeley Segmentation Dataset
PRI	Probability Rand Index
BDE	Border Displacement Error
VOI	Variation of Information
GCE	Global Consistency Error
AFVR	Adaptive Feature Vector Representation
K-S Test	Kolmogorov-Smirnov Test
GLCM	Gray Level Co-occureence Matrix
DooG	Difference Of Offset Gaussian filter
SOM	Self-Organizing Maps
MFHIST	Multi-Faceted Hierarchical Image Segmentation Taxonomy
NaN	Not a Number

CHAPTER 1

Introduction : Aspects of Human Visual System(HVS) important for Image Segmentation and Texture Characterization

Computer vision has been one of the most popular topics for research activity in the past as well as present decade. Computer vision has two words, namely *Computer* and *Vision*, i.e. the way computer performs multiple tasks as our Human Visual System (HVS) starting from -

1) Basic understanding of the camera captured input images and their inherent feature set similar to human's early vision stage to -
2) Enforcement of constraints to segregate the meaningful objects or segments present in the image on the basis of the existing or transformed features pertaining to mid-level vision and further processing to -
3) Drawing simple inferences from image pixel values and combining information to deduce a global analysis or characterization of the image that acts as foundation for classification and recognition, an integral part of high-level vision.

The computer vision is an inter-disciplinary subject which aims to transform data using models to decouple and transform the input data to capture maximum information with the help of mathematics, statistics, physics, biology and learning theory. Computer vision has a wide variety of applications both old and new. Some of the early existing applications are robotic navigation, industrial inspection, process and quality control, military purpose, medical diagnosis, remote sensing, image retrieval. Some of the new areas of applications are object recognition, building image mosaics, automatic surveillance and understanding human activity, neuroimaging. Most of these applications tackle with similar underlying problem: automatic segmentation and automatic characterization of the visual appearance of the objects irrespective of different application domains. In the present market full of competition, reliable and efficient way to tackle these challenges is a key point to achieve success and ensure high quality standards among companies dealing in computer vision related applications. Traditional approaches are subjective in nature as they use human expertise for analyzing the visual appearance in doing segmentation or characterization of a given image. This is quite tedious, difficult to cater to large volumes of tasks and may be prone to human error. The technology developments demand online or real time analysis, hence there is need to develop automated automated computer vision systems for better productivity and reliability. The techniques for deciphering and encoding the visual aspect of the objects employ two very important features - color and texture, even though other features like gray-scale, shape are also been used. Color-texture describes a material and distinguishes from other in a more intuitive and human perceptible manner.

For applications like scene classification, content based image retrieval, automated tagging, motion of a robot in a real world using computer vision system, understanding the images captured by the visual sensors is quite important. Objects in a colorful world have more chances to be correctly identified and recognized as compared to the objects in the gray world. In most of the computer vision tasks, there is no guarantee of the presence of a known object. Hence the gray world algorithms using a limited few dozens of gray shades fail to solve the problem of distinguishing or separating the objects present in the image. Color is a higher-order cue useful in finding relations in a scene content. Texture is an aggregate feature which gives a measure of variation present in an image. Thus representation of image content in terms of combination of low-level features such as color and texture add more information associated with the visual surface of the object present in the image and increase the effectiveness of automatic segmentation and characterization process. Foundation fundamentals for understanding the thesis can be appreciated by going through books on color image processing by Gonzalez and Woods

(2007), Trussell *et al.* (2005); on texture by researchers Tuceryan and Jain (1998), Rao and Lohse (1993), Zhang and Tan (2002) and on statistics by Fisher and Yates (1947), Kurz and Benteftifa (2006), Dallal (2012), Cohen and Lea (2004), Jain (1981). We are motivated by the multifaceted approach in formation of building blocks and the process of learning by humans in interpreting the visual information.

The chapter is organized as follows. Section 1.1 briefs about the aspects of HVS in terms of colorfulness captured by humans, the spatial resolving capacity to visualize the environment and the ability of unified approach for various image modalities. Section 1.2 discusses the role of color and texture as two fundamental descriptors for outdoor natural images. Section 1.3 briefly discusses the cognitive process to achieve perceptual grouping, which plays an important role in object segmentation and characterization. Section 1.4 mentions our objectives and contributions. Section 1.5 details about the thesis organization.

1.1 Human Visual System

This section makes an attempt to capture the various facets of Human Visual System (HVS) that can be considered pivotal for automation of color image segmentation and characterization, a crucial part of image processing, and an integral part of computer vision.

1.1.1 Significance of light in object perception by humans

We can see and perceive extremely broad range of intensities when HVS combines the two types of photo-receptors present in the retina - cones and the rods. The colorfulness captured by HVS is result of stimulating combination of achromatic and chromatic information. The activation of rods and cones depends on the level of illumination levels. There are three types of human vision - monochromatic (scotopic), photopic and mesopic visions as mentioned by Trémeau *et al.* (2008). The photo receptors (rods/cones) get proportionately functional depending on the type of vision. For low illumination level, rods are active. For normal daylight cones are active and for dark-surround or dim-surround, both rods and cones are active. The automated model for image segmentation should be able to perform an image analysis with contribution from both color and gray-scale features. One of the issues is that we're able to perceive very large range of intensities, but not at the same time. So we cannot see in very dark rooms and very light regions of the room at the same time. Our eyes adapt to these extreme cases and is able to discriminate between objects as different levels to certain extent after a certain period of time. Weber's Law helps in adjustment depending on the background light, large changes are required for dark background and small for lighter ones.

1.1.2 Significance of visual acuity in texture perception by humans

HVS has the ability to process visual information and construct a representation of the surrounding environment. Both eyes and brain work in tandem for capturing the region of interest (ROI) and continually assembles a higher resolution image in order to visualize the fine details. This spatial resolving capacity which enables the eye to see the fine details is termed as visual acuity as mentioned by Kalloniatis and Luu (2007). Consider the first row in Figure 1.1[1] where the text appearing in the van is considered to be ROI. Similarly in the second row of Figure 1.1, the ROI is the door present inside the room. The resolution is increased till the ROI under consideration is clearly seen. Unlike the HVS, the automated visual model considers a digital image as input similar to a single frame snapshot camera. Maximum information needs to be extracted from the static data obtained from the output of an image capturing device.

The visual attention process primarily focuses on the region of interest and selectively extracts the low-level cues such as intensity, color, texture, shape, size, orientation or their various combinations. Texture and shape are high level spatial features. Con-

[1] Courtesy: Photography by Abhra DasGupta

Figure 1.1: Visual Acuity : ROI for first row is Text, ROI for second row is Door

siderable research efforts have been put to identify collection of powerful set of feature spaces. The low level stimuli when spatially organized, guides the appearance of the image or part of the image. The spatial organization of the cues is local and considers neighborhood of a predefined size. HVS adjusts the amount of such information captured depending on the location (far or near scene) and purpose(fine or coarse). The automated model in similar lines localizes windows of appropriate size for either detailing to precision or capturing the abstraction and estimate the captured content depiction. The localization considers neighborhood of particular size which conceptualizes the spatial compactness aspect of the feature space. This bottom-up approach is considered to be very helpful for simulating the visual acuity tasks. There are three basic types of acuity tasks - target detection, target localization and target recognition as mentioned by Kalloniatis and Luu (2007).

1.1.3 Significance of a unified approach by humans for analysis of diverse objects

The word target used in context of visual acuity is very general in the perspective of human vision system. Though the visibility of a human being is limited (wavelength 700nm to 400nm), nevertheless with the advancement in technology, one can acquire knowledge from the various types of image capturing devices, also pertaining to non-

(a) (b) (c) (d)

Figure 1.2: Different image sources a) Photographic image b)Medical image c) Aerial Image d) Microscopic image

visible electromagnetic spectrum. Different modalities of an image captured by different cameras produce different information covering the electromagnetic spectrum - gamma, XRay, optical, infrared, radio waves, etc. Hence the target consists of multi-domain images as shown in Figure 1.2 ranging from a photographic image (as shown in benchmark dataset by Martin *et al.* (2001)), medical images[1], high resolution aerial images[2] to microscopic images[3]. The cognitive vision of HVS has an extraordinary capability of perceptually grouping and drawing inference from any domain specific images. This is possible because of the domain knowledge (prior information) obtained by learning capability of humans in various fields.

1.2 Color-texture as fundamental descriptors for natural images

Natural images exhibit both color and texture characteristics as shown in Figure 1.3, taken from benchmark dataset by Martin *et al.* (2001). The two most fundamental descriptors of region - color and texture simulate the human intuitive psychophysical approach in performing generic vision functionality of detection, localization, recognition and understanding. Color is crucial to many pattern recognition tasks as it adds additional information to the intensity gray shades. Especially when we consider na-

[1] https://www.healthtap.com/

[2] https://earth.esa.int/ers

[3] Source: U.S. Centers for Disease Control and Prevention (CDC). "HIV and Tuberculosis." Atlanta, Georgia; March 19, 2013

Figure 1.3: Natural images taken from Berkeley dataset

ture which is an abode of unimaginable variations in color, our human visual system does justice in recognizing the thousands of color shades present as compared to our ability to perceive only two-dozen gray shades. Texture is an aggregate feature unlike color which is a pixel feature. Hence the texture feature analyzes locally the variation in spatial content in the neighborhood of the pixel. Texture perception has three high level features - repetition, orientation and complexity. The texture characterization is one of the most fundamental step in target detection. The target is mostly perceived as a spatial region. Target detection in natural images aims at presence or absence of particular aspect of the stimuli present in an object such as presence or absence of tonal or color, texture, orientation or shape, etc. The image or its subregion should be large enough for target detection, if not, abstraction of the image can roughly guide the target detection. The target detection has been contributed in Chapter 4 in terms of identifying the presence of tonal and texture regions present in a given image. In Chapter 6, the target detection aims at presence of orientation characterization in terms of row, column and interaction specific.

Target localization can be considered as a low-level partitioning or segmentation process of the natural scenery captured in the digital image which guides towards target recognition and scene understanding. Hence target detection should not be a limiting factor.

1.3 Role of perceptual grouping in object segmentation and characterization process

Segmentation divides the image into regions mainly using the concept of homogeneity at the feature level. The low features extracted needed to be grouped into a scene representation based on perceptual grouping principles, popularly known as Gestalt principles or Gestalt laws referred in a paper by Narasimhan (1963).

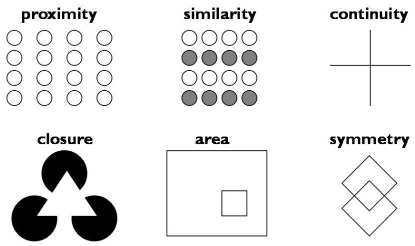


Figure 1.4: Gestalt Principles

The main Gestalt laws in Gestalt psychology describes the way human brain achieves perceptually grouping as referred in Figure 1.4. These laws are - similarity, proximity, closure, symmetry, simplicity, continuity. Similarity property groups similar features together which is accomplished using various distance measures. Proximity groups spatially close pixels. Closure fills in the missing gaps to give a complete picture of the entities. Figure 1.4 closure diagram shows three gaps between the rounded structures. Due to the closure concept, one tends to perceive the diagram as a triangle with three circles at its three corners. These concepts are used in image processing especially for image segmentation algorithms. The closure law helps in fine tuning and perceiving the object structure by filling in the missing pieces arising due to noise or occlusion. This can be found in Section 3.6, which provides details of post processing steps consisting of spatially constrained region merge. Grouping is an instance of the continuity principle in the sense that the oriented units tend to be integrated into perceptual wholes if they are aligned with each other. Figure 1.4 continuity diagram can be seen in either of the two perspectives- two lines meeting at a point or four lines meeting at a point. The first option is perceived as the most usual form. The two lines aligned vertically are grouped as a single vertical line and another two lines aligned horizontally are grouped as a single horizontal line. The mind has the capacity to perceive symmetry and the law helps in reducing the redundant information in visual forms. The mind tends to break a complex picture into simple possible forms say in terms of basic orientation, symmetry, regularity. This concept has been found to be very useful in texture characterization. Law of simplicity focuses on the whole of an object or scene rather than on its parts. From the Gestalt psychology it can be inferred that entity as a whole cannot be simply deduced from its parts put together, rather its respective parts gets influenced by the inherent perceptual structure of the whole. The whole (eg. car) carries a greater meaning than its individual components (tyres, axle, paint). A cognitive process takes place in viewing the whole while comprehending the parts.

1.4 Objectives and Contributions

The understanding of HVS and its perceptual processing capabilities, such as perceptual grouping, spatial resolving capacity and drawing inference in an unified manner, for different input modalities, provides an ever-present motivation and definite guidelines in the study of image processing and computer vision. The HVS also lays foundation, for future line of work in these areas. Certain visually related concepts such as, study of color, texture, brightness, contrast and resolution are quite significant, in development of the model of the visual system, where lots have been achieved, yet far from being complete. The Section 1.1 on HVS has been added for completeness in a holistic manner, as a motivation for automation in the vision research, though it has no bearing on our contributions. The domain of natural image segmentation is found to be a natural choice for the vision researchers, as it involves a substantial level of subjectivity. Natural scenic images, give rise to the crucial aspect of representation of various surfaces in the scene, that differs in both color and texture. Natural images fall under many object categories varying in color such as fruits, flowers, animals to a complex camouflage wildlife scene. The texture defines local spatial distribution property, of a region consisting of surface discontinuities, due to contrast in chromatic or achromatic components. The object surfaces exhibit different textural aspect, such as constant variations as in calm water, random variations in bushes, or regular horizontal or vertical variations as in bricked wall, striped animals or fabrics. For any given image, its feature set has a collective weightage of both color and texture features, that complement each other rather than compete. Recent trend shows the motivation of computer vision researchers, in integration of the two most fundamental descriptors - color and texture, to simulate the humans intuitive psychophysical approach, in order to find a more reliable segmentation and characterization, as mentioned in survey paper by Ilea and Whelan (2011). Our objective is to propose an efficient methodology for unsupervised color-texture image segmentation, and an effective automatic statistical based texture characterization tool for outdoor natural images. The following subsections briefs about our contributions - New taxonomy in Section 1.4.1, Unsupervised segmentation and its components in Section 1.4.2, Texture characterization and its components in Section 1.4.3, Framework and its approaches in Section 1.4.4.

1.4.1 Multi-Faceted Hierarchical Image Segmentation Taxonomy (MFHIST)

In our thesis, we present survey on both the aspects of color image segmentation and texture analysis. A comprehensive study on the evolution of color-texture image segmentation is prepared after referring to sixteen survey papers by Fu and Mui (1981), Haralick and Shapiro (1985), Shaw and Lohrenz (1992), Pal and Pal (1993), Reed and Dubuf (1993), Skarbek and Koschan (1994), Lucchese and Mitra (2001), Cheng *et al.* (2001), Zhang and Tan (2002), Vergés-Llahí (2005), Zhang (2006), Busin *et al.* (2008), Trémeau *et al.* (2008), Dey *et al.* (2010), Ilea and Whelan (2011), Vantaram *et al.* (2012); distributed over a span of four decades. We have referred four surveys for texture analysis by Reed and Dubuf (1993), Zhang and Tan (2002), Ilea and Whelan (2011) and Vantaram *et al.* (2012), over a span of last three decades. A gap in the existing segmentation taxonomy has been found, which has been mentioned later in literature survey Section 2.3. There is a need to enhance the characteristics of taxonomy to have apt representation as well as integration of image representation, feature descriptors and knowledge-driven model appropriate for mid-level and high level analysis. A multi-tier taxonomy called Multi-Faceted Hierarchical Image Segmentation Taxonomy (MFHIST) is recommended in the thesis to assist a researcher for a systematic categorization of most of these algorithms found in abundance.

1.4.2 Unsupervised Segmentation

Our segmentation approach is unsupervised, a fusion of edge based and region based. The feature space used here are various color representation spaces for proposed hybrid color space. The texture understanding has been achieved using orthogonal polynomials. The fusion aspects of color and texture, the recommended segmentation, performance aspects are briefly mentioned below. Depending on the feature set, it can produce null segments.

Hybrid Color Space

Researchers have experimented with more than one color spaces to suit their segmentation requirement. The subjective nature of the color space is used in an ad-hoc manner and is not suitable for all images. We propose a new hybrid color space, which is objective in nature, built from appropriate color feature spaces, which has couple of aspects embedded in it - color purity and color space with highest number of dimensions. The color components, when considered in fusion increase the dimension of the color dataset, which is then reduced using PCA to get Hybrid Color Space (HCS).

Texture understanding with Orthogonal Polynomials(OP)

Researchers have employed various simple to complex texture analysis methods to capture the texture understanding, specifying initial assumptions about scales, orientations. Challenge lies in capturing the texture, in a computationally simple manner for natural outdoor images that has randomness in texture patterns. Orthogonal Polynomial operators(OP) mentioned by Ganesan and Bhattacharyya (1995), is to make use of a wholesome bank of filters which is capable of interpreting the textural variations found in natural images without having any prior knowledge about the specific orientation and scale. A set of orthogonal polynomials is first used to obtain point-spread operators for a fixed window size, which subsequently can be used to generate a complete set of difference operators. Orthogonal effects are produced on applying these operators on the color planes of an image. Correspondence between orthogonal effects (main effects and interaction effects) and texture characterization has been identified and exploited. The micro texture can be characterized by the fact that interaction effects dominate over the main effects.

Color-Texture Feature Integration

We propose adaptive feature vector representation where number of features may vary from image to image depending on which color-texture feature variant is used to elicit local/subjective characteristics. For color texture feature integration, we propose two variants of set of filter banks of tensor products obtained from OP, i.e. OP3(order 3) and OP5(order 5), which are applied on HCS.

Clustering based Segmentation

An iterative K-means methodology has been developed to obtain appropriate number of clusters. The two straightforward stopping criteria used are region population size and maximum number of clusters obtained. The third stopping criterion is a crucial one - Kolmogorov-Smirnov Test which exploits the dissimilarity in the local color-texture coherence of the clusters as the objective function in deciding further segmentation. To achieve object segmentation, we implement two steps namely object labeling and region merge to obtain all the distinct objects in terms of color-texture characteristic from the image. Proposed image type, cluster and region metadata are generated as a by-product of the segmentation. Two variants of segmentation methodology are proposed - OP3-HCS and OP5-HCS.

The quantitative evaluation of the segmentation results is done using four metrics namely, Probability Rand Index(PRI), Variation of Information(VOI), Global Consistency Error(GCE) and Boundary Displacement Error(BDE), which are popular and most widely used metrics for benchmarking the performance of the segmentation algorithms. Quantitatively, the higher PRI and lower VOI, GCE and BDE are considered to be better for segmentation accuracy. Comparisons are drawn with several other color-texture segmentation algorithms taken from literature. It may be noted that our proposed method does not involve any smoothing process or any preprocessing step. This unsupervised segmentation method exhibits a balanced nature between over and under segmentations for extracting meaningful content, based on the analysis of the above four metrics. Berkeley Segmentation Dataset(BSD) has been used for our experiments as this dataset is meant for object segmentation.

1.4.3 Texture Characterization

The natural outdoor image is a complex one, with major components tonal, texture and their mixtures. The bottom up methods may not be potential to characterize the image, in a perceptual manner for content development. Clustering is an amalgamation process which is a bottom-up and may not be potential to extract holistic characteristics of the image. Attempt needs to be made for extracting globally representative characteristics. Most of the texture characterization is done either keeping a specific domain (synthetic or natural images specific to a category) or is application specific (image retrieval). We develop a methodology for building the essential knowledge for developing objective texture characterization based on inferential statistics called Analysis of Variance (ANOVA) free from such subjectivities. A schematic texture characterization has been evolved based on ANOVA 2-way and 3-way methodologies with replication, for building adequate knowledge for a natural image.

Progressive Sampling

There are various sampling techniques such as simple random technique, systematic random sampling. The characterization technique we are using is ANOVA with replication, hence there is a need for proposed new progressive sampling, for minimizing sampling error in terms of pseudoreplication. Progressive sampling implements the sample selection along the edges, so that maximum variation is achieved quickly among the samples. In case of image sampling, spatial pseudoreplication may arise, if sampling density is high. To overcome the problem of pseudoreplication, we have devised control parameters to achieve the appropriate sampling density. The collection of the features obtained from image samples is rich in capturing both luminance and chrominance information. 2-way ANOVA model has been administered for 2-D samples, whereas 3-way ANOVA model has been considered for 3-D HSV color space.

ANOVA Grading

Grading implies a differentiation of the texture and non-texture property. We propose a *grading scheme* represented as string code representation, which will reflect the specific orientation using appropriate character codes for significant or non-significant main and interaction effects. For 2-way ANOVA, three hypothesis tests and in case of 3-way ANOVA, seven hypothesis tests are performed. Depending upon the p-value for each hypothesis; grade characters are assigned for high, medium and no significance respectively. Based on these characters, string of length 3 for 2-way ANOVA and a string of length 7 for 3-way ANOVA is generated.

Knowledge Representation

Given region will be assigned a character string based on ANOVA grading system, by concatenating codes in a sequence obtained from various color feature sets. This is referred as string code which is in lossless form. Further meaningful phrase is attached to this string named as semantic code using proposed polling method. The collection of all these string codes is referred as knowledge representation of the image. This has been demonstrated on various benchmark datasets namely BSD, Vistex, Corel and Prague.

1.4.4 Framework

The framework usually supports a specific objective and serves as a guide to make use of the components as required. We have developed a conceptual framework, which acts as an analytical tool for color-texture image segmentation and texture characterization, that can work with several tonal-texture variations present in natural outdoor images. The framework consists of two approaches - segmentation approach and block approach. Segmentation approach has two sub-components in sequence - unsupervised segmentation and texture characterization. The block approach will go through block generation and then texture characterization. The output of the framework is a collection of string and semantic codes and is referred as knowledge representation.

Segmentation Approach

A natural outdoor image when given as an input to the framework, can be segmented into objects or regions, which can be further processed for texture characterization. The knowledge representation of the segments, after texture characterization becomes more informative with additional inferential statistical ANOVA grading, to the already available descriptive statistics metadata computed during segmentation stage.

Block Approach

The natural patterns are usually repetitive and may contain similar textural arrangements. A systematic way of texture analysis is provided by further partitioning of one image block into two block sets - set of 4 and set of 9. The string characterization of the image block and its block subsets provide a hierarchical aspect to the problem of texture analysis.

1.4.5 Publications

1. Tilottama Goswami, Arun Agarwal, C.Raghavendra Rao, Hybrid region and edge based unsupervised color-texture segmentation for natural images, International Journal of Information Processing(IJIP), 9(1), pp 77-92, 2015.

2. Tilottama Goswami, Arun Agarwal, C. Raghavendra Rao, Color Texture features Integration for Unsupervised Image Segmentation based on Orthogonal Polynomial Operators and Hybrid Color Space, ICISP-2014, pp 52-59, Elsevier publication

3. Tilottama Goswami, Arun Agarwal, C.Raghavendra Rao, Statistical Learning for Texture Characterization, Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP-2014), ACM (online)

The contributions in Chapter 3 and 4 have been published in [1](Goswami *et al.* (2014*a*)) considering only OP3-HCS segmentation methodology, and in [2](Goswami *et al.* (2015)) considering both the OP3-HCS and OP5-HCS segmentation variants. The texture characterization related contributions mentioned in Chapter 5 has been published in [3] (Goswami *et al.* (2014*b*)).

1.5 Thesis Organization

The thesis is organized in seven chapters. The Chapter 2 is on literature survey which is organized using a proposed new multi-tier taxonomy on image segmentation. Chapter 3 provides step-by-step discussion on our proposed framework of color-texture feature integration for unsupervised image segmentation based on two variants OP3-HCS and OP5-HCS. In Chapter 4, the quantitative evaluation of the proposed image segmentation approach is demonstrated. The comparative study of our proposed color-texture image segmentation with other state-of-art and recent methods is also mentioned here. Chapter 5 discusses our contributions on progressive sampling and ANOVA grading system, which are considered to be fundamental steps in our proposed statistical based learning for texture characterization. Chapter 6 gives an overview of the framework bringing together color-texture segmentation and texture characterization of any image, mostly natural outdoor images. The thesis summarizes and concludes with future work in Chapter 7.

CHAPTER 2

Literature Survey based on new Multi-Faceted Hierarchical Image Segmentation Taxonomy

The color and texture information has strong implication with human perception in describing, characterizing and recognizing the image content. This psychophysical perception motivated the researchers from computer vision community to explore more on these fundamental descriptors in a collective manner in order to globally and locally extract the image characteristics. The natural images, in particular, pose to be a challenging task because they consist of significant in-homogeneities both in color and texture. The color-texture feature extraction and integration is a part of low-level image processing. Nevertheless, it attracted substantial interest among the vision researchers in the area of low-level image segmentation and mid-level image analysis tasks such as texture analysis and characterization. There are various association of both image segmentation and texture analysis with high level computer vision and pattern recognition tasks like image classification, image retrieval, texture classification, object classification, object recognition, scene understanding and action recognition. This chapter provides an overview of two main directions of research related to these key features, namely image segmentation and texture analysis. The chapter mentions various groupings in context with surveys, taxonomy, color-texture feature integration, segmentation approaches and texture analysis. To ensure a good clarity and understanding, we would like to provide appropriate vocabulary for these context based groupings. Literature surveys are grouped into *classification*. The proposed multifaceted hierarchical taxonomy is structured in layers for each facet called *category levels* and its corresponding choices to be made as *sub-categories*. The way, the color-texture features extraction and integration is done can be grouped as *trends*. The segmentation approaches are grouped into *types* and *sub-types*. Texture analysis is grouped into various *methods*. Any process adopted by a researcher's work in the area of image segmentation and texture analysis is mentioned as *technique* and thus can be categorized using the proposed Multi-Faceted Hierarchical Image Segmentation Taxonomy (MFHIST).

After going through exhaustive list of survey papers on image segmentation, we made an attempt to study, comprehend and summarize the areas of focus, scope and state-of-art techniques in image segmentation and texture analysis. In this chapter, we provide a comprehensive overview with the goals to: 1) Make an effort to provide a contrasting analysis of papers distributed over a span of four decades. A multi-tier taxonomy is proposed to assist a researcher for a systematic categorization of the segmentation algorithms found in abundance. These segmentation methods are developed to cater to various application scope starting from low-level to high-level image processing algorithms. 2) Provide an overview of the texture analysis methods to give an idea about the texture framework.

The remainder of this chapter is organized as follows. Section 2.1 describes the organization of all the literature surveys on image segmentation being considered for the thesis. Section 2.2 focuses on the evolution of segmentation taxonomies from 1970s till current. Section 2.3 discusses the proposed new multi-tier taxonomy, MFHIST and its facets. The section also mentions some of the recent as well as popular state-of-art techniques in the area of color-texture image segmentation. Section 2.4 focuses on the texture analysis, the concepts and its state-of-art methodologies. Section 2.5 gives an illustration of populating MFHIST with some of the research works. Finally, Section 2.6 summarizes our observations and concludes with future work.

2.1 Organization of literature surveys

A comprehensive study on the evolution of color-texture image segmentation is prepared after referring to the 16 survey papers. The referred surveys spanning over a time period of four decades are listed in Table 2.1. We have divided the time period of four decades starting from 1975 till date into four time periods: 2 survey papers from 1975-1985, 4 survey papers from 1985-1995, 4 survey papers from 1995-2005, 6 survey papers from 2005-2015 (as shown in the time-line diagram below in Figure 2.1).

The literature surveys on image segmentation can be grouped into three broad classifications:

1. Survey on image segmentation in general
2. Survey on image segmentation specific to an application
3. Survey on image segmentation specific to a certain approach or technique.

(a)

Figure 2.1: Time-line Diagram

The classification-1 discusses about the evolution of image segmentation in general, such as survey done by Fu and Mui (1981), Haralick and Shapiro (1985), Shaw and Lohrenz (1992), Pal and Pal (1993), Lucchese and Mitra (2001), Cheng *et al.* (2001), Vergés-Llahí (2005), Zhang (2006), Busin *et al.* (2008), Trémeau *et al.* (2008) and Vantaram *et al.* (2012).

Under the classification-2, application specific surveys are mentioned, such as survey on medical image analysis by McInerney and Terzopoulos (1996), on remote sensing by Dey *et al.* (2010), on industrial applications by Elena González and Saetta (2013). It can be seen that the application oriented surveys, though not very high in number play an important role in catering to specific community and researchers group.

Under the classification-3, surveys are focused on a particular approach which became the area of focus during that period. The survey on early approach of thresholding

Table 2.1: List of Literature Surveys on Image Segmentation

S.N.	**Year**	**Author**	**Attributes**
1	1981	Fu and Mui (1981)	Grayscale
2	1985	Haralick and Shapiro (1985)	Grayscale
3	1992	Shaw and Lohrenz (1992)	Grayscale
4	1993	Pal and Pal (1993)	Grayscale
5	1993	Reed and Dubuf (1993)	Grayscale Texture
6	1994	Skarbek and Koschan (1994)	Color
7	2001	Lucchese and Mitra (2001)	Color
8	2001	Cheng *et al.* (2001)	Color
9	2002	Zhang and Tan (2002)	Grayscale Texture
10	2005	Vergés-Llahí (2005)	Color
11	2006	Zhang (2006)	Color
12	2008	Busin *et al.* (2008)	Color
13	2008	Trémeau *et al.* (2008)	Color Texture
14	2010	Dey *et al.* (2010)	Color Texture Shape
15	2011	Ilea and Whelan (2011)	Color Texture
16	2012	Vantaram *et al.* (2012)	Color Texture Contour

techniques is mentioned by Sahoo *et al.* (1988). Reed and Dubuf (1993), Tuceryan and Jain (1998) and Zhang and Tan (2002) have done a survey on texture analysis and texture segmentation methods. A great deal of activity on recent statistical approaches to level set segmentation is referred in survey paper by Cremers and Rousson (2007). In the present decade more emphasis has been given to hybrid combinations in various aspects, such as combination of color-texture features and their integration is mentioned by Ilea and Whelan (2011), segmentation techniques with both region and boundary information integration by Freixenet *et al.* (2002).

This survey does not discuss the mathematical details behind the techniques for simplicity. Our aim is to provide a broader picture of segmentation in terms of how it has evolved. We perceive segmentation as multi-faceted design approach and propose a multi-level taxonomy as discussed in the next section.

2.2 Evolution of Existing Segmentation Taxonomy

Firstly, we list down the existing segmentation taxonomy found in the literature surveys across the decades. The inter-relationship of taxonomies is shown in Figure 2.2. The mapping among the taxonomies is shown pictorially as how they have evolved as re-grouping, sub-grouping and sometimes totally new group emerged. In recent decade, the segmentation taxonomy which has been grouped into three broad types - 1) Feature space based or spatially blind, 2) Image domain based or spatially guided and 3) Physics based. All other taxonomies found earlier surveys form a sub-type of the above mentioned three types.

Feature space segmentation approaches consider that the connected regions, which are not homogeneous in certain feature space such as color or texture, cannot be merged into single region. They usually neglect the spatial relationships among feature space. There is no guarantee that the regions obtained from clustering will show spatial compactness. Earlier in 1980's Haralick and Shapiro (1985) grouped the segmentation

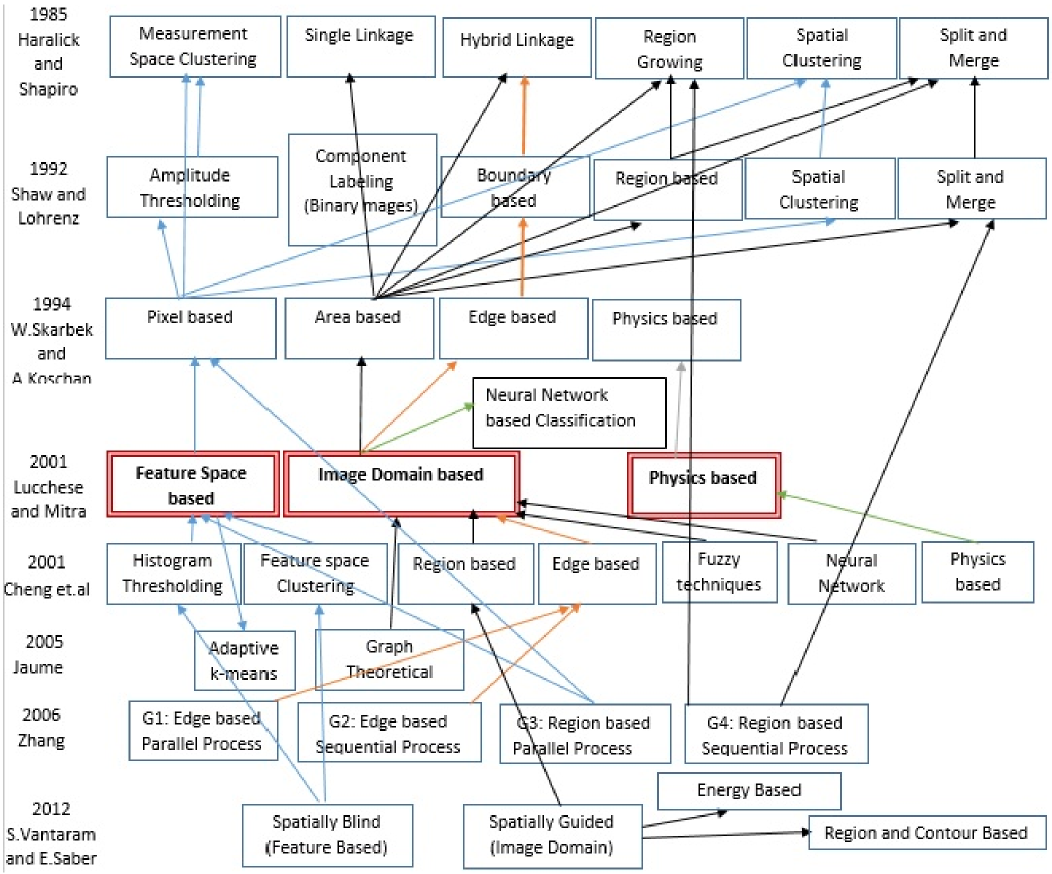


Figure 2.2: Evolution of Image Segmentation Taxonomy over the decades

algorithms in six types, of which *Measurement space clustering* and *Spatial clustering* groups got merged to *Pixel based* type in survey by Skarbek and Koschan (1994); which finally got renamed as *Feature space based* by Lucchese and Mitra (2001). In 2012, Vantaram *et al.* (2012) further renamed it as *Spatially blind*. Similarly the types *Amplitude thresholding* and *Spatial clustering* mentioned by Shaw and Lohrenz (1992) got merged into *Feature space based*, named by Lucchese and Mitra (2001). Some of the techniques which fall into the feature space based approach are clustering and thresholding. K-means, fuzzy C-means, mean-shift, neural network based data clustering - self organizing map are some of the clustering techniques.

Domain specific segmentation or spatially guided approaches consider spatial compactness to be another desirable property apart from homogeneity, where pixels are clustered exclusively on the basis of their spatial relationship or surface coherence. Initially in 1985, four types are mentioned by Haralick and Shapiro (1985) - *Single linkage*, *Hybrid linkage*, *Region growing*, *Split and merge*. These were grouped as *Area based* approaches in 1994 by Skarbek and Koschan (1994). These *Area based* along with *Edge based* and *Neural network based* later were merged into a new segmentation type called *Image domain based* in 2001 by Lucchese and Mitra (2001). During the same period, Cheng *et al.* (2001) merged the upcoming *Fuzzy techniques* and *Region based* as a sub-type to this group. Later in the year 2005, Vergés-Llahí (2005) created a new sub-type called *Graph theoretical* to highlight its contribution in the field of image segmentation under the *Image domain* group. Recently in 2012, Vantaram *et al.* (2012) conveyed the importance of yet another two more rising concepts as sub-types - *Energy based* and *Region and contour based*, which got added to the group.

Physics based segmentation is based on reflection model which is highly related to the nature of the materials. Healey (1989) grouped materials into two classes - optically homogeneous (metals, crystal, glass) and optically inhomogeneous (plastic, paint, paper). The segmentation method considers the dichromatic reflection model for analyzing color in an image and has been used by Maxwell and Shafer (1995). The physics based segmentation is normally not in much use due to its limitations.

Most of the aforementioned surveys are single-tier taxonomy except surveys by Dey *et al.* (2010) and Vantaram *et al.* (2012) who have proposed multiple aspects of categorization. Dey *et al.* (2010) provided survey on object segmentation with remote-sensing perspective based on segmentation control, feature homogeneity measure, image operations. The taxonomy proposed does not consider the supervised and unsupervised control specification in the taxonomy. On a similar note, Vantaram *et al.* (2012) proposed taxonomy based on - high level and low-level classification. The image type considered in this survey is limited to color and grayscale. One would need a more generic grouping in terms of image modality (X-Ray, ultrasound, photo) and dataset type (real or synthetic). Vantaram *et al.* (2012) mentions human interaction as the only way for achieving supervised segmentation. There are other model-driven or knowledge driven top-down approaches which can fit into the supervised modes. There is a need to adequately capture semi-supervised control approaches (such as paper by Zhu (2006)) in the taxonomy. The main difference between fully-supervised and semi-supervised segmentation is the type of learning the algorithm adapts. We propose to fill this gap by separating the user interaction from knowledge driven and data driven aspects in our taxonomy in the next section. Vantaram *et al.* (2012) mentions top-down and bottom-up approaches in miscellaneous category of low-level taxonomy, which is not very clearly mentioned in the paper.

The tremendous growth of image segmentation methods makes the design of taxonomy more challenging as more and more new recent techniques and approaches are emerging for segmentation based computer vision applications. The segmentation scope is not restricted to only the low level analysis but goes beyond. We have proposed a systematic multi-tier taxonomy which guides towards the appropriate approaches and techniques to be used depending on the scope of the application. A multi-tier segmentation taxonomy is proposed in the forthcoming section.

2.3 MFHIST - A Proposed Image Segmentation Taxonomy

After going through each of the above mentioned surveys and their corresponding taxonomies, we present our views in a way, we look at the multi-faceted image segmentation. Each facet can be considered as a design task for the segmentation. The facets are mostly independent but at the same time complement each other. Every facet has multiple mutually exclusive choices. We see the task of segmentation as a collection of these choices made in an appropriate hierarchical manner. The hierarchy of facets is represented as layers of category levels, as shown in Figure 2.3. The various facets of the segmentation are - scope specification, requirement specification, control specification, feature specification, image representation and segmentation approach specification.

In accordance to this view, we propose a taxonomy, which is a general method of grouping the segmentation algorithms, and is also relevant beyond image segmentation scope and extends to applications related to image analysis and computer vision. Depending on scope of the application - low level, mid level or high level, the segmentation approach remains simple or becomes complex as it goes from low to high level. One such example - region segmentation is low level, but object segmentation is a mid level analysis and object recognition and scene classification fall under high level image understanding. The proposed taxonomy not only captures the overall intention of the segmentation requirement, but also the extent of the level to which it has to be used. The six category-levels are discussed further below.

1. Category Level-1: Refers to **scope specification** of the research paper, whether it is meant to be used at a low-level analysis, mid-level or high-level image understanding. For higher-level, more detailed and different specifications are required which was not getting captured in the earlier mentioned taxonomies.

2. Category Level-2: Refers to **requirement specification** of the image segmentation in three ways - 1) Specific application or generalized, 2) Data source used

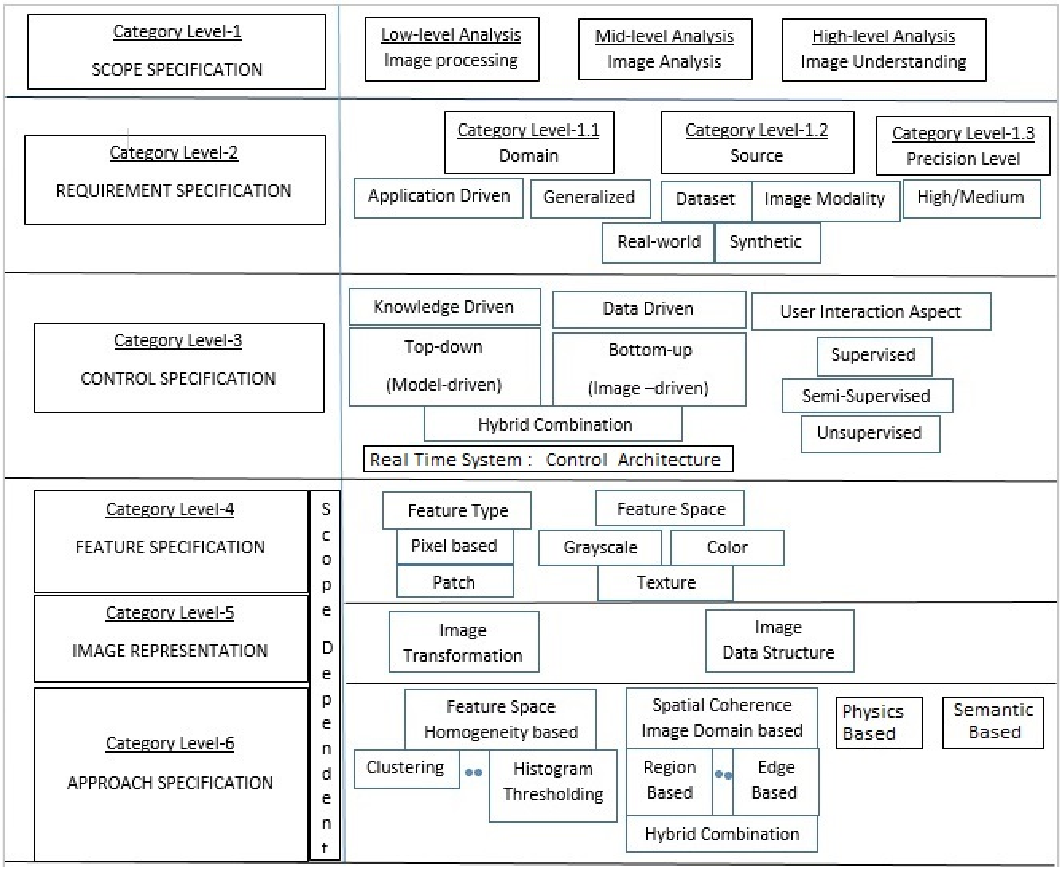


Figure 2.3: Our proposed Multi-Faceted Hierarchical Image Segmentation Taxonomy (MFHIST)

for segmentation and 3) Precision level to be achieved.

3. Category Level-3: Refers to **control specification** which specifies both the extent of the user control and prior information about the image separately. Top-down approach is knowledge driven where a priori information is provided in some form, to guide the segmentation. In contrast, bottom-up information is extracted from the given image, which is intrinsic or embedded into it, hence meant to be unsupervised. The bottom-up approach is problem independent and works irrespective of complexity of images. Top-down is supervised. It has been found in literature that hybrid combination of top-down and bottom-up approaches provide a more robust segmentation and better accuracy.

4. Category Level-4: Refers to the **feature specification**. Color, texture, shape,

intensity values are some of the widely used features. The feature detection is more complicated in case of computer vision applications and high level image analysis and also is tightly coupled with the image representation.

5. Category Level-5: Refers to the **image representation** that represents the transformed feature space, from which the features need to be extracted. The representations can have two configurations - single-scale and multi-scale configurations. Image data structures like tree, graph are some of the commonly used efficient data structures found in the literature.

6. Category Level-6: Refers to the **segmentation approach specification** that can be directly mapped to the types and sub-types of segmentation approaches, mentioned in existing taxonomy. The physics based segmentation is normally not in much use due to its limitations. The other two basic approaches are feature based and image domain based. Image domain considers spatial coherence and there are basically two popularly used sub-types - region based and edge based. Recent emerging computing paradigm, such as granular computing and rough sets can be added to the taxonomy under image domain approach type. We also propose to add a new high-level type for semantic based segmentation approach to cater to high level computer vision tasks.

Each of the above mentioned, hierarchical category levels corresponding to various facets such as scope, requirement, control, feature, image representation and approach specifications are discussed in details below.

2.3.1 Scope Specification

Every segmentation algorithm is developed considering the scope of it, i.e the extent to which it might be of use. The scope specification can be divided into three sub-categories - 1) low-level image processing 2) mid level image analysis 3) high-level image understanding.

Image processing techniques is a part of low-level processing which include extraction of maximum information from image in terms of features, segment the images into disjoint homogeneous regions. The focus is mainly on developing a general methodology to segment the image upto a satisfactory level without human intervention, based on appropriate feature-space. To cite examples, few research papers catering to this scope are on color histogram by Bhoyar and Kakde (2010), non-parametric clustering with Gaussian blurring mean-shift by Carreira-Perpinán (2006), voting framework based on Bayesian flooding and region merging by Panagiotakis *et al.* (2014), graph-based segmentation by Felzenszwalb and Huttenlocher (2004). The segments obtained at this stage can be further improved and modified to aid the interpretation using image analysis tasks.

Image analysis is considered to be at a higher level process for extraction of meaningful objects from image; inspired by human visual aspect of perceptual grouping. The algorithms developed for image analysis tends to be more complex and domain-specific - such as medical imaging, remote-sensing. Content based image retrieval, texture classification, are some of the related areas. The papers on image analysis not only extracts features, segments but also characterizes to provide additional knowledge information to assist in reducing the semantic gap. To cite examples, few research papers catering to this scope are on perceptual features for CBIR by Abbadeni (2011), texel extraction by Ahuja and Todorovic (2007), color-texture characterization by Sanda *et al.* (2013), semi-automatic segmentation approach for extracting object contours by Adamek and O'Connor (2006), etc.

Image understanding algorithms as mentioned by Rosenfeld and Kak (1982), are high-level processes and it includes methods for acquiring, extracting, analyzing and understanding images in the form of decisions for recognition and interpretation purposes. Such algorithms can be found in papers for object class recognition such as by Shotton *et al.* (2009), for class based segmentation such as by Borenstein and Ullman (2004), building classification in satellite imagery such as by Belgiu and Drăgut (2014), for defect detection applicable to ceramic tiles industry by Xie and Mirmehdi (2007*b*).

