IMAGE QUALITY STATISTICS AND THEIR USE IN STEGANALYSIS AND COMPRESSION

by

İsmail Avcıbaş

B.S. in E.E., Uludağ University, 1992

M.S. in E.E., Uludağ University, 1994

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor
of
Philosophy

Boğaziçi University

2001

IMAGE QUALITY STATISTICS AND THEIR USE IN STEGANALYSIS AND COMPRESSION

APPROVED BY:

Prof. Bülent Sankur……..……
(Thesis Supervisor)

Assoc. Prof. Lale Akarun….....……

Prof. Emin Anarim……..……

Assoc. Prof. Nasir Memon..……

Assoc. Prof. Yücel Yemez……...……

DATE OF APPROVAL…….

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I would like to cordially thank my supervisor, Professor Bülent Sankur, for welcoming me in his research group, for his continuous help and guidance, for so many fruitful discussions and sharing his ideas. I am still amazed by the diversity and the spectrum of his scientific working knowledge. I cannot emphasize enough how much I benefited from his knowledge, wisdom and personality. It was an honor to know and work with him.

I wish to express my sincere thanks to Professor Nasir Memon for several good reasons. He has been a good friend in so many ways. He has traveled very long distances and participated in my thesis jury. He welcomed me to VIP Lab at Polytechnic University and introduced me to the challenges in visual communications. His supervision, support and confidence in me, is truly appreciated.

I would also like to thank Professors Emin Anarim, Lale Akarun, Yücel Yemez for serving on my thesis committee.

I express my gratitude to graduate studies supervisor Professor Yorgo İstefanopulos and Professor Emin Anarım for their help on a very critical moment in my PhD study.

Grants from TUBİTAK BDP Program and NSF (9724100) are highly acknowledged, as part of this research was made possible.

I was very fortunate in having good teachers. Professor Khalid Sayood introduced me to data compression. I have benefited too much from the information theoretic lectures tutored by Professor Hakan Deliç. I should mention Professor Erdal Panayırcı for the course on pattern recognition.

I would like to express my thanks to the Electronics Engineering Department at Uludağ University for allowing me to do my PhD at Boğaziçi University. Special thanks goes to the chairman Professor Ali Oktay for his support and the very first encouragement;

and to a former Professor at Uludağ University, Haluk Gümüşkaya, as he helped me in developing first interests in signal processing.

Thanks are also due: To Professors Haldun Hadimioğlu, Onur G. Güleryüz and Dr. Kris Popat. To colleagues and friends at Electronics Engineering Department of Uludağ University. To past and present crews of the labs, BUSIM at Boğaziçi and VIP at Poly. To Tahsin Torun, İbrahim Dede, Mahmut Çiftçi, Seçkin Arı, Hasan H. Otu, Ali Oztürk, Veysi Öztürk, Şaban Serin, for being there at the right time when I needed them most.

It is the biggest pleasure to express my gratitude to my beloved parents and sister, as it is their never-ending loving support and encouragement leading me here.

Last, but certainly not least, I wish to convey my insurmountable gratitude to my wife, for her sincere love, for patience and encouragement during my doctoral studies.

ABSTRACT

IMAGE QUALITY STATISTICS AND THEIR USE IN STEGANALYSIS AND COMPRESSION

We categorize comprehensively image quality measures, extend measures defined for gray scale images to their multispectral case, and propose novel image quality measures. The statistical behavior of the measures and their sensitivity to various kinds of distortions, data hiding and coding artifacts are investigated via Analysis of Variance techniques. Their similarities or differences have been illustrated by plotting their Kohonen maps. Measures that give consistent scores across an image class and that are sensitive to distortions and coding artifacts are pointed out.

We present techniques for steganalysis of images that have been potentially subjected to watermarking or steganographic algorithms. Our hypothesis is that watermarking and steganographic schemes leave statistical evidence that can be exploited for detection with the aid of image quality features and multivariate regression analysis. The steganalyzer is built using multivariate regression on the selected quality metrics. In the absence of the ground-truth, a common reference image is obtained based on blurring. Simulation results with the chosen feature set and well-known watermarking and steganographic techniques indicate that our approach is able to reasonably accurately distinguish between marked and unmarked images.

We also present a technique that provides progressive transmission and near-lossless compression in one single framework. The proposed technique produces a bitstream that results in progressive reconstruction of the image just like what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after each layer of the successively refinable bitstream is decoded. Experimental results for both lossless and near-lossless cases are presented, which are competitive with the state-of-the-art compression schemes.

ÖZET

İSTATİSTİKSEL İMGE KALİTESİNİN STEGO-ANALİZDE KULLANIMI VE KODLAMA

İmge nitelik ölçütlerini detaylı bir şekilde sınıflandırdık, gri seviye imgeler için tanımlanmış olan imge nitelik ölçütlerini çok bantlı imgelere genişlettik. Bu ölçütleri değişik bozulumlar, imge kodlama ve damgalama uygulamalarında ortaya çıkan bozulumlar için istatistiksel olarak karşılaştırdık. Bu ölçütler arasındaki benzerlik ve farklılıklar Kohonen haritaları kullanılarak görselleştirildi. Değişinti analizlerine dayanarak imge kodlama, damgalama ve diğer bozulumlara tutarlı ve duyarlı tepki veren ölçütler saptandı.

Damgalanmış imgelerde damga varlığının sezimi ile ilgili yöntemler sunulmaktadır. Belirli bir damgalama yönteminin imgede istatistiksel ve yapısal izler bırakacağı varsayımından yola çıkarak, bu izlerin uygun özniteliklerin seçimi ve çoklu bağlanım analizi ile damga varlığının seziminde kullanılabileceği gösterilmiştir. İyi bilinen damgalama yöntemleri ve zengin bir imge kümesi üzerinde elde edilen görece yüksek doğru sezim yüzdeleri, önerilen yöntemlerin başarılı olduğunu kanıtlamaktadır.

Yeni bir yitimsiz kodlama yöntemi ile, yitimlerin istenen sınırlar içinde kalmasının sağlandığı, dolayısıyla sınırlı yitimli olarak adlandırılan bir kodlama yöntemi geliştirildi. Yöntemimizin bir diğer önemli özelliği, herhangi bir geçişten sonra elde edilen bit dizgisinden geri çatılan imgede pikseller üzerindeki en büyük hatanın sınırlandırılmış olmasıdır. Bu özelliği aşamalı kodlayıcılar sınıfında ön plana çıkan dalgacık tabanlı kodlayıcılar sağlayamamaktadırlar. Yitimsiz ve yitimsize yakın kodlamaya ilişkin deneysel sonuçlar, önerilen yöntemin başarımının bilinen en iyi kodlayıcılar kadar iyi hatta bazı imgeler için daha da iyi olduğunu göstermektedir.

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS / ABBREVIATIONS

$\|\mathbf{a}\|$	Norm of vector $\mathbf{a}$
$\langle \mathbf{a}, \mathbf{b} \rangle$	Inner product of vectors $\mathbf{a}$ and $\mathbf{b}$
$\mathbf{C}$	Multispectral image
$\hat{\mathbf{C}}$	Distorted multispectral image
C_k	k^{th} band of $\mathbf{C}$
$C_k(i, j)$	$(i, j)^{\text{th}}$ pixel of the k^{th} band image
$\mathbf{C}(i, j)$	$(i, j)^{\text{th}}$ multispectral (with K bands) pixel vector
C_k^l	l^{th} block of the k^{th} band image
$C^m(i, j)$	m^{th} scale output of the Derivative of Gaussian operator
$C_w(i, j)$	Pixels in the wxw neighborhood of $C(i, j)$
DCT^{-1}	2-D inverse DCT
$D(p\|q)$	Relative entropy measure between p and q
$d(\cdot,\cdot)$	Distance metric
d_r^k	Distance in r^{th} multiresolution in the k^{th} band
$E(i, j, \sigma_m)$	Edge map at scale σ_m and position (i, j)
E_p	Expectation with respect to p
f	Cover signal
g	Stego signal
$G_m(x, y)$	Derivative of gaussian operator
$H(.)$	Entropy funtion
H^2	Squared Hellinger norm
H_0	Null hypothesis
H_A	Alternative hypothesis
$H(\rho)$	Band-pass filter in polar coordinates modeling HVS
$H(m, n)$	Gaussian blur filter

$(L_m, R_m]$	Interval associated with hypothesis m
$M(u,v)$	Magnitude spectra
p_i	p.m.f. of the pixel at position i
p^*	optimum p.m.f. w.r.t. a constraint
$\{p_{i-j}\}_{j=1}^{N}$	Set of neighboring pmf's of current pixel at position i
$Q(i, j)$	Edge stability map at position (i, j)
$R(p)$	Average bit rate with respect to p
$\{\beta_j\}_{j=1}^{N}$	Set of real-valued linear prediction coefficients
χ_x^m	Indicator function $I\{x \in (L_m, R_m]\}$
Δ_l	l^{th} largest deviation
ε_k	Error over all the pixels in the k^{th} bands, $C_k - \hat{C}_k$
σ^2	Variance
$\varphi(u,v)$	Phase spectra
Γ	Test set
$\Gamma_k(u,v)$	DFT of the k$^{\text{th}}$ band image
$\Gamma_k^l(u,v)$	DFT of the l^{th} block of the k^{th} band image
μ	Mean
T	Training set
$\Omega_k(u,v)$	DCT of the k$^{\text{th}}$ band image
w	Watermark signal
Ψ	Set of features in steganalysis of watermarking techniques
Ω	Set of features in steganalysis of stego techniques
Ξ	Set of features in pooled steganalysis

ANOVA	ANalysis Of VAriance
C1	Normalized Cross-Correlation
C2	Image Fidelity
C3	Czekanowski Correlation

C4	Mean Angle Similarity
C5	Mean Angle-Magnitude Similarity
CALIC	Context Adaptive Lossless Image Compression
CIE	International Commission of Illumination
D1	Mean Square Error
D2	Mean Absolute Error
D3	Modified Infinity Norm
D4	L*a*b* Perceptual Error
D5	Neighborhood Error
D6	Multiresolution Error
DCT	Discrete Cosine Transform
DFT	Discrete Cosine Transform
E1	Pratt Edge Measure
E2	Edge Stability Measure
H1	HVS Absolute Norm
H2	HVS L2 Norm
H3	Browsing Similarity
H4	DCTune
HVS	Human Visual System
i.i.d.	independent and identically distributed
IQM	Image Quality Measure
ITU	International Telecommunications Union
JND	Just Noticeable Difference
JPEG	Joint Pictures Experts Group
KS	Kolmogorov-Smirnov
LSB	Least Significant Bit
MAP	Maximum A Posteriori
ML	Maximum Likelihood
MSE	Mean Square Error
p.d.f.	probability distribution function
p.m.f.	probability mass function
PSNR	Peak SNR
RGB	Red Green Blue

S1	Spectral Phase Error
S2	Spectral Phase-Magnitude Error
S3	Block Spectral Magnitude Error
S4	Block Spectral Phase Error
S5	Block Spectral Phase-Magnitude Error
SNR	Signal to Noise Ratio
SOM	Self Organizing Map
SPIHT	Set Partitioning In Hierarchical Trees
SRC	Spearman Rank Correlation
VQEG	Video Quality Experts Group
Z1	Rate Distortion Measure
Z2	Hellinger distance
Z3	Generalized Matusita distance
Z4	Spearman Rank Correlation

1. INTRODUCTION

1.1. Motivation

There has been an explosive growth in multimedia technology and applications in the past several years. Efficient representation for storage, transmission, retrieval and security are some of the biggest challenges faced.

The first concern addressed in this thesis is the efficient compression of data. Visual information is one of the richest but also most bandwidth consuming modes of communication. However, to meet the requirements of new applications such as mobile multimedia, interactive databases (encyclopedias, electronic newspaper, travel information, and so on) powerful data compression techniques are needed to reduce the bit rate, even in the presence of growing communications channels offering increased bandwidth. Other applications are in remote sensing, education and entertainment.

Not only reducing the bit rate but functionalities of progressive transmission or progressive decoding of the bit stream became more important features of compression schemes. A typical application is data browsing. A user may want to visualize the picture at a lower quality to save transmission time. Another application is tele-radiology where a physician can request portions of an image at increased quality (including lossless reconstruction) while accepting unimportant portions at much lower quality, thereby reducing the overall bandwidth required for transmitting an image.

Although most applications require high compression ratios, this requirement is in general in conjunction with the desire for high quality in the resulting content. Guaranteeing a certain level of quality after compression has become a prime concern for content providers, as the quality in the resulting content is the most important factor in the success of an application in the market place.

A second concern in the thesis is the understanding of image quality metrics. Compression, transmission and sensor inadequacy lead to artifacts and distortions affecting

image quality. Identifying the image quality measures that have highest sensitivity to these distortions would help systematic design of coding, communication and imaging systems and of improving or optimizing the picture quality for a desired quality of service.

The third concern is image security and secret communication. Given the proliferation of digital images, and given the high degree of redundancy present in a digital representation of an image (despite compression), there has been an increased interest in using digital images as cover-objects for the purpose of data hiding. Since unlimited number of copies of an original can be easily distributed or forged, the protection and enforcement of intellectual property rights is an another important issue. A digital watermark is intended to complement cryptographic processes, and is an imperceptible signal added to digital content that can be later detected or extracted in order to make some assertion about the content. Although the main applications for digital watermarking appear to be copyright protection and digital rights management, watermarks have also been proposed for secret communication, that is, steganography. However, despite this obvious and commonly observed connection to steganography, there has been very little effort aimed at analyzing or evaluating the effectiveness of watermarking techniques for steganographic applications. Instead, most work has focused on analyzing or evaluating the watermarking algorithms for their robustness against various kinds of attacks that try to remove or destroy them. If digital watermarks are to be used in steganography applications, detection of their presence by an unauthorized agent defeats their very purpose. Even in applications that do not require hidden communication, but only robustness, we note that it would be desirable to first detect the possible presence of a watermark before trying to remove or manipulate it.

The steganalysis is especially important, as there is a number application areas interested in information hiding, such as

- Military and intelligence agencies require secret communications.
- Terrorists and criminals also place great value on secret communications.
- Law enforcement and counter intelligence agencies are interested in understanding these technologies and their weaknesses, so as to detect and trace hidden messages.

- Schemes for digital elections and digital cash make use of anonymous communication techniques.

1.2. Approaches

The focus of this thesis is on three challenges of visual communications. One is the efficient representation of image data for storage and transmission, which is the art and science of identifying models for different types of structures existing in image data to obtain compact representations. Second is image quality and the third is multimedia security. These are diverse research areas, but also are integral parts of visual communications as a whole. Findings in one field are readily used in the others. For example incorporation of a good image model in a compression scheme decreases the bit rate to represent the image. Still this image model can be used in image data hiding to get a true idea of the data-hiding capacity. Good image models together with quality metrics incorporating human visual system are indispensable in the design of both image coding and watermarking systems as we want visually pleasing, compactly represented and robust images resistant to various attacks.

Image Quality: There is a wealth of research on subjective and/or objective image quality measures to reliably predict either perceived quality across different scenes and distortion types or to predict algorithmic performance computer vision tasks. Our approach is different from the companion studies, in that set of objective image quality measures has been statistically analyzed to identify the ones most sensitive and discriminative to compression, watermarking, blurring and noise distortions. The identified measures are important in the design and performance evaluation of compression, watermarking and imaging systems. Given that an imaging system introduces blurring, by incorporating the most discriminative measure to blurring, among other constraints, and using it as an objective function in the design process would alleviate this problem. Similarities, differences and mutual correlations of image quality measures have been illustrated by plotting Kohonen's self organizing maps. In case one measure is not tractable in the design process, still another mutually correlated and simpler measure can be used. Also, if a combination of measures is required, the redundant measures can be easily eliminated.

