

Feature extraction for image quality prediction

Citation for published version (APA):

Kayargadde, V. (1995). *Feature extraction for image quality prediction*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven. <https://doi.org/10.6100/IR435553>

DOI:

[10.6100/IR435553](https://doi.org/10.6100/IR435553)

Document status and date:

Published: 01/01/1995

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Feature Extraction For Image Quality Prediction



Vishwakumara Kayargadde

Feature Extraction
For
Image Quality Prediction

cover: prominent edges from the Terrace image detected using the algorithm described in Chapter 2 of this thesis.

Feature Extraction For Image Quality Prediction

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Eindhoven,
op gezag van de Rector Magnificus, prof.dr. J.H. van Lint,
voor een commissie aangewezen door het College van Dekanen
in het openbaar te verdedigen
op dinsdag 25 april 1995 om 16.00 uur

door

Vishwakumara Kayargadde

geboren te Bayar dorp, India

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. J.A.J. Roufs

prof.dr. H. Bouma

en de copromotor

dr.ir. J.B.O.S. Martens

The research described in this thesis was carried out in the Vision Research Group of the Institute for Perception Research (IPO), Eindhoven, The Netherlands. IPO is a joint venture between the Philips Research Labs, Eindhoven and the Eindhoven University of Technology.

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Kayargadde, Vishwakumara

Feature extraction for image quality prediction /
Vishwakumara Kayargadde. - Eindhoven : Eindhoven
University of Technology
Thesis Technische Universiteit Eindhoven. - With ref.
ISBN 90-386-0006-2
Subject headings: image quality / feature extraction.

Copyright © 1995 by Vishwakumara Kayargadde

Druk: Wibro dissertatiedrukkerij, Helmond

*mātru pitru guru
charanāravindayoh
susamarpitham*

Acknowledgements

I am deeply indebted to a number of people whose contributions to this thesis have been invaluable. In particular, I am grateful to: Prof. Jacques Roufs, for his encouragement, constant support and valuable advice; Dr. Jean-Bernard Martens, for the stimulating discussions during the research years and while writing the manuscript, that have contributed significantly to the style and the contents of this thesis; Prof. Herman Bouma, for reading the manuscript and suggesting a number of improvements; Dr. Huib de Ridder, for the invigorating discussions on psychophysical methods and multidimensional scaling that have helped me immensely in the analysis of experimental data.

I would like to thank Prof. Vanwormhoudt, Universiteit Gent, Prof. Tom Harrington, University of Nevada, and Neville Lobo, for reviewing the manuscript and for their valuable and constructive comments. I am especially grateful to Prof. Yegnanarayana, IIT Madras, for being a source of inspiration and enlightenment.

I am indebted to Prof. J. O. Ramsay, McGill University, for generously providing the MULTISCALE program. Thanks are also due to Dr. Hugo van Leeuwen for providing special T_EX style files which have been used in the preparation of this thesis.

Special thanks are due to all those who participated in my psychophysical experiments. I am also extremely grateful to all my colleagues at IPO for the excellent working atmosphere during my tenure at IPO.

Contents

1	Introduction	1
1.1	Image Quality	1
1.1.1	Definition	1
1.1.2	Measurement	2
1.1.3	Scope	3
1.2	Image Quality Prediction	4
1.3	Aim of this Thesis	7
2	Estimation of Edge Parameters, Image Blur, and an Objective Measure for Perceived Blur Using Local Derivatives	11
2.1	Introduction	12
2.2	Polynomial Transforms	16
2.3	Polynomial Coefficients of a Blurred Image	19
2.4	Polynomial Coefficients of an Edge	20
2.4.1	Hermite Coefficients of a Step Edge	21
2.4.2	Hermite Coefficients of a Blurred Edge	23
2.5	Estimation of Edge Parameters	25
2.5.1	Noiseless Case	25
2.5.2	Effect of Independent Errors in Polynomial Coefficients on Estimates	26
2.5.3	Effect of Image Independent, Additive Noise on Estimates	27
2.6	Blur Estimation Algorithm	31

2.6.1	Selecting Edge Candidates	32
2.6.2	Estimation of Blur Parameter at Edge Locations	33
2.6.3	Estimation of Uniform Image Blur	36
2.7	Multiscale Blur Estimation Algorithm	38
2.8	Results of Blur Estimation	41
2.9	Experiment: Perceived Effect of Blur on Natural Images	51
2.10	Estimation of Blur-index	55
2.10.1	Results of Blur-index Estimation	57
2.11	Conclusions	59
3	Estimation of Noise Variance and an Objective Measure for Perceived Noise Using Local Energy	61
3.1	Introduction	62
3.2	Noise Model Parameter Estimation	65
3.2.1	White Noise	66
3.2.2	Nonwhite Noise	66
3.3	Noise Variance Estimation Algorithm	68
3.3.1	White Noise	70
	A. Using Gradient Energy	70
	B. Using Gradient Amplitude	72
	C. Using Residual Energy	72
3.3.2	Nonwhite Noise	72
3.4	Results of Noise Parameter Estimation	73
3.5	Psychophysical Experiments to Measure Noisiness	78
3.5.1	Equipment and Subjects	78
3.5.2	Relevance of the PDF of the Noise to Noisiness	79
3.5.3	Experiment 1: Scaling Noisiness	80
3.5.4	Experiment 2: Comparing Noisiness	81
3.5.5	Experiment 3: Effect of Noise SD on Noisiness	86
3.5.6	Experiment 4: Effect of Local Luminance on Noisiness	88

3.5.7	Experiment 5: Effect of White Noise on Natural Images	91
3.5.8	Experiment 6: Effect of Noise Correlation Length on Noisiness	91
3.6	Noise-index: A Measure for Noisiness	96
3.7	Estimation of Noise-index	97
3.7.1	Results of Noise-index Estimation	100
3.8	Conclusions	103
4	Subjective Characterization of Images Degraded by Both Blur and Noise	105
4.1	Introduction	106
4.2	Experimental Setup	108
4.3	Results	109
4.3.1	MDS Analysis of Dissimilarities	109
4.3.2	Results of Scaling Attributes	117
4.3.3	MDS Analysis of Scaled Attributes	119
4.4	Conclusions	127
5	Estimation of Perceptual Attributes From Objective Measures	131
5.1	Introduction	132
5.2	Estimation of Objective Measures for Perceptual Attributes . .	134
5.2.1	Estimation of Blur-index	134
5.2.2	Estimation of Noise-index	135
5.3	Mapping From Psychometric Space to Perceptual Space	135
5.4	Estimation of Perceived Quality	142
5.5	Conclusions	146
6	An Algorithm to Estimate an Objective Measure for Global Brightness Contrast	147
6.1	Introduction	147
6.2	Estimation of Contrast-index	150

6.3	Results	155
6.4	Conclusions	158
7	Epilogue	159
	References	165
	Summary	175
	Samenvatting	179

Chapter 1

Introduction

1.1 Image Quality

In today's information age, images are widely used as an effective means of communicating information. Images produced by TVs, satellites and medical imaging systems are typical examples. In these imaging systems, the image quality as perceived by the human observer is often the deciding factor of the overall quality of the system. The perceptual quality of the images produced by such systems depends to a large extent on the choices made while designing them. To obtain the desired image quality and to have a cost-effective design cycle, such decisions must be based on a knowledge of image quality and the factors which influence it. These demands have made a systematic study of perceptual image quality essential.

1.1.1 Definition

Every person has a notion of image quality. This notion may however depend on the context. It is difficult to find a general definition of image quality that is applicable in all contexts. Roufs & Bouma (1980) defined perceptual image quality as "the degree of excellence of the image". The term 'subjective image quality' is often used instead of 'perceptual image quality'. Although both terms aim at the same concept, use of the word 'subjective' may sometimes lead to confusion (Roufs, 1992). The term subjective quality may imply that the impressions are personal and may differ widely across subjects. However, it is known that subjects are able to make consistent judgements of image quality and that judgements of different subjects coincide to a considerable extent. The term 'subjective image quality' may also mean that aesthetic components play a role. To avoid such confusion, the term 'perceptual image quality' is preferred (Roufs, 1992).

Perceptual image quality expresses the overall impression of an observer and hence is a global psychological attribute of an image. In addition to perceiving

the global attribute image quality, human observers of images also perceive several other (basic) perceptual attributes of images: sharpness, brightness, brightness contrast, noisiness, etc. These basic attributes affect the overall impression of image quality. In general, it is relatively easier to study and understand the factors influencing these basic attributes than to directly study those affecting perceived quality. For example, the decrease in sharpness of an image may be mainly due to blurring. Similarly, an increase in the noisiness of an image may be directly related to a increase in noise variance. In addition to sharpness, noisiness and brightness contrast, we will also come across another related basic perceptual attribute, unsharpness, in this thesis. The perceived lack of sharpness of an image is referred to as unsharpness or perceived blur. Thus, unsharpness is an attribute that implies a meaning that is opposite to that of the attribute sharpness. Another concept often used in image quality research, similarly to the global attribute perceptual image quality, is (global or overall) perceptual impairment. Perceived global impairment implies a concept that is opposite to that of perceptual quality (de Ridder, 1992; Nijenhuis, 1993). It therefore means the 'perceived degree of degradation of the image'. The perceptual quality of an image can decrease due to many reasons, for example due to noise in the image or blurring of the image. The physical processes, such as blur and noise, that lead to a decrease in the perceptual quality of images by causing physical damage to images are called (physical) impairments or degradations.

Since a strict definition of image quality is not available, operational definitions are often used. The operational definitions may vary depending on the context. At this point, it is important to introduce a distinction between two types of contexts in which image quality is used. First, there is the quality related to performing a task based on an image. Examples are: reading from a video display unit (VDU), detection of a target from an image, such as a tumor in a CT image or a tank in an aerial photograph. The image quality in such environments is called the performance-oriented quality (Hunt & Sera, 1978). This is different from the quality of an image in an entertainment environment such as TV or film. In the case of TV or film, the quality is mainly concerned with appreciation or involvement, hence the name appreciation-oriented quality is used. These two kinds of qualities may also influence each other (Roufs & Boschman, 1991). In this thesis, we are concerned with appreciation-oriented quality, although the feature extraction methods developed here could also be used in applications involving performance-oriented quality.

1.1.2 Measurement

To be useful, image quality has to be measurable. The measurements must be reliable and reproducible. Traditionally, image quality has been measured us-

ing psychophysical methods, where humans (subjects) participate as observers and judges of the images (stimuli). Such experiments involve four main components: stimuli, viewing conditions, methodology (instructions) and subjects. The stimulus generates a sensation in the subject. The strength of that sensation is expressed as a response by the subject according to the instruction (Roufs, 1992). For example, the instruction may be to judge an attribute of the stimulus, for instance sharpness, using a scale from 1 to 10. The reliability and reproducibility of the data depend on the chosen scenes and subjects, as well as on the viewing conditions and instructions. Recommendations on these were made by the CCIR (1986).

To ensure the validity and generality of the experimental results, special attention has to be paid to the following issues. The scenes have to be carefully chosen, bearing in mind that the variability of data between scenes may be considerably higher than that between subjects (Roufs, 1992; Roufs *et al*, 1994). Although the problem of scene selection is largely unsolved, it can be circumvented by using many scenes of different types. The number and type of subjects have to be chosen in light of the fact that differences exist between experienced and naïve subjects. Using subjects who are very familiar with either the scene or the kind of distortions may cause misleading results (Westerink, 1989). The methodology depends on the type of environment: performance-oriented or appreciation-oriented. Scaling according to categories using adjectives or numbers is one of the faster and more reliable methods to express the sensation. For example, for appreciation-oriented quality, the CCIR recommends the use of a category scale with adjectives: Excellent, Good, Fair, Poor, Bad. A numerical category scale has also been shown to be very reliable for scaling quality and its attributes (Roufs *et al*, 1990; Ridder & Majoor, 1990).

Inter-laboratory tests performed by international organizations such as the CCIR have shown that the results of scaling image quality are reliable and reproducible. An excellent example is the COST scaling experiment, where four video coding algorithms were compared on a category scale, using adjectives in five European languages (Allnatt *et al*, 1983). The results for the five countries were strikingly similar.

1.1.3 Scope

The notion of image quality is at least as old as photography. Its importance increased with the advent of television. The perceived quality of an image is influenced by several factors. The physical parameters of the imaging system that is used to produce an image are among the most important factors that influence the perceptual image quality. Because of this, an important task of

imaging system designers is to make the choices regarding the physical parameters of the system in such a way that the perceived quality of the images produced by the system is maximum. These choices have to be made within the given constraints on the system, such as constraints on maximum bandwidth, bit rate or display size. To make these choices, the designer must be aware of the relations between the physical parameters and the perceived image quality. For example, in image display system design, it is important to know the relation between perceived quality and physical parameters: bandwidth, resolution, display size, screen luminance, sampling lattice, etc., in order to make an optimal choice of parameters. In recent years, there has been a growing need to communicate or store large amounts of image data using limited bandwidth or bits. These bit-rate reductions are attained using image-coding or compression algorithms. The designers of image-compression algorithms aim at attaining bit rate reduction with no apparent loss of quality or with minimum loss. Therefore it is important to know which physical properties of the image influence quality the most. All these and many other problems regarding imaging system design require a thorough understanding of image quality and its relation to the physical parameters that influence it.

The scope of a good understanding of image quality is large. Decisions on physical parameters that influence the design of an imaging system (with a specified quality and cost) can be simplified, which means a much shorter and cost-effective design process. Existing systems can be made to offer better image quality at comparable cost. New imaging systems can be visualized and realized faster and at lower design cost.

1.2 Image Quality Prediction

Although the results of the psychophysical methods of measuring image quality are reliable, the experiments are very time-consuming and expensive. Special skills and experience as well as expensive equipments are needed to carry out such experiments. Because of these limitations, it is difficult to incorporate subjective tests in the design of imaging systems. This is especially true in a fast design cycle, where the designer would like to know, almost immediately, the quality of the image produced using a system or an algorithm. Such difficulties have created a demand for objective measures of image quality.

An objective measure for image quality must satisfy two important constraints. Firstly, it must correlate well with the perceived quality of the image. Therefore, equal intervals in the objective measure must correspond to approximately equal perceived differences in image quality. Secondly, it must be estimated from the image or from the physical parameters of the imaging system. The estimates may be given by simple equations or complicated models

implemented as computer programs or algorithms.

Several image quality measures are available in literature. Most of the early work on image quality metrics attempted to obtain objective measures for image display sharpness. The most successful among them are based on the Modulation Transfer Function (MTF). Examples of such approaches are Granger & Cupery (1972), Higgins (1977), Carlson & Cohen (1980), and Barten (1989,1990,1991,1993). These measures essentially compute a weighted sum of the product or the ratio of the MTF of the display and the modulation threshold function of the eye. Hultgren (1990) showed that many of these objective sharpness measures can be unified into one general framework. Although these MTF-based methods provide reliable estimates of display sharpness or quality, they cannot be successfully applied to images with multiple impairments, such as those encountered in coded images. This is mainly because the distortions in digitally coded images cannot be easily described using MTFs.

In applications such as image coding, image fidelity is often used as a quality criterion. The measures assume the existence of an 'original' ¹ image and a 'degraded' or 'processed' image of the same scene. A distance function between the two images is used to estimate quality. Such situations occur typically in image coding or transmission, where one would like to estimate the quality of the decoded picture, using the original image as a reference. Such an estimate may be used, for example, to pick the best coding scheme out of many competing schemes or to rank them in terms of their output image quality. Early attempts towards this goal were based purely on physical measures like root-mean-square error (Pratt, 1991). Although these measures perform well for certain distortions and scenes, in general they correlate poorly with the perceived quality of the image (Marmolin, 1986). This is especially true when images contain multiple impairments.

Image quality and its attributes are determined by the physical parameters of the image as well as by the properties of the human visual system. Therefore, knowledge about the human visual system plays an important role in understanding as well as predicting image quality. Most of the later methods for image-quality-prediction make use of knowledge about the human visual system. Limb (1979) incorporated the threshold and masking properties of the visual system in the quality metric. Lukas & Budrikis (1982) proposed a quality measure based on a spatio-temporal model of threshold vision that incorporates filtering and masking. Zetzsche & Hauske (1989) proposed a multiple channel model for quality prediction based on a 'ratio of Gaussians' pyramid and orientation selective filtering. Daly (1992, 1993) proposed a method to predict

¹By original image, we mean, the image of a scene from which all other images (of the same scene) are generated. Note that an original image of a scene, obtained using an imaging process is itself impaired (see Fig. 1.1.)

the visible difference between the original and processed image and uses it as image fidelity. A brief review of image quality metrics based on image fidelity is given by Ahumada (1993).

Knowledge about the human visual system also helps in designing better imaging systems (Lubin, 1993). For example, knowledge about invisibility of certain distortions can save many bits while coding images (Daly, 1992, 1993; Girod, 1992, 1993). The properties of the human visual system have been widely applied to design better image processing or coding algorithms (Anderson & Netravali, 1976; Budrikis, 1972; Mannos & Sakrison, 1974; Gray *et al.*, 1993). Many models of the human visual system have emerged to help build better imaging systems (Pearlman, 1978; Stockham Jr., 1972; Hall & Hall, 1977; Granrath, 1981; Watson & Ahumada, 1989).

Many models of image quality have been developed for specific applications in image coding. Many of these exploit the spatial or temporal masking properties of the human visual system (Girod, 1992; Watson *et al.*, 1986). Netravali & Prasada (1977) used the spatial masking property of the visual system in amplitude quantization of digitized pictures. Safranek & Johnston (1989) and Safranek *et al.* (1990) used an empirically derived masking model to optimize a sub-band image coder. Girod (1989) proposed a non-linear spatio-temporal model of human threshold vision and used it to predict the coding gain that can be achieved by incorporating masking effects in video coding.

A comparison of some early measures of image quality can be found in (Beaton, 1983). However, a comparison of recent measures with respect to their predictive power is not available. Hence, the value of most of the above-mentioned models remains to be established. An assumption common to most models is that the image quality is a uni-dimensional attribute shared by all subjects. This point of view is questionable. In our view, a better understanding and prediction of image quality can be attained by acknowledging that image quality is a multi-dimensional percept that is influenced by several basic perceptual attributes such as sharpness, noisiness, brightness contrast, etc., (Roufs & Bouma, 1980; Roufs, 1989). Marmolin & Nyberg (1975) identified some underlying dimensions of image quality using multi-dimensional scaling (MDS) techniques. They found that sharpness, noisiness and contrast were the three most important dimensions of image quality in their experiment. Other studies have also found similar dimensions, both for natural images (Goodman & Pearson, 1979) and for processed medical images (Escalante *et al.*, 1994; Escalante, 1992).

The multi-dimensional nature of image quality has to be taken into account while estimating quality. So far, very few quality measures have made use of the multi-dimensional nature of image quality. Nakayama *et al.* (1980) use the multi-dimensional nature of image quality in their model. According to their

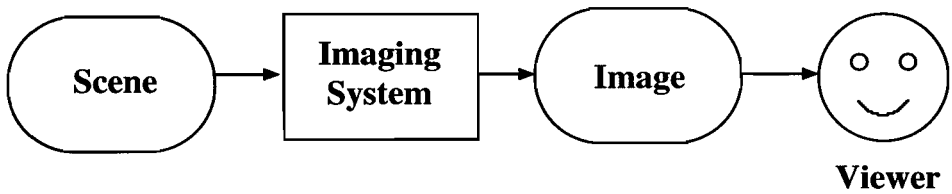


Figure 1.1: An image communication system

model, the overall image quality is given as a weighted sum of the underlying quality dimensions, and each underlying dimension, such as sharpness, is estimated separately. The multi-dimensional nature of image quality has also been used by Ridder, (1992) and Nijenhuis, (1993), to find metrics for image quality in terms of its attributes.

1.3 Aim of this Thesis

A schematic diagram of an image communication system is shown in Fig. 1.1. The imaging system captures the scene and produces an image which is viewed by a human observer. The imaging system may be simple, consisting of just a camera and a display or a printer, or may also contain additional components such as a processor, encoder-decoder, transmitter-receiver, etc. The imaging system is characterized by several *physical* parameters: band-width, resolution, display size, contrast parameter gamma, maximum luminance, minimum luminance, etc. It also introduces physical distortions such as blur, noise, etc., which can also be described using physical parameters such as the spread of the blurring kernel, the noise variance, etc. The human observer perceives several basic attributes of the image, such as sharpness, brightness contrast, overall brightness, noisiness, etc. All these basic attributes combine to form the overall impression of image quality. The image quality and its attributes are perceived notions and hence are described in a psychological or a *perceptual* domain. **The aim of this thesis is to contribute towards estimating the perceived quality of an image from the image itself.**

The approach used in this thesis is based on the multi-dimensional nature of image quality. It is based on the notion that the images are located in a perceptual space spanned by the important basic image attributes such as sharpness, noisiness, brightness contrast, etc. Quality is obtained by combining these basic attributes. The combination rules can be derived from the perceptual space. The perceptual image quality can therefore be estimated by first estimating the image attributes and then combining them. In this thesis we

consider three important basic attributes of image quality: sharpness, noisiness and global brightness contrast.

Each perceptual attribute is predominantly influenced by one physical parameter, although second-order influences from secondary parameters may also exist. For example, sharpness is mainly influenced by the spread of the blurring kernel (Westerink & Roufs, 1989; Westerink, 1991) and noisiness is mainly determined by the noise standard deviation. The sensitivity of these perceptual attributes to changes in the physical parameters is measurable (Watt & Morgan, 1983). Based on these sensitivity measurements, objective measures for the perceptual attributes can be derived from the physical parameters (Nijenhuis, 1993). Therefore, by estimating the perceptually relevant physical parameters from the image, the objective measures for the basic attributes of image quality can be estimated. For objective measures derived in this way (by taking the sensitivity of the visual system into account), equal intervals in the objective measure correspond to approximately equal perceived differences in the attribute. The objective measures derived in this way are thus psychometric measures (Hunt, 1978).

The physical parameters that predominantly influence the attributes are estimated from the image using computational algorithms. **The computational algorithms are based on image features that are perceptually relevant.** First we consider the basic attribute sharpness (**Chapter 2**). The computational algorithm to estimate the physical parameter influencing sharpness, the spread of the blurring kernel, is based on the edges in the image. Then we consider the basic attribute noisiness (**Chapter 3**). The algorithm to estimate the physical parameters influencing noisiness, the standard deviation of noise and the noise correlation length, are based on the uniform regions in the image. The relevant information from these features is extracted using local operators that have many properties in common with the human visual system. We use local derivative operators to estimate the blur and noise parameters.

Since our aim is to develop computational algorithms to estimate parameters influencing perceived image quality, it is compelling for us to use a transform that performs operations similar to those performed by the visual system. The local operators we use in the computational algorithms belong to an image description technique motivated by the properties of the human visual system called the Hermite transform (Martens, 1990a, 1990b). The Hermite transform performs a local image decomposition, where the image is localized by multiplying with a Gaussian window. The local image is decomposed in terms of many basic functions which satisfy certain properties such as orthogonality with respect to the window function. The Hermite transform facilitates detection as well as analysis of the perceptually relevant features in the image, such as uniform regions, edges, lines and corners. This is attained by local analysis

which makes the relevant information of the features explicit. An important property of the operators that are used to compute the Hermite transform of an image is that they are equal to the derivatives of a Gaussian. The human visual system is known to perform operations similar to the Gaussian derivative operations (Young, 1985). The Hermite transform is also in agreement with other image representations models such as the theory of scale-space representation (Koenderink, 1984, 1990; Witkin, 1984; Marr & Hildreth, 1980; Marr, 1982) and is related to the wavelet representation (Mallat, 1989; Mallat & Zhong 1992).

The parameter estimation algorithms presented in this thesis have wider application than just image-quality prediction. For example, estimates of the spread of the blurring kernel can be used to deblur the image (Demoment, 1989), or to estimate the depth in a scene (Pentland, 1987). Similarly, noise variance estimates can be used in segmentation or restoration of images (Rosenfeld & Kak, 1982).

The multi-dimensional perceptual space of images degraded by blur and noise is constructed using data collected in psychophysical experiments (**Chapter 4**). Using this perceptual space, the important basic attributes are identified, and the relations between them, such as dimensional orthogonality, are established (Roufs, 1989; Ashby & Townsend, 1986). The relation between the quality and its attributes is also established in the perceptual space. The number of variables involved in a psychophysical experiment are limited. Hence, the space obtained using the experimental data will only be a subspace of the entire perceptual space. This has to be borne in mind while comparing spaces obtained from different experiments involving different variables. The objective measures for the basic attributes sharpness and noisiness are estimated using the algorithms developed in Chapters 2 and 3. Based on these estimates, a notion of a psychometric space spanned by the objective (psychometric) measures for perceptual attributes is introduced (**Chapter 5**). The relation between the perceptual space and the psychometric space is established and by using that relation quality as well as other attributes of images are predicted.

Similar to sharpness and noisiness, global brightness contrast is an important attribute of image quality, and an objective measure for it is very useful for optimizing the perceived quality of images produced by imaging systems. A simple computational algorithm to estimate an objective measure for global brightness contrast is presented in **Chapter 6**.

Finally, the main conclusions of this thesis are listed in **Chapter 7**. Some issues that are closely related to the research reported here, but not addressed in this thesis, have been pointed out. Some areas that will require more research when the computational approach is extended to estimate other attributes, have also been identified.

As stated earlier, the main aims of this thesis are to find objective measures for perceptual attributes of image quality and to develop computational algorithms to estimate those objective measures. We also develop methods to estimate perceptual image quality using the estimates of the objective measures. Using these computational algorithms and methods, we intend to show how the objective measures for perceptual attributes can be derived from the image itself and show how the quality of the image can be estimated from the objective measures. We have chosen three important attributes, sharpness, noisiness and global brightness contrast, for this purpose, although the approach presented here can also be extended to other attributes. The main emphasis of this thesis is thus on computational methods for image quality. It is, however, not the intention of this thesis to explain the perceptual significance of the objective measures proposed, nor is it the intention to provide an explanation for the performance of a certain objective measure, although knowledge about these will help in building better computational models for image quality in future.

We have chosen to demonstrate the computational algorithms and methods for image quality estimation developed in this thesis using the simplest situation, the still, black and white images. Although, to be of greater practical merit, the objective measures and image quality estimation methods must also be applicable to moving colour images, no explicit attempts have been made here to extend the methods developed here to colour or moving images. This is mainly because of the following reasons. Firstly, it is our intention to test how well the approach used here performs with the simplest of the situations before attempting to extend it to more complex situations of colour and moving images. Secondly, it is beyond the scope of this thesis to extend and test the methods developed here to more complex situations of colour and video. In our view, the methods and framework presented here are general in nature and can be extended to more complex situations of colour and video.

Chapter 2

Estimation of Edge Parameters, Image Blur, and an Objective Measure for Perceived Blur Using Local Derivatives

Abstract

A method is presented for detecting blurred edges in images and for estimating the following edge parameters: position, orientation, amplitude, mean value and edge slope. The method is based on a local image-decomposition technique called a polynomial transform. The information that is made explicit by the polynomial transform is well suited for detecting image features such as edges and for estimating feature parameters. By using the relationship between the polynomial coefficients of a blurred feature and those of the *a priori* assumed (unblurred) feature in the scene, the parameters of the blurred feature can be estimated. The performance of the proposed edge-parameter estimation method in the presence of image noise has been analysed. An algorithm is presented for estimating the spread of a position-invariant Gaussian blurring kernel, using estimates at different edge locations over the image. First a single-scale algorithm is developed where one polynomial transform is used. A critical parameter of the single-scale algorithm is the window size, which has to be chosen *a priori*. Since the reliability of the estimate for the spread of the blurring kernel depends on the ratio of this spread to the window size, it is difficult to choose a window of appropriate size *a priori*. The problem is overcome by a multiscale blur-estimation algorithm where several polynomial transforms at

¹A substantial part of this chapter has been published in a paper in CVGIP Graphical Models and Image Processing (Kayargadde & Martens, 1994a)

²Some results of the blur estimation algorithm were presented at the International Workshop on Image Processing, Budapest (Kayargadde & Martens, 1994b), and the results of a comparison with psychophysical data were reported at ECVP-94 (Kayargadde & Martens, 1994c)

different scales are applied, and the appropriate scale for analysis is chosen *a posteriori*. By applying the blur estimation algorithm to natural and synthetic images with different amounts of blur and noise, it is shown that the algorithm gives reliable estimates for the spread of the blurring kernel even at low signal-to-noise ratios.

2.1 Introduction

Blur is a widespread and important degrading factor in images. In the case of images of natural scenes on television or in print, blur degrades the appreciation-oriented quality. In the case of medical or astronomical images, blur often affects the performance-oriented quality of the images, because it makes detection, classification or diagnosis more difficult. Image blur can have many causes: camera defocus, low-pass filtering performed to obtain reduced data rates while coding, pre-filtering before sampling, etc. Blurring is an inherent property of image formation systems and cannot be avoided. In certain operations such as sampling, some blurring may even be desirable to reduce sampling artifacts such as aliasing (Nijenhuis, 1993). Image blur can often be approximated by a convolution between the original scene and a blurring kernel.

There are many applications that need a quantitative description of the amount of blur. The prediction of image quality is one such application. By identifying the subjective dimensions of image quality (Marmolin & Nyberg, 1975) and estimating the degradation parameters such as the spread of the blurring kernel, the noise variance etc. that influence these dimensions, the quality can be predicted (Nakayama, Kurosu, Honjyo & Nishimoto, 1980). Because sharpness is an important quality dimension, which is heavily influenced by physical blur (Westerink & Roufs, 1989), blur estimation is one of the first steps in automatic image quality prediction. Image deblurring is another application that needs an estimate of the blurring kernel (Demoment, 1989; Hummel, Kimia & Zucker, 1987). In many deblurring applications, the nature of the blurring kernel is modelled and the parameters involved have to be estimated. For instance, in this chapter a Gaussian blurring kernel will be assumed. Image deblurring must be preceded by an estimation of the blur parameters. Depth estimation is another application where blur estimates are used (Pentland, 1987). By measuring the size of the blurring kernel (caused by camera defocus) at different positions in the image, a depth map can be drawn for the entire scene. Depth maps are very useful in tasks such as object recognition and scene interpretation in computer vision (Horn, 1986). The distance between a point in the scene and the lens, i.e., the depth, is related to the size of the blurring kernel and to the parameters of the lens system. Since

the parameters of the lens system are known, the problem of depth estimation is essentially reduced to one of estimating blur. It has been shown that the human visual system can also estimate the (relative) size of a blurring kernel with high accuracy (Hamerly & Dvorak, 1981; Watt & Morgan, 1983).

Gaussian blur is one of the most commonly encountered descriptions of blur. It has been shown that small amounts of blur due to camera defocus can be well approximated by a two-dimensional (2-D) Gaussian kernel (Pentland, 1987). In systems involving many blurring operations, the net effect, in light of the central limit theorem, can be described by a Gaussian kernel. Gaussian blur also underlies the process of generating image representations such as the scale-space representation (Koenderink, 1984; Witkin, 1984) and the Gaussian pyramid representation (Burt & Adelson, 1983). In the scale-space representation, images are parameterized by the scale parameter which is equal to the size of the Gaussian blurring kernel. Hence, estimating the spread of the blurring kernel is equivalent to locating the image in the scale space. In view of the above reasons, we have chosen to concentrate on Gaussian blur in our present work. An additional reason to those stated above is that the case of Gaussian blur is easy to tackle mathematically.

The spread of the blurring kernel has been shown to be an important determinant of the perceived sharpness in images (Westerink & Roufs, 1989). The sensitivity of the perceived blur to changes in the spread of the blurring kernel has been measured by Watt and Morgan (1983). Based on their data on just noticeable differences (JND) of blur, Nijenhuis (1993) derived an empirical relation for an objective measure for blur based on the spread of the blurring kernel. The objective measure for blur derived in this way has been shown to be a good correlate of perceived (un)sharpness, over a large range of blur spread (Nijenhuis, 1993). Thus, an objective measure for blur in an image can be estimated from the estimate of the spread of the blurring kernel.

When an image is blurred, the regions of sharp transition in the image such as edges and lines undergo a relatively large change, while uniform regions remain essentially unchanged. Hence, these features are the best candidates for estimating blur. Accordingly, edges are used in the present study to estimate the spread of a Gaussian blurring kernel.

As mentioned before, reliable estimates of edge parameters are essential for many applications in computer vision. Estimates of edge parameters also have applications in image coding, such as coding using oriented edges (Kunt, Ikonomopoulos & Kocher, 1985; Kunt, 1988) and coding using the local dimensionality of the image (Martens, 1990b). Many authors have developed methods to estimate the parameters of step edges (Lyvers & Mitchell, 1988; Lyvers, Mitchell, Akey & Reeves, 1989; Tabatabai & Mitchell, 1984; Chen & Medioni, 1989; Huertas & Medioni, 1986). Special attention has been paid to

the estimation of edge location (Tabatabai & Mitchell, 1984; Chen & Medioni, 1989; Huertas & Medioni, 1986). Lyvers *et al.* (1989) and Tabatabai & Mitchell (1984) used moment-based edge operators to estimate the parameters of a step edge. However, edges that occur in real images are blurred (Nalwa & Binford, 1986; Petrou & Kittler, 1991). Lyvers *et al.* (1989) used look-up tables to correct for the errors in the estimates caused by such deviations from the step edge model. Previous work on the estimation of edge blur has been mainly aimed at depth estimation (Pentland, 1987; Lai, Fu & Chang, 1993). The algorithm of Pentland (1987) is based on the observation that the Laplacian of an image in the vicinity of an edge is a linear function of the distance from edge and hence the slope can be estimated by linear regression. Although the details of the algorithm are somewhat obscure, it appears that an accurate and separate estimate of the edge orientation is required before this regression can be performed. The algorithm of Lai *et al.* (1993) fits an edge profile with three parameters to the image. An accurate and separate estimate of edge position is required for their algorithm.

We present a method for simultaneously estimating all parameters of an edge that is blurred by a Gaussian kernel. Our approach is inspired by the knowledge of the early stages of the human visual system. The visual system is thought to extract relevant information, mainly the location, orientation and contrast of luminance transitions, for further image analysis. This so-called primal sketch (Marr & Hildreth, 1980) is constructed by local processing with receptive fields of different form and size (Koenderink, 1984). To interpret the form of these receptive fields, a new image decomposition technique called a polynomial transform was developed by Martens (1990a). Polynomial transforms have been applied to image coding (Martens, 1990b), image deblurring (Martens, 1990c) and recently to noise reduction in images (Escalante Ramirez & Martens, 1992). Polynomial transforms have also been applied to detect and classify one-dimensional (1-D) patterns such as edges and lines (Martens, 1990b). The algorithm for estimating edge parameters that we present in this chapter is also based on the polynomial transform. Recent psychophysical experiments (Georgeson & Freeman, 1993; Georgeson & Freeman, 1994) indicate that the human visual system possibly uses an approach that is very similar to the algorithm for estimating blur proposed in this chapter.

The blur estimation algorithm presented in this chapter consists of two steps. Firstly, the regions of locally one-dimensional (1-D) edges in the image are detected. Secondly, the edge parameters are estimated at those locations. Both edge detection and estimation are carried out using polynomial transforms. *An important assumption made here is that there exist at least some locations in the imaged scene where the luminance distribution is locally an ideal step edge.* Scene refers to the viewer or camera centered (monocular) 2-D view of the luminance distribution of the 3-D scene. Note that the computa-

tional assumption made relates to the scene and not the image. An image is the result of a imaging process that inherently contains blurring. The above computational assumption is true for most natural scenes, although exceptions to this could be found. We assume that the blurring kernel is circularly symmetric, but using the theory we develop here, we can verify whether this assumption holds. Indeed, by pooling the estimates of edge blur for different orientations, we can check the circular symmetry of the blurring kernel. Similarly, by comparing the estimates of the edge blur at different positions in the image, the uniformity of blur over the image can be checked.

The scale of the polynomial transform has to be chosen *a priori* when using a single-scale algorithm where one polynomial transform is used. When an image contains edges with different amounts of blur, a single scale will not be optimal for estimating the edge parameters of all edges in the image. In order to avoid this *a priori* choice, a multiscale algorithm is presented, in which the image is analysed by polynomial transforms of different window sizes.

In applications such as image quality prediction and image deblurring, it is useful to have an estimate for the spread of the Gaussian blurring kernel over the entire image. When the blur is uniform over the entire image, image blur can be estimated by statistically combining the estimates of blur parameters at different edge locations over the image. If the blurring is due to a space-invariant operation such as camera defocus or low-pass filtering, then the blur is uniform over the image. In natural images, the assumption of uniform blur will be satisfied only in a loose sense, since in general all points in the scene will not be focused to the same extent. However, when the depth-of-field of the lens system used in the imaging device is large compared to the range of distances involved in the scene, it can be assumed that the blur is uniform over the entire image for most scenes.

This chapter is organized as follows. In Section 2.2, we briefly review polynomial transforms and introduce the relevant notation. The relation between the polynomial coefficients of a blurred signal and the unblurred signal is derived in Section 2.3. In Section 2.4, we first derive the polynomial coefficients of a step edge and later derive those of a blurred edge. In Section 2.5, we show how the parameters of a blurred edge can be estimated from its polynomial coefficients. In Section 2.6, we give a description of the single-scale blur estimation algorithm. The multiscale blur estimation algorithm is described in Section 2.7. The results of the blur estimation algorithm on different images are given in Section 2.8. A psychophysical experiment to scale unsharpness of natural images is described in Section 2.9. In Section 2.10, we present an algorithm to estimate a measure for perceived blur, called blur-index, and correlate the results of the algorithm with scaled unsharpness of images. Finally, the conclusions are summarized in Section 2.11.

2.2 Polynomial Transforms

Image analysis using polynomial transforms involves two steps. In the first step, the image is localized by multiplying with a window function. This windowing takes place at several positions over the entire image. The window positions constitute a sampling lattice S . In the second step, the image within every window is described as a sum of weighted polynomials. The polynomials that are orthogonal with respect to the square of the window function are used as the basis functions for the polynomial expansion (Martens, 1990a). For example, when the Gaussian window is used, the Hermite polynomials are used for the expansion. The mapping from the input image to the coefficients of the polynomials, referred to as polynomial coefficients, is called a forward polynomial transform. By interpolating the polynomial coefficients with windowed polynomials, the original image can be recovered. This mapping of the polynomial coefficients to an output image is called the inverse polynomial transform.

We first give a description of local image decompositions (Martens, 1993) and later discuss the specific case of polynomial transforms.

Given a window function $w(x, y)$, the input image $f(x, y)$ is decomposed into a sum of windowed images $f(x, y) \cdot w(x - p, y - q)$, i.e.,

$$f(x, y) = \frac{1}{h(x, y)} \sum_{(p, q) \in S} f(x, y) \cdot w(x - p, y - q), \quad (2.1)$$

where the periodic weighting function

$$h(x, y) = \sum_{(p, q) \in S} w(x - p, y - q) \quad (2.2)$$

is assumed to be different from zero for all (x, y) . We choose a set of basis functions such that they are orthonormal with respect to $w^2(x, y)$, i.e.,

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w^2(x, y) \varphi_{m, n-m}(x, y) \varphi_{l, k-l}(x, y) dx dy = \delta_{nk} \delta_{ml}, \quad (2.3)$$

for $n, k = 0, \dots, \infty, m = 0, \dots, n$ and $l = 0, \dots, k$ (Szegő, 1959; Martens, 1990a). We now compute a set of coefficients:

$$f_{m, n-m}(p, q) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) a_{m, n-m}(x - p, y - q) dx dy, \quad (2.4)$$

at all positions (p, q) in the sampling lattice S , for $n = 0, 1, \dots, \infty$ and $m = 0, \dots, n$. The functions

$$a_{m, n-m}(x, y) = \varphi_{m, n-m}(x, y) \cdot w^2(x, y) \quad (2.5)$$

are referred to as the analysis functions of order m along x and order $n - m$ along y . Expanding $f(x, y)$ in the basis functions $\varphi_{m, n-m}(x - p, y - q)$ and using the coefficients in Eq. (2.4) leads to a series expansion that converges in a weighted (by $w^2(x - p, y - q)$) quadratic sense, i.e.,

$$\lim_{N \rightarrow \infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w^2(x - p, y - q) \left[f(x, y) - \sum_{n=0}^N \sum_{m=0}^n f_{m, n-m}(p, q) \varphi_{m, n-m}(x - p, y - q) \right]^2 dx dy = 0, \quad (2.6)$$

for all sampling positions (p, q) in S . This result may also be expressed by saying that the image part within each window $f(x, y) \cdot w(x - p, y - q)$ is decomposed into a sum of orthonormal (unweighted) functions $w(x - p, y - q) \varphi_{m, n-m}(x - p, y - q)$.

The mapping from the image $f(x, y)$ to the coefficients $f_{m, n-m}(p, q)$ for all orders $m, n - m$ and positions $(p, q) \in S$ specifies the analysis stage or the forward transform. The inverse operation of synthesizing the image from these coefficients can also be performed and is called the inverse transform (Martens, 1990a).

In the particular case of polynomial transforms, the basis functions $\varphi_{m, n-m}(x, y)$ are polynomials, chosen so that they are orthonormal with respect to $w^2(x, y)$, as given by Eq. (2.3).

Several choices have to be made in selecting a specific polynomial transform. Firstly, the type of window function has to be selected. Based on psychophysical insights formulated in the scale-space theory (Koenderink, 1984; Witkin, 1984) and the evidence that the early stages of the human visual system use receptive fields that can be modelled as Gaussian derivatives (Young, 1985; Young, 1987), the Gaussian window is one of the best candidates for the window function. Additional motivation for using a Gaussian window in our application has been discussed in Chapter 1. In the case of estimating Gaussian blur, using a Gaussian analysis window also makes the problem mathematically tractable. Secondly, the size of the window function, also referred to as the spatial scale of the polynomial transform, has to be chosen. For example, when a Gaussian window is used, the spread of the window σ has to be set. In our application, we have adopted two approaches: a single-scale algorithm and a multiscale algorithm. In the single-scale algorithm, a polynomial transform of fixed window size is used. Hence, the size of the window has to be chosen *a priori*. A window that is small compared to the size of the blurring kernel will not contain a sufficiently large region of the blurred edge to make a reliable estimate. However, a window that is very large compared to the size of the blurring kernel will not be able to detect and locate edges accurately,

thus reducing the accuracy of the estimates. Moreover, increasing the window size also reduces the probability of detecting isolated edges. In the multiscale algorithm, polynomial transforms with windows of different sizes are applied, and the appropriate size of the window is selected *a posteriori*. Thirdly, the sampling lattice S has to be chosen so that the entire image is covered. This is attained by ensuring that the weighting function $h(x, y)$ does not reach values close to zero. We choose a rectangular sampling grid with a spacing that is a multiple of the sampling distance. Finally, we should also choose the highest order of the polynomial coefficients that have to be explicitly computed for the application at hand. It will be clear from our analysis that we need coefficients only up to order $n = 3$ in our application.

We use the Gaussian window for reasons discussed above. A useful property of 2-D Gaussian windows is that they are separable, i.e., $w(x, y) = w(x)w(y)$. The analysis functions are then also separable, i.e., $a_{m,n-m}(x, y) = a_m(x)a_{n-m}(y)$ and can be implemented efficiently. The orthonormal polynomials that are associated with the Gaussian window function

$$w^2(x, y) = \frac{1}{\pi\sigma^2} \exp \left[\frac{-(x^2 + y^2)}{\sigma^2} \right] \quad (2.7)$$

are the Hermite polynomials. The resulting polynomial transform is called the Hermite transform (Martens, 1990a). The analysis functions associated with the Hermite transform are

$$a_{m,n-m}(x, y; \sigma) = \frac{1}{\sqrt{2^n m! (n-m)!}} H_m \left(\frac{x}{\sigma} \right) H_{n-m} \left(\frac{y}{\sigma} \right) w^2(x, y), \quad (2.8)$$

where $H_n(x)$ is the Hermite polynomial of degree n in x . It is worth noting that the functions $a_{m,n-m}(-x, -y; \sigma)$ are equal to derivatives of a Gaussian (Martens, 1990a). Gaussian derivatives are used for detecting local features in images such as edges and lines (Bevington & Mersereau, 1987) and in psychophysical modelling of the human visual system (Marr & Hildreth, 1980).

The following energy measures derived from the polynomial coefficients will prove useful in our analysis. The n -th order local energy measure is defined as

$$E_n \triangleq \sum_{m=0}^n f_{m,n-m}^2 \quad (2.9)$$

for $n = 1, 2, \dots, \infty$. We will use these energy measures to estimate the parameters of a blurred edge.

The first-order energy E_1 is of particular interest in Chapter 3. The energy E_1 given by $E_1 = f_{0,1}^2 + f_{1,0}^2$ can be derived using two derivative filters: $a_{0,1}(x, y)$

and $a_{1,0}(x, y)$. The first-order energy E_1 is the square of the local gradient, i.e. $E_1 = f_1^2$, where the local gradient magnitude $|f_1| = \sqrt{f_{0,1}^2 + f_{1,0}^2}$.

Another measure that will be of interest in Chapter 3 is the overall local-energy measure

$$E_R \triangleq \sum_{n=1}^{\infty} E_n = \sum_{n=0}^{\infty} \sum_{m=0}^n f_{m,n-m}^2 - f_{0,0}^2, \quad (2.10)$$

i.e., is the sum of local energies of all orders (except zero). This is also called the residual energy (Martens, 1994), because it is the residual of signal energy minus the square of the local mean. The residual energy E_R can be derived by making use of the generalized Parseval's theorem for polynomial coefficients, without explicitly computing the polynomial coefficients (Martens, 1990a). It follows directly from the theorem that

$$E_R = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f^2(x, y) w^2(x - p, y - q) dx dy - \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) w^2(x - p, y - q) dx dy \right]^2, \quad (2.11)$$

which is a more efficient way of determining E_R (Martens, 1994).

2.3 Polynomial Coefficients of a Blurred Image

Let the image $f(x, y)$ be the result of blurring the scene $f'(x, y)$ by a blurring kernel $b(x, y)$. This blurring operation is denoted by the following convolution expression,

$$f(x, y) = f'(x, y) * b(x, y). \quad (2.12)$$

The forward polynomial transform of the blurred image can be interpreted as a multi-rate filter-bank with filters $a_{m,n-m}(-x, -y)$ (Martens, 1990a). The polynomial coefficients of the blurred image are therefore given by

$$\begin{aligned} f_{m,n-m}(p, q) &= [(f'(x, y) * b(x, y)) * a_{m,n-m}(-x, -y)]_{(p,q)} \\ &= [f'(x, y) * a'_{m,n-m}(-x, -y)]_{(p,q)}, \end{aligned}$$

where

$$a'_{m,n-m}(x, y) = a_{m,n-m}(x, y) * b(x, y) \quad (2.13)$$

is the blurred analysis function. In other words, the polynomial coefficients of an image blurred by a kernel $b(x, y)$, obtained using analysis functions

$a_{m,n-m}(x, y)$, are the same as the polynomial coefficients of the unblurred image obtained using blurred analysis functions $a'_{m,n-m}(x, y)$.

The above result is very useful in solving many problems involving blur. In this chapter we show how the above property can be used to derive the relationship between the Hermite coefficients of blurred and unblurred signals. We also exploit this relation to estimate the parameters of a blurred edge and the parameters of the image blurring kernel.

In our application, we are specifically interested in Gaussian blur for reasons discussed in the Introduction to this chapter. Hence, in the remainder, we will concentrate on the relation between the Hermite coefficients of a Gaussian blurred signal and those of the unblurred signal.

Specifically, if $b(x, y)$ is a Gaussian blurring kernel given by

$$b(x, y) = \frac{1}{\pi\sigma_b^2} \exp \left[\frac{-(x^2 + y^2)}{\sigma_b^2} \right] \quad (2.14)$$

then, by using the expression for the Fourier transform of $a_{m,n-m}(x, y; \sigma)$ (Martens, 1990a), it can be easily shown that

$$a'_{m,n-m}(x, y; \sigma) = \left(\frac{1}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right)^n a_{m,n-m}(x, y; \sigma\sqrt{1 + (\sigma_b/\sigma)^2}). \quad (2.15)$$

where $a_{m,n-m}(x, y; \sigma)$ is defined by Eq. (2.8). It follows from the equation above and from Eq. (2.4) that the Hermite coefficients of the blurred signal $f_{m,n-m}(p, q; \sigma)$ are given by

$$f_{m,n-m}(p, q; \sigma) = \left(\frac{1}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right)^n f'_{m,n-m}(p, q; \sigma\sqrt{1 + (\sigma_b/\sigma)^2}). \quad (2.16)$$

In other words, the Hermite coefficients of an image after blurring with a Gaussian kernel with parameter σ_b , obtained using a window of spread σ , are the same as the Hermite coefficients of the unblurred image obtained using a window of spread $\sigma\sqrt{1 + (\sigma_b/\sigma)^2}$, weighted by a factor $(1/\sqrt{1 + (\sigma_b/\sigma)^2})^n$.

2.4 Polynomial Coefficients of an Edge

In the previous section, we derived the relation between the polynomial coefficients of a blurred image and those of the unblurred image. In this section, we first derive the polynomial coefficients of a step edge. By using the relations derived in the previous section, we later derive the polynomial coefficients of a blurred edge.

We use the following *Gaussian* edge model

$$\text{edge}(x, y; V_e, \Delta V, \sigma_b, d, \theta) = V_e + \frac{\Delta V}{2} \operatorname{erf} \left[\frac{x \cos \theta + y \sin \theta - d}{\sigma_b} \right], \quad (2.17)$$

for a blurred edge with mean signal value V_e , height ΔV , blur parameter σ_b , distance from the origin d , and orientation θ . The above edge model is the result of a step edge with parameters V_e , ΔV , d , and θ blurred by a Gaussian blurring kernel of spread σ_b . In the limit $\sigma_b \rightarrow 0$, the above model reduces to a *step edge*. A schematic diagram of a step edge in (x, y) is shown in Fig. 2.1. Figure 2.2 shows the cross-section of a blurred edge along with the step edge by blurring of which the blurred edge was obtained.

2.4.1 Hermite Coefficients of a Step Edge

The polynomial coefficients of a step edge for a window $w(x, y)$ centered at the origin are given by (Martens, 1990b)

$$f'_{m, n-m} = f'_n \sqrt{\frac{n!}{m!(n-m)!}} \cos^m \theta \sin^{n-m} \theta, \quad (2.18)$$

where

$$f'_0 = V_e + \frac{\Delta V}{2} \Omega_0(d) \quad (2.19)$$

$$f'_n = \frac{\Delta V}{2} \Omega_n(d) \quad (2.20)$$

with

$$\Omega_n(d) = \int_d^{+\infty} \varphi_n(x) w^2(x) dx - \int_{-\infty}^d \varphi_n(x) w^2(x) dx, \quad (2.21)$$

for $n = 0, 1, \dots, \infty$, $m = 0, \dots, n$, where $\varphi_n(x)$ is the orthonormal polynomial of degree n over $w^2(x)$. From Eq. (2.18) and Eq. (2.9), the n -th order energy of a step edge is given by

$$E'_n = |f'_n|^2 \quad (2.22)$$

or equivalently, $|f'_n| = \sqrt{E'_n}$.

Specifically, for a Gaussian window of spread σ , centered at the origin, it can be shown that

$$f'_n = \frac{\Delta V \cdot \sigma}{\sqrt{2n}} a_{n-1}(d; \sigma), \quad (2.23)$$

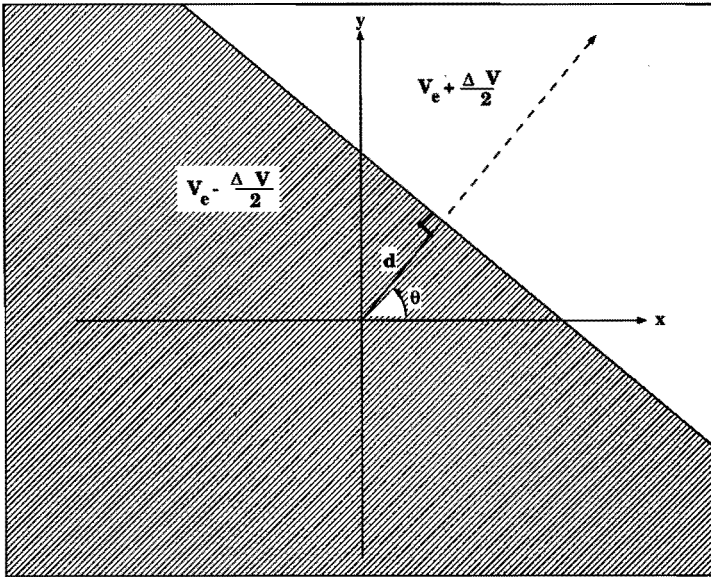


Figure 2.1: A *step edge* in (x, y) with height ΔV , mean value V_e , and orientation θ located at a distance d from the origin.

for $n = 1, \dots, \infty$, where $a_n(-d; \sigma)$ is the n -th order derivative of a Gaussian (Martens, 1990b). From the above expression we obtain

$$\begin{aligned}
 f'_0 &= V_e - \frac{\Delta V}{2} \operatorname{erf} \left[\frac{d}{\sigma} \right] \\
 f'_1 &= \frac{\Delta V}{\sqrt{2\pi}} \exp \left[- \left(\frac{d}{\sigma} \right)^2 \right] \\
 f'_2 &= \frac{\Delta V}{\sqrt{2\pi}} \frac{d}{\sigma} \exp \left[- \left(\frac{d}{\sigma} \right)^2 \right] \\
 f'_3 &= \frac{\Delta V}{\sqrt{2\pi}} \frac{1}{\sqrt{6}} \left(2 \left(\frac{d}{\sigma} \right)^2 - 1 \right) \exp \left[- \left(\frac{d}{\sigma} \right)^2 \right]
 \end{aligned} \tag{2.24}$$

up to order three. The corresponding energies can be computed using Eq. (2.22). Note that the coefficients f'_n are independent of the edge orientation θ .

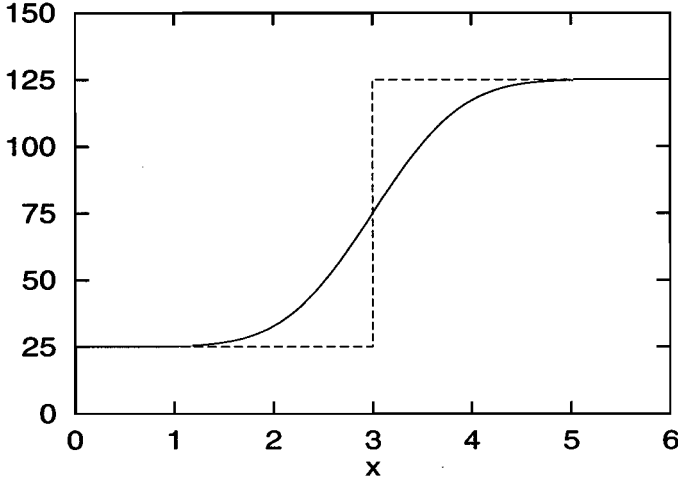


Figure 2.2: The cross-section of a blurred edge along the direction of the edge (drawn curve), for an edge with height $\Delta V = 100$, mean value $V_e = 75$, located at a distance $d = 3$ from the origin and blurred with a Gaussian blurring kernel with $\sigma_b = 1$. The cross-section of the step edge by blurring of which the blurred edge was obtained is also shown (dotted curve).

