AN ACTIVE CONTOUR SEGMENTATION USING THE PERSPECTIVE BOX CONCEPT FOR FIGURE-GROUND OBJECT DETECTION

MUHAMMAD KHAIRI ISMAIL

UNIVERSITI TEKNOLOGI MALAYSIA

AN ACTIVE CONTOUR SEGMENTATION USING THE PERSPECTIVE BOX CONCEPT FOR FIGURE-GROUND OBJECT DETECTION

MUHAMMAD KHAIRI ISMAIL

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

JANUARY 2012

Dedicated to my dear parents, Haji Ismail Abdullah and Mrs Che Nah Disa, my beloved brothers and sister,my research supervisor , supportive colleagues, and friends.Thank you very much for the motivation, support and understanding.

ACKNOWLEDGEMENT

I wish to extend my grateful thanks to everyone who has contributed directly or indirectly to the preparation of this research. I would like to express my gratitude to my supervisor, Assoc. Prof. Dr. Ali Selamat, for his generous time, commitment and advice. Throughout my research work, he has encouraged me to develop independent thinking and research skills. He stimulated my analytical thinking and greatly assisted me with scientific writing. His support and passion towards this research has encouraged me to complete this thesis as presented here.

Special thanks to my examiners for my master's first assessment and UTM lecturers had provided valuable comments and suggestions for my research direction. Thanks are due to Muhammad Tarmizi Lockman, Zhi-Sam Lee, Imam Much Ibnu Subroto, Ng Choon Ching, Siti Dianah Abdul Bujang, Siti Nurkhadijah Aishah Ibrahim for their help and valuable comments. In addition, I would like to present my sincere appreciation to Professor Richard L. Spear for the valuable suggestions in improving the thesis.

A great gratitude also goes to the Ministry of Science, Technology & Innovation (MOSTI), Malaysia and Research Management Center Universiti Teknologi Malaysia (UTM), for providing financial support under Vot 79227 for this research. Last but not least, I would like to express my great gratitude and sincere thanks to my beloved family, friends, for all their constant encouragement, inspiration and patience throughout my research work.

ABSTRACT

Image segmentation has been widely employed as a vital pre-processing phase in content-based image retrieval system, object tracking system, camera surveillance system, etc. Image segmentation procedure helps visual system to detect and to recognise main objects in a digital image scene. The active contour approach is a common technique used in the initial processes to detect objects. According to previous researches, the active contour approach used in region-based segmentation procedures had shown impressive results in segmenting an image scene into different object categories. However, when dealing with a complex image scene, the active contour approach was unable to segment the figure-ground and background of the given objects. Therefore, this thesis proposes an alternative method in image segmentation based on the perspective box concept, to detect main objects. The proposed concept introduces vanishing points to model an image scene based on the visual attention concept. The vanishing points are used to identify the relevant image information or attention points in the scene by using a Hough transform. Once the points of attention in the objects have been identified, the region of interest will be minimised in the active contour segmentation process. Then, the statistical region merging algorithm is used to construct multiple layers of the image object in order to produce the bounding box coordinate in XML codes. For benchmarking purposes, the codes were compared with PASCAL Visual Object Classes Challenge 2010 (VOC 2010) dataset by using three performance parameters namely precision, recall and F-measure. The results have shown that the average detection rate is more than 50%. Therefore, the proposed approach outperforms the active contour segmentation technique. In addition, procedures of the perspective box concept can be carried out automatically without any manual intervention or reliance on intelligent systems.