Watermark and Stego-mark Detection: Most of the work in watermarking has focused on analyzing or evaluating the watermarking algorithms for their robustness against various kinds of attacks that try to remove or destroy them. However, despite the obvious and commonly observed connection to steganography, there has been very little effort aimed at analyzing or evaluating the effectiveness of watermarking techniques for steganographic applications. General detection techniques as applied to steganography have not been devised and methods beyond visual inspection and specific statistical tests for individual techniques are not present in the literature.

Our approach addresses exactly an automatic detection of the presence of watermark or steganographic marks in images. Hiding information in digital media requires alterations of the signal properties that introduce some form of degradation, no matter how small. These degradations can act as signatures that could be used to reveal the existence of a hidden message. For example, in the context of digital watermarking, the general underlying idea is to create a watermarked signal that is perceptually identical but statistically different from the host signal. A decoder uses this statistical difference in order to detect the watermark. However, the very same statistical difference that is created can potentially be exploited to determine if a given image is watermarked or not. We show that addition of a watermark or message leaves unique artifacts, which can be detected using the most discriminative image quality measures identified from their statistical analysis and multivariate regression analysis. We show that selected image quality measures form a multidimensional feature space whose points cluster well enough to do a classification of marked and non-marked images.

Lossless Progressive Compression: Lossless or reversible compression refers to compression approaches in which the reconstructed data exactly matches the original. Near-lossless compression denotes compression methods, which give quantitative guarantees on the nature of the loss that is introduced. Near-lossless compression is potentially useful in remote sensing, medical imaging, space imaging and image archiving applications, where the huge data size could require lossy compression for efficient storage or transmission. However, the need to preserve the validity of subsequent image analysis performed on the data set to derive information of scientific or clinical value puts strict constraints on the error between compressed image pixel values and their originals. In

such cases, near-lossless compression can be used as it yields significantly higher compression ratios compared to lossless compression and at the same time, the quantitative guarantees it provides on the nature of loss introduced by the compression process are more desirable compared to the uncertainties that are faced when using lossy compression.

Another pillar of this thesis is the proposal of a novel image compression scheme that provides progressive transmission and near-lossless compression in one single framework. We formulate the image data compression problem as one of asking the optimal questions to determine, respectively, the value or the interval of the pixel, depending on whether one is interested in lossless or near-lossless compression. New prediction methods based on the nature of the data at a given pass are presented and links to the existing methods are explored. The trade-off between non-causal prediction and data precision is discussed within the context of successive refinement. Context selection for prediction in different passes is addressed. Experimental results for both lossless and near-lossless cases are presented, which are competitive with the state-of-the-art compression schemes.

1.3. Contributions

The major contributions of this thesis can be highlighted as follows:

- We have presented collectively a set of image quality measures in their multispectral version and categorized them. Image quality measures have been statistically analyzed to identify the ones most sensitive and discriminative to compression, watermarking, blurring and noise distortions. We have also pointed out the image features that should be taken more seriously into account in the design of more successful coding, imaging and data hiding systems. The correlation between various measures has been depicted via Kohonen's Self-Organizing Map. The placement of measures in the two-dimensional map has been in agreement with one's intuitive grouping.

- By using statistically most significant image quality measures, we have developed steganalysis techniques both for conventional LSB like embedding

used in the context of a passive warden model and for watermarking which can be used to embed secret messages in the context of an active warden. The techniques we present are novel and to the best of our knowledge, the first attempt at designing general purpose tools for steganalysis.

- We proposed a novel technique that unifies progressive transmission and near-lossless compression in one single bit stream. The proposed technique produces a bitstream that results in progressive reconstruction of the image just like what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after each layer of the successively refinable bitstream is decoded. Furthermore, the compression performance provided by the proposed technique is superior or comparable to the best-known lossless and near-lossless techniques proposed in the literature. The originality of the method consists in looking at the image data compression as one of asking the optimal questions to determine the interval in which the current pixel lies.

1.4. Outline

We present a set of image quality measures and analyze them statistically with respect to coding, blurring and noise distortions in the second section. The measures are categorized into pixel difference-based, correlation-based, edge-based, spectral-based, context-based and HVS-based (Human Visual System-based) measures. We conduct a statistical analysis of the sensitivity and consistency behavior of objective image quality measures. The mutual relationships between the measures are visualized by plotting their Kohonen maps. Their consistency and sensitivity to coding as well as additive noise and blur are investigated via analysis of variance of their scores.

Using the discrimination power concept of image quality measures, we develop steganalysis techniques for watermarking and steganographic applications in section three. We first describe the steganography problem in terms of prisoner's problem. Next, we show that the distance between an unmarked image and its filtered version is different from the distance between a marked image and its filtered version. This is the critical finding on

which we build the steganalyzer by using a small subset of image quality measures and multivariate regression analysis. Later, we give the design principles and extensively describe the experiments to test the performance of the steganalyzer with a variety of best known watermarking and mostly cited steganographic algorithms on a rich image set.

Key problem in lossless compression is accurate probability mass function estimation. We describe probability mass function estimation methods, data models, and our novel approach to the design of an embedded lossless and near-lossless image compression scheme in section four. This section is compact and the knowledge it requires from sections two and three is at minimal.

Section five concludes the thesis and explores directions for future work.

2. STATISTICAL EVALUATION OF IMAGE QUALITY MEASURES

2.1. Introduction

Image quality measures are figures of merit used for the evaluation of imaging systems or of coding/processing techniques. We consider several image quality metrics and study their statistical behavior when measuring various compression and/or sensor artifacts.

A good objective quality measure should well reflect the distortion on the image due to, for example, blurring, noise, compression, sensor inadequacy. One expects that such measures could be instrumental in predicting the performance of vision-based algorithms such as feature extraction, image-based measurements, detection, tracking, segmentation etc. tasks. Our approach is different from companion studies in the literature focused on subjective image quality criteria, such as in [1, 2, 3]. In the subjective assessment of measures characteristics of the human perception becomes paramount, and image quality is correlated with the preference of an observer or the performance of an operator on some specific task.

In the image coding and computer vision literature, the raw error measures based on deviations between the original and the coded images are overwhelmingly used [4, 5, 6], Mean Square Error (MSE) or alternatively Signal to Noise Ratio (SNR) varieties being the most common measures. The reason for their widespread choice is their mathematical tractability and it is often straightforward to design systems that minimize the MSE. Raw error measures such as MSE may quantify the error in mathematical terms, and they are at their best with additive noise contamination, but they do not necessarily correspond to all aspects of the observer's visual perception of the errors [7, 8], nor do they correctly reflect structural coding artifacts.

For multimedia applications and very low bit rate coding, quality measures based on human perception are being more frequently used [9, 10, 11, 12, 13, 14]. Since a human observer is the end user in multimedia applications, an image quality measure that is based

on a human vision model seems to be more appropriate for predicting user acceptance and for system optimization. This class of distortion measures gives in general a numerical value that will quantify the dissatisfaction of the viewer in observing the reproduced image in place of the original (though Daly's VPD map [13] is a counter example to this). The alternative is subjective tests where subjects view a series of reproduced images and rate them based on the visibility of artifacts [15, 16]. Subjective tests are tedious and time consuming and the results depend on various factors such as observer's background, motivation, etc., and furthermore actually only the display quality is being assessed. Therefore an objective measure that accurately predicts the subjective rating would be a useful guide when optimizing image compression algorithms.

Recently there have been ITU (International Telecommunications Union) efforts to establish objective measurement of video quality. Thus in the context of distribution of multimedia documents, video programming in particular, in-service continuous evaluation of video quality is needed. This continuous video quality indicator would be an input to the network management, which must guarantee a negotiated level of service quality. Obviously such a quality monitoring can only be realized with objective methods [17, 18]. It must be pointed out, however, that subjective assessment, albeit costly and time-consuming, if not impractical, is accurate. Objective methods, on the other hand, can at best try to emulate the performance of subjective methods, utilizing the knowledge of the human visual system.

Similarly for computer vision tasks, prediction of algorithmic performance in terms of imaging distortions is of great significance [19, 20]. In the literature the performance of feature extraction algorithms, like lines and corners [19], propagation of covariance matrices [20], quantification of target detection performance and ideal observer performance [21, 22, 23] have been studied under additive noise conditions. It is of great interest to correlate coding and sensor artifacts with such algorithmic performance. More specifically one would like to identify image quality metrics that can predict accurately and consistently the performance of computer vision algorithms in distorted image records, the distortions being due to compression, sensor inadequacy etc.. An alternative use of image quality metrics is in the inverse mapping from metrics to the nature of distortions [24]. In other words given the image quality metrics, one tries to reconstruct the distortions (e.g.,

blur, noise, etc. amount in a distortion space) that could have resulted in the measured metric values.

Thus in this study we aim to study objective measures of image quality and to investigate their statistical performance. Their statistical behavior is evaluated first, in terms of how discriminating they are to distortion artifacts when tested on a variety of images using Analysis of Variance (ANOVA) method. Second, the measures are investigated in terms of their mutual correlation or similarity, this being put into evidence by means of Kohonen maps.

Some 26 image quality metrics are described and summarized in Table 2.1. These quality metrics are categorized into six groups according to the type of information they are using. The categories used are:

- Pixel difference-based measures such as mean square distortion;
- Correlation-based measures, that is, correlation of pixels, or of the vector angular directions;
- Edge-based measures, that is, displacement of edge positions or their consistency across resolution levels;
- Spectral distance-based measures, that is Fourier magnitude and/or phase spectral discrepancy on a block basis;
- Context-based measures, that is penalties based on various functionals of the multidimensional context probability;
- Human Visual System-based measures, measures either based on the HVS-weighted spectral distortion measures or (dis)similarity criteria used in image database browsing functions.

We define several distortion measures in each category. The specific measures are named as D1, D2.. in the pixel difference category, as C1, C2 .. in the correlation category etc. for ease of reference in the results and discussion sections.

Table 2.1. List of Symbols and Equation Numbers of the Quality Metrics

SYMBOL	DESCRIPTION	EQUATION
D1	Mean Square Error	2.3
D2	Mean Absolute Error	2.1
D3	Modified Infinity Norm	2.3
D4	L*a*b* Perceptual Error	2.4
D5	Neighborhood Error	2.5
D6	Multiresolution Error	2.6
C1	Normalized Cross-Correlation	2.7
C2	Image Fidelity	2.8
C3	Czenakowski Correlation	2.9
C4	Mean Angle Similarity	2.10
C5	Mean Angle-Magnitude Similarity	2.11
E1	Pratt Edge Measure	2.12
E2	Edge Stability Measure	2.13
S1	Spectral Phase Error	2.14
S2	Spectral Phase-Magnitude Error	2.15
S3	Block Spectral Magnitude Error	2.16
S4	Block Spectral Phase Error	2.17
S5	Block Spectral Phase-Magnitude Error	2.18
Z1	Rate Distortion Measure	2.19
Z2	Hellinger distance	2.20
Z3	Generalized Matusita distance	2.21
Z4	Spearman Rank Correlation	2.22
H1	HVS Absolute Norm	2.23
H2	HVS L2 Norm	2.24
H3	Browsing Similarity	2.26
H4	DCTune	

2.2. Image Quality Measures

We define and describe the multitude of image quality measures considered. In these definitions the pixel lattices of images A, B will be referred to as $A(i, j)$ and $B(i, j)$, i, j = 1...N, as the lattices are assumed to have dimensions NxN. The pixels can take values from the set $\{0,...,G\}$ in any spectral band. The actual color images we considered had G = 255 in each band. Similarly we will denote the multispectral components of an image at the pixel position *i, j,* and in band *k* as $C_k(i, j)$, where $k = 1,...,K$. The boldface symbols $\mathbf{C}(i, j)$ and $\hat{\mathbf{C}}(i, j)$ will indicate the multispectral pixel vectors at position (i,j). For example for the color images in the RGB representation one has **C**(i,j) = [*R(i,j) G(i,j) B(i,j)*]^T. These definitions are summarized as follows:

$C_k(i,j)$	$(i,j)^{\text{th}}$ pixel of the k^{th} band of image C
$\mathbf{C}(i,j)$	$(i,j)^{\text{th}}$ multispectral (with K bands) pixel vector
$\mathbf{C}$	A multispectral image
C_k	The k^{th} band of a multispectral image C
$\varepsilon_k = C_k - \hat{C}_k$	Error over all the pixels in the k^{th} band of a multispectral image C.

Thus for example the power in the k'th band can be calculated as $\sigma_k^2 = \sum_{i,j=0}^{N-1} C_k(i,j)^2$. All these quantities with an additional hat, i.e., $\hat{C}_k(i,j)$, $\hat{\mathbf{C}}$ etc., will correspond to the distorted versions of the same original image. As a case in point, the expression $\left\|\mathrm{C}(i,j) - \hat{\mathrm{C}}(i,j)\right\|^2 = \sum_{k=1}^{K}\left[C_k(i,j) - \hat{C}_k(i,j)\right]^2$ will denote the sum of errors in the spectral components at a given pixel position i, j. Similarly the error expression in the last row of the above table expands as $\varepsilon_k^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[C_k(i,j) - \hat{C}_k(i,j)\right]^2$. In the specific case of RGB color images we will occasionally revert to the notations $\{R,G,B\}$ and $\{\hat{R},\hat{G},\hat{B}\}$.

2.2.1. Measures Based on Pixel Differences

These measures calculate the distortion between two images on the basis of their pixelwise differences or certain moments of the difference (error) image.

2.2.1.1. Minkowsky Metrics. The L_γ norm of the dissimilarity of two images can be calculated by taking the Minkowsky average of the pixel differences spatially and then chromatically (that is over the bands):

$$\varepsilon^\gamma = \frac{1}{K}\sum_{k=1}^{K}\left\{\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left|C_k(i,j) - \hat{C}_k(i,j)\right|^\gamma\right\}^{1/\gamma} \tag{2.1}$$

Alternately the Minkowsky average can be first carried over the bands and then spatially, as in the following expression:

$$\varepsilon^{\gamma} = \frac{1}{N^2}\left[\sum_{i,j=0}^{N-1}\left[\frac{1}{K}\sum_{k=1}^{K}\left|C_k(i,j)-\hat{C}_k(i,j)\right|\right]^{\gamma}\right]^{1/\gamma}. \qquad (2.2)$$

In what follows we have used the pixel-wise difference in the Minkowsky sum as given in Eq. (2.1). For γ = 2, one obtains the well-known Mean Square Error (MSE) expression, denoted as *D1*:

$$D1 = \frac{1}{K}\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left\|\mathbf{C}(i,j)-\hat{\mathbf{C}}(i,j)\right\|^2 = \frac{1}{K}\sum_{k=1}^{K}\varepsilon_k^2 . \qquad (2.3)$$

An overwhelming number of quality results in the literature is in fact given in terms of the Signal to Noise Ratio (SNR) or the Peak SNR (PSNR), which are obtained, respectively, by dividing the image power by *D1*, and by dividing the peak power G^2 by D1. Though the SNR and the PSNR are very frequently used in quantifying coding distortions, their shortcomings have been pointed out in various studies [13]. However, despite these oft-cited criticisms of the MSE-based quality measures there has been a recent resurgence of the SNR/PSNR metrics [17, 18]. For example studies of the Video Quality Expert Group (VQEG) [17] have shown that the PSNR measure is a very good indicator of subjective preference in video coding.

