Computational Aesthetics and Visual Preference - An Experimental Approach

Florian Hoenig

March 1, 2006

Dedicado a Karla...

Acknowledgements

I wish to thank all the people who supported me during the course of this work and who had given both positive and negative comments in a constructive manner.

In particular I want to thank my supervisor Josef Scharinger for acceptance and support of this topic, which had aroused from personal interest and which is inherently hard to elaborate in a scientific way. Special thanks to Gerhard Widmer for guidance and helpful suggestions regarding machine learning experiments.

I'd to also thank Todor Georgiev and Mark Grundland, who had the patience to answer many questions and gave inspiring comments during early stages of this thesis.

Kurzfassung

"Computational Aesthetics" ist ein Begriff, der haeufig von Wissenschaftlern fuer den quantitativen Zugang zur Aesthetik verwendet wurde. Ein Ueberblick sowohl ueber die vergangenen als auch die neuesten Werke wird dargelegt und deren hauptsaechlichen Ideen von *Komplexitaet* und *Ordnung* aufgezeigt.

Da diese jedoch kaum auf Erkenntnisse der menschlichen Wahrnehmungsforschung basieren, wird der Begriff der *Komplexitaet* im Rahmen eines derzeit allgemein anerkannten Modelles zur visuellen Wahrnehmung neu ausgelegt. Hinsichtlich der experimentellen Auswertung dieses Ansatzes wird eine Hypothese formuliert, welche besagt, dass visuell-aesthetische Wahrnehmung nicht voellig von Komplexitaet im primaeren visuellen Kortex (Zellanreiz) unabhaengig ist.

Zur Ueberpruefung werden *Multivariate Gabor Filter Reaktionen* als Schaetzwerte fuer diese Zellaktivitaeten vorgelegt und zusaetzlich einfache Bildeigenschaften wie beispielsweise *Seitenverhaeltnis* und *Aufloesung* herangezogen, um die Wechselbeziehungen umfassender zu pruefen. Diese werden schliesslich mit menschlicher Bewertung von Fotos verglichen und Ergebnisse hieraus zeigen statistische Signifikanz auf, die jedoch niedrig ausfaellt. Zusaetzliche Experimente nach Methoden des maschinellen Lernens wurden durchgefuehrt, denen es jedoch misslingt die menschliche Preferenz vorherzusagen.

Obgleich diese Ergebnisse nur schwach mit aesthetischer Wahrnehmung in Zusammenhang stehen, so wird dadurch dennoch weitere Forschungstaetigkeit und naehere Begutachtung von Bildmerkmalen und deren Zusammenhang mit Wahrnehmungswirkungen und visuellen (aesthetischen) Vorlieben angeregt.

Abstract

Computational Aesthetics is a term which has been frequently used by scientists interested in quantitative approaches towards the concept of aesthetics during the last century. A review of both past and recent works attempting to quantify aesthetic preference of various stimuli is given, which shows that aesthetics was described as some function of *complexity* and *order* in many theories.

Since most measures were hardly relating to knowledge of human perception, *complexity* is reinterpreted in the context of a currently accepted model of visual perception and a hypothesis is formulated which states that human visual preference is not independent of complexity (cell excitation) at the very first stage of visual processing.

An estimate representative for cell activity in early visual processing is presented: *Multivariate Gabor Filter Responses*. Additionally, image properties such as *aspect ratio*, *resolution* and *JPEG compressibility* are used to sanity-check any correlations.

The estimate calculated, compared against human preference ratings of photographs, shows statistically significant but low correlations. However, the machine learning experiments performed, fail to predict any better than one would by taking the mean value of the data.

Even though these results only loosely relate to aesthetic perception, it's motivating further research and closer inspection of image features and their relation to perceptual properties and visual (aesthetic) preference.

Contents

List of Figures

List of Tables

Part I

Theory

Chapter 1

Introduction

"There is no must in art because art is free"
–Wassily Kandinsky

Starting with this quotation, it'd like make clear what this thesis is *not* about. It is not trying to build a theory of art. Also I am aware that the topic of rational aesthetics is very controversial from an art perspective and also *really* easily leads to wrong scientific claims. However, it has been addressed by scientists over and over and is getting increased attention in some scientific fields at the time of writing.

1.1 Motivation

Philosophers have discussed aesthetics for ages. Although the Greek origin of the word is $\alpha\iota\sigma\theta\eta\tau\iota\kappa\eta$, meaning "a perceiver", it is now widely accepted in almost any encyclopedia to be defined as "the philosophical study of beauty and taste". Kant had also described aesthetics as a reinforcing supplement to logic ideas or concepts [20], hinting that objects are of *higher value* to us if they are beautiful (in addition to the values of meaning and function). No doubt that ingestion is of higher pleasure if the food is tasty.

In a similar fashion, aesthetics plays a major role in *design*, complementing *function* and improving the products value in many ways. This fact is commonly known and we experience this aesthetics in the every day usage of many design objects such as cars, furniture, consumer electronics and so on. Good aesthetic design supports our understanding of complex functional objects, unifies their perception in a socioeconomic context (e.g. commercials) and helps seamlessly integrating them into an environment (probably the best example would be architecture).

Now, heavy use of computer aided design tools introduces a certain kind of uniformity in the aesthetics produced. It appears as implicit aesthetics, which is not planned or taken care of. This 'byproduct' of how functional design and planning is performed by computational tools can be widely observed in e.g. architecture, food packaging, leaflets, etc. and the number of objects produced by these tools is growing enormously. Particularly, the Internet delivers a flood of media created by both private individuals and professionals who do not necessarily put effort into *aesthetic design*. All of which leads to a phenomenon of *aesthetic pollution*.

However, this part of the design process offers a problem for computer science to improve software tools in such a way that they are also aware of aesthetics, even if there is no human (artist) involved. This introduces one major motivation for finding computational methods which are capable of making aesthetic decisions. A paradigm which was called *computational aesthetics* by various researchers. This signification will be used throughout the whole work.

To identify solvable aesthetic problems in this context, one must point out the differences between objects of design and objects of art. The latter differ from the former by the lack of functional requirements which allows for unconstrained aesthetic possibilities. In other words, there are no determined objective demands for their aesthetics. For scientific research it makes no sense to bind this freedom of art in any way. That means for aesthetic research it is incidental to focus on the more determined aspects. However, since objects of art are aesthetically more versatile, they offer a more explorative basis for analysis. Also, computer generated art has been a rather popular topic for scientists interested in aesthetic computation in history and present, likely because it has often turned out as the only test-bed for developed aesthetic theories and measures.

On the bottom line, currently observed rational aesthetic research puts emphasis on application and focuses on aesthetic problems in design, which most importantly offers immediate application. There are essential questions: Can tools be created that assist with creating *beauty* as easily as they already do with purely functional creation? Can machines learn to perform aesthetic decisions in a similar fashion as human perception does?

1.2 Structure

Motivated by these questions, this thesis tries to point out a mathematical and computational perspective of aesthetic phenomena and tries to elaborate and explore it with findings in *visual perception*. It is laid out in the following way:

Chapter 2 historically reviews attempts of building aesthetics theories and approaches towards quantification of certain aesthetic aspects of objects. Works are ranging from pure *Gedanken* experiments and metaphors to concrete psychologically grounded studies. This summary contributes to a better understanding of how scientists were involved with aesthetics and how technological and mathematical advances were accompanying research development. Digesting past work, Chapter 3 filters out the essential aspects observed, which seem contributory to *computational aesthetics*. Finding in visual perception are compared and matched with methods in computer vision and a model is postulated, which situates various aesthetic phenomena in the various logical layers of perception. Finally, a hypothesis is formulated which relates a part of aesthetics experiences to the most known layer: early vision and the primary visual cortex.

Chapter 5 describes two implementations of estimates of *visual complexity* in an early vision way of understanding. Both are compared against aesthetic preference ratings of photographs collected from users of a huge online photo community site Chapter 4 describes this data set.

Finally the results and contributions are concluded and open research problems and directions are pointed out.

Chapter 2

Computational Aesthetics

"Pay attention only to the form; emotion will come spontaneously to inhabit it.
A perfect dwelling always find an inhabitant."
–Andre Paul Guillaume Gide

This chapter will try to outline a scientific direction that has been continuingly reappearing throughout the last century. Today a trend again shows resumed interest of computer scientists for quantifiable aspects of the phenomenons of beauty, because mathematical and computational tools are becoming more powerful and potentially more capable to imitate human perceptual behavior.

The following sections give a historical review of the origin of these ideas and summarize and categorize important works in that paradigm.

2.1 Historical Overview

2.1.1 Origin

There is no know scientific theory of aesthetics to date and therefore, no matter what methodology used, research is inherently empirical and experimental. Due to this fact, it seems right to date the origin of computational aesthetics back to the pioneering work of Gustav Theodor Fechner's

"Vorschule der Aesthetik" [11]. Fechner defined the term *experimental aesthetics* which continuously provided a foundation for collecting and evaluating human preference data. His experiments included tests on aesthetic *rules* like the golden ratio and his methodologies, as well as the research questions he had pointed out, are today still considered current in the field of empirical or *Experimental Aesthetics*.

George David Birkhoff, who was aware of Fechner's work, however was the first mathematician trying to quantify the aesthetic value of an object in his book *Aesthetic Measure* [8] published in 1933. Since it involves computational methods, this work is often regarded as the actual beginning of *computational aesthetics*, as also described in a recently published historical summary by Greenfield [17].

2.1.2 Time Line

Besides the origin of the idea of quantifiable aesthetics, which is debatable, its historical development displays interest from many fields of research. Along with the advent of new technologies and insights, new methods and approaches have emerged for this problem. The following time line is laid out to give an overview on scientists' involvement with this topic:

1876 G.T. Fechner's "Vorschule der Aesthetik" defined the term *Experimental Aesthetics*, which is groundwork for today's modern empirical research of aesthetics [11].

1933 G.D. Birkhoff came up with the first attempt to directly quantify aesthetic value as a mathematical function of *order* and *complexity* in his book "Quantifying Aesthetics". He started the idea of *aesthetic measures* [8].

1954 Max Bense published *Aesthetica I - Metaphysische Beobachtungen am Schoenen*, his first book about aesthetics and mentioning Birkhoff [3].

1958 Abraham Moles published the first book on information theory and aesthetics, being a contemporary of Max Bense [26].

1965 Max Bense integrated Claude Shannon's information theory with Birkhoff's formula, using the term *information aesthetics* in his book "Aesthetica - Einfuehruung in die neue Aesthetik" [4]. The book represents the sum up of his earlier four-volume "Aesthetica" series. Much of his work was based on his students' research.

1969 Until then, Bense published more articles on aesthetics and another book dedicated to his *information aesthetics* work. [6].

1971 Berlyne summarized the state of psychological research of aesthetics [7] which has been basis for many further investigations since then.

1978 Stiny and Gips formulated the first algorithms-oriented approach towards aesthetics and the automated criticism and generation of art [36].

1993 Scha and Bod published a Dutch article which, although having no new research contributions, was actually the first known published work that summarized former work under the name *Computational Aesthetics* [29].

1994 Michael Leyton founded the `International Society for Mathematical and Computational Aesthetics` (IS-MCA), identifying the need for aesthetics in design and pointing out research directions for major areas. The IS-MCA is the first known group or institution dedicated to *Computational Aesthetics.*

1994 Shumeet Baluja et al. trained a neuronal network to learn human aesthetic preference of images for automated evolution, and some results were indeed quite interesting [2].

1997 Frank and Franke published "Aesthetische Information" [14] as an attempt to finally define *information aesthetics* as a field and not let it die. It is an almost complete summary of follow-up works based on Bense et al.

1998 Machado and Cardoso published a paper named "Computing Aesthetics", picking up the quest for measure of aesthetic value again, using more recent computer science methods. They also made the first link between aesthetics and fractals. [25]. However, there is not direct reference to Birkhoff in the paper.

1998 David Chek et al. published "Aesthetic Measure for Screen Design", concluding successful objective measurement of screen layout aesthetics building upon Birkhoff's formula [10].

1999 Tomas Staudek also picked up interest in Birkhoff's formula and tried some extensions; this appeared in a technical report [33].

2000 Sudweeks and Simoff used Chernoff faces as metaphors to let users encode aesthetic values and laid out an interface for aesthetic evaluation [37, 31].

2002 Gary Greenfield publishes work on images generated by genetic algorithms under the name *simulated aesthetics* [16].

2002 A Dagstuhl workshop on "Aesthetic Computing" was held by Paul Fishwick and others. It set out goals to define an area of research where art directly supports science. An "Aesthetic Computing Manifesto" was published in [12].

2003 Tomas Staudek presents a GUI tool called Arthur for generating shape-based images and evaluating them with measures defined by information aesthetics [34].

2003 A paper called *Universal aesthetic of fractals* [32] appeared to be the first direct comparison of fractal dimension and human aesthetic preference.

2004 Staudek follows up research on aesthetics presenting a similar fractal approach as above in the journal article called: *Personality Characteristics and Aesthetic Preference for Chaotic Curves* [35].

2004 Tali Lavie and Noam Tractinsky committed some studies on web site aesthetics, being aware of the field of experimental aesthetics, but do not refer to Birkhoff roots [24].

2005 A Eurographics workshop on *Computational Aesthetics in Graphics, Visualization and Imaging* set out the goal to once again try to define the meaning of *computational aesthetics* and Gary Greenfield published the latest known summary on the origins of this term [17].

2.2 Theoretical Efforts

The following sections describe and categorize historical works on this new aesthetics, which tries to be rational and quantify the phenomenons of beauty or perceptual pleasure. Admittedly, categorization is a vague term, since the amount of works is very limited and little consensus between different theorists is observed.

2.2.1 Rational Aesthetics

In 1933, the American mathematician George David Birkhoff wrote the first quantitative theory of aesthetics in his book *Aesthetic Measure* [8]. He was known for many contributions to mathematical analysis in dynamics and linear differential equations. Interestingly, he formulated the *Ergodic Theorem* which is about measure-preserving transformations and came from problems in statistical physics. Birkhoff had also insights from physics and also collected a wide range of experiences from art, which finally caused his interest in aesthetics.

Being aware of most philosophical views and descriptions of aesthetics, he described the aesthetic experience in three phases:

1. First, effort is required to focus attention on an object. (Complexity)

2. A feeling of value that rewards this effort (Aesthetic Measure).

3. Realization of certain harmony that the object contains (Order).

From these observations he constructed the formula (2.1) which should describe this aesthetic relationship which he says is commonly known as "unity in variety".

$$AestheticMeasure = \frac{Order}{Complexity} \tag{2.1}$$

His definition or meaning of *complexity* is based upon attention and it's physical correlative, what he calls "automatic motor adjustments". If for example attention is fixed on a complex polygon, the eye follows the lines of the object and these adjustments raise the feeling of *effort*. Therefore he counted lines of polygons and took the number as an estimate for the object's complexity.

Birkhoff quoted the psychologist Theodor Lipps, who agreed with the idea that more or less complete identification with the perceived object enhances the aesthetic experience. This identification lies in the amount of associations a person can make with the percept.

He distinguished between two basic type of associations. First, there are *formal* ones like symmetry, proportion, balance and so on, where some of them contribute to positive and others to negative emotions. All other associations, like *meaning*, he called *connotative*.

The properties of an object causing those associations were called *formal elements of order* and *connotative elements of order* respectively.

From this he formulated the problem in a mathematical fashion as follows:

> Within each class of aesthetic objects, to define order O and the complexity C so that their ratio $M = \frac{O}{C}$ yields the aesthetic measure of any object of the class.

With this definition, Birkhoff did not actually make too big claims about the scope of the formula's validity (in contrast to the book's title). Instead he stated that it will be only applicable to properly restricted classes of objects which allow direct intuitive comparison. Nobody would compare a vase with a melody, but portraits of the same person are readily compared and sorted in order of aesthetic preference.

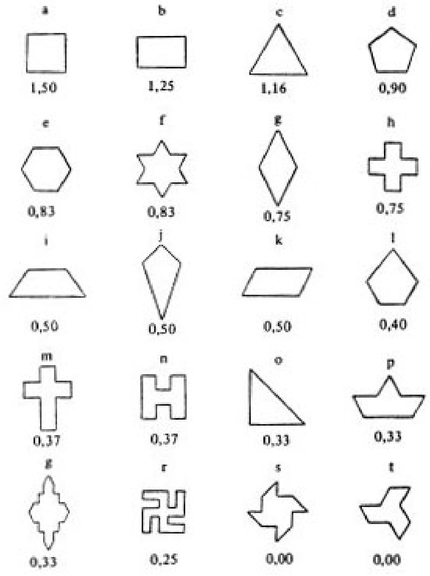

Figure 2.1: Birkhoff's Aesthetic Measure of Polygons

He also said that one needs to define a suitable definition of complexity within that class. Moreover, only formal elements of order can be considered, since connotative elements are beyond measurement. Finally, he said that the measure represents the aesthetic judgement of an "idealized observer", meaning someone who is familiar with that class of aesthetic objects (probably to eliminate dependencies on diverse connotative associations).

The rest of his work focused on defining measures of complexity and order for various classes of aesthetic objects including vases, ornaments, polygonal forms, music and poetry.

Birkhoff's work was the origin of many of the following developments.

2.2.2 Information Aesthetics

The last revival of a significantly large scientific movement, namely *information aesthetics*, was a book by Frank and Franke published in 1997 [14]. It set goal to preserve works centered around the ideas of Max Bense and Abraham Moles (see 2.1.2) which originated earlier in the century and were followed up by a number of researchers and computer artists until then. The authors of this book have tried to situate this *theory of aesthetic information* into the field of cybernetics, the theory of control and communication of machines and living entities. It is more or less a complete review of this movement. However, only parts of it are relevant for the interest in computational aesthetics.

Max Bense was one of the two actual founders, was a German mathematician, physicist and philosopher. His teaching included mathematical logic, scientific theory and the philosophy of technology. He tried to bring information theory together with the ideas of Birkhoff's measure and published the book "Aesthetica. Einfuehrung in the neue Aesthetik" [4] in 1965, based on his previous work and the research of his students Gunzenhaeuser and Frank (as reported in [17]). Both of them where students of Max Bense.

Contrary to his predecessors and contemporaries as well as most philosophers, Bense tried to define an objective and universal aesthetic. In the foreword of one of his younger books, "Einfuehrung in die informationstheoretische Aesthetik" [6], he speaks about an objective, scientific, material and abstract aesthetic, relying on the usage of mathematics, semiotic, physics, information and communication theory. He also clearly expresses it's segregation from philosophical aesthetics and meta-physics. Bense wrote:

> "Therefore this aesthetic can be seen as objective and material, not using speculative, but rather rational methods. It's interest lies primarily and ultimately on the object itself; the relationship to it's consumer, buyer, view, critic, etc. recedes."

In an article in *rot journal* [5], Bense summarized his many terms and definitions. He used the word *abstract aesthetics* for his theory which was applicable to any kinds of aesthetic objects, such as architecture, design, painting, film, music, theater, and so on. He took Birkhoff's attempt of quantification and derived two classes of numerical aesthetics:

macro aesthetics Bense interpreted Birkhoff's intentions as being so called macro aesthetics, since it's components can be perceived and measured straightforwardly. Further, it is geometrically oriented.

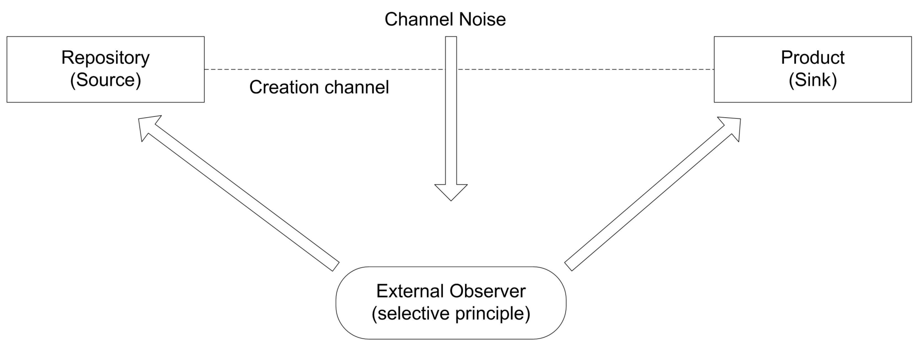

Figure 2.2: Creative model of Max Bense.

micro aesthetics however tries to measure the way an object was probabilistically "selected from a finite repertoire of material elements". In contrary, this is of statistical nature.