Table 2.2: Scope Specification

Low-level Image Processing	**Mid-level Image Analysis**	**High-level Image Understanding**
Bhoyar and Kakde (2010)	Abbadeni (2011)	Shotton *et al.* (2009)
Carreira-Perpinán (2006)	Ahuja and Todorovic (2007)	Borenstein and Ullman (2004)
Panagiotakis *et al.* (2014)	Sanda *et al.* (2013)	Belgiu and Drăgut (2014)
Felzenszwalb and Huttenlocher (2004)	Adamek and O'Connor (2006)	Xie and Mirmehdi (2007*b*)

The Table 2.2 gives a demonstration for categorization of few research papers on image segmentation with different scope/target applications.

2.3.2 Requirement Specification

Application specific segmentation algorithms are found large in number and are especially designed to work with images pertaining to that particular area as compared to segmentation of any image in general.

Application specific algorithms have prior knowledge about the input images and are quite popular for their effectiveness in terms of accuracy, in contrast to a generalized segmentation algorithms without having any knowledge about the image. Application specific algorithms are usually complex in nature and time consuming. A generalized segmentation can be considered as an time efficient preprocessing tool with some deterrence in accuracy, which can guide to higher order complex algorithms for application specific purpose in an effective manner. All the research papers mentioned in Table 2.2 are considered once again in Table 2.4 for demonstrating the filling up the requirement specification in terms of domain, source of images and pre-requisite precision level. It can be observed that the low level image processing methodologies developed by Felzenszwalb and Huttenlocher (2004), Carreira-Perpinán (2006), Bhoyar and Kakde (2010), Panagiotakis *et al.* (2014) have generalized approach whereas the mid-level analysis and high level understanding methodologies are mostly applica-

REQUREMENT SPECIFICATION	Domain	Source: Dataset	Source: Image Modality	Precision Level
Bhoyar and Kakde (2010)	Generalized	Real-world Natural Images BSD	Color	Medium
Abbadeni (2011)	Application Specific Texture Retrieval	Real-world Natural Textures Brodatz Images	Grayscale	Medium
Shotton et. al. (2007)	Application Specific Semantic Segmentation and Editing of Photographs	Real-world MSRC-21 class Corel- 7 class Sworby-7 class TV shows	Color	Medium
Carreira-Perpinán (2006)	Generalized	Real-world Hand Cameraman	Color Grayscale	Medium
Ahuja and Todorovic (2007)	Application Specific Texel detection and Segmentation	Real-world 2.1D textures	Grayscale	Medium
Borenstein and Ullman(2004)	Generalized Learning Class based Segmentation	Real-world Horse, Cars, Human faces	Color	Medium
Panagiotakis et al. (2014)	Generalized Texture Segmentation	Real-world Natural Images Prague Mosaic Textures	Color	Medium
A.T.Sanda et al. (2013)	Application Specific Color-Texture Characterization	Real-world Natural Textures Vistex Images	Color	Medium
Belgiu and Dragut (2014)	Application Specific Object based image analysis for classifying buildings	Real-world QuickBird WorldView-2	VHR Satellite Imagery	High
Felzenszwalb and Huttenlocher(2004)	Generalized	Real-world COIL images and Synthetic	Color Grayscale	Medium
Adamek and O'Connor(2006)	Application Specific	Real-world Personal images	Color	Medium
Xie and Mirmehdi (2007)	Application Specific Defect Segmentation for ceramic tile production	Real-world Ceramic Tiles Vistex images	Color	High

Figure 2.4: Requirement Specification

tion specific. Nevertheless, the paper by Borenstein and Ullman (2004) mentions their proposed class based segmentation is general. Usually, the experiments are carried out by the research community on standard datasets.

2.3.3 Control Specification

The main thrust for going ahead with semi-automatic or interactive algorithms is to cater to the critical applications which require error free preprocessing segmentation task. This involves user interaction explicitly or implicitly. Image segmentation approaches can be broadly grouped under control facet as supervised/semi-automatic and unsupervised/automatic, depending on whether the segmentation is guided by training using prior information about the objects or object-classes. This facet is tightly coupled

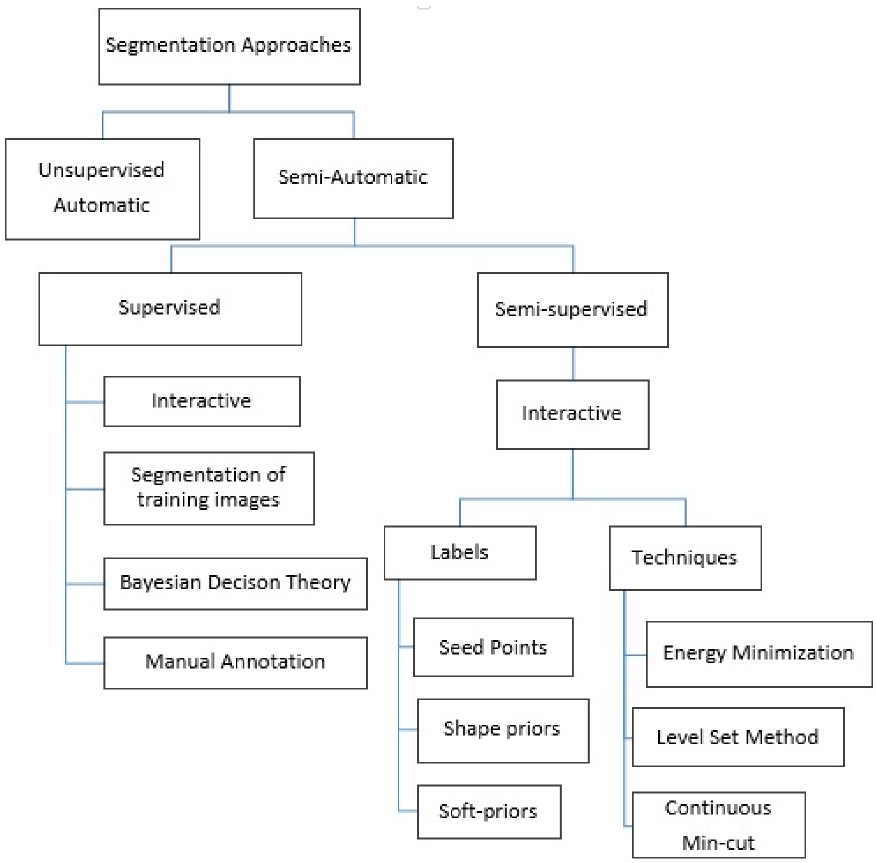


Figure 2.5: Supervised Taxonomy

with choices made under various sub-categories. Only the supervised methodologies are presented in a tree structure as shown in Figure 2.5.

Data classification techniques are supervised and need training data, whereas data clustering techniques are unsupervised and train themselves using the present data i.e. data driven.

Automatic segmentation provides an efficient solution to most of the applications that need quick, coarse and region based segments. Automatic or unsupervised does not have any prior knowledge about the objects or the object-classes and is mostly a bottom-up segmentation. Bottom-up segmentation is dependent on low level cues such as color values, intensities, texture, contours and their various combinations. They are also referred as data-driven algorithms and are mostly used in general segmentation techniques. The main challenge with unsupervised is to achieve accurate meaningful segments or objects present in an image.

The high-level information can be entered into the system in various possible ways using the top-down approach. Top-down approach is model driven and starts with certain assumptions about the target object to be segmented. They are mostly applicable for domain specific segmentation tasks, such as face extraction, medical image analysis, etc. The semantic gap can be filled only by providing object level information prior to the segmentation. The supervised segmentation solution if invoked by human operator is called interactive segmentation. One such way is assistance provided by human operator for marking foreground and background objects through interactions. These information can be templates as mentioned in Table 2.3 to supervise domain or application specific object detection and recognition, which gives rise to the need of supervised segmentation. The main difference between fully-supervised and semi-supervised segmentation is the type of learning the algorithm adapts. Some of the semi-automatic approaches use complex descriptors such as bounding-box, interest point detectors, figure-ground labeling for supervision, as shown in Figure 2.6. Such algorithms have been found to be quite effective when top-down segmentation approach is further combined with bottom-up approach Borenstein and Ullman (2004).

Table 2.3: What do we understand by prior knowledge?

S.N.	Prior knowledge	Usage
1	Deformable shape templates	Automatic image interpretation
2	Probability distributions on shape variables	Object shape
3	Gibbs Random Field	Image prior to model and enforce spatial coherence

Figure 2.6: Semi-automatic approaches(left to right): Bounding-box descriptor, Interest point detectors, Figure-ground labeling

The Table 2.4 shows the control strategy used for the set of the algorithms being used for demonstrating the proposed taxonomy. Unsupervised learning as mentioned by Ahuja and Todorovic (2007) is guided on the basis of the statistical properties embedded in the input image. On the other hand, supervised learning using the training set as mentioned in paper by Borenstein and Ullman (2004) is based on the training data which includes both input and desired results for accurate learning process.

Mobile robot systems bridges the high-level vision techniques with real-world environments. /textbfReal-time AI includes approximate data, intermediate representations, approximate reasoning methodologies, and approximate control strategies. In such a complex dynamic environment - data, domain knowledge and control representations are not fixed, but are continuously affected by approximate processing, due to uncertainty, imprecision and sometimes incomplete information present in the environment encountered by the robot as mentioned by Decker *et al.* (1990). The control architecture used for approximate reasoning are blackboard system, extension of blackboard system and whiteboard system; usually independent of domain. The blackboard system is viewed as cooperative functioning of various knowledge sources, each one performing non-deterministically. It is an aid to cooperative problem solving, in which a global

Table 2.4: Control Specification

Research Papers	Control	User Interaction
Charles Thorpe and Shafer (1988)	Whiteboard CODGER system	Unsupervised
Decker *et al.* (1990)	Blackboard	Unsupervised
Agarwal and Reddy (1990)	Extended Blackboard ARGUS system	Unsupervised
Ray and Sharma (1990)	Extended Blackboard RBD system	Unsupervised
Felzenszwalb and Huttenlocher (2004)	Bottom-up	Unsupervised
Borenstein and Ullman (2004)	Hybrid	Training
Adamek and O'Connor (2006)	Top-down	Supervised
Carreira-Perpinán (2006)	Bottom-up	Unsupervised
Ahuja and Todorovic (2007)	Bottom-up	Unsupervised Learning
Xie and Mirmehdi (2007*b*)	Bottom-up	Unsupervised Training
Shotton *et al.* (2009)	Bottom-up	Unsupervised Training
Bhoyar and Kakde (2010)	Bottom-up	Unsupervised
Abbadeni (2011)	Bottom-up	Unsupervised
Sanda *et al.* (2013)	Bottom-up	Unsupervised
Panagiotakis *et al.* (2014)	Bottom-up	Unsupervised
Belgiu and Drăgut (2014)	Bottom-up	Unsupervised

or local view of problem solving process, determines the goals to be achieved, and processes that may achieve the goals. It constitutes the objects and their instances, current assertions and rules. Ray and Sharma (1990) proposed RBD-Reasoning Blackboard for Data fusion system, which is an implementation of a shell to build expert system. Agarwal and Reddy (1990) proposed a skeletal system for tackling the problem of information interpretation, the archival representation for a general understander system called ARGUS based on Distributed Problem Solving (DPS) strategy. This helps to tackle the problem of information interpretation in AI research in the sub-areas like speech and vision in a distributed system. This is an extended version of blackboards permitting distributed control and parallel asynchronous execution of various processes. The archival representation for a general understander system is mentioned in paper by B.N.S. Chakravarthy and Reddy (1988). In earlier blackboard distributed control system, a single master process is aware of the capability of each AI module, but may not know the internal working of each. The master thus becomes the major AI module and the main bottleneck. In contrast to this blackboard control, the paper by Charles Thorpe and Shafer (1988) proposes whiteboard distributed control architecture called CODGER-Communication Database with Geometric Reasoning, where the individual modules in the mobile robot system is autonomous, and are free to decide when and how to communicate and when to execute. Charles Thorpe and Shafer (1988) developed a mobile robot system which allows to interact the vision techniques with real-world environments, integrating the perception and navigation capabilities. Color vision is used for locating roads in color images and 3-D vision used for obstacle detection and terrain modeling. The authors propose a distributed architecture articulated around a knowledge database, which coordinates control and information flow between the high-level symbolic processes running on computers, and lower-level control on real-time hardware.

The thesis focuses on unsupervised segmentation, hence it is beyond the present scope to further elaborate on the supervised and semi-supervised, real-time distributed control methodologies. Nevertheless, this taxonomy is capable of accommodating, in general all segmentation algorithms irrespective of the extent of user interaction aspects or not. Detailed discussion on high-level image understanding is not in the present scope.

2.3.4 Feature Specification

Color and texture features of a color image are the two most significant attributes that complement each other. Color is crucial to many pattern recognition and computer vision tasks, as its higher dimension aspect adds additional information as compared to the intensity gray shades. A review on *Color Spaces* can be found in Appendix A. The color features are usually computed from pixel-based image representation. The feature values can be extracted from 3-dimensional pixel of any given color image as mentioned by Gonzalez and Woods (2007). Researchers Mirmehdi and Petrou (2000), Hedjam and Mignotte (2009) have experimented with more than one color spaces for segmentation to suit their segmentation requirements. Goswami *et al.* (2014*a*) proposed Hybrid Color Space (HCS), taking advantage of various color space components.

Texture is considered to be a major hindrance for all segmentation techniques. In the vision research community, this led to separate discussions on texture analysis from the color representation and analysis perspective. The texture feature is an area feature, a spatially cohesive group of pixels satisfying a statistical/structural/model based homogeneity predicates as referred by Zhang and Tan (2002). The textural feature extraction in images, texture retrieval, texture classification, texture segmentation are some of the important related topics. Most of the texture processing approaches assume the images to be already having some predefined textures and experiments are carried out on textured datasets. Natural images have random textures and interpreting the textural variations without having any prior knowledge about the specific orientation and scale is very difficult. Compared to only-color or only-texture information, the collective use of color and texture information links with human perception; and works more accurately in describing the feature set in practical scenarios.

Literature surveys on segmentation algorithms shows a gradual shift of choosing image feature descriptors from only color (by Cheng *et al.* (2001), Lucchese and Mitra (2001)) or only texture (by Reed and Dubuf (1993), Tuceryan and Jain (1998), Zhang and Tan (2002)) to color-texture integration (by Vantaram *et al.* (2012), Ilea and Whelan (2011)). Recent trend shows the motivation of computer vision researchers in integration of the two most fundamental descriptors - color and texture to simulate the humans intuitive psychophysical approach in order to find a more reliable segmentation.

For higher level image understanding, tremendous growth has happened in the recent decade to capture the appearance of objects in a compact form called the patch representation. For object descriptors, usually the patch type features are used. In one of the recent papers by Ciurte *et al.* (2014) on semi-supervised segmentation method for ultrasound images, initial labels are defined as soft priors. Ciurte *et al.* (2014) has mentioned that the use of patches as image features is more appropriate for achieving independence in the imaging systems (medical diagnosis) as they are flexible and accurate to deal with multiple segmentation scenarios. The concept of patch features were first introduced for texture synthesis by Efros and Leung (1999). In case of semantic based segmentation, higher-level image descriptors better represent the queries usually formulated by a user. For object class recognition, focus has shifted to patch descriptors based on Scale Invariant Feature Transform (SIFT) developed by Lowe (2004), interest point detectors to capture the appearance of objects in a compact form. A number of recent approaches have used fragments or patches to perform object detection and recognition. Borenstein and Ullman (2004) uses shape fragments to segment class objects from the background. A recent survey on both high-level and low-level image feature descriptors can be found by in a paper by Kumar and Sreekumar (2014).

For object classification system, detailed description of both image representation (classified as sparse and dense) and appropriate feature descriptors can also be found in thesis by Schroff (2009). The sparse representation consists only a subset of image regions called regions of interest using edge, corner and region detectors. The appropriate feature descriptors used are patch based - SIFT and shape descriptors. On the other hand, the dense image representation is a widely used alternative to region detectors. In this case sub-image or the whole image can be described by feature descriptor computed for every image-pixel or sampled on a dense grid. The advantage is uniform textured regions are represented well in dense as compared to sparse. The sparse and dense representations along with the high level feature descriptors are illustrated in the Figure 2.7. The interesting points on any object is usually extracted from the training image to provide the feature description of that object, and subsequently used for training purpose. Table 2.5 gives a brief idea about the feature specifications on few of the segmentation papers.

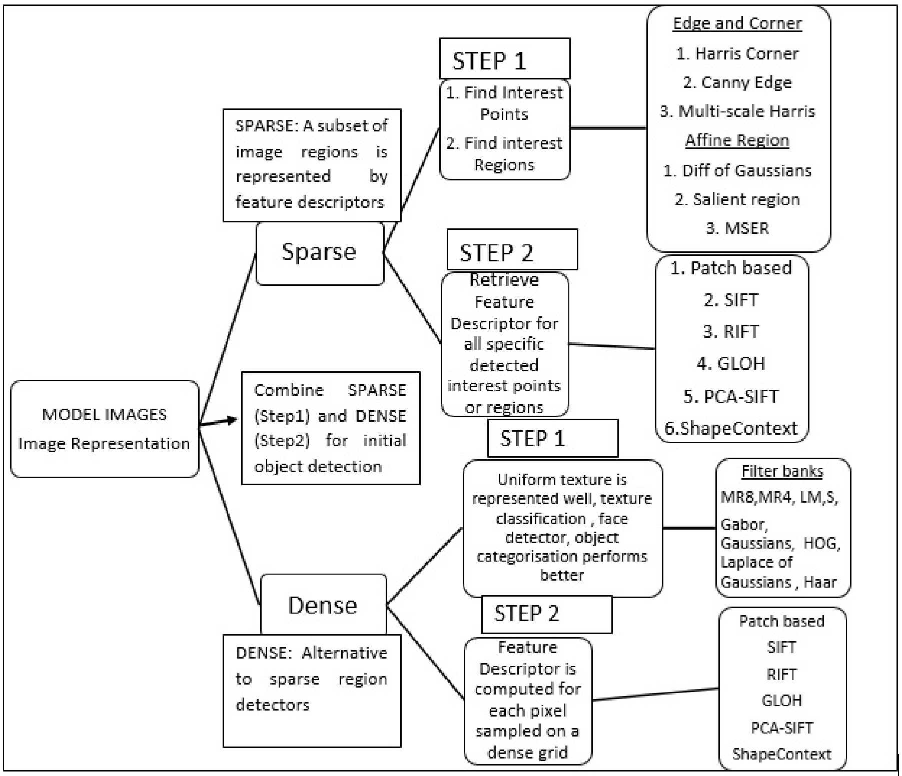


Figure 2.7: High Level Feature Descriptors

Table 2.5: Feature Specification

Research Papers	**Feature type**	**Feature Space**
Bhoyar and Kakde (2010)	Pixel	Color
Abbadeni (2011)	Pixel	Grayscale Texture
Shotton *et al.* (2009)	Pixel	Color, Texture, Layout, Location, Edge
Carreira-Perpinán (2006)	Pixel	Color
Ahuja and Todorovic (2007)	Pixel	Color, Shape, Area, Topology
Borenstein and Ullman (2004)	Patch	Color
Panagiotakis *et al.* (2014)	Pixel	Color, Texture
Sanda *et al.* (2013)	Pixel	Color, Texture
Belgiu and Drăgut (2014)	Pixel	Color, Shape
Felzenszwalb and Huttenlocher (2004)	Pixel	Color
Adamek and O'Connor (2006)	Pixel	Color, Shape
Xie and Mirmehdi (2007*b*)	Patch	Color, Texture, Shape
Ciurte *et al.* (2014)	Patch	Intensity

2.3.5 Image Representation

Image representation plays a crucial role in many computer vision applications in modeling human vision and also serves purpose of dimensionality reduction for enabling better analysis and understanding of the image contents. A recent review on image representation is done by Tian (2013). It represents the transformed feature space, from which the features need to be extracted. The representations can be further grouped as single-scale and multi-scale configurations as mentioned by Vantaram *et al.* (2012). The methods used to characterize and represent the features whether color, texture, shape etc are mentioned in this level. Tree structured image data representation gives a hierarchical overview of the image contents and have been found in papers by Chow and Rahman (2007), tree partitioning by Wang *et al.* (2008). Bag-of-visual-word representation has been in recent use for image annotation and retrieval as can be found in a paper by Yang *et al.* (2007). As our work primarily deals with color and texture features, we have provided a brief essential background on color representation in Appendix A. An elaborate survey on texture analysis and characterization can be found subsequently in Section 2.4. The concepts of texel, texem and texton are discussed in Section 2.4.1. The Table 2.6 gives a glimpse of image representation in the form of color space, texture, or data structure used for few papers.

Table 2.6: Representation Specification

Research Papers	**Representation**
Lee (1996)	2D Gabor Wavelet
Salembier and Garrido (2000)	Binary Partition Tree
Yang *et al.* (2007)	Bag-of-visual-word
Abbadeni (2011)	Autocorrelation
Ahuja and Todorovic (2007)	Texel
Borenstein and Ullman (2004)	Shape representation
Panagiotakis *et al.* (2014)	L*a*b*, Wavelet transform
Sanda *et al.* (2013)	Clifford algebra, Gabor filter
Felzenszwalb and Huttenlocher (2004)	Graph-based
Xie and Mirmehdi (2007*b*)	Texem

The effectiveness of usage of hybrid global and local feature representations along with the need for semantic representations are some of the interesting and challenging topics. The image and feature representations are highly dependent on each other.

2.3.6 Approach Specification

We propose to add a new high-level type for semantic based segmentation approach, apart from the already mentioned three approach types, in the earlier surveys as shown in Figure 2.8. These cater to high level computer vision and understanding tasks, such as scene classification, object recognition etc. With the emergence of semantic based segmentation, a good mix of supervised and semi-supervised modes of segmentation algorithms are developed to achieve high precision results. The present taxonomy does not provide a place to capture information regarding image representation aspects, feature descriptors and appropriate model for mid-level and high level analysis.

Recent techniques like granular computing such as by Liu *et al.* (2014) and rough sets by Rocio *et al.* (2014) have also been applied in the field of color-texture image segmentation. Information is captured in granules which reflects existing domain knowledge. The information granule is dependent on underlying formalism of granular computing. The details about granular computing, its characteristics and applications can be found in articles by Pal *et al.* (2015). Granular computing has been used to develop segmentation algorithm for remotely sensed images by Shankar (2007), for medical image by Roselin and Thangavel (2012), for natural images by Liu *et al.* (2014). Recent emerging computing paradigm of information processing such as granular computing and rough sets are also paving their way in achieving application independence in context with image processing and understanding. This new approach, named Granular and Rough computing paradigm, can be added to the taxonomy under image domain approach type.

Image segmentation approaches thus can be broadly grouped into four types -

1. Feature space based
2. Spatial coherence image domain, based on homogeneity or based on discontinuity of pixel properties
3. Physics based
4. Semantic based

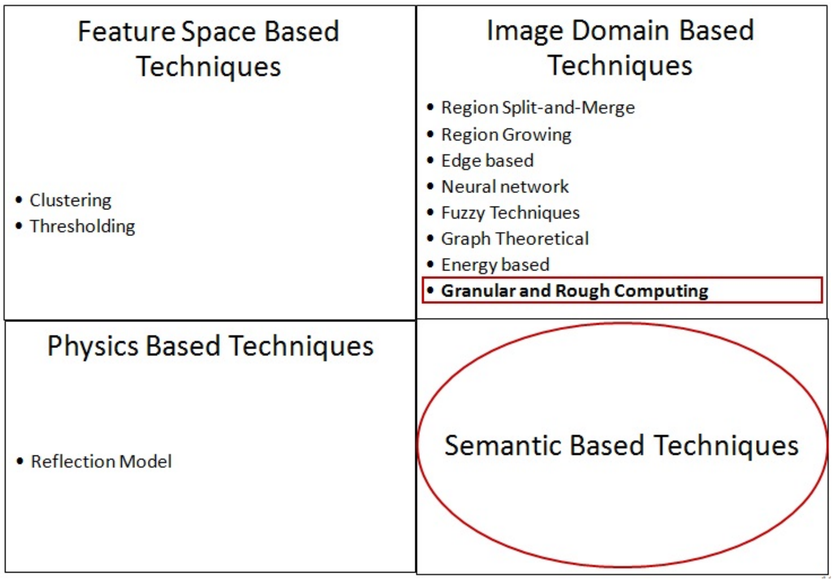


Figure 2.8: Image Segmentation Types and Sub-types

The literature on color image segmentation during early period till 2005 were mostly based on extension of gray level image segmentation approaches with different color feature spaces. Color images are considered to be a special case of multi-spectral image.

Surveys by Fu and Mui (1981), Haralick and Shapiro (1985), Shaw and Lohrenz (1992), Pal and Pal (1993), Reed and Dubuf (1993), Zhang and Tan (2002) give detailed descriptions of various monochrome image segmentation and texture characterization approaches. The surveys by Skarbek and Koschan (1994), Lucchese and Mitra (2001), Cheng *et al.* (2001), Vergés-Llahí (2005), Zhang (2006), Busin *et al.* (2008) give an insight about color image segmentation. The color-texture image segmentation is exclusively studied in surveys done by Trémeau *et al.* (2008), Dey *et al.* (2010), Ilea and Whelan (2011) and Vantaram *et al.* (2012).

Details about the segmentation techniques are briefly stated in Table 2.7. The advantages and disadvantages of each approach is given in Table 2.8. Homogeneity based approaches are called region techniques which consist of thresholding, region split and merge, region growing and clustering. Region based segmentation methods usually use clustering schemes extensively to group the coherent regions with similar features. Some popular techniques used are graph based clustering (methods such as NTP by Wang *et al.* (2008), NCut by Shi and Malik (2000), SC by Yu and Shi (2003), GBIS by Felzenszwalb and Huttenlocher (2004)), modeling probability density function using mixture model based clustering (methods such as DCM Nikou *et al.* (2010), GBMS Carreira-Perpinán (2006), mean-shift Dorin and Peter (2002)). The discontinuity based approaches refer to edge detection or contour techniques. The edge detection technique has been extensively utilized for monochrome image segmentation.

Considering the present scope of our work, we have discussed some of the popular and recent works on unsupervised image segmentation based on color-texture features, in the next section.

2.3.7 State-of-Art Techniques in Color-Texture Segmentation

In this section, we will discuss some of the state-of-art color-texture unsupervised segmentation algorithms which have been used for comparative analysis with our proposed segmentation approach. The Table 2.9 gives an insight about the segmentation approaches and corresponding techniques used for few of these state-of-art image segmentation algorithms.

Table 2.7: Details about the Segmentation approaches

Approach	Description
Clustering	Clustered regions are formed in feature space without any spatial cohesiveness. Feature vectors are iteratively assigned to the closest cluster center based on similarity measure. High similarity pixels are clustered in one group.
Thresholding	Regions are formed by identifying peaks, valleys or shapes from the corresponding histogram. Pixels are allocated to classes depending on the range of pixel values.
Split-and-merge	The initial image is considered as single region and attempts to divide it into uniform regions based on homogeneity criteria. Merges similar region to create a large homogeneous region
Region growing	Starts with a single or small set of pre-defined pixels and iteratively collects pixels based on homogeneity criterion. Regions formed are spatially connected and compact.
Edge based	Edge filters or operators are applied to image and the pixels are classified as edge or non-edge based on local discontinuity criterion. Edges are connected to obtain the object contours. Edge linking, model fitting are some of the techniques used
Neural network	Used for data clustering and classification. Capable of parallel processing and can incorporate any degree of non-linearity.
Fuzzy	Fuzzy membership function can represent the real-world uncertainty factors arising in feature extraction, image processing. It models the uncertainty using fuzzy operators and inference rules.
Graph Theoretical	Image is represented as undirected graph. Edges are weighted as a function of similarity between the nodes. Capitalize on various cost functions for global energy minimization to yield optimized segmentation. Least-cost path is defined as the best contour.
Energy based	Related to the field of probability theory and region characteristics are modeled through MRFs and also to variational approach using shape priors.
Granular and Rough Computing	Information is compressed and summarized in form of granules. Uses rough set theory as mathematical framework. The concepts of upper and lower approximation is mapped to an image region in terms of granules.
Hybrid	Combination of any of the above approaches

Table 2.8: Advantages and Disadvantages of the Segmentation approaches

Approach	**Advantages**	**Challenges**
Clustering	Easy to implement and good for classification	Need to know the number of clusters. Usually converges to local minima
Thresholding	Easy and popular for monochrome images	Selection of thresholds, image dependent
Split-and-merge	Applicable where region description is more appropriate such as textures	Meaningful regions may not be homogeneous due to varying illumination, color and texture
Region growing	spatial compactness, need initialization	sequential processing is computationally intensive with considerable memory requirements.
Edge based	Use local filter bank that is good for separating object contours. Less complex algorithms than the region based	Choice of a relevant filter, susceptible to noise
Neural network	Good for classification.	Contains very complex set of interdependencies. Training time is long.
Fuzzy	Good for representing uncertainty, each pixel has a membership value to which it belongs(region or boundary)	Selection of appropriate fuzzy membership function, involves high computation
Graph Theoretical	Supervised techniques using human interaction reduces complexity and proves to be good for developing tools for various imaging applications	Computational complexity is high.
Energy based	Image energy drives the active contour model to match the image features. Matches very well with the desired object contours.	High level of supervision with significant computational complexity.
Granular and Rough Computing	Information granulation is appropriate for object extraction, uncertainty can be handled for objects	Granules are context-dependent. Selecting appropriate granule size to capture the information correctly.
Hybrid	Exploits the strength of each approach	May increase computational complexity

Table 2.9: Approach - Few State-of-Art Techniques

Approach	**Feature:Technique**	**Research Paper**
Clustering	Grayscale:CLUSTER	Hoffman and Jain (1987)
Clustering	Color Texture:3D color histograms, multiscale perceptual representation	Mirmehdi and Petrou (2000)
Clustering	Color:Mean-shift	Dorin and Peter (2002)
Clustering	Color Texture: SOM Classifier, Gabor filters, Adaptive Spatial K-means	Ilea and Whelan (2008)
Clustering	Color, Texture: Superpixel segmentation, Huffman coding compression based Clustering and Texture Merging	
Thresholding	Grayscale: Bimodal distributions	Perez (1987)
Thresholding	Color:Histogram	Bhoyar and Kakde (2010)
Region growing	Color Texture: Multiscale region growing	Deng and Manjunath (2001)
Graph	Color, Texture: Normalized-cut, K-means clustering, DOOG filters	Shi and Malik (2000)
Graph	Efficient Graph based	Felzenszwalb and Huttenlocher (2004)
Graph	Interactive graph cuts-based	Boykov and Lea (2006)
Hybrid Region and Edge	Grayscale:Watershed Transform	Hill *et al.* (2003)
Hybrid Region and Edge	Color Texture: Oriented Watershed Transform and agglomerative clustering	Arbeláez and Fowlkes (2011)
Hybrid Energy and Region Growing	Color Texture : Modal energy of deformable surfaces	Krinidas and Pitas (2009)
Hybrid Graph and Region Split-and-merge	Color Texture : Region Adjacency graph, Markov Clustering and Region merging	Hedjam and Mignotte (2009)

A recent review Ilea and Whelan (2011) identified following three major *trends* in the extraction and integration of color and texture features -

1. Implicit color-texture feature integration which extracts the texture features from single or multiple color channels and the segmentation process usually includes coarse-to-fine strategy
2. Extract color and texture in succession as serial processes
3. Extract color and texture features on separate channels and then combine in segmentation processes.

Each of the three trends and the research works associated with these, mostly which have been mentioned in the survey by Ilea and Whelan (2011) is mentioned below.

(i) Color-Texture Feature Integration - First Trend

Under the first *trend* of color-texture feature integration algorithms, the algorithm developed by Panjwani and Healey (1993) is considered to be one of representative work on implicit color texture integration. The paper considers region based segmentation approach based on color Gaussian Markov Random Field (GMRF) considering spatial interactions within and between the color bands. Parameters are estimated by ML method. It is a two step segmentation process. At first step region splitting is done on the image iteratively in square blocks till a uniformity criterion is reached. In the second step region merge is done using agglomerative clustering for regions with similar characteristics. Experiments are performed on natural images but no comparative analysis is provided. Such implicit integration has been found in papers by Hoang *et al.* (2005), Shi and Funt (2007), Goswami *et al.* (2014*a*). Hoang *et al.* (2005) uses Gaussian color model (wavelength-Fourier space) for color space. Shi and Funt (2007) provides a compact color pixel representation based on quaternion which takes care of inter and intra channel relationships. Initially compressed feature vector is obtained from Quaternion-PCA by training procedure. Then K-means clustering takes place followed by merging of regions with similar texture characteristics. The number of clusters is user defined and the performance is highly dependent on the parameters. The experiments are performed on 12 images and has been visually evaluated. A popular integration scheme is referred in CTM, compressed based texture merging by Allen *et al.* (2008). CIE L*a*b* color space is used. Color-texture features are extracted at pixel level within a 7 x 7 window for each color plane. For computational ease, the feature vector dimension is reduced to fixed value 8 by PCA method. Coding-based clustering process is used for segmentation to accommodate image content defined by Gaussian mixtures. The authors analyzed the performance of their proposed segmentation algorithm and did a comparative study with other state-of-art algorithms mean-shift Dorin and Peter (2002) and normalized cut Shi and Malik (2000). The authors Dorin and Peter (2002) discusses the fact that feature based image analysis is a global natured paradigm which can represent the input reliably for performing low-level tasks supported by independent high-level information. Such approaches lead towards a more generic application independent segmentation goal. The significant features map to denser regions in feature space which form clusters. Thus segmentation is the process to delineate the clusters. But in case of lesser feature support for cluster information, additional domain specific spatial parameters and post processing guide towards more accurate results. In such cases nonparametric clustering methods are used to analyze randomly structured feature space which can be grouped into two - hierarchical clustering Jain and Dubes (1988) and density estimation. The hierarchical methods are computationally expensive and merging criteria is not very intuitive. Dorin and Peter (2002) proposes approach for mode detection and clustering based on mean shift feature-space procedure which depends on kernel bandwidth. Unlike the usual approach of focusing on local features and their variations in data, authors Shi and Malik (2000) aims at extracting the global characterization of the image. The segmentation is considered as graph partitioning problem and uses the normalized cut as the global criterion for image segmentation. Normalized cut is measure of disassociation among subgroups of a graph computed by generalized eigenvalue system. On a 100 x 120 image it takes about 2 minutes on Intel Pentium machines.

(ii) Color-Texture Feature Integration - Second Trend

Under the second *trend* of color-texture feature integration algorithms, color and texture are not considered to be dependent in image formation process and their extraction is considered as serial process. Segmentation takes place as multi-phase approach considering coarse-to-fine partitioning technique. One such representative segmentation algorithm is proposed Mirmehdi and Petrou (2000). The author introduces multi-scale perceptual image tower emulating the human perception of looking at the image from different distances. A weighted sum of Gaussian kernels are applied to each color plane to extract the texture. Further in the next stage core color clusters are extracted followed subsequently by probabilistic segmentation process that assigns pixels hierarchically starting from coarsest to finest high resolution. Experiments have been performed on Vistex mosaic images. The results have been compared with Ma and Manjunath (1997). Another widely popular segmentation method belonging to this trend is the one proposed by Deng and Manjunath (2001). CIE L*u*v* color space is used. Color quantization and spatial segmentation are done in succession. The algorithm enforces spatial coherence of class labels using J value criterion. Segmentation is defined as multi-scale region growing process. Experiments are done on 2500 images from Corel database without parameter tuning. Krinidas and Pitas (2009) introduced Modal Image Segmentation (MIS) based on deformable energy function. Coarse image representation is done using color quantization. Agglomerative merging is applied at the final stage. Hedjam and Mignotte (2009) proposed Hierarchical graph-based Markovian Clustering (HMC), where initially image is segmented with fixed large number of K classes using MRF and subsequently the image is modeled as Region Adjacency Graph(RAG). Clusters are formed using edge linking.

(iii) Color-Texture Feature Integration - Third Trend

Under the third *trend* of color-texture feature integration algorithms, feature integration extract color and texture features on independent channels. Both color and texture features are separately modeled and their contributions are optimized during integration in the segmentation process. The authors Ilea and Whelan (2008) proposed color-texture segmentation framework referred as CTex. The first step is color segmentation which involves filtering using anisotropic diffusion to remove noise and better the color coherence. Dominant colors are first extracted to identify the number of clusters using unsupervised Self Organizing Map (SOM) network. The texture features are separately extracted from luminance component based on Gabor filters. The color and texture are then integrated using adaptive spatial K-means framework. CTex is experimented on Berkeley natural images dataset and compared with Deng and Manjunath (2001) quantitatively using only probabilistic rand index(PRI) measure. A hybrid approach of region and edge based image segmentation model is proposed by Mignotte (2012). A novel idea of de-texturing step is used at pre-processing step where it converts texture image into a noisy color image which at a later stage is efficiently regularized. This simplifies the clustering based segmentation model. The L*a*b* color space is considered. For feature extraction model, a texton is characterized based on color and gradient magnitude histograms for a window size 7. The number of classes for k-means algorithm is adaptively set for each image depending upon the complexity metric computed for an image. Bhoyar and Kakde (2010) proposes Just Noticeable Difference histogram based color segmentation and has done a quantitative evaluation using PRI metric. Modeling probability density function of pixel attributes such as intensity and texture with mixture models is one common way to cluster data. Authors Nikou *et al.* (2010) propose hierarchical Bayesian model which imposes spatial smoothness constraints for mixture model-based segmentation. It assumes the mixing proportions to follow Dirichlet compound multinomial distribution. The model is applied to Berkeley images and compared to Gaussian blurring mean-shift algorithm in Carreira-Perpinán (2006) and normalized cut Shi and Malik (2000).

The evolution of segmentation approaches is tightly coupled with, one of the most widely used visual attributes (apart from color) - texture in many areas of image analysis and pattern recognition. The various concepts of texture found in literature, the methodologies for texture extraction and characterization has been taken up separately in the forthcoming section of texture analysis.

2.4 Texture Analysis

Texture plays a very important role as special high level feature in the area of image processing, analysis and computer vision. The texture analysis is the basis of many image engineering tasks starting from low level texture characterization, texture segmentation to high level image processing such as texture classification, texture based object detection. The definition of texture is ill-defined as mentioned in paper by Zhang and Tan (2002), and hence the scope in which the texture is defined affects the choice of the texture analysis methods. This has been an active research topic since many decades, initially started with gray-scale textures and moved towards hybrid combination of color-texture features. We have referred Reed and Dubuf (1993), Zhang and Tan (2002) surveys for gray-scale texture and Ilea and Whelan (2011), Vantaram *et al.* (2012) for color-texture. A texture consists of primitives of varying sizes, granularity, orientation, pattern (regular or random). Images with multi-textural primitives are very difficult to model. Natural images consist mainly of such scenarios, for which new concepts have emerged to capture the not so regular patterns for texture based segmentation, which are discussed in next section.

2.4.1 Concepts : Texels, Textons, Texems

Image segmentation has gradually taken a new turn from a unsupervised image segmentation approach towards a semi-supervised learning approach for improving the segmentation accuracy in terms of human perception and in a meaningful manner. The sophistication inculcated in the segmentation algorithm use relatively new concepts such as texels mentioned by Todorovic and Ahuja (2009), textons mentioned by Zhu *et al.* (2002) and texems mentioned by Xie and Mirmehdi (2007*a*) to capture the smallest textural variations present in the image. This takes one step towards high level image understanding rather than keeping the segmentation process as simple as in the low level processing. Texel, abbreviation of TEXture ELements can be considered as the smallest area present in a specific textured region which repeats itself in some regular fashion.

Three texel properties as described in Todorovic and Ahuja (2009) are : 1) geometric (substructure) and photometric 2) structural 3) placement. Texton is more sophisticated than texel. It is usually learned from training images serve as a mini-template consisting a number of image bases having geometric and photometric configurations. Texton is still considered to be a vague concept and sound mathematical foundation has not yet been laid. Texem is abbreviation of texture exemplar. Texems are exemplar image patches. A texem model is developed as a mixture representation for color images using multi-scale analysis and by capturing neighborhood pixel interactions. The segmentation using texems posed as a potential tool in providing psycho-visual semantic perception of the scenic image. Our work on texture is presently scoped on statistical measure based on OP and ANOVA; rather than on the concepts of texel, texton and texem, which are based on aspects such as topology, geometric and semantics. The next section describes the various early and recent groups of methods for texture feature extraction, alongwith their characteristics .

2.4.2 Texture Feature Extraction Methodologies

Texture segmentation mostly was considered to be an independent sub-process for texture feature extraction followed by segmentation algorithm as mentioned in a survey by Reed and Dubuf (1993). Texture feature extraction has been grouped as 1) Feature based, 2) Model based and 3) Structural as shown in Figure 2.9. Model is considered as sub-classification of feature based as the model parameters are basically texture features. Structural based methods assumes that there exists primitive elements and hence its usage is very limited.

An elaborate survey paper on invariant texture analysis is presented in Zhang and Tan (2002). The basic texture analysis can be grouped into three broad methodologies - 1) Statistical methods 2) Structural methods 3) Model based methods according to Zhang and Tan (2002). The paper primarily focuses on invariant texture analysis which has received quite an attention as it has proved to be highly desirable from the aspect of

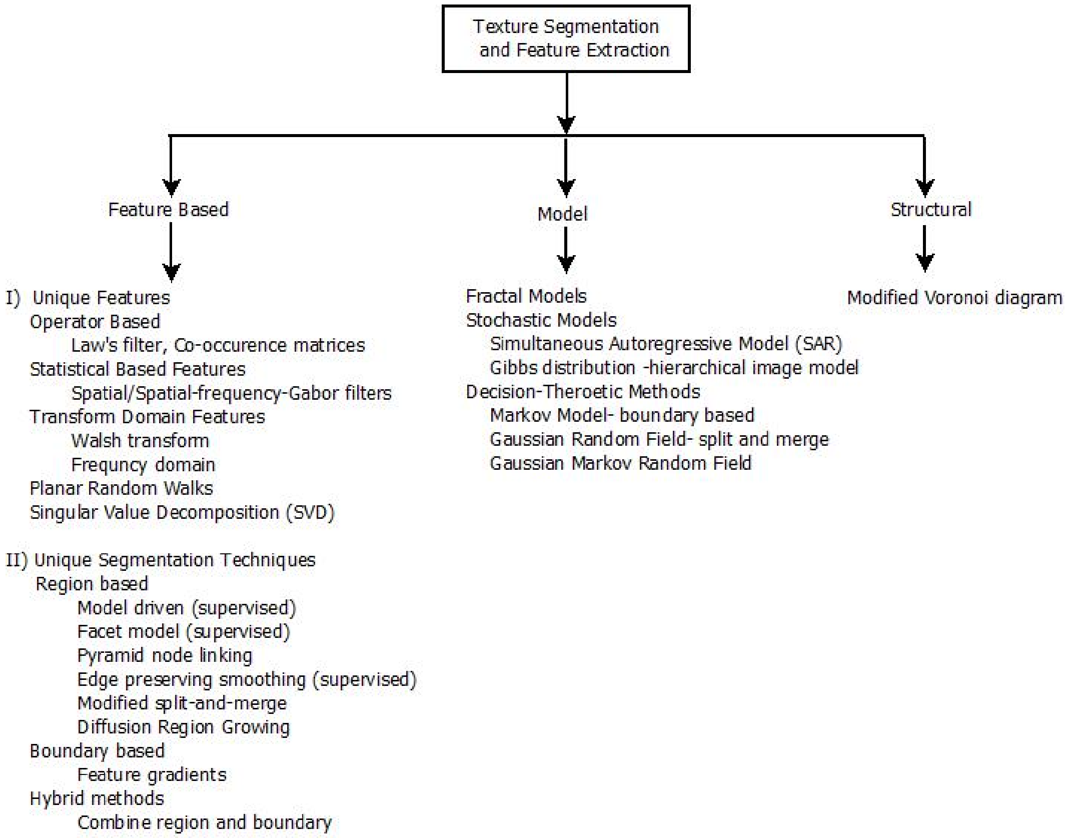


Figure 2.9: Taxonomy of earlier approaches for texture feature extraction -1993 (Reed. et. al)

automating a human observer. The grouping of texture analysis though remains same as a decade ago mentioned in Reed and Dubuf (1993). Nevertheless, many new and recent texture segmentation and texture analysis methods have been mentioned in this survey. In most of the recent work, the model coefficients are transformed to achieve the invariance. Based on both of these surveys and some recent work found in the literature we have presented texture analysis methods and their corresponding techniques in Figure 2.10. Characteristics of each of the texture analysis methods is mentioned in Table 2.10 in terms of its key application areas, advantages and disadvantages.

Earlier texture analysis methods mainly focused on first-order and second-order statistics for textures. The trend moved towards textured image being modeled using probability or a linear combination of a set of basis functions. In case of natural images,

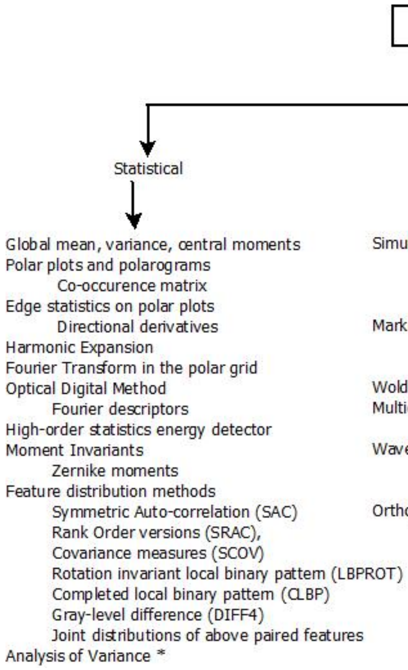


Figure 2.10: Recent Texture Analysis Methods

Table 2.10: Characteristics of the Texture Analysis Methods

Methods	**Application Area**	**Advantages**	**Disadvantages**
Statistical	Image classification, texture discrimination	Applicable to all images in an efficient manner based on first order, second order and higher order statistics	Less intuitive, prior information required to adequately characterize textured regions
Model	Texture classification, segmentation and synthesis	Powerful tool mostly based on probability model where model coefficients used for texture characterization	Estimation of model coefficients is challenging, choice for the correct model has to be made for a selected texture
Structural	Texture classification	Structural primitives texels based on usually multi-scale segmentation, suits for macro-texture	Limited to synthetic or regular textured image patches

it is almost impossible to do a structural textural analysis due to its inherent randomness such as in grass, leaves. The paper Todorovic and Ahuja (2009) showed a mixture of results for both mosaic textured images and real-world images. The author claims that their segmentation is successful on low-contrasted regions. Quantitative evaluation on Berkeley dataset has not been evaluated by the authors as its annotation is meant to be only for object segmentation. Hence natural image segmentation performance cannot be compared with any paper related to structural based texture segmentation methods. Feature distribution methods use local window statistics and hence is not in a position to capture the overall global directionality. These methods do not fair well for strongly ordered texture and is used for classification purpose. In the thesis, we have proposed ANOVA based texture characterization Goswami *et al.* (2014*b*) which does global and local texture characterization.

