Software Requirements Specification

Under the supervision of
Mrs. Arti Gautam
(Project Guide)

April 20, 2018

A Salient Region Extraction Based on Color and Texture Features

Submitted By:

Aakarsh
Shitiz Kumar Aggarwal

Table of Contents

List of Figures

1.0. Introduction

1.1. Purpose

Natural images have several semantic objects. Our objective is to divide these natural images into salient regions according to the human visual perception of images. One salient region should be compact, complete and significant enough, and neither a small region nor a fragmentary region can be one meaningful region. Furthermore, in terms of colors and textures, the salient regions should also have distinguishing features from their neighbor regions or other regions. Provided that each region is characterized by its feature vectors, they are the syntheses of color and texture features.
In order to achieve this we firstly convert the RGB image into intensity image. This image is segmented according to any simple algorithm. The multi resolution Gabor texture features of the segmented regions are calculated which are similar to human visual multichannel perception of texture. Also we use RGB image to calculate the dominant color of the selected regions. Now these parameters are used in the novel saliency metric that we propose. All the regions having the saliency values greater than that of a particular threshold are selected. This way salient region is extracted from a natural image.

1.2. Scope of Project

1. **Medical imaging:**
   Medical imaging is creating the image of human organs for clinical purposes or medical science. Proper image segmentation is most important part for the classification of various internal organs or tumors. For identifying the abnormality we must be able to classify the images. For that we must extract the salient regions from the image. Other applications of this are:
   - Locate tumors and other pathologies
   - Measure tissue volumes
   - Computer-guided surgery
   - Diagnosis
   - Treatment planning
   - Study of anatomical structure

2. **Locating objects in satellite images:**
   The images from satellite may be used for mapping or forestry. Whatever may be the purpose but we must be able to identify what the image contains. In order to classify an image into road or forest or agricultural land we must be able to extract the portions of different object from the images.

3. **Face recognition systems:**
   A facial recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a facial database. It is typically used in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems.

4. **Iris recognition system:**
   Iris recognition is an automated method of biometric identification that uses mathematical pattern-recognition techniques on video images of the iris of an individual's eyes, whose complex random patterns are unique and can be seen from some distance. Not to be confused with another, less prevalent, ocular-based technology, retina scanning, iris recognition uses camera technology with subtle infrared illumination to acquire images of the detail-rich, intricate structures of the iris. Digital templates encoded from these patterns by mathematical and statistical algorithms allow unambiguous positive identification of an individual.

5. **Machine vision:**
   Machine vision (MV) is the process of applying a range of technologies to provide imaging-based automatic inspection, process control and robot guidance in industrial applications. While the scope of MV is broad and a comprehensive definition is difficult to distill a "generally accepted definition of machine vision is '... the analysis of images to extract data for controlling a process or activity.'"

1.3. References

- Sangoon lee, Crawford M.M., “Unsupervised image classification using Hierarchical clustering with a Bayesian similarity measure” PP. 312-320, (IEEE 2005)
- Fredembach C., Schroder M., Susstrunk S., “Eigen regions for image classification” PP. 1645-1649 (IEEE 2004)
- John F Canny., “a computational approach to edge detection. Pattern analysis and machine intelligence” PP. 679-698 (IEEE 1986)
- Mingxin Zhang, Zhaowei Shang, JunyiShen. Title in “A novel unsupervised color image fuzzy classification scheme based on salient region”, PP. 4-5 .(IJCSNS 2008)
- Mingxin Zhang, Zhaogan Lu, JunyiShen. Title in “A Robust Salient Region Extraction based on Color and Texture Features”, PP. 1-4.

- Olivier Chapelle, Patrick Haffner and Vladimir Vapnik. Title in “SVMs used for Histograms Based Image classification”. PP 1-3 (AT and T labs)
- G. Das and S. Ray. Title in “A compact feature representation and image indexing in Content-Based Image Retrieval.” , pages 387–391 In Proceedings of Image and Vision Computing New Zealand 2005 Conference (IVCNZ 2005)
- Timor Kadir, Andrew Zisserman, and Michael Brady. Title in “An affine invariant salient region detector”, PP 1-4.
- Myron Flickner, HarpreetSawhney, Wayne Niblack. Title in “Query by Image and Video Content. The QBIC system” PP 1-10 (IBM research center 1995)
- F.Mai,Y Hung, H.Zhong and W.Sze. Title in “ A Hierarchical approach for fast and Robust Ellipse Extraction.” PP 2512-2524, (Pattern Recognition 2008)
- Takuya Maekawa, Takahiro Hara, ShojuoHishio. Title in “Image classification for Web Browsing”, PP. 2.
- Caelli T, Reye D. On “The classification of image regions by color, texture and shape. Pattern Recogn” 1993;26:461– 470.
- Berens J. “Image indexing using compressed color histograms.” Ph.D. Thesis, University of East Anglia, Norwich, 2003.