This *numerical micro aesthetics*, based on Birkhoff's notion and interpreted with Shannon is formulated as:

$$MicroAesthetics = \frac{Redundancy}{Entropy} \tag{2.2}$$

Here entropy represents complexity and statistical redundancy represents order, and is measured in the binary representations of artworks. In order to further apply this information theoretical measure to the creative process, Bense constructed an alternative model to the classical sender/receiver/channel model of communication, the *creative model* (see figure 2.2).

One of it's main differences is the addition of an external observer, who selectively communicates elements from the finite material repertoire into an innovative state, the product. Bense generalizes the word *material* as being discrete, differentiable and manipulatable. An aesthetic repertoire is such a set of elements from which aesthetic states can be created using some manipulative principles, and therefore each aesthetic state is repertoire dependent. Further, the demands for those manipulative processes cannot be deterministic or they will not create true aesthetic states. This so called aesthetic state can then solely be recognized in the final result created by these non deterministic and weakly determined processes.

$$Rep = \{\langle E_1, p_1\rangle \, \langle E_2, p_2\rangle \, ..., \langle E_n, p_n\rangle\} \tag{2.3}$$

Bense then describes a creative process as a sequence of selections from this repertoire (2.3) of elements E (colors, words, etc.) and their probabilities p.

Max Bense was a theorist and his works very hardly ever tested, other than in experiments which tried to synthesize art according to those rules and principles. He also constructed ways to fit his theory into higher levels using semiotics, which try to describe the connection between this information theoretic analysis and meaning.

Andre Abraham Moles was a contemporary of Max Bense and equally influential to the movement of *information aesthetics*, was a French communication theorist. He initially studied engineering and later lead the Hermann Scherchen Laboratory for Electro-Acoustic Music where his interest in music research must have developed. Finally, he was director of the Institute for Social Psychology at the University of Strasbourg.

Contrary to Max Bense, his approach to aesthetics was focusing on foundations of human perception, even though he as well used information theory as a tool. In his book *Theorie de l'Information et Perception Esthetique* published in 1958 [26] he had among others referred to Birkhoff, Fechner and Berlyne, and also Max Bense's students. These references emphasize the link to *experimental aesthetics*, psychophysiology and the school of Max Bense. However, contrary to Max Bense, Moles used mostly auditory senses and stimuli (music) for demonstration and reasoning of his theories.

Moles built his aesthetic theory centered around the term *originality*, which he uses synonymous to quantity of unpredictability, the quantity of information that he took from information theory. If a message is certain, it will not excite the receiver and therefore not adjust his behavior. If on the other hand a message is unexpected, the receiver will adjust his behavior. A message can be however original on the semantic level as well as on the aesthetic. He distinguished information values of messages, having both semantic and aesthetic parts. The existence of the latter, as he said gets clear when exhausting the semantic part of a message. An example Moles gave was that if one can learn a theatrical piece by heart, there is nothing unexpected and original left if watched again. Still everybody would enjoy it. Hence, there must be more than information on the semantic level. This space for aesthetic information lies in the degree of uncertainty of how symbols are communicated. For example spoken language is not exactly defined by it's acoustic signal, since each symbol (e.g. phonemes) can have a certain peculiarity, which depends on personal preference of the transmitter. This aesthetic part of the message is untranslatable and doesn't have a universal repertoire common to everybody. Instead, the semantic part does follow a universal logic and is translatable into another language or system. Within this model, Moles make a very clear distinctions between semantic and aesthetic information which implies that aesthetics is not about meaning. However, even though aesthetic information is purely depending on personal characteristics of transmitter and receiver and therefore hard to measure, he hypothesizes that it as well follows the laws of information theory.

Moles tried to develop methods to extract the aesthetic part of messages and measure it separately form their semantic value. He concluded that if one can extract the semantic information (notation) of a musical piece by destroying the aesthetic parts (e.g. orchestration, instrumentation, an so forth), one must be able to isolate the aesthetic parts by destroying the semantic information. The technique he inspected was *reversing*, which as he said keeps the aesthetic aspects of the message.

Measuring the relative amount of aesthetic information can be performed experimentally with an approximation procedure that involves the following three steps:

1. Performing filtering techniques which eliminate semantical information and sustain the aesthetic value.

2. Determining the maximum possible information by estimating the communication channel's capacity, which is the amount of possible symbols and their possible arrangements. Discrete symbols are resulting from

perceptual sensitivity thresholds; that is, spatially, acoustically and temporally.

3. Determining a message's redundancy by randomly destroying parts until (aesthetic) understanding is completely destroyed.

The last part lacks any methods on how to measure retained "aesthetic understanding". However, assuming it can be measured the results are amounts of aesthetic information or originality. Additionally, Moles describes the phenomenons of *banality* and *overwhelming* by lower and upper complexity thresholds of an individual, depending on her or his a priori knowledge of the message. To derive from that what is aesthetically interesting, there must be ideal levels of complexity which make the individual not apprehend the whole and neither make it loose track of it. An interval where the originality of the message is not exhausted.

Moles' apprehension of aesthetics was colliding with Bense's in many points. It is subjective and dynamic, depending on psychological parameters of individuals rather than properties of the objects in question. Both theories were lacking practical outcomes however.

2.2.3 Cybernetic Aesthetics

Frank and Franke summarized the field of *information aesthetics* [14] and tried to make the original, conflicting theories of their predecessors more consistent. Frank did not agree with Moles' distinction between aesthetic and semantic information, arguing that a composer wouldn't produce any aesthetics then (which would be the task of the interpreter). To Frank, aesthetic information lies in the process of obtaining simpler levels of consciousness from information already known. For example building words from characters and sentences from word sequences, where the understanding of each separate word is more complex than the general picture of the sentence. For Frank, the reason for the existence of this aesthetic information from a cybernetic viewpoint is seen as lack of explanation of human behavior only by means of survival, knowledge gaining and behavioral adjustment. The additional necessity for aesthetics comes from the desire for diversity.

Considering this, the measurement of information constructed by Frank is derived from Shannon entropy, with some adjustments to subjectivity. The probabilities of certain *situations* or signs to appear, are defined as the subjective probabilities resulting from what was already learned about such situations. Frank states that the subjective probabilities w_k converge (through the statistical learning process of the individual) towards global probabilities p_k. $1/w_k$ are usually smaller than $1/p_k$ and therefore subjective information is usually greater (Eqn. 2.4). This process reduces inter-subjectivity to a minimum, which results in virtually universal aesthetics as stated by Bense earlier.

$$\sum_k p_k \log_2 \frac{1}{w_k} > \sum_k p_k \log_2 \frac{1}{p_k} \tag{2.4}$$

In order for a piece of art, design, nature or mental structure to have aesthetic value it must be perceivable at multiple levels of abstraction (e.g. characters, words, sentences, etc.) and form unity at a higher level which can be realized in different ways (e.g. melodic interpretation of a musical piece). For the appreciation of beauty of an object, one is overwhelmed by the complexity on at least one level and able to fully capture it at a higher level. For instance, it is impossible to capture the full complexity of an orchestra piece in the frequency domain (I_{n-j}, $(n \geq j \geq 1)$). On a higher level (I_n) such as harmonic changes, information is reduced such that it *fits* into the viewer's capacity of short time memory (K). In this transition, the aesthetic *effect* lies.

However this most recent summary of this movement [14] didn't give any examples on how to measure the named subjective probabilities in order to calculate information complexity. Claimed outcomes of this new theory remain mainly in the field of educational cybernetics. There, aesthetic information describes the information of texts which content is already known, and research is in particular committed to it's learning supportive aspects. With the death of Max Bense, most of his students turned away from their research in aesthetics.

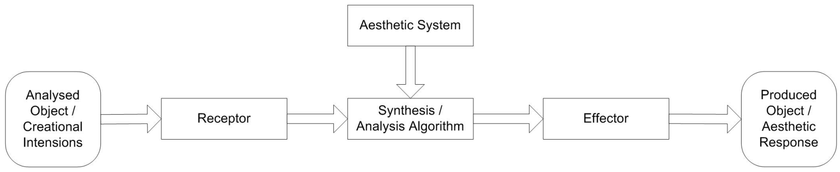

Figure 2.3: Algorithmic Design and Criticism System by Stiny and Gips

2.2.4 Algorithmic Aesthetics

The term *algorithmic aesthetics* first appeared in an equally titled book by Stiny and Gips [36]. To them, *aesthetics* meant the philosophy of criticism, and further how this criticism can be used to produce or design. They tried to describe algorithmic concepts necessary for criticism and design of artworks in hope for gaining a better understanding of aesthetics and providing a computational framework for aesthetic reasoning. A recent implementation of this concept was done by Tomas Staudek, who wrote a computer program called *Arthur* which can build and evaluate information-theoretical aesthetic measures according to these modules described [34].

The first idea is to build a dichotomy of design and criticism or creation and evaluation respectively. The model described depicted in figure 2.3, taking account of these two processes consists of a set of functionally separated algorithms and their input/output relationships. In the center of these modules is the *aesthetic system*, which is built from four algorithms: the *interpretation* and the *reference* algorithm, the *evaluation* and the *comparison* algorithm. This set of algorithms contains the aesthetic *knowledge* of the model and is meant as a metaphor to either the artist's or the art critic's conception of aesthetics. As the model is only a structural description of various possible algorithms, it might be explained best by one example, such as an aesthetic system implementing Birkhoff's measure of polygons:

1. The *receptor* could be for example a scanner, reading a polygonal shape from a sheet of paper and converting it into vector graphics, which will be the *representation* of the object in question.

2. The *interpretation* algorithm take this representation as an input and outputs a vector of properties as defined by Birkhoff (number of lines,

number of symmetries, and so forth). The *reference* algorithm provides a boolean predicate whether a given interpretation is consistent with a representation, which in this case would be obsolete.

3. The *evaluation* algorithm would be the simple Birkhoff formula outputting the ratio M and the *comparison* algorithm is basically only the $<, >, =$ operators.

4. The *analysis* algorithm could for example coordinate inputs of several polygons, call the aesthetic system and send outputs to the effector, which could be a simple *printf* to a screen.

Stiny and Gips contributed models for algorithm design but did hardly contribute to any new insight towards quantification of certain aspects of aesthetics. However, their work is interesting and historically relevant, since it *summarized and categorized* many theories and findings of computational aesthetics. They were also the first who used the word aesthetics most directly connected to computer science methods.

2.3 Practical Attempts

Contrary to the last sections, the following sections describe practical findings where algorithms and implementations are available and which are based on more recent scientific methods.

2.3.1 Aesthetics and Compression

Machado and Cardoso, who were aware of Moles' work, published a more modern and practical implementation of a direct aesthetic measure [25]. They started off with two main assumptions:

1. Aesthetic is the study of form rather than content, and for estimating the visual aesthetic value of an artwork, the latter is dispensable.

2. The visual aesthetic value depends on visual perception and is therefore mostly hardwired and universal.

Using the example of fractal images, which can look rather complex but are mostly represented by a simple and compact formula, they build a dichotomy of complexity. Firstly, the retinal image representation is highly complex but (analogously to the visual system's preprocessing) the internal representation is simple. An image of high aesthetic value is such, that it is highly complex (*image complexity*, IC) but easy to process (*processing complexity*, PC). This ratio (2.5) is similar to Birkhoff's aesthetic measure (Eqn. 2.1).

$$M = \frac{IC}{PC} \tag{2.5}$$

As an extension, due to the assumption that while viewing an image one finds increasingly complex levels of detail, processing complexity at several moments in time is taken account of (Eqn. 2.6). This represents a similar concept like Frank's aesthetic transition between multiple levels of detail (Section 2.2.3).

$$(PC(t_1) - PC(t_0)) \tag{2.6}$$

Implementation of these estimates (IC and PC) was done using JPEG and fractal image compression using several compression parameters for the latter, in order to simulate the different moments in time. The estimates where defined as the RMS Error over the compression ratio, which should represent *compressibility*.

Evaluation of the estimator was done using a psychological test for drawing appreciation, which was designed to evaluate an individual's level of artistic aptitude, reacting to aesthetic principles such as: rhythm, symmetry, balance, proportion, etc. The algorithm achieved an average score higher than fine art graduates.

Along with this idea, a few immediate questions are left open. First of all, the image complexity (IC) is argued in a way such that it implies *objective* complexity similar to Max Bense's, which is independent from visual processing. Secondly, the method of evaluation reacts to rhythm, symmetry and so forth. Since fractal image compression does directly take advantage of such redundancy (exploiting self similarities), the high test score seems obvious.

2.3.2 Aesthetics and Fractals

Significant findings of relationships between aesthetic preferences and quantifiable phenomenons have been reported in the area of fractal analysis. Spehar et at. have performed tests of human aesthetic preference of several types of fractal images (natural, artistic and artificial) in connection with fractal dimension [32]. The images used for analysis were (1) natural fractals, such as trees, mountains and waves - (2) mathematical fractals and (3) cropped regions of paintings by Jackson Pollock.

The measure calculated for comparison was *fractal dimension*, which is a property widely used for the analysis of fractal signals and images and describes the scaling relationship between patterns at different magnifications. A rough definition of the fractal dimension as given in [32] is such exponent D in $n(\epsilon) = \epsilon^{-D}$, with $n(\epsilon)$ being the minimum number of open sets of diameter ϵ required to cover the set. For lines $D = 1$, for planes $D = 2$ and for fractals patterns $1 < D < 2$. For approximately calculating the fractal of images (binary bitmaps), a standard algorithm know as *box-counting* is applied. A virtual mesh of different scales ϵ is laid over the image and the number of squares occupied $n(\epsilon)$ is counted for each ϵ. When plotting $-\log(\epsilon)$ against $\log(n)$, fractal patterns should result in a straight line of gradient D.

For Jackson Pollock's paintings, this scaling relationship was observed in a range of $0.8mm$ up to $1m$, showing the patterns to be fractal over the whole image.

Human values judgements of the three classes of fractal of various dimensions was collected using *forced-choice paired comparison*, where each possible pair of two images in the set is presented next to each and the proportional frequency of each image being chosen is taken as a measurement for preference.

Results have been compared to aesthetic preference tests on random patterns, to show that they do not depend purely on image density. Figure 2.4

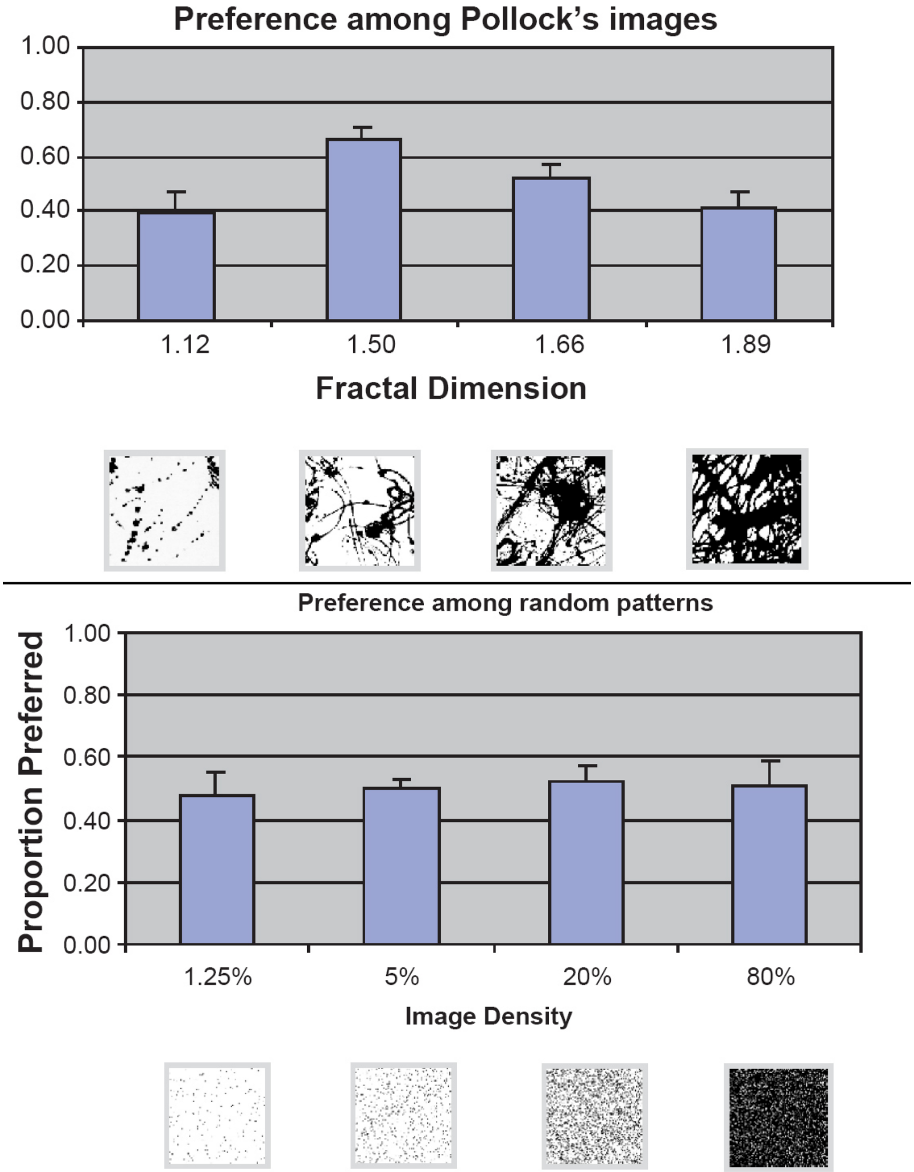

Figure 2.4: Aesthetic preference of Jackson Pollock's paintings as a function of fractal dimension versus aesthetic preference of random patterns as a function of image density (*figures taken from the original paper*).

shows increased preference for patterns of particular fractal dimensions.

This measure can be seen as some kind of estimate of complexity or order of a pattern respectively and therefore is not far off the history of quantitative aesthetics.

In a more recent psychological study about fractals and human perception, the term aesthetics was raised once again:

> "The aesthetic appeal of fractals can also be considered within the framework of more traditional theories of aesthetics. Although the nomenclature varies between different research disciplines, aesthetic appeal is frequently presented in terms of a balance between predictability and unpredictability of the stimulus...It is possible that the intricate structure and apparent disorder of fractal patterns might provide the required degree of unpredictability while the underlying scale invariance establishes an order and predictability." [38]

2.3.3 Generative Aesthetics

An area of research and a testbed for metrics of aesthetic features is (automatic) computer generated art. Rules that try to define a certain style and certain aesthetics are incorporated into *evolutionary algorithms'* fitness functions and generate whatsoever media (generally images or music). Even though research almost never involves a scientific evaluation of such a system's outputted aesthetic values, it can be empirically verified with artists' judgment. Since at least in the freedom of the art world this is sufficient, many researchers have committed themselves to developing new techniques of computer generated artworks.

Basic problems in generating images by evolutionary algorithms employ the choice of image representation (pixels, shapes, etc.) and the evaluation of the fitness function. In many cases pixel bitmaps are not well suited as an encoding for each individual image, so most of the time parametric representation are preferred. The fitness function is usually desired to represent aesthetic appeal or value of the particular image. In interactive evolution, a

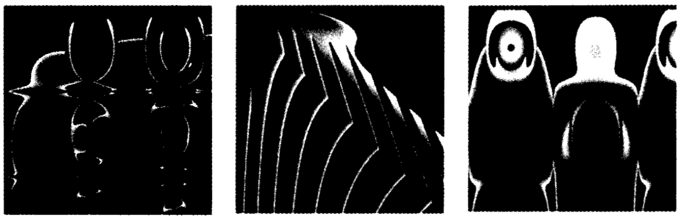

Figure 2.5: Images that resulted from artificial evolution *taken from the original publication*

human evaluator can rate images in order of his aesthetic preference and the algorithm can evolve new images from this judgement. Automatic evolution on the other hand requires some objective function to at least simulate this aesthetic choice.

Baluja et al. [2] have attempted to capture an individual user's aesthetic preference during interactive image evolution with a neuronal network, which was then used to simulated the users choices in automated evolution.