One of the key areas of focus is invariant texture classification. In such cases statistical methods of texture analysis are applicable as the texels are statistically similar and their spatial arrangement too is statistically uniform, which makes it more flexible

Table 2.11: List of Invariant Texture Analysis Methods

S.N.	Invariant Perception	Texture Analysis Methods
1	**Rotation Invariance**	**Statistical methods**: Global mean, Variance, Central moments, Polarograms, Texture edge statistics, Harmonic expansion, Zernike moments, Feature distribution methods **Model based** : RISAR, HMM, Polar 2-D Gabor filter, Daubechies four-tap wavelet filter coefficients **Structural method** Perimeter and Compactness of a primitive
2	**Scaling Invariance**	**Statistical methods**: Fourier transform on polar grid **Structural method**: Roundness or compactness , Iterative morphological decomposition(IMD)
3	**Rotation Scale Invariance**	**Statistical methods**: Optical digital method of Fourier descriptors
4	**Translation Invariance**	**Statistical methods**: Global mean, variance and central moments, High-order statistics energy detector, Zernike moments, Fourier spectrum **Structural method**: Roundness or compactness

and reasonable for texture based image processing. The list of invariant texture analysis methods have been separately listed in Table 2.11.

Texture classification methods use classifiers such as multivariate linear discriminant function, neural nets, K-nearest neighbor classifier, BP-neural network, rule-based network classifier. Few of the statistical as well as model based methods have orthogonality such as Zernike polynomials, Harmonic components, Wold-like model, Orthogonal polynomial operators. Wold-like model is broken down into three mutually orthogonal components; as mentioned by Zhang and Tan (2002). Some of the popular state-of-art techniques for texture analysis are mentioned in the following section.

2.4.3 State-of-Art Techniques in Texture Analysis

The texture analysis techniques have been used in several domains such as segmentation, content based retrieval, classification and shape formation from texture. Various methods of texture analysis and characterization are reported in literature. Haralick's co-occurrence matrix method is one of the early attempts for global description of textural images. The features are derived for the purpose of texture classification. Edge detection technique guides the concept of texture analysis, texture discrimination and texture based segmentation. Texture perception has three high level features - repetition, orientation and complexity as mentioned in Rao and Lohse (1993). Some of the widely used texture operators are Gabor filter design as mentioned in papers by Jain and Farrokhnia (1991), Manjunath and Ma (1996), statistical texture primitive - local binary pattern (LBP) operator by Ojala *et al.* (2002), DCT texture features by Gonzalez and Woods (2007). Several papers have also addressed the evaluation of texture-only based segmentation algorithms by Ilea *et al.* (2010), Todorovic and Ahuja (2009), based on multi-scale hierarchical approach. Most of the gray scale texture operators have been directly applied on color images by either combining gray level texture features with color features or they derive the texture features by computing separately on each of the color channels. Signal processing based texture analysis use a wide range of such filters namely Gabor mentioned by Gabor (1946), Gabor wavelet by Manjunath and Ma (1996). Model based method like Gaussian Markov random field models, mentioned in paper by Chellappa *et al.* (1999) is also popular to describe texture. Tamura *et al.* (1978) proposed textural features using six computational forms which correspond to human visual perception. Ahuja and Rosenfeld (1981) compared and summarized mosaic models and statistical texture models. They favored the mosaic model as they provide a hierarchical character for image modeling of textured images. The authors Vaisey and Gersho (1992), define texture as region with high luminance variance. Ganesan and Bhattacharyya (1995) proposed texture representation using positional number system for small regions and has been further used for classification. Experiments were conducted on Brodatz textural album for grayscale natural textured images. Rao and Lohse (1996) proposed a texture naming system that classifies textures into meaningful, hierarchical structured groups. Won and Park (1997) used simple block transform based coding schemes for detecting textures. Ojala *et al.* (2002) makes use of Local Binary Pattern (LBP) which is a grayscale invariant texture primitive statistic. Ahuja and Todorovic (2007) identifies basic elements of image texture called texture elements or texels. The paper proposed texel extraction using graylevel contrasts specifically for homogeneous thin planar structured objects based on learning algorithm. The method is applied for texture segmentation. Abbadeni (2011) proposes a perceptual model to estimate textural feature set using psychometric method. The computational features are applied for content based image retrieval on grayscale images. From the last couple of decades the texture analysis has been extended to color images. The authors Mäenpää and Pietikäinen (2004) performed three experiments with two different texture sets and compared the color indexing methods to grayscale and color texture methods. Parametric stochastic models have been proposed for color texture characterizations which are basically extension of graylevel. Panjwani and Healey (1995) provided Markov random field model for color textures in terms of spatial interaction within and between the color planes. A recent survey of methods and applications for automatic characterization can be found in Gonzälez *et al.* (2013) with special focus given to visual appearance of materials used in industrial applications such as granite, fabric, ceramic etc. Sanda *et al.* (2013) proposed Clifford Algebra and Gabor filter for color image texture characterization to represent multi-component images and Wouwer *et al.* (1999) presents a scheme for colored texture image classification based on wavelet correlation signatures containing the energies of each color plane and cross correlation between different planes. Most of the texture characterization is done either keeping a specific categorical data (synthetic or natural images) or is application specific (segmentation on specific benchmark dataset or image retrieval). Our paper by Goswami *et al.* (2014*b*) proposes a novel method of texture characterization based on inferential statistics called ANOVA, which is not application specific. MFHIST can be further refined to capture the texture specific aspects, concepts in future. The next section gives a demonstration of MFHIST being filled with details for some of the popular works.

2.5 Illustration demonstration of MFHIST

We have considered few research papers on image processing, focused mainly on related areas of segmentation and characterization. We show how to use the proposed MFHIST for all these papers in the Figure 2.11. The first aspect being considered after going through a paper on segmentation related work is the scope. Taking an example, the objective of the paper by Goswami *et al.* (2014*a*) is low level image processing related to unsupervised segmentation. The scope column is filled first with *Low Level*. The next details to be filled in the taxonomy is related to the requirement specifications which are filled up as follows : Domain-*General*, Source of images taken- *BSD natural images* and the Precision level acceptance- *Medium* respectively. The third category level to be filled up captures the control specification in two aspects - *Driven By* and *User Interaction*. The paper proposes unsupervised segmentation, hence the control is driven by *Data driven* and the user interaction aspect is mentioned as *Unsupervised*. The feature specification is captured next into the MFHIST based on feature type - *Pixel* and the feature spaces considered - *Color, Texture*. The next category level tells about the image representation facet in terms of any transformation and data struc-

Paper	Scope	Requirement Specification			Control Specification		Feature		Representation	Segmentation Approach
		Domain	Source	Precision	Driven By	User Interaction	Type	Space	Transformation and Data Structure	
1975-1985										
Haralick et. al. (1973)	Mid Level	Sandstone, Land-use	Aerial	Medium	Data Driven	Training	Pixel	Grayscale Texture	-	Statistical texture analysis
Nevatia (1977)	Low Level	General	Few Images	Medium	Data Driven	Unsupervised	Pixel	Color	-	Edge based
Ohlander et. al. (1978)	Low Level	General	Few Natural , Aerial	Medium	Data Driven	Unsupervised	Pixel	Color, Grayscale Texture	-	Histogram threshold, Edge texture operator
Ohta et. al. (1980)	Low Level	General	Few Images	Medium	Data Driven	Unsupervised	Pixel	Color	Karhunen Loeve Transform	Spatially Blind Region based
1985-95										
Perez (1987)	Low Level	Application Specific	Few Images	Medium	Data Driven	Unsupervised	Pixel	Grayscale	-	Threshold
Thorpe et. al. (1988)	High Level	Application Specific	Real-world	High	White -board	Mobile Robot	Pixel	Color, Texture	Hough Transform	Robert edge based texture
Jain and Farrokhnia (1991)	Low Level	Natural and Artificial Textures	Brodatz	Medium	Data Driven	Unsupervised	Pixel	Grayscale Texture	-	Texture Segmentation - Gabor Filters
Panjwani and Healey (1995)	Low Level	General	Natural images	Medium	Data Driven	Unsupervised	Pixel	Color, Texture	-	MRF modeling, Clustering
1995-2005										
Shi and Malik (2000)	Low Level	Static and Motion	Natural images	Medium	Data Driven	Unsupervised	Pixel	Color, Texture	Weighted Graph	Graph-based DOOG filters
Comaniciu and Meer(2002)	Low Level	Parameter tuning	Real, Synthetic	Medium	Data Driven	Unsupervised	Pixel	Grayscale, Color	Region Adjacency Graph	Spatial Guided, Mean Shift Clustering
Hill et. al. (2003)	Mid Level	Application specific	Few Images	Medium	Data Driven	Unsupervised	Pixel	Grayscale Texture	Watershed Transform	Hybrid Region and Edge
Borenstein and Ullman (2004)	High Level	General	Real-world House, Car, Face	Medium	Hybrid	Training	Patch	Color	Shape Representation	Figure-Ground Class based segmentation
2005-2015										
Adamek and O'Connor (2006)	Mid Level	Application Specific	Real-world	Medium	Model Driven	Supervised	Pixel	Color, Shape	Binary Partition Tree	Interactive Object Segmentation
Ahuja and Todorovic (2007)	Mid Level	Application Specific	Real-world 2.1D	Medium	Data Driven	Unsupervised Learning	Pixel	Color, Shape Area, Topology	Segmentation Tree	Texel Detection, Segmentation – Region Topology
Goswami et. al. (2015)	Low Level	General	BSD Natural	Medium	Data Driven	Unsupervised	Pixel	Color, Texture	Hybrid Color Space (HCS)	Hybrid Region and Edge OP-HCS
Goswami et. al. (2014b)	Mid Level	General Natural images	Vistex, Corel, BSD	Medium	Data Driven	Unsupervised	Pixel	Color, Texture	-	ANOVA Statistical Learning

Figure 2.11: Populating MFHIST with papers from four decades

ture involved. The paper proposes Hybrid Color Space transformation. Finally the last column regarding the segmentation approach type is filled up. The paper proposes segmentation technique which falls under hybrid region and edge (under image domain type). MFHIST provides broader coverage about a paper in a systematic and concise manner, which can be further detailed out, if required.

2.6 Summary and Conclusions

The survey trend shows that in the past four decades, the image segmentation has moved from application specific techniques to more generalized approaches that strive for application independence. The hybrid combinations seems to be the future in the area of image segmentation whether in the feature space or in the segmentation technique itself. The Figure 2.12 illustrates the gradual shift towards more sophisticated algorithms, texture models getting more general and matured, giving more attention to invariant and 3-D texture analysis. The focus on optimal color-texture integration along with appropriate feature extraction helps in achieving the application independence, as shown in Figure 2.13.

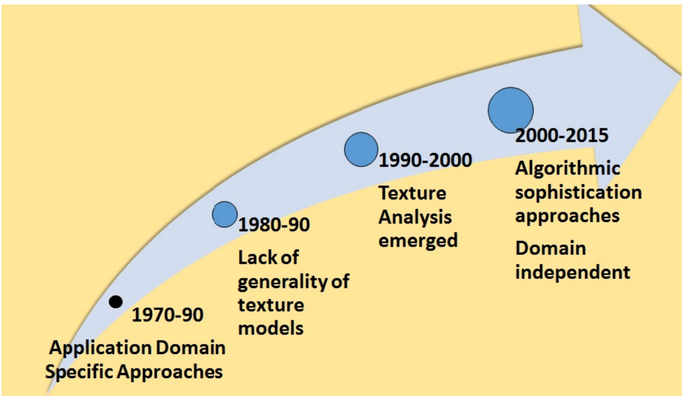


Figure 2.12: Evolution of Image Segmentation

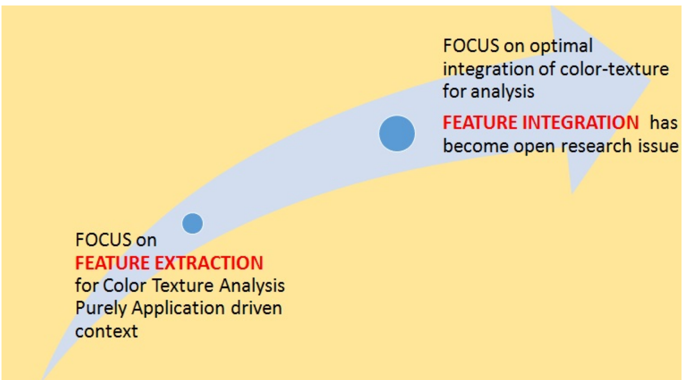


Figure 2.13: Color-Texture Transition Analysis

An abundance of various automated and semi-automated techniques can be found in literature, that cater to wide range of image analysis and image understanding applications. The present taxonomy found in the literature, is usually found to be focused on the segmentation technique categorization, rather than portraying the overall aspects of its purpose and multiple segmentation facets involved in the process. The proposed taxonomy MFHIST fills in this gap and assists the researchers in categorizing their works according to scope specification catering to low-level, mid-level and high-level image processing; and other facets such as requirement, control, feature, representation, approach specifications. Recently, the hybrid combinations in control specification, feature specification and approach specification are in practice as they give more robust results taking advantage of the combinations. The emerging segmentation approaches and techniques based on new paradigm can also be further mapped to approach specification types. MFHIST does not presently address the texture analysis in details, and can be taken up further in future.

From the survey, it is evident that in spite of abundance in segmentation algorithms, not a single algorithm can segment all images to a satisfactory level, especially in cases of natural outdoor scenes having large variation in color-texture mix. As a part of low-level image processing, the image segmentation is heading towards completely automatic approaches. Segmentation, being a psychophysical perception expects the segments obtained to be significant, in terms of the application that is using the segmentation results for high level understanding. Thus the expectation of being able to automatically extract meaningful segments from all types of images, irrespective of any specific application in mind, is a challenging task. Sophisticated/complex algorithms are applicable for highly textured images. Such algorithms are an over-kill for simple tonal/color images. For this purpose, we build a framework for color-texture segmentation and characterization of mostly, natural outdoor images. The framework also proposes image types as one of the useful metadata, based on statistical analysis. The segmentation and characterization complement each other and generate useful metadata, which can help in representing and manipulating the uncertainties at a later stage of the vision and understanding system, in an application independence perception. As the present study confines to mid-level framework component and produces appropriate knowledge representation in form of metadata and string codes; the work on image understanding and analysis can be taken up in future.

CHAPTER 3

Color-Texture Integration for Image Segmentation

An extensive literature study on exclusive case for unsupervised color-texture image segmentation methodologies can be found in Chapter 2 - Literature Survey. The Chapter proposes a generic color-texture feature integration framework. The contribution considers both color and texture as a joint phenomenon upon which segmentation has been found quite effective. The image segmentation considers color and texture as mutually dependent attributes The edge based approach has been exploited in the methodology in capturing the textural variation or discontinuities present in the natural images. The grouping of coherent regions with similar color-texture features can be further achieved using clustering which is homogeneity based region technique. The proposed image segmentation for natural images is unsupervised and falls into the category of hybrid region and edge based segmentation technique.

The rest of this chapter is organized as follows. Section 3.1 gives a brief overview of our proposed segmentation methodology. Section 3.2 presents proposed Hybrid Color Space(HCS). Section 3.3 emphasizes the role of OP in texture understanding. It gives a brief overview of two types of orthogonal polynomial operators (OP3 and OP5) and their role in texture extraction. Section 3.4 explains color-texture feature integration and adaptive feature vector representation for proposed two variants - OP3-HCS and OP5-HCS. Section 3.5 discusses the proposed image types, iterative region splitting segmentation algorithm and metadata generated in the process. Section 3.6 describes spatially constrained region merge technique. Section 3.7 summarizes and concludes with future work.

3.1 OP-HCS Segmentation

We have proposed generic hybrid color-texture feature integration methodology, based on tensor products obtained from orthogonal polynomials (OP) applied on hybrid color space (HCS). The orthogonal polynomial operators have been previously employed as edge detector by Ganesan and Bhattacharyya (1997), texture analysis by Ganesan and Bhattacharyya (1995) in grayscale image, as color edge detector by Krishnamoorthi and Bhattacharya (1998), in image coding technique by Krishnamoorthi and Kannan (2009), and in texture discrimination by Suguna and Anandhakumar (2011). Ganesan and Bhattacharyya (1997) proposed a framework based on a class of orthogonal polynomials, that has been used to generate a complete set of difference operators. One specialty is that, in case of untextured images, the operator separates out both the responses of Gaussian noise and edges found in the image. Two statistical descriptors called pronum and prospectrum are computed based on polynomial operators of size 3, for detecting presence of both textures and edges present in the grayscale image. The paper presents the simple computational procedure for constructing the set of difference operators, some of which are widely known edge detectors such as Roberts, Prewitt, Sobel and Marr's LOG edge operators, details of which can be found in Goswami (2011). Ganesan and Bhattacharyya (1995) describes the texture characterization based on orthogonal effects, which is a combination of main and interaction effects. Krishnamoorthi and Bhattacharya (1998) proposes a framework for color edge detector in terms of orthogonal polynomials, and is based on identifying the zero crossings in second derivatives of the color image region. The authors have reported experimental results with few untextured color images. Krishnamoorthi and Kannan (2009) presents a novel coding scheme using orthogonal polynomials on monochrome images. Suguna and Anandhakumar (2011) proposes a feature extraction for texture discrimination. The proposed approach builds a model for texture class and texture classification based on orthogonal operators of size 5. Experiments were conducted on grayscale Brodatz album.

There is no repeated work found in literature that has experimented on OP-HCS for color-texture feature integration. The OP-HCS segmentation process can be seen in Figure 3.1. The proposed feature extraction model outputs tensor responses. The integration is done by performing multidimensional filtering of the hybrid color planes with the bank of 9 or 25 tensor products which act as filters, depending on the OP size chosen as an external parameter setting. The feature space data is clustered repeatedly into two clusters using region based k-means clustering. Further segmentation is controlled by couple of factors; firstly cluster accuracy test which is assessed by two sample Kolmogorov-Smirnov test(as mentioned in paper by Massey (1951)) on histogram analysis of tensor responses for each of the two clusters, secondly further segmentation is stopped when very small sized cluster or maximum number of clusters is reached. The initial segmentation region map favors over-segmentation. In order to guide the segmentation to accurate partition, spatially constrained region merge technique based on Mahalanobis distance similarity measure is implemented. The contributions of this chapter has been published by Goswami *et al.* (2014*a*) and Goswami *et al.* (2015). The next section discusses our proposed HCS.

3.2 Hybrid Color Space

Color space seems to be a natural feature space as it tends to form clusters. The challenge lies in selection of an appropriate color space which can segment the image appropriately.

Busin *et al.* (2008) mentions in the paper that a specific color space is not suitable for all the image segmentation problems. Researchers have experimented with more than one color spaces to exploit the specific characteristics such as - two color spaces have been used by Mirmehdi and Petrou (2000), six by Hedjam and Mignotte (2009) to exploit the specific characteristics of color spaces. Mirmehdi and Petrou (2000) used two color spaces at different stage of segmentation. The kernels are applied on opponent color space $O_1O_2O_3$, and subsequently the image has been converted to CIELuv. The authors Hedjam and Mignotte (2009) report using six color spaces RGB, HSV, CIELUV, YIQ, XYZ and CIELAB. The color textural measure is done for each channel

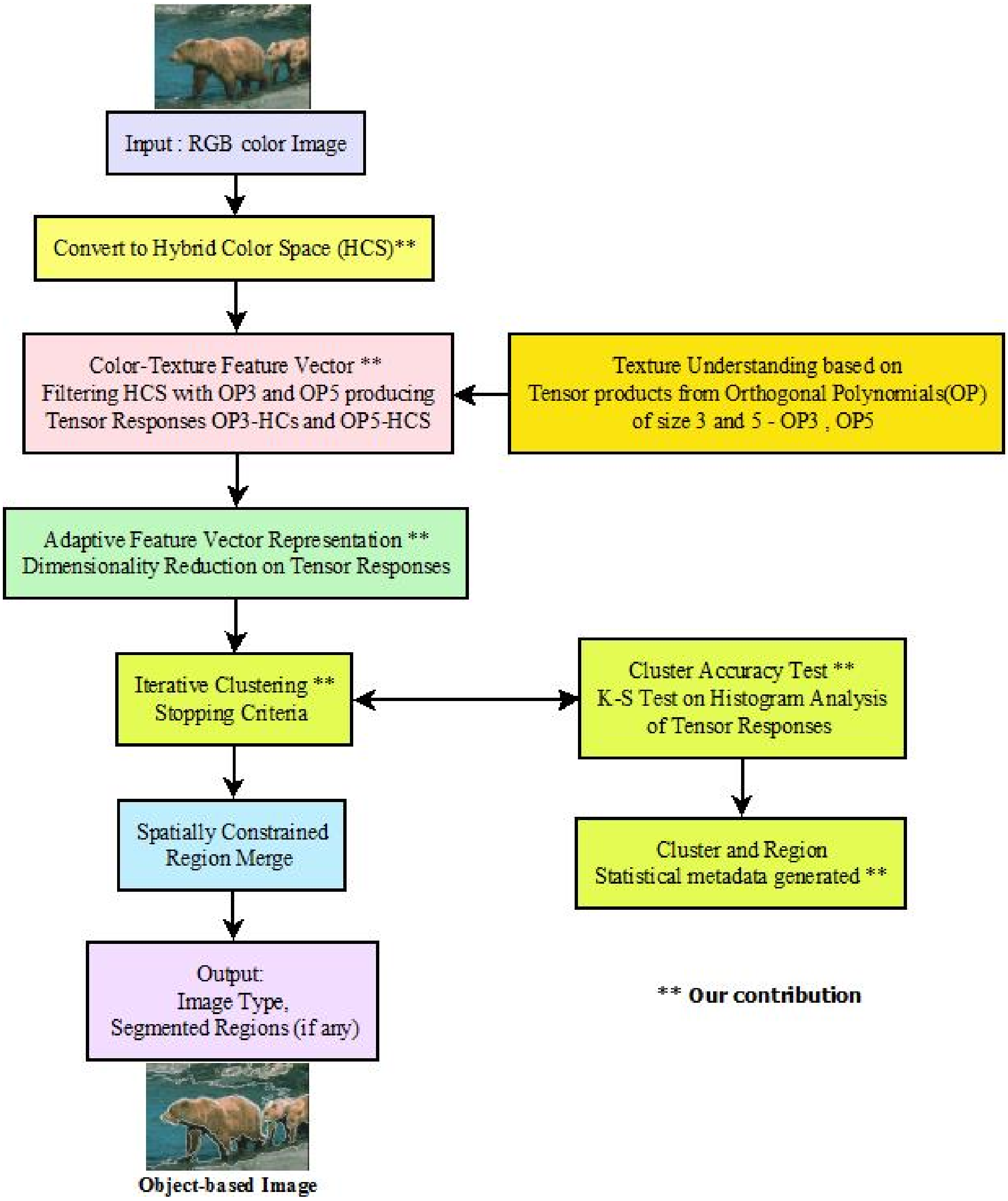


Figure 3.1: Proposed Segmentation Framework

of the color spaces that has been normalized between 0 and 255. In paper by Ilea and Whelan (2008), filtering is applied separately to RGB and YIQ. The paper by Mignotte (2010) proposes Probabilistic Rand Index Fusion (PRIF) which fuses 60 segmentations obtained from ten different color spaces (RGB, HSV, YIQ, XYZ, L*a*b*, LUV, $I_1 I_2 I_3$, $H_1 H_2 H_3$, YC_bC_r, TSL), 3 different number of classes and 2 bins for color histogram. The researchers are using the color spaces to suit their application needs. This ad-hoc approach is thus subjective in nature.

The motivation behind proposing a new hybrid color space is its objective nature and it has couple of aspects embedded in it - color purity and color space with highest number of dimensions (CMYK-subtractive color space). The color purity is obtained from chrominance dominant color components a*, b* and H, S taken from perceptual spaces L*a*b and HSV. HSV values are normalized and rounded to have a fixed range between 0 and 255. CMYK, the only color space represented in four higher dimensions is also included in HCS. The actual process for subtractive color reproduction is complicated and cannot be comprehensively modeled by any equation as observed by Trussell *et al.* (2005). Hence these systems are usually characterized by look up tables to capture their input-output relationships empirically.

Algorithm 1: Conversion of RGB Image to HCS

Input: Img_RGB
Output: Img_HCS
$P\_1D\_Img \leftarrow Reshape3D\_Img\_RGB$
$Transform\_To\_Lab$
$ab \leftarrow Extract\_a\_b\_Planes$
$Transform\_To\_HSV$
$hs \leftarrow Extract\_H\_S\_Planes$
$cmyk \leftarrow Transform\_To\_CMYK$
$imageData\_8D \leftarrow [(ab)(hs)(cmyk)]$
$Img\_HCS \leftarrow PCA(imageData\_8D, ReduceTo3D)$
return Img_HCS

The method considered here for conversion from RGB to CMYK is a non-linear process unlike the conventional simpler version of conversion, i.e. $RGB = (1 - CMY)$. The Algorithm 1 describes the steps for conversion from RGB to HCS. The implementation is done using MATLAB - Image and Signal processing toolbox, which is widely used programming development by research community in the area of digital image processing. All the three representations from color spaces L*a*b*, HSV and CMYK increase the dimension of the color dataset from 3(R,G,B) to 8(a,b,H,S,C,M,Y,K). In order to reduce and approximate the color data, PCA one of the most widely used dimensionality reduction technique is applied. Applying PCA to the 8 components helps in determining the independent axis components which have non-redundant and uncorrelated information. Latent is a vector returned by the PCA containing the eigenvalues of the covariance matrix obtained from all the 8 color dataset. The dimension of the color feature dataset is reduced to three using PCA and we refer to it as Hybrid Color Space (HCS). Color is a 3-dimensional entity, hence the color feature dataset is reduced to three, which also makes it easier to visualize.

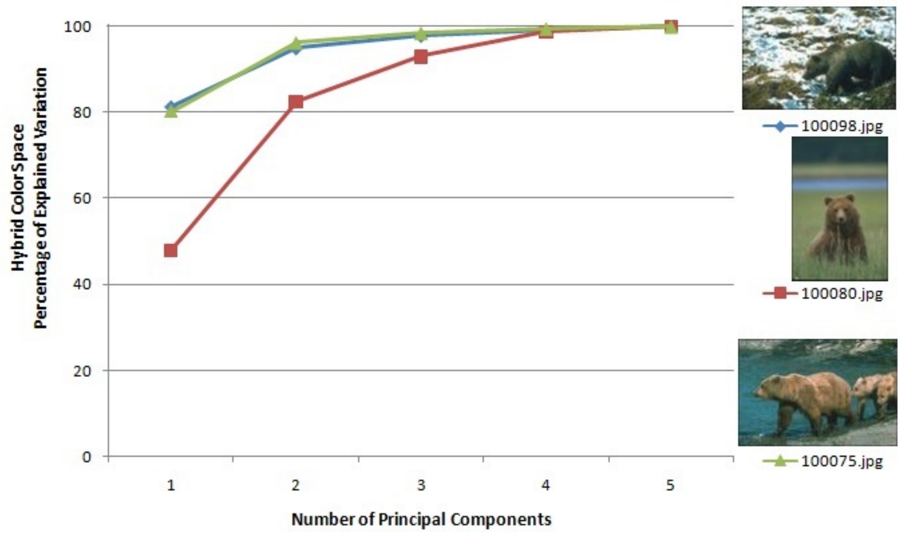
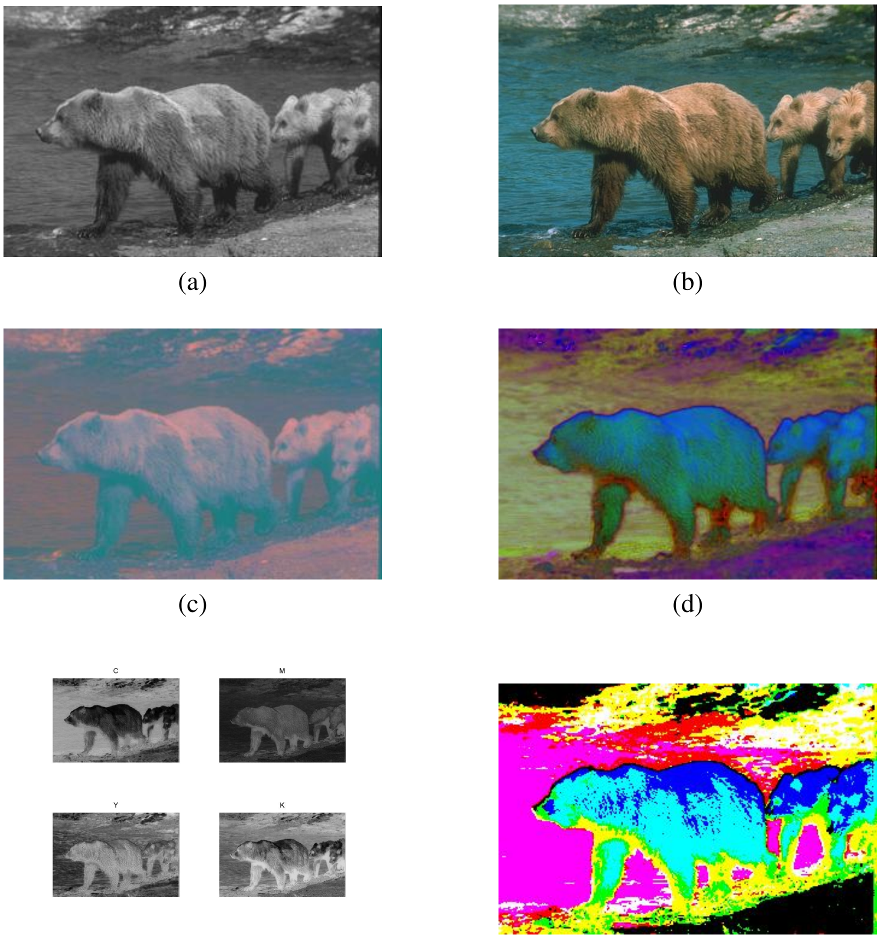


Figure 3.2: HCS - Percentage of explained variation

The Figure 3.2 shows the graph depicting the percentage of explained variation for PCA based HCS planes for three BSD images - 100098.jpg, 100080.jpg and 100075.jpg. It can be observed that the plot slowly grows as the number of principal components increases and reaches 100 percent. Using three principal components, the image 100098.jpg reaches 97.8%, 100080.jpg reaches 93.0% and 100075.jpg reaches 98.4% respectively. The first three principal components for HCS feature space usually contributes to 90% or more, thus retaining the color based visual information. The Figure 3.3 shows an image 100075.jpg from BSD dataset being viewed in different planes such as 2 dimensional planes (gray-scale) and 3 dimensional (RGB, HSV, LAB, HCS). CMYK is four dimensional which cannot be visualized, so 2D planes for each C,M,Y,K are shown. The next section describes the process of texture feature extraction.



(a) (b) (c) (d) (e) (f)

Figure 3.3: Image planes a) Gray (2D) b) RGB (3D) c) LAB (3D) d) HSV (3D) e) CMYK (4D) f) HCS(3D)

3.3 Texture understanding with Orthogonal Polynomials (OP)

The present section focuses on extraction of the different types of textures available. Texture, a contrasting feature represents variations in color or intensity. The task of detecting edges gives multiple directions to various concepts. Firstly, edge detection can be indirectly applied to segmentation task. The discontinuity based approaches include high-emphasis spatial frequency filtering, gradient operators such as first difference operators and second difference Laplacian operators, heuristic search and dynamic programming, relaxation, line and curve fitting as mentioned in survey paper by Tuceryan and Jain (1998). A detailed description about texture analysis methods is provided in Section 2.4 of the literature survey chapter. The next two subsections provide brief overview of the required information pertaining to orthogonal polynomial operators and its role in texture extraction.

3.3.1 Orthogonal Polynomial Operators: A Review

The polynomials play a very crucial role in numerical analysis as mentioned by Collins (2003). If a problem can be solved by a polynomial, then numerical analysis of appropriate methods are often used to resemble that polynomial to achieve the accuracy to maximum extent. Thus the theory of interpolation is based on polynomial approximation. Orthogonal polynomial is a very special and important class of polynomials [1]. The concept of orthogonality, focuses on two things being perpendicular to each other. Orthogonal polynomials form a complete set, that helps to express any arbitrary polynomial in terms of linear combinations of the set elements, which makes basis functions used for interpolation. A set of orthogonal polynomials is first used to obtain point-spread operators for a fixed window size, which subsequently can be used to generate a complete set of difference operators.

[1] http://ads.harvard.edu/books/1990fnmd.book/

The generation of the bank of filters, or set of difference operators has been detailed out in Step 1 and Step 2 as mentioned below. The following steps are mentioned in one of the publications by (Goswami *et al.* (2014*a*)) based on the contribution of this thesis.

Step 1. Point spread operator or Orthogonal basis matrix is a matrix representation M, using a set of orthogonal polynomials $u_0(t), u_1(t), ..u_{n-1}(t)$ of degrees $0, 1, 2, ..n-1$

$$M = \begin{bmatrix} u_0(t_1) & u_1(t_1) & \dots & u_{n-1}(t_1) \\ u_0(t_2) & u_1(t_2) & \dots & u_{n-1}(t_2) \\ \vdots & \vdots & \vdots & \vdots \\ u_0(t_n) & u_1(t_n) & \dots & u_{n-1}(t_n) \end{bmatrix}$$

There are two ways of obtaining point spread operator of size n. Statistical Tables by Fisher and Yates (1947) provide directly the matrix elements for size n, usually odd numbers like 3, 5, or 7 are used. The other way is to derive it mathematically. The mathematical explanation for the construction is presented in paper by Ganesan and Bhattacharyya (1995). We consider two point spread operators M1 and M2 for order 3 and 5 respectively.

(i) Orthogonal Polynomial Operators of order 3

Point spread operator M1 of size 3 is represented by set of 3 orthogonal polynomials $\hat{P}_0, \hat{P}_1, \hat{P}_2$.

$$M1 = [\hat{P}_0, \hat{P}_1, \hat{P}_2] = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}$$

Step 2. Construct a complete set of tensor products O^n_{ij} or orthogonal basis functions from $M1$ and $M2$ by vector outer product. O^n_{ij} is the difference operator for $n \times n$ window size. The set $O^n_{ij}(0 \leq i, j \leq n-1)$ can be computed as $O^n_{ij} = \hat{u}_i \bigotimes \hat{u}_j{}^t$ where $\hat{u}_i$ is the $(i+1)^{st}$ column vector of $|M|$. O_{00} is local averaging operator and remaining are difference operators. For example, O_{00} is obtained by vector product of first column of $M1$ with itself.

$$O_{00} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bigotimes \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Similarly O_{02} is obtained by vector product of first column of $M1$ with its third column.

$$O_{02} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bigotimes \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix}$$

A complete set of 9 difference operators or tensor products are obtained by $M1 \bigotimes M1$. μ stands for mean, L stands for linear component, Q stands for quadratic component and O_{ij}s are tensor products. O_{00} is the mean operator. There are total six pure linear and pure quadratic operators, three linear 2 L (O_{01} and O_{10}), LL (O_{11}), and three quadratic 2 Q (O_{02} and O_{20}) and QQ (O_{22}). LQ (O_{12}) and QL (O_{21}) are combination of linear and quadratic components. The complete set of OP3 operators along with their component types are mentioned as follows.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\mu \;\; O_{00} \qquad\quad L \;\; O_{01} \qquad\quad L \;\; O_{10}$$

$$\begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$Q \;\; O_{02} \qquad\quad Q \;\; O_{20} \qquad\quad LL \;\; O_{11}$$

$$\begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

$$LQ\ \ O_{12} \qquad\qquad QL\ \ O_{21} \qquad\qquad QQ\ \ O_{22}$$

(ii) Orthogonal Polynomial Operators of order 5

Point spread operator M2 of size 5 is represented as

$$M2 \ = \begin{bmatrix} 1 & -2 & 2 & -1 & 1 \\ 1 & -1 & -1 & 2 & -4 \\ 1 & 0 & -2 & 0 & 6 \\ 1 & 1 & -1 & -2 & -4 \\ 1 & 2 & 2 & 1 & 1 \end{bmatrix}$$

A complete set of 25 difference operators or tensor products are obtained by $M2 \bigotimes M2$.The set of 25 tensor products are shown in Figure 3.4. The first one is averaging operator and rest all others are difference operators.

3.3.2 Role of OP in texture extraction

The most common method to capture the textured image information is using signal processing techniques that usually filters the image with specific filters or bank of filters and exploit the filter responses to create texture features as mentioned in paper by Jain and Farrokhnia (1991). To capture textures, an image is supposed to be first convolved with bank of filters tuned to various orientations and spatial frequencies. This section develops a framework based on orthogonal polynomial(OP) operators as filters for its simplicity and computational efficiency, without specifying any scale and orientation in advance. The framework needs to mention the order of OP as external parameter and does not require any other parameter settings for choosing scale, spatial frequencies and orientation. The study confines to adopt OP operators of order 3 and 5.

1	1	1	1	1	-2	-2	-2	-2	-2	2	2	2	2	2	-1	-1	-1	-1	-1	1	1	1	1	1
1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	2	2	2	2	2	-4	-4	-4	-4	-4
1	1	1	1	1	0	0	0	0	0	-2	-2	-2	-2	-2	0	0	0	0	0	6	6	6	6	6
1	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-2	-2	-2	-2	-2	-4	-4	-4	-4	-4
1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1
-2	-1	0	1	2	4	2	0	-2	-4	-4	-2	0	2	4	2	1	0	-1	-2	-2	-1	0	1	2
-2	-1	0	1	2	2	1	0	-1	-2	2	1	0	-1	-2	-4	-2	0	2	4	8	4	0	-4	-8
-2	-1	0	1	2	0	0	0	0	0	4	2	0	-2	-4	0	0	0	0	0	-12	-6	0	6	12
-2	-1	0	1	2	-2	-1	0	1	2	2	1	0	-1	-2	4	2	0	-2	-4	8	4	0	-4	-8
-2	-1	0	1	2	-4	-2	0	2	4	-4	-2	0	2	4	-2	-1	0	1	2	-2	-1	0	1	2
2	-1	-2	-1	2	-4	2	4	2	-4	4	-2	-4	-2	4	-2	1	2	1	-2	2	-1	-2	-1	2
2	-1	-2	-1	2	-2	1	2	1	-2	-2	1	2	1	-2	4	-2	-4	-2	4	-8	4	8	4	-8
2	-1	-2	-1	2	0	0	0	0	0	-4	2	4	2	-4	0	0	0	0	0	12	-6	-12	-6	12
2	-1	-2	-1	2	2	-1	-2	-1	2	-2	1	2	1	-2	-4	2	4	2	-4	-8	4	8	4	-8
2	-1	-2	-1	2	4	-2	-4	-2	4	4	-2	-4	-2	4	2	-1	-2	-1	2	2	-1	-2	-1	2
-1	2	0	-2	1	2	-4	0	4	-2	-2	4	0	-4	2	1	-2	0	2	-1	-1	2	0	-2	1
-1	2	0	-2	1	1	-2	0	2	-1	1	-2	0	2	-1	-2	4	0	-4	2	4	-8	0	8	-4
-1	2	0	-2	1	0	0	0	0	0	2	-4	0	4	-2	0	0	0	0	0	-6	12	0	-12	6
-1	2	0	-2	1	-1	2	0	-2	1	1	-2	0	2	-1	2	-4	0	4	-2	4	-8	0	8	-4
-1	2	0	-2	1	-2	4	0	-4	2	-2	4	0	-4	2	-1	2	0	-2	1	-1	2	0	-2	1
1	-4	6	-4	1	-2	8	-12	8	-2	2	-8	12	-8	2	-1	4	-6	4	-1	1	-4	6	-4	1
1	-4	6	-4	1	-1	4	-6	4	-1	-1	4	-6	4	-1	2	-8	12	-8	2	-4	16	-24	16	-4
1	-4	6	-4	1	0	0	0	0	0	-2	8	-12	8	-2	0	0	0	0	0	6	-24	36	-24	6
1	-4	6	-4	1	1	-4	6	-4	1	-1	4	-6	4	-1	-2	8	-12	8	-2	-4	16	-24	16	-4
1	-4	6	-4	1	2	-8	12	-8	2	2	-8	12	-8	2	1	-4	6	-4	1	1	-4	6	-4	1

Figure 3.4: M2 $\bigotimes$ M2 : OP5 Tensor Products

Figure 3.5 shows all the nine orthogonal operators of order 3 along with their characteristics. Orthogonal effects β_{ij}, $\beta_{ij}(0 \leq i, j \leq 2; i, j \neq 0)$ corresponding to O_{ij} are produced on applying these operators on the color planes of an image. Orthogonal effects correspond to texture characterization as they are combination of two disjoint subsets, main effects and interaction effects as mentioned by Ganesan and Bhattacharyya (1995). Main effects are the average effect of each factor over all levels of other factor. $\beta_{01}, \beta_{02}, \beta_{10}, \beta_{20}$ correspond to main effects. β_{01} and β_{10} effects are Prewitt edge operator responses. Interaction effects are effect of one factor at different levels of the other factor. $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$ correspond to interaction effects. β_{12} and β_{21} effects are Sobel edge operator responses and β_{22} is Laplace edge operator response. Thus one can see that using set of difference operators which are basis matrices produced by point spread function $M1 \bigotimes M1$ gives us some extra responses apart from some of the well known edge operators. Tonality is represented by the fact that main effects dominate and interaction effects are negligible. The micro texture can be characterized by the fact that interaction effects dominate over the main effects. Similarly in case of OP5,

$\begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$ O_{00} **Average**	$\begin{matrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{matrix}$ O_{01} Linear **Prewitt** *Main Effects*	$\begin{matrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{matrix}$ O_{02} Quadratic *Main Effects*
$\begin{matrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix}$ O_{10} Linear **Prewitt** *Main Effects*	$\begin{matrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{matrix}$ O_{11} Linear *Interaction Effects*	$\begin{matrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{matrix}$ O_{12} Linear & Quadratic **Sobel** *Interaction Effects*
$\begin{matrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{matrix}$ O_{20} Quadratic *Main Effects*	$\begin{matrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{matrix}$ O_{21} Linear & Quadratic **Sobel** *Interaction Effects*	$\begin{matrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{matrix}$ O_{22} Quadratic **Laplacian** *Interaction Effects*

Figure 3.5: OP3 Operators and their characteristics

the natural variations can be captured by the superset of 25 difference operators. The next section describes the color-texture feature integration in order to construct adaptive feature vector representation for image segmentation.

3.4 Color-Texture Feature Integration - proposed variants OP3-HCS and OP5-HCS

The proposed HCS color features and texture feature extraction based on OP needs to be integrated. We follow the first trend of implicit color-texture feature integration as discussed earlier in literature survey, because both color and texture complement each other and need to be processed together. The texture features are extracted from single or multiple color channels. The set of orthogonal polynomial operators are directly used with HCS, to analyze the responses or effects for segmentation purpose. We propose two variants of OP - OP3 of order 3 and OP5 of higher order 5, which are applied on Hybrid Color Space (HCS) for color texture feature integration. In case of OP3-HCS, each pixel is represented as 27 dimensional feature vector. Similarly, in case of OP5-HCS, the color and texture features are integrated by convolving each of the three HCS planes with 25 set of tensor products. Each pixel in the image is thus represented as 75 dimensional feature vector in the color-texture feature space.

When the operators OP3 are applied on the HCS color image (such as BSD 100075.jpg), the illustration is shown in Figure 3.6 in terms of histogram (Fig.3.6 c,e,g) and tensor responses (Fig.3.6 d,f,h) for each of the three planes. On a similar note, when the operators OP5 are applied on the HCS color image (such as BSD 100075.jpg), the illustration is seen in Figure 3.7 in terms of histogram (Fig.3.7 c,e,g) and tensor responses (Fig.3.7 d,f,h) for each of the three planes. The very first responses of figures d,f,h of both the representations depict the mean response and the objects with tonal contrasts are seen quite prominently. The rest of the other responses capture the contrasting textures. It can be observed that the responses are rich in both color-texture related information implicitly. The size of OP (3 or 5) is an external parameter provided by the user (parameter 1).