1.4. Overview of Document

The research work of content-based image retrieval (CBIR) over the last few years has shown that retrieving images through matching images solely on that basis of global similarities is often too crude to produce satisfactory results. On the other hand, semantic object-based image retrieval is still far too rudimentary and fragile to produce reliable results. Intermediate-level processing between high and low-level processing for content-based image retrieval is required. Therefore, it is necessary to identify the perceptually salient and semantically meaningful regions in images. However, it is difficult to isolate the meaningful region of interest from the scene without a prior knowledge. In a common case, the regions with many abrupt changes or some unpredictable characteristics often attract the human's attention, are considered as the salient regions of images in this paper. Thus, salient regions of one image are those regions that could present the main contents of the image, which were detected according to local features as such as colors, textures and shapes. Moreover, we believe that these salient regions are potentially more effective for image indexing, retrieval and classification. There are many approaches like hierarchical clustering based approach (HCBA), and the Eigen regions based strategy. These all are the unsupervised approaches. Unsupervised approaches increase the complexity. They are more complicated and take long time to give results. Also the new regions which cannot be detected by these approaches will continue to go unnoticed. Actually, as the complexity of

practical image analysis, ones could not deal with every region for different sizes, and it's very important to exclude those implicit regions smaller than the specified size. As the regions to present the main content of one image, the salient regions should be larger than the given region sizes and with rich semantic information to human being.
Consequently, we propose a novel salient region detection method, which could take account of the implicit region sizes and their saliency relative to other regions. The edges of the region are detected as the local gradient i.e. local maxima in the given region of the images. The detected local gradients are edges. These edges are not always continuous so overcoming this we will detect the continuous edged which will definitely give us the regions in the image. Further all the regions smaller than a threshold region say N_0 are ignored as they are no implicit regions. Also the regions which are inside the larger regions are found so that they can be merged.
For clustering the images low level histogram based technique is used so that the features of low and high level image analysis can be combined without losing the simplicity. Although color histograms are very old technique they are simple and fast which is necessity of image indexing and retrieval.

2.0. Overall Description

2.1 System Environment

2.1.1. Java 1.6 or Higher

Java platform, standard Edition or Java SE is a widely used platform for programming in the Java language. It is the Java Platform used to deploy portable applications for general use. In practical terms, Java SE consists of a virtual machine, which must be used to run Java programs, together with a set of libraries (or "packages") needed to allow the use of the file systems, networks, graphical interfaces, and so on, from within those programs.

2.1.2 Development Tools (.Net beans 6.1 or Eclipse)

- NetBeans and Eclipse refers to both a platform framework for both desktop applications, and an integrated development environment (IDE) for developing with Java, JavaScript, PHP, Python, Ruby, Groovy, C, C++, Scala, Clojure, and others.
- The NetBeans IDE is written in Java and runs everywhere where a JVM is installed, including Windows, Mac OS, Linux, and Solaris. A JDK is required for Java development functionality, but it is not required for development in other programming languages.
- The NetBeans Platform allows applications to be developed from a set of modular software components called modules. Applications based on the NetBeans Platform (including the NetBeans IDE) can be extended by third party developers.
- The NetBeans IDE is an open-source integrated development environment. NetBeans IDE supports development of all Java application types (Java SE including JavaFX, Java ME,

Web, EJB and Mobile applications) out of the box. Among other features are an Ant-Based project system, Maven support, refractorings, Version Control (supporting CVS, Subversion, Mercurial and Clearcase)

2.2 Functional Requirements Specification

Use case diagrams are used to gather the requirements of a system including internal and external influences. These requirements are mostly design requirements. Hence, when a system is analyzed to gather its functionalities, use cases are prepared and actors are identified.