In their approach, images where represented through symbolic prefix-order expressions, taking x- and y-coordinates as arguments. For example: $sub(log(x), avg(sqrt(y), log(x)))$. The genetic algorithm's cross-over operator was defined in a way that a random node in each of two recombining composition trees is chosen and the subtrees are copied from the source to the target. For preserving diversity in the search, a constant mutation rate of one mutation per individual per generation is applied, which simply swaps function arguments. The interactive evolution was realized with a graphical user interface, where a test person could pick preferences throughout a desired amount of generations. This is a very typical setup for interactive image evolution.

The difference in their research was automation, where a neural network was trained upon *raw* pixel data of 400 user-ranked images. The outcomes where, obviously only visually evaluated. Some results of automatic evolution controlled by the trained neuronal network are shown in figure 2.5.

The succeeding chapter will commit to a digestion of the above works, focus on visual perception and aesthetics, which will lead to the experiments.

Chapter 3

Perceptual Models

The historical survey of computational aesthetics in the last chapter has shown a variety of approaches towards models of quantitative aesthetics spread out over many disciplines. From the concepts of aesthetics, as inherited by philosophers, initial psychological experiments have led research in a direction of objectivity, where mathematicians and computer scientists have picked it up and tried to formalize it. Later, information theory has played an important role in this area and various concepts of complexity where considered being connected to aesthetics.

This chapter will present a digestion of the basic concepts that were considered in the past and interpret them in the context of current research of human visual perception. From this, a new hypothesis is built which connects aesthetic experiences with various complexity measures in different parts of vision processing.

3.1 Aesthetic Dimensions

Most authors of aesthetic theories by themselves did not really develop a solid methodology for a measurable, quantitative aesthetics. However, they did make clear some aspects of it in which objectivity may be found. It can be observed that these concepts where trying to describe aesthetic values as a particular state or configuration of certain properties of the objects or their perception, that is closely related to the concept of *complexity*. This ultimate state of perception was generally approached from two sides:

1. minimize complexity
2. maximize order

In an information theoretical sense, these two directions seem ambiguous, because order (redundancy) implies the absence of complexity. However, the use of these heuristics can be interpreted taking up two positions.

On the one hand these theorists tried to confine the range of possible levels of complexity in which aesthetic states may reside, while knowing little about what increases perceptual complexity and little of what reduces it. This could be providing remedy for not knowing any canonical measures. In other words, by measuring complexity and redundancy, an estimate can be provided wherever there is no analytical way to determine complexity.

On the other hand, the concept of aesthetics was very often decomposed into more tangible factors. For example in Birkhoff's measure, *complexity* was the effort of motorial adjustments in the observer. *Order* was the presence of several perceptual forms which are known a priori, and therefore less complex (more predictable). This demonstrates the idea of complexity measures on different levels of abstraction.

Together this leads to a multivariate interpretation of perceptual complexity.

3.1.1 Complexity

The term complexity has defined in various contexts by different sciences. In mathematics, it represents the difficulty of solving a problem. More specifically in theory of computation, it is the study of how much time and memory (space) an algorithm requires to solve the problem. Similarly, information theory deals with complexity of strings, which can be seen as the difficulty to compress them. A frequently used indicator is the *information entropy* or in algorithmic information theory, the length of the shortest program which generates the string.

All these uses of the term complexity are defined for a particular scenario, which is quantifiable and well defined. Such as for example its applications in binary communication or compression. Complexity as a measure of aesthetics however is lacking the clear task. Some of the authors, like Max Bense, tried to find an objective, universal aesthetics that is a pure statistics of the objects themselves. Birkhoff, seeming to also have slight tendencies towards objectivity, however defined complexity as something related to personal effort. Later developments of information aesthetics where applying complexity measure to cybernetic concepts, especially learning. Three views of complexity can be observed in the last chapter.

- "Objective" complexity: Complexity is measured as a function of statistical or geometrical properties of an object.
- Complexity based on a communication model. Usually the subject is the receiver and his environment the transmitter. The receiver's memory is seen as the receivers repertoire of symbols and complexity is based on unpredictability of messages for the subject.
- A computational model of complexity. The subject's perception is seen as a "machine" which requires various processing complexities for different message structures.

3.1.2 Order and Form

If aesthetics solely depended on single measures of complexity, a plain white canvas would be the most beautiful and the pure visual *noise* the ugliest picture (or vice versa). This is obviously untrue. Therefore, scientists have frequently used the term *order* in the course of finding explanations of when certain levels of complexity are more appealing.

Max Bense took Birkhoff's formula, using statistical redundancy in place of *order*, reasoning that it represents identifiability, the *known*. To him a creative process was an "ordering process".

Moles associated the concept of order with form, the problem figure and ground separation or signal versus noise respectively. To him it was represented by *redundancy*, which is caused by a perceiver's *a priori* knowledge of a received stimulus and keeps complexity down to an interesting or aesthetically pleasant level. More precisely he related order to the degree of predictability and internal coherence, expressed by *autocorrelation*.

In Birkhoff's aesthetics the role of order was to perceptually *reward* the effort of focusing attention on something complex. He assumed that there exist elements of order such as *symmetry*, *rhythm*, *repetition*, etc. which psychologically cause a positive tone of feeling, and also elements that cause negative tones, such as *ambiguity* or *undue repetition*.

The recent approach done by Machado and Cardoso, who tried to apply *fractal image compressibility* as an element of order in their aesthetic measure, assumed that self-similarities can be more easily perceived. They follow a similar argumentation as Birkhoff, however using computational methods for measuring.

Similarly, in the article *Universal aesthetic of fractals* [32] a direct comparison of *fractal dimension* and human aesthetic preference has shown another type of redundancy. The scaling relationship.

Another aspect of order is found in color research. Color perception is far from being trivial and further it is sometimes regarded as one of the most important factors for aesthetics. Antal Nemcsics has developed a color order system named Coloroid [27]. In essence, it is a color-space that is supposed to be uniform in aesthetic distances rather than in perceptual differences. This should allow measuring of color harmony, an element of visual order.

Additionally, empirical work on concepts of order (e.g. symmetry, equilibrium, rhythm, etc.) are found in Arnheim's *Art and Visual Perception* [1]. In this book he defined an analogy of visual patterns to physical systems and derived a set of laws, which are relevant to perception of art. His work is commonly taught in art schools and could be a guide to identify perceptual elements of order.

On the bottom line on can see that many authors put the term *order* into an important role in aesthetics, and which is most of the time interpreted as an opposing force to *complexity*. Even though this seems ambiguous, it showed the repetitive interest in *what* reduces complexity. Is it a color gradient or a well recognized shape?

3.1.3 Learning and Dynamics

Another concept observed which was related to aesthetics is learning and the dynamics of perception. From human intuition, we can immediately agree that what we think is beautiful is connected to our experiences, i.e. what we have learned. One inspiring example would be listening to a song. Sometimes when hearing a particular song for the first time it can seem uninteresting and even unpleasant. After a few temporally distributed repetitions of hearing it, it suddenly becomes beautiful.

In Moles' information theoretical model of aesthetics, the concept of memory represents the important role of influencing perceived redundancy and therefore also the quantity of aesthetic information. Following the fact that human memory (i.e. the repertoire of elements for communication) changes dynamically, he introduces the term *differential information*. Also later cybernetical aesthetics relied on this concept of differential aesthetics.

Even if it seems natural to search for aesthetics in the dynamics of subjective perception and learning, an inter-subjective aesthetics could be found in "hard wired" parts of perception, such as in parts of *early vision*.

The following section will describe key findings in visual perception and their computer vision equivalent in order to then come up with a perceptual and computational interpretation of complexity.

3.2 Visual Perception

Visual Perception allows an individual to gain information about the environment by detecting objects and structures in patterns of light reflected by surfaces. This process appears to occur nearly instantaneously in the brain, but is relatively complex to model. In *Visual Perception* [41], Steven Yantis describes some key progress of vision research during the last century, which addresses several problems along the path of visual processing. Consistently to how computer vision [13] has divided it's model and problems, it can be categorized as follows:

- Early Vision

  Encoding of elementary sensory input

- Mid-Level Vision

  Perceptual constancy - correction for viewing conditions

  Perceptual organization - grouping of low level features into surfaces and objects

- High-Level Vision

  Object recognition - categorize and identify objects

- Attention

- Awareness

For some of these problems, such as in early vision, there is already a good foundation of knowledge, but for the higher level aspects of vision there is only very limited insight. In any case, two important principles about vision are known. First is functional specialization, which occurs throughout the whole path of visual processing. This starts at the retina, which consists of four distinct types of photosensitive cells, arranged in ways to suit different tasks such as for example the rods are relied on for night vision. At the next higher level, there are distinct types of optical nerve bundles, specialized for spatial, color and temporal codings respectively.

The other principle known is distributed coding, which is observed in many stages as well. One example would be the representation of an edge in the visual field. Rather than a single cell responding to the occurrence of one particular edge, it is a population of cells sensitive to different orientations that describe or detect the edge together.

These principles help in understanding vision as well as implementations of computational models for image analysis and computer vision.

3.2.1 Early Vision

Starting from the visual patterns of light projected on the eye's retinal array of cells, early vision models a representation of simple features such as color, orientation, motion and depth. These features tend to be local and very tightly connected to the retinal image. In other words, local spots on the retina having corresponding cells in the primary visual cortex. This is called *retinotopic mapping*. Most of these aspects have been extensively studied and are well know.

Most importantly, at the very first stage of vision processing the retinal ganglion cells (RGCs) perform difference computation on the array of brightness values from the retina at specific locations in all directions. The applied contrast sensitivity function has the shape of a Gaussian, which is also used in image analysis for edge detection [15]. The output of the RGCs connects (via an area called lateral geniculate nucleus, which is rather unknown) to the primary visual cortex (named V1), where the cortical cells specify whether a given position contains an edge, basically throwing brightness information away. The resulting representation is a description of edge locations and orientations.

Another feature the primary visual cortex encodes is *scale*. It consists of cells which are tuned to various spatial frequencies, or in other words, different sizes of receptive fields. This is of importance, considering the mathematical fact that any visual pattern can be represented by sums of different frequencies (e.g. Fourier analysis).

The third important aspect is the coding of depth information available from linear perspective, overlap, relative size and relative motion in the retinal image. It is know that this information, especially stereoscopic depth cues, are incorporated very early in the visual cortex and combines with the above features to achieve object detection.

Early vision is also dealing with color. As for object recognition, the most important feature seems to be shape, which can even help to guess it's color. Knowing an object's color alone however doesn't tell anything about what the object is or what it's other properties could be (according to [41]). Still it's an important feature that helps detecting boundaries of objects and material surfaces. What is very well know about color perception are the spectral sensitivities of the retinal cones and how they combine. Together with the opponent process theory of color perception [19], a range of phenomenons (visual illusions) can be explained, such as for example:

Color adaption After looking at a red patch of color for a few minutes, looking at a white surface it appears green.

Simultaneous color contrast A gray patch will appear slightly green when surrounded by bright red color.

Mach Band effect At the boundary of two patches of highly different intensities (steps), the region close to the edge looks lighter on the dark side and darker on the bright side than their patches intensity respectively.

The last spatial feature which caught a lot of attention from vision researchers is motion. It is known that there exist complex cells in the primary visual cortex which are sensitive to direction of motion as a function of orientation. Even more complex cells respond to motion in their receptive field as a function of *both* orientation and scale. Retinal image motion is caused by motion of the eyes, motion of the observer and motion of the observed object. A still image most of the time is not present. This arises the need for a system which compensates for this motion in order to perceive stable images. This is an open issue for both computer vision and human vision research.

A currently accepted model of early human vision which is also used as a basis for computer vision (texture description in particular) is a multiresolution model (see figure 3.1). In this model the retinal image is split into several sub-bands and each of which is convolved by linear filters of several orientations, which look similar to derivatives of a Gaussian. The resulting maps are subjected to nonlinearity in order to take account of complex spatio-temporally sensitive cells. The output (including added noise) is then passed on to the next higher level of vision.

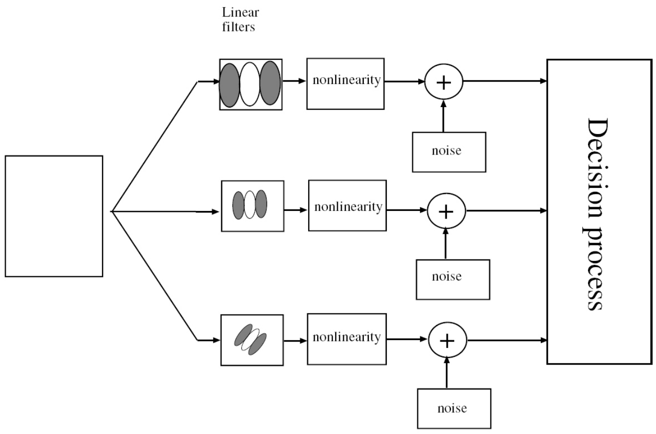

Figure 3.1: Simple multiresolution model of early vision (with the number of filters simplified to three).

3.2.2 Mid-Level Vision

Retinal images in most situations not just measure properties of the objects viewed, but also irrelevant and transient factors such as illumination, distance and orientation. Despite of these constantly changing conditions, most of the time the visual system is still able to independently tell the correct object properties like color, size and shape. In vision this is called the problem of perceptual constancy and is only very little known. It has been also addressed in research of image understanding and computer vision [9].

The second problem in mid-level vision is the problem of perceptual organization. In particular grouping, segmentation and completion. Since features described in early vision are local, they need some form of integration in order to detect and classify objects. Even worse, in real scenes most objects are occluded by others, but human vision is still able to capture the shape from local features of these objects in many cases. This hint for some complex mechanism incorporated in higher levels of vision which perform grouping in a similar way as described by Gestalt theory [39].

Gestalt psychologists have developed a set of rules of when local features are grouped, as shown in figure 3.2. These rules can be easily agreed, but finding a way to implement this theory in computer vision is hard, because the examples used to demonstrate the rules are extreme cases and there is no

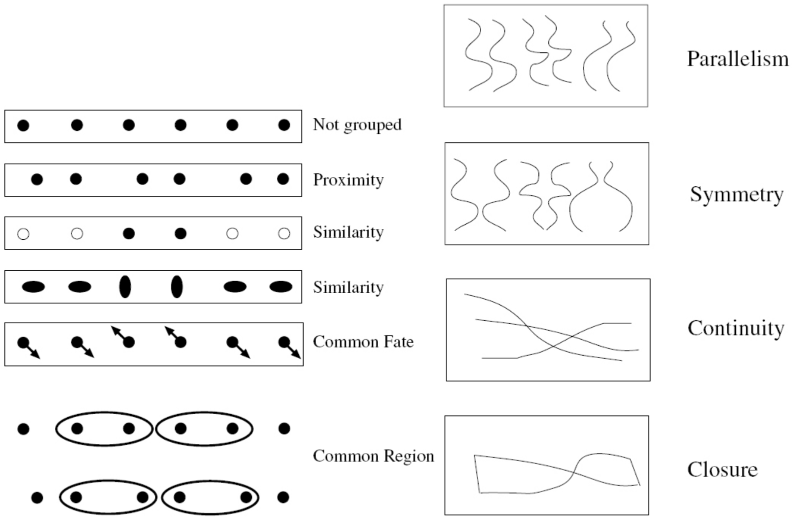

Figure 3.2: Grouping rules from Gestalt theory (from Gordon, Theories of Visual Perception, page 67)

laws on when which rule applies. Especially the rule describing that if features form a recognizable *figure*, this grouping is preferred. It introduces the problem of *figure and ground* separation, which in many cases is ambiguous. Still, these Gestalt rules give some insight into perceptual organization.

A closely related phenomenon of vision is perceptual completion. The disks and lines shown in figure 3.3 are rather seen as completed, occluded by a white triangle. The triangle doesn't exist however in a sense of physical contrast to the background. Some part in the visual system fills in additional contrast for the white triangle to the background. Vision has found evidence for this phenomenon in the secondary visual cortex (V2) [41].

Computational Models of mid-level vision mainly address the problem of segmentation which can be formulated as follows: Given a shape or structure, determine whether a pixel belongs to it. The difficulty lies in the fact that the shape is not a priori known by the machine. However, there has been effort to build image segmentation algorithms that work similar to the laws of Gestalt grouping [28].

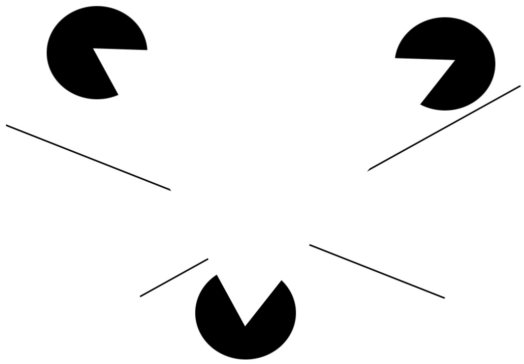

Figure 3.3: Kanizsa Triangle, can be seen even if it's brightness is the same as the background.

3.2.3 High-Level Vision

The next higher level of visual perception is the recognition of objects. It's difficulties are obvious, considering the almost infinite number of retinal images a simple object can produce. Obviously, the visual system can still recognize a large set of objects invariant to rotation, lighting conditions and so forth. There are two major ways of thinking in vision research trying to explain this concept.

First group of models for object recognition employ the idea of *template matching*. In it's trivial form it would require all possible views of an object to be stored in memory, which appears highly implausible and there is no neurological evidence for that. An extended model is, that vision normalizes the object for size, viewpoint, etc. and matches it against a few stored canonical views, which together describe the object. This requires much less memory and results in a more plausible theory.

The second group of approaches of object recognition describe that it is based on the statistics of features or parts the objects consist of. For example the presence of parallelism and certain colors and shapes can describe a certain class of objects. A major part-based model of recognition is called *structure description*, where objects are build from a set of three-dimensional primitives, e.g. boxes and cylinders. Once such a description is extracted form a retinal image, it is viewpoint independent and easy to match against memorized complex objects of similar descriptions.

Vision research suspects that the ”correct” model might be again a mixture of both approaches. In computer vision however, object recognition based on template matching is popular, but recognition based on structural descriptions has been used as well.

3.2.4 Attention

Sensory input apprehended coherently at any given time is much too diverse for a human to process. When running through a crowd of hundreds of people looking for one particular person, vision couldn't process each of the peoples faces, nor the cloth they're wearing, their hairstyle and so forth. In spite of the heavy load of visual information present, the visual system can select what is important and what enters awareness. This process is called *perceptual selection*.

One theory in visual perception states that object recognition is performed at a stage called *feature integration*. Low-level features obtained by early visual processing (orientation, color, etc.) are connected to build objects, but this can only happen in one location at a time. This theory idea is accepted in vision research and also neurological evidence has been found.

In computational models, a very popular approach is called *saliency maps* first described by Koch and Ullman [21]. The idea is centered around a retinotopic map which encodes the locations of visual stimuli that *stand out*, are *salient*. This map provides bottom-up information for guidance of visual attention.

3.2.5 Awareness

The last stage of visual perception research, and also the most unknown is awareness. Views on how consciousness is achieved from attended perceptions are very abstract and most of the questions are still left to philosophers. Tractable problems are limited to experiments of what neural activity corresponds to what we can express or communicate (e.g. verbally). It is a search for correlations between neural and neurochemical activities and conditions that cause states of awareness. Since this a new area of investigations and outcomes are only a few, there obviously is nothing coexistent in computer vision at the time of writing.

3.3 Towards Aesthetics

The forefather of empirical aesthetics, G. T. Fechner [11] described the new aesthetics research as being necessarily *bottom-up*, while top-down approaches are mostly left to philosophers. The major difference is that bottom-up explanations of aesthetic phenomena are built on top of stimuli. Rules combining them into higher level descriptions are investigated, to which aesthetic preference can be related. The top-down approach starts from the phenomenological concept of beauty and tries to break it down into situations where it holds or doesn't. According to Fechner, this top-down approach still helps guiding the preferable bottom-up path. Birkhoff's aesthetic measure can be also interpreted in this manner, approaching aesthetics from *two directions*.

Taking the bottom-up path of explaining visual aesthetics, the most plausible way seems to start with the visual stimulus entering the perceptual system through the eye, going through various stages of processing until awareness is achieved. How does this process relate to aesthetics?