For γ = 1 one obtains the absolute difference denoted as *D2*. For γ = ∞ power in the Minkowski average the maximum difference measure

$$\varepsilon^{\infty} = \max_{i,j}\sum_{k=1}^{K}\frac{1}{K}\left|C_k(i,j)-\hat{C}_k(i,j)\right| = \max_{i,j}\|\mathbf{C}(i,j)-\hat{\mathbf{C}}(i,j)\| \qquad (2.4)$$

is obtained. Recall that in signal and image processing the maximum difference or the infinity norm is very commonly used [6]. However given the noise-prone nature of the maximum difference, this metric can be made more robust by considering the ranked list of pixel differences $\Delta_l\left(\mathbf{C}-\hat{\mathbf{C}}\right)$, $l = 1 \ldots N^2$, resulting in a modified Minkowsky infinity metric, called D3. Here $\Delta_l\left(\mathbf{C}-\hat{\mathbf{C}}\right)$ denotes the l^{th} largest deviation among all pixels [25].

Thus $\Delta_1\left(\mathbf{C}-\hat{\mathbf{C}}\right)$ is simply the error expression ε^{∞} above. Similarly Δ_2 correspond to the second largest term etc. Finally a modified maximum difference measure using the first *r* of Δ_m terms can be constructed by computing the root mean square value of the ranked largest differences, $\Delta_1 \ \ldots \ \Delta_r$.

$$D3=\sqrt{\frac{1}{r}\sum_{m=1}^{r}\Delta^2{}_m\left(\mathbf{C}-\hat{\mathbf{C}}\right)} \tag{2.5}$$

2.2.1.2. MSE in Lab Space. The choice of the color-space for an image similarity metric is important, because the color-space must be uniform, so that the intensity difference between two colors must be consistent with the color difference estimated by a human observer. Since the RGB color-space is not well-suited for this task two color spaces are defined: 1976 CIE $L^*u^*v^*$ and the 1976 CIE $L^*a^*b^*$ color spaces [26]. One recommended color-difference equation for the Lab color-space is given by the Euclidean distance [27]. Let

$$\Delta L^*(i,j)=L^*(i,j)-\hat{L}^*(i,j) \tag{2.6}$$

$$\Delta a^*(i,j)=a^*(i,j)-\hat{a}^*(i,j) \tag{2.7}$$

$$\Delta b^*(i,j)=b^*(i,j)-\hat{b}^*(i,j) \tag{2.8}$$

denote the color component differences in $L^*a^*b^*$ space. Then the Euclidean distance is:

$$D4=\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left[\Delta L^*(i,j)^2+\Delta a^*(i,j)^2+\Delta b^*(i,j)^2\right]. \tag{2.9}$$

Note that (2.37) is intended to yield perceptually uniform spacing of colors that exhibit color differences greater than JND threshold but smaller than those in Munsell book of color [27]. This measure applies obviously to color images only and cannot be generalized to arbitrary multispectral images. Therefore it has been used only for the face images and texture images, and not in satellite images.

2.2.1.3. Difference over a Neighborhood. Image distortion on a pixel level can arise from differences in the gray level of the pixels and/or from the displacements of the pixel. A distortion measure that penalizes in a graduated way spatial displacements in addition to gray level differences, and that allows therefore some tolerance for pixel shifts can be defined as follows [28, 29]:

$$D5 = \sqrt{\frac{1}{2(N-w)^2}\sum_{i,j=w/2}^{N-w/2}\left[\min_{l,m\in w_{i,j}}\left\{d\left(\mathbf{C}(i,j),\hat{\mathbf{C}}(l,m)\right)\right\}\right]^2 + \left[\min_{l,m\in w_{i,j}}\left\{d\left(\hat{\mathbf{C}}(i,j),\mathbf{C}(l,m)\right)\right\}\right]^2} \quad (2.10)$$

where $d(\cdot,\cdot)$ is some appropriate distance metric. Notice that for w=1 this metric reduces to the mean square error as in D1. Thus for any given pixel $\mathbf{C}(i,j)$, we search for a best matching pixel in the d-distance sense in the *wxw* neighborhood of the pixel $\hat{\mathbf{C}}(i,j)$, denoted as $\hat{\mathbf{C}}_w(i,j)$. The size of the neighborhood is typically small e.g., 3x3 or 5x5, and one can consider a square or a cross-shaped support. Similarly one calculates the distance from $\hat{C}(i,j)$ to $C_w(i,j)$ where again $C_w(i,j)$ denotes the pixels in the *wxw* neighborhood of coordinates (*i*,*j*) of $C(i,j)$. Note that in general $d\left(\mathbf{C}(i,j),\hat{\mathbf{C}}_w(i,j)\right)$ is not equal to $d\left(\hat{\mathbf{C}}(i,j),\mathbf{C}_w(i,j)\right)$. As for the distance measure $d(\cdot,\cdot)$, the city metric or the chessboard metric can be used. For example city block metric becomes

$$d^{city}\left(\mathbf{C}(i,j),\hat{\mathbf{C}}(l,m)\right) = \frac{\left(|i-l|+|j-m|\right)}{N} + \frac{\left\|\mathbf{C}(i,j)-\hat{\mathbf{C}}(l,m)\right\|}{G} \quad (2.11)$$

where ||.|| denotes the norm of the difference between $\mathbf{C}$(i, j) and $\hat{\mathbf{C}}$(i, j) vectors. Thus both the pixel color difference and search displacement are considered. In this expression *N* and *G* are one possible set of normalization factors to tune the penalties due to pixel shifts and pixel spectral differences, respectively. In our measurements we have used the city block distance with 3x3 neighborhood size.

2.2.1.4 Multiresolution Distance Measure. One limitation of standard objective measures of distance between images is that the comparison is conducted at the full image resolution. Alternative measures can be defined that resemble image perception in the

human visual system more closely, by assigning larger weights to low resolutions and smaller weights to the detail image [30]. Such measures are also more realistic in machine vision tasks that often use local information only.

Consider the various levels of resolution denoted by $r \geq 1$. For each value of *r* the image is split into blocks b_1 to b_n where *n* depends on scale *r*. For example for *r* = 1, at the lowest resolution, only one block covers the whole image characterized by its average gray level *g*. For *r* = 2 one has four blocks each of size $(\frac{N}{2} \text{ x } \frac{N}{2})$ with average gray levels g_{11}, g_{12}, g_{21} and g_{22}. For the r^{th} resolution level one would have than 2^{2r-2} blocks of size $(\frac{N}{2^{r-1}} \text{ x } \frac{N}{2^{r-1}})$, characterized by the block average gray levels g_{ij}, $i, j = 1,...,2^{2r-2}$. Thus for each block b_{ij} of the image C, take g_{ij} as its average gray level and $\hat{g}_{ij}$ to corresponds to its component in the image $\hat{C}$ (For simplicity a third index denoting the resolution level has been omitted). The average difference in gray level at the resolution *r* has weight $\frac{1}{2^r}$. Therefore the distortion at this level is

$$d_r = \frac{1}{2^r} \frac{1}{2^{2r-2}} \sum_{i,j=1}^{2^{r-1}} \left| g_{ij} - \hat{g}_{ij} \right| \tag{2.12}$$

where 2^{r-1} is the number of blocks along either the *i* and *j* indices. If one considers a total of *R* resolution levels, then a distance function can simply be found by summing over all resolution levels, *r = 1,.., R*, that is $D\left(C, \hat{C}\right) = \sum_{r=1}^{R} d_r$. The actual value of *R* (the number of resolution levels) will be set by the initial resolution of the digital image. For example for a 512x512 images one has *R* = 9. Finally for multispectral images one can extend this definition in two ways. In the straightforward extension, one sums the multiresolution distances d_r^k over the bands:

$$D6 = \frac{1}{K} \sum_{k=1}^{K} \sum_{r=1}^{R} d_r^k \tag{2.13}$$

where d_r^k is the multiresolution distance in the k^{th} band. This is the multiresolution distance definition that we used in the experiments. Alternatively the Burt pyramid was constructed to obtain the multiresolution representation. However in the tests it did not perform as well as the pyramid described in [30].

A different definition of the multiresolution distance would be to consider the vector difference of pixels:

$$D(\mathbf{C},\hat{\mathbf{C}}) = \sum_{r=1}^{R} d_r' \quad \text{with} \quad d_r = \frac{1}{2^r}\frac{1}{2^{2r-2}}\sum_{i,j=1}^{2^{r-1}}\left[\left(g_{ij}^R - \hat{g}_{ij}^R\right)^2 + \left(g_{ij}^G - \hat{g}_{ij}^G\right)^2 + \left(g_{ij}^B - \hat{g}_{ij}^B\right)^2\right]^{1/2} \quad (2.14)$$

where, for example, g_{ij}^R is the average gray level of the *ij*'th block in the "red" component of the image at the (implicit) resolution level *r*. Notice that in the latter equation the Euclidean norm of the differences of the block average color components *R*, *G*, *B* have been utilized.

Notice that the last two measures, that is, the neighborhood distance measure and the multiresolution distance measure have not been previously used in evaluating compressed images.

2.2.2. Correlation-Based Measures

2.2.2.1 Image Correlation Measures. The closeness between two digital images can also be quantified in terms of correlation function [5]. These measures measure the similarity between two images, hence in this sense they are complementary to the difference-based measures: Some correlation based measures are as follows:

Structural content

$$C1 = \frac{1}{K}\sum_{k=1}^{K}\frac{\sum_{i,j=0}^{N-1} C_k(i,j)^2}{\sum_{i,j=0}^{N-1}\hat{C}_k(i,j)^2}, \quad (2.15)$$

normalized cross-correlation measure

$$C2 = \frac{1}{K}\sum_{k=1}^{K}\frac{\sum_{i,j=0}^{N-1} C_k(i,j)\hat{C}_k(i,j)}{\sum_{i,j=0}^{N-1} C_k(i,j)^2}. \tag{2.16}$$

A metric useful to compare vectors with strictly non-negative components as in the case of images is given by the Czekanowski distance [31]:

$$C3 = \frac{1}{N^2}\sum_{i,j=0}^{N-1}\left(1 - \frac{2\sum_{k=1}^{K}\min\left(C_k(i,j), \hat{C}_k(i,j)\right)}{\sum_{k=1}^{K}\left(C_k(i,j) + \hat{C}_k(i,j)\right)}\right). \tag{2.17}$$

The Czenakowski coefficient (also called the percentage similarity) measures the similarity between different samples, communities, quadrates.

Obviously as the difference between two images tends to zero $\varepsilon = C - \hat{C} \to 0$, all the correlation-based measures tend to 1, while as $\varepsilon^2 \to G^2$ they tend to 0. Recall also that distance measures and correlation measures are complementary, so that under certain conditions, minimizing distance measures is tantamount to maximizing the correlation measure [32].

2.2.2.2. Moments of the Angles. A variant of correlation-based measures can be obtained by considering the statistics of the angles between the pixel vectors of the original and coded images. Similar "colors" will result in vectors pointing in the same direction, while significantly different colors will point in different directions in the **C** space. Since we deal with positive vectors $\mathbf{C}, \hat{\mathbf{C}}$, we are constrained to one quadrant of the Cartesian space. Thus the normalization factor of $2/\pi$ is related to the fact that the maximum difference attained will be $\pi/2$. The combined angular correlation and magnitude difference between two vectors can be defined as follows [33, 31]:

$$\chi_{ij} = 1 - \left[1 - \frac{2}{\pi} \cos^{-1} \frac{\left\langle \mathbf{C}(i,j), \hat{\mathbf{C}}(i,j) \right\rangle}{\left\| \mathbf{C}(i,j) \right\| \left\| \hat{\mathbf{C}}(i,j) \right\|} \right] \left[1 - \frac{\left\| \mathbf{C}(i,j) - \hat{\mathbf{C}}(i,j) \right\|}{\sqrt{3 \cdot 255^2}} \right] \tag{2.18}$$

We can use the moments of the spectral (chromatic) vector differences as distortion measures. To this effect we have used the mean of the angle difference (*C4*) and the mean of the combined angle-magnitude difference (*C5*) as in the following two measures:

$$C4 = \mu_\chi = 1 - \frac{1}{N^2} \sum_{i,j=1}^{N} \left(\frac{2}{\pi} \cos^{-1} \frac{\left\langle \mathbf{C}(i,j), \hat{\mathbf{C}}(i,j) \right\rangle}{\left\| \mathbf{C}(i,j) \right\| \left\| \hat{\mathbf{C}}(i,j) \right\|} \right), \tag{2.19}$$

$$\mathrm{C5} = \frac{1}{\mathrm{N}^2} \sum_{\mathrm{i,j=1}}^{\mathbf{N}} \chi_{\mathrm{ij}}\,, \tag{2.20}$$

where μ_χ is the mean of the angular differences. These moments have been previously used for the assessment of directional correlation between color vectors.

2.2.3. Edge Quality Measures

According to the contour-texture paradigm of images, the edges form the most informative part in images. For example, in the perception of scene content by human visual system, edges play the major role. Similarly machine vision algorithms often rely on feature maps obtained from the edges. Thus, task performance in vision, whether by humans or machines, is highly dependent on the quality of the edges and other two-dimensional features such as corners [9, 34, 35]. Some examples of edge degradations are: Discontinuities in the edge, decrease of edge sharpness by smoothing effects, offset of edge position, missing edge points, falsely detected edge points etc [32]. Notice however that all the above degradations are not necessarily observed as edge and corner information in images is rather well preserved by most compression algorithms.

Since we do not possess the ground-truthed edge map, we have used the edge map obtained from the original uncompressed images as the reference. Thus to obtain edge-

based quality measures we have generated edge fields from both the original and compressed images using the Canny detector [36]. We have not used any multiband edge detector; instead a separate edge map from each band has been obtained. The outputs of the derivative of gaussians of each band are averaged, and the resulting average output is interpolated, thresholded and thinned in a manner similar to that in [12]. The parameters are set[1] as in [36].

In summary for each band *k=1...K*, horizontal and vertical gradients and their norms, G_x^k, G_y^k and $N^k = \sqrt{G_x^{k\,2} + G_y^{k\,2}}$ are found. Their average over bands is calculated and thresholded with $T = \alpha(T_{max} - T_{min}) + T_{min}$, where $T_{max} = \frac{1}{K}\sum_k \max(N^k)$ and $T_{min} = \frac{1}{K}\sum_k \min(N^k)$, $\alpha = 0.1$. Finally they are thinned by using interpolation to find the pixels where the norms of gradient constitute the local maxima.

2.2.3.1. Pratt Measure. A measure introduced by Pratt [32] considers both edge location accuracy and missing / false alarm edge elements. This measure is based on the knowledge of an ideal reference edge map, where the reference edges should have preferably a width of one pixel. The figure of merit is defined as:

$$E1 = \frac{1}{\max\{n_d, n_t\}} \sum_{i=1}^{n_d} \frac{1}{1 + a d_i^2} \tag{2.21}$$

where n_d and n_t are the number of detected and ground-truth edge points, respectively, and d_i is the distance to the closest edge candidate for the i^{th} detected edge pixel. In our study the binary edge field obtained from the uncompressed image is considered as the "ground truth", or the reference edge field. The factor $\max\{n_d, n_t\}$ penalizes the number of false alarm edges or conversely missing edges.

