Modeling the Perception of Visual Complexity in Texture Images and Painting Images

(テクスチャー画像および絵画に対する複雑さの知覚モデルの構築)

by

Xiaoying GUO

Graduate School of Engineering

Hiroshima University

A thesis submitted for the degree of

Doctor of Engineering

September 2013

To my family, for their love

Abstract

This thesis devotes to building a relationship between human visual complexity perception and objective image features.

Visual complexity of images is a worth-well investigation. It has a wide range of applications from computer science (emotion semantic image retrieval, digital watermarking and image analysis, etc.) to arts (the design of product surface, wallpapers, painting appreciation and selection, etc.). Various methods for computing image complexity have been developed depending on different applications, such as information theory, fractal dimension, quad tree method, etc. However, these measures are not sufficiently related to human visual perception of complexity. Intuitively, visual complexity is influenced by various factors perceived by humans from the image, not directly related to simple objective measures like distribution of spatial frequencies.

In this thesis, we have investigated the perception of visual complexity in texture images and painting images from the point of human visual perception. The motivation of this research was to build a relationship between human "Kansei" (visual complexity) and computable image features. "Kansei" is a Japanese word with meaning of emotion and sensibility. Kansei Engineering is an approach to connect human sensibility to computing application. The main achievements proposed in this thesis are listed as below.

1) A new method of estimating visual complexity of texture images has been proposed.

A texture has certain features that can affect viewers' complexity perception. We firstly conducted a set of psychophysical experiments to identify these features. By the experiments, we have identified that five important features affect human visual complexity of textures, namely, regularity, roughness, directionality, density, and understandability. Visual complexity is a function of not only each individual characteristic but also of interactions between them. Then a set of methods was designed for objectively measuring the features of regularity, roughness, directionality, and density.

We proposed in particular a new method for estimating understandability of a texture by naming the textures. We discovered that understandability is affected by two factors of a texture: the maximum number of similar names assigned to a specific type and the total number of types.

Multiple linear regression was performed as a mapping function to bridge the relationship between visual complexity perception and five texture features. A series of statistical analyses was performed to test the fitness and correlation between prediction from the proposed model and subjective complexity evaluation given by humans. Compared with the conventional measures based on information theory and fuzzy pattern, the proposed method considers human visual perception, and it predicts the visual complexity of a texture corresponding with the subjective visual impression.

2) A novel framework to assess visual complexity of painting images has been developed.

We proposed a framework to assess visual complexity of paintings. This framework provides a machine learning scheme for investigating the relationship between human visual complexity perception and low-level image features. Since the global and local characteristics of paintings affect human's holistic

impression and detail perception, we studied theoretical and empirical concepts from psychology and art theory to extract the features that represent the global and local characteristics of paintings. Inspiration for these features was from a questionnaire survey we conducted to identify the factors that affect human's complexity assessments of paintings. Then we conducted feature selection, by which we looked into the role that each image feature plays in assessing visual complexity, and then obtained the feature combination that yields the best performance. From a computational perspective, we need to obtain a prediction of visual complexity from all input image features (some responses on complexity that corresponding with human assessment). But it is difficult to predict a specific value for visual complexity. Instead we introduced a machine learning method to classify the visual complexity into three classes: low complexity, middle complexity, and high complexity. All features were combined by a support vector machine for classification.

Experimental results indicated that the proposed work can predict the visual complexity perception of paintings with the accuracy of 88.13%, which is highly close to the assessments of visual complexity given by humans. Compared with the conventional measure of complexity, our approach considers human visual perception and performs more efficiently in assessing visual complexity of painting images. Furthermore, we applied the proposed method to architecture images. The experimental results showed the validity of our method in architecture images.

Keywords: visual complexity, texture images, painting images, kansei engineering, affective engineering, image features, support vector machine, color complexity, understandability.

Acknowledgements

I would like to express my deep appreciation and gratefulness to those who help me achieve this work.

First of all, my sincere gratitude to ***Professor Akira Asano***, who's my first supervisor in Hiroshima University. His attitude to research encourage me to make progress step by step. Although he moved to Kansai University, he still keep discussion with me monthly. Without these discussions, this work could not be a success.

I am very grateful to ***Professor Takio Kurita***, who is my supervisor in the later two years of Ph.D study. I admire his rich knowledge, unique thinking and kindly personality. I would like to thank him for providing me a position of research assistant, which practices me in the ability of both research and communication.

I sincerely appreciate ***Associate Professor Chie Muraki Asano*** at Yasuda Women's University. Her patient guidance and support always encourage me in researches. Her useful comments and encouragements make me complete this interesting and challenge topic.

I am very grateful to the members of the dissertation committees, ***Professor Koichi Harada***, ***Professor Kazufumi Kaneda*** and ***Associate Professor Toru Tamaki***, for their constructive feedbacks and outstanding suggestions.

I would like to express my gratitude to ***China Scholarship Council*** (CSC) for supporting me financial assistance in Japan, which makes me fully concentrate on my Ph.D study.

There are many friends and lab members who helped me in my life and study in Japan. Firstly, I would sincerely thank ***Liang Li***, who is the first person I met in Japan. He guided me a lot in life, and taught me how to do research in the beginning of my Ph.D study. And I would thank ***Kavitha*** for her useful suggestions when I wrote this thesis. I am also grateful to my lab members and my friends, ***Lei yang, Zhibin Zhang, Zhouxin Yang, Pengliang Wu, Zhiming Gu, Yi Yuan, Yilin Du, Qianru Xu, Xuezhen Li***, whom I shared my fruitful discussions and joyful moments both inside and outside the Lab.

Forever, my deepest love and appreciation to my parents. Their endless supports, encourages and trust always push me move ahead. In all, illimitably grateful to my parents.

Last thanks to my husband ***Feng*** for his encouragements, patience, and help in my master study and Ph.D study periods.

Last, but nor least, I would like to acknowledge all who directly and indirectly help me in completion of this work.

Xiaoying Guo

Hiroshima University

September 2013

Contents

Abstract i

Acknowledgements v

List of Figures xiii

List of Tables xvii

1 Introduction 1

1.1 Background and Motivation 1

1.2 Objectives 3

1.3 Significance and Possible Applications 4

1.4 Structure of the Dissertation 7

I Basic Theories of Visual Perception and Complexity 9

2 Visual Perception and Complexity 11

2.1 Visual Perception 11

2.1.1 Gestalt theory of visual perception 11

2.1.1.1 Gestalt principles of perception 12

2.2 Visual Complexity and Related Works 14

2.2.1 Definitions of Complexity 14

2.2.2 Visual complexity of images 15
2.2.3 Visual complexity and aesthetic 16
2.2.4 Visual complexity and websites 17
2.2.5 Visual complexity and watermarking 18
2.2.6 Visual complexity and information theory 19
2.3 Measuring Visual Complexity 19
2.3.1 Information theory 20
2.3.1.1 Entropy and fuzzy approach 20
2.3.1.2 Image composition complexity 20
2.3.1.3 Compressed file size of image 21
2.3.2 Fractal dimension 22
2.3.3 Quad tree method 22
2.3.4 Region of Interest method 23
2.3.5 Visual attention method 23
2.4 Affective Engineering and Kansei Engineering 24

II Modeling the Perception of Visual Complexity in Texture Images 25

3 Visual Complexity of Texture Images 27
3.1 Introduction 27
3.1.1 Texture perception 27
3.1.2 Visual complexity of textures 28
3.2 Texture Features that Affect Visual Complexity 29
3.2.1 Experiments and results 29
3.2.1.1 Experiment set-ups 29
3.2.1.2 Semantic differential method 30

3.2.1.3 Experiment 1: visual complexity assessment 31
3.2.1.4 Experiment 2: paired comparison evaluation 32
3.2.2 Results and Statistical analysis 33
3.2.2.1 Correlation analysis between texture characteristics and complexity 33
3.2.2.2 The importance of perceptual characteristics for visual complexity 40
3.2.2.3 Multidimensional scaling map of texture characteristics and complexity 41
3.3 Summary 42

4 Measures of Texture Features **43**
4.1 Introduction 43
4.2 Feature Measures 44
4.2.1 Regularity 44
4.2.1.1 Autocorrelation function 44
4.2.1.2 Regularity extracted from the ACF 45
4.2.1.3 Results 45
4.2.2 Roughness 46
4.2.2.1 Roughness calculation 47
4.2.2.2 Results 48
4.2.3 Directionality 49
4.2.3.1 Directionality calculation based on texture edges 49
4.2.3.2 Results 50
4.2.4 Density 50
4.2.4.1 Density calculation 51
4.2.4.2 Results 52
4.2.5 Understandability 52

4.2.5.1 Experiment: naming the textures 53
4.2.5.2 Results 57
4.3 Discussion 58
4.4 Summary 59

5 Estimation of Visual Complexity Based on Texture Features 61
5.1 Introduction 61
5.2 Modeling the Visual Complexity of Textures 63
5.2.1 Simple Linear Regression 64
5.2.2 Multiple Linear Regression 64
5.2.3 Least Square method 64
5.2.4 Subjective Visual Complexity of Textures 65
5.2.4.1 Experiment 65
5.2.4.2 Consistency 67
5.2.4.3 Results 67
5.2.5 Modeling visual complexity by MLR 67
5.2.6 Statistical test of the proposed model 69
5.2.7 Validation test of the prediction 71
5.2.8 Comparison with other methods of measuring complexity 73
5.2.9 Discussions 74
5.3 Summary 76

III Modeling the Perception of Visual Complexity in Painting Images 77

6 Objectively Estimating Visual Complexity of Painting Images 79
6.1 Introduction 79
6.1.1 Hypothesis 80

6.1.2 Our work . . . 81
6.2 Subjective Complexity Assessment . . . 82
6.2.1 Experiments . . . 83
6.2.1.1 Respondents . . . 83
6.2.1.2 Apparatus and stimuli . . . 83
6.2.1.3 Procedure . . . 83
6.2.2 Results . . . 84
6.3 Feature Extractions . . . 85
6.3.1 Global features . . . 87
6.3.1.1 Color features . . . 88
6.3.1.2 Content features . . . 92
6.3.2 Local features . . . 94
6.3.2.1 Image segmentation . . . 94
6.3.2.2 Color features . . . 98
6.3.2.3 Distributions features . . . 99
6.4 Objective Measure of Complexity . . . 102
6.4.1 Support Vector Machine . . . 102
6.4.1.1 Linear SVM classifier . . . 102
6.4.1.2 Nonlinear SVM classifier . . . 103
6.4.1.3 Solution of SVM formulation . . . 105
6.4.1.4 SVM kernel functions . . . 106
6.4.1.5 Multiclass SVM . . . 106
6.4.2 Complexity classification using SVM . . . 107
6.4.2.1 Selecting kernel functions . . . 107
6.4.2.2 Scaling data . . . 108
6.4.2.3 Parameters selection . . . 108
6.4.3 Feature selection . . . 109

6.4.4 Classification performance . . . 109
6.4.5 Comparisons with other measures . . . 110
6.4.6 Validity of our method in other images . . . 110
6.4.7 Application: local complexity of images . . . 111
6.5 Discussion . . . 111
6.6 Summary . . . 114

7 Conclusions and Future Work **121**
7.1 Summary of Achievements . . . 121
7.2 Limitations and Future work . . . 125

Appendix **127**

Bibliography **129**

Publication List **141**

List of Figures

1.1 An example host image: it is much difficult to insert watermark information in yellow rectangle region than in red rectangle region. 6
1.2 Scope of the research and structure of chapters in the dissertation. 8

2.1 Examples of Gestalt principle: similarity. 12
2.2 Example of Gestalt principle: proximity. 13
2.3 Example of Gestalt principle: continuity. 13
2.4 Example of Gestalt principle: closure. 13
2.5 Example of Gestalt principle: area. 13
2.6 Example of Gestalt principle: symmetry. 13
2.7 Berlyne theory [17]: relationship between complexity and pleasure. 17

3.1 Twenty texture images used in the experiments. 30
3.2 The screen used in the experiment and the distance from the respondents to the screen. 30
3.3 The 7-point rating scale used in the first experiment. 33
3.4 Relationships between regularity of textures and visual complexity. 37
3.5 Relationships between directionality of textures and visual complexity. . . . 38
3.6 Relationships between roughness of textures and visual complexity. 38
3.7 Relationships between understandability of textures and visual complexity. 39
3.8 Multidimensional Scaling map. 41

4.1 One-dimensional ACF along the x and y directional of texture d47 in Fig. 3.1. 46
4.2 Relationship between the regularity calculated by the ACF and that evaluated subjectively. 47
4.3 A smooth texture and a rough texture. 47
4.4 Correlation between the calculated and evaluated (subjective) values of roughness. 49
4.5 Relationship between the calculated and evaluated (subjective) values of directionality. 51
4.6 Relationship between the calculated edge density and the evaluated visual density. 52
4.7 The webpage used in the experiment. When the stop button is pressed, the viewing time will show in the right box. 54
4.8 Frequencies of choices A, B and C in the questionnaire. 56
4.9 Two examples of naming textures. The center images are the experimental images $d26$ and $d13$, and the surrounding names are given by the respondents. 57

5.1 Mapping structure between low-level features and high-level feeling. 63
5.2 50 texture images used in the experiment. 66
5.3 Subjective complexity and calculated complexity of 40 texture samples. . . 70
5.4 Correlation between subjective complexity and calculated complexity of modeling samples. 70
5.5 Subjective complexity and predicted complexity of testing samples. 71
5.6 Correlation between subjective complexity and predicted complexity of testing samples. 73

6.1 Framework of the proposed method. 82

6.3 The frequencies of the options (from A to H) of two questions mentioned by 30 respondents in Part II (questionnaire part) in the experiment. 85
6.4 The framework of composing image features for paining complexity assessment. 87
6.5 The model of Gaussian pyramid. 90
6.6 Representation of RAG: spatial view of the image. (a) Segmented regions. (b) Region Adjacency Graph. 96
6.7 (a) One example painting image in our dataset. (b)Initial watershed segmentation. (c) Initial mean shift segmentation. (d) The finial segmentation after region merging. 97
6.8 A linear SVM classification example with a hyperplane that maximizes the separating margin between two classes (indicated by the green points and red triangles). Support vectors are the elements that lie on the boundary hyperplane of the two classes. 103
6.9 A linear SVM classifier is not complex enough to classify the samples into two classes. 104
6.10 Mapping the input pattern into a higher dimensional space where it exhibits linear hyperplane. 105
6.2 Examples that are labeled as LC, MC, and HC. The painting images on the first row are labeled as low-complex, the painting images on the second row are labeled as mid-complex, and the painting images on the bottom row are labeled as high-complex. 116
6.11 Classification accuracies of 24 feature combinations. 117
6.12 The label under the images are the prediction from our method. The red frames indicate the wrong classifications of image's complexity. 118

6.13 Architecture images that are labeled as LC, MC, and HC. The images on the first row are labeled as low-complex, the images on the second row are labeled as mid-complex, and the images on the bottom row are labeled as high-complex. 119
6.14 Local complexity estimation of two images that are divided into 9 regions and 16 regions respectively. 120

List of Tables

3.1 Summary of verbal descriptions from 30 respondents who participated in Experiment 1. 32
3.2 Average scores of subjective perception complexity, regularity, density, directionality , roughness and understandability. 34
3.3 Ranking of average scores for 20 texture images in ascending order. 35
3.4 Correlation matrix of perceptual characteristics of textures and visual complexity. 36
3.5 Factor loadings without rotation. 40

4.1 Experimental data from the experiment of naming the textures. 55
4.2 Correlation coefficients between NT, MNDT, RT and understandability. . . 57

5.1 Average scores of complexity of 50 texture samples. 68
5.2 Results of Goodness of fit test and F test. 72
5.3 Comparisons between the proposed method and other related methods. . . . 73

6.1 Results of survey in the experiment. 85
6.2 All features extracted from the painting images (including global features and local features). 101
6.3 Performance comparisons of various SVM kernel. 107
6.4 Confusion matrix of predict classification. 110

6.5 Correlations between subjective complexity and objective measures of complexity. 110
6.6 Classification accuracies of the proposed method used for painting images and architecture images. 111

Chapter 1

Introduction

1.1 Background and Motivation

The rapid development of digital technologies and internet make people easy to see and feel the world. At the same time, it raises a group of interesting topic like allowing a computer to sense the world as that sensed by humans. This keeps the same objective with affective engineering.

Affective engineering (also called kansei engineering in Japan) is an approach to connect human sensibility to computing application. "Kansei" is a Japanese word; that can be explained as the meaning of feeling, impression, and emotion and so on. Kansei research is an attempt to enumerate and quantify those subjective feelings, such as visual aesthetic perception, pleasing feeling and complexity perception and so on. A picture has certain factors that can affect viewers' impressions. These factors are called kansei factors [1]. Kansei Engineering is a methodology that encompasses human-oriented emotions with low-level computable visual features. The challenge in kansei research is to identify the kansei factors, and associate them with the feeling brought forth in the viewer.

In this thesis, we focus on the perception of visual complexity in texture images and painting images. The motivation of this research is to build the gap between human visual complexity and computable image features. By employing the methods of kansei engineering, we respectively propose a new model of estimating visual complexity in texture images and a novel framework of evaluating visual complexity in painting images. In all, we achieve the purpose that endowing the computer an ability of estimating image complexity of texture images and painting images as that perceived by humans.

The notion of image complexity is prevalent and essential in the fields of computer science, neurology and psychology. It is widely used by computer engineers to construct information systems and tools for analysis, estimation, visualization and recognition of images. It is also concerned by neuroscientists and psychologists, who are interested in the mechanism of object perception and the process of learning and memory [2].

As is known to us, visual complexity is a very basic and important aspect of visual aesthetic. Usually, images are visually judged as simplicity / complexity. If we are able to measure and combine the features of an image to match complexity as judged by humans, then it is useful in real application such as image compression, watermarking, and it would be helpful for analyzing the effect of visual complexity on human's aesthetic judgments [3]. Visual complexity will also help to inspire practitioners to make many more exciting projects and help the public better understand the importance and beauty of visualizations of complexity [4].

Complexity has been defined by various ways depending on different applications. But there is no uniform definition. Human beings are able to form a coherent perception of complexity and identify a complex image or object at a glance [5]. Although there is no agreed understanding of image complexity, there is a global agreement in classifying images by complexity for a majority of people.

Various measures of image complexity have been proposed [6, 7, 8, 9, 10, 11, 12, 13, 14], including information theory, pattern measure, fractal dimension, quad tree method, region of interest method and so on. These measures provide the computational methods of measuring complexity. However, intuitively, visual complexity is usually affected by visual content (various image features) of images, such as the color features, distributions of objects, understanding and so on. It is not directly related to simple objective measures like distribution of spatial frequencies. Hence, the previous measures of complexity are not sufficiently considering human visual perception. In this work, we attempt to investigate visual complexity from the point of human visual perception to make a computer sense the visual complexity of images as that sensed by a majority of people.

In this thesis, we mainly consider two types of digital images: textures and paintings. Textures are around our lives. It is an important cue of surface property. Visual perception of texture has a wide range of applications from computer science to arts. Many researches have been done regarding investigating visual perception of textures, such as, coarseness, contrast, regularity and line-likeness, etc. However, with respect to the high-level kansei feeling, visual complexity of textures, seldom studies have been conducted. Therefore, we firstly investigate the visual complexity perception of textures. Moreover, with the rapid development of digital technologies and internet, people have more opportunities to appreciate the paintings without going to museums. More and more users select preferable paintings from the internet. If they consider selecting images only by visual feeling (e.g. aesthetic) instead of specific keywords (e.g. flowers), visual complexity has some information of composing the feeling [15, 16]. Hence, presenting an objective index of complexity that fits human feeling to the users is useful.

1.2 Objectives

The overall objective of this study is to map the relationship between visual complexity perception of human (subjective feeling) and the features extracted from the images (objective features). Addressing this problem requires solving a number of sub-problems including subjective assessment of complexity, measures of image features and the mapping functions between visual complexity and image features, etc. A wide range of knowledge from psychology to computer science is also required to better understand the problems. Specially, we investigate the visual complexity both in texture images and painting images. The detail objectives in this thesis are:

- To identify the features that affect human visual complexity perception of images (including textures and paintings) from the point of view of human visual perception.
- To design a group of new feature extraction methods for better matching the human visual perception of image features.
- To develop the new model of assessing image complexity based on image features to correspond well with the subjective visual perception of complexity.

1.3 Significance and Possible Applications

Estimating the image complexity is useful for a wide range of applications from psychology to computer science. Here, we will briefly list several important applications of image complexity.

1) **Visual perception mechanism**

Learning the visual complexity perception is helpful for both psychologist and artist. For psychologists who are interested in the visual perception mechanism, research on complexity will assist them to study the visual perception principles and processes, and then help them to analyze the possible reasons that influence the complexity perception. Moreover, for artists, visual complexity is an additional way to analyze the visual aesthetic. Visual complexity is a basic aspect of aesthetics perception. From Berlyne' theory of aesthetic perception [17], it was identified that viewer's aesthetic perception of an object is connected to the complexity of this object. If we can measure and combine computational features of an image to match visual feeling as perceived by humans, then it would be the first step in being able to predict the effect of human feeling of an image used in the art design and other real applications.

2) **Image retrieval**

Image complexity can be used in image retrieval. In [6], the complexity of image is regarded as a universal property that related to similarity. The authors presented a new method of measuring similarity of the natural images based on the joint complexity of the images, and applied it to content-based image retrieval.

It was pointed out that there is a trend towards dealing with a higher level of multimedia semantics: cognitive level and affective level [18]. Image retrieve is now addressed by using keywords (e.g. flowers, sea et al.), textural descriptions and similarity criteria (like Google Search by image). It is very often that humans always select a special image from a pool of images that contain the same objective criteria (for example the same subject). However, to account for the way humans look for the information, there is an additional way to retrieve the image, that is considering the affective content of the images, for example, Emotional Semantic Image Retrieval (ESIR) [19]. The affective of complexity contributes more in developing the retrieval level. Especially, in some applications like selecting prefer paintings or other arts, designing advertisements, affective and visual perception are much more critical [20]. Suppose the customers consider choosing some painting images only by visual feeling (e.g. aesthetic) instead of specific keywords (e.g. flowers), visual complexity necessarily has some information of composing the feeling [15, 21].

3) **Estimation of watermarking capacity**

Image complexity can also be used in watermarking [22] for copy protection [23]. It is regarded as an estimator of watermarking capacity. Determining the capacity of watermark in a digital image means identifying how much information can be hidden in image. More watermarking information can be transmitted in a complex image compare with a flat image [24]. Hence, knowing the watermark capacity is useful for selecting an appropriate watermark inserted into the image.

In an image, the capacity of watermark depends on each region in an image. There

Figure 1.1: An example host image: it is much difficult to insert watermark information in yellow rectangle region than in red rectangle region.

are two important requirements of watermark: invisibility and robustness [25]. In order to satisfy the property of invisibility, some flat regions in the image are not suitable for watermarking. For example in the Fig 1.1, the yellow rectangle (sky region) is totally flat; it is difficult to hide some information in this region. However, it is much easier to insert some information in the red rectangle region. Hence, this brings two questions: one is to estimate an appropriate watermarking capacity for image; the other one is to find the suitable regions that can be well inserted into watermark information.

4) **Image analysis**

Complexity-based methods have been applied to image analysis. In [26], the complexity-based method was applied to earth observation imagery, and in [27], approximation of Kolmogorov complexity was used for image classification. In addition, image complexity was applied to analyze the remote sensing image and it was used in the image classification in [28].

Moreover, image complexity measures have been also used to other image processing tasks like, automatic target recognition [29] and pattern recognition tasks [30].

1.4 Structure of the Dissertation

This dissertation consists of eight chapters from chapter 1 to 7.

Chapter 1 introduces the research background, the motivation, the objectives of this study, and the possible applications. The organization of this research is shown as Fig. 1.2.

Part I only includes **Chapter 2**. In this chapter, the basic theories of visual perception, the related works of visual complexity, and the previous methods of measuring complexity are introduced.

Part II includes **Chapter 3**, **Chapter 4**, and **Chapter 5**. In this part, a new method of measuring visual complexity perception in texture images is investigated.

Chapter 3 identifies five important factors that affect human visual complexity perception of textures from two experiments, namely, regularity, understandability, roughness, directionality, and density.

Chapter 4 presents the objective methods for measuring regularity, roughness, directionality, and density. Particularly, a new method to measure humans' understandability of a texture image is introduced.

Chapter 5 applies Multiple Linear Regression to build the model of texture's visual complexity based on texture features.

Part III contains **Chapter 6**. In this part, we propose a novel framework to assess visual complexity of paintings.

Chapter 6 illustrates an experiment of subjective assessment of complexity and identifies the factors that affect human visual complexity of painting images, and describes a set of feature extraction methods which globally and locally represent the above factors. Then these features are combined by a Support Vector Machine for predicting the visual complexity of painting images.

Chapter 7 summaries the presented works and discusses the perspectives.

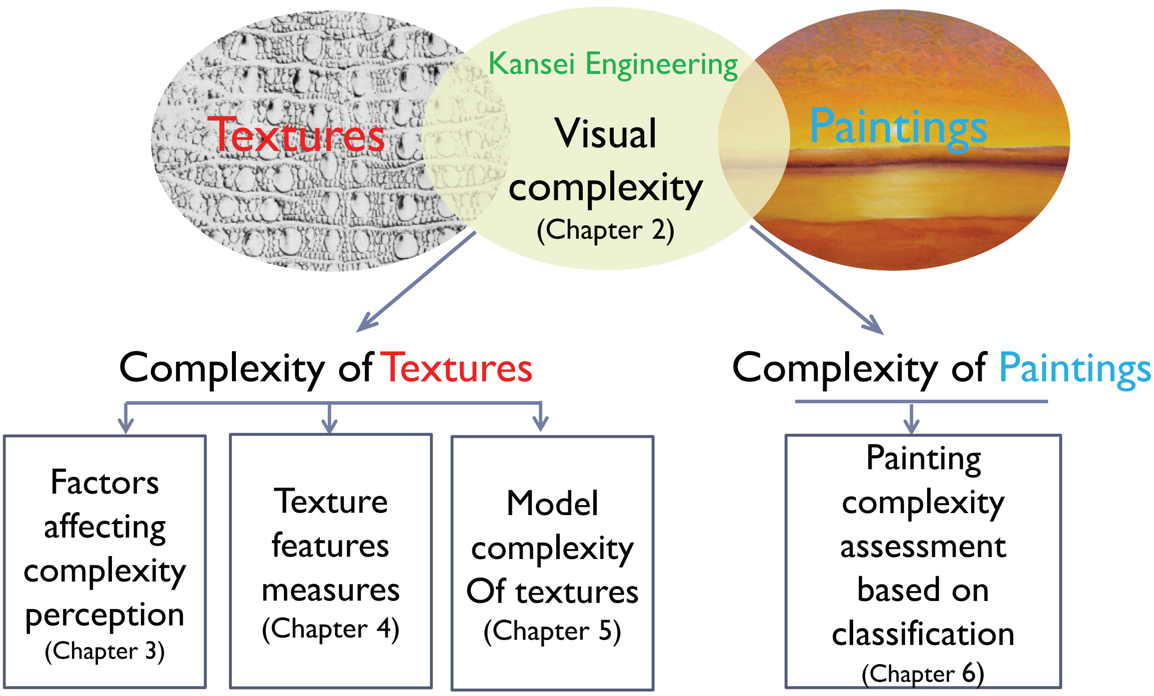


Figure 1.2: Scope of the research and structure of chapters in the dissertation.

Part I

Basic Theories of Visual Perception and Complexity

Chapter 2

Visual Perception and Complexity

2.1 Visual Perception

Vision is perhaps the most essential aspect of human senses. A large part of human brain devotes to vision. Visual perception is the ability to interpret the surrounding environment by processing information that is contained in visible light [31]. It is a function of our eyes (physiological reaction) and brain (psychological process). A number of structures within the eye contribute to visual perception, with the most important being the rods and cones in the back of the eye which respond to light, sending signals along the optic nerve to the brain so that the brain can interpret them. Another structure of interest in the eye is the pupil, which dilates and contracts to control the amount of light which enters the eye [32]. Then the brain works to make sense of the images it sees. The famous Gestalt theory explains us how the brain deals with visual input, and the ways in which the brain smooths out and normalizes images to make sense of them [32].