2.4.2 Hermite Coefficients of a Blurred Edge

In this subsection we make use of the results derived in section 2.3 to obtain the Hermite coefficients of a blurred edge from those of a step edge. Using Eq. (2.16), Eq. (2.18) and Eq. (2.23) we obtain the Hermite coefficients of a Gaussian blurred edge, where the blurring kernel is given by Eq. (2.14). Hermite coefficients of a blurred edge can also be written in the form of Eq. (2.18), i.e.,

$$f_{m,n-m} = f_n \sqrt{\frac{n!}{m!(n-m)!}} \cos^m \theta \sin^{n-m} \theta. \quad (2.25)$$

Specifically, up to order three, we obtain

$$\begin{aligned} f_0 &= V_e - \frac{\Delta V}{2} \operatorname{erf} \left[\frac{d/\sigma}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right] \\ f_1 &= \left(\frac{1}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right) \frac{\Delta V}{\sqrt{2\pi}} \exp \left[-\frac{(d/\sigma)^2}{1 + (\sigma_b/\sigma)^2} \right] \end{aligned} \quad (2.26)$$

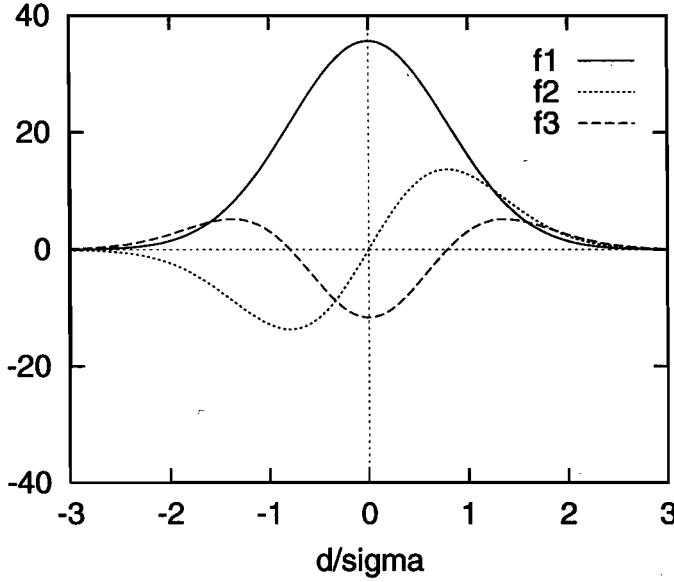


Figure 2.3: Coefficients f_1 , f_2 and f_3 of an edge, as a function of d/σ . The edge parameters are $\sigma_b/\sigma = 1/2$ and $\Delta V = 100$ (arbitrary units).

$$f_2 = \left(\frac{1}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right)^2 \frac{\Delta V}{\sqrt{2\pi}} \frac{d/\sigma}{\sqrt{1 + (\sigma_b/\sigma)^2}} \exp \left[-\frac{(d/\sigma)^2}{1 + (\sigma_b/\sigma)^2} \right]$$

$$f_3 = \left(\frac{1}{\sqrt{1 + (\sigma_b/\sigma)^2}} \right)^3 \frac{\Delta V}{\sqrt{2\pi}} \frac{1}{\sqrt{6}} \left(\frac{2(d/\sigma)^2}{1 + (\sigma_b/\sigma)^2} - 1 \right) \exp \left[-\frac{(d/\sigma)^2}{1 + (\sigma_b/\sigma)^2} \right]$$

The coefficients of higher orders can also be derived, but are not needed here. The resulting energies are given by $E_n = |f_n|^2$. Note that, when $\sigma_b = 0$, Eq. (2.26) reduces to Eq. (2.24). Figure 2.3 shows f_1 , f_2 and f_3 of a blurred edge of height $\Delta V = 100$ (in arbitrary units) and $\sigma_b/\sigma = 1/2$ as a function of d/σ . Note that when f_2 is maximum, f_3 equals zero. This can also be seen by solving the equation

$$\frac{\partial f_2}{\partial d} = 0 \quad (2.27)$$

to obtain $(d/\sigma)^2 = (1 + (\sigma_b/\sigma)^2)/2$, which corresponds to the positions $d = \pm\sigma/\sqrt{[1 + (\sigma_b/\sigma)^2]/2}$ where $f_3 = 0$.

2.5 Estimation of Edge Parameters

2.5.1 Noiseless Case

The edge model is completely defined by five parameters : mean value V_e , height ΔV , distance from the window centre d , orientation θ and blur parameter σ_b . Using the relationships derived in the previous section, we can estimate the parameters of an edge from the polynomial coefficients. We need at least five independent equations to determine all five edge parameters. The mean value V_e appears only in the equation for $f_{0,0}$, hence this equation is needed for estimating V_e . In choosing the rest of the equations, we make use of the fact that the signal-to-noise ratio of the coefficients decreases with increasing order of the polynomial coefficients (Martens, 1990c). We choose the other four equations so that they involve coefficients of the lowest order possible. It is easy to see that we need coefficients up to order three to solve for all the five edge parameters. From Eq. (2.26) we derive the following relationships

$$\frac{f_2}{f_1} = \frac{1}{1 + (\sigma_b/\sigma)^2} \frac{d}{\sigma} \quad (2.28)$$

$$\frac{f_3}{f_1} = \frac{1}{1 + (\sigma_b/\sigma)^2} \frac{1}{\sqrt{6}} \left(\frac{2(d/\sigma)^2}{1 + (\sigma_b/\sigma)^2} - 1 \right), \quad (2.29)$$

between the polynomial coefficients of an edge. By solving the above set of equations we obtain

$$\frac{\sigma_b}{\sigma} = \left[\left(\frac{2f_2^2}{f_1^2} - \frac{\sqrt{6}f_3}{f_1} \right)^{-1} - 1 \right]^{1/2} \quad (2.30)$$

$$\frac{d}{\sigma} = \frac{f_2}{f_1} \left(\frac{2f_2^2}{f_1^2} - \frac{\sqrt{6}f_3}{f_1} \right)^{-1}. \quad (2.31)$$

By substituting the above two results in the expression for f_1 , we obtain

$$\Delta V = f_1 \sqrt{2\pi} \left(\frac{2f_2^2}{f_1^2} - \frac{\sqrt{6}f_3}{f_1} \right)^{-(1/2)} \exp \left[\frac{f_2^2}{f_1^2} \left(\frac{2f_2^2}{f_1^2} - \frac{\sqrt{6}f_3}{f_1} \right)^{-1} \right]. \quad (2.32)$$

The mean value V_e is obtained by substituting the above parameter values in $f_{0,0} = f_0$. All edge parameters except the orientation θ can be estimated just from the coefficients up to order three. The edge orientation θ can be obtained from the ratio $f_{0,1}/f_{1,0} = \tan \theta$. The signs of d and ΔV are determined from the signs of the first and the second order coefficients.

To use Eq. (2.30) for blur estimation, we need f_1, f_2 and f_3 . We first compute their magnitudes from energies E_1, E_2 and E_3 computed using Eq. (2.9) and Eq. (2.22). The signs of f_1, f_2 and f_3 are obtained using the relation between the signs of the horizontal and the vertical coefficients of order 1, 2, and 3: $f_{0,1}, f_{1,0}, f_{0,2}, f_{2,0}, f_{0,3}$ and $f_{3,0}$. From Eq. (2.25), $f_{1,0} = f_1 \cos \theta$ and $f_{0,1} = f_2 \sin \theta$. Hence, the signs of first order coefficients $f_{1,0}$ and $f_{0,1}$ depend on the sign of ΔV as well as that of $\cos \theta$ and $\sin \theta$ respectively, whereas the sign of f_1 is given by the sign of ΔV alone. Therefore, the sign of f_1 can be determined by finding the signs of $\cos \theta$ and $\sin \theta$, which is equivalent to finding to which quadrant the θ belongs. Based on these facts, the following steps are used to determine the sign of f_1 :

$$\begin{array}{llll} \text{if } f_{0,1} \cdot f_{1,0} > 0 & \text{then} & \text{if } f_{0,1} + f_{1,0} > 0 & \text{then } f_1 \text{ positive} \\ & & & \text{else } f_1 \text{ negative} \\ & \text{else} & \text{if } f_{0,1} - f_{1,0} > 0 & \text{then } f_1 \text{ positive} \\ & & & \text{else } f_1 \text{ negative} \end{array} \quad (2.33)$$

From Eq. (2.25), $f_{2,0} = f_2 \cos^2 \theta$ and $f_{0,2} = f_2 \sin^2 \theta$. Hence, the sign of f_2 is given by the sign of $f_{2,0} + f_{0,2}$. From Eq. (2.29), the sign of f_3 is opposite to that of f_1 when $|d| < \sigma / \sqrt{[1 + (\sigma_b/\sigma)^2]/2}$ and the same as that of f_1 otherwise. The sign of the ratio $f_{0,3}/f_{0,1}$ or $f_{3,0}/f_{1,0}$ can be used to decide this. We use the most reliable ratio of the two, as follows:

$$\begin{array}{llll} \text{if } |f_{1,0}| > |f_{0,1}| & \text{then} & \text{if } f_{3,0}/f_{1,0} > 0 & \text{then } f_3/f_1 \text{ positive} \\ & & & \text{else } f_3/f_1 \text{ negative} \\ & \text{else} & \text{if } f_{0,3}/f_{0,1} > 0 & \text{then } f_3/f_1 \text{ positive} \\ & & & \text{else } f_3/f_1 \text{ negative} \end{array} \quad (2.34)$$

Having obtained the coefficients f_1, f_2 , and f_3 , the edge parameters can be computed using Eq. (2.30), Eq. (2.31) and Eq. (2.32).

2.5.2 Effect of Independent Errors in Polynomial Coefficients on Estimates

The independent errors in the measured energies occur mainly due to the quantization of filter coefficients and the quantization of polynomial coefficients. For small independent errors in the energies it is possible to express the error in the estimates of edge parameters in terms of the partial derivatives. For example, the error in the estimate of σ_b can be expressed as

$$\begin{aligned} \Delta \sigma_b &= \frac{\partial \sigma_b}{\partial f_1} \Delta f_1 + \frac{\partial \sigma_b}{\partial f_2} \Delta f_2 + \frac{\partial \sigma_b}{\partial f_3} \Delta f_3 \\ &= S_1(\sigma, \sigma_b, d, \Delta V) \Delta f_1 + S_2(\sigma, \sigma_b, d, \Delta V) \Delta f_2 + S_3(\sigma, \sigma_b, d, \Delta V) \Delta f_3 \end{aligned} \quad (2.35)$$

where $S_1(\sigma, \sigma_b, d, \Delta V)$, $S_2(\sigma, \sigma_b, d, \Delta V)$ and $S_3(\sigma, \sigma_b, d, \Delta V)$ are sensitivity factors expressed in terms of edge parameters and window spread. Similarly, estimation errors in other edge parameters can also be expressed in terms of their partial derivatives.

Figure 2.4 shows the error in σ_b as a function of different parameters for the case of $\Delta f_1 = \Delta f_2 = \Delta f_3 = 0.5$. These figures are obtained by plotting the analytical expressions for $\Delta\sigma_b$ given by Eq. 2.35 against its parameters. Figure 2.4a plots the error in the estimate of σ_b as a function of σ (for fixed σ_b , d and ΔV). The error is minimum when the window spread σ is approximately equal to the blur parameter σ_b . Figure 2.4b plots the error in the estimate of σ_b as a function of the distance from the window centre d (for fixed σ_b , σ and ΔV). The error in the estimate of σ_b increases almost linearly with distance. Figure 2.4c plots the error in the estimate of σ_b as a function of the edge height ΔV (for fixed σ_b , σ and d). The error in the estimate of σ_b decreases with increasing ΔV since the signal-to-noise ratio at the edge increases with ΔV .

2.5.3 Effect of Image Independent, Additive Noise on Estimates

The statistical nature of the estimates of edge parameters in the presence of image noise has to be studied before using the method on real images. In many practical situations, the noise can be modelled as an additive Gaussian random process which is independent of the image (Andrews & Hunt, 1977). We assume that the noise in the image can be modelled as image-independent, stationary, additive Gaussian stochastic process with zero-mean. When the additive noise is a Gaussian random process, the polynomial coefficients are jointly Gaussian random variables whose covariance matrix can be easily computed once the noise autocorrelation function is known (Escalante Ramírez & Martens, 1992). To study the statistical properties of the edge parameters, we need to obtain their probability density functions (PDFs).

The edge orientation θ is given by $\tan^{-1}(f_{0,1}/f_{1,0})$. The polynomial coefficients $f_{1,0}$ and $f_{0,1}$ are independent Gaussian random variables with equal variance (Escalante Ramírez & Martens, 1992). Hence, the angle is a ratio of two independent Gaussian random variables with equal variance, with an arc tangent transformation. The PDF of the angle can be derived (Lyvers & Mitchell, 1988). The angle so estimated is unbiased. Because of the non-linear nature of the expressions for the other edge parameters, it is not possible to obtain analytical expressions for their PDFs. Instead we use computer simulations to study their nature.

In the computer simulations, an edge of known height, orientation, mean value and position is blurred with a 2-D Gaussian kernel with known blur parameter, σ_b . In each trial of the simulation, the following steps are performed.

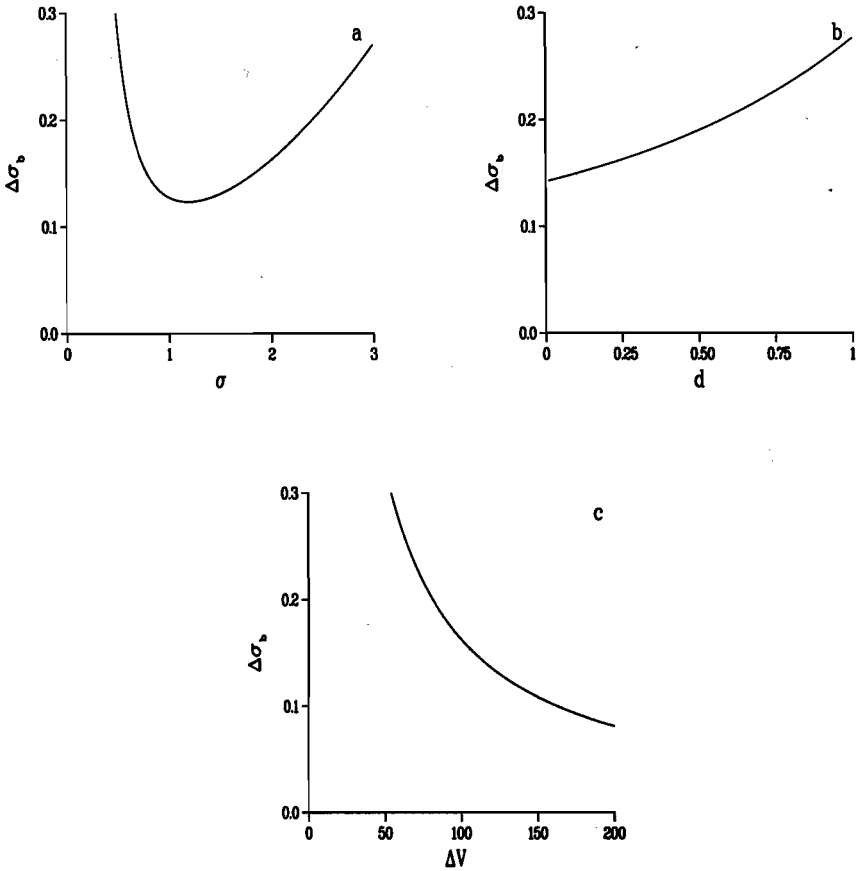


Figure 2.4: Error in the estimate of σ_b due to independent errors in f_1, f_2 , and f_3 . Plot (a) shows the error as a function of σ , for $\sigma_b = 1.0$ pixel-width, $d = 0.25$ pixel-width and $\Delta V = 100$ (arbitrary units) Plot (b) shows the error as a function of d , for $\sigma = 2$, $\sigma_b = 1.0$, and $\Delta V = 100$. Here d is expressed in units of sampling interval. Plot (c) shows the error as a function of ΔV , for $\sigma = 2$, $\sigma_b = 1.0$, and $d = 0.25$. In each of the 3 plots, $\Delta f_1 = \Delta f_2 = \Delta f_3 = 0.5$ (arbitrary units).

First, noise with Gaussian PDF of known variance is added to the blurred edge. Following this, the noisy blurred edge is analysed using a polynomial transform to obtain the coefficients. Finally, the edge parameters are computed. The signal-to-noise ratio (SNR) at an edge can be defined as

$$\text{SNR} = \frac{\Delta V}{\sigma_n} \quad (2.36)$$

where ΔV is the signal difference corresponding to the edge height and σ_n is the standard deviation of the noise.

The probability density functions of θ , ΔV , d/σ , and σ_b/σ , obtained by simulations at three different SNRs, are shown in Fig. 2.5. In the figure edge height $\Delta V = 100$ (arbitrary units) and the noise standard deviation $\sigma_n = 5, 10$, and 20 (arbitrary units). The dotted vertical lines show the values that the parameters should have when there is no noise in the edge. The variance of all the estimates increases with decreasing SNR. The estimate of the angle θ is unbiased, as was also deduced from the analytical expression. This also serves as a check on the validity of the simulations. The PDF of the estimate of edge height remains symmetrical with increasing SNR, indicating no significant bias. The PDF of the estimate of edge location d/σ is symmetrical at high SNRs, but deviates slightly from symmetry when the SNR drops to 5. The bias due to this asymmetry is small, for example, when $\text{SNR} = 5$ and d/σ is minimum, i.e., $d/\sigma = 0$, we obtain a bias of 0.001, and for the same SNR when $d/\sigma = 0.5$ we obtain a bias of 0.07. The PDF of the estimate of σ_b/σ remains symmetrical about the mean at high SNRs, but becomes increasingly asymmetrical at low SNRs. At SNRs above 10 the bias in the estimate of σ_b/σ is less than 5%. At low SNRs the blur parameter σ_b is underestimated. This is due to the nonlinear nature of Eq. (2.30). The operation of squaring f_2 in Eq. (2.30) causes the bias. It will be shown in the next section that the bias in the estimate of σ_b is proportional to $(\sigma_n/f_1)^2$.

In the analysis so far we have dealt with images in the continuous spatial domain. Since we are dealing with discrete images, the effect of sampling needs to be looked into. The effect of sampling on polynomial coefficients has been studied (Martens, 1990c). When the window spread σ is chosen so that $\sigma > 3T/2$ where T is the sampling interval, the effect of sampling on the signal-to-noise ratio of the polynomial coefficients is negligible. In our application we choose $\sigma > 3T/2$ in order to keep the effects of sampling on polynomial coefficients minimal. Once this has been guaranteed, the effect of sampling on the parameter estimation is minimum, since the estimates are based on the coefficients.

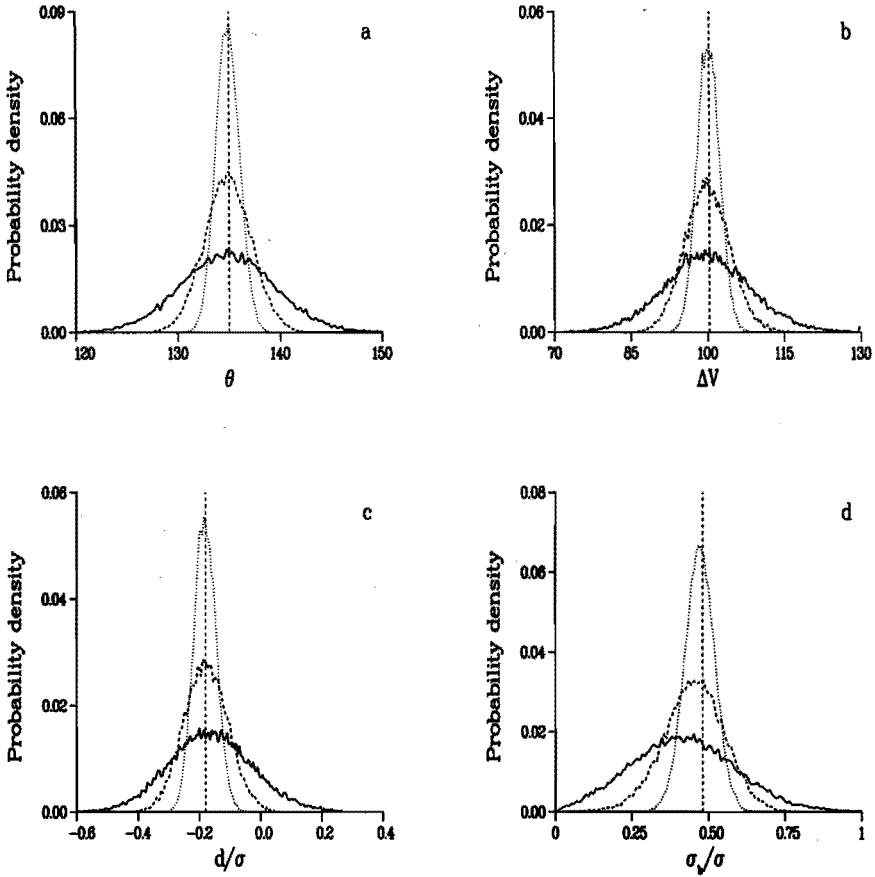


Figure 2.5: The PDFs of the edge parameter estimates. Panel (a) shows PDF of angle θ , (b) PDF of height ΔV , (c) PDF distance of d/σ and (d) PDF of σ_b/σ . The dotted, dashed and solid lines correspond to SNRs 20, 10 and 5 respectively. The dashed vertical lines show the estimates for a noiseless edge. The σ of the analysis window is 2. Each curve is the result of 32000 simulation trials.

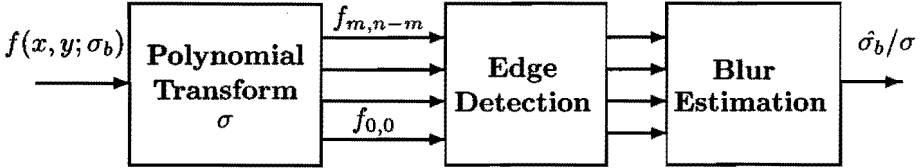


Figure 2.6: Single-scale blur estimation algorithm

2.6 Blur Estimation Algorithm

In this section we show how the relations derived in the previous section on the estimation of edge parameters can be used to estimate the amount of Gaussian blur in natural images. We adapt the following computational assumption about the *scene*:

We assume that there are at least a few locations in the captured *scene* where the luminance distribution is locally a one-dimensional step edge.

A one-dimensional (1-D) step edge is constant in some direction and varies as a step function in the perpendicular direction (see Fig. 2.1). The above computational assumption is true for most natural scenes. A block diagram of the proposed algorithm is shown in Fig. 2.6. Since the relations derived in the previous section hold for an edge, we first have to detect the locations of locally edge-like patterns in the image. This task is performed by the edge detector shown in Fig. 2.6. At each identified edge location the required edge parameters σ_b/σ , d/σ , or ΔV can be estimated from the Hermite coefficients using Eq. (2.30), Eq. (2.31) or Eq. (2.32) respectively. Thus, we have a set of observations obtained from many different edge locations over the image. In the case of uniform blur, the goal is to obtain an estimate of σ_b for the entire image by statistically combining the blur estimates at all edge points.

2.6.1 Selecting Edge Candidates

Since the use of a strict definition of an edge is somewhat impractical, it is customary to adopt an operational definition. We define edge positions as those locations where a significant locally one-dimensional transition occurs between distinct image values. Such transition regions are characterized by a local maximum in the gradient magnitude $|f_1|$ in the direction θ , where θ is the angle of the edge as defined in Fig. 2.1. Therefore edges can be detected by locating local maximum in $|f_1|$ in the direction θ , where $|f_1| = \sqrt{f_{1,0}^2 + f_{0,1}^2}$ and $\theta = \tan^{-1}(f_{0,1}/f_{1,0})$. Local maximum in $|f_1|$ are detected by finding the points where $|f_1|$ is larger than its two closest neighbours whose positions are in the direction indicated by θ at that point. The edge locations detected in this way are reorganized into chains of edge locations (Mallat & Hwang, 1992). A chain of edge locations is a series of consecutive, connected edge positions, where each edge location acts as a link in the chain. Thus an edge-chain is characterized by parameters chain-length and average f_1 over all edge positions in the chain. The chaining of the edge locations is based on the fact that the f_1 varies smoothly across the edge curve, and that the orientation θ of the edge at any location on the curve is perpendicular to the tangent of the edge curve at that location. Among the edge-chains that are detected in this way, only those chains whose average $|f_1|$ is greater than a given threshold and whose chain-length is greater than a given threshold are considered to be reliable and prominent edge candidates. Applying thresholds on the average $|f_1|$ over the chain and the chain-length helps to decrease the probability of false alarm of edges. Edges with high f_1 and large segment length are relatively more prominent in an image and hence are assumed to play a greater role in determining the perceived blur.

The blur estimation method is based on a locally 1-D edge. Not all positions selected by the above detection process are locally 1-D. For instance, corners are also detected (see for example, top rows of Fig. 2.15 and 2.16). A 1-D pattern is constant in some direction and varies in the perpendicular direction. Typical 1-D patterns are edges, lines, etc. The Hermite coefficients of a locally 1-D pattern are equal to zero for all orders greater than zero in one direction, and non-zero for at least one order greater than zero in the perpendicular direction. From the edge locations detected above, the 1-D edges are detected by rejecting all those locations where the pattern is locally 2-D. A 2-D pattern varies in two (orthogonal) dimensions. Examples of 2-D patterns are corners, curves with large curvature, etc. We use a 2-D energy measure E_{2D} for rejecting the 2-D patterns (Martens, 1995). The E_{2D} indicates the extent to which a pattern is locally two-dimensional. For example, E_{2D} up to order 2 is given by

$$E_{2D} = \left[\left(\sqrt{(f_{2,0} - f_{0,2})^2 + 2f_{1,1}^2} - |f_{2,0} + f_{0,2}| \right) / 2 \right]^2 \quad (2.37)$$

Ideally, locally 1-D patterns are characterized by $E_{2D} = 0$. For example, 1-D edges and lines have E_{2D} equal to zero, whereas locally 2-D patterns such as corners and steep curves have a high value of E_{2D} . We detect locally 1-D edges by putting a threshold on the ratio E_{2D}/E_1 . Using this threshold, 2-D patterns such as corners which have a high value of E_{2D} are rejected. For our estimate of the blur to be reliable, the pattern on which the estimate is based has to be a 1-D edge. In other words, the cost associated with a false alarm is much higher than that associated with a miss. Hence, we keep a sufficiently low threshold on E_{2D}/E_1 to be certain that the locations identified really correspond to 1-D edges.

2.6.2 Estimation of Blur Parameter at Edge Locations

Having identified the edge locations in the image we now have to estimate the edge parameters at each location. When the distance between the centre of the window and the edge (d/σ) is small, the windowed signal contains most of the region of the edge transition, i.e., the region of the edge from which most reliable estimates of the blur can be obtained. However, when d/σ is large, the windowed signal contains only the tail of the edge and estimates based on that are bound to be less reliable, especially in the presence of noise. When d/σ is small, the dependence of the estimated blur on the blur model is also small since the transition region is approximately linear for all realistic blur models. For these reasons, we are interested in making sure that the window is positioned sufficiently close to the centre of the edge. Since the windowing takes place at regular intervals in the image, an edge location that is far from the centre of one window will be close to the centre of another window. Hence, if the window spread σ is sufficiently large compared to the sampling distance, most edge locations will be included at least in one window.

To obtain a reliable estimate of the blur spread, we want to ensure that the edge is sufficiently close to the centre of the window. The position of an edge relative to the window centre can be verified using the ratio f_3/f_1 . By applying a threshold on the ratio

$$\frac{f_3}{f_1} < \beta \quad (2.38)$$

we reduce the range of allowed distances, d/σ . The result of using Eq. (2.38) in the $(d/\sigma, \sigma_b/\sigma)$ plane is shown in Fig. 2.7. The area of the region can be varied by varying β in the range $\beta > -1/\sqrt{6}$. In Fig. 2.7, we have chosen $\beta = -1/2\sqrt{6}$. When $-1/\sqrt{6} < \beta < -1/2\sqrt{6}$ the reliable region is bounded in the $(d/\sigma, \sigma_b/\sigma)$ -plane to $d/\sigma < 1/2$ and $\sigma_b/\sigma < 1$. For example, in Fig. 2.7, when β is decreased below $-1/2\sqrt{6}$, the solid curve moves towards the origin, thus shrinking the region. When β is increased above $-1/2\sqrt{6}$, the solid curve

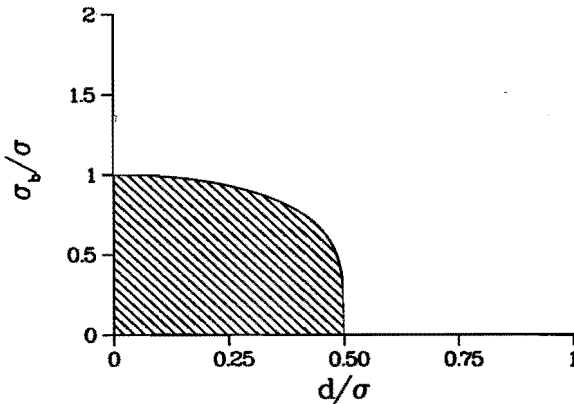


Figure 2.7: The region of reliable edge positions in the $(d/\sigma, \sigma_b/\sigma)$ plane. The solid line is given by the equation $f_3/f_1 = \beta$. The hatched region shows the region of allowed solutions for $\beta = -1/2\sqrt{6}$. The figure is symmetric around the ordinate.

moves outwards (away from the origin), thus expanding the region. At every edge location identified we ensure that (2.38) is satisfied before Eq. (2.30) is used to estimate blur. Thus, we ensure that only those edges that give reliable estimates of edge parameters are used to estimate image blur.

Figure 2.8 shows the region covered by periodic windows of a fixed scale σ , and period T , in the (d, σ_b) plane. These bounds can be varied by varying the ratio T/σ and the threshold β in Eq. (2.38). In Fig. 2.8, $T/\sigma = 2/3$ and $\beta = -1/2\sqrt{6}$. Edges that fall within the non-overlapping regions are analysed by only one window that lies nearest to it. Edges that lie in the overlapping regions could be analysed by more than one window. However, since in our implementation the edges are detected by locating local maximum in $|f_1|$, the edges lying in the overlapping regions are analysed by only one window that is located closest to it. The T/σ ratio and the threshold β have to be chosen so that the repetitive windows span the entire d -axis. Blurred edges whose parameters are outside the shaded region will not be analysed by a polynomial transform with a fixed scale. This calls for a multiscale algorithm which is discussed in Section 2.7.

At an edge location, observation \tilde{r} is corrupted by noise, i.e. $\tilde{r} = \sigma_b + \tilde{n}$, where σ_b is the parameter to be estimated, \tilde{r} is the measured quantity given

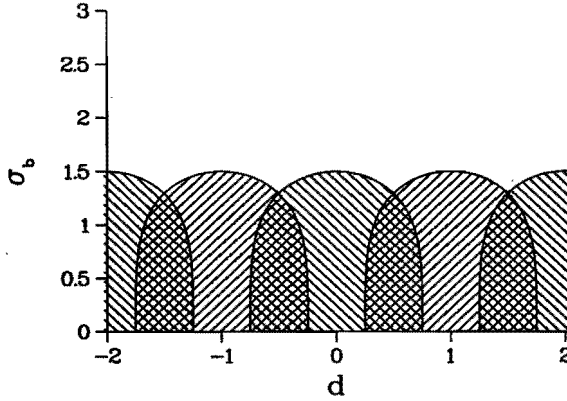


Figure 2.8: The region covered by a periodic window in the (d, σ_b) plane. Both d and σ_b are expressed in units of T . In this figure $\sigma = 3T/2$ and $\beta = -1/2\sqrt{6}$. The solid line is the boundary of the region analysed by the window located at a fixed position.

by (see Eq. (2.30)),

$$\tilde{r} = \sigma \left[\left(\frac{2\tilde{f}_2^2}{\tilde{f}_1^2} + \frac{\sqrt{6}\tilde{f}_3}{\tilde{f}_1} \right)^{-1} - 1 \right]^{1/2} \quad (2.39)$$

and \tilde{n} is the noise. To obtain an estimate of σ_b and to study its statistical properties we need to know the PDF, $p_{\tilde{r}|\sigma_b}(\tilde{r}|\sigma_b)$. As we noted in the previous section, it is not possible to obtain analytical expressions for this PDF. The computer simulations shown in Fig. 2.5d demonstrate that the expected value of \tilde{r} is lower than that of σ_b , i.e. $E[\tilde{r}] = \sigma_b - b$, indicating bias. The bias is mainly due to the squaring of f_2 in Eq. (2.30). Simulations similar to that shown in Fig. 2.5d also show that the deviation of the expected value of f_2^2 , $E[f_2^2]$, from its noise-free value is proportional to σ_n^2 .

If the bias b can be estimated, we can obtain an unbiased estimate of σ_b by adding b to $E[\tilde{r}]$. The bias b was studied as a function of the input noise variance σ_n^2 , the window spread σ , and the edge parameters ΔV , d , θ and σ_b , using simulations. Figure 2.9 shows the result of simulations. From simulations

it was found that b is approximately proportional to $(\sigma_n/f_1)^2$, i.e.,

$$b \approx k \left(\frac{\sigma_n}{f_1} \right)^2, \quad (2.40)$$

where k is a constant, independent of the edge parameters, window spread and noise variance. For $\text{SNR} > 5$ and $\sigma > 3T/2$, the constant k is approximately equal to 0.53.

An unbiased estimate of the blur parameter at an edge location i can be obtained by adding the bias b to r_i . From Eq. (2.40), it follows that

$$\hat{\sigma}_b(r_i) = r_i + k \left(\frac{\sigma_n}{f_{1,i}} \right)^2. \quad (2.41)$$

, where $\hat{\sigma}_b(r_i)$ is the unbiased estimate of σ_b at i -th edge location. At an edge location, f_1 is a measured quantity. If we know the noise variance σ_n^2 , we can obtain an unbiased estimate of the blur parameter at each edge location using the above equation. Noise variance can be either assumed to be known or can be estimated from uniform regions in the image (Chapter 3).

The variance of the estimation error, $\text{Var}[\hat{\sigma}_b(\tilde{r}) - \sigma_b]$, also increases with $(\sigma_n/f_1)^2$. The standard deviation of the estimation error as a function of $\sigma_n/|f_1|$ is plotted in the bottom panel of Fig. 2.9. From the figure we see that, for a given amount of noise and image blur, the variance of the estimate decreases with $|f_1|$. From Eq. (2.26) we note that $|f_1|$ increases with the edge height ΔV , decreases with d/σ and with σ_b/σ . Therefore, when the image contains additive noise, to obtain estimates with low variance, σ has to be chosen such that the ratio σ_b/σ is minimum. Edges with high contrast that are situated close to the centre of the analysis window give estimates with lower variance.

2.6.3 Estimation of Uniform Image Blur

In an image with space-invariant Gaussian blur we can estimate the blur parameter σ_b more reliably. The observation space R consists of N (independent) observations from N edge locations over the image. Each observation r_i is biased as well as corrupted by noise, i.e.,

$$r_i = \sigma_b - b_i + n_i, \quad i = 1, 2, \dots, N, \quad (2.42)$$

where σ_b is a constant. Based on these observations, we want to derive the estimate $\hat{\sigma}_b(R)$.

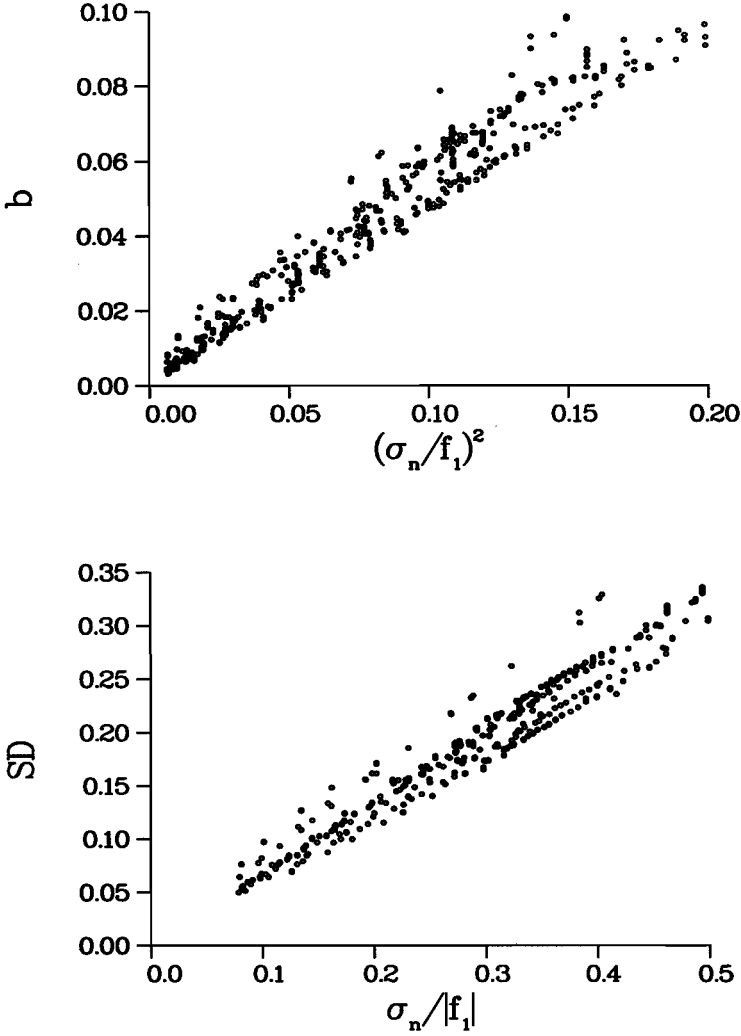


Figure 2.9: Bias and standard deviation in the estimate of edge blur obtained using simulations, for different values of the parameters: $\sigma, \sigma_b, \Delta V, d, \theta$ and σ_n . Top panel shows the bias in the estimate of edge blur. Bottom panel shows the standard deviation of the individual errors in the estimate of edge blur. Each point in the plots is the result of 32000 simulation trials.

We assume that the noise variance, σ_n^2 , is constant over the image. Thus, from Eq. 2.42 and from Eq. 2.40 we obtain

$$r_i = \sigma_b - \frac{K}{f_{1,i}^2} + n_i, \quad i = 1, 2, \dots, N, \quad (2.43)$$

where the two parameters, σ_b and K , are constant over the image. Hence, when the blur and the noise are uniform over the image, the bias in the estimate of σ_b across the edges in the image varies only with $|f_{1,i}|$, the square-root of the first-order energy at the i -th edge location. An estimate for σ_b (and K) can be obtained by minimizing the weighted error

$$\min_{\sigma_b, K} \sum_i \left[\sigma_b - \left(r_i + \frac{K}{f_{1,i}^2} \right) \right]^2 w(|f_{1,i}|), \quad (2.44)$$

where the sum is over all edge locations. Weighting is necessary because the reliability of r_i increases with $|f_{1,i}|$, as discussed in the previous subsection. We choose the weight w_i proportional to $|f_1|$. Note that, in this case, there is no need to know the noise variance to obtain the estimate $\hat{\sigma}_b(R)$.

2.7 Multiscale Blur Estimation Algorithm

In the previous sections the single-scale blur estimation algorithm was discussed. When using a single-scale algorithm, the window size σ has to be chosen *a priori*. Using an *a priori* chosen window size σ , (i.e., the single-scale algorithm) reliable estimates for blur parameter can only be made within a limited range for the parameter σ_b (see Fig. 2.7). To increase this range we can analyse the image with multiple window sizes and choose the appropriate window size *a posteriori*.

The single-scale algorithm in Fig. 2.6 can be easily extended to a multiscale algorithm as shown in Fig. 2.10. Given a blurred image it is first analysed at a scale σ_1 . All the edges whose parameters are within the bound given in Fig. 2.8 will be estimated at the first level, resulting in estimates $\hat{\sigma}_b/\sigma_1$ at those edge positions. Edges with parameters that fall outside the shaded region will not be estimated at the first level and have to be estimated at a higher level. The zero-order coefficient image at the first level is taken as the input to the second level of the analysis. Since the zero-order Hermite filter is Gaussian, the effective blur parameter at the input to the second level is $\sqrt{\sigma_b^2 + \sigma_1^2}$. The analysis at the second level takes place at scale σ_2 , with $\sigma_2 > \sigma_1$. The range of parameters of edges that can be analysed by the second level is bounded in the (d, σ_b) plane (see Fig. 2.11 for example, where $\sigma_2 = 2\sigma_1$). Those edges whose parameters are within these new bounds (excluding those that have already

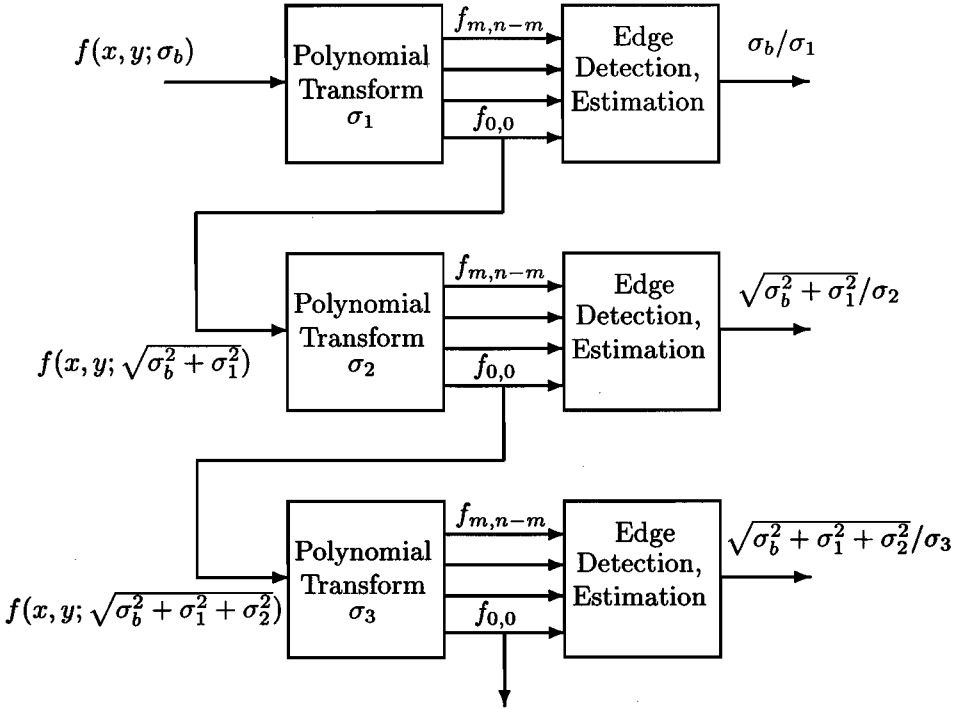


Figure 2.10: Multi-scale blur estimation algorithm

been analysed at the first level) will be estimated at the second level resulting in estimates $\sqrt{\hat{\sigma}_b^2 + \sigma_1^2}/\sigma_2$ at selected edge positions. Those edges that are not estimated either at the first or at the second level will have to be estimated at higher levels. In this way we proceed to estimate the parameters of all edges with the appropriate window size.

The successive levels in the multiscale algorithm can be chosen so that $\sigma_i = D\sigma_{i-1}$. We choose $D = 2$ in the present implementation. This can be implemented by using windows whose spread keeps doubling, or by using the well-known pyramid structure (Burt & Adelson, 1983). In the pyramid structure, when a subsampling factor of D is used, i.e. $T_i = DT_{i-1}$, the same

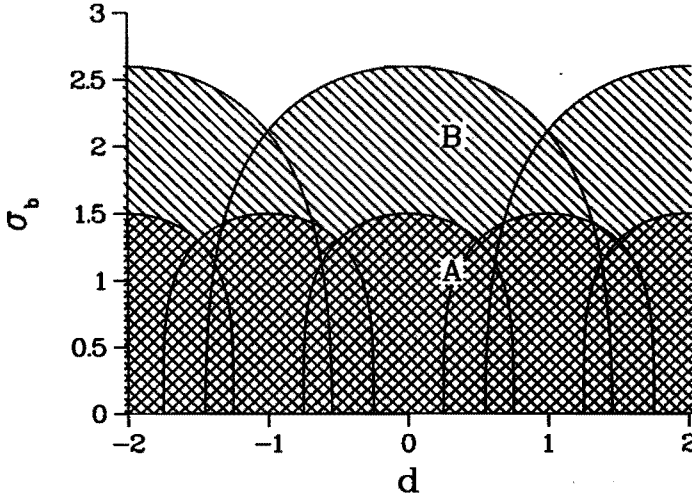


Figure 2.11: Region covered by the multi-scale periodic windows in the (d, σ_b) plane. Both d and σ_b are expressed in units of T . In this figure $\sigma_2 = 2\sigma_1$, $T_2 = 2T_1$, $\sigma = 3T/2$ and $\beta = -1/2\sqrt{6}$. The hatched region shows the region covered by the second layer of the multi-scale algorithm. The crossed region shows the region covered by the first layer (same as Fig. 2.8). The edge A with $\sigma_b = 1$ is analysed reliably at level one, whereas the edge B with $\sigma_b = 2$ has to be analysed at level two.

window function can be used at every level. The lowest σ that can be used is $\sigma = 3T/2$, due to the effects of sampling on the SNR of the polynomial coefficients (Martens, 1990c). Figure 2.11 shows the range of parameters of edges analysed at the first and the second level of a multiscale algorithm with $D = 2$. By going further to higher scales, the entire (d, σ_b) plane can be covered. Note that at the second level the effective blur parameter is $\sqrt{\sigma_b^2 + \sigma_1^2}$. Hence, in Fig. 2.11 the reliable region boundary at second level ($\sigma_2 = 2\sigma_1$) is obtained by computing Eq. (2.38) at the second level and by subtracting the contribution of σ_1 from the effective blur $\sqrt{\sigma_b^2 + \sigma_1^2}$.

Edges for analysis at each level of the multiscale algorithm are chosen so that they are estimated most reliably at that level. This can be seen from the reliable region for each level in the (d, σ_b) plane (see Fig. 2.11). The error

in the estimate of σ_b is mainly due to two sources, one due to noise in the image (see subsection 2.5.3) and the other due to independent errors in the polynomial coefficients (see subsection 2.5.2). The variance in the estimate of σ_b due to image noise decreases with $|f_1|$ and hence with increasing window size σ . However, the error in the estimate due to independent errors in the polynomial coefficients increases when σ is increased beyond σ_b (Fig. 2.4a). Using a window with σ that is small compared to σ_b also results in large errors in the estimate due to independent errors in the polynomial coefficients (see Fig. 2.4a). Considering both forms of estimation errors, the most reliable estimates of σ_b are given by a window with spread σ in the range $\sigma_b < \sigma < 2\sigma_b$. From this bound it follows that, using a window of spread σ , reliable estimates of σ_b can be made in the range $\sigma/2 < \sigma_b < \sigma$. However, because of the effects of sampling on the SNR of the polynomial coefficients (Martens, 1990c), the lowest σ that can be used is $\sigma = 3T/2$, where T is the sampling interval. The reliable region for each level in the (d, σ_b) plane is based on this bound on σ_b for that level. For example, in Fig. 2.11, edge A with parameters $\sigma_b = 1$ and $d = 1/4$ is estimated most reliably at the first level. Instead, if edge A is estimated at the second level, the estimation error will increase substantially. This is because at the second level the error due to independent errors in the coefficients will increase substantially since σ_b/σ changes from $2/3$ to $1/3$, whereas the decrease in the error due to image noise will be small. Similarly, edge B with parameters $\sigma_b = 2$ and $d = 1/4$ is estimated most reliably at the second level.

2.8 Results of Blur Estimation

In this section we present the results of applying the blur estimation algorithm to synthetic and natural images. The exercise on grey-scale images is mainly aimed at comparing the estimates of blur-spread given by the algorithm against the input blur-spreads. Figure 2.12 shows four original images used to test the algorithm. The two images at the top, called 'Mondrian'(left) and 'Text', are synthesized on the computer, while the two at the bottom, called 'Terrace' (left) and 'Wanda', are digitized images of natural scenes. The images of natural scenes were captured with a camera (which by itself introduces some blur), and were later electronically digitized. Each image shown in Fig. 2.12 contains 512 by 512 pixels, with 8 bits per pixel accuracy. The synthetic images Mondrian and Text, in which the edges are ideally sharp (i.e., step functions) are chosen to test the performance of the algorithm in ideal conditions and to compare it with that for real images of natural scenes.

Images with different amounts of blur were generated from the original images by filtering with 2-D separable binomial windows of order B . A binomial window of order B is approximately equivalent to a Gaussian window of



Figure 2.12: Original images used to test the algorithm. Images at the top are 'Mondrian'(left) and 'Text' and at the bottom are 'Terrace' (left) and 'Wanda'.

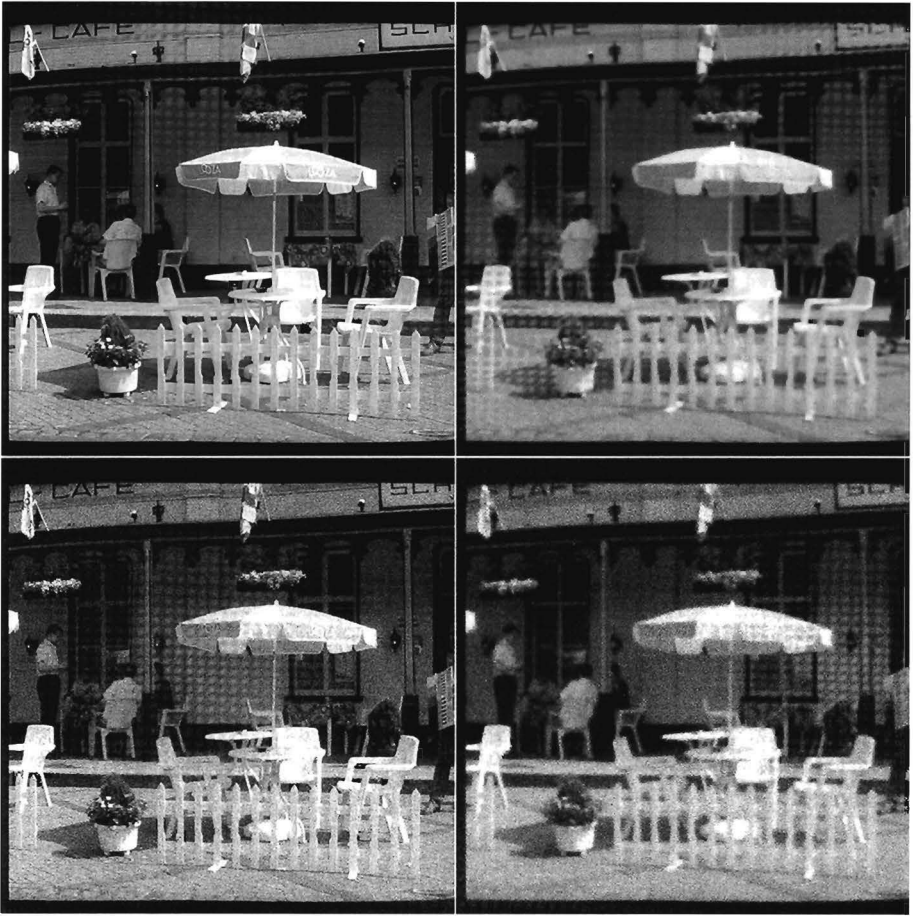


Figure 2.13: Images of Terrace scene with blur, noise, and blur as well as noise. The image on the top left is the ‘original’, and the image on the top right is obtained by blurring the original with a binomial window of order $B = 8$ (i.e., $\sigma = 2$). The image on the bottom left is obtained after adding Gaussian density noise with SD $\sigma_n = 14$ and the image on the bottom right is obtained by first blurring with a binomial window of order $B = 8$ and then adding noise of SD $\sigma_n = 14$.

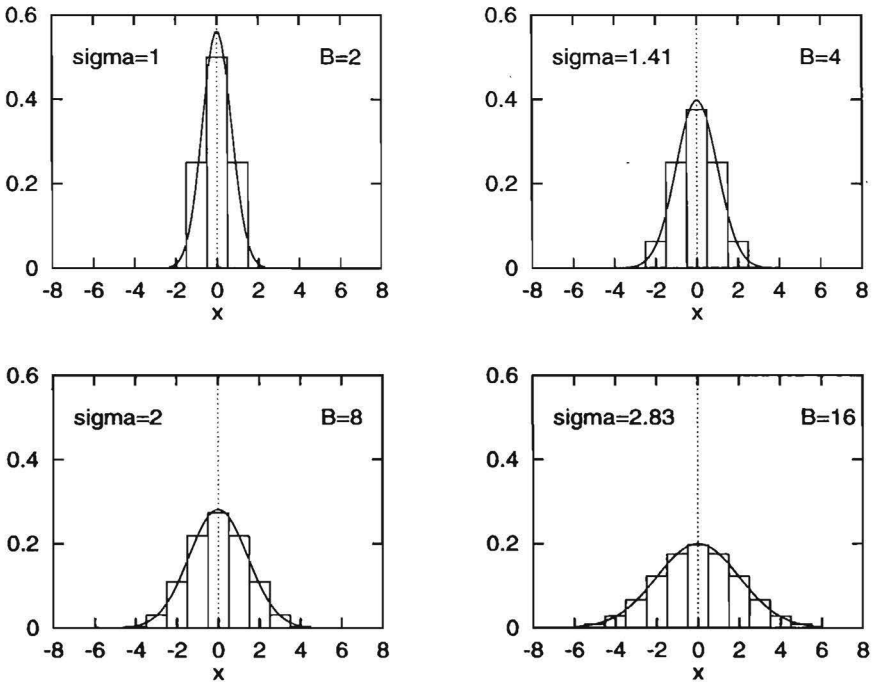


Figure 2.14: Approximation of Gaussian windows using binomial windows, for windows of order $B = 2, 4, 8$, and 16 . The coefficients of binomial windows are shown using a bar graph and the corresponding Gaussian windows, with $\sigma = \sqrt{B/2}$, are shown using drawn curves.

spread $\sigma_b = \sqrt{B/2}$ (Martens, 1990b) (see Fig. 2.14). For example, $B = 8$ is approximately equivalent to $\sigma_b = 2$. To test the behaviour of the algorithm in the presence of noise, noisy images were used. Noisy images were generated by adding zero-mean white noise with Gaussian PDF of known standard deviation, σ_n , to each pixel of the blurred image. Figure 2.13 shows examples of images with blur, noise, and blur and noise.

Figure 2.15 shows the results of edge detection, using the single-scale algorithm described in Section 2.6. The two input images are Mondrian and Terrace blurred with a binomial window of order $B = 2$ which corresponds approximately to $\sigma_b = 1.0$. The top row shows the locations of detected edges. While detecting the edges, the threshold on $|f_1|$ of edge chains and that on chain length is selected so that a certain percentage (e.g. 5 %) of the image area is judged to be edge regions. From these edges, 1-D edges are detected by

thresholding the ratio E_{2D}/E_1 . A fixed value of 0.01 is used as threshold on E_{2D}/E_1 for all the results reported here. The bottom row shows the locations of 1-D edges used for blur estimation. Edge locations for blur estimation are chosen from the locations of 1-D edges using the threshold β in Eq. (2.38). Results reported here are generated using a fixed value for $\beta = -1/2\sqrt{6}$. The figure shows that the 1-D edge detector correctly rejects higher-dimensional patterns such as corners as well as 0-D patterns (i.e., uniform regions). Using the edges chosen in Fig. 2.15, the blur estimation algorithm gives an estimate of $\sigma_b = 1.03$ and $\sigma_b = 1.16$ with SDs 0.18 and 0.39 for Mondrian and Terrace respectively (approximate expected value $\sigma_b = 1$ for both images). The higher value of the blur parameter, σ_b , in the case of the natural scene is mainly because the original image of that scene is blurred in the process of camera recording and digitization, whereas the original synthetic image has ideally sharp transitions.