This scaling factor provides a relative weighting between smeared edges and thin but offset edges. The sum terms penalize possible shifts from the correct edge positions. In

[1] At http://robotics.eecs.berkeley.edu/~sastry/ee20/cacode.html

summary the smearing and offset effects are all included in the Pratt measure, which provides an impression of overall quality.

2.2.3.2. Edge Stability Measure. Edge stability is defined as the consistency of edge evidences across different scales in both the original and coded images [37]. Edge maps at different scales have been obtained from the images by using the Canny [36] operator at different scale parameters (with the standard deviation of the Gaussian filter assuming values σ_m = 1.19, 1.44, 1.68, 2.0, 2.38 (The output of this operator at scale m is thresholded with T^m, where $T^m = 0.1(C_{\max} - C_{\min}) + C_{\min}$. In this expression $C_{\max}$ and $C_{\min}$ denote the maximum and minimum values of the norm of the gradient output in that band. Thus the edge map at scale σ_m of the image C is obtained as

$$E(i,j,\sigma_m) = \begin{cases} 1 & C^m(i,j) > T^m \quad at \quad (i,j) \\ 0 & otherwise \end{cases} \tag{2.22}$$

where $C^m(i,j)$ is the output of the Derivative of Gaussian operator at the m^{th} scale. In other words using a continuous function notation one has $C^m(x,y) = C(x,y) ** G_m(x,y)$ where

$$G_m(x,y) = \frac{1}{2\pi\sigma_m^4} xy \exp\left\{-\frac{x^2+y^2}{2\sigma_m^2}\right\} \tag{2.23}$$

and ** denotes two dimensional convolution. An edge stability map $Q(i,j)$ is obtained by considering the longest subsequence $E(i,j,\sigma_m),...,E(i,j,\sigma_{m+l-1})$ of edge images such that

$$Q(i,j) = l \quad \text{where} \quad l = \arg\max_{l} \bigcap_{\sigma_m \le \sigma_k \le \sigma_{m+l-1}} E(i,j,\sigma_k) = 1. \tag{2.24}$$

The edge stability index calculated from distorted image at pixel position i,j will be denoted by $\hat{Q}(i,j)$. We have used five scales to obtain the edge maps of five band-pass filtered images. Finally a fidelity measure called Edge Stability Mean Square Error (ESMSE) can be calculated by summing the differences in edge stability indexes over all

edge pixel positions, n_d, that is the edge pixels of the ground-truth (undistorted) image at full resolution.

$$E2 = \frac{1}{n_d} \sum_{i,j=0}^{n_d} \left(Q(i,j) - \hat{Q}(i,j) \right)^2 \tag{2.25}$$

For multispectral images the index in (2.51) can be simply averaged over the bands. Alternatively a single edge field from multiband images [38, 39] can be obtained and the resulting edge discrepancies measured as Eq. (2.51).

A property complementary to edge information could be the surface curvature [40], which is a useful feature for scene analysis, feature extraction and object recognition. Estimates of local surface types [41], based on the signs of the mean and Gaussian curvatures, have been widely used for image segmentation and classification algorithms. If one models a gray level image as a 3-D topological surface, then one can analyze this surface locally using differential geometry. A measure based on the discrepancy of mean and Gaussian curvatures between an image and its distorted version has been used in [42]. However this measure was not pursued further due to the subjective assignment of weights to the surface types and the fact that this measure didn't perform particularly well in preliminary tests.

2.2.4. Spectral Distance Measures

In this category we consider the distortion penalty functions obtained from the complex Fourier spectrum of images [10].

Let the Discrete Fourier Transforms (DFT) of the kth band of the original and coded image be denoted by $\Gamma_k(u,v)$ and $\hat{\Gamma}_k(u,v)$, respectively. The spectra are defined as:

$$\Gamma_k(u,v) = \sum_{m,n=0}^{N-1} C_k(m,n) \exp\left[-2\pi i m \frac{u}{N} \right] \exp\left[-2\pi i n \frac{v}{N} \right], \; k = 1...K \tag{2.26}$$

Spectral distortion measures, using difference metrics as for example given in (2.1-2.3) can be extended to multispectral images. To this effect considering the phase and magnitude spectra, that is

$$\varphi(u,v) = \arctan\big(\Gamma(u,v)\big), \tag{2.27}$$

$$M(u,v) = \big|\Gamma(u,v)\big|, \tag{2.28}$$

the distortion occurring in the phase and magnitude spectra can be separately calculated and weighted. Thus one can define the spectral magnitude distortion

$$S = \frac{1}{N^2}\sum_{u,v=0}^{N-1}\left|M(u,v) - \hat{M}(u,v)\right|^2 \tag{2.29}$$

the spectral phase distortion

$$S1 = \frac{1}{N^2}\sum_{u,v=0}^{N-1}\left|\varphi(u,v) - \hat{\varphi}(u,v)\right|^2 \tag{2.30}$$

and the weighted spectral distortion

$$S2 = \frac{1}{N^2}\left(\lambda\sum_{u,v=0}^{N-1}\left|\varphi(u,v) - \hat{\varphi}(u,v)\right|^2 + (1-\lambda)\sum_{u,v=0}^{N-1}\left|M(u,v) - \hat{M}(u,v)\right|^2\right) \tag{2.31}$$

where λ is to be judiciously chosen e.g., to reflect quality judgment. These ideas can be extended in a straightforward manner to multiple band images, by summing over all band distortions. In the following computations, λ is chosen so as to render the contributions of the magnitude and phase terms commensurate, that is $\lambda = 2.5\text{x}10^{-5}$.

Due to the localized nature of distortion and/or the non-stationary image field, Minkowsky averaging of block spectral distortions may be more advantageous. Thus an image can be divided into non-overlapping or overlapping *L* blocks of size *b* x *b*, say

16x16, and blockwise spectral distortions as in (2.14-2.15) can be computed. Let the DFT of the l^{th} block of the k^{th} band image $C_k^l(m,n)$ be $\Gamma_k^l(u,v)$:

$$\Gamma_k^l(u,v) = \sum_{m,n=0}^{b-1} C_k^l(m,n)\exp\left[-2\pi i m\frac{u}{b}\right]\exp\left[-2\pi i n\frac{v}{b}\right] \tag{2.32}$$

where $u,v = -\frac{b}{2}...\frac{b}{2}$ and $l = 1,...,L$,or in the magnitude-phase form

$$\Gamma_k^l(u,v) = \left|\Gamma_k^l(u,v)\right| e^{j\phi_k^l(u,v)} = M_k^l(u,v)e^{\phi_k^l(u,v)}. \tag{2.33}$$

Then the following measures can be defined in the transform domain over the l^{th} block.

$$J_M^l = \frac{1}{K}\sum_{k=1}^{K}\left(\sum_{u,v=0}^{b-1}\left(\left|\Gamma_k^l(u,v)\right| - \left|\hat{\Gamma}_k^l(u,v)\right|\right)^{\gamma}\right)^{1/\gamma} \tag{2.34}$$

$$J_\varphi^l = \frac{1}{K}\sum_{k=1}^{K}\left(\sum_{u,v=0}^{b-1}\left(\left|\phi_k^l(u,v)\right| - \left|\hat{\phi}_k^l(u,v)\right|\right)^{\gamma}\right)^{1/\gamma} \tag{2.35}$$

$$J^l = \lambda J_M^l + (1-\lambda)J_\varphi^l \tag{2.36}$$

with λ the relative weighting factor of the magnitude and phase spectra. Obviously measures (2.36)-(2.38) are special cases of the above definitions for block size b covering the whole image. Various rank order operations on the block spectral differences J_M and / or J_φ can prove useful. Thus let $J^{(1)},...,J^{(L)}$ be the rank ordered block distortions, such that for example $J^{(L)} = \max_l\{J^l\}$. Then one can consider the following rank order averages: Median block distortion $\frac{1}{2}\left(J^{\left(\frac{L}{2}\right)} + J^{\left(\frac{L+1}{2}\right)}\right)$, Maximum block distortion, $J^{(L)}$; and Average block distortion: $\frac{1}{L}\sum_{i=1}^{L} J^{(i)}$. We have found that median of the block

distortions is the most effective averaging of rank ordered block spectral distortions and we have thus used:

$$S3 = \underset{l}{median}\ J_M^l \tag{2.37}$$

$$S4 = \underset{l}{median}\ J_\phi^l \tag{2.38}$$

$$S5 = \underset{l}{median}\ J^l \tag{2.39}$$

In this study we have averaged the block spectra with γ=2 and as for the choice of the block size we have found that block sizes of 32 and 64 yield better results than sizes in the lower or higher ranges.

2.2.5. Context Measures

Most of the compression algorithms and computer vision tasks are based on the neighborhood information of the pixels. In this sense any loss of information in the pixel neighborhoods, that is, damage to pixel contexts could be a good measure of overall image distortion. Since such statistical information lies in the context probabilities, that is the joint probability mass function (p.m.f.) of pixel neighborhoods, changes in the context probabilities should be indicative of image distortions.

A major hurdle in the computation of context distortions is the requirement to calculate the high dimensional joint probability mass function. Typical p.m.f. dimensions would be of the order of *s* = 10 at least. Consequently one incurs “curse of dimensionality problems”. However as detailed in [43, 44], this problem can be solved by judicious usage of kernel estimation and cluster analysis. A modification of the kernel method is to identify the important regions in a *s*-dimensional space X^s by cluster analysis and to fit region-specific kernels to these locations. The result is a model that well represents both mode and tail regions of p.m.f.’s, while combining the summarizing strength of histograms with generalizing strength of kernel estimates.

In what follows we have used the causal neighborhood of pixels i.e., $C_k(i,j)$, $C_k(i-1,j)$, $C_k(i,j-1)$, $C_k(i-1,j-1)$, k = 1, 2, 3. Hence we have derived s = 12 dimensional p.m.f.'s obtained from 4-pixel neighborhoods in the 3-bands.

2.2.5.1. Rate-Distortion Based Distortion Measure. A method to quantify the changes in context probabilities is the relative entropy, defined as

$$D(p\|\hat{p}) = \sum_{\mathrm{x}\in X^s} p(\mathrm{x}) \log \frac{p(\mathrm{x})}{\hat{p}(\mathrm{x})} \tag{2.40}$$

where X^s denotes a s-pixel neighborhood and $\mathbf{x} = [\mathrm{x}_1,...,\mathrm{x}_s]$ a random vector. Furthermore p and $\hat{p}$ are the p.m.f.'s of the original image contexts and that of the distorted (e.g., blurred, noisy, compressed etc.) image. The relative entropy is directly related to efficiency in compression and error rate in classification. Recall also that the optimal average bit rate is the entropy of x

$$H(\mathbf{X}) = -\sum_{\mathbf{X}\in X^s} p(\mathbf{X}) \log p(\mathbf{X}) = R(p). \tag{2.41}$$

If instead of the true probability, a perturbed version $\hat{p}$, that is the p.m.f. of the distorted image, is used, then the average bit rate $R(\hat{p})$ becomes

$$R(\hat{p}) = -\sum_{\mathbf{X}\in X^s} p(\mathbf{X}) \log_2 \hat{p}(\mathbf{X}) = H(\mathbf{X}) + D(p\|\hat{p}). \tag{2.42}$$

The increase in the entropy rate is also indicative of how much the context probability differs from the original due to coding artifacts. However we do not know the true p.m.f. *p*, hence its rate. We can bypass this problem by comparing two competing compression algorithms, in terms of the resulting context probabilities $\hat{p}_1$ and $\hat{p}_2$. If $\hat{p}_1$ and $\hat{p}_2$ are the p.m.f.'s resulting from the two compressed images, then their difference in relative entropy

$$Z1 = D(p\|\hat{p}_1) - D(p\|\hat{p}_2) = R(\hat{p}_1) - R(\hat{p}_2) \tag{2.43}$$

is easily and reliably estimated from a moderate-size sample by subtracting the sample average of $-\log \hat{p}_2$ from that of $-\log \hat{p}_1$ [44]. The comparison can be carried out for more than two images compressed to different bit rates in a similar way, that is comparing them two by two since the unknown entropy term is common to all of them.

As a quality measure for images we have calculated Z1 for each image when they were compressed at two consecutive bit rates, for example, $R(\hat{p}_1)$ at the bit rate of quality factor 90 and $R(\hat{p}_2)$ at the bit rate of quality factor 70, etc. Alternatively the distortion was calculated for an original image and its blurred or noise contaminated version.

2.2.5.2. f-divergences. Once the joint p.m.f. of a pixel context is obtained, several information theoretic distortion measures [45] can be used. Most of these measures can be expressed in the following general form

$$d(p,\hat{p}) = g\left[E_p\left[f\left(\frac{\hat{p}}{p} \right) \right] \right] \tag{2.44}$$

where $\frac{\hat{p}}{p}$ is the likelihood ratio between, $\hat{p}$, the context p.m.f. of the distorted image, p the p.m.f. of the original image, and E_p is the expectation with respect to p. Some examples are as follows:

Hellinger distance, $f(x) = \left(\sqrt{x} - 1\right)^2$, $g(x) = \frac{1}{2}x$

$$Z2 = \frac{1}{2}\int \left(\sqrt{\hat{p}} - \sqrt{p}\right)^2 d\lambda, \tag{2.45}$$

generalized Matusita distance, $f(x) = \left|1 - x^{1/r}\right|^r$, $g(x) = x^{1/r}$

$$Z3 = \sqrt{\int \left| p^{1/r} - \hat{p}^{1/r} \right|^r d\lambda} \qquad r \geq 1. \tag{2.46}$$

Notice that integration in (2.44)-(2.45) are carried out in *s*-dimensional space. Also we have found according to ANOVA analysis that the choice of $r = 5$ in the Matusita distance, yields good results. Despite the fact that the p.m.f.'s do not reflect directly the structural content or the geometrical features in an image, they perform sufficiently well to differentiate artifacts between the original and test images.

2.2.5.3. Local Histogram Distances. In order to reflect the differences between two images at a local level, we calculated the histograms of the original and distorted images on the basis of 16x16 blocks. To this effect we considered both the Kolmogorov-Smirnov (KS) distance and the Spearman Rank Correlation (SRC).

For the KS distance we calculated the maximum deviation between the respective cumulatives. For each of the 16x16 blocks of the image, the maximum of the KS distances over the K spectral components was found and these local figures were summed over all the blocks to yield $\sum_{u=1}^{b} \max_{k=1..K} \{KS_u^k\}$ where $\mathrm{KS}_\mathrm{b}^\mathrm{k}$ denotes the Kolmogorov-Smirnov distance of the block number u and of the k^{th} spectral component. However the KS distance did not turn out to be effective in the ANOVA tests. Instead the SRC measure had a better performance. We again considered the SRC on a 16x16 block basis and we took the maximum over the three spectral bands. The block SRC measure was computed by computing the rank scores of the "gray" levels in the bands and their largest for each pixel neighborhood. Finally the correlation of the block ranks of the original and distorted images is calculated.

$$Z4 = \sum_{u=1}^{b} \max_{k=1..K} \{SRC_u^k\} \tag{2.47}$$

where SRC_u^k denotes the Spearman Rank Correlation for the u^{th} block number and k^{th} spectral band.

2.2.6. Human Visual System Based Measures

Despite the quest for objective image distortion measure it is intriguing to find out the role of HVS based measures. The HVS is too complex to be fully understood with present psychophysical means, but the incorporation of even a simplified HVS model into objective measures reportedly [7, 46, 10, 14] leads to a better correlation with the subjective ratings. It is conjectured therefore that in machine vision tasks they may have as well some relevance.