2.1.1 Gestalt theory of visual perception

Gestalt theory first arose in 1890 as a reaction to the prevalent psychological theory of the time - atomism. Atomism examined parts of things with the idea that these parts could

Figure 2.1: Examples of Gestalt principle: similarity.

then be put back together to make wholes [33]. Because there are many aspects in Gestalt psychology, we will not cover all aspects. Instead we will discuss the Gestalt principles that explain our perception.

2.1.1.1 Gestalt principles of perception

Figure and ground The concepts figure and ground explain how we use elements of the scene which are similar in appearance and shape and group them together as a whole. Similar elements (figure) are contrasted with dissimilar elements (ground) to give the impression of a whole.

Similarity, proximity or contiguity Similarity: If the distances between elements are the same, the ones that are physically similar will be grouped together, according to the principle of similarity. Therefore, green and red dots in the following figure seem to be organized in columns (in Fig. 2.1 (a)) and in rows (in Fig. 2.1 (b)).

Proximity: The principle of proximity or nearness enables us to group what we see according to closeness. Visual stimuli that are close together are grouped together. In the figure below, the circles are seen as arranged in pairs, shown as Fig. 2.2.

The principle of continuity predicts the preference for continuous figures. We perceive the Fig. 2.3 as two crossed lines instead of 4 lines meeting at the center.

Figure 2.2: Example of Gestalt principle: proximity.

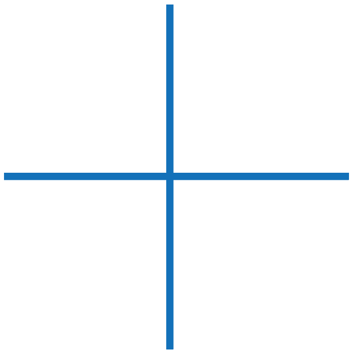


Figure 2.3: Example of Gestalt principle: continuity.

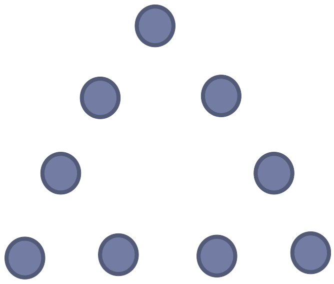


Figure 2.4: Example of Gestalt principle: closure.

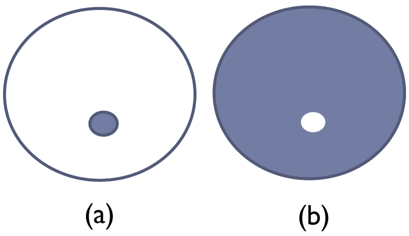


Figure 2.5: Example of Gestalt principle: area.

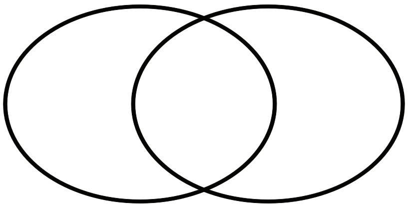


Figure 2.6: Example of Gestalt principle: symmetry.

Closure, area, symmetry The principle of closure applies when we tend to see complete figures even when part of the information is missing. For example Fig. 2.4, even though the circle to the left is not joined together, we still perceive a triangle due to the principle of closure.

The principle of area states that the smaller of two overlapping figures is perceived as figure while the larger is regarded as ground, shown as Fig. 2.5.

Symmetry: Grouping on the basis of symmetry refers to the perception of the more natural, balance, and symmetrical figure as the same unit. The Fig. 2.6 below show that perceptual organization follows the symmetrical pattern.

2.2 Visual Complexity and Related Works

2.2.1 Definitions of Complexity

In recent decades, visual complexity has become an important and appealing issue. It has been defined in many ways, but there is no uniform definition.

The earliest study of complexity derives from the nineteenth century. The definition of complexity was found to depend on the shape (lines and details) of the objects. Later (1957) the complexity was evaluated by the number of image features such as crossing lines, corners, turns and so on [34, 35].

The well known definition of complexity is Kolmogorov's complexity (1965). It is defined as: the length of the shortest description as a measure of the objects complexity. In other words, It is natural to call an object "simple" if it has at least one short description, and to call it "complex" if all of its descriptions are long. However, kolmogorov complexity is not computable.

Later in Webster's dictionary (1986), a complex object is defined as "an arrangement of parts, so intricate as to be hard to understand or deal with". The pioneers investigated in empirical aesthetics and visual perception have regarded complexity as one-dimensional concept. However, other studies have found visual complexity rely on two kinds of features. Nicki and Moss (1975) [36] suggested that there might be two kinds of factors that affect visual complexity. One is "perceptual" complexity ,which is related with the number and variety of elements; the other one is "cognitive" complexity, which is related with the amount of associations or cognitive tags elicited by stimuli. Different from the front, other researchers conceived that the complexity as a multi-dimensional concept. They consider complexity as the amount of elements in stimulus, such as, lines, angles, turns and so on, and complexity is influenced by the degree of asymmetry, incongruity and disorganization. Nadal et al. [37] identified seven complexity dimensions, respectively, asymmetry, amount of elements, element heterogeneity, variety of colors, and three-dimensional appearance. Although many researches have focused on the complexity, the concept of complexity refers to different aspects.

In recent decades, a number of researchers engaged themselves in psychology and computer science have defined the concept of visual complexity in their studies. Peters and Strickland [29] defined that image complexity is related to the number of objects and segments n image. Scha and Bod [38] described complexity as being largely a function of the number of elements that an image consists of and their order of placement in the image. Heylighen [39] suggested that the perception of complexity is correlated with the amount of variety in the visual stimulus. Heaps and Handel [40] defined complexity as "the degree of difficulty in providing a verbal description of an image." Similarly, Li [41] suggested that complexity is related to certain measures of difficulty concerning the object or the system. Rigau et. al [42] indicated image complexity is related to image intensity. In [43], Mario et. al considered that complexity is represented by a fuzzy interpretation of edges in an image. They emphasized that image complexity is related to how much attention been paid to detect and recognize objects by a person.

The above definitions of complexity demonstrate that there are different methods for computing image complexity depending on different applications. However, there is no any agreement on the definition of complexity.

2.2.2 Visual complexity of images

In the field of psychology, Olive et al. [5] investigated the perceptual dimensions of the visual complexity of scenes. In this study, 34 participants used the method of hierarchical grouping to classify indoor scenes. The results showed that visual complexity was represented by several dimensions such as the number of objects, clutter, openness, symmetry, organization, and the variety of colors. Pieters et al. [44] investigated the visual complexity of advertising. They distinguished two types of visual complexity (feature complexity and design complexity) in advertising and proposed a objective measure for each. Saleem et al. [45] studied the visual complexity of three dimensional (3D) shapes and introduced an approach based on view similarity to determine the perceived shape complexity. Purchase et al. [46] explored the visual complexity of images. They attempt to investigate whether the visual complexity could be quantified to match the human's perception of complexity. By an empirical study, they concluded that the subjective notion of complexity is consistent both to an individual and to a group, but that it does not easily relate to the most obvious computational metrics.

2.2.3 Visual complexity and aesthetic

Visual complexity has been known to be a significant predictor of preference for artistic works [47]. In the research of visual aesthetics, visual complexity is argued as a primary cue on judgments of visual appeal.

The investigation of visual complexity and beauty can date back to the ancient Greeks [37]. However, it was not until Fechner's (1876) [48] work that the research of visual complexity were systematically studied. In the early nineteenth century, some pioneers [49, 50, 51] studied the influence of order and complexity to perceived beauty, they pointed that both order and complexity contribute positively to the appreciation of beauty.

In 1971, Berlyne [17] presented his influential framework that the study of complexity's influence on the appreciation of beauty based on firm psychological and neurobiological grounds. Berlyne found that preference and interest increase linearly with visual complexity until an optimum level of arousal is reached. With increasing complexity over the optimum point, pleasure begins to decline. Berlyne stated that the moderate complexity has the biggest affective appraisal. That means the relationship of complexity and appeal represents an inverted U-shape, as can be seen in Fig. 2.7.

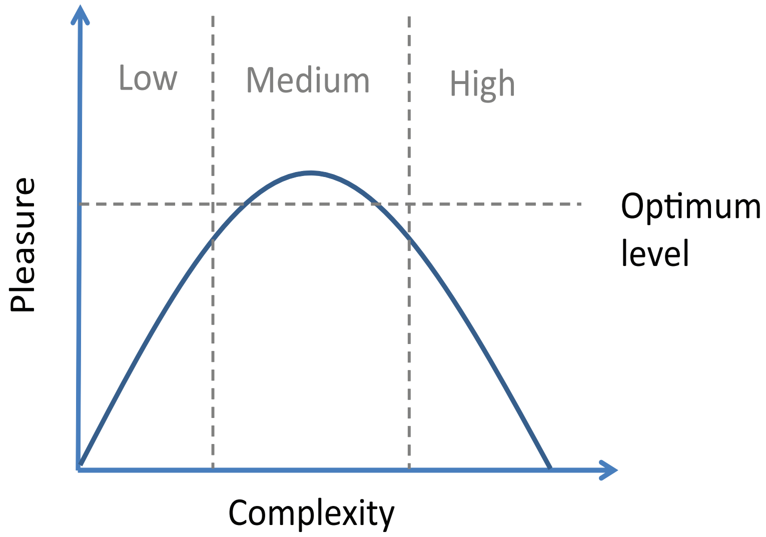


Figure 2.7: Berlyne theory [17]: relationship between complexity and pleasure.

2.2.4 Visual complexity and websites

In terms of factors that influence website appeal, many researchers believe that visual complexity is one of the most important factors. Visual complexity is an important cue to assess the visual appeal of images (especially the art works, such as, paintings, wallpapers et al.). In the real applications such as websites and interfaces of softwares, visual complexity also has a part to play in how much a user like the interface and how appeal the websites will be. Hence, investigating visual complexity is good for the study of visual understanding and visual aesthetic.

Michailidou [52] studied the visual complexity of web pages and defined ViCRAM, a framework that describes the visual complexity of web pages. These studies have made progress to some extent in measuring visual complexity by using information theory and pattern methods. Tuch et al. [53] investigated the first impression of a website's complexity and perceived aesthetics by the experiment. They found there is a linear relationship between complexity level of website and its visual appeal. The websites with low VC and high PT were perceived as highly appealing.

Similarly, Eleni et al. [54] introduced an investigation into user perception of the visual complexity and aesthetic appearance of Web pages. The visual complexity and aesthetic influence the use's interaction with the web pages. The experimental results showed that visual complexity of a page is negatively related with use perception of how organized, clear clean and beautiful a page looks. The authors suggested that this result should be used by designer for balancing the aesthetic appearance of a web page and its visual complexity.

In addition, Reinecke et al. [55] demonstrated a method to predict uses' first impression of website aesthetic based on website colorfulness and visual complexity. They collected ratings of visual complexity, colorfulness and visual appeal of 450 websites online. Based on these experimental data, they proposed computational methods that measure the perceived visual complexity and colorfulness of a website.

2.2.5 Visual complexity and watermarking

Digital watermarking is the process of embedding information into digital multimedia content (audio, image, video and printed materials) such that the information (which we call the watermark) can later be extracted or detected by computers and digital devices. The inserted information should be imperceptible to human senses. In the issue of watermarking, the capacity of information inserted into the host multimedia content is very essential.

Image complexity can be used for estimating watermarking capacity. Generally it is assumed that complex images have more watermarking capacity than simple ones [22]. Hence, if the complexity of the host image is measured, the inserted watermark could be appropriately determined. In [22], Yaghmaee et al. proposed a method to estimate the capacity of images based on the images' complexities. They introduced the idea of Region Of Interest (ROI) to estimate the image complexity. Their experimental results show the proposed capacity estimation method is better in robustness and image quality than the previous work.

Besides, the local complexity of the image is also of importance. In order to achieve one of the requirements of watermarking, invisibility, it is necessary to estimate the regional complexity of the image. The most complexity region of the image holds the highest watermarking capacity. If the watermark information is inserted into the most complexity region, the inserted information could be well hidden. In this dissertation, we extend our method to measure the local complexity of the image. The result will appear in Chap. 7.

2.2.6 Visual complexity and information theory

In information theory, entropy is a measure of the uncertainty in a random variable [56]. In terms of entropy, we usually refer to the Shannon entropy, which quantifies the expected value of the information contained in a message [57]. Primarily, we regard a system holding the large entropy as a complex system. As a result, information entropy becomes a classical and basic way of measuring complexity. In [6], the authors presented a new method of measuring the complexity of the images based on Independent Component Analysis (ICA), and used this method to estimate the joint complexity of images that measure the similarity of the natural images. ICA in this paper was used to estimate the entropy of images. In addition to this, many other derivations from Shannon entropy are used for evaluating image complexity, such as, fuzzy entropy, Kolmogorov-Sinai entropy [58] and so on. We will introduce these measures in the following section.

2.3 Measuring Visual Complexity

Investigations of measuring visual complexity have been done for several decades. Up to now, there are different kinds of methods for measuring visual complexity, such as, information theory, fractal, fuzzy entropy, patterns and so on. In the following contents, we will introduce some measures for computing image complexity.

In the beginning of this chapter, we introduced the famous definition of complexity: Kolmogorov complexity. It is originated with concepts of complexity of description, randomness and a priori probability. However, the weakness of Kolmogorov complexity is that it is not computed easily and it is not directly related to the mechanisms of visual perception. For the image complexity estimation, however, a computable algorithm is necessary.

2.3.1 Information theory

Andrienko et al. [7] developed a complexity measure based on mean information gain and applied it to two-dimensional (2D) patterns. Patel and Holt [59] compared a pattern measure proposed by Klinger and Salingaros [8] with respondents' perceptions of the complexity of background image scenes; the results showed that a high and positive correlation existed between mathematical measures and the subjects' perceptions. Furthermore, Rigau et al. [42] proposed a new framework for investigating the complexity of an image by using information theory.

2.3.1.1 Entropy and fuzzy approach

In [9], Cardaci et al. proposed a fuzzy mathematical model of visual complexity based on fuzzy measures of entropy. The authors selected fuzzy entropic distance functions described in [60] and [61]. Besides, the authors also conducted an experiment to obtain the perceived time that can be used as an indicator of the complexity of an image. The results showed that the model fits well with the subjective measure of complexity based on perceived time.

Mario et al. [10] proposed a novel fuzzy approach to determine the complexity of an image based on analysis of edge level percentage. They classified the the images into three levels: Little Complex, More or Less Complex and Very Complex. Then the membership value belong to that class is calculated by the interval mapping functions.

2.3.1.2 Image composition complexity

Image composition complexity (ICC) is first proposed in [42]. In this method, Jensen-Shannon divergence was used to express the image composition complexity of an image.

This method can be regarded as the spatial heterogeneity of an image from a given partition. The Jensen-Shannon divergence applied to an image is defined by

$$\begin{aligned} JS(X,\hat{Y}) &= H(X) - \sum_{i=1}^{R} \frac{n_i}{N} H(X_i) = H(X) - H(X|\hat{Y}) = I(X,\hat{Y}), \\ H(X) &= -\sum_{i=1}^{N} p_i log p_i, \end{aligned} \tag{2.1}$$

where R is the number of regions of the image, X_i is the random variable associated with region i, which is the intensity histogram of ith region, n_i is the pixel number of region i, N is the total number of the whole image, p_i is the probability distribution, and $H(X)$ id the Shannon entropy.

Since ICC is based on the image segments, the partition of the image is very important for calculating the complexity of the image. Thus, given an segmented image, we can calculate the heterogeneity of an image using JS-divergence applied to the probability distribution of each region.

2.3.1.3 Compressed file size of image

Another possible way to assess complexity of an image is the utilization of digital image compression. The file size after compression (JPEG, GIF) estimates human subjective assessment of complexity. Hence a larger file size indicates a higher complexity. Donderi [11] verified that the objective indication of visual complexity od an image can be measured by JPEG compressed file size. In [47], it was found that GIF file size of an image is correlated with human perception of visual complexity.

Compression based method should be the simplest method of measuring the complexity of an image. However, it is abstract and difficult to explain why some images look more complex than others. Moreover, these algorithms do not sufficiently use the mechanisms of information processing in human vision.

2.3.2 Fractal dimension

In [12], Cutting et al. investigated the relationship between fractal curves and complexity. They found that fractal dimension together with number of segments, and recursion depth provide a good prediction of complexity. Fractal dimension (FD) has often been applied as a parameter of complexity, related to, for example, surface roughness, or for classifying textures or line patterns [62]. In [47], fractal dimension was used as an estimator of both aesthetic and complexity in arts. The experimental results showed that FD accounts for more of the variance in judgments of perceived beauty in visual art than measures of visual complexity alone, particularly for abstract and natural images. The most often method of measuring fractal dimension is box-counting dimension [63].

2.3.3 Quad tree method

A quadtree is a tree data structure in which each internal node has exactly four children. Quadtree is most often used in binary image by recursively subdividing it into four quadrants or regions, but it can also be used in the gray images. In the gray images, there is a criterion of subdividing, that is if the intensity variance of the current block is lower than a given value, the block will not be divided into four sub-blocks. It means that the flat block contains less information and is less complex. Therefore, a quadtree can be used for measuring image complexity [13].

$$Complexity = \sum_{i=1}^{N}(n_i \times 2^i), \tag{2.2}$$

where, i is the level number in a quadtree with N levels, and n_i is the number of nodes in level i.

2.3.4 Region of Interest method

Region of Interest (ROI) of an image usually attracts more attention than other regions of the image. Hence, it is better to analyze the image complexity from the points of ROI. In [22], the author proposed a method of measuring complexity based on image ROI. Image was divided into several blocks. Each block is one region. In each ROI, five features were extracted, including intensity ($M_{intensity}$), contrast ($M_{contrast}$), location ($M_{location}$), edginess ($M_{edginess}$) and texture ($M_{texture}$). Here, M means the computing value of each feature. After all features were extracted from ith region, the image complexity is calculated by following equation:

$$IM(S_i) = M_{intensity}(S_i)^2 + M_{contrast}(S_i)^2 + M_{location}(S_i)^2 + M_{edginess}(S_i)^2 + M_{texture}(S_i)^2, \quad (2.3)$$

where, IM is the important measure (IM) of the ith region. The region having the highest value is perceptually most important region.

The IM score of each region is calculated by the above equation. Then the algorithm sums all IM(S_i) of all regions. Finally, the mean score of sub regions is taken as an estimator of image complexity.

2.3.5 Visual attention method

In [14], Silva et al. proposed a method of measuring image complexity based on visual attention. The experimental result showed attentional behavior is a good estimator of image complexity. In this paper, the author verified the assumption that the heatmaps generated by a computational model of attention vary from each other according to image complexity. The heatmaps are obtained by conducting the method in [64].

The above measures do not fully consider the knowledge of information process in human visual. In order to understand and measure visual complexity, it is vitally important to develop a measure that is theoretically informed and can account for some of the processes involved in the perception of complexity [47].

2.4 Affective Engineering and Kansei Engineering

Emotions play important roles in human intelligence, rational decision making, social interaction, perception, and so on. Emotional skills, especially the ability to recognize and express emotions is essential for natural communication with humans. Rosalind pointed in the book of “Affective Computing” that emotions play a pivotal role in functions considered essential to intelligence [65]. This new understanding about the role of emotion in humans indicates a need to rethink the role of emotion in computers. The purpose of affective computing is try to give emotional abilities to computers.

Affective engineering includes many things, such as endowing a computer the ability to recognize and express emotions, developing some algorithms that responds human’s emotions.

Affective engineering is also called kansei engineering in Japan. “Kansei” is a Japanese word, that can be explained as the meaning of feeling, impression, and emotion and so on. Kansei research is an attempt to enumerate and quantify those subjective feelings, such as visual aesthetic perception, pleasing feeling and complexity perception and so on. A picture has certain factors that can affect viewers’ impressions. These factors are called kansei factors [1] in Japan. The challenge in kansei research is to identify the kansei factors, and associate them with the feeling brought forth in the viewer.

The goal of this paper is to endow a computer the ability of recognizing complexity of an image as human’s emotion. In order to achieve this goal, we employ the methods of kansei engineering to obtain the subjective visual complexity evaluations and map the relationship between the objective image features with the subjective complexity.

Part II

Modeling the Perception of Visual Complexity in Texture Images

Chapter 3

Visual Complexity of Texture Images

3.1 Introduction

Evaluation of visual complexity aims at investigating humans' "*Kansei*" of the complexity of the visual scene. It can be extended to include aesthetics, visual psychology, and cognitive systems. In addition, research into visual complexity is useful in understanding the mechanism of human perception and is of interest to real applications such as image compression and information theory [66].

Texture is a part of our daily life - it's in natural and artificial things all around us. Visual scenes are composed of numerous textures, objects, and colors. Although scenes are visually complex, human beings are able to form a coherent perception of complexity and identify a complex image or object at a glance [5]. This provokes the question of how human beings extract information from visual scenes and which characteristics of images affect humans' perception of visual complexity.

3.1.1 Texture perception

Many studies regarding the visual perception of textures have been conducted. Tamura et al. [67] developed six texture properties (coarseness, contrast, directionality, line-likeness, regularity, and roughness) that correspond to human visual perception. In addition, they developed computational measures and compared them with the psychological measurements given by respondents. The comparison indicated strong correlations between certain properties. Amandasun and King [68] investigated five properties of texture corresponding to human visual perception: coarseness, contrast, busyness, complexity, and texture strength. Fujii et al. [69] presented a set of measures for texture parameters that correspond to the perception of visual texture. Li et al. [70] proposed a method of evaluating the visual impression of gray-scale textures using morphological manipulation. They used this method to investigate the influences of global features and local features on the visual perception of similarity.

Perceptual dimensions of textures have also been proposed. Rao and Lohse [71, 72] proposed that three dimensions were important to texture perception: (1) repetitive vs. non-repetitive; (2) high contrast and non-directional vs. low-contrast and directional; and (3) granular, coarse and low-complexity vs. granular, coarse and high-complexity. Cho et al. [73] extended the number of perceptual dimensions of texture to four: coarseness, regularity, contrast, and lightness. However, few studies have been conducted on the visual perception of complexity in textures.

3.1.2 Visual complexity of textures

A texture image is consist of repeated or random arrangement of simple elements. The arrangement of these elements creates different visual impressions. Usually, textures are visually described by Kansei words such as “simple” and “complex.” However, the criteria for judging a texture as simple or complex are unclear.

Many studies of visual perception have featured texture images [67, 68, 69, 71, 72]; however, few researches have been carried out into the visual complexity of texture images. Motivated by the above considerations, we aim to identify the characteristics of texture images that affect humans’ perception of visual complexity.

3.2 Texture Features that Affect Visual Complexity

3.2.1 Experiments and results

In order to achieve the objectives, two experiments involving visual complexity assessment and paired comparison evaluation were carried out.

The first experiment was conducted on the purpose of identifying the characteristics that evoke the visual complexity impression of the texture images. In this experiment, the respondent was asked to evaluate the visual complexity of each texture and describe the characteristics of textures that affect their evaluation. From the summaries of descriptions given by respondents, we obtained the main characteristics of textures that affect respondents' visual complexity. In order to analyze the relationship between these characteristics and visual complexity, the second experiment was performed by conducting a series of paired comparisons.

3.2.1.1 Experiment set-ups

Respondents For each experiment, 30 respondents from Hiroshima University with a background in information engineering, education, management, or social economics participated in the experiment. Although some of the respondents were engaged in image science, they were unaware of the purpose of this study. Their ages ranged from 20 to 35 years old. All respondents had normal or corrected-to-normal vision. All respondents are native Chinese speakers.

Apparatus and stimuli Twenty texture images were selected for the experiments (Fig. 3.1). The sample images were obtained from a standard source, Brodatz's album [74], which has been widely used in the fields of texture analysis and visual perception.

In these two experiments, sample images were placed on a single screen and shown to respondents one by one in a random order. The screen was part of a 46-inch plasma

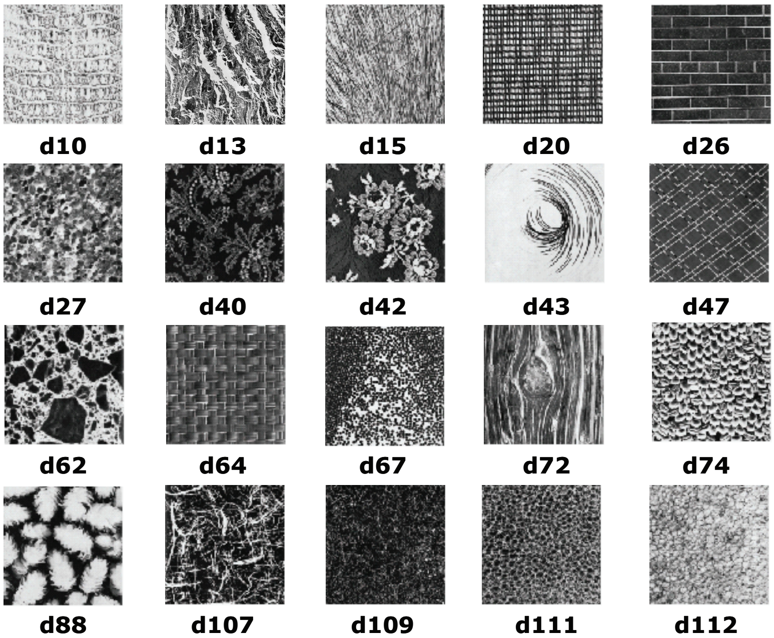


Figure 3.1: Twenty texture images used in the experiments.

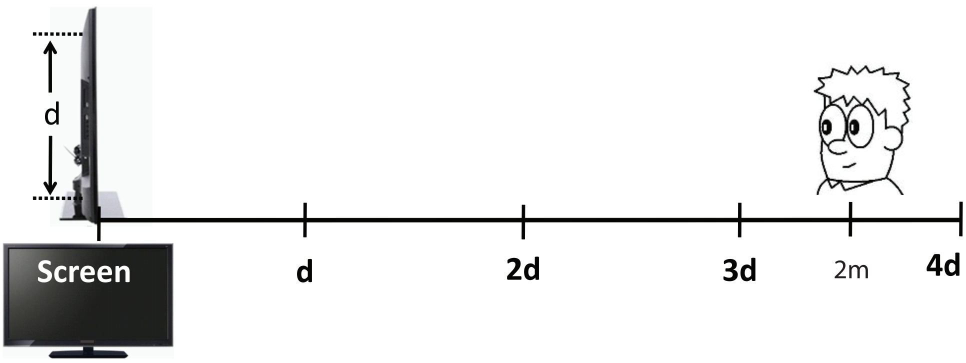


Figure 3.2: The screen used in the experiment and the distance from the respondents to the screen.

display (the type of display is Panasonic TH-P46V1). The experiments were conducted in a laboratory with normal illumination. The respondents were required to sit at 2 meters from the screen. The status of the screen and viewing distance are shown in Fig. 3.2.

3.2.1.2 Semantic differential method

The semantic differential method is the most popular way to acquire kansei information from respondents' verbal reaction [75]. Respondents describe their feelings about the scenes or products on a rating scale (usually a 5-point or a 7-point Likert scale). The scales are often designed from a pair of adjectives, for instance, simple and complex. In this paper, a 7-point Likert scale was used in the first experiment and a 5-point Likert scale was used in the second experiment.