Figure 2.16 shows the results of edge detection using a single-scale algorithm, in the presence of image noise. The two input images are Mondrian and Terrace blurred with a binomial window of order $B = 2$ corresponding to $\sigma_b = 1.0$, and contain added noise with $\sigma_n = 10$. Using the edges chosen in Fig. 2.16, the blur estimation algorithm gives an estimate of $\sigma_b = 0.97$ and $\sigma_b = 1.17$ with SDs 0.30 and 0.39 for Mondrian and Terrace respectively. In Fig. 2.15 and Fig. 2.16, the low contrast edges are not detected because we have chosen the threshold on $|f_1|$ of edge chains such that the probability of false alarm is low.

Figure 2.17 shows the result of the single-scale blur estimation algorithm for images with different amounts of blur. These results also show the effect of parameters σ and β on the estimate. The dashed line (\square) shows the estimates given by an algorithm with $\sigma = 2$ and $\beta = -0.04$. Since β is high, edges with large σ_b/σ ratio were also considered in the estimate. The drawn line (\diamond) shows the estimates obtained when $\beta = -1/2\sqrt{6} = -0.204$, but with σ chosen such that it is always higher than the expected σ_b , i.e., such that the σ_b/σ ratio for most edges in the images is below 1. At low input σ_b , estimates obtained using both settings are almost equal for all scenes, although some differences are present for the Wanda scene. It is clear from the figure that the original images of both natural scenes are blurred. The effects of using a large window are clear from the results for the Text scene, which has many line-like structures. When the window is large, it encloses more than one edge within it (i.e., line in the case of Text) and hence the estimate based on a single edge within a window fails. In the Text scene this happens when σ exceeds the width of the character stroke. However, when the Text images are analysed using lower σ but with higher β (to obtain a sufficient number of edge locations on which to base the estimate), the estimate is closer to the expected estimate (\square). The expected estimate is shown by the drawn straight line. The lengths

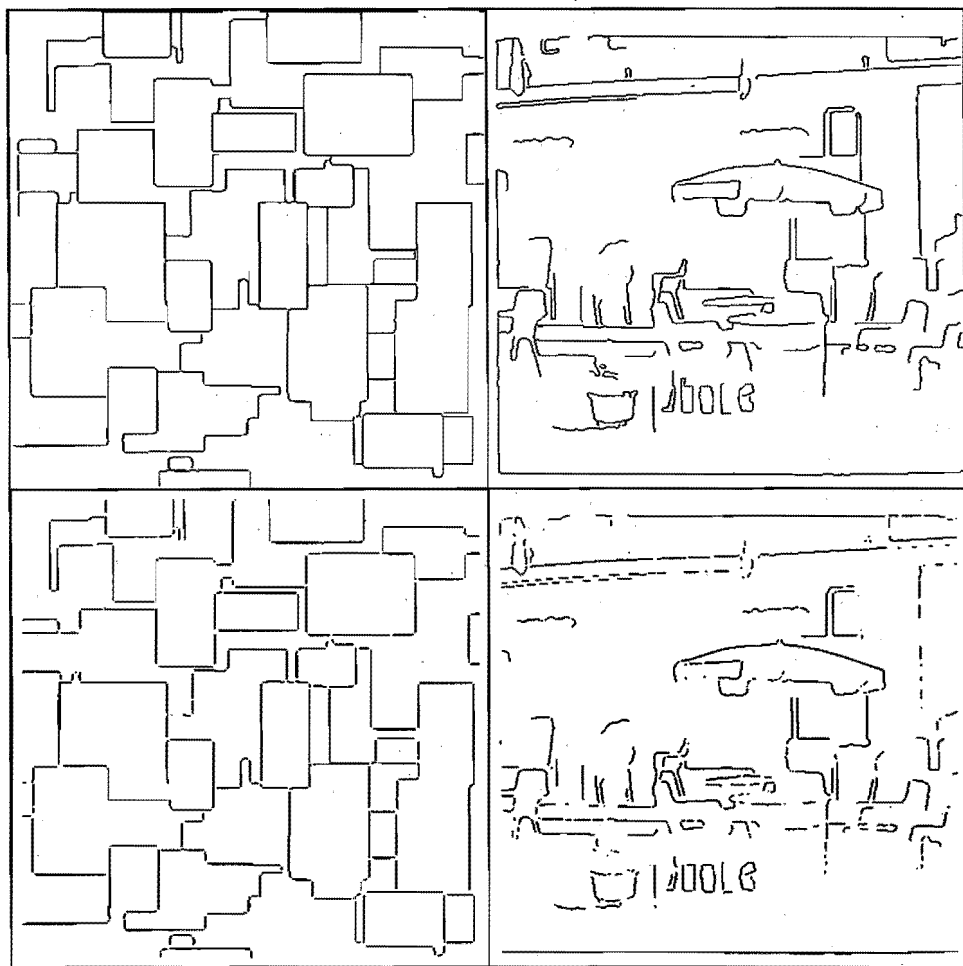


Figure 2.15: Result of edge detection. The input images are Mondrian and Terrace blurred with $\sigma_b = 1$. The top row shows the locations of detected edges and the bottom row shows the locations of 1-D edges used for blur estimation. Notice that the corners, which are 2-D features, have been correctly rejected in the images at the bottom. These results are obtained at scale $\sigma = 2.0$.

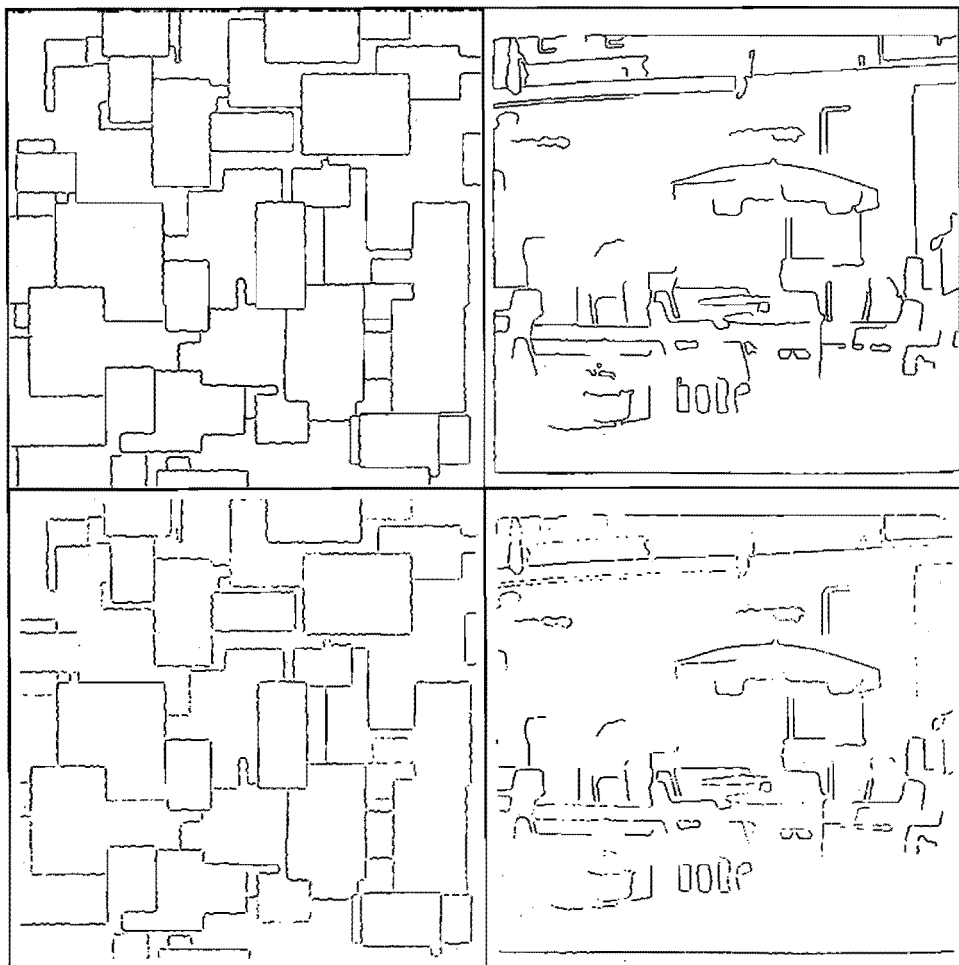


Figure 2.16: Result of edge detection in images containing noise. The input images are Mondrian and Terrace blurred with $\sigma_b = 1$ containing noise with $\sigma_n = 10$. The top row shows the locations of detected edges and the bottom row shows the locations of 1-D edges used for blur estimation. These results are obtained at scale $\sigma = 2.0$.

of the error bars in all figures are equal to twice the standard error of the mean (SEM). However, in most cases, the SEM is small and the error bars in the plots are barely visible.

Figure 2.18 shows the result of the blur estimation algorithm for images with different amounts of blur and noise. All estimates are obtained using a single polynomial transform of window size $\sigma = 2$ and $\beta = -0.204$. Points joined by a line correspond to images that have the same amount of blur, but contain different amounts of noise. In each plot, the point in the lower left corner corresponds to the original image. The parallel curves demonstrate that the algorithm gives reliable estimates of the blur parameter in the presence of image noise. In the case of Mondrian, consistent estimates of σ_b are obtained in the presence of noise, for all levels of blur except for the case $\sigma_b = 0$ (the lower most curve). The algorithm overestimates the blur of a noisy image with zero input blur. The $\sigma_b = 0$ situation does not occur in real images, since real images are blurred due to the imaging process. In the case of Text also the estimates are consistent in the presence of noise but are slightly lower than those for Mondrian. In the case of Terrace, consistent estimates of σ_b are obtained in the entire range tested. In the case of Wanda we obtain consistent estimates when the variance of the input noise is small, but when the input noise is high the estimates become less reliable. This is because the image Wanda does not contain high-contrast edges. When an image does not contain high-contrast edges, reliability of the blur estimate decreases rapidly as the input noise increases because the SNR at the edges decreases drastically.

Figure 2.19 shows how the estimate of σ_b varies as a function of the scale of the polynomial transform, σ . The input images are Terrace with $\sigma_b = 0.7$, $\sigma_n = 10$ (lower curve) and Mondrian with $\sigma_b = 1.4$, $\sigma_n = 5$ (upper curve). We have used $\sigma > 3T/2$ to keep the effects of sampling on the polynomial coefficients to a minimum. There is a small variation in the estimated value of σ_b^2 as the scale varies, but it is within the SD of the estimate. For both images the SD of the estimate does not vary much with the scale in the range of σ tested.

Figure 2.20 shows the result of the multiscale algorithm on Terrace as σ_n increases. These results are obtained using a three-level algorithm with $\sigma_1 = 2$, $\sigma_2 = 4$, and $\sigma_3 = 8$, which has a pyramid structure, with $\beta = -1/2\sqrt{6}$. The aim here is to merely demonstrate the use of the proposed multi-scale blur estimation algorithm using a typical example. Therefore, the results of the multi-scale algorithm shown here are intended only as a demonstration of the multi-scale algorithm to estimate blur in images of natural scenes.

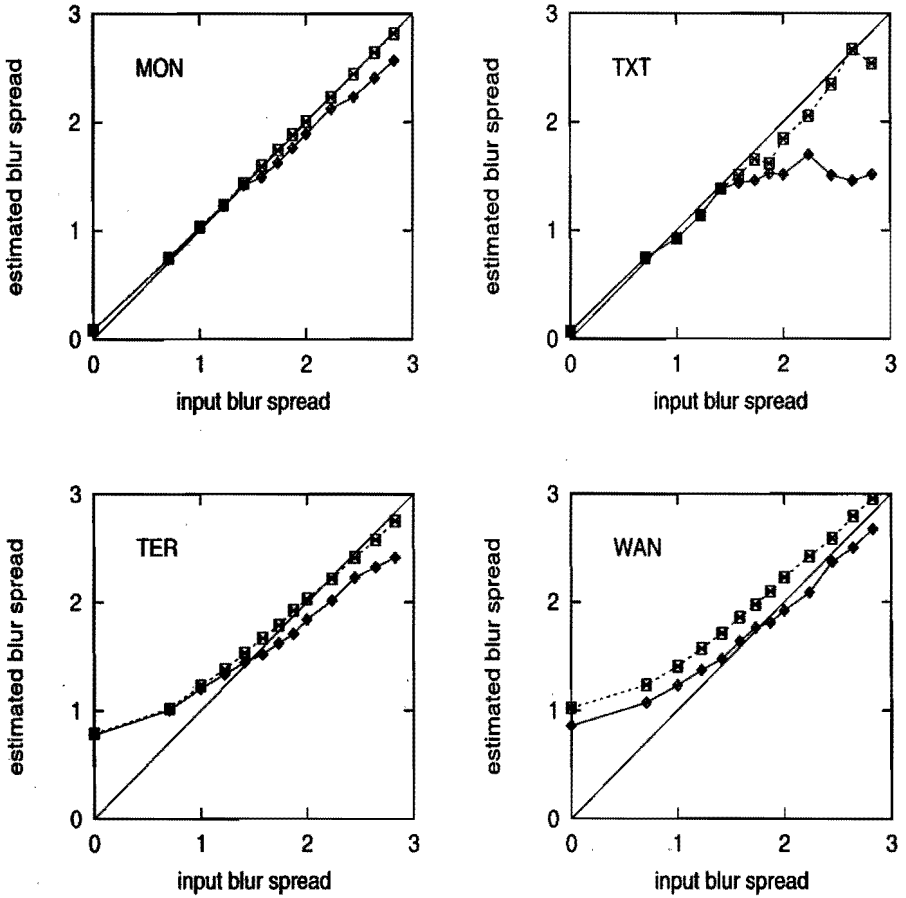


Figure 2.17: Results of blur estimation using a single-scale algorithm on images with different amounts of blur. The inputs are original images of the Mondrian, Text, Terrace and Wanda scenes, blurred with binomial filters of order $B = 0$ (original), to 16 (i.e., $\sigma_b \approx 0$, to 2.83). The dashed line (\square) shows the results obtained with $\sigma = 2.0$ and $\beta = -0.04$. The drawn line (\diamond) shows the results obtained with $\beta = -0.204$ but with different values of σ . For each scene, the images with input $\sigma_b < 1.5$ (first 5 images) were analysed at $\sigma = 2.0$, those with $1.5 < \sigma_b \leq 2$ (next 4 images) were analysed at $\sigma = 2.83$ and those with $\sigma_b > 2$ (last 4 images) were analysed at $\sigma = 4.0$

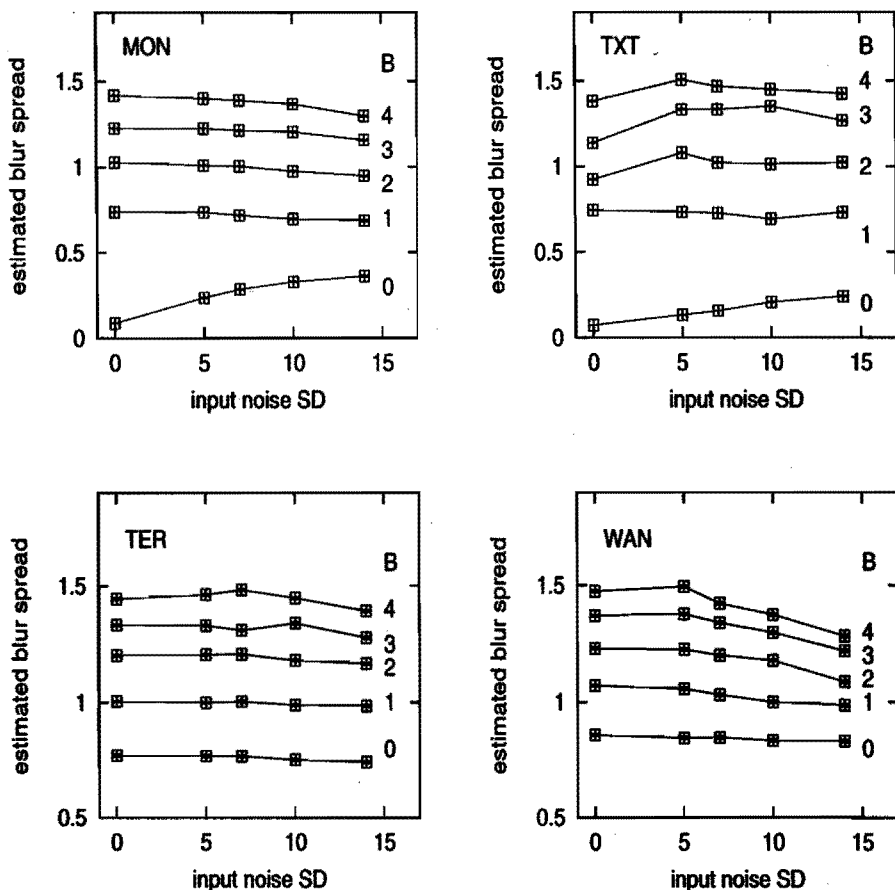


Figure 2.18: Results of blur estimation on images with different amounts of blur and noise. The input images are: Mondrian, Text, Terrace and Wanda, blurred with binomial filters of order $B = 0$ (original), 1, 2, 3, and 4 (i.e., $\sigma_b \approx 0, 0.71, 1, 1.23$, and 1.41) and containing noise with $\sigma_n = 0, 5, 7, 10$ and 14 . These results were obtained using a single-scale algorithm with $\sigma = 2.0$ and $\beta = -0.204$.

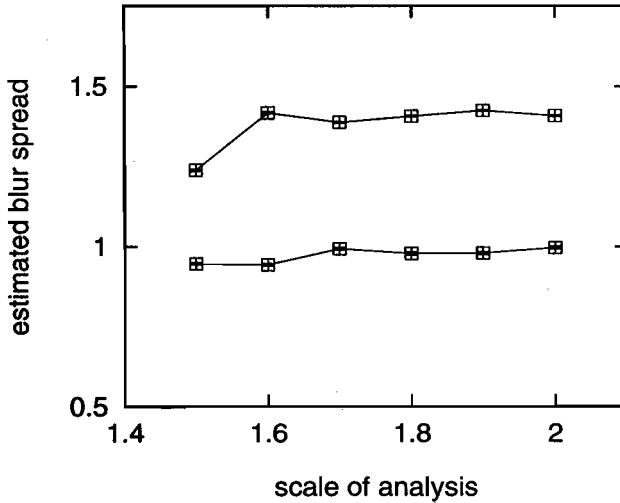


Figure 2.19: Blur estimate as a function of scale, σ . The input images are Terrace with $\sigma_b = 0.7, \sigma_n = 10$ (lower curve) and Mondrian with $\sigma_b = 1.4, \sigma_n = 5$ (upper curve).

2.9 Experiment: Perceived Effect of Blur on Natural Images

In this section, we describe a psychophysical experiment that is aimed at studying the effect of blur on the perceived unsharpness of the natural images. The results of this experiment will be used to test the estimates of a measure for perceived blur obtained using the computational algorithm.

Stimuli: In this experiment, black and white still images are used as stimuli. The parameter that is varied is the spread of the blurring kernel σ_b . Two natural scenes were used: Terrace and Wanda (see Fig 2.12). Stimuli were generated by blurring the ‘original’ images of the scenes with binomial filters of given order B . A binomial filter of order B is approximately equal to a Gaussian window of parameter $\sigma_b \approx \sqrt{B/2}$ (see Fig. 2.14). Six values of B were used: 1, 2, 3, 4, 6 and 8. The original images of the scenes were also used in the experiment. Thus there were seven stimuli per scene. The images had a size of 512x512 pixels with 8 bits per pixel accuracy, but in the experiment only the central region of 450x450 pixels (0.24 m by 0.24 m or 9.8° by 9.8°) was displayed. The viewing distance was six times the image height. The images were displayed on a BARCO-CCID-7351B high-resolution non-interlace monitor using a VME-bus based system (with MEN-A301 graphics

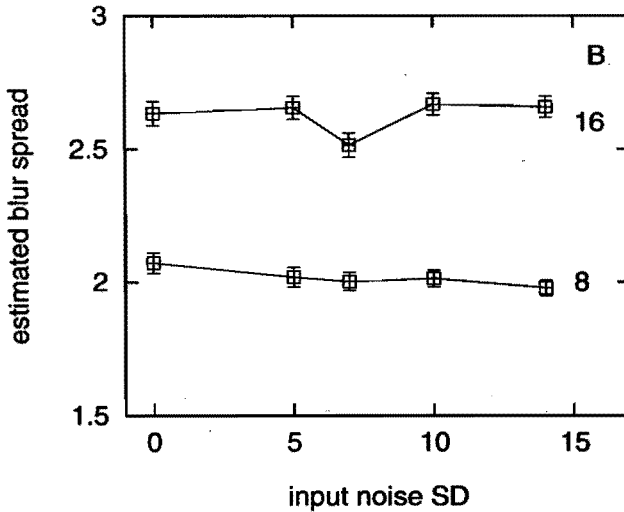


Figure 2.20: Result of the multi-scale algorithm on noisy, blurred images. The input images are Terrace with $B = 8$, (i.e., $\sigma_b = 2.0$, lower curve) and $B = 16$ (i.e., $\sigma_b = 2.83$, upper curve) containing noise with $\sigma_n = 0, 5, 7, 10$ and 14 . These results were obtained using a three-scale algorithm with $\sigma_1 = 2$, $\sigma_2 = 4$, and $\sigma_3 = 8$. The lengths of the error bars are equal to twice the standard error of mean (SEM).

board). The display characteristics were calibrated to have a $\gamma = 2.5$ (Poynton, 1993). The luminance of the monitor for an all-white screen was 60 cd/m^2 .

Method: The viewing conditions satisfied the CCIR recommendation 500 (CCIR, 1986). Images were displayed for a fixed duration of 5 sec, with at least a 2 sec adaptation time between two presentations, during which an adaptation field of 10.7 cd/m^2 was displayed. The luminance of the adaptation field was approximately equal to the average of the mean luminance of the two scenes. The subjects were instructed to judge the unsharpness of the images using a 10-point numerical category scale ranging from one to ten. Subjects took part first in a session with images of the Terrace scene and then with those of the Wanda scene. Each stimulus was repeated four times. Thus there were $7 \times 4 = 28$ presentations per session. The order of presentation within each session was random. Before the start of the each session, subjects took part in a trial session with 8 stimuli covering the entire range, in order to adjust the sensitivity of their scale.

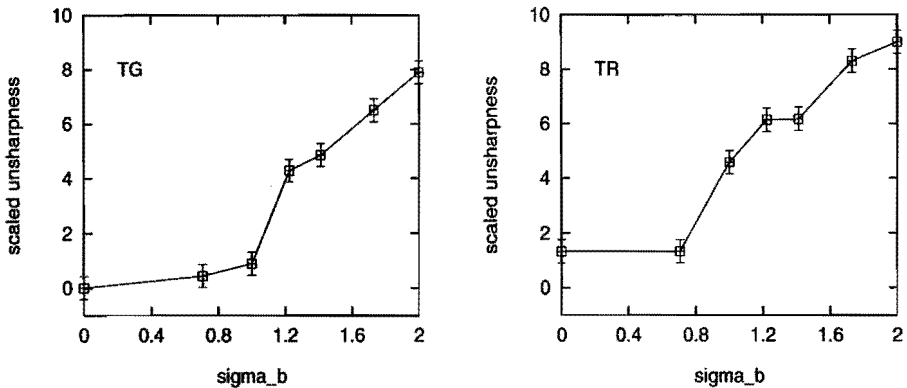


Figure 2.21: Results of scaling unsharpness plotted against input blur-spread σ_b , for images of Terrace scene with blur, for subject TG and TR. The lengths of the error bars are equal to twice the standard error of mean (SEM).

Subjects: Eight subjects between 22 and 38 years of age took part in the experiment. All were familiar with the use of numerical a category scaling procedure and the concept of unsharpness in images. All had normal or corrected-to-normal visual acuity between 1.5 and 2, measured on a Landolt chart.

Results: The numerical category data from the experiment were transformed into an interval scale on a psychologically linear scale using Thurstone's law of categorical judgement. The class I model involving replications over trials within one subject with constraints of condition D was applied (Torgerson, 1958). In condition D the number of parameters are limited by assuming that the dispersions of the category boundaries as well as the stimulus locations are constant, and that the correlation between their momentary positions is also constant.¹

Figure 2.21 shows the results for two subjects, for the Terrace scene. The figure shows that there are no significant differences between the trends in

¹However, when a stimulus is judged to be in a boundary category in all the repetitions, the condition D constraints lead to a misleading position for that stimulus. This is because such stimuli do not satisfy the assumptions of condition D. To obtain meaningful positions in the psychological continuum for such stimuli as well, a correction was applied before applying the Thurstone's model (de Ridder, 1994). Let us assume that there are n categories (labelled 1 to n), and m repetitions per stimulus. Then the correction applied was as follows: (1) two additional categories were added: 0 and $n + 1$, one at each boundary, (2) the scores of the stimuli which had full entry m in the boundary category (either 1 or n) were distributed symmetrically among its two neighbouring categories, as follows, if category n had m entries, it was corrected by placing 1 in category $n - 1$, replacing $m - 2$ in category n , and placing 1 in category $n + 1$. Similar correction was applied if m entries occurred in category 1.

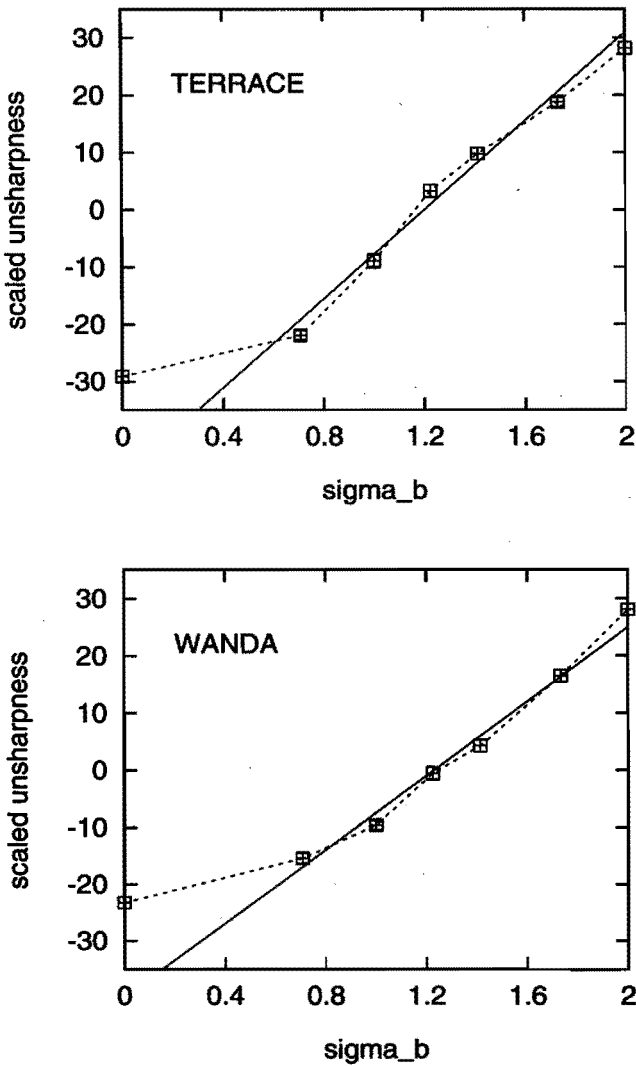


Figure 2.22: Average results of scaling unsharpness plotted against input blur-spread σ_b , for images of the Terrace and Wanda scenes with blur. Each point is the result of averaging over 8 subjects.

the data between subjects. This was also the case with the data for other subjects. Hence, the Thurstone corrected data were averaged over subjects. Before averaging over subjects, the data for each subject were normalized using the z-score transform (Hays, 1988). Finally, the z-scores averaged over the subjects were normalized so that the resulting average standard error of mean

(SEM) was equal to one. The results averaged over all subjects is shown in Fig. 2.22. These results show that the subjects are able to integrate the blur over the image very consistently. Notice that the original images of both scenes, (the lower-most point in each plot) are deviating from the drawn line.

2.10 Estimation of Blur-index

It has been shown that, for small amounts of blur, the spread of the blurring kernel σ_b is a good correlate of the perceived blur (or unsharpness) of images (Westerink & Roufs, 1989; Westerink, 1991). Nijenhuis (1993) derived an empirical relation for an objective measure for blur from the spread of the blurring kernel, based on the sensitivity of the visual system to changes in the σ_b (See Introduction to this chapter). The relation is given by

$$S_b = 1 - \frac{1}{[1 + (\sigma_{bi}/\sigma_{b0})^2]^{1/4}} \quad (2.45)$$

where S_b is the objective measure for blur ($0 \leq S_b \leq 1$), σ_{bi} is the spread of the image blurring kernel and σ_{b0} is the intrinsic blur in the early-visual pathway. In this thesis, we refer to the objective measure S_b as *blur-index*. Equal intervals in S_b correspond to approximately equal perceived differences in unsharpness. Hence, S_b is a psychometric quantity (Hunt, 1978). In Figure 2.23, blur-index S_b is plotted as a function of σ_{bi}/σ_{b0} . At very low values, $\sigma_{bi}/\sigma_{b0} < 0.25$, S_b is very low and insensitive to changes in σ_{bi}/σ_{b0} . In the range $0.5 < \sigma_{bi}/\sigma_{b0} < 2$, the relation increases rapidly and almost linearly. At higher values of σ_{bi}/σ_{b0} , S_b saturates. Using this relation, the objective measure for blur of an image can be estimated from the estimate of the spread of the blurring kernel from the image. In this section we present an algorithm to estimate the blur-index S_b of an image from the image itself.

Fig. 2.24 shows the schematic diagram of the algorithm to estimate blur-index S_b . The luminance image is obtained from the grey-scale image using the Gamma-characteristics of the monitor (Poynton, 1993),

$$L = \begin{cases} L_{max} (g/g_{max})^\gamma, & L_{max} (g/g_{max})^\gamma > L_{min} \\ L_{min}, & L_{max} (g/g_{max})^\gamma \leq L_{min} \end{cases} \quad (2.46)$$

where L is the luminance, L_{max} is the maximum luminance of the monitor, L_{min} is the minimum luminance of the monitor, g is the grey-value, g_{max} is the maximum grey-value, and γ is the exponent. In our case, $g_{max} = 255$, $L_{max} = 60 \text{ cd/m}^2$, $L_{min} = 0.2 \text{ cd/m}^2$, and $\gamma = 2.5$. The optical blur of the visual system is approximated by blurring the luminance image with an approximately Gaussian window. We use a Gaussian with spread $\sigma_{b0} = 0.65 \text{ min arc}$. The

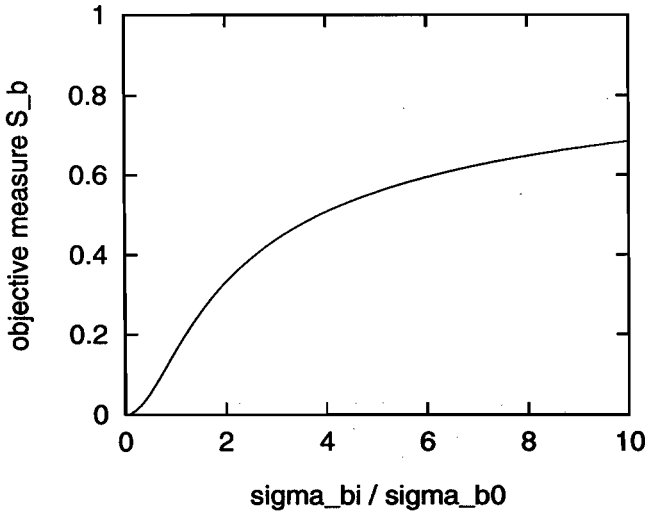


Figure 2.23: An objective measure for blur S_b as a function of the ratio of spread of the blurring kernel to the intrinsic blur σ_b/σ_{b0} (Nijenhuis, 1993).

resulting luminance image is transformed into an approximate psychometric brightness image using a pointwise transformation. This transformation reflects the non-linear compressive transformation that takes place in the early stage of the visual system, expressed by the Weber-Fechner or the Stevens relation. This non-linear compressive transformation in the visual system is thought to take place prior to the spatial filtering operations that compute relevant image information such as edge location (Morgan *et al.*, 1984). We use the CIE 1976 relation for psychometric lightness as the compressive transformation (Hunt, 1978),

$$B = \begin{cases} 116 \left(\frac{L}{L_{max}} \right)^{1/3} - 16, & \frac{L}{L_{max}} > 0.008856 \\ 903.3 \left(\frac{L}{L_{max}} \right), & \frac{L}{L_{max}} \leq 0.008856 \end{cases} \quad (2.47)$$

where $0 \leq B \leq 100$. Although, strictly speaking the CIE relation given above expresses psychometric lightness and hence is valid only for secondary light sources (reflecting bodies), we have adapted it here because it is one of the few standardized relations available to express the non-linear transformation. In our application, the exact nature and the exponent of the compressive transformation is found to be not very critical. This is another reason why we use the CIE relation, which has an exponent of $1/3$, although other exponents as high as 0.5 are also possible (Roufs, Blommaert & de Ridder, 1991). Noise is added to the psychometric brightness image obtained after the compressive

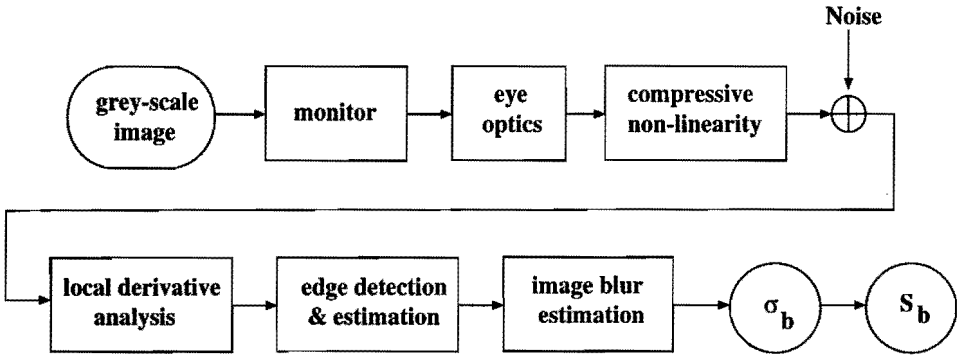


Figure 2.24: Blur-index estimation algorithm

transformation. This represents the internal noise in the early visual system (See Chapter 3). We add stationary, Gaussian distributed noise, with white density spectrum with standard deviation 3.0 units (i.e., $SD = 3\%$) (See Chapter 3). The resulting image is used as the input to the blur estimator. The later steps are the same as those in the blur estimation algorithm, described in Section 2.6.

2.10.1 Results of Blur-index Estimation

In this section, the results of the psychophysical experiment will be correlated with the blur-index \hat{S}_b estimated using the algorithm. The \hat{S}_b for each stimulus was estimated by applying the algorithm shown in Figure 2.24. The blur-index S_b was computed from the estimate of $\hat{\sigma}_b$ using the equation, $S_b = 1 - \sqrt{\sigma_{b0}/\hat{\sigma}_b}$, where $\hat{\sigma}_b$ is the estimate obtained using the algorithm (Figure 2.24) and σ_{b0} is the intrinsic blur of the visual system. This equation is obtained from Eq. (2.45), by noting that $\hat{\sigma}_b$ estimated in Figure 2.24 is the combined result of image blur σ_{bi} and the optical blur of the visual system σ_{b0} , i.e., $\sigma_b^2 = \sigma_{bi}^2 + \sigma_{b0}^2$. We use $\sigma_{b0} = 0.65$ min arc. The results of correlating the estimated \hat{S}_b with the scaled unsharpness averaged over all subjects are given in Figure 2.25. Note that the estimated blur-index of the original image of the Wanda scene (the lowermost point) is higher than that of the Terrace scene, indicating that the latter is sharper. The coefficient of correlation between the estimated blur-index and the measured unsharpness are 0.992 and 0.985 for the Terrace and Wanda scenes respectively. These results show that the blur-index S_b obtained using the algorithm is a good correlate of the perceived unsharpness.

Most objective measures for perceived sharpness found in the literature

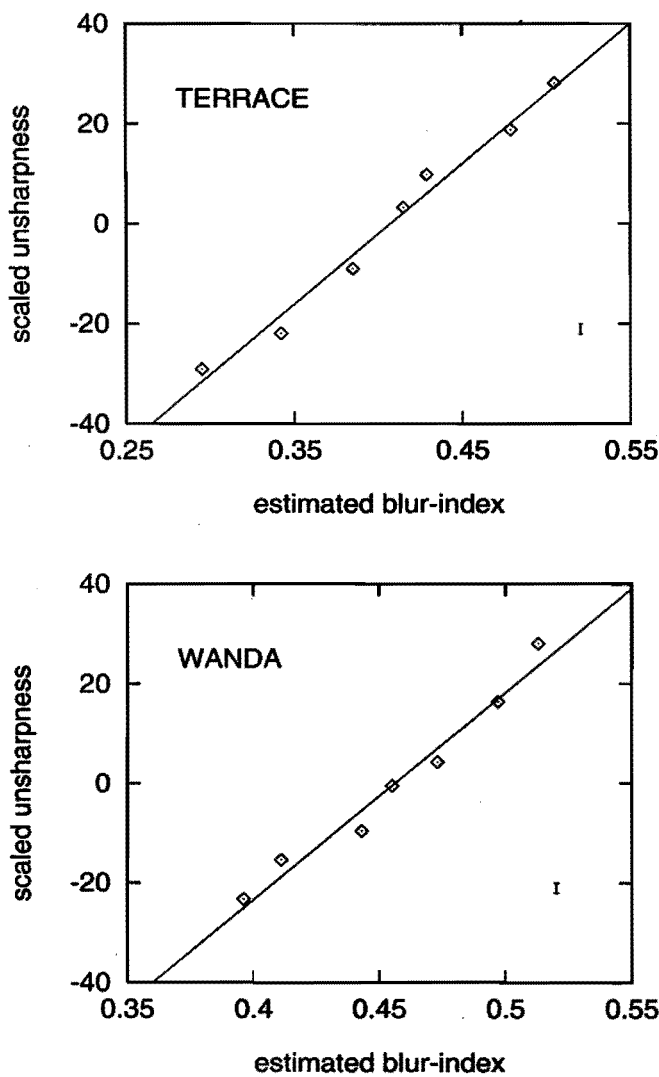


Figure 2.25: Perceived unsharpness Vs. estimated blur-index \hat{S}_b , for the Terrace and Wanda scenes. Notice the different scales on the horizontal axes.

are based on the modulation transfer function (MTF) of the display and the modulation threshold function of the eye (Carlson & Cohen; Barten, 1990). They aim at evaluating the sharpness of images produced by displays and hence are often called display-sharpness measures. The MTF is a spatial-frequency domain equivalent of the spread of the blurring kernel σ_b . The MTF-based objective measures for sharpness are thus the spatial-frequency domain methods for deriving S_b from σ_b . The main difference between the MTF-based methods and the method presented here is that, the method presented here estimates the sharpness of an image by first estimating the blur-spread σ_b from the image itself and then computing S_b from the estimated σ_b , whereas the MTF-based methods compute sharpness from the specified MTF.

2.11 Conclusions

In this chapter, a new method for estimating the parameters of a blurred edge has been presented. The parameter estimation method presented is based on a local image decomposition technique called a polynomial transform. We have used the polynomial transform for a Gaussian window. By studying the behaviour of the estimates in the presence of noise, it is shown that the method gives reliable estimates of edge parameters.

An algorithm to estimate the spread of the Gaussian image blurring kernel has been presented. The algorithm has two stages, a detection stage where locally 1-D edges are detected, followed by an estimation stage. The blur estimation algorithm can be realized as a single-scale algorithm or as a multiscale algorithm. In the case of uniform image blur, the spread of the blurring kernel is estimated by statistically combining the estimates of the blur parameter at different edge locations over the image. It is shown that the algorithm gives reliable estimates of the spread of the Gaussian blurring kernel at low SNRs by applying it to natural as well as synthetic images with different amounts of blur and noise.

An algorithm to estimate an objective measure for perceived blur has been presented. Using subjective experiments, it is shown that the measure for blur estimated using the algorithm correlates well with the perceived unsharpness in images.

Chapter 3

Estimation of Noise Variance and an Objective Measure for Perceived Noise Using Local Energy

Abstract

An algorithm for estimating the variance of additive white noise in images is presented. It is based on the first-order local energy in uniform regions of the image. The performance of the new method is compared with a recently developed method based on the overall local energy (Martens, 1994), and it is shown that the method presented here provides equally reliable estimates at approximately the same computational complexity. A new algorithm is also presented for estimating the parameters of nonwhite noise in images. It is based on comparing the first-order local energies for two windows of different size, i.e. two scales.

For image quality applications, it is necessary to have a measure that correlates well with the perceived noise in images. It must also be possible to estimate such a measure from the noisy image using computational algorithms. In an effort to achieve this, the effect of noise on images is studied using psychophysical experiments. It is shown that the perceived noise (or noisiness) of an image is independent of the probability density function (PDF) of the noise. It is also shown that the local luminance of uniform regions in an image does not influence the noisiness significantly. Subjective experiments show that the two important parameters that influence the noisiness are the noise standard deviation (SD) and the noise correlation length (CL). Based on these findings about the noisiness in the uniform regions, an objective measure for noisiness of an image, called noise-index, is proposed. The noise-index is based on the noise SD in the case of white noise, and in the case of nonwhite noise it is based on the SD of the white noise by filtering of which the nonwhite noise has been obtained. Both of these SDs can be estimated from the image using the algorithms presented in this

chapter. Using the results of the psychophysical experiments it is shown that the proposed measure for noisiness correlates well with the measured noisiness.

3.1 Introduction

Noise in an image affects the quality of the image. Noise not only decreases the perceptual quality of the images but can also degrade the performance of the task for which the images are intended (van Overveld, 1994). For example, a diagnosis based on X-ray or computed tomography (CT) images contaminated by noise is likely to be less reliable because of the loss of details caused by the noise. Many applications require a quantitative description of noise and a method to measure it. Image restoration and segmentation (Demoment, 1989; Rosenfeld & Kak, 1982) are typical examples. Image quality prediction is an emerging application that requires a perceptually relevant quantitative description of noise. By identifying the subjective dimensions of image quality (Marmolin & Nyberg, 1975) and by estimating the perceptually relevant physical parameters of degradations (such as, the spread of the blurring kernel (Westerink & Roufs, 1989) or the noise variance) the image quality can be predicted (Nakayama, Kurosu, Honjyo & Nishimoto, 1980).

In many practical imaging situations, noise can be modeled approximately as an additive Gaussian random process which is independent of the image (Andrews & Hunt, 1977), i.e.,

$$f(x, y) = f'(x, y) + \tilde{n}(x, y), \quad (3.1)$$

where $f'(x, y)$ is the uncorrupted image and $\tilde{n}(x, y)$ is an image-independent, stationary, stochastic process. The aim is to estimate the parameters of $\tilde{n}(x, y)$, given the image $f(x, y)$.

Many researchers have developed methods to estimate the variance of white noise in images (Besl & Jain, 1988; Lee, 1981; Meer., Jolin & Rosenfeld, 1990; Canny, 1986; Bracho & Sanderson, 1985; Voorhees & Poggio, 1987; Lee & Hoppel, 1989). All methods are based on the observation that in uniform regions of the image the variations are mainly due to noise. Thus, the noise variance is estimated by computing a local measure in the uniform regions of the image and deriving an estimate of the image noise variance from this measure. Although the methods are based on the uniform regions, they do not explicitly detect uniform regions prior to estimating noise. In fact, detection of uniform regions in noisy images itself needs an estimate of noise variance to set thresholds on uniformity (Besl & Jain, 1988). The need to detect uniform regions is circumvented by making use of *a priori* knowledge about the

probability of occurrence of locally uniform regions for the class of images to which the method is applied. For example, for most natural images (excluding texture) it can be assumed that a large percentage of the image area is locally uniform. Using this assumption, the noise variance is estimated by: (i) computing a local measure at all positions over the entire image, and (ii) using only those values of the local measure that are most likely to have arisen from the uniform regions.

In the literature, there are mainly two classes of noise estimation methods, one based on overall energy measures and the other based on gradient-amplitude (first-order derivative) measures. Examples of methods based on the overall energy are found in Besl & Jain (1988) and Lee (1981). Besl & Jain (1988) use the root-mean-square error of the local planer fit as their measure. The measure is discarded if the local gradient amplitude is higher than a threshold, where the threshold is based on the percentage of image area assumed to be uniform. The image noise variance is computed from the mean of the root-mean-square-error measure over all the undiscarded positions in the image. Other methods estimate the image noise variance from the average of the smallest local variances measured in a set of image blocks, implicitly assuming that such measurements are most likely to have arisen from a uniform region. For example in Lee (1981), the average of the 5 smallest variances, for 7x7 blocks, is used. Sophisticated and robust versions of such approaches can be found in Meer. *et al.* (1990) and Lee & Hoppel (1989).

Examples of the estimates based on gradient-amplitude measures are found in Canny (1986), Bracho & Sanderson (1985) and Voorhees & Poggio (1987). In these methods, the local gradient amplitude is used as a measure. Canny (1986) used the output of his edge detector as the local measure, which is approximately equal to the local gradient. Noise was estimated from the lowest 80-th percentile of the amplitude histogram of the edge detector responses. Bracho and Sanderson (Bracho & Sanderson, 1985) compute the noise variance by finding the peak in the histogram of the image gradient amplitude, which is Rayleigh distributed in uniform regions of the image, when the input noise has a Gaussian distribution. Extension of their method to textured images can be found in Voorhees & Poggio (1987), where a Rayleigh distribution is fitted only to the rising portion of the histogram.

For most of the methods mentioned above the accuracy of the estimate depends on the detailed way in which the algorithm has been implemented. For histogram-based methods, the range and the number of bins used in the histogram and the form of the smoothing function have to be chosen. The reliability of the estimate depends considerably on these choices, especially for small noise variances. Recently, a comparative study of many of the above-mentioned noise-variance-estimation methods was reported (Olsen, 1993). It

was concluded that the most reliable estimates of noise variance result from the methods based on the average local variance, such as the one in Besl & Jain (1988).

So far only two kinds of measures have been used for noise estimation, the gradient amplitude and the overall energy. Recently it has been shown that other local measures that are derived by quadratic filtering of the image could also be used for noise variance estimation (Martens, 1994). It has been shown using simulations that an algorithm based on the overall local energy provides reliable estimates of the noise variance (Martens, 1994).

When an image is distorted by additive noise, the uniform regions in the image undergo maximum visually noticeable change. The visibility of noise is reduced near steep transients in the image such as edges and lines, due to masking (Girod, 1989; Girod, 1993). Hence, uniform regions in the image are the perceptually relevant parts of the image to be used in noise estimation. The noise estimation methods presented in this chapter are based on the locally uniform features in the image. Except for Martens (1994), all noise estimation methods so far use rectangular windows. We use smooth overlapping windows in the analysis. A motivation for the methods used in this chapter comes from knowledge about the visual system. There is evidence that the early stages of the human visual system use overlapping receptive fields that can be modeled as derivatives of a Gaussian (Young, 1985; Young, 1987), and that such information is gathered at multiple spatial scales. The methods presented in this chapter are based on local derivatives of the image.

In this chapter, a new noise variance estimation algorithm based on gradient energy is developed. Unlike the previous methods which are based on the gradient amplitude, this algorithm is based on the *gradient energy* and its probability density function (PDF) in the uniform regions. An algorithm based on the gradient amplitude, similar to the one in Bracho & Sanderson (1985) is also implemented for comparison purposes. These noise estimation methods based on gradient measures are compared with the method based on overall energy (Martens, 1994).

Most algorithms found in the literature assume that the input noise is white. However, in real images noise is nonwhite, often filtered (i.e., blurred) by a low-pass filter. In this chapter, the autocovariance function of the nonwhite noise is assumed to be Gaussian, characterized by two parameters, the variance and the correlation length. We develop an algorithm for simultaneously estimating both of these parameters.

The perceptual attribute of image quality that is most strongly influenced by noise in an image is the noisiness or the perceived noise in the image (Marmolin & Nyberg, 1975), although it may also affect other attributes such as

sharpness. The perceptual effects of noise in images have also been described in the literature using alternative terms such as visibility of noise (Marmolin & Carlström, 1985) and annoyance of noise (Escalante-Ramírez, 1992). In this chapter the perceptual effect of noise in images is referred to as *noisiness*. For the image quality applications, a measure that correlates well with the noisiness has to be estimated. Noisiness of an image is dependent on many parameters of the noise and of the image, such as: the noise variance, the bandwidth of the noise (Huang, 1965), the average luminance of the image (Marmolin & Carlström, 1985), etc. In this chapter, the influence of these parameters on the noisiness is investigated using psychophysical experiments. Most importantly, it is shown experimentally that the noisiness of images with white noise depends only on the noise SD and is largely independent of the PDF of the noise. Using the results of the psychophysical experiments, a measure for the perceived noisiness in images, called noise-index, is presented. It uses the estimates of the image noise variance.

This chapter is organized as follows. In Section 3.2 the noise estimators for white and nonwhite noise are derived. In Section 3.3, a description of the noise estimation algorithms is given. The performance of the noise variance estimation algorithms are discussed in Section 3.4. The psychophysical experiments on noisiness are described in section 3.5. Based on the results of the psychophysical experiments, a measure for noisiness, called the noise-index, is proposed in Section 3.6. An algorithm to estimate the noise-index and the results of that algorithm are given in Section 3.7.

3.2 Noise Model Parameter Estimation

The aim here is to estimate the parameters of $\tilde{n}(x, y)$ in Eq. (3.1) using the polynomial coefficients in the uniform regions of the image. We refer to Section 2.2 of Chapter 2, for a discussion on polynomial transforms, as well as the relevant notation. When $\tilde{n}(x, y)$ is a Gaussian random process, the polynomial coefficients are jointly Gaussian random variables. In a uniform region they have zero mean except for $f_{0,0}$. Let the input noise $\tilde{n}(x, y)$ be signal-independent, zero mean, stationary, Gaussian stochastic process. In addition to these properties, if the noise autocovariance function $R(x, y)$ satisfies certain symmetry conditions (Escalante Ramirez & Martens, 1992), then the first-order coefficients $f_{1,0}$ and $f_{0,1}$ are independent Gaussian random variables with equal variance. The probability density function (PDF) of the gradient energy E_1 and that of the gradient amplitude $|f_1| = \sqrt{E_1}$, in a uniform region of the image, are then given by (Papoulis, 1991),

$$P(E_1) = \frac{1}{2\beta^2} \exp\left(-\frac{E_1}{2\beta^2}\right) U(E_1), \quad (3.2)$$

$$P(|f_1|) = \frac{|f_1|}{\beta^2} \exp\left(-\frac{|f_1|^2}{2\beta^2}\right) U(|f_1|), \quad (3.3)$$

where $U(\cdot)$ is the unit step function and $\beta^2 = R_{1,0;1,0} = R_{0,1;0,1}$ is the cross-correlation of the polynomial coefficients, i.e.,

$$R_{m,n-m;l,k-l} = [R(x, y) * a_{m,n-m}(x, y) * a_{l,k-l}(-x, -y)]_{x=y=0}. \quad (3.4)$$

The mean and the variance of E_1 are $2\beta^2$ and $4\beta^4$ respectively. Note that in the uniform regions of the image $|f_1|$ has a Rayleigh density with peak density at $|f_1| = \beta$ (Papoulis, 1991).

3.2.1 White Noise

If the noise is white with variance σ_n^2 , i.e., $R(x, y) = \sigma_n^2 \delta(x, y)$, then we obtain from Eq. (3.4),

$$R_{1,0;1,0} = \sigma_n^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} a_{1,0}^2(x, y) dx dy, \quad (3.5)$$

which can be further simplified if the filter function $a_{1,0}(x, y)$ is separable. If $w^2(x, y)$ is a Gaussian window with parameter σ , i.e.,

$$w^2(x, y) = \frac{1}{\pi\sigma^2} \exp\left[-\frac{(x^2 + y^2)}{\sigma^2}\right], \quad (3.6)$$

then $a_{1,0}(x, y)$ is separable (Martens, 1990a) and we obtain from Eq. (3.5)

$$\beta = \frac{\sigma_n}{2\sqrt{\pi}\sigma}. \quad (3.7)$$

Instead, if $w^2(x, y)$ is a binomial window of order B , then $a_{1,0}(x, y)$ is again separable (Martens, 1990a) and we obtain

$$\beta = \frac{\sigma_n B}{2B - 1} \left[\frac{(2B)!}{(2^B B!)^2} \right]^2. \quad (3.8)$$

In both cases, β is proportional to σ_n .

3.2.2 Nonwhite Noise

Let the noise autocovariance function be

$$R(x, y; \sigma_n; l_c) = \frac{\sigma_n^2}{2\pi l_c^2} \exp\left[-\frac{(x^2 + y^2)}{2l_c^2}\right]. \quad (3.9)$$

This autocovariance function is the same as that of a white noise with variance σ_n^2 filtered (or blurred) by a Gaussian filter of parameter $\sigma = l_c$ (as in Eq. (3.6)). The aim is to determine both the noise variance σ_n^2 and the correlation length l_c . The SD of the nonwhite noise is given by

$$\eta = \sqrt{R(0, 0; \sigma_n; l_c)} = \frac{\sigma_n}{\sqrt{2\pi} l_c}. \quad (3.10)$$

The relation between the polynomial coefficients of a blurred and the unblurred image has been derived in Chapter 2. Using those relations and substituting Eq. (3.9) in Eq. (3.4), it can be shown that, for nonwhite noise with autocovariance function given in Eq. (3.9), the correlation between the polynomial coefficients for a Gaussian analysis window of spread σ becomes

$$\begin{aligned} R_{m,n-m;l,k-l}(\sigma_n, l_c; \sigma) \\ &= [R(x, y; \sigma_n; l_c) * a_{m,n-m}(x, y; \sigma) * a_{l,k-l}(-x, -y; \sigma)]_{x=y=0} \\ &= \left(\frac{\sigma}{\sigma_e}\right)^{n+k} R_{m,n-m;l,k-l}(\sigma_n, 0; \sigma_e) \end{aligned} \quad (3.11)$$

where $\sigma_e = \sqrt{\sigma^2 + l_c^2}$ is the effective spread. In other words, the cross-correlation coefficient of order $m, n-m; l, k-l$ of a noisy image with correlation length l_c analyzed at scale σ is the same as that of a white noise image with variance σ_n^2 analyzed at scale $\sigma_e = \sqrt{\sigma^2 + l_c^2}$, weighted by a factor $(\sigma/\sigma_e)^{n+k}$.

We need to find $\beta = \sqrt{R_{1,0;1,0}}$. Using Eq. (3.11) and Eq. (3.7),

$$\beta = \frac{\sigma_n}{2\sqrt{\pi}\sigma_e} \frac{\sigma}{\sigma_e} = \frac{\sigma_n}{2\sqrt{\pi}} \frac{\sigma}{l_c^2 + \sigma^2} \quad (3.12)$$

is obtained. This relation can also be derived directly using Eq. (3.4) and Eq. (3.9). Note that when $l_c = 0$, Eq. (3.12) reduces to Eq. (3.7). Also for $l_c \ll \sigma$, (3.12) approaches Eq. (3.7).