2.2.6.1. HVS Modified Spectral Distortion. In order to obtain a closer relation with the assessment by the human visual system, both the original and coded images can be preprocessed via filters that simulate the HVS. One of the models for the human visual system is given as a band-pass filter with a transfer function in polar coordinates [46]:

$$H(\rho) = \begin{cases} 0.05e^{\rho^{0.554}} & \rho < 7 \\ e^{-9\left[\left|\log_{10}\rho - \log_{10} 9\right|\right]^{2.3}} & \rho \geq 7 \end{cases} \tag{2.48}$$

where $\rho = \left(u^2 + v^2\right)^{1/2}$. Image processed through such a spectral mask and then inverse DCT transformed can be expressed via the $U\{\cdot\}$ operator, i.e.,

$$U\{C(i,j)\} = DCT^{-1}\left\{H\left(\sqrt{u^2+v^2}\right)\Omega(u,v)\right\} \tag{2.49}$$

where $\Omega(u,v)$ denotes the 2-D Discrete Cosine Transform (DCT) of the image and DCT^{-1} is the 2-D inverse DCT.

Some possible measures [5, 47, 48, 49] for the K component multispectral image are

normalized absolute error

$$H1 = \frac{1}{K}\sum_{k=1}^{K}\left(\sum_{i,j=0}^{N-1}\left|U\{C_k(i,j)\} - U\{\hat{C}_k(i,j)\}\right| / \sum_{i,j=0}^{N-1}\left|U\{C_k(i,j)\}\right|\right), \tag{2.50}$$

L2 norm

$$H2 = \frac{1}{K}\sum_{k=1}^{K}\left\{\frac{1}{N^2}\sum_{i,j=0}^{N-1}\left|U\{C_k(i,j)\}-U\{\hat{C}_k(i,j)\}\right|^2\right\}^{1/2}, \qquad (2.51)$$

normalized mean square HVS error

$$H = \frac{1}{K}\sum_{k=1}^{K}\left(\sum_{i,j=0}^{N-1}\left[U\{C_k(i,j)\}-U\{\hat{C}_k(i,j)\}\right]^2 / \sum_{i,j=0}^{N-1}\left[U\{C_k(i,j)\}\right]^2\right). \qquad (2.52)$$

2.2.6.2. A Distance Metric for Database Browsing. The metric proposed in [14, 50] based on a multiscale model of the human visual system, has actually the function of bringing forth the similarities between image objects for database search and browsing purposes. This multiscale model includes channels, which account for perceptual phenomena such as color, contrast, color-contrast and orientation selectivity. From these channels, features are extracted and then an aggregate measure of similarity using a weighted linear combination of the feature differences is formed. The choice of features and weights is made to maximize the consistency with similarity.

We have adopted this database search algorithm to measure discrepancies between an original image and its distorted version. In other words an image similarity metric that was conceived for browsing and searching in image databases was adapted to measure the similarity (or the difference) between an image and its distorted version.

More specifically, we exploit a vision system designed for image database browsing and object identification to measure image distortion. The image similarity metric in [14] uses 102 feature vectors extracted at different scales and orientations both in luminance and color channels. The final (dis)similarity metric is

$$H3 = \sum_{i=1}^{102}\omega_i d_i \qquad (2.53)$$

where ω_i are their weights as attributed in [50] and d_i are the individual feature discrepancies. We call this metric "browsing metric" for the lack of a better name. For example the color contrast distortion at scale *l* is given by

$$d_\mu = \frac{1}{N_l N_l} \sum_{i,j=0}^{N_l} \left(K(i,j) - \hat{K}(i,j) \right)^2 \tag{2.54}$$

where $N_l \text{x} N_l$ is the size of the image at scale *l*. $K(i,j)$ and $\hat{K}(i,j)$ denote any color or contrast channel of the original image and of the coded image at a certain level *l*. The lengthy details of the algorithm and its adaptation to our problem are summarized in [14, 50]. Finally note that this measure was used only for color images, and not in the case of satellite three-band images.

The last quality measure we used that reflects the properties of the human visual system was the DCTune algorithm [56]. DCTune is in fact a technique for optimizing JPEG still image compression. DCTune calculates the best JPEG quantization matrices to achieve the maximum possible compression for a specified perceptual error, given a particular image and a particular set of viewing conditions. DCTune also allows the user to compute the perceptual error between two images in units of JND (just-noticeable differences) between a reference image and a test image. This figure was used as the last metric (*H4*) in Table 2.1.

2.3. Goals and Methods

2.3.1 Quality Attributes

Objective video quality model attributes have been studied in [17, 18]. These attributes can be directly translated to the still image quality measures in the multimedia and computer vision applications.

Prediction Accuracy: The accurate prediction of distortion, whether for algorithmic performance or subjective assessment. For example when quality metrics are shown in box plots as in Fig. 2.1., an accurate metric will possess a small scatter plot.

Prediction Monotonicity: The objective image quality measure's scores should be completely monotonic in their relationship to the performance scores.

Prediction Consistency: This attribute relates to the objective quality measure's ability to provide consistently accurate predictions for all types of images and not to fail excessively for a subset of images.

These desired characteristics reflect on the box plots and the F scores of the quality metrics, as detailed in the sequel.

2.3.2. Test Image Sets and Rates

All the image quality measures are calculated in their multiband version. In the study of the quality measures in image compression, we used the two well-known compression algorithms: The popular DCT based JPEG [51] and wavelet zero-tree method, Set Partitioning in Hierarchical Trees (SPIHT), due to Said and Pearlman [52]. The other types of image distortions are generated by the use of blurring filters with various support sizes and by the addition of white Gaussian noise at various levels.

The rate selection scheme was based on the accepted rate ranges of JPEG. It is well-known that the JPEG quality factor Q between 80-100 corresponds to visually imperceptible impairment, Q between 60-80 is the perceptible but not annoying distortion, Q between 40-60 becomes slightly annoying, Q between 20-40 is annoying, and finally 01-20 is the Q range where the quality is very annoying. Thus each image class was compressed with 5 JPEG Q factors of 90, 70, 50, 30 and 10. For each class the average length of compressed files was calculated and the corresponding bit rate (bit/pixel) was accepted as the class rate. The same rate as obtained from the JPEG experiment was also used in the SPIHT algorithm.

The test material consisted of the following image sets: 1) Ten three-band remote sensing images, which contained a fair amount of variety, i.e., edges, textures, plateaus and

contrast range, 2) Ten color face images from Purdue University Face Images database [53], 3) Ten texture images from MIT Texture Database (VISTEX)[2].

2.3.3. Analysis of Variance

The Analysis of Variance (ANOVA) [54] was used as a statistical tool to put into the evidence merits of quality measures. In other words ANOVA was used to show whether the variation in the data could be accounted for by the hypothesized factor, for example the factor of image compression type, the factor of image class etc.

The purpose of a one-way ANOVA is to find out whether data from several groups have a common mean. That is, to determine whether the groups are actually different in the measured characteristic. In our case each "compression group" consists of quality scores from various images at a certain bit rate, and there are k = 5 groups corresponding to the 5 bit rates tested. Each group had 30 sample vectors since there were 30 multispectral test images (10 remote sensing, 10 faces, 10 textures). Similarly three "blur groups" were created by low-pass filtering the images with 2-D Gaussian-shaped filters with increasing support. Finally three "noise groups" were created by contaminating the images with Gaussian noise with increasing variance, that is σ^2 = 200, 600, 1700. This range of noise values spans the noisy image quality from the just noticeable distortion to annoying degradation.

The purpose of two-way ANOVA is to find out whether data from several groups have a common mean. One-way ANOVA and two-way ANOVA differ in that the groups in two-way ANOVA have two categories of defining characteristics instead of one. Since we have two coders (i.e., JPEG and SPIHT algorithms) two-way ANOVA is appropriate. The hypotheses for the comparison of independent groups are:

H_{01}: $\mu_{11} = \mu_{12} = ... = \mu_{1k}$ means of the groups w.r.t. first factor are equal,

H_{A1}: $\mu_{1i} \neq \mu_{1j}$ means of the groups w.r.t. first factor are not equal.

H_{02}: $\mu_{21} = \mu_{22} = ... = \mu_{2k}$ means of the groups w.r.t. second factor are equal,

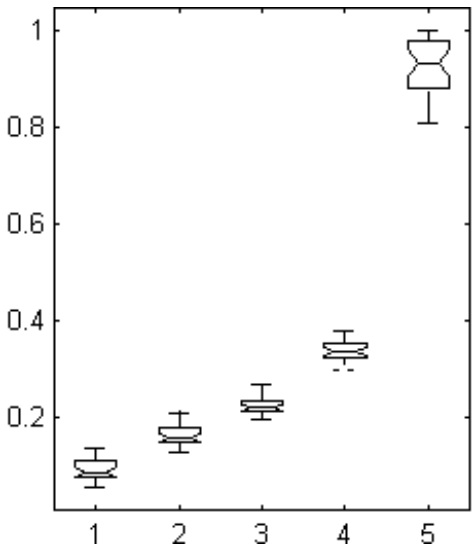
[2] http://www-white.media.edu/vismod/imagery/VisionTexture/vistex.html

H_{A2}: $\mu_{2i} \neq \mu_{2j}$ means of the groups w.r.t. second factor are not equal.

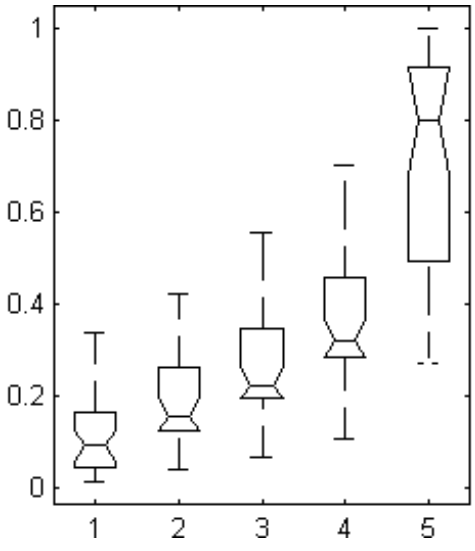
It should be noted that the test statistic is an *F* test with *k-1* and *N-k* degrees of freedom, where *N* is the total number of compressed images. ANOVA returns the p-value for the null hypothesis that the means of the groups are equal [54]. A low *p*-value (high *F* value) for this test indicates evidence to reject the null hypothesis in favor of the alternative. In other words, there is evidence that at least one pair of means are not equal. We have opted to carry out the multiple comparison tests at a significance level of 0.05. Thus any test resulting in a *p*-value under 0.05 would be significant, and therefore, one would reject the null hypothesis in favor of the alternative hypothesis. This is to assert that the difference in the quality metric arises from the image coding artifacts and not from random fluctuations in the image content.

To find out whether the variability of the metric scores arises predominantly from the image quality, and not from the image set, we considered the interaction between image set and the distortion artifacts (i.e., compression bit rate, blur etc.). To this effect we considered the F-scores with respect to the image set as well. As discussed in sub-section 2.3.1. and shown in Tables 2.2.-2.3, metrics that were sensitive to distortion artifacts were naturally sensitive to image set variation as well. However for the “good” measures that identified the sensitivity to image set variation was always inferior to the distortion sensitivity.

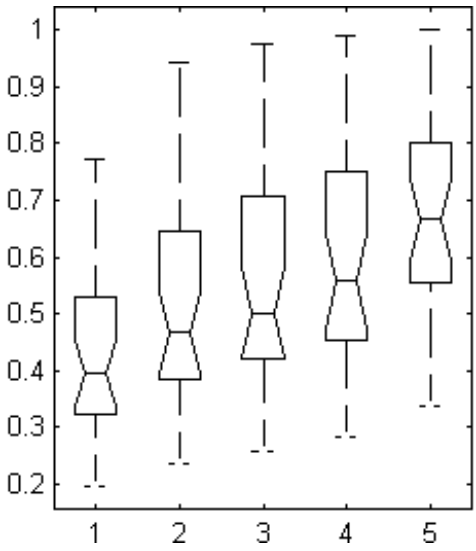
Boxplot is a garphical way of looking at the distribution of the data in different groups. Boxplot produces a box and whisker plot for each group. The box has lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box to show the extent of the rest of the data. Outliers are data with values beyond the ends of the whiskers. If the *F*-value is high, there will be little overlap between the two or more groups. If the *F*-value is not high, there will be a fair amount of overlap between all of the groups. In the box plots, a steep slope and little overlap between boxes, as illustrated in Figure 2.1, are both indicators of a good quality measure. In order to quantify the discriminative power of a quality measure, we have normalized the difference of two successive group means normalized by respective variances, i.e.,



a) H2 JPEG F=2291, p=0



b) D1 JPEG F=104.6, p=0



c) C4 SPIHT F=7.91, p=0

Figure 2.1. Box plots of quality measure scores. a) good measure, b) moderate measure, c) poor measure. The F-scores as well as the significance level p are given.

$$Q_{r,r+1} = \frac{\mu_r - \mu_{r+1}}{\sqrt{\sigma_r \sigma_{r+1}}} \tag{2.55}$$

$$Q = Ave\{Q_{r,r+1}\} \qquad r = 1,...,k-1 \tag{2.56}$$

where μ_r denotes the mean value of the image quality measure for the images compressed at rate *r* and σ_r is the standard deviation, *k* is the number of different bit rates at which quality measures are calculated. A good image quality measure should have high *Q* value, which implies little overlap between groups and/or large jumps between them hence high discriminative power of the quality measure. It should be noted that the Q values and the F-scores yielded totally parallel results in our experiments. In Figure 2.1. we give box plot examples of a good, a moderate and a poor measure. For the box plot visualization the data has been appropriately scaled without any loss of information. The horizontal axis corresponds to bitrate variation and the vertical axis is the normalized IQM scores.

2.3.4. Visualization of Quality Metrics

Since we would like to visualize the quality metrics data, we organize them as vectors and feed them to a SOM (Self-Organizing Map) algorithm. The elements of the vectors are the corresponding quality scores. For example, consider the MSE error (D1) for a specific compression algorithm (e.g., JPEG) at a specific rate. The corresponding vector **D1** is M dimensional, where M is the number of images, and it reads as:

$$\mathbf{D1}(\text{JPEG, bitrate}) = [D1(1|\text{ JPEG, bitrate}) \ldots. \ D1(M|\text{ JPEG, bitrate}]^T \qquad (2.57)$$

There will be 5 such vectors, one for each bit rate considered. Overall for training of SOM we utilize 30 images x 5 bit rates x 2 compressors x 26 metrics = 7800 vectors.

Recall that the self-organizing map (SOM) is a tool for visualization of high dimensional data. It maps complex, non-linear high dimensional data into simple geometric relationships on a low dimensional array and thus serves to produce abstractions. Among the important applications of the SOM one can cite the visualization of high dimensional data, as the case in point, and discovery of categories and abstractions from raw data.