ABSTRAK

Pembahagian imej telah digunakan secara meluas bagi fasa penting dalam pra-pemprosesan untuk sistem dapatan semula berdasarkan kandungan imej, sistem pengesanan objek, sistem pengawasan kamera, dan lain-lain. Prosedur pembahagian imej membantu sistem visual untuk mengesan dan mengenalpasti objek utama dalam imej digital. kaedah garis bentuk aktif adalah satu teknik yang biasa digunakan dalam proses awal bagi mengesan objek. Menurut kajian sebelumnya, kaedah garis bentuk aktif yang digunakan dalam prosedur pembahagian berasaskan rantau telah menunjukkan hasil yang memberangsangkan dalam memisahkan babak imej kepada beberapa kategori objek yang berbeza. Walau bagaimanapun, apabila berurusan dengan babak imej yang kompleks, pendekatan garis bentuk aktif tidak dapat memisahkan *figure-ground* dan latar belakang dalam objek yang diberikan. Oleh itu, tesis ini mencadangkan satu kaedah alternatif dalam pembahagian imej berdasarkan konsep kotak perspektif, untuk mengesan objek utama. Konsep yang dicadangkan memperkenalkan titik lenyap bagi model babak imej yang berdasarkan konsep tumpuan visual. Titik lenyap ini telah digunakan untuk mengenal pasti maklumat imej yang berkaitan atau titik tumpuan di tempat kejadian dengan menggunakan pengubah Hough. Apabila titik tumpuan dalam objek telah dikenal pasti, rantau yang penting akan dapat dikurangkan untuk proses pembahagian garis bentuk aktif. Kemudian, algoritma penggabungan rantau statistik digunakan untuk membina pelbagai lapisan dalam imej objek untuk menghasilkan koordinat kotak batasan dalam kod XML. Untuk tujuan penanda aras, kod tersebut dibandingkan dengan set data Pertandingan Pengkelasan Objek Tahun 2010 anjuran PASCAL menggunakan tiga parameter prestasi iaitu kepersisan, perolehan kembali dan Ukuran-F. Keputusan menunjukkan kadar purata ketepatan adalah melebihi 50%. Oleh itu, pendekatan

yang dicadangkan mengatasi teknik pembahagian garis bentuk aktif. Di samping itu, prosedur konsep kotak perspektif boleh dijalankan secara automatik tanpa sebarang campur tangan pengguna atau pergantungan ke atas sistem pintar.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	**ii**
	DEDICATION	**iii**
	ACKNOWLEDGEMENT	**iv**
	ABSTRACT	**v**
	ABSTRAK	**vi**
	TABLE OF CONTENTS	**viii**
	LIST OF TABLES	**xiv**
	LIST OF FIGURES	**xv**
	LIST OF ABBREVIATIONS	**xix**
	LIST OF SYMBOLS	**xxi**
	LIST OF APPENDICES	**xxiv**
1	**INTRODUCTION**	**1**
	1.1 Introduction	1
	1.2 Problem Background	3
	1.3 Problem Statement	5
	1.4 Hypothesis	7
	1.5 Aim	7
	1.6 Objectives of Study	8
	1.7 Scopes of Study	8

1.8 Significance of the Research 9
1.9 Contribution of the Work 10
1.10 Research Plan 10
1.11 Summary and Thesis Organization 11

2 LITERATURE REVIEW 13
2.1 Introduction 13
2.2 Image Segmentation Concept 14
2.3 Visual Strategies on Scene Understanding 14
2.3.1 Visual Attention 17
2.3.2 Visual Perception 18
2.3.3 Visual Literacy 20
2.4 Images Visual Components 20
2.5 Introduction to the Image Segmentation Problem 22
2.5.1 Top-Down Approach 23
2.5.2 Bottom-up Approach 24
2.6 Image Segmentation Process 26
2.6.1 Edge Based Segmentation 27
2.6.1.1 The Prewitt Edge Detection 28
2.6.1.2 Canny Edge Detection 29
2.6.1.3 Sobel Edge Detection 30
2.6.1.4 Roberts Edge Detection 31
2.6.1.5 Phase Congruency Edge Detection 32
2.6.2 Region-Based Segmentation 34
2.6.2.1 Color Segmentation 35
2.6.2.2 Texture Segmentation 36
2.6.2.3 Shape Segmentation 37
2.6.3 Model Based Object Segmentation 38
2.6.3.1 The Mumford Shah Model 38