To estimate both σ_n^2 and l_c we need at least two equations. These are obtained by analyzing the image at two scales σ_1 and σ_2 . Thus from Eq. (3.12) we obtain two simultaneous equations

$$\begin{aligned} \beta_1 &= \frac{\sigma_n \sigma_1}{2\sqrt{\pi}(l_c^2 + \sigma_1^2)}, \\ \beta_2 &= \frac{\sigma_n \sigma_2}{2\sqrt{\pi}(l_c^2 + \sigma_2^2)}, \end{aligned} \quad (3.13)$$

one for each scale. By solving the above equations, the following estimates

$$\sigma_n = \frac{2\sqrt{\pi}\beta_1\beta_2(\sigma_2^2 - \sigma_1^2)}{\beta_1\sigma_2 - \beta_2\sigma_1} \quad (3.14)$$

$$l_c = \left[\frac{\sigma_1\sigma_2(\beta_2\sigma_2 - \beta_1\sigma_1)}{\beta_1\sigma_2 - \beta_2\sigma_1} \right]^{1/2} \quad (3.15)$$

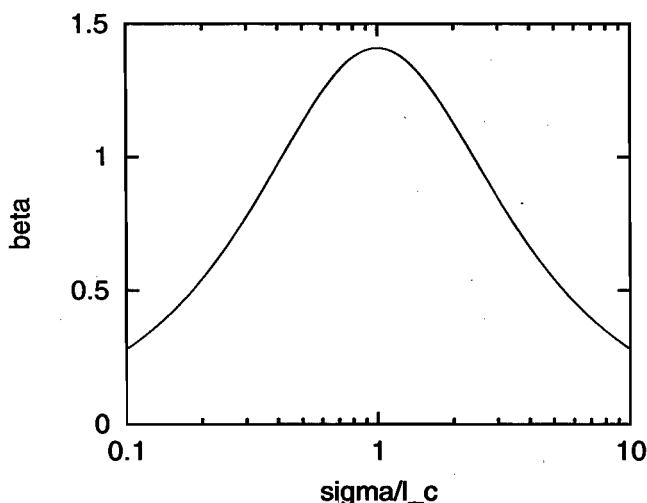


Figure 3.1: The parameter β for nonwhite noise as function of σ/l_c , plotted on a log axis, at $\sigma = 1$.

are obtained, where $\sigma_2 > \sigma_1$.

Figure 3.1 shows how β varies as a function of σ/l_c , on a log axis. The parameter β is maximum when $\sigma = l_c$, and decreases on either side. Provided that the two scales σ_1 and σ_2 are chosen such that they are sufficiently far apart on the log axis and preferably both σ_1 and σ_2 greater than l_c , then the value of β s at those two scales of analysis will be sufficiently distinct to give reliable estimates of σ_n and l_c . For example, this can be attained by choosing $\sigma_2 = 2\sigma_1$.

3.3 Noise Variance Estimation Algorithm

In this section, noise variance estimation algorithms that make use of $P(E_1)$ and $P(|f_1|)$ are described. The estimates derived in the previous section are based on the locally uniform features in the image. To use those methods in images, we make use of the following computational assumption: we assume that

there are at least a few locations in the scene, where the luminance distribution is locally zero-dimensional (0-D)

A 0-D feature is constant in all directions, i.e., uniform. This assumption is valid for most natural scenes. In fact, in most natural scenes, a large percentage

of the image area is locally uniform. An exception to the above assumption are textured images. To obtain an estimate of the noise variance using the PDF of E_1 in the uniform regions, the uniform regions in the image have to be detected first. However, detecting uniform regions itself needs an estimate of the noise variance (see Introduction to this chapter). To circumvent the detection problem and yet be able to estimate the noise variance from the uniform regions in the image, we make use of the following *a priori* knowledge about the class of natural images. For most natural images, a large part of the image areas is locally uniform. This forms our second computational assumption: we assume that,

there are a sufficiently large number of locations in the scene, where the luminance distribution is locally zero-dimensional (0-D)

By sufficiently large we mean large enough to represent the total image area adequately. The first computational assumption is essential to use the 0-D features for noise estimation and the second assumption is needed to circumvent their explicit detection. In the uniform regions, the variations are mainly due to noise and hence the histogram of E_1 in those regions will be close to $P(E_1)$. Steep transitions such as edges, which occur less frequently in the image, result in large values of E_1 , and hence contribute mostly to the tail of the histogram (see Fig. 3.2 for example). Therefore, for images that satisfy the second computational assumption, at low values of E_1 the histogram of E_1 for the entire image will not differ much from $P(E_1)$ in the uniform regions, given by Eq. (3.2). The same holds for the histogram of $|f_1|$. Therefore, the noise can be estimated from the initial part of the histogram of either E_1 or $|f_1|$ for the entire image, without having to explicitly detect the uniform regions.

The original images of natural scenes 'Wanda' and 'Terrace' with which the noise estimation algorithms have been tested, have been shown in Fig. 2.12. The images are of size 512x512 pixels with 8 bits per pixel. Figure 3.2 shows the histograms of E_1 and $|f_1|$ for natural image Wanda with Gaussian distributed noise of standard deviation (SD) $\sigma_n = 10$. The PDFs given by Eq. (3.2) and Eq. (3.3) are also plotted for comparison. For small argument values the histograms do not differ much from their derived PDFs. Thus the parameter β in Eq. (3.2) and in Eq. (3.3) can be determined from the initial part of the histogram of either E_1 or $|f_1|$ for the entire image. The noise variance is computed from the estimate of β .

The noise estimation method based on the residual energy (Martens, 1994) is also implemented for comparison purposes.

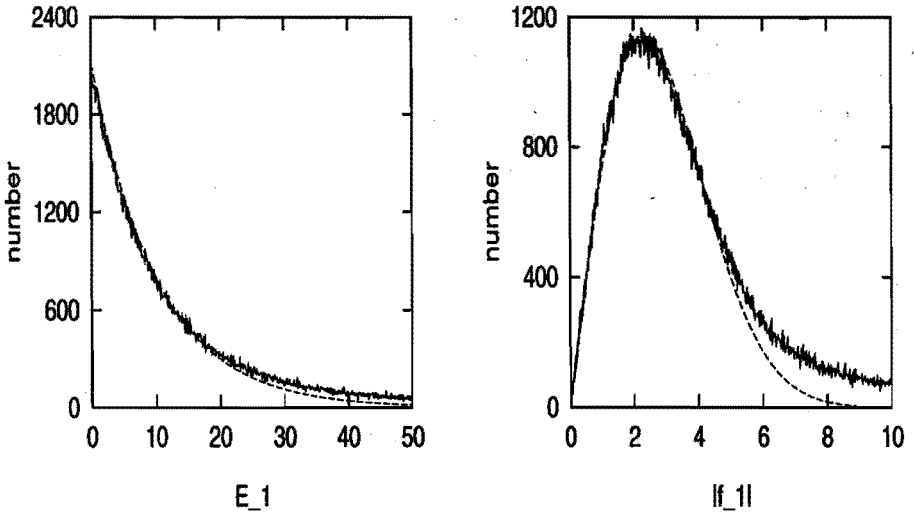


Figure 3.2: Histograms of E_1 and $|f_1|$ for natural image Wanda containing noise with Gaussian PDF of SD $\sigma_n = 10$. Dotted lines show the PDF of E_1 and $|f_1|$ that were fitted to the histograms. The analysis window is binomial of order $B = 4$. The estimate of σ_n from the histogram of E_1 is $\hat{\sigma}_n = 10.95$. The mode of the smoothed histogram of $|f_1|$ is at 2.32 and the resulting estimate $\hat{\sigma}_n = 11.2$.

3.3.1 White Noise

A. Using Gradient Energy

The noise variance is estimated by estimating β in Eq. (3.2). First, the local derivatives $f_{1,0}$, $f_{0,1}$ and the gradient energy $E_1 = f_{1,0}^2 + f_{0,1}^2$ are computed at each sampling position in the image. Next, the histogram of E_1 , $h(E_1)$ for the entire image is computed. An estimate of β is obtained by minimizing the weighted error

$$\min_{\beta, K} \sum_{E_1} \left[h(E_1) - K \exp\left(-\frac{E_1}{2\beta^2}\right) \right]^2 w(E_1). \quad (3.16)$$

where

$$w(E_1) = 1 \text{ for } E_1 < T \quad (3.17)$$

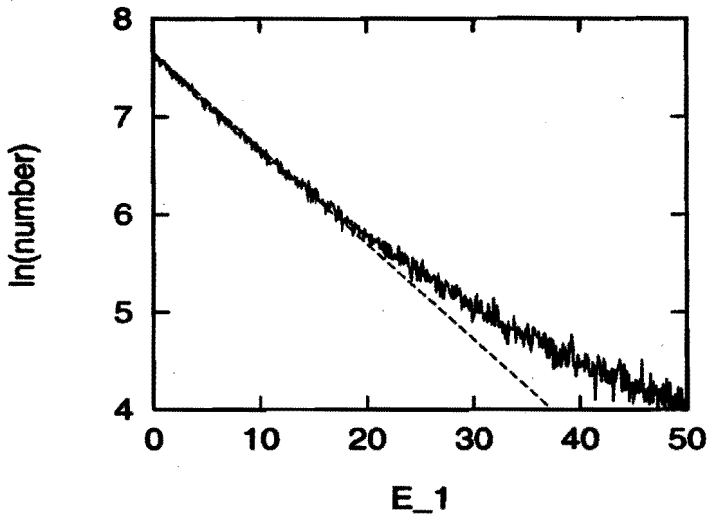


Figure 3.3: Log of the histogram of E_1 for natural image Wanda with added noise of SD $\sigma_n = 10$. The dotted line shows the $\log P(E_1)$ that was fitted to the histogram, only in the range $E_1 = 0$ to 8. The analysis window is binomial of order $B = 4$. The estimated SD from this histogram $\hat{\sigma}_n = 10.93$.

is derived from the estimated value of β using either Eq. (3.7) or Eq. (3.8), depending on which window is used for the analysis.

Alternatively, β can be estimated by linear regression of $\log(P(E_1))$ to E_1 , i.e., by minimizing the weighted error

$$\min_{\beta, K} \sum_{E_1} \left[\log(h(E_1)) + \frac{E_1}{2\beta^2} + C \right]^2 w(E_1). \quad (3.18)$$

where $w(E_1)$ is again given by Eq. (3.17). Figure 3.3 shows a plot of $\log(P(E_1))$ vs E_1 for the natural image Wanda.

There is no significant difference between the results obtained using the method of Eq. (3.16) and those obtained using Eq. (3.18), except when the noise variance is very small, $\sigma_n < 1$. At low noise variances, the linear regression method of Eq. (3.18) is more robust than the method of Eq. (3.16). Hence the method in Eq. (3.18) is used for all estimates that are based on the gradient energy E_1 that are reported in this chapter. In the present implementation, a histogram of 256 bins is used.

B. Using Gradient Amplitude

From Eq. 3.3 we know that the gradient amplitude $|f_1|$ has a Rayleigh density with peak at β . Hence, β can be estimated from the peak of the histogram of $|f_1|$. This method is similar to the one used in (Bracho & Sanderson, 1985). First, the local derivatives $f_{1,0}$, $f_{0,1}$ and the gradient amplitude $|f_1| = \sqrt{f_{1,0}^2 + f_{0,1}^2}$ are computed at each sampling position in the image. Next, the histogram of $|f_1|$ for the entire image is computed. Since the histogram is noisy, it is smoothed before the position of the peak is estimated. In the present implementation uniform windows are used for smoothing, although other smoothing windows such as the binomial could also be considered. The noise variance is computed from the position of the peak (i.e, β), using either Eq. (3.7) or Eq. (3.8) depending on which window is used for the analysis. In the present implementation, a histogram of 1024 bins is used. The range of $|f_1|$ covered by the histogram is decided from the image such that it corresponds to a preset percentage (for example 66 %) of the cumulative histogram of $|f_1|$.

C. Using Residual Energy

Similar to $|f_1|$, the residual energy E_R can also be used to estimate the noise variance (Martens, 1994). When the input noise is white, the PDF of the residual energy in the uniform regions is also peaked. The mode of the histogram is related to the noise variance and can be used to compute the noise variance. For most natural images, the mode of the histogram of the residual energy E_R for the entire image does not differ substantially from that of the PDF in the uniform regions. Hence, the noise variance can be computed from the mode of the histogram of the residual energy for the entire image. An example of the histogram of the residual energy E_R in a natural image with noise is shown in Fig. 3.4. Since the histogram is noisy, it is smoothed before computing the position of the peak. In the present implementation uniform windows are used for smoothing, although other smoothing windows such as binomial could also be considered. In the present implementation, a histogram of 1024 bins is used. The range of E_R covered by the histogram is decided from the image such that it corresponds to a preset percentage (for example 66 %) of the cumulative histogram of E_R . In the implementation, a binomial window is used, although other windows, such as a uniform window, could also be used (Martens, 1994).

3.3.2 Nonwhite Noise

When the input noise is nonwhite with a Gaussian autocovariance function (with certain spread), the image has to be analyzed at two scales in order for us to be able to estimate both σ_n^2 and l_c . At each scale, β is computed using the

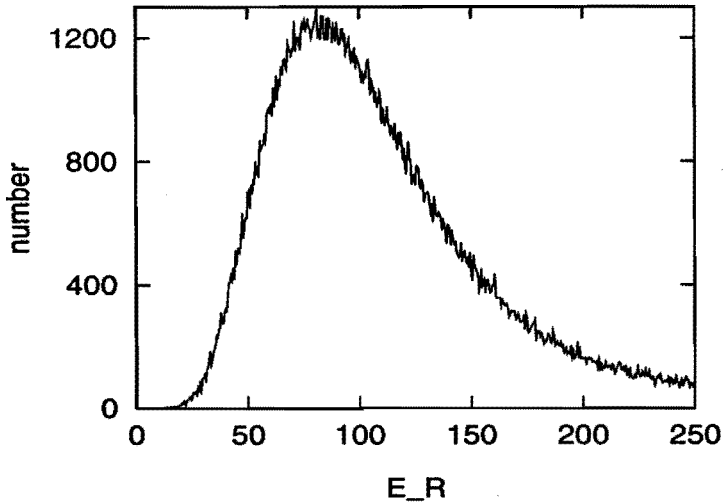


Figure 3.4: Histogram of the residual energy E_R for the natural image Wanda with added noise of SD $\sigma_n = 10$. The analysis window is binomial of order $B = 4$. The estimated mode of the smoothed histogram is 82.0 and the resulting estimate of noise SD $\hat{\sigma}_n = 10.4$.

method described in Eq. (3.18). From the resulting estimates of β at two scales (β_1 and β_2), the noise variance σ_n^2 and the correlation length l_c are computed using Eq. (3.14) and Eq. (3.15). Although not used here, the method based on $|f_1|$ described in subsection 3.3.1.B could also be used to estimate β_1 and β_2 .

The two scales of analysis σ_1 and σ_2 have to be chosen *a priori*. Our simulations show that reliable estimates are obtained for a large range of σ_1 and σ_2 , provided the σ_1 is not very small or σ_2 is not very large compared with the correlation length l_c and when the ratio σ_2/σ_1 is in the range between 1.5 and 2 (See also Fig. 3.1 and Section 3.2.2). The scales can be chosen *a posteriori* using a multi-scale algorithm, similar to that in Chapter 2. However, a multi-scale algorithm has not been attempted here since it is beyond the scope of this chapter.

3.4 Results of Noise Parameter Estimation

Results of applying the noise variance estimation algorithms to synthetic and natural images with varying amounts of noise are shown in Fig. 3.5. The synthetic image was uniform with grey value 163 and the natural images were

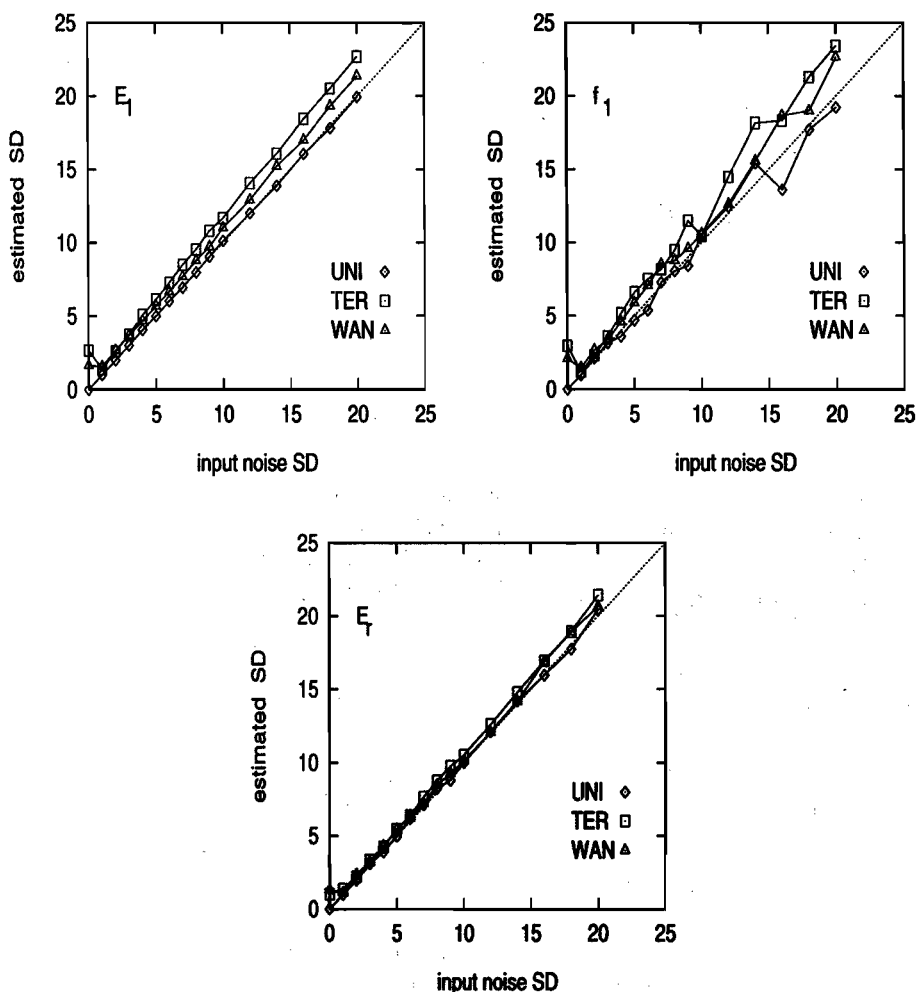


Figure 3.5: Results of noise variance estimation using E_1 , $|f_1|$ and E_R , for a uniform image and natural images Wanda and Terrace, with varying amounts of noise. The dotted line shows the ideally expected estimate. The analysis window was binomial of order $B = 4$.

Wanda and Terrace. For the uniform image, estimation algorithms based on E_1 and E_R both give reliable estimates in the entire range tested. The algorithm based on the peak of the histogram of $|f_1|$, gives less reliable estimates. For the natural images, algorithms based on E_1 and E_R both give consistent estimates in the entire range tested, although estimates based on E_1 are approximately 5% higher than those given by E_R . This is mainly because of the contributions

of the image structure (i.e., non-uniform regions) to the histogram of E_1 at low values of E_1 , which causes the histogram to deviate from the PDF. Note that the estimates show that the 'original' images of both natural scenes are corrupted by noise. These estimates are obtained using a binomial window of order $B = 4$.

From these simulations it can be concluded that the estimates based on the histogram of the first-order energy E_1 for the entire image provide satisfactory estimates of the noise SD. Estimates based on the mode of the histogram of the residual energy E_R in the entire image provide slightly more reliable estimates of the noise SD. The difference between the two methods is mainly that the estimates based on E_1 are about 5% higher than those provided by E_R . If the same window is used to analyze the image in both methods, then the computational complexity to compute E_1 or E_R per pixel is the same. Hence, the difference is only in the complexity of computing the noise variance from the histogram of the energy. The histogram of E_R has to be smoothed before estimating its mode. When $|f_1|$ is used, β is estimated from $h(|f_1|)$ using Eq. (3.18). The choice between these two methods will be mainly decided by the type of intermediate result that is available. For example, if the intermediate results consist of first order derivatives of the image, then the method based on E_1 can be used. This happens for example in edge detection (Canny, 1986) and in segmentation based on gradient information (Bracho & Sanderson, 1985). The method based on the residual energy will be preferred in applications such as noise reduction (Martens, 1994; Lee, 1981). Results of Fig. 3.5 also show that, although the estimates based on the mode of the histogram of $|f_1|$ are less reliable, they can nevertheless be used as a first approximation.

Estimates based on E_1 and $|f_1|$ assume that the PDF of the input noise is Gaussian. Simulations were performed to test the deviations of the estimated noise variance when the input noise PDF deviates from a Gaussian. Images containing noise with three different PDFs: Gaussian, triangular and uniform, were generated by adding noise to a uniform image of grey value 163. Results of the simulation using E_1 are shown in Figure 3.6. As can be seen from the figure, the estimated noise SDs deviate only slightly for triangular and uniform PDFs. The estimated SD of triangular PDF noise is about 2% higher and that of uniform PDF noise is about 4% higher than the expected value of SD (see Fig. 3.6). Similar simulations using natural images (not shown here), also show approximately the same results.

Figure 3.7 shows the results of estimating the SD σ_n and the correlation length l_c of nonwhite noise in uniform, Wanda and Terrace images. Images with different amount of input noise SD σ_n and correlation length l_c were generated from the original images by adding nonwhite noise. The nonwhite noise was generated by filtering the white noise with binomial filters. Estimates were

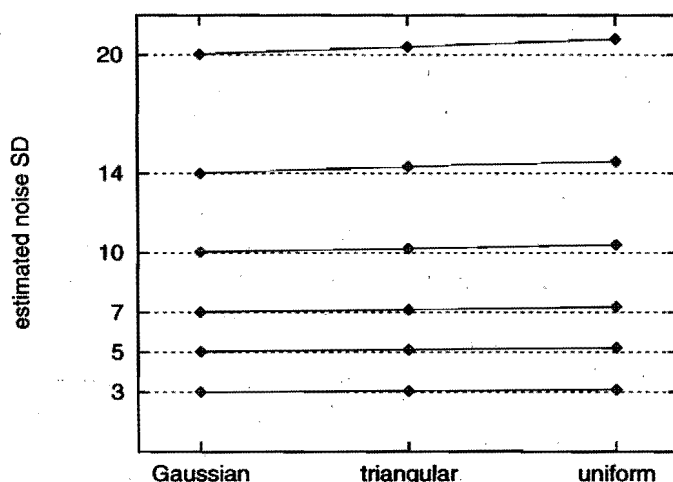


Figure 3.6: Results of noise variance estimation using E_1 in a uniform image with input noise of different PDF. The points joined by a line correspond to images containing the same input noise SD but different PDF. The dotted lines show the expected noise SD. Each point represents the mean of 10 trials. The analysis window was binomial with $B = 4$.

derived by analyzing the images at two scales, using binomial windows of order $B_1 = 2$ and $B_2 = 8$. For the uniform image, the estimates are close to the expected estimates. For the natural image Terrace, the correlation length is overestimated at low values of input noise SD. For the natural image Wanda, the correlation length is underestimated at low input noise SD and high input correlation lengths. For both natural images the noise SD σ_n is overestimated. This is mainly because of the image structure that interferes with the noise at those correlation lengths. Similar results were obtained also when $B_1 = 4$ and $B_2 = 8$ was used.

To study the dependence of the scales of analysis σ_1 and σ_2 on the estimates of nonwhite noise parameters, simulations were performed with different ratios of σ_2/σ_1 . Figure 3.8 shows how the estimates vary with the ratio of the two scales of analysis. When the ratio is close to 1, the estimates are less stable. When the ratio is around 2 the estimates are more stable. The uniform image and the Wanda image show a broad region of stable estimates. The Terrace image shows only a narrow region of stable estimates. A ratio of 2 between scales is also motivated from psychophysical evidence that the human visual system performs spatial analysis using a number of distinct spatial frequency-tuned channels, at approximately one octave apart (Wilson & Bergen, 1979).

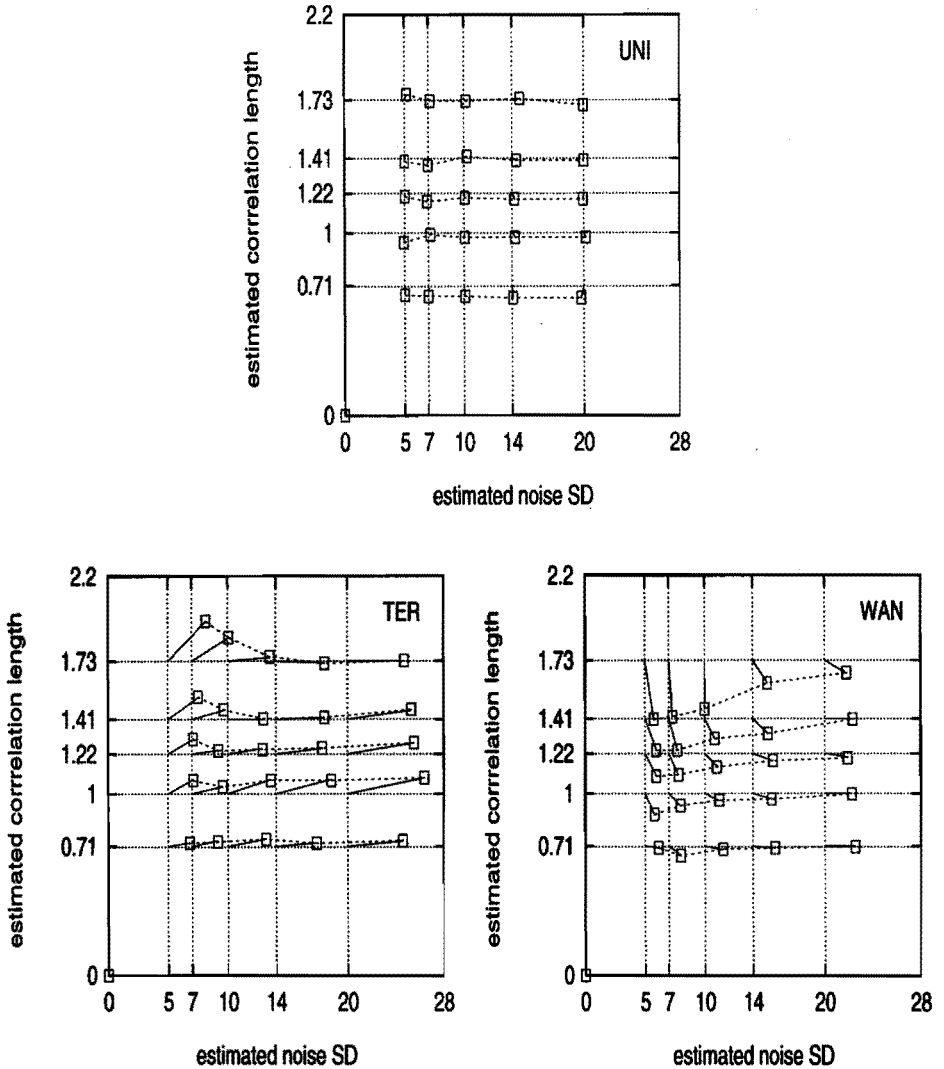


Figure 3.7: Results of nonwhite noise estimation with input noise of different variance and correlation length, for a uniform image, Wanda image and Terrace image. The points joined by a line correspond to images containing the same input noise correlation length l_c but different SD σ_n . The intersections of the dotted lines represent the expected estimates. The analysis windows were binomial with $B_1 = 2$ and $B_2 = 8$.

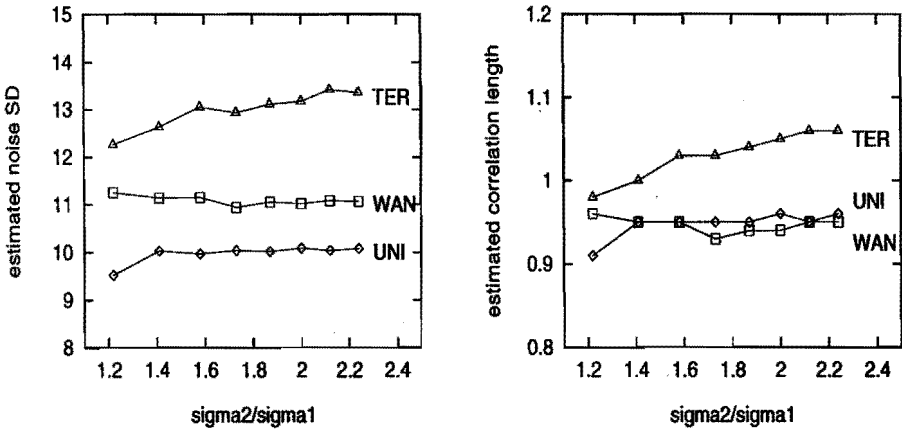


Figure 3.8: Effect of the ratio of analysis scales σ_2/σ_1 on the estimate of nonwhite noise parameters, for uniform image, Wanda and Terrace, with added noise of SD $\sigma_n = 10$ filtered with a binomial window of order $B = 2$ ($l_c \approx 1$). The lower scale of analysis window was fixed at binomial order $B_1 = 2$, i.e., $\sigma_1 = 1.0$ and the upper scale varied from $B_2 = 3$ to $B_2 = 10$, i.e., $\sigma_2 = 1.2$ to 2.2

3.5 Psychophysical Experiments to Measure Noisiness

Perceived noise in an image influences the perceived quality of the image. It has been identified to be one of the important dimensions of image quality (Marmolin & Nyberg, 1975; Goodman & Pearson, 1979). In this thesis we refer to the perceived effect of noise in images as *noisiness*. Several researchers have studied noisiness and its effect on image quality. (Marmolin & Carlström, 1985; Huang, 1965; Hunt & Sera, 1978). Their studies indicate that the main factors that influence noisiness of images are the noise SD and the noise bandwidth (or the noise correlation length). In uniform regions of the image, the mean luminance of the uniform region may also affect the noisiness. Using psychophysical experiments, we investigate how each of the above parameters influences the noisiness in images. We also investigate whether the form of the PDF of the noise influences the noisiness in images.

3.5.1 Equipment and Subjects

In all psychophysical experiments described in this chapter, images were displayed on a BARCO-CCID-7351B high-resolution non-interlace monitor using a VME-bus based system (with MEN-A301 graphics board). The rela-

tion between the grey level and the luminance on the monitor is specified by the Gamma characteristics of the monitor (Poynton, 1993). It is given by $L = L_{max}(g/255)^\gamma$ where L is the luminance, L_{max} is the maximum luminance of the monitor, g is the grey value and γ is the exponent. The Gamma characteristics of the monitor was calibrated to have a $\gamma = 2.5$, with a maximum monitor luminance of 60 cd/m^2 . The viewing conditions satisfied the CCIR recommendation 500 (CCIR, 1986) except for the viewing distance which was 1.4 meters for all experiments. At this viewing distance, each pixel on the monitor extended approximately $1.3 \text{ min arc} \times 1.3 \text{ min arc}$ retinal angle, and the viewing distance to image height (d/h) ratio varied from 5.5 to 8 depending on the height of the images used in different experiments. In all experiments except Experiment 2, each stimulus was displayed for a fixed duration of 5 sec. In Experiment 2, each stimulus was displayed for 8 sec. Between two presentations, a uniform adaptation field was displayed, at least for 2 sec. The luminance of the adaptation field was approximately equal to that of the average luminance of the stimuli.

All subjects that took part in the experiments were between 22 and 38 years of age. All were familiar with the use of numerical category scaling procedures and the concept of noisiness in images. All had normal or corrected-to-normal visual acuity between 1.5 and 2, measured on a Landolt chart. Before the start of each experiment, subjects took part in a trial session with a number of stimuli representing the entire range of the stimuli in the experiment, in order to adjust the sensitivity of their scale.

3.5.2 Relevance of the PDF of the Noise to Noisiness

Two psychophysical experiments were performed to study the influence of the probability density function (PDF) of the noise on the noisiness of images. The question we sought to answer using these experiments was: Suppose a subject is provided with two images of the same scene, containing white noise of equal variance but with different PDFs, will the subject perceive one image to be more noisy than the other? Three forms of PDF are considered in the experiment: Uniform, triangular and Gaussian. The Gaussian PDF is one of the most commonly used descriptions of noise (Demoment, 1989). The PDF of quantization noise is uniform. The PDF of the sum of two random variables with uniform PDF is triangular.

Two experiments have been used to address the question. The first experiment uses a numerical category scaling technique. The second experiment is based on a method of paired comparisons, where two images are compared for noisiness. These two methods complement each other, in the sense that the category scaling is very efficient for studying a large range of differences and the

method of paired comparisons is more reliable for studying small differences.

3.5.3 Experiment 1: Scaling Noisiness

In this experiment the numerical category scaling (de Ridder & Majoor, 1990) method is used to infer the dependence of the noisiness on the PDF of the noise. The numerical category scaling has been shown to be an efficient and reliable technique for evaluating image quality and its attributes (de Ridder & Majoor, 1990).

Stimuli: The variables in this experiment were the PDF of the noise and the noise SD. Three PDFs: Gaussian, triangular and uniform, were used. Five levels of standard deviations were used: 7.1, 10, 14.1, 20 and 25. Noise of all combinations of three PDFs and five SDs were considered to obtain a total of $3 \times 5 = 15$ stimuli. Stimuli were generated by adding noise of a given PDF and SD to every pixel of a uniform image. The uniform image had a grey value of 163, which resulted in an average luminance of 20 cd/m^2 when displayed on the monitor. The stimuli had a size of 335×335 pixels with 8 bits per pixel accuracy. Each stimulus was repeated 4 times in the experiment. To have independent noise samples in the four repetitions, four replicas of each stimulus with independent noise samples were generated. Thus, a total of $3 \times 5 \times 4 = 60$ images were created. The images had a size of 0.17 m by 0.17 m on the monitor (7° by 7° in viewing angle).

Method: The subjects were instructed to judge the noisiness of the images using a 10-point numerical category scale ranging from one to ten. All 15 stimuli were presented four times, each time with independent noise samples. Thus there were $15 \times 4 = 60$ presentations. The order of presentation was random. Ten subjects took part in the experiment.

Results: The numerical category data from the experiment were transformed into a psychologically linear (interval) scale using Thurstone's law of categorical judgement. This was carried out in the same way as was earlier described in Session 2.9 of Chapter 2.

Figure 3.9 shows the results for four subjects. The trends in the data for the other six subjects were also very similar. Since there were no significant differences between the trends in the data of the subjects, the Thurstone-corrected data were averaged over subjects. Before averaging over subjects, the data for each subject were normalized using the z-score transform (Hays, 1988). Finally, the z-scores averaged over the subjects were normalized so that the resulting average standard error of mean (SEM) is equal to one. The average result for 10 subjects are shown in Fig. 3.10. In the top plot of Fig. 3.10, the points joined by a drawn curve refer to the same input noise SD but a different PDF.

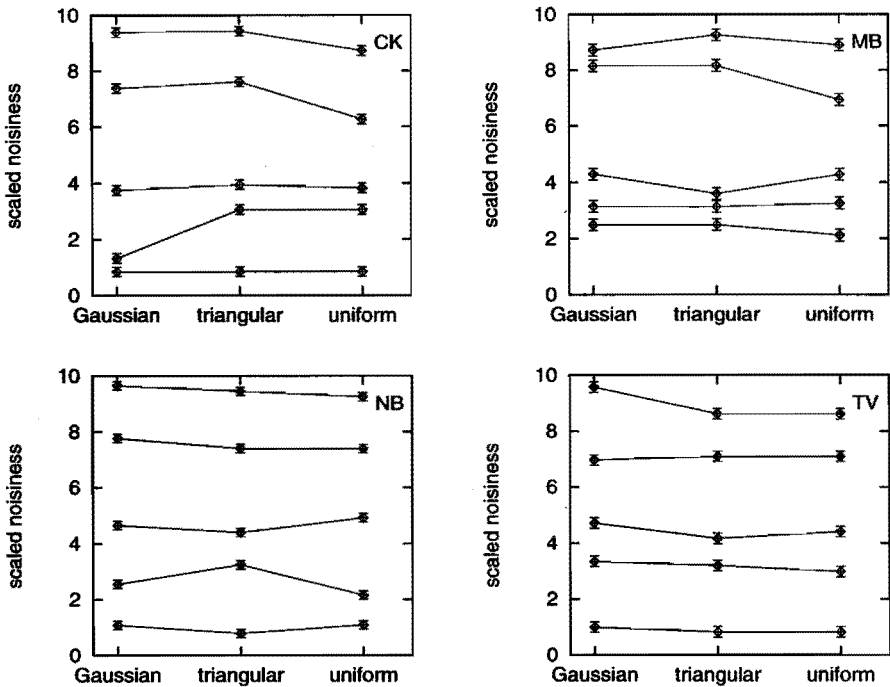


Figure 3.9: Results of Experiment 1 plotted as a function of the noise PDF, with the noise SD σ_n as the parameter, for four subjects: CK, MB, NB and TV. Points joined by a line correspond to stimuli containing noise with equal SD but of different PDF. The lengths of the error bars are equal to twice the standard error of the mean.

There is no systematic dependence of noisiness on the PDF. Hence, it can be concluded from the results of this experiment that the noisiness is not very dependent on the noise PDF in the range tested.

The results as a function of noise SD are plotted in the bottom panel of Fig. 3.10. This figure also shows that the noisiness is not very dependent on the noise PDF. Figure 3.10 also shows that, in the range tested, the noisiness is approximately linearly related to noise SD.

3.5.4 Experiment 2: Comparing Noisiness

The results of the Experiment 1, in which noisiness is scaled using numerical category scaling, showed that the PDF of the noise has no significant influence on the noisiness, at least not in the range that was tested. However, using

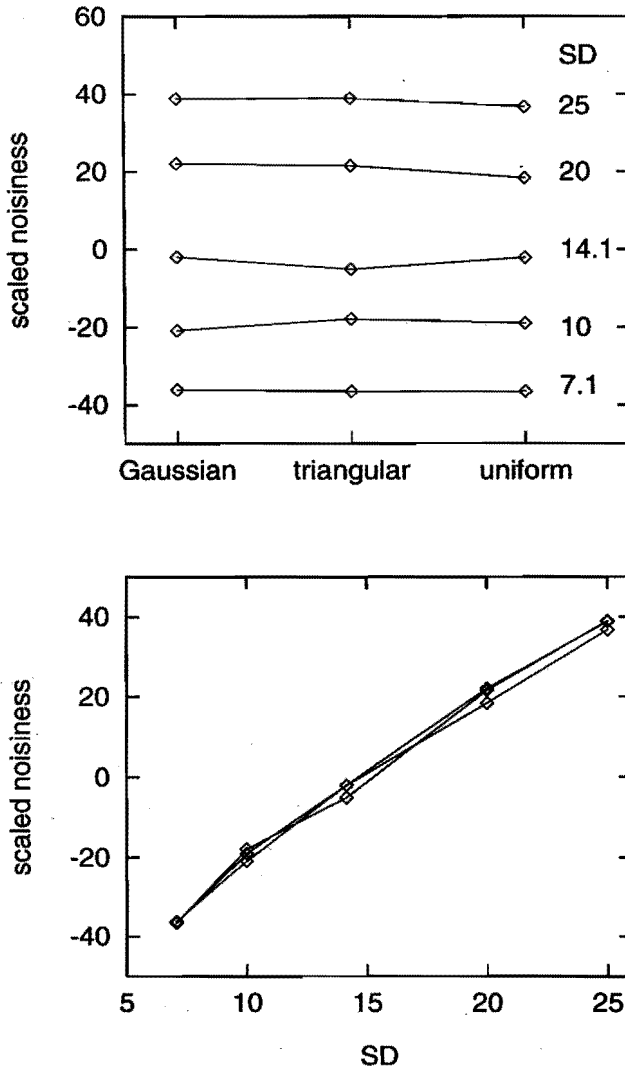


Figure 3.10: Top: Results of Experiment 1 plotted as a function of the noise PDF, with the noise SD σ_n as the parameter. Points joined by a line correspond to stimuli containing noise with equal SD but of different PDF. Each point is the average over 10 subjects. Bottom: The same results plotted as a function of the noise SD σ_n .

numerical category scaling it is in general difficult to discriminate stimuli with very small perceived differences (order of 1 jnd or less) (Sjoberg, 1987). On the other hand, the method of paired comparisons can reveal small differences, but is more laborious. In general, the task of comparing images for one attribute is less taxing for the subjects than scaling the attribute of an image. Hence, more reliable results could be obtained at an added cost of having to perform many more trials. Thus, a second experiment was carried out to test more rigorously whether the PDF of the noise influences the noisiness and to confirm the results of the first experiment. This experiment deals with the comparison of the noisiness of two images.

Stimuli: In this experiment, each stimulus was an image pair. Each image in the pair was 240 pixels wide and 470 pixels high, with 8 bits per pixel. Thus two images could be displayed simultaneously side by side on the monitor (one on the left, and the other on the right) with a separation of 20 pixels between them. The two images in a pair had noise with the same SD but with the same or different PDF. There were three different sessions in this experiment. Stimuli in each session had three PDFs. From these three PDFs, all three possible pairs were considered: (Gaussian, triangular), (triangular, uniform), and (uniform, Gaussian). In the experiment each stimulus pair was presented 20 times. In addition, one pair of identical stimuli (Gaussian, Gaussian) were also used for reference purposes. It was presented 10 times. To have independent noise samples in all the presentations, 20 replicas of each stimulus pair were generated with independent noise samples. Thus, a total of $3 \times 2 \times 20 + 2 \times 10 = 140$ images (70 pairs) were generated per session.

The stimuli in three sessions differ either in the image to which the noise has been added to obtain the stimuli or the domain in which the noise has been generated. The stimuli in the first session were generated by adding noise of SD, $\sigma_n = 20$ to a uniform image of grey value 163. This uniform image had a luminance of 20 cd/m^2 when displayed on the monitor (i.e. for Gamma=2.5 and maximum monitor luminance of 60 cd/m^2). The stimuli in the second session were generated by adding noise in the luminance domain. Noise with SD $\sigma_n = 4 \text{ cd/m}^2$ was added to a uniform image of 20 cd/m^2 . In the grey level domain, this approximately corresponded to SD $\sigma_n = 13.5$ added to a uniform image of grey level 163. Note that now the PDFs have their defined form in the luminance domain. The corresponding PDFs in the grey scale domain will have a slightly skewed form due to the non-linear Gamma transformation. The stimuli in the third session were generated by adding noise with SD $\sigma_n = 10$ to the grey values of the natural image Wanda. The images had a size of 0.12 m by 0.25 m on the monitor (4.8° by 10.2°), with 1.25 cm separation.

Method: Each stimulus (image pair) was displayed for a fixed duration of 8 sec. The subjects were instructed to indicate which image in the presented pair

appeared most noisy. They had to choose one of the two. Each PDF pair in each session was presented 20 times, 10 times with an image with one type of PDF on the left of the monitor and 10 pairs with the same on the right side of the monitor. The identical pair was presented 10 times. Thus, for each session there were 70 presentations. Every presentation had images with independent noise samples. The subjects took part first in session one, then in session two and finally in session three. The order of presentation within each session was random. Seven subjects took part in the experiment.

Results: Table 3.I shows the results. Each entry (i, j) against stimulus pair (A,B) shows that: i times the subject found the stimulus A most noisy and j times the stimulus B. In the last row of each session, each entry (i, j) against stimulus pair (Gaussian, Gaussian) shows that: i times the subject found the stimulus displayed on the left of the monitor most noisy and j times the one on the right. These results are interpreted by using the χ^2 goodness-of-fit test (Chatfield, 1983). The null hypothesis used is H_0 : the subject perceives two images containing noise with equal SD but of different PDF to be equally noisy. Then, the alternative hypothesis is, H_1 : they are different in noisiness, i.e., the subject perceives one to be more noisy than the other. If H_0 holds, the expected probability of the subject choosing either image in a pair to be most noisy is equal, i.e., $P_1 = P_2 = 1/2$. Therefore out of 20 presentations, the expected frequency of each image being chosen is $20 \times (1/2) = 10$. Then the χ^2 statistic is given by (Chatfield, 1983),

$$\chi^2 = \frac{(i - 10)^2}{10} + \frac{(j - 10)^2}{10} \quad (3.19)$$

and $i + j = 20$. In our experiment, the number of categories are two, hence the number of degrees of freedom are $(2 - 1) = 1$. For this case, from the χ^2 distribution tables (Chatfield, 1983), if χ^2 computed in Eq. (3.19) is less than 3.84, then the results will not be significant at 5% level and hence the null hypothesis can be accepted. If χ^2 is greater than 3.84, then the results will be significant at 5% level and hence the null hypothesis has to be rejected. This happens when either i or j in the data pair (i, j) in Table 3.I is greater than or equal to 15, for the case of 20 presentations. For the case of ten presentations, the null hypothesis fails when either i or j is greater than or equal to 9. In Table 3.I, the underlined data show the cases where the null hypothesis fails. The last column shows the number of subjects for which the null hypothesis holds. Note that for the identical pair, (Gaussian, Gaussian) presented ten times, the hypothesis holds for all subjects in all the three sessions. This shows that the experimental set-up (i.e., the monitor) was not biased for noisiness judgement.

From the results, the following conclusions can be drawn: (1) for natural images, the perceived noisiness is independent of the noise PDF, at least up

Table 3.I: Results of Experiment 2. Each entry $i - j$ against stimulus pair A-B shows that: i times the subject found the image A most noisy and j times the image B. N is the number of times the stimulus pair was presented. The underlined data show the stimulus pair for which the null hypothesis H_0 fails. M is the number of subjects (out of seven) for which the H_0 holds.

1. Uniform image(163) + noise with $\sigma_n = 20$ in the grey level domain									
stimulus pair	N	Subject							M
		CK	CT	GJ	MB	NB	TR	TV	
Gaussian-uniform	20	12-8	9-11	10-10	<u>20-0</u>	<u>20-0</u>	<u>5-15</u>	<u>15-5</u>	3
Gaussian-triangular	20	<u>15-5</u>	12-8	9-11	<u>20-0</u>	<u>18-2</u>	11-9	8-12	4
triangular-uniform	20	11-9	14-6	8-12	13-7	<u>15-5</u>	8-12	14-6	6
Gaussian-Gaussian	10	5-5	6-4	5-5	5-5	8-2	7-3	4-6	7
2. Uniform image($20cd/m^2$) + noise with $\sigma_n = 4cd/m^2$									
stimulus pair	N	Subject							M
		CK	CT	GJ	MB	NB	TR	TV	
Gaussian-uniform	20	<u>15-5</u>	14-6	8-12	<u>17-3</u>	<u>15-5</u>	11-9	13-7	4
Gaussian-triangular	20	9-11	9-11	9-11	13-7	<u>16-4</u>	12-8	7-13	6
triangular-uniform	20	11-9	12-8	9-11	<u>16-4</u>	11-9	10-10	6-14	6
Gaussian-Gaussian	10	5-5	5-5	8-2	3-7	5-5	6-4	8-2	7
3. Natural image Wanda + noise with $\sigma_n = 10$ in the grey level domain									
stimulus pair	N	Subject							M
		CK	CT	GJ	MB	NB	TR	TV	
Gaussian-uniform	20	11-9	6-14	12-8	13-7	<u>16-4</u>	12-8	10-10	6
Gaussian-triangular	20	12-8	6-14	12-8	12-8	13-7	12-8	9-11	7
triangular-uniform	20	9-11	14-6	12-8	9-11	12-8	10-10	14-6	7
Gaussian-Gaussian	10	6-4	2-8	5-5	6-4	5-5	3-7	5-5	7

to SD $\sigma_n = 10$, (2) for uniform regions, at large noise SDs (of the order of $\sigma_n = 20$), most subjects do not perceive the noisiness of images with noise of different PDF to be different. However, at high SDs some subjects do perceive the noisiness of the Gaussian PDF noise to be more than that of the uniform PDF noise, (3) there is no significant difference between the results of the experiments with noise added in the grey level domain and noise added in the luminance domain. For the case of noise in a uniform background, the results for the subjects 'MB' and 'NB' differ significantly from those of the other subjects. This may be because the subjects were identifying the stimuli with a particular PDF and rated them consistently to be most noisy, without actually comparing the noisiness of the two images in the pair. The important

practical implication of the results of this experiment is that, for natural images, the perceived noisiness is independent of the type of noise PDF but depends only on the noise SD, especially for $\sigma_n < 10$, the range that is of practical importance (for instance, in applications such as dithering of images).

3.5.5 Experiment 3: Effect of Noise SD on Noisiness

The results of Experiment 1, plotted as a function of noise SD in Fig 3.10, show that the noisiness is approximately linearly related to the noise SD σ_n , in the range tested. However, in Experiment 1, noise SD σ_n was varied only from 7 to 25, in the grey level domain. To obtain a measure for noisiness that is valid for a much wider range, it is necessary to study the relation between noisiness and SD also at very low SDs. The following experiment was conducted to achieve this goal.

Stimuli: The variable in this experiment is the noise SD. The stimuli were generated by adding white noise with Gaussian PDF, in the luminance domain. Noise with given SD was added to each pixel of a uniform image of size 335x335 pixels. There were two scenes. In the first scene, the luminance of the uniform image was 20cd/m^2 and in the second 30cd/m^2 . There were 20 stimuli per scene, whose SD varied from 0 to 6.17cd/m^2 . This approximately corresponds to SD range of 0 to 20 in the grey level domain, for the scene with 20cd/m^2 luminance. The highest noise SD chosen was dictated by the criteria that the histogram of the noisy image may (at most) be slightly clipped at the ends. The images had a size of 0.17 m by 0.17 m on the monitor (7° by 7°) and 8 bits per pixel accuracy.

Method: The subjects were instructed to judge the noisiness of the images using a 10-point numerical category scale ranging from one to ten. There were two sessions, one for 20cd/m^2 scene and another for 30cd/m^2 scene. Each stimuli were presented four times. Thus there were $20 \times 4 = 80$ presentations per session. The order of presentation was random. Eight subjects took part in the experiment.

Results: The numerical category data from the experiment were transformed in the same way as in Experiment 1 described in subsection 3.5.3. Since there were no significant differences between the trends in the data of the subjects, the Thurstone-corrected data were averaged over subjects in the same way as in Experiment 1. The average result over eight subjects is shown in Fig. 3.11. It can be seen from the figure that at both luminances, there is a threshold SD below which the perceived noisiness is approximately the same as that of the (noiseless) uniform image. This threshold is higher at higher average luminance. Discussion about the dashed curves is deferred until Section 3.6.

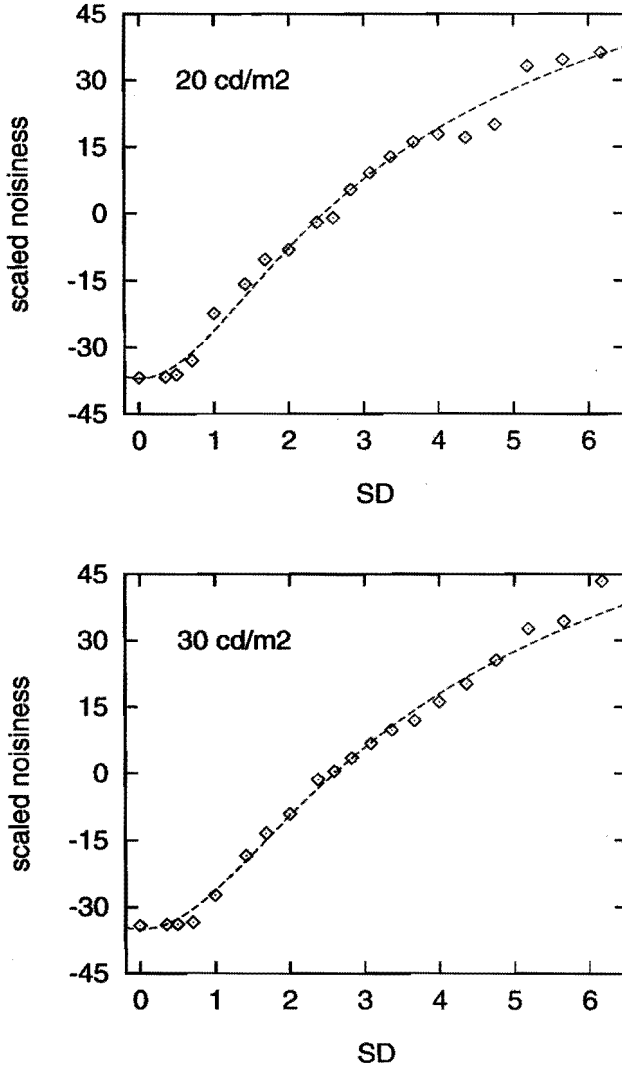


Figure 3.11: Results of Experiment 3 plotted as a function of the noise SD σ_n . Each point is the average over 8 subjects. The drawn lines are given by Eq. 3.20.

3.5.6 Experiment 4: Effect of Local Luminance on Noisiness

The influence of the average luminance of the image on the visibility of noise has been studied in (Marmolin & Carlström, 1985; Girod, 1989). Marmolin & Carlström (1985) found that the perceived noisiness is proportional to the ratio of the noise SD σ_n to the average luminance. They used a monitor with a Gamma of 1.0. Girod (1989) studied the visibility thresholds of noise in uniform images, as a function of the luminance of the uniform images. He found that, for a monitor Gamma of 2.2, when the background luminance increases from zero, the just visible variance of the noise first decreases steeply, reaches a minimum at around a grey value of 82 and then increases slowly. To obtain a local measure of noisiness, it is important to study how the local luminance of the region can influence the noisiness. This experiment is aimed at studying how the noisiness in uniform regions of an image varies with the local luminance. In the experiment the stimuli are generated such that only the local luminance changes while the average luminance of the image remains constant.

Stimuli: The schematic diagram of the image used in this experiment is shown in Fig. 3.12. The image was of size 384x384 pixels with 8 bits per pixel accuracy. Each sub-block was uniform with a size of 128x128 pixels. The grey values for one stimulus are given in Fig. 3.12. Stimuli were generated by adding Gaussian-distributed white noise of given SD to each pixel in the central sub-block of the image. Four SDs were used in the experiment: SD = 5, 7.07, 10 and 14.1. Stimuli with a different local luminance (in the central block) were obtained from the image by swapping the central sub-block with one of the corner sub-blocks, before adding noise. Thus, the average luminance of the stimuli was kept constant. Five local luminances were used in the experiment: 1.9, 5.2, 10.7, 18.7 and 29.5 cd/m^2 , corresponding to grey values 64, 96, 128, 160 and 192 respectively. These five luminance levels correspond to the central and the four corner sub-blocks in Fig 3.12. All combinations of 5 luminance levels and 4 SDs were considered to obtain a total of $4 \times 5 = 20$ stimuli. The images had a size of 0.2 m by 0.2 m (8.2° by 8.2°) when displayed on the monitor. Thus the central square region was 2.7° by 2.7° in size.

Method: The subjects were instructed to judge "the noisiness in the central square region of the image" using a 10-point numerical category scale ranging from one to ten. Each stimulus was repeated four times. The order of presentation was random. Eight subjects took part in the experiment.