Let the data vectors be denoted as $\mathbf{X} = [x_1,\ldots, x_M]^T \in R^M$, where *M* is the number of images considered (*M* = 30 in our case). With each element in the SOM array, a

parametric real vector $\mathbf{m}_i = [\mu_{i1},...,\mu_{iM}]^T \in R^M$ is associated. The image of an input vector $\mathbf{X}$ on the SOM array is defined by the decoder function $\mathrm{d}(\mathbf{X},\mathbf{m}_\mathrm{i})$, where $d(.,.)$ is a general distance measure. The image of the input vector will have the array index c definedas $\mathrm{c} = \arg\min_{\mathrm{i}} \mathrm{d}(\mathbf{X},\mathbf{m}_\mathrm{i})$. A critical part of the algorithm is to define the $\mathbf{m}_i$ in such a way that the mapping is ordered and descriptive of distribution of $\mathbf{X}$. Finding such a set of values that minimize the distance measure resembles in fact the standard VQ problem. In contrast, the indexing of these values is arbitrary, whereby the mapping is unordered. However if the minimization of the objective functional based on the distance function is implemented under the conditions described in [55], then one can obtain ordered values of $\mathbf{m}_i$, almost as if the $\mathbf{m}_i$ were lying at the nodes of an elastic net. With the elastic net analogy in mind, SOM algorithm can be constructed as

$$\mathbf{m}_\mathrm{i}(\mathrm{t}+1) = \mathbf{m}_\mathrm{i}(\mathrm{t}) + \alpha(\mathrm{t})[\mathbf{X}(\mathrm{t}) - \mathbf{m}_\mathrm{i}(\mathrm{t})] \tag{2.58}$$

where $\alpha(\mathrm{t})$ is a small scalar, if the distance between units c and i in the array is smaller than or equal to a specified limit (radius), and $\alpha(\mathrm{t}) = 0$ otherwise. During the course of ordering process, $\alpha(t)$ is decreased from 0.05 to 0.02, while radius of neighborhood is decreased from 10 to 3. Furthermore scores are normalized with respect to the range.

The component planes j of the SOM, i.e., the array of scalar values μ_{ij} representing the j'th components of the weight vectors $\mathbf{m}_i$ and having the same format as the SOM array is displayed as shades of gray.

2.4. Statistical Analysis of Image Quality Measures

Our first goal is to investigate the sensitivity of quality measures to distortions arising from image compression schemes. In other words to find out the degree to which a quality measure can discriminate the coding artifacts and translate it into a meaningful score. We establish similarly the response sensitivity of the measures to such other distortion effects as blur and noise. Our second goal is to establish how various quality measures are related to each other and to show the degree to which measures respond to

(dis)similarly to coding and sensor artifacts. As the outcome of these investigations we hope to extract a small subset of measures that hopefully satisfies the above desiderata.

2.4.1. ANOVA Results

The two-way ANOVA results of the image quality measures for the data obtained from all image classes (Fabrics, Faces, Remotes) are listed in Table 2.2. In these tables the symbols of quality measures *D1*, *D2*...*H3*, *H4* are listed in the first column while the F-scores of JPEG compression, of SPIHT compression, of blur and of noise distortions are given, respectively, in the succeeding four columns. The first factor tested is the bitrate variation and the second factor is the image set variation.

The metric that responds most strongly to one distortion type is called the "fundamental metric" of that distortion type [24]. Similarly the metric that responds to all sorts of distortion effects is denoted as the "global metric". One can notice that:

- The fundamental metrics for JPEG compression are H2, H1, S2, E2, that is, HVS L2 norm, HVS absolute norm, spectral phase-magnitude, and edge stability measures. These measures are listed in decreasing order of F-score.

- The fundamental metrics for SPIHT compression are E2, S2, S5, H2, that is, edge stability, spectral phase-magnitude, block spectral phase-magnitude, and HVS L2 norm.

- The fundamental metrics for the BLUR effect are S1, E2, S2, H1, that is, spectral phase, edge stability, spectral phase-magnitude, HVS absolute norm. Notice the similarity of metrics between SPIHT and blur due, in fact, to the blurring artifact encountered in wavelet-based compression.

- The fundamental metric for the NOISE effect is, as expected, D1, the mean square error.

- Finally the image quality metric that is sensitive to all distortion artifacts are, in order, E2, H1, S2, H2, S5, that is, edge stability, HVS absolute norm, spectral phase-magnitude, HVS L2 norm, block spectral phase-magnitude.

Table 2.2. ANOVA results (F-scores) for the JPEG and SPIHT compression distortions as well as additive noise and blur artifacts. For each distortion type the variation due to image set is also established

	JPEG		SPIHT		BLUR		NOISE	
ANOVA2	Bitrate F Score	Imageset F Score	Bitrate F Score	Imageset F Score	Blur F Score	Imageset F Score	Noise F Score	Imageset F Score
D1	104.6	42.59	39.23	13.28	43.69	2.06	9880	17.32
D2	108.5	67.45	29.56	15.93	33.94	17.76	6239	20.4
D3	63.35	29.37	53.31	48.53	38.55	24.13	1625	11.15
D4	89.93	1.99	13.75	3.71	27.87	0.96	166.4	9.88
D5	20.26	80.71	14.09	68.22	6.32	55.11	1981	43.51
D6	76.73	5.94	37.52	11.22	412.9	45.53	44.61	4.38
C1	1.35	124.6	12.05	325.5	5.61	107.2	3.82	6.17
C2	12.26	93.83	15.18	82.87	11.19	39.77	58.04	45.63
C3	82.87	83.06	24.96	22.42	30.92	1.71	567.5	52.01
C4	45.65	47.36	7.91	5.94	16.48	0.77	198.8	19.03
C5	91.42	38.17	27.51	5.28	52.57	2.44	704	10.8
E1	26.24	3.64	77.86	137	125.8	21.09	87.76	27.87
E2	176.3	92.75	212.5	200.4	768.7	23.41	158.5	24.84
S1	150.5	102.2	104	68.17	1128	60.04	47.29	38.42
S2	191.3	98.42	161	101.8	572.2	17.95	107.1	4.83
S3	145.6	56.39	38.58	26.97	24.28	6.39	2803	8.59
S4	129.1	63.26	128	46.85	215	11.17	56.04	55.1
S5	146.1	71.03	144.1	61.65	333.6	27.84	78.04	26.53
Z1	1.69	141.8	21.36	14	35.9	62.5	44.89	110.9
Z2	7.73	114.7	11.41	77.68	10.17	1.80	3.03	11.36
Z3	17.63	223	23.22	181.4	17.26	8.31	14.71	21.12
Z4	9.4	23.58	9.84	32.41	8.45	14.74	24.99	3.31
H1	371.9	0.09	107.2	40.05	525.6	69.98	230.7	19.57
H2	2291	5.46	132.9	22.82	47.28	101.7	624.3	21.32
H3	123	1.2	27.45	7.6	67.31	6.77	117.3	0.50
H4	78.83	7.14	25.2	95.72	12.55	2.11	29.06	6.69

To establish the global metrics, we gave rank numbers from 1 to 26 to each one metric under the four types of distortion as in Table 2.2. For example for JPEG the metrics are ordered as H2, H1, S2, E2, etc. if we take into consideration their F-scores. Then we summed the rank numbers and the metrics for which the sum of the scores were the smallest were declared as the global metric, that is the one that qualifies well in all discrimination tests.

The metrics that were the least sensitive to image set variation are D4, H3, C4, C5, D6 etc.. However it can be observed that these metrics show in general poor performance in discriminating distortion effects. On the other hand for the distortion sensitive metrics, even though their image set dependence is higher than the so-called "image independent" metrics, more of the score variability is due to distortion than to image set change. This can be observed based on the higher F-scores for distortion effects as compared to image set related F-scores. These observations are summarized in Table 2.3. where one-way ANOVA results are given for each image class (Fabrics, Faces, Remote Sensing) separately, and two-way ANOVA results are presented for the combined set. In the two bottom rows of Table 2.3. the metrics that are least sensitive to the coder type and to the image set are given.

Table 2.3. One-way ANOVA results for each image class and two-way ANOVA results for the distortions on the combined and image set independence

1-way ANOVA	IMAGE SET	JPEG	SPIHT	BLUR	NOISE
	Fabrics	H4,H2,E2,S4	E1,S1,E2,S2	S1,S5,E2,S4	D1,D2,D5,D3
	Faces	H2, D1,S3,H1	H4,D3,H2,C1	S2,H1,S1,E2	D1,S3,D2,D3
	Remote Sensing	H2,H4,S4,S5	S2,S5,S4,S1	D6,S5,S4,S1	D1,D2,C3,C5
2-way ANOVA	Combined Set	H2,H1,S2,E2	E2,S2,S5,H2	S1,E2,S2,H1	D1,D2,S3,D5
	Image Set Independence	H1,H3	D4,C5	C4,D4	H3,Z4
	Coder Type Independence	D2,D1,Z4,D3			

We also investigated the metrics with respect to their ability to respond to bit rate and coder type. More specifically the first factor tested was bitrate variation and the second was coder type variation. For this analysis the scores of the JPEG and SPIHT compressors were combined. It was observed in Table 2.4. that:

- The metrics that were best in discriminating compression distortion as parameterized by the bit rate, whatever the coder type, that is JPEG or SPIHT, were H2, H1, S2, S5 (HVS L2 norm, HVS absolute norm, spectral phase-magnitude, block spectral phase-magnitude etc.

- The metrics that were capable of discriminating the coder type (JPEG versus SPIHT) were quite similar, that is, D6, H2, H4 and H1 (Multiresolution error, HVS L2 norm, DCTune, HVS L1 norm).

- Finally the metrics that were most sensitive to distortion artifacts, but at the same time, least sensitive to image set variation were C5, D1, D3, S3, D2, C4..., (Mean angle-magnitude similarity, Mean square error, Modified infinity norm, Block spectral magnitude error, Mean absolute error, Mean angle similarity...).

These metrics were identified by summing the two rank scores of metrics, the first being the ranks in ascending order of distortion sensitivity, the second being in descending order the image set sensitivity. Interestingly enough almost all of them are related to the mean square error varieties. Despite its many criticisms, this may explain why mean square error or signal-to-noise ratio measures have proven so resilient over time.

Table 2.4. ANOVA scores for the bit rate variability (combined JPEG and SPIHT scores) and coder variation

	JPEG+SPIHT	
ANOVA2	Bitrate	Coder
D1	89.79	0.75
D2	74.98	2.72
D3	71.55	1.21
D4	70.52	43.85
D5	17.07	0.0005
D6	85.22	118.8
C1	2.66	45.47
C2	12.28	18.27
C3	56.48	1.56
C4	31.3	2.43
C5	78.98	2.23
E1	42.69	11.61
E2	122.4	26.28
S1	99.12	5.29
S2	140.1	12.37
S3	92.99	9.27
S4	115.5	39.1
S5	124.8	43.09
Z1	4.28	41.6
Z2	9.54	0.83
Z3	12.87	0.56
Z4	9.39	6.64
H1	278.6	52.87
H2	493	87.21
H3	97.94	16.19
H4	21.13	57.72

As expected the metrics that are responsive to distortions are also almost always responsive to the image set. Conversely the metrics that do not respond to the image set variation are also not very discriminating with respect to the distortion types. The fact that the metrics are sensitive, as should be expected, to both the image content and distortion artifacts does not eclipse their capability to potential as quality metrics. Indeed when the metrics were tested within more homogeneous image sets (that is only within Face images or Remote Sensing images etc.) the same high-performance metrics scored consistently higher. Furthermore when one compares the F-scores of the metrics with respect to bit rate variation and image set variation, even though there is a non-negligible interaction factor, one can notice that the F-score due to bit rate is always larger than the F-score due to Image sets.

2.4.2. Self Organizing Map of Quality Measures

Our second investigation was on the mutual relationship between measures. It is obvious that the quality measures must be correlated with each other as most of them must respond to compression artifacts in similar ways. On the other hand one can conjecture that some measures must be more sensitive to blurring effects, while others respond to blocking effects, while still some others reflect additive noise.

Self Organizing Map (SOM) [55] is a pictorial method to display similarities and differences between statistical variables, such as quality measures. We have therefore obtained spatial organization of these measures via Kohonen's self-organizing map algorithm. The input to the SOM algorithm was vectors whose elements are the scores of the measure resulting from different images. More explicitly, consider one of the measures, D1, and a certain compression algorithm, e.g., JPEG. The instances of this vector will be 60-dimensional, one for each of the images in the set. The first 30 components consist of 30 images compressed with JPEG, the next 30 juxtaposed components of the same images compressed with SPIHT. Furthermore there will be five such vectors, one for each one of the bit rates.

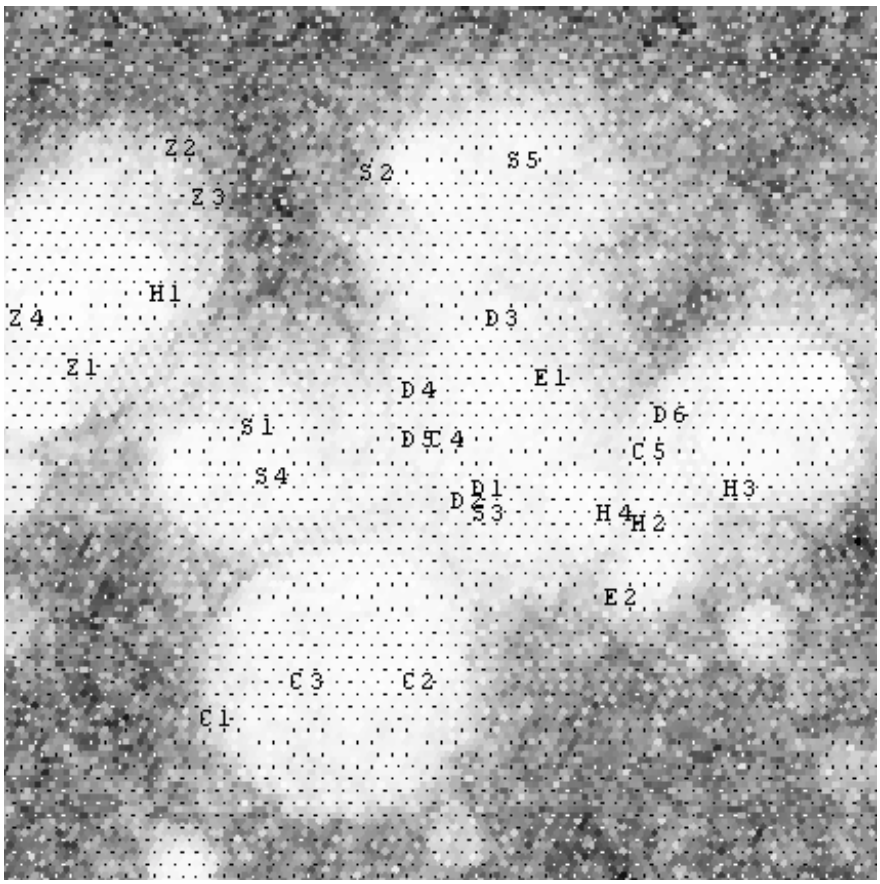


Figure 2.2. SOM of distortion measures for JPEG and SPIHT

The SOM organization of the measures in the 2-D space for pooled data from JPEG and SPIHT coders is shown in Figure 2.2. The map consists of 70 X 70 cells. These maps are useful for the visual assessment of possible correlation present in the measures. One would expect that measures with similar trends and that respond in similar ways to artifacts would cluster together spatially. The main conclusions from the observation of the SOM and the correlation matrix are as follows:

- Clustering tendency of pixel difference based measures (D1, D2, D4, D5) and spectral magnitude based method (S3) is obvious in the center portion of the map, a reflection of the Parseval relationship, that is, distortion in image energy in spatial domain matches in the same way in the frequency domain. However notice that spectral phase measures (S2, S5) stay distinctly apart from these measures.

- The human visual system based measures (H2, H3, H4), multiresolution pixel-difference measure (D6), E2 (edge stability measure) and C5 (mean angle-magnitude measure) are clustered in the right side of the map. The correlation of the multiresolution distance measure, D6 with HVS based measures (H2, H3, H4) is not surprising since the idea behind this measure is to mimic image comparison by eye more closely, by assigning larger weight to low resolution components and less to the detailed high frequency components.