3.3.1 Computational Model Postulate

Remembering the aesthetic theories in the last chapter, the following categories of aesthetic measures occurred:

- Aesthetics as a function of various object properties and/or perceptual properties (*Birkhoff, Machado*).

- Aesthetics emerging from certain (pleasant) levels of complexity and redundancy in a model of communication (*Moles, Bense*).

- Aesthetics as a function of different, perceived complexities at various levels of abstraction (*Frank and Franke*).

Together with the bottom-up idea of Fechner and the last section on visual perception, one can derive the following postulate.

Visual aesthetics can be described as a function of *complexities* at *multiple and distinct levels of visual perception processing*. Complexity can be estimated either *directly* (e.g. disorder or energy), or *inversely* by determining known elements of order (e.g. redundancy, form).

This formulation of the problem is based on two concepts which require further explanation and refinement. First of all, conditions need to be set under which a visual aesthetic experience is found and what is exactly meant by this. Further it requires a definition of perceptual complexity, based on known concepts of visual perception at different logical levels of abstraction. The following section will address those problems.

3.3.2 Visual Aesthetic Experience

The prototypical visual aesthetic experience, as formulated by Kubovy in the encyclopedia of psychology [23], is such that:

- Attention is firmly fixed upon heterogenous but interrelated components of a visual pattern.

- Is not disturbed by awareness of other objects or events.

- Attention is not fixed upon an object's relation to ourselves, to the artist, or to culture.

- Feelings or emotions are evoked by the visual pattern.

- Experience is coherent, "hangs together".

- While in this state, the image is not perceived as a material object.

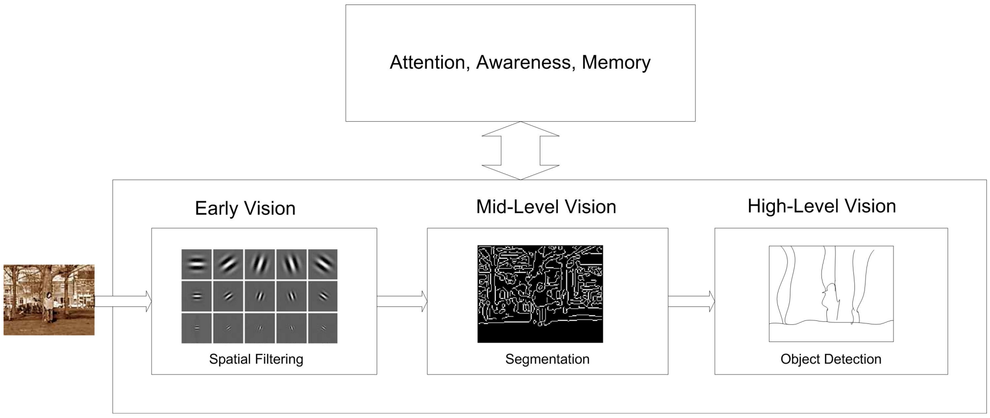

Figure 3.4: Simplified model of vision.

Assuming a group of people focusing on objects of the same class under conditions similar to those above, an inter-subjective comparison of aesthetic preference could be built. This assumption that aesthetics is not purely a social construct will remain for the rest of this thesis.

3.3.3 Perceptual Complexity

Provided the visual system can be seen as a "machine" processing the retinal input, visual complexity can be defined as a function of time and energy consumed, in order to process the image. To build according estimates for different levels of abstraction, a processing model of vision is constructed.

Figure 3.4 shows a simplified model of the human visual system, where the representations of the visual stimulus are given for each phase. In early vision it is given as outputs of spatial filters of different orientations and scale. Mid-level vision deals with texture regions and shapes segmented from the image, providing building elements of objects. Finally, high-level vision deals with objects, as they were identified from shapes and features. For different image statistics, each of these representations have various levels of complexity. For instance a cartoon drawing might be rather simple at the early stages, however contain many diverse shapes and objects. Contrary to that, a photo of one tree standing on a wide meadow might be complex in texture, but the image is containing only a few shapes and only one or two objects.

In order to build a computational model of aesthetics, definitions of complexity need to be constructed for each of these levels of abstraction.

Complexity in early vision could be seen as function of cell spiking rates, which are modeled as a set of spatial filters. The filter output might be taken as the energy required by the cells to fire at this rate, and the total energy of filter responses as an estimate for complexity. However, other estimates are possible. This complexity can be interpreted as *image quality*.

On the segmented image, simple complexity measures such as shape count, and shape complexity (symmetry, etc.) seem a good starting point. How well this represents visual perception however depends on the quality of the segmentation algorithm. On this level, the aesthetic interpretation could be the *quality of shape and forms*.

As for object recognition, finding even a simple complexity estimate seems more challenging. Object identification is tightly connected to a priori knowledge. Also there is very little known from visual perception on how this process works in the visual system. Therefore object complexity seems impossible to measure on a perceptual level. The simplest estimate would be however *object count*. At this stage, the interpretation from an aesthetic point of view would be *composition*.

3.3.4 Deriving a Hypothesis

An aesthetic model for predicting visual preference would need to calculate complexities at each stage of vision. Taking into account that abstract art usually excludes the high-level object detection mechanisms of vision, it's aesthetic values could be found within the earlier stages of vision. First logical step towards such a model is to test, whether the lowest level image features correlate significantly with human preference.

Hypothesis *Visual, aesthetic preference* of images of similar classes is not purely subjective and depends to some extent on the *amount of cell activity in V1*.

The following chapters will provide an implementation of an estimate of this cell activity (complexity), which will be later used to verify our hypothesis against aesthetic preference ratings of a photo database described in the next chapter.

Part II

Experiments

Chapter 4

Photography Database

This chapter describes the data set used in this thesis experiments, which was obtained from the photography platform *photo.net*, which contains an enormously large amount of images numerically rated for aesthetic preference by it's users.

> As for October 2003: "680,000 gallery photos, with 40,000 new photos submitted per month and 2.7 million photo ratings, with 210,000 new ratings per month".

These images cover a wide range of artistic qualities from simple family snapshots to professional photographs and fine arts. Such a database seems an interesting starting-point to find relationships between those ratings given and image features.

The raw data collected is analyzed using basic statistics and interpreted. In the last section, some pre-filtering is described based on the initial interpretations and irrelevant images are thrown out.

4.1 Legal Issues

In the context of this purely academic work, the site's terms of use and the individual copyright holders were fully respected. Since the data was collected using automated computer programs which the site doesn't allow in general, they were built in a way such that they do not cause more site traffic than a typical human would cause using a web-browser. Some of the site's photos are printed in this thesis to demonstrate some principles and outcomes of this work. This is done in hope that each author will grant permissions to do so.

4.2 Obtaining Data

For extracting photographs and their related user-ratings, a Perl script was developed which emulates a user browsing the site and looking through the galleries. Images where stored in JPEG format as downloaded from the site and ratings and image ID's where stored in comma separated text files to be loaded into Matlab.

4.3 Data Statistics

The total number of photographs collected is 1958 (with a portion of 147 black and white images), which resulted from a one week run of the download script. The script was running slow in order to respect the site's policies and was then terminated for time reasons. Selection of the images from the site was purely random (randomly selecting photo ID numbers) and includes both newly submitted entries and photographs which have been online from months to years. At the time of data collection there were no content categories (such as 'People', 'Architecture', 'Nude', etc.) assigned to the photographs, so category selection was also random. All images have about the same resolution.

4.3.1 Image Properties

Table 4.1 shows basic statistical properties of the data set's images and figure 4.1 shows their histograms. A very typical resolution for web presentation can be observed, lying between 0.2 and 0.5 mp. However, a few outliers exist. As for the image aspect ratios, there are three major classes: (1) landscape

Property	Min	Max	Mean	Median	StdDev
Resolution (megapixel)	0.0764	1.6031	0.3375	0.3158	0.1085
Aspect Ratio	0.4080	5.9823	1.2063	1.3180	0.4026
Compression Ratio	0.2399	89.9076	16.9542	14.2045	12.0483

Table 4.1: Image Statistics

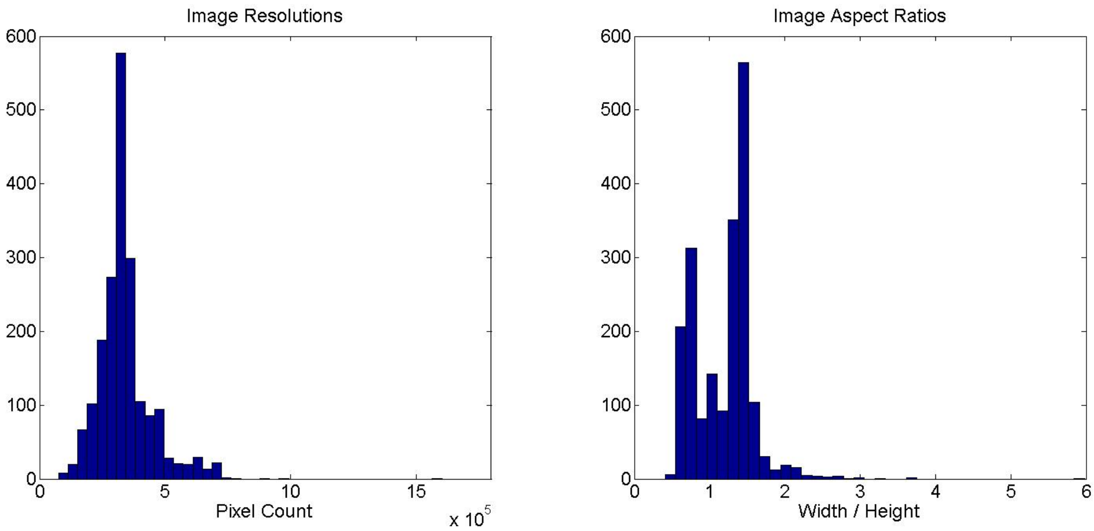

Figure 4.1: Histograms of resolution and aspect ratios.

format, (2) rectangular and (3) portrait format. They can be seen in the histogram in figure 4.1 as three peaks >> 1, 1 and << 1. Most frequent ratios seem to be landscape formats around 1.5. The outliers are panorama photos and other non-typical photo artworks.

4.3.2 Rating Statistics

Ratings for the site's gallery photos are given in two categories: 'originality' and 'aesthetics'. Instructions for what these categories mean are given to the user on the web-site (*http://www.photo.net/gallery/photocritique/standards/*):

> "Aesthetics: Give a picture a high rating for aesthetics if you like the way it looks, if it attracts and holds your attention, or if it conveys an idea well ... Some ways of judging photographs cut across all genres. Does the photo have an interesting composi-

> tion? An effective balance of colors? But most genres also have specific standards of appraisal. For example, with portraits, give a high rating if you think the subject has been captured effectively and if the subject's personality is communicated ..."
>
> "Originality: Give a photo a high rating for originality if it shows you something unexpected or a familiar subject in a new, insightful, striking, humorous, or creative way. Originality can be very subtle. A photo does not have to be taken on Mars to be original. For example, can a portrait be original? Hundreds of millions of portraits have already been taken. But a portrait can be original if it uses expression, lighting, posing, or setting in a creative, sensitive or humorous way to reveal the character or situation of another person."

The two ranking scales provided are expressed both numerical and nominal, and range from 1 to 7 or "Very Bad" to "Excellent" respectively. The instructions given to the users on the site are stated as follows:

> "You can use the photo ratings system to critique photos along two dimensions: 'Aesthetics' and 'Originality' On each dimension, rate the photos from 'Very Bad' (1) to 'Excellent' (7). Photos should be rated relative to other photos on photo.net. Excellent/7 does not mean 'one of the best photos of all time'; it simply means one of the best photos of its type on photo.net. Your most frequent rating should be a 4, and most of your ratings should be 3 to 5, with progressively fewer being rated 6 and 7 (or 2 and 1)."

For the sample of 1958 photos collected from the site, these ratings employ the statistics shown in table 4.2 and figure 4.2. The rating values are averages of the users' individual ratings per image.

The range of rating counts per image is relatively large, because some images attract more attention than other and image with lots of high ratings are shown more frequently on the site's "top photos" gallery. The range of

Property	Min	Max	Mean	Median	StdDev
Number of ratings (N)	3	280	23.4219	10	36.3457
'Aesthetics' Rating (A)	2.3300	6.8400	5.2148	5.1700	0.8954
'Originality' Rating (O)	2.6700	6.8700	5.1703	5.1000	0.8628

Table 4.2: Rating Statistics

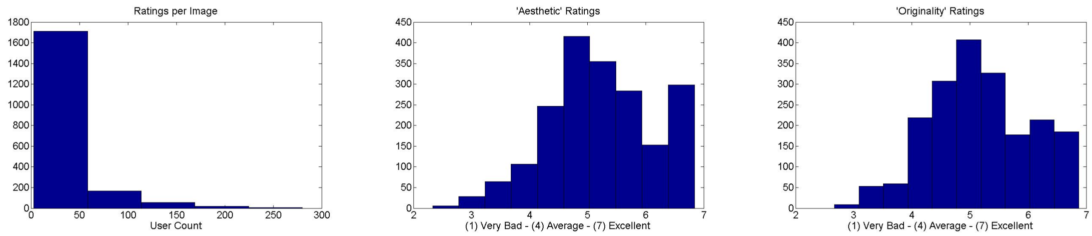

Figure 4.2: Rating Histograms

the (A) and (O) ratings does not cover the full theoretical range and there is nothing below 'Bad' and there is a majority of ratings centered around 5 ('Good').

Not surprisingly, the three rating properties show high correlation, since the rating instruction described above are handled *very* freely and are not controlled. These relationships are shown in table 4.3 and figure 4.3. However, a combination of these value can be used and interpreted as an overall *image preference* in the context of this online photo community.

Pearson	N	A	O
N	1	0.6237	0.6311
A	0.6237	1	0.9305
O	0.6311	0.9305	1

Table 4.3: Rating Statistics

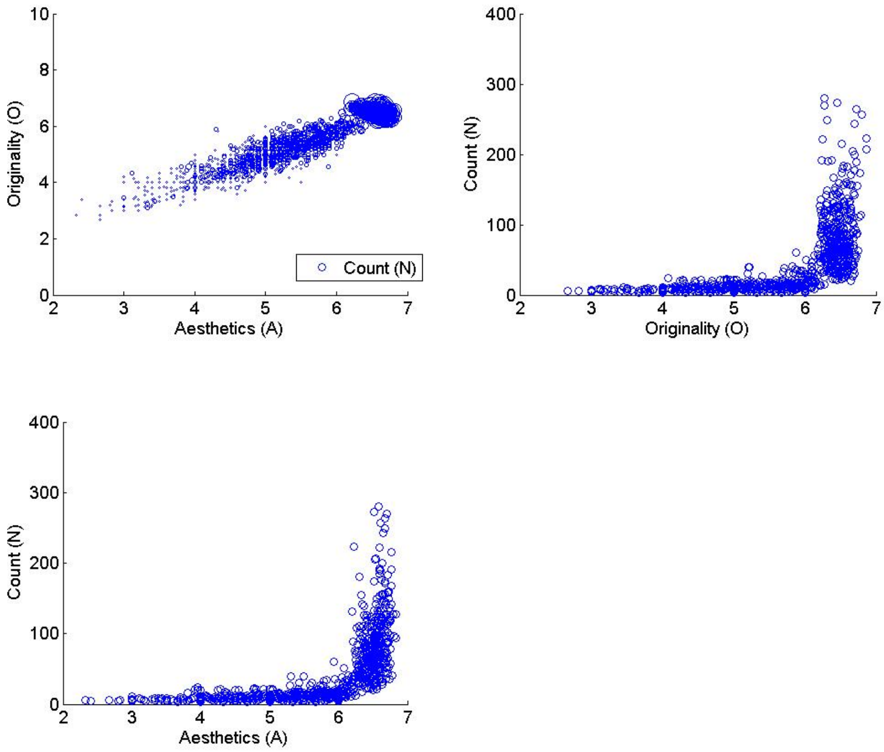

Figure 4.3: Scatter plots of relationships between A, O and N rating data.

Chapter 5

Visual Preference Experiments

In this chapter, two estimates of the amount of neural activity in early vision are described. The first estimate is based on the thoughts in chapter 3 and the second was used before in the context of visual aesthetics [25] and is used for comparison. They both exploit properties of perception which are based on findings in neurophysiology and should theoretically relate to the way aesthetics is experienced.

5.1 Multivariate Gabor Filters

Gabor filters are a linear filtering technique used in many computer vision tasks, such as face recognition or texture segmentation. They became popular because of their similarity to various cell behaviors of the human primary visual cortex.

A straightforward model for complexity in early vision stages would be the amount of cell activity, estimated by the *mean absolute gabor filter responses*, using a filter-bank that covers similar receptive fields as real cells do. Such a complexity estimate is built in this section, based on Gabor filters of various orientations and scales.

The Gabor function in the spatial domain is a combination of a complex sinusoid *carrier* and a gaussian *envelope* as defined in [22] as:

$$g_{\xi,\eta,\gamma,\lambda,\theta,\varphi,\sigma}(x,y) = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2})\cos(2\pi\frac{x'}{\lambda} + \varphi) \quad (5.1)$$

$$x' = (x - \xi)\cos\theta - (y - \eta)\sin\theta \quad (5.2)$$

$$y' = (x - \xi)\sin\theta + (y - \eta)\cos\theta \quad (5.3)$$

The parameters are (x, y) and $\xi, \eta, \gamma, \lambda, \theta, \varphi, \sigma$, where (x, y) is the position in the visual field (or image respectively) and the rest is described as follows:

(ξ, η) The position of the receptive field's center relative to the image position in image coordinates.

σ The standard deviation of the Gaussian, reflecting the size of the receptive field.

γ The aspect ratio of the elliptical form of the Gaussian (receptive field), found to be $0.23 < \gamma < 0.92$ in real cells and 0.5 is a typical value used in computational models.

λ The wavelength of the cosine factor which determines the preferred spatial frequency the cell is tuned to.

θ The orientation of the cell, which determines the response to oriented edges.

φ The phase offset of the cosine, which determines the symmetry of the filter with respect to the center of the receptive field.

A set of Gabor filters (see example in figure 5.1) at six orientations and five scales is constructed, utilizing that function. The image is then convolved with this filter bank resulting in a 5×6-dimensional complex valued feature vector for each image pixel. The sum of absolute values of complex numbers is taken over all orientations for each scale, and the mean value of each scale's responses is calculated. This is defined as:

Figure 5.1: An example of a Gabor filter in spatial (left) and frequency domain (right)

$$m_{\sigma_i} = \sum_j abs(\overrightarrow{I}_{\theta_j}) \tag{5.4}$$

5.1.1 Implementation

The filter response features defined above were implemented in Matlab using the *simple gabor toolbox* which is freely downloadable and helps in creating filter banks and calculating their filter responses for an image or signal.

Sourecode 5.1.1

```
function IC = build_icg(filename, b, s, o)

 % read image
 [I, map] = imread(filename);

 % if image is 8bit indexed, convert to rgb
 if length(map) > 0
     I = ind2rgb(I,map);
     I = cast(fix(I*255),'uint8');
 end

 if length(I(1,1,:)) > 1
     I = rgb2gray(I);
```

```
 end

 I = double(I)./255;

 % compute gabor filter responses
 bank=sg_createfilterbank(size(I), b , f, o);
 r=sg_filterwithbank(I, bank);

for i=1:o
 r(i) = abs(sum(r.freq{i}.resp));
end

 % image complexity
 IC = [mean(r)];
```

The filter bank is created by the `sg_createfilterbank` command and builds $f \times o$ frequency domain Gabor filters for the following parameters: $\sigma_i = \{b, \frac{b}{\sqrt{2}}, \frac{b}{\sqrt{2}^2}, ..., \frac{b}{\sqrt{2}^{f-1}}\}$ and $\theta_j = \{0, \frac{\pi}{n}, \frac{2\pi}{n}, ..., \frac{o\pi}{n}\}$.

This is passed on to the `sg_filterwithbank` command which filters the image I in the frequency domain and stores responses for each scale and orientation in r. Each response matrix is the size of the original image and contains complex values. Their real and imaginary part can be used to determine the magnitude and phase response of the filter. For each orientation, the absolute sum of the complex response values is taken and averaged over the image, resulting in an o-dimensional feature vector.