The purposes of use case diagrams can be said to be as follows:

1. Used to gather the requirements of a system.
2. Used to get an outside view of a system.
3. Identify the external and internal factors influencing the system.
4. Shows the interactions among the requirements and actors.

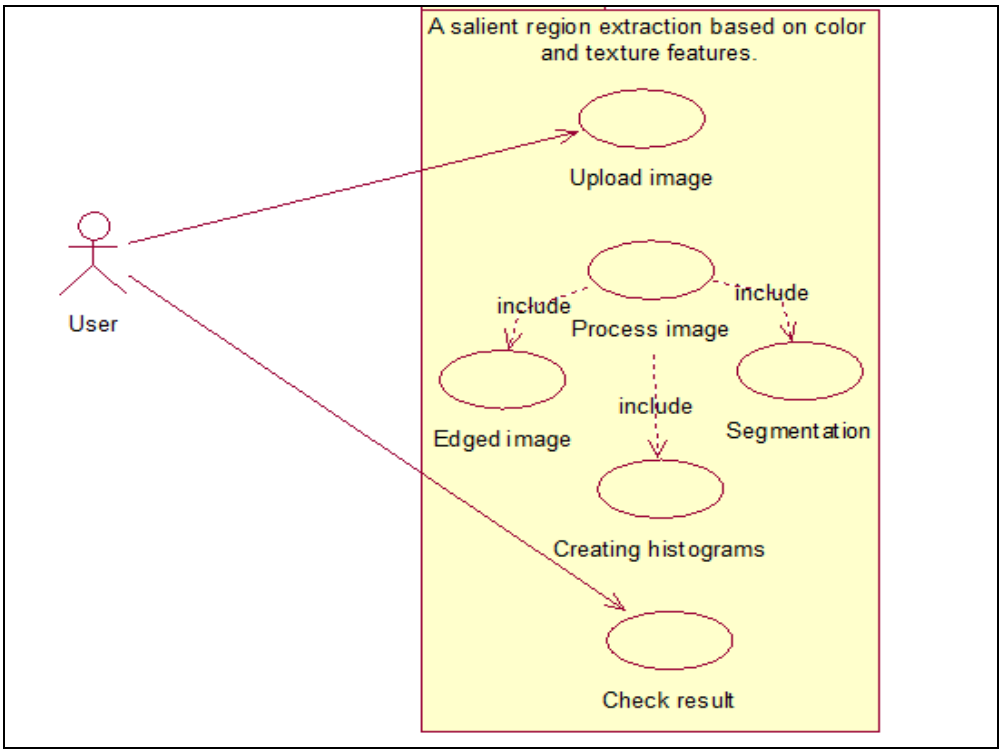
As per our application, we have identified the actors and functionalities to be represented as use case. They are as follows:

1. Actors- User
2. Functionalities-
    - Upload image
    - Process image
    - Edged image
    - Creating histograms
    - Segmentation
    - Check results

These diagrams are used at a very high level of design. This high level design is refined again and again to get a complete and practical picture of the system. A well-structured use case also describes the pre-condition, post condition, and exceptions. These extra elements are used to make test cases when performing the testing.

Use case diagrams specify the events of a system and their flows. But use case diagram never describes how they are implemented. Use case diagram can be imagined as a black box where only the input, output, and the function of the black box is known.

This section outlines the use cases for each of the active readers separately. The reader, the author and the reviewer have only one use case apiece while the editor is main actor in this system.



Brief Description
The use case diagram of our software depicting the functions being performed along with the actors interacting with the software.

2.3 Class Diagram

Class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing, and documenting different aspects of a system but also for constructing executable code of the software application.

Class diagram describes the attributes and operations of a class and also the constraints imposed on the system. The class diagrams are widely used in the modeling of object oriented systems because they are the only UML diagrams, which can be mapped directly with object-oriented languages.

Class diagram shows a collection of classes, interfaces, associations, collaborations, and constraints. It is also known as a structural diagram. The purpose of class diagram is to model the static view of an application. It is also considered as the foundation for component and deployment diagrams.