Results: The numerical category scaling data from the experiment were transformed in the same way as in Experiment 1 described in subsection 3.5.3. Since there were no significant differences between the trends in the data of the subjects, the Thurstone-corrected data were averaged over subjects in the same

96	144	160
112	64	176
192	80	128

Figure 3.12: A schematic diagram of the stimuli used in the Experiment 4. The image size is 384x384 pixels. Each sub-block has a size of 128x128 pixels. The noise is added only to the central sub-block. The numbers show the grey values of the sub-blocks when the central block has grey value 64. The other four possible grey values at the central sub-block are obtained by swapping the central sub-block with one of the corner sub-blocks, before adding noise.

way as in Experiment 1. The average result over eight subjects is shown in Fig. 3.13. In Fig. 3.13, top panel, points joined by a curve contain stimuli with noise of the same SD σ_n but with different local luminance. The results do not show a strong influence of the local luminance on the noisiness. However, there is a weak tendency for noisiness to decrease with increasing local luminance. The two lowermost curves corresponding to SD, $\sigma_n = 5$ and $\sigma_n = 7.07$ show a slight decrease in noisiness as the local luminance increases. However, the curve corresponding to $\sigma_n = 10$ shows no systematic variations. In the uppermost curve corresponding to $\sigma_n = 14$, the two stimuli with local luminance 64 and 96 are perceived as being more noisy than those with local luminance 128 and 160. As illustrated by the error bar, the noisiness judgement for the stimuli with local luminance 192 is less reliable.

The results as a function of the noise SD σ_n are replotted in Fig. 3.13, bottom panel. This figure, similar to Fig. 3.10, also shows that the noisiness is approximately linearly related to the noise SD in the range tested.

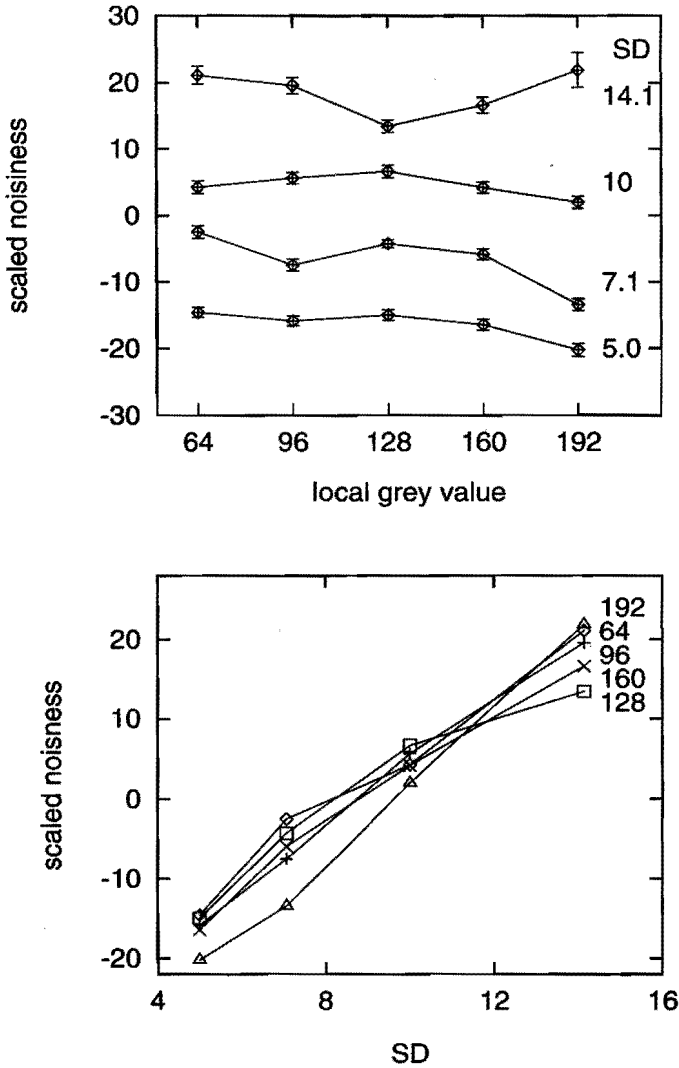


Figure 3.13: Top: Results of Experiment 4 plotted as the function of local grey value, with noise SD as the parameter. Points joined by a line correspond to stimuli containing noise with equal SD but having different local luminance. Each point is the average over 8 subjects. The lengths of the error bars are equal to twice the standard error of the mean. The plot at the bottom shows the same results plotted as a function of noise SD, where local grey value is the parameter.

3.5.7 Experiment 5: Effect of White Noise on Natural Images

In the experiments so far we have studied the effect of white noise on the noisiness of uniform regions. The following experiment is aimed at studying the effect of white noise on the noisiness of the natural images. The variable in this experiment is the noise SD σ_n . The results of this experiment will be used to test the correlation between noisiness and the proposed measure for noisiness.

Stimuli: Two natural scenes were used in this experiment: Terrace and Wanda. Stimuli were generated by adding white noise with a Gaussian PDF to the ‘original’ images of the scenes. Six levels of the noise SD σ_n were used: 3.5, 5, 7.1, 10, 12.25 and 14.1. The original images of the scenes were also used in the experiment. Thus there were seven stimuli per scene. The images had a size of 512x512 pixels with 8 bits per pixel accuracy, but in the experiment only the central region of 450x450 pixels (0.24 m by 0.24 m or 9.8° by 9.8°) was displayed.

Method: The subjects were instructed to judge the noisiness of the images using a 10-point numerical category scale ranging from one to ten. Subjects took part first in a session with images of the Terrace scene and then with those of the Wanda scene. Each stimulus was repeated four times. Thus there were $7 \times 4 = 28$ presentations per session. The order of presentation in each session was random. Eight subjects took part in the experiment.

Results: The numerical category scaling data from the experiment were transformed in the same way as in Experiment 1 described in subsection 3.5.3. Since there were no significant differences between the trends in the data between subjects, the Thurstone-corrected data were averaged over subjects in the same way as in Experiment 1. The average result over eight subjects is shown in Fig. 3.14. Notice that the original images of both scenes (the lowermost point in each plot) deviate from the drawn line.

3.5.8 Experiment 6: Effect of Noise Correlation Length on Noisiness

So far we have studied the effect of white noise on the noisiness. The noise bandwidth or the correlation length also strongly influences the noisiness (Huang, 1965). This experiment was aimed at studying how the noisiness varies as a function of the noise correlation length.

Stimuli: Two natural scenes were used in this experiment: Terrace and Wanda. Stimuli were generated by adding filtered noise to the ‘original’ images of the scenes. Filtered noise was obtained by filtering white noise with a Gaussian PDF with standard deviation σ_n by means of a binomial filter of order

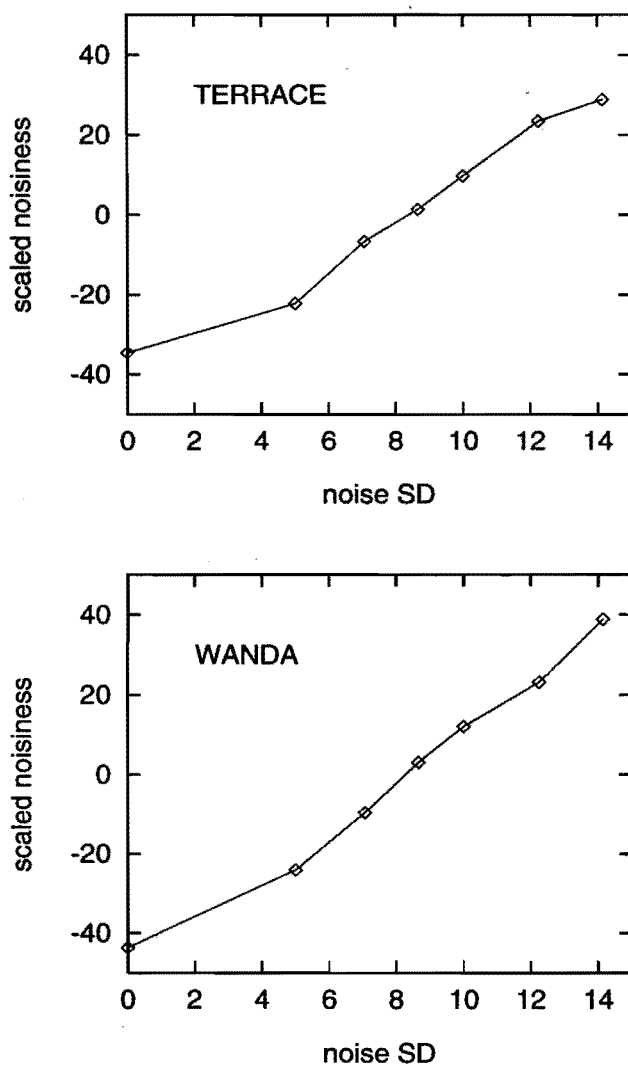


Figure 3.14: Results of Experiment 5, for natural scenes Wanda and Terrace. Each point is the average over 8 subjects. The lengths of the error bars are equal to twice the standard error of the mean.

B. The SD of the filtered noise is equal to $\eta = \sigma_n / \sqrt{2\pi}l_c$ (see. Eq. (3.10)), where $l_c \approx \sqrt{B/2}$ (Martens, 1990a). Two values for the η were used: 5 and 7.07. Seven values of the correlation length l_c were used: 0, 0.71, 1.0, 1.22, 1.41, 1.73, 2.0, corresponding to binomial filters of order $B = 0$ (unfiltered), 1, 2, 3, 4, 6 and 8. The standard deviation of the white noise to be filtered, σ_n , was chosen such that the filtered noise would have the desired standard deviation η . The original images were also used in the experiment. Thus, considering all levels of η and l_c , for both scenes, a total of $(2 \times 7 + 1) \times 2 = 30$ stimuli were generated. The images had a size of 512x512 pixels with 8 bits per pixel accuracy, but in the experiment only the central region of 450x450 pixels (0.24 m by 0.24 m or 9.8° by 9.8°) was displayed.

Method: The subjects were instructed to judge the noisiness of the images using a 10-point numerical category scale ranging from one to ten. Each stimulus was repeated four times. Thus there were $30 \times 4 = 120$ presentations. The order of presentation was random, except that two stimuli of the same scene were not allowed to appear one after the other. Six subjects took part in the experiment.

Results: The numerical category data from the experiment were transformed in the same way as in Experiment 1. described in subsection 3.5.3. Since there were no significant differences between the trends in the data between subjects, the Thurstone-corrected data were averaged over subjects in the same way as in Experiment 1. The average result over 6 subjects is shown in Fig. 3.15. For both scenes the noisiness increases linearly with l_c at low values of l_c . At higher values of l_c , the noisiness saturates. The linear range is larger at higher η . The linear range is also scene-dependent. At $\eta = 7.07$, the linear range goes up to about $l_c \approx 2.2$ min arc for Wanda and up to about $l_c \approx 1.6$ min arc for Terrace. From the results of this experiment, it can be concluded that, at low values of l_c , i.e., $l_c < 1.6$ min arc, the noisiness is proportional to the correlation length l_c .

Figure 3.16 shows the plot of the data as a function of σ_n , the SD of the white noise by filtering of which the nonwhite noise has been obtained. This plot shows the difference between the scenes. For Wanda scene, the curves at $\eta = 5$ and $\eta = 7.07$ merge well whereas for Terrace scene, they remain separate. Again in this plot, at large values of σ_n , the noisiness saturates for both scenes. From these results, we concluded that, as a first approximation, at low values of correlation length l_c , the white noise SD σ_n is the predominant factor in determining the noisiness of images with nonwhite noise.

Huang (1965) studied the objectionability of the noise as a function of the width of the noise power density spectrum. He found that, at a constant noise power, when the noise bandwidth decreases from a large value, the objectionability of the noise first increases, reaches a maximum and then decreases.

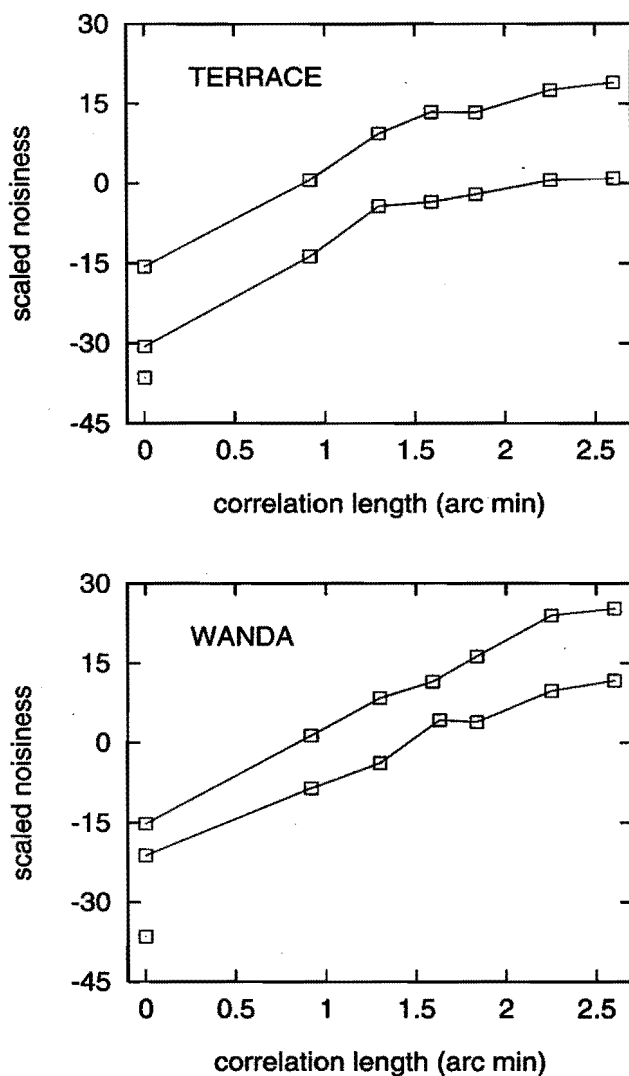


Figure 3.15: Results of Experiment 6, plotted as a function of the noise correlation length l_c . The parameter is the SD of the nonwhite noise η . The points joined by a line correspond to stimuli with identical noise η but different correlation length. For each scene, the upper curve corresponds to $\eta = 7.07$ and the lower to $\eta = 5$. The lowest points on the lower left of each plot show the result for the original images of the scenes.

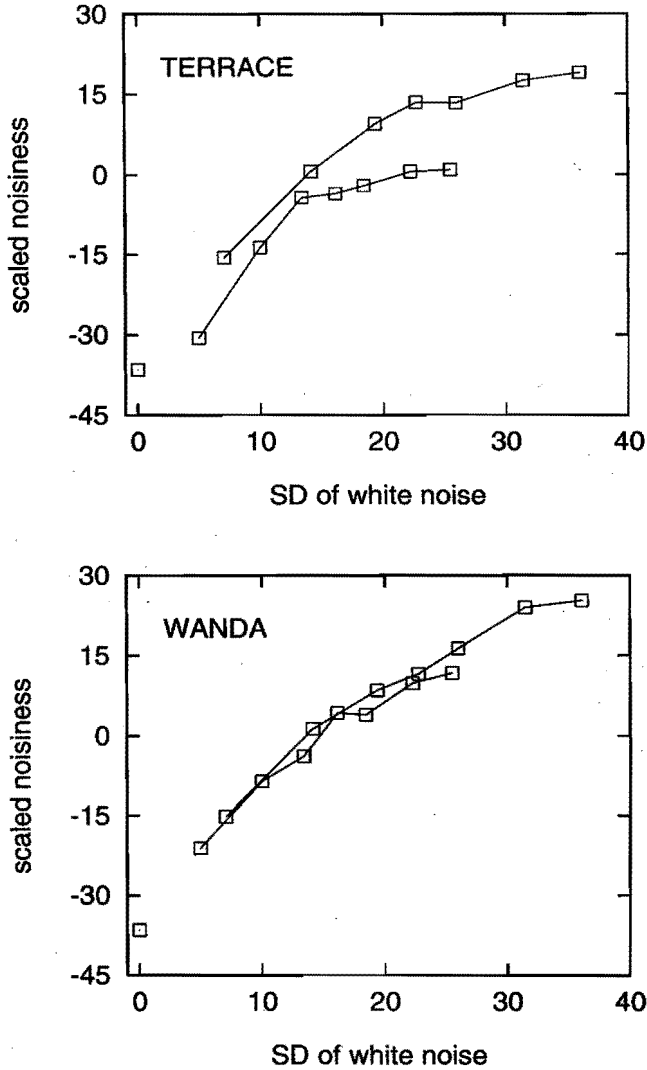


Figure 3.16: Results of Experiment 6, plotted as a function of σ_n , the SD of the white noise by filtering of which the non-white was obtained. The parameter is the SD of the nonwhite noise η . The points joined by a line correspond to stimuli with identical noise η but different correlation length. For each scene, the upper curve corresponds to $\eta = 7.07$ and the lower to $\eta = 5$.

Since the noise power density spectrum (and hence the noise bandwidth) and the auto-correlation function are related via the Fourier transform (Papoulis, 1991), Huang's findings can be reformulated as follows: when the correlation length l_c increases from a value close to zero, the objectionability of the noise first increases, reaches a maximum and then decreases. Our findings are in agreement with those of Huang. The range of l_c in our experiment is small. Our stimuli covered only the initial linear region and part of the saturation region. Stimuli with large correlation length were not included here because of their limited practical importance. Huang also found that his isopreference curves were scene-dependent.

3.6 Noise-index: A Measure for Noisiness

In this section we propose a measure for noisiness based on the results of the psychophysical experiments described in the previous sections. To be useful, the noise measure must satisfy two requirements. Firstly, it has to correlate well with the noisiness in natural images. Secondly, it has to be possible to estimate the measure from the noisy image using computational algorithms.

Using experiment 1 and 2, it was shown that the noisiness is independent of the noise PDF and depended only on the variance, for the ranges used. Using Experiment 4, it was shown that the local luminance of the uniform region also does not have a strong influence on the perceived noisiness, in a reasonably chosen range. Results of Experiment 6 showed that, for small values of l_c , the noisiness of nonwhite noise is mainly determined by the SD of the white noise by filtering of which the nonwhite noise is generated. Hence, the main factor that influences the noisiness in images is the SD of the white noise.

We are interested in deriving the relation between the (perceived) noisiness and the physical parameter noise SD. Nijenhuis (1993) derived an empirical relation for a measure for perceived blur using the spread of the blurring kernel. It is given by $S_b = 1 - 1/[1 + (\sigma_{bi}/\sigma_{b0})^2]^{1/4}$ where S_b is the blur-index, σ_{bi} is the spread of the image blurring kernel and σ_{b0} is the intrinsic blur in the early-visual pathway. Using this relation, the blur-index of an image can be estimated by estimating the spread of the blurring kernel from the image, as was done in Chapter 2. We fitted a similar empirical relation to the noisiness vs noise SD data obtained in Experiment 3. It is given by,

$$S_n = 1 - \frac{1}{[1 + (\sigma_{ni}/\sigma_{n0})^2]^{1/4}} \quad (3.20)$$

where S_n , similarly to the blur-index is called the noise-index ($0 \leq S_n \leq 1$), σ_{ni} is the SD of the noise in the image and σ_{n0} is the equivalent SD of the

intrinsic noise in the early-visual pathway. The fitted relations are shown by the dashed curves in Fig. 3.11. The best fitting σ_{n0} were 1.7 cd/m^2 for the scene with 20 cd/m^2 average luminance and 2.0 cd/m^2 for 30 cd/m^2 average luminance scene. Similar to blur-index S_b , equal intervals in S_n correspond to approximately equal perceived differences in noisiness. Hence, S_n is also a psychometric quantity (Hunt, 1978).

Figure 3.17, shows the perceived noisiness in Experiment 3 plotted against the noise-index S_n computed using Eq. 3.20. The linear fit in the figure shows that, S_n is a good correlate of noisiness in the range tested. Figure 3.18 shows the same for Experiment 1 and Experiment 4.

Based on the above findings we propose that the noise PDF and the local luminance do not significantly influence the noisiness, and that the noisiness is proportional to the noise-index, given by the Eq. 3.20. In the present model we are interested only in small values of l_c . For nonwhite noise with small values of l_c , Eq. 3.20 is also applicable, provided the SD σ_n of the white noise by filtering of which the nonwhite noise was generated is used as the relevant physical parameter.

The above findings are applicable to noisiness in the uniform regions. For image quality prediction, it is useful to have one measure for the entire image. In natural images, the noisiness of an image is decided mainly by the noisiness in the uniform regions of the image, since the visibility of noise is less at other parts of the image such as edges and lines due to masking (Girod, 1993). Hence, the noisiness of the entire image can be obtained by combining the noisiness over the uniform regions in the image. When the noise is stationary over the image, the noise SD and the correlation length are constant over the image. Hence the noise measure is obtained from the estimate of noise SD for the entire image. Therefore, Eq. (3.20) is assumed to hold for the entire image, where SD is estimated from the uniform regions of the image.

3.7 Estimation of Noise-index

Fig. 3.19 shows the schematic diagram of the algorithm to estimate noise-index. The initial stages in the algorithm are the same as those in the blur-index estimation algorithm. For a discussion on those stages we refer to Section 2.10. The values of the parameters of the optics stage and the (intrinsic) noise stages are also the same as those used in the blur-index estimation algorithm. The later steps are the same as those in the noise estimation algorithm, described in Section 3.3.1.A or 3.3.2, depending on whether input noise is white or non-white.

The noise-index was computed from the estimate of σ_n using the equation,

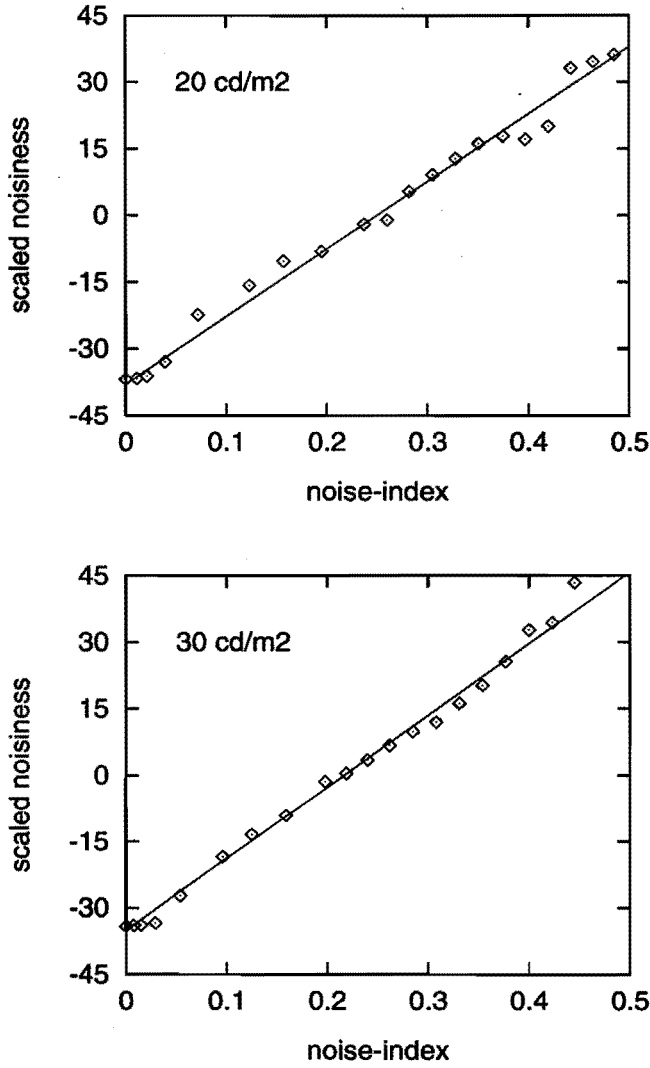


Figure 3.17: Results of Experiment 3 plotted as a function of the noise-index given by Eq. 3.20. The fit at $20\text{cd}/\text{m}^2$ was obtained with $\sigma_{n0} = 1.7$ and at $30\text{cd}/\text{m}^2$ with $\sigma_{n0} = 2$.

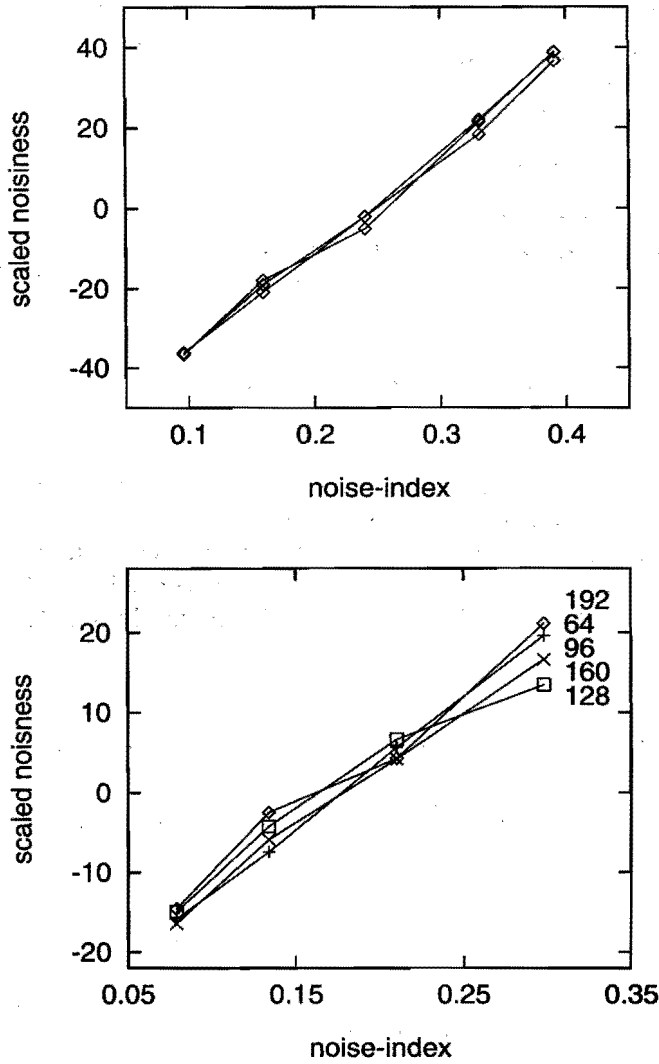


Figure 3.18: Results of Experiment 1 (top) and 4 (bottom) plotted as a function of the noise-index given by Eq. 3.20. The fit for Experiment 1 was obtained with $\sigma_{n0} = 3.2$ grey levels and for Experiment 4 with $\sigma_{n0} = 2.8$, grey levels. The parameter in the top plot is the noise PDF and the parameter in the bottom plot is the local grey value.

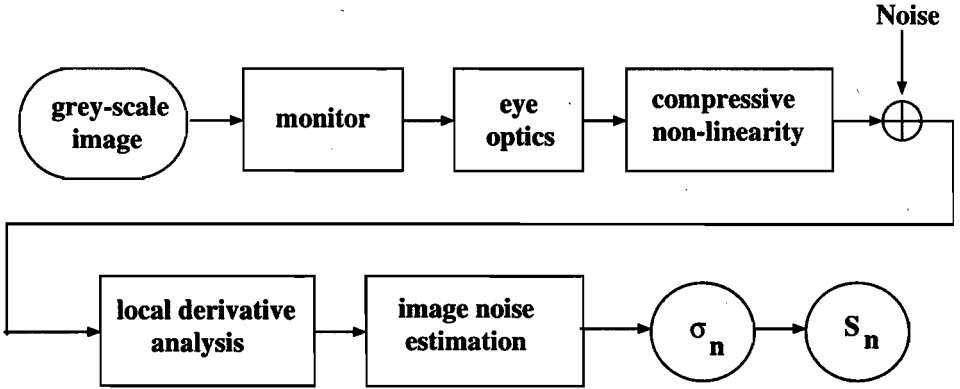


Figure 3.19: Noise-index estimation algorithm

$S_n = 1 - \sqrt{\sigma_{n0}/\hat{\sigma}_n}$, where $\hat{\sigma}_n$ is the estimate obtained using the algorithm (Figure 3.19) and σ_{n0} is the intrinsic noise of the visual system. This equation is obtained from Eq. (3.20) by noting that $\hat{\sigma}_n$ estimated in Figure 3.19 is the combined result of image noise σ_{ni} and the intrinsic noise of the visual system σ_{n0} , i.e., $\sigma_n^2 = \sigma_{ni}^2 + \sigma_{n0}^2$. We use $\sigma_{n0} = 3$ in the brightness domain which has a range of 0 to 100.

3.7.1 Results of Noise-index Estimation

The results of correlating the measured noisiness in psychophysical experiments 5 and 6 with the estimated noise-index will be reported here. The noise SD for each stimulus was estimated by applying the algorithm shown in Figure 3.19.

The results for Experiment 5 are given in Fig. 3.20. The coefficients of correlation between the estimated noise SD and the measured noisiness are 0.989 and 0.998 for Wanda and Terrace scenes respectively. The results for Experiment 6 are correlated with the noise-index in Fig. 3.21. In Experiment 6, the noise added to the images was nonwhite, i.e., generated by filtering white noise of SD σ_n with binomial filters. Therefore, the estimate of the white noise SD $\hat{\sigma}_n$ given by the algorithm (Fig. 3.19 and Section 3.3.2) is used while computing S_n . The coefficients of correlation between S_n and the measured noisiness are 0.98 and 0.99 for Wanda and Terrace scenes respectively. Note that although the nonwhite noise estimation algorithm overestimates the noise SD σ_n of natural images (Fig. 3.7), the estimates correlate rather well with the perceived noisiness. This indicates that the judgement of noisiness by subjects is also influenced by the image content or detail that has a similar structure to that of the noise. The overestimates of σ_n given by the nonwhite noise

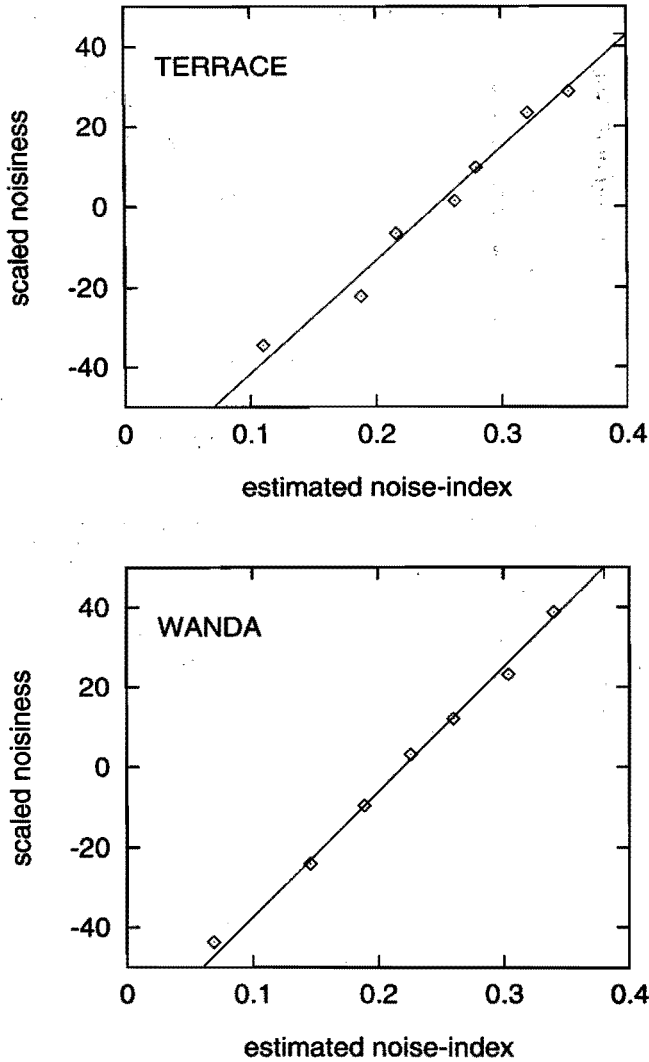


Figure 3.20: Results of Experiment 5 (see Fig. 3.14) plotted against the estimated noise-index, for natural images Wanda and Terrace. Each point is the average over 8 subjects. The analysis window used of estimation was binomial with $B = 4$.

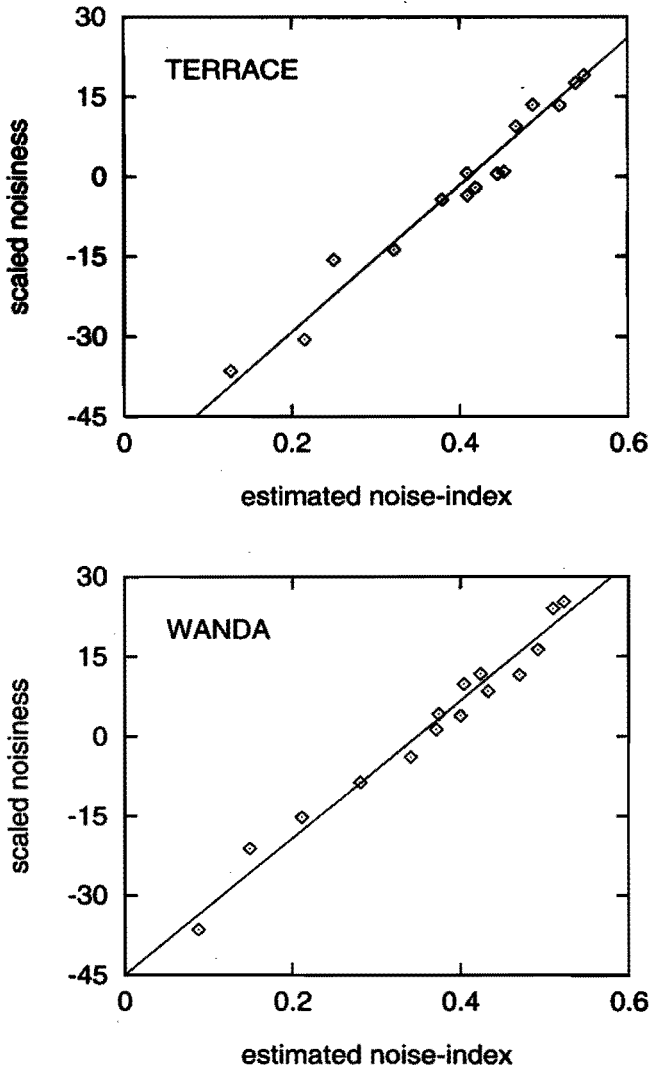


Figure 3.21: Results of Experiment 6 with nonwhite noise (see Fig. 3.16) plotted against the estimated noise-index, for natural images Wanda and Terrace scenes. Each point is the average over 8 subjects. The analysis windows used for estimation were binomial with $B_1 = 2$ and $B_2 = 8$.

estimation algorithm, discussed in Section (3.4), were also for the same reason.

These results show that the estimated noise-index S_n is a good correlate of noisiness of images, impaired with either white noise or nonwhite noise with low correlation length l_c .

3.8 Conclusions

An overview of the existing algorithms for estimating the noise variance is given. A new algorithm for estimating the variance of white noise, based on the first-order local energy, is presented. It has been shown that the new algorithm provides reliable estimates of noise variance in synthetic as well as natural images. The new method is compared with the method based on the overall local energy (Martens, 1994) and is shown to provide similar performance at approximately the same computational complexity. It is also shown to provide more reliable estimates than the algorithm based on the peak of the histogram of the gradient amplitude $|f_1|$ (Bracho & Sanderson, 1985). A new algorithm to estimate nonwhite noise with a Gaussian autocovariance function has been presented. By applying the algorithm on synthetic and natural images, it is shown that the algorithm provides reliable estimates except in some cases for natural images where the noise SD is overestimated.

To derive a measure for perceived noise, the perceptual effect of noise in images is studied using psychophysical experiments. It is shown that the form of the PDF of the noise is not significant for the perceived noise. It is also shown that the local luminance of the uniform regions in the image does not significantly influence the noisiness. The two important parameters that influence the noisiness are the SD and the correlation length. Based on the psychophysical experiments, a measure for perceived noise, called noise-index, is proposed. The analytical equation for the noise-index is similar to that for the blur-index. In the case of white noise, the noise-index is determined the noise SD. In the case of nonwhite noise, noise-index is determined by the SD of the white noise by filtering of which the nonwhite noise has been obtained. An algorithm to estimate the noise-index from an image has been presented. It is shown that the estimates of noise-index given by the algorithm correlate well with the image noisiness, measured using the psychophysical experiments, both for white noise as well as for nonwhite noise.

Here we address possible improvements upon some issues discussed in this chapter. Similar to the curve fitting method based on the first-order energy E_1 , curve fitting to the histogram of the gradient amplitude $|f_1|$ can be used to estimate noise variance. It is expected that such a curve fitting method on $|f_1|$ will provide more reliable estimates than the peak-based method, since it

pools information from a larger region of the histogram. The nonwhite noise estimation algorithm can be extended to a multi-scale algorithm, where the scales can be chosen *a posteriori*. In the experiment on the effect of the PDF on noisiness, we have considered only three forms of PDFs. However, other PDFs such as Poisson may also be of importance in images, and may have to be investigated. In our experiment with nonwhite noise as well as that of Huang (Huang, 1965), some scene dependence has been observed. Further experiments will be needed to investigate this issue. We have studied the effect of noise PDF, local luminance, noise SD, and correlation length on the noisiness of images. An important parameter that is known to influence the image noisiness is the average luminance of the image (Marmolin & Carlström, 1985). Another parameter that influences the image noisiness is the γ of the display (Girod, 1989; van Overveld, 1994). Presently no attempt has been made to incorporate the effects of these global image parameters into the noisiness measure.

Chapter 4

Subjective Characterization of Images Degraded by Both Blur and Noise

Abstract

The notion of a perceptual space spanned by a number of basic perceptual image attributes is very useful to subjectively characterize images. In this chapter, images degraded by blur and noise are subjectively characterized by locating them in a multi-dimensional perceptual space. The perceptual space is constructed using multi-dimensional scaling (MDS) techniques. The data collected in a series of psychophysical experiments, involving images of one synthetic scene and two natural scenes, is used as input to the MDS techniques. Two different MDS approaches, one making use of the perceived dissimilarity between the images and another making use of the scaled perceptual attributes of the images, are used. In the latter approach, three important perceptual attributes, i.e., unsharpness, noisiness and quality are considered. It is shown that both approaches lead to very similar stimulus configurations in the perceptual space. For each of the three scenes used, a 2-D perceptual space is shown to adequately represent the images. The perceptual spaces obtained for all the three scenes considered are very similar. Using the directions of the attribute vectors in the perceptual space, it is shown that the attributes unsharpness and noisiness are not orthogonal, but are separated by an angle of approximately 80° between them, indicating a weak interaction between the two. It is also shown that the impairment vector (the direction of which is opposite to the direction of quality vector) lies between the unsharpness and noisiness vectors, with an angle of approximately 33° between unsharpness and impairment vectors.

4.1 Introduction

Perceptual image quality research has attracted considerable attention in the past decade. Image quality is a global attribute of images that is affected by many impairments such as blur, noise, etc. Images may be simultaneously impaired by several impairments. Most approaches to understanding and predicting image quality treat image quality as a unidimensional global attribute, that is directly determined by the physical image. Although such approaches are successful when images are degraded by one impairment, they fail when images are multiply impaired. When images are multiply impaired, it is useful both for analyzing and for modeling image quality to consider image quality as being determined by several perceptual attributes such as sharpness, noisiness, brightness contrast, etc. This has led to the multi-dimensional approach to image quality, where images are treated as being located in a multi-dimensional perceptual space, spanned by the basic perceptual attributes of the images. The task of locating the images (stimuli) in such a multi-dimensional space is carried out by multi-dimensional scaling (MDS) techniques (Green, Carmone & Smith, 1989; Kruskal & Wish, 1978). The input to the MDS methods are data of psychophysical experiments. The MDS method consists of a series of mathematical algorithms to determine the location of the stimuli in the multi-dimensional space from this input. The problem is similar to locating cities in a country (i.e., constructing the map), given the distances between all the cities (Kruskal & Wish, 1978).

The MDS space can be derived from dissimilarity (or similarity) judgements between stimuli or from judgements of several attributes for each stimulus (Jackson, 1978; Schiffman, Reynolds & Young, 1981). Depending on the nature of the distance measure used to determine the position of the stimuli in the space, the MDS techniques are divided into two classes: metric and non-metric. The metric MDS methods assume that the perceived dissimilarities can be represented on a ratio scale (ratio-scaled data) and hence that the distances between the stimuli in the MDS space are proportional to the dissimilarity judgements. The non-metric methods assume that the perceived dissimilarities can be represented only on an ordinal scale (ordinal-scaled data) and hence that there is only a monotonic relation between dissimilarities and the distances in the space. In most metric methods, Euclidean distances are used, although some methods also allow for general Minkowski distances to be used. In this chapter, only metric MDS methods with Euclidean distances are used. MDS has been used in many areas of research that deal with the psychophysical evaluation of stimuli varying in multiple aspects (Green *et al.*, 1989). Multi-dimensional scaling can also be used to test the dimensional orthogonality of perceptual attributes (Ashby & Townsend, 1986).

Marmolin & Nyberg (1975) used MDS techniques to understand image

quality in a multidimensional framework. They used the data of psychophysically evaluated impaired images as input to the MDS method. The physical variables were the modulation transfer function (MTF) of the system, the noise power spectrum, the noise level and the gamma value (Poynton, 1993). By visually inspecting the location of the images in the MDS space, they attributed sharpness, noise and contrast to 3 dimensions of their 4-dimensional space. The fourth unidentified dimension was found to be scene-dependent. Goodman & Pearson (1979) studied the MDS space of television images impaired by DPCM-quantized errors, echo, noise and bandwidth limitation. They also found a 4-dimensional space and its dimensions were associated with sophisticated names that describe concepts that are similar to visibility of echo, noise, sharpness and visibility of objects. In both the studies mentioned above, the dimensions were labelled on the basis of visual inspection of the stimulus locations in the MDS space. Escalante-Ramírez (1992) and Escalante-Ramírez, Martens & de Ridder (1994), studied the perceptual quality of noise-reduced CT images using MDS techniques. In addition to obtaining an MDS space using the dissimilarity data, they also obtained an MDS space using the scaling data of the most relevant attributes: general preference, annoyance of noise, sharpness and visibility of structures. Although both approaches resulted in similar configurations, the latter approach helped to rigorously identify the dimensions of the perceptual space. A 2-D space was appropriate to represent the images. The first dimension was associated with the annoyance of noise. The second dimension could not be associated with one specific attribute, but it was found to correlate with sharpness and visibility of structures. This method of labelling dimensions is often more reliable than the method based only on the visual inspection of stimuli in the MDS space. It can also be used to draw stronger conclusions from the MDS space, for example about dimensional orthogonality.

Linde (1981) studied the interaction between blur and noise using similarity data. The data analysis as well as the results of MDS showed interaction between the two percepts. Physically constant noise standard deviation (SD) intervals increased perceptually with increasing blur and physically constant blur intervals decreased perceptually with increasing noise. However, the extent of the interaction could not be quantified from their analysis. The interaction was also found to be scene-dependent.

In order to relate perceptual image quality to its attributes, it is essential to understand their interaction and how they combine to give the overall impression of quality. The aim of this chapter is to perceptually characterize images degraded by blur and noise, by locating them in a multidimensional perceptual space. By locating the images in the MDS space and by identifying the directions of the important perceptual attributes in the space, the inter-relation between the attributes can be understood. Specifically, the aim

is to test whether the attributes unsharpness and noisiness are dimensionally orthogonal (Ashby & Townsend, 1986). Knowledge about their interaction is essential in order to predict image quality from the estimates of these individual attributes.

The MDS techniques will be used to locate the images in a perceptual space spanned by the basic attributes. Dissimilarity data collected in psychophysical experiments will be used for this purpose. The underlying perceptual dimensions of the MDS space will be identified by using an MDS technique based on the psychophysically scaled image attributes. The data for this purpose will be collected using numerical category scaling of important perceptual attributes and quality of images.

This chapter is organized as follows. The experimental setup is described in section 4.2. The results of analyzing the dissimilarity data using MDS techniques are reported in section 4.3. The results of scaling attributes and its analysis using MDS is also reported in section 4.3. In section 4.4, the implications of the results obtained are discussed.

4.2 Experimental Setup

Stimuli: Three scenes were used in this experiment: two natural scenes Wanda and Terrace and a synthetic scene Mondrian. The original images of these scenes are shown in Fig. 2.12. The stimuli were derived from the original images by first blurring the images with a binomial window of variable order B and then adding zero mean Gaussian-distributed white noise of given SD σ_n to every pixel of the image. A binomial window of order B is approximately equal to a Gaussian window of spread $\sigma_b \approx \sqrt{B/2}$ (Martens, 1990b). Four levels of σ_b corresponding to binomial filters of order $B = 0, 2, 4$, and 8 were used. Four levels of σ_n were used: $0, 7, 10$ and 14 . The pixel value of noisy images was truncated to be between zero and 255 . However, because the noise SDs were small, the grey value of the pixels of noisy images rarely exceeded 255 or fell below 0 . All combinations of the four levels of blur and four levels of noise were considered to obtain a total of 16 stimuli per scene. Although the original images were 512×512 pixels in size, only a region of interest of 240×470 pixels (0.13 m by 0.25 m or 5° by 9.5°) was displayed in the experiment. This enabled simultaneous display of two images side by side on the monitor, one on the left and the other on the right. The images were displayed on a BARCO-CCID-7351B high-resolution non-interlace monitor using a VME-bus based system (with MEN-A301 graphics board). The display characteristics were calibrated so as to have a $\gamma = 2.5$, a maximum luminance $L_{max} = 60 \text{ cd/m}^2$ and a minimum luminance $L_{min} = 0.2 \text{ cd/m}^2$ (See Eq. (2.46)). The viewing conditions satisfied the CCIR recommendation 500 (CCIR, 1986). The viewing

distance was 1.5 m, which corresponds to a height/distance (h/d) ratio of 6. Images were displayed for a fixed duration of 5 seconds, with at least a 2 seconds adaptation period between two presentations. The adaptation field was uniform and had a luminance of 9 cd/m^2 , which is approximately equal to the average of the mean luminance of the original images of all scenes.

Method: There were two experimental sessions for each scene. In the first session, dissimilarity scores for all pairs of stimuli were assembled. All combinations of 16 stimuli of the same scene were considered, to obtain a total of 256 pairs of stimuli per scene. During the experiment, each pair of stimuli was displayed on the monitor in random order, with one image on the left side of the monitor and the other on the right. The subjects were instructed to rate the dissimilarity between the two stimuli using an integer score ranging from 0 to 10. Subjects were urged to base their score only on how dissimilar or different they perceive the images to be and not to base their score on any preference, quality or emotional criteria. A score of 10 was used to indicate the largest dissimilarity and a score of 0 implied no perceived difference. Before the start of a session, subjects took part in a trial session with 15 presentations covering the entire range, in order to adjust the sensitivity of their scale.

In the second session, attributes of the images were judged using a numerical category scaling procedure (de Ridder & Majoor, 1990). There were three subsessions, each evaluating a different attribute of the image: unsharpness, noisiness and quality. The 16 stimuli were repeated four times in each subsession. The order of presentation was random. The subjects were instructed to judge the attributes of the image using an 11-point numerical category scale, ranging from 0 to 10. The stronger the judged attribute, the higher the score. Before the start of the session, subjects took part in a trial session with 12 stimuli covering the entire range, in order to adjust the sensitivity of their scale.

Subjects: Five subjects between 25 and 39 years of age took part in the first session of the experiment. Two additional subjects, from the same age group, took part in the second session, making a total of 7 subjects. All subjects were familiar with numerical category scaling and the concepts of image quality, unsharpness and noisiness. All had normal or corrected-to-normal visual acuity between 1.5 and 2, measured on a Landolt chart.

4.3 Results

4.3.1 MDS Analysis of Dissimilarities

The first session resulted in a dissimilarity matrix of size 16×16 for each subject and each scene. The entry $\delta_{i,j}$ at position (i, j) in the matrix is the judged

dissimilarity score between stimuli i and j . The upper and the lower triangular matrices were averaged in order to obtain a single distance between stimuli i and j , i.e., $\overline{\delta_{i,j}} = (\delta_{i,j} + \delta_{j,i})/2$. The resulting lower triangular matrix (without the diagonal) was used for further processing. These data were used to study the distribution of the stimuli in a multidimensional space.

Two MDS techniques were used to analyze the data: SINDSCAL and MULTISCALE. SINDSCAL is a computer program that implements the individual difference model for MDS (Carroll & Chang, 1970). This model assumes that all subjects share a common perceptual space and that the individual subjects differ in the weight that they attribute to the different orthogonal dimensions. Data for SINDSCAL involves several dissimilarity matrices, one for each subject. SINDSCAL is a metric MDS method. A metric method arrives at a configuration of the stimuli in a given dimensional space, such that the distances between the stimuli in that space are proportional (as closely as possible) to the dissimilarity data between them. In other words, a metric method assumes that the input data are ratio-scaled dissimilarities (Kruskal & Wish, 1978). The goodness-of-fit for a SINDSCAL solution is indicated by the variance accounted for (VAF) by the solution or the 'loss' of variance ($= 1 - \text{VAF}$). The dimensionality of the SINDSCAL solution is decided on the basis of how the VAF or loss changes with the number of dimensions. The output of SINDSCAL consists of two spaces, a group stimulus space and a subject weight space. The group stimulus space is a common psychological space for all subjects, with N points, corresponding to the N stimuli. The subject weight space shows the weight each subject assigns to the dimensions. It contains M points, corresponding to the M subjects. The individual psychological space for each subject is obtained by scaling the group stimulus space by the subject weight factor along each dimension, derived from the subject weight space.

MULTISCALE is based on the maximum likelihood (ML) estimation principle (Ramsay, 1977). The main advantage of MULTISCALE is that it can be used as an exploratory as well as confirmatory data analysis tool. MULTISCALE makes explicit assumptions about the distribution of errors and about the transformation of dissimilarity. Many statistical tests of significance on the data and on the MDS solution can be performed by varying these assumptions in MULTISCALE. For example, the chi square test of significance can be used to decide the dimensionality of the MDS solution for a given data set. The confidence regions indicating the relative precision of the stimulus coordinates in the MDS solution map are very useful in confirmatory analysis. It has been shown that MULTISCALE provides very reliable results, often better than SINDSCAL, especially for noisy data (Ramsay, 1977). In a comparative study of MDS methods, MULTISCALE has also been shown to produce the best results among the metric methods tested (Spence & Lewandowsky, 1989).

The metric of the Cartesian coordinate system can be specified in MULTISCALE: identity, diagonal or full metric. In the identity metric case, all dimensions have equal weight, hence the solution is rotation invariant. When the metric is diagonal, each orthogonal dimension can have a different weight, and the solution is not rotation invariant. This case is similar to that of SINDSCAL, in the sense that it results in a common perceptual space for all subjects and a subject weight space. The full metric case is the most general one, where the angles between the dimensions is also allowed to vary. MULTISCALE allows for different assumptions about the transformation of dissimilarities into distances, the distribution of errors, and the variance component model for subjects and stimuli. The explicit assumptions can be specified by the user. We assumed a power transformation between the scaled dissimilarities and the distances. This transformation captures some of the non-linearity observed between dissimilarities and distances, such as Weber's law. The distribution of errors was assumed to be log-normal. According to this assumption, larger judgements produce larger variability (errors) and smaller judgements produce smaller variability, which reflects a very general finding in psychophysics. The stimulus variance was assumed to be equal for all stimuli and the subject variance was allowed to be subject-dependent. When this assumption is used, MULTISCALE provides within-subject standard error estimates. This is an index for the relative lack of precision of each subject's data in comparison to the estimated model.

Both SINDSCAL and MULTISCALE locate the stimuli in a multi-dimensional space. To interpret the solution of the MDS, *the perceptual meaning of the dimensions has to be understood*. This can not be done from the MDS solution alone. In the next section, it will be shown how direct scaling of perceptual attributes of the stimuli can be used in interpreting the dimensions.

A separate SINDSCAL analysis was performed for each scene. The dissimilarity matrices of all subjects for the given scene were input to SINDSCAL. For each scene, several runs were looked at, each with a different starting configuration, and the configuration with the least number of dimensions that adequately accounted for the variance was retained. Configurations with dimensions 1 up to 5 were tried for each scene. The loss as a function of the dimension of the configuration is shown in Fig. 4.1. This plot shows that, for all scenes, by going from a 1-D solution to a 2-D solution, there is a substantial decrease in the loss, whereas by going from a 2-D solution to a 3-D solution, the decrease in loss is relatively less. The same is also true for other dimensions larger than 2. Thus, from the point of view of accounted variance, the 2-D configuration gives an adequate fit for all three scenes. Examining the 3-D and the higher dimensional solutions did not give any additional insight into the configuration, since there were no systematic trends in the configuration in those dimensions. Based on these observations, we conclude that a 2-D configuration gives an

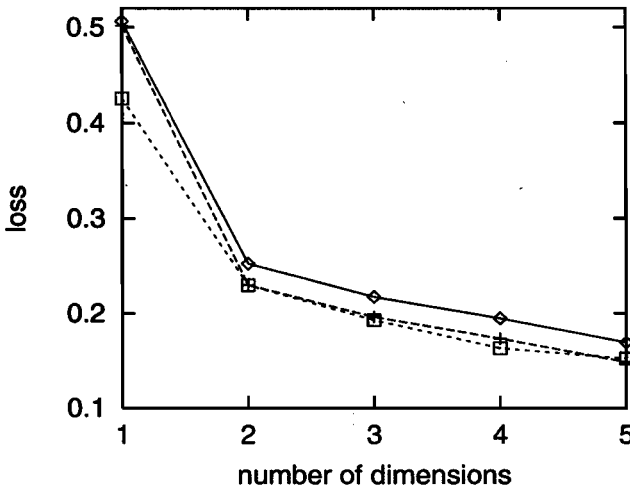


Figure 4.1: Loss as a function of dimension of the configuration, of analyzing the dissimilarity data using SINDSCAL, for the Mondrian (◇), Terrace (+) and Wanda (□) scenes.

adequate fit for all three scenes.