2.6.3.2 Region Growing Algorithm 39
2.6.3.3 Statistical Region Merging (SRM) 40
2.6.3.4 Active Contours Segmentation 42
2.6.4 Scene Based Segmentation 43
2.6.4.1 Perspective Box Basic Concept 45
2.6.5 Standard Hough Transform (SHT) 47
2.6.6 Perspective Box in Figure-Ground Object Analysis 50
2.7 Bounding Box Detection System 51
2.7.1 Figure-Ground Object Detection Evaluation 52
2.7.2 Evaluation Measures 53
2.7.2.1 Precision 53
2.7.2.2 Recall 54
2.7.2.3 F-Measure (F1 Score) 54
2.8 Image Benchmarking Results 55
2.8.1 Constructing the Graphs 56
2.9 Summary 57

3 RESEARCH METHODOLOGY 58
3.1 Introduction 58
3.2 Overview Methodology 59
3.2.1 Dataset Selection 59
3.2.2 Research Design 59
3.2.3 System Implementation and Analysis 60
3.3 Detail of Research Flowchart 60
3.3.1.Step 1:Conventional Figure-Ground Object Segmentation 61
3.3.2 Step 2: Redesign the Current Approach to Adapt the Perspective Box Concept 63

3.3.3 Step 3: Creation of the Specification Element with Concept Modeling 64
3.3.4 Step 4: Using Statistical Region Merging (SRM) in the Region Segmentation Process 66
3.3.5 Step 5: Design and Create New Procedure on Prototyping with MATLAB Programming 66
3.4 The Designing and Modeling Phase 67
3.5 The Result and Benchmarking Phase 67
3.6 Summary 69

4 PERSPECTIVE BOX CONCEPT 70
4.1 Introduction 70
4.2 Assumptions and Ideas 71
4.3 System Design and Modeling 74
4.3.1 Perspective Box for Figure-ground Object Detection 76
4.3.1.1 Image Pre-Processing: Edge Detection 76
4.3.1.2 Module 1: Vanishing Point Estimation 78
4.3.2 Module 2: Converting Images into Statistical Region Merging 83
4.3.3 Module 3: Create an Initial Region Mask for the Segmentation Process 86
4.3.4 Module 4: Bounding Box Detection 92
4.4 Summary 94

5	**EXPERIMENTAL RESULTS AND DISCUSSION**	**95**
	5.1 Introduction	95
	5.2 Evaluations Procedure with Dataset Selection	96
	5.2.1 Ground Truth Dataset	98
	5.3 Experiments on Testing Dataset	99
	5.3.1 Edge and Vanishing Point Detection Result	100
	5.3.2 The Segmentation Results	104
	5.4 Experiments on Evaluation Dataset	108
	5.4.1 Image Benchmarking Result	109
	5.4.2 Bounding Box and Object Detection Evaluation	109
	5.4.2.1 XML Script Input	110
	5.4.3 Detection Result	111
	5.4.4 Performance Measure Results	114
	5.4.4.1 Quality Measures	114
	Active Contour Segmentation Result	115
	Active Contour with Statistical Region Merging (SRM) Segmentation Results	116
	Perspective Box Segmentation Results	118
	5.4.4.2 Performance Measures using Recall Versus Precision Graph	119
	5.5 Summary	121
6	**CONCLUSION**	**122**
	6.1 Introduction	122
	6.2 Summary of the Thesis	123
	6.3 Research Significance	124
	6.4 Thesis Contributions	127

6.5 Future Work and Other Issues 127

6.6 Conclusion 128

REFERENCES **129**

APPENDICES A - F **136**

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	The components of image visualization (Dunstan and Bernard, 1979)	22
5.1	A summary of result on vanishing point detection	103
5.2	Figure-ground object detection result on testing dataset	106
5.3	Figure-ground object detection result on evaluation dataset	112
6.1	Objective versus outcome	125