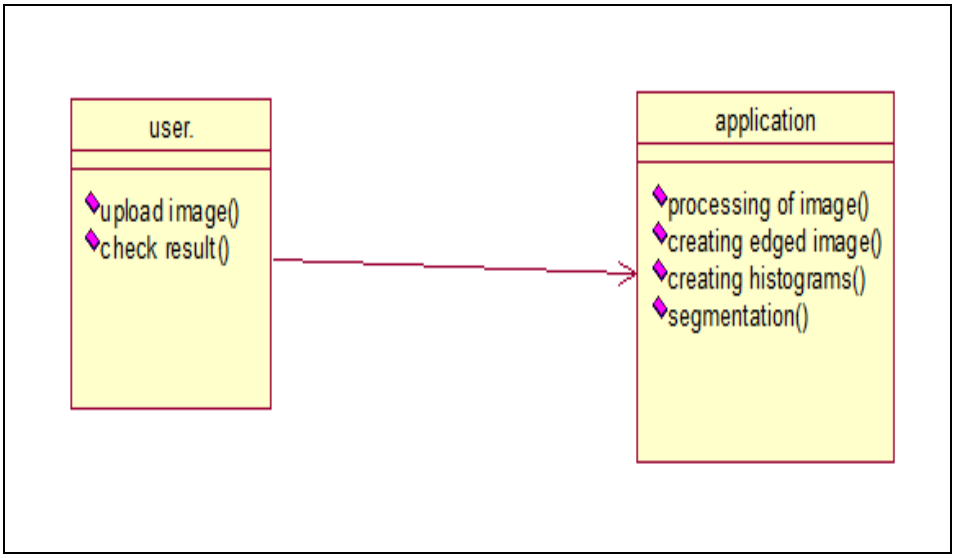


Figure 2.3 Class Diagram

Figure 2.3 depicts the class diagram of our software where two entities are interacting with each other and have their own defined operations which they perform.

3. REQUIREMENT ANALYSIS

3.1 Data Flow Diagrams

Data flow diagram is graphical representation of flow of data in an information system. It is capable of depicting incoming data flow, outgoing data flow and stored data. The DFD does not mention anything about how data flows through the system.

Levels of DFD:

1. Level 0- Highest abstraction level DFD is known as Level 0 DFD, which depicts the entire information system as one diagram concealing all the underlying details. Level 0 DFDs are also known as context level DFDs.
2. Level 1- The Level 0 DFD is broken down into more specific, Level 1 DFD. Level 1 DFD depicts basic modules in the system and flow of data among various modules. Level 1 DFD also mentions basic processes and sources of information.
3. Level 2-At this level, DFD shows how data flows inside the modules mentioned in Level 1.

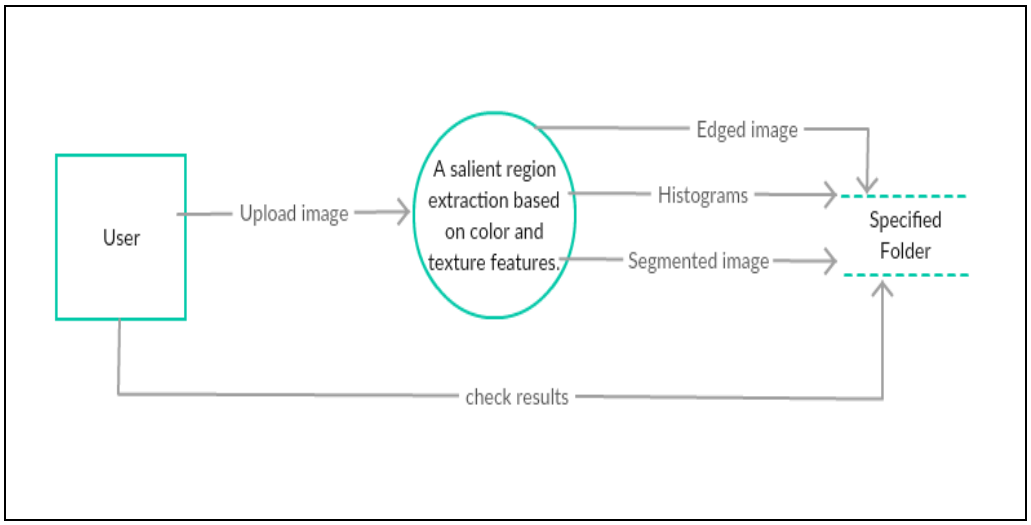


Figure 3.1 DFD- Level 0

Figure 3.1 shows the level 0 data flow diagram where the overview of the whole working of software is defined. It explains the overall functions, input and output of the software.