AFVR - Adaptive Feature Vector Representation

The increase in number of features add to the computational complexity. Hence, one of the challenges is knowing about the sufficient number of features, fixed or adaptive in order to represent the image in totality. Feature vector dimensionality reduction using PCA is found to be very popular in the literature. In the literature, one can find feature vector dimensionality reduction using PCA in Hoang *et al.* (2005) and Allen *et al.* (2008). Hoang *et al.* (2005) reduces feature vector to 4-dimension and Allen *et al.* (2008) reduces to 8 feature vector. Allen *et al.* (2008) found that experimentally 8-dimensional feature vector space satisfactorily captures the information regarding textures from natural images. Our work got motivated towards adaptive feature vector representation where the number of features vary from image to image. The contribution is mentioned in paper by Goswami *et al.* (2014*a*). We justify using the adaptive mode of feature selection where the percentage of explained variation is fixed to be at the most 99% (parameter 2). The feature vector thus obtained, is referred as Adaptive Feature Vector Representation(AFVR). The number of features selected is adaptive, and varies from image to image depending upon its complexity and at the same time, it proves to

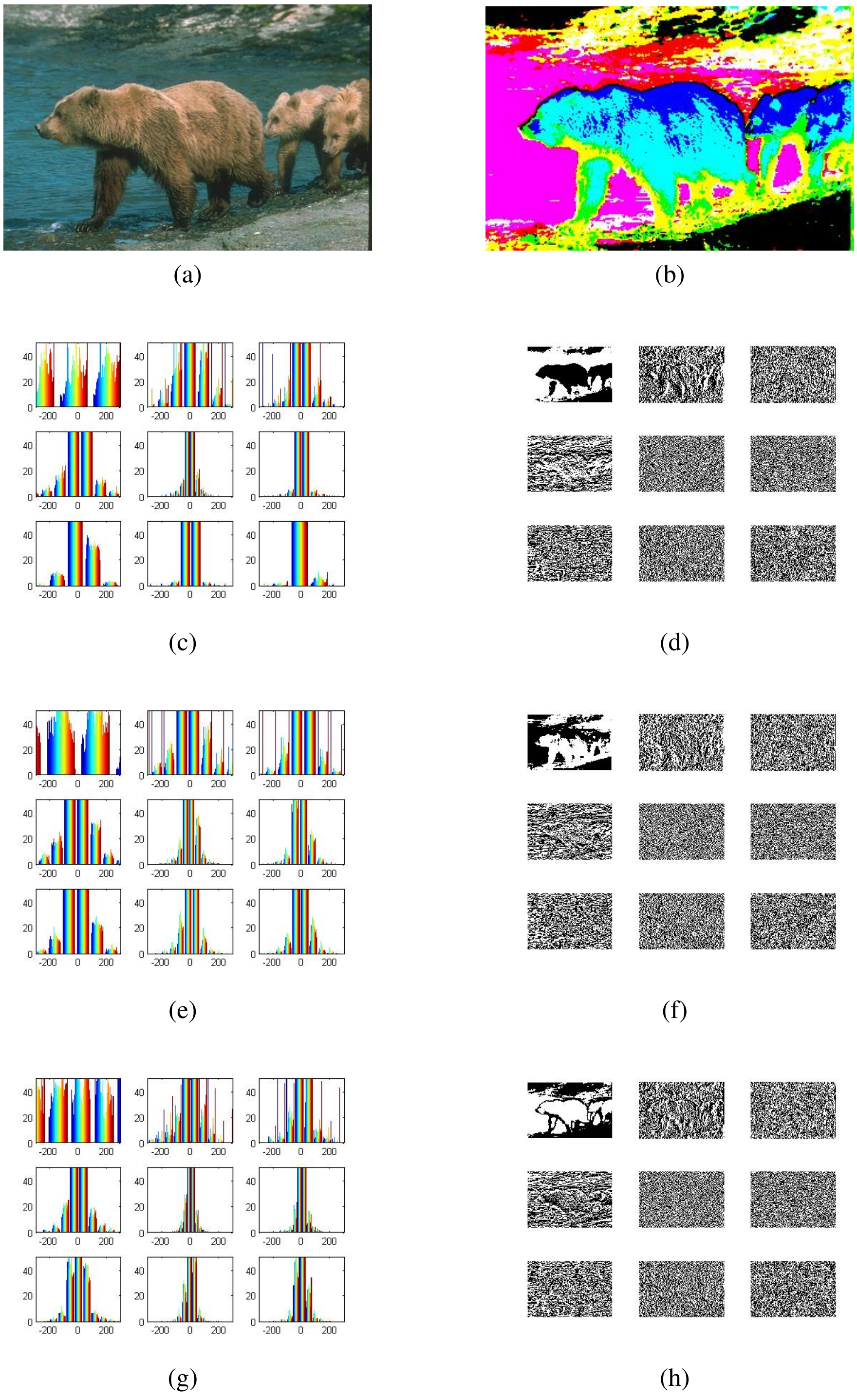


Figure 3.6: OP3 for knowledge representation of HCS image (a) Original image (b) HCS image (c) OP3-HCS Plane1 Histogram (d) OP3-HCS Plane1 tensor responses (e) OP3-HCS Plane2 Histogram (f) OP3-HCS Plane2 tensor responses (g) OP3-HCS Plane3 Histogram (h) OP3-HCS Plane3 tensor responses

(a) (b) (c) (d) (e) (f) (g) (h)

Figure 3.7: OP5 for knowledge representation of HCS image (a) Original image (b) HCS image (c) OP5-HCS Plane1 Histogram (d) OP5-HCS Plane1 tensor responses (e) OP5-HCS Plane2 Histogram (f) OP5-HCS Plane2 tensor responses (g) OP5-HCS Plane3 Histogram (h) OP5-HCS Plane3 tensor responses

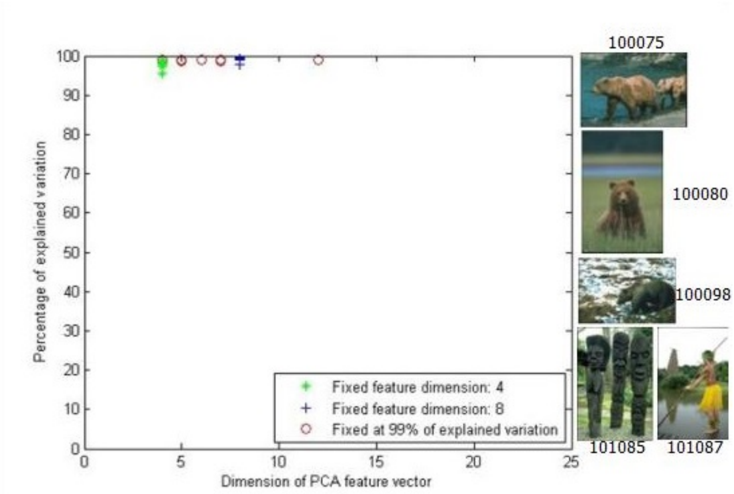
be time efficient for the subsequent steps in image segmentation, as mentioned in Table 3.1.

Table 3.1: Adaptive Feature Vector Dimensions

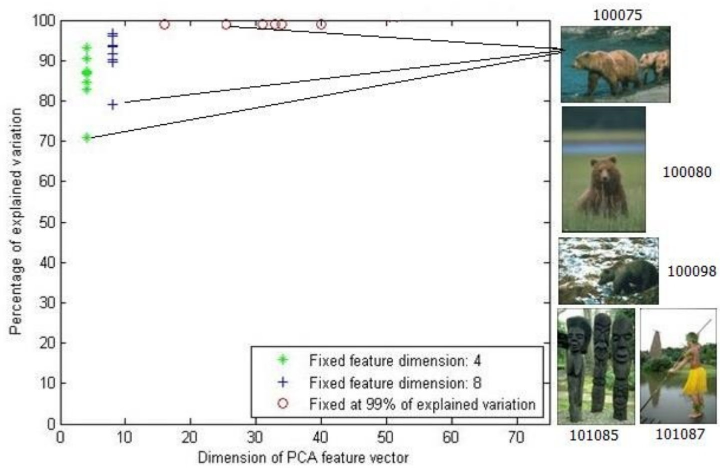
BSD Images	AFVR Number of Features	Time (secs)
101085	40	24.7
101087	40	24.7
100098	33	20.0
100075	26	18.9
100080	16	9.4

We apply PCA technique to the tensor responses produced from OP3/OP5 to reduce the feature dimensions. Latent vector and scores are returned from the PCA, from which we can calculate the *percentage of explained variation* Scores are the data formed by transforming the original data into the space of the principal components. Latent vector are the variance of the columns of score. The illustration of the experimental results is shown in Figure 3.8, based on 5 BSD images. OP3-HCS produces 27 feature vector for each image pixel and reducing to both four and eight PCA features, usually attain 99% explained variation as shown in Figure 3.8 (a). It is clear from the figure that all the green symbols (for 4 dimensions) and blue symbols (for 8 dimensions), achieve more than 90% of explained variation. The red symbols denote the number of features required for the explained variation fixed at 99%.

On the other hand OP5-HCS produces 75 feature vectors for each image pixel, but after reducing to 4 dimensions, few images did not even achieve 90% of explained variation, as can be seen in Figure 3.8 (b). To site an example, for the image 100075.jpg, the percentage of explained variation almost dropped to 70% using 4 feature vectors from OP5-HCS, shown by green symbol. When the features are reduced to 8, the percentage of explained variation did not cross 80%, as can be seen by blue symbol. Thus fixed length feature vector of dimensions four or eight, does not guarantee the capture of maximum possible image information for some images, based on OP3-HCS



(a)



(b)

Figure 3.8: Adaptive versus Fixed Feature Vector Representation: Illustrations for (a) OP3-HCS (max 25 dimensions) (b) OP5-HCS (max 75 dimensions)

and OP5-HCS. Hence we propose Adaptive Feature Vector Representation (AFVR), where the percentage of explained variation is fixed to 99%, seen as red symbol. Thus, for the 100075.jpg, 26 PCA feature vectors are considered in case of OP5-HCS to reach 99% as compared to 5 features for OP3-HCS. The algorithm for adaptive feature vector representation is shown in Algorithm 2. The next section proceeds further with iterative clustering process on the AFVR obtained for a given image.

Algorithm 2: Adaptive Feature Vector Representation

Input: Img_HCS
$TensorProducts\_Matrix$
Output: $Adaptive\_FV$
$TensorResponses$
for $Each\_HybridColorPlane\_In\_Img\_HCS$ **do**
  for $EachTensorProducts\_Matrix$ **do**
    $TensorResponses \leftarrow imfilter(HybridColorPlane, TensorProducts\_Matrix)$
  end for
end for
$[score, latent] \leftarrow PCA(TensorResponses)$
$variation \leftarrow CummulativeSum(latent)/sum(latent) * 100$
$numFeatures \leftarrow length(find(variation <= VariationInPercent))$
$Adaptive\_FV \leftarrow PCA(TensorResponses, numFeatures)$
return $Adaptive\_FV,$
$TensorResponses$

3.5 Clustering based segmentation

We propose to use clustering, a region based segmentation technique to partition the image into separate regions based on color-texture homogeneity. K-means clustering is one of the most popular methods found in literature. It is known that, one of challenging tasks found in employing this method, is fixing the number of clusters, needed to be known a-priori. Fixing K - the number of clusters (K), usually can be one way

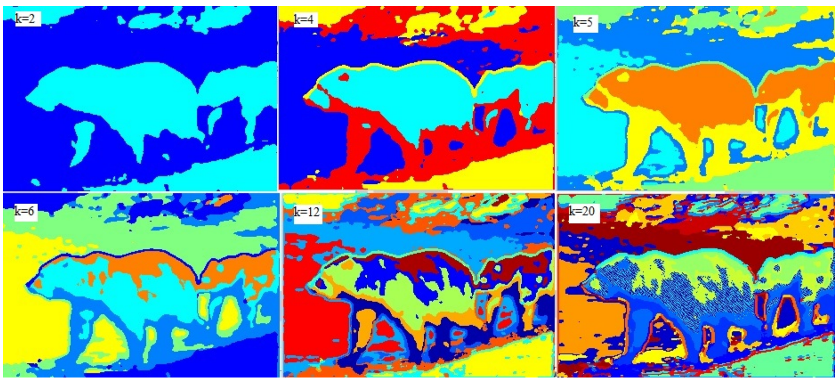


Figure 3.9: Segmentation demo of K-means clustering with different number of K values. First row: K=2, K=4, K=5 Second row: K=6, K=12, K=20

of achieving the segments. Figure 3.9 illustrates image segmentation using K-means clustering with different number of clusters. It can be observed that the first row, first image is segmented with K=2, and the result is under-segmented. On the other hand, in the second row, last image is segmented with K=20, and the result is over-segmented. A blind clustering based machine learning approach may lead to meaningless segments, mostly in case of natural outdoor images. A multi-step learning approach is expected to give effective and meaningful segments. One such approach can be generating and compiling useful metadata for images to be segmented. The proposed image types is one such metadata generated, based on some statistical analysis, which is mentioned in next Section 3.5.1. In order to arrive at an appropriate number of clusters, iterative K-means along with intelligent stopping criteria is proposed and implemented as discussed in the Section 3.5.2. The cluster metadata obtained in the segmentation process is mentioned in Section 3.5.3.

3.5.1 Image Types

Let us digress briefly, and discuss difficulty in segmenting natural outdoor images before we proceed further with iterative region splitting procedure. The natural image object segmentation pose to be quite challenging due to the intrinsic characteristics found in the images in terms of various ratio of color and texture mix. The image contains a single object or many different objects which can vary from being tonal to complex color-textured as illustrated in Figure 3.10. In tonal images, the presence of

(a) (b) (c)

Figure 3.10: Natural images for object segmentation

single object or multiple objects are very prominent due to certain factors such as large object size, the presence of salient contrast in color (as can be seen in Figure 3.10a)) or color-texture variation (as seen in Figure 3.10b,c). This makes the object detection demonstrable to certain extent.

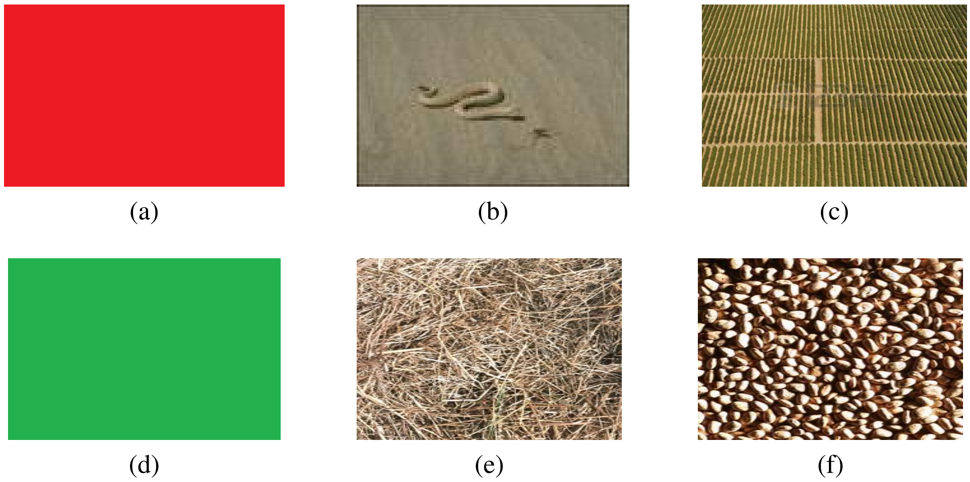


Figure 3.11: (a,d) Pure toned images, (b,e)Homogeneous color based texture, (c,f)Heterogeneous color based texture

The poorly performed segmented images usually consist of homogeneous color based textured images, as camouflage makes it difficult to segment the object in totality as shown in Figure 3.11(b,e). Segmentation also becomes difficult in case of heterogeneous color based texture as shown in Figure 3.11(c,f), where perceptual grouping becomes a challenge when segmented, and outputs large number of meaningless regions.

The BSD dataset is meant for object segmentation in images catering to outdoor natural scenes as mentioned by Todorovic and Ahuja (2009). Our unsupervised algorithm is tested on the BSD for quantitative evaluation and comparisons. We do not want to limit our work to only BSD, and aim to cater to an overall variations of color-texture present in the nature irrespective of whether an object is present or not. We have considered various other datasets such as Vistex textured images, Corel categorical(food, building) images, Prague(mosaic textured images) and found the difficult segmentation scenarios. Some border cases have also been tested such as pure toned colors. Researchers deploying segmentation techniques mostly assume the image to have objects.

Sometimes meaningless pseudo segments are obtained when the technique fails in perceptual grouping. One can overcome, such type of failures by setting some informative flags to the image as metadata, whether to subject it to segmentation or not. For this purpose, we have developed a schematic methodology to categorize the images into types, which serves the purpose of image metadata as well. In all, we have identified four types of images. Our segmentation algorithm implements in identifying such images which are as follows.

1. Type-1: Object-based Image: Image can be segmented into meaningful objects with prominent color-texture variations.
2. Type-2: Heterogeneous color based texture Image: Image can be segmented into large number of small regions, where perceptual grouping is a challenge.
3. Type-3: Homogeneous color based texture Image: Image is not having many color variations. If there is any object present, it poses a camouflage problem with the similar shaded background.
4. Type-4: Pure toned Image: Single colored image, which cannot be segmented into objects.

Out of the four image types listed above, segmentation is not possible for three image types except the *objects-based image* type. Our proposed segmentation methodology is able to find the object (snake) from similar shaded homogeneous color based texture image 196073.jpg (Figure 3.11(b)). Once it is able to extract a prominent region, it labels the image as object-based image. But in case of Vistex grass image (Figure 3.11(e)), the category is correctly identified as homogeneous color based texture. Our work further aims to build some knowledge about these images, so that such knowledge will have an added advantage by providing input to appropriate higher order complex algorithms for image understanding. Hence, such images once identified needs to be characterized using texture characterization, which will be discussed in the next chapter. In case the segments are obtained from the *objects-based image* type;

the segmented regions can also be further processed for texture characterization. The iterative clustering based on OP-HCS is discussed in forthcoming subsection.

3.5.2 Iterative Region Splitting on OP-HCS

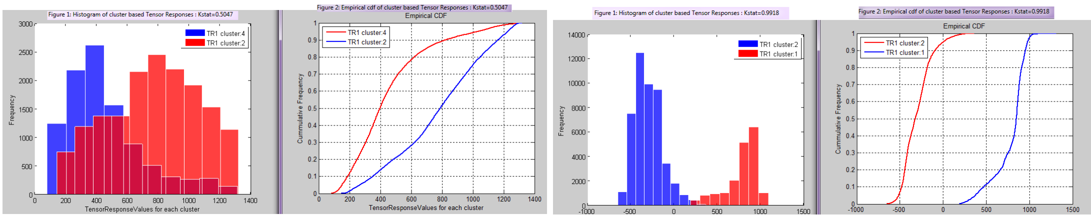
Considering these many different cases of image types, it is evident that for unsupervised segmentation for natural images, knowing the number of cluster will vary from one image to another. The proposed segmentation approach in an unsupervised manner has considered the popular K-means clustering in an iterative manner, partitioning with K=2 clusters in each iteration. No prior knowledge about initial number of clusters is known, hence a stopping criteria is also employed in each iteration to decide for proceeding with further segmentation or not.

Initially for checking the border cases, first a check is done to test if the image is plain color, i.e. without any object. The variance of each input RGB color plane is computed and if all the three planes have variance less than one, the image is considered to be pure single color tonal image. In such cases, Otsu thresholding will also fail in finding any contrast. Region splitting step is not possible in such case.

Intelligent stopping criteria: Kolmogorov-Smirnov Test

The two straightforward stopping criteria are region population size and maximum number of clusters obtained. The values of the parameters are set experimentally. If the new cluster area as compared to the total image area in pixels, is equal or less than 20 percent (parameter 3), stop further segmentation. The upper bound of the number of maximum clusters reached is 15 (parameter 4). The third stopping criterion is a crucial one, which exploits the dissimilarity in the local color-texture coherence of the clusters, as the objective function in deciding further segmentation.

Firstly, compute the mean of the histogram of the tensor responses of the latest cluster obtained, and compare with mean of each previous cluster tensor responses. Next, assess if they have non-overlapping support, by setting a difference of mean threshold cutoff. If at this step, the clusters have overlapping support and do not contain a contrasting mean. then the image is similar shaded textured image. The quantitative cluster similarity assessment between local color-texture distributions is evaluated only for the clusters, whose tensor responses have non-overlapping support using the two sample Kolmogorov-Smirnov(K-S Test) metric. K-S Test metric has been used in paper by Ilea and Whelan (2008) for finding similarity between local and global color-texture distributions of the clusters based on multi-resolution approach. In this work, K-S Test is used differently as a stopping criteria. The test statistic K-S statistic is the maximum difference between the curves, which are empirical cumulative distribution function (c.d.f) of the two clusters. The tensor responses of clusters with overlapping support,



(a) (b)

Figure 3.12: Quantitative cluster similarity assessment (Red colored region marks overlap) (a)Overlapping supports, KSstatistic=0.5047 (b)Non-overlapping supports, KSstatistic=0.9918

returns K-S statistic value near to zero, as the cluster properties are identical. When the clusters have distinguishable feature characteristics, they have non-overlapping support and the test statistic difference is maximum as shown in Figure 3.12. The flowchart for iterative segmentation is shown in Figure 3.13.

The K-S statistic is the maximum difference between the curves, which are empirical c.d.f of the two clusters. The tensor responses of clusters with overlapping support returns K-S statistic value near to zero, as the cluster properties are identical. When the clusters have distinguishable feature characteristics, they have non-overlapping support and the test statistic difference is maximum. We have empirically chosen the difference mean threshold equal to 900 (parameter 5) and threshold for K-S statistic equal to 0.5 (parameter 6) after trial and error, which happens to be one time setting for all the

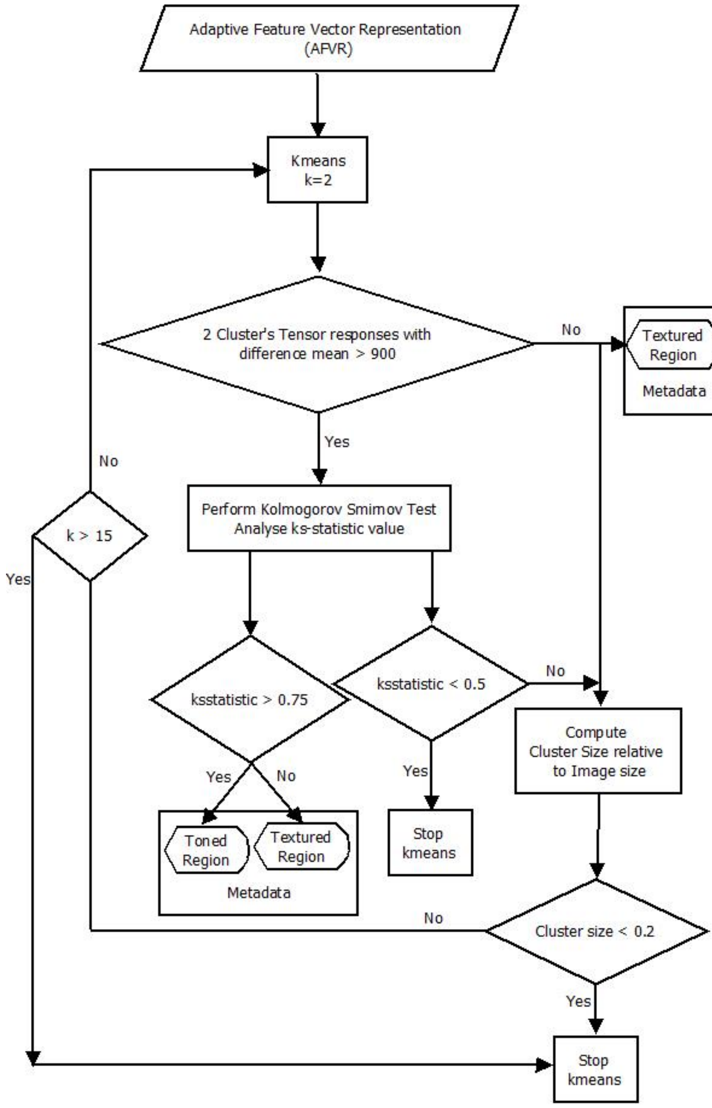


Figure 3.13: Flowchart of Iterative Segmentation

images. As the number of clusters increase, the third stopping criterion becomes time consuming, if it includes only K-S Test. Hence to reduce the time complexity, assessing first non-overlapping support clusters is an inexpensive option. Finally, we achieve an over-segmented output image after the clustering without any post processing.

The sequence of stopping criteria used while segmentation process, associates metadata with each segmented region as toned or textured. This metadata can be considered as target detection as mentioned in Section 1.2, where the target is presence or absence of texture characteristics present in a segmented image. Tonality feature of an image is

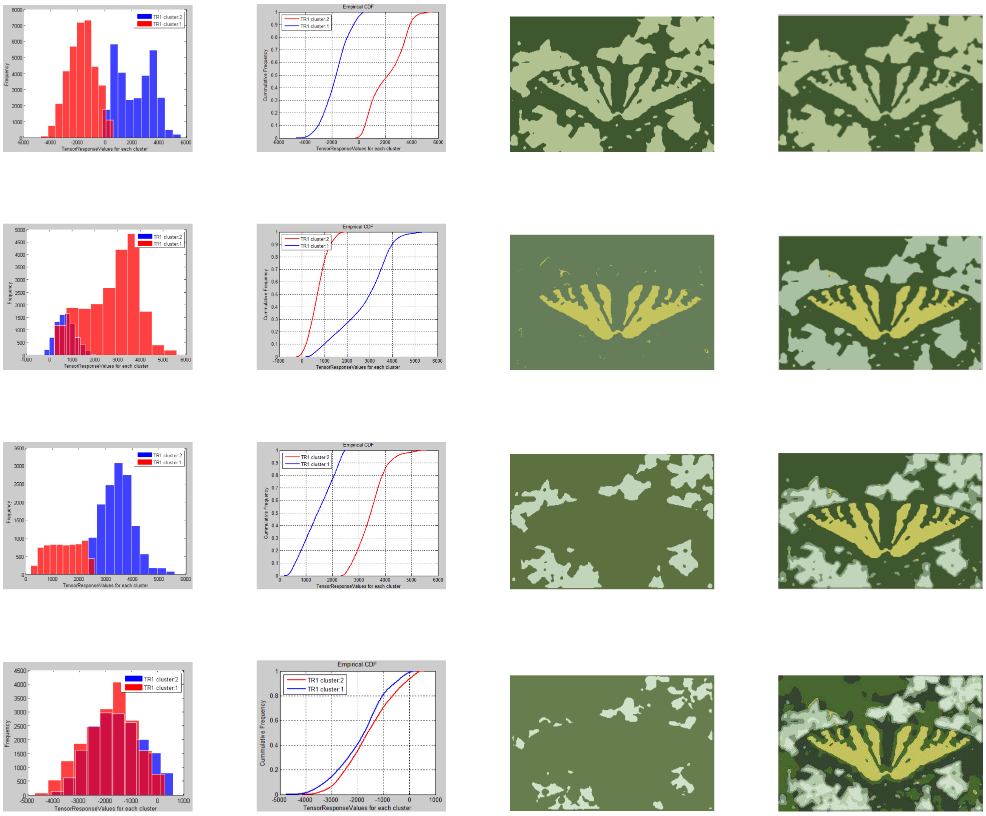


Figure 3.14: Segmentation Steps for each iteration OP3-HCS
Column 1: Histogram of cluster based tensor responses of two K-means clusters;
Column 2: Empirical cdf of cluster based tensor responses of two K-means clusters;
Column 3: New cluster formed;
Column 4: Segmented image

revealed, when the K-S statistic is close to 1, in the range from 0.8 to 1.0, or else they are textured. Textured segments having K-S statistic less than 0.5, is not considered as a candidate for further segmentation, as they are found similar in color and texture. The algorithm for iterative clustering is shown in Algorithm 3. In Figure 3.14 the stopping criteria values in each iteration are mentioned as follows:

Row 1: Mean difference:3910, K-S statistic value:0.9579, "Toned cluster"

Row 2: Mean difference:2040, K-S statistic value:0.7797, "Toned cluster"

Row 3: Mean difference:2028, K-S statistic value:0.9828, "Toned cluster"

Row 4: Mean difference:888.6, K-S statistic value:0.10631, "Textured cluster"

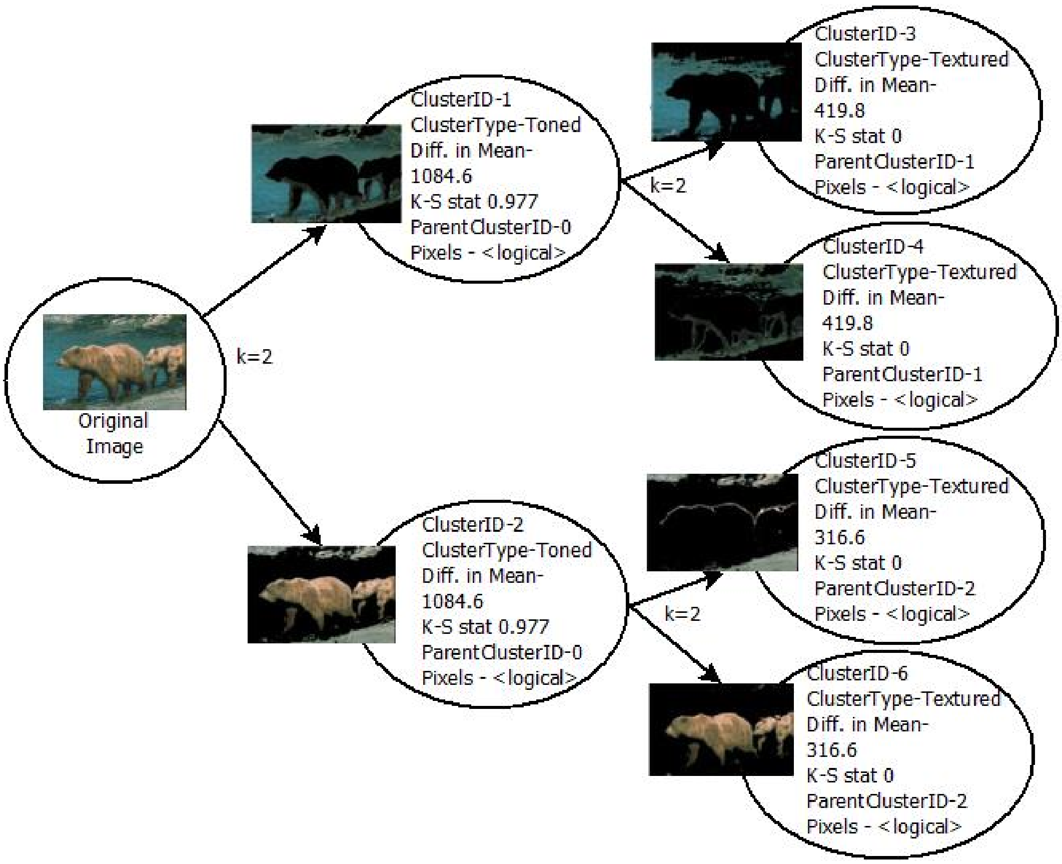
Algorithm 3: Clustering based Segmentation

```
Input: Adaptive_FV; TensorResponses; RegionSizeInPercent;
       MaximumClustersAllowed; MeanThreshold; KSstatisticThreshold
       N_AsNumberOfClusters
Output: Clustered_Output_Img;ClusterMetadata
do_Kmeans ← 1; do_KSTest ← 0; check_ClusterSize ← 0
while NumberOfClusters <= MaximumClustersAllowed & do_Kmeans = 1 do
    Kmeans_output ← Perform_Kmeans_Clustering(Adaptive_FV, 2)
    if (Kmeans_Fail) then
        Display('No_Clusters_found')
    else
        Clustered_Output_Img ←
        Populate_With_2KmeansIndex(2N, 2N − 1)
        [Mean1, Mean2] ←
        Mean_Histogram_Analysis(TensorResponses, 2KmeansIndex)
        check_ClusterSize ← 1
        if Difference(Mean1,Mean2) > MeanThreshold then
            do_KSTest ← 1
        else
            ClusterMetadata ←
            Populate(Cluster_Description,'Textured')
        if do_KSTest = 1 then
            KS_StatisticIndex ←
            Perform_KSTest(TensorResponses(2KmeansIndex))
            if KSstatistic >= 0.75 then
                do_Kmeans ← 1
                ClusterMetadata ←
                Populate(Cluster_Description,'Toned')
            else
                do_Kmeans ← 1
                ClusterMetadata ←
                Populate(Cluster_Description,'Textured')
            if KSstatistic < 0.5 then
                Stop_Kmeans; do_Kmeans ← 0
        if Check_ClusterSize = 1 then
            Cluster_size ← ComputeClusterSize_RelativeToImage
            if Cluster_size < RegionSizeInPercent then
                Stop_Kmeans; do_Kmeans ← 0
```

The cluster metadata generated at each iteration is depicted in the next subsection.

3.5.3 Cluster Metadata

Cluster metadata is by-product of the segmentation. Cluster metadata is feature-based and does not consider spatial compactness. Each pair of clusters obtained from the K-means (K=2), is annotated with a cluster identification number, cluster type (toned or textured) depending on the difference in means and K-S statistic test measure. The parent cluster is also recorded from which the particular cluster is formed. The pixels forming the cluster is stored to extract the binary image of the cluster. The illustration of the cluster metadata structure is shown in Figure 3.15 for BSD image 100075.jpg. Finally, the region generation from the segmentation, region merging and region metadata generation are post processing steps, which are discussed in the next section.



(a)

Figure 3.15: Cluster Metadata

3.6 Post processing - Spatially constrained Region Merge

Clustering shows only the clusters with different homogeneity constraint and does not take care of spatial proximity. To achieve object segmentation, we implement two steps, namely object labeling and region merge to obtain all the distinct objects in terms of color-texture characteristic from the image. The regions which are not adjacent, or far apart are considered as separate entity, though they are homogeneous in feature space.

(i) Object labeling

Spatial constraint has been imposed using 8-connected components technique, in such a manner that if clusters with same label exist at different spatial location, then those clusters will be labeled with different labels. All regions are obtained finally. This step provides region-level segmentation. The segmented image after applying the above explained technique consists of a number of small regions, which needs to be post processed. If the number of regions are more than 100, then the image can be categorized as multi-color highly textured image.

(ii) Region Merge

In order to increase the segmentation accuracy, we have devised a strategy which merges the over-segmented image. The Gestalt laws in Gestalt psychology plays an important role in describing the way human brain achieves perceptually grouping as mentioned earlier in Section 1.3. The region merge mostly is based on the closure law, where very small regions get fused to large region in proximity with high similarity. A two step method is applied for region merge with spatial constraints. Firstly, we propose to identify the very small regions. Each cluster may consist of more than one region of various sizes. Calculate deviation of each region from region with maximum area in terms of number of pixels. The regions with 90% deviation or more (parameter 7) are added to the merge list. Secondly, the very small regions of each cluster, are adaptively fused or absorbed by their neighboring cluster in an agglomerative manner, in which the pair of regions showing the highest similarity is merged first. The Mahalanobis distance measure differs from other distance measures, as it takes into consideration the correlation of the data set and thus is scale invariant. The data set considered is the selected feature space of adaptive dimensions for that image. This distance measure of merging regions is also adopted in paper by Nguyen *et al.* (2000). The similarity between region R_i and R_j is given by

$$\varphi_{i,j} = (\mu_i - \mu_j)^T [\Sigma_i + \Sigma_j]^{-1} (\mu_i - \mu_j) \tag{3.1}$$

where μ_i , μ_j are the mean vectors and Σ_i , Σ_j are the covariance matrices computed from feature vectors of regions R_i and R_j. The advantage of this measurement is that, the uncertainty represented by covariance matrices Σ_i and Σ_j of the mean vectors μ_i and μ_j is taken into consideration. The regions R_i and R_j are merged if the value of $\varphi_{i,j}$ is under the threshold (parameter 8) set at 1000. The final regions or objects are obtained after post processing steps. The cluster labeling and object labeling along with all the objects obtained are shown in Figure 3.16.

(iii) Region Metadata

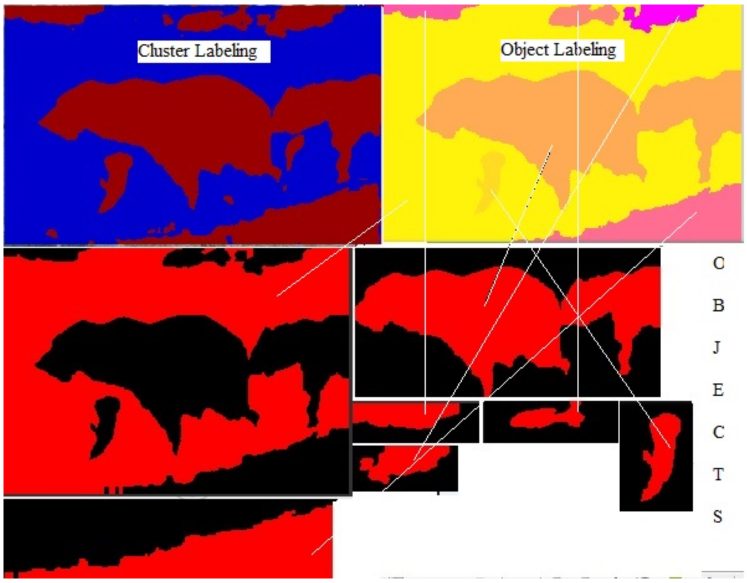
Each region obtained after post processing is associated with metadata. The region information basically consists of descriptive statistics - first order statistics (mean intensity) and second order statistics (GLCM properties). Gray level co-occurrence matrix mentioned by (Haralick *et al.* (1973)), is one of the early methods used for global description of textural images. Once GLCM is created, one can derive several statistics about the texture of the image. These statistics are listed below.

a) Contrast : Returns a measure of the intensity difference between a pixel and its neighbor over the whole image or segmented region. Contrast is zero for a constant image.

b) Correlation : Returns a measure of how correlated a pixel is to its neighbor over the whole image. Correlation is 1 or -1 for a perfectly positively or negatively correlated image and is NaN for a constant image.

c) Energy : Returns the sum of squared elements in the GLCM. Energy is 1 for a constant image.

(a)

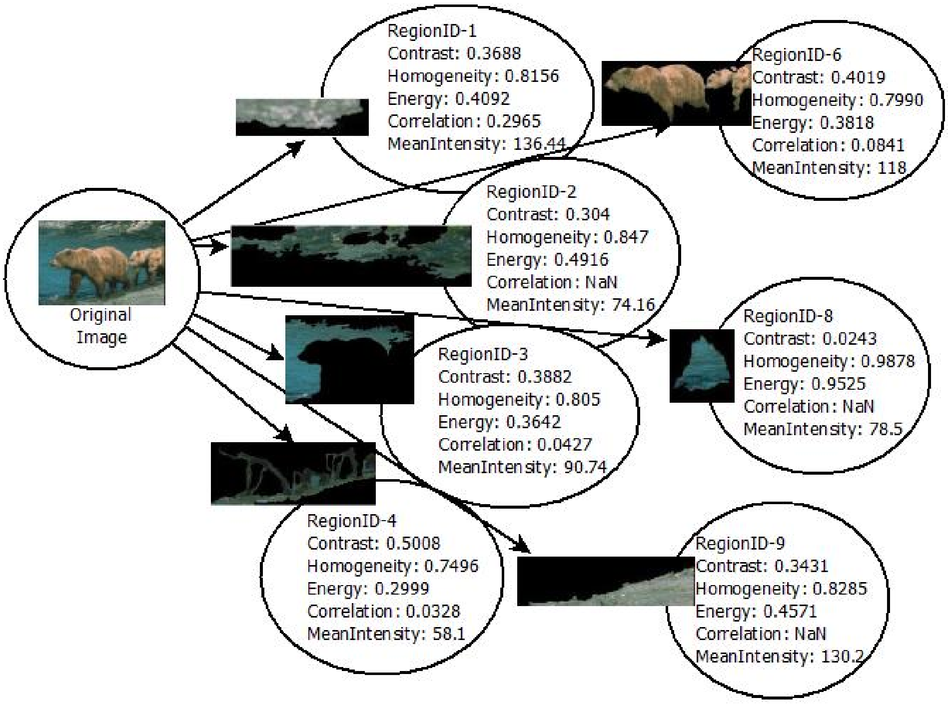


(b)

Figure 3.16: (a) Original Image (b) Cluster labeling and Object labeling after merging small regions

d) Homogeneity : Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for a diagonal GLCM.

A window size of 11x11 has been considered on grayscale intensity values of the image for GLCM computation. These measures serve as a metadata for the regions produced after object segmentation and region merge, and is illustrated in Figure 3.17. Descriptive statistics is just a single number depicting the whole region. These descriptive statistics GLCM based measures do not reveal anything about the nature of basic orientation of the textures present. This is taken up in the next chapter - image texture characterization based on simple perceptual variations in row, column or diagonal specific. The algorithm for post-processing is shown in Algorithm 4. The overall algorithm



(a)

Figure 3.17: Region metadata based on Descriptive statistics : First order statistics - Mean of gray scale intensity , Second order statistics - 4 GLCM measures for 11x11 sub-region

for unsupervised color-texture image segmentation based on OP-HCS can be found in Algorithm 5.

Algorithm 4: Post Processing

Input: $Adaptive\_FV$
$SimilarityThreshold$
$SmallRegionThreshold$
$RegionMergeThreshold$
$Clustered\_Output\_Img$

Output: $Region\_Merged\_Img$
$RegionMetadata$

1: $Region\_Output\_Img \leftarrow SpatiallyConstrained\_RegionLabeling(Clustered\_Output\_Img)$
2: $Region\_Merged\_Img \leftarrow RegionMerge(Adaptive\_FV, Region\_Output\_Img, SmallRegionThreshold, SimilarityThreshold)$
3: $RegionMetadata \leftarrow PopulateRegionMetadata(Region\_Merged\_Img)$
4: **return** $Region\_Merged\_Img, RegionMetadata$

Algorithm 5: OP-HCS Unsupervised Segmentation Method

Input: Img_RGB
OP_size
$FeatureVariationInPercent$
$RegionSizeInPercent$
$MaximumClustersAllowed$
$MeanThreshold$
$KSstatisticThreshold$
$SmallRegionThreshold$
$RegionMergeThreshold$
$SimilarityThreshold$

Output: $Region\_Merged\_Img$
$ClusterMetadata$
$RegionMetadata$

1: $Img\_HCS \leftarrow ConvertImageFromRGBToHCS(Img\_RGB)$
2: $TensorProducts \leftarrow FilterBankBasedOnOrthogonalPolynomial(OP\_size)$
3: $Adaptive\_FV \leftarrow GenerateAdaptiveFeatureVectorRepresenation(Img\_HCS, TensorProducts, FeatureVariationInPercent)$
4: $[Clustered\_Output\_Img, ClusterMetadata] \leftarrow PerformIterativeClustering(Adaptive\_FV, TensorResponses, RegionSizeInPercent, MaximumClustersAllowed, MeanThreshold, KSstatisticThreshold, N\_AsNumberOfClusters)$
5: $[Region\_Merged\_Img, RegionMetadata] \leftarrow PostProcessing(Adaptive\_FV, SimilarityThreshold, SmallRegionThreshold, RegionMergeThreshold, Clustered\_Output\_Img)$
6: **return** $Region\_Merged\_Img$
$ClusterMetadata$
$RegionMetadata$

3.7 Summary and Conclusions

We have established two variants of our method OP3-HCS and OP5-HCS based on the use of filter banks of tensor products, obtained from orthogonal polynomials of order 3 (OP3) and order 5 (OP5), that are applied and experimented on Hybrid Color Space (HCS) for color-texture segmentation. The distinctive features of our segmentation include adaptive feature vector representation (AFVR), iterative K-means clustering using cluster similarity test based on histogram analysis and K-S Statistic test. The segmentation tree is constructed in a hierarchical fashion, where the clusters and regions obtained at each level, can be labeled as *toned* or *textured* depending on test statistic threshold. The parameter settings used is one time initialization. The choice of parameters adds flexibility between under-segmentation and over-segmentation. The inherent simplicity of the generation of color-texture integration using OP-HCS (OP3 or OP5), exemplifies satisfactory clustering performance quantitatively, which has been discussed in the next chapter.

CHAPTER 4

Quantitative Evaluation of the Proposed Image Segmentation Approach

The proposed image segmentation methodology needs to be evaluated qualitatively and quantitatively, mainly for the natural outdoor images. Berkeley Segmentation Dataset(BSD), a collection of 300 outdoor natural images has been used for our experiments, as it is very popular among vision researchers for object segmentation. The other datasets available for natural images can be found in Appendix B. BSD provides groundtruths based on 5-7 human experts, which proves to be useful, for quantitative evaluation and perform further comparisons with other state-of-art algorithms available in the literature. The performance metrics used are PRI, BDE, VOI and GCE. All these four measures are mostly reported by researchers in the present decade as it collectively gives a rough estimate of various aspects such as area consistency, boundary error, global consistency measures with the groundtruths. The code for computing these four quantitative measures have been taken from Berkeley website [1]. It is a standard practice to use the same program for assessment, so that all the parameter settings remain same for all the algorithms being assessed. In this code, the benchmark images are normalized to have longer side equal to 314 pixels. So before performing any processing on the input image, we have normalized it to have longer side equal to 314 pixels.

The rest of this chapter is organized as follows. The four performance metrics used are briefly stated in Section 4.1. Section 4.2 discusses the experimental setup. Further sections Section 4.3 to Section 4.10 exhibits the various strengths of our methodology and its limitations in few cases. Section 4.3 experiments with HCS and RGB color-only features. Section 4.4 experiments OP with other color spaces to justify the use of HCS

[1] `http://www.eecs.berkeley.edu/yang/software/lossysegmentation/`

components. Section 4.5 compares the OP3 and OP5 with Gabor filters for segmentation. Section 4.6 discusses the influence of multiple ground truths on the quantitative evaluation of the proposed methodology, when most experts favor over-segmentation or under-segmentation. Section 4.7 performs an analysis with HMC, a state-of-art method which uses 6 color spaces. Section 4.8 reflects both the strength and weakness of our methodology, for images under heterogeneous highly textured image type. Section 4.9 shows visual comparisons for a fixed set of images with 12 methodologies found in literature. Section 4.10 makes an comparative analysis quantitatively with 13 state-of-art methods. Section 4.11 summarizes and concludes with future work.

4.1 Performance Evaluation Measures

Lots of research efforts have been made in development of algorithms for quantitative measures in the area of color-texture segmentation. The evaluation of segmentation accuracy is not trivial in case of natural image segmentation as it includes a lot of subjectivity. Till date, the most common and easy method to test the effectiveness is to visually compare against a group of results obtained from segmentation results. This process is tedious and infeasible to compare hundreds of images for all possible segmentation results. Another way is supervised evaluation where assessment is done by comparing with a reference segmented image called ground-truth data provided by multiple human subjects. The perceptual grouping of objects vary for humans, hence one ground-truth cannot be considered as the only solution especially in color texture natural images. In real-time applications, no ground-truths are available and in such cases unsupervised or stand-alone evaluation methods are employed to assess the empirical goodness of the algorithm which matches a comprehensive set of characteristics of good segmentation desired by humans. The criteria for good segmentation proposed by Haralick and Shapiro (1985) are as follows:

1. Regions should be uniform and at the same time homogeneous with respect to one or more characteristics.

2. Regions that are adjacent should have significant differences with respect to the characteristics on which they are homogeneous.

3. Inside of the regions should be without holes.

4. Boundaries should be smooth and not ragged.

The third point is usually not applicable for highly textured images (bushes) when unsupervised segmentation is performed. A prior knowledge about the semantic of the object may satisfy the third condition. For natural images, the fourth point may not be applicable.