3.2.1.3 Experiment 1: visual complexity assessment

The first experiment was briefly described to the respondents, and then the texture samples were displayed two times. On the first time, each respondent viewed all samples one by one with no time constraint. On the second time, for each texture sample, the respondents were asked to score complexity on a 7-point Likert scale by using their own knowledge and judgment. The 7-point scale ranged from 1 (very simple) to 7 (very complex). After scoring, the respondents were asked to verbally describe the characteristics of textures that affect their evaluation of visual complexity perception. Each verbal description was recorded and classified into the corresponding criterion. Table 3.1 summarizes all the criteria given by the respondents and the frequency of criteria that they used to perceive the complexity of texture images. The scores of complexity from the respondents are listed in the first column of Table 3.2.

In Table 3.1, the frequency indicated the strength of the criteria that the respondents used to perceive the complexity of textures (i.e., most commonly used, often used, or seldom used). Ninety percent of the respondents regarded regularity to be the main characteristic influencing their complexity assessment. In addition, high frequencies were recorded for understandability (almost 67%), density (approximately 57%), and directionality (50%). Specifically, roughness was not verbally described by the respondents in the first experiment; however, roughness was defined as a combination of contrast and coarseness in Tamura's research [67]. Hence, we adopted roughness instead of contrast and different primitives described in this experiment. Hence, it is concluded that regularity, understandability, density, directionality and roughness are the main characteristics of textures that affect human visual perception of complexity in textures.

Table 3.1: Summary of verbal descriptions from 30 respondents who participated in Experiment 1.

Descriptions	*Frequency*
Regularity	**27**
Understandability	**20**
Density	**17**
Directionality	**15**
Contrast	10
Different texture primitives	9
Structure	8
Symmetry	6
Nonuniform	4

3.2.1.4 Experiment 2: paired comparison evaluation

The method used in this experiment was paired comparison evaluation, which is widely used in the field of psychology [76]. The method of paired comparison is perhaps the most straightforward way of presenting items for comparative judgment.

Five pairs of adjectives were used for the paired comparison evaluation: namely irregular versus regular, low density versus high density, nondirectional versus directional, smooth versus rough, and understandable versus abstract. The five pairs of comparisons were defined as follows. (1) Regularity: irregular versus regular. Regularity was defined as variation in the placement rule of texture primitives (or texture elements), in agreement with the definition of regularity in Tamura's research. (2) Density: low density versus high density. Density was used for testing whether the perceived primitives and edges were dense or sparse. (3) Directionality: nondirectional versus directional. The directionality of texture was related to primitive shape and the global placement rule, in agreement with Tamura et al. (4) Roughness: smooth versus rough. This property is fundamentally related to touch; however, when we observe each texture, we can decide if the texture feels

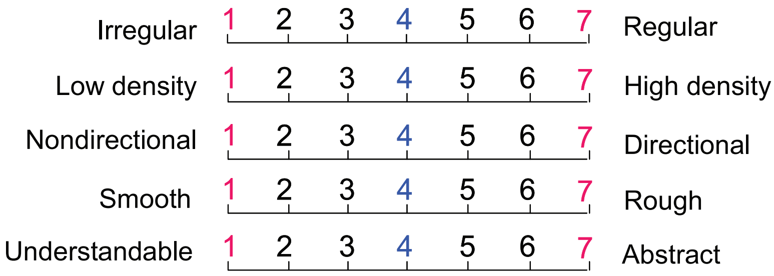


Figure 3.3: The 7-point rating scale used in the first experiment.

rough or smooth. (5) Understandability: understandable versus abstract. This was related to respondents' prior knowledge and experience.

This experiment was conducted under the same conditions as those in the first one. After an introduction to the experiment and a brief explanation, the respondents were instructed to view all the images one by one. For each paired comparison, the respondents scored the corresponding perception on a 7-point Likert scale. The scale and its anchor-point phrases are shown in Fig. 3.3. The order of the presentation of the samples was randomized to avoid any order effect. Table 3.2 shows the average scores for the complexity and paired comparisons evaluated by the respondents. The texture images were sorted in ascending order by average evaluation score, and these results are shown in Table 3.3.

3.2.2 Results and Statistical analysis

From the visual complexity assessment experiment, we obtained the major characteristics of textures that affect visual complexity perception. From the paired comparison evaluation experiment, we got the subjective evaluations of different characteristics respectively. In the following contents, we will conduct several statistical analyses methods to analyze the experimental data.

3.2.2.1 Correlation analysis between texture characteristics and complexity

Correlation analysis method is used to test whether there is a linear relationship between two quantitative variables. The most common correlation analysis is the Pearson correla-

Table 3.2: Average scores of subjective perception complexity, regularity, density, directionality , roughness and understandability.

	Com.	*Reg.*	*Den.*	*Dir.*	*Rou.*	*Und.*
$d10$	4.27	4.07	4.47	4.67	4.47	3.70
$d13$	5.80	1.70	5.00	3.27	6.07	5.03
$d15$	4.77	2.47	5.70	4.23	4.93	4.43
$d20$	4.07	6.10	6.17	6.30	4.10	3.67
$d26$	1.43	6.53	2.53	6.63	2.87	1.07
$d27$	5.80	2.43	5.27	1.87	5.20	5.20
$d40$	3.53	3.43	3.40	3.40	3.40	2.17
$d42$	2.93	4.17	2.50	3.60	3.10	1.50
$d43$	3.07	2.53	1.83	4.37	2.33	3.90
$d47$	2.47	6.53	2.80	6.60	2.13	2.00
$d62$	4.37	2.43	3.73	2.23	4.70	4.80
$d64$	3.03	6.13	5.07	6.30	3.93	1.73
$d67$	3.30	3.53	5.33	3.03	3.40	3.50
$d72$	5.27	2.73	4.73	4.87	5.30	3.23
$d74$	4.10	4.27	5.37	3.23	3.53	2.93
$d88$	3.07	4.60	3.17	3.47	2.83	2.57
$d107$	5.50	1.57	4.80	1.38	5.20	6.07
$d109$	5.47	2.33	5.07	1.43	5.00	6.10
$d111$	5.50	2.77	6.47	2.07	4.40	4.93
$d112$	5.63	2.83	5.97	2.03	4.60	5.40

Com., *Complexity*; *Reg.*, *Regularity*; *Den.*, *Density*;
Dir., *Directionality*; *Rou.*, *Roughness*; *Und.*, *Understandability*

Table 3.3: Ranking of average scores for 20 texture images in ascending order.

Com.	*Reg.*	*Den.*	*Dir.*	*Rou.*	*Und.*
$d26$	$d107$	$d43$	$d107$	$d47$	$d26$
$d47$	$d13$	$d42$	$d109$	$d43$	$d42$
$d42$	$d109$	$d26$	$d27$	$d26$	$d64$
$d64$	$d62, d27$	$d47$	$d112$	$d88$	$d47$
$d88, d43$	$d15$	$d88$	$d111$	$d42$	$d40$
$d67$	$d43$	$d40$	$d62$	$d67$	$d88$
$d40$	$d72$	$d62$	$d67$	$d40$	$d74$
$d20$	$d111$	$d10$	$d74$	$d74$	$d72$
$d74$	$d112$	$d72$	$d13$	$d64$	$d20$
$d10$	$d40$	$d107$	$d40$	$d20$	$d67$
$d62$	$d67$	$d13$	$d88$	$d10$	$d10$
$d15$	$d42$	$d64, d109$	$d42$	$d111$	$d43$
$d72$	$d10$	$d27$	$d15$	$d112$	$d15$
$d109$	$d74$	$d67$	$d43$	$d62$	$d62$
$d111$	$d88$	$d74$	$d10$	$d15$	$d111$
$d107$	$d20$	$d15$	$d72$	$d109$	$d13$
$d112$	$d64$	$d112$	$d20, d64$	$d107, d27$	$d27$
$d13, d27$	$d26, d47$	$d20$	$d47$	$d72$	$d112$
		$d111$	$d26$	$d13$	$d107$
					$d109$

Com., *Complexity*; *Reg.*, *Regularity*; *Den.*, *Density*;
Dir., *Directionality*; *Rou.*, *Roughness*; *Und.*, *Understandability*

Table 3.4: Correlation matrix of perceptual characteristics of textures and visual complexity.

	Reg	*Den*	*Dir*	*Rou*	*Und*
Com	−0.794∗	0.590∗	−0.711∗	0.879∗	0.877∗
Und	−0.838∗	0.484	−0.763∗	0.712∗	
Rou	−0.753∗	0.480	−0.491		
Dir	0.716∗	−0.386			
Den	−0.177				

* p < 0.01

tion coefficient, which is sensitive only to a linear relationship between two variables. It is defined as:

$$\rho(x, y) = \frac{cov(x, y)}{\sigma_x \sigma_y}, \tag{3.1}$$

where, x and y are the two sets of variables, σ_x and σ_y are standard deviations, and $cov(x, y)$ means covariance.

We used a Pearson correlation analysis to investigate the correlation between characteristics of textures and visual complexity. The results of the analysis are shown in Table 3.4.

Table 3.4 shows that complexity is strongly correlated with understandability ($r = 0.877, p < 0.01$), which indicates that prior knowledge and experience considerably affect human perception of complexity; this is in agreement with the definition of complexity in Webster's dictionary, i.e., a complex object is one that is difficult to understand or deal with. Interestingly, although roughness was not mentioned in the first experiment, it shows a high correlation ($r = 0.879, p < 0.01$) with the perception of complexity. This might be partly because the respondents perceived roughness to be associated with the imagination of texture images (which relates to understandability); this is demonstrated by the

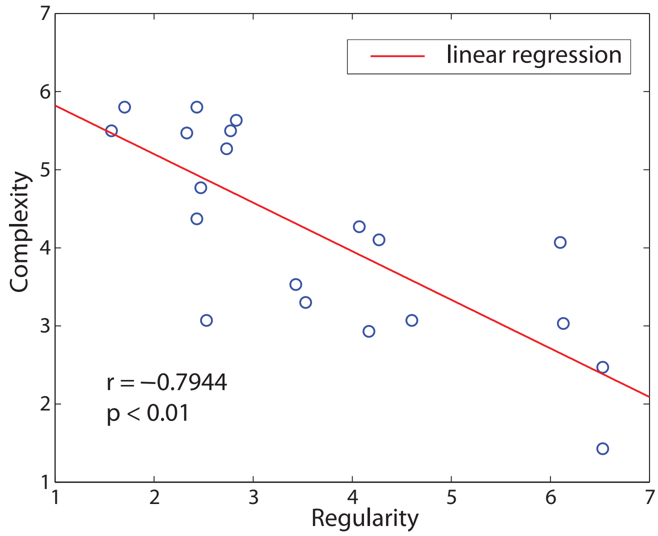


Figure 3.4: Relationships between regularity of textures and visual complexity.

correlation between roughness and understandability ($r = 0.712, p < 0.01$).

Figure 3.4 and 3.5 clearly show that regularity and directionality have negative correlation with visual complexity perception. On the contrary, roughness and understandability have a strong positive correlation with visual perception of complexity (shown as Fig. 3.6 and 3.7). For instance, texture d107 was perceived to be fairly complex because it was the most irregular and had the least directional texture. In addition, texture d13 was perceived to be very complex because of its characteristics such as irregular placement, rough feeling, and hard to understand.

Visual complexity is a function of not only each individual characteristic but also interactions between them, which is demonstrated by the correlation coefficients of perceptual characteristics in Table 3.4. The correlation between regularity and understandability is high ($r = -0.838, p < 0.01$). In general, textures characterized by regular placement are easy to understand, leading to a perception of less visual complexity. Similarly, the correlation between directionality and understandability is also very high ($r = -0.763, p < 0.01$). An interaction exists between roughness and understandability, regularity and roughness, and directionality and roughness. Therefore, it is suggested that the respondents used a

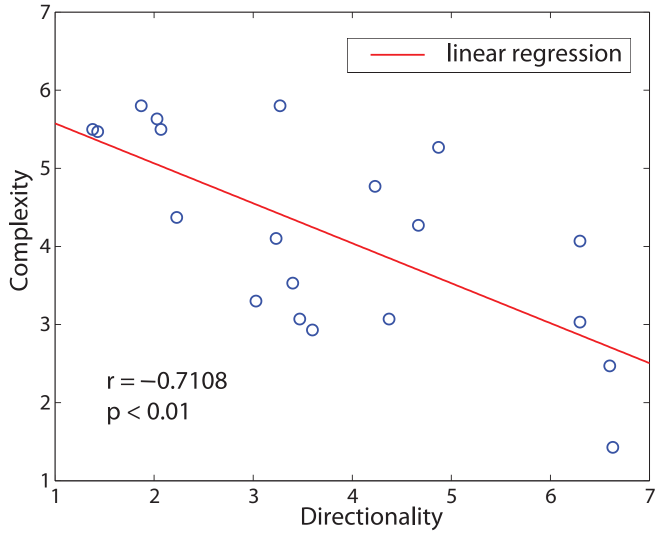


Figure 3.5: Relationships between directionality of textures and visual complexity.

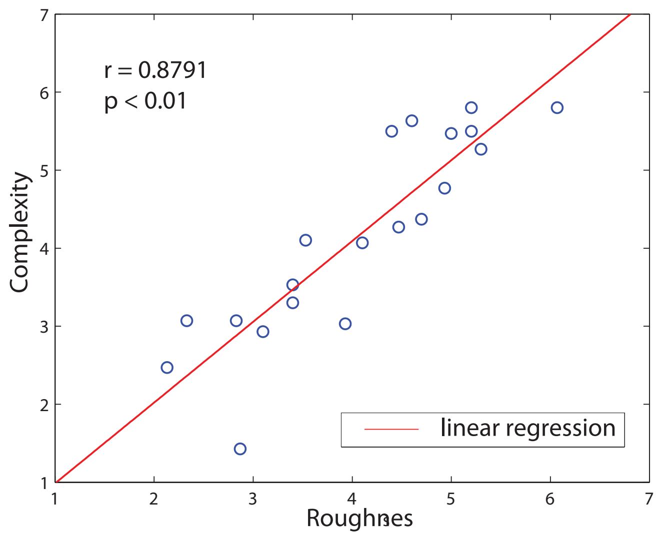


Figure 3.6: Relationships between roughness of textures and visual complexity.

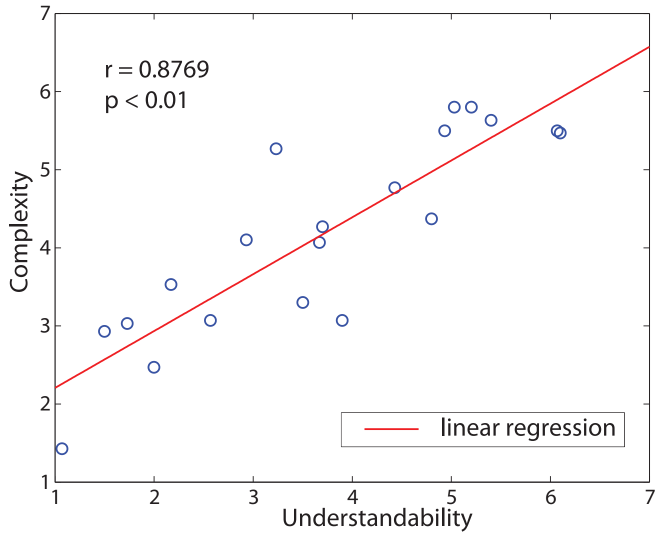


Figure 3.7: Relationships between understandability of textures and visual complexity.

different combination of these characteristics while evaluating the visual complexity of textures.

In some cases, one or two characteristics of textures dominated the respondents' evaluation of visual complexity. In the experiments, textures d42 and d26 were evaluated as having similar level of visual complexity, although d42 was more irregular and less directional. For d42, its characteristic of understandability led the respondents to assess its complexity as being similar to that of d26. For texture d43, although it was perceived as being smoother, having lower density, and being more directional, its abstract understandability property resulted in it being perceived to be more complex than d42. In these cases, understandability dominated the respondents' evaluation. Moreover, the high correlations between understandability and three other salient characteristics also demonstrated that prior knowledge and experience have significantly affected the visual perception of complexity of texture images.

Table 3.5: Factor loadings without rotation.

Variables		*Factor Loadings*	
		*Factor*1	*Factor*2
Irregular	*Regular*	−0.989	0.124
Low density	*High density*	0.377	0.918
Nondirectional	*Directional*	−0.828	−0.004
Smooth	*Rough*	0.711	0.430
Understandable	*Abstract*	0.834	0.263
Contribution(%)		**60.16**	**22.25**
Accumulative Contribution(%)		60.16	82.41

3.2.2.2 The importance of perceptual characteristics for visual complexity

For each texture sample, the evaluated values of five paired adjectives were statistically standardized. To investigate the importance of perceptual characteristics for visual complexity perception, these values were analyzed using factor analysis.

Factor analysis is a kind of multivariate analysis technology. In particular, it seeks to discover if the observed variables can be explained largely or entirely in terms of a much smaller number of variables called factors. The factor loadings, also called component loadings in PCA, are the correlation coefficients between the variables (rows) and factors (columns). It is similar with the correlation coefficient.

Principal component analysis was employed for defining a set of factors. Finally, two factors were extracted. The results of factor loadings are shown in Table 3.5.

Table 3.5 shows that regularity, directionality, roughness, and understandability contribute considerably to explaining factor 1, and density contributes significantly to explaining factor 2. Accumulative contributions of two factors show that 82.41% of human perceptions of visual complexity can be explained by these two factors. Factor 1, which has a contribution of over 50%, plays a particularly influential role in affecting human visual

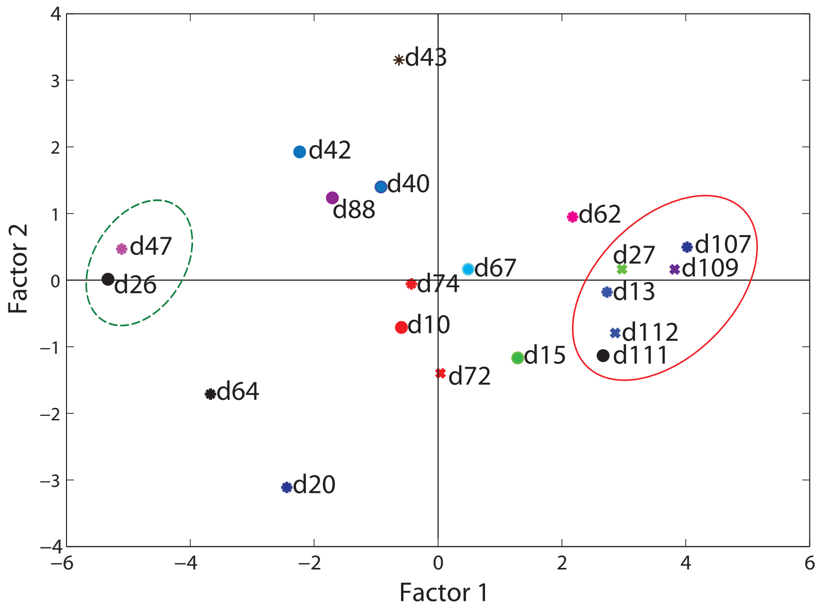


Figure 3.8: Multidimensional Scaling map.

perception of complexity.

3.2.2.3 Multidimensional scaling map of texture characteristics and complexity

Multidimensional scaling (MDS) is a set of related statistical techniques often used in information visualization for exploring similarities or dissimilarities in data. MDS is a special case of ordination. An MDS algorithm starts with a matrix of itemitem similarities, and then assigns a location to each item in N-dimensional space, where N is specified a priori. For sufficiently small N, the resulting locations may be displayed in a graph or 2D visualization techniques such as scatter plots [77].

After applying factor analysis, we used method of MDS to diagrammatically display the results of the evaluation in the environment of Matlab 2011b. Two dimensions (defined using factor analysis) were used to create the MDS map, as shown in Fig. 3.8.

The MDS map provides the following graphical representation of the texture samples: along the horizontal axis, the samples are mapped according to factor 1 (regularity, directionality, roughness, and understandability), with the simpler (regular, understandable, and directional) samples on the left and the more complex (irregular, abstract, and nondirectional) samples on the right. Along the vertical axis, the samples are positioned according to factor 2 (density), with the lower density samples at the top and the higher density samples toward the bottom.

MDS uses similarities and dissimilarities among the complexity evaluations given by respondents and provides a representation of visual complexity perception in a 2D map. As seen in Fig. 3.8, the left dotted circle and right solid circle correspond to respondents' definitions of a simple group and a complex group, respectively. In addition, the points within both circles are clustered horizontally, which appears to suggest that visual perception of complexity is deterministically affected by factor 1 (regularity, directionality, roughness, and understandability).

3.3 Summary

The purpose of this chapter is to identify the image features that affect human visual complexity of textures. In order to achieve this objective, we performed two experiments involving visual complexity assessment and paired comparison evaluation. In addition, the techniques of correlation analysis, factor analysis, and multidimensional scaling were employed to further analyze the experimental results.

In summary, five important texture characteristics that evoke visual complexity perception of texture images have been identified: regularity, understandability, roughness, directionality, and density. Visual complexity was a function of not only each individual characteristic but also the interactions between them. Regularity, understandability, directionality, and roughness were shown to be the most influential characteristics affecting visual complexity evaluation. Moreover, in the case of texture images with similar level of regularity or directionality, understandability dominates the evaluation of visual complexity.

Chapter 4

Measures of Texture Features

4.1 Introduction

In Chapter 3, we have identified five low-level characteristics that are used by humans to perceive the visual complexity of textures; namely, regularity, roughness, directionality, density, and understandability. In this chapter, we aim to develop a set of methods to objectively measure these characteristics. This will facilitate the mapping between objective texture characteristics and human visual complexity perception. It will also make the emotional based image retrieval possible in the real application.

In this chapter, we develop a set of objective methods to measure four texture characteristics (regularity, roughness, directionality, and density). Regularity is the property of variations across a whole texture. We estimate regularity using an autocorrelation function that extracts the periodicity of texture variations. Roughness is defined by variations within a small area. Hence, we estimate roughness using the changes in a small region. Directionality is defined by the orientation of edges, because image edges have a significant influence on visual perception. Therefore, directionality is measured using the main orientation of edges in different directions. Similarly, density is measured using the edge density. The results of the correlation analyses show significant correlations between the objective measures and subjective evaluations of these texture characteristics. Particularly, we propose a new method for estimating the fifth texture characteristic (understandability). The experimental results demonstrate that it is possible to evaluate the understandability of a texture using its names. In addition, these results show that the understandability of a texture can be estimated from two factors: (1) the maximum number of similar names belonging to a specific type and (2) the total number of types for a texture. The larger the number of similar names for a texture, the more understandable it is. The larger the number of types for a texture, the less understandable it is.

4.2 Feature Measures

4.2.1 Regularity

Regularity is the property of variations in a placement rule across a whole texture [67]. For a regular texture, the extraction of texture periodicity is very important in analyzing the regularity [78]. In this study, we employed an autocorrelation function (ACF) to extract the periodicity of the texture and measure the regularity that corresponds to human visual perception.

4.2.1.1 Autocorrelation function

The ACF is a a 2D function defined as

$$\Phi(\Delta x, \Delta y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \frac{p(x, y)p(x + \Delta x, y + \Delta y)}{MN}, \tag{4.1}$$

where $p(x, y)$ is the input signal and $p(x + \Delta x, y + \Delta y)$ is the shift vector of the input. M and N refer to the signal size in the horizontal and vertical directions, respectively. For an image, $p(x, y)$ denotes the gray value at the position of (x, y), and $p(x + \Delta x, y + \Delta y)$ denotes the neighborhood gray value at a position shift from (x, y) by a distance of Δx in the horizontal direction and Δy in the vertical direction. Usually, the ACF is normalized as

$$\phi(\Delta x, \Delta y) = \Phi(\Delta x, \Delta y) / \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} p(x, y)^2 / MN, \tag{4.2}$$

where the denominator in Eq. (4.2) is the maximum value of the ACF and represents the power of the image. After normalization, the ACF takes a maximum value of 1.0 at the origin [69].

4.2.1.2 Regularity extracted from the ACF

The ACF of an image can be used to detect repetitive patterns of texture elements. For regular textures, the autocorrelation function will have peaks and valleys. Peaks in the autocorrelation function of a regular-texture image characterize the texture periodicity [78].

From the calculated ACF, regularity can be extracted by measuring the amplitude of the maximum peak in the x and y directions [69](see Fig. 4.1). In Fig. 4.1, ϕ_1 related to the periodicity of the texture, and it reflects the regularity of the texture. Therefore, regularity is represented by ϕ_1. By comparing the heights of the maximum peaks in both directions, we found the maximum peak in the x direction.

By applying the ACF to the sample images, we calculated regularity by measuring the amplitude of the maximum peak.

4.2.1.3 Results

We compared the calculated values of regularity with the subjective values of regularity to verify the hypothesis that regularity calculated by the ACF relates to the perceived regularity. Figure 4.2 shows a high correlation (r = 0.840, $p < 0.01$) exists between the calculated regularity and that evaluated subjective.

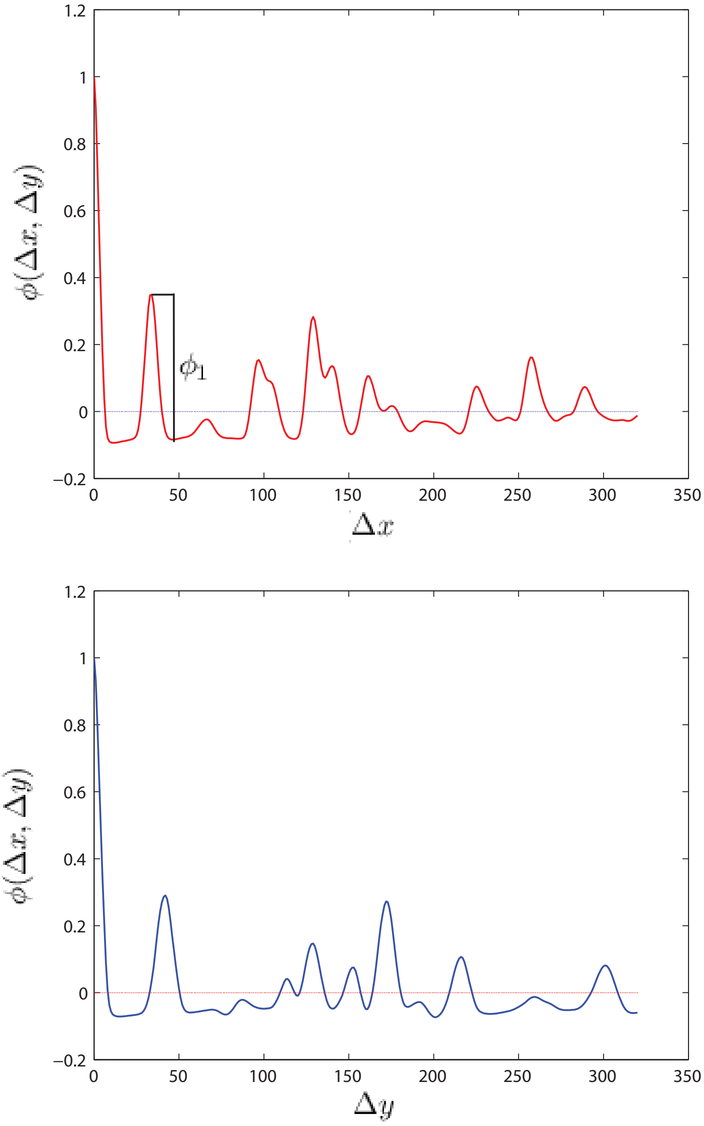


Figure 4.1: One-dimensional ACF along the x and y directional of texture d47 in Fig. 3.1.

4.2.2 Roughness

Roughness was originally meant for tactile textures, not for visual ones [67]. However, when we observe textures such as those shown in Fig. 4.3, we can identify the rougher texture. Visual roughness is affected by the texture's features. For a rough texture, the grey values change quickly in a local region. Conversely, for a smoother texture, the gray values change slowly in a local region.