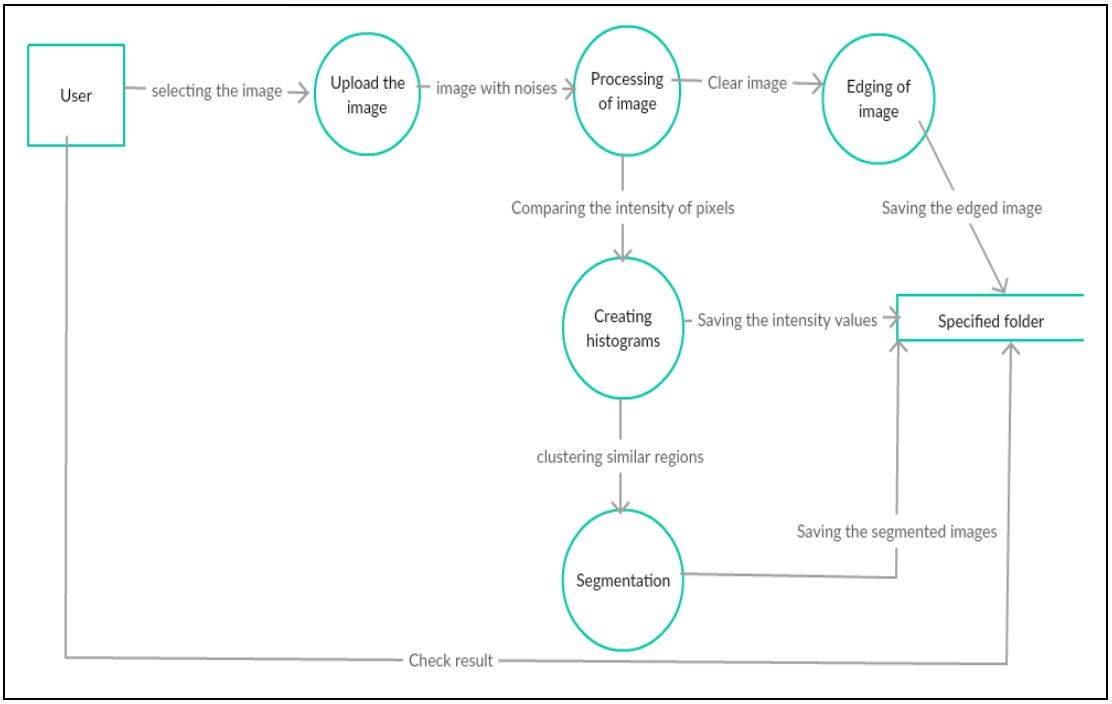


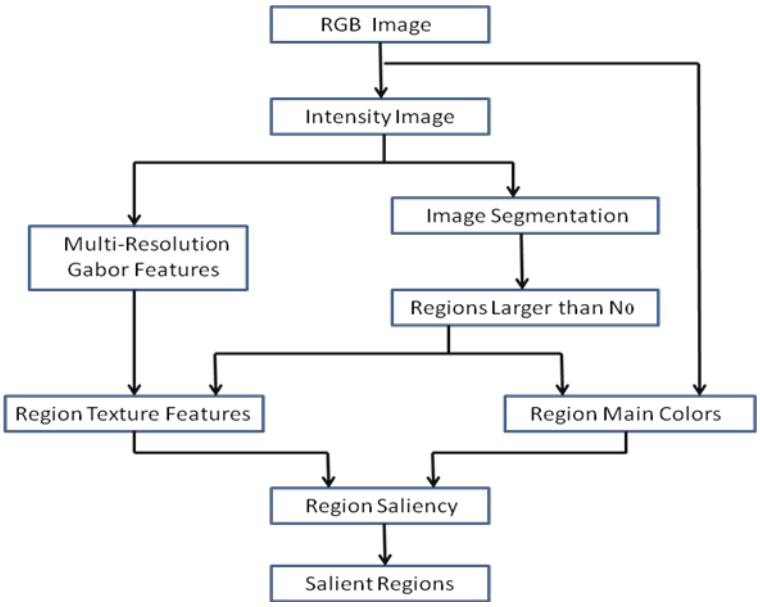
Figure 3.2 DFD- Level 1

Figure 3.2 gives the detailed description about the flow of data in our software. It explains the inputs and the outputs of our main functions and how the final images are stored.