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1	Images without dominant themes for figure-ground object and background image	4
1.2	A dog picture demonstrates the principle of dominance in visual perception (Steven Lehar, 2003)	6
2.1	Vehicle plate number	15
2.2	Rectangle shape background with letters and numbers and flower textures	15
2.3	Image scenes with multiple figure-ground objects and background overlays (Everingham, M. et al., 2010)	16
2.4	Image with abstract and artistic view (www.flickr.com)	19
2.5	Visual concepts between image visual components	21
2.6	Segmentation results using the top-down approach (Bjorkman, M. and Eklundh, J.O., 2005)	23
2.7	VOCUS: the visual attention system for bottom up approach using an attention center (Frintrop, S., 2006)	25
2.8	An original image with definite ground truth data selection in an evaluation process (Everingham, M. et al., 2010)	26

2.9	Masks used by Prewitt operator (Raman, M. and Sobel, J. S.,2006)	28
2.10	Gradient mask value with convolution direction (Canny, J.,1986)	29
2.11	Masks used by sobel operator (Jagadish, H. P. and Shambhavi, D. S.,2010)	30
2.12	Mask for Roberts operator (Roushdy, M.,2006)	31
2.13	A color segmentation using color space thresholding. (Bruce, J., 2000)	35
2.14	A multi-scale aggregation on shape elements for filter responses (Galun, M., 2003)	36
2.15	A Shape segmentation using level sets (H. E. Abd El Munim, 2005)	37
2.16	Basic form of an active contour (Bakoš M., 2007)	42
2.17	An Image Scene with 3D information on multiple figure-ground objects (Everingham, M. et al., 2010)	44
2.18	Image Scene with the perspective view in 3D	46
2.19	Vanishing points detection on an image plane using standard Hough transforms on a Gaussian sphere (F.A. van den Heuvel, 1998)	48
2.20	Figure-ground object projections in perspective box	50
2.21	The differences between (a) ground truth rectangles and (b) detection rectangles	53
2.22	XML Code Format (input)	56
2.23	XML Code Format (output)	56
3.1	Flowchart of the Research	62
4.1	The perspective box concept based on three dimensions	72
4.2	The perspective box with segmentation components and elements	73
4.3	Procedures for figure-ground object detection framework	74

4.4	Pseudo-code for figure-ground object detection	75
4.5	Edge detection using phase congruency	76
4.6	Lines clustering for experiment purposes	78
4.7	Pseudo-code for vanishing point detection	79
4.8	Vanishing points in an image dataset	80
4.9	Vanishing point candidate estimation using a standard Hough Transform (SHT)	81
4.10	Vanishing points detection using SHT in MATLAB	82
4.11	Figure-ground object candidates in perspective box concept	83
4.12	Pseudo-code for statistical region merging (SRM) (Battiato, S., 2006)	84
4.13	Image region averaged with random colors in SRM	85
4.14	Image regions averaged in original colors in SRM	85
4.15	A segmentation box coordinate for *p1, p2, p3* and *p4*	86
4.16	Pseudo-code for initial region mask declaration	87
4.17	A level set initial mask based on the vanishing point that is set as the attention center (Vc)	89
4.18	Perspective box concepts with initial R_{inside} and $R_{outside}$ regions for the segmentation process	89
4.19	Pseudo-code of the proposed segmentation method	90
4.20	Pseudo-code for bounding box detection process	92
4.21	Figure-ground object based on rectangle bounding boxes detection	93
4.22	A perspective box concept prototype system	94
5.1	Example of images from PASCAL VOC2010 dataset (M. Everingham et al., 2010)	97
5.2	Testing Dataset with object segmentation and detection	100

5.3	(a) An original image with (b) line detection using the MATLAB imfilter function	101
5.4	Vanishing point detection using phase congruency	102
5.5	Edge detection comparisons	102
5.6	Figure-ground object with object background (white = a figure-ground object candidate, black = background)	104
5.7	(a) (b) (c) Region growing segmentation	105
5.8	(a) (b) (c) An active contour segmentation using vanishing point as initial level set box coordinate	105
5.9	(a) (b) (c) An active contour segmentation with perspective box concept	106
5.10	Experiment on perspective box concept using testing dataset	108
5.11	An active contour segmentation for image benchmarking result	115
5.12	An active contour with SRM segmentation for image benchmarking result	116
5.13	A perspective box concept segmentation for image benchmarking result	118
5.14	Performance Measures using Recall Vs. Precision Graph	119