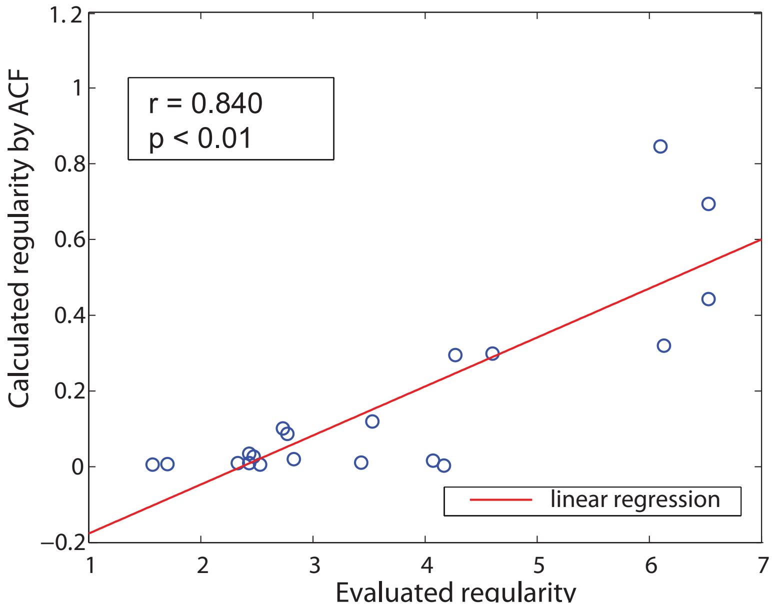


Figure 4.2: Relationship between the regularity calculated by the ACF and that evaluated subjectively.

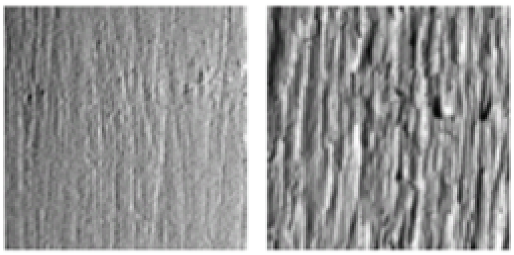


Figure 4.3: A smooth texture and a rough texture.

4.2.2.1 Roughness calculation

According to the results of our experiments on visual roughness, we emphasized the effect of local changes in the texture. Homogeneity refers to the closeness of distribution of elements in the gray level concurrence matrix (GLCM). It is used to measure local changes in the texture. The larger the homogeneity, the more smooth and flat the texture is. Hence, we use homogeneity to estimate the smoothness of a texture, shown as

$$H = \sum_{g_1} \sum_{g_2} \frac{p(g_1, g_2)}{(k + |g_1 - g_2|)}, \tag{4.3}$$

where H indicates the smoothness of a texture, $p(g_1, g_2)$ is the GLCM, and the effect of k is to avoid the denominator being equal to zero.

For the measure of roughness, we introduce Eq. (4.4). We suppose the smoothness of a white image is 1. For each texture, compared with the white image, it is visually rough. So the relative roughness can be estimated by

$$R = 1 - kH, \tag{4.4}$$

where R is the roughness of a texture. 1 is supposed as the value of smoothness of a white image.

The values of H and R are affected by the value of k. In our experiment, $k = 1$, 0.1, 0.01, and 0.001 yielded different values of correlation coefficients (r) between subjective roughness and calculated roughness, respectively 0.681, 0.718, 0.721, and 0.722. Note that the absolute value of r slowly increases as the k decreases. Finally, we fixed k equal to 0.001 as follows:

$$\Delta r = r_k - r_{k'} < 0.001, (k = k^{'}/10), \tag{4.5}$$

where r is the correlation coefficient, When Δr satisfied the above condition, we fixed the k value.

4.2.2.2 Results

We plotted the correlation between the calculated and evaluated values of roughness (Fig. 4.4). Note that the calculated roughness is related to the visual roughness estimation by the subjects.

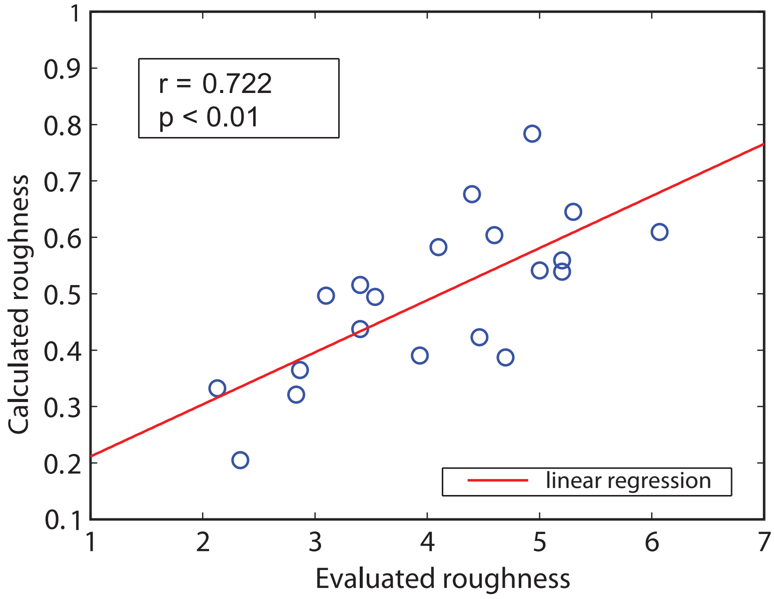


Figure 4.4: Correlation between the calculated and evaluated (subjective) values of roughness.

4.2.3 Directionality

For an image, the edges have a large influence on human visual perception and the orientation of the edges reflects the orientation of the textures. Subjects' visual judgments on directionality are vulnerable to the effects of linear edges.

4.2.3.1 Directionality calculation based on texture edges

The line-likeness and orientation of edges help us characterize the directionality of a texture. In this study, we applied line-likeness measurement on the edges of a texture to calculate the line-likeness of a texture, from which the directionality of the texture could be extracted. The detailed procedures follow:

a) Detect the edges of the texture image using the Canny algorithm.

b) Calculate the line-likeness using equation (6) [67].

$$F_{dir} = \sum_{i}^{n} \sum_{j}^{n} P_{Dd}(i, j) cos|(i - j) * 2\pi/n| / \sum_{i}^{n} \sum_{j}^{n} P_{Dd}(i, j), \tag{4.6}$$

where P_{Dd} is the $n \times n$ local directional co-occurrence matrix of points at distance d. This matrix is defined as the relative frequency with which two neighboring cells separated by a distance d along the edge direction occur on the image. Variables i and j are the direction codes in matrix P_{Dd}. In this experiment, 8 directions were included in matrix P_{Dd}, and $d = 6$ yielded the best correlation coefficient r.

c) After the line-likeness was calculated for each direction, the maximum value of the line-likeness was regarded as the directionality of the texture. That is because the maximum line-likeness reflects the existence of many lines in that orientation, which creates impression of the visual directionality.

4.2.3.2 Results

We compared and plotted the relationship between the calculated and evaluated (subjective) values of directionality (Fig. 4.5). The correlation ($r = 0.746$, $p < 0.01$) indicates that the line-likeness of edges well characterizes the directionality of a texture. It also indicates that the orientations of the edges have a large impression on the perceived visual directionality.

4.2.4 Density

The visual density of a texture refers to the density of visual information in the texture. Normally, the visual information of a texture is contained in the edges of primitives or regions of color change. Vaidyanathan and Lynch [79] hold the view that texture can be modeled as an arrangement of visible edges that in turn amalgamate into patterns. Hence, it is intuitive to relate the pixel number of edges in an image to the number and spatial arrangement of primitives. The pixel number of edges in a fixed-size region tells us how busy the region is.

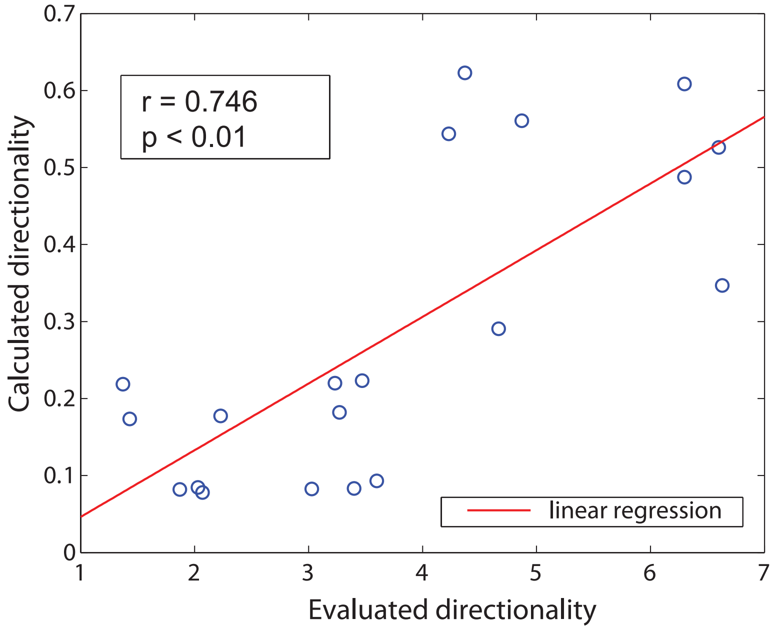


Figure 4.5: Relationship between the calculated and evaluated (subjective) values of directionality.

4.2.4.1 Density calculation

Respondents visually perceive a texture with many edges as dense. Therefore, we used the method of measuring the edge density to represent the visual texture density. The edge density can be determined by the ratio between the pixel number of the extracted edges and the pixel number of the whole texture as follows:

$$\rho_{den} = N_{edges}/N_{img}, \tag{4.7}$$

where ρ_{den} is the edge density of a texture. N_{edges} is the pixel number of the extracted edges and N_{img} is the pixel number of the whole texture.

In this study, we compared several edge detection algorithms (i.e., Roberts, Prewitt, Sobel, and Canny), and we finally adopted the Canny algorithm, because it gave the best results for edge detection.

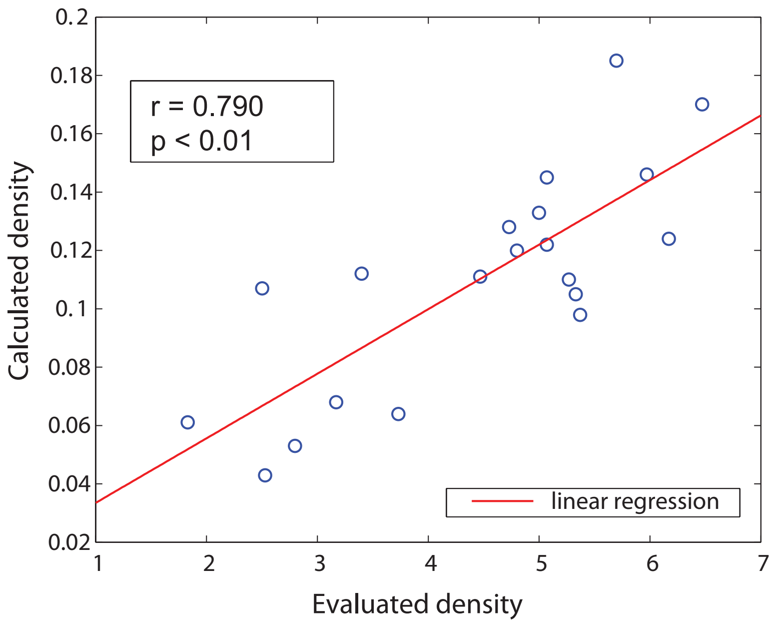


Figure 4.6: Relationship between the calculated edge density and the evaluated visual density.

4.2.4.2 Results

To verify whether the edge density can be used to represent the visual density of a texture, we compared and plotted the correlation between the edge density and the visual density (Fig. 4.6). The high correlation ($r = 0.790$, $p < 0.01$) between the edge density and human's visual density suggests that the edges in a texture significantly influence visual perception. More edges indicate more objects and primitives, and more visual information in the texture.

4.2.5 Understandability

It was indicated in Chap. 3 that the understandability of a texture is related to a human's prior knowledge and experience. So it is difficult to estimate the understandability of textures using mathematical methods. As already known, naming is a straightforward approach that reflects one's knowledge and experience. Therefore, we analyzed understandability using an experiment of naming textures. For each texture, we asked the respondents to give a name, and then we sorted the names into different types, including names that were similar, although not identical. We attempted to estimate the understandability of textures by the types and number of names.

4.2.5.1 Experiment: naming the textures

This experiment used the naming of textures to measure the respondents' understandability. In addition, the reaction time between the respondent viewing the texture and giving its name was also recorded.

The respondent was given a brief introduction, which included how to operate the webpage and the criteria for naming a texture. On the webpage, a button controlled the display of texture and enabled starting and stopping of the timer: once the button was pressed, after 500ms, the texture appeared and simultaneously the timer began. Once the respondent thought of a name for the texture, the button was pressed quickly; the texture disappeared and respondent immediately gave the name. The names given by the respondents were not constrained to any predetermined list, but were given according to respondents' visual perceptions and understanding. However, the names were required to be nouns rather than descriptive adjectives. For instance, the name of a flower-like texture should be "flower" or a word group including flower, not a description of something beautiful or nice. If the respondent could not give a possible name, he or she could answer, "I do not know." Before the experiment began, the respondent was required to use the test webpage to become familiar with its operation (The webpage used in the experiment is shown as Fig. 4.7).

After the introduction of the experiment, the respondent was asked to concentrate on the webpage. The respondent clicked the button to view the texture and clicked the button again when he or she thought of a name for the texture. In addition, after naming each texture sample, two questions were asked and the respondent recorded his or her answer on a 5-point Likert scale. The two questions were "How difficult was it to name this image?" with a score from 1 (very easy) to 5 (very difficult), and "How familiar were you with this

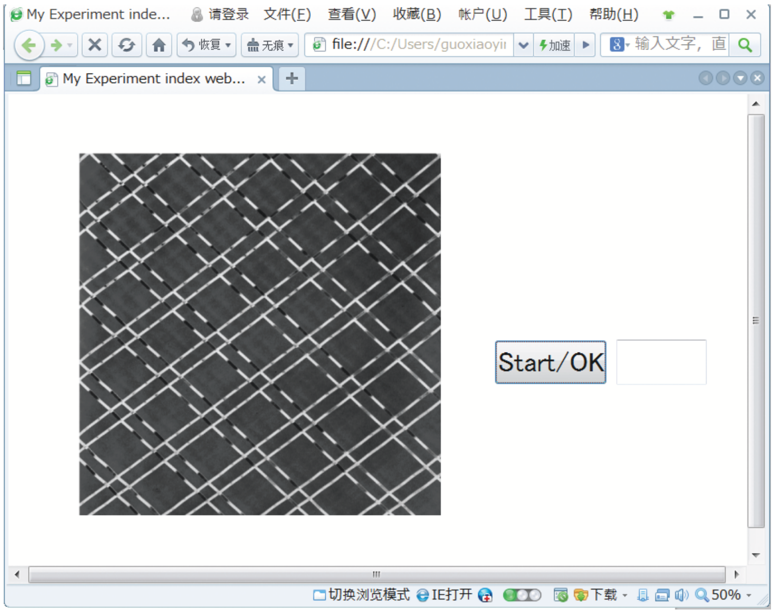


Figure 4.7: The webpage used in the experiment. When the stop button is pressed, the viewing time will show in the right box.

image?" with a score from 1 (very familiar) to 5 (very unfamiliar). All texture samples were tested one by one using the above procedure. Finally, the respondent was required to fill out a questionnaire. The questions in the questionnaire are shown in the Appendix A.

In this experiment, we obtained data on names of each texture, reaction time (RT), difficulty score of naming (DSN), familiarity score of the textures (FST), and the questionnaire answers. We categorized the names into different types, including names that were similar, although not identical. We also ranked the number of similar names belonging to a specific type, and chose the maximum number of different types (MNDT) as a salient property for each texture. The number of types (NT), MNDT, RT, DSN, and FST are shown in Table 4.1.

In the experiment, the respondents were required to fill out a questionnaire (shown in the Appendix A). The questionnaire was used to analyze the reasons that influenced respondents' understanding of the textures. We plotted the frequencies of their choices in Fig. 4.8. For question (1), the frequencies of choices A and C are equal (93.3%). However, for question (2), the frequency of choice C is much higher than that of choice A, which

Table 4.1: Experimental data from the experiment of naming the textures.

	NT	$MNDT$	$RT(s)$	DSN	FST
$d10$	7	17	5.91	2.93	2.93
$d13$	9	11	4.96	2.53	2.47
$d15$	7	17	5.90	2.90	2.90
$d20$	4	18	6.05	3.43	3.13
$d26$	1	30	2.33	1.13	1.07
$d27$	10	12	8.27	3.40	3.43
$d40$	3	21	5.61	2.50	2.37
$d42$	1	30	3.16	1.37	1.40
$d43$	8	15	5.89	2.73	2.67
$d47$	5	16	5.69	3.07	2.77
$d62$	6	10	5.93	3.07	3.03
$d64$	3	21	4.62	2.20	2.17
$d67$	6	17	7.23	3.00	3.07
$d72$	5	25	3.78	2.00	1.90
$d74$	6	13	6.26	2.53	2.67
$d88$	6	10	4.23	2.37	2.20
$d107$	9	10	6.67	3.97	3.70
$d109$	8	4	10.32	4.17	4.20
$d111$	13	9	6.08	3.13	3.13
$d112$	6	10	5.79	3.17	2.93

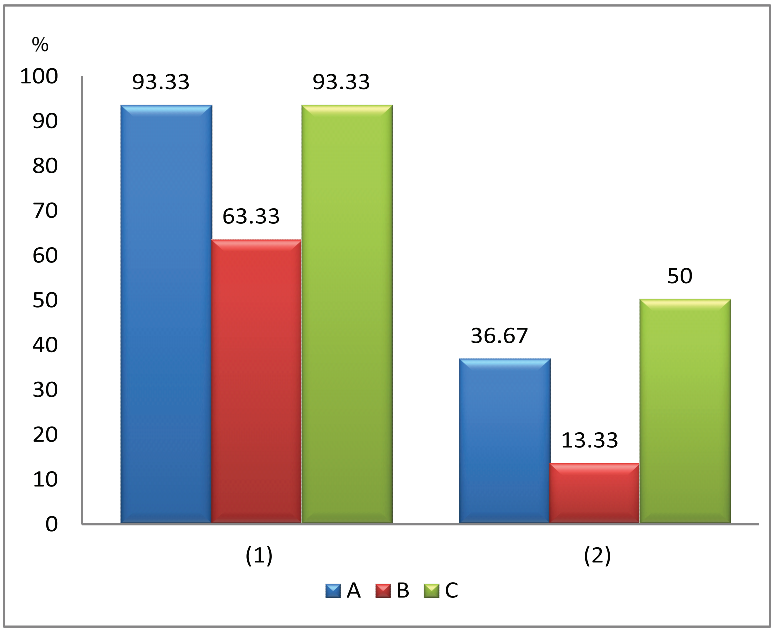


Figure 4.8: Frequencies of choices A, B and C in the questionnaire.

indicates that imagination has more influence on the respondents' understanding of the textures than having seen similar content before.

We demonstrated that the understandability of a texture was related to a human's prior knowledge and experience. Naming a texture is a good way to reflect one's prior knowledge and experiences. We hypothesized that the understandability of a texture could be estimated by the names given to it by respondents. For an easy understood texture, most respondents have a common understanding of it, and they will give the same or similar names for it. In contrast, for an abstract texture, respondents have a unique understanding of it, and they will give a wide variety of names for it. We show an two examples of naming for $d26$ and $d13$ like Fig. 4.9. For $d26$, all names given by the respondents are almost the same. Whereas, for $d13$, the names vary from each other. Some respondents even have no idea about this texture.

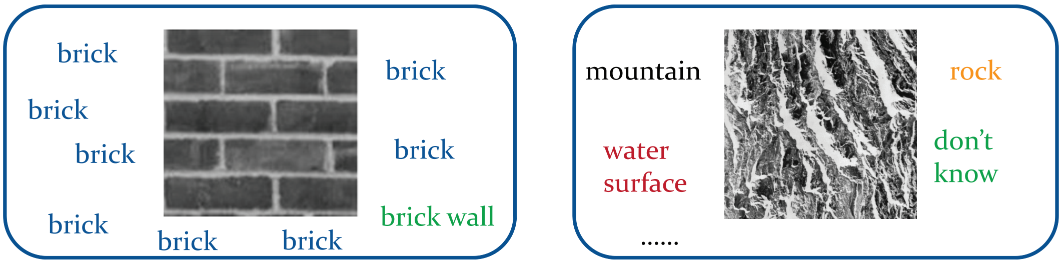


Figure 4.9: Two examples of naming textures. The center images are the experimental images $d26$ and $d13$, and the surrounding names are given by the respondents.

Table 4.2: Correlation coefficients between NT, MNDT, RT and understandability.

Correlations coefficients	*understandability*
NT	0.796
MNDT	−0.797
RT	0.710

4.2.5.2 Results

In the second experiment, we obtained data involving NT, MNDT, RT, DSN, FST, and so on. Correlation analyses were used to analyze the relationships between MNDT and understandability, NT and understandability, and RT and understandability. The correlation coefficients between them were respectively −0.797, 0.796, and 0.710, respectively, shown in Table 4.2. The analysis results showed both MNDT and NT were significantly related to understandability. The larger the number of similar names for a texture, the more understandable it is. The more types the texture evokes, the less understandable it is.

The analysis results indicated that it is possible to evaluate the understandability of a texture by the names given to it. In addition, it showed that understandability can be estimated from two factors: the maximum number of similar names belonging to a specific type, and the total number of types for a texture. The understanding of a texture is also related to the reaction time: it takes longer for us to think of names for less understandable textures.

4.3 Discussion

Visual complexity is a function of not only each individual characteristic but also of interactions between them, which is demonstrated by the correlation coefficients of perceptual characteristics in Table 3.4 in Chapter 3. The correlation between regularity and understandability is high ($r = -0.838$, $p < 0.01$). In general, textures characterized by regular placement are easy to understand, leading to a perception of less visual complexity. Similarly, the correlation between directionality and understandability is also very high ($r = -0.763$, $p < 0.01$). Interactions exist between roughness and understandability, regularity and roughness, and directionality and roughness. Therefore, it is suggested that respondents used a different combination of these characteristics while evaluating the visual complexity of a texture.

A set of objective methods was employed to measure the characteristics of a texture. For regular textures, the autocorrelation function will have peaks and valleys. As has been previously established, peaks in the autocorrelation function of a regular-texture image characterize the texture periodicity [69]. For the regularity of a texture, we adopted the ACF to extract the periodicity of the texture and measure the regularity corresponding to human visual perception. The validity of ACF analysis for determining the regularity of a texture was examined by comparing the calculated regularity with subjective scores of regularity. The high correlation ($r = 0.840$, $p < 0.01$) between them showed the validity of ACF analysis.

Edges have a significant influence on human visual perception [79]. Respondents perceive a texture with many edges as dense. Therefore, we used the method of measuring the edge density to represent the visual texture density. Similarly, the perception of directionality is easily affected by the orientations of the edges. Hence, we regarded the maximum line-likeness of edges in different directions as the main direction of the texture. The results showed that the calculation of the maximum line-likeness of edges represented the visual direction of the texture. Similarly, the correlation between objective methods and subjective evaluations was validated.

We measured visual roughness mathematically based on the gray changes in a small region. The more changes in the local region, the rougher the texture is. We compared the calculated roughness and subjective roughness. The result showed that the calculated roughness was related to the subjective roughness. However, human visual roughness is also influenced by the imagined tactile feeling of a texture, not just by its objective features. In other words, viewers use knowledge stored in memory to attribute meaning to an image of a texture [80]. When they look at a texture, they can imagine whether it will feel rough or smooth. If one regards the texture as the surface of a product, then he or she imagines how the surface would feel.

We introduced a new approach to estimate the understandability of a texture by naming it. In the experiment of naming the textures, all respondents were Chinese, which ensured that they all had the same cultural background. In addition, both DSN and FST were obtained from the respondents. As has been shown, an easy to name texture is familiar and easy to understand. In our experiment, the correlations between DSN and understandability ($r = 0.811$, $p < 0.01$) and FST and understandability ($r = 0.826$, $p < 0.01$) also verified this result. In addition, MNDT and CT were significantly related to understandability. Therefore, the understandability of a texture can be estimated from two aspects: the maximum number of similar names belonging to a specific type and the total number of types for a texture. The larger the number of similar names for a texture, the more understandable it is. The more types the texture evokes, the less understandable it is.

4.4 Summary

In this chapter, computational measures of the four characteristics (regularity, roughness, directionality, and density) were developed and correlated with human visual perceptions of them. For regularity, we adopted the autocorrelation function to extract the periodicity of a texture and measure the regularity. The result showed that the computational regularity was significantly correlated with the subjective regularity. For roughness, we measured the gray changes within a small region to mathematically calculate visual roughness. The comparison showed that the calculated roughness was related to the subjective roughness. The subjective perception of directionality was easily affected by edges. Hence, we regarded the maximum line-likeness of edges in different directions as the main direction of a texture. The results showed that the calculation of the maximum line-likeness of edges represented the subjective direction of a texture. For density, we employed the method of calculating the edge density. A comparison between the calculated density and subjective density indicated that the subjective density was vulnerable to the influence of edges.

In particular, we introduced a new method to measure humans' understandability of a texture. The experimental results showed that it is possible to evaluate the understandability of a texture by naming it. In addition, the results showed that understandability could be estimated from two factors of a texture: its maximum number of similar names belonging to a specific type and its total number of types. The larger the number of similar names for a texture, the more understandable it is. The more types the texture evokes, the less understandable it is.

Chapter 5

Estimation of Visual Complexity Based on Texture Features

5.1 Introduction

Understanding of visual perception of textures has a wide range of applications from computer science (image processing and image segmentation) to arts (the design of product surfaces, packages, carpets, and wallpapers) [81]. For instance, understanding the visual perception of textures is helpful in the process of image retrieval and helps artists choose a texture for artwork or a product that evokes a particular emotion.

Textures are often evaluated by visual simplicity/complexity. Although assessing visual complexity of a texture is a highly subjective task, a majority of humans with similar background may have the same visual complexity perception towards certain textures.

Perception of visual complexity in textures is very important for visual understanding and visual aesthetic. If we are able to measure and combine computational features of a texture to match visual complexity as judged by human subjects, then it would be the first step in being able to predict the effect of visual complexity of textures used in the product design and its effect on human visual aesthetic evaluation[82]. However, the measures of complexity mentioned in Chapter 2 are primarily on the basis of information theory, pattern and fuzzy theory, not sufficiently related to human visual impression of complexity. Intuitively, the visual complexity is affected by various factors perceived by humans. Motivated by this, we attempted to investigate visual complexity of textures from human visual perception.

As explained in section 2.4, kansei is a Japanese word with the meaning of feeling, impression, and emotion and so on. Kansei engineering is a method of connecting human sensibility with engineering applications. It can "measure" feelings and shows their relationships to certain objective properties. In this work, we employed the methods of kansei engineering to first obtain the subjective visual complexity evaluations, and then map the relationship between the objective texture characteristics with the subjective complexity. The key issues of modeling the visual complexity of textures are including three steps: 1) Identifying the texture characteristics that affect human visual perception of complexity. 2) Designing a set of computational methods for these texture characteristics. 3) Mapping these texture characteristics with visual complexity.

We have finished the first two steps in Chapter 3 and Chapter 4. In Chapter 3, we have identified that visual complexity of a texture is affected by five important texture characteristics: regularity, roughness, directionality, density and understandability. Moreover, a set of methods have been developed to measure these characteristics associated with visual complexity perception in Chapter 4. The measures obtained by these methods were highly related to the evaluations given by the respondents on each texture characteristic [83].