Due to the complexity and highly subjective representation of natural images, statistical comparative evaluation with multiple ground-truths for a natural image became more popular and gained large acceptance in computer vision community. We have focused on the metrics that quantify the quality of color-texture image segmentation methods in accordance to the ground-truth data. These metrics have been mentioned in details by Ilea and Whelan (2011) such as local refinement error, global consistency error, local consistency error, boundary displacement error, variation of information, rand index, probabilistic rand index, normalized probabilistic rand, earth mover distance, distance distribution signatures, hamming distance, precision and recall, F-measure and ROC curves.

The set of four metrics namely, Probability Rand Index (PRI) by Unnikrishnan *et al.* (2005), Variation of Information (VOI) by Meilă (2005), Global Consistency Error (GCE) by Martin *et al.* (2001) and Boundary Displacement Error (BDE) by Freixenet *et al.* (2002) are popular and most widely used metrics for benchmarking the performance of the segmentation algorithms. Quantitatively, the higher PRI and lower VOI, GCE and BDE are considered to be better for segmentation accuracy. PRI is the fraction of pair of pixels whose labeling is consistent between the segmented output and averaging of multiple groundtruths. BDE considers only the displacement in boundary pixels. VOI metric is quite different method of comparing segmentations which is based on information theoretical concept. The average conditional entropy of one segmentation given the other is measured by metric called variation of information. Global consistency error tolerates and perceives the under segmentation and over segmentation as refinement of each other. The Chapter 4 uses these set of the above mentioned four metrics for comparison of our segmentation results with state-of-art algorithms.

Recently a contest has been held by ICPR [1] with an aim to unify the assessment methodology used in the image segmentation community. The unsupervised segmentation method by authors is tested against Prague benchmark dataset consisting of random mosaics filled with natural color textures and is evaluated using 21 benchmark criteria. The criteria is categorized into four groups - region based, pixel-based, consistency error measure and clustering comparison. The region based criteria are correct detection, over-segmentation, under-segmentation, missed error, noise error. The pixel-wise weighted average criteria are namely omission error, commission error, weighted average class accuracy, recall, precision, type 1 error , type 2 error, mean class accuracy estimate, mapping score, root mean score proportion estimation error, comparison index and F-measure The consistency error consists of local consistency error and global consistency error. Finally the last group clustering comparison criteria consists of variation of information, Mirkin metric, Van Dongen metric. The next section discusses the experimental setup of our proposed methodology, lists of parameters used and their purpose.

4.2 Experimental Setup

In our experiment, we have tested both OP3-HCS and OP5-HCS for segmentation on the Berkley Standard Dataset (BSD300). Our experiments have one external parameter $P1$ i.e OP of order 3 or 5. Rest of the seven are internal parameters or control parameters ($P2 - P8$), which have been set initially before running our segmentation program. The same values are set for all BSD300 images, and are not tuned specifically for any

[1] http://mosaic.utia.cas.cz/icpr2014/

particular image. No training has been provided for parameter tuning. The purpose of the control parameters and their predetermined values is mentioned as follows. $P1$ is set to 5 or 3, it controls the OP size for the texture framework. $P2$ is the maximum percent variation captured in features, set to 99$\%$ and it controls the adaptive feature vector representation. $P3$ is relative region size in percentage and is set to 0.2. $P4$ is maximum number of clusters allowed which is set to 15. $P5$ is threshold for mean distribution difference with value 900. $P6$ is threshold for K-S statistic with value 0.5. The parameters $P3, P4, P5, P6$ controls the stopping criteria for the iterative K-means clustering. $P7$ helps in identifying very small region for region merge with deviation threshold set to 90$\%$. The threshold parameter $P8$ is used for fusing small regions having similarity less than 1000. The algorithm is provided without any preprocessing step, image smoothing or initial training strategy for parameter tuning that could have significantly given a definite advantage. The following sections addresses various aspects of our methodology based on experimental evidences qualitatively and quantitatively.

4.3 Analysis of HCS and RGB color-only features for image segmentation

The experimental results of segmentation with color-only feature vector is shown in Table 4.1. HCS color space proved better than RGB with our proposed segmentation method in all the four measures.

Table 4.1: Statistics PRI, BDE, VOI, and GCE for segmentation of BSD300 dataset using only color spaces without texture. Higher PRI and lower values of BDE,VOI and GCE represent better segmentation.

Color Space	PRI	BDE	VOI	GCE	Time(sec)
RGB	0.6358	13.8854	2.2805	0.2182	36
HCS	**0.6649**	**13.6600**	**2.1471**	**0.1904**	40

Table 4.2: Statistics PRI, BDE, VOI, and GCE for segmentation of BSD300 dataset, comparing OP3 and OP5 with various color spaces. Higher PRI and lower values of BDE,VOI and GCE represent better segmentation.

OPsize	Color Space	PRI	BDE	VOI	GCE	Time(sec)
OP3	**HCS**	**0.7457**	10.0457	2.4113	0.2857	38
OP5	**HCS**	0.7431	9.6806	2.8832	**0.2496**	48
OP3	**RGB**	0.6818	12.1180	**2.3539**	0.2604	37
OP5	**RGB**	0.7377	**9.3723**	3.0320	0.2894	43
OP3	L*a*b*	0.6790	12.1082	2.3695	0.2698	37
OP5	L*a*b*	0.7311	9.9270	2.8926	0.3099	46
OP3	HCS(without CMYK)	0.6646	13.9238	2.3967	0.2641	37
OP5	HCS(without CMYK)	0.7245	10.4567	2.8332	0.2953	49

4.4 Analysis of color space experiments in combination with OP3 and OP5

Nevertheless, Table 4.2 justifies the advantage of color-texture feature integration in terms of PRI and BDE, to a considerable extent as compared to having only color features. PRI metric increases from 0.6358 (only RGB) to 0.7377 (OP5-RGB), and from 0.6649 (only HCS) to 0.7457 (OP3-HCS) and 0.7431 (OP5-HCS), depending on size of OP. Experiments are conducted separately on HCS, RGB, L*a*b* and HCS without CMYK (PCA on color channels a*, b* and H, S) along with OP3 and OP5. We can definitely see that HCS using both color dominant channels and CMYK, performs very well with both OP3 and OP5, and is competitive in terms of PRI and GCE measures as compared to other color spaces. PRI measure is usually better in case of OP5, as it captures more contrasting variations. On the other hand VOI measures are better with OP3 as compared with OP5, because VOI favors under-segmentation and the randomness and uncertainty is under control using OP3, as the contrasting variations captured are coarser and not at finer level as in case of OP5. So we can conclude, that feature integration OP3-HCS is ideal for coarser level image segmentation and can be applied for high-level recognition tasks. The objects which are intervened in feature space, need finer analysis with OP5-HCS. Thus non-linearity gets captured using larger window size and may be an ideal option for texture characterization.

4.5 Comparison of OP3 and OP5 with bank of Gabor filters considering RGB and HCS color spaces

We have also experimented with Gabor filter bank, built from five scales corresponding to five centre frequencies and four orientations, on three color channels. The Gabor filter scale and orientation parameter settings are taken from Jain and Farrokhnia (1991). The 20 Gabor filters when convoluted with 3 color planes of RGB and HCS, result into 60 filtered responses. These responses are then subjected to our iterative segmentation process. The results are shown in Table 4.3. OP3 and OP5 have performed considerably

Table 4.3: Statistics PRI,BDE,VOI and GCE for segmentation of BSD300 dataset, comparing Gabor filter bank with OP based filters. Higher PRI and lower values of BDE,VOI and GCE represent better segmentation.

Filter Bank	Color Space	PRI	BDE	VOI	GCE	Time(sec)
OP3	HCS	**0.7457**	10.0457	2.4113	0.2857	38
OP5	HCS	0.7431	9.6806	2.8832	0.2496	48
OP5	RGB	0.7377	**9.3723**	3.0320	0.2894	43
Gabor	HCS	0.6291	13.9447	2.3209	0.2485	38
Gabor	RGB	0.6340	14.6331	**2.2793**	**0.2167**	37

better than Gabor filters, for RGB and HCS color spaces in terms of PRI and BDE, and results of VOI and GCE are found to be comparable. Time taken does not have much of a difference. Thus we can see that filter banks generated from the orthogonal polynomials are wholesome, and better captures the tonal and textured variations in an image using our proposed framework which is directly reflected by steep hike in PRI.

4.6 Interpretation of the quantitative evaluation in terms of multiple ground truths

Nature segmentation by humans are highly contextual sensitive. Same scenes are biased toward over-segmentation or under-segmentation because human poses some domain knowledge.

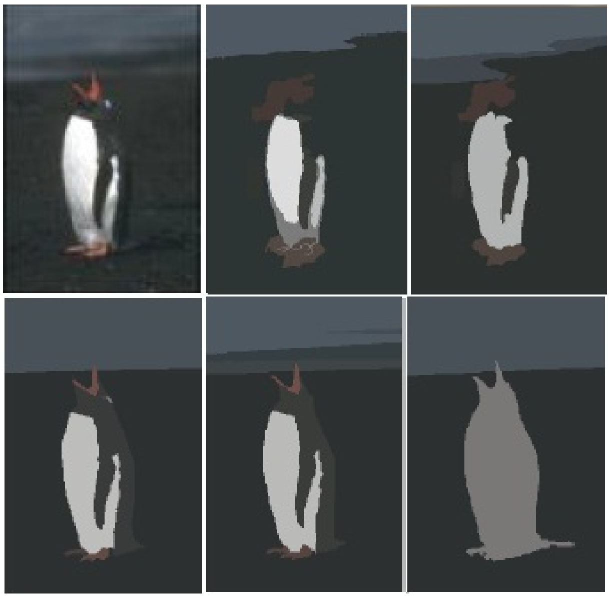


Figure 4.1: BSD Image 106025: Row1- Original, OP3-HCS (6 regions), OP5-HCS(11 regions), Row2,Row3 [L-R] show ground-truth images with 10,11,3,12,13,9 regions

Table 4.4: Statistics PRI,BDE,VOI and GCE for segmentation of image 106025 : experts favoring over-segmentation.

OPsize	Color Space	Regions	PRI	BDE	VOI	GCE	Time(sec)
OP3	HCS	6	0.76834	7.1577	1.5144	0.22383	37
OP5	HCS	11	0.81108	5.7093	1.2339	0.15188	41

We have considered two case studies to compare the effects of OP3-HCS and OP5-HCS, in scenarios where human experts differ. Let us take the case of the BSD image 106025.jpg. Multiple ground truths exist, mostly which prefer over-segmentation. There are six ground truths available with regions 10, 11, 3, 12, 13 and 9 respectively as shown in Figure 4.1. We can see that experts preferred to segment in more than 6 regions, hence over-segmentation using OP5-HCS gave better results in all the four quantitative measures as compared to OP3-HCS. Our segmentation results are shown in Table 4.4.

In another case 0f BSD 12003.jpg, most of experts favor under-segmentation. There are five ground truths with varying number of regions - 6, 6, 98, 6 and 7 for each expert respectively as shown in Figure 4.2. We can see that many experts preferred to segment

Figure 4.2: BSD Image 12003: Row1- Original, OP3-HCS (27 regions), OP5-HCS(109 regions), Row2,Row3 [L-R] show ground-truth images with 6,6,98,6,7 regions

Table 4.5: Statistics PRI,BDE,VOI and GCE for segmentation of image 12003 :experts favoring undersegmentation.

OPsize	Color Space	Regions	PRI	BDE	VOI	GCE	Time(sec)
OP3	HCS	27	0.78262	4.7287	2.2114	0.12343	37
OP5	HCS	109	0.72783	6.5664	3.8819	0.11216	49

in less than 8 regions except one, hence segmentation using OP3-HCS produced better results as shown in Table 4.5, as compared to OP5-HCS. Analysis shows that OP5-HCS is biased towards over-segmentation as compared to OP3-HCS.

4.7 Analysis of multiple color spaces as compared to RGB

This section cites the importance of multiple color spaces, as compared to RGB in color-texture segmentation methods. Hedjam and Mignotte (2009) proposes a hierarchical graph based Markovian clustering (HMC) using fixed number of 6 classes. The Table 4.6 shows results performed with only one color space (RGB), and comparison with our proposed technique of feature extraction on RGB (i.e. RGB+OP5). OP5-RGB performs much better in three metrics BDE, VOI, GCE and with PRI value at par. The

Table 4.6: Statistics PRI,BDE,VOI and GCE for comparison with Hedjam and Mignotte (2009). Higher PRI and lower values of BDE,VOI and GCE represent better segmentation.

Technique	Color Space	PRI	BDE	VOI	GCE
HMC	RGB	**0.7419**	10.4570	3.8315	0.3073
Ours-OP3	**RGB**	0.6818	12.1180	**2.3539**	**0.2604**
Ours-OP5	RGB	0.7377	**9.3723**	3.0320	0.2894
HMC	6 color spaces	**0.7835**	**9.5700**	3.9900	0.2900
Ours-OP5	HCS	0.7431	9.6806	2.8832	**0.2496**
Ours-OP3	HCS	0.7457	10.0457	**2.4113**	0.2857

paper Hedjam and Mignotte (2009) has reported better results using six color spaces (RGB, HSV, LUV, YIQ, XYZ and LAB), the results are shown in the same table. Our proposed HCS along with OP outperforms in terms of two metrics, VOI and GCE. PRI reported in HMC is much better, nevertheless BDE is at par for HCS-OP5. It can be further noted that use of multiple color spaces proved to be better than single color space(RGB) for PRI measure, irrespective of technique(HMC or OP-HCS) used.

4.8 Study of Heterogeneous highly textured images in BSD images

Section 3.5.1 discusses the image types, that can be identified by our segmentation framework. When the 300 images of BSD are run using OP5-HCS, the method identifies 20 images as heterogeneous highly textured images and rest are identified as object-based image. The OP5-HCS favors over-segmentation, hence it was able to extract the heterogeneous highly textured image type and OP3-HCS could not identify except 242078.jpg (rest are object-based images). The Table 4.7 shows that out of these 20 images, 9 images show very good segmentation results, based on PRI as illustrated in Figure 4.3.

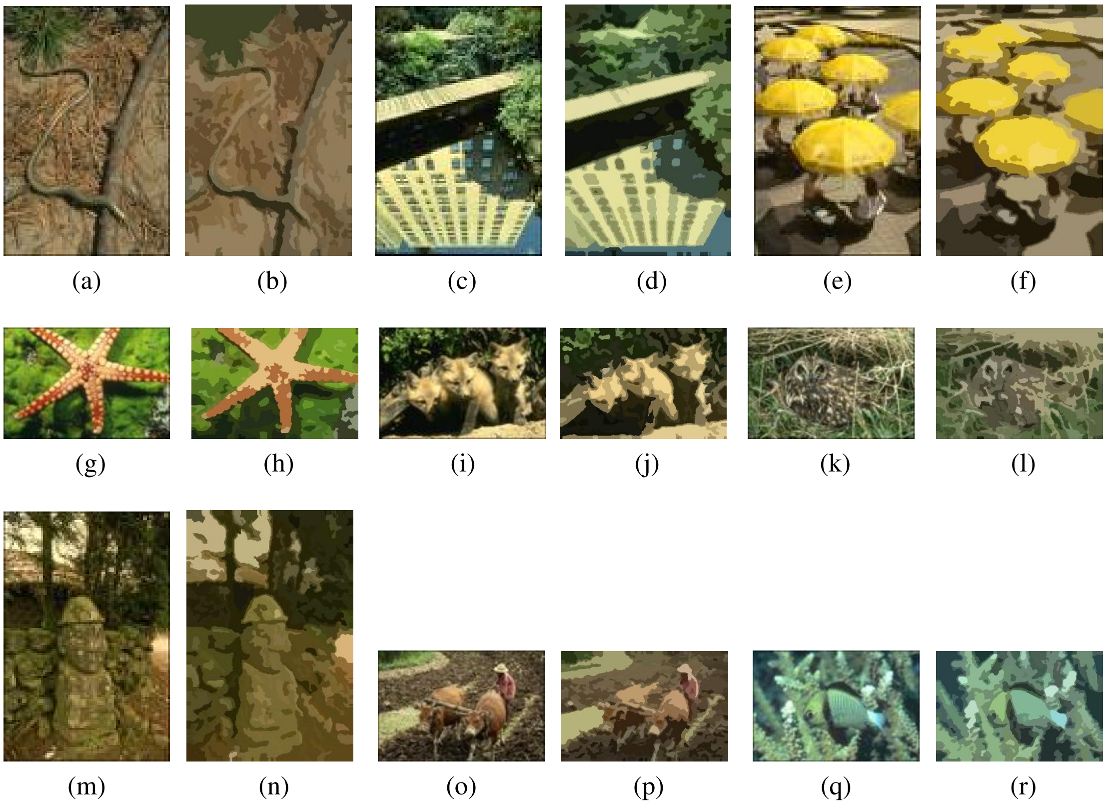


Figure 4.3: BSD Highly textured images with **GOOD** segmentation measures, using OP5-HCS variant . Each set of image shows the original and its corresponding segmented regions. BSD Images: [a-b] 175032, [c-d] 148026, [e-f] 242078, [g-h] 12003, [i-j] 159008, [k-l] 8143, [m-n] 55073, [o-p] 202012, [q-r] 209070

Table 4.7: Performance of Heterogeneous highly textured images identified by OP5-HCS in BSD dataset, which proved to be **Good**, based on both **PRI** and **BDE** measures.

BSD Images	PRI	BDE	VOI	GCE
175032	0.71182	8.1844	4.8858	0.16775
159008	0.71819	14.2584	4.3353	0.23482
202012	0.72783	7.2976	5.4051	0.19531
12003	0.72798	6.5848	3.8726	0.11102
148026	0.77654	11.2297	3.7893	0.14401
8143	0.77928	5.7752	5.3623	0.26369
55073	0.81095	8.826	4.69	0.23055
242078	0.82272	5.286	3.8533	0.26833
209070	0.83113	5.279	4.2911	0.261

The Table 4.8 shows that 11 show poor results due to mostly camouflage problem as illustrated in Figure 4.4.

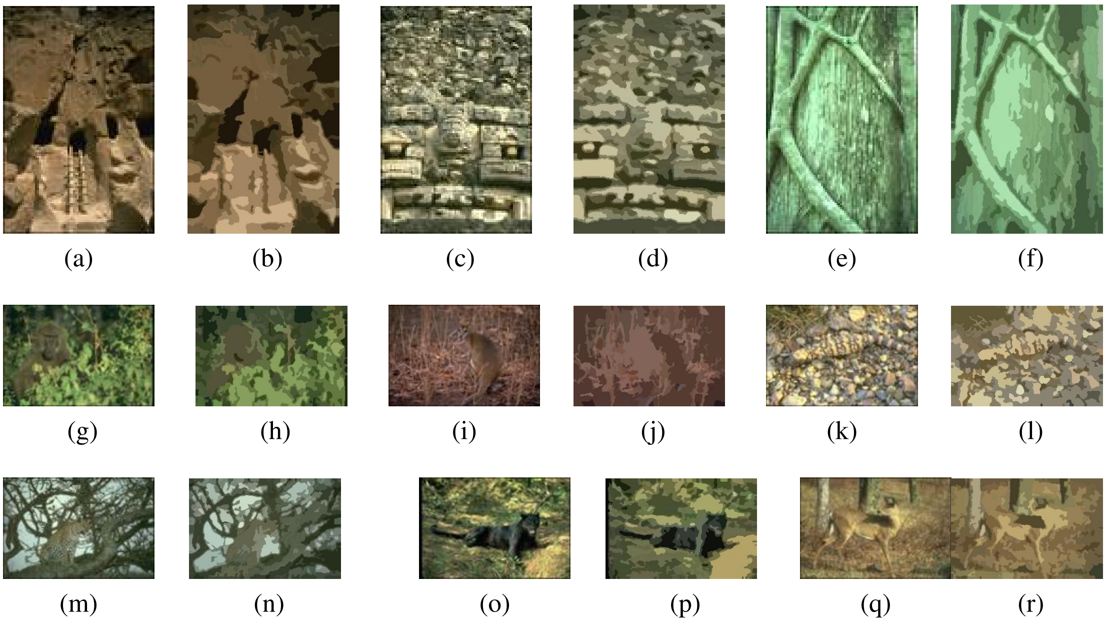
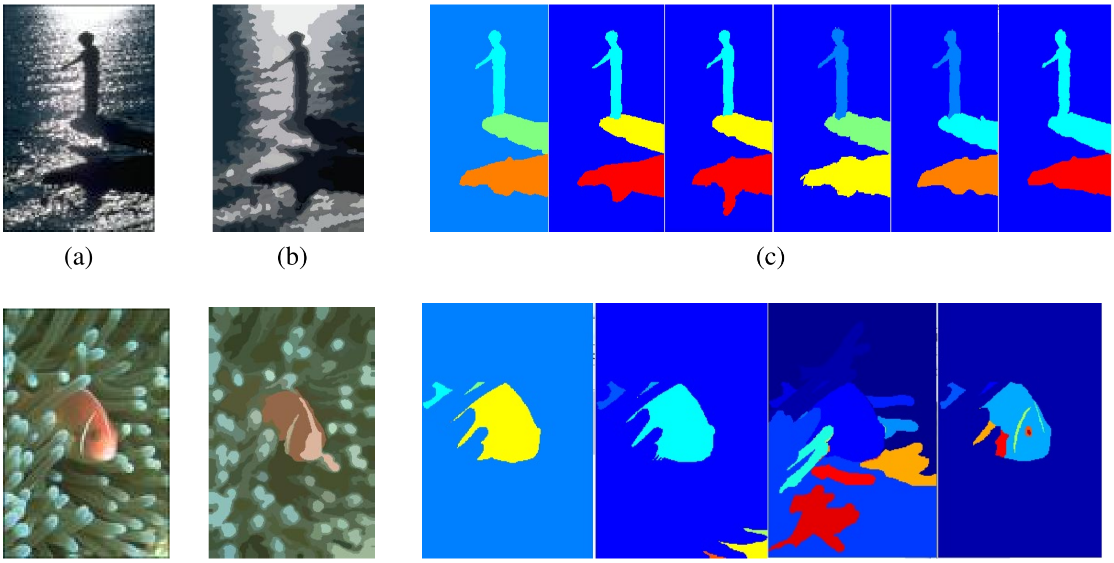


Figure 4.4: BSD Highly textured images with **POOR** segmentation measures, using OP5-HCS. Each set shows original and its segmented regions. BSD Images: [a-b] 54005, [c-d] 33039, [e-f] 311081, [g-h] 16052, [i-j] 69040, [k-l] 87065, [m-n] 134035, [o-p] 304034, [q-r] 236017

Table 4.8: Performance of Heterogeneous highly textured images identified by OP5-HCS in BSD, which are found to be **Poor** mainly based on **PRI** and **BDE**.

BSD Images	PRI	BDE	VOI	GCE
210088	0.32496	21.6236	5.4629	0.10464
26031	0.35743	13.4835	5.4509	0.15675
304034	0.3986	16.3628	5.4847	0.064347
54005	0.40791	15.9174	5.1302	0.14464
69040	0.40855	19.7584	5.066	0.18739
16052	0.53515	18.5165	4.6433	0.14968
311081	0.58365	20.3227	5.1591	0.0829
87065	0.58717	12.5842	5.1108	0.17965
134035	0.59109	12.106	5.8837	0.19724
236017	0.64137	7.6809	4.8321	0.25506
33039	0.75863	18.0801	5.6035	0.14608

Two images *26031.jpg* and *210088.jpg* are considered to be special cases, where the human experts have identified one prime object from the scene, and ignored the textural



(a) (b) (c)

(d) (e) (f)

Figure 4.5: BSD Highly textured images with **POOR** segmentation measures due to experts biased towards main object. Each set of image shows the original and its corresponding segmented regions and the experts segmentation. BSD Images: [a-c] 26031, [d-f] 210088

variations as can be seen in Figure 4.5, which degrades the performance measures. The segmentation output is shown as regions, depicted by nearby color of the segments, obtained from the original image. This gives a true picture of the spatial coherence obtained due to homogeneity in the color-texture features.

4.9 Visual Comparisons with few State-of-art segmentation methods

The survey by Vantaram *et al.* (2012) does a visual comparison with different segmentation algorithms. We have considered these segmentation results from 12 papers, taken from the survey paper and performed a subjective comparison, with our proposed two variants OP3-HCS and OP5-HCS as shown in Figure 4.6. Figure 4.6 a) and b) illustrates the results of our proposed OP3-HCS and OP5-HCS segmentation respectively. The results for the algorithm on mean-shift clustering, guided by edge detection known

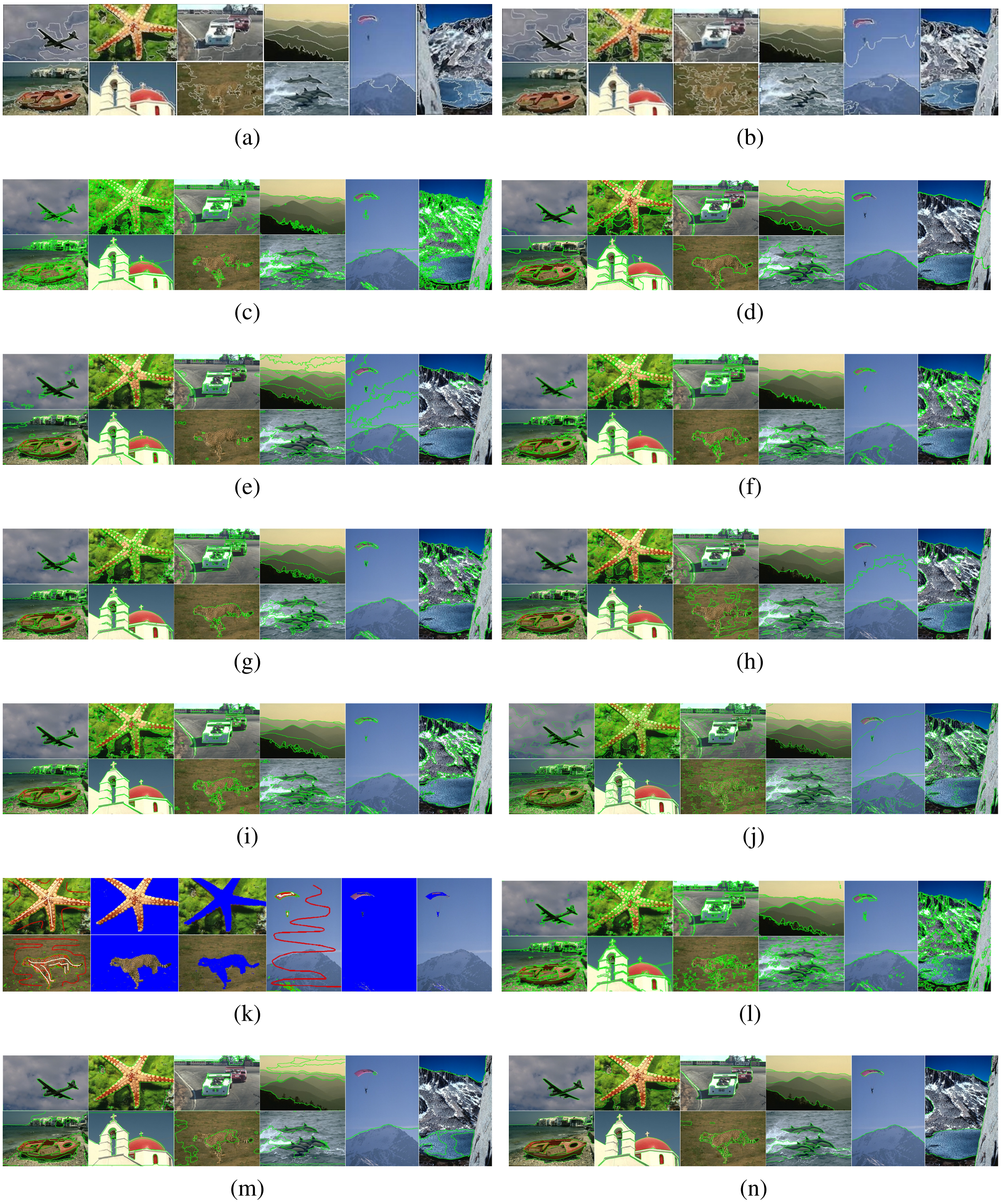


Figure 4.6: Visual Comparison of Segmentation Results a) OP3-HCS, b) OP5-HCS c) EDISON d) CTM e) DCGT f) GSEG g) MAPGSEG h) JSEG i) GRF j) LevelSet k) Interactive Graph Cuts l) Graph based m) gpb-OWT-ucm n) Hoang

as EDISON system, is developed by Christoudias *et al.* (2002). The Figure 4.6 c) shows that some images are over-segmented, and some are under-segmented. The Figure 4.6

d) shows the results of algorithm known as CTM by Allen *et al.* (2008), which are considerably good. Figure 4.6 e) shows the results of a region growing algorithm, known as Dynamic Color Gradient Thresholding(DCGT) by Balasubramanian *et al.* (2008). The Figure 4.6 f) depicts the segmentation results of gradient segmentation, known as GSEG proposed by Ugarriza *et al.* (2009). The algorithm MAPGSEG is an extension of GSEG developed by Vantaram *et al.* (2010), and reduces the computational expense by half, which is shown in Figure 4.6 g). Next Figure 4.6 h) shows JSEG algorithm results by Deng and Manjunath (2001). Figure 4.6 i) shows results obtained using energy based segmentation techniques, based on Gibbs random field (GRF) and spit-and-merge technique, by Vantaram and Saber (2011). The level set algorithm is based on active contour, whose results are obtained using toolbox by Sumengen (2005), shown in Figure 4.6 j). Figure 4.6 k) results are obtained using graph-based segmentation techniques, proposed by Boykov and Jolly (2001), which is interactive, and the image is divided into foreground and background image content. The researchers Felzenszwalb and Huttenlocher (2004) developed graph-based technique, based on tree structure, shown in Figure 4.6 l). Segmentation based on watershed protocol, using hybrid region and edge is used by Arbeláez and Fowlkes (2011), in developing oriented watershed transform (OWT), and the algorithm is known as gPb-owt-ucm whose results are shown in Figure 4.6 m). Hoang *et al.* (2005) segmentation method is robust to shading and highlighting effects, whose results are shown in Figure 4.6 n). Our segmentation is biased towards over-segmentation(plane, and is able to capture the texture-tonal variations(star fish, mountains,church), except in case in cases where camouflage appears (tiger).

4.10 Comparative analysis with state-of-art methods

Finally, an important remark of our experimental investigation is comparison with some state-of-art methods present in the literature. We have considered two variants of our proposed framework, OP3-HCS and OP5-HCS. For comparison, we present quantitative comparisons in Table 4.9 with 13 color-texture segmentation algorithms.

Table 4.9: Statistics PRI, BDE, VOI and GCE for segmentation of BSD300 dataset for comparing OP3-HCS and OP5-HCS with 11 segmentation methods. Higher PRI and lower values of BDE,VOI and GCE represent better segmentation. Abbreviations:MDS-Multi-Dimensional Scaling, CTM - Compression based Texture Merging, NTP - Normalized Tree Partitioning, GBIS - Graph Based Image Segmentation, SC - Spectral Clustering, GBMS - Gaussian Blurring Mean Shift, NCuts - Normalized Cuts, JND - Just Noticeable Difference histogram, DCM- Dirichlet Compound Multinomial distribution, HMC - Hierarchical graph-based Markovian Clustering, MIS - Modal Image Segmentation. Best segmentation results are reported in bold.

Segmentation Methods	PRI	BDE	VOI	GCE
HUMANS Allen *et al.* (2008)	0.8750	4.9900	1.1000	0.0790
OP3-HCS Goswami *et al.* (2014*a*)	0.7457	10.0457	2.4113	0.2857
OP5-HCS Goswami *et al.* (2015)	0.7431	9.6806	2.8832	0.2496
MDS Mignotte (2012)	**0.8110**	7.951	2.0040	0.2050
MIS Krinidas and Pitas (2009)	0.7900	**7.8200**	**1.9300**	0.1900
HMC Hedjam and Mignotte (2009)	0.7835	9.5700	3.9900	0.2900
CTM Allen *et al.* (2008)	0.7620	9.8960	2.0240	0.1880
MeanShift Dorin and Peter (2002)	0.7550	9.7001	2.4770	0.2598
NTP Wang *et al.* (2008)	0.7521	16.3000	2.4954	0.2373
GBMS Carreira-Perpinán (2006)	0.7362	14.8480	3.5701	0.4405
SpectralClustering(SC) Yu and Shi (2003)	0.7357	15.4000	2.6339	0.2469
NCuts Shi and Malik (2000)	0.7229	9.6038	2.9329	0.2182
JND Bhoyar and Kakde (2010)	0.7190	-	-	-
GBIS Shi and Malik (2000)Wang *et al.* (2008)	0.7139	16.6700	3.3949	**0.1746**
DCM Nikou *et al.* (2010)	0.7080	15.4700	3.5376	0.4339
Bit-stream based Mecimore and Creusere (2009)	0.7031	-	3.4970	0.3107
Our Ranking considering each measure independently (OP3-HCS)	$\frac{7}{14}$	$\frac{7}{14}$	$\frac{4}{14}$	$\frac{9}{14}$
Our Ranking considering each measure independently (OP5-HCS)	$\frac{7}{14}$	$\frac{7}{14}$	$\frac{7}{14}$	$\frac{8}{14}$

They are Normalized Tree Partitioning (NTP) by Wang *et al.* (2008), Normalized Cut (NCut) by Shi and Malik (2000), Hierarchical graph-based Clustering by Hedjam and Mignotte (2009), multi-class spectral clustering (SC) by Yu and Shi (2003) and graph based image segmentation(GBIS) Felzenszwalb and Huttenlocher (2004), modeling pdf using Dirichlet compound multinomial distribution (DCM) by Nikou *et al.* (2010), Gaussian blurring mean-shift (GBMS) by Carreira-Perpinán (2006), mean-shift

(MS) by Dorin and Peter (2002), clustering multivariate mixed data (CTM) by Allen *et al.* (2008), bit domain based by Mecimore and Creusere (2009) and Just Noticeable Difference (JND) method by Bhoyar and Kakde (2010), deformable energy function by Krinidas and Pitas (2009) and combining contour and texture approach by Mignotte (2012). Our proposed method does not involve any smoothing process or any pre-processing step. PRI and BDE favor over-segmentation, whereas VOI and GCE favor under-segmentation. In order to achieve segmentation accuracy, higher PRI should reflect lower BDE. This balance act is reflected from our PRI and BDE values obtained. The comparative analysis justifies that our proposed method strikes a good balance between over-segmentation and under-segmentation. Though GBIS Wang *et al.* (2008) shows best result in GCE, our method performs better than it in rest of the three metrics PRI, BDE and VOI by a considerable extent. We expect to improve our GCE by developing a more intuitive region merge technique so that the refinement of segments from coarse-to-fine is captured properly. We prefer OP3-HCS for segmenting BSD 300 images with higher PRI as compared to OP5-HCS. We found 12 images with poor PRI

Figure 4.7: BSD Segmentation results with PRI < 0.5 using OP3-HCS variant

measure segmented by OP3-HCS less than 50 percentage, which are shown in Figure 4.7. Based on OP5-HCS, 18 images were found to have PRI below 50 percent. More segmentation results are shown with PRI greater than 0.75 in Figure 4.8 for OP3-HCS variant, and Figure 4.9 for OP5-HCS variant. are considered to be better for segmentation accuracy. Our segmentation model on an average takes 43 seconds to segment an image of BSD300 dataset on P600 Intel Pentium CPU, 2.00 GHz, 2.00 GB RAM when run on Windows system. Our code is written using Matlab.

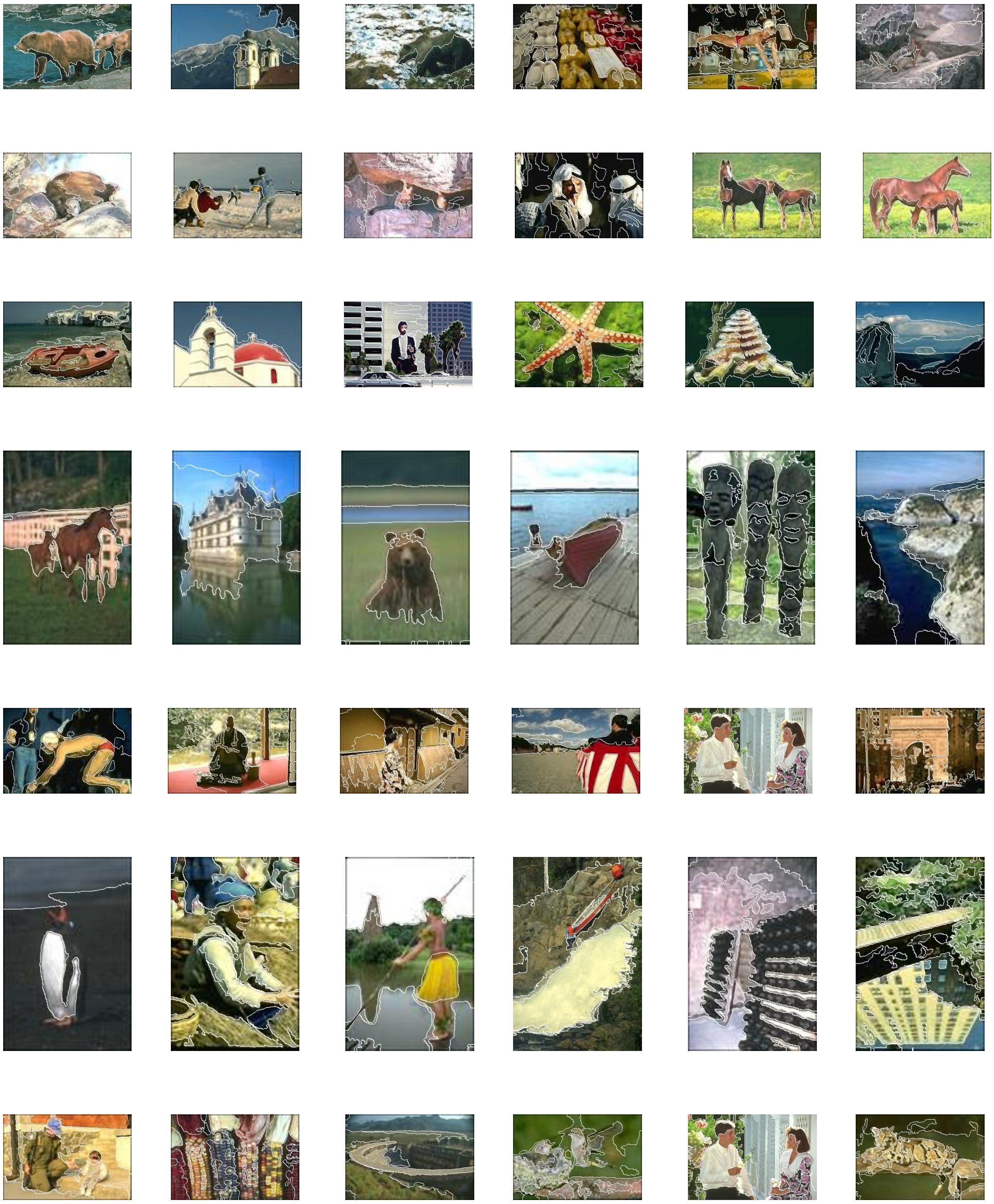


Figure 4.8: Few BSD Segmentation Results using OP3-HCS variant

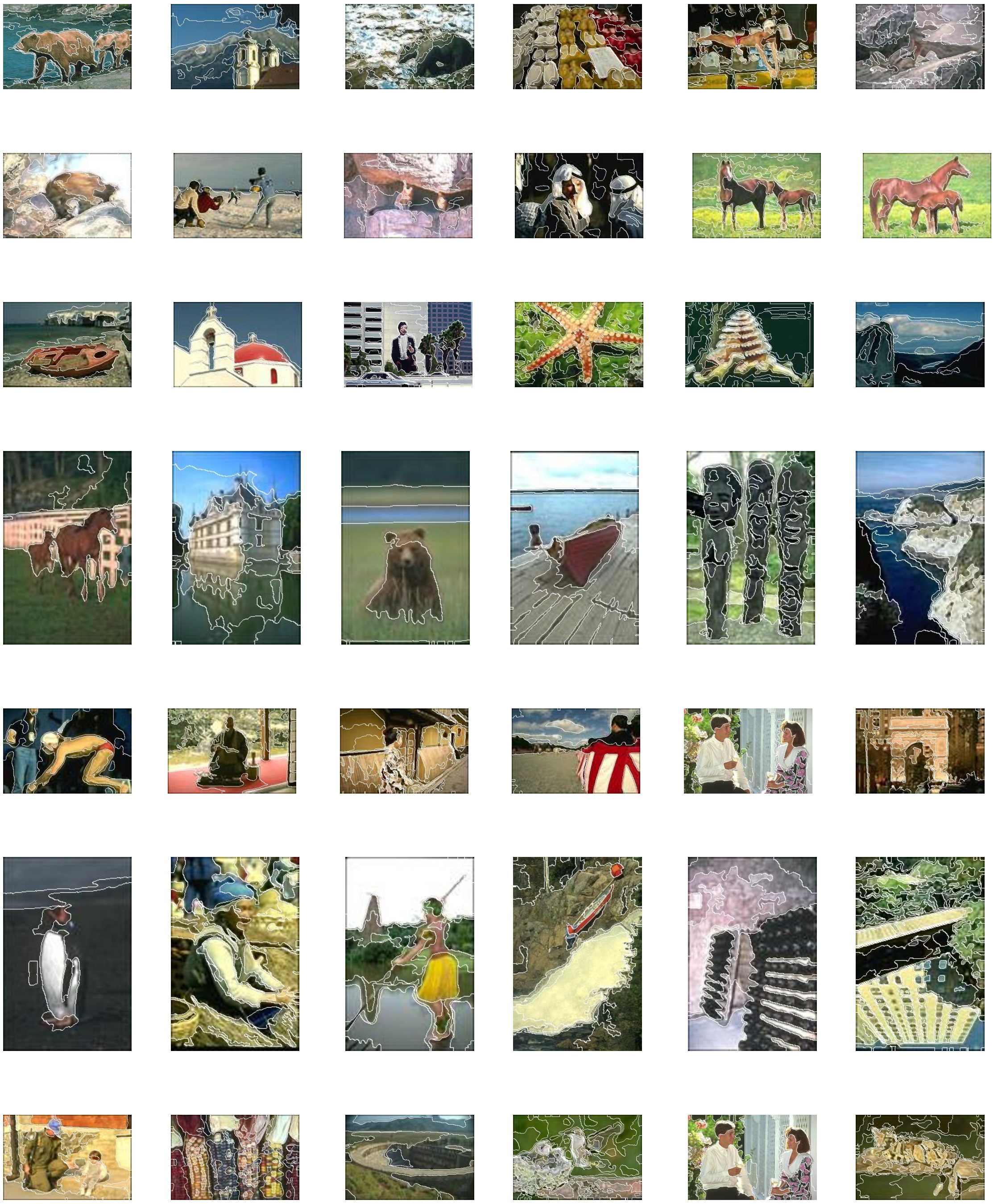


Figure 4.9: Few BSD Segmentation Results using OP5-HCS variant

4.11 Summary and Conclusions

Experimental results demonstrate the effectiveness of the hybrid region and edge based color-texture unsupervised segmentation, based on the four most popular metrics used for segmentation evaluation, namely PRI, BDE, VOI and GCE. The two variants of OP-HCS, namely OP3-HCS and OP5-HCS segmentation are both successfully tested quantitatively on 300 natural images taken from Berkeley Standard Dataset. No training is given to the process for parameter tuning. The internal parameter settings are fixed for all images, except the external parameter, which determines the OP size 3 or 5. Segmentation results exhibit the effectiveness by achieving on an average 74 percent on PRI which is quite encouraging without involving any preprocessing, training or smoothing processes. The experimental findings are as follows - 1) HCS proved to be better than RGB color space. 2) OP3-HCS is better for coarser level image segmentation and can be applied for high-level recognition tasks. 3) OP5-HCS is biased towards over-segmentation as non-linearity gets captured using larger window size, and may be an ideal option for texture characterization. 4) OP3 and OP5 have performed considerably better than Gabor for most of the color spaces in terms of PRI and BDE and results of VOI and GCE are found to be comparable. 5) OP5-HCS is recommended for identifying heterogeneous highly textured image types as it favors over-segmentation. Our method is able to segment considerably well in some cases of highly textured images where the color variations are prominent. It has limitations and shows poor results due to mostly camouflage problem. 6) Visual comparisons with few BSD images with some recent and popular state-of-art algorithms, justifies that our method is capable of separating texture-tonal mix, prefers over-segmentation, and is comparable with some except that CTM, GSEG, MAPGSES, gpb-OWT-ucm and Hoang perform much better in the tiger image having camouflage. 7) OP3-HCS and OP5-HCS are compared to 13 recent and popular state-of-art methods - four graph based clustering methods, four clustering based on probability density function and mixture models, bit domain method, fusing contour and gradient features method, deformable energy function and a color histogram method. We understand that the list is not exhaustive; still it gives us a comparative estimate with other segmentation methods. The four metric measures are found to be quite encouraging especially performed in a time-efficient manner (approx. 43 seconds). In future work, intelligent region merge based on statistical metadata will help in improving the VOI and GCE to a great extent.