LIST OF ABBREVIATIONS

2D	-	Two Dimensional
3D	-	Three Dimensional
AC	-	Active Contour
BU	-	Bottom Up
DetEval	-	Detection Evaluation Software
GT	-	Ground Truth
LMS	-	Long Medium Shot
LS	-	Long Shot
MATLAB	-	MATrix LABoratory
MRF	-	Markov Random Field
PASCAL	-	Pattern Analysis, Statistical Modelling and Computational Learning
PB	-	Perspective Box Concept
PDE	-	Partial Differential Equation
RG	-	Region Growing
ROI	-	Region of Interest
SRM	-	Statistical Region Merging
SVM	-	Support Vector Machine
TD	-	Top Down
VOC	-	Visual Object Recognition Challenge
VOCUS	-	Visual Object Detection System with a Computational Attention System

WTA	-	Winner Take All
WWW	-	World Wide Web
XML	-	Extensible Markup Language

LIST OF SYMBOLS

$(\boldsymbol{S_I})$	-	The set of nearby pixel pairs
(x)	-	The signal location in the Fourier transform equation
(x; y)	-	The continuous set of all possible lines candidate coordinate in Gaussian sphere
$\overline{R}_a$	-	An average color an in SRM region
$\sum$	-	The value of sum
$B^{'}$	-	The background image
I	-	The set of figure-ground object region in image scene
K_{xi}	-	The width value in vector
K_{yi}	-	The height value in vector
O_x	-	The main figure-ground object
$R_{\|I\|}$	-	The set of regions in image scene
R_{inside}	-	The mean value of everything inside the image region
$R_{out\ side}$	-	The mean value of everything outside the image region
R_x	-	The image region that contained a figure-ground object and a background image.
V_c	-	The attention point in perspective box concept
max	-	The maximum values
[i]	-	The number of pixels
|E(x)|	-	The complex vectors of Local Energy
$\emptyset(\mathrm{x}, \mathrm{t})$	-	The level set function

$\in$	-	The element of values
π	-	Pi
$\cup$	-	Union
$\lfloor \rfloor$	-	The quantity that is equal to itself
a	-	The constant values in PDE equation,vector mask values
A_n	-	Amplitude
b	-	The constant values in PDE equation, vector mask values
B_i	-	The blue color value in SRM
C	-	The evolving curve in image region
C_1	-	The averages of inside curve in image region
C_2	-	The averages of outside curve in image region
C_i	-	The ceiling coordinate in perspective box concept
D_o	-	The detected rectangle box in perspective box concept
f	-	The function that is differentiated
g	-	The image function
G_i	-	The green color value in SRM
G_o	-	The ground truth rectangle box in perspective box concept
G_x	-	The vertical kernel in edge detection
G_y	-	The horizontal kernel in edge detection
H	-	The value of height
k	-	The constant value for equation
L	-	The value of length
noD	-	The number of detected rectangles in image benchmarking process
noGT	-	the number of ground truth rectangles in image benchmarking process
Oi	-	The region boundaries
$Ø_n$	-	The phase angle in equation
P	-	The precision values
p_0	-	The pixels value belong to image regions

Q	-	The set that functions independently with random values
r	-	The recall values
R_i	-	The red color value in SRM
T	-	The value of energy
U_0	-	The constant value for processing image
v	-	The coarseness in segmentation process
$\theta(theta)$	-	The angle of orientation by the equation
V_p	-	The vanishing point value
V_x	-	The vanishing point value for *x* coordinate
V_y	-	The vanishing point value for *y* coordinate
W	-	The width value
W(x)	-	The weight factor of frequency spread
W_L	-	left wallpaper coordinates in perspective box concept
W_R	-	right wallpaper coordinates in perspective box concept
ɾ	-	A discontinuous set in the image domain
μ	-	The factors that manage the quality
μ	-	The piecewise that smooth image with sharp edges
Ωi	-	The region in the image that represents an object
R	-	An image region
fo	-	The detected rectangles for figure-ground objects
$p1$	-	A point corresponds to the upper left in perspective box concept
$p2$	-	A point corresponds to the lower left in perspective box concept
$p3$	-	A point corresponding to the lower right in perspective box concept
$p4$	-	A point corresponding to the upper right in perspective box concept
ε	-	A constant value in equation to avoid division by zero
$\rho(rho)$	-	The distance from the origin to the line along a vector