Figure 4.2 shows the denormalized two-dimensional group-stimulus-spaces obtained using SINDSCAL. The variance explained by the first and the second dimension is, 49 % and 26 %, 41 % and 33 %, and 52 % and 26 % , for the Mondrian, Terrace and Wanda scenes respectively. For the 2-D configuration, the correlation between the scores of the subject and the computed scores varied from 0.82 to 0.91 (mean 0.87), from 0.84 to 0.91, (mean 0.86) and from 0.80 to 0.93 (mean 0.88), over subjects, for the Mondrian, Terrace and Wanda scenes respectively. Figure 4.3 shows the subject-weight-spaces corresponding to the stimulus-spaces given in Fig. 4.2. In the subject-weight-space, with the exception of subject 4 for the Mondrian scene, all subjects are located close to each other, and hence could be considered to belong to one group. The average subject weight for the first and the second dimension were, 0.7 and 0.5, for Mondrian, 0.64, and 0.58 for Terrace, and 0.71 and 0.5 for Wanda respectively. The denormalized group stimulus spaces shown in Fig. 4.2 were obtained by scaling the normalized solution given by SINDSCAL with the square root of the average subject weights, along both dimensions. In Fig. 4.2, the MDS configurations for all three scenes are very similar, indicating no strong scene dependence. There are however, some small differences, especially between the synthetic scene Mondrian and the natural scenes, Terrace and Wanda. A better picture of these differences will emerge from the MULTISCALE solution

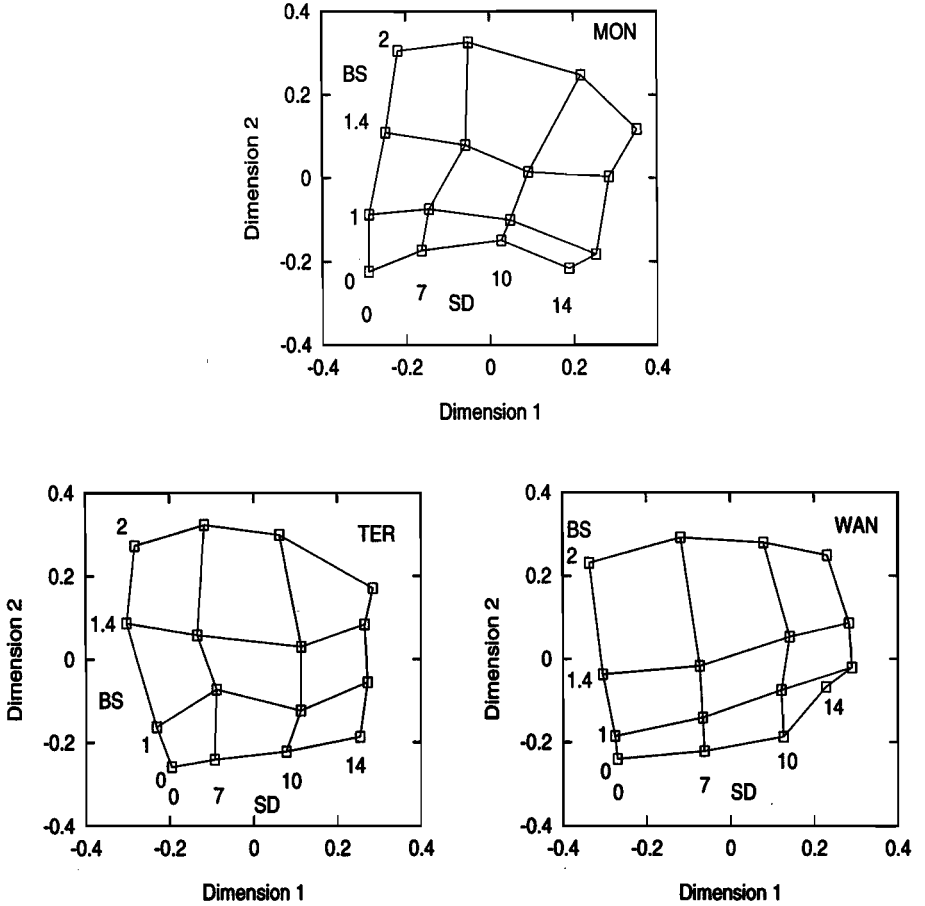


Figure 4.2: Results of analyzing the dissimilarity data using SINDSCAL. Denormalized stimulus spaces for the Mondrian, Terrace and Wanda scenes. The stimulus parameters, the noise standard deviation σ_n (SD) and the spread of the blurring kernel σ_b (BS), are also shown. The corresponding subject-weight-spaces are shown in Fig. 4.3.

(Fig. 4.4) and hence a discussion about this is deferred until we discuss the MULTISCALE solution.

A separate MULTISCALE analysis was also performed for each scene. The dissimilarity matrices of all subjects for the given scene were input to MULTISCALE. Configurations with dimensions 1, 2, 3 and 4 were tried with the identity as well as with the diagonal metric of the Cartesian coordinate sys-

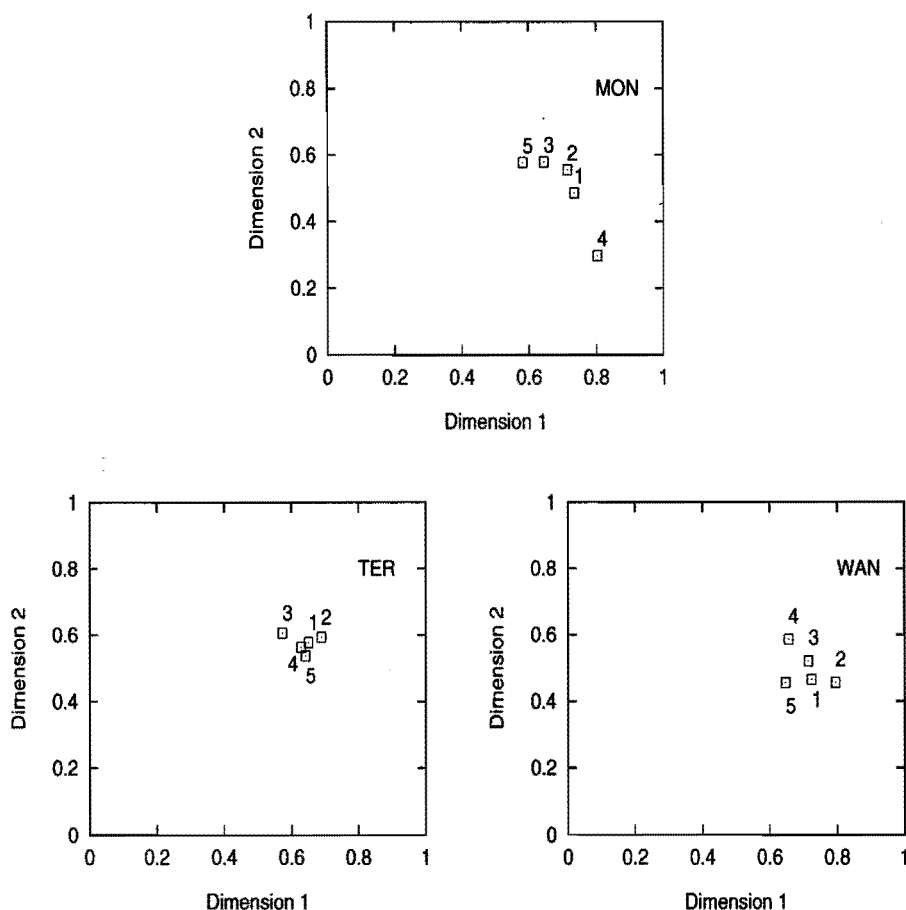


Figure 4.3: Results of analyzing the dissimilarity data using SINDSCAL. Subject weight spaces, for five subjects: VK(1), TV(2), TR(3), JV(4), and HR(5), for the Mondrian, Terrace and Wanda scenes.

tem. Using an asymptotic chi-square goodness of fit, for the identity-metric case and diagonal-metric case, a 2-D configuration with identity-metric was found to provide an adequate fit for all three scenes. The 2-D solution with identity-metric was considered adequate over the diagonal-metric case, mainly because the identity metric case has a lower number of parameters than the diagonal-metric case. Also the diagonal-metric solution was very similar to that obtained using the identity-metric, and the subject weights for all subjects were approximately equal. For the 2-D solution, the within-subject multiple correlation varied from 0.77 to 0.91 (mean 0.85), from 0.77 to 0.88 (mean 0.85), and

from 0.79 to 0.92 (mean 0.86), for the Mondrian, Terrace and Wanda scenes respectively. The 2-D solutions are shown in Fig. 4.4. Note that for the identity metric case, the configuration is rotation invariant. MULTISCALE chooses the direction of the first dimension such that it is along the direction where the points vary most, the second perpendicular to it, etc. Hence, although the configuration for scene Terrace is not aligned in the same way as that for the other two scenes, it can be rotated. The MDS configurations in Fig. 4.4 for all three scenes are very similar. There are however some small differences between the synthetic scene Mondrian and the natural scenes. In the case of Mondrian, it can be seen that the physically constant noise standard deviation (SD) intervals increase perceptually with increasing blur and physically constant blur-spread intervals decrease perceptually with increasing noise. This is similar to the findings of Linde (1981). However, this is much less pronounced in the case of the natural images Terrace and Wanda. Similar observations can also be made about the SINDSCAL solutions, although in the case of the SINDSCAL solution, the scene differences are relatively less prominent.

The 95 % confidence regions for the stimulus locations generated by MULTISCALE are shown in Fig. 4.5. For all scenes, the stimulus with $B=4$ and $SD=10$ has a relatively large confidence region, indicating a relatively small precision in the estimate of its location. The confidence regions for all other images are almost equal in size, for all scenes. The approximate circular shape of the confidence regions indicate that the precision in the stimulus locations is equal in both dimensions of the MDS space.

Comparing the configurations obtained via SINDSCAL and MULTISCALE shows that the main differences are in the positions of the corner stimuli. Similar trends can also be seen from the results of Monte Carlo simulations (Ramsay, 1977), where noisy data has been analyzed using both SINDSCAL and MULTISCALE. In their study, comparing the stimulus space obtained using SINDSCAL with the ideal configuration showed that larger errors occur in the location of the stimuli that are located away from the centre of the configuration. We compared the configuration obtained using SINDSCAL with that obtained with MULTISCALE. The correlation was done using MULTISCALE, where rotation of the configurations was allowed to obtain the best match. The denormalized solution of SINDSCAL (Fig. 4.2) was used for correlation with the MULTISCALE configuration (Fig. 4.4). The correlation coefficients between the configurations were 0.992, 0.985 and 0.986 for the Mondrian, Terrace and Wanda scenes respectively. The high correlation coefficients indicate that the configurations obtained using SINDSCAL and MULTISCALE are very similar.

The stimuli in the experiment varied in two aspects, viz, the spread of the blurring kernel σ_b and the SD of the noise σ_n . MDS analysis of the dissimilarity

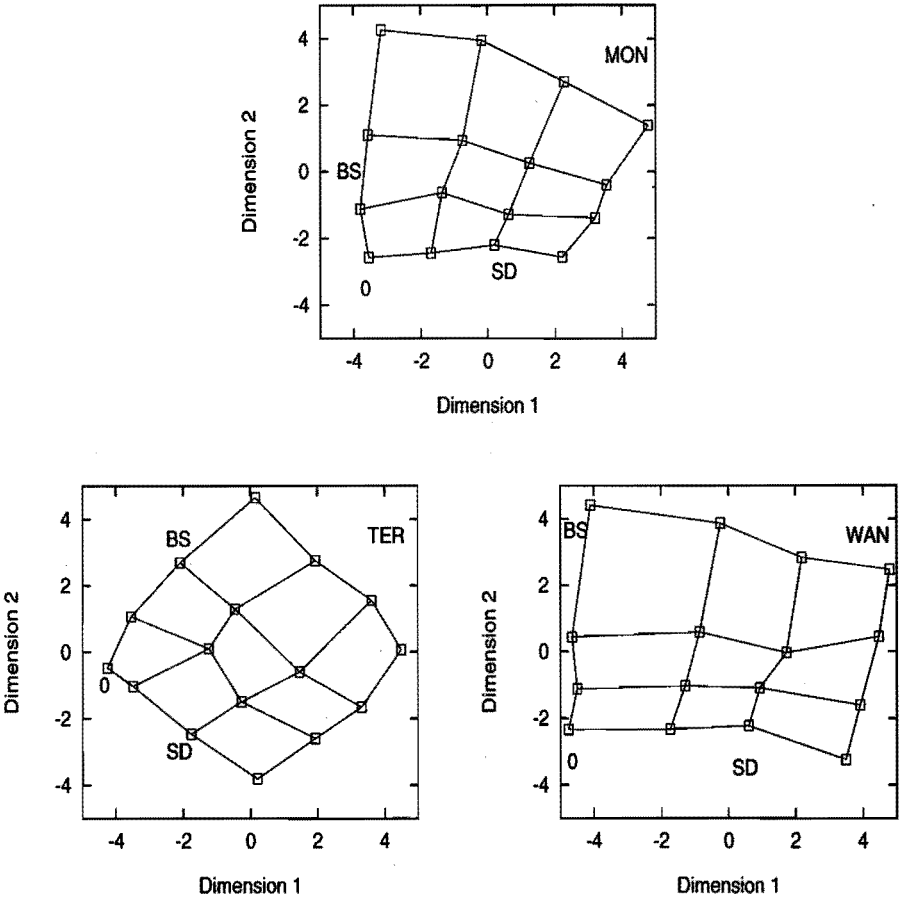


Figure 4.4: Results of analyzing the dissimilarity data using MULTISCALE, for the Mondrian, Terrace and Wanda scenes. The directions along which the stimulus parameters, the noise standard deviation σ_n (SD) and the spread of the blurring kernel σ_b (BS) increase are also shown.

data showed that a 2-D solution space is adequate to represent the images. This shows that, varying the two physical parameters of the stimuli σ_b and σ_n has led to variations in only two perceptual dimensions. Although at a first glance this may appear to be obvious, i.e., variation in one physical parameter leading to variation in (only) one perceptual dimension, it need not be true in general. Changing one physical parameter may sometimes lead to changes in more than one perceptual dimension. For example, changing the gamma of an imaging system can cause changes in the global brightness contrast as well as changes

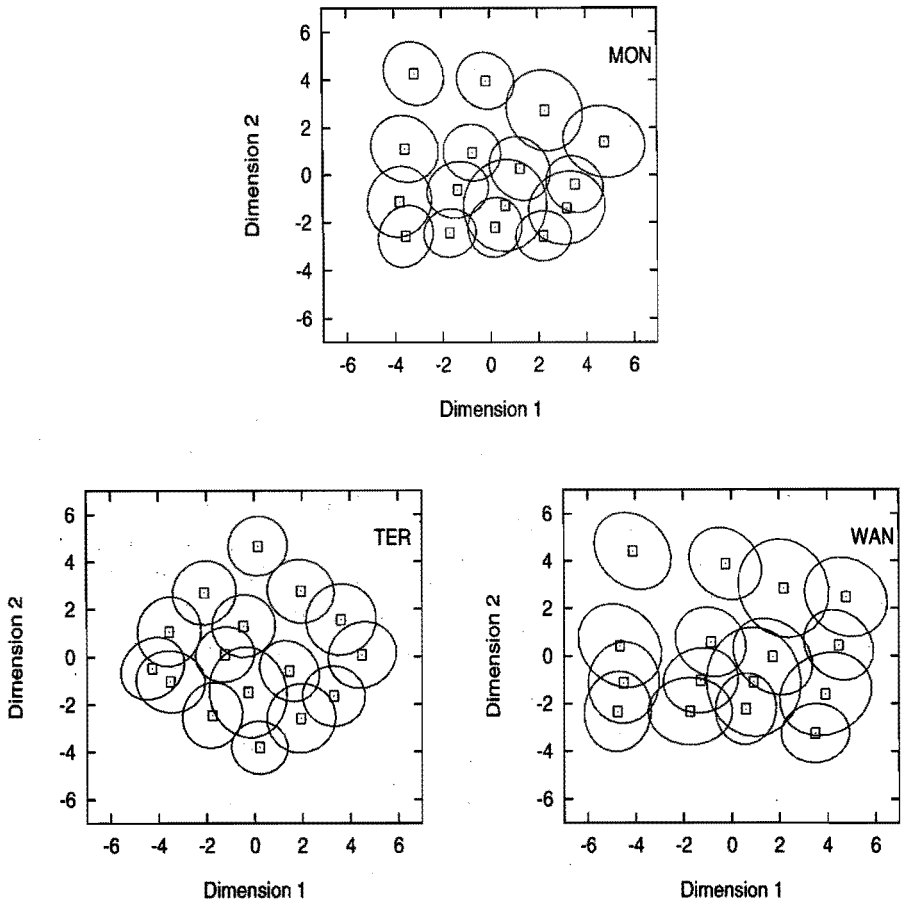


Figure 4.5: Confidence regions of estimated stimulus coordinates using MULTISCALE, for the Mondrian, Terrace and Wanda scenes. The stimulus coordinates shown using (□) are the same as those shown in Fig. 4.4.

in the perceived sharpness of images (Roufs, 1989).

4.3.2 Results of Scaling Attributes

In the second session, the scaling of three attributes, unsharpness, noisiness and quality resulted in numerical category data on an 11-point scale. For each scene, the 16 stimuli were judged 4 times, for all 3 attributes, by each of the 7 subjects. The numerical category data from the experiment were transformed

into an interval scale on a psychologically linear continuum using Thurstone's law of categorical judgement, in the same way as was earlier described in Session 2.9 of Chapter 2.

Since there were no significant differences between the trends in the data of the subjects, the Thurstone transformed data were averaged over subjects. Before averaging over subjects, the Thurstone transformed data for each subject were normalized using the z-score transform (Hays, 1988). Finally, the z-scores averaged over the subjects were normalized so that the resulting average standard error of mean (SEM) was equal to one. The average results for 7 subjects are shown in Figures 4.6, 4.7, 4.8, 4.9, 4.10 and 4.11.

Fig. 4.6 shows the results of scaling unsharpness as a function of blur spread, σ_b . Points joined by a drawn curve contain stimuli with the same input noise SD but different blur spread. The four different curves correspond to noise SD 0, 7, 10 and 14. From this figure it can be seen that the perceived unsharpness increases almost linearly with blur spread for all noise levels. The same results have been replotted in Fig. 4.7 as a function of noise SD σ_n . This figure shows that the perceived unsharpness is dependent on the noise SD. At high blur spread perceived unsharpness decreases with noise SD, at low blur spread perceived unsharpness increases with noise SD, and at intermediate noise SD perceived unsharpness is independent of noise SD. In other words, adding noise to sharp images makes them appear blurred, whereas adding noise to very blurred images makes them appear slightly less blurred.

Fig. 4.8 shows the results of scaling noisiness as a function of noise SD, σ_n . Points joined by a drawn curve contain images with the same blur spread but different input noise SD. The four different curves correspond to binomial filters of order $B = 0, 2, 4$, and 8 , which are approximately equivalent to $\sigma_b \approx 0, 1, 1.4$ and 2 pixels. From this figure it can be seen that the perceived noisiness increases almost linearly with noise SD for all blur spread levels. The same results have been replotted in Fig. 4.9 as a function of the blur spread σ_b . This figure shows that at low values of the noise SD, the noisiness is almost independent of blur spread and at high values of the noise SD the noisiness increases with blur spread. In other words, blurred images with noise always appear more noisy than sharp images with noise, especially for large amounts of noise.

The results of scaling quality are shown in Fig. 4.10 and 4.11, as a function of blur spread, σ_b and noise SD, σ_n , respectively. From these figures, it can be seen that the quality decreases with both σ_b and σ_n , due to blur and noise. In Fig. 4.10, points joined by a drawn curve contain images with the same input noise SD but different blur spread. In Fig. 4.10 the slope of the curves decreases with increasing noise SD σ_n . Similarly, in Fig. 4.11, the slope of the curves decreases with increasing blur spread σ_b .

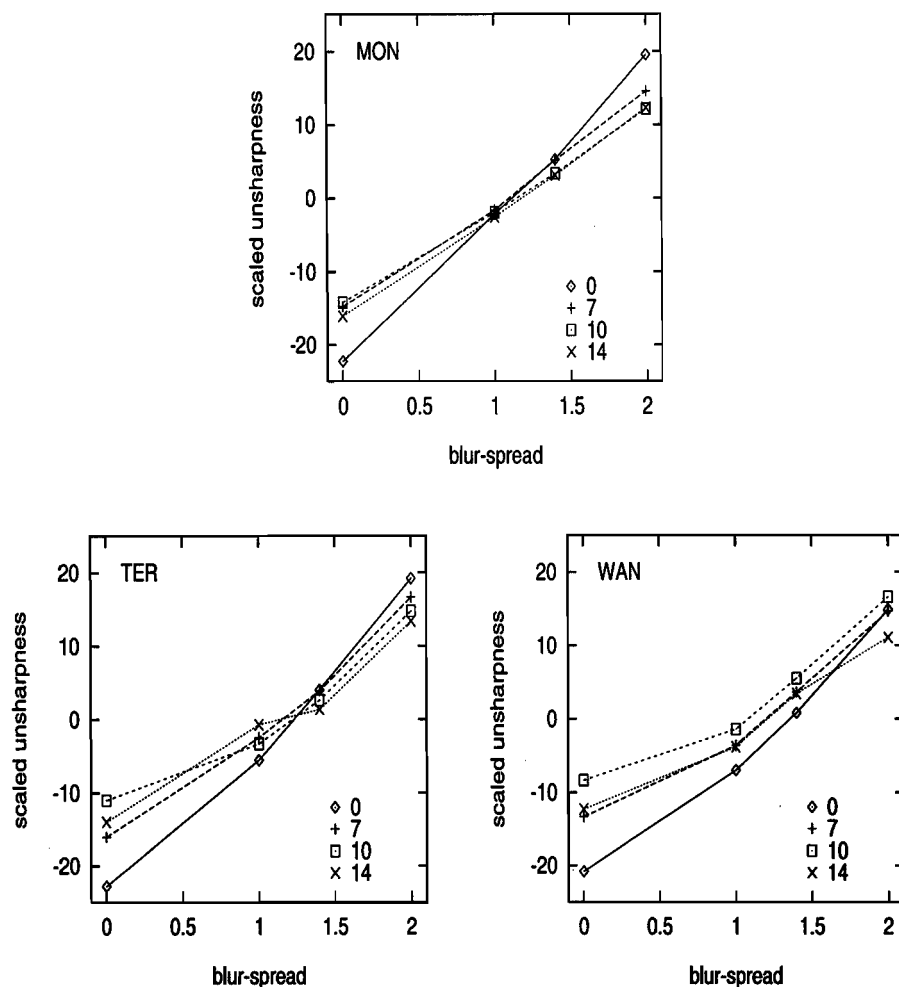


Figure 4.6: Results of scaling unsharpness, plotted as a function of the spread of the blurring kernel σ_b , with noise SD σ_n as the parameter, for the Mondrian, Terrace and Wanda scenes.

4.3.3 MDS Analysis of Scaled Attributes

In the previous subsection we looked at the mean scores of the scaled attributes. These data can also be used to locate the stimuli in a multi-dimensional perceptual space, similar to the way in which MULTISCALE and SINDSCAL were used for the dissimilarity data. The perceptual space that will be arrived at using the scaled attributes and the one arrived at earlier using dissimilarities

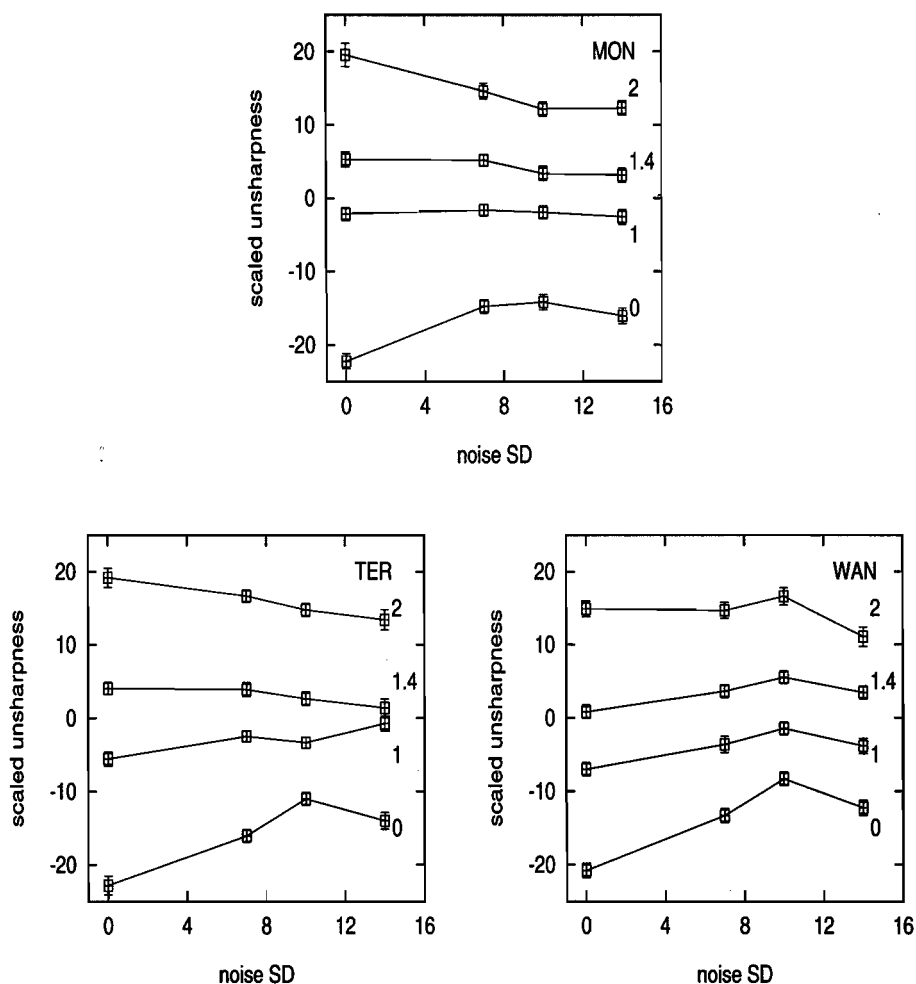


Figure 4.7: Results of scaling unsharpness, plotted as a function of the noise SD σ_n , with blur-spread σ_b as the parameter, for the Mondrian, Terrace and Wanda scenes.

can then be compared. The perceptual space that is constructed using the data of scaled attributes is based on the assumption that the chosen attributes are the most relevant ones and that they span the entire perceptual space. Hence, it is assumed that any other attribute can be expressed in terms of these basic attributes. The main difference between the MDS techniques based on the dissimilarities and those based on the scaled attributes data is that, in the latter case, the direction of the attribute vectors given by the method assists in a

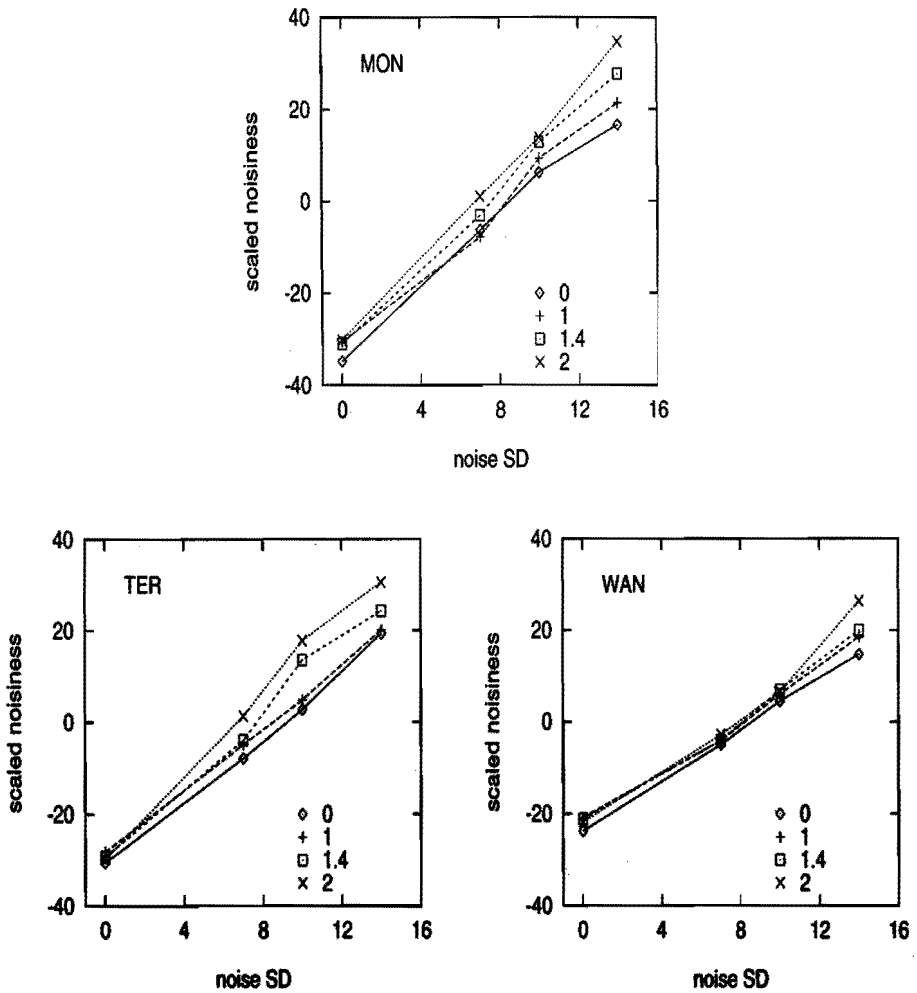


Figure 4.8: Results of scaling noise, plotted as a function of the noise SD σ_n , with blur-spread σ_b as the parameter, for the Mondrian, Terrace and Wanda scenes.

better understanding of the perceptual meaning of the dimensions.

The multi-dimensional model MDPREF (Green *et al.*, 1989) is used to analyze the data of the scaled attributes. This model represents both the attribute vectors and the stimulus vectors simultaneously in a multi-dimensional space. It is a metric MDS model. Using principle component analysis, it decomposes the input matrix of N attributes \times n stimuli into two matrices, one of N attributes \times r dimensions and another of r dimensions \times n stimuli, ac-

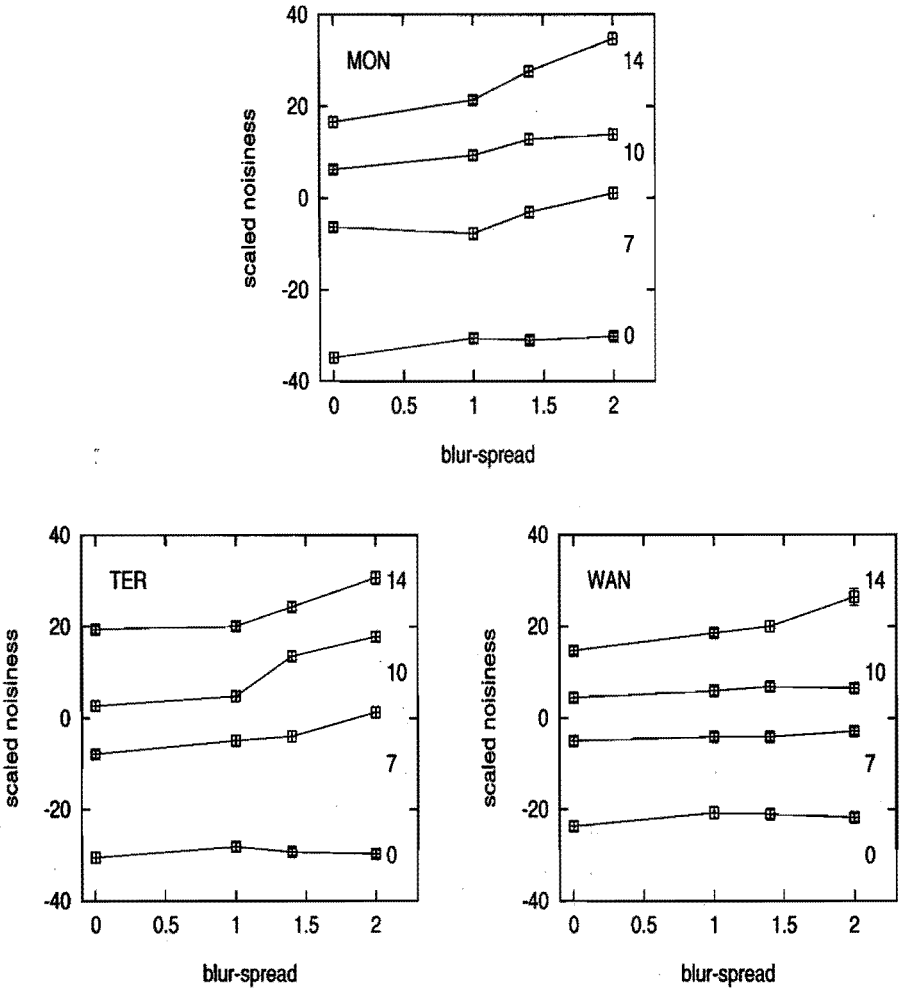


Figure 4.9: Results of scaling noise, plotted as a function of the blur spread σ_b , with noise SD σ_n as the parameter, for the Mondrian, Terrace and Wanda scenes.

cording to a least-square error criterion. The first matrix defines the attribute vectors and the second matrix represents coordinates of the stimuli in the r -dimensional perceptual space. The perpendicular projections of the stimuli on the attribute vectors represent the strengths of the corresponding attributes of the stimuli. These projections are the least-square estimates of the data in the input matrix ($N \times n$). The goodness-of-fit measure in MDPREF is the cumulative proportion of variance accounted for by all dimensions of a certain

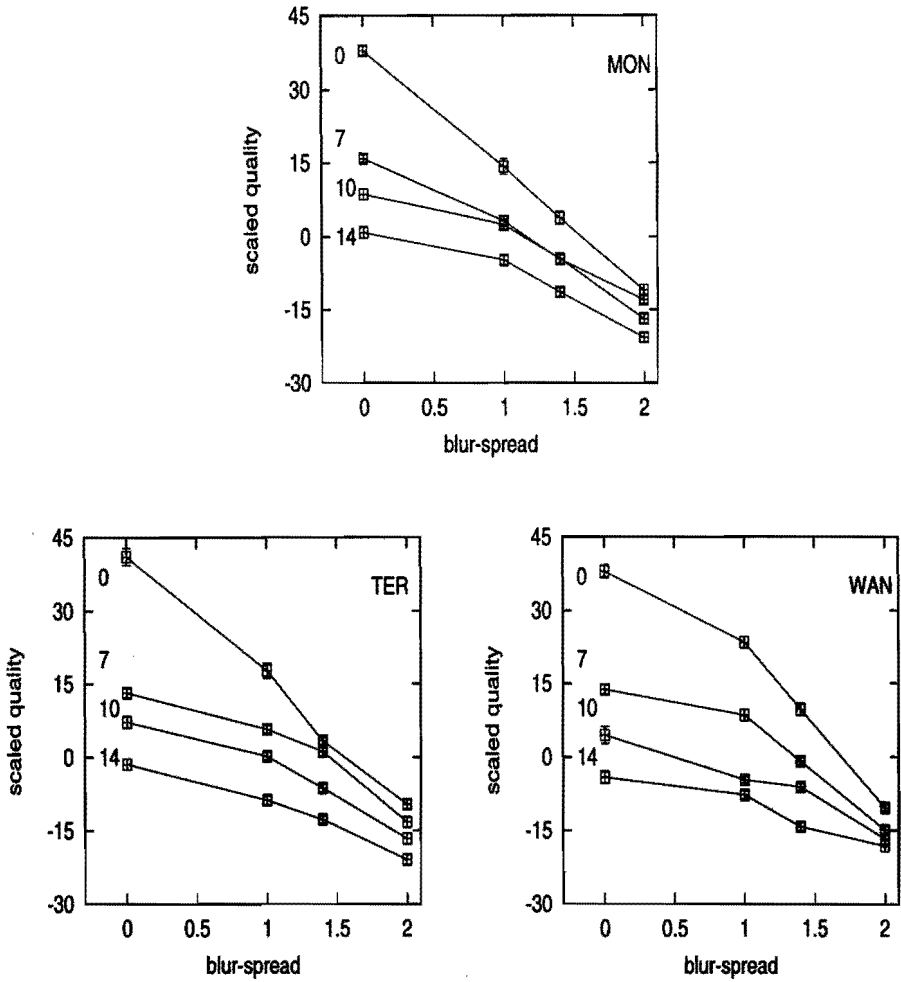


Figure 4.10: Results of scaling quality, plotted as a function of the spread of the blurring kernel σ_b , with noise SD σ_n as the parameter, for the Mondrian, Terrace and Wanda scenes.

solution.

The data of the scaled attributes, unsharpness, noisiness and quality are averaged over all seven subjects and input to MDPREF. Separate analysis is performed for each scene. Solutions up to 3-D were looked at. The proportion of variance accounted for by dimensions 1, 2 and 3 respectively are 0.697, 0.294 and 0.001 for the Mondrian, 0.718, 0.271 and 0.012 for the Terrace and 0.72, 0.266 and 0.014 for the Wanda scene. Since for all three scenes, 99 %

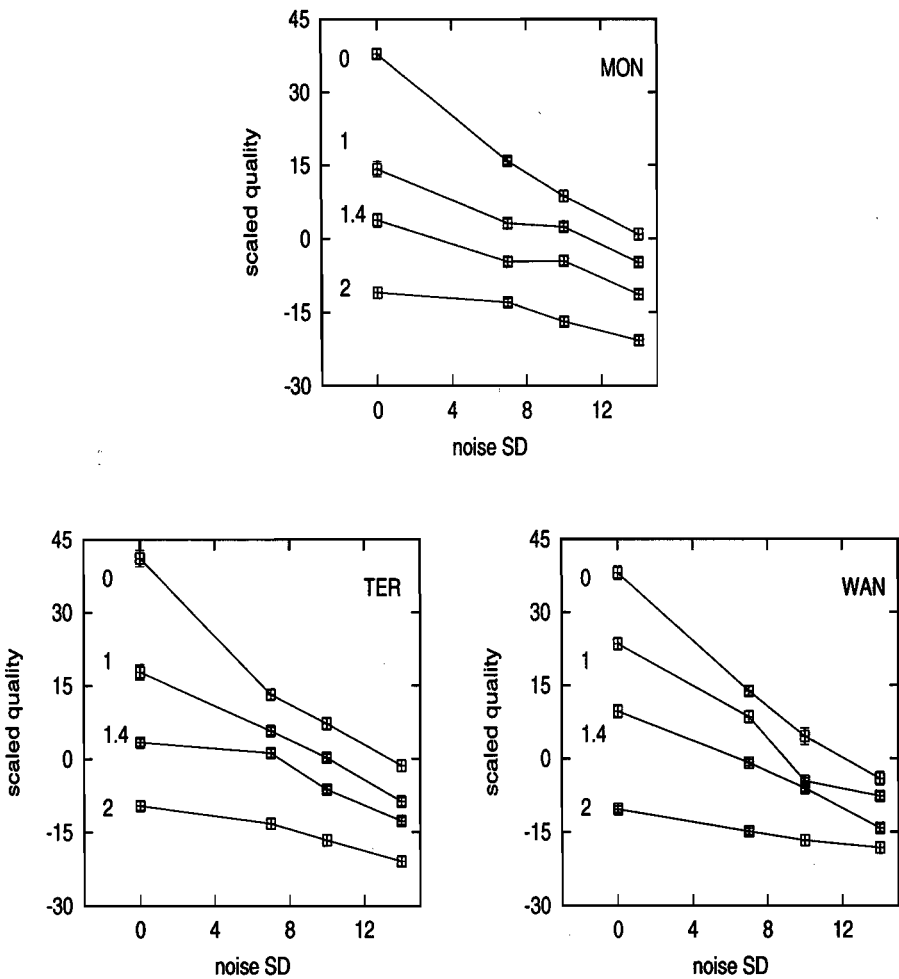


Figure 4.11: Results of scaling quality, plotted as a function of the noise SD σ_n , with blur-spread σ_b as the parameter, for the Mondrian, Terrace and Wanda scenes.

of the cumulative variance is explained by a 2-D solution, the 2-D solution was accepted as the best fitting solution for all the scenes. The 2-D solutions are shown in Fig. 4.12. The vectors show the directions of the attributes in the 2-D perceptual space. The perpendicular projections of a stimulus onto the attribute vectors correspond to the scaled attributes of that stimulus. In MDPREF, just as in MULTISCALE, the first dimension lies along the direction in which the points vary most, the second dimension is perpendicular to it,

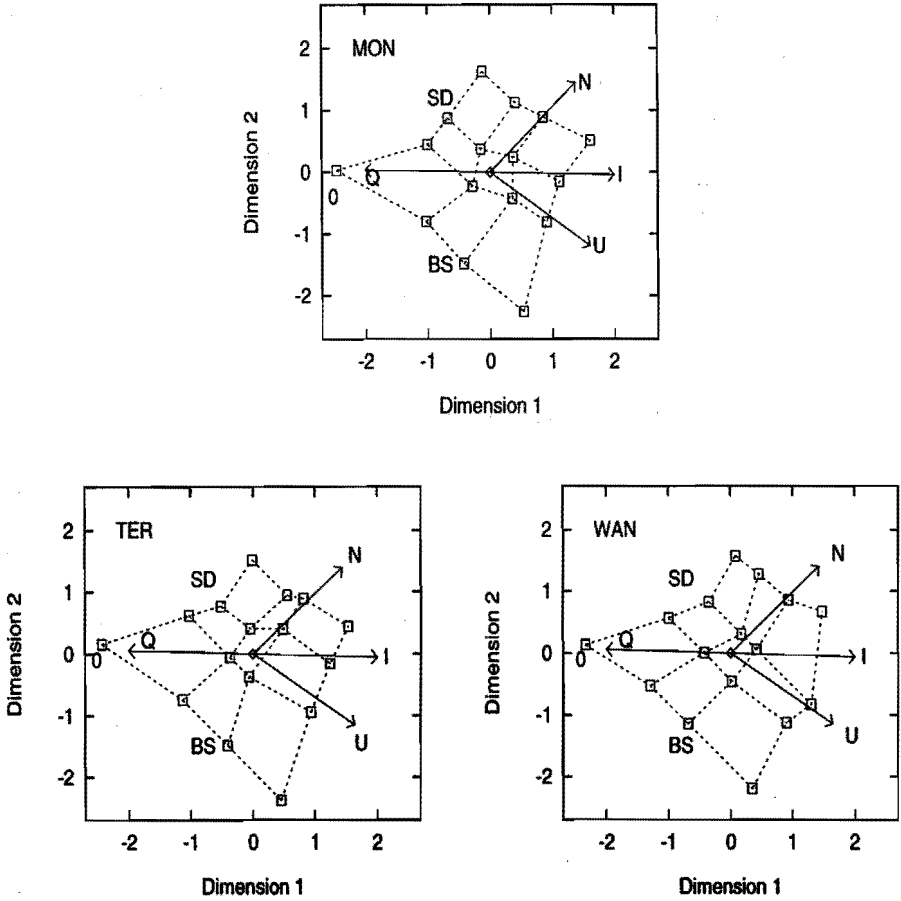


Figure 4.12: Results of analyzing the scaled attributes using MD-PREF, for the Mondrian, Terrace and Wanda scenes. The directions of attributes unsharpness (U), noisiness (N), quality (Q) and impairment ($I = -Q$) in the 2-D space are indicated by vectors.

etc. The relative angle θ between the vectors unsharpness and noisiness is 83° , 79° and 79° for the Mondrian, Terrace and Wanda scenes respectively. These results indicate that the attributes unsharpness and noisiness are not orthogonal. However, since the angles are close to 90° the interaction is weak. Here again, it is worth noting that the perceptual spaces for all three scenes are very similar.

The relation between the attribute vectors and the quality vector is also important for estimating quality. To quantify these relations, it is useful to

define a vector called *impairment* vector whose direction is opposite to that of quality, i.e. $\text{Impairment} = -\text{Quality}$ (de Ridder, 1992; Nijenhuis, 1993). In our experiment, the impairment vector lies between the unsharpness and noisiness vectors. The relative angle θ_u between the impairment vector and the unsharpness vector is 35.5° , 33.6° and 33.0° for the Mondrian, Terrace and Wanda scenes respectively. The relative angle θ_n between the impairment vector and the noisiness vector is 47.7° , 45.7° and 46.4° for the Mondrian, Terrace and Wanda scenes respectively. Note that, since the space is two-dimensional, $\theta_u + \theta_n = \theta$. For all the scenes, the impairment vector is closer to the unsharpness vector than to the noisiness vector. This implies that, in our experiments, impairment (or decrease in quality) is influenced more by unsharpness than by noisiness.

The positions of stimuli and the directions of attribute vectors given by MDPREF are best in the sense that the projection of stimuli onto the axes are the least square estimates of the input data. MULTISCALE, however, does not provide an estimate of the relative precision (i.e., standard deviation) of the estimates. To obtain an estimate of the relative precision of the directions of attributes given by MDPREF, we projected the stimulus positions given by MDPREF onto vectors of varying direction and correlated the projections with the scaled attributes. Figure 4.13 shows the correlation coefficients as a function of the direction of the vector, for attributes unsharpness, noisiness and quality. The angle at which the correlation is maximum is the optimum orientation of the attribute vector. These orientations closely match the orientations of attributes obtained using MDPREF, although there are some small differences. The differences can be seen better in the bottom right plot of Fig. 4.13, which shows a close-up of correlation coefficient versus direction of vector plot, for scaled quality. All plots in Fig. 4.13 are characterized by broad peaks, indicating a relatively low sensitivity of the direction of the attribute estimates.

The configurations obtained from the scaled attributes data using MDPREF closely resemble those obtained from dissimilarity data using MULTISCALE and SINDSCAL. The MDPREF configurations were compared with those obtained using MULTISCALE, in a similar way as the SINDSCAL and MULTISCALE configurations were compared earlier. The resulting correlation coefficients were, 0.968, 0.964 and 0.963 for the Mondrian, Terrace and Wanda scenes respectively. The MULTISCALE configuration rotated to best-fit with that of the MDPREF configuration is shown in Fig. 4.14, along with the MDPREF vectors.

Our findings are in agreement with the qualitative results obtained by Linde (1981), namely that the attributes unsharpness and noise are not orthogonal. However, unlike in their findings, we found, within the investigated range, no significant scene dependence of the perceptual spaces. The angle between the

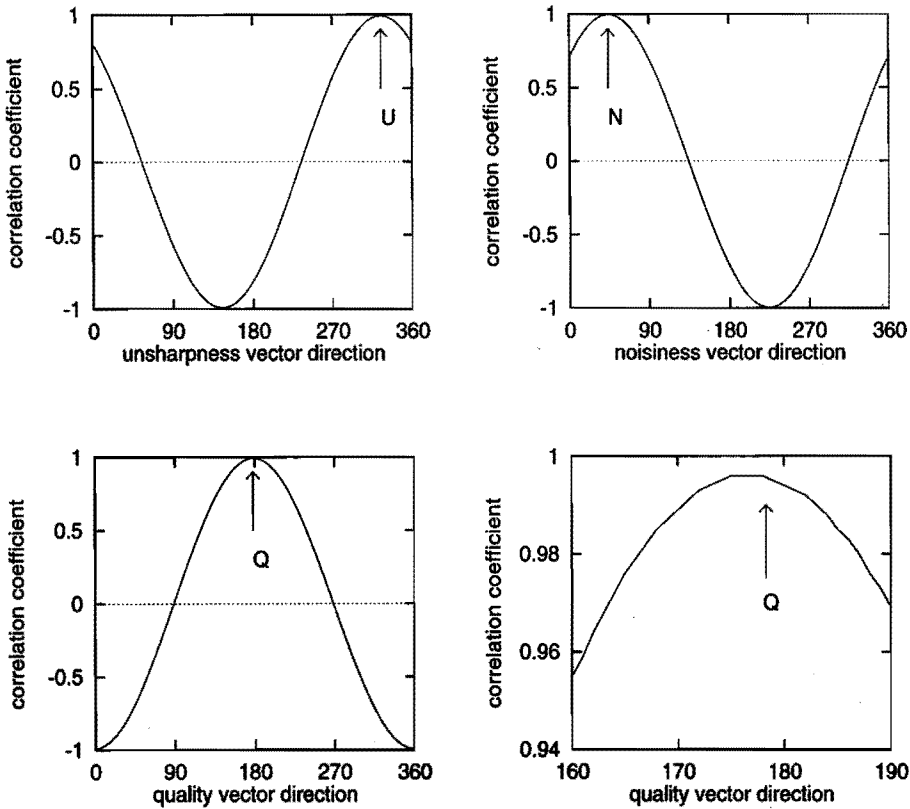


Figure 4.13: Correlation between the scaled attributes and the projections of stimulus locations obtained using MDPREF, onto vectors in different directions, as a function of the vector direction, for the Terrace scene. The directions of attributes unsharpness (U), noisiness (N) and quality (Q) given by MDPREF are indicated by arrows. The bottom right plot shows a close-up of the bottom left plot.

attributes unsharpness and noisiness was about 80° for all three scenes.

4.4 Conclusions

Images degraded by blur and noise have been characterized by locating them in a multi-dimensional perceptual space. This is carried out using two different approaches, the first based on the perceived dissimilarity between the images and the second based on the perceptual attributes of the images.

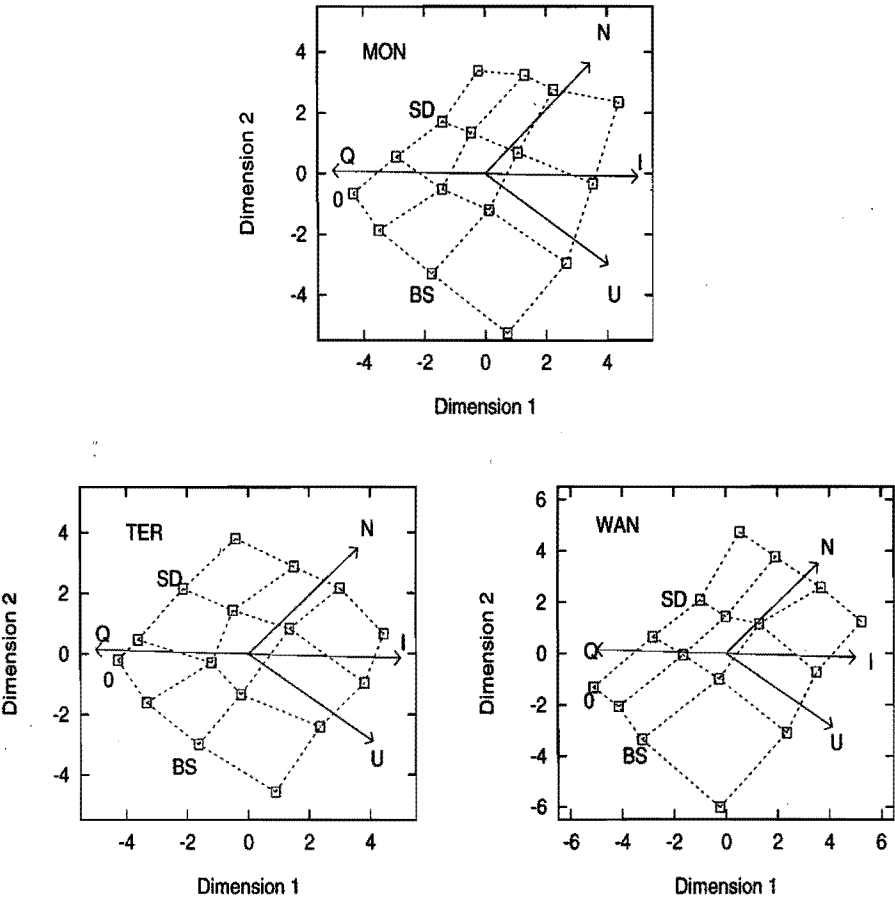


Figure 4.14: The configuration given by MULTISCALE that is rotated to best match the configuration given by MDPREF, for the Mondrian, Terrace and Wanda scenes. The directions of attributes unsharpness (U), noisiness (N), quality (Q) and impairment (I) are indicated by vectors

The first approach is based on the assumption that the underlying perceptual dimensions can be obtained from the perceived differences between the images. Therefore, the scaled dissimilarities between all stimulus pairs are used as the input to a MDS analysis. Two different MDS techniques, SINDSCAL and MULTISCALE, are used. A 2-D solution is found to provide an adequate fit using both techniques. Both techniques resulted in very similar stimulus configurations in the perceptual space.

To understand the perceptual significance of the dimensions of the MDS space, a second approach is used. This approach is based on the assumption that the underlying dimensions of the perceptual space can be obtained by constructing a MDS space directly from scaling the most relevant attributes. The scaling data of three important attributes, unsharpness, noisiness and quality, is used as input to a MDS technique (MDPREF). A 2-D solution is found to be the best fitting one using MDPREF. The configuration found using scaled attributes (MDPREF) closely matches that found using dissimilarity data (MULTISCALE and SINDSCAL). The perceptual spaces for all three scenes considered are very similar. Using the direction of the perceptual attributes unsharpness and noisiness, it is shown that the two attributes are not orthogonal. Nevertheless, the interaction is found to be weak in the range tested. The angle between the attributes unsharpness and noisiness is about 80° . The impairment vector, whose direction is opposite to the direction of quality vector, is found to lie between the unsharpness and noisiness vectors, with an angle of approximately 33° between unsharpness and impairment vectors.

Chapter 5

Estimation of Perceptual Attributes From Objective Measures

Abstract

Reliable and economic methods to assess image quality are essential for designing better imaging systems. Although reliable psychophysical methods are available to assess perceptual image quality with the help of human subjects, the cost of performing such experiments prevents their use for evaluating large amounts of image material. This has led to an increasing demand for objective methods to estimate image quality. The perceived quality of an image is determined by many underlying perceptual attributes, such as sharpness and noisiness. The relationship between these attributes and the overall quality can be represented in a multi-dimensional perceptual space, spanned by the important perceptual attributes of the images. In this perceptual space, the images are represented by points, and the strengths of their perceptual attributes are modelled by the projections of these image positions onto the attribute axes. Two important image attributes, unsharpness and noisiness, are considered in this chapter. Each of these attributes is predominantly influenced by one physical parameter of the image. The value of the objective measures for these attributes can be computed from the estimates of the corresponding physical parameters from the image, where the physical parameters are estimated from the image using computational algorithms. Similarly to the perceptual space, a psychometric space spanned by the objective (psychometric) measures for the attributes is introduced in this chapter. The mapping between the psychometric space and the perceptual space is estimated and used to map the image positions in the psychometric space to image positions in the perceptual space. The perceptual attributes of an image are hence estimated by: estimating the physical parameters, computing the objective measures for

the attributes i.e., locating the image position in the psychometric space, mapping the image position from psychometric space to the perceptual space and projecting the position onto the attribute axes. We show that the attributes and the quality thus estimated correlate well with the perceived attributes and quality.

5.1 Introduction

The technical quality of an imaging system is determined mainly by the perceptual quality of the images produced by it. Therefore, it is extremely important to evaluate the perceptual quality of the images produced by a system. Often, the perceptual quality of large amounts of image material has to be assessed with an affordable cost and within a limited time. The evaluation of processed or coded images generated with different parameter settings in order to decide the best parameter setting is a typical example. New subjective evaluation techniques with better efficiency and reliability have been developed for this purpose (Roufs, 1992; de Ridder & Majoor, 1990). However, the use of such techniques to evaluate large amounts of image material is hindered due to the amount of time and the costs required to perform the subjective evaluations. This has led to an increasing demand for objective methods for image quality prediction. Such an objective estimate must correlate well with the perceived quality of the image. A reliable objective method for image quality estimation is very useful for researchers to optimize the quality performance of image coding or processing algorithms, for example, to obtain the maximum quality with a given bit rate, or to obtain the minimum bit rate with a given quality.

Most of the early work on image quality metrics was restricted to objective measures for image display sharpness. Sharpness is an important attribute that affects the quality of the image. However, image quality is not only influenced by sharpness but also by other attributes: noisiness, brightness contrast, etc. Most popular sharpness measures were based on the modulation transfer function (MTF) of the display and the modulation threshold function of the eye. Examples of such approaches are (Barten, 1989; Barten, 1990; Carlson & Cohen, 1980). Such measures are derived from the display system characteristics, such as the MTF. In applications such as image coding, it is useful to derive the objective measure from the image instead of the system. Such measures are based on image fidelity. They assume the existence of an 'original' image and a 'degraded' or 'processed' image of the same scene. A distance function between the two images is used as the fidelity measure. Most measures make use of knowledge about the human visual system while arriving at the distance function. A large number of image quality measures of this kind have been developed, with different degrees of success. Examples of such measures

include (Limb, 1979; Lukas & Budrikis, 1982; Marmolin, 1986; Zetzsche & Hauske, 1989; Daly, 1992). A brief review of image quality metrics of this kind is given by Ahumada (1993). Recent image quality estimation methods based on models of the human visual system can be found in (Watson, 1993). A rigorous comparisons of the predictive power of these measures is not available, except for comparison of some early measures (Beaton, 1983).