In this chapter, we mainly focus on the third step mentioned above. A new model of estimating subjective visual complexity of a texture based on texture characteristics is proposed. The structure of the proposed model is shown in Fig. 5.1. Five low-level features are extracted from the texture images, and the high-level of feeling (visual complexity), is acquired from the subject' evaluation. Multiple Linear Regression (MLR) is employed as a mapping function to map five texture characteristics and visual complexity. In order

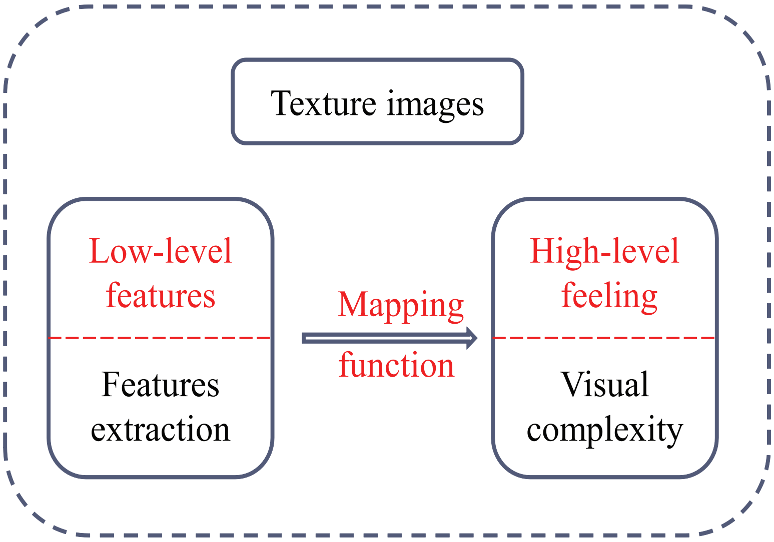


Figure 5.1: Mapping structure between low-level features and high-level feeling.

to validate this model, F-test and correlation analyses are performed to test the fitness and correlation between proposed model and subjective complexity evaluation by humans. The results show that the prediction from the proposed model is highly related to the human visual complexity perception. Furthermore, the results of the comparison show that the proposed method is efficient in predicting visual complexity than the traditional measurements of complexity (Shannon entropy) and the method in [11].

5.2 Modeling the Visual Complexity of Textures

It has been identified that five characteristics affect human visual complexity of a texture. And the relationship between each characteristic and visual complexity was analyzed in Chapter 3 by using correlation analysis (Person correlation). The results of the analysis are shown in Table 3.4. It was obviously shown that the visual complexity perception is highly linear related to these characteristics. Hence, in this paper, we assumed that the relationship between visual complexity and texture characteristics can be built by a mapping function of Multiple Linear Regression. The model of estimating visual complexity of texture is introduced in this section. Let us first learn about the Simple Linear Regression.

5.2.1 Simple Linear Regression

A linear regression model attempts to explain the relationship between two or more variables using a straight line [84]. Simple linear regression (SLR) is the simplest model among all regression models, and it is trying to "explain" the values of a variable y using the values of another variable x, these two variables being assumed to entertain a linear relation:

$$y = ax + b + \varepsilon(x), \tag{5.1}$$

where $\varepsilon(x)$ is a random "noise" that depends a priori on x.

5.2.2 Multiple Linear Regression

A linear regression model that contains more than one predictor variable is called a MLR model [84]. MLR [85] addresses just about the same problem with SLR, however, the difference is that, the "response variable" y is supposed to be explained not by just one variable x, but by several variables x_p. If we slightly change the foregoing notations, we assume that the linear relation between $f(x)$ and the x_p is as equation (5.2):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p + \varepsilon(x) = \beta X + \varepsilon(x), \tag{5.2}$$

where β is the coefficient matrix, X is the set of variables and $\varepsilon(x)$ is a random noise.

5.2.3 Least Square method

MLR is based on least squares (LS): the model is fit such that the sum-of-squares of differences of observed and predicted values is minimized. How to minimize? Denote by X the $N \times (P + 1)$ matrix with each row an input vector (with a 1 in the first position), and let y be the N-vector of outputs in the training set [86]. Then the residual sum-of-squares is written as

$$RSS(\beta) = (y - \beta X)^T (y - \beta X). \tag{5.3}$$

Differentiating with respect to β we obtain

$$\frac{\partial RSS}{\partial \beta} = -2X^T (y - X\beta). \tag{5.4}$$

Assuming that X has full column, and hence $X^T X$ is positive definite, we set the first derivative to zero

$$X^T (y - X\beta) = 0. \tag{5.5}$$

to obtain the unique solution

$$\hat{\beta} = (X^T X)^{-1} X^T y. \tag{5.6}$$

5.2.4 Subjective Visual Complexity of Textures

In this section, subjective complexity evaluations is represented. Subjective complexity evaluations are obtained from an experiment.

5.2.4.1 Experiment

In this experiment, we extended 20 images to 50 images that are from Brodaz database, shown as Fig. 5.2. Thirty respondents were asked to score complexity on a 7-point Likert scale by using their own knowledge and judgment. The 7-point scale ranged from 1 (very simple) to 7 (very complex). The experiment set-ups and experiment procedure are the same as the experiments in Chapter 3.

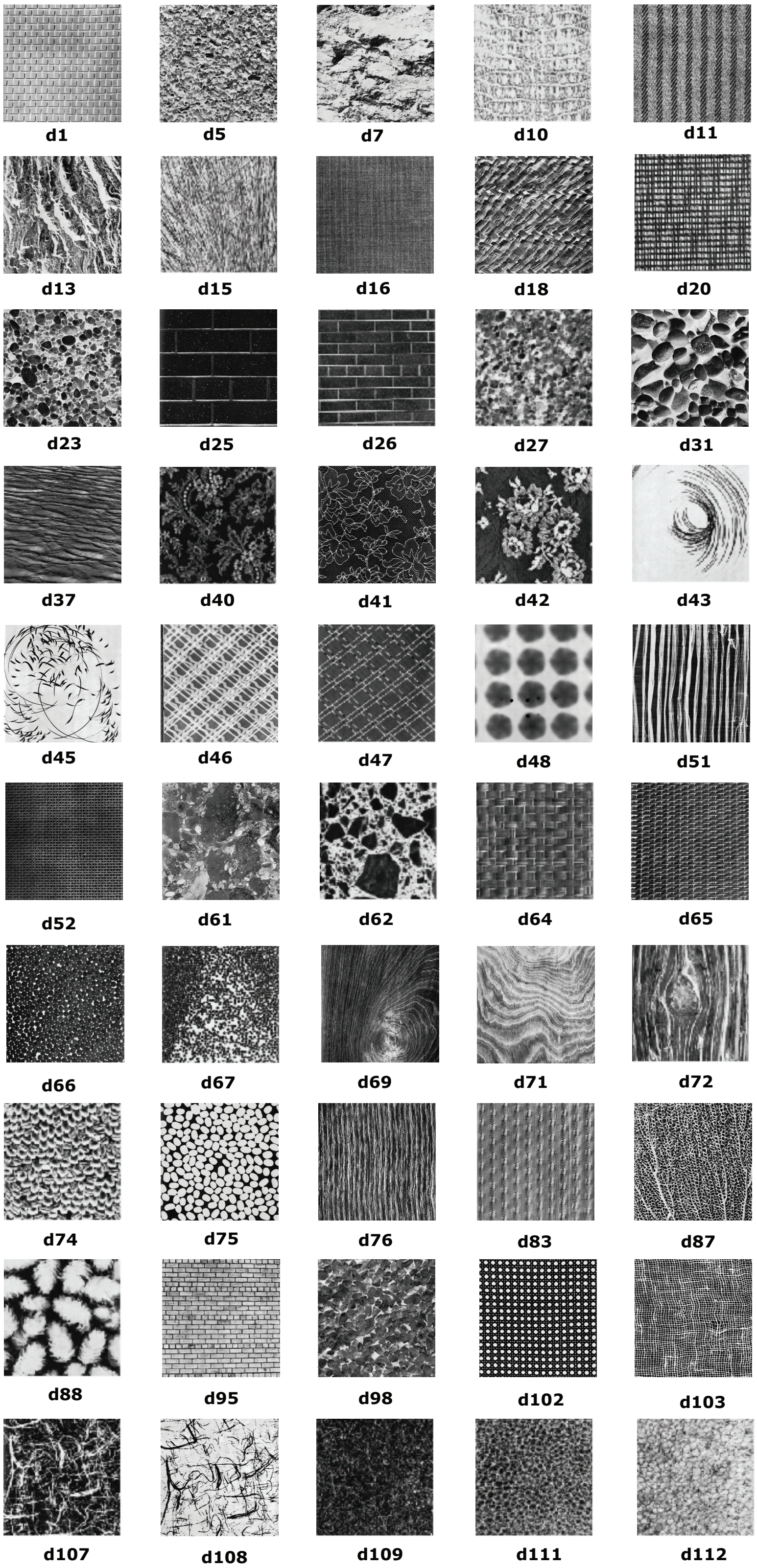


Figure 5.2: 50 texture images used in the experiment.

5.2.4.2 Consistency

In the experiment, there is a problem should be considered, that is the consistencies within respondents and between respondents. For the former one consistency, each image was shown twice so that we can ensure that the respondents can keep consistency in the evaluation of visual complexity. For the later one consistency, it was evaluated by calculating the standard deviation (SD) for each image. The range of SD is from 0.7 to 1.5, which also guarantees the consistency between respondents.

5.2.4.3 Results

By the experiment, we acquired the subjective evaluations (scores) for each texture. Considering that both the consistencies within respondents and between respondents in this experiment are guaranteed, hence, we calculate the average scores of visual complexity for 50 textures, which are shown in Table 5.1. In this table, it shows that the texture d26 is visually regarded as the simplest texture, and the textures d13 and d27 are visually perceived as the most complex textures.

5.2.5 Modeling visual complexity by MLR

MLR is used to map the visual complexity of a texture and five texture characteristics. In the Eq. 5.2, y is the vector of visual complexity of a texture. X is the texture feature vectors (including regularity, roughness, directionality, density and understandability). β is the coefficient matrix and $\varepsilon(x)$ is the error vector. The Least Square algorithm is adopted to estimate the matrix β, which yields parameter estimates such that the sum of squares of errors is minimized. Then visual complexity y can be calculated by Eq. (5.8) for a given texture image.

We employ 40 texture samples evaluated by 30 respondents as stimulated data to build

Table 5.1: Average scores of complexity of 50 texture samples.

TN	SV	TN	SV	TN	SV	TN	SV	TN	SV	TN	SV	TN	SV	TN	SV
$d1$	2.27	$d5$	5.40	$d7$	4.53	$d10$	4.27	$d11$	3.73	**d13**	**5.80**	$d15$	4.77	$d16$	3.00
$d18$	4.80	$d20$	4.07	$d23$	4.87	$d25$	2.13	**d26**	**1.43**	**d27**	**5.80**	$d31$	4.00	$d37$	4.47
$d40$	3.53	$d41$	3.07	$d42$	2.93	$d43$	3.07	$d45$	3.40	$d46$	3.13	$d47$	2.47	$d48$	2.07
$d51$	3.53	$d52$	3.60	$d61$	6.13	$d62$	4.37	$d64$	3.03	$d65$	4.07	$d66$	3.67	$d67$	3.30
$d69$	4.07	$d71$	4.53	$d72$	5.27	$d74$	4.10	$d75$	2.80	$d76$	4.60	$d83$	3.93	$d87$	5.47
$d88$	3.07	$d95$	2.33	$d98$	4.60	$d102$	2.87	$d103$	4.60	$d107$	5.50	$d108$	5.07	$d109$	5.47
$d111$	5.50	$d112$	5.63												

TN: *Texture Number*; SV: *Subjective Value of Visual Complexity* the model. By the LS algorithm in Matlab R2011b, we obtain the coefficient matrix β.

$$\beta = \begin{pmatrix} 3.957 \\ -0.255 \\ 0.788 \\ -0.048 \\ -0.482 \\ 0.656 \end{pmatrix}^T \tag{5.7}$$

Then, we get the model of visual complexity of a texture, which is shown as (5.8).

$$Y = 3.957 + \begin{pmatrix} -0.255 \\ 0.788 \\ -0.048 \\ -0.482 \\ 0.656 \end{pmatrix}^T \begin{pmatrix} regularity \\ roughness \\ directionality \\ density \\ understandability \end{pmatrix} \tag{5.8}$$

According to the Eq. (5.8), we calculate the complexity of the textures and compare them with the subjective complexity evaluated by the respondents. Figure 5.3 plots the results of subjective complexity and calculated complexity of 40 texture samples. Figure 5.4 shows the correlation between subjective complexity and calculated complexity of 40 texture samples.

5.2.6 Statistical test of the proposed model

The goodness-of-fit of this model is represented by its R^2 value (0.787). The value of R^2 indicated that the proposed model explains 78.7% of variance in subjective evaluation of visual complexity.

F-test is used to estimate the significance of the proposed model. The results are shown

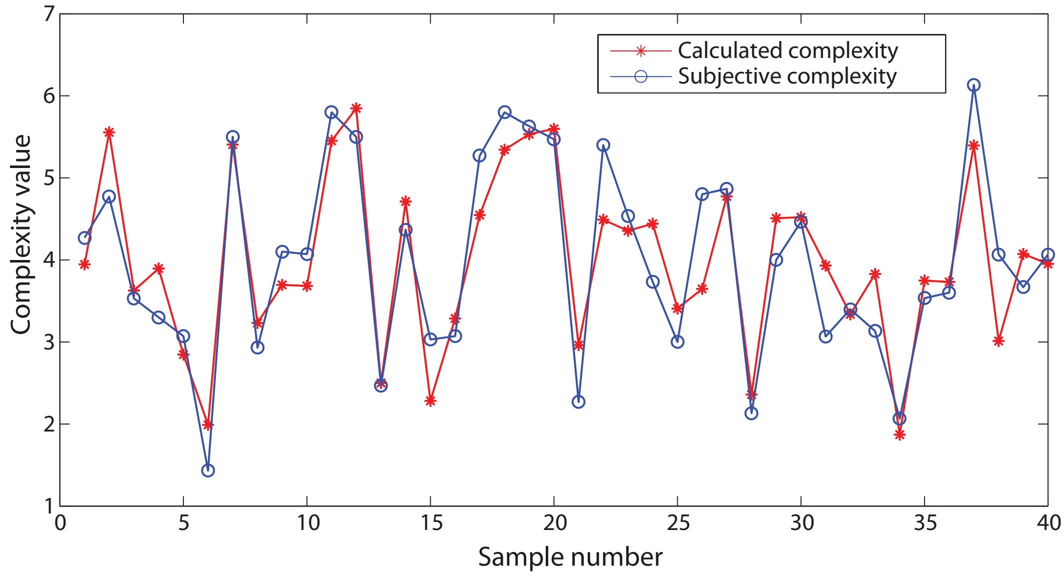


Figure 5.3: Subjective complexity and calculated complexity of 40 texture samples.

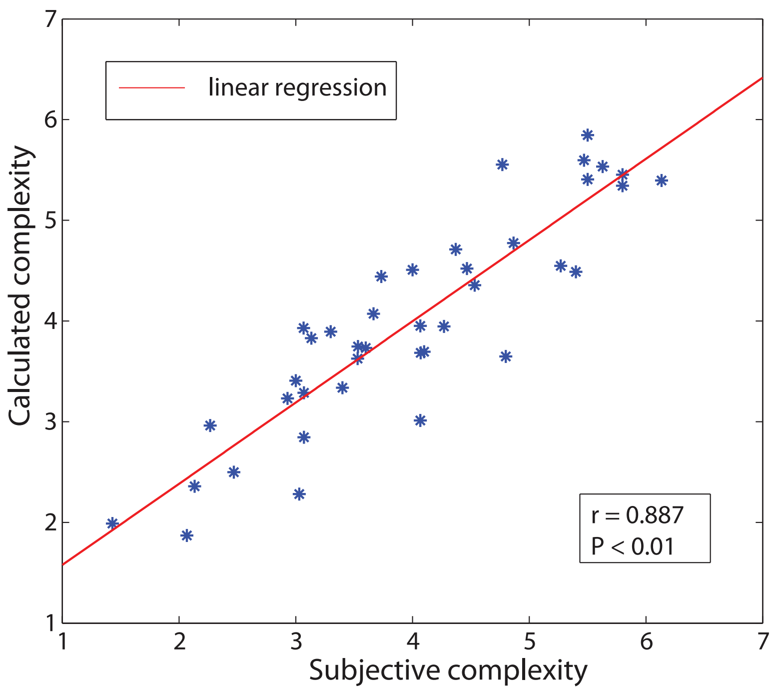


Figure 5.4: Correlation between subjective complexity and calculated complexity of modeling samples.

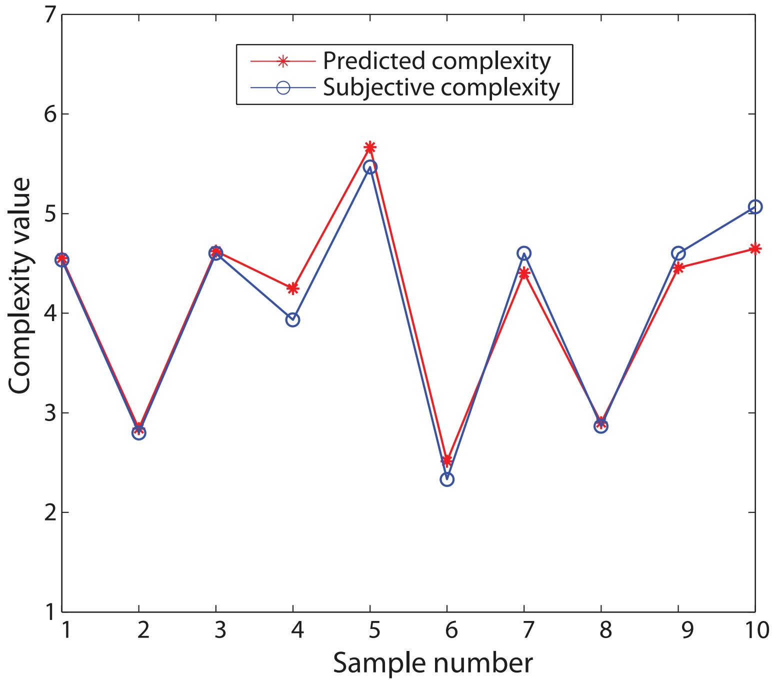


Figure 5.5: Subjective complexity and predicted complexity of testing samples.

in Table 5.2. In the table, it is obvious that $F > F_{0.01}(5, 34)$, which indicates the proposed model of visual complexity is very significant (P < 0.01).

5.2.7 Validation test of the prediction

Although a statistically significant regression formula of computing complexity has been found based on five texture characteristics, we have to test the validity of its prediction. We tested the prediction of the model using other 10 texture samples, and compared the predicted complexity with the subjective evaluations given by respondents. The values of predicted complexity and subjective complexity are plotted in Fig. 5.5. Their correlation coefficient is 0.951, which verifies that the proposed model is efficient and the predictions are highly close to the subjective evaluations, which is shown in Fig. 5.6.

Table 5.2: Results of Goodness of fit test and F test.

Source of Variation	*Degrees of Freedom*	*Sum of Squares*	*Mean Squares*	*F Statistic*	*P value*	R^2
Regression	5	41.245	8.249	25.150	**0.000**	**0.787**
Error	34	11.152	0.328			
Total	39	52.397				

$F_{0.01}(5, 34) = 3.610$

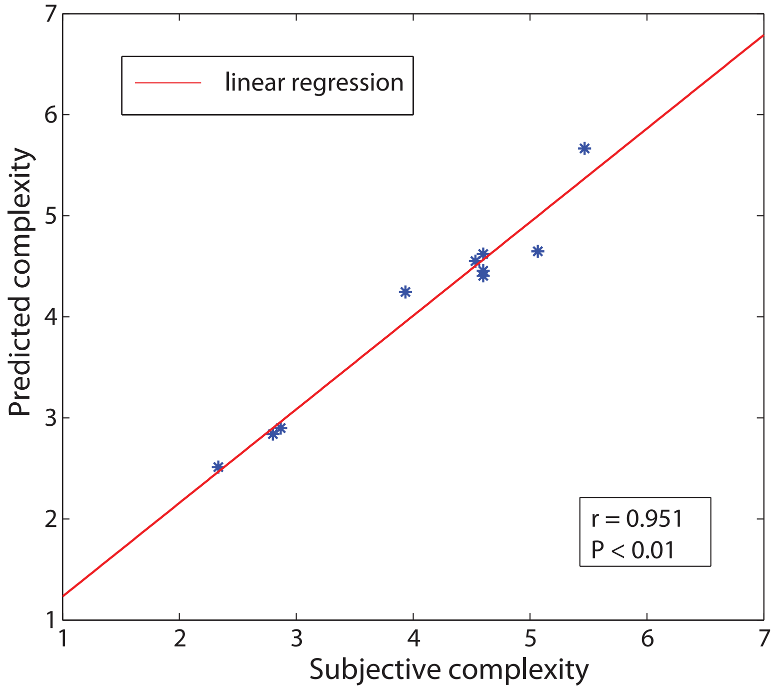


Figure 5.6: Correlation between subjective complexity and predicted complexity of testing samples.

Table 5.3: Comparisons between the proposed method and other related methods.

	Correlation coefficients
The proposed method	**0.951**
Shannon entropy	0.432
method in [11]	0.504

5.2.8 Comparison with other methods of measuring complexity

In the related works, we summarized that the methods of measuring complexity were primarily based on the information theory and fuzzy theory, not sufficiently related to visual perception. Compared with these measures, our proposed model is slightly better because visual complexity measured in our method is considering the human visual perception.

To analyze the performance of our method, we compared it with the classical complexity measure (Shannon Entropy) and the method in [11] (evaluating complexity by the size of compressed digital image files). We applied these methods to 10 testing images, and then evaluated the performance of each method. We respectively calculated the correlation coefficients between different calculated complexity values and subjective complexity evaluations. The comparison results are shown in Table 5.3. According to the results, it is obvious that the complexity calculated by our proposed method is highly close to the subjective complexity, which indicated that the proposed method can predict the visual complexity of a texture as that judged by the subjects.

5.2.9 Discussions

Nowadays, visual textures are used in many fields of applications, not only in image processing tasks, but also in the design of product surfaces and packages, shops or private living rooms (such as carpets and wallpapers). For customers, the visual impression is an essential factor for them to select the appropriate products; for designers, learning the user's visual preferences is also of importance. Hence, there is a need to analyze the human perception associated with visual textures.

In this chapter, we attempted to derive a computational model for an important aspect of visual perception: visual complexity. In Chapter 3, we identified and determined that there are five important characteristics which affect the human visual complexity: regularity, roughness, directionality, density and understandability. All the characteristics have a strong linear relationship with the visual complexity. In Chapter 4, we have designed a set of methods for these characteristics. In this chapter, we employed MLR to map the visual complexity of a texture with five texture characteristics.

LS algorithm was used to estimate the regression coefficients matrix β, which yielded parameter estimates such that the sum of squared errors are minimized. In the matrix β (Eq. 5.7), it is obvious that the coefficients of roughness and understandability are higher than the coefficients of the other features. This indicates that these two features make more contribution in predicting the visual complexity by using Eq. 5.8. Interestingly, the coefficient of directionality is very small. It is partly because the visual directionality is prone to be regarded as the regularity. Furthermore, the density is also of importance in predicting the visual complexity. This agrees with the situation that the more edges (patterns and elements) there are in a texture, the more complex it is.

Some statistical analyses methods were applied to estimate the significance of the regression model. Firstly, goodness-of-fit test was performed on the simulated data to test the fitness of the proposed model with human visual perception. The high value of goodness-of-fit ($R^2 = 0.787$) showed that the fitting was very good, and it also indicated that the results obtained through this model were well correlated with human visual perception. Secondly, F-test was employed to inspect the significance of the proposed model. The test results suggested that the proposed model was statistically significant ($P < 0.01$). Finally, correlation analysis was employed respectively on the predicted values and the subjective evaluations of simulated samples and test samples. The correlation coefficients are respectively 0.887 and 0.951, which shows the strong correlation between the predicted values and the subjective evaluations.

Many researches on measuring visual complexity have been done in the field of computer science (introduced in section 2.3). However, the primarily measures were based on information theory and the fuzzy pattern and so on, not sufficiently related to human visual impression of complexity. In fact, when a viewer looks at the images, the feeling of visual complexity is aroused by the visual factors (image features) in the images. Hence, to overcome the previous methods that not sufficiently consider visual perception, we propose a new way of measuring visual complexity of a texture using texture characteristics, on the basis of human visual perception of complexity. In order to evaluate the performance of the proposed method and the related works, we compared this method with Shannon entropy and the method in [11]. The comparison results were listed in Table 5.3. It was shown that the complexity predicted by our method is highly close to subjective perception.

5.3 Summary

In this chapter, we proposed a model to estimate the visual complexity of a texture while considering the human visual perception. The proposed model of visual complexity was built based on five texture characteristics, namely, regularity, roughness, directionality, density and understandability. MLR was used to map these low-level texture features with the high-level visual complexity perception. The results of statistical analysis showed that the proposed regression model can significantly predict the visual complexity perception of a texture. Compared with the method of Shannon entropy and the method in [11], the prediction from our method is more close to the subjective complexity perception.

This investigation contributes to the computation of visual complexity of a texture considering human visual perception. It aims to build the relationship between visual complexity perception of the human (high-level feeling, also called as Kansei) and the computational visual features (low-level features) extracted from the texture images. It will also facilitate image retrieval based on subjective feelings. This kind of relationship can be a complicated issue. Our method is not a full solution, just an attempt in this new and interesting Kansei research.

Part III

Modeling the Perception of Visual Complexity in Painting Images

Chapter 6

Objectively Estimating Visual Complexity of Painting Images

6.1 Introduction

Nowadays, the rapid developments of digital technologies and internet have changed the modern life a lot. It makes people have more opportunities to appreciate the paintings without going to museums [87]. More and more users select preferable paintings from the internet. If they consider selecting images only by visual feeling (e.g. aesthetic) instead of specific keywords (e.g. flowers), visual complexity has some information of composing the feeling [15, 16]. Hence, presenting an objective index of complexity that fits human feeling to the users is useful. Proposing an objective measure of complexity is not only useful for the emotion-based image retrieval, but also is possible to use for estimating watermarking capacity of images [22].

Assessing visual complexity is a highly subjective task. However, there is a natural intuition that a majority of people with the same background may have a global agreement on classifying visual complexity towards certain paintings. In this work, we overcome this challenge by introducing a machine learning method, aiming to classify the complexity of paintings into three classes: low complexity (LC), middle complexity (MC), and high complexity (HC).

Support Vector Machine (SVM) is a learning tool originated in modern statistical learning theory [88]. It uses the knowledge of geometry to exactly calculate the optimal separating hyperplane for classifying the samples in the feature space. As a result, SVM performs well in the classification problems.

In this chapter, we firstly conduct an experiment to analyze the factors that affect human visual complexity perception of paintings and obtain the subjective assessments of paintings, and then extract a group of image features to globally and locally represent these factors. These features are combined using an SVM to build the model of visual complexity assessment of paintings. Moreover, in order to look into the role that each feature plays, we conduct the feature selection step. Experimental results indicate that the prediction of complexity from the proposed work is highly close to the assessments given by the respondents. Compared with the conventional measure of complexity, our approach considers human visual perception and performs more efficiently in assessing visual complexity of painting images. Furthermore, we also extend the proposed method to architecture images, and apply it to the local complexity estimation of images.

6.1.1 Hypothesis

The methods mentioned in Chapter 2 evaluate the complexity of images mainly on the basis of information theory and the fuzzy theory, regardless of human visual impression of complexity. In this research, we attempt to compute the complexity of paintings from the point of human visual perception. Consequently, we assume that the perception of visual complexity is influenced by various characteristics (image features) intuitively perceived by humans.

In order to verify this assumption, we design a group of methods to extract image features that represent both the global and local characteristics of paintings. Inspiration for these features is from a questionnaire survey we conducted to identify the factors that affect human's complexity assessments of paintings. Furthermore, feature selection is conducted, by which we analyze the role that each image feature plays in assessing visual complexity. Then the selected features are combined by an SVM for classifying the visual complexity of a painting into three classes: low complexity, middle complexity, and high complexity. Moreover, we extend the proposed work to architecture images and testify its validity.