3.2 Hardware Requirements

1. Pentium 4 Processor.
2. 1 GB RAM minimum.
3. Other basic hardware.

3.3 Modules in the project:



1. **RGB Image:**

   The first step of segmenting the image is the RGB image. In RGB model the primary colors are red, green and blue. It is an additive model in which colors are produced by adding all the components. In RGB model white imply that all the colors are present while black implies the absence of color. RGB model is represented by the unit cube. With one corner located at the origin of 3 dimensional spaces. The axes are named as RGB with values ranging from 0 to 1. Origin (0,0,0) is considered as the black and opposite corner (1,1,1) is white.

2. **Intensity Image:**

   The next step includes the calculation of the intensity of each dominant region from the image. The intensity can be calculated using the HSI model. In this model the luminance, or also known as intensity is decoupled from the color information already gathered, which is described using hue channel and saturation channel. Hue and Saturation of color responds closely to the way human perceive color. This model is suitable for interactive manipulation of color images where changes occur for each variable shift that corresponds to that operator inputs.

3. **Multi-Resolution Texture Feature:**

Texture is an important property of surfaces which characterizes the nature of the surfaces. An important task in image processing and machine vision is the task of segmenting regions of different texture in an image. Texture segmentation is the process of partitioning an image into regions based on their texture. The texture segmentation is being inspired by multi-channel operation of the human visual system for interpreting texture. A typical definition of texture is, "a spatial arrangement of local (grey-level) intensity attributes which are correlated within areas of the visual scene corresponding to surface regions."

This lead to the conclusion that people are sensitive to three texture properties: repetition, directionality and complexity. So, we have focused on the multi-channel Gabor filtering approach for identifying different texture regions. The multi-channel filtering approach is actually a multi-resolution decomposition process comprising of series of filters of different range of frequency and orientation. The filters with smaller bandwidth are desirable because they allow making finer distinction among different textures.

4. **Image Segmentation:**

As the name suggests, image segmentation is the process of partitioning digital image into different segments based on its salient regions present in the image and the dominant colors of the image. The image segmentation has three methods: manual, automatic and semi-automatic. The goal of image segmentation is clustering of pixels into salient image regions or simplifying the representation of the image into something that is more meaningful and easier to analyze. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics.

The result of image segmentation is a set of segments that collectively cover the entire image, or set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristics or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to some characteristics.

5. **Dominant Color Feature:**

Although people can distinct any acute differences in color spectrum, the human visual system cannot simultaneously perceive a large number of colors. Therefore, dominant colors can account for the spatially varying image characteristics in different implicit meaningful regions, and are invariant to region scales. For every salient region achieved its color values in terms of red, green and blue color components which are then classified into k groups via the adaptive clustering algorithms (ACA). The adaptive clustering algorithm (ACA) is an iterative algorithm that segments the image into k classes, where each class is characterized by a spatially varying characteristic function. The key of adapting to the local image values is that the ACA estimates the characteristic function by averaging the image values corresponding to each class over a sliding window whose size decreases as the algorithm converges. Then the average values of each group are considered as dominant colors, and their corresponding percentage of occurrence in current regions are also further figured out according to region size and the color dominant pixel.

6. **Salient Region Extraction:**

One salient region should be compact, complete and significant enough, and neither a small region nor a fragmentary region can be one meaningful region. Furthermore, in terms of color and texture, the salient region should also have distinguishing features from their neighbor regions or other regions. Provided that each region is characterized by its feature vectors, their synthesis of color and texture features. According to the feature vectors of all the regions, the saliency of one region is given as the sum of its Euclidean distance from all the other regions. More than one salient region should be found to meet for the requirements of image analysis, so the set of salient regions are selected as the first several regions according to their saliency values or those regions whose saliency values are greater than one given special threshold. Meanwhile, in order to reduce the competition complexity of the saliency values, some smaller fragmentary regions must be excluded from the candidate regions, as these regions are not considered as meaningful regions for incoming image processing, i.e., regions smaller than N_0.