- The three correlation based measures (C1, C2, C3) are together in the lower part of the map while the two spectral phase error measures (S2, S5) are concentrated separately in the upper part of the map.

- It is interesting to note that all the context-based measures (Z1, Z2, Z3, Z4) are grouped in the upper left region of the map together with H1 (HVS filtered absolute error).

- The proximity between the Pratt measure (E1) and the maximum difference measures (D3) is meaningful, since the maximum distortions in reconstructed images are near the edges. The constrained maximum distance or sorted maximum distance measures can be used in codec designs to preserve the two dimensional features, such as edges, in reconstructed images.

2.5. Conclusions

We have presented collectively a set of image quality measures in their multispectral version and categorized them. Statistical investigation of 26 different measures using a ANOVA analyses has revealed that local phase-magnitude measures (*S2 or S5)*, HVS-filtered *L1*, *L2* norms, edge stability measure are most sensitive to coding, blur and artifacts, while the mean square error (*D1*) remains as the measure for additive noise. One can state that combined spectral phase-magnitude measures and HVS filtered error norms should be paid more attention in the design of coding algorithms and sensor evaluation. On the other hand the pixel-difference based measures remain still to be the measures responsive to distortions and least affected by image variety.

The Kohonen map of the measures has been useful in depicting similar ones, and identifying the ones that are sensitive possibly to different distortion artifacts in compressed images. The correlation between various measures has been depicted via Kohonen's Self-Organizing Map. The placement of measures in the two-dimensional map has been in agreement with one's intuitive grouping.

Future work will address subjective experiments and prediction of subjective image quality using the above salient measures identified. Another possible avenue is to combine various "fundamental" metrics for better performance prediction.

3. STEGANALYSIS USING IMAGE QUALITY METRICS

3.1. Introduction

Steganography refers to the science of "invisible" communication. Unlike cryptography, where the goal is to secure communications from an eavesdropper, steganographic techniques strive to hide the very presence of the message itself from an observer. Although steganography is an ancient subject, the modern formulation of it is often given in terms of the *prisoner's problem* [57, 58, 59] where Alice and Bob are two inmates who wish to communicate in order to hatch an escape plan. However, all communication between them is examined by the warden, Wendy, who will put them in solitary confinement at the slightest suspicion of trouble. Specifically, in the general model for steganography, we have Alice wishing to send a *secret message m* to Bob. In order to do so, she "embeds" *m* into a *cover-object c*, to obtain the *stego-object s.* The stego-object *s* is then sent through the public channel.

The warden Wendy who is free to examine all messages exchanged between Alice and Bob can be *passive* or *active.* A passive warden simply examines the message and tries to determine if it potentially contains a hidden message. If it appears that it does, then she takes appropriate action, else she lets the message through without any action. An active warden, on the other hand, can alter messages deliberately, even though she does not see any trace of a hidden message, in order to foil any secret communication that can nevertheless be occurring between Alice and Bob. The amount of change the warden is allowed to make depends on the model being used and the cover-objects being employed. For example, with images, it would make sense that the warden is allowed to make changes as long as she does not alter significantly the subjective visual quality of a suspected stego-image.

It should be noted that the main goal of steganography is to communicate securely in a completely undetectable manner. That is, Wendy should not be able to distinguish in any sense between cover-objects (objects not containing any secret message) and stego-objects (objects containing a secret message). In this context, "*steganalysis*" refers to the body of

techniques that are designed to distinguish between cover-objects and stego-objects. It should be noted that nothing might be gleaned about the contents of the secret message *m.* When the existence of hidden message is known, revealing its content is not always necessary. Just disabling and rendering it useless will defeat the very purpose of steganography. In this paper, we present a steganalysis technique for detecting *stego-images*, i.e. still images containing hidden messages, using image quality metrics. Although we focus on images, the general techniques we discuss would also be applicable to audio and video data.

Given the proliferation of digital images, and given the high degree of redundancy present in a digital representation of an image (despite compression), there has been an increased interest in using digital images as cover-objects for the purpose of steganography. The simplest of such techniques essentially embed the message in a subset of the LSB (least significant bit) plane of the image, possibly after encryption [60]. It is well known that an image is generally not visually affected when its least significant bit plane is changed. Popular steganographic tools based on LSB like embedding vary in their approach for hiding information. Methods like *Steganos* and *Stools* use LSB embedding in the spatial domain, while others like *Jsteg* embed in the frequency domain. Other techniques include the use of quantization and dithering. For a good survey of steganography techniques, the reader is referred to [60]. What is common to these techniques is that they assume a passive warden framework. That is they assume the warden Wendy will not alter the image. We collectively refer to these techniques as *conventional* steganography techniques or for brevity, more simply as steganography techniques.

Conventional steganography techniques like LSB embedding techniques are not useful in the presence of an active warden as the warden can simply randomize the LSB plane to thwart communication. In order to deal with an active warden Alice must embed her message in a robust manner. That is, Bob should be able to accurately recover the secret message *m* despite operations like LSB randomizing, compression, filtering, rotation by small degrees etc. performed by the active warden Wendy. Indeed, the problem of embedding messages in a robust manner has been the subject of active research in the image processing community under the name of *digital watermarking* [61, 62, 63].

A digital watermark is an imperceptible signal added to digital content that can be later detected or extracted in order to make some assertion about the content. For example, the presence of her watermark can be used by Alice to assert ownership of the content. Given the proliferation of content in digital form, recent years have seen an increasing interest in digital watermarking and in the past few years, many different watermarking algorithms have been proposed for different applications. Although the main applications for digital watermarking appear to be copyright protection and digital rights management, watermarks have also been proposed for secret communication, that is, steganography. Essentially digital watermarks provide a means of image-based steganography in the presence of an active warden since modifications made by the warden will not affect the embedded watermark as long as the visual appearance of the image is not significantly degraded.

However, despite this obvious and commonly observed connection to steganography, there has been very little effort aimed at analyzing or evaluating the effectiveness of watermarking techniques for steganographic applications. Instead, most work has focused on analyzing or evaluating the watermarking algorithms for their robustness against various kinds of attacks that try to remove or destroy them. However, if digital watermarks are to be used in steganography applications, detection of their presence by an unauthorized agent defeats their very purpose. Even in applications that do not require hidden communication, but only robustness, we note that it would be desirable to first detect the possible presence of a watermark before trying to remove or manipulate it. This means that a given signal would have to be first analyzed for the presence of a watermark. Based on this analysis there could then be attempts made to remove the watermark.

In this thesis, we develop steganalysis techniques both for conventional LSB like embedding used in the context of a passive warden model and for watermarking which can be used to embed secret messages in the context of an active warden. In order to distinguish between these two models, we will be using the terms watermark and message when the embedded signal is in the context of watermarking and conventional steganography, respectively. Furthermore, we simply use the terms marking or embedding when the context of discussion is general to include both watermarking and steganography.

The techniques we present are novel and to the best of our knowledge, the first attempt at designing general purpose tools for steganalysis. General detection techniques as applied to steganography have not been devised and methods beyond visual inspection and specific statistical tests for individual techniques [64, 65, 66, 67] are not present in the literature. Since too many images have to be inspected visually to sense hidden messages, the development of a technique to automate the detection process will be very valuable to the steganalyst.

Our approach is based on the fact that hiding information in digital media requires alterations of the signal properties that introduce some form of degradation, no matter how small. These degradations can act as signatures that could be used to reveal the existence of a hidden message. For example, in the context of digital watermarking, the general underlying idea is to create a watermarked signal that is *perceptually identical but statistically different* from the host signal. A decoder uses this statistical difference in order to detect the watermark. However, the very same statistical difference that is created could potentially be exploited to determine if a given image is watermarked or not. In this thesis, we show that addition of a watermark or message leaves unique artifacts, which can be detected using Image Quality Measures [68, 69, 70, 71, 72].

The rest of this section is organized as follows. In Section 3.2., we discuss the selection of the image quality measures to be used in the steganalysis and the rationale of utilizing concurrently more than one quality measure. We then show that the image quality metric based distance between an unmarked image and its filtered version is different as compared to the distance between a marked image and its filtered version. Section 3.3. describes the regression analysis that we use to build a composite measure of quality to indicate the presence or absence of a mark. Statistical tests and experiments are given in Section 3.4. and, finally, conclusions are drawn in Section 3.5.

3.2. Choice of Image Quality Measures

As stated in the introduction, the main goal of this is to develop a discriminator for message or watermark presence in still images, using an appropriate set of IQMs. Image quality measurement continues to be the subject of intensive research and experimentation

[4, 5, 2, 73, 47]. Objective image quality measures are based on image features, a functional of which, correlates well with subjective judgment, that is, the degree of (dis)satisfaction of an observer [13]. The interest in developing objective measures for assessing multimedia data lies in the fact that subjective measurements are costly, time-consuming and not easily reproducible. Objective measures are also utilized in performance prediction of vision algorithms against quality loss due to sensor inadequacy or compression artifacts [24]. In this work, however, we want to exploit image quality measures, not as predictors of subjective image quality or algorithmic performance, but as steganalysis tools, that is, as detection features of watermarks or hidden messages.

A good IQM should be accurate, consistent and monotonic in predicting quality as already mentioned in second section. In the context of steganalysis, however, *prediction accuracy* can be interpreted as the ability of the measure to detect the presence of hidden message with minimum error on average. Similarly, *prediction monotonicity* signifies that IQM scores should ideally be monotonic in their relationship to the embedded message size or watermark strength. Finally, *prediction consistency* relates to the quality measure's ability to provide consistently accurate predictions for a large set of watermarking or steganography techniques and image types. This implies that the spread of quality scores due to image variety should not eclipse the score differences arising from message embedding artifacts.

The steganalysis technique we develop is based on regression analysis of a number of *relevant* IQMs. Hence, we seek IQMs that are sensitive specifically to watermarking and steganography effects. In other words, those measures for which the variability in score data can be explained better because of treatment rather then as random variations due to the image set. The idea behind detection of watermark or hidden message presence is to obtain a consistent distance metric for images containing a watermark or hidden message and those without, *with respect to a common reference processing*. The reference processing we have used was low-pass filtering based on a Gaussian kernel. The filter was chosen as a Gaussian smoothing filter

$$H(m,n) = Kg(m,n) \tag{3.1}$$

where

$$g(m,n) = (2\pi\sigma^2)^{-1} \exp\{-(m^2 + n^2)/2\sigma^2\} \tag{3.2}$$

is the 2-D Gaussian kernel and

$$K = (\sum_m \sum_n |g(m,n)|^2)^{-1/2} \tag{3.3}$$

is the normalizing constant. We experimentally choose $\sigma = 0.5$, implemented via a 3x3 filter. The reason why Gaussian blurring works fine as a common reference is that it gives us the local mean which is also the maximum likelihood (ML) estimate of the image under Gaussian assumption [74]. Under Laplacian distribution assumption the median would have been the ML estimate. Therefore the blurred image minus the original yields, in fact, the maximum likelihood estimate of the additive watermark. In fact we have tested both the mean and median filters as the ML estimates of the image and we have found out that the former performs slightly better in the detection tests.

Most watermarking techniques or steganographic message embedding techniques, whether by spread-spectrum or quantization modulation or LSB insertion, can be represented as a signal addition to the cover image, as shown in Figure 3.1. Let f be the cover image, $g = f + w$ be the stego-image, and w the inserted watermark. Let H be the ML operator for the estimate of the watermark sequence. For the two ML estimators that we have tested, H obviously corresponds to the subtraction from the received stego-image of its local mean or median. In the absence of any watermark or stego-signal $Hg = \hat{f}$ corresponds to the high-frequency content $\hat{f}$ of the image, while for a marked signal it yields $Hg = \hat{f} + \hat{w}$ where $\hat{w}$ denotes the ML estimate of the mark. The image quality metrics, in fact, are trained to differentiate between these two signals $\hat{f}$ and $\hat{f} + \hat{w}$. Figure 3.2. gives an instance of the watermarked-nonwatermarked class separability based on a scatter diagram of the three image quality metrics used.

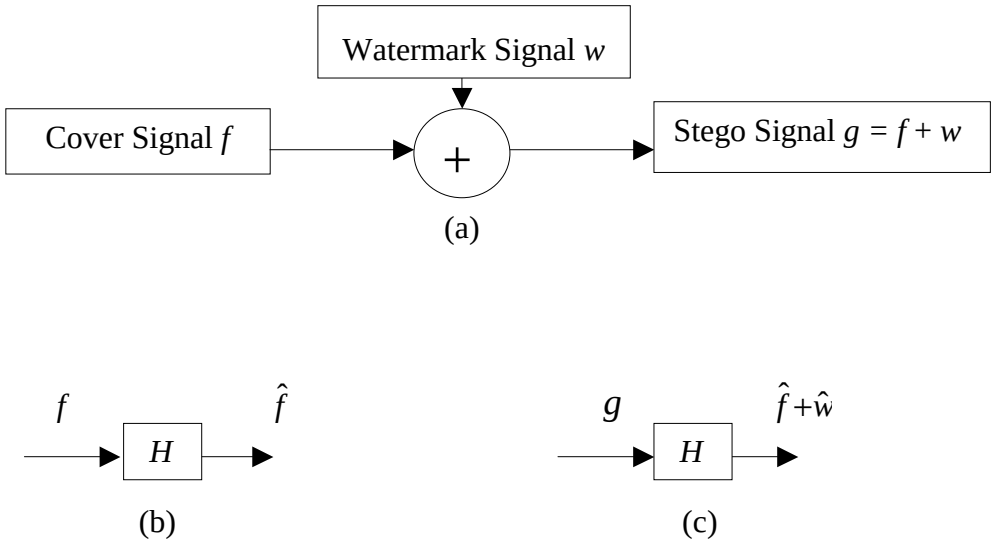


Figure 3.1. Schematic descriptions of (a) watermarking or stegoing, (b) filtering an un-marked image, (c) filtering a marked image.

As for the selection of quality measures we have gleaned out the ones that serves well the purpose of our steganalysis. The rationale of using several quality measures is that different measures respond with differing sensitivities to artifacts and distortions. For example, some measures like mean square error respond more to additive noise, others such as spectral phase or mean square HVS-weighted (Human Visual System) error are more sensitive to pure blur, while the gradient measure reacts to distortions concentrated around edges and textures. Recall that some watermarking algorithms inject *noise* in block DCT coefficients, others in a narrow-band of global DCT or Fourier coefficients, still others operate in selected localities in the spatial domain. Since we want our steganalyzer to be able to work with a variety of watermarking and steganography algorithms, a multitude of quality features are needed so that the steganalyzer has the chance to probe all features in an image that are significantly impacted by the embedding – watermarking or steganographic – process.