For the experiments, the filter bank parameters were 5 scales and 6 orientations, starting at frequency $b = 0.35$ cycles/pixel. This value for the maximum frequency was chosen a little lower than suggested in [22], since preliminary experiments done had shown that too high frequencies show very similar responses over all images.

The frequency domain visualization of this bank is shown in figure 5.2 and the summed responses over all orientations are shown at each scale for

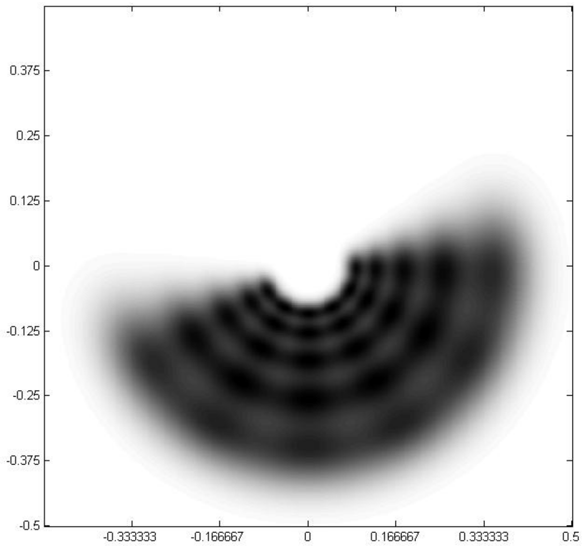

Figure 5.2: Gabor Filter Bank in the frequency domain.

an example photo in figure 5.3.

5.1.2 Results

A sample of photos from the data set ordered for their estimated complexity (mean absolute summed filter response) is shown in figure 5.4 and the histogram of values in figure 5.5.

Looking at figure 5.4, it appears difficult to order the images even manually and some pairs are nearly ambiguous in complexity. Still, one can agree that there is some sort of increasing complexity from top to bottom. It can be observed that large uniform image regions tremendously decrease and small patterns of a certain scale seem to increase the estimated value.

5.2 Jpeg Compressibility

Machado and Petrou [25] presented an estimator of image complexity based on *compressibility*, taking it as a metaphor for human perception, which is

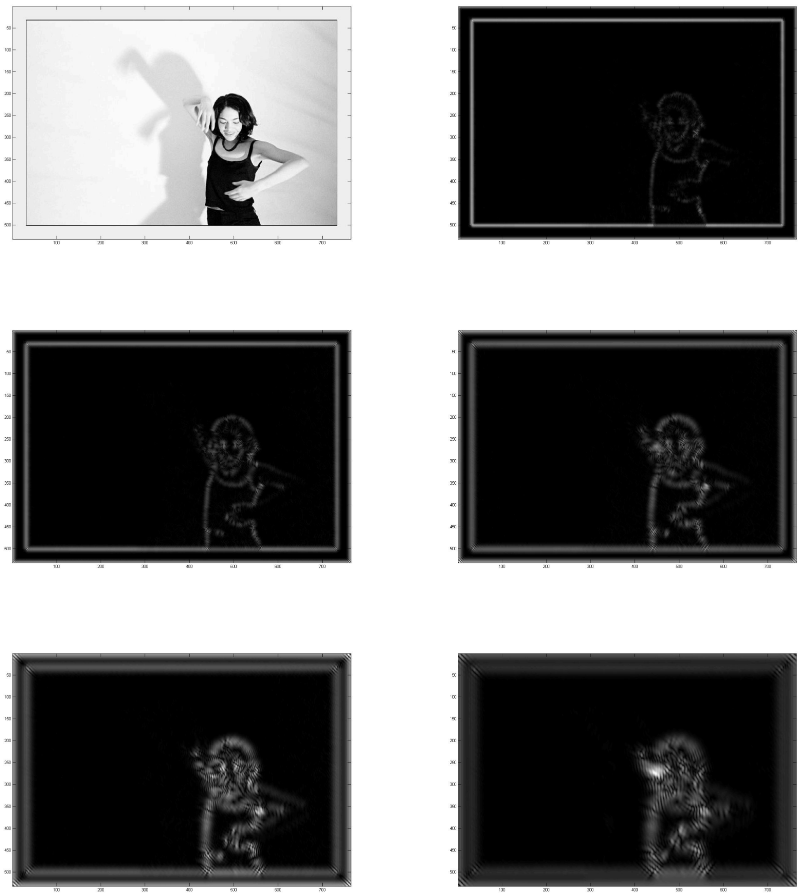

Figure 5.3: Gabor Filter Responses at five scales summed over six orientations.

performing compression by throwing out irrelevant information when building an internal representation of an image. This reduction is emulated by using the *lossy* JPEG compression algorithm [18] which also was designed to discard information irrelevant to human perception while keeping essential visual detail. More technically, the image complexity measure increases the lower the achieved compression ratio and the higher the error in accuracy of the visual representation gets. This measure (I chose to name it ICM here) is expressed by the following term:

$$ICM = \frac{RMSError}{CompressionRatio} \tag{5.5}$$

Figure 5.4: Nine photos ranging from the lowest to the highest 'complexity' values.

$RMSError$ meaning Root Mean Squared Error is used as a measure for the loss in accuracy and $CompressionRatio$ is the ratio between the number of pixels in the original image and the compressed JPEG stream. Expanding the above expression that leads to the following formula:

$$ICM = \frac{\sqrt{E[(I-J)^2]}}{\frac{\|I\|}{\|jpegstream(J)\|}} \tag{5.6}$$

For simplicity only the luminance channel of the photographs was used and color information was discarded, since human visual perception is by far more sensitive to luminance than it is to chrominance.

5.2.1 Implementation

The following `Matlab` code will describe how this measure was calculated from the image set. It was called for each image and a set of JPEG quality parameters: $\{\{5, 10, 20, 30, 40, 50, 60, 70, 80, 90\}$.

Sourecode 5.2.1

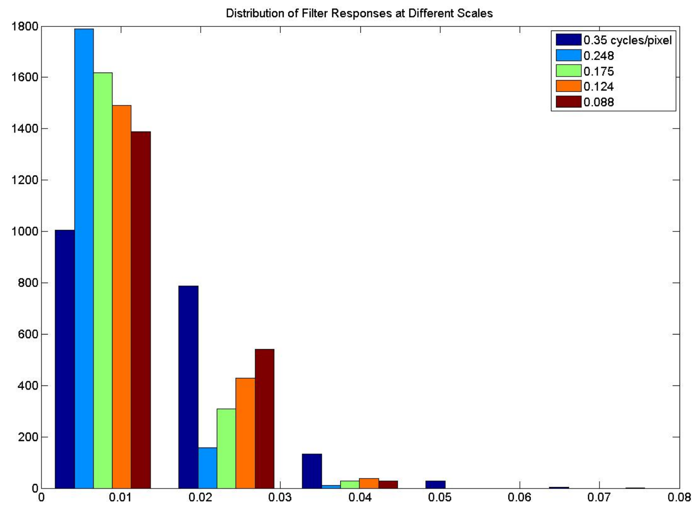

Figure 5.5: Distribution of 'complexity' values.

```
function IC = build_ic(filename, quality)

% read image
 [I, map] = imread(filename);

% if image is 8bit indexed, convert to rgb
 if length(map) > 0
    I = ind2rgb(I,map);
    I = cast(fix(I*255),'uint8');
end

% now convert to grayscale
 if length(I(1,1,:)) > 1
    I = rgb2gray(I);
end

% write to lossy jpeg and read back in
 imwrite(I, 'temp.jpg', 'jpeg', 'mode',
          'lossy', 'quality', quality);
 J = imread('temp.jpg');
```

```
% compute RMS_error
 K = I-J;\\
 L = double(K);\\
 rms = sqrt(sumsqr(L(:))/length(L(:)));

% compute compression ratio
 info = imfinfo('temp.jpg');
 c = info.FileSize;\\
 ratio = length(I(:))/c;

% image complexity
 IC = rms / ratio;
```

Both `imwrite` and `imread` use the `Matlab` internal, standard conform JPEG coding and decoding implementation. The JPEG algorithm transforms the input image into the frequency domain using DCT (discrete cosine transform). Quantization, the step that throws away perceptually irrelevant information, components close to zero in the frequency domain are discarded. These components are calculated for 8×8 or 16×16 pixel regions. The quantified blocks are then compressed using Huffman coding, which removes binary coding redundancy.

The objective was to perform a test following the idea presented in the original paper.

5.2.2 Results

First task was to test the behavior of different JPEG quality parameters in relation to the resulting image complexity values. It turned out that at parameter 50 there was the widest distribution of complexity values over the set of images and at 70, there is a low. This is probably due to the fact that most images in the database where already JPEG compressed before, since they where intended for Web presentation and commonly used compression parameters are in the upper third of the parameter's values. Note that the quality parameter value is purely arbitrary and doesn't reflect any percentage. Also it's scale can vary through different implementations of the algorithm. At the low end (< 50), the algorithm probably hits a performance boundary. This is shown in figure 5.6.

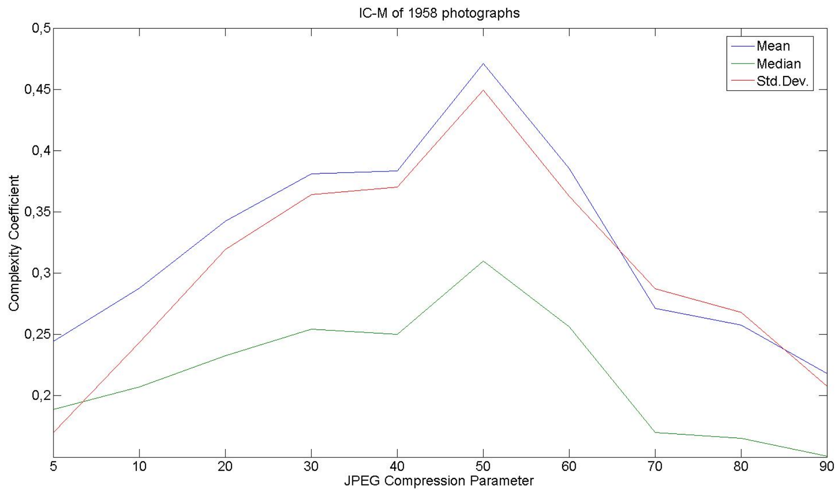

Figure 5.6: The estimator seems to be most sensitive when using JPEG compression quality parameter 50.

The distribution of measures shown in figure 5.7 displays a big number of low valued image complexities and very few high values (JPEG parameter 50). However, the distribution of values does hardly change, when other parameter values are used. Also the resulting ordered sequence of images is very similar throughout different JPEG parameters for the estimator and very similar to the ordering resulting from the Gabor filter based estimator.

The following sections will compare the estimates described above and analyze their relationship to the human image preference data shown in the last chapter.

5.3 Correlations with Image Preference

This section describes the core of the experiments that were done on the image set. Different image properties and features that were calculated are

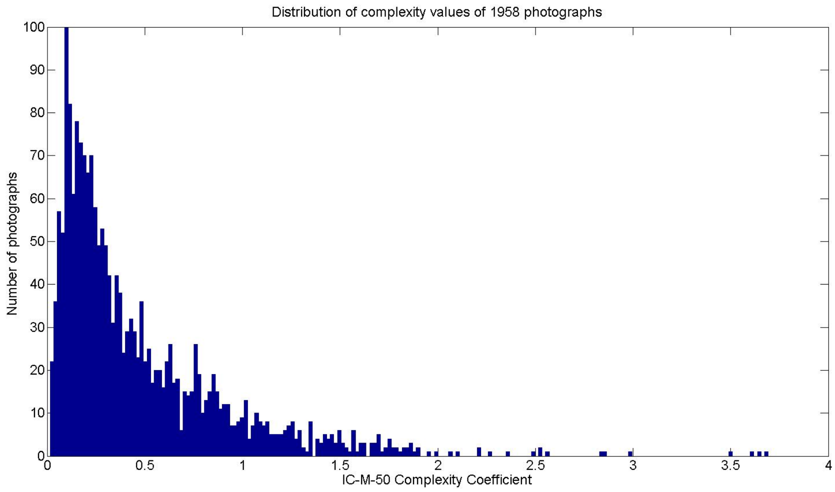

Figure 5.7: Histogram of complexity values.

statistically analyzed and interpreted.

Chapter 4 describes human image preference as three variables in the data-set. The *originality* rating, the *aesthetic* rating and the amount of users rating on one particular image. All of them show significant statistical correlation and cannot be seen as independent ratings. However, a general *image preference* can be observed within them.

First step in analysis is a simple correlation test between each of the variables together with the estimates described above over the whole set of photographs. The correlation matrix, together with the statistical significance is shown in table 5.1.

There is significant (at $\alpha = 0.01$) positive correlation between the preference ratings and the features described in this chapter, even though it is low. It is highest in G3 which is the response at filter scale 0.179 cycles/pixel. Correlation with aesthetics and originality ratings are very similar (due to their close relationship in the data set). However the number of ratings (N) relates more vaguely.

Besides the correlation between the image preference ratings and the estimates (G1-5 and JC), there are some other interesting relationships in the data set. Most unexpected is that the original compression ratio (CR) correlates with the number of ratings that were given for each photograph.

Analytically, this is left unexplained. By knowing the site's customs, it may be guessed that users usually submitting popular photos tend to use higher JPEG quality settings.

There is also a small significant negative correlation between the originality rating (O) and the aspect ratio of the image (AR). Potentially, this is due to 'strange' width to height ratios that pop out and are considered more original by the users. However, this is solely an interpretation. Even more surprising is the relatively high correlation between compression ratio and aspect ratio, since the Jpeg algorithm operates on eight-by-eight image blocks and width to height ratio is discarded.

Moreover, there is a high negative correlation between the original image (Jpeg) compression ratio (CR) and the Gabor filter estimates (G1-5), because the more diverse frequency components in an image, the harder it is for the Jpeg algorithm to compress.

Since Gabor filters extract local responses to frequencies at each image location but estimates described above take the mean of these responses, it discards the spatial information. This results in extracting similar information than Jpeg does with the discrete cosine transform, which does not incorporate spatial information anyway. Hence, the similarity of the estimate to the *jpeg compressibility* described in section 5.2 (JC). Still, the Gabor based estimates show a higher correlation with the preference ratings (A, O and N).

Another important observation is that Jpeg stream size (the Huffman coded quantified image DCT components of each block) and JC still highly correlates with the Gabor estimates. Why the removal of binary coding redundancy does not affect this correlation to a higher extent is left to inspect outside of this thesis' context.

The effect of image aspect ratio (AR) on the Gabor estimates is significant. It's relatively high for filters tuned to high frequencies and gets lower for low frequencies. The explanation for this lies in the implementation of the frequency domain filter (`sg_filterwithbank`), which does a Fourier Transform of the input image resulting in a matrix of components of the same dimensions. For this reason, there are either more components of horizontal and vertical frequencies respectively and the mean response is influenced by this. In other words, two image of different aspect ratios and the same texture have different mean filter responses.

5.4 Machine Learning Experiments

Next step in the experiments was to find out whether using the above variables, more complex models can be built and used to predict the image preference ratings.

The tool used for the experiments (WEKA) is described in [40]. It provides a general framework for various data mining tasks and algorithms and allows for quick start data analysis.

In particular, the algorithm used for analysis is the a model tree learner named MP5, which is provided by the WEKA toolkit. A model tree algorithm initially builds a regular decision tree and the numerical classes are predicted by a linear model at each leaf. For sanity check, other algorithms were applied as well, but they produced results of the same magnitude and proportions and were therefore omitted in this section.

The following sections describe the set of experiments committed and interpret what was learned from different subsets of attributes in the data set. For prediction, all target of the three target classes (A, O and N) were used. For some attribute sets they all produced highly similar errors and were left out of the description. Primary class of interest is the *aesthetics* rating (A), and O and N are left out in this section because O shows equal results to A and predicting N doesn't make much sense. However, the full test runs are shown in appendix A.

5.4.1 Learned Prediction Models

Obviously most interesting in the scope of this work was to check whether it was possible to predict visual preference ratings solely from the features described above. For comparison, to additionally check how simpler image properties are performing, M5 pruned model trees (using smoothed linear models) were learned from the following subsets of attributes:

G1-5 The five Gabor filter response estimates.

AR, CR, RES Aspect ratio, original Jpeg compression ratio and image resolution.

AR, RES Since Jpeg compression ratio relates to psychophysical properties, it provides a rather complex measure and is therefore omitted in this subset. The test should show where the simplest features can already predict the preference classes.

G1-5, AR, CR, RES This attribute set was chosen to check whether the images can be divided into categories to improve prediction with G1-5.

G1-5, AR, RES Same as above, but without compression ratio.

The resulting error metrics resulting from ten-fold cross-evaluation of the models are shown in table 5.2.

Although they cause very little absolute error, they all perform only a few percent better than the trivial predictor (simply taking the mean value of A over all instances). It can be also observed that the correlation between the predicted values and the actual values increases slightly as compared to each variable separately.

5.5 Discussion

Two experiments were done to analyze the data. First, a simple correlation test and second, a machine learning experiment. The results from the learned models in the last section are insufficient for prediction of users' visual preference ratings, but the correlations shown in table 5.1 are significant. Due to that fact, the experiments did provide some interesting, however small insights into structure, quality and relationships in the data set. These outcomes are discussed next.

5.5.1 Originality and Aspect Ratio

From all the correlations between the variables, an unexpected and interesting insight is the significant negative correlation between aspect ratio and originality ratings. Cropping images to different proportions is considered an aesthetic decision and is discussed heavily in the user comments section below the particular photographs on the site. This might potentially be a reason for the correlation.

5.5.2 Aesthetics Hypothesis

As for the hypothesis stated in section 3.3.4, considering the significant correlations between the presented estimates and the ratings, it cannot be rejected taking two assumptions:

1. The mean Gabor filter response presented *does* simulate the average amount of cell activity in V1.

2. The ratings obtained from the photo community site *do* to some extend reflect aesthetic preference other than preference of solely content.

There are two issues which weaken the possibility to take these assumptions: First, even though there is evidence from vision research that cell excitation in the primary visual cortex corresponds to edge local detection in the field of vision (Gabor filters), the model used in the experiments is very simplified and filter bank parameters were chosen based on values used in other applications (e.g. texture description and segmentation). Second, the preference ratings are collected in a totally uncontrolled way and are most likely full of noise. Especially because for photography content (i.e. signs, associations) is very important and cannot be ignored.

So at this point interpreting the correlation as being related to aesthetic perception seems very vague if not wrong. However, looking for support of this hypothesis after finishing the experiments, some related literature was found. Gretchen Schira reported findings of significant correlation between Gabor filter responses and human aesthetic preference of texture images in a more controlled experimental setup [30]. Schira used more sophisticated methods from experimental psychology to obtain the preference ratings. This included tests which removed textured images from the set that the subjects had associations with to validate the correlation. However, correlation was still significant on the set *with* those images, which is a result in favor of assumption (2) above.

5.5.3 Conclusion

On the bottom line, it should be stated that this work is highly preliminary. Sill, the results are motivating to further inspect the relationship between known properties of the human visual system and visual aesthetic preference.

	O	A	N	CR	AR	RES	G1	G2	G3	G4	G5	JC
O	1	0.930	0.631	0.0465	-0.0643	0.0038	**0.136**	**0.155**	**0.165**	**0.145**	**0.122**	0.080
A		1	0.623	0.051	-0.039	-0.021	**0.133**	**0.154**	**0.164**	**0.140**	**0.110**	0.074
N			1	0.126	-0.0192	-0.040	0.0578	0.068	0.065	0.042	0.021	0.035
CR				1	0.162	-0.040	-0.402	-0.426	-0.442	-0.454	-0.442	-0.421
AR					1	-0.721	-0.153	-0.149	-0.088	-0.091	-0.10	0.075
RES						1	0.019	0.025	-0.013	-0.0138	-0.0168	-0.066
G1							1	0.916	0.945	0.928	0.870	0.783
G2								1	0.952	0.875	0.781	0.842
G3									1	0.960	0.877	0.872
G4										1	0.950	0.813
G5											1	0.713
JC												1

	O	A	N	CR	AR	RES	G1	G2	G3	G4	G5	JC
O	-	1	1	0	1	0	1	1	1	1	1	1
A		-	1	0	0	0	1	1	1	1	1	1
N			-	1	0	0	0	1	1	0	0	0
CR				-	1	0	1	1	1	1	1	1
AR					-	1	1	1	1	1	1	1
RES						-	0	0	0	0	0	1
G1							-	1	1	1	1	1
G2								-	1	1	1	1
G3									-	1	1	1
G4										-	1	1
G5											-	1
JC												-

Table 5.1: Correlation coefficient (Pearson) matrix between data set parameters and estimates (top). Significance matrix at $\alpha = 0.01$ (bottom). All 1958 images were used.