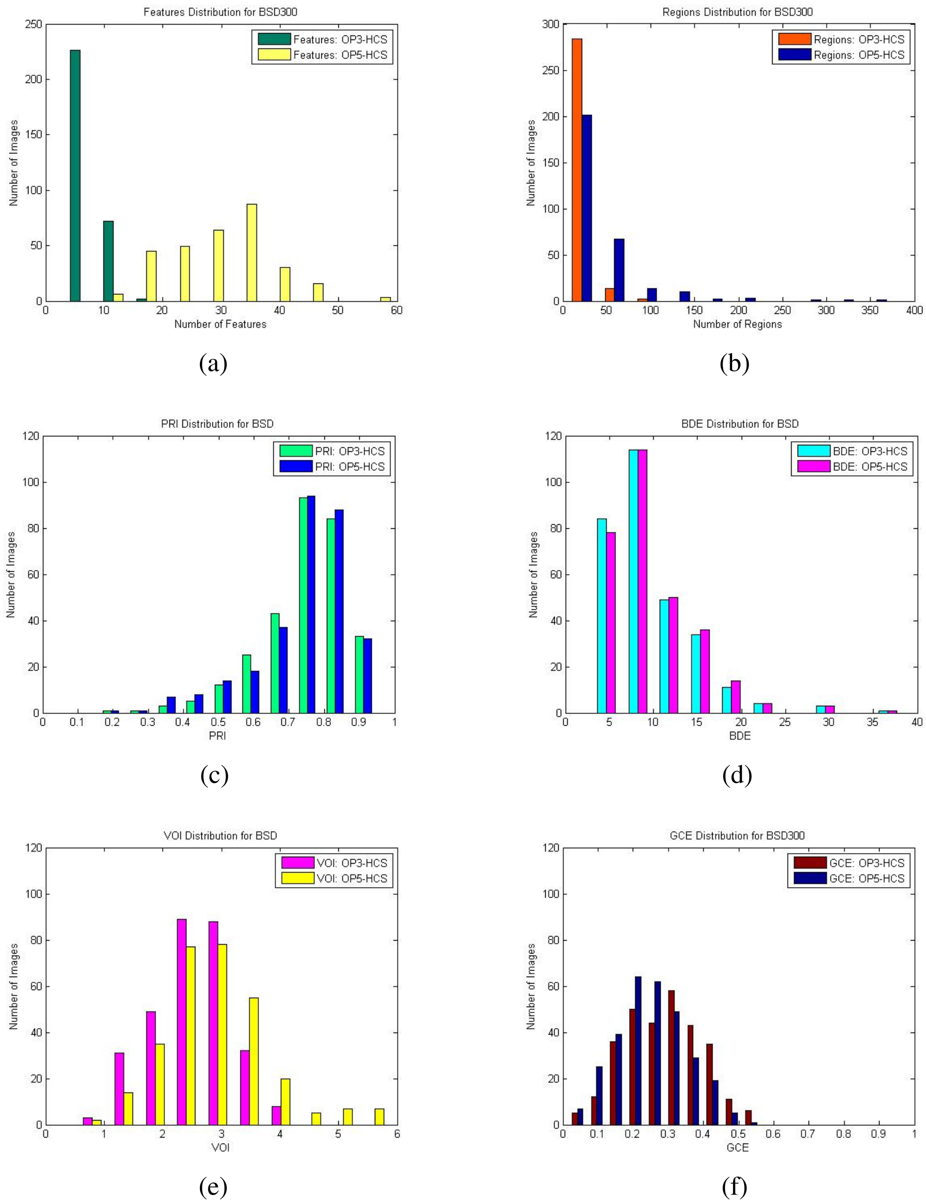


Figure 4.10: Histogram Distributions for BSD 300 images: a) Feature distribution, b) Region distribution, c) PRI distribution, d) BDE distribution, e) VOI distribution, f) GCE distribution

The distribution of BSD 300 images in terms of the adaptive number of features, number of regions obtained and the four quantitative measures are illustrated in the Figure 4.10, when segmented using OP3-HCS and OP5-HCS respectively. It can be clearly noted that OP5-HCS generates more number of regions.

The annotation of Berkeley dataset is meant only for object segmentation. The approach taken by our proposed segmentation method in this chapter is biased towards over-segmentation to extract the object regions. The image segmentation is one aspect of knowing the components present in the image. Another aspect can be, once the object segments are obtained, there is a need to study the local intrinsic characteristics of the objects present. Texture characterization can be effective in building knowledge about the textural content in natural images. In real world scenario, not all images will consist of an object which is very prominently seen. In such cases, the texture characterization can be performed both considering the whole image(globally) and also in terms of collection of image blocks(locally). The next chapter of the thesis proposes a methodology for texture characterization.

CHAPTER 5

Texture Characterization

In this chapter, a novel method is proposed for image texture characterization. Characterization is governed by simple perceptual variations in relative orientations in terms of either no variations present or variations present as row specific, column specific or diagonal specific. This generalization is obtained by modeling the input as a whole/image blocks depending on the broader or narrow coverage respectively. Most of the texture characterization found in literature is done either keeping a specific domain (synthetic or natural images) or is application specific (image retrieval). Need arises to develop techniques free from such subjectives. This chapter develops a methodology for building the essential knowledge for developing objective texture characterization. Texture is a region property not location property. It has the complexity depending on size, shape, topology of region and so on. As this texture is a relative property to its neighborhood pixels, the contrast should speak about its texture and properties. Estimation of a contrast on the associated analysis is well established and demonstrated in the domain of design and analysis of experiments in agricultural statistics using ANOVA methodology as mentioned by Kurz and Benteftifa (2006). Our method is not biased towards any domain or application; rather it acts as a pre-processing for guiding towards locating both non-textural and orientation specific textured image blocks. The proposed method quantifies the texture-tonal characterization of an image/image blocks using statistical ANOVA grading system. Once the grading for abstraction is assigned for both image as a whole and also for image blocks, the decision as to which higher-level algorithms need to be implemented on which block will become easier. The proposed method can be considered to be a three stage process - progressive sampling, image partitioning in blocks, ANOVA analysis and grading.

The rest of this chapter is organized as follows. Section 5.1 gives a brief overview of 2 models of ANOVA - 2-way and 3-way ANOVA with replication. Section 5.2 presents the role of ANOVA in image processing. Section 5.3 emphasizes the need of proposed strategic progressive sampling methodology and its characteristics such as sample size, sampling density and sampling features. Section 5.4 discusses proposed ANOVA grading system for 2-way and 3-way which is later used for knowledge representation of any input image in terms of both string and semantic codes. Section 5.5 experimentally shows the impact of image size on grading. Finally Section 5.6 summarizes and concludes with future work.

5.1 Overview of Analysis of Variance (ANOVA)

ANOVA is related to variance of samples, hence the name *ANalysis Of VAriance*. The procedure ANOVA is like the t test that helps to determine whether the sample means differ enough to conclude that they do not differ just by chance, but there is some real difference among the populations they represent. The ANOVA can be applied to any number of groups (2-way ANOVA, 3-way ANOVA), rather than just two as in the case of the t test. Statistically, three basic types of effects are tested in a two-way ANOVA : main effect for independent variable or factor A, main effect for independent variable or factor B, and effect for the interaction of factors A and B. Groups are levels of a factor. The samples are assumed to be independent and are of same size. The variances of the populations is assumed to be equal. In case of significant interaction, the effect of one factor depends on the level of the other factor. For 2-way ANOVA, there are three null hypotheses. The three null hypotheses are: i) H0A - The population means of the first factor are equal. ii) H0B - The population means of the second factor are equal. iii) H0AB - There is no interaction between the two factors.

The pictorial description of within-variance and between-variance can be seen in Figure 5.1. One can observe that there are two sources of variation - means of groups and spread of data in each group. Each of these variations have degrees of freedom associated with it. In case of between-groups variation, the degree of freedom is one less than number of groups. In case of within-groups variation degrees of freedom is first calculated for each group i.e. one less than number of observations in each group. Finally add the degrees of freedom together. For example, for three groups the within-groups degrees of freedom are $(N1-1)+(N2-1)+(N3-1)$. There is an F-test for each of the hypotheses, and the F-test is the mean square for each main effect and the interaction effect (effect variance) divided by the within-group variance. Each of the variances computed to analyze the main effects are like the between-group variances. Each F-test is based on its own F-Ratio. F-ratio is the ratio of the variance between the group means relative to the amount of variation in the sample (i.e., variation within each of the groups). The main effect of each factor along with the interaction effect thus has separate hypothesis and a separate F-test and each test has its own p-level. A p-value is the probability of observing data as or more extreme as the actual outcome when the null hypothesis is true.

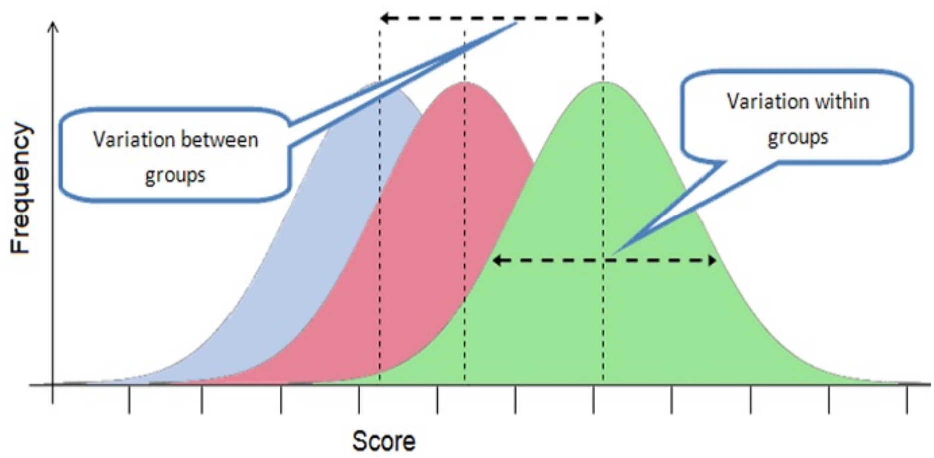


Figure 5.1: Sources of Variation : Within-variance and Between-variance

The statistic analysis used in ANOVA is called F-statistic which is the ratio of total between-groups variation and within-groups variation. Larger difference between groups as compared to within-groups reflects large value of F-statistic and hence likely to exist real differences. ANOVA is more interested in testing the null hypothesis which tells that samples from groups come from same population.

Each F-ratio is evaluated for significance. It tests whether the differences are statistically significant or not. The p-value here plays a crucial role in determining the statistical significance of the F-statistic. Calculate p-values from the F distribution, corresponding to specified F statistic and degrees of freedom. The critical values are then calculated for the same F distribution corresponding to the specified alpha (significance) level. If any p-value is near zero, the associated null hypothesis remains doubtful. A sufficiently small p-value for H0A suggests there is a main effect due to factor A. A sufficiently small p-value for H0B suggests that there is a main effect due to factor B. A sufficiently small p-value for H0AB suggests that there is an interaction between factors A and B. The choice of a limit for the p-value to determine whether a result is statistically significant is left to the researcher. From the literature, it is seen as a common practice to declare a result statistically significant if the p-value is less than 0.05 (as mentioned by Dallal (2012)). The following section provides overview of ANOVA models adapted in this thesis.

5.1.1 Model 1: 2-way ANOVA with replication

Experiments with 2 factors with m and n levels respectively and each level with N experiments, response can be represented by y in the following equation.

$$y_{ijl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijl} \tag{5.1}$$

$i = 1, 2, ..., m; j = 1, 2, ..., n; l = 1, 2, ...N.$ where

i, j indicate the level of the first and second factor respectively,

l indicates repetition (known as replicates) of that combination i, j,

μ represents the general effect,

α_i represent the effect of first factor at level i,

β_j is the effect of second factor at level j,

$(\alpha\beta)_{ij}$ is the interaction effect of the two factors at levels i, j respectively,

ϵ_{ijl} refers to chance or error term.

From this model one can perform the following tests -

1. First factor effects at different levels are same.
2. Second factor effects at different levels are same.
3. There are no interaction effects among both the factors.

5.1.2 Model 2: 3-way ANOVA with replication

Experiments with 3 factors with m , n and q levels respectively and each level with N experiments, response can be represented by y in the following equation.

$$\begin{aligned} y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{jk} + (\gamma\alpha)_{ki} + \\ (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl} \end{aligned} \tag{5.2}$$

$i = 1, 2, ..., m; j = 1, 2, ..., n; k = 1, 2, ...q; l = 1, 2, ...N.$ where apart from the first two factors and the interactions which are same as mentioned in the previous model, the third factor at level k is introduced here. The extra terms in the equation are -
γ_k represents the effect of third factor at level k,
$(\beta\gamma)_{jk}$ is the interaction effect of the second and third factors at levels j, k respectively,
$(\gamma\alpha)_{ki}$ is the interaction effect of the third and first factors at levels k, i respectively,
$(\alpha\beta\gamma)_{ijk}$ is the three factor interaction at i, j, k levels respectively,
ϵ_{ijkl} refers to chance or error term.
Apart from the three tests mentioned in previous model, the following are the additional tests one can perform-

1. Third factor effects at different levels are same.
2. There are no interaction effects among the first and third factors.
3. There are no interaction effects among the second and third factors.
4. There are no interaction effects among all the three factors.

In the context of texture characterization, we have adapted two ways to capture the texture orientation which is discussed in details in Chapter 5. Firstly characterization is

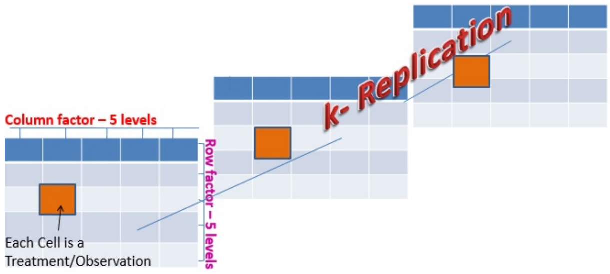


Figure 5.2: ANOVA- Factors, Treatments and Replication

based on 2-way ANOVA considering two factors (independent variables)- row and column of an image with n levels (sample size) for large number of sampled patches which are treated as replicates as shown in Figure 5.2 and secondly based on 3-way ANOVA using three factors row, column and 3D color space. In our case 2-way ANOVA is a way of studying the effects (main effects) of row and column factors separately and sometimes together called their interaction effects of row and column. For significant interaction, the effect of row is different for different levels of column. The row significant orientation means that the row differences in sampled values would contribute to main effect of row. The column significant orientation means that the column differences in sampled values would contribute to main effect of column. Whenever the interaction effect is significant in a two-way ANOVA, the main effects will not be interpretable at all, hence the significance (or non-significance) of the main effects will usually be ignored in such case.

In the present context of our thesis, the rows of the sampled image patches are considered as first factor, columns as second factor and the sampled patches are treated as replicates. The pixel value (scalar) in the sampled patch is response for the corresponding row and column of the patch which are analyzed by using 2-way ANOVA. 3-way ANOVA analysis has been adapted by considering pixel values as a vector (which indicates the third factor, the color descriptor) of the considered HSV color space, in our case. The components of the vector are taken as levels. The next section discusses the role of ANOVA in image processing.

5.2 Role of ANOVA for statistical image processing

The texture perception in humans interprets the scene, gathers more and more scenic observations in order to perform an efficient texture discrimination and segmentation for the purpose of high level image understanding. This gives rise to the need for summarization of all the observations. In this chapter, the automatic texture characterization of an image involves the process of image abstraction in terms of capturing its orientation. The need of statistics arise when a meaning needs to be captured using just few measures for a large collection of observations on many different cases. There are two types of statistics - descriptive statistics and inferential statistics(advanced branch of statistics as mentioned by Cohen and Lea (2004). A measure of central tendency, the arithmetic mean is one of the most obvious descriptive statistic that summarizes all the observations with a single number. The need of inferential statistics arises when it is not practical to measure facts of each and every observation. A subset of observations then is used for summarization and can be considered as a sample taken from a large population to make a guess(or inference) about the population. The distinction between sample and population becomes prominent in inferential statistics. A good sampling technique ensures that the sample is a good representation of the population. Lack of proper sampling can impact statistical inferences. This chapter makes an attempt to model the texture characterization mathematically based on analysis of variance(ANOVA), an inferential statistical method to explain the texture-tonal orientation present in an image.

In processing large images, there are chances that the noise variance can change significantly over various regions of the image. Hence to capture the image data correctly, it needs to be scanned by an appropriate sized localized mask. Window masking is needed and it needs to be adapted. This makes the hypothesis tests more accurate. Though the statistical model is linear, the actual operations involving imaging data are nonlinear as mentioned by Kurz and Benteftifa (2006). Even though in some areas the assumption of a noise-free image might be true, for most other applications the noise term has an important corrupting effect. This demands for a robust data modeling procedure that includes the noise term explicitly. As ANOVA, which has a noise parameter is a data model that includes the noise term in the parameterization of the observations. The class of test statistics used is primarily based on different forms of the linear model involving ANOVA techniques in the framework of experimental designs. One can make a statement pertaining to the influence of parameters and their interactions on image pixel characteristics. From this statement one can induce a hypothesis for statistical testing. Initially important features or factors are taken into account and depending on the results of any statistical test based on the linear model, the factors are interpreted. The hypotheses are next introduced that best fits the objective, the selected factors and the observable data. Null hypotheses play an important role in testing the significance of differences in samples. Finally, the importance is attached to the results by means of confidence. The usual null hypothesis of an experiment is that there are no differences among the effects of the parameters (factors). The proper method of sampling plays a crucial initial step to reach the correct statistical inference about the image characterization. The next section gives a detailed description of the proposed progressive sampling.

5.3 Progressive Sampling

The technique used for the proposed characterization is ANOVA with replication, hence there is a need for proposed new progressive sampling, for minimizing sampling error in terms of pseudoreplication. To obtain the objects in the image, we use the simplest form of segmentation, global thresholding using Otsu's method as found in paper by Otsu (1979) which segregates the pixels into two categories, foreground pixels and background pixels based on grayscale intensity. Figure 5.3 illustrates a binary image segregating the image 100075.jpg into two classes using Otsu's thresholding. Binary images are thus produced from color images by segmentation. We get all the connected components from the binary image which represents various regions. For each connected component, the region boundary is extracted. The area near the boundary of the region often has high inter-class variance. The proposed progressive sampling is flexible and iterative in nature. The positional sampling (patch of pixels) has been done along the boundary region and is moved gradually from peripheral to the center of the region. The position of each patch of pixels is guided by the boundary region, but the values taken for analysis are extracted from respective feature set such as grayscale or color. The samples are also referred as image patches in this Chapter. The detailed algorithm for progressive sampling is mentioned in Algorithm 6. The sampling parameter settings depend on certain aspects which have been discussed in the following subsection. The parameter initialization of the values are set accordingly.

Figure 5.3: Binary image obtained by Otsu thresholding

The three aspects of progressive sampling are mentioned in subsections below. Section 5.3.1 discusses the *sample size* to be used. Section 5.3.2 mentions the control parameters used for appropriate *sampling density* for choosing the replicates. Section 5.3.3 mentions the four types of *sampling features* used for sampling.

5.3.1 Sample size

It has been learnt from the past research work that at least 20 pixels of observable data is good enough for obtaining detailed resolution. Mask of size 5x5, 7x7 is commonly used in image processing. Odd sized patches are considered with same number of rows and columns. Patch of 5x5 window size has been applied for both global abstraction at the

Algorithm 6: Progressive sampling

```
Input: Img_features, nSample_size, nInsideRegion_pxl,
       nGap_Pixels, nSampleCoverage
Output: nPatchCount
        Patches(1..nPatchCount)
for i = 1 → Maximum_connected_components do
  nBoundaries ← Parent_region_boundaries
  for i = 1 → nBoundaries do
    nIteration ← 0
    SampledPixels ← 0
    window_size ← nSample_size
    while SampledPixels < nSampleCoverage do
      nIteration ← nIteration + 1
      (x, y) ← start_pt_boundary
      Update_start_pt ← Penetrate_nInsideRegion * nIteration
      for Each_encircling_boundary do
        if Sample_not_considered_earlier then
          nPatchCount ← nPatchCount + 1
          Sample_Img_features_in_8_directions
          Patches(nPatchCount) ← Samples
          Update_start_pt ← nGap_Pixels
          SampledPixels ← (window_size)^2 * nPatchCount
        end if
      end for
    end while
  end for
end for
return nPatchCount,
Patches(1..nPatchCount)[, ]
```

whole image level and for local abstraction at the image block levels. The determination of patch size is very crucial for achieving accurate results.

5.3.2 Sampling density

We use ANOVA models with replication. The mathematical models are discussed in Section 3.3. The definition of replicates is that replicates are independent of one another. Hence image patches sampled are non-overlapping and independent. Overlapping patches will be representing dependent replicates. Hence it is desirable to avoid the overlapping patches. Non-overlapping patches are assumed to be independent replicates for building the methodology in the rest of the chapter. There are chances that the hypothesis test interpretation in terms of significance may become wrong in case of pseudoreplication. In case of image sampling, spatial pseudoreplication may arise if sampling density is high. To overcome the problem of pseudoreplication, we have devised three control parameters to achieve the appropriate sampling density. Too high sampling density will pose the problem of pseudoreplication and too low will not capture variety in observations, if present.

- Gap pixels (nGap_Pixels): Gap pixels guide us to the next sample along the boundary.
- Inside region pixels (nInsideRegion_pxl): Inside pixels guide us to the next sample towards the center of the region.
- Sampling coverage (nSampleCoverage): Sampling coverage is the ratio of the number of pixels obtained from sample collection to that of total number of pixels present in the image.

Increase in number of gap pixels, number of inside region pixels along with low value of sample coverage is ideally suited for sampling purpose to have a reliable ANOVA grading analysis.

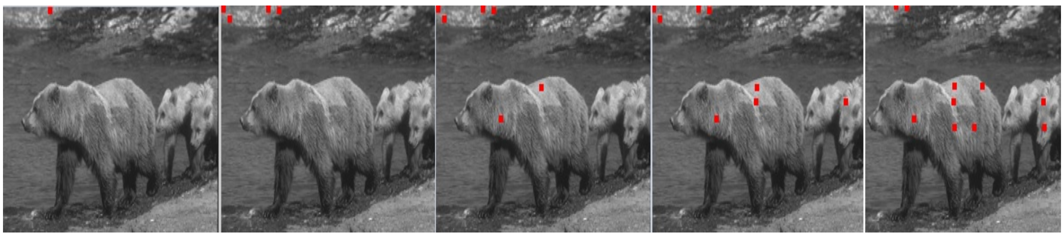


Figure 5.4: Positioning of Samples in capturing the maximum Variance in the image (100075.jpg)

Progressive sampling implements the sample selection in such a manner that maximum variation is achieved quickly as the sampling points are guided by the boundaries obtained from the thresholded binary image. On the other hand, overlapping pixel values mostly consist of repeated sample values. Overlapping sampling will have less variance as it does not capture the true variation present in the image. The list of overlapping samples will be an exhaustive list, till it finds out the maximum variance in the samples. Figure 5.4 gives a glimpse of the sampling strategy governed by proposed progressive sampling for initial ten samples for the image 100075.jpg. The positioning of the samples, are guided by the boundaries, obtained from binary image based on Otsu's thresholding, which enables directly access to different samples across the boundaries obtained from thresholded image. Thus the samples have a wide coverage. Both the

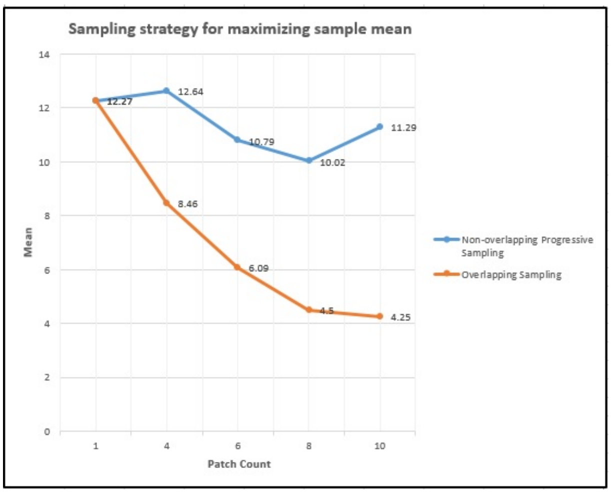


Figure 5.5: Role of Sampling Strategy in capturing the maximum Mean in the image (100075.jpg)

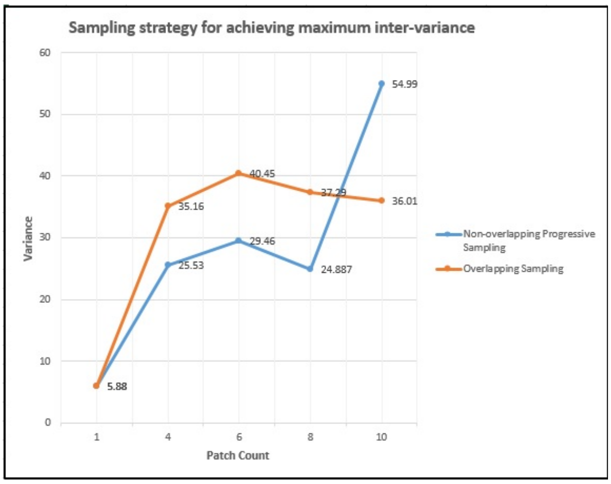


Figure 5.6: Role of Sampling Strategy in capturing the maximum Variance in the image

Figure 5.5 and Figure 5.6 are useful to understand the sampling pattern and illustrates how the proposed progressive sampling strategy for initial ten samples for the image 100075.jpg is effective. Though the samples in progressive sampling maintained the sample mean from inital value 12.27 to 11.29, depicted by blue line in the Figure 5.5. It can be observed that there is a drastic fall in the average value in case of overlapping samples from 12.27 to 4.25 (depicted by orange line). The variance in the samples is illustrated in Figure 5.6 which shows that at the tenth sample, the progressive sampling reached an area in the image where the variance has maximised from 5.08 to 54.99 as compared to 36.01 in case of overlapping samples. This experiment has been performed for other images to justify the sampling pattern. In future, the sample selection can be further refined, based on threshold for sample mean and sample variance collectively.

5.3.3 Sampling features

The position of each patch of pixels is guided by the boundary region, but the values taken for analysis are extracted from respective selected feature set such as grayscale or color. We have considered four types of features for sampling:

- 2-D sample: Intensities from grayscale image

  The samples are grayscale values obtained from the corresponding grayscale image of the input color image. We apply 2-way ANOVA model with replication in this case.
- 2-D sample: Values from 2-dimensional color difference map for color images generated from L*a*b* representation of the image.

  The samples can also be taken from a 2-D color difference map. The RGB image is first converted to L*a*b*, as it is a well-known fact that L*a*b* is device independent and perceptually uniform. Each pixel's L*a*b* value is compared with L*a*b* values of 8 of its neighbors by computing the CIEDE2000 color-difference formula as mentioned by Sharma *et al.* (2005). CIEDE2000 is considered the best uniform color-difference model in sync with subjective visual perception. The pixel position in a new 2-D matrix of image size initialized with zero is then updated with the maximum difference. The 2-D color difference map thus generated is used for color related progressive sampling for specified odd sized patches. The time complexity to generate the contrast color map is high (approx 2 min) as compared to the other features which are straightforward such as grayscale value, color plane. If the map is generated as one-time initialization, the ANOVA analysis at a subsequent stage will be efficient.

- 2-D sample: Values from each color plane. We have considered hue, saturation and value planes from HSV color space. The texture-tonal mix abstraction is characterized in terms of estimation of color attributes extracted from color spaces. We have selected the HSV color space because of couple of reasons. HSV is device independent color representation format. Its separates the luma (image intensity or value component) from the chroma (hue and saturation components - color information) thus providing a good estimate of the variance captured both in terms of purity of color and strength of light.
- 3-D sample: Values from any specific color space. All color planes of HSV are considered in this chapter.

We performed characterization experiments with our proposed HCS used in segmentation as mentioned in Chapter 3, and found that it behaves in similar manner as HSV in capturing the contrast. The contrast is captured most of the time, but not always. HSV being a pure data form is a lossless representation and more apt for contrast analysis, as compared to HCS. HCS is not a pure data form as, PCA is applied to various color components and reduced to three dimensions pertaining to usually capturing 90% color information. The sampling features considered for ANOVA, consist of the color planes in pure data form and lossless in nature such as H,S,V, L*a*b* color contrast and grayscale intensities. HCS, on the other hand exhibits its strength in clustering, as it possess distinguishable feature characteristics along with OP3 and OP5 in capturing texture-tonal mix in a given image. The collection of the features obtained from

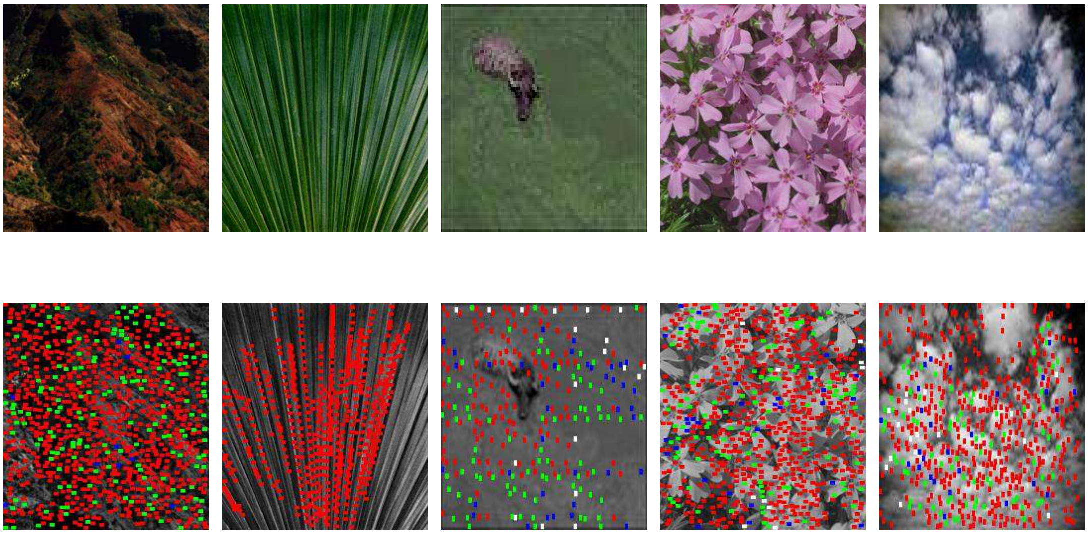


Figure 5.7: First row: Original images, Second row: Progressive positional sampling of grayscale image, color changes with each iteration inside the region red-green-blue-white

2-D sample is rich in capturing both luminance and chrominance information. 2-way ANOVA has been administered for 2-D samples, whereas 3-way ANOVA has been considered for 3-D color space. The number of image patches, N vary from image to image and refers to the total number of replicate observations in each combination. Sampling iteration penetrates inside the region. For depicting the process, a color coded image is shown in second row of Figure 5.7. The color code changes from red to green, blue and white with every iteration respectively. There is a stop criterion after four iterations. The penetration level represented by iteration can be changed suitably. The next section discusses the proposed ANOVA grading system to quantify the characterization.

5.4 Proposed ANOVA Grading System

The collection of patches either in image or block are treated as replicates of a 2-way data with row effects, column effects or interaction effects. These analysis of variance will produce corresponding 3 p-values. Grading of the image/block has been proposed based on these p-values and is represented by a *Characterization code type-1* referred here as *string code*. The following sections demonstrate the data analysis part pertaining to ANOVA. Characterization consists of both analysis and grading. Grading implies a differentiation of the texture and non-texture property. The grading is represented as string code representation, which will reflect the specific orientation using appropriate character codes for significant or non-significant main and interaction effects. The standard level of significance used to justify a claim of a statistically significant effect is 0.05. The term statistically significant has become synonymous with p-value=0.05 as mentioned in Refer Fisher and Yates (1947)). If p-value is between 0.1 and 1.0, there is certainly no reason to suspect the hypothesis tested, as mentioned by Dallal (2012). The following subsections discusses the step-by-step procedure for achieving the characterization codes. Section 5.4.1 briefs about how ANOVA is implemented and how to interpret the results for 2-way and 3-way models. Section 5.4.2 discusses about the grades to be assigned to the source of variability - row, column and color planes. Section 5.4.3 briefs about the formation process of the 22-length characterization string code. Section 5.4.4 briefs about the proposed polling method used to obtain another form of characterization code called semantic code. Section 5.4.5 shows example of string codes and corresponding semantic codes. Section 5.4.6 shows association of natural images with semantic code.

5.4.1 Implementation of ANOVA

The ANOVA table is produced when we use the MATLAB in-built function for performing analysis of variance using anova2 (for 2-way) or anovan (for 3-way). ANOVA table has six columns as shown in Table 5.1. The p-value for the F-statistic is the very important column for interpreting the result as discussed in previous section. The purpose of each of the six columns id mentioned below.

1. Source of the variability: Observed variance in a particular variable is partitioned into components attributable to different sources of variation
2. Sum of squares (SS) due to each source is a variation index. SS stands for Sum

Table 5.1: ANOVA Table

Sources of Variation	SS	df	MS	F	p-value
Main Effects	..	..	..	..	..
Interaction Effects	..	..	..	..	..
Error or Residual	..	..	..	..	..
Total	..	..	..	..	..

of squared deviation between each of the set of values and mean of these values. $SS = \Sigma(value - mean)^2$

3. Degrees of freedom (df) associated with each source. It is the number of values in the final calculation of a statistic that are free to vary.

4. Mean squares (MS), which is the ratio SS/df. They are used to determine, whether factors are significant.

5. F statistics which are the ratios of the mean squares. The test statistic used for ANOVA is the F-statistic and is calculated by taking the Mean Square(MS) of the variable divided by Mean Square of the Error (MSE).

6. p-values for the F statistics. The p-value tests the null hypothesis that, data collected from all groups are drawn from populations with identical means. A significant p-value indicates, at least one group mean, is significantly different from others. In case of interaction, p-value tests the null hypothesis which states that there is no interaction between the groups.

For 3-way ANOVA implementation we have used MATLAB function $anovan$. The Algorithm 7 shows as how to obtain the p-values for all the main and interaction effects for 3-way ANOVA model.

5.4.2 Grading for 2-way and 3-way ANOVA

Grading of source of variability has been arrived depending on the corresponding p-value obtained from ANOVA analysis as follows -

1. ‘H’ for highly significant when p-value is less than 0.05 (inclusive)
2. ‘M’ for medium significant when p-value is between 0.05 and 0.1 (inclusive)
3. ‘N’ for not significant when p-value is greater than 0.1

Statistical Inference	Main Effects		Interaction Effects
	Row	Column	Row Col
Not Significant	N	N	N
Row Significant	H/M	N	N
Column Significant	N	H/M	N
Row and Column Significant	H/M	H/M	N
Interaction Significant	X	X	H/M
X → Don’t care			
String Code : [SSS] , first two → main effects and third → interaction effect			

Figure 5.8: Proposed 2-way ANOVA Grading

Wherever 2-way ANOVA analysis is administered, the image or corresponding block will be coded with a string generated by grading scores of first factor, second factor and their interaction thus forming a 3 letter string [SSS] as shown in Figure 5.8. The letter ‘X’plays a special role in formulating the proposed ANOVA grading. ‘X’

Algorithm 7: ANOVA Implementation

Input: $ColorImg\_Data$
$RowLevels$
$ColLevels$
$PlaneLevels$

Output: $AnovaTable$

$(p, table, stats) \leftarrow compute_anova(ColorImg\_Data,$
$RowLevels, ColLevels, PlaneLevels)$
$***MainEffects***$
$p\_valueRow \leftarrow table(2,7)$
$p\_valueCol \leftarrow table(3,7)$
$p\_valueCplanes \leftarrow table(4,7)$
$***2WayInteractions***$
$p\_valueRowCol \leftarrow table(5,7)$
$p\_valueRowCplanes \leftarrow table(6,7)$
$p\_valueColCplanes \leftarrow table(7,7)$
$***3WayInteractions***$
$p\_valueRowColCplanes \leftarrow table(8,7)$

in ‘XXH’ means don’t care condition, which means it can be either ‘H/M/N ’. This is because the main effects are ignored if interaction effect is significant. The possible options are ‘NNN’, ‘XXH’, ‘HNN’, ‘NHN’, ‘HHN’. Visual interpretation of ‘NNN’ for is toned, ‘HNN’ is row significant, ‘NHN’ is column significant and ‘XXH’ is diagonal oriented texture. In case of 3-way ANOVA, a string code of size 7 is generated based on grading scores of main effects(3), 2-factor interaction effects(3) and 3-factor interaction effect in the predefined order. The string format appears like [SSS-SSS-S] as shown in the Figure 5.9.

Statistical Inference for 3-way	Main Effects			2-way Interaction Effects			3-way Interaction Effects
	Row	Column	Color Plane	Row-Col	Row-Color Plane	Col-Color Plane	Row-Col-Color Plane
Not Significant	N	N	N	N	N	N	N
Row Significant	H/M	N	X	N	H/M	N	N
Column Significant	N	H/M	X	N	N	H/M	N
Row and Column Significant	H/M	H/M	X	N	X	X	N
Interaction Significant	X	X	X	H/M	X	X	H/M
X → Don’t care							
String Code : [SSS-SS-S] , first three→ main effects and next two → 2-way interaction effect, and lastly 3-way interaction effect							

Figure 5.9: Proposed 3-way ANOVA Grading

5.4.3 Characterization Code Type-1 : String Code

Any image can be represented by 22 length string format which is in lossless form. The string is characterized by concatenating codes in a sequence obtained from 6 feature sets such as HSV 3-D color space, 2-D grayscale intensity values, 2-D LAB color contrast,2-D hue plane, 2-D saturation plane and 2-D value plane. Code of length 7 is obtained from 3-way ANOVA grading of HSV color space [SSS-SSS-S], codes of length 3 obtained from each graylevel [SSS], L*a*b* color contrast [SSS], color plane Hue [SSS], Saturation [SSS], Value [SSS], thus arriving at a *string code* of total code length of 22. This type of lossless representation thus becomes a structured string which can facilitate for knowledge discovery at later scope.

5.4.4 Characterization Code Type-2: Semantic Code

The semantic meaning can be extracted from the *string code* using our proposed polling method. The variance in contrast may or may not be captured well in all the 6 feature sets at a time. In the polling method, each feature set is analyzed based on its position in the string code of 22 code length, and if at least one feature set shows significant variation in a particular orientation such as significant main effects in row, column, row and column and significant interaction and no orientation (due to no contrast or very small size) effects respectively, the image is given a *semantic code* such as *ROW*, *COLUMN*, *ROW AND COLUMN*, *INTERACTION*, *COLOR SHADE* or *NO GRADE* respectively. The Algorithm 8 shows the proposed polling method depending on the code value H/M/N present at a particular position in the code. INTERACTION effect if found significant, other effects are ignored. COLOR SHADE depicts that no prominent orientation could be captured in terms of row and column, but the variations are due to colors present in the image. The semantic code in such case is denoted by *COLOR SHADE*. When the ANOVA fails to gather enough samples for the variation analysis, the semantic code is denoted by *'NO GRADE'*. The *semantic code* provides an overall idea about the image orientation by polling method, and does not reflect the contribution of each feature set in the code as in case of lossless string code. It can be observed from the Table 5.2 that there can exist more than one different string code for same semantic code.

5.4.5 Images Graded with Semantic Code

The characterization codes provide useful statistical metadata for a given natural image about its basic orientation characteristics whether it is row, column, row and column or interaction specific. This knowledge representation can be considered as target detection as mentioned in Section 1.2, where the target is presence or absence of specific orientation present in the image. In this section, we have reported experimental results for images taken from BSD and Vistex dataset to demonstrate the knowledge representation as shown in Figure 5.10. It has been observed that the grades will vary, if the image size is further reduced, which has been illustrated in the next section.

Algorithm 8: Polling method-Convert Lossless string code to Lossy Semantic code

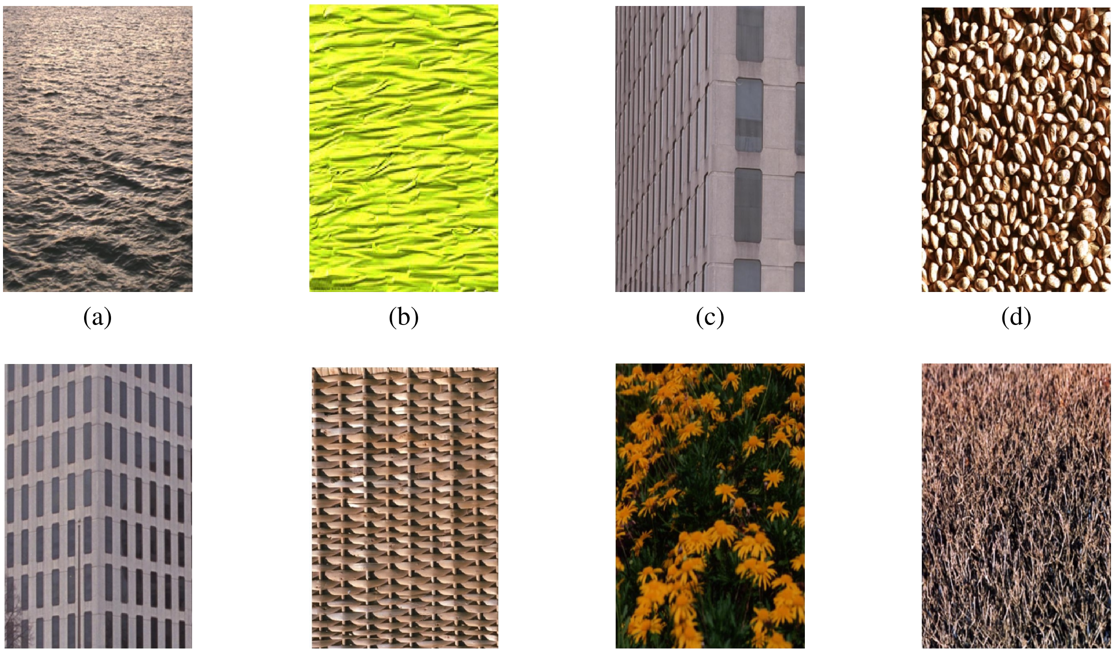
Input: $StringCode$
Output: $SemanticCode$
$InteractionFlag, RowFlag, ColFlag, ColorShade \leftarrow 0$
if $(code(4) ==' H'||code(7) ==' H'||code(10) ==' H'||code(13) ==' H'||code(16) ==' H'||code(19) ==' H'||code(22) ==' H'||code(4) ==' M'||code(7) ==' M'||code(10) ==' M'||code(13) ==' M'||code(16) ==' M'||code(19) ==' M'||code(22) ==' M')$ **then**
    $InteractionFlag \leftarrow 1$
end if
if $(code(1) ==' H'||code(5) ==' H'||code(8) ==' H'||code(11) ==' H'||code(14) ==' H'||code(17) ==' H'||code(20) ==' H'||code(1) ==' M'||code(5) ==' M'||code(8) ==' M'||code(11) ==' M'||code(14) ==' M'||code(17) ==' M'||code(20) ==' M')$ **then**
    $RowFlag \leftarrow 1$
end if
if $(code(2) ==' H'||code(6) ==' H'||code(9) ==' H'||code(12) ==' H'||code(15) ==' H'||code(18) ==' H'||code(21) ==' H'||code(2) ==' M'||code(6) ==' M'||code(9) ==' M'||code(12) ==' M'||code(15) ==' M'||code(18) ==' M'||code(21) ==' M')$ **then**
    $ColFlag \leftarrow 1$
end if
if $(code(3) ==' H' and\ all\ other\ codes\ are\ 'N'$ **then**
    $ColorShade \leftarrow 1$
end if
if $not(InteractionFlag)\ and\ not(RowFlag)\ and\ not(ColFlag)\ and\ not(ColorShade)$ **then**
    $SemanticCode \leftarrow' NO\ GRADE',\ return$
else if $InteractionFlag$ **then**
    $SemanticCode \leftarrow' INTERACTION\ SPECIFIC',\ return$
else if $RowFlag\ and\ ColFlag$ **then**
    $SemanticCode \leftarrow' ROW\ AND\ COLUMN\ SPECIFIC',\ return$
else if $RowFlag\ and\ \ ColFlag$ **then**
    $SemanticCode \leftarrow' ROW\ SPECIFIC',\ return$
else if $RowFlag\ and\ ColFlag$ **then**
    $SemanticCode \leftarrow' COLUMN\ SPECIFIC',\ return$
else if $ColorShade$ **then**
    $SemanticCode \leftarrow' COLOR\ SHADE',\ return$
end if
return $SemanticCode$

Table 5.2: Few Characterization Code Representations

String Code	Semantic Code
NNHNNHNNNNNHNNNNHHMNNN	INTERACTION SPECIFIC
NNHNNHNNNNNNNNNNHHMNNN	INTERACTION SPECIFIC
NNHNNMNNNNNNNNNNHHHNNN	INTERACTION SPECIFIC
NNHNNHHNNNNMNNNNHHNNNN	INTERACTION SPECIFIC
NNHNHNHNNNNNNNNNHHNHNN	INTERACTION SPECIFIC
NNHNHNNNNNHHNNNNHNNNNN	ROW AND COLUMN SPECIFIC
HNHNHNNHNNNNNHNNHNNHHN	ROW AND COLUMN SPECIFIC
HHHNHHNHNNNHNHNNHNNHNN	ROW AND COLUMN SPECIFIC
MNHNNHNHMNNNNHMNNNNMHN	ROW AND COLUMN SPECIFIC
MNHNHHNHMNHNNHMNNNNNNN	ROW AND COLUMN SPECIFIC
NMHNNNNNNNNNNNNNNNNNNN	COLUMN SPECIFIC
NNHNNNNNNNNNNNNNNNNNMN	COLUMN SPECIFIC
NNHNNHNNNNNHNNNNNHNNNN	COLUMN SPECIFIC
NNHNNNNNNNNNNNNNNHNNNN	COLUMN SPECIFIC
NNHNNNNNNNNNNNNNNNNMNN	ROW SPECIFIC
NNHNNNNMNNNNNMNNNNNHNN	ROW SPECIFIC
NNHNNNNNNNNNNNNNMNNNNN	ROW SPECIFIC
NNHNHNNNNNHNNNNNHNNNNN	ROW SPECIFIC
NNHNNNNNNNNNNNNNNNNNNN	COLOR SHADE

5.5 Impact of Image Size in Grading

The number of samples collected using progressive sampling depends on the image size. Larger the image, more are the samples collected. The same image when reduced to half, the grades change for all the cases as shown in Figure 5.11. Experiments have been conducted on image 100075.jpg of initial size. The patch count decreases to 3 size is reduced to one-eighth; and to zero when each of the block from sets of 4 and 9, are reduced to one-eighth. Similarly the semantic code also changes from column specific to row when halved; and to column when reduced to one-eighth.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 5.10: Semantic grading on Vistex categorical textured images - (a-b)Row specific, (c-d)Column specific, (e-f)Row and Column specific, (g-h)Interaction specific

		4 Blocks Partitioning				9 BlocksPartitioning								
Image Size	Whole	B4-1	B4-2	B4-3	B4-4	B9-1	B9-2	B9-3	B9-4	B9-5	B9-6	B9-7	B9-8	B9-9
Original	619	59	63	73	67	15	28	21	22	16	16	25	24	9
One-half	87	4	3	1	3	0	0	0	0	0	0	0	0	0
One-fourth	3	0	0	0	0	0	0	0	0	0	0	0	0	0
One-eighth	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(a)

		4 Blocks Partitioning							9 BlocksPartitioning					
Image Size	Whole	B4-1	B4-2	B4-3	B4-4	B9-1	B9-2	B9-3	B9-4	B9-5	B9-6	B9-7	B9-8	B9-9
Original	ROW	ROW	ROW	COLOR SHADE	ROW&COL	COLUMN	ROW&COL	ROW	COLUMN	COLOR SHADE	COLOR SHADE	COLUMN	COLOR SHADE	ROW
One-half	COLUMN	ROW	COLOR SHADE	NO GRADE	COLOR SHADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE
One-fourth	COLOR SHADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE
One-eighth	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE	NO GRADE

(b)

Figure 5.11: Impact of Image Size in Texture characterization a) Patch Count varies b) Semantic Code varies

5.6 Summary and Conclusions

The characterization technique we are using is ANOVA with replication, hence experimental evidence has proven that the proposed progressive positional sampling is suitable for minimizing sampling error in terms of pseudoreplication. The characterization code based on ANOVA Grading has two aspects - lossless and lossy. The lossless code is a full length 22 length string code and the lossy one reflects the compact semantic representation of the string code. It has been evident from the experiments that the ANOVA grading is dependent on the image size. One of the applications related to this semantic coding scheme can be image classification. It can also be made useful in domain specific applications, which need to identify horizontal and vertical orientations present in the image, such as manuscripts and stone inscriptions. The characterization can be applied to an image in couple of aspects such as whole image, regular blocks of image and image segments. All these possibilities are put together in the proposed framework is mentioned in the next chapter.