LIST OF APPENDICES

APPENDIX NO.	TITLE	PAGE
A	Evaluation Dataset	136
B	Perspective Box Concept Detection Result	143
C	Ground Truth Dataset (XML input)	150
D	Image Benchmarking Result (XML output)	158
E	Performance Measures Result	161
F	List of Publications	165

CHAPTER 1

INTRODUCTION

1.1 Introduction

Figure-ground object detection is an important and challenging vision task. Figure-ground object and background segmentation are widely employed in many computer vision tasks, such as object detection, identification, image editing, graphics rendering and image retrieval. However it is still an open problem due to the complexity of object scene and images. When begin analyzing this complexity, the problem can be divided into several subtasks. First of all, the visual system has to recognize the image scene to find its visual content and image understanding before the entire figure-ground-ground object takes an account. Then the second step is to search for the dominant object based on the visual attention point, perception and literacy of the image scene in order for the system to interact and react adequately to the given detection tasks. Dominance in an image scene is the relationship between multiple variant regions in which one region grabs the attention better than others in influencing some visual traits. The dominance term describes how the mind organizes visual data around the stronger visual content in the gestalts to clarify the figure-ground object features.

Image features are inequal in their differential relevance in computing the similarities among images. Different persons or the same person from different perspectives may view the same image but with different attention, perception and understanding. Each class of figure-ground objects and background images that have the same color, shape and textures will obtain different attention by human observers and this problem must be addressed when attempting to determine the region of interest (ROI) in figure-ground object detection while the perspective box concept tries to adapt these scenarios into segmentation and detection problems. There are several fundamental problems associated with image segmentation between the figure-ground object and background image with objects that do not have dominant features based on visual information (Steven Lehar, 2003 and Chunhui Gu et. al., 2009).

The existing image segmentation and classification methods usually concentrate on color, textures; shape and other variation algorithm to cluster the visual information without knowledge of visual attention and image literacy (Frintrop S., 2006). These visual information extractions assist the semantic understanding of an image. It also provides improved browsing and retrieval facilities to the recognition and identification systems to analyze figure-ground object especially in the new concept on perspective segmentation for natural images (Heitz G. and Koller D., 2008). The anticipated results of this research are to expand our knowledge of the region based segmentation process to classify figure-ground object and background image on the perspective box concept. The perspective box concept is a visual attention system based on image segmentation and bounding box figure-ground object detection. The image segmentation starts with the vanishing point detection on image scene to detect attention point on image scene that 3D reconstruction structure of the scene. This step also involves the pre-processing step to detect visual component on image scenes and then all the lines are clustered and edged to detect the intersection point between the image plane and the detected image region using active contour segmentation via the vanishing point to detect figure-ground object.

Most successful figure-ground object detection relies on binary classification, deciding only if an object is present or not in actual image scene with the object location. To perform localization, perspective box adopts a sliding window approach on bounding box to detect the figure-ground object. The bounding box detection is an important task for the automatic understanding of images scene and to separate figure-ground objects from the background. The detection analyzes the spatial relations of different figure-ground objects in an image scene to other detected objects. The bounding box is the tightest rectangle which includes the image region and described by the x-y coordinates of the lower-left corner of the image region, followed by the x-y coordinates of the upper-right corner of the region. Since the bounding box in perspective box optimizes the same quality procedure as the benchmarking process on VOC 2010 dataset that is based on bounding box detection, the same performance and detection scores could have been achieved to evaluate the result based on precision, recall and F-measure. Then from the evaluation and benchmarking process, the graph is generated to show the effectiveness this perspective box in solving the segmentation issues on attention and image understanding of the scene segmentation process.