Most methods for image quality estimation treat image quality as a uni-dimensional attribute. However, image quality is a multi-dimensional percept that is influenced by many basic perceptual attributes: sharpness, noisiness, brightness contrast, etc. Although it has been recognized for over a decade that image quality is multi-dimensional (Marmolin & Nyberg, 1975), very few quality measures have made use of this idea. Nakayama, Kurosu, Honjyo & Nishimoto (1980) have modelled image quality as a weighted sum of underlying quality dimensions. Escalante-Ramírez (1992) derived a perceptual space using multi-dimensional scaling (MDS) analysis and investigated the relation between image quality and its important attributes. de Ridder (1992) and Nijenhuis (1993) showed that a Minkowski metric with exponent close to 2 is a useful combination rule for the underlying dimensions of image quality.

The image quality estimation method presented in this chapter is also based on the multi-dimensional nature of image quality. According to this approach, images are located in a M-dimensional perceptual space spanned by important perceptual image attributes: sharpness, noisiness, brightness contrast, etc. In such a (metric) perceptual space, the distance between the points (images) is proportional to the perceived dissimilarity between them, and the different perceptual attributes correspond to different directions in the perceptual space. The projection of an image onto an attribute axis is proportional to the perceived strength of that attribute for that image. Therefore, by locating images in a perceptual space, their perceived attributes can be modelled. The relation between different attributes of image quality can also be derived from this space (Chapter 4). For example, the orthogonality of different attributes can be verified (Chapter 4). Because of these properties, the perceptual space is very useful in understanding and predicting perceived image quality, especially when the images are multiply impaired. The perceptual space can be constructed using MDS analysis of psychophysical data on images (Chapter 4) (Escalante-Ramírez, Martens & de Ridder, 1994).

Most perceptual attributes are *predominantly* influenced by one perceptually relevant physical parameter of the image. For example, perceived sharpness is mainly determined by the spread of the blurring kernel σ_b (Westerink, 1991)(Chapter 2), while perceived noisiness is mainly determined by the noise standard deviation σ_n (Chapter 3), etc. The relevant physical parameters can be estimated from the image using computational algorithms. The objective

measures for the perceptual attributes can be computed from these estimates by taking into account of the sensitivity of the visual system (see Chapters 2, and 3). For objective measures derived in this way, equal intervals in the objective scale correspond to approximately equal perceived difference in the attribute. The objective measures are thus psychometric quantities (Hunt, 1978).

The perceptual space is spanned by the important perceptual image attributes. Similarly, it is useful to consider a psychometric space spanned by the objective (psychometric) measures for the perceptual attributes. The objective measures are computed from the estimates of the corresponding perceptually relevant physical parameters such as noise SD and spread of the blurring kernel. The relation between these two spaces (i.e., the mapping or the transformation) can be derived. Using such a mapping, any image in the psychometric space with estimated objective measures can be mapped to a position in the perceptual space. The perceptual attributes of the image can then be estimated by projecting the image onto the attribute axes. Note that the psychometric space is spanned by the objective (psychometric) measures that have been derived from the physical parameters, taking the sensitivity of the visual system into account, whereas the perceptual space is derived from psychophysical data and hence is a psychological space.

In this chapter, we first show that the objective measures for attributes estimated using computational algorithms are good correlates of the perceived attributes of images with multiple impairments. Two important attributes, unsharpness and noisiness, will be considered. The psychophysical data from Chapter 4 will be used for this purpose. The mapping between the psychometric and the perceptual space is derived using the estimates of objective measures obtained using the computational algorithms and the MDS space obtained in Chapter 4. Using this mapping, the images are projected from the psychometric space to the perceptual space. The strengths of the perceptual attributes of the images are estimated by projecting the images onto the attribute axes. The attributes estimated using this method are correlated with the subjectively measured attributes of the images. We show that the estimated attributes and quality of images correlate well with the perceived attributes and quality.

5.2 Estimation of Objective Measures for Perceptual Attributes

5.2.1 Estimation of Blur-index

In Chapter 2, it has been shown that, for blurred images, the blur-index S_b based on the spread of the blurring kernel σ_b is a good correlate of the perceived unsharpness. The blur-index S_b can be estimated from the blurred image using the computational algorithm described in Chapter 2. A schematic diagram

of the algorithm to estimate the blur-index was shown in Fig. 2.24. For a description of the steps in the algorithms we refer to Section 2.10. In this Section we show that the blur-index S_b is a good correlate of unsharpness, also for images with multiple impairments.

The results of scaling unsharpness in images containing blur as well as noise were reported in Chapter 4. The estimated blur-index S_b of those images is compared to the scaled unsharpness in Fig. 5.1. The resulting correlation coefficients are, 0.97, 0.97 and 0.94 for the Mondrian, Terrace and Wanda scenes respectively. The correlation coefficients in this case are lower than those for images with only blur as impairment (Chapter 2). Nevertheless, these results show that the blur-index S_b is a fair correlate of perceived unsharpness also in images with multiple impairments.

5.2.2 Estimation of Noise-index

In Chapter 3, it has been shown that, for images containing noise, the noise-index S_n based on the standard deviation (SD) is a good correlate of the perceived noisiness. The noise-index S_n can be estimated from the noisy image using the computational algorithm described in Chapter 3. A schematic diagram of the algorithm to estimate the noise-index was shown in Fig. 3.19. For a description of the steps in the algorithms we refer to Section 3.7. In this Section we show that the noise-index S_n is a good correlate of noisiness, also for images with multiple impairments.

The results of scaling noisiness in images containing blur as well as noise were reported in Chapter 4. The estimated noise-index S_n of those images is compared to the scaled noisiness in Fig. 5.2. The resulting correlation coefficients are, 0.98, 0.98 and 0.99 for the Mondrian, Terrace and Wanda scenes respectively. These results show that the estimated noise-index S_n is a good correlate of the perceived noisiness in images with multiple impairments.

5.3 Mapping From Psychometric Space to Perceptual Space

A perceptual space is spanned by a number of perceptual attributes. Most attributes are predominantly influenced by one physical parameter of the image, although secondary influences from other parameters may also exist. For example, for images degraded by blur and noise, the spread of the blurring kernel and the SD of the noise are the two perceptually relevant physical parameters. The objective measures for the perceptual attributes unsharpness and noisiness are determined by these physical parameters. It is therefore useful to model a space spanned by the objective measures for the perceptual attributes. We call

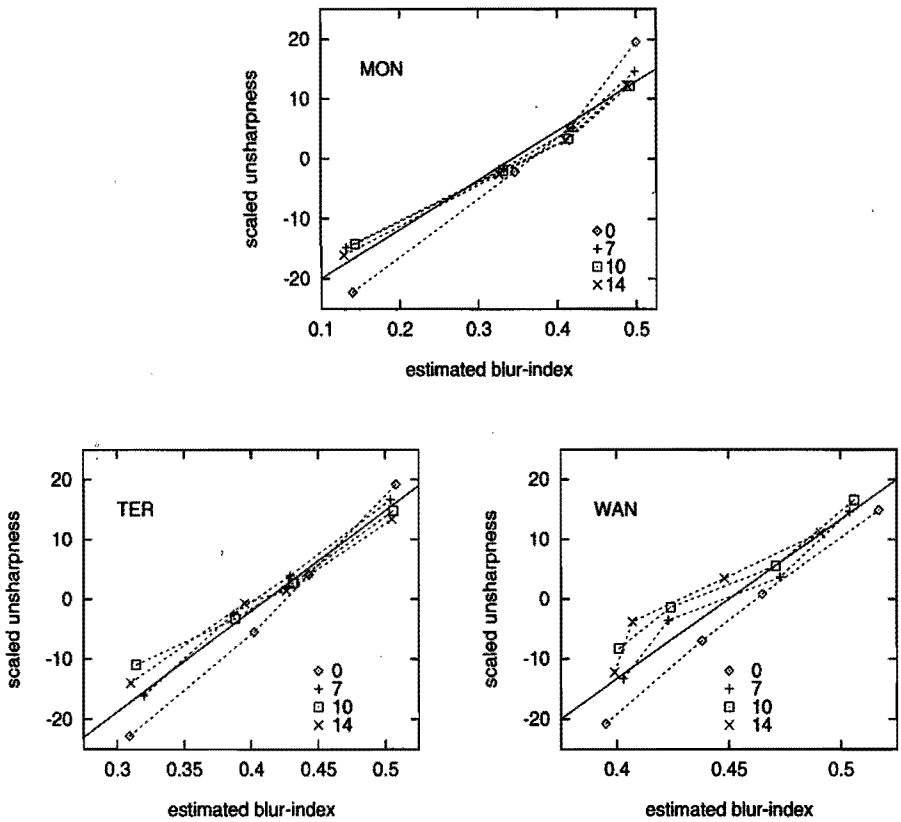


Figure 5.1: Results of correlation between scaled unsharpness and estimated blur-index S_b , for images with blur and noise. Scenes: Mondrian, Terrace and Wanda. The parameter is the SD of noise σ_n .

such a space a psychometric space. The dimensions of this space are objective (psychometric) measures based on the physical parameters, derived by taking the sensitivity of the visual system to changes in the physical parameters into account. For example, the psychometric space for images degraded by blur and noise is spanned by two parameters, the blur-index S_b and the noise-index S_n . The relation between the perceptual space and the psychometric space can be used to understand the process of image quality perception. Using the coordinates of the images in both spaces, the transformation or mapping between the two spaces can be derived. For this mapping, different kinds of transformations could be allowed: scaling, rotation about the origin, affine transformation, nonlinear warping, etc.

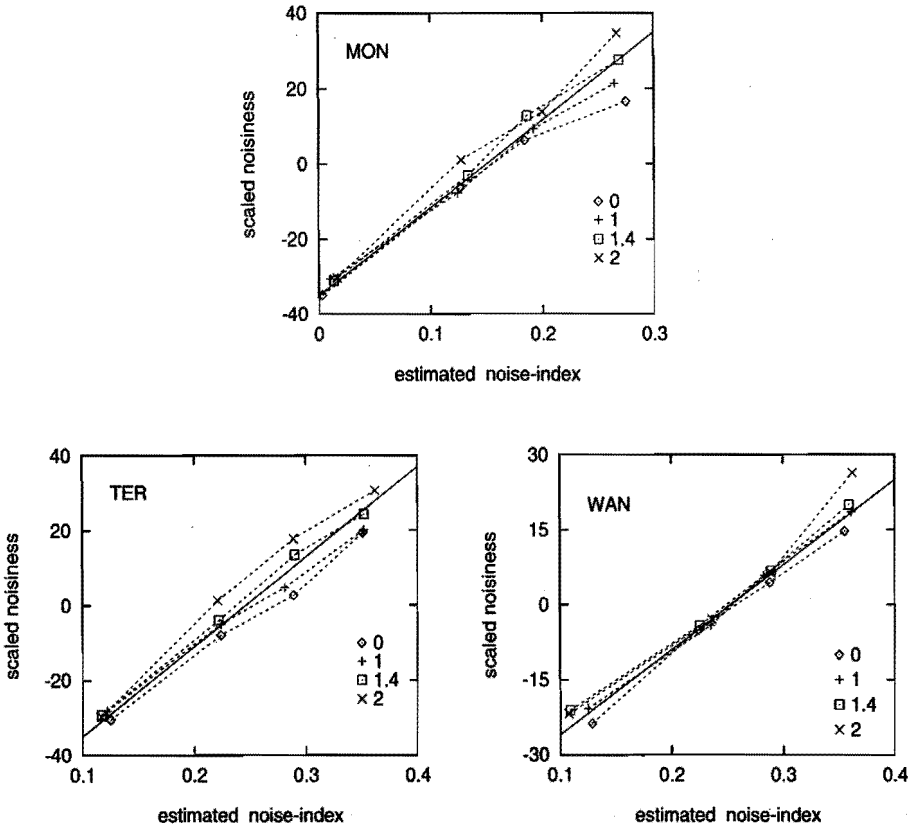


Figure 5.2: Results of correlation between perceived noisiness and estimated noise-index S_n , for images with blur and noise. Scenes: Mondrian, Terrace and Wanda. The parameter is the blur-spread σ_b .

In Chapter 4, the perceptual space of images degraded by blur and noise was derived using MDS analysis. A 2-D perceptual space spanned by the attributes unsharpness and noisiness was found to be appropriate to describe the data. Therefore, each image in the perceptual space is specified using two coordinates. Each image in the corresponding psychometric space is also specified by two parameters, the blur-index S_b and the noise-index S_n . The coordinates of images in the psychometric space are given by (S_b, S_n) , where it is assumed that the axes S_b and S_n are orthogonal. The parameters S_b and S_n are derived from the image using computational algorithms described in Sections 2.10 and 3.7, respectively.

We use MULTISCALE (Ramsay, 1991) to find the mapping between the

psychometric space and the perceptual space. The coordinates of the images in the psychometric space are input to MULTISCALE as a reference matrix Y . Matrix Y has M rows corresponding to M images and two columns corresponding to two coordinates. The matrix of coordinates of the images in the perceptual space X is transformed using a transformation matrix A such that the transformed coordinates XA best matched the reference matrix Y , where the best match is obtained by maximizing the *inner product* $\text{tr}(A^t X^t Y)$. MULTISCALE also gives a correlation measure for the match (Ramsay, 1991). Scaling, rotation and affine transformation (oblique axis) are permitted to obtain the best match, i.e., the matrix A is only required to be nonsingular. The transformation matrices A obtained for the Mondrian, Terrace and Wanda scenes are given below:

$$A = \begin{bmatrix} 0.034 & 0.009 \\ -0.008 & 0.055 \end{bmatrix} A = \begin{bmatrix} 0.024 & 0.018 \\ -0.027 & 0.023 \end{bmatrix} A = \begin{bmatrix} 0.027 & 0.001 \\ -0.006 & 0.017 \end{bmatrix} \quad (5.1)$$

The respective correlation coefficients for the match are 0.93, 0.98 and 0.99.

Our aim is to map the images from the psychometric space to the perceptual space. The transformation matrix $G = (A^T)^{-1}$ was used for this purpose, i.e.,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{bmatrix} \begin{bmatrix} S_n \\ S_b \end{bmatrix} \quad (5.2)$$

where x and y are the coordinates of the image in the perceptual space, along dimension one and two respectively. Figure 5.3 shows the results of such mapping for three scenes, along with the MDS solution. The match between the mapped coordinates and the coordinates of the images obtained using MDS is good for the two natural scenes, Terrace and Wanda, with high correlation coefficients 0.98 and 0.99 respectively. However, for the synthetic scene, Mondrian, the correlation coefficient is only 0.93 and the match is worse, especially for images with no blur, i.e., the lowermost row.

One possible explanation for the difference observed in the results for the synthetic scene, Mondrian, is that while judging dissimilarities, subjects classify images into those that are multiply impaired and those that are impaired by only one impairment. It is relatively easy for subjects to perform such classification in the case of Mondrian, since it has edges that are close to ideally sharp transitions and uniform regions that are close to perfectly noiseless regions. On the other hand, such a classification is relatively difficult in the case of images of natural scenes, since the original images of natural scenes are already blurred and are impaired by noise. To test this, a separate analysis was performed for the Mondrian scene in which the images with no blur were not used. The results of the analysis and the mapping are shown in Figure 5.4. Now the correlation coefficient increased from 0.93 to 0.96. A similar analysis was

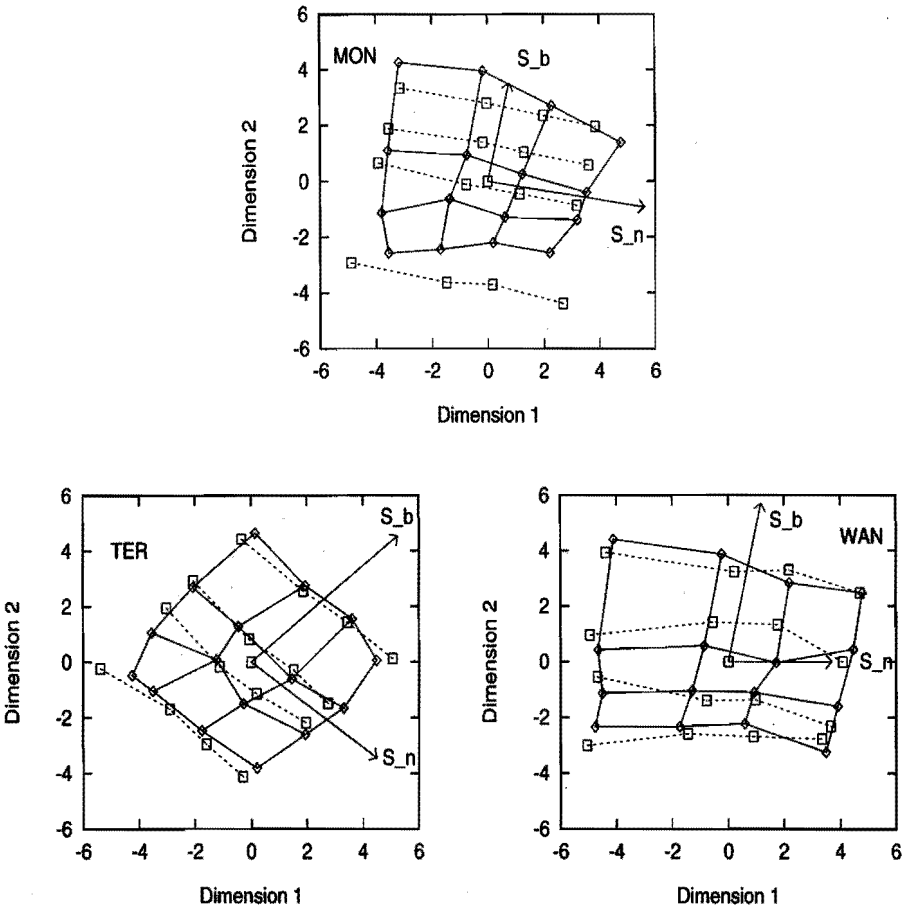


Figure 5.3: Results of correlation between points in the perceptual space and the points mapped onto it from the psychometric space, for the Mondrian, Terrace and Wanda scenes. The points joined by the drawn curves show the positions of images obtained using MDS analysis of dissimilarities (MULTISCALE solution, same as Fig. 4.4). The points joined by the dotted line show the positions of images obtained by mapping their positions from the psychometric space, using the mapping G . The vectors S_b and S_n show the mapped directions of S_b and S_n axes from the psychometric space.

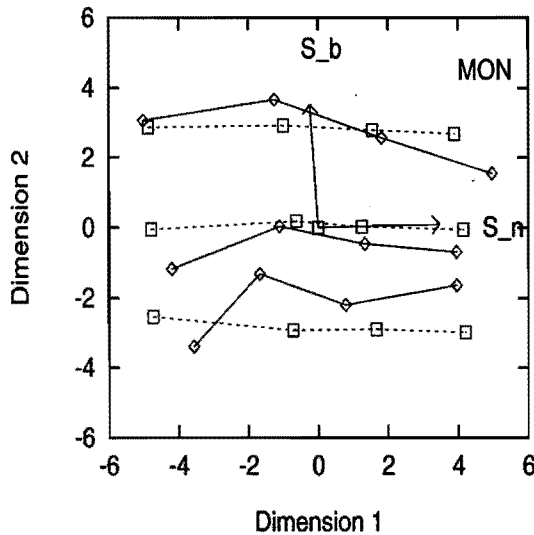


Figure 5.4: Results of correlation between points in the perceptual space and the points mapped onto it from the psychometric space, for the Mondrian scene, where images with no blur have not been considered. The points joined by the drawn curves show the positions of images obtained using MDS analysis of dissimilarities (MULTI-SCALE solution). The points joined by the dotted line show the positions of images obtained by mapping their positions from the psychometric space, using the mapping G . The vectors S_b and S_n show the mapped directions of S_b and S_n axes from the psychometric space.

also performed with only those images of the Mondrian that contained both impairments, and a result similar to that shown in Fig 5.4 was obtained. Figure 5.4 shows that although there is a considerable improvement in the match, the transformation is unable to completely capture the variations in the perceptual space. A main reason for this is that, while mapping, only rotation, scaling and affine transformations were allowed.

The transformation G is a combined result of scaling, rotation and affine transformation. To study each of these transformations separately, the transformation matrix G can be decomposed into four matrices, corresponding to the four different operations, where the scaling operation has been decomposed further into two operations, equal scaling on both axes and unequal scaling. Thus the four operations are: unequal scaling, affine transformation, equal scaling, and rotation, in this order. The matrix decomposition corresponding

to these operations can be written as

$$G = \begin{bmatrix} g_{1,1} & g_{1,2} \\ g_{2,1} & g_{2,2} \end{bmatrix} = \begin{bmatrix} \cos \omega & \sin \omega \\ -\sin \omega & \cos \omega \end{bmatrix} \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \cos \phi & \sin \phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}. \quad (5.3)$$

Here, ω is the angle by which the configuration is rotated, s is the equal scaling factor on both axes, ϕ is the angle between the axes in the new coordinate system and λ is the weight on the objective measure of the second attribute. By solving the above equations, the parameters ω , ϕ , s and λ can be computed from the coefficients of the matrix G . Using this procedure, we obtained $\phi = 87^\circ, 78^\circ$ and 78° for the Mondrian, Terrace and Wanda scenes. These angles are close to the angles between the attribute unsharpness and noisiness ($\theta = 83^\circ, 79^\circ$ and 79°) obtained using MDS analysis for these scenes (Chapter 4). This indicates that the angle between the mapped directions of axes S_b and S_n onto the perceptual space very closely represent the angle between the directions of the attributes unsharpness and noisiness. The weights λ are 1.58, 0.85 and 0.66 for the Mondrian, Terrace and Wanda scenes respectively. These weights indicate that, for natural scenes the second objective measure (i.e, blur-index) has to be compressed relative to the first while mapping, whereas for the synthetic scene it has to expanded. This can also be seen graphically from the relative lengths of vectors S_b and S_n for each scene, in Fig. 5.3. For the solution of the Mondrian scene shown in Fig. 5.4, the angle $\phi = 87^\circ$ and weight $\lambda = 0.99$. The weight λ in this case is closer to the values obtained for natural scenes. This indicated that the images of the Mondrian scene with multiple impairment behave very similar to the natural images, although images with only one impairment may deviate in their behaviour from that of the natural images.

In the transformation G derived above, the rotation and equal scaling operations do not have any perceptual significance. Hence, the two significant operations are: the unequal scale factor on the objective measures and the affine transformation. Based on this, we propose that any image in the psychometric space with coordinates (S_n, S_b) can be mapped onto the perceptual space (x, y) using the transformation,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \cos \phi & \sin \phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} S_n \\ S_b \end{bmatrix} \quad (5.4)$$

The two parameters in this transformation are the angle ϕ between the axes in the perceptual space and the weighting factor of the second dimension λ . These can be estimated using the procedure described above. For natural images we used $\phi \approx 80^\circ$ and the average $\lambda \approx 0.75$.

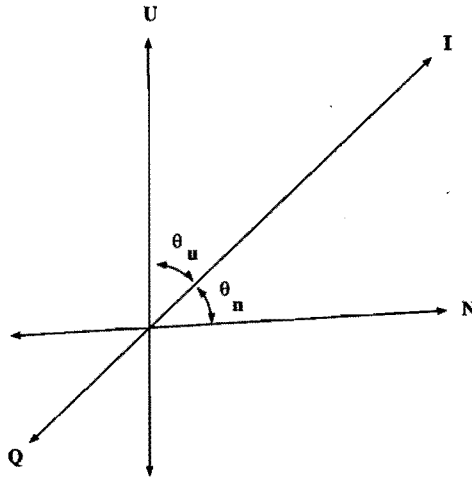


Figure 5.5: Relation between the directions of attributes quality (Q), impairment (I), unsharpness (U) and noisiness (N), in the perceptual space. The relative angles θ_u and θ_n are also indicated.

5.4 Estimation of Perceived Quality

The approach to image quality estimation to be developed in this Section is based on the assumption that the perceptual attributes in a perceptual space can be considered as vectors, each specified by a direction in the perceptual space. In Chapter 4, the orientation of the perceptual attributes in the perceptual space were identified using MDS analysis. The perceptual quality of an image was modelled to be proportional to the projection of that image onto the quality axis. In this Section we show how this method of projection can be used to estimate quality and other attributes of images.

It was found in Chapter 4 that the perceptual attributes unsharpness and noisiness are separated by an angle $\theta \approx 80^\circ$. The direction of overall impairment (i.e., -quality) was in between those of the two attributes. A schematic diagram indicating the relative angle θ_u between impairment and unsharpness, and the relative angle θ_n between impairment and noisiness, is shown in Fig 5.5.

Any image in the psychometric space can be mapped to the perceptual space using the transformation given by Eq. (5.4). We propose that the direction of vectors representing any basic perceptual attribute in the perceptual space can be estimated by mapping the direction of vectors representing the corresponding relevant physical parameter in the psychometric space. For example, the direction of noisiness can be estimated by mapping the S_n -axis

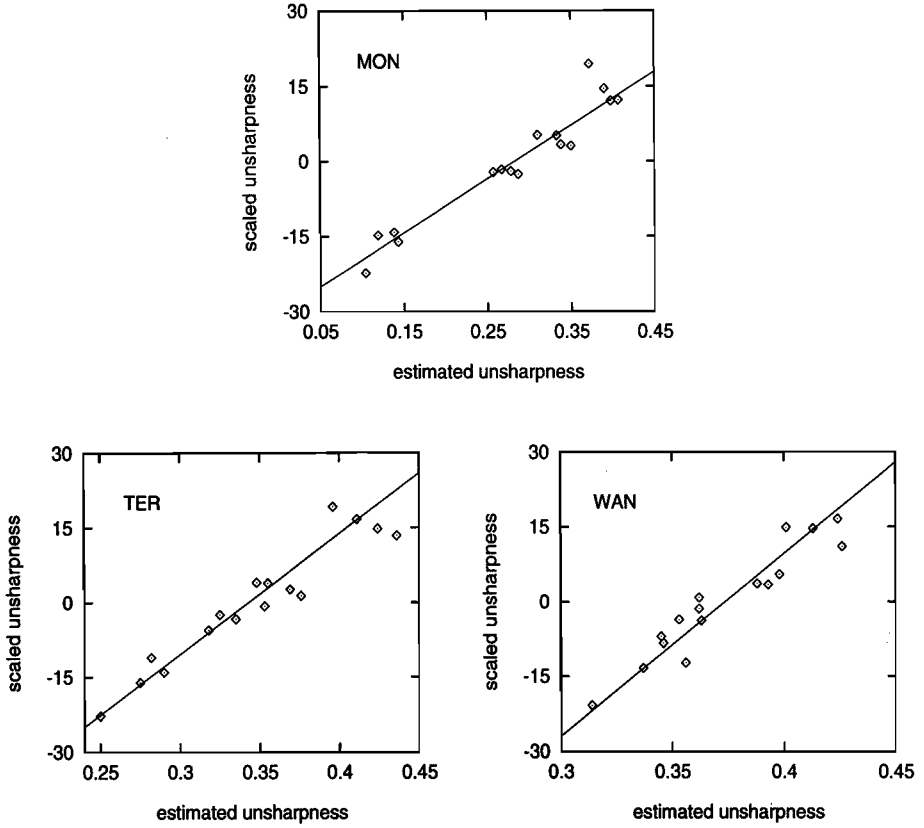


Figure 5.6: Results of correlation between perceived unsharpness and its estimate obtained by projecting the image onto the unsharpness axis, for the Mondrian, Terrace and Wanda scenes.

using Eq. (5.4). This can also be verified by finding the directions of the attribute based on the perceptual attributes measured on the images. We also propose that the directions of the vectors representing the basic attributes in the perceptual space are invariant to the range of variations of attributes in the experiment. The direction of global attributes such as quality, that are derived from the basic attributes, may vary between experiments. Hence, the angle between the quality axis and the other basic attribute axes may change, but not the angles between the basic attribute axes.

Using the mapping from the psychometric space to the perceptual space and the assumptions stated above, the directions of the basic attributes in the perceptual space can be estimated. The strengths of the attributes in the

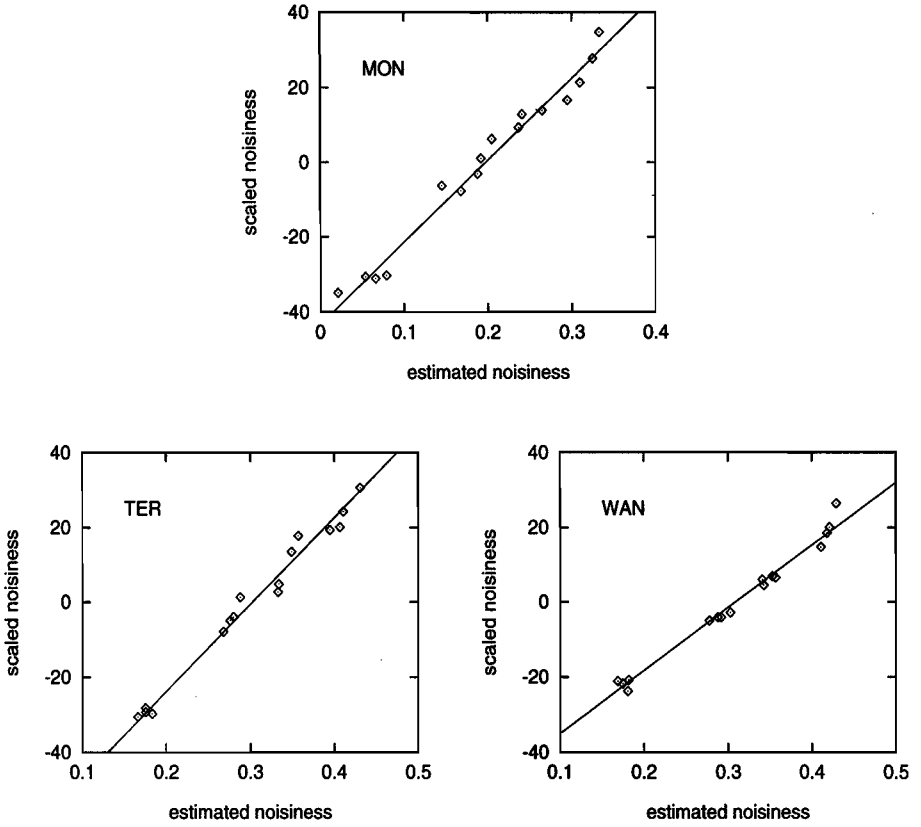


Figure 5.7: Results of correlation between perceived noisiness and its estimate obtained by projecting the image onto the noisiness axis, for the Mondrian, Terrace and Wanda scenes.

images can be estimated by projecting the stimulus position onto the axes. Having estimated the direction of the basic attributes in the perceptual space, the direction of quality can be specified with respect to the basic attributes. For example, in our case, using θ_u or θ_n , with respect to unsharpness or noisiness. Thus the direction of quality can also be estimated. The quality of the images is given by their projection onto the quality axis.

First we show the results of estimating the attributes obtained using the projection of stimuli onto the attribute axes. In Fig. 5.6, perceived unsharpness is plotted against the strength of unsharpness estimated using the projection method described above. Figure 5.7 shows the same for noisiness. These results were obtained using Eq. (5.4), with $\phi = 80^\circ$ and $\lambda = 0.75$. The correlation

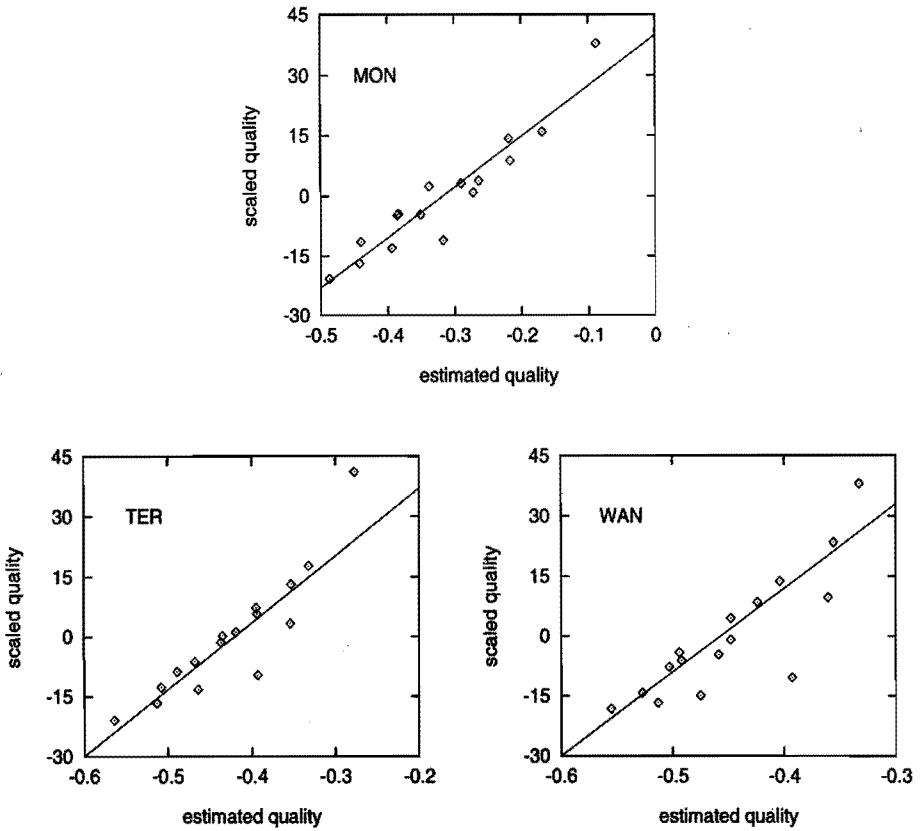


Figure 5.8: Results of correlation between perceived quality and quality estimates obtained by projecting the image onto the quality axis, for the Mondrian, Terrace and Wanda scenes.

coefficients for the unsharpness estimates (Fig. 5.6) are 0.96 , 0.96 and 0.95 for the Mondrian, Terrace and Wanda scenes respectively. The correlation coefficients for the noisiness estimates (Fig. 5.7) are 0.99, 0.99 and 0.99 for the Mondrian, Terrace and Wanda scenes respectively.

In Fig. 5.8, perceived quality is plotted against the quality estimated using the projection method described above. For all the three scenes, $\theta_u = 34^\circ$ and $\theta_n = 80^\circ - 34^\circ = 46^\circ$ was used. This is the average value of angles for the three scenes obtained from the perceptual space constructed using MDS analysis (Chapter 4). The correlation coefficients for quality estimates are 0.95, 0.91 and 0.86 for the Mondrian, Terrace and Wanda scenes respectively.

Note that the estimates reported in Figs. 5.6, 5.7, and 5.8, are obtained

with fixed values of parameters λ , ϕ , and θ_n , for all scenes. The results reported here show that the concept of a mapping between the psychometric space and the perceptual space is very useful for quality prediction. The results also show that the attribute and quality prediction by projecting the images onto the attribute axes in the perceptual space is a very promising method, and leads to reliable predictions for attributes and quality.

5.5 Conclusions

Image quality is determined by many underlying perceptual attributes such as sharpness, noisiness and brightness contrast. The relationship between these attributes and the quality can be studied in the perceptual space spanned by the important basic attributes of the images. Each of the perceptual attributes is predominantly influenced by one perceptually relevant physical parameter of the image. These physical parameters can be estimated from the image using computational algorithms. We have shown that the blur-index estimated using the algorithm proposed in Chapter 2 correlates well with the perceived unsharpness of images with both blur and noise. Similarly, the noise-index estimated using the algorithm proposed in Chapter 3 is shown to correlate well with the perceived noisiness in images with both blur and noise. In this chapter, a psychometric space spanned by the objective measures for perceptual attributes is introduced. The mapping between the psychometric and the perceptual space is shown to be very useful to estimate the perceptual attributes and the quality of images. It is shown that the estimates of quality and other attributes obtained using this new method correlate well with the perceived quality and the attributes of images.

Chapter 6

An Algorithm to Estimate an Objective Measure for Global Brightness Contrast

Abstract

Imaging system designers require objective measures that correlate well with perceived image quality and its attributes. Image quality as a function of one of its important perceptual attributes, global brightness contrast, has a maximum. Therefore, an objective measure for global brightness contrast is very useful to optimize the quality of images produced by imaging systems. In this chapter, an objective measure for global brightness contrast is presented. It is based on the slope of the cumulative histogram of an approximate psychometric brightness image. By applying the algorithm on images with different contrast and by using the subjectively scaled brightness contrast data, we show that the proposed objective contrast measure is a good correlate of the perceived global brightness contrast of natural images. We also show that the contrast measure derived here can be used to optimize the perceptual quality of images produced by imaging systems.

6.1 Introduction

Global brightness contrast is an important basic attribute of perceptual image quality. Subjects are able to make consistent judgements of this image attribute (Roufs, 1992; van Overveld, 1994). The perceptual image quality of natural scenes as a function of their perceived global brightness contrast has a maximum (Roufs & Goossens, 1988). In imaging system design, it is important to ensure the optimum quality of the images produced by the system. To attain this, the global brightness contrast of the output image has to correspond to the optimum quality. This optimization needs an objective measure that correlates well with the global brightness contrast of the image.

Perceptual image quality is determined by many underlying attributes: unsharpness, noisiness, brightness contrast, etc (Marmolin & Nyberg, 1975). Image quality can be estimated by first estimating the underlying attributes and then combining them. In Chapter 2 we have proposed an algorithm to estimate an objective measure for image unsharpness. In Chapter 3 we have proposed an algorithm to estimate an objective measure for image noisiness. In this chapter, we present an algorithm to estimate an objective measure that correlates well with the perceived global brightness contrast of images.

The luminance reproduction characteristics of an imaging system are an important factor influencing the brightness contrast of the output image. The overall luminance reproduction characteristics of an imaging system are given by the combination of the luminance reproduction characteristics of its units: the camera, the image processor, the display, etc. The relation between the input and output of most of these units can be approximately described as power functions, although some units may have approximately linear or linearized input-output characteristics. The exponent of such a power relation 'gamma' (Poynton, 1993) is a widely used physical parameter to specify the input-output characteristics of the units. For example, for a display monitor the relation between the input voltage and the output luminance is given by a power function with approximate $\gamma = 2.5$. When the input-output relation, i.e., the luminance reproduction function, is plotted on log-log axes, the slope of the linear part of the curve equals gamma.

In image quality applications, the interest is to obtain an estimate for the global brightness contrast of the image. Subjects are able to make consistent and accurate judgements of global brightness contrast of images, although the local brightness contrast may vary considerably across any given image (Roufs, 1992). There are many definitions and methods to measure local brightness contrast (Peli, 1990; Lillesaeter, 1993; Hurvich & Jameson, 1966). However, the way in which these local contrasts combine to form a global impression of brightness contrast is still largely unknown, especially for images of natural scenes. The gamma parameter of an imaging system indicates only the ability of the system to reproduce the input luminance contrast. It however is not an adequate measure for the global brightness contrast of the image of a scene. Roufs, Koselka & Tongeren (1994) showed that an objective measure called effective gamma correlates closely with the global brightness contrast of natural images. They derived the effective gamma by a linear regression of the most frequently occupied region of the luminance reproduction curve of the imaging system.

Our initial attempts towards developing an objective measure for global brightness contrast were based on local images features. We first estimated a local contrast measure at prominent locations in the image and then combined

them to form an estimate of a global contrast measure. This is similar to the way in which the (global) image blur spread is computed from (local) edge blur spread in Chapter 2. We chose edges as the prominent local features and estimated local contrast measures at all prominent edge locations in the image. The local contrast measures at edges were computed using the edge parameter estimation algorithm described in Chapter 2. Prominent edges were also selected in the same way as they were selected for blur estimation in Chapter 2. Several definitions were tried for local contrast measures, including the widely used Michelson contrast given by $C_l = (S_{max} - S_{min}) / (S_{max} + S_{min})$, where S_{max} and S_{min} are the signal levels on the high and the low side of the edge respectively. These measures were tried on luminance images as well as on approximate psychometric brightness images. The mean of the local contrast measure computed over all selected edge locations was taken as an estimate of global contrast measure. Although the perceived contrast increased monotonically with a measure derived in this way, images of two different scenes perceived to be of equal contrast had a very different value for the global contrast measure. Thus, our attempts based on (local) image features, similar to those presented in Chapter 2 and Chapter 3, did not lead to useful results towards developing a measure for global brightness contrast. The reasons for this may be manyfold. First of all, it is difficult to find a reliable objective measure for local brightness contrast (Cannon Jr, 1984; Peli, Yang, Goldstein & Reeves, 1991). Secondly, the selection of local features that play a prominent role in deciding the global brightness contrast is very difficult. Thirdly, even when a reliable local contrast measure is found, thorough knowledge about how the local measures have to be combined to derive a global measure does not exist. Finally, the global brightness contrast is influenced mainly by the first-order variation in brightness across the image, whereas the local edge contrast does not provide any information about this sort of variation across the image. The first-order variation in brightness across the image is influenced by the areas of different brightness in the image and the contrast between them.

Because of the abovementioned difficulties in arriving at a measure for contrast based on local features, we sought to derive a heuristic measure for global brightness contrast that captures the first-order variation in brightness across the image. The main motivation for our heuristic approach comes from the histogram equalization methods used in image processing (Pratt, 1991) as well as from the concept of an effective gamma (Roufs *et al.*, 1994).

In this chapter, we present an algorithm to estimate an objective measure for the global brightness contrast of an image, from the image itself. We call this measure a *contrast-index*. The contrast-index we propose here is based on the cumulative histogram of the approximate psychometric brightness image. Thus it is not based on first estimating the local contrast measures and then combining them to derive a global measure, instead it directly derives a global

measure from the image. Since the contrast-index is based on the cumulative histogram, it indirectly captures the effect of areas of different brightnesses on the impressions of global brightness contrast. Using the contrast-index, the global brightness contrast of any image can be estimated from the image itself. Since the contrast-index correlates well with the perceived global brightness contrast and because quality as a function of global brightness contrast has a maximum, the contrast-index can also be used to optimize the perceptual quality of images produced by imaging systems.

We first describe the contrast-index estimation algorithm. The results of psychophysical experiments involving scaling of global brightness contrast and quality of images (Roufs *et al.*, 1994) will be used to show that the contrast-index proposed here correlates well with the scaled brightness contrast of images.

6.2 Estimation of Contrast-index

The contrast-index of an image is estimated from an approximate psychometric brightness image. Figure 6.1 shows a schematic diagram of the contrast-index estimation algorithm. The steps in this algorithm are similar to the blur-index and noise-index algorithms presented earlier in Chapters 2 and 3. In the initial stage of the algorithms, the grey-scale image is transformed into a luminance image using the monitor characteristics, as defined in Eq. (2.46) of Chapter 2. The luminance image is blurred according to the eye-optics and then transformed into an approximate psychometric brightness image. This transformation from luminance to approximate psychometric brightness is the same as the one used in the algorithm to estimate blur-index and noise-index. We refer to Section 2.10 of Chapter 2 for a description of this transformation and the details of the eye-optics and the intrinsic noise. The cumulative histogram of the approximate psychometric brightness image is computed and used for further estimation.

Figure 6.2 shows the histograms of approximate psychometric brightness images for three scenes, at image processor gamma $\gamma_p = 0.4$ and 1.6. The three scenes are, stad, an indoor scene containing a sculpture and indoor plants, ter, an outdoor scene of a cafe and Wanda03 (w03), a female portrait. For all scenes, the dynamic range (spread or width) of the psychometric brightness histogram is small (≈ 50) when $\gamma_p = 0.4$ and is large (≈ 80) when $\gamma_p = 1.6$. The perceived global brightness contrast of images with $\gamma_p = 1.6$ is also considerably higher than that of images with $\gamma_p = 0.4$. These histograms show that the dynamic range of the psychometric brightness in an image is an important factor in determining its perceived global brightness contrast. This fact has been exploited for many years by image contrast enhancement algorithms based on histogram

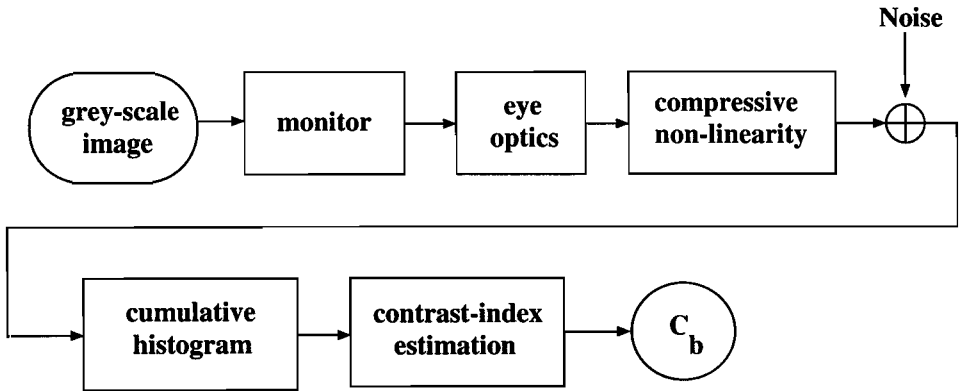


Figure 6.1: Contrast-index estimation algorithm

equalization or modification (Pratt, 1991; Gonzalez & Wintz, 1987; Woods & Gonzalez, 1981). We use the width of the psychometric brightness histogram as a first order measure for the dynamic range of brightness in the image.

Let us assume, for the purposes of explanation, that the probability density function (PDF) of the approximate psychometric brightness image is uniform (Papoulis, 1991), i.e.,

$$f(b) = \begin{cases} \frac{1}{b_2 - b_1} & b_1 \leq b \leq b_2 \\ 0 & \text{otherwise.} \end{cases} \quad (6.1)$$

The cumulative distribution is then a ramp function,

$$F(b) = \begin{cases} 0 & b < b_1 \\ \frac{1}{b_2 - b_1} [b - b_1] & b_1 \leq b \leq b_2 \\ 1 & b > b_2. \end{cases} \quad (6.2)$$

The cumulative histogram increases linearly from 0 to 1, when b increases from b_1 to b_2 . The slope of the linear part of this cumulative distribution is $1/(b_2 - b_1)$. It is inversely proportional to the width of the histogram ($b_2 - b_1$). Therefore, the width of the histogram can be obtained from the estimate of the slope of the linear part of the cumulative distribution. Figure 6.3 shows plots of uniform PDFs and their cumulative distribution functions, for two different widths of the PDF. Note that the PDF with the larger width has a shallower slope in the cumulative distribution. Note also that the starting position of the linear part of the cumulative distribution depends on the starting position of the uniform PDF, b_1 .

In the above treatment, we assumed a uniform PDF to relate the width of the PDF to the slope of the cumulative distribution. Similar relations also

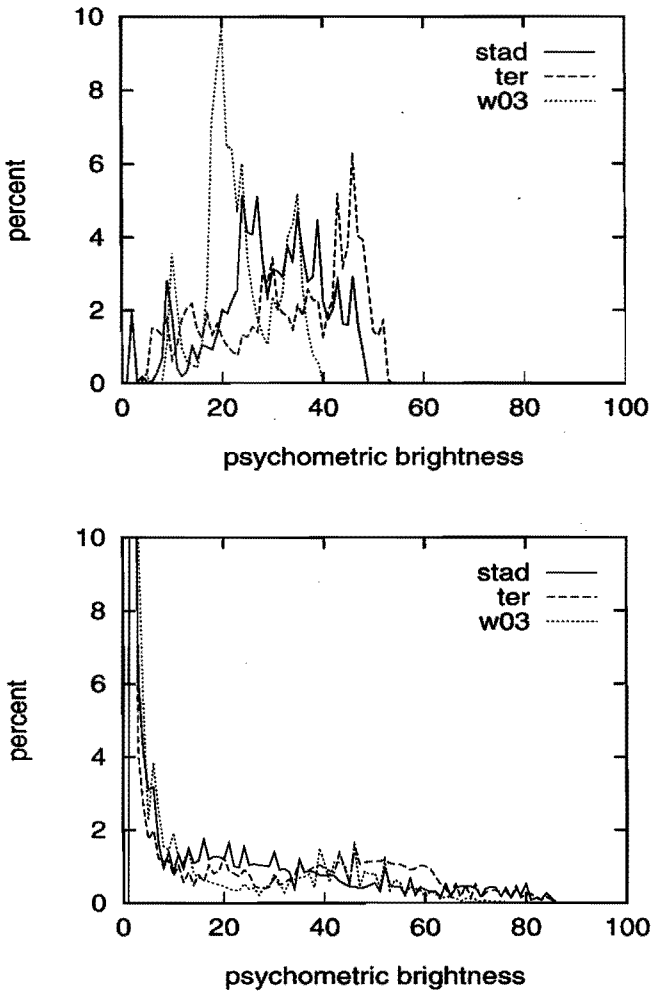


Figure 6.2: Histograms of approximate psychometric brightness images for 3 scenes, stad, ter and W03. Top, at $\gamma_p = 0.4$ and bottom, at $\gamma_p = 1.6$.

hold when the PDF is a Gaussian. When the PDF is a Gaussian, the slope of the linear part of the cumulative distribution is inversely proportional to the spread of the Gaussian (Papoulis, 1991). The cumulative distribution in the range 0.16 to 0.84 is almost linear. Hence, the inverse of the slope of the (linear part of the) cumulative distribution can be used as a first order approximation for the width of the histogram, for a wide variety of PDFs.

The cumulative histograms for three scenes, at image processor gamma

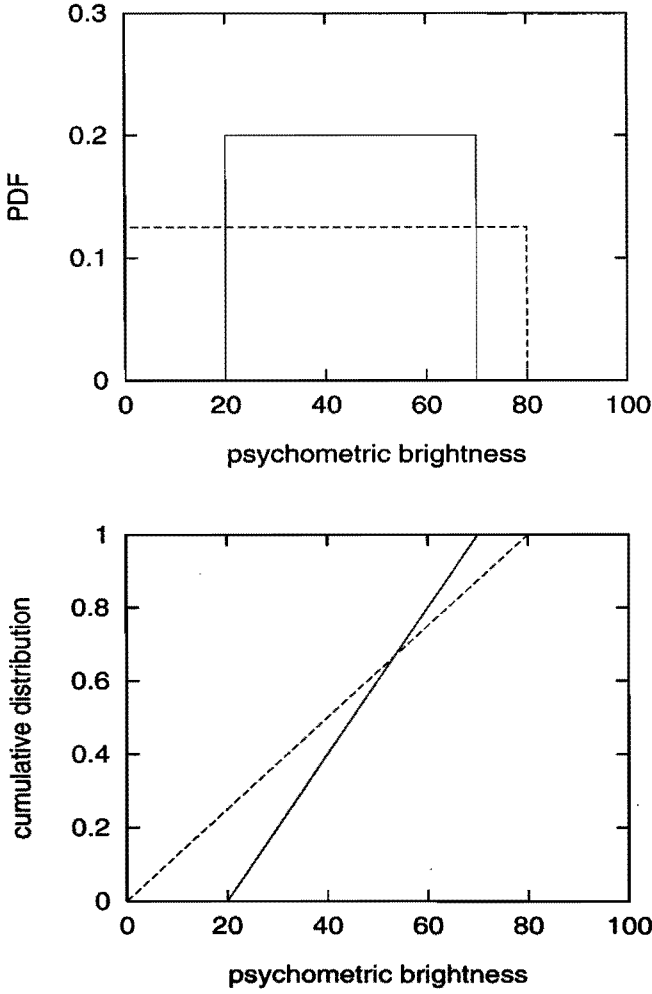


Figure 6.3: Uniform PDFs and their corresponding cumulative distributions, at two different widths of PDF.

$\gamma_p = 0.4$ and 1.6 are shown in Fig. 6.4. Large parts of these cumulative histograms are approximately linear. In the algorithm, the slope of the linear part of the cumulative histograms is estimated. An estimate for the slope ε is obtained by minimizing the error

$$\min_{\varepsilon, K} \sum_i [\varepsilon b_i + K - F(b_i)]^2, \quad (6.3)$$

where $F(b_i)$ is the value of the cumulative histogram of psychometric brightness

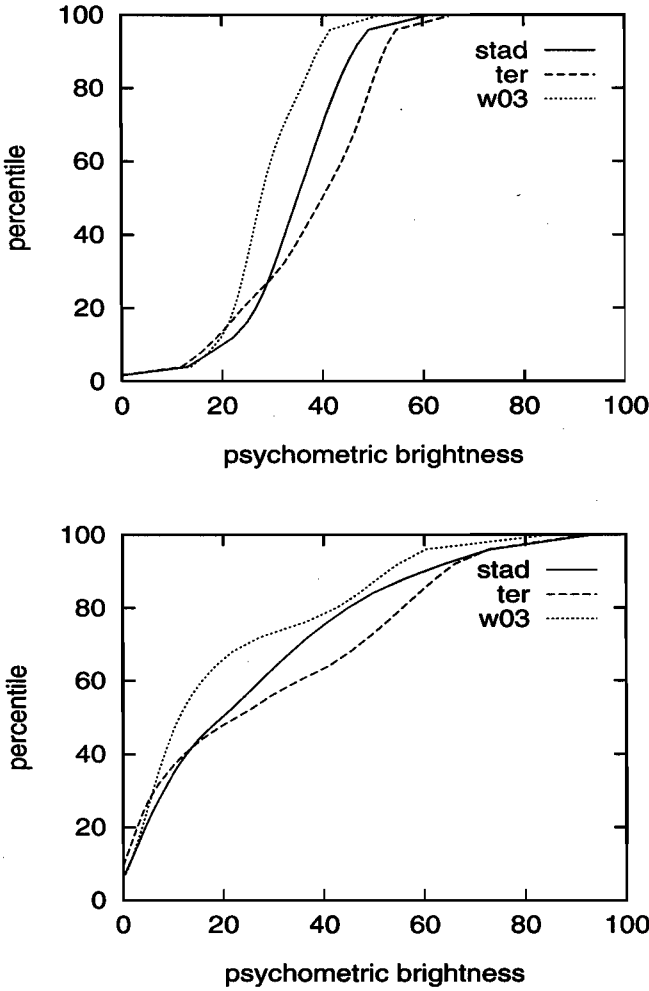


Figure 6.4: Cumulative histograms of approximate psychometric brightness images for 3 scenes. Top at $\gamma_p = 0.4$ and bottom at $\gamma_p = 1.6$.

b_i , at its i -th sampling position. In our implementation, only those samples of $F(b_i)$ for which $b > 5$ are used. A first order approximation to the dynamic range of approximate psychometric brightness in an image is given by the inverse of the slope ε . Therefore we use $C_b = 1/\varepsilon$ as an index for global brightness contrast.

6.3 Results

The contrast estimation algorithm was applied to images used in Experiment 1, reported by Roufs *et al.* (1994). In their experiment, the images were obtained by scanning photographic slides with a slide scanner with exponent 0.4. The contrast of the resulting digital images was varied with an image processor whose exponent γ_p could be varied. The resulting images were displayed on a monitor with exponent $\gamma = 2.5$, $L_{max} = 149 \text{ cd/m}^2$ and $L_{min} = 0.2 \text{ cd/m}^2$. In the experiment γ_p was varied from 0.4 to 2.2, in steps of 0.3. The algorithm that varied the γ_p was designed such that the mean luminance of the displayed image is the same for all images of a scene. The mean luminance was held constant to avoid the effects of variation of mean luminance on perceived contrast and quality. In the psychophysical experiment, the subjects judged the brightness contrast of the images in one session and the quality in another session, both using a 10-point numerical category scale ranging from one to ten.

In Fig. 6.5, the results of the experiment are plotted against the estimates given by the algorithm, for grey-scale images of three natural scenes. The contrast and quality scores are averaged over three subjects. Although there are small differences between the different scenes, the results show that the estimated contrast-index C_b is a fair correlate of perceived global brightness contrast. The plot of perceived quality vs contrast-index C_b shows that the quality is maximum when the contrast-index is approximately 0.906.