6.1.2 Our work

The objective of this work is to propose a computational measure of estimating visual complexity of paintings. In order to achieve this purpose, we conduct three steps:

1) Assessing the visual complexity of paintings and identifying the factors that affect human visual complexity by a psychophysical experiment.

2) Extracting a series of features to globally and locally represent these factors in a painting.

3) Using a machine learning method to build the relationship between the visual complexity perceived by humans and the features extracted from the paintings.

The framework is shown as Fig. 6.1. We will describe each step in the following sections.

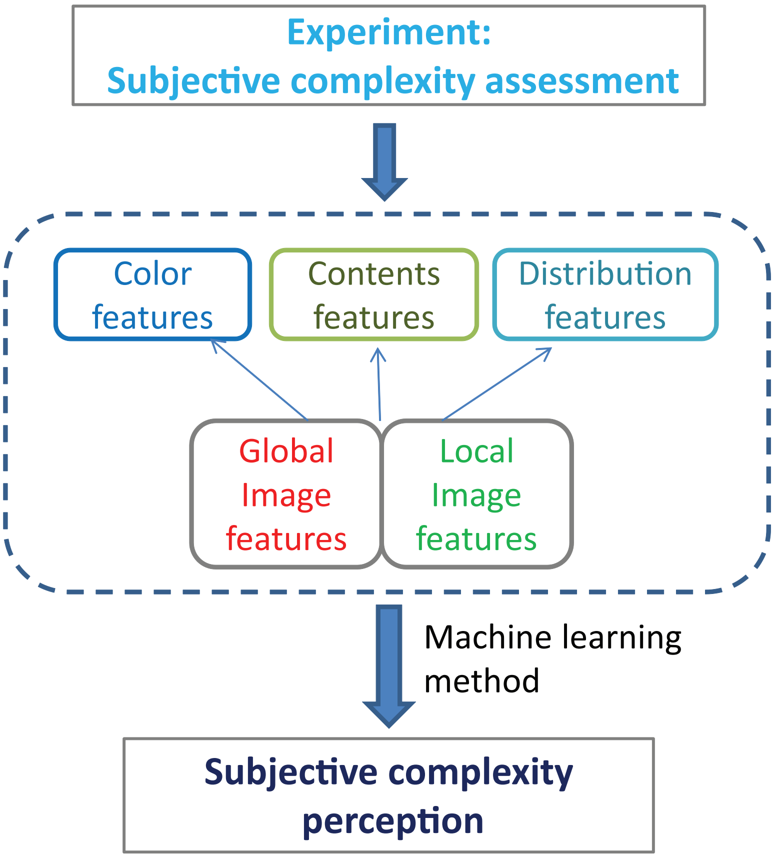


Figure 6.1: Framework of the proposed method.

6.2 Subjective Complexity Assessment

It is clear that the purpose of this work is to make a computer to assess the visual complexity of a painting as that evaluated by a majority of people. However, one challenge occurs when we conducted this research, which is that judging visual complexity of images is always considered to be subjective. Especially for evaluating the paintings, different people have different ideas towards the same painting. However, there is a natural intuition that a majority of people with same background may have a global agreement on classifying visual complexity towards certain paintings. Hence, one solution to this challenge is to classify complexity into three levels aiming to reducing the variance of ratings from human beings. Another one is to ignore other psychological aspects in human feelings, and just get the intuition feedback of visual complexity.

6.2.1 Experiments

6.2.1.1 Respondents

Thirty two respondents participated in the experiment. All of them were from Hiroshima University and with the bachelor's degree or above. None of the respondents in this experiment was in major of art or art-related. Their ages ranged from 21 to 30 years old. All respondents had normal or corrected-to-normal vision.

6.2.1.2 Apparatus and stimuli

Fifty digital images from the dataset of PaintingDb [89] are used in this experiment. All images were resized to the same height (300px) and randomly displayed on a 46-inch plasma display one by one. The respondents were required to sit at 2m from the screen. All respondents were asked whether they have ever view the paintings when they rated the painting images. For each painting image used in our experiment, no one respondent viewed any of the painting images before starting our experiment. This ensures that the subjective assessment are not related to the the respondents' preferences and painting's fame.

6.2.1.3 Procedure

The experiment procedure includes two parts. Part I is complexity-rating, and Part II is a questionnaire (The questionnaire is shown in Appendix A).

After the brief introduction of the experiment, Part I was done before Part II. In Part I, all images were displayed twice. On the first display, the respondent was required to view all images one by one with no time constraint. On the second display, the respondent was asked to score complexity on a 7-point Likert scale according to their perception. The 7-

point Likert scale ranged from 1 (very simple) to 7 (very complex). In Part II, we provided a questionnaire with the list of possible factors that affect complexity assessment (the options are listed in the left column of Table 6.1). The questionnaire is shown in Appendix A. The respondent was asked to select two factors which are the most important for them to assess the complexity of a painting.

6.2.2 Results

The subjective complexity assessments of 50 paintings were obtained in Part I from 32 respondents. From the assessments of complexity, we calculated the histogram of the scores over all images. The maximum labels in the histogram are "3" and "5", which are selected as the threshold for labeling images as "LC", "MC", and "HC". A painting image is marked as LC if its score is lower than 3. A painting image is marked as HC if its score is equal to or greater than 5. And a painting image in-between is marked as MC. Figure 6.2 shows some examples that marked as LC, MC, and HC.

From the questionnaire results in Part II, answers can be ranked according to the frequency of options mentioned by the respondents. We summarized the answers from the first question of the questionnaire. The results are shown in Table 6.1. We also plotted the frequency of options from two questions in a questionnaire, shown in Fig. 6.3.

Table 6.1: Results of survey in the experiment.

Options	*Frequency*
Distribution of compositions	**28**
Colors	**26**
Contents	**22**
Understandable or not	19
Color variation	18
Contrast	16
Familiarity	10
Symmetry	6

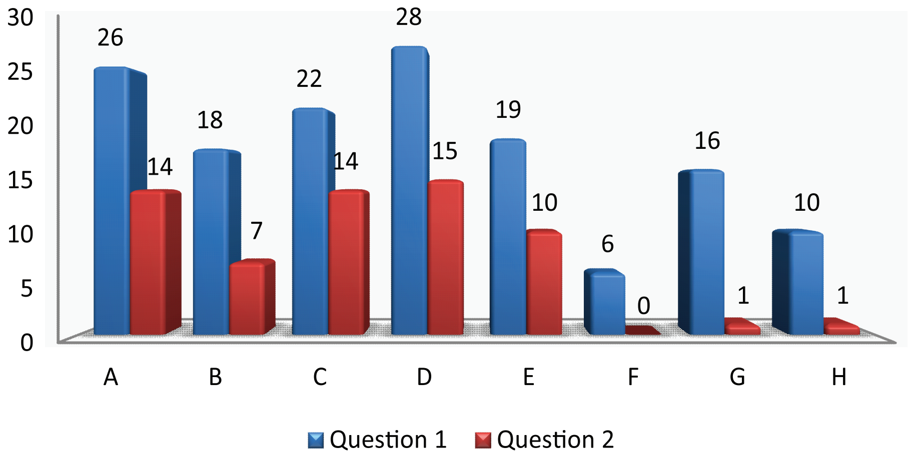


Figure 6.3: The frequencies of the options (from A to H) of two questions mentioned by 30 respondents in Part II (questionnaire part) in the experiment.

6.3 Feature Extractions

The purpose of this research is to build a method to quantitatively evaluate the visual complexity of a painting based on its features. Hence, extracting the features is a vital step in this research.

In general, we extract the features mainly concerning three factors identified in the experiment of subjective complexity assessment: distributions, colors, and contents (shown as Fig. 6.4). The meanings of these factors are discussed. "Distribution of composition" includes both the global organization of a painting and the distribution of local parts. Distribution features are analyzed though the characteristics and the spatial relationships between the segments in the painting. "Colors" refer to the palette of the artist, is very important [87], including color variances, color contrast between each parts in the painting. "Contents" represent the objects and element parts in a painting. More objects there are in the painting, more complexity the painting is. Content features are extracted by analysis the edges and interests of points in the painting. Synthesizing the results in the experiment and the common sense in art or the intuition, we extract a series of features to represent the above three factors, and then evaluate whether these features are useful or not.

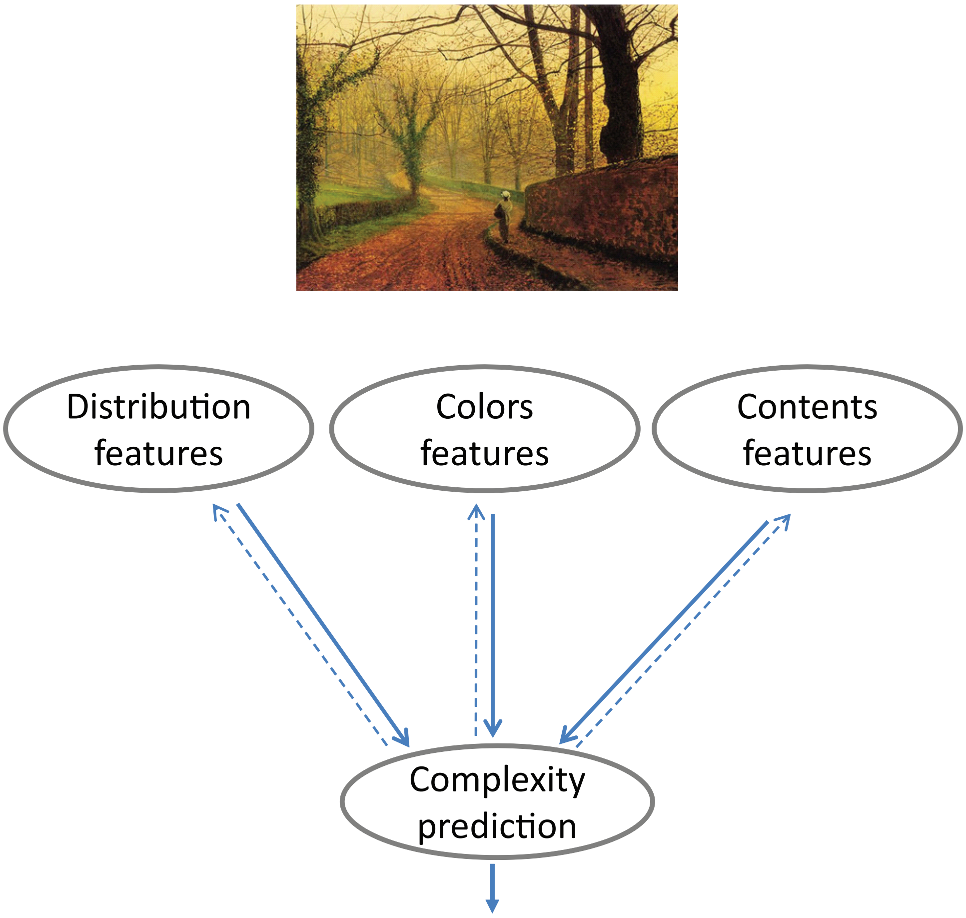


Figure 6.4: The framework of composing image features for paining complexity assessment.

All these features are separated into two categories: global features and local features. Global features refer to the characteristics of the first impression when human beings see a painting. On the other hand, the local features reflect the regional information of this painting.

6.3.1 Global features

As we all know, no matter in art or in the daily life, it turns out that when viewing something, people firstly get a holistic impression of it and then go into segments and details [90]. Hence, the global features influence the human's first perception of a painting. Each global feature is shown and explained as follows.

6.3.1.1 Color features

Color is defined as “that characteristic of a visible object or light source by which an observer may distinguish differences between two structure-free fields of the same size and same shape, such as may be caused by differences in the spectral compositions of the light concerned in the observation” [91]. Just as in [20], spectral power exists in the physical word, but color exists in human eyes and brain. Color is an inherent subjective perception when human view the world.

Color complexity measure Colors are the basic elements of a painting. If the color in a painting is complex, the painting is also of complexity. To measure the complexity of the color in a painting, we employed the Color Complexity Measure (CCM) in [92]. CCM is defined as follows

$$\psi(i, j) = \int \int G_\alpha(\| c(x, y) - \bar{c} \|) dx dy, \tag{6.1}$$

where G_α denotes the Gaussian weighting function and $\bar{c}$ is an average color value within a local window size $\Omega_{(i,j)}$ centered at (i, j). $\| \, \|$ denotes the color difference measure. Large CCM value reflects that the color variation in the local neighbor is high and small CCM value means that the pixel is located in the homogenous region.

The average color value within a local window is calculated by

$$\bar{c} = \frac{1}{N} \sum_{x,y \in \Omega_{(i,j)}} c(x, y), \tag{6.2}$$

where N is the number of pixels in the local window $\Omega_{(i,j)}$.

The color difference measure is important to represent the human visual perception of colors. Generally, the color difference is evaluated using the distance between two color points in a color space. Euclidean distance is usually used for measuring distances. In this part, we adopted the CIELab color space which describes all the colors visible to the human eye. But in this color space, that small Euclidean distance between two color points is proportional to the difference that human visual system perceives. It was identified in [92] that a larger Euclidean distance has no meaning but only large difference in human visual system. Hence, we employed the color difference proposed in [92]. It is defined as

$$D(c(i,j), c(x,y)) = 1 - exp\left[-\frac{E(c(i,j), c(x,y))}{\gamma}\right], \tag{6.3}$$

where γ is the normalized factor and $E(c(i,j), c(x,y))$ is the Euclidean distance in CIELab color space.

$$\begin{aligned} &E(c(i,j), c(x,y)) = \\ &\sqrt{(L_{ij} - L_{xy})^2 + (a_{ij} - a_{xy})^2 + (b_{ij} - b_{xy})^2}, \end{aligned} \tag{6.4}$$

Hence, we employed Eq. (6.1) to calculate the color complexity of local neighbor region centered at pixel of (i, j). For the whole image, we calculate the mean CCM of all pixels.

A Gaussian pyramid is a series of images, which are regenerated by Gaussian smoothing and downsampling from the original image. Figure 6.5 shows the model of Gaussian pyramid. The first level of the pyramid, called level 1, is usually defined as the original image. The second level is called level 2, it is generated from level 1 by firstly conducting Gaussian smoothing and then downdampling. The level $k + 1$ is generated from level k using the same method. The higher the level, the smaller the image size.

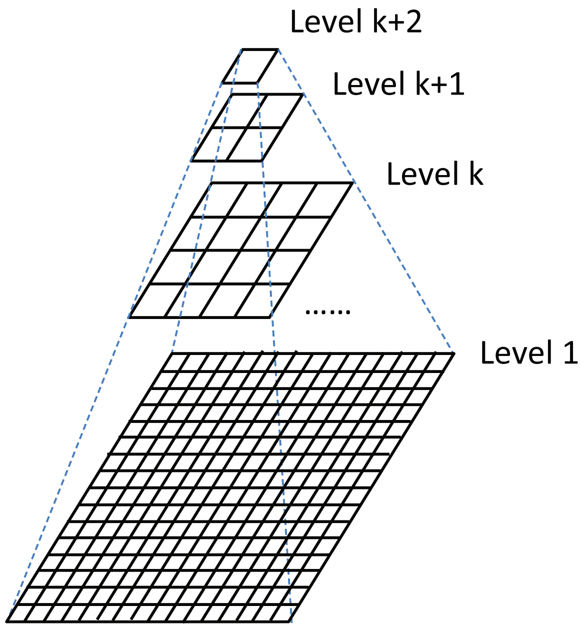


Figure 6.5: The model of Gaussian pyramid.

The Gaussian pyramid consists of low-pass filtered, reduced density (i.e., downsampled) images of the preceding level of the pyramid. Some features hidden in this resolution are extracted in another resolution. In this study, we applied Gaussian pyramid on the painting image and calculated four features respectively on the four levels of pyramid.

$$
\begin{aligned}
& f_i = CCM \mid G_j(x, y), \\
& \qquad i = 1, 2, 3, 4. \\
& \qquad j = 0, 1, 2, 3.
\end{aligned} \tag{6.5}
$$

where i is the feature in different level (j) of pyramid. Hence, four CCM features are extracted.

$f_1 - f_4$: Color complexity in four levels of image pyramid are extracted. Colors are basic elements of a painting. If the colors in a painting are complex, the painting is also of complexity. To measure the complexity of the colors in a painting, we employed the method of Color Complexity Measure (CCM) [92]. The image features extracted from the different levels of image pyramids have different effects for image classification of visual complexity. Therefore, we applied Gaussian pyramid on the painting images and calculated four features respectively on the four levels of pyramid.

Basic color features In colorimetry, the Munsell color system [94] is a color space that specifies colors based on three color dimensions: hue, value (lightness), and chroma (color purity). It is based on rigorous measurements of human subjects visual responses to color. In fact every color has three sides to its personality: hue, value, and chroma. A painter trying to mix a color on their palette to accurately match a color in their subject needs to consider all of these.

Hue and lightness in HSL (hue, saturation and lightness) can be equal to hue and value in Munsell color system. Saturation and chroma both represent the purity of the color, but there is a little of difference between them. Different from saturation, chroma doesn't have upper limit and the maxima of chroma for different hues are different [87]. However, saturation is the most closest to chroma in color spaces. Hence, we use HSL color space to instead of Munsell color system.

1. Hue

   Hue is the most obvious characteristic of a color [95]. In artist sense, the average hue and saturation more or less reflect the colorful keynote of that painting, relative the viewer's perception [87]. The average hue is calculated by

   $$f_5 = \frac{1}{MN} \sum_n \sum_m I_H(m, n), \tag{6.6}$$

   where M and N are the number of rows and columns of the image, $I_H(m, n)$ is the hue value at the pixel (m, n).

2. Saturation

   Saturation, also called chroma, is the purity of a color [95]. The saturation of a color is a measure of how intense it is. High saturation colors look rich and full. Low saturation colors look dull and grayish. The saturation of a painting somehow influence the human's feeling evoked by the painting. The average saturation is calculated by

$$f_6 = \frac{1}{MN} \sum_n \sum_m I_s(m, n), \tag{6.7}$$

where M and N are the number of rows and columns of the image, $I_H(m, n)$ is the saturation value at the pixel (m, n).

3. Brightness

   Brightness is also called value in Munsell color system. It is a measure of how light or dark a color is in the whole painting, and it reflects the tone of a painting. The average brightness of a painting can be calculated as:

$$f_7 = \frac{1}{MN} \sum_n \sum_m L(m, n), \tag{6.8}$$

   where M and N are the number of rows and columns of the image, $L(m, n) = (I_R(m, n) + I_G(m, n) + I_B(m, n))/3$.

$f_5 - f_7$: The average hue (f_5), saturation (f_6) and brightness (f_7) of a painting are exacted based on HSL color space. Hue is the most obvious characteristic of a color [95]. Saturation measures the intensity of a color. Lightness reflects the tone of a painting. These features vitally affect human visual complexity perception of paintings.

6.3.1.2 Content features

Points of interest Symmetry plays a revelent role in perception problems [96]. It is an interesting property in detecting the points of interest. The more the points of interest in an image, the more complex the image is perceived to be [9]. Discrete symmetry transform (DST) is an algorithm to measure the local symmetry, which searches the area of interest in active vision. In this study, we employed DST method to extract the points of interest in the painting images. DST computes local symmetry of an image based on a measure of axial moments of a body around it center of gravity. It is defined as [9]

$$DSTd_{i,j} = (1 - std_k(T^k_{i,j}))E_{i,j}, \tag{6.9}$$

where std returns the standard deviation, for $k = 0, ..., n-1$, of the first order moment relative to an axis with orientation $\alpha_k = k\pi/n$.

$$T^k_{i,j} = \sum_{(x,y)\in C_r} \mid (x-i)sin\alpha_k - (y-j)cos\alpha_k \mid g_{x,y}, \tag{6.10}$$

where C_r is the disk center at (i, j) of radius r. $E_{i,j}$ measure the local smoothness of the image.

$$E_{i,j} = \sum_{(x,y)\in\delta C_r (s,t)\in\delta C_{r+1}} \mid g_{x,y} - g_{s,t} \mid, |x-s| + |y-t| = 1, \tag{6.11}$$

where δC_r stands for the circular edge of C_r. If the image is locally flat, $E_{i,j} = 0$.

Points of interest (P) is extract by the following way,

$$P_{ij} = \begin{cases} DST_{ij}, & DST_{ij} \geq \mu + 2\sigma \\ 0 & otherwise \end{cases} \tag{6.12}$$

where μ and σ are respectively the mean and standard deviation of DST values calculated from the whole image.

By the above DST method, we extracted the points of interest in the painting images.

$$f_8 = P, \tag{6.13}$$

f_8: Symmetry plays a relevant role in perception problems [96]. It is an interesting property in detecting the points on interest. The more the points of interest is in an image, the more complex the image is perceived to be [9].

Edge density An image with more edges contains more objects, which will result in a greater perceived complexity [9]. The edge density can be determined by the ratio between the pixel number of the extracted edges and the pixel number of the whole image as follows:

$$f_9 = N_{edges}/N_{img}, \tag{6.14}$$

where f_9 is the edge density of an image. N_{edges} is the pixel number of the extracted edges and N_{img} is the pixel number of the whole image. In this part, we used Canny algorithm to extract the edges.

f_9: Edge density of a painting. The edge density can be determined by the ratio between the pixel number of the extracted edges and the pixel numbers of the whole image.

6.3.2 Local features

Local features represent the detail information of the paintings, which may be more attractive for the viewers' deep viewing. To extract the local features in the painting, we firstly segment the image into parts, and then analyze the characteristics in segments.

6.3.2.1 Image segmentation

In this study, an initial segment is firstly required to partition the image into small regions for merging. By comparison with the method of watershed, we chose the method of mean shift for initial segment because it creates less over segmentation (shown in Fig 6.6 (b) and (c)). We used a free software, EDISON System [97], to obtain the initial segmentation map. After the initial segmentation, the image is subdivided into many small regions.

In human visual perception, some regions with similar color or spatially adjacency should be merged into one region. Hence we need to represent these regions by some feature descriptors and define a rule for region merging. Here, we employed the method of color histogram similarity [98] to calculate the similarity between two adjacent regions. Each color histogram is quantized into 16 levels and then total 4096 bins in each region. The color histogram similarity ρ is calculated between two adjacent regions (regions P and Q) using the Bhattacharyya coefficient. The Bhattacharyya coefficient is an approximate measurement of the amount of overlap between two statistical samples. The coefficient can be used to determine the relative closeness of the two samples. It is defined as:

$$\rho(P, Q) = \sum_{u=1}^{4096} \sqrt{Hist_P^u * Hist_Q^u}, \tag{6.15}$$

where $Hist_P$ and $Hist_Q$ are the normalized histogram of adjacent regions P and Q, and u means the uth bin of them. The higher the Bhattacharyya coefficient between P and Q is, the higher the similarity between them is. After calculating the similarity between two neighbor regions, we used the Region Adjacency Graph to store the similarity of the pair of regions.

Region Adjacency Graph: Region Adjacency Graph (RAG) is a way of "spatially" representing the image. The method to represent the RAG consists of associating a vertex at each region and an edge at each pair of adjacent regions [99]. Fig. 6.6 (b) give the represent of the region adjacency graph. If the regions P and Q are adjacent in Fig. 6.6 (a), there is an edge connecting the regions P and Q; likewise regions 1 and 2 are adjacent, hence, there is an edge between them. $S_{P,Q}$ shows the similarity of the regions P and Q.

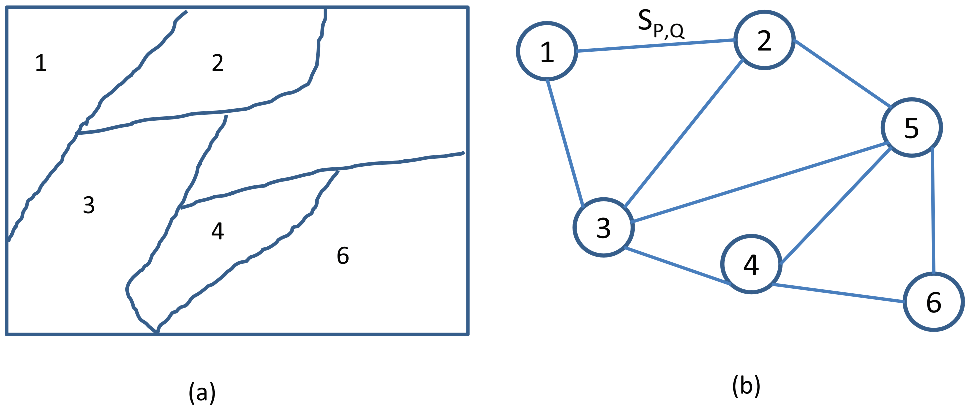


Figure 6.6: Representation of RAG: spatial view of the image. (a) Segmented regions. (b) Region Adjacency Graph.

The RAG can be used to merge adjacent regions that have a color distribution sufficiently closed. In this work, the merging rule is defined as: if the similarity (Bhattacharyya coefficient) between adjacent regions P and Q is the maximal one among all the similarities, we will merge the regions P and Q together. Following this merging rule, we got the finial segmentation map. In this experiment, we iterate twice for merging the adjacent regions. The initial mean shift segmentation and finial segmentation images are shown in Fig. 6.7.

(a)

(b)

(c)

(d)

Figure 6.7: (a) One example painting image in our dataset. (b)Initial watershed segmentation. (c) Initial mean shift segmentation. (d) The finial segmentation after region merging.

It can be easily understood that human vision is sensitive to the large segments in the images. Hence, in this experiment, we mainly consider the visual influences of the first largest segment (FLS) and the second largest segment (SLS). The local features are extracted based on the FLS and the SLS.

6.3.2.2 Color features

Color features are important not only for global impression, but also for the local details. In this part, we represent the local color features using the average hue, saturation and lightness for the top two largest segments and color contrast between the largest segments and their neighbor segments.

Hue, saturation and lightness of the top two largest segments Totally 6 features are calculated based on HSL color space in this part, shown as below. The hue, saturation and lightness are calculated as:

$$f_{9+i} = \frac{1}{Area_{R_i}} \sum\nolimits_{(m,n)\in R_i} I_H(m,n), i = 1,2 \tag{6.16}$$

$$f_{11+i} = \frac{1}{Area_{R_i}} \sum\nolimits_{(m,n)\in R_i} I_S(m,n), i = 1,2 \tag{6.17}$$

$$f_{13+i} = \frac{1}{Area_{R_i}} \sum\nolimits_{(m,n)\in R_i} I_L(m,n), i = 1,2 \tag{6.18}$$

where $Area_{R_i}$ is the area of the segment R_i. $I_H(m,n)$, $I_S(m,n)$ and $I_L(m,n)$ are the hue value, saturation value and lightness value of the pixel (m,n).

$f_{10} - f_{15}$: Hue (f_{10}), saturation (f_{12}) and lightness (f_{14}) of the FLS. Hue (f_{11}), saturation (f_{13}) and lightness (f_{15}) of the SLS.

Color contrast between segments The largest segment usually attracts most human visual attention. Hence, the color contrast between the largest segment and its neighbor segments may have some visual influence on the perception of visual complexity. In this part, we analyze the contrast of hue, saturation and lightness between the largest segment and its neighbor segments. The calculations are listed as follows.

Hue, saturation and lightness contrast of the largest segment are calculated as:

$$f_{16} = max\left|H_{largest} - H_i\right|, i \in \Omega_{nei} \tag{6.19}$$

$$f_{17} = max\left|S_{largest} - S_i\right|, i \in \Omega_{nei} \tag{6.20}$$

$$f_{18} = max\left|L_{largest} - L_i\right|, i \in \Omega_{nei} \tag{6.21}$$

where Ω_{nei} is the set of the neighbor segments around the largest segment. $H_{largest}$ is the hue value of the largest segment, and H_i is the hue value of the *ith* neighbor segment. $S_{largest}$ is the saturation value of the largest segment, and S_i is the saturation value of the *ith* neighbor segment. $L_{largest}$ is the lightness value of the largest segment, and L_i is the lightness value of the *ith* neighbor segment.