1.2 Problem Background

Currently, many photographers snap photos with digital cameras and upload their precious moments and art work to their own blogs or websites. The quick changing digital camera technology and how this technology helps photographers to capture emotion in photos has become a new scenario as well as becoming a new complexity in visual image retrieval because of the limitation in intelligent system to determine what the photo is all about and the objects that are residing inside the image scene. In the context of visual literacy, there are always different meanings in a single photo, but we may be seeing the similarities only (Frintrop S., 2006). A figure-ground object can be recognized because the brain can process the visual

content and then compares it with visual projections through the eye to objects that are previously stored in our brains (Lamme V. A. F., 1995).

Everything on the eyes and vision can be recognized as always following the rules of light reflection and this idea will help us to comprehend what and how the photographer sees things before he/she snaps the perfect photo. Learning from these ideas, an image segmentation concept can be redesigned to determine object detection with visual attention problem and deal with the image segmentation problem, see Figure 1.1. The current figure-ground and background segmentation procedure typically fails when dealing with images as in Figure 1.1, especially when the detection system needs to extract and cluster the image with the similar features in multiple object segmentation.

Figure 1.1: Images without dominant themes for figure-ground object and background image

1.3 Problem Statement

This study intends to come up with a concept to provide insights into solving the image segmentation process and figure-ground object detection. The research question is:

How can one produce reliable image segmentation processes that are able to be used for figure-ground object detection using concept of visual literacy, attention and human perception?

In order to answer the main issue raised here, the following issues need to be addressed and discussed:

i. How have the previous works solved the problem of image segmentation and figure-ground object detection?
ii. It is well known that figure-ground object detection consists of partitioning an image into significant object regions with detection problem of the undetected figure-ground object, such as non segment image region with image background overlay. How can this be overcome?
iii. What is the problem of the existing object segmentation and detection methods like selecting the best figure-ground object in image scene? How can this be countered?
iv. What is the most suitable image segmentation and object detection for visual system?
v. How can one perform figure-ground object detection on visual literacy, attention and perception without relying on the intelligent system?
vi. How can one test the bias of image benchmarking process and the performance of figure-ground object detection?

In visual literacy and perception, wherever an interesting point with visual content draws our inspiration to an image scene, it is often not just the particular

element that sparks our brain visual cortex; it is usually more of the totality of the visual element and its surrounding environment. Visual elements with perceptual organization have visual entities on which the detection processes can operate. We then have options with regard to what these entities should be: points, curves or regions (Chunhui Gu et. al., 2009). An individual visual element and the whole surrounding objects are important both separately and together, and are essential to the understanding of how gestalts influence our design choices. The gestaltism is the psychology term that Max Wertheimer (Lehar, S. ,2003) introduced for the essence or shape of an entity's complete form within the context of a visual component, and these gestalts can be classified as proximity, similarity, figure-ground object, symmetry, common fate and closure.

Figure 1.2 A dog picture demonstrates the principle of dominance in visual perception (Lehar, S., 2003)

The well-known figure description as shown in Figure 1.2 illustrates a Dalmatian dog sniffing under a stand of trees. At the first sight, the picture only shows incongruent white and black blobs and then the observer will make a demand for the other component parts, such as the dog's head, feet and then he/she will see the entire picture that contains a dog and a tree. The interaction and attention on

dominant region is gained before the observer grasps the visual details in image scene and figure-ground object detection.

1.4 Hypothesis

In this research, the proposed perspective box concept is employed to improve the figure-ground object detection in terms of the segmentation performance and detection effectiveness. Therefore, several hypotheses have been made:

i. The preprocessing method applied will increase the effectiveness of figure-ground object detection.
ii. The perspective box can improve the segmentation process without using the intelligent system.
iii. By using the perspective box, statistical region merging region and active contour segmentation method can solve the visual attention, perception and literacy problems.
iv. The use of benchmarking procedures in detection process will help to evaluate the results.