Class = A	G1-5	AR,CR,RES	AR,RES	G1-5,AR,CR,RES	G1-5,AR,RES
Correlation coefficient	0.1951	0.3682	0.3192	0.3363	0.2576
Mean absolute error	0.7086	0.6651	0.6828	0.6753	0.7009
Root mean squared error	0.879	0.8334	0.8486	0.8444	0.8706
Relative absolute error	97.7602%	91.7548%	94.2016%	93.1634%	96.6942%
Root relative squared error	98.1074%	93.0156%	94.7083%	94.2487%	97.1661%

Table 5.2: Prediction errors resulting models built using five subsets of the available attributes, predicting the numerical class A. Although they cause very little absolute error, they all perform only a few percent better than the trivial predictor (simply taking the mean value of A over all instances).

Chapter 6

Summary, Conclusions and Outlook

To end this thesis, the following last chapter will summarize and conclude what has been done and point out open problems and future research questions.

6.1 Summary

Initially, the importance of aesthetics for the human environment as stated by philosophers was described and the current problem of aesthetic pollution resulting from computer aided design was pointed out. This led to the motivation that computers should *learn certain aesthetic decisions*, to provide more assistance to humans in design processes.

Chapter 2 gave a historical overview of how scientists tried to approach this problem in the past. Various theories originated around the beginning of the last century and were followed by more modern practical approaches towards the present time. Theories included usage of mathematical and technological findings, like for instance Shannon's information theory, but however have shown no significant results. Only recently, small but more scientific contributions to a computational aesthetic methodology were made. Admittedly, past works have explored a fundamental part of philosophical and psychological problems related to a quantitative approach.

While the historical review was summarizing works which described various kinds of sensory input, the succeeding parts focused primarily on visual aesthetics. Aesthetic measures were identified as being properties of either (1) the object, (2) communication between object and subject or (3) the internal perceptual structures of the subject. Most of all those properties were measures of complexity.

With a preference for (3), current findings of visual perception research and computer vision were described, which model a visual stimulus at various levels of abstractions and representations. Upon these representations, possible estimates for perceptual processing complexity were explored and focus was pointed on the very first stage: early vision. Since there is hardly any subjective influence at this particular stage of perception, the hypothesis derived from this model stated that aesthetics is to some degree dependent on these inter-subjective properties. This hypothesis formed the basis for the successive experiments on the photo database.

This database was built from an online photo community site that offers a feature to its users, enabling them to mutually rate photo entries. The rating categories were *aesthetics* and *originality*, which did however show a very high positive correlation and were therefore interpreted as a *general image preference* in the context of this community's aesthetic common sense.

About two thousand image-rating pairs, accompanied by other basic image properties, such as resolution or aspect ratio were passed through (1) simple correlation significance tests and (2) machine learning experiments.

For testing the hypothesis stated at the end of chapter 3, two estimates for complexity in an early vision methodology were introduced: The *Jpeg compressibility* and *multivariate Gabor filter responses*. Both of these estimates were used in the context of quantifiable aesthetic in other works before and had shown a statistically significant, but low positive correlation with the image preference ratings.

To add another level of verification, a M5P pruned model tree was learned for various subsets of the image and rating attributes, in hope for gaining more insight into the structures and relationships within the database. Results included interesting findings but were only preliminary.

6.2 Contributions

There were two major contributions in this work.

First was a historical overview of works adding to the formation of a paradigm which is frequently called *computational aesthetics* to date. It's theoretical outcomes were interpreted and put in the new light of current research in visual perception and computer vision, which should help identifying solvable problems in the area of aesthetics.

The second contribution was an experiment done on a large database of photographs and which was the first know attempt to analyze human visual preference of such data. Results did include significant positive correlations of the presented estimates with human preference ratings.

6.3 Conclusions and Future Research

Led by the motivation for computational tools which can decide aesthetic problems, aesthetic research has been redefined in a new technological context during the previous century and new theoretical concepts were formed. I have sketched the essential concepts and pointed out their relevance for aesthetic quantification approaches.

The results of the experiments are satisfying. They do not claim to be a description or explanation of the concept of aesthetics, but encourage further experiments. Since they only inspected relationships with early vision properties and very important properties such as *color*, *form* and *composition* were omitted, experiments done using higher levels of visual features are of interest.

On the path towards applications, emphasis should be put on *objects of design* and their difference to *objects of art*, which lack functional requirements. Most significantly research should focus on aesthetics in *form* rather than *content* and find objectivity in psychophysical models of human perception. In contrast, any pure theoretical outcome or reasoning about the values of *Art* is rather pointless, taking into account the philosophical problems one will encounter.

However, this work's results are encouraging a new discipline which seems justified and might catch increasing attention by researchers from now on.

Bibliography

[1] Rudolf Arnheim. *Art And Visual Perception.* University Of California Press, the new version edition, 1974.

[2] Shumeet Baluja, Dean Pomerleau, and Todd Jochem. Towards automated artificial evolution for computer-generated images. *Connection Science*, 6(2 & 3):325–354, 1994.

[3] Max Bense. *Aesthetica I - Metaphysiche Beobachtungen am Schoenen.* Dt. Verlag-Anst., 1954.

[4] Max Bense. *Aesthetica. Einfuehrung in die neue Aesthetik.* Agis-Verlag, Baden-Baden, 1965.

[5] Max Bense. *Einfuehrung in die informationstheoretische Aesthetik.* Rowohlt, Reinbek bei Hamburg, 1969.

[6] Max Bense. kleine abstrakte aesthetik. *Edition rot*, page text 38, January 1969.

[7] D.E. Berlyne. *Aesthetics and Psychobiology.* Appleton-Century-Crofts, New York, 1971.

[8] George D. Birkhoff. *Aesthetic Measure.* Harvard University Press, Cambridge Massachusetts, 1933.

[9] Michael J. Black, David J. Fleet, and Yaser Yacoob. Robustly estimating changes in image appearance. *Comput. Vis. Image Underst.*, 78(1):8–31, 2000.

[10] David Chek, Ling Ngo, and John G. Byrne. Aesthetic measures for screen design. In *OZCHI '98: Proceedings of the Australasian Conference on Computer Human Interaction*, pages 64–, Washington, DC, USA, 1998. IEEE Computer Society.

[11] Gustav Theodor Fechner. *Vorschule der Aesthetik*, volume 1. Druck und Verlag von Breifkopf & Haertl, Leipzig, 1876.

[12] Paul A. Fishwick. Aesthetic computing manifesto. *Leonardo*, 36(4):255–256, 2003.

[13] David A. Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.

[14] Helmer G. Frank and Herbert W. Franke. *Aesthetische Information*. Academia Libroservo, 1997.

[15] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[16] Gary Greenfield. Simulated aesthetics and evolving artworks: A coevolutionary approach. *Leonardo, MIT Press*, 35(3):283–289, June 2002.

[17] Gary Greenfield. On the origins of the term "computational aesthetics". In L. Neumann, M. Sbert, B. Gooch, and W. Purgathofer, editors, *Computational Aesthetics in Graphics, Visualization and Imaging*, number 1, pages 1–4, 2005.

[18] Joint Photographic Experts Group. http://www.jpeg.org/. Official Website.

[19] Hurvich, Leo, Jameson, and Dorthea. An opponent-process theory of color vision. *Psychological Review*, 64:384–390, 1957.

[20] Immanuel Kant. *Kritik der Urteilskraft*. W. Weischedel, 1790.

[21] C. Koch and S. Ullman. Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4:219–227, 1985.

[22] P. Kruizinga and N. Petkov. Non-linear operator for oriented texture. *IEEE Transactions on Image Processing*, 8(10):1395–1407, 1999.

[23] Michael Kubovy. *Encyclopedia of Psychology*, chapter Visual Aesthetics. New York: Oxford Univ. Press, 2000.

[24] Talia Lavie and Noam Tractinsky. Assessing dimensions of perceived visual aesthetics of web sites. *Int. J. Hum.-Comput. Stud.*, 60(3):269–298, 2004.

[25] P. Machado and A. Cardoso. Computing aesthetics. *Lecture Notes in Computer Science*, 1515:219, 1998.

[26] Abraham Moles. *Thorie de l'Information et Perception Esthtique*. Flammarion, Paris, 1958.

[27] Antal Nemcsics. The coloroid color order system. *Color Research and Application*, 5:113–120, 1980.

[28] Joerg Ontrup, Heiko Wersing, and Helge Ritter. A computational feature binding model of human texture perception. *Cognitive Processing*, 5(1):31–44, 2004.

[29] Remko Scha and Rens Bod. Computationele esthetica. *Informatie en Informatiebeleid*, 11(1):54–63, 1993.

[30] Gretchen Schira. Analysis of digital image properties and human preference. In *Proceedings of the ACADIA 2002 - Thresholds Between Physical And Virtual*, Cal Poly, October 2002. Association for Computer-Aided Design in Architecture (ACADIA).

[31] Simeon J. Simoff and Fay Sudweeks. The metaphor of the face as an interface for communicating non-quantitative information. In *AUIC*, pages 95–102, 2000.

[32] Branka Spehar, Colin W. G. Clifford, Ben R. Newell, and Richard P. Taylor. Chaos and graphics - universal aesthetic of fractals. *Computers & Graphics*, 27(5):813–820, 2003.

[33] Tomas Staudek. On birkhoffs aesthetic measure of vases. Technical Report 06, Faculty of Informatics, Masaryk University, 1999.

[34] Tomas Staudek. Computer-aided aesthetic evaluation of visual patterns. In *ISAMA-BRIDGES Conference Proceedings*, pages 143–149, 2003.

[35] Tomas Staudek and Vaclav Linkov. Personality characteristics and aesthetic preference for chaotic curves. *Special Edition of the Journal of Mathematics & Design*, 4(1):297–304, June 2004.

[36] George Stiny and James Gips. *Algorithmic Aesthetics: Computer Models for Criticism and Design in the Arts*. University of California Press, 1978.

[37] Fay Sudweeks and Simeon J. Simoff. Quantifying beauty: An interface for evaluating universal aesthetics. In J. Gammack, editor, *Proceedings of the Western Australian Workshop on Information Systems Research (WAWISR)*, pages 262–267, Perth, 1999. Murdoch University.

[38] R. P. Taylor, B. Spehar, J. A. Wise, C. W. G. Clifford, B. R. Newell, C. M. Hagerhall, T. Purcell, and T. P. Martin. Perceptual and physiological responses to the visual complexity of fractal patterns. *Journal of Nonlinear Dynamics, Psychology, and Life Sciences*, 9(1):89–114, 2005.

[39] M. Wertheimer. Untersuchungen zur lehre von der gestalt, ii. *Psychologische Forschung*, 4:301–350, 1923.

[40] Ian H. Witten and Eibe Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, October 1999.

[41] Steven Yantis, editor. *Key Readings in Visual Perception*. Taylor & Francis Group, 2001.

Appendix

Appendix A

WEKA Results

The following output listings were generated by the M5P algorithm implementation coming with the WEKA framework. It shows the models learned for the data described in chapter 5.

A.1 Gabor Estimates Only

Model A.1.1

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   6
               A
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 G3 <= 0.007 : LM1 (833/119.806%)
 G3 >  0.007 : LM2 (1125/120.168%)

LM num: 1
 A =
    -19.5458 *  G1
    + 74.164 *  G3
    + 95.7658 *  G4
    - 78.3746 *  G5
    + 4.9053

LM num: 2
 A =
    -0.2069 *  G1
    + 52.5417 *  G3
    - 51.5637 *  G4
    + 5.449

Number of Rules : 2

Time taken to build model: 3.84 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.1951
Mean absolute error                      0.7086
Root mean squared error                  0.879
Relative absolute error                 97.7602 %
Root relative squared error             98.1074 %
Total Number of Instances             1958
```

Model A.1.2

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   6
               O
               G1
               G2
               G3
               G4
               G5
```

```
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 G3 <= 0.007 : LM1 (833/117.942%)
 G3 >  0.007 : LM2 (1125/119.891%)

LM num: 1
O =
    -18.2825 *  G1
    + 84.2174 *  G3
    + 57.7877 *  G4
    - 54.3201 *  G5
    + 4.8597

LM num: 2
```

Model A.1.3

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   6
               N
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)
LM1 (1958/155.166%)
```

```
O =
    -0.1861 *  G1
    + 42.1181 *  G3
    - 40.5445 *  G4
    + 5.3712

Number of Rules : 2

Time taken to build model: 3.86 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.185
Mean absolute error                      0.6917
Root mean squared error                  0.8483
Relative absolute error                 98.0477 %
Root relative squared error             98.2747 %
Total Number of Instances             1958
```

```
LM num: 1
 N =
    1712.5082 *  G3
    - 1392.4309 *  G4
    + 22.7114

Number of Rules : 1

Time taken to build model: 2.27 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.0465
Mean absolute error                     23.0785
Root mean squared error                 36.3929
Relative absolute error                 98.9596 %
Root relative squared error            100.0728 %
Total Number of Instances             1958
```

A.2 Image Properties Only

Model A.2.1

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   4
               A
               CR
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 CR <= 8.613 :
|   RES <= 0.232 : LM1 (196/113.016%)
|   RES >  0.232 :
|   |   AR <= 0.671 : LM2 (69/90.741%)
|   |   AR >  0.671 : LM3 (264/100.039%)
 CR >  8.613 :
|   AR <= 1.472 :
|   |   RES <= 0.229 :
|   |   |   RES <= 0.214 :
|   |   |   |   AR <= 1.353 : LM4 (47/104.712%)
|   |   |   |   AR >  1.353 :
|   |   |   |   |   RES <= 0.18 : LM5 (33/28.471%)
|   |   |   |   |   RES >  0.18 : LM6 (26/80.327%)
|   |   |   RES >  0.214 : LM7 (97/112.519%)
|   |   RES >  0.229 :
|   |   |   RES <= 0.263 :
|   |   |   |   RES <= 0.258 : LM8 (129/107.374%)
|   |   |   |   RES >  0.258 : LM9 (104/113.013%)
|   |   |   RES >  0.263 :
|   |   |   |   RES <= 0.567 :
|   |   |   |   |   CR <= 12.834 : LM10 (113/116.797%)
|   |   |   |   |   CR >  12.834 :
|   |   |   |   |   |   AR <= 1.079 : LM11 (164/115.986%)
```

```
|   |   |   |   |   |    AR >  1.079 :
|   |   |   |   |   |   |    AR <= 1.303 : LM12 (107/102.845%)
|   |   |   |   |   |   |    AR >  1.303 : LM13 (29/88.422%)
|   |   |   |    RES >  0.567 : LM14 (162/112.621%)
|    AR >  1.472 :
|   |    RES <= 0.191 :
|   |   |    AR <= 1.554 : LM15 (39/127.014%)
|   |   |    AR >  1.554 : LM16 (128/107.712%)
|   |    RES >  0.191 : LM17 (251/114.048%)

LM num: 1
 A =
    0.0001 *  CR
    + 0.3359 *  AR
    + 3.2218 *  RES
    + 3.7913

LM num: 2
 A =
    0.0001 *  CR
    + 0.0644 *  AR
    - 0.4125 *  RES
    + 5.0259

LM num: 3
 A =
    0.0001 *  CR
    + 0.0136 *  AR
    + 0.0006 *  RES
    + 5.1927

LM num: 4
 A =
    -0.0011 *  CR
    + 1.3076 *  AR
    + 14.3069 *  RES
    + 1.4969

LM num: 5
 A =
    -0.0011 *  CR
    + 0.3184 *  AR
    - 0.9064 *  RES
    + 5.9786

LM num: 6
 A =
    -0.0011 *  CR
    - 2.3348 *  AR
    - 1.0468 *  RES
    + 9.6547

LM num: 7
 A =
    -0.0011 *  CR
    - 1.0624 *  AR
    - 1.1694 *  RES
    + 7.0236

LM num: 8
 A =
    0.0004 *  CR
    - 0.073 *  AR
    + 8.7145 *  RES
    + 3.2027

LM num: 9
 A =
    0.0004 *  CR
    - 0.073 *  AR
    - 1.337 *  RES
    + 5.2855

LM num: 10
 A =
    0.0014 *  CR
    - 0.0566 *  AR
    - 0.1171 *  RES
    + 5.3546

LM num: 11
 A =
    0.0079 *  CR
    - 0.1002 *  AR
    - 1.3882 *  RES
    + 6.2985

LM num: 12
 A =
    -0.0068 *  CR
    + 0.1911 *  AR
    + 0.2688 *  RES
    + 5.2717

LM num: 13
 A =
    0.0008 *  CR
    + 7.0269 *  AR
    + 0.953 *  RES
    - 4.1302

LM num: 14
 A =
    0.0007 *  CR
    - 0.0767 *  AR
    - 1.3799 *  RES
    + 6.1785

LM num: 15
 A =
    -0.0006 *  CR
    - 0.1762 *  AR
    - 0.6681 *  RES
    + 6.1291

LM num: 16
 A =
    -0.0006 *  CR
    - 0.2906 *  AR
    - 0.6681 *  RES
    + 5.9001

LM num: 17
 A =
    -0.0114 *  CR
    + 5.7747 *  AR
    + 31.1127 *  RES
    - 9.9169

Number of Rules : 17

Time taken to build model: 3.83 seconds

=== Cross-validation ===
```

```
=== Summary ===

Correlation coefficient                  0.3682
Mean absolute error                      0.6651
Root mean squared error                  0.8334
Relative absolute error                 91.7548 %
Root relative squared error             93.0156 %
Total Number of Instances             1958
```

Model A.2.2

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   4
              O
               CR
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 CR <= 9.695 :
|   RES <= 0.298 : LM1 (300/110.518%)
|   RES >  0.298 : LM2 (301/93.914%)
 CR >  9.695 :
|   AR <= 1.32 :
|   |   RES <= 0.639 :
|   |   |   CR <= 12.834 : LM3 (105/112.056%)
|   |   |   CR >  12.834 :
|   |   |   |   AR <= 1.084 :
|   |   |   |   |   RES <= 0.424 : LM4 (98/114.006%)
|   |   |   |   |   RES >  0.424 :
|   |   |   |   |   |   AR <= 0.978 : LM5 (76/103.99%)
|   |   |   |   |   |   AR >  0.978 : LM6 (27/100.653%)
|   |   |   |   AR >  1.084 :
|   |   |   |   |   RES <= 0.262 : LM7 (45/99.16%)
|   |   |   |   |   RES >  0.262 : LM8 (120/111.166%)
|   |   RES >  0.639 : LM9 (131/116.089%)
|   AR >  1.32 :
|   |   RES <= 0.203 :
|   |   |   AR <= 1.501 :
|   |   |   |   RES <= 0.146 : LM10 (19/132.527%)
|   |   |   |   RES >  0.146 : LM11 (68/81.179%)
|   |   |   AR >  1.501 :
|   |   |   |   RES <= 0.166 : LM12 (93/121.708%)
|   |   |   |   RES >  0.166 : LM13 (83/100.776%)
|   |   RES >  0.203 :
|   |   |   RES <= 0.208 : LM14 (150/104.133%)
|   |   |   RES >  0.208 : LM15 (342/108.71%)

LM num: 1
O =
    0.0001 *  CR
    - 0.0077 *  AR
    + 1.2712 *  RES
    + 4.6511

LM num: 2
O =
    0.0001 *  CR
    + 1.0447 *  AR
    + 0.0082 *  RES
    + 4.3191

LM num: 3
O =
    0.0008 *  CR
    - 0.0303 *  AR
    - 0.0647 *  RES
    + 5.3222

LM num: 4
O =
    0.0003 *  CR
    + 1.8365 *  AR
    + 3.3777 *  RES
    + 3.0255

LM num: 5
O =
    0.0129 *  CR
    - 0.3211 *  AR
    - 2.4429 *  RES
    + 7.0162