CHAPTER 6

Framework Demonstration and Observations

A framework is often a structure indicating the interrelation between its components. The framework usually supports a specific objective, and serves as a guide to make use of the components as required and applicable from case to case. We have developed a conceptual framework, which acts as an analytical tool for color-texture image segmentation and image characterization, which can work with several tonal-texture variations present in natural outdoor images.

This chapter will discuss the two major output of our framework - segment characterization and block characterization. The methodology for the segmentation has been broadly covered in Chapter-4 where our major contributions on two variants of color-texture feature integration OP3-HCS and OP5-HCS, iterative cluster based object segmentation, and local knowledge representation in terms of cluster and region metadata have been discussed in details. Chapter-5 provides methodology for texture characterization of outdoor natural images. The proposed progressive sampling and ANOVA grading system are major contributions of that chapter.

Our framework makes an attempt to connect our research project's goals and contributions under one umbrella. We have developed a conceptual framework as a part of low and mid-level image analysis component having two components:

1 Segmentation approach: Color-texture image segmentation and segment characterization.

2 Block approach: Block generation and blocks characterization

More advanced methodologies can then be applied, based on the metadata produced by the framework, for higher level analysis in future. A schematic diagram of the framework has been presented in Figure 6.1 for a BSD image (100075.jpg). Segmentation

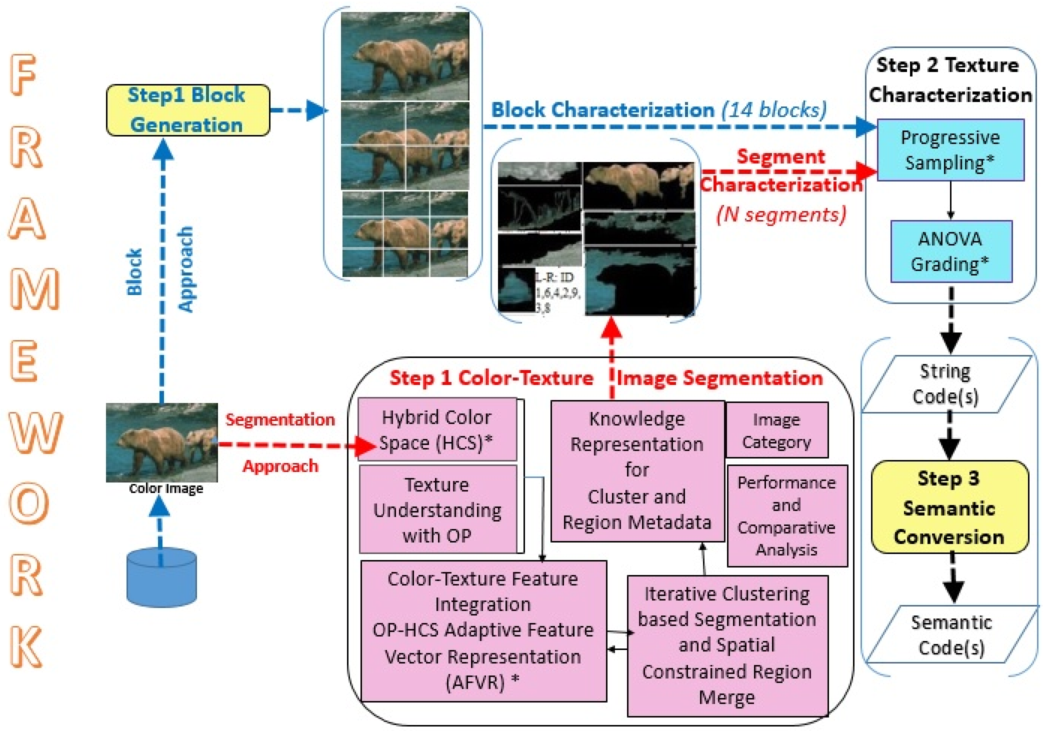


Figure 6.1: Framework results for a natural scene

approach has two sub-components in sequence - unsupervised segmentation and texture characterization. The block approach will go through block generation and then texture characterization. The output of the framework is a collection of string and semantic codes and is referred as image representation. The chapter is organized as follows. Section 6.1 describes segmentation approach and Section 6.2 describes block approach. Section 6.3 demonstrates the results and observations obtained from the framework. Section 6.4 summarizes and concludes with future work.

6.1 Segmentation Approach

A natural outdoor image when given as an input to the framework, can be segmented into objects or regions, which can be further processed for texture characterization. The knowledge representation of the segments, after texture characterization becomes more informative with additional inferential statistical ANOVA grading, to the already available descriptive statistics metadata computed during segmentation stage.

Each segment obtained from image segmentation as mentioned in Chapter 4, is considered separately for segment characterization. Experimentally it has been observed that the window size has an impact on the characterization code. Patch or sample count, decreases significantly on increasing the window size as shown in Figure 6.2. Each color code represents a segment, in the figure. For every case, it can be seen that, the patch count decreases, for increasing window size. Few big segments have been illus-

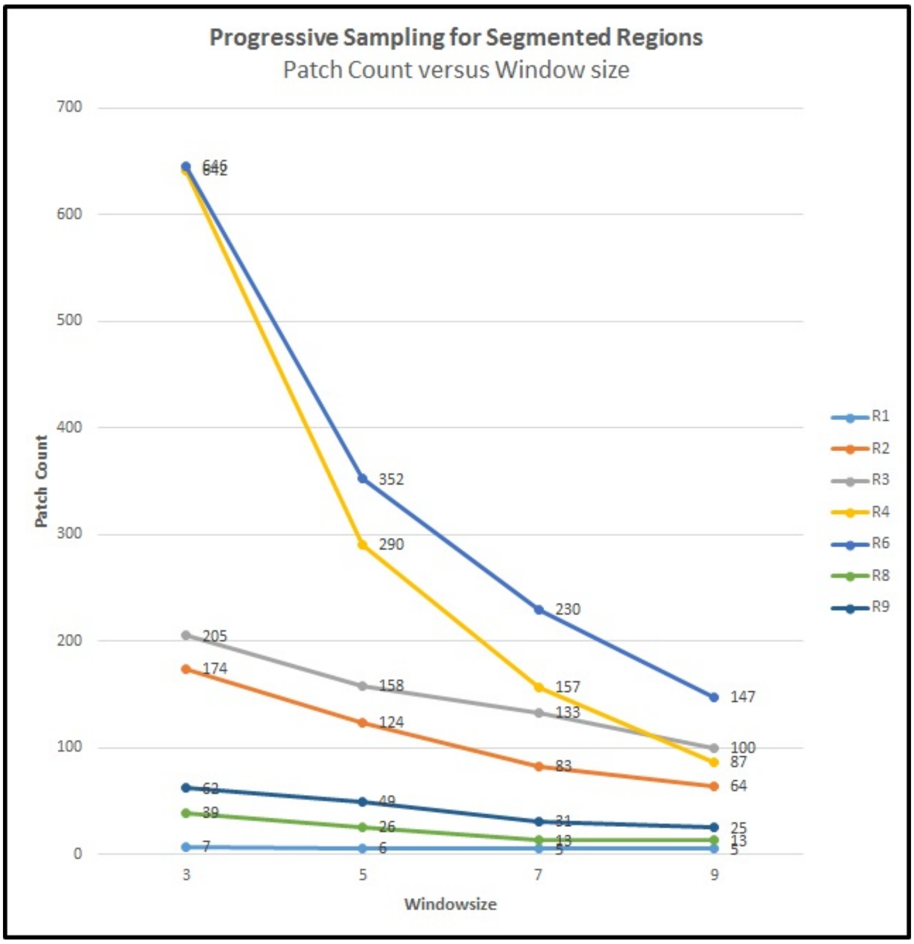


Figure 6.2: Progressive Sampling for segmented regions: Patch count decreases with increase in windowsize

trated for segment sampling in Figure 6.3. Illustrations show sampling for each of the four big segments with window size 3x3 in Figure 6.3(a,c,e,g)and with window size 7x7 as seen in Figure 6.3(b,d,f,h). In our experiments, we do not consider the window size as 3x3, as it gives 9 data points, which are not enough for analysis. Just for the sake

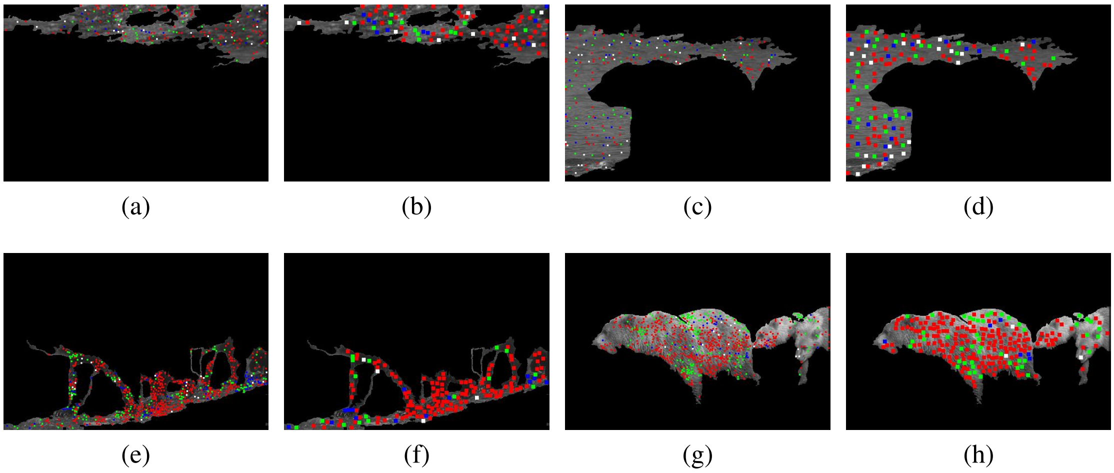


Figure 6.3: Each set of images contain sampling of segmented region with windowsize 3 and 7. Each image is of size 321x481

of illustration, window sizes 3x3 and 7x7 have been considered. The color code of the samples in each image, changes to depict the progressive sampling, that has been earlier discussed in Chapter 5. The framework not only characterizes image segments, but also whole image and image blocks, depending upon the input image, as will be discussed in the next section. Hence to have a one time parameter setting for each and every case, window size of 5x5 has been found to be appropriate for progressive sampling in both the cases of block and segment characterization. The knowledge representation of the segments becomes more informative with additional inferential statistical data as mentioned in Figure 6.4. The codes obtained from ANOVA grading for each of the six feature sets are plugged onto the region metadata already computed initially during segmentation process, based on descriptive statistics as previously shown in Chapter 4. The *nograde* is assigned for very small regions as enough sampling is not obtained. The metadata for Region-6 is shown in Table 6.1. In case, no segments are produced after segmentation, still the framework generates metadata for the input image using block approach as discussed in the forthcoming section.

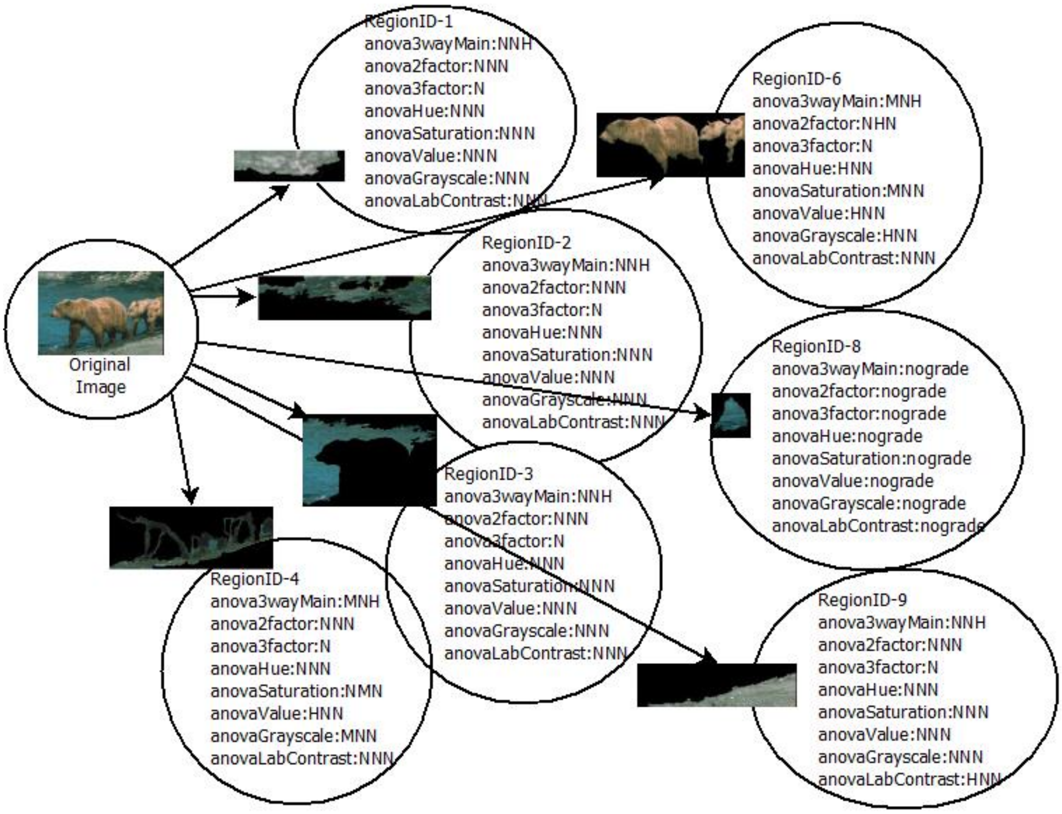


Figure 6.4: Region metadata based on ANOVA-inferential statistics of 100075.jpg

Table 6.1: Region Metadata for Region 6 of 100075.jpg

Region-6	**Metadata**	**Value**
1	Contrast	0.4019
2	Homogeneity	0.7990
3	Energy	0.3818
4	Correlation	0.0841
5	Mean Intensity	118
6	Anova3wayMain	MNH
7	Anova2factor	NHN
8	Anova3factor	N
9	AnovaHue	HNN
10	AnovaSaturation	MNN
11	AnovaValue	HNN
12	AnovaGrayscale	HNN
13	AnovaLabContrast	NNN
14	Semantic code	ROW

6.2 Block Approach

The objective of extracting the blocks from the image is two-fold - 1) natural visual patterns are usually repetitive, and may contain similar textural arrangements. 2) image and block level hierarchical textural metadata characterization provides the inherent complexity of the image in totality. The blocks are obtained from image by simple partitioning method. A systematic way of texture analysis is provided by further partitioning of one image block into two block sets - set of 4 and set of 9. The whole image is considered as one block. The string characterization of the image block and its block subsets provide a hierarchical aspect to the problem of texture analysis. Figure 6.5 shows a Vistex grass image (referred in two papers by Sanda *et al.* (2013) and Wouwer *et al.* (1999))The proposed block illustrations are as follows :

1. Whole image considered as 1 block
2. Image partitioned into 4 blocks
3. Image partitioned into 9 blocks

(a) (b) (c)

Figure 6.5: Vistex dataset - Grass image a) Block-1 b) Block-4 c)Block-9

Usually when an image is highly textured, or when it is difficult to extract an object from it, the global image characterization helps in capturing some idea about its inherent orientation. The image taken from Vistex categorical texture dataset of grassland is highly textured and no meaningful segments are produced, only characterization from block approach results are obtained from the framework.

There is an impact of image size in grading as discussed earlier in Section 5.5. Hence if number of blocks for partitioning the image, is further increased, the patch count would be very less and the texture characterization won't be robust and effective. As most of the selected dataset images are not so large, we have experimented with two set of block image sets - set of 4 and set of 9. However one can do extensive experimentation for different sizes, and carry out computational intelligence methods, for fixing the block size for each image size. The extent to which, the number of blocks, an image can be partitioned can be application and domain specific. Thus the proposed statistical learning for the texture-tonal characterization, is very flexible and can prove to be an eligible mid-level image analysis component, without making any assumptions about the nature of texture present. The proposed progressive guided sampling replicates will be minimal, for producing robust statistical analysis. The next section demonstrates the results obtained from the framework.

6.3 Demonstration and Observations

This section provides the framework results in terms of output, obtained from segmentation approach and block approach. The experimental setup is referred in Section 6.3.1. The outdoor natural textures have been characterized globally at an whole image level, and the observations are discussed in Section 6.3.2. The texture characterization metadata collected at the segment level and block level, collectively represent the knowledge representation of a given image, which has been illustrated in Section 6.3.3.

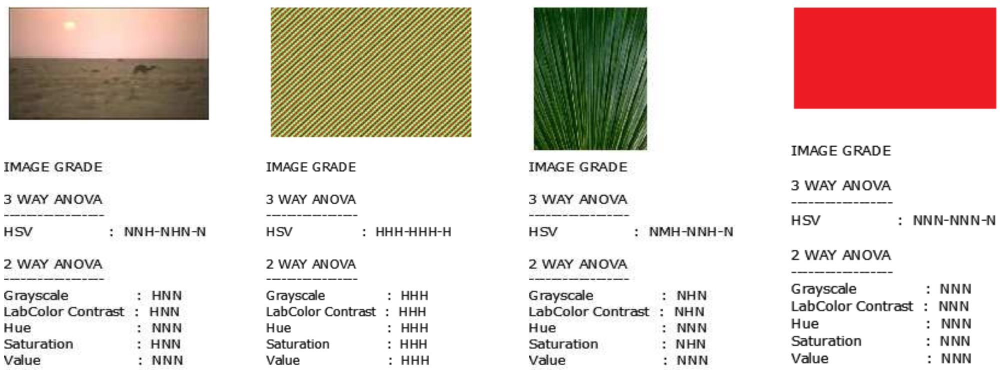
6.3.1 Experimental Setup

The experimental setup consists of a collection of images, taken from different benchmark datasets. The Berkeley dataset Martin *et al.* (2001) has been selected for panoramic images, Vistex (refer MITMediaLab (2005)) for categorical natural images, Prague (refer Haindl and Mikes (2014)) for textural mosaics and Corel (refer Tao (2009)) for both synthetic and natural textures. The parameter settings for the experiments performed are one-time initialization. The patch window size is set to 5x5, nGap_Pixels is set to 50, nInsideRegion_pxl is also set to 50 and nSampleCoverage is set to 40 percent.

The total time taken for the ANOVA analysis for images is measured to be average 40 seconds, including the global and local block characterization. In case of color contrast feature map, time taken is approximately 2 minutes to generate the map, but the ANOVA analysis is fast. Our program has been executed on P600 Intel Pentium CPU, 2.00 GHz, 2.00 GB RAM and the code is written in Matlab running on Windows system.

6.3.2 Demonstration of Orientation characteristics present in Nature Images

Outdoor natural images consist of random patterns based on its visual characteristics, which is a mixture of tonal-texture variations. The global abstraction of clouds, terrain, flowers, water etc are studied using ANOVA, available in Vistex dataset.



(a) (b) (c) (d)

Figure 6.6: ANOVA Grade System for 2-way and 3-way analysis reflects the nature of orientation a) Row significant b) Interaction significant c) Column significant d) Non-Significant(Toned)

An image whether synthetic or natural can be characterized using ANOVA. It can be observed from the Figure 6.6 that, the grading results may depend on the sample

(a) (b) (c) (d)

Figure 6.7: Main Effects: (a-b)Significant Row Main Effects, (c-d)Significant Column Main Effects

Figure 6.8: Significant Both Row and Column Main Effects

features, selected for the ANOVA analysis. In case of a natural images as shown in Figure 6.6 a), feature spaces consisting of grayscale, L*a*b* color contrast and the saturation plane, captures the row significance (HNN), whereas hue and value plane does not (NNN). The 3-way ANOVA too captures row significance in the 2-factor interaction part (NHN). In case of synthetic image as seen in Figure 6.6 b), the highly textured nature is captured by all the features and by both ANOVA models. In Figure 6.6 c), the vertical pattern of the leaf structure, is captured as column specific (NHN) by grayscale, L*a*b* color contrast and saturation planes. Also the 3-way ANOVA, captures the column specific nature. The image in Figure 6.6 d), is a pure toned image and does not have any variation, either in orientation or among color planes. Unlike the synthetic image, nature has a complex mix of color texture variations. We observe that each feature set has its own strength in terms of color and contrasts. Appropriate set of feature space selection, leads to a more reliable texture characterization. It can be observed that water bodies and cloud, demonstrate significant row main effects in Figure 6.7 (a-b), zebra skin and leaf exhibits significant column main effects in Figure 6.7 (c-d). The

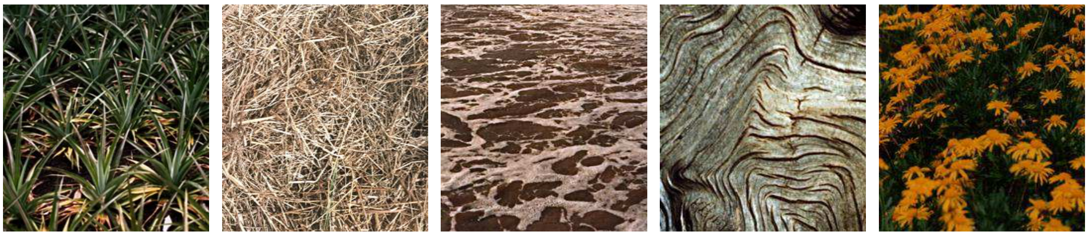
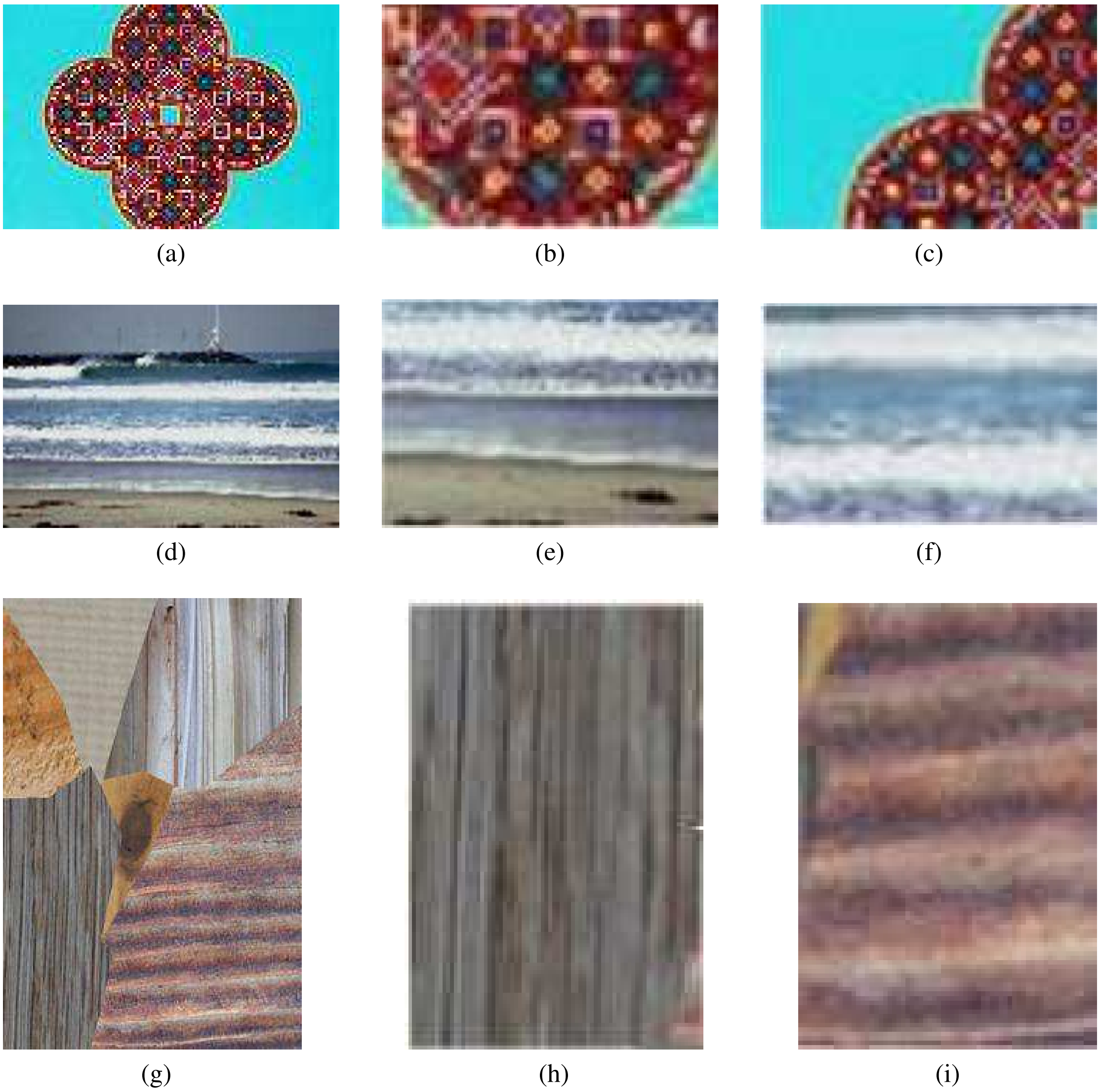


Figure 6.9: Significant Interaction Effects

leaves and water currents show both significant row and column main effects in Figure 6.8. The wood, twigs, flower bed and foamy water body exhibits significant interaction effects as shown in Figure 6.9. In another illustration, as can be seen in Figure 6.10, we have displayed the whole image grading and couple of its block samples grading, which characterizes the image content. The Figure 6.10 lists images from different datasets - images of top two rows, are from Corel, and the last row image is taken from Prague. We can see the variation in orientation if present, can be captured in blocks as in case of the last Prague image.

6.3.3 Framework Workflow and Knowledge Representation

The knowledge representation of an image obtained from our framework, includes the segmentation metadata and block-level metadata in terms of ANOVA grading - a collection of string codes and semantic codes, as mentioned earlier in Chapter 5. For the grass image from Vistex, no meaningful segments can be realized; thus the knowledge representation system generated by our framework, constitutes collection of fourteen string codes only for the blocks as shown in Table 6.2. and the workflow is shown in Figure 6.13 a). The BSD image 100075.jpg when segmented, could separate out the tonal-texture portions in bear, water and the land segmented areas. Out of fourteen segmented regions, seven regions (Region IDs-1,2,3,4,6,8,10) could be graded, as the region sizes are large enough for progressive sampling and variation analysis. The workflow in represented in Figure 6.13 b) and the knowledge representation of 100075 image is illustrated in Figure 6.11.



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 6.10: ANOVA Grading Results: (a) Synthetic Image: Global level-Interaction effect, (b) Block level-Column significant,(c) Block level-Row and Column significant,(d) Natural Image(waves): Global level-Row significant (e) Block level-Row significant,(f) Block level-Row significant (g) Textural Mosaic Image: Global level-Column significant, (h) Block level-Column significant,(i) Block level-Row significant

The BSD image 175032.jpg when segmented, produced 54 regions out of which 44 were not meaningful, had small region size and ANOVA fails to grade them. Only ten segments could be graded. The knowledge representation of this image obtained from the framework is shown in Figure 6.12 a). The BSD image 148026.jpg when segmented, produced 36 regions out of which 29 were not meaningful and no grade could

Table 6.2: Knowledge representation for Vistex-Grass image

Region	String Code	Semantic Code
Block-1	NNHNHHHNNNMHMNNNNNNNNHH	INTERACTION SPECIFIC
Block4-1	NNHNHNNNHNNNNNHNHNNNNHN	ROW AND COLUMN
Block4-2	NNHNNNNNNNNNNNNNNNNNNNM	INTERACTION SPECIFIC
Block4-3	NNHNHHNNNNNHNNNNNNNNNHN	ROW AND COLUMN
Block4-4	NNHNHNNNNNHNNNNNNNHNNHN	ROW AND COLUMN
Block9-1	NNHNHMNNMNNNNNMNNNNNNHH	INTERACTION SPECIFIC
Block9-2	NNHNNHNNNNNHNNNNMNNHNN	ROW AND COLUMN
Block9-3	NNHNNNNNHNNNNNHNNNNNNHN	COLUMN SPECIFIC
Block9-4	NNHNNNNNHNNNNNHNNNNNNNN	COLUMN
Block9-5	NNHNHNNNNNNNNNNNNNHNNNN	ROW AND COLUMN
Block9-6	NNHNNHNNHNNMNNHNHHNNHN	ROW AND COLUMN
Block9-7	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-8	NNHNNHNNNNNNNNNNMNNHNN	ROW AND COLUMN
Block9-9	NNHNHNNNNNHNNNNNNNNNNNN	ROW SPECIFIC

be assigned. In this case too only seven segments could be graded. The knowledge representation of this image obtained from the framework is shown in Figure 6.12 b). The image 87065.jpg when segmented produced 60 segmented regions, out of which only a handful of 5 could be graded as shown in 6.12 c). Thus it may be observed that, for highly textured images, texture characterization using block approach is more appropriate, gives compact knowledge representation and may facilitate higher level image processing in future.

Region	String Code	Semantic Code
Block-1	NNHNHNNNNNNNNNNNNHNNHNN	ROW SPECIFIC
Block4-1	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Block4-2	NNHNNNNNNNNNNNNNHNNNNN	ROW SPECIFIC
Block4-3	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block4-4	NNHNNNNNNNNNNNNNHNNHMN	ROW AND COLUMN
Block9-1	NNHNNNNNNNNNNNNNNNNNNMN	COLUMN SPECIFIC
Block9-2	NNHNNNNNNNNNNNNNHNNNMN	ROW AND COLUMN
Block9-3	MNHNNNNNNNNNNNNNHNNNNN	ROW SPECIFIC
Block9-4	NNHNNNNNNNNNNNNNNMNNNN	COLUMN SPECIFIC
Block9-5	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-6	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-7	NNHNNNNNNNNNNNNNNNNNNHN	COLUMN SPECIFIC
Block9-8	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-9	NNHNNNNNNNNNNNNNNNNNMNN	ROW SPECIFIC
Segment1-ID1	NNHNHNNNNNNNNNNNNHNNHNN	ROW SPECIFIC
Segment2-ID2	NNHNHNNMNNNNNMNNHNNNMN	ROW AND COLUMN
Segment3-ID3	NNHNNHNNNNNHNNNNNHNNHN	COLUMN SPECIFIC
Segment4-ID4	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Segment5-ID6	NNHNHNNMNNNNNMNNHNNNMN	ROW AND COLUMN
Segment6-ID8	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Segment7-ID10	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE

(a)

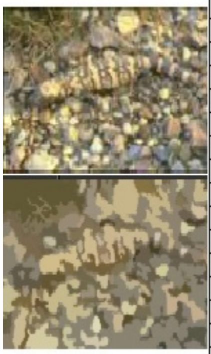
Figure 6.11: Knowledge representation for 100075.jpg

Region	String Code	Semantic Code
Block-1	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block4-1	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Block4-2	NNHNNNNNNNNNNNNNNNNNNMN	COLUMN SPECIFIC
Block4-3	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block4-4	NNHNNHNNNNNHNNNNNHNNNN	COLUMN SPECIFIC
Block9-1	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-2	NNHNNNNNNNNNNNNNNNNNNMN	COLUMN SPECIFIC
Block9-3	NNHNMNNNNNNNNNNNNNNNNNN	ROW SPECIFIC
Block9-4	NNHNHNNNNNHNNNNNNNNNNN	ROW SPECIFIC
Block9-5	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Block9-6	NNHNHNNNNNMNNNNNNNNHNN	ROW SPECIFIC
Block9-7	NNHNNHNNNNNHNNNNNNNNHN	COLUMN SPECIFIC
Block9-8	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-9	NNHNNNNNNNNNNNNNNNNNMNN	ROW SPECIFIC
Segment1-ID1	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment2-ID3	NNHNNHNNNNNHNNNNNHNNHN	COLUMN SPECIFIC
Segment3-ID5	NNHNHNNNNNHNNNNNHNNHNN	ROW SPECIFIC
Segment4-ID7	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment5-ID11	NNHNNHNNNNNHNNNNNMNMNN	ROW AND COLUMN
Segment6-ID15	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment7-ID17	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment8-ID33	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Segment9-ID37	NNHNNNNNNNHNNNNNNNNNNN	ROW SPECIFIC
Segment10-ID45	NNHNNHNNNNNNNNNNNNHNNHN	COLUMN SPECIFIC

(a)

Region	String Code	Semantic Code
Block-1	NNHNHNNMNNHNNMNNHNNNNN	ROW SPECIFIC
Block4-1	NNHNHNNNNNHNNNNNHNNHNN	ROW SPECIFIC
Block4-2	NNHNNNNNNNNNNNNNNMNNNNN	ROW SPECIFIC
Block4-3	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Block4-4	NNHNNNNNNNNNNNNNNNNNNHN	COLUMN SPECIFIC
Block9-1	NNHNHMNNNNMNNNNNNNNNNN	ROW AND COLUMN
Block9-2	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Block9-3	NNHNNHNNNNNHNNNNHNNNNN	ROW AND COLUMN
Block9-4	NNHNNNNNNNNNNNNNNNNNNHN	COLUMN SPECIFIC
Block9-5	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-6	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-7	NNHNNNNNNNNNNNNNNMNNNNN	ROW SPECIFIC
Block9-8	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-9	NNHNHMNNNNHNNNNNNNNNNN	ROW AND COLUMN
Segment1-ID1	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment2-ID5	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Segment3-ID7	HNHNHNNHNNHNNHNNHNNHNN	ROW SPECIFIC
Segment4-ID8	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Segment5-ID9	NNHNHNNNNNHNNNNNHNNNNN	ROW SPECIFIC
Segment6-ID10	NNHNHNNHNNHNNHNNNNNNNN	ROW SPECIFIC
Segment7-ID11	NNHNNNNHNNNNNHNNNNNNNN	ROW SPECIFIC

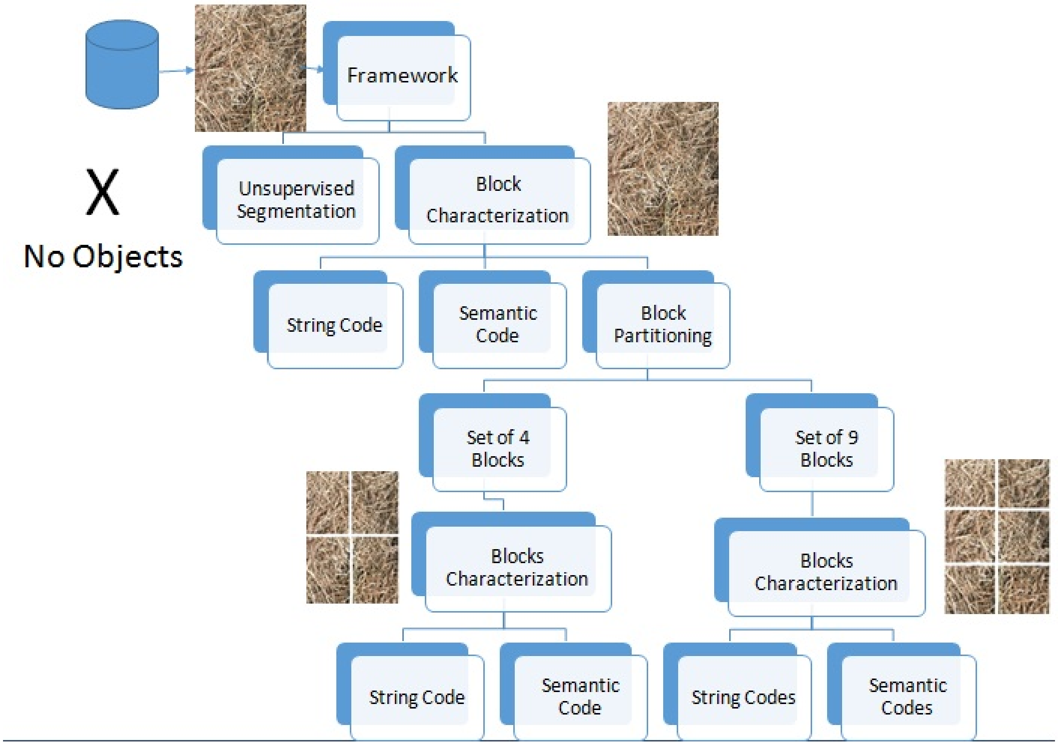
(b)



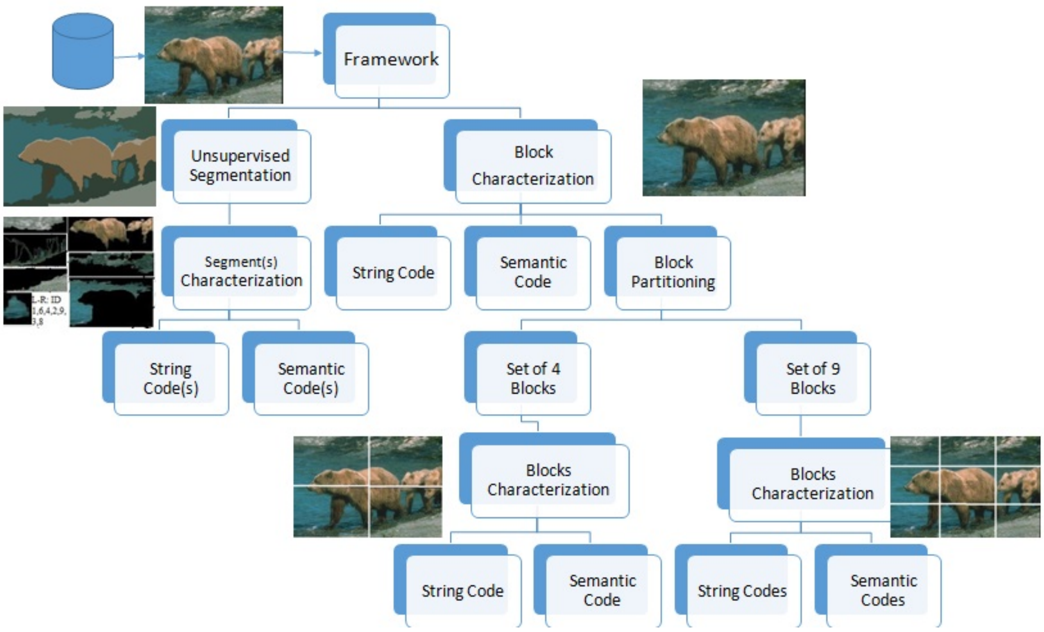
Region	String Code	Semantic Code
Block-1	HHHNHHNNNNHHNNNNHHMNHN	INTERACTION SPECIFIC
Block4-1	NNHNHHNNNNHNNNNNHHNNHN	ROW AND COLUMN
Block4-2	NNHNNNNNNNNNNNNNHNNNNN	ROW SPECIFIC
Block4-3	NNHNNNNNNNNNNNNNHHNNMN	ROW AND COLUMN
Block4-4	NNHNHNNNNNMNNNNNHNNNNN	ROW SPECIFIC
Block9-1	NNHNNNNNNNNNNNNNNNNNNNN	COLOR SHADE
Block9-2	NNHNHNNNNNHNNNNNHNNNNN	ROW SPECIFIC
Block9-3	NNHNHNNNNNMNNNNNHNNMNN	ROW SPECIFIC
Block9-4	NNHNNNNNNNNNNNNNNHNHMN	ROW AND COLUMN
Block9-5	NNHNHNNNNNHNNNNNNNNNHN	ROW AND COLUMN
Block9-6	NNHNHNNNNNMNNNNNHNNNNN	ROW SPECIFIC
Block9-7	NNHNHNNNNNNNNNNNNHNNNN	ROW AND COLUMN
Block9-8	NNHNNNNNNNNNNNNNHNNNNN	ROW SPECIFIC
Block9-9	NNHNHNNNNNMNNNNNHNNNNN	ROW SPECIFIC
Segment1-ID1	NNHNNNNNNNNNNNNNNNNNHNN	ROW SPECIFIC
Segment2-ID2	NNHNNHNNNNNNNNNNNHNNNN	COLUMN SPECIFIC
Segment3-ID19	NHHNNNNNHNNNNNHNNNNNNN	COLUMN SPECIFIC
Segment4-ID28	NNHNNMNNNNNNNNNNNNMNNNN	COLUMN SPECIFIC
Segment5-ID33	NNHNNNNNNNHNNNNNNNNNNN	ROW SPECIFIC

(c)

Figure 6.12: Knowledge representation a) 175032 b) 148026 c) 87065



(a)



(b)

Figure 6.13: Framework Workflow a) Vistex - Grass b) BSD 100075

6.4 Summary and Conclusions

The framework serves the dual purpose of image segmentation and texture characterization of the outdoor natural images. Segment characterization adds inferential statistical data in the form of *string code* and its corresponding *semantic code* to the region metadata, which has been initially populated during unsupervised segmentation. This additional information makes the metadata more informative for future use. In case not a single distinct object can be extracted from the image by segmentation; the framework processes the texture characterization using block approach, to know the basic orientation embedded in the image as whole and the blocks. The proposed statistical modeling of the image, as a whole or blocks is simple, computationally inexpensive and also needs minimal storage. The string characterization of the image and its block subsets provide a hierarchical aspect to the problem of texture analysis. The characterization code is rotation and scale variant. Higher complex order algorithms for color and texture characterization, classification or segmentation can be applied at later stage according to row significant, column significant or row-column interaction specific to obtain the texture primitives. Our method can be considered as pre-processing step for guiding towards localizing the textures, if at all present. Formalization of this method of localizing the texture can be done in our future work. The ANOVA analysis is done considering various types of feature maps such as grayscale intensities, color contrast map (L*a*b*), each color plane H,S,V taken separately and finally the HSV as 3-D color vector. We have experimented with two models of ANOVA for 5x5 window size, 2-way ANOVA with replication for 2-D features and 3-way ANOVA with replication for all color planes of a color space. Significant factor interaction is interpreted as highly textured. From the experimental evidence we can conclude that appropriate features selection leads to a more reliable texture characterization, if pseudoreplication is under control. We recommend the proposed framework as a mid-level image analysis component which generates valuable metadata, stage by stage in both segmentation and block approaches which can be further used for vision tasks.

CHAPTER 7

Conclusion and Future Directions

Computer vision and recognition systems need sufficient flexibility for processing the uncertainties coming from any of the low levels of image processing such as image segmentation, and also should be able to retain as much information as possible at each level. There is a need of a framework combining the efforts of low-level segmentation and mid-level characterization process, for compiling rich and quality metadata that can build appropriate knowledge representation about the image for further assisting high level image understanding. By doing so, one can avoid making a crisp bad segmentation at an earlier stage, and rather pass the knowledge gathered, through the higher processing levels.

An abundance of various automated and semi-automated techniques can be found in literature, that cater to wide range of image analysis and understanding applications. The present taxonomy found in the literature, is usually focused on the segmentation technique categorization, rather than portraying the overall aspects of its purpose and multiple segmentation facets involved in the process. The proposed taxonomy MFHIST fills in this gap by assisting the researchers in categorizing their works, according to multiple facets such as scope, control, feature, representation, approach specifications. Recently, the hybrid combinations in control specification, feature specification and approach specification are in practice as they give more robust results. MFHIST does not presently address the texture analysis in details, and can be taken up as a future work. The taxonomy would also help in studying the trend analysis on the evolution of image segmentation approaches towards application independence.

From the survey, it is evident that in spite of abundance in segmentation algorithms, not a single algorithm can segment all images to a satisfactory level, especially in cases of natural outdoor scenes having large variation in color-texture mix. Sophisticated/complex algorithms prove to be useful for highly textured images, but at the same time poses to be an over-kill for simple tonal/color images. For this purpose, we build a framework for unsupervised segmentation and characterization for mostly, natural outdoor images. The framework proposes image types as one of the useful metadata, based on statistical analysis such as object based, homogeneous color based textured image, heterogeneous color based textured image and pure toned image.

Our framework is capable of separating toned and textural segments from a given image, and can also perform texture characterization. The segmentation and characterization complement each other and generate useful metadata, which can help in representing and manipulating the uncertainties at a later stage of the vision and understanding system, in an application independence perception. The image understanding is not in the present scope of the thesis. The proposed methodologies do not consider any a-priori information, pre-processing, smoothing operations, thus work on the full range of color and intensity levels present in the given image. The role of color-texture as fundamental descriptors proved to be quite effective. The mathematical concepts used in texture extraction framework, based on orthogonal polynomial (OP) operators is found be efficient and reliable in capturing textural and non-textural regions in natural images. The performance measure in case of unsupervised segmentation significantly improved, when the proposed chrominance dominated Hybrid Color Space (HCS) is used in combination with OP. Experiments have been conducted with two variants of OP-HCS, i.e. OP3-HCS and OP5-HCS. OP5-HCS favors over-segmentation and captures finer details.