1.5 Aim

The aim of this research is to improve the segmentation and detection process based on the perspective box concept by using the region based with figure-ground

object detection and the image benchmarking approach. This thesis reveals how the perspective box concept is applicable to different figure-ground detection and retrieval scenarios without relying on the intelligent system or manual interaction and intervention.

1.6 Objectives of Study

The following are the objectives of this research:

i. To design an improved segmentation process especially in the pre-processing stage.
ii. To develop an effective figure-ground object detection system based on the perspective box concept using hybrid technique.
iii. To evaluate the segmentation and detection systems with the comparison results on the bounding box detection and benchmarking prosedures using ground truth (human segmentation) dataset and detection result dataset.

1.7 Scopes of Study

The scopes of this study are defined as follows:

i. This research is focusing on testing and designing the concept, and procedures that can be adopted into image segmentation and figure-ground detection system using the MATLAB programming language.

ii. A hundreds of images are used as testing and evaluation datasets by utilizing image database provided by PASCAL VOC 2010 which contains multiple figure-ground objects that rely in image scene act as ground truth dataset.

iii. The evaluation process begins with a comparison study on image benchmarking process between the ground truth dataset and bounding box detection result and then the performance result is simplified and visualized using the scatter chart on recall, precision and F-measure that elaborates the figure-ground object detection result for each image scene.

1.8 Significance of the Research

The significances of this research are as follows:

i. It enhances the research area of segmentation process with region based segmentation on image scene such as figure-ground object detection system.

ii. It demonstrates the importance of visual perception, attention and capability to understand the image scene in constructing and designing figure-ground object segmentation and detection system.

1.9 Contribution of the Work

This research contributes better ideas on improving the segmentation process especially for region based segmentation and demonstrates the result performance in image pre-processing step in order to complete certain levels of achievement. The main work is to design better and effective detection system based on the perspective box concept that can applied on statistical region merging (SRM), vanishing point detection and active contour segmentation.

1.10 Research Plan

This research is carried out within six semesters. The first part of the project focuses on understanding the general views of image processing tool in MATLAB programming. Then it highlights on understanding the recent algorithms, concepts and procedures that have been applied by other researchers. Most of the time spent in the first and second semesters were used to explore and gather relevant information from textbooks and published journals. Figure-ground object detection methods are vital in order to comprehend different methods that can be used in solving similar problems. The research requires a better understanding of image segmentation and detection process to improve the visual system performance.

The second part of the project involves implementing figure-ground object and background segmentation of images based on the perspective box concept. This technique will be used to learn how we understand the image perspective based on the Hough transform geometrical model. The experience includes classification, comparison and benchmarking processes that are conducted to monitor the

performance of detection process and figure-ground object clustering process. Finally, the report including the experimental result and conclusion is prepared.

1.11 Summary and Thesis Organization

The introduction to detection system with image segmentation includes the problems, hypothesis, scopes, contribution, aim, objectives and research plan.In order to enhance the method of figure-ground object detection using the perspective box, several chapters are constructed to arrange the contents of the thesis. The contents of each chapter are as follows:

i. Chapter 1 presents a general introduction of the thesis, which includes introduction, problem background, problem statement, aim, research objectives, research scope, significance of the research, contribution of the work, result, project plan as well as an outline and thesis organization.
ii. Chapter 2 offers a review of the relevant and related literature on visual perception, visual attention, image segmentation procedures and figure-ground image detection. It also provides an overview of the vanishing point detection in the perspective box concept. This chapter also clarifies how these image preprocessing approaches are applied in our project with a brief description of the segmentation and detection procedures.
iii. Chapter 3 discusses the methodology used in this research, which consists of image pre-processing, line clustering, vanishing point detection, the region based segmentation process, the figure-ground object detection procedure on bounding box detection and the image benchmarking process.
iv. Chapter 4 elaborates on the modeling process on the perspective box concept that starts from image segmentation process which is applied into detection procedure.

v. Chapter 5 presents the experiment and its results together with other experiments that have been conducted based on the procedure and perspective box concept that is based on the image benchmarking process on precision, recall and F-measures.

vi. Chapter 6 signifies the conclusions as well as suggestions for future research.