Figure 6.6 shows the results of applying the algorithm to images of an experiment identical to that described above, except that in this experiment the images were displayed on a monitor with exponent gamma = 2.0. The figure shows that the estimated contrast-index again correlates well with the perceived contrast. The quality is maximum when the contrast-index is 0.914, which is approximately equal to the optimum value observed in the previous experiment.

These results show that the effective dynamic range over which the brightness in an image varies plays an important role in determining the perceived global brightness contrast of the image. The contrast-index derived here is essentially a first-order approximation to the effective brightness range in the image. We derived the effective dynamic range of brightness based on the histogram or frequency of occurrence of the approximate psychometric brightness in the image. When the γ_p is low, the dynamic range of brightness is small and the image appears to have a low contrast. At this contrast, the visibility of details in the bright regions of the image is poor. On the other hand, when the γ_p is high, the dynamic range of brightness is large and the image appears to have high contrast. At high contrast, the visibility of details in the dark regions of the image deteriorates. The optimum perceived quality is observed at inter-

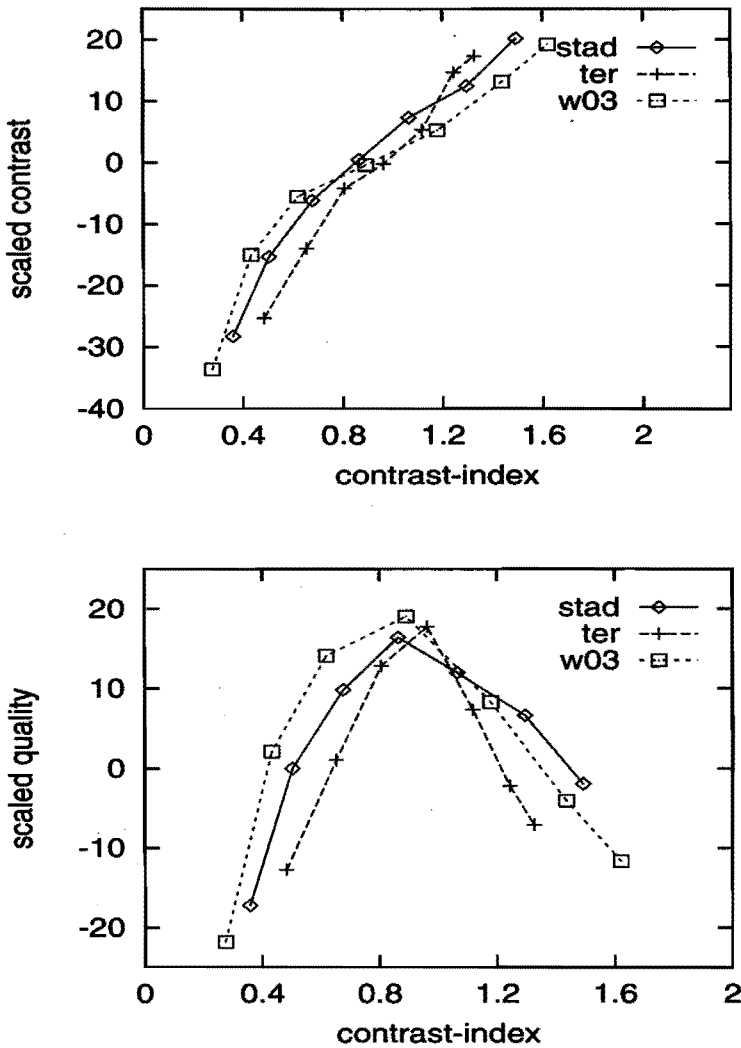


Figure 6.5: Scaled contrast and quality vs estimated global contrast-index C_b , for 3 scenes. Data from Experiment 1 in Roufs *et al.* (1994). The contrast and quality scores are average over three subjects. Display monitor exponent $\gamma = 2.5$.

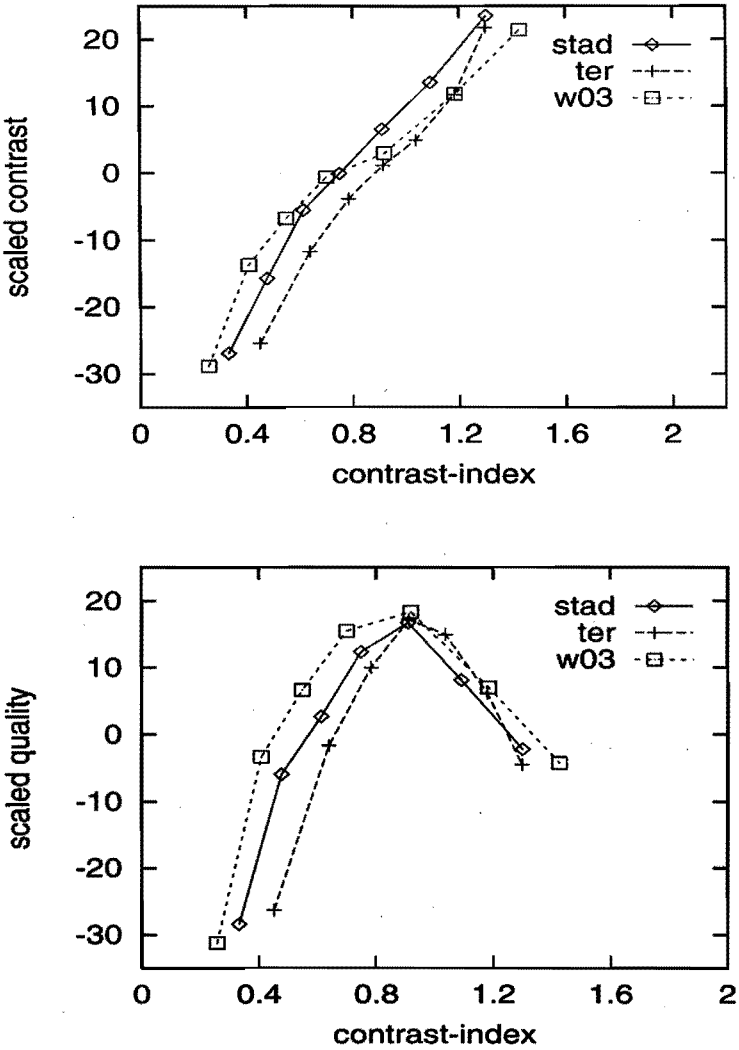


Figure 6.6: Scaled contrast and quality vs estimated global contrast-index for 3 scenes. Data from an experiment similar to that of Experiment 1, in Roufs *et al.* (1994), except that the display monitor exponent γ here is 2.0. The contrast and quality scores are average over three subjects.

mediate values of γ_p , where probably the most natural brightness distribution, causing the most natural brightness contrast, appears. At intermediate values of γ_p , the visibility of details in both light and dark regions of the image is good and this may also be another factors in determining the optimum in perceived quality as a function of perceived global brightness contrast. In addition to this, many other factors may also contribute to the optimum in perceived quality. For example, images at very low contrast values are perceived to be misty and blurred (Roufs, 1989).

6.4 Conclusions

In this chapter, we have presented an objective measure for global brightness contrast of an image. The measure, called contrast-index C_b , can be derived from the image using a simple algorithm. It is based on the slope of the cumulative histogram of an approximate psychometric brightness image, and hence mainly incorporates the area effects on the global brightness contrast. By applying the algorithm to images of natural scenes with different contrast, we have shown that the contrast-index C_b is a good correlate of global brightness contrast. The scaled quality results on the images show that the quality is maximum when the contrast-index C_b is approximately equal to 0.91.

Encouraging results have been obtained by testing the proposed contrast-index estimation algorithm using psychophysical data on three scenes, at two different display monitor exponents. More psychophysical data will be required to rigorously test the usefulness of the proposed contrast-index in imaging system design.

Chapter 7

Epilogue

The perceived quality of images produced by an imaging system plays a significant role in deciding the overall quality of the imaging system. Because of this, imaging system designers are often posed with the problem of attaining maximum quality with given constraints on the system. An example of this is to compare the quality of images produced for different parameter settings of a coding or processing algorithm and to choose the best parameter settings. Another often encountered problem is how to compare the quality of images produced by different algorithms and to choose the best algorithm. Until now these problems have been solved mostly using subjective experiments. Subjective experiments, however, are time-consuming and expensive. Hence, for an efficient and economic solution to the above-stated problems, an objective measure that correlates well with the perceived quality of images is essential.

This thesis has dealt with finding objective measures for perceptual image quality and its attributes, and with developing computational algorithms to estimate these measures from the image itself. In this thesis, the perceptual image quality prediction problem has been addressed from a multi-dimensional point of view. According to this approach, the overall impression of image quality is determined by several underlying perceptual attributes. Examples of these attributes are unsharpness, noisiness, global brightness and global brightness contrast.

Each basic attribute is predominantly affected by one physical parameter of the image. For example, the perceived unsharpness is mainly affected by the spread of the blurring kernel and the noisiness by the standard deviation of the noise (although sometimes in extreme conditions, other parameters also appear to be important). Therefore, estimating these physical parameters from the image is an important step towards computational algorithms to image quality prediction. Two important basic attributes were considered in this thesis: unsharpness in Chapter 2, noisiness in Chapter 3. Computational algorithms based on explicit computational assumptions were used to estimate the physical parameters from the image. Algorithms were also developed to estimate objective measures that correlate well with unsharpness and noisiness.

These objective measures were derived from the physical parameters, based on the sensitivity of the visual system to the changes in the strengths of physical parameters. The estimates of objective measures given by these algorithms were shown to correlate well with the corresponding perceived attributes of the images, even when images contained multiple impairments.

Understanding the relationship between the basic perceptual attributes, and knowing how the basic attributes combine to form the overall impression of perceptual image quality, is essential for estimating objective measures for image quality from the objective measures for the underlying basic attributes. These relationships can be expressed in the perceptual space spanned by the basic attributes. In such a (metric) space, the distance between two image positions is proportional to their dissimilarity and the projection of an image position onto an attributes axis is modelled to be proportional to the strength of the attribute for that image. The perceptual space is constructed using multi-dimensional scaling (MDS) techniques on psychophysical data. In Chapter 4 we constructed the perceptual space of images degraded by blur and noise. In the 2-D perceptual space thus obtained, we found that the attributes unsharpness and noisiness are not orthogonal but are separated by an angle of 80° . The orientation of the quality axis relative to the attributes unsharpness and noisiness was also determined.

Similar to the perceptual space, a psychometric space spanned by the objective measures for the perceptual attributes was introduced in Chapter 5. As stated earlier, these objective measures were derived from the estimates of relevant physical parameters, taking the sensitivity of the visual system into account. An important property of these objective measures is that equal intervals in the objective measures correspond to approximately equal perceived differences in the attributes. Hence, they are psychometric measures. The location of images in the psychometric space was obtained from the estimates of these objective measures. The mapping from the psychometric space to the perceptual space was derived and used to estimate the location of the images in the perceptual space. After locating the images in the perceptual space, their quality was estimated by projecting the stimulus positions onto the quality axis. The concept of a psychometric space and the mapping from it to the perceptual space have helped us in unravelling the relationship between the psychometric and the perceptual spaces. Use of these relationships have also shown encouraging results for image quality prediction.

Imaging system designers can readily use the algorithms developed here to estimate the objective measures, blur-index, noise-index and contrast-index from an image. For example the blur-index can be estimated from an image using the algorithm shown in Fig. 2.24. The luminance image on the display is obtained from the grey-scale image using the luminance transfer characteristics

of the monitor that is used to display the image. For a calibrated monitor, the luminance transfer function can be well approximated by Eq. (2.46). For some monitors, the luminance image may also have to be filtered with the blurring kernel of the monitor, although for monitors with a relatively high resolution this filtering is not needed, since the blur due to the monitor is relatively small compared to that due to the eye-optics. For further computation the image is specified in retinal angles (in minutes of arc). A unit distance (in meters) on the monitor corresponds to $2 \tan^{-1}(1/2d)$ min arc in retinal angle, where d is the viewing distance (also in meters). The luminance image thus obtained is blurred using eye-optics to obtain the retinal image. A Gaussian with spread σ_{b0} in the range 0.5 to 0.75 min arc is a good approximation for optical blur of the eye. The retinal image is transformed into an approximate psychometric brightness image using the transformation shown in Eq. (2.47). The psychometric brightness image computed in this way has values between 0 and 100. Intrinsic noise is then added to this image. White noise with Gaussian PDF and SD, σ_{n0} in the range 2 to 3 units is a good approximation to the intrinsic noise. Finally, the blur estimation algorithm involving local derivative analysis, edge detection and blur estimation is applied to this image to obtain an estimate of blur-spread $\hat{\sigma}_b$, specified in min arc. Blur-index S_b of the image is computed from $\hat{\sigma}_b$ using the equation $S_b = 1 - \sqrt{\sigma_{b0}/\hat{\sigma}_b}$. Similarly, the noise-index S_n and contrast-index C_b can also be estimated from an image using the algorithms shown in Fig. (3.19) and (6.1) in Chapters 3 and 6, respectively.

For images containing only blur as impairment, quality is proportional to the blur-index and for images containing only noise, quality is proportional to the noise-index. The image quality of images degraded by both blur and noise can be predicted from the blur-index and noise-index of the image using the mapping from the psychometric space to the perceptual space. The transformation from the psychometric space to the perceptual space given by Eq. (5.4) is used on the blur-index S_b and noise-index S_n , to obtain the co-ordinates of the image (x, y) in perceptual space. Our results showed that $\phi \approx 80^\circ$ and $\lambda \approx 0.75$ are reasonable values for the parameters in this mapping. Having obtained (x, y) , the image quality is predicted by projecting the image position onto the quality axis. The direction of the quality axis is specified with respect to the unsharpness and noisiness axes using θ_u and θ_n . The direction of unsharpness is along the y axis. Thus, predicted quality $Q_p = \sqrt{x^2 + y^2} \cos[\tan^{-1}(y/x) + \theta_u - 90]$. Our results on limited number of scenes gave a value of $\theta_u \approx 33^\circ$. More experiments will be needed to confirm this value of θ_u .

In the subjective experiments reported in this thesis we have used only three scenes, but each belonged to a different category, a synthetic scene, an outdoor scene and a portrait. The perceptual spaces for the three scenes are found to

be very similar. To generalize our findings about the perceptual space, more data of subjective experiments involving a large number of scenes and subjects is required. Generalizing the results found in this thesis or establishing the classes of images to which the findings apply, is important for transferring the knowledge gained here into imaging system design.

The computational algorithms in this thesis have demonstrated that important physical parameters of images can be reliably estimated by extracting features such as edges, by local processing of the image. A similar approach can also be used to estimate other physical parameters of images, for example, to estimate parameters influencing visibility of structure in sampled and interpolated images (Nijenhuis 1993).

The computational methods developed in this thesis are applicable to black and white still images of natural scenes. Although extension to colour images has not been attempted here, use of the algorithms on the luminance component of the colour images is expected to provide reliable estimates. Extension to moving images (i.e., video) is important and should be considered in future research.

Our main aim in this thesis was to derive objective measures for global perceptual attributes, i.e., deriving one measure for the entire image. For example, we derived objective measures for image unsharpness, image noisiness and image or global brightness contrast. In addition to these global measures, in applications such as image coding and restoration, it is sometimes also essential to have objective measures for local perceptual attributes, such as local unsharpness, local noisiness and local contrast. To derive objective measures for local perceptual attributes, estimates of local parameters of the image are required. The edge parameter estimation method developed in Chapter 2 will be very useful for this purpose. An objective measure for local blur can be estimated from the local edge blur derived using the edge parameter estimation method presented in Chapter 2. A measure for local blur can be derived from the edge blur, by taking the sensitivity of the visual system into account, in the same way as was done while arriving at the blur-index, in Chapter 2. The important physical parameters that influence the local brightness contrast, the edge height, the background value and the edge slope, can also be obtained from the edge parameter estimation method presented in Chapter 2. However, the relation between these physical parameters of the edge and the local contrast has to be experimentally determined to arrive at an objective measure for local brightness contrast. To find a measure for local noisiness, the noise parameters have to be estimated locally. In the uniform regions, this can be done by first detecting the locally uniform regions and then estimating the noise parameters in the uniform regions, by using the techniques described in Chapter 3, separately in each uniform region. As discussed in the Introduc-

tion to Chapter 3, the detection of uniform regions and the noise parameter estimation are inter-connected problems and image segmentation prior to local noise estimation will be required to obtain local estimates of noise parameters. Deriving local measures for noisiness in and around non-uniform regions, such as edges, is even more complex (Girod, 1993).

In this thesis a 2-D *subspace* of the entire N-dimensional (N-D) perceptual space was constructed. The two attributes that we considered are unsharpness and noisiness. Considering only two attributes made the problem tractable. When more attributes are allowed to vary in the experimental setup, psychophysical data collection is expected to become a main bottleneck. For example, in our experiment there were two variables, each with four levels, resulting in 4^2 stimuli. Considering all pairs for dissimilarity judgement resulted in $16^2 = 256$ pairs to be judged. Allowing approximately 10 seconds per viewing and judgement, needed 45 minute for an experiment. Similarly, when 3 attributes are varied $(4^3)^2 = 4096$ pairs have to be judged and for 4 variables it becomes 65,536 pairs. The number of pairs increases as a power function, k^{2n} , where n is the number of variables and k is the number of levels per variable. Therefore, while considering more variables at a time, ways have to be sought to keep the data collection problem tractable. Some partial solutions to this problem, such as discarding pairs that would provide redundant information have been suggested in the literature (Berkum, 1985).

In the blur estimation algorithm presented here, image blur is obtained by a weighted sum of edge blur at prominent edge locations in image, implicitly assuming that blur at prominent edges plays a greater role in deciding the overall blur of an image. In our algorithm prominence is based on the edge height and edge segment length, and the weight is based on the edge height. In reality, however, the prominence may also be decided by the region of attention in the images, driven by higher order features. For example, the prominent higher-order features in a portrait would be the eyes, nose, etc. One way of locating prominent regions in images would be to incorporate information about eye movements. Estimation based on the prominence of image features will be essential, especially while estimating the parameters of video images. These considerations will form the main research goals while developing methods to predict the quality of video images.

We have considered simple distortions in images, blur and noise, that occur in conventional analog imaging systems as well as in digital systems. Nevertheless, the results in this thesis have shown that the problem of image quality prediction is not trivial. Distortions in processed or coded images are relatively complex. An example is the blocking artifact (Rabbani & Jones, 1991). A much deeper understanding of image perception will be required to identify and to arrive at estimates of appropriate physical parameters of images with

such impairments. The issue of integrating local information to form a global impression will play an even more important role while predicting their perceptual attributes and quality. The problem of how local perceptual information, such as local blur or noisiness, combines to form a global impression of the attribute is largely unsolved. More research effort will have to be directed in the future towards solving these underlying problems.

References

- Ahumada, A. J. Jr. (1993). Computational image-quality metrics: A review, *Society for Info. Display, SID 93 Digest*, 305-308.
- Allnatt, J. W., Gleiss, N., Kretz, F., Sciarappa, A. & Zee, E. van der (1983). Definition and the validation of methods of subjective assessment of visual telephone picture quality CSE2T, *Raporti Technici*, XI, 59-65.
- Anderson, G. L. & Netravali, A. (1976). Image restoration based on a subjective criterion, *IEEE Syst. Man Cybern.*, **SMC-6**, 845-853.
- Andrews, H.C. & Hunt, B.R. (1977). *Digital Image Restoration*, Prentice-Hall, Inc, New Jersey.
- Ashby, F. G. & Townsend, J. T. (1986). Varieties of perceptual independence, *Psychological Review*, **93**, 154-179
- Barten, P. (1989). The effects of picture size and definition on perceived image quality, *IEEE Trans. Electr. Dev.*, **36**, 1865-1869.
- Barten, P. (1990). Evaluation of subjective image quality with the square-root integral method, *J. Opt. Soc. Am.*, **7**, 2024-2031.
- Barten, P. (1991). Evaluation of the effect of noise on subjective image quality, *SPIE Human Vision, Visual Processing, and Digital Display II*, SPIE **1453**, 2-15.
- Barten, P. (1993). Effects of quantization and pixel structure on the image quality of color matrix displays, *Journal of the Society for Info. Display SID*, **2**, 147-153.
- Beaton, R. (1983). Quantitative models of image quality, *Proc. Human Factors Society-27th Annual Meeting*, 41-45.
- Berkum, E.E.M. van (1985). Optimal paired comparison designs for factorial experiments, Ph. D. Thesis, Eindhoven University of Technology.
- Besl, P. & Jain, R. (1988). Segmentation through variable-order surface fitting, *IEEE Trans. Pattern Anal. Machine Intell.*, **10**, 167-192.
- Bevington, J. & Mersereau, R. (1987). Differential operator based edge and line detection, *Proc. ICASSP*, 249-252.

- Bracho, R. & Sanderson, A. (1985). Segmentation of images based on intensity gradient information, *Proc. CVPR-85, IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 341-347.
- Budrikis, Z. (1972). Visual fidelity criterion and modeling, *Proc. IEEE*, **60**, 771-779.
- Burt, P. & Adelson, E. (1983). The Laplacian pyramid as a compact image code, *IEEE Trans. Commun.* **31**, 532-540.
- Cannon Jr, M. W. (1984). A study of stimulus range effects in free modulus magnitude estimation of contrast, *Vision Res*, **24**, 1049-1055.
- Canny, J. (1986). A computational approach to edge detection, *IEEE Trans. Pattern Anal. Machine Intell.*, **8**, 679-698.
- Carlson, C. R. & Cohen, R. W. (1980). A simple psychophysical model for predicting the visibility of displayed information, *Proc. Society for Info. Display SID*, **21**, 229-246.
- Carroll, J. D. & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition, *Psychometrika*, 283-319.
- CCIR, (1986). Method for the subjective assessment of the quality of television pictures, Recommendation 500-3. In: Recommendations and reports of the CCIR. International Telecommunication Union, Geneva.
- Chatfield, C. (1983). *Statistics for Technology*, 3rd Ed., Chapman & Hall, London.
- Chen, J. & Medioni, G. (1989). Detection, localization and estimation of edges, *IEEE Trans. Pattern Anal. Machine Intell.* **11**, 191-198.
- Daly, S. (1992). Visible differences predictor: an algorithm for the assessment of image fidelity, in B. E. Rogowitz, (Ed.), *SPIE Human Vision, Visual Processing and Digital Display III*, Proc. SPIE **1666**, 2-15.
- Daly, S. (1993). Visible differences predictor: an algorithm for the assessment of image fidelity, in A. B. Watson (Ed.), *Digital Images and Human Vision*, A Bradford book, MIT Press, London. 179-206.
- Demoment, G. (1989). Image reconstruction and restoration: Overview of common estimation structures and problems, *IEEE Trans. Acoust., Speech, Signal Processing* **37**, 2024-2036.
- Escalante-Ramírez, B. (1992). Perceptually-assessed digital processing of medical images, Ph. D. Thesis, Eindhoven University of Technology.
- Escalante Ramírez, B. & Martens, J. B. (1992). Noise reduction in computerized tomography images by means of polynomial transform, *Journal of Visual Communication and Image Representations* **3**, 272-285.
- Escalante-Ramírez, B., Martens, J. B. & de Ridder, H. (1994). Multi-dimensional perceptual quality of noise-reduced Computed-Tomography

- images, in H.L. Kundel (Ed.), *Medical Imaging 1994: Image Perception, Proc. SPIE-Int. Society for Optical Eng.*, Newport Beach, California, **2166**, 119-130.
- Georgeson, M. A. & Freeman, T. C. A. (1993). Perceived edge blur in one-dimensional and two-dimensional images, *Perception*, **22** Supplement, 21.
- Georgeson, M. A. & Freeman, T. C. A. (1994). A short sharp shock: perceived edge blur in two-dimensional images, *Perception*, **23** Supplement, 76.
- Girod, B. (1989). The information theoretical significance of spatial and temporal masking in video signals, *Human Vision, Visual Processing and Digital Display*, Proc. SPIE **1077**, 178-187.
- Girod, B. (1992). Psychovisual aspects of image communication, *Signal Processing* **28**, 239-251.
- Girod, B. (1993). "What's wrong with mean-squared error?," in A. B. Watson (Ed.), *Digital Images and Human Vision*, A Bradford book, MIT Press, London. 207-220.
- Gonzalez, R. C. & Wintz, P. (1987). *Digital Image Processing*, 2nd Ed., Addison-Wesley, Massachusetts.
- Goodman, J. S. & Pearson, D. E. (1979). Multidimensional scaling of multiply-impaired television pictures, *IEEE Trans. Systems, Man and Cybern.*, **SMC-9**, 353-356.
- Granger, E. M. & Cupery, K. N. (1972). An optical merit function (SQF) which correlates with subjective image judgements, *Phot. Sci. Eng.*, **15**, pp 221-230.
- Granath, D. (1981). The role of human visual models in image processing, *Proc. IEEE*, **69**, 552-561.
- Gray, R., Cosman, P. & Oehler, K. (1993). Incorporating visual factors into vector quantizers for image compression in A. B. Watson (Ed.), *Digital Images and Human Vision*, A Bradford book, MIT Press, London, 35-52.
- Green, P. E., Carmone, F. J. Jr & Smith, S. M. (1989). *Multidimensional Scaling, Concepts and Applications*, Allyn and Bacon, Boston.
- Hall, C. & Hall, E. (1977). A nonlinear model for the spatial characteristics of the human visual system, *IEEE Trans. Syst. Man Cybern.* **SMC-7**, 161-170.
- Hamerly, J. R. & Dvorak, C. A. (1981). Detection and discrimination of blur in edges and lines, *J. Opt. Soc. Am.* **71**, 448-452.
- Hays, W. L. (1988). *Statistics*, 4th Ed., Holt, Rinehart and Winston Inc., New York.
- Higgins, G. C. (1977). Image quality criteria, *J. App. Photo. Eng.*, **3**, 53-60.
- Horn, B. (1986). *Robot Vision*, MIT Press, Cambridge, MA.

- Huang, T. S. (1965). The subjective effect of two-dimensional pictorial noise, *IEEE Trans. Info. Theory*, **11**, 43-53.
- Huertas, A. & Medioni, G. (1986). Detection of intensity changes with sub-pixel accuracy using Laplacian-Gaussian masks, *IEEE Trans. Pattern Anal. Machine Intell.* **8**, 651-664.
- Hultgren, B. (1990). Subjective quality factor revisited (A unification of objective sharpness measures), *Human Vision and Electronic Imaging: Models, Methods and Applications*, Proc. SPIE **1249**, 12-22.
- Hummel, R., Kimia, B. & Zucker, S. (1987). Deblurring Gaussian blur, *Comput. Vision Graphics Image Process.* **38**, 66-80.
- Hunt, B. R. & Sera, G. F. (1978). Power-law stimulus-response models for measures of image quality in nonperformance environments, *IEEE Trans. Syst. Man Cybern.*, **SMC-8**, 781-791.
- Hunt, R. W. G. (1978). Colour Terminology, *Color research and application*, **3**, 79-87.
- Hurvich, L. & Jameson, D. (1966). *The perception of brightness and darkness*, Allyn and Bacon, Inc., Boston.
- Jackson, J. E. (1978). Multidimensional scaling, *Photographic Science and Engineering* **22**, 97-101.
- Kayargadde, V. & Martens, J. B. (1994a). Estimation of edge parameters and image blur using polynomial transforms, *CVGIP: Graphical models and image processing*, **56**, 442-461.
- Kayargadde, V. & Martens, J. B. (1994b). Estimation of edge parameters and image blur using local derivatives, *Journal on Communications*, **XLV**, pp 33-35.
- Kayargadde, V. & Martens, J. B. (1994c). Estimation of perceived blur in images, *Perception*, **23** Supplement, 11.
- Koenderink, J. (1984). The structure of images, *Biolog. Cybern.* **50**, 363-370.
- Koenderink, J. & van Doorn, A. (1990). Receptive field families, *Biolog. Cybern.* **63**, 291-297.
- Kunt, M., Ikonomopoulos, A. & Kocher, M. (1985). Second-generation image-coding techniques, *Proc. IEEE* **73**, 549-574.
- Kunt, M. (1988). Directional image coding in the context of a visual model, in *Working Models in Human Perception*, Academic Press, London.
- Kruskal, J. B. & Wish, M. (1978). *Multidimensional scaling*, Sage University Paper Series on Quantitative Applications in the Social Sciences, series no. 07-011, Beverly Hills and London: Sage Publications.
- Lai, S., Fu, C. & Chang, S. (1993). A generalized depth estimation algorithm with a single image, *IEEE Trans. Pattern Anal. Machine Intell.* **14**,

405-411.

- Lee, J. S. (1981). Refined filtering of image noise using local statistics, *Comput. Vision, Graphics and Image Proc.*, **15**, 380-389.
- Lee, J. S. & Hoppel, K. (1989). Noise modeling and estimation of remotely-sensed images, *Proc. 1989 IEEE International Conference on Geoscience and Remote Sensing, Vancouver 2*, 1005-1008.
- Lillesaeter, O. (1993). Complex contrast, a definition for structured targets and backgrounds *J. Opt. Soc. Am. A* **10**, 2453-2457.
- Limb, J. (1979). Distortion criteria of the human viewer, *IEEE Trans. Syst. Man Cybern.*, **SMC-9**, 779-793.
- Linde, L. (1981). Similarity of distorted pictures: on the interaction between edge blur and random noise, *FOA Report C 53004-H9*, Swedish National Defense Research Institute, Stockholm.
- Lubin, J. (1993). The use of psychophysical data and models in the analysis of display system performance, in A. B. Watson (Ed.), *Digital Images and Human Vision*, A Bradford book, MIT Press, London. 149-162.
- Lukas, F. & Budrikis, Z. (1982). Picture quality prediction based on a visual model, *IEEE Trans. Commun.*, **COM-30**, 1679-1692.
- Lyvers, E. & Mitchell, O. (1988). Precision edge contrast and orientation estimation, *IEEE Trans. Pattern Anal. Machine Intell.* **10**, 927-937.
- Lyvers, E., Mitchell, O., Akey, M. & Reeves, A. (1989). Subpixel measurements using a moment-based edge operator, *IEEE Trans. Pattern Anal. Machine Intell.* **11**, 1293-1308.
- Mallat, S. (1989). Multifrequency channel decompositions of images and wavelet models, *IEEE Trans. Acoust., Speech, Signal Processing* **37**, 2091-2110.
- Mallat, S. & Hwang, W. L. (1992). Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory* **38**, 617-643.
- Mallat, S. & Zhong, S (1992). Characterization of signals from multiscale edges, *IEEE Trans. Pattern Anal. Machine Intell.* **14**, 710-732.
- Mannos, J. & Sakrison, D. (1974). The effects of a visual fidelity criterion on the encoding of images, *IEEE Trans. Inform. Theory*, **IT-20**, 525-536.
- Marmolin, H. & Nyberg, S. (1975). Multidimensional scaling of subjective image quality, *FAO report, C 30039-H9*, Swedish National Defense Research Institute, Stockholm.
- Marmolin, H. & Carlström, A. (1985). Visibility of noise in homogeneous and textured areas, *FAO report, C 50010-H2*, Swedish National Defense Research Institute, Stockholm.

- Marmolin, H. (1986). Subjective MSE measures, *IEEE Trans. Syst. Man Cybern.*, **SMC-16**, 486-489.
- Marr, D. & Hildreth, E. (1980). Theory of edge detection, *Proc. Roy. Soc. London B* **207**, 187-217.
- Marr, D. (1982). *Vision*, W. H. Freeman and Company, New York.
- Martens, J. B. (1990a). The Hermite transform-theory, *IEEE Trans. Acoust., Speech, Signal Processing* **38**, 1595-1606.
- Martens, J. B. (1990b). The Hermite transform-applications, *IEEE Trans. Acoust., Speech, Signal Processing* **38**, 1607-1618.
- Martens, J. B. (1990c). Deblurring digital images by means of polynomial transforms, *Comput. Vision Graphics Image Process.* **50**, 157-176.
- Martens, J. B. (1993). Adaptive image processing by means of polynomial transforms, in B. E. Rogowitz (Ed.), *SPIE Human Vision, Visual Processing and Digital Display III*, SPIE **1666**, 276-287.
- Martens, J. B. (1994). Adaptive contrast enhancement by residue-image processing, Accepted for publication in *Signal Processing*.
- Martens, J. B. (1995). In preparation.
- Meer., P., Jolin, J. & Rosenfeld, A. (1990). A fast parallel algorithm for blind estimation of noise variance, *IEEE Trans. Pattern Anal. Machine Intell.*, **12**, 216-223.
- Nakayama, T., Kurosu, M., Honjyo, K. & Nishimoto, T. (1980). Evaluation and prediction of displayed image quality, *Proc. Society for Info. Display, SID 80 Digest*, 180-181.
- Nalwa, V. & Binford, T. (1986). On detecting edges, *IEEE Trans. Pattern Anal. Machine Intell.* **8**, 699-714.
- Netravali, A. & Prasada, B. (1977). Adaptive Quantization of picture signals using spatial masking, *Proc. IEEE*, **65**, 536-548.
- Nijenhuis, M. (1993). Sampling and Interpolation of static images: a perceptual view, Ph. D. Thesis, Eindhoven University of Technology.
- Ohtsuka, S., Inoue, M. & Watanabe, K. (1989). Quality evaluation of pictures with multiple impairments based on visually weighted errors, *Proc. Society for Info. Display SID*, **30**, 3-8.
- Olsen, S. I. (1993). Estimation of noise in images: an evaluation *CVGIP: Graph. Models Image Proc.*, **55**, 319-323.
- van Overveld, W. M. C. J. (1994). The effect of gamma on subjective quality and contrast of X-ray images, in H.L. Kundel (Ed.), *Medical Imaging 1994: Image Perception, Proc. SPIE-Int. Society for Optical Eng.*, Newport Beach, California, **2166**, 96-104.

- Papoulis, A. (1991). *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, Inc. New York.
- Pearlman, W. (1978). A visual system model and a new distortion measure in the context of image processing, *J. Opt. Soc. Am.*, **68**, 374-386.
- Peli, E. (1990). Contrast in complex images, *J. Opt. Soc. Am. A* **7**, 2032-2040.
- Peli, E., Yang, J., Goldstein, R. & Reeves, A. (1991). Effect of luminance on suprathreshold contrast perception, *J. Opt. Soc. Am. A* **8**, 1352-1359.
- Pentland, A. (1987). A new sense for depth of field, *IEEE Trans. Pattern Anal. Machine Intell.* **9**, 523-531.
- Petrou, M. & Kittler, J. (1991). Optimal edge detectors for ramp edges, *IEEE Trans. Pattern Anal. Machine Intell.* **13**, 483-491.
- Poynton, C. (1993). "Gamma" and its disguises: The nonlinear mappings of intensity in perception, CRTs, film and video, *SMPTE Journal*, **102**, 1099-1108.
- Pratt, W. K. (1991). *Digital Image Processing*, 2nd Ed., John Wiley and Sons, New York.
- Rabbani, M. & Jones, P. W. (1991). *Digital Image Compression Techniques*, SPIE Tutorial Text, SPIE-The International Society for Optical Engineering, Washington.
- Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling, *Psychometrika*, **42**, 241-266.
- Ramsay, J. O. (1991). *MULTISCALE Manual*, McGill University.
- de Ridder, H. (1992). Minkowski-metrics as a combination rule for digital-image-coding impairments, in B. E. Rogowitz, (Ed.), *SPIE Human Vision, Visual Processing and Digital Display III*, Proc. SPIE **1666**, 16-26.
- de Ridder, H. (1994). Personal communication.
- de Ridder, H. & Majoor, G. M. M. (1990). Numerical category scaling: An efficient method for assessing digital image coding impairments, *SPIE Human Vision and Electronic Imaging: Models, Methods and Applications*, Proc. SPIE **1249**, 65-77.
- Rosenfeld, A. & Kak, A. (1982). *Digital Picture Processing*, 2nd Ed., vol. 1 and 2, Academic Press, New York.
- Roufs, J. A. J. (1989). Brightness contrast and sharpness, interactive factors in perceptual image quality, in B.E. Rogowitz, (Ed.), *SPIE Human Vision, Visual Processing and Digital Display* SPIE **1077**, 66-72.
- Roufs, J. A. J. (1992). Perceptual image quality: concept and measurement, *Philips Journal of Research*, **47**, 35-62.
- Roufs, J. A. J., Blommaert, F. & de Ridder, H. (1991). Brightness-luminance relations: future developments in the light of the past, Proc. of 22nd CIE

- session, Melbourne, July 2-11, CIE Pub. no 91, 135-148.
- Roufs, J. A. J. & Boschman, M. C. (1991). Visual comfort and performance, in J. A. J. Roufs (Ed.), *Vision and Visual Dysfunction*, Vol 15, J. R. Cronly-Dillon (Ed.), *The Man-Machine Interphase*, MacMillan Press, London. pp 24-40.
- Roufs, J. A. J. & Bouma, H. (1980). Towards linking perception research and image quality, *Proceedings of the Society for Info. Display SID* **21**, 247-270.
- Roufs, J. A. J. & Goossens, A. (1988). The effect of gamma on perceived image quality, *Proc. 1988 Int. Display Conf.* San Diego, CA IDRC Digest, 27-31.
- Roufs, J. A. J., Koselka, V. & Tongeren, A. van (1994). Global brightness contrast and the effect on perceptual image quality, in Rogowitz and Allebach, (Ed.), *SPIE Human Vision, Visual Processing and Digital Display V*, SPIE, **2179**, 80-89.
- Roufs, J. A. J., de Ridder, H. & Westerink, J. (1990). Perceptual image quality metrics, In *Applied Vision Topical Meeting San Francisco*, Applied Vision Technical Digest Series, Washington: Optical Society of America.
- Safranek, R. & Johnston, J. (1989). A perceptually tuned sub-band image coder with image-dependent quantization and post-quantization data compression, *Proc. ICASSP 89*, **3**, 1945-1948.
- Safranek, R., Johnston, J. & Rosenholtz, R. (1990). A perceptually tuned sub-band coder, *SPIE Human Vision and Electronic Imaging: Models, Methods and Applications*, *Proc. SPIE* **1249**, 284-293.
- Schiffman, S. S., Reynolds, M. L. & Young, F. W. (1981). *Introduction to Multidimensional Scaling, Theory, Methods and Applications*, Academic Press, New York.
- Sjoberg, L. (1987). Psychometric considerations in the dimensional analysis of subjective picture quality, *DISPLAYS Technology and Applications* **8**, 210-212.
- Spence, I. & Lewandowsky, S. (1989). Robust multidimensional scaling, *Psychometrika* **54**, 501-513.
- Stockham, Jr. T. (1972). Image processing in the context of a visual model, *Proc. IEEE*, **60**, 828-842.
- Szegő, G. (1959). *Orthogonal Polynomials*, American Mathematical Society, Colloquium Publication.
- Tabatabai, A. & Mitchell, O. (1984). Edge location to subpixel values in digital imagery, *IEEE Trans. Pattern Anal. Machine Intell.* **6**, 188-201.
- Torgerson, W. S. (1958). *Theory and methods of scaling*, John Wiley and Sons, New York.

- Voorhees, H. & Poggio, T. (1987). Detecting texons and texture boundaries in natural images, *Proc. I IEEE International Conference on Computer vision*, London 250-258.
- Watson, A. B. (1993). *Digital Images and Human Vision*, A. B Watson (Ed.), A Bradford Book, The MIT Press, Cambridge, Massachusetts.
- Watson, A. B., Ahumada, A. J. Jr. & Farrel, J. E. (1986). Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays, *J. Opt. Soc. of Am. A* **3**, 300-307.
- Watson, A. B. & Ahumada, A. J. Jr. (1989). A hexagonal orthogonal oriented pyramid as a model of image representation in visual cortex, *IEEE Trans. Biomedical. Eng.* **BME-36**, 97-106.
- Watt, R. J. & Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision, *Vision Res.* **23**, 1465-1477.
- Westerink, J. (1989). Influences of subject expertise in quality assessment of digitally coded images, *Society for Info. Display, SID International Symposium*, May 15-19, Baltimore, Digest of Technical Papers **20**, 124-127.
- Westerink, J. (1991). Perceived sharpness in static and moving images, Ph. D. Thesis, Eindhoven University of Technology.
- Westerink, J. & Roufs, J. A. J. (1989). Subjective image quality as a function of viewing distance, resolution and picture size, *SMPTE Journal* **98**, 113-119.
- Wilson, H. & Bergen, J. (1979). A four mechanism model for threshold spatial vision, *Vision Res.*, **19**, 19-32.
- Witkin, A. (1984). Scale-space filtering: A new approach to multi-scale description, in *Image Understanding 1984*, ch. 3, 79-95.
- Woods, R. E. & Gonzalez, R. C. (1981). Real-time digital image enhancement, *Proc. IEEE* **69**, 643-654.
- Young, R. (1985). The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line-weighting profiles, General Motors Research Technical Report, GMR-4920.
- Young, R. (1987). The Gaussian derivative model for spatial vision. I. Retinal mechanisms, *Spatial Vision* **2**, 273-293.
- Zetzsche, C. & Hauske, G. (1989). Multiple channel model for the prediction of subjective image quality, *Human Vision, Visual Processing and Digital Display*, *Proc. SPIE* **1077**, 209-216.

Summary

The need to have fast, reliable and economic methods to evaluate images that have been generated, processed or coded in new ways has lead to an increasing demand for techniques that objectively evaluate the perceptual quality of images. This thesis addresses the problem of perceptual image quality prediction, treating image quality as a multidimensional percept, consisting of several underlying perceptual attributes such as sharpness, noisiness and global brightness contrast. Each of these attributes is predominantly influenced by one physical parameter of the image. Therefore, by estimating those relevant physical parameters from the image, and by relating the estimates to the attributes, the attributes can be predicted. The image quality can then be predicted by finding the relation between the attributes and the quality. The goal of the thesis is to develop computational algorithms to estimate image feature parameters from degraded images, and to develop methods to predict image quality based on these estimates.

Computational algorithms are developed to estimate two important image parameters: the spread of the image blurring kernel (Chapter 2) and the standard deviation of the noise in an image (Chapter 3). These physical parameters of the image are known to be the main determinants of perceptual attributes unsharpness and noisiness. The image parameter estimation algorithms are based on perceptually relevant features of the image. For example, blur in an image causes a large perceived change near sharp transitions such as edges and lines. Similarly, noise in an image is most visible in uniform regions of the image. The detection and analysis of these perceptually relevant features is carried out using local derivatives of the image. Specifically, local derivatives of a Gaussian, which have many properties in common with some receptive fields in the human visual system, are used. The blur estimation algorithm detects the pronounced one-dimensional edges in the image and estimates the blur from the relation between the local derivatives of a blurred and the unblurred edge. The noise estimation algorithm is based on the statistics of the energy in the first-derivative in uniform regions of the image. The design, analysis and results of these algorithms are discussed in detail in this thesis. The

results show that the algorithms provide robust estimates of the required image parameters even at very low SNRs.

The perceived attributes of images depend not only on the value of the physical parameters that influence them, but also on the sensitivity of the visual system to changes in the values of the physical parameters. Therefore, objective measures for unsharpness and noisiness are obtained by transforming the physical parameters, where the transformations take the sensitivities into account. The objective measure obtained by transforming blur-spread is called blur-index and that obtained from noise standard deviation is called noise-index. Algorithms have been developed to estimate the blur-index and the noise-index from an image. It has been shown that the estimates of objective measures obtained using these algorithms correlate well with the strengths of the perceived attributes of images, even when images have multiple impairments.

Results of several subjective experiments involving stimuli varying in blur and noise are reported in Chapter 4. Multidimensional scaling (MDS) techniques have been used to analyze these data and to construct the perceptual space of images degraded by blur and noise. The perceptual space is spanned by the important perceptual attributes of images and can be used to understand the relation between the perceived attributes and quality. Using the directions of attributes derived in the perceptual space, it is shown that the attributes unsharpness and noisiness are not orthogonal, but interact weakly, with an angle of approximately 80° between them. It is also shown that the impairment vector, whose direction is opposite to the direction of the quality vector, lies between the unsharpness and noisiness vectors, with an angle of approximately 33° between the unsharpness and the impairment vectors.

Similar to the perceptual space spanned by the basic image attributes, a notion of a psychometric space spanned by the objective (psychometric) measures has been introduced in Chapter 5. The mapping between these two spaces has been derived and used to map images from the psychometric space to the perceptual space. The positions of images in the psychometric space are obtained from the estimates of the objective measures. The strengths of the perceptual attributes are given by the projections of the image positions in the perceptual space onto the attribute axes. Therefore, the perceptual attributes of an image can be estimated by estimating the objective measures for the attributes using the algorithms, i.e., locating the image position in the psychometric space, mapping the image position from psychometric space to the perceptual space, and projecting the position onto the attribute axes. It is shown that the attributes and the quality thus estimated correlate well with the perceived attributes and perceived quality of images.

Similar to unsharpness and noisiness, the global brightness contrast is also an important perceptual attribute of images. An algorithm to estimate a measure for global brightness contrast, called a contrast-index, has been presented in Chapter 6. The contrast-index estimation algorithm is based on the cumulative histogram of an approximate psychometric brightness image. The algorithm essentially captures the effective dynamic range of brightness in an image.

The main findings of this thesis have been summarized in Chapter 7. The computational algorithms developed in this thesis have demonstrated that important physical parameters of images can be reliably estimated from image features such as edges and uniform regions extracted by local processing of the image. The results obtained in this thesis show that the proposed objective measures for unsharpness and noisiness, based on edge features and uniform regions, are good correlates of perceived unsharpness and noisiness. The results also show that the approach used in this thesis, based on the multidimensional nature of the image quality and the mapping between the perceptual space and the psychometric space, is a promising one.

Samenvatting

De behoefte aan snelle, betrouwbare en zuinige methoden voor de evaluatie van op nieuwe manieren gegenereerde, bewerkte of gecodeerde beelden heeft geleid tot een toenemende vraag naar technieken waarmee de perceptieve kwaliteit van beelden objectief geëvalueerd kan worden. Dit proefschrift behandelt het probleem van het voorspellen van perceptieve beeldkwaliteit, waarbij beeldkwaliteit beschouwd wordt als een multidimensionaal percept dat bestaat uit verschillende onderliggende percepten zoals scherpte, ruizigheid en globaal helderheidscontrast. Elk van deze attributen wordt hoofdzakelijk beïnvloed door één fysische parameter van het beeld. Daarom kunnen de attributen voorspeld worden door het schatten van de relevante fysische parameters uit het beeld en het relateren van deze schattingen aan de attributen. De perceptieve beeldkwaliteit kan dan voorspeld worden door het vinden van de relatie tussen de attributen en de kwaliteit. Het doel van dit proefschrift is het ontwikkelen van computationele algoritmen voor het schatten van de parameters van beeldkenmerken uit gedegradeerde beelden en het ontwikkelen van methoden om de beeldkwaliteit te voorspellen op basis van deze schattingen.

Er zijn computationele algoritmen ontwikkeld voor het schatten van twee belangrijke beeldparameters: de helling van de stapresponsie van een filter dat onscherpte in het beeld introduceert (Hoofdstuk 2) en de standaardafwijking van de ruis in een beeld (Hoofdstuk 3). Het is bekend dat deze fysische beeldparameters hoofdzakelijk bepalend zijn voor de perceptieve attributen onscherpte en ruizigheid. De schattingsalgoritmen voor beeldparameters zijn gebaseerd op perceptief relevante kenmerken van het beeld. Onscherpte in een beeld veroorzaakt bijvoorbeeld een grote waargenomen verandering in de buurt van scherpe overgangen zoals randen en lijnen. Evenzo is ruis in een beeld het duidelijkst zichtbaar in uniforme gebieden van het beeld. De detectie en analyse van deze perceptief relevante kenmerken wordt uitgevoerd met behulp van lokale afgeleiden in het beeld. In het bijzonder worden lokale afgeleiden van een Gaussiaan gebruikt, die veel eigenschappen gemeen hebben met sommige receptieve velden in het menselijke visuele systeem. Het schattingsalgoritme voor onscherpte detecteert de uitgesproken ééndimensionale randen

in het beeld en schat de onscherpte uit de relatie tussen de lokale afgeleiden van een onscherpe en de scherpe rand. Het schattingsalgoritme voor ruis is gebaseerd op de statistiek van de energie in de eerste afgeleide in uniforme gebieden van het beeld. Het ontwerp, de analyse en de resultaten van deze algoritmen worden in detail besproken in dit proefschrift. De resultaten tonen aan dat de algoritmen robuuste schattingen van de benodigde beeldparameters opleveren, zelfs bij zeer lage signaal-ruisverhoudingen.

De waargenomen attributen van beelden hangen niet alleen af van de waarde van de fysische parameters die hen beïnvloeden, maar ook van de gevoeligheid van het visuele systeem voor veranderingen in de waarden van die fysische parameters. Daarom worden objectieve maten voor waargenomen onscherpte en ruizigheid verkregen door het transformeren van de fysische parameters, waarbij de transformaties rekening houden met de gevoeligheden. De maat die wordt verkregen door transformatie van de helling van de stapresponsie van het onscherpte-filter wordt de onscherpte-index genoemd, en de maat verkregen uit de standaardafwijking van de ruis wordt de ruis-index genoemd. Er zijn algoritmen ontwikkeld voor het schatten van de onscherpte-index en de ruis-index uit een beeld. Er is aangetoond dat de schattingen van de objectieve maten die met deze algoritmen verkregen worden goed correleren met de waargenomen sterkte van de perceptieve attributen van beelden, zelfs wanneer beelden meervoudige degradaties bevatten.

In hoofdstuk 4 worden de resultaten vermeld van een aantal subjectieve experimenten, waarbij stimuli gebruikt zijn die variëren in onscherpte en ruis. Multidimensionale schalingstechnieken (MDS) zijn gebruikt om deze data te analyseren en om de perceptieve ruimte te construeren van beelden die door onscherpte en ruis gedegradieerd zijn. De perceptieve ruimte wordt opgespannen door de belangrijke perceptieve attributen van beelden en kan gebruikt worden om de relatie tussen waargenomen attributen en kwaliteit te begrijpen. Gebruik makend van de richtingen van attributen zoals afgeleid in de perceptieve ruimte wordt aangetoond dat de attributen onscherpte en ruis niet loodrecht op elkaar staan, maar een zwakke interactie hebben met een hoek van ongeveer 80° ertussen. Er wordt ook aangetoond dat de degradatievector, wiens richting tegengesteld is aan de kwaliteitsvector, tussen de onscherpte- en ruizigheidsvectoren in ligt, met een hoek van ongeveer 33° tussen de onscherpte- en de degradatievector.

Analoog aan de perceptieve ruimte opgespannen door de fundamentele beeldattributen wordt het begrip van een psychometrische ruimte, opgespannen door de objectieve (psychometrische) maten voor de perceptieve attributen, geïntroduceerd in Hoofdstuk 5. De afbeelding tussen deze twee ruimten wordt afgeleid en gebruikt om beelden af te beelden van de psychometrische naar de perceptieve ruimte. De posities van beelden in de psychometrische ruimte

worden verkregen uit de schattingen van de objectieve maten. De sterkten van perceptieve attributen worden gegeven door projectie van de posities van beelden in de perceptieve ruimte op de attribuut-assen. De perceptieve attributen van een beeld kunnen daarom geschat worden door het schatten van de objectieve maten voor de attributen gebruik makend van de algoritmen, dus het bepalen van de positie van het beeld in de psychometrische ruimte, het afbeelden van de positie van het beeld van de psychometrische naar de perceptieve ruimte, en het projecteren van deze positie op de attribuut-assen. Er wordt aangetoond dat de op deze manier geschatte attributen en de kwaliteit goed correleren met de waargenomen attributen en kwaliteit van beelden.

Net als perceptieve onscherpte en ruizigheid is ook het globale helderheidscontrast een belangrijk perceptief attribuut van beelden. In Hoofdstuk 6 wordt een algoritme gepresenteerd voor het schatten van een maat voor globaal helderheidscontrast, de contrastindex genaamd. Het schattingsalgoritme voor de contrastindex is gebaseerd op het cumulatieve histogram van een bij benadering, psychometrische helderheidsbeeld. In wezen bepaalt het algoritme het effectieve dynamische bereik van de helderheid in een beeld.

De belangrijkste bevindingen van dit proefschrift worden nog eens in onderling verband gebracht in Hoofdstuk 7. De in dit proefschrift ontwikkelde computationele algoritmen hebben aangetoond dat belangrijke fysische parameters van beelden betrouwbaar geschat kunnen worden uit beeldkenmerken zoals randen en uniforme gebieden, die door lokale bewerking uit het beeld geëxtraheerd worden. De resultaten verkregen in dit proefschrift laten zien dat de voorgestelde objectieve maten voor onscherpte en ruizigheid, gebaseerd op kenmerken van randen en uniforme gebieden, goed correleren met waargenomen onscherpte en ruizigheid. De resultaten tonen tevens aan dat de in dit proefschrift gebruikte aanpak, gebaseerd op de multidimensionale aard van beeldkwaliteit en de afbeelding tussen perceptieve en psychometrische ruimte, veelbelovend is.

Curriculum Vitae

Vishwakumara Kayargadde was born in Bayar village, India, in 1963. He obtained the B.Sc.Ed. in Science Education from the Regional College of Education, Myrore, in 1985 and the M.E. in Electrical Communication Engineering, from the Indian Institute of Science (IISc), Bangalore, in 1989. He was the recipient of the Alumni Medal and the H.R. Bapu Medal from the Dept of ECE at IISc in 1989. He has worked as a Research Engineer at the Central Research Labs., BEL, Bangalore, in the area of spread spectrum system analysis. In 1990, he was awarded a scholarship from the Eindhoven International Institute (EII) to participate in a postgraduate research program at EII. While at EII, he conducted research at the Philips Research Labs., Eindhoven, on the applicability of Multi-resolution Filtering Techniques for Medical Image Enhancement and developed algorithms for Automatic Text Discrimination. Since 1991, he has been working as a Research Assistant in the Vision Research Group, at the Institute for Perception Research (IPO), Eindhoven, on a project entitled 'Objective Correlates of Subjective Image Quality Dimensions'.

Stellingen

behorende bij het proefschrift

Feature Extraction For Image Quality Prediction

van Vishwakumara Kayargadde

1. The widely used objective measure for quality, the root mean square difference between the original and the processed image, is good only for images impaired by additive white noise but not for those with other impairments such as blur.

This thesis.

Limb J. (1979) Distortion criteria of the human viewer, *IEEE Trans. Syst. Man Cybern.*, SMC-9, 778-793.

2. When the headlights of an automobile are on, the visibility of the front indicator lights is relatively poor. This is mainly due to the glare from the headlights and a higher Weber fraction. The visibility of the indicator lights can be improved considerably by slightly reducing the glare or the intensity of the headlights when the indicator lights are on.
3. Science and common sense have to be clearly distinguished in a classroom.
4. The design of the Dutch ten-cent coin needs reconsideration because of its poor ergonomic properties, especially in relation to its use by the elderly.
5. Cooking is an unique example of human creativity that gives rise to psychological attributes stemming from all five senses, although the weight given to the different attributes varies widely across cultures.

6. The degree of dissatisfaction due to an action, is often a function of the distance between the desired or the expected result and the achieved result. Since it is often not possible to make this distance zero, one way to remain immune to dissatisfaction is to assign a negligible weight to this distance.

The complete works of Swami Vivekananda (1962) Vol 1, 11th Ed., Advaita Ashrama, Calcutta.

7. It took the extermination of quite a few species for mankind to realize that the consequences of wiping out any one species from earth's ecosystem are considerably greater than those of knocking out a brick from a wall.
8. While robots are frequently being designed to function like humans, imitating robots by humans has become a common form of street entertainment.