$f_{16} - f_{18}$: The contrast of hue (f_{16}), saturation (f_{17}) and lightness (f_{18}) between the FLS and its neighbor segments.

6.3.2.3 Distributions features

In the experiment of visual complexity assessment (Section 6.2), it is found that distribution of the components in the painting image is one of the important aspects that affect visual complexity perception. This aspect refers to the distribution of the segments, the number of all segments and the shapes of the major segments.

Number of all segments In the image segmentation procedure, the painting image is merged according to the color similarity, and it is finally segmented into small regions. Usually, the number of all segments in the painting image has a positive relationship with its complexity. The larger the number of all segments is, the more complex the image is.

The number of all segments in whole image is

$$f_{19} = N_R, \tag{6.22}$$

where N_R refers to the number of all segments in the image.

Areas of top two largest segments The area of the two top largest segments will also influence human visual complexity perception. The greater the area of the FLS is, the more homogenous the region is. This will create a gentle visual perception.

$$f_{19+i} = Area_{R_i}, i = 1, 2 \tag{6.23}$$

where $Area_{R_i}$ refer to the areas of the top two largest segments.

Shape complexity of first two largest segment The contours of the regions with different shapes (rectangle, circle or fractal) will arouse different visual perception. In this study, we use Perimetric Complexity [100] to measure the shape complexity of the top two largest segments.

Shape complexity of the first two largest segments:

$$f_{21+i} = \frac{P_{R_i}^2}{4\pi A_{R_i}}, i = 1, 2 \tag{6.24}$$

where P_{R_i} and A_{R_i} are the perimeter and area of the largest region and the second largest region.

In all, we extracted 23 global and local features to represent the factors identified in section 6.2. They are shown in Table 6.2.

Table 6.2: All features extracted from the painting images (including global features and local features).

Features	*Definitions*	*Groups*
f_1	*Color complexity measure of* 4 *level Gaussian pyramid*	*Global Features*
f_2		
f_3		
f_4		
f_5	*Mean hue of the whole image*	
f_6	*Mean saturation of the whole image*	
f_7	*Mean brightness of the whole image*	
f_8	*Points of interest of the whole image*	
f_9	*Edge density of the whole image*	
f_{10}	*Mean hue of the largest segment*	*Local Features*
f_{11}	*Mean saturation of the largest segment*	
f_{12}	*Mean brightness of the largest segment*	
f_{13}	*Mean hue of the second largest segment*	
f_{14}	*Mean saturation of the second largest segment*	
f_{15}	*Mean brightness of the second largest segment*	
f_{16}	*Maximal hue contrast of the largest segment and its neighbor segments*	
f_{17}	*Maximal saturation contrast of the largest segment and its neighbor segments*	
f_{18}	*Maximal brightness contrast of the largest segment and its neighbor segments*	
f_{19}	*Number of all segments*	
f_{20}	*Area of the largest segment*	
f_{21}	*Area of the second largest segment*	
f_{22}	*Shape complexity of the largest segment*	
f_{23}	*Shape complexity of the second largest segment*	

6.4 Objective Measure of Complexity

In this study, we treat the complexity evaluation as a three-class pattern classification problem. We apply a classifier to represent these three classes as LC, MC, and HC. Suppose that $x \in R^n$ denotes an extracted feature of the image, y denotes its class label (i.e. $y \in \{1, 2, 3\}$ respectively referring to LC, MC and HC), and $X_i = \{x_1, x_2..., x_{23}\}$ is the set of all features extracted from an image. Let $\{(X_i, Y_i), i = 1, 2, ...k\}$ denote a set of k training samples. Hence, the problem is considered to construct a classifier that exactly classifies a new testing sample X into one of the three classes. In this section, we introduce the application of SVM in effectively classifying the visual complexity of painting images.

6.4.1 Support Vector Machine

Let us first introduce the simple SVM classifier.

6.4.1.1 Linear SVM classifier

The simple case of SVM classifier should be the linear SVM classifier. A linear classifier has the form

$$f(x) = W^T(x) + b, \tag{6.25}$$

where W is known as the weight vector and b is the bias. In 2D the discriminant is a line and in 3D the discriminant is a is a plane, and in nD it is a hyperplane.

For the binary classification problem, given training data (X_i, y_i) for $i = 1, ...N$, with $X_i \in R^d$ and $y_i \in -1, 1$, learn a classifier $f(x)$ such that

$$f(x) \begin{cases} \geq 0 & y_i = +1 \\ < 0 & y_i = -1 \end{cases}$$

where $f(x) \geq 0$ means a correct classification. Here we show a linear classification example, such as Fig. 6.8.

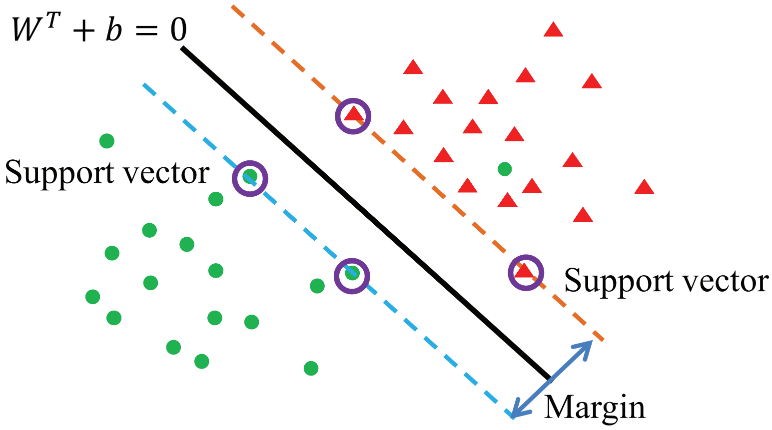


Figure 6.8: A linear SVM classification example with a hyperplane that maximizes the separating margin between two classes (indicated by the green points and red triangles). Support vectors are the elements that lie on the boundary hyperplane of the two classes.

In Fig. 6.8, the better solution of the classification is to enlarge the margin, even though one sample is misclassified. Mathematically, the optimization problem becomes learning a classifier $f(x)$ such that

$$min(\| W \|^2 + C \sum_{i}^{N} \varepsilon_i),$$

$$\text{subject to} \qquad y_i(W^T X_i + b) \geq 1 - \varepsilon_i, \text{ for } i = 1...N, \tag{6.26}$$

where C is a regularization parameter. Usually it is assigned a positive value by users. A small C allows constraints to be easily ignored, otherwise, a large C makes constraints hard to ignore.

6.4.1.2 Nonlinear SVM classifier

Many real world data cannot be separated linearly in a reasonable way because that linear classifiers are not complex enough to solve some problems like Fig. 6.9. The solution is to extend the linear SVM to nonlinear classifier by using a nonlinear operator $\phi(x)$ to map the input pattern X into a higher dimensional space H where it exhibits linear patterns, shown as Fig. 6.10. The nonlinear SVM classifier is defined as

$$f(x) = W^T \phi(x) + b, \tag{6.27}$$

where for the transformed data $\phi(x)$, the classifier is linear, but nonlinear in terms of original data x. The mapping function $\phi(x)$ maps input $x \in X$ to F (feature space). For nonlinear SVM classifier, the parameters of the decision function $f(x)$ are determined by the following minimization

$$min(\| W \|^2 + C \sum_i^N \varepsilon_i),$$

$$\text{subject to} \qquad y_i(W^T \phi(X_i) + b) \geq 1 - \varepsilon_i, \text{ for } i = 1...N. \tag{6.28}$$

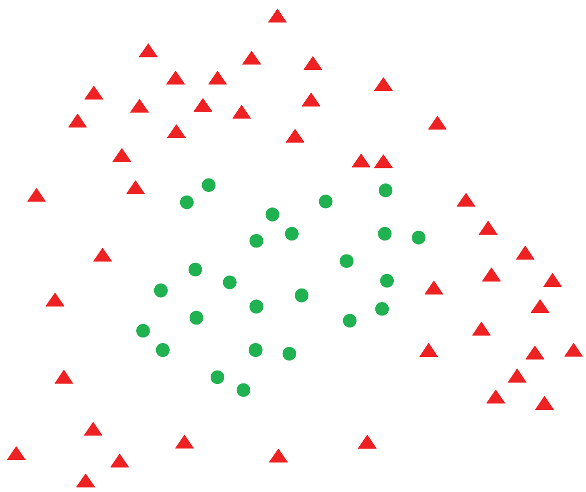


Figure 6.9: A linear SVM classifier is not complex enough to classify the samples into two classes.

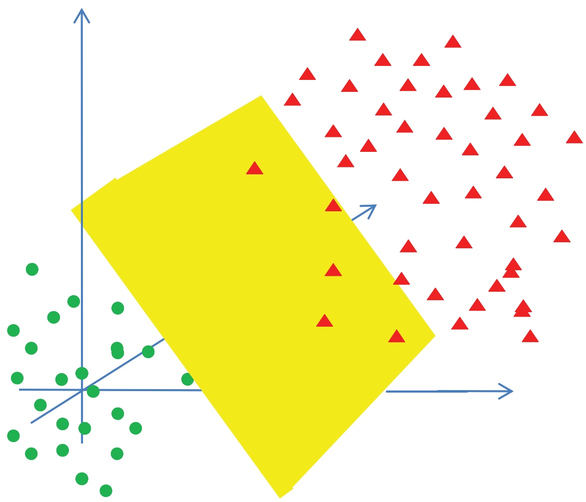


Figure 6.10: Mapping the input pattern into a higher dimensional space where it exhibits linear hyperplane.

6.4.1.3 Solution of SVM formulation

In the high dimensionality of the vector variable W, it is usually to solve Eq. (6.28) though its Lagrangian dual problem [101]:

$$F(\alpha) = \frac{1}{2}\alpha^T Q\alpha - e^T\alpha,$$

$$\text{subject to} \qquad 0 \leq \alpha_i \leq C, \text{i} = 1...N, \tag{6.29}$$

where $Q_{ij} \equiv y_i y_j \phi(x_i)^T \phi(x_i)$ and e is the vector of all ones. Here,

$$K(x_i, x_i) \equiv \phi(x_i)^T \phi(x_i), \tag{6.30}$$

is called the kernel function.

6.4.1.4 SVM kernel functions

We have found an appropriate mapping function $\phi(x)$ that allows for a linear separation in the high dimensional feature space H. In the process of mapping, a kernel function makes us directly calculate the inner products between the transformed feature vectors from the original data without actually having to consider $\phi(x)$ [102], just as the Eq. (6.30) Kernel functions help us to simplify the representation of data to be analyzed with data analysis methods. Frequently used kernel functions are

$$\text{Linear Kernel:} \quad K(x_i, x_j) = x_i^T x_j + c. \tag{6.31}$$

$$\text{Polynomial Kernel:} \quad K(x_i, x_j) = (\gamma x_i^T x_j + c)^d. \tag{6.32}$$

$$\text{Radial basis function (RBF) kernel:} \quad K(x_i, x_j) = exp\left(-\gamma \parallel x_i - x_j \parallel^2\right), \gamma > 0. \tag{6.33}$$

$$\text{Sigmoid Kernel:} \quad K(x_i, x_j) = tanh(\gamma x_i^T x_j + c). \tag{6.34}$$

where γ, c and d are kernel-specific parameters.

6.4.1.5 Multiclass SVM

Basically, SVMs are binary classifies, which means that they can be used as a decision function that will return yes or no for a given input data point. However, they can also be used in multiple class classification. Multiclass SVM aims to assign labels to instances by using SVM, where the labels are drawn from a finite set of several elements. In this study, we employed LIBSVM for classification. It implements “one versus one” method to classify the multiple labels [103]. If k is the number of classes, it should generate $k(k-1)/2$ binary models, each of which involves only two classes of training data. In classification, LIBSVM uses a voting strategy: each binary classification is considered to be a voting where votes can be cast for all data points, in the end a point is designated to be in a class with the maximum number of votes.

Table 6.3: Performance comparisons of various SVM kernel.

Types of kernel	*CA(%)*
linear	76.25
Polynomial	38.44
RBF	76.25
Sigmoid	**88.13**

6.4.2 Complexity classification using SVM

In this study, we employ SVM classifiers to classify image complexity into three classes. The tool that we used for implementing a multiclass SVM is LIBSVM-3.1 on the platform of Matlab 2011b.

6.4.2.1 Selecting kernel functions

It is widely acknowledged that a key factor in an SVM's performance is the choice of kernel function. The linear kernel is the simplest kernel function. The RBF kernel is actually a commonly used one in most of real applications because of its efficient performance. Polynomial kernel creates more hyperparameters than the RBF kernel. We tested the performances of four kernel functions (linear, polynomial, RBF and sigmoid functions), and finally we chose sigmoid function because it yielded the best performance, shown in Table 6.3. From this table, it is obviously shown that the sigmoid kernel function yields the best classification accuracy with 88.13%. In addition, a sigmoid kernel function is equivalent to a two-layer, perceptual neural network, which is similar to human visual perception mechanism.

6.4.2.2 Scaling data

Scaling before applying SVM is a very important step. The main advantage of scaling is to avoid attributes in greater numeric ranges dominating those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the calculation. Because some kernel values usually depend on the inner products of feature vectors (for examples, the linear kernel and the polynomial kernel), large attribute values might cause numerical problems [103]. In this experiment, all data were standardized by the function

$$X_i^{STD} = \frac{X_i - X_{mean}}{\sigma}, \tag{6.35}$$

where X_i denotes the each feature to be standardized, X_i^{STD} denotes the standardized score of X_i, X_{mean} denotes the mean value of dataset X, and σ denotes the standard deviation of X. All training data and testing data were scaled using the same method.

6.4.2.3 Parameters selection

The sigmoid kernel was used to process 1280 training data and 320 testing data. The training data were labeled as 1, 2, and 3 that respectively denoted to low complexity, middle complexity, and high complexity.

In order to realize the classification, parameters selection is very critical because SVM algorithm depends on several parameters. One of them is c, called penalty parameter, controls the tradeoff between margin maximization and error minimization. Another one is kernel parameter γ, which determines the kernel function and relates to the input range of the training data set.

In this experiment, we performed "grid.py" in LIBSVM to obtain the optimal parameters c and γ. The values of the parameters are respectively "8" and "13". We chose these values because they create the best classification accuracy.

6.4.3 Feature selection

In order to obtain the feature combination that yields the best performance, we replicated the experiment by 23 times while reducing one feature for each time. By this way, we can also look into the role that each feature plays. A total of 24 classification accuracies are shown in Fig. 6.11. In this figure, "None" marked in the horizontal axis means that we used all features for classification. The subsequent marks mean the deletion of the current feature in each classification. By analyzing this figure, we conclude that there are two aspects of influences for the classification accuracy (CA). One is that the deletions of some features improve or keep the CA, such as f_3, f_{12}, f_{14} and f_{15}. Another is that the deletions of some features severely decrease the CA, such as f_1, f_2, f_8, f_{10} and f_{16}. Since the deletions of f_3, f_{12}, f_{14} and f_{15} keep or even improve the CA, we decided to remove these indifferent features. Finally, we employed the selected 19 image features for classification.

6.4.4 Classification performance

The classification performance of the proposed method is shown in a confusion matrix (Table 6.4). The classification accuracy is 88.13%. We also calculated the Kappa coefficient (K), which estimates the agreement between subjective evaluations and predicted results. The K is equal to 0.8805. According to the interpretation of Kappa, the K value obtained indicates that the complexity predicted from our method keeps accordance with human visual complexity assessment.

Figure 6.12 gives the images with LC, MC, and HC predicted by our method. From this figure, the red frames indicate the wrong classification samples of image's complexity. The red frame listed in the first row is predicted as low complexity. However, the respondents assessed it as high complexity because it is hard to understand. The red frame listed in the second row is predicted as high complexity by the proposed method. The respondents believed, however, it is middle complex because it is a familiar scene for the respondents.

Table 6.4: Confusion matrix of predict classification.

	LC	*MC*	*HC*	
LC	63	1	3	67
MC	0	30	32	62
HC	1	30	189	191
	64	32	224	320

Table 6.5: Correlations between subjective complexity and objective measures of complexity.

	Proposed method	*Method in* [11]
Correlations	0.8901	0.8341

6.4.5 Comparisons with other measures

We compared the proposed method with the conventional method in [11]. Correlations (Pearson correlation) were calculated between subjective complexity assessments and objective measures of complexity. In Table 6.5, correlation comparisons show that the proposed method is more efficient in assessing visual complexity of paintings.

6.4.6 Validity of our method in other images

Moreover, we testified the validity of our method in assessing complexity in other images, for example, the architecture images [104]. Some example images are shown in Fig. 6.13. We used 122 training samples and 75 testing samples in the classification. The experimental results are shown in Table 6.6. Although the classification accuray of architecture images (73.33%) is not as high as that of painting images (88.13%), in a way, it indicates the validity of the proposed method in the architecture images.

Table 6.6: Classification accuracies of the proposed method used for painting images and architecture images.

	Painting images	*Architecture images*
Classification Accuracy	88.13%	73.33%

6.4.7 Application: local complexity of images

Our proposed method can also be used in evaluating the local complexity of images. This application is helpful for watermarking capacity estimation which has been illustrated in Chapter 2. Local complexity evaluation can assist to determine the appropriate region for watermarking. Usually, it is better to insert information in higher complexity regions. The more complex the local region is, the larger information can be inserted into.

We tested several images to analyze the validity of the proposed method in evaluating the local complexity of the image. Similarly, we classified the local regions into three levels of complexity: LC, MC, and HC. Here, we show two examples of local complexity estimation. These two images are respectively separated into 3×3 and 4×4 regions, shown as Fig. 6.14 (a) and Fig. 6.14 (b).

6.5 Discussion

Figure 6.3 illustrates the frequencies of options mentioned by the respondents in Part II. In this figure, the blue bar shows the result of question 1 and the red bar shows the result of question 2. According to the answers to question 1, it is obvious that 87.5% (28 in 32) of the respondents regarding “Distribution of compositions” as the most important factor that influences their complexity perception of paintings. Colors is supposed as the second important reason, with frequency of 81.25%. In addition, Contents is also of importance by the frequency of 68.75%. Other factors like understandability, symmetry, contrast and familiarity more or less affect visual complexity of paintings. For question 2, we asked the respondents to give two options that influences their evaluation of complexity most. Similarly, the options that frequently mentioned in question 1 are also with high frequencies in question 2. The top three frequently mentioned options in both questions are "Distribution of compositions", " Colors", and "Contents".

Interestingly, besides the options listed in the questionnaire, another factor that influences the judgment of complexity is mentioned by many respondents, that is "Color tone of image". It is believed by the respondents of this experiment that an image with warm-toned color is more complex than an image with cold-toned color on the basis of same level of contents and distributions. Hence, we concluded that three important factors affect respondents' assessment of complexity: "Distribution of compositions", " Colors", and "Contents".

We studied theoretical and empirical concepts from psychology and art theory to design 23 perceived features to represent the three factors that affect visual complexity of paintings: distribution of compositions, colors and contents. These features were combined by an SVM for classifying the complexity of paintings into three levels: HC, MC, and LC. In order to select the best combination of features for classification, we conducted the feature selection step by reduplicating the classification while abandoning one feature for each time. Figure 6.11 shows the classification accuracies of different feature combinations. It is obvious that the deletions of f_1, f_2, f_8, f_{10}, and f_{16} severely decrease the classification accuracy, which indicates that these features play pivotal roles in predicting visual complexity. We will explain the meanings of these features from human visual perception and psychology theory.

f_1 and f_2 represent the color complexity of first two levels of image pyramids. Compared with f_3 and f_4, f_1 and f_2 contain much useful information for predicting visual complexity. Color complexity is measured by color variation in the local regions of a painting. The higher the color variation is in the local neighbor, the more complex the region is. Hence, the painting is also of complexity. As has been identified in our experiment, color is one of important factors that affect human's perception of visual complexity in paintings.

f_8 is the feature of points of interest. It represents the content of the paintings. The more points of interests are in the painting, the more complex the painting is.

f_{10} and f_{16} are related to the hue information of the largest segment. It is said that the hue count of an image is a measure of the image's simplicity [106]. In the painting theory, the hue reflects the colorful keynote of a painting [105]. It represents the colorful keynote of the painting. The more colorful keynotes the painting has, the more complex the painting is perceived to be.

Definitely, it is shown in Fig. 6.11 that the deletions of some features do not impact the accuracy, such as f_{14} and f_{15}. By contrast, some deletions even improve the accuracies. They are f_3 and f_{12}. So we decided to remove these four features. Finally, 19 features were finally chosen for the classification.

We compared the proposed method with the conventional method in [11]. Correlations (Pearson correlation) were calculated between subjective complexity assessments and objective measures of complexity. In Table 6.5, the correlation comparisons show that the proposed method is more efficient in assessing visual complexity of paintings.

Moreover, we tested the validity of our method in assessing complexity in other images, for example, the architecture images. We used 122 training samples and 75 testing samples in the classification. The results in Table 6.6 indicate the validity of the proposed method in the architecture images.

Furthermore, we extended the proposed method in estimating the local complexity of images. The experimental results show that the proposed method is expected to be used for estimating the local complexity of images. This can be used for the application of watermarking capacity estimation.

6.6 Summary

In this chapter, we first conducted an experiment to analyze the factors that affect human visual complexity perception of paintings and obtain the subjective assessments of paintings. Three main factors that affect human visual complexity perception are identified: distribution of compositions, colors, and contents. Then we studied theoretical and empirical concepts from psychology and art theory to deign 23 features to globally and locally represent these factors. These features were combined by an SVM for classifying the complexity of paintings into three levels: HC, MC, and LC. In order to select the best combination of features for classification, we conducted the feature selection step by repeating the classification while abandoning one feature for each time. By the feature selection method, we looked into the role of each feature plays, and then removed those features which were not efficient for improving the classification accuracy. Finally, 19 features were selected and applied to the classification. Experimental results indicated that the proposed work could predict the visual complexity perception of paintings with the accuracy of 88.13%, which was highly close to the assessments given by humans.

In this part, we proposed a new framework to assess visual complexity of paining images. This framework provides a machine learning scheme for exploring human visual complexity and image features. Compared with the conventional measure of complexity, our work considers human visual perception and performs efficiently in assessing visual complexity of painting images.

This study is an attempt to assess the visual complexity of paintings by a machine learning method. It aims to build the relationship between high-level human feeling and low-level image features. Although this work is efficient in predicting human visual complexity of painting, it is far from enough. Since human visual perception is influenced by different aspects, we have to do our best to make this method more robust.

We showed a slight extension of our approach to architecture images, and applied the proposed method to evaluate the local complexity of images. In the future work, we will further investigate the validity of our approach in assessing complexity in other images and applications like watermarking capacity estimation.

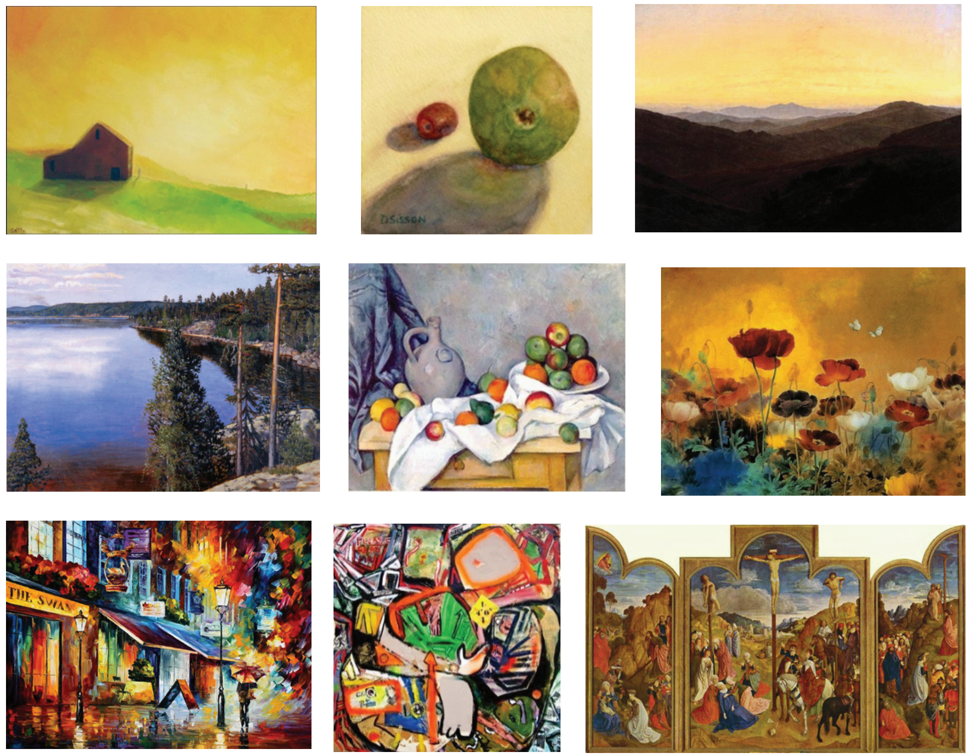


Figure 6.2: Examples that are labeled as LC, MC, and HC. The painting images on the first row are labeled as low-complex, the painting images on the second row are labeled as mid-complex, and the painting images on the bottom row are labeled as high-complex.

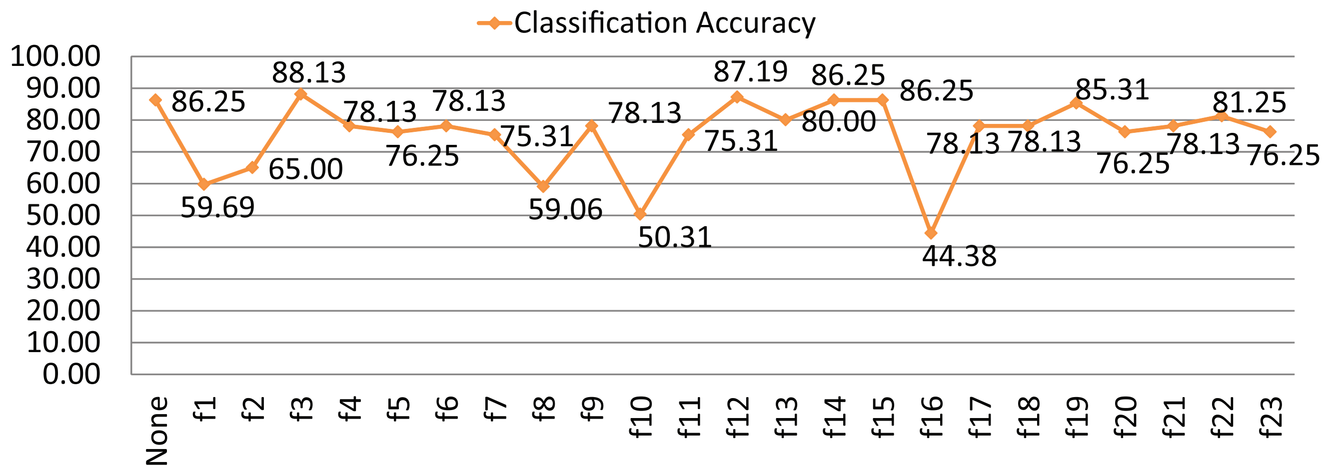
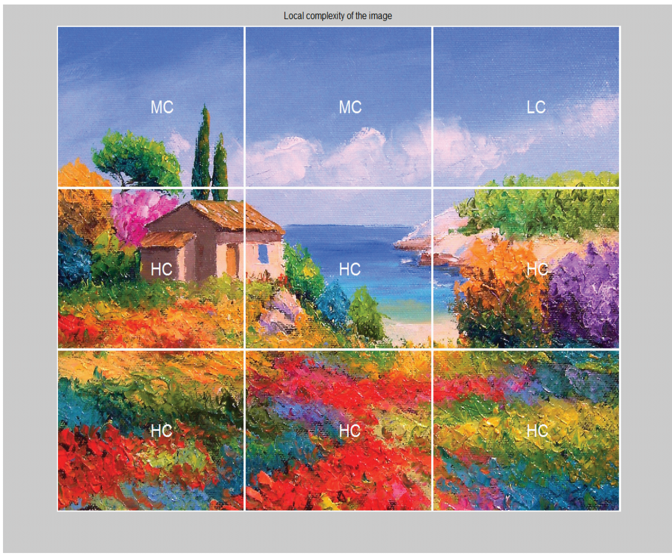


Figure 6.11: Classification accuracies of 24 feature combinations.