LM num: 6
O =
    0.0119 *  CR
    - 0.6276 *  AR
    - 0.0647 *  RES
    + 5.5844

LM num: 7
O =
    0.0003 *  CR
    - 0.5566 *  AR
    - 1.2363 *  RES
    + 6.8448

LM num: 8
O =
    0.0003 *  CR
    - 0.2829 *  AR
    - 0.5854 *  RES
    + 5.7986

LM num: 9
O =
    0.0001 *  CR
    - 0.06 *  AR
    - 1.3293 *  RES
    + 6.135

LM num: 10
O =
    -0.023 *  CR
    + 1.1637 *  AR
    + 2.5275 *  RES
    + 4.1457

LM num: 11
O =
    -0.0018 *  CR
    + 0.3974 *  AR
    + 0.6485 *  RES
    + 5.4323

LM num: 12
```

```
O =
    -0.0022 *  CR
    - 0.1399 *  AR
    + 2.7436 *  RES
    + 5.2478

LM num: 13
O =
    -0.0023 *  CR
    - 0.1449 *  AR
    - 1.0385 *  RES
    + 5.5139

LM num: 14
O =
    -0.009 *  CR
    - 14.2682 *  AR
    + 0.8653 *  RES
    + 26.2088
```

```
LM num: 15
O =
    -0.0004 *  CR
    + 4.3369 *  AR
    + 9.9925 *  RES
    - 3.449

Number of Rules : 15

Time taken to build model: 3.66 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3684
Mean absolute error                      0.6419
Root mean squared error                  0.8034
Relative absolute error                 90.9777 %
Root relative squared error             93.0766 %
Total Number of Instances             1958
```

Model A.2.3 `=== Run information ===`

```
Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   4
               N
               CR
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 CR <= 13.872 :
|   CR <= 6.714 : LM1 (361/57.193%)
|   CR >  6.714 :
|   |   AR <= 0.774 : LM2 (179/101.032%)
|   |   AR >  0.774 :
|   |   |   RES <= 0.301 : LM3 (328/119.018%)
|   |   |   RES >  0.301 :
|   |   |   |   CR <= 8.461 :
|   |   |   |   |   CR <= 7.247 :
|   |   |   |   |   |   AR <= 0.972 : LM4 (2/21.456%)
|   |   |   |   |   |   AR >  0.972 : LM5 (3/23.33%)
|   |   |   |   |   CR >  7.247 :
|   |   |   |   |   |   RES <= 0.422 :
|   |   |   |   |   |   |   RES <= 0.327 : LM6 (2/2.146%)
|   |   |   |   |   |   |   RES >  0.327 : LM7 (5/57.484%)
|   |   |   |   |   |   RES >  0.422 : LM8 (6/15.53%)
|   |   |   |   CR >  8.461 : LM9 (70/160.452%)
 CR >  13.872 :
|   RES <= 0.203 :
|   |   AR <= 1.503 : LM10 (86/221.866%)
|   |   AR >  1.503 :
|   |   |   RES <= 0.159 :
|   |   |   |   AR <= 1.785 : LM11 (13/212.954%)
|   |   |   |   AR >  1.785 : LM12 (47/81.357%)
|   |   |   RES >  0.159 : LM13 (63/62.051%)
|   RES >  0.203 :
|   |   AR <= 1.286 :
|   |   |   RES <= 0.599 :
|   |   |   |   AR <= 1.075 :
|   |   |   |   |   RES <= 0.413 : LM14 (78/236.514%)
|   |   |   |   |   RES >  0.413 : LM15 (90/177.352%)
|   |   |   |   AR >  1.075 :
|   |   |   |   |   RES <= 0.278 : LM16 (29/224.754%)
|   |   |   |   |   RES >  0.278 :
|   |   |   |   |   |   CR <= 27.947 : LM17 (55/129.658%)
|   |   |   |   |   |   CR >  27.947 : LM18 (27/20.469%)
|   |   |   RES >  0.599 : LM19 (94/81.355%)
|   |   AR >  1.286 :
|   |   |   RES <= 0.21 : LM20 (134/79.435%)
|   |   |   RES >  0.21 :
|   |   |   |   AR <= 1.398 :
|   |   |   |   |   RES <= 0.254 : LM21 (32/145.162%)
|   |   |   |   |   RES >  0.254 :
|   |   |   |   |   |   RES <= 0.267 : LM22 (93/39.32%)
|   |   |   |   |   |   RES >  0.267 : LM23 (30/52.258%)
|   |   |   |   AR >  1.398 : LM24 (131/122.936%)

LM num: 1
 N =
    0.0444 *  CR
    - 0.1976 *  AR
    - 0.3771 *  RES
    + 11.9717

LM num: 2
 N =
    0.0299 *  CR
    - 0.1976 *  AR
    - 0.3771 *  RES
    + 16.9766

LM num: 3
 N =
    0.0299 *  CR
    - 19.1059 *  AR
    - 124.9748 *  RES
    + 71.5524

LM num: 4
 N =
    -20.1322 *  CR
    - 84.1088 *  AR
    - 0.3771 *  RES
    + 289.8639
```

```
LM num: 5
 N =
    -20.1322 *  CR
    - 79.5678 *  AR
    - 0.3771 *  RES
    + 281.4134

LM num: 6
 N =
    -14.3716 *  CR
    + 7.694 *  AR
    + 29.2775 *  RES
    + 135.0447

LM num: 7
 N =
    -14.3716 *  CR
    + 6.1845 *  AR
    + 21.657 *  RES
    + 139.8589

LM num: 8
 N =
    -14.3716 *  CR
    - 2.3697 *  AR
    - 24.3499 *  RES
    + 160.9351

LM num: 9
 N =
    3.0936 *  CR
    - 2.3697 *  AR
    - 0.3771 *  RES
    - 3.8284

LM num: 10
 N =
    0.0064 *  CR
    + 70.3405 *  AR
    + 458.2842 *  RES
    - 110.421

LM num: 11
 N =
    0.0064 *  CR
    - 204.7775 *  AR
    - 690.8801 *  RES
    + 495.732

LM num: 12
 N =
    0.0064 *  CR
    - 15.2615 *  AR
    + 19.2695 *  RES
    + 55.0449

LM num: 13
 N =
    0.0064 *  CR
    - 10.7102 *  AR
    - 72.6475 *  RES
    + 47.6259

LM num: 14
 N =
    0.0064 *  CR
    + 103.7884 *  AR
    + 132.7803 *  RES
    - 76.2693

LM num: 15
 N =
    0.0064 *  CR
    - 66.4992 *  AR
    - 20.2359 *  RES
    + 106.3058

LM num: 16
 N =
    -0.1988 *  CR
    - 237.6376 *  AR
    - 83.2825 *  RES
    + 370.576

LM num: 17
 N =
    -0.1398 *  CR
    - 11.5888 *  AR
    - 19.5517 *  RES
    + 48.3087

LM num: 18
 N =
    -0.1752 *  CR
    - 1.6988 *  AR
    - 2.0227 *  RES
    + 26.4644

LM num: 19
 N =
    0.0064 *  CR
    - 8.6824 *  AR
    - 27.6311 *  RES
    + 44.8584

LM num: 20
 N =
    0.0064 *  CR
    + 7.8749 *  AR
    + 31.0254 *  RES
    - 5.0363

LM num: 21
 N =
    0.0064 *  CR
    - 469.0257 *  AR
    + 41.4595 *  RES
    + 651.063

LM num: 22
 N =
    0.0064 *  CR
    + 31.0551 *  AR
    + 59.272 *  RES
    - 44.5507

LM num: 23
 N =
    0.0064 *  CR
    + 389.6995 *  AR
    + 139.4588 *  RES
    - 530.4973

LM num: 24
```

```
 N =
    0.0064 *  CR
    + 172.1326 *  AR
    + 742.1708 *  RES
    - 393.4539

Number of Rules : 24

Time taken to build model: 2.22 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.4234
Mean absolute error                     19.1404
Root mean squared error                 32.9358
Relative absolute error                 82.0734 %
Root relative squared error             90.5666 %
Total Number of Instances             1958
```

A.3 Aspect Ratio and Resolution

Model A.3.1

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   3
               A
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 AR <= 1.472 :
|    AR <= 0.751 :
|   |    RES <= 0.696 : LM1 (314/110.924%)
|   |    RES >  0.696 : LM2 (99/109.269%)
|    AR >  0.751 :
|   |    RES <= 0.265 :
|   |   |    RES <= 0.214 :
|   |   |   |    AR <= 1.338 : LM3 (73/126.809%)
|   |   |   |    AR >  1.338 : LM4 (81/95.852%)
|   |   |    RES >  0.214 :
|   |   |   |    AR <= 1.333 : LM5 (82/133.302%)
|   |   |   |    AR >  1.333 : LM6 (305/106.238%)
|   |    RES >  0.265 : LM7 (473/115.131%)
 AR >  1.472 :
|    RES <= 0.189 : LM8 (228/124.874%)
|    RES >  0.189 : LM9 (303/114.605%)

LM num: 1
 A =
    0.0596 *  AR
    + 0.6106 *  RES
    + 4.8377

LM num: 2
 A =
    0.1771 *  AR
    - 1.0731 *  RES
    + 5.6873

LM num: 3
 A =
    0.456 *  AR
    + 13.5839 *  RES
    + 2.3468

LM num: 4
 A =
    0.4219 *  AR
    + 0.5817 *  RES
    + 5.2745

LM num: 5
 A =
    0.0133 *  AR
    - 0.9472 *  RES
    + 5.4953

LM num: 6
 A =
    2.3119 *  AR
    - 0.4341 *  RES
    + 1.9081

LM num: 7
 A =
    -0.0125 *  AR
    - 0.0109 *  RES
    + 5.5194

LM num: 8
 A =
    -0.007 *  AR
    + 3.4995 *  RES
    + 4.731

LM num: 9
 A =
    5.4868 *  AR
    + 18.9257 *  RES
    - 7.2349

Number of Rules : 9

Time taken to build model: 3.38 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3192
Mean absolute error                      0.6828
Root mean squared error                  0.8486
Relative absolute error                 94.2016 %
Root relative squared error             94.7083 %
Total Number of Instances             1958
```

Model A.3.2

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   3
              O
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 AR <= 1.472 :
|   AR <= 0.751 :
|   |   RES <= 0.696 : LM1 (314/107.395%)
|   |   RES >  0.696 : LM2 (99/102.781%)
|   AR >  0.751 :
|   |   RES <= 0.306 :
|   |   |   RES <= 0.213 :
|   |   |   |   AR <= 1.353 : LM3 (76/121.301%)
|   |   |   |   AR >  1.353 : LM4 (71/105.312%)
|   |   |   RES >  0.213 :
|   |   |   |   AR <= 1.298 : LM5 (168/121.311%)
|   |   |   |   AR >  1.298 :
|   |   |   |   |   AR <= 1.339 : LM6 (165/109.935%)
|   |   |   |   |   AR >  1.339 : LM7 (202/96.042%)
|   |   RES >  0.306 : LM8 (332/119.942%)
 AR >  1.472 :
|   RES <= 0.189 : LM9 (228/122.752%)
|   RES >  0.189 : LM10 (303/110.048%)

LM num: 1
O =
    -0.0029 * AR
    + 0.8143 * RES
    + 4.7753

LM num: 2
O =
    -0.0029 * AR
    - 0.7623 * RES
    + 5.5308

LM num: 3
O =
    0.3663 * AR
    + 12.7201 * RES
    + 2.6101

LM num: 4
O =
    0.3884 * AR
    + 1.035 * RES
    + 5.1903

LM num: 5
O =
    -0.0794 * AR
    - 4.5281 * RES
    + 6.6802

LM num: 6
O =
    -10.0277 * AR
    + 0.3978 * RES
    + 18.0864

LM num: 7
O =
    6.7125 * AR
    + 22.2926 * RES
    - 9.6042

LM num: 8
O =
    -0.0216 * AR
    - 0.0081 * RES
    + 5.5268

LM num: 9
O =
    -0.0076 * AR
    + 3.1785 * RES
    + 4.7248

LM num: 10
O =
    5.4983 * AR
    + 15.9301 * RES
    - 6.6867

Number of Rules : 10

Time taken to build model: 3.36 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3179
Mean absolute error                      0.6624
Root mean squared error                  0.8186
Relative absolute error                 93.8899 %
Root relative squared error             94.8315 %
Total Number of Instances             1958
```

Model A.3.3

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   3
               N
               AR
               RES
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 AR <= 0.774 : LM1 (454/110.359%)
 AR >  0.774 :
|   RES <= 0.302 :
|   |   RES <= 0.203 :
|   |   |   AR <= 1.505 : LM2 (188/206.628%)
|   |   |   AR >  1.505 : LM3 (232/141.255%)
|   |   RES >  0.203 :
|   |   |   AR <= 1.333 :
```

```
|   |   |   |   |   AR <= 1.238 :
|   |   |   |   |   |   AR <= 0.992 : LM4 (19/32.778%)
|   |   |   |   |   |   AR >  0.992 : LM5 (51/233.923%)
|   |   |   |   |   AR >  1.238 : LM6 (153/118.922%)
|   |   |   |   AR >  1.333 :
|   |   |   |   |   RES <= 0.24 :
|   |   |   |   |   |   RES <= 0.216 : LM7 (261/87.837%)
|   |   |   |   |   |   RES >  0.216 :
|   |   |   |   |   |   |   AR <= 1.446 :
|   |   |   |   |   |   |   |   RES <= 0.233 : LM8 (88/53.954%)
|   |   |   |   |   |   |   |   RES >  0.233 : LM9 (32/124.94%)
|   |   |   |   |   |   |   AR >  1.446 : LM10 (33/170.19%)
|   |   |   |   |   RES >  0.24 : LM11 (144/40.993%)
|   RES >  0.302 :
|   |   RES <= 0.378 :
|   |   |   AR <= 1.104 : LM12 (71/257.071%)
|   |   |   AR >  1.104 :
|   |   |   |   AR <= 1.234 : LM13 (40/75.738%)
|   |   |   |   AR >  1.234 : LM14 (11/199.108%)
|   |   RES >  0.378 : LM15 (181/181.632%)

LM num: 1
 N =
    -0.2938 *  AR
    - 10.7995 *  RES
    + 24.0231

LM num: 2
 N =
    47.4721 *  AR
    + 304.3807 *  RES
    - 74.1286

LM num: 3
 N =
    -23.5484 *  AR
    - 247.6171 *  RES
    + 98.1085

LM num: 4
 N =
    26.6807 *  AR
    - 66.3294 *  RES
    + 15.6531

LM num: 5
 N =
    11.7435 *  AR
    - 66.3294 *  RES
    + 51.4181

LM num: 6
 N =
    -137.0262 *  AR
    - 463.5345 *  RES
    + 321.3195

LM num: 7
 N =
    2.6144 *  AR
    + 7.6041 *  RES
    + 8.5891

LM num: 8
 N =
    -12.8441 *  AR
    + 7.6041 *  RES
    + 29.9874

LM num: 9
 N =
    -608.9958 *  AR
    + 7.6041 *  RES
    + 865.6222

LM num: 10
 N =
    2.6144 *  AR
    + 7.6041 *  RES
    + 15.3378

LM num: 11
 N =
    434.6337 *  AR
    + 993.4485 *  RES
    - 827.4862

LM num: 12
 N =
    110.0404 *  AR
    - 0.2006 *  RES
    - 50.9857

LM num: 13
 N =
    70.206 *  AR
    + 97.682 *  RES
    - 90.2892

LM num: 14
 N =
    146.2832 *  AR
    + 206.8588 *  RES
    - 192.0886

LM num: 15
 N =
    1.1619 *  AR
    - 0.2006 *  RES
    + 29.5203

Number of Rules : 15

Time taken to build model: 1.8 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.2883
Mean absolute error                     20.6787
Root mean squared error                 34.8553
Relative absolute error                 88.6694 %
Root relative squared error             95.8449 %
Total Number of Instances             1958
```

A.4 Gabor Estimates with Image Properties

Model A.4.1

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   9
               A
               CR
               AR
               RES
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 G3 <= 0.007 : LM1 (833/117.814%)
 G3 >  0.007 :
|   CR <= 13.023 : LM2 (624/110.248%)
|   CR >  13.023 : LM3 (501/114.147%)

LM num: 1
 A =
    0.0101 * CR
    - 0.2262 * AR
    - 0.483 * RES
    - 29.6078 * G1
    - 0.3383 * G2
    + 91.044 * G3
    + 111.6876 * G4
    - 80.9177 * G5
    + 5.079

LM num: 2
 A =
    0.0162 * CR
    - 0.012 * AR
    + 0.5176 * RES
    - 0.6846 * G1
    - 18.8175 * G2
    + 63.2379 * G3
    - 18.7667 * G4
    - 20.3143 * G5
    + 4.851

LM num: 3
 A =
    0.0213 * CR
    - 1.5839 * AR
    - 2.1723 * RES
    - 43.2484 * G1
    - 0.2516 * G2
    + 159.5122 * G3
    - 50.1608 * G4
    + 7.5234

Number of Rules : 3

Time taken to build model: 3.84 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3363
Mean absolute error                      0.6753
Root mean squared error                  0.8444
Relative absolute error                 93.1634 %
Root relative squared error             94.2487 %
Total Number of Instances             1958
```

Model A.4.2

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   9
              O
               CR
               AR
               RES
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 G3 <= 0.007 :
|   CR <= 16.162 :
|   |   G3 <= 0.003 :
|   |   |   G2 <= 0.002 : LM1 (39/77.504%)
|   |   |   G2 >  0.002 : LM2 (27/74.526%)
|   |   G3 >  0.003 :
|   |   |   G5 <= 0.007 :
|   |   |   |   CR <= 8.672 : LM3 (57/87.638%)
|   |   |   |   CR >  8.672 : LM4 (37/99.252%)
|   |   |   G5 >  0.007 : LM5 (155/88.653%)
|   CR >  16.162 :
|   |   RES <= 0.261 :
|   |   |   RES <= 0.187 : LM6 (81/131.947%)
|   |   |   RES >  0.187 : LM7 (240/113.427%)
|   |   RES >  0.261 : LM8 (197/116.734%)
 G3 >  0.007 :
|   CR <= 13.015 :
|   |   RES <= 0.298 : LM9 (337/111.829%)
|   |   RES >  0.298 : LM10 (284/100.256%)
|   CR >  13.015 :
|   |   AR <= 1.313 : LM11 (212/111.68%)
|   |   AR >  1.313 :
|   |   |   RES <= 0.203 :
|   |   |   |   AR <= 1.502 : LM12 (48/70.62%)
|   |   |   |   AR >  1.502 : LM13 (43/90.762%)
|   |   |   RES >  0.203 : LM14 (201/101.781%)

LM num: 1
O =
    0.0039 * CR
    - 0.0139 * AR
    - 0.0186 * RES
    - 1.6574 * G1
    - 0.2671 * G2
```

```
    - 16.7517 *  G3
    - 21.7614 *  G4
    + 11.475 *  G5
    + 4.8343

LM num: 2
O =
    0.0346 *  CR
    - 0.0139 *  AR
    - 0.0186 *  RES
    - 1.6574 *  G1
    + 398.6072 *  G2
    - 26.2569 *  G3
    - 28.8075 *  G4
    + 16.7768 *  G5
    + 3.3906

LM num: 3
O =
    0.0148 *  CR
    - 0.0139 *  AR
    - 0.0186 *  RES
    - 1.6574 *  G1
    - 0.2671 *  G2
    + 9.4291 *  G3
    + 2.8998 *  G4
    - 8.4236 *  G5
    + 4.8033

LM num: 4
O =
    0.0186 *  CR
    - 0.0139 *  AR
    - 0.0186 *  RES
    - 1.6574 *  G1
    - 0.2671 *  G2
    + 9.4291 *  G3
    + 2.8998 *  G4
    - 8.4236 *  G5
    + 5.0903

LM num: 5
O =
    0.0038 *  CR
    - 0.0139 *  AR
    - 0.0186 *  RES
    - 1.6574 *  G1
    - 0.2671 *  G2
    + 9.4291 *  G3
    + 2.8998 *  G4
    - 6.7959 *  G5
    + 4.7927

LM num: 6
O =
    0.0008 *  CR
    - 0.1032 *  AR
    - 0.809 *  RES
    - 9.8119 *  G1
    - 4.8005 *  G2
    + 149.5487 *  G3
    + 4.0517 *  G4
    - 3.7904 *  G5
    + 5.123

LM num: 7
O =
    0.0008 *  CR
    - 0.0601 *  AR
    - 0.3306 *  RES
    - 56.6464 *  G1
    - 4.8005 *  G2
    + 192.9436 *  G3
    + 4.0517 *  G4
    - 3.7904 *  G5
    + 4.7611

LM num: 8
O =
    0.0122 *  CR
    - 0.6017 *  AR
    - 1.2748 *  RES
    - 40.7026 *  G1
    - 228.3318 *  G2
    + 232.3482 *  G3
    + 96.6076 *  G4
    - 52.9204 *  G5
    + 6.0898

LM num: 9
O =
    0.0016 *  CR
    - 0.0303 *  AR
    - 0.0145 *  RES
    - 17.8267 *  G1
    - 1.3188 *  G2
    + 53.367 *  G3
    - 1.6467 *  G4
    - 16.6323 *  G5
    + 5.0118

LM num: 10
O =
    0.0017 *  CR
    + 0.5398 *  AR
    - 0.0145 *  RES
    - 0.7086 *  G1
    - 36.8837 *  G2
    + 92.5372 *  G3
    - 49.8528 *  G4
    - 0.8704 *  G5
    + 4.7963

LM num: 11
O =
    0.002 *  CR
    - 1.3225 *  AR
    - 2.0872 *  RES
    - 23.0826 *  G1
    + 60.9601 *  G2
    + 10.9675 *  G3
    - 2.8853 *  G4
    + 7.7439

LM num: 12
O =
    0.0046 *  CR
    + 1.0516 *  AR
    + 2.39 *  RES
    - 25.5141 *  G1
    - 0.1987 *  G2
    + 72.112 *  G3
    - 2.383 *  G4
    + 3.8466
```

```
LM num: 13
O =
    0.0284 *  CR
    - 3.1527 *  AR
    - 15.4078 *  RES
    - 61.6952 *  G1
    + 81.7599 *  G2
    + 50.4442 *  G3
    - 2.383 *  G4
    + 12.6564

LM num: 14
O =
    0.0195 *  CR
    + 2.2048 *  AR
    + 7.1228 *  RES
    - 54.7315 *  G1
    - 0.1987 *  G2
    + 111.2034 *  G3
    - 2.383 *  G4
    - 0.2317

Number of Rules : 14

Time taken to build model: 3.99 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3713
Mean absolute error                      0.6429
Root mean squared error                  0.8026
Relative absolute error                 91.1313 %
Root relative squared error             92.9827 %
Total Number of Instances             1958
```