The thesis makes an attempt to model the nature mathematically and explain the texture-tonal orientation based on ANOVA. The global abstraction of an image is captured by proposed texture characterization depending on main effects and interaction effects. The block approach of the framework consists of the whole image and its corresponding set of 4-blocks and 9-blocks. Location of H/M in the code generated by texture characterization, communicates about the orientation aspect in terms of row, column and interaction specific, depending on its position in the 22-length string code. Output from segmentation approach is unstructured, as the number of segments produced is not fixed and varies from image to image. Whereas, output from block approach is structured to fixed number of string codes, which is fourteen in number, considering all the blocks. The ANOVA grading gets impacted by variation in window size used for sampling, and also by the block size used in block generation. The number of samples reduce with increase in window size and decrease in block size. The block grading along with segment grading gives a comprehensive compact representation about the image and its component objects or regions. The structured knowledge representation acquired from the block approach characterization, can be recommended as a preprocessing step, to understand the image in a holistic manner. In future, the collection of descriptive and inferential statistical region metadata, can be used for intelligent region merge.

The formalization task of detecting and localizing the texels can be carried out in future. The proposed framework is expected to be useful in the field of computer vision for cases where prior knowledge about the scene is unknown. The framework for color-texture segmentation and characterization of natural outdoor images can be further extended, as a guide towards image content characterization for domain specific applications.

APPENDIX A

Color Spaces

The appendix gives a compact overview of different types of color spaces and their subcategories which will have quite a lot of references for color feature space selection in both Chapter 3 and Chapter 6.

HVS measures electromagnetic spectrum ranging from approximately 400 to 700 nm and humans perceive the visible part of this spectrum as color. Color space represent colors as coordinate system in which the image pixel values are mapped as mentioned by Koschan and Abidi (2008). The authors Tkalčič and Tasič (2003) propose three categories of color spaces - HVS based, application specific and standardized perceptual uniform color spaces as described in the next subsections. The appendix section will brief about colors from each category which has been used in this thesis such as RGB, HSV, CMYK and CIELAB (or L*a*b*). The following sections A.1 provide details of RGB, HSV of the HVS based color space, A.2 provides CMYK of application based color space and A.3 provides details of CIELAB of perceptual uniform color spaces. For ready references, other related details can be found in literature by Klinker *et al.* (1990), Tkalčič and Tasič (2003),Koschan and Abidi (2008).

A.1 Human Visual System based color spaces

HVS based color spaces are motivated purely by the principles of human visual system, the manner in which we see, the ability to see and perceive. There are collection of theories behind the development of colors, which are as follows -

1. Maxwell, Young and Helmholtz's Trichromatic Theory
2. Hering's Opponent color theory
3. Isaac Newton's color circle

In trichromatic theory, three types of photoreceptors are considered that are sensitive to red, green, blue region of spectrum. This is motivated by HVS, which has 3 cones for wavelength sensitivity(L,M,S). Most of the devices capture images in an LMS detector i.e RGB color space. According to Hering's opponent color theory, opponent color vector consists of achromatic component (white-black) and two chromatic components (R-G and Y-B). This transformation is done in postreceptors retina cells, called ganglion cells that respond to opponent color stimulus. One such example is Ohta color space, obtained from Karhunen-Loeve transformation of RGB. In Isaac Newton's color circle, color is described based on hue and saturation for describing colors. Motivated by this color circle, another property of brightness is added to the hue and saturation, thus forming the phenomenal or perception based color space. The phenomenal color space is intuitive and is considered as highest level for representation of colors, similar to human visual processing. Some examples of phenomenal color space are Munsell, HSL (Hue Saturation Lightness), HIS(Hue Intensity Saturation), HSV(Hue Saturation Value). Its application can be found in PaintShop/ Photoshop. The main disadvantage of the HVS based color spaces are that they are device dependent, as they depend on specific sensitivity function and there does not exist any perceptual uniformity.

Most devices capture images in **RGB** color space. The color consists of three components: R, G and B. The value of these components is the sum of respective sensitivity functions and the incoming light, as mentioned by Tkalčič and Tasič (2003); is shown below:

$$R = \int_{300}^{830} S(\lambda)R(\lambda)d(\lambda) \tag{A.1}$$

$$G = \int\limits_{300}^{830} S(\lambda)G(\lambda)d(\lambda) \tag{A.2}$$

$$B = \int\limits_{300}^{830} S(\lambda)B(\lambda)d(\lambda) \tag{A.3}$$

where $S(\lambda)$ is the light spectrum, $R(\lambda)$, $G(\lambda)$ and $B(\lambda)$ are the sensitivity functions for the R, G and B sensors respectively. The RGB values depend on specific sensitivity function of capturing device which makes it device-dependent. It is additive primary

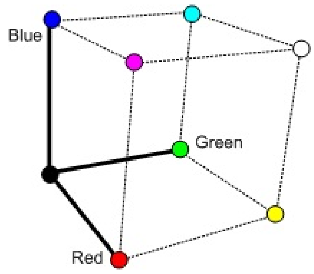


Figure A.1: RGB coordinate system

model where red, green and blue are mixed in different proportions to reproduce enormous number of colors as mentioned in Figure A.1 [1]). The problem with RGB color space is that high correlation exists between R, G, B components and is non-intuitive. Other color spaces can be obtained by color space transformations from RGB. RGB color space is most widely used in industrial applications.

HSV stands for hue, saturation and value. HSV is user-friendly color space than RGB. The value indicates the color intensity, and is separated from the color information. Hue and saturation are dependent color attributes and is similar to human's color perception. With respect to coordinate system, RGB when projected along the diagonals of white to black, a hexacone is formed which forms the HSV pyramid as shown in Figure A.2. It represents cylindrical coordinate system, but the colors are defined inside the hexacone. H ranges from 0 to $360\,\text{deg}$. S stands for purity of color which refers to

[1]Source: `http://badripatro-iitb.blogspot.in/2012/06/about-yuv-video.html`

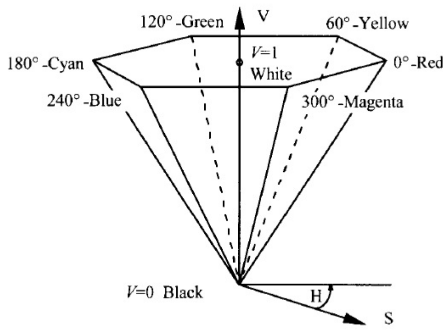


Figure A.2: HSV coordinate systems

amount of white added to color and ranges from 0 (pure color) to 1 (no white). Value ranges from 0(black) to 1(white).

Tkalčič and Tasič (2003) mentions the method of Gonzales and Woods for calculating HVS as shown below: H is hue, S is saturation and V is value.

$$H = cos^{-1}\left(\frac{0.5(R-G)+(R-B)}{\sqrt{(R-G)^2+(R-G)(G-B)}}\right) \tag{A.4}$$

$$S = 1 - \left(\frac{3}{R+G+B}\right)min(R,G,B) \tag{A.5}$$

$$V = \frac{R+G+B}{3} \tag{A.6}$$

The increase in value makes the colors brighter. When the saturation value is zero, hue is undefined. When the value is zero(black color), hue and saturation have no importance. Lighter color lies in the top side of the hexacone.

A.2 Application specific color spaces

Color spaces exclusively used for printing systems(CMYK), TV systems (YUV, YIQ), photography (Kodak PhotoYCC) are application dependent. The conversion equations obtained from RGB can be found in papers by Tkalčič and Tasič (2003) and Koschan and Abidi (2008).

CMYK is application specific and is used in printing technology. It represents subtractive primary colors such as cyan, magenta, yellow. Additionally black Karbon is used for printing black ink for economical reasons and reliable output. CMYK is the only color space with highest dimensions. With respect to coordinate system, CMY coordinate system is similar to RGB color cube as shown in Figure A.3.

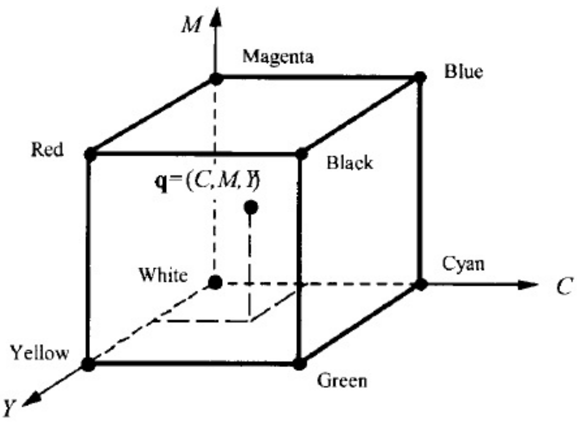


Figure A.3: CMYK coordinate systems

Two types of conversions to CMY are possible from RGB. Simple transformation from RGB color space is shown below. This is not advisable, as it does not give proper results.

$$C = 1 - R \tag{A.7}$$

$$M = 1 - G \tag{A.8}$$

$$Y = 1 - B \tag{A.9}$$

The transformation from CMY to CMYK is computed using following equations

$$K = min(C, M, Y) \tag{A.10}$$

$$C = \frac{C - K}{1 - K} \tag{A.11}$$

$$M = \frac{M - K}{1 - K} \tag{A.12}$$

$$Y = \frac{Y - K}{1 - K} \tag{A.13}$$

The one applied in practical applications for conversion from RGB to CMY use complicated 3-D interpolation of lookup tables or complicated polynomial arithmetic.

A.3 International Commission on Illumination(CIE) based uniform color spaces

CIE is an international organization devoted to standarization of color science. CIE standarizes XYZ, as reference tristimulus values perceptible by an human observer. CIE XYZ color space is device independent, and any transformations from XYZ is also considered as device independent. CIE proposed color spaces such as CIE L*u*v*, CIE L*a*b* has perceptual linearity. It means that the Euclidean distance measure between two colors, is in sync with human visual perception.

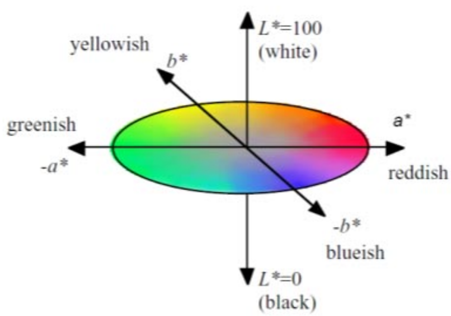


Figure A.4: CIELAB coordinate systems

CIELAB / L*a*b* - CIELAB color space is based on opponent color space and is derived from CIE XYZ. L* stands for lightness and a*, b* are color opponent dimensions. a* represents red-green opponent channel and b* represents yellow-blue opponent channel. The transformation from CIE XYZ to CIELAB is device independent. L*a*b* is computed using the following three equations -

$$L^* = 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16 \tag{A.14}$$

$$a^* = 500\left[\left(\frac{X}{X_n}\right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n}\right)^{\frac{1}{3}}\right] \tag{A.15}$$

$$b^* = 200\left[\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n}\right)^{\frac{1}{3}}\right] \tag{A.16}$$

The property of uniform color space is very useful when similar colors need to be compared. The perceptually linear color difference between two colors are computed as follows -

$$\delta E^*_{ab} = \sqrt{(\delta L^*)^2 + (\delta a^*)^2 + (\delta b^*)^2} \tag{A.17}$$

APPENDIX B

Benchmark Datasets

The computer vision community has proposed a large variety of publicly available datasets consisting of color-texture natural and synthetic images. We have used BSD for segmentation purpose. For texture characterization, images are considered from various datasets as Vistex, Corel, Prague and BSD. The details of the datasets are mentioned below.

Berkeley Segmentation Dataset and Benchmark (BSD) [1] is a large dataset of 300 natural images where each image has been segmented by a group of expert human subjects. The multiple ground-truths serve the purpose of providing benchmark for comparing different segmentation and boundary detection algorithms found in the literature. It consists of 200 images for training and 100 for test set.

MIT- Vistex [2] stands for Vision Texture database which is a collection of texture images which represent real world conditions for computer vision community.

Corel database [3] is a photo gallery consisting of 80 concept groups where each category is semantically homogeneous. For example grasslands, water, leaf, buildings are some examples of categories found in this dataset. Each image has atleast one object. This dataset is very popular for content based image retrieval. Segmentation and texture characterization can be applied to such images.

The Prague Texture Segmentation Data Generator and Benchmark [4] consists of computer generated texture mosaics and their corresponding ground truths. The specialty with this dataset is that a user can generate synthetic or mosaic scenes based

[1] Source: [2001] http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/
[2] Source: [1995] http://vismod.media.mit.edu/pub/VisTex/Images/Reference/
[3] https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval
[4] Source: [2008] http://mosaic.utia.cas.cz/

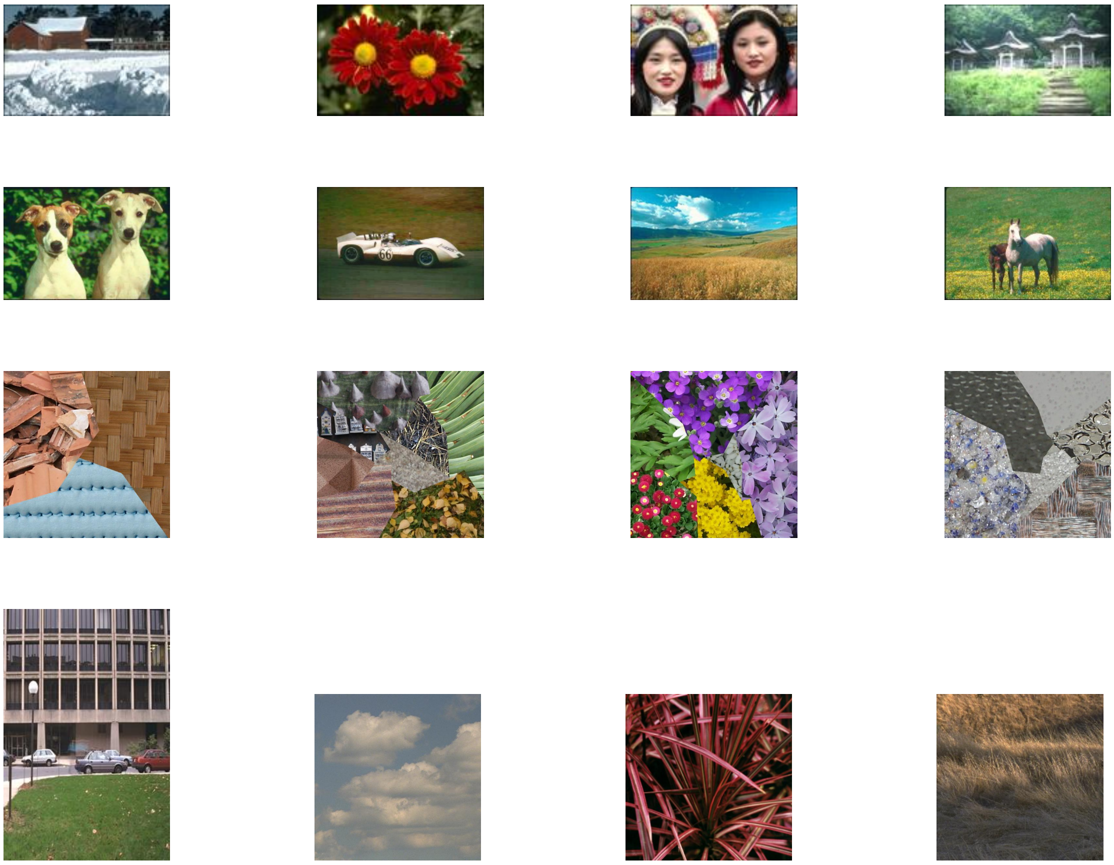


Figure B.1: Sample images from Benchmark Datasets - Row 1: BSD, Row 2: Corel, Row 3: Prague, Row 4: Vistex

on Voronoi polygon random generator where the ground truth is unambiguous. The Prague database consists of more than 1000 color textures. The ground-truth is not affected by human annotations produced for natural images as in BSD and is independent of the subjectivity. Prague dataset is useful in identifying the transitions between textures. Haindl and Mikes (2014) allows the user to online evaluate their segmentation algorithms with 27 different metrics.

REFERENCES

1. **Abbadeni, N.** (2011). Computational perceptual features for texture representation and retrieval. *IEEE Trans. on Image Processing*, **20**(1).

2. **Adamek, T.** and **N. E. O'Connor** (2006). Interactive object contour extraction for shape modeling. *1st International Workshop on Shapes and Semantics, Matsushima, Japan*.

3. **Agarwal, A.** and **P. Reddy** (1990). A distributed vision system using argus. *Expert Systems for Management and Engineering*, 236–245.

4. **A.Ghosh**, **N.R.Pal**, and **S.K.Pal** (1991). Neural network, gibbs distribution and object extraction. *Intelligent Robotics*, 95–106.

5. **Ahuja, N.** and **A. Rosenfeld** (1981). Mosaic models for textures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **3**(1), 1–8.

6. **Ahuja, N.** and **S. Todorovic** (2007). Extracting texels in 2.1d natural textures. *ICCV*, **12**(2), 1–8.

7. **Allen, Y.**, **J. Wright**, **Y. Ma**, and **S. Sastry** (2008). Unsupervised segmentation of natural images via lossy data compression. *Computer Vision and Image Understanding*, **110**(2), 212–225.

8. **Arbeláez, P.** and **C. Fowlkes** (2011). Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intel.*, **33**(5), 898–916.

9. **Bajcsy, R.**, **S. Lee**, and **A. Leonardis** (1990). Color image segmentation and color constancy. *Proc. SPIE vol. 1250 Perceiving, Measuring and Using Color*, **1(2)**, 245–255.

10. **Baker, D.**, **S. Hwang**, and **J. Aggarwal** (1989). Detection and segmentation of man-made objects in outdoor scences: concrete bridges. *Journal of the Optical Society of America A*, **6**(6), 938–951.

11. **Balasubramanian, G.**, **E. Sabera**, **V. Misic**, **E. Peskin**, and **M. Shaw** (2008). Unsupervised color image segmentation using a dynamic color gradient thresholding algorithm. *Proc. SPIE 6806, 68061H*.

12. **Belgiu, M.** and **L. Drăgut** (2014). Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, **96**, 67–75.

13. **Belongie, S.**, **C. Carson**, **H. Greenspan**, and **J. Malik** (1998). Color and texture-based image segmentation using em and its application to content-based image retrieval. *6th Int. Conf. on Computer Vision*, **43**(1), 675–682.

14. **Belongie, S.**, **C. Carson**, **H. Greenspan**, and **J. Malik** (2002). Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**(8), 1026–1038.

15. **Bhoyar, K.** and **O. Kakde** (2010). Color image segmentation based on jnd color histogram. *International Journal of Image Processing (IJIP),2010*, **3**(6).

16. **B.N.S. Chakravarthy, A. A.** and **P. Reddy** (1988). Argus: Archival representation for a general understander system. *Intl. Computer Science Conference, Hong Kong*.

17. **Borenstein, E.** and **S. Ullman** (2004). Learning to segment. *ECCV 2004, LNCS 3023*, 315–328.

18. **Boykov, Y. Y.** and **M. P. Jolly** (2001). Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images. *8th IEEE Inter. Conf. Comp. Vis. (ICCV), New York*, **1**, 105–112.

19. **Boykov, Y. Y.** and **G. F. Lea** (2006). Graph cuts and efficient n-d image segmentation. *Inter. J. Comp. Vis.*, **70**(2), 109–131.

20. **Busin, L.**, **N. Vandenbroucke**, and **L. Macaire** (2008). Color spaces and image segmentation. *Advances in Imaging and Electron Physics*, **151**, 65–168.

21. **Carreira-Perpinán, M.** (2006). Fast nonparametric clustering with gaussian blurring mean-shift. *Proceedings of the 23rd International Conference on Machine Learning*, 153–160.

22. **Chapron, M.** (1992). A new chromatic edge detector used for color image segmentation". *Proc. 11th Int. Conf. on Pattern Recognition*, **3**, 311–314.

23. **Charles Thorpe, T. K., Martial H. Hebert** and **S. A. Shafer** (1988). Vision and navigation for the carnegie-mellon navlab. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **10**(3).

24. **Chellappa, R.**, **R. Kashyap**, and **B. Manjunath** (1999). Model-based texture segmentation and classification. *Proc. C. Chen and L. Pau and P. Wang eds. Handbook of Pattern Recognition and Vision, World Scientific, Singapore, 2nd Edition*, 249–282.

25. **Cheng, H. D.**, **X. H. Jiang**, **Y. Sun**, and **J. L. Wang** (2001). Colour image segmentation: advances and prospects. *Pattern Recognition*, **34**(12), 2259–2281.

26. **Chow, T. W. S.** and **M. K. M. Rahman** (2007). A new image classification technique using tree-structured regional features. *Neurocomputing*, **70**(4-6), 1040–1050.

27. **Christoudias, C. M.**, **B. Georgescu**, and **P. Meer** (2002). Synergism in lowlevel vision. *16th IEEE Int. Conf. Pattern Recogn., New York*, **4**, 150–155.

28. **Ciurte, A.**, **X. Bresson**, **O. Cuisenaire**, **N. Houhou**, and **S. Nedevschi** (2014). Semi-supervised segmentation of ultrasound images based on patch representation and continuous min cut. *PLoS ONE 9(7): e100972*, **9**(7), 1–14.

29. **Cohen, B. H.** and **R. B. Lea** (2004). Essentials of statistics for the social and behavioral sciences.

30. **Collins, G. W.** (2003). Fundamental numerical methods and data analysis. `http://ads.harvard.edu/books/1990fnmd.book/`.

31. **Comaniciu, D.** and **P. Meer** (2002). Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**(5), 603–619.

32. **Cremers, D.** and **M. Rousson** (2007). International journal of computer vision. *International Journal of Computer Vision*, **72(2)**, 195–215.

33. **Cumani, A.** (1996). Edge detection in multispectral images. *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, **53**(1), 40–51.

34. **Dallal, G.** (2012). The little handbook of statistical practice.

35. **Decker, K.**, **V. Lesser**, and **R. Whitehair** (1990). Extending a blackboard architecture for approximate processing. *The Journal of Real-Time Systems*, **2**(1/2).

36. **Deng, Y.** and **B. Manjunath** (2001). Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30.

37. **Derin, H.** and **H. Elliott** (1987). Modelling and segmentation of noisy and textured images using gibbs random fields,. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **9**(1), 39–55.

38. **Dey, V.**, **Y. Zhang**, and **M. Zhong** (2010). A review on image segmentation techniques with remote sensing perspective. *IAPRS, ISPRS TC VII Symposium*, **XXXVIII,Part 7A**, 31–42.

39. **Dorin, C.** and **M. Peter** (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machachine Intelligence*, **24**(5), 603–619.

40. **Efros, A. A.** and **T. K. Leung** (1999). Texture synthesis by non-parametric sampling. *ICCV*.

41. **Elena González, M. X. A., Francesco Bianconi** and **S. A. Saetta** (2013). Automatic characterization of the visual appearance of industrial materials through colour and texture analysis: An overview of methods and applications. *Advances in Optical Technologies*.

42. **Felzenszwalb, P. F.** and **D. P. Huttenlocher** (2004). Efficient graph-based image segmentation. *International Journal on Computer Vision*, **59**(2), 167–181.

43. **Fisher, R. A.** and **F. Yates** (1947). Statistical tables for biological, agricultural and medical research.

44. **Freixenet, J.**, **X. M. noz**, **D. Raba**, **J. Martí**, and **X. Cufí** (2002). Yet another survey on image segmentation: Region and boundary information integration. *Computer Vision ECCV 2002*, 21–25.

45. **Fu, K. S.** and **J. K. Mui** (1981). A survey on image segmentation. *Pattern Recognition*, **13**, 3–16.

46. **Gabor, D.** (1946). Theory of communications. *Journal of Institution of Electrical Engineers*, **93**(26), 429–459.

47. **Ganesan, L.** and **P. Bhattacharyya** (1995). A statistical design of experiments approach for texture description. *Pattern Recognition*, **28**, 99–105.

48. **Ganesan, L.** and **P. Bhattacharyya** (1997). Edge detection in untextured and textured images-a common computational framework. *IEEE Trans. Sys. Man Cybern. B.*, **27**(5), 823–834.

49. **Golam, S.**, **A. Abraham**, and **D. Laurence** (2001). Texture classification based on dct and soft computing. *FUZZ-IEEE'01*.

50. **Gonzalez** and **Woods** (2007). Digital image processing, 3rd edition.

51. **Gonzälez, E.**, **F. Bianconi**, **M. X. Älvarez**, and **S. A. Saetta** (2013). Automatic characterization of the visual appearance of industrial materials through colour and texture analysis: An overview of methods and applications. *Advances in Optical Technologies*, 11.

52. **Goswami, T.** (2011). Edge detection in images- an overview. *Proceedings of the National Conference On Emerging Trends in Soft Computing,In Association with Indian Society for Rough Sets and Computer Society of India, Pune*, 51–58.

53. **Goswami, T.**, **A. Agarwal**, and **C. R. Rao** (2014*a*). Color texture features integration for unsupervised image segmentation based on orthogonal polynomial operators and hybrid color space. *Eighth International Conference on Image and Signal Processing, ICISP-2014, Elsevier*, 52–59.

54. **Goswami, T.**, **A. Agarwal**, and **C. R. Rao** (2014*b*). Statistical learning for texture characterization. *Indian Conference on Computer Vision, Graphics and Image Processing, (ICVGIP 2014), ACM*.

55. **Goswami, T.**, **A. Agarwal**, and **C. R. Rao** (2015). Hybrid region and edge based unsupervised color-texture segmentation for natural images. *International Journal of Information Processing*, **9**(1), 77–92.

56. **Guo, Z.**, **D. Zhang**, and **D. Zhang** (2010). A completed modeling of local binary pattern operator for texture classification. *Image Processing, IEEE Transactions on*, **19**(6), 1657–1663.

57. **Haindl, M.** and **S. Mikes** (2014). Prague texture segmentation benchmark dataset. `http://mosaic.utia.cas.cz/`.

58. **Han, J.** and **K.-K. Ma** (2007). Rotation-invariant and scale-invariant gabor features for texture image retrieval. *Image and Vision Computing*, **25**, 1474–1481.

59. **Haralick, R.**, **K. Shanmugan**, and **I. Dinstein** (1973). Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics SMC*, **3**, 610–621.

60. **Haralick, R. M.** and **L. G. Shapiro** (1985). Survey - image segmentation techniques. *Computer vision, Graphics and Image Processing*, **29**, 100–132.

61. **Healey, G.** (1989). Using color for geometry-insensitive segmentation. *Opt. Soc. Am.*, **22**(1), 920–937.

62. **Hedjam, R.** and **M. Mignotte** (2009). A hierarchical graph-based markovian clustering approach for the unsupervised segmentation of textured colour images. *Proceedings of the International Conference on Image Processing (ICIP09)*, 1365–1368.

63. **Hill, P. R.**, **C. N. Canagarajah**, and **D. R. Bull** (2003). Image segmentation using a texture gradient based watershed transform. *IEEE Trans. Image Proc.*, **12**(12), 1618âĂŞ1633.

64. **Ho, W.** and **C. Osborne** (1991). Texture segmentation using multilayered back propagation. *Proceedings Int. Joint Conf. Neural Net*, 981–986.

65. **Hoang, M.**, **J. Geusebroek**, and **A. Smeulders** (2005). Colour texture measurement and segmentation. *Signal Processing*, **85**(2), 265–275.

66. **Hoffman, R.** and **A. K. Jain** (1987). Segmentation and classification of range images. *IEEE PAMI*, **9**(5), 608–620.

67. **Huntsberger, T.**, **C. Jacobs**, and **R. Cannon** (1985). Iterative fuzzy image segmentation. *Pattern Recognition*, **18(2)**, 131–138.

68. **Ilea, D. E.** and **P. F. Whelan** (2008). Ctex an adaptive unsupervised segmentation algorithm based on colourâĂŞtexture coherence. *IEEE Transactions on Image Processing*, **17**(10), 1926–1939.

69. **Ilea, D. E.** and **P. F. Whelan** (2011). Image segmentation based on the integration of colour-texture descriptors-a review. *Pattern Recognition*, **44**, 2479–2501.

70. **Ilea, D. E.**, **P. F. Whelan**, and **O. Ghita** (2010). Unsupervised image segmentation based on the multi-resolution integration of adaptive local texture descriptors. *Proceedings of the Fifth International Conference on Computer Vision Theory and Applications (VISAPP2010)*, 17–21.

71. **Jain, A.** and **J. Dubes** (1988). Algorithms for clustering data. *Prentice Hall*.

72. **Jain, A. K.** (1981). Advances in mathematical models for image processing. *Proceedings of the IEEE*.

73. **Jain, A. K.** and **F. Farrokhnia** (1991). Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, **24**(12), 1167–1186.

74. **John, C.** (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**(6), 679–698.

75. **Kalloniatis, M.** and **C. Luu** (2007). Visual acuity. `http://webvision.med.utah.edu/book/part-viii-gabac-receptors/visual-acuity/`.

76. **Kassand, M.**, **A. Witkin**, and **D. Terzopoulos** (1988). Snakes: Active contour models. *International Journal of Computer Vision*, **1**(4), 321–331.

77. **Klinker, G.**, **S. Shafer**, and **T. Kanade** (1990). A physical approach to color image understanding. *Int. Journal of Computer Vision*, **4(1)**, 7–38.

78. **Kohonen, T.** (1995). Self-organizing map. *Springer, Berlin, Heidelberg (Third, Extended Edition, 2001)*.

79. **Koschan, A.** and **M. Abidi** (2005). Detection and classification of edges in color images. *IEEE Signal Processing Magazine*, **22**(1), 64–73.

80. **Koschan, A.** and **M. Abidi** (2008). Digital color image processing.

81. **Krinidas** and **Pitas** (2009). Color texture segmentation based on the modal energy of deformable surfaces. *IEEE Transactions on Image Processing*, **18**(7), 1613–1622.

82. **Krishnamoorthi, R.** and **P. Bhattacharya** (1998). Color edge extraction using orthogonal polynomials based zero crossings scheme. *Information Sciences*, **112**, 51–65.

83. **Krishnamoorthi, R.** and **N. Kannan** (2009). A new integer image coding technique based on orthogonal polynomials. *Image and Vision Computing*, **27**(8), 999–1006.

84. **Kumar, R. M.** and **K. Sreekumar** (2014). A survey on image feature descriptors. *International Journal of Computer Science and Information Technologies*, **5**(6), 7668–7673.

85. **Kurz, L.** and **M. H. Benteftifa** (2006). Analysis of variance in statistical image processing.

86. **L.A.Zadeh** (1965). Fuzzy sets. *Information and Control*, **8**(3), 338–353.

87. **Lee, T. S.** (1996). Image representation using 2d gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence - PAMI*, **18**(10), 959–971.

88. **Liu, H.**, **L. Li**, and **C. an Wu** (2014). Color image segmentation algorithms based on granular computing clustering. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, **7**(1), 155–168.

89. **Lowe, D.** (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, **60**(2), 91–110.

90. **Lucchese, L.** and **S. K. Mitra** (2001). Colour image segmentation:a state-of-the-art survey. *Proceedings of the Indian National Science Academy 67A(2)*, 207–221.

91. **Ma, W.** and **B. Manjunath** (1997). Edge flow: A framework of boundary detection and image segmentation. *In Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR 1997)*, 744–749.

92. **Mäenpää, T.** and **M. Pietikäinen** (2004). Classification with color and texture: jointly or separately ? *Pattern Recognition*, **37**, 1629–1640.

93. **Manjunath, B.** and **W. Ma** (1996). Texture features for browsing and retrieval of image data. *IEEE Trans, On Pattern Anal. Machine Intell.*, **18**(8), 837–842.

94. **Manjunath, B. S.**, **T. Simchony**, and **R. Chellappa** (1990). Stochastic and deterministic networks for texture segmentation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, **38**(6), 1039–1049.

95. **Martin, D.**, **C. Fowlkes**, **D. Tal**, and **J. Malik** (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *8th International Conference on Computer Vision*, **2**, 416–423.

96. **Massey, F. J.** (1951). The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, **46**(253), 68–78.

97. **Maxwell, B.** and **S. Shafer** (1995). Physics-based segmentation: Looking beyond color. (CMU-RI-TR-95-37).

98. **McInerney, T.** and **D. Terzopoulos** (1984). Stochastic relaxation, gibss distributions and the bayesion restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **6**, 721–741.

99. **McInerney, T.** and **D. Terzopoulos** (1996). Deformable models in medical image analysis:a survey. *Medical Image Analysis*, **1(2)**, 91–108.

100. **McLachlan, G.** and **D. Peel** (2000). Finite mixture models. *John Wiley and Sons*.

101. **Mecimore, I.** and **C. Creusere** (2009). Unsupervised bitstream based segmentation of images. *Proceedings of the Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop*, 643–647.

102. **Meilă, M.** (2005). Comparing clusterings: an axiomatic view. *Proceedings of the 22nd International Conference on Machine Learning*, 577–584.

103. **Mignotte, M.** (2010). A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation. *IEEE Transactions on Image Processing*, **19**(6), 1610–1624.

104. **Mignotte, M.** (2012). Mds-based segmentation model for the fusion of contour and texture cues in natural images. *Computer Vision and Image Understanding*, **116**(9), 981–990.

105. **Mirmehdi, M.** and **M. Petrou** (2000). Segmentation of colour textures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **22**(2), 142–159.

106. **MITMediaLab** (2005). Mitmedialab, vision texture-vistex database. `http://aiweb.techfak.uni-bielefeld.de/content/bworld-robot-control-software/`.

107. **Narasimhan, R.** (1963). Syntactic descriptions of pictures and gestalt phenomena of visual perception, technical report no. 142.

108. **Nevatia** (1977). A color edge detector and its use in scene segmentation. *IEEE Trans. System Man Cybernet*, **7**(11), 820–826.

109. **Nguyen, H.**, **M. Worring**, and **A. Dev.** (2000). Detection of moving objects in video using a robust motion similarity measure. *IEEE Transactions on Image Processing*, **9**(1), 137–141.

110. **Nikou, C.**, **C. Likas**, and **N. Galatsanos** (2010). A bayesian framework for image segmentation with spatially varying mixtures. *IEEE Transactions on Image Processing*, **19**(9), 2278–2289.

111. **Ohlander, R.**, **K. Price**, and **D.R.Reddy** (1978). Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*, **8**, 313–333.

112. **Ohlander, R. B.** (1975). Analysis of natural scenes. *Doctoral Dissertation,CMU*.

113. **Ohta, Y.**, **T. Kanade**, and **T. Sakai** (1980). Color information for region segmentation. *Computer Graphics and Image Processing*, **13**, 222–241.

114. **Ojala, T.**, **M. Pietikäinen**, and **T. Mäenpää** (2002). Multiresolution gray scale and rotation invariant texture analysis with local binary patterns. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**(7), 971–987.

115. **Otsu, N.** (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, **9**(1), 62–66.

116. **Pal, N. R.** and **S. K. Pal** (1993). A review on image segmentation techniques. *Pattern Recognition*, **26**(9), 1277–1294.

117. **Pal, S. K.**, **S. K. Meher**, and **A. Skowron** (2015). Data science, big data and granular mining. *Pattern Recognition Letters*.

118. **Panagiotakis, C.**, **I. Grinias**, and **G. Tziritas** (2014). Texture segmentation based on voting of blocks, bayesian flooding and region merging. *ICPR 2014*.

119. **Panjwani, D.** and **G. Healey** (1995). Unsupervised segmentation of textured color images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **17**(10), 939–954.

120. **Panjwani, D. K.** and **G. Healey** (1993). Unsupervised segmentation of textured colour images using markov random field models. *Proc. CVPR'93*, 776–777.

121. **Pappas, T.** (1992). An adaptive clustering algorithm for image segmentation. *IEEE Trans. on Signal Processing*, **40**(4), 901–913.

122. **Perez, A.** (1987). An iterative thresholding algorithm for image segmentation. *IEEE PAMI*, **9**(6), 742–750.

123. **Perez, F.** and **C. Koch** (1994). Toward color image segmentation in analog vlsi:algorithm and hardware. *International Journal of Computer Vision*, **12(1)**, 17–42.

124. **Priese, L.** and **V. Rehrmann** (1993). On hierarchical color segmentation and applications. *Proc. CVPR'93*, **1(2)**, 633–634.

125. **Raafat, H. M.** and **A. K. Wong** (1988). A texture information-directed region growing algorithm for image segmentation and region classification. *Computer Vision, Graphics, and Image Processing*, **43**(1), 1–21.

126. **Rao, A.** and **G. Lohse** (1993). Identifying high level features of texture perception. *Graphical Models and Image Processing*, **55**, 218–233.

127. **Rao, A.** and **G. Lohse** (1996). Towards a texture naming system: Identifying relevant dimensions of texture. *Vision Res.*, **36**(11), 1649–1669.

128. **Ray, S.** and **C. Sharma** (1990). Reasoning blackboard for data fusion. *Expert Systems for Management and Engineering*, 197–208.

129. **Reed, T.** and **J. Dubuf** (1993). A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, **57**(3), 359–372.

130. **Rocio, A.**, **Lizarraga-Morales**, **E. R. Sanchez-Yanez**, **V. Ayala-Ramirez**, and **F. E. Correa-Tome** (2014). Integration of color and texture cues in a rough setâĂŞbased segmentation method. *Journal of Electronic Imaging*, **23**(2), 1–12.

131. **Roselin, R.** and **K. Thangavel** (2012). Mammogram image segmentation using granular computing based on rough entropy. *Proceedings of the International Conference on Pattern Recognition, Informatics and Medical Engineering*.

132. **Rosenfeld, A.** and **A. Kak** (1982). Digital picture processing. *Academic Press, ISBN 0-12-597301-2*.

133. **Saber, E.**, **A. Tekalp**, **R. Eschbach**, and **K. Knox.** (1995). Annotation of natural scenes using adaptive color segmentation. *In Proc. of the SPIE - The Int'l Soc. for Optical Eng., Image and Video Proc.,volume III*, 72–80.

134. **Sahoo, P. K.**, **S. Soltani**, and **A. K. C. Wong.** (1988). A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, **41**, 233–260.

135. **Salembier, P.** and **L. Garrido** (2000). Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans Image Process.*, **9**(4), 561–576.

136. **Sanda, A.**, **E. Ezin**, **P. Gouton**, and **J. Tossa** (2013). Clifford algebra and gabor filter for color image texture characterization. *Proceedings of the 2013 International Conference on Signal-Image Technology and Internet-Based Systems*, 212–216.

137. **Schroff, F.** (2009). Semantic image segmentation and web-supervised visual learning. *PhD thesis from University of Oxford*.

138. **Shafer, S.** (1985). Using color to separate reflection components. *Color research and application*, **10(4)**, 210–218.

139. **Shankar, B. U.** (2007). Novel classification and segmentation techniques with application to remotely sensed images. *Transactions on rough sets VII*.

140. **Sharma, G.**, **W. Wu**, and **E. N. Dalal** (2005). The ciede2000 color difference formula: implementation notes, supplementary test data, and mathematical observations. *Color Research and Application*, **30**(1), 21–30.

141. **Shaw, K. B.** and **M. C. Lohrenz** (1992). A survey of digital image segmentation algorithms. *Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, Mississippi*.

142. **Shi, J.** and **J. Malik** (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(8), 888–905.

143. **Shi, L.** and **B. Funt** (2007). Quaternion color texture segmentation. *Computer Vision and Image Understanding*, **107**(1-2), 88–96.

144. **Shotton, J.**, **J. Winn**, **C. Rother**, and **A. Criminisi** (2009). Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *International Journal of Computer Vision*, **81**(1), 2–23.

145. **Skarbek, W.** and **A. Koschan** (1994). Colour image segmentation - a survey. *Technical report, Institute for Technical Informatics, Technical University of Berlin,*.

146. **S.K.Pal** (1989). Fuzzy sets and decision making in color image processing. *Artificial Intelligence and Applied Cybernatics*, 89–98.

147. **Suguna, R.** and **P. Anandhakumar** (2011). A novel feature extraction technique for texture discrimination using orthogonal polynomial operators. *European Journal of Scientific Research,ISSN 1450-216X*, **51**(4), 549–555.

148. **Sumengen, B.** (2005). A matlab toolbox implementing level set methods. `http://barissumengen.com/level_set_methods/`.

149. **Tamura, H.**, **S. Mori**, and **T. Yamawaki** (1978). Texural features corresponding to visual perception. *IEEE Trans. Syst. Man Cybern.*, **8**(6), 460–472.

150. **Tao, D.** (2009). The corel database for content based image retrieval. `https://sites.google.com/site/dctresearch/Home/content-based-image-retrieval/`.

151. **Terzopoulos, D.** and **K. Fleischer** (1988). Deformable models. *The Visual Computer*, **4**(6), 306–331.

152. **Tian, D. P.** (2013). A review on image feature extraction and representation techniques. *International Journal of Multimedia and Ubiquitous Engineering*, **8**(4), 385–396.

153. **Tkalčič, M.** and **J. Tasič** (2003). Colour spaces: perceptual, historical and applicational background. *EUROCON 2003*, **1**, 304–308.

154. **Todorovic, S.** and **N. Ahuja** (2009). Texel-based texture segmentation. *IEEE 12th International Conference on Computer Vision (ICCV)*, 841–848.

155. **Tominaga, S.** (1987). Expansion of color images using three perceptual attributes. *Pattern Recognition Letters*, **6**, 77–85.

156. **Tominaga, S.** (1990). A colour classification method for color images using a uniform color space. *Proc. 10th. Int. Conf. on Pattern Recognition*, **1**, 803–807.

157. **Trémeau, A.**, **S. Tominaga**, and **K. N. Plataniotis** (2008). Color in image and video processing: Most recent trends and future research directions. *Hindawi Publishing Corporation EURASIP Journal on Image and Video Processing*, 26.

158. **Trussell, H. J.**, **E. Saber**, and **M. Vrhel** (2005). Color image processing. *IEEE Signal Processing Magazine*, **22**(1), 14–22.

159. **Tuceryan, M.** and **A. Jain** (1998). Texture analysis - handbook of pattern recognition and computer vision. *in: C.H.Chen, L.F.Pau, P.S.P.Wang(Eds.) World Scientific Publishing*.

160. **Ugarriza, L. G.**, **E. Saber**, **S. R. Vantaram**, **V. Amuso**, **M. Shaw**, and **R. Bhaskar** (2009). Automatic image segmentation by dynamic region growth and multiresolution merging. *IEEE Trans. Image Proc.*, **18**(10), 2275–2288.

161. **Unnikrishnan, R.**, **C. Pantofaru**, and **M. Hebert** (2005). A measure for objective evaluation of image segmentation algorithms. *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **3**.

162. **Vaisey, J.** and **A. Gersho** (1992). Image compression with variable block size segmentation. *IEEE Trans. Signal Processing*, **40**(8), 2040–2060.

163. **Vandenbroucke, N.**, **L. Macaire**, and **J. Postaire** (2003). Color image segmentation by pixel classification in an adapted hybrid color space-application to soccer image analysis. *Computer Vision and Image Understanding*, **90**(2), 190–216.

164. **Vantaram**, **S. Rao**, **Saber**, and **Eli** (2012). Survey of contemporary trends in color image segmentation. *Journal of Electronic Imaging*, **21**(4), 040901–1–040901–28.

165. **Vantaram, S. R.** and **E. Saber** (2011). An adaptive bayesian clustering and multivariate region merging-based technique for efficient segmentation of color images. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Czech Republic*, 1077–1080.

166. **Vantaram, S. R.**, **E. Saber**, **S. A. Dianat**, **M. Q. Shaw**, and **R. Bhaskar** (2010). Multiresolution adaptive and progressive gradient-based color image segmentation. *J. Elect. Imag.*, **19**(1).

167. **Vergés-Llahí, J.** (2005). Survey on color segmentation. *Thesis: Color Constancy and Image Segmentation Techniques for Applications to Mobile Robotics*.

168. **Wang, J.**, **Y. Jia**, **X.-S. Hua**, **C. Zhang**, and **L. Quan** (2008). Normalized tree partitioning for image segmentation. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, AK, USA*, 1–8.

169. **Won, C.** and **D. Park** (1997). Image block classification and variable block size segmentation using a model-fitting criterion. *Optical Eng.*, **36**(8), 2204–2209.

170. **Wouwer, G.**, **P. Scheunders**, **S. Livens**, and **D. V. Dyck** (1999). Wavelet correlation signatures for color texture characterization. *Pattern Recognition*, **32**(3), 443–451.

171. **Xie** and **Mirmehdi** (2007*a*). Color image segmentation using texems. *Annals of the British Machine Vision Association (BMVA)*, **6**, 1–10.

172. **Xie, X.** and **M. Mirmehdi** (2007*b*). Texems: Texture exemplars for defect detection on random textured surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(8), 1454–1464.

173. **Yang, J.**, **Y.Jiang**, **A. G. Hauptmann**, and **C. Ngo** (2007). Evaluating bag-of-visual-words representations in scene classification. *In Proc. Workshop on MIR*.

174. **Yokoya, N.** and **M. D. Levine** (1989). Range image segmentation based on differential geometry: A hybrid approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(6), 643–649.

175. **Yu, S. X.** and **J. Shi** (2003). Multiclass spectral clustering. *IEEE Internatioanl Conference on Computer Vision (ICCV)*, 313–319.

176. **Zhang, H.**, **J. E. Fritts**, and **S. A. Goldman** (2008). Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, **110**, 260–280.

177. **Zhang, J.** and **T. Tan** (2002). Brief review of invariant texture analysis methods. *Pattern Recognition*, **35**(3), 735–747.

178. **Zhang, Y. J.** (2006). An overview of image and video segmentation in the last 40 years. *Advances in Image and Video Segmentation, IRM Press, Chapter 1 (1-15).*

179. **Zhu, S.**, **C. Guo**, **Y. Wu**, and **Y. Wang** (2002). What are textons? *Proc. of 7th European Conf. on Computer Vision, Copenhagen, Denmark,*, 793–807.

180. **Zhu, X.** (2006). Semi-supervised learning literature survey. *Technical Report 1530, Computer Sciences, University of Wisconsin-Madison.*