Figure 6.12: The label under the images are the prediction from our method. The red frames indicate the wrong classifications of image's complexity.

Figure 6.13: Architecture images that are labeled as LC, MC, and HC. The images on the first row are labeled as low-complex, the images on the second row are labeled as mid-complex, and the images on the bottom row are labeled as high-complex.



(a)

(b)

Figure 6.14: Local complexity estimation of two images that are divided into 9 regions and 16 regions respectively.

Chapter 7

Conclusions and Future Work

This chapter summarizes the proposed methods and discusses the future works.

7.1 Summary of Achievements

This dissertation can be divided into three parts: Part I only includes Chapter 2, it reviewed the basic knowledge of visual complexity and the related research works; Part II devoted to developing a new method of measuring subjective visual complexity of texture images; Part III extended the object of study to the painting images, and proposed a novel method of assessing visual complexity perception of painting images.

In Chapter 2, we presented the notion of visual perception and the classical principles of visual perception: Gestalt theory. We also introduced the concept of affective engineering that emphasizes the topic of endowing a computer the ability of recognizing human's emotion on images or other scenes. Moreover, we reviewed the different definitions of complexity depending on the different applications. Although complexity has been defined in various ways, there is no uniform definition. Then we presented the possible applications of complexity and its related works in different fields, such as, website designing, aesthetic perception, watermarking and so on. Because of the different understanding of complexity, the methods for measuring complexity become non-uniform. We summarized the previous methods for calculating complexity, including information theory, fractal dimension, quad tree method, visual attention method, etc.

Part II includes three chapters. In Chapter 3, we identified five important characteristics of texture images that affect visual complexity perception. The five characteristics are regularity, understandability, roughness, directionality, and density. Visual complexity is a function of not only each individual characteristic but also of interactions between them. In order to identify the texture features that affect subjective visual complexity, two psychophysical experiments were conducted involving visual complexity assessment and paired comparison evaluation. Thirty respondents participated in the experiments and evaluated the visual complexity of textures and described the criteria that they used to perceive complexity. The respondents marked five pairs of comparisons on a 7-point Likert scale. The techniques of correlation analysis, factor analysis, and multidimensional scaling were employed to further analyze the experimental results. The statistical analyses showed that the first four characteristics (regularity, understandability, roughness, and directionality) had significant effects on visual complexity perception. In the case of textures with similar level of regularity or directionality, understandability dominates the evaluation of visual complexity.

In Chapter 4, we designed a set of methods to objectively determine the five characteristics identified in Chapter 3. It helps determine the relationships between objective (calculated) and subjective (perceived) texture characteristics. This will facilitate the mapping between objective texture characteristics and human visual complexity perception. Computational measures for regularity, roughness, directionality and density were developed and correlated with human visual perceptions of them. For regularity, we adopted the autocorrelation function to extract the periodicity of a texture and measure the regularity. The result showed that the computational regularity was significantly correlated with the subjective regularity. For roughness, we measured the gray changes within a small region to mathematically calculate visual roughness. The comparison showed that the calculated roughness was related to the subjective roughness. The subjective perception of directionality was easily affected by edges. Hence, we regarded the maximum line-likeness of edges in different directions as the main direction of a texture. The results showed that the calculation of the maximum line-likeness of edges represented the subjective direction of a texture. For the density, we employed the method of calculating the edge density. A comparison between the calculated density and subjective density indicated that the subjective density was vulnerable to the influence of edges.

In particular, we introduced a new method to measure human understandability of a texture. The experimental results showed that it was possible to evaluate the understandability of a texture by naming it. In addition, the results showed that understandability could be estimated from two factors of a texture: its maximum number of similar names belonging to a specific type and its total number of types. The larger the number of similar names is for a texture, the more understandable it is. The more types the texture evokes, the less understandable it is.

In Chapter 5, we proposed a new approach of measuring visual complexity of a texture using texture characteristics, on the basis of human visual perception of complexity. Since we have identified five characteristics of textures that affect visual complexity and designed the computational methods to objectively calculate them, we employed Multiple Linear Regression as a mapping function that bridges a set of texture characteristics with visual complexity. In order to evaluate the performance of the proposed method and the related works, we compared with the method of Shannon entropy and the method in [11], the prediction of complexity from our method is closer to the subjective complexity perception. This method contributes to the computation of visual complexity of a texture considering human visual perception. It achieves the purpose that building the relationship between visual complexity perception of the human (high-level feeling) and the computational visual features (low-level features) extracted from the texture images.

Part III only contains Chapter 6. In this chapter, we firstly conducted an experiment to identify the factors that affect visual complexity perception of paintings. The experiment includes two Parts: complexity-rating and questionnaire. From Part I, we obtained the subjective assessment of complexity in painting images, and we got the main factors that affect human visual complexity perception in paintings from Part II, they are distribution of compositions, colors and contents.

After we obtained the main factors that affect visual complexity perception of painting images, we studied theoretical and empirical concepts from psychology and art theory to deign 23 features to globally and locally represent these factors. Then these features were combined by a multiclass SVM for classifying the complexity of paintings into three levels: HC, MC, and LC. In order to select the best combination of features for classification, we conducted the feature selection step by repeating the classification while abandoning one feature for each time. By the feature selection method, we looked into the role of each feature plays, and then removed those features which were not efficient for improving the classification accuracy. Finally, 19 features were seected and applied to the classification. Experimental results indicated that the proposed work can predict the visual complexity perception of paintings with the accuracy of 88.13%, which is highly close to the assessments given by humans. Compared with the conventional measure of complexity, our approach considers human visual perception and performs more efficiently in assessing visual complexity of painting images.

This thesis devotes to the computation of visual complexity in texture images and painting images. It achieves in building the relationship between visual complexity perception of human and the computational visual features in images. This kind of relationship is complicated. Our work is not a fill solution, just an attempt in this new and interesting kansei research.

7.2 Limitations and Future work

There are some potential limitations in the proposed methods descried as follows:

1) In this work, we obtained the subjective complexity assessments from the psychophysical experiments. The experimental samples (texture samples and painting samples) used in the experiments were limited. In the future work, a large number of experimental samples should be utilize in order to obtain more accurate experimental data.

2) Similarly, the respondents participated in each experiment was around thirty, and all respondents are Chinese students. However, amount of respondents and respondents of different nationalities are necessary to obtain more accurate experimental data and make the proposed model of visual complexity more robust.

3) We have proposed a new method of estimating human understanding of textures by an experiment of naming the textures, whereas, a computational method is also expected to objectively estimate human understandability. That is also what we try to achieve in the future work.

In this study, we have proposed the methods of measuring visual complexity of texture images and that of painting images. As we have illustrated in the first chapter, image complexity has many applications, for example, watermarking capacity estimation. We will further investigate the application of image complexity in the field of watermarking and other fields like aesthetic research etc.

Appendix A

1. The questions in the experiment of naming the textures

In the experiment of naming the textures, we asked the respondents to fill out a questionnaire, which was helpful for analyzing the reasons that influenced their understanding of textures. The questions in the questionnaire were as follows:

1. Which reasons helped you easily understand each image?

   A. You have seen the image content or similar content before.

   B. The image is simple and regularly composed.

   C. You can easily imagine the image.

   D. ________________________(if you have any other reasons)

2. Which one of the above reasons (A, B, C,or D)influenced you most? __________

2. The questionnaire in experiment of subjective assessment of paintings

The questionnaire in the experiment of subjective perception of complexity is shown as follows:

Please answer the following two questions.

1) which factors affect your judgment of visual complexity of paintings(multiple choices)?

   A. The colors

   B. The strength of the color changes

   C. Content (elements, objects, people)

   D. Distribution of composition (Regular, or not)

E. Understandability-Abstract

F. Symmetry

G. Contrast

H. Familiarity

I. _______________ (if you have any other reasons)

2) Which two factors that are most important?_____

Bibliography

[1] S. Tanaka, M. Inoue, M. Ishiwaka, S. Inoue, “A method for extracting and analyzing kansei factors from pictures,” IEEE Workshop on Multimedia Signal Processing, pp. 251–256 (1997).

[2] V. Chikhman, V. Bondarko, M. Danilova, A. Goluzina, and Y. Shelepin, “Complexity of images: experimental and computational estimates compared,” Perception, 2012, vol. 41, pp. 631–647.

[3] X. Guo, C. M. Asano, A. Asano, and T. KURITA, “Modeling Visual Complexity of Textures Associated with Texture Characteristics,” 2012 International Conference of Information Science and Computer Applications, 2012, pp. 121–126.

[4] M. Lima, “Visual Complexity Mapping Patterns of Information,” Princeton Architectural Press, 37 East Seventh Street, New York, 2011.

[5] A. Olive, M. L. Mack, M. Shrestha, and A. Peeper, “Identifying the perceptual dimensions of visual complexity of scenes,” Proceeding of the 26th Annual Meeting of the Congnitive Society, 2004, pp. 1041–1046.

[6] J. Perki and A. Hyvrinen, “Modelling image complexity by independent component analysis, with application to content-based image retrieval,” in 19th International Conference on Artificial Neural Networks: Part II, Limassol, Cyprus, 2009, pp. 111, Springer Berlin / Heidelberg.

[7] Y. A. Andrienko, N. V. Brilliantov, and J. Kurths, "Complexity of two-dimensional patterns," The European Physical Journal B, 2000, 15(3), pp. 539–546.

[8] A. Klinger and N. A. Salingaros, "A Pttern Measure," Environment and Planning B: Planning and Design, 27(4), 2000, pp. 537–547.

[9] M. Cardaci, V. D. Gesu, M. Petrou, and M. E. Tabacchi, "A fuzzy approach to the evaluation of image compexity," Fuzzy Sets and Systems, 2009, pp. 1474–1484.

[10] M. Murguia, A. Saenz, R. Rodriguez, "A Fuzzy Approach on Image Complexity Measure," Comp. y Sist. v.10 n.3 Mxico ene./mar. 2007, pp. 268–284.

[11] D. C. Donderi, "An information theory analysis of visual complexity and dissimilarity," Perception, 2006, 35(6), pp. 823–835.

[12] E. James Cutting, J. Jeffrey Garvin, "Fractal curves and complexity," Perception & Psychophysics, July 1987, Volume 42, Issue 4, pp 365–370.

[13] M. Jamzad, F. Yaghmaee, "Achieving higher stability in watermarking according to image complexity," Scientia Iranica Journal 13(4), 404–412 (2006)

[14] M.P.D. Silva, V. Courboulay, and P. Estraillier, "Image complexity measure based on visual attention," in Proc. ICIP, 2011, pp.3281–3284.

[15] B. Guy, "Art & complexity: an exploration of aesthetics," Ph.D. thesis, University: Nottingham Trent University, 2010.

[16] C. Stickel, M. Ebner and A. Holzinger, "The XAOS Metric - Understanding Visual Complexity as Measure of Usability," USAB 2010: 278–290.

[17] D. Berlyne, "Studies in the New Experimental Aesthetics. Washington," DC: Hemisphere Pub. Corp., 1974.

[18] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In. Multimedia, pages 83–92. ACM, 2010.

[19] W. Wang and Q. He. "A survey on emotional semantic image retrieval," 15th IEEE Int. Conf. on Image Processing, pp. 117–120, 2008.

[20] N. Bianchi-Berthouze, "Subjective perception of natural scenes: the role of color," Proc. SPIE 5008, Color Imaging VIII: Processing, Hardcopy, and Applications, 1 (January 20, 2003).

[21] C. Stickel, M. Ebner, and A. Holzinger, "The xaos metric - understanding visual complexity as measure of usability," in USAB'10, 2010, pp. 278–290.

[22] F. Yaghmaee and M. Jamzad, "Estimating Watermarking Capacity in Gray Scale Images Based on Image Complexity," EURASIP Journal on Advances in Signal Processing 2010.

[23] Y. K. Seong, Y. Choi, and T. Choi, "Scenebased Watermarking Method for Copy Protection Using Image Complexity and Motion Vector Amplitude," In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), May 2004, 409–412.

[24] F. Zhang and H. Zhang, "Wavelet domain watermarking capacity analysis, in Electronic Imaging and Multimedia Technology IV," vol. 5637 of Proceedings of SPIE, pp. 657–664, Beijing, China, February 2005.

[25] H. C. Kim, K. R. Kwon, J. J. Kim, "Adaptive Image Watermarkimg Using a Stochastic Multiresolution Modeling," International Technical Conference on Circuits Systems, Computers and Communications, 2002, 7, pp. 173–176.

[26] D. Cerra, A. Mallet, L. Gueguen, M. Datcu, "Complexity based analysis of earth observation imagery: an assessment," In: ESA-EUSC 2008: Image Information

Mining: pursuing automation of geospatial intelligence for environment and security (March 2008).

[27] M. Li, Y. Zhu, "Image classification via lz78 based string kernel: A comparative study," In: PAKDD, pp. 704–712 (2006).

[28] X. Yang, and C. Zhou, "Analysis of the Complexity of Remote Sensing Image and Its Role on Image Classification," In: Proceedings of the IEEE Geoscience and Remote Sensing Symposium (IGARSS 2000), July 2000, 2179–2181.

[29] I. Peters and R.N. Strickland, "Image complexity metrics for automatic target recognizers," in Automatic Target Recognizer System and Technology Conference. 1990, pp. 1–17, Citeseer.

[30] B. Auffarth, Y. Muto, and Y. Kunii, "An artificial system for visual perception in autonomous Robots," In: Proceedings of the IEEE International Conference on Intelligent Engineering Systems, September 2005, 211–216.

[31] http://en.wikipedia.org/wiki/Visual_perception.

[32] http://www.wisegeek.com/what-is-visual-perception.htm.

[33] Gestalt principles of perception.
http://www.usask.ca/education/coursework/skaalid/theory/gestalt/gestalt.htm.

[34] F. Attneave, "Physical determinants of the judged complexity of shapes," Journal of Experimental Psychology, 1957, 53, pp. 221–227.

[35] P. C. Vitz, T. C. Todd, "A model of the perception of simple geometric figures," Psychological Review 1971, 78, pp. 207–228.

[36] R. M. Nicki, & V. Moss, "Preference for non-representational art as a function of various measures of complexity," Canadian Journal of Psychology, 29, pp. 237–249, 1975.

[37] M. Nadal, E. Munar, G. Marty, C. J. Cela-Conde, "Visual Complexity and Beauty Appreciation: Explaining the Divergence of Results," Empirical Studies of The Arts, vol. 28, no. 2, pp. 173–191, 2010.

[38] R. Scha and R. Bod, "Computationele Esthetica," Informatie en Informatiebeleid, vol. 11 ,no. 1, pp. 54–63, 1993.

[39] F. Heylighen, "The Growth of Structural and Functional Complexity during Evolution, in F. Heylighen, J. Bollen & A. Riegler(eds.)," The Evolution of Complexity, Kluwer Academic, Dordrecht, 1997, pp. 17–44.

[40] C. Heaps and S. Handel, "Similarity and features of natural textures," Journal of Experimental Psychology: Human Perception and Performance, 1999, 25(2), pp. 299–320.

[41] W. Li, "On the Relationship between Complexity and Entropy for Markov Chains and Regular Languages," Complex Syst., 5, pp. 381–399, 1991.

[42] J. Rigau, M. Feixas, and M. Sbert, "An information-theoretic framework for image complexity," in Proceedings of the Euro-graphics Workshop on Computational Aesthetics in Graphics,Visualization and Imaging, 2005.

[43] M. Mario, D. Chacon, and S. Corral, "Image complexity measure: a human criterion free approach," in Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS'05), pp. 241–246, June 2005.

[44] R. Pieters, M. Wedel, R. Batra, "The Stopping Power of Advertising: Measures and Effects of Visual Complexity," Journal of Marketing, 74 (5), 2010, pp. 48–60.

[45] W. Saleem, A. Belyaev, D. Wang, H. P. Seidel, "On visual complexity of 3D shapes," Computers & Graphics, 35 (3), pp. 580–585.

[46] H.C. Purchase, E. Freeman, and J. Hamer, “Predicting Visual Complexity,” In: Predicting Perceptions. The 3rd International Conference on Appearance, 17-19 April 2012, Edinburgh, Scotland.

[47] A. Forsythe, M.Nadal, N. Sheehy, C. J. Cela-Conde, and M. Sawey, “Predicting beauty: Fractal dimension and visual complexity in art,” British journal of psychology (London, England:1953), 2011, pp. 49–70.

[48] G. T. Fechner, “Vorschule der aesthetik. Leipzig,” Germany: Breitkoff & Hartel, 1876.

[49] G. D. Birkhoff, “Aesthetic measure,” Cambridge, MA.: Harvard University Press, 1932.

[50] H. J. Eysenck, “The empirical determination of an aesthetic formula,” Psychological Review, 48, 1941, pp. 83–92.

[51] H. J. Eysenck, “The experimental study of the Good Gestalt-A new approach,” Psychological Review, 49, 1942, pp. 344–363.

[52] E. Michailidou, “ViCRAM: visual complexity rankings and accessibility metrics,” SIGACCESS Access. Comput, iss. 86, pp. 24–27, 2006.

[53] A. N. Tuch, E. E. Presslaber, M. Stcklin, K. Opwis, J. A. Bargas-Avila, “The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments,” International Journal of Human - Computer Studies vol. 70 issue 11 November, 2012. p. 794–811.

[54] E. Michailidou, S. Harper, and S. Bechhofer, “Visual complexity and aesthetic perception of web pages,” In Proceedings of the 26th annual ACM international conference on Design of communication (SIGDOC’08). ACM, New York, NY, USA, pp. 215–224.

[55] K. Reinecke, T. Yeh, L. Miratrix, Y. Zhao, M. Rahmatri, J. Liu, and K. Gajos, "Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness," In Proceedings of CHI '13, 2013. To appear.

[56] S. Ihara, *Information theory for continuous systems*, World Scientific Pub Co Inc (October 1993).

[57] http://en.wikipedia.org/wiki/Entropy_(information_theory).

[58] V. Benci, C. Bonanno, S. Galatolo, G. Menconi, F. Ponchio, "Information, complexity and entropy: a new approach to theory and measurement methods," 2001.

[59] L. N. Patel and P. Holt, "Testing a computational model of visual complexity in background images," In Advanced Concepts for Intelligent Systems, Baden, 2000, pp. 119–123.

[60] A. De Luca andS. Termini, "A Definition of a Non-probabilistic Entropy in the Setting of Fuzzy Sets Theory," Information and Control, Vol. 20, p 301, 1972.

[61] V. Di Ges'u and S. Roy, "Fuzzy measures for image distance," Advances in Fuzzy Systems and Intelligent Technologies (F. Masulli, R. Parenti, G. Pasi Eds.), Shaker Publishing (NL), 2000.

[62] K. Sandau, H. Kurz, "Measuring fractal dimension and complexity–an alternative approach with an application," J Microsc. 1997 May;186(Pt 2):164–76.

[63] W. Chen, X. X. Chen, and L. Zhou, "Box-Counting Dimension Analysis of Edge Images Detected with Classical Edge Detector," ICMTMA '10 Proceedings of the 2010 International Conference on Measuring Technology and Mechatronics Automation, Volume 01, pp. 756–758.

[64] M.P.D. Silva, V. Courboulay, and P. Estraillier, “Evaluation of preys / predators systems for visual attention simulation,” in VISAPP 2010 - International Conference on Computer Vision Theory and Applications, Angers, 2010, pp. 275–282, INSTICC.

[65] R. W. Picard, “Affective Computing,” Cambridge, MA:MIT Press, 1997.

[66] X. Guo, C. M. Asano, A. Asano, and T. KURITA, “Visual Complexity Perception and Texture Image Characteristics,” 2011 International Conference on Biometrics and Kansei Engineering (ICBAKE 2011), Takamatsu, Japan, pp. 260–265, 2011.

[67] H. Tamura, S. Mori, and T. Yamawaki, “Textural Features Corresponding to Visual Perception,” IEEE Transactions On Systems, Man, and Cybernetics, 1978, 8(6), pp. 460–473.

[68] M. Amandasun and R. King, “Textural features corresponding to textural properties,” IEEE Transactions On Systems, Man, and Cybernetics, 1989, 19, pp. 1264–1274.

[69] K. Fujii, S. Sugi and Y. Ando, “Textural properties corresponding to visual perception based on the correlation mechanism in the visual system,” Psychological Research, 2003, 67(3), pp. 197–208.

[70] L. Li, A. Asano, and C. Muraki Asano, “Statistical analysis of human visual impressions on morphological image manipulation of gray scale textures,” Optical Review, vol. 17, no. 2, pp. 90–96, 2010.

[71] A. R. Rao and G. L. Lohse, “Towards a texture naming system: identifying relevant dimensions of texture,” Vision Research, 1996, 36(11), pp. 1649–1669.

[72] A. R. Rao and G. L. Lohse, “Identify high-level features of texture perception,” Graphical Models and Image Processing, 1993, 55, pp. 218–233.

[73] R. Y. Cho, V. Yang, and P. E. Hallett, "Reliability and dimensionality of judgments of visually textured materials," Attention, Perception & Psychophysics, vol. 62, no. 4, pp. 735-752, 2000.

[74] P. Brodatz, *Texturs*, New York: Dover, 1966.

[75] M. Kokubun, M. Konishi, M. Kuwasumi, and H. Iguchi: Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conference of the IEEE. (2000) 1592.

[76] C. H. Coombs, R. M. Dawes, and A. Tversky, "Mathematical Psychology: An Elementary Introduction," Englewood Cliffs, N. J. :Prentice-Hall, 1970, pp. 235–300.

[77] http://en.wikipedia.org/wiki/Multidimensional_scaling.

[78] H. C. Lin, L. L. Wang, and S. N. Yang, "Extracting periodicity of a regular texture based on autocorrelation functions," Pattern Recognition Letters, 18, pp. 433–443 (1997).

[79] G. Vaidyanathan and P. M. Lynch, "Texture direction analysis using edge counts," Southeastcon'89. Proc. Energy and Information Technologies in the Southeast., IEEE. 1989, vol.2, pp. 733–738.

[80] R. V.Egmond, P. Lemmens, T. N. Pappas, and H. de Ridder, "Roughness in sound and vision," in [Human Vision and Electronic Imaging XIV], Rogowitz, B. E. and Pappas, T. N., eds., Proc. SPIE 7240, 72400B112 (Jan. 2009).

[81] W. Groissboeck, E. Lughofer, S. Thumfart, "Associating visual textures with human perceptions using genetic algorithms," Information Sciences, 2010, 180 (11), pp. 2065–2084.

[82] X. Guo, C. Muraki, A. Asano, and T. Kurita, "Modeling the Perception of Visual Complexity in Texture Images," International Journal of Affective Engineering, Vol.12, No.2 (2013) in press.

[83] X. Guo, C. Muraki, A. Asano, T. Kurita and L. Li, "Analysis of Texture Characteristics Associated with Visual Complexity Perception," OPTICAL REVIEW Vol. 19, No. 5 (2012), pp. 306–314.

[84] http://www.weibull.com/DOEWeb/simple_linear_regression_analysis.htm.

[85] http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_regression_linear_multiple.htm.

[86] Trevor Hastie, Robert Tibshirani, Jerome Friedman, *The Elements of Statistical Learning*, Springer, December 2008.

[87] C. Li and T. Chen. "Aesthetic Visual Quality Assessment of Paintings, " IEEE Journal of Selected Topics in Signal Processing, 2009, 3(2), pp.236–253.

[88] V. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.

[89] PaintingDb,http://paintingdb.com/.

[90] R. Arnheim, "Art and Visual Perception: a Psychology of the Creative Eye(expand and revised edition)," University of Califormia Press, 1974.

[91] G. Wyszecki and W.S. Stiles, "Color Science: Concepts and Methods," Quantitative Data and Formulae, Wiley, New York, 1982.

[92] K. J. Yoon and I. S. Kweon, "Color image segmentation considering of human sensitivity for color pattern variations," Proc. SPIE. 2001. pp. 269–278.

[93] Gaussian pyramid, http://www.wisegeek.com/what-is-a-gaussian-pyramid.htm.

[94] http://en.wikipedia.org/wiki/Munsell_color_system.

[95] http://www.greatreality.com/color/ColorHVC.htm.

[96] V. Di Ges and C. Valenti, "The Discrete Symmetry Transform in Computer Vision," Technical Report DMA 011 95, Palermo Univ., 1995.

[97] http://coewww.rutgers.edu/riul/research/code/EDISON/index.html.

[98] J.F. Ning, L.Zhang, D. Zhang, C.K. Wu, “Interactive image segmentation by maximal similarity based region merging,” Pattern Recognition 43(2): 445–456 (2010).

[99] A. Tremeau, “Regions adjacency graph applied to color image segmentation,” IEEE Transactions on Image Processin, 9(4), pp.735–744.

[100] A. B. Watson, “Perimetric complexity of binary digital images,” The Mathematica Journal. 2011. pp. 1–40.

[101] H. T. Lin and C. J. Lin, “A Study on Sigmoid Kernels for SVM and the Training of non-PSD Kernels by SMO-type Methods,” http://www.csie.ntu.edu.tw/ cjlin/papers/tanh.pdf.

[102] B. Vanschoenwinkel and B. Manderick, “Appropriate Kernel Functions for Support Vector Machine Learning with Sequences of Symbolic Data,” Deterministic and Statistical Methods in Machine Learning 2004,pp. 256–280.

[103] C. C. Chang and C. J. Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[104] Photo Quality Dataset, http://mmlab.ie.cuhk.edu.hk/CUHKPQ/Dataset.htm.

[105] Understanding Hue, Value and Chromatic Intensity, http://www.finkeldeistudio.com/art-how-to-articles/78.

[106] Y. Ke, X. Tang, and F. Jing, “The design of high-level features for photo quality assessment,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, June 2006, pp. 419–426.

List of Publications

Refereed Journal Papers

1. Xiaoying Guo, Chie Muraki Asano, Akira Asano, Takio Kurita, and Liang Li, "Analysis of Texture Characteristics Associated with Visual Complexity Perception," **Optical Review**, Vol. 19, No. 5 (2012) pp. 306–314.

2. Xiaoying Guo, Chie Muraki Asano, Akira Asano, and Takio Kurita, "Modeling the Perception of Visual Complexity in Texture Images," **International Journal of Affective Engineering**, Vol.12, No.2 (2013) pp. 223–231.

Refereed International Conference Papers

1. Xiaoying Guo, Chie Muraki Asano, Akira Asano, and Takio Kurita, "Visual Complexity Perception and Texture Image Characteristics," 2011 International Conference on Biometrics and Kansei Engineering (**ICBAKE 2011**), Takamatsu, Japan, pp. 260–265, 2011.

2. Xiaoying Guo, Chie Muraki Asano, Akira Asano, and Takio Kurita, "Analysis of the Perception of Visual Complexity in Texture Images, " 2012 International Conference on Kansei Engineering and Emotion Research (**KEER 2012**), Penghu, Taiwan, pp. 216–224, 2012.

3. Xiaoying Guo, Chie Muraki Asano, Akira Asano, and Takio Kurita, “Modeling Visual Complexity of Textures Associated with Texture Characteristics,” 2012 International Conference of Information Science and Computer Applications (**ICISCA 2012**), Bali, Indonesia, November 19–20, pp. 121–126, 2012.

4. Xiaoying Guo, Takio Kurita, Chie Muraki Asano, and Akira Asano, “Visual Complexity Assessment of Painting Images,” 2013 International on Image Processing (**ICIP 2013**), September 15-18, Melbourne, Australia, Accepted.