Model A.4.3

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   9
              N
              CR
              AR
              RES
              G1
              G2
              G3
              G4
              G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 CR <= 13.872 : LM1 (956/112.511%)
 CR >  13.872 :
|   G3 <= 0.004 : LM2 (241/120.061%)
|   G3 >  0.004 :
|   |   RES <= 0.203 :
|   |   |   AR <= 1.503 : LM3 (78/217.833%)
|   |   |   AR >  1.503 :
|   |   |   |   RES <= 0.159 :
|   |   |   |   |   AR <= 1.756 :
|   |   |   |   |   |   RES <= 0.15 : LM4 (6/190.483%)
|   |   |   |   |   |   RES >  0.15 : LM5 (3/159.85%)
|   |   |   |   |   AR >  1.756 : LM6 (20/92.501%)
|   |   |   |   RES >  0.159 :
|   |   |   |   |   G3 <= 0.007 :
|   |   |   |   |   |   G4 <= 0.008 : LM7 (15/10.612%)
|   |   |   |   |   |   G4 >  0.008 : LM8 (8/2.816%)
|   |   |   |   |   G3 >  0.007 :
|   |   |   |   |   |   G4 <= 0.011 :
|   |   |   |   |   |   |   G3 <= 0.008 : LM9 (2/12.874%)
|   |   |   |   |   |   |   G3 >  0.008 : LM10 (4/29.49%)
|   |   |   |   |   |   G4 >  0.011 : LM11 (19/14.44%)
|   |   RES >  0.203 : LM12 (606/168.661%)

LM num: 1
 N =
    0.9416 *  CR
    - 0.1838 *  AR
    - 0.3469 *  RES
    + 1383.829 *  G3
    - 16.5627 *  G4
    - 916.7874 *  G5
    + 6.5013

LM num: 2
 N =
    0.3324 *  CR
    - 8.911 *  AR
    - 15.6885 *  RES
    - 57.232 *  G1
    + 90.7634 *  G2
    + 187.5226 *  G3
    - 58.2722 *  G4
    + 24.775

LM num: 3
 N =
    0.1783 *  CR
    + 63.3412 *  AR
    + 355.4139 *  RES
    - 103.3643 *  G1
    + 29.9426 *  G2
    + 1201.6177 *  G3
    - 583.7115 *  G4
    - 88.5961

LM num: 4
 N =
    1.5342 *  CR
    - 180.6128 *  AR
    - 832.4562 *  RES
    - 787.4585 *  G1
    + 2509.7528 *  G2
    + 3281.1118 *  G3
    - 588.3181 *  G4
    + 438.9559

LM num: 5
 N =
    1.5342 *  CR
    - 190.2723 *  AR
    - 832.4562 *  RES
    - 787.4585 *  G1
    + 2509.7528 *  G2
    + 3281.1118 *  G3
```

```
    - 588.3181 *  G4
    + 449.6841

LM num: 6
 N =
    1.2486 *  CR
    - 89.2656 *  AR
    - 655.8923 *  RES
    - 434.673 *  G1
    + 2509.7528 *  G2
    + 2263.2524 *  G3
    - 588.3181 *  G4
    + 244.2721

LM num: 7
 N =
    0.491 *  CR
    - 16.4145 *  AR
    - 209.2682 *  RES
    + 202.8236 *  G1
    + 3171.4506 *  G2
    - 472.042 *  G3
    - 866.3516 *  G4
    + 119.4071 *  G5
    + 62.849

LM num: 8
 N =
    0.5693 *  CR
    - 16.4145 *  AR
    - 209.2682 *  RES
    + 202.8236 *  G1
    + 3171.4506 *  G2
    - 472.042 *  G3
    - 950.9705 *  G4
    + 155.7484 *  G5
    + 60.8532

LM num: 9
 N =
    1.1513 *  CR
    - 16.4145 *  AR
    - 209.2682 *  RES
    + 202.8236 *  G1
    + 3100.9717 *  G2
    - 428.713 *  G3
    - 588.3181 *  G4
    + 55.903

LM num: 10
 N =
    1.609 *  CR
    - 16.4145 *  AR
    - 209.2682 *  RES
    + 202.8236 *  G1
    + 3100.9717 *  G2
    - 428.713 *  G3
    - 588.3181 *  G4
    + 45.626

LM num: 11
 N =
    0.8988 *  CR
    - 16.4145 *  AR
    - 209.2682 *  RES
    + 202.8236 *  G1
    + 3100.9717 *  G2
    - 428.713 *  G3
    - 435.9192 *  G4
    + 54.023

LM num: 12
 N =
    0.0221 *  CR
    - 71.9616 *  AR
    - 80.0116 *  RES
    - 42.0082 *  G1
    + 29.9426 *  G2
    + 1230.2123 *  G3
    - 65.4314 *  G4
    + 133.1807

Number of Rules : 12

Time taken to build model: 2.66 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.3177
Mean absolute error                     19.8784
Root mean squared error                 34.8033
Relative absolute error                 85.2378 %
Root relative squared error             95.7019 %
Total Number of Instances             1958
```

A.5 Gabor Estimates with Aspect Ratio and Resolution

Model A.5.1

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   8
               A
               AR
               RES
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)
```

```
 G3 <= 0.007 :
|   G3 <= 0.003 :
|   |   RES <= 0.269 : LM1 (139/126.94%)
|   |   RES >  0.269 : LM2 (93/109.028%)
|   G3 >  0.003 : LM3 (601/115.143%)
 G3 >  0.007 :
|   RES <= 0.27 :
|   |   RES <= 0.203 :
|   |   |   AR <= 1.505 : LM4 (118/120.757%)
|   |   |   AR >  1.505 :
|   |   |   |   RES <= 0.191 : LM5 (79/102.385%)
|   |   |   |   RES >  0.191 : LM6 (33/106.456%)
|   |   RES >  0.203 :
|   |   |   AR <= 1.333 :
|   |   |   |   AR <= 1.306 :
|   |   |   |   |   AR <= 0.938 : LM7 (15/98.211%)
|   |   |   |   |   AR >  0.938 : LM8 (38/87.549%)
|   |   |   |   AR >  1.306 : LM9 (29/114.284%)
|   |   |   AR >  1.333 : LM10 (321/105.65%)
|   RES >  0.27 : LM11 (492/111.094%)

LM num: 1
 A =
    -0.0366 *  AR
    - 0.082 *  RES
    - 41.1738 *  G1
    - 0.4761 *  G2
    + 19.4303 *  G3
    + 6.5118 *  G4
    - 5.4062 *  G5
    + 5.0168

LM num: 2
 A =
    -0.0451 *  AR
    - 1.1261 *  RES
    - 8.3961 *  G1
    - 513.5986 *  G2
    + 836.9331 *  G3
    + 6.5118 *  G4
    - 205.0447 *  G5
    + 5.6541

LM num: 3
 A =
    -0.0095 *  AR
    - 0.0189 *  RES
    - 0.8344 *  G1
    - 0.4761 *  G2
    + 3.0285 *  G3
    + 86.6845 *  G4
    - 86.4154 *  G5
    + 5.2251

LM num: 4
 A =
    1.6007 *  AR
    + 11.8681 *  RES
    - 1.2406 *  G1
    - 87.9909 *  G2
    + 87.7722 *  G3
    - 9.24 *  G4
    - 1.915 *  G5
    + 1.4094

LM num: 5
 A =
    -0.2121 *  AR
    - 0.3975 *  RES
    - 6.705 *  G1
    - 1.6111 *  G2
    + 12.9493 *  G3
    + 86.5777 *  G4
    - 104.5776 *  G5
    + 6.0075

LM num: 6
 A =
    -0.3801 *  AR
    - 1.0449 *  RES
    - 11.9418 *  G1
    + 3.0246 *  G2
    + 7.6416 *  G3
    + 31.9033 *  G4
    - 31.8848 *  G5
    + 5.9094

LM num: 7
 A =
    0.6771 *  AR
    + 7.7309 *  RES
    - 5.7646 *  G1
    - 4.4641 *  G2
    + 19.2413 *  G3
    + 39.1505 *  G4
    - 41.3375 *  G5
    + 2.7039

LM num: 8
 A =
    0.391 *  AR
    - 2.5485 *  RES
    - 5.7646 *  G1
    - 4.4641 *  G2
    + 19.2413 *  G3
    + 21.3509 *  G4
    - 25.6851 *  G5
    + 5.9644

LM num: 9
 A =
    0.0841 *  AR
    - 3.8884 *  RES
    - 71.5614 *  G1
    - 4.4641 *  G2
    + 100.646 *  G3
    - 1.8661 *  G4
    - 5.2688 *  G5
    + 6.5132

LM num: 10
 A =
    -0.0477 *  AR
    - 0.0921 *  RES
    - 51.9346 *  G1
    - 2.6147 *  G2
    + 101.532 *  G3
    - 1.8661 *  G4
    - 25.9955 *  G5
    + 5.4313

LM num: 11
 A =
```

```
    0.5295 *  AR
    - 0.0055 *  RES
    - 0.1978 *  G1
    - 1.1692 *  G2
    + 58.2467 *  G3
    - 58.0008 *  G4
    - 0.6234 *  G5
    + 5.1472

Number of Rules : 11
```

Model A.5.2

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   8
              O
               AR
               RES
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 G3 <= 0.007 :
|   AR <= 0.751 :
|   |   RES <= 0.69 : LM1 (128/95.903%)
|   |   RES >  0.69 :
|   |   |   AR <= 0.666 : LM2 (19/98.968%)
|   |   |   AR >  0.666 : LM3 (43/63.574%)
|   AR >  0.751 :
|   |   RES <= 0.261 : LM4 (459/118.465%)
|   |   RES >  0.261 : LM5 (184/119.087%)
 G3 >  0.007 :
|   AR <= 1.313 : LM6 (566/114.832%)
|   AR >  1.313 :
|   |   RES <= 0.203 :
|   |   |   AR <= 1.499 : LM7 (74/110.019%)
|   |   |   AR >  1.499 :
|   |   |   |   RES <= 0.168 : LM8 (54/123.81%)
|   |   |   |   RES >  0.168 : LM9 (76/94.551%)
|   |   RES >  0.203 : LM10 (355/106.649%)

LM num: 1
O =
    1.4908 *  AR
    + 0.5328 *  RES
    - 4.5184 *  G1
    + 6.9534 *  G2
    + 0.2958 *  G3
    + 16.8174 *  G4
    - 42.0802 *  G5
    + 3.9119

LM num: 2
O =
    -1.1455 *  AR
    - 0.0818 *  RES
    - 6.7763 *  G1
    + 135.2442 *  G2
    - 169.3316 *  G3
    + 104.7745 *  G4
    - 18.4461 *  G5
    + 5.1789

LM num: 3
O =
    0.2193 *  AR
    + 1.2845 *  RES
    - 6.7763 *  G1
    + 324.2723 *  G2
    - 244.866 *  G3
    + 72.6884 *  G4
    - 18.4461 *  G5
    + 3.3593

LM num: 4
O =
    -0.0081 *  AR
    + 0.0218 *  RES
    - 33.2348 *  G1
    - 0.3838 *  G2
    + 132.0524 *  G3
    + 1.1278 *  G4
    - 2.2771 *  G5
    + 4.6987

LM num: 5
O =
    -0.0081 *  AR
    + 0.0678 *  RES
    - 3.0244 *  G1
    - 81.8654 *  G2
    + 13.7957 *  G3
    + 1.1278 *  G4
    - 3.5696 *  G5
    + 5.5322

LM num: 6
O =
    0.5157 *  AR
    - 0.0043 *  RES
    - 14.158 *  G1
    - 0.9039 *  G2
    + 73.1204 *  G3
    - 50.939 *  G4
    + 5.1084

LM num: 7
O =
    1.6752 *  AR
    + 3.7457 *  RES
    - 10.3846 *  G1
```

```
Time taken to build model: 3.91 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.2576
Mean absolute error                      0.7009
Root mean squared error                  0.8706
Relative absolute error                 96.6942 %
Root relative squared error             97.1661 %
Total Number of Instances             1958
```

```
    - 65.4491 *  G2
    + 36.1206 *  G3
    + 29.7989 *  G4
    + 2.8204

LM num: 8
O =
    -0.4388 *  AR
    + 4.6396 *  RES
    - 7.5029 *  G1
    - 6.5962 *  G2
    + 26.7359 *  G3
    - 8.4358 *  G4
    - 6.5075 *  G5
    + 5.5053

LM num: 9
O =
    -0.4388 *  AR
    - 0.771 *  RES
    - 7.5029 *  G1
    - 7.6696 *  G2
    + 26.7359 *  G3
    - 8.4358 *  G4
    - 4.9343 *  G5
    + 5.983

LM num: 10
O =
    2.0708 *  AR
    + 7.4206 *  RES
    - 57.0249 *  G1
    - 2.9621 *  G2
    + 83.7966 *  G3
    - 3.2374 *  G4
    + 0.5798

Number of Rules : 10

Time taken to build model: 3.92 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.2739
Mean absolute error                      0.6753
Root mean squared error                  0.8327
Relative absolute error                 95.7171 %
Root relative squared error             96.4697 %
Total Number of Instances             1958
```

Model A.5.3

```
=== Run information ===

Scheme:       weka.classifiers.trees.M5P -M 4.0

Instances:    1958
Attributes:   8
               N
               AR
               RES
               G1
               G2
               G3
               G4
               G5
Test mode:    10-fold cross-validation

=== Classifier model (full training set) ===

M5 pruned model tree:
(using smoothed linear models)

 AR <= 0.774 : LM1 (454/108.858%)
 AR >  0.774 :
|    RES <= 0.302 :
|   |    RES <= 0.203 :
|   |   |    AR <= 1.505 : LM2 (188/201.565%)
|   |   |    AR >  1.505 : LM3 (232/139.586%)
|   |    RES >  0.203 :
|   |   |    AR <= 1.333 :
|   |   |   |    AR <= 1.238 :
|   |   |   |   |    AR <= 0.992 : LM4 (19/30.855%)
|   |   |   |   |    AR >  0.992 : LM5 (51/233.923%)
|   |   |   |    AR >  1.238 : LM6 (153/116.141%)
|   |   |    AR >  1.333 : LM7 (558/90.978%)
|    RES >  0.302 : LM8 (303/206.757%)

LM num: 1
 N =
    -0.285 *  AR
    - 10.2337 *  RES
    - 750.3682 *  G1
    - 786.4892 *  G2
    + 2753.8453 *  G3
    - 20.3092 *  G4
    - 917.6692 *  G5
    + 27.201

LM num: 2
 N =
    -2.6361 *  AR
    + 300.2804 *  RES
    - 6.0556 *  G1
    - 8612.2528 *  G2
    + 10330.0307 *  G3
    - 3422.0245 *  G4
    - 5.7332 *  G5
    - 2.3234

LM num: 3
 N =
    -24.5507 *  AR
    - 232.5259 *  RES
    - 6.0556 *  G1
    - 151.9638 *  G2
    + 3012.0372 *  G3
    - 3000.4704 *  G4
    - 5.7332 *  G5
    + 102.9735

LM num: 4
 N =
    14.855 *  AR
    - 59.3827 *  RES
    - 2940.5249 *  G1
    - 1112.8887 *  G2
    + 4540.6842 *  G3
    + 2762.0139 *  G4
    - 1633.2883 *  G5
    + 34.8058

LM num: 5
 N =
```

```
    3.1154 *  AR
    - 59.3827 *  RES
    - 1683.5463 *  G1
    - 582.8758 *  G2
    + 2719.0795 *  G3
    + 1464.8068 *  G4
    - 852.5929 *  G5
    + 62.5453

LM num: 6
 N =
    -139.1694 *  AR
    - 417.9206 *  RES
    - 192.7147 *  G1
    + 1384.2314 *  G2
    - 417.101 *  G3
    + 1087.1662 *  G4
    - 918.7866 *  G5
    + 308.5812

LM num: 7
 N =
    125.0287 *  AR
    + 368.8584 *  RES
    - 17.53 *  G1
    - 1055.8894 *  G2
    + 1636.0457 *  G3
    - 57.977 *  G4
    - 679.0221 *  G5
    - 247.1502

LM num: 8
 N =
    34.7764 *  AR
    - 0.1967 *  RES
    - 1576.8897 *  G1
    + 3499.6302 *  G2
    + 130.1668 *  G3
    - 80.9779 *  G4
    - 5.7332 *  G5
    + 5.8476

Number of Rules : 8

Time taken to build model: 2.39 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                  0.2371
Mean absolute error                     21.4863
Root mean squared error                 35.4074
Relative absolute error                 92.1322 %
Root relative squared error             97.363  %
Total Number of Instances             1958
```

Curriculum Vitae

Florian Hoenig

Fuxstr. 19
4600 Wels
Phone:+43-6504105070
E-Mail:hoenig@ingenium.at
Date of Birth: 02.10.1979 in Grieskirchen, Austria

Education

1986–1990	Volksschule Vogelweide Wels, Austria
1990–1998	Realgymnasium Brucknerstrasse Wels, Austria
2003–2004	ERASMUS Exchange Student, Universiteit Leiden, The Netherlands
1999-current	Computer Science Student, JKU Linz, Austria

Work Experience

1998–1999	Zivildienst Lebenhilfe Oberoesterreich in Wels, Austria
1999–2006	Software Developer, Allerstorfer-Hoenig-Valenti GnbR, Wels, Austria.

Languages German (native), English (fluent), Spanish (basic)

Eidesstattliche Erklaerung

Ich erklaere mich an Eides statt, dass ich die vorliegende Arbeit selbststaendig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benuetzt und die den benutzten Quellen woertlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Datum Unterschrift