In order to identify specific quality measures that are useful in steganalysis, we use ANOVA [54] test. ANOVA helped us to distinguish measures that are most consistent and accurate vis-à-vis the effects of watermarking and of steganography. Various quality measures are subjected to a statistical test to determine if the fluctuations of the measures

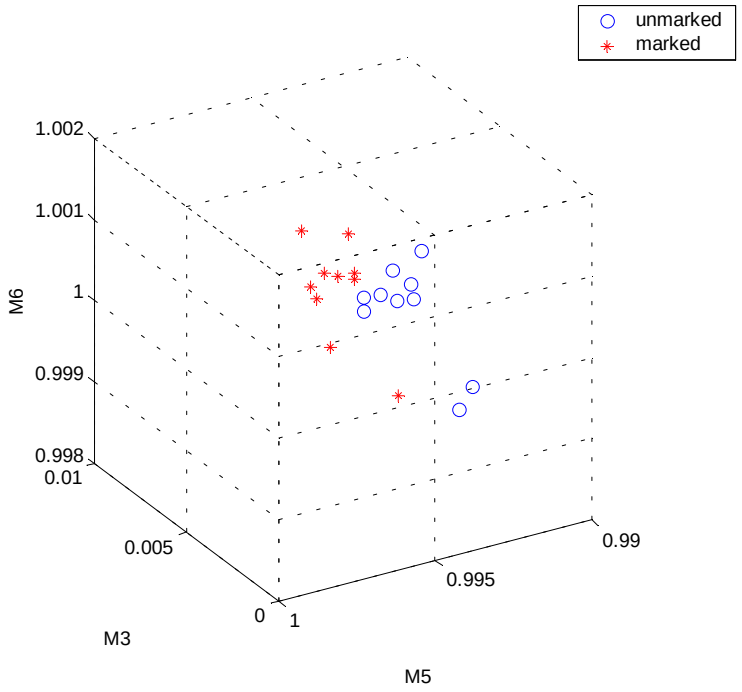


Figure 3.2. Scatter plots of the three Image Quality Measures (M3: Czenakowski measure, M5: Image Fidelity, M6: Normalized Cross-correlation).

result from image variety or whether they arise due to treatment effects, that is, watermarking or stegoing.

We performed three different ANOVA tests: The first was for watermarking, the second for steganography, and the last one for both watermarking and steganography.

For watermarking, the first group consisted of the IQM scores computed from plain images and their filtered versions. The remaining three groups consisted of the IQM scores computed from watermarked images by Digimarc [75], PGS [76] and COX [77] techniques, respectively, and their filtered versions. The data given to the ANOVA algorithm consisted of four vectors, each of dimensions N, where $N = 12$ is the number of images used in the test from the training set. More specifically, consider a typical quality measure, say $M(\mu_i)$, where the parametric dependence upon the watermarking algorithm is shown with μ_i, $i = 0 \ldots 3$, for plain images, Digimarc, PGS and COX techniques, respectively. The N-dimensional vector M reads as:

$$M(\mu_i) = [M(1|\mu_i) \ldots M(N|\mu_i)]^T . \quad (3.4)$$

For steganography, the first group consisted of the IQM scores computed from plain (non-marked) images and their filtered versions. The remaining three groups consisted of the IQM scores computed from stegoed images by Steganos [78], Stools [79] and Jsteg [80], respectively, and their filtered versions.

For the joint watermarking and steganography analysis, the first group consisted of the IQM scores computed from plain images and their filtered versions. The remaining six groups consisted of the IQM scores computed from watermarked images by Digimarc, PGS and COX technique, stegoed images by Steganos, Stools and Jsteg, respectively, and their filtered versions.

In Table 3.1. we give ANOVA results with respect to watermarking, steganography and combined techniques. The measures that have relatively higher discriminative power, measures that catch the statistical evidence of watermarking or steganography, are shown in bold. These measures, in fact, sense the statistical difference between the populations of marked and non-marked images and hence they can be used to separate the two classes.

The implications of the result are two fold. One is that, using these features a steganalysis tool can be designed to detect the watermarked or stegonagraphically marked images, as we show in Section 3.3., using multivariate regression analysis. The other is that, current watermarking or steganographic algorithms should exercise more care on those statistically significant image features to eschew detection. For instance, the relative ordering of the statistically significant IQMs for watermarking and steganographic algorithms are different. While the Minkowsky measures were not statistically significant for steganographic algorithms, they were for the watermarking algorithms. Minimizing the Mean Square Error (MSE) or the Kullbeck-Leibler distance between the original (cover) image and the stego image is not necessarily enough to achieve covert communication as the evidence can be caught by another measure such as spectral measures. The selected subset of image quality measures in the design of steganalyzer with respect to their statistical significance were as follows:

Table 3.1. One-way ANOVA tests for watermarking, steganography and pooled waterwarking and steganography

	Watermark		Stego		Watermark&Stego	
Image Quality Measures	F	p	F	p	F	p
Minkowsky Metric $\gamma = 2$	**6.06**	0.01	0.56	0.58	5.28	0.00
Minkowsky Metric $\gamma = 1$	**3.28**	0.05	0.57	0.58	**3.07**	0.03
Maximum Difference	0.13	0.93	0.31	0.74	0.25	0.93
Sorted Maximum Difference	0.14	0.93	0.07	0.92	0.13	0.98
Czenakowski	**4.63**	0.02	1.08	0.37	**4.66**	0.01
Structural Content	0.62	0.61	0.15	0.86	0.58	0.71
Cross Correlation	2.08	0.14	0.21	0.81	0.74	0.60
Image Fidelity	2.67	0.08	0.40	0.68	1.14	0.37
Angle Mean	1.95	0.17	**4.20**	0.04	**3.40**	0.02
Angle Standard Deviation	0.45	0.72	3.27	0.08	2.36	0.08
Spectral Magnitude	**5.50**	0.03	0.02	0.98	**4.35**	0.01
Spectral Phase	**5.49**	0.03	0.02	0.98	4.34	0.01
Weighted Spectral Distance	1.12	0.37	0.06	0.94	0.66	0.65
Median Block Spectral Magnitude	0.79	0.51	0.001	0.99	0.44	0.81
Median Block Spectral Phase	0.47	0.72	**3.95**	0.05	**4.24**	0.02
Median Block Weighted Spectral Dist.	0.45	0.72	**3.96**	0.05	**4.22**	0.02
Normalized Absolute Error (HVS)	0.16	0.92	1.16	0.35	0.74	0.61
Normalized MS ERROR (HVS)	**3.30**	0.05	**4.93**	0.02	**2.69**	0.05
HVS Based L2	0.19	0.90	0.46	0.64	0.47	0.79

Watermarking: Mean Absolute Error *D2*, Mean Square Error D_1, Czekanowski Correlation Measure C_3, Image Fidelity C_2, Cross Correlation C_1, Spectral Magnitude Distance S, Normalized Mean Square HVS Error H. We denote this feature set as $\Psi = \{D_1, D_2, C_1, C_2, C_3, S, H\}$ for future reference in the experiments in Section 3.4.

Steganography: Angle Mean C_4, Median Block Spectral Phase Distance S_4, Median Block Weighted Spectral Distance S_5, Normalized Mean Square HVS Error H. We denote this feature set as $\Omega = \{C_4, S_4, S_5, H\}$.

Pooled Watermarking And Steganography: Mean Absolute Error D_2, Mean Square Error D_1, Czekanowski Correlation Measure C_3, Angle Mean C_4, Spectral Magnitude Distance S, Median Block Spectral Phase Distance S_4, Median Block Weighted Spectral Distance S_5, Normalized Mean Square HVS Error H. We denote this feature set as $\Xi = \{D_1, D_2, C_3, C_4, S_4, S_5, H\}$.

3.3. Regression Analysis of the Quality Measures

The steganalysis we propose is based on the observation in Section 3.2. that an embedded and filtered image would differ significantly from a non-embedded but filtered image. In other words, both the embedded and non-embedded images are compared against the common reference of their filtered images. It has been observed that filtering an image with no message causes changes in the IQMs differently than the changes brought about on embedded images. This differential behavior is in part because watermarking or steganographic embedding is not in general a global operation, but is local in nature. The watermark or message signal is either injected locally, e.g., on a block basis, or the signal is subjected to a perceptual mask. In any case, we consistently obtained statistically different quality scores from filtered-and-embedded images and from filtered-but-not-embedded sources. For the hypothesis test, we used various measured quality scores, which are either due to the difference between an originally non-embedded image and its filtered version, or due to the difference between embedded image and its filtered version. In conclusion the selected statistically significant IQMs form a multidimensional feature space whose points cluster well enough to do a classification of embedded and non-embedded images.

In the design phase of the steganalyzer, we regressed the normalized IQM scores to, respectively, -1 and 1, depending upon whether an image did not or did contain a message. Similarly, IQM scores were calculated between the original images and their filtered versions. In the regression model [54], we expressed each decision label y in a sample of n observations as a linear function of the IQM scores x's plus a random error, ε:

$$\begin{aligned} y_1 &= \beta_1 x_{11} + \beta_2 x_{12} + \ldots + \beta_q x_{1q} + \varepsilon_1 \\ y_2 &= \beta_1 x_{21} + \beta_2 x_{22} + \ldots + \beta_q x_{2q} + \varepsilon_2 \\ &\vdots \\ y_n &= \beta_1 x_{n1} + \beta_2 x_{n2} + \ldots + \beta_q x_{nq} + \varepsilon_n \end{aligned} \tag{3.5}$$

In this expression, x_{ij} denotes the IQM score, where the first index indicates the *i*'th image and the second one the quality measure. The total number of quality measures considered is denoted by q. The β's denote the regression coefficients. The complete statement of the standard linear model is

$$y = \mathbf{X}_{nxq}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad \textit{such that} \quad \begin{cases} rank(\mathbf{X}) = q \\ E[\boldsymbol{\varepsilon}] = \mathbf{0} \\ Cov[\boldsymbol{\varepsilon}] = \sigma^2 \mathbf{I} \end{cases}. \tag{3.6}$$

The corresponding optimal MMSE linear predictor $\boldsymbol{\beta}$ can be obtained by

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\left(\mathbf{X}^T\mathbf{y}\right). \tag{3.7}$$

Once the prediction coefficients are obtained in the training phase, these coefficients can be used in the testing phase. Given an image in the test phase, first it is filtered and the *q* IQM scores are obtained using the image and its filtered version. Then using the prediction coefficients, these scores are regressed to the output value. If the output exceeds the threshold 0 then the decision is that the image is embedded, otherwise the decision is that the image is *not* embedded. That is

$$\hat{y} = \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_q x_q \tag{3.8}$$

for $\hat{y} \geq 0$ the image contains watermark, and for $\hat{y} < 0$ it does not. The schematic diagram of the steganalyzer is given in Figure 3.3.

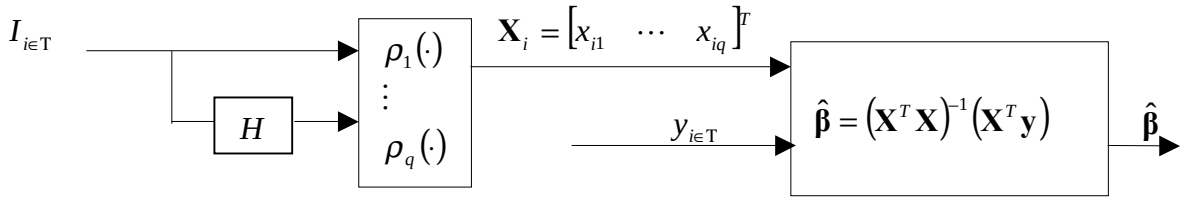
3.4. Simulation Results

The watermarking techniques we used were the following: 1) Photoshop plug-in Digimarc [75], Cox et. al.'s technique [77], and the technique from Swiss Federal Institute of Technology, PGS [76]. One reason for the selection of these techniques was their free availability on the Internet and the fact that they were all popularly known algorithms. A more important reason was that with these techniques it was possible to embed watermarks

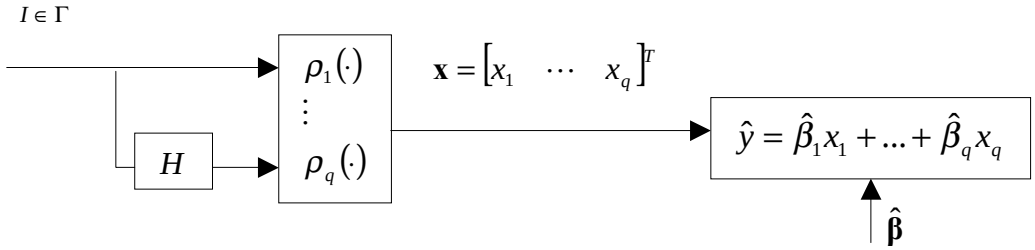
at different strengths, which was instrumental to extract the sensitive IQMs. The steganographic tools were Steganos [78], S-Tools [79] and Jsteg [80]. These tools were among the most cited ones for their pleasing results with respect to steganographic applications. We used the image database for the simulations. The database contained a variety of images including computer generated images, images with bright colors, images with reduced and dark colors, images with textures and fine details like lines and edges, and well-known images like Lena, peppers etc. We performed eight experiments.

The first three experiments involved watermarking only, namely: 1) The steganalysis of individual watermarking algorithms, Digimarc, PGS and Cox et. al. for admissible watermark strengths; 2) The steganalysis of pooled watermarking algorithms at admissible watermark strengths; 3) In the third experiment the steganalyzer was trained on images watermarked by Digimarc, and tested on images watermarked by PGS and Cox et. al.

The next three experiments involved steganography only: 4) The steganalysis of individual steganography algorithms, Steganos, Stools and Jsteg for different embedded



(a)



(b)

Figure 3.3. Schematic description of (a) training, and (b) testing.

message sizes; 5) The steganalysis of pooled steganography algorithms for different message sizes; 6) In the sixth experiment the steganalyzer was trained on images embedded with Steganos and Stools, and tested on images embedded with Jsteg.

The final two experiments involved both watermarking tools and steganography algorithms. 7) The seventh experiment was the steganalysis of pooled three steganographic and three watermarking algorithms for admissible levels of watermark strength and for different message lengths. 8) In the last eighth experiment steganalyzer was trained on images embedded with Steganos, Stools, watermarked by Digimarc and tested on images embedded with Jsteg and watermarked by Cox et. al. The aim of the last three experiments was to see the generalizing ability of the steganalyzer in case an image was to be analyzed unknown to it in the learning phase. In experiments 1, 2 and 3 the feature set was Ψ which was defined in Section II, for the experiments 4, 5 and 6 the feature set was Ω, while the feature set was Ξ for the remaining experiments 7 and 8. In the training phase of every experiment, the feature sets were regressed to –1 and +1.

The organizations of the training and testing samples for the experiments are given in Tables 3.2.-3.12. The images in the training and test sets are denoted by numbers. More specifically the training set is $T = \{1,\ldots,12\}$ and the test set is $\Gamma = \{13,\ldots,22\}$. There were four levels of watermark strength for Digimarc and PGS. We used the original settings of Cox's technique; modified the 1000 most significant coefficients in spectral domain. The embedded message sizes were 1/10 and 1/40 of the cover image size for Steganos and Stools, while the message sizes were 1/100 of the cover image size for Jsteg.

Table 3.2. Training and test samples for Digimarc and PGS for experiment 1

	Level 1	Level 2	Level 3	Level 4
Training samples	1,2,3	4,5,6	7,8,9	10,11,12
Test samples	13,14,15	16,17	18,19,20	21,22

Table 3.3. Training and test samples for Cox for experiment 1

	1000 coefficients
Training samples	1...12
Test samples	13…22

Table 3.4. Training and test samples for pooled watermarking algorithms for experiment 2 (L1: Level 1 etc.)

	Digimarc				PGS				COX
Levels	L1	L2	L3	L4	L1	L2	L3	L4	
Train	1	2	3	4	5	6	7	8	9,10,11,12
Test	13	14	15	16	17	18	19	20	21,22

Table 3.5. Training and test samples for experiment 3: Train on Digimarc, test on PGS and COX

Training	Digimarc			
WM Levels	L1	L2	L3	L4
Training samples	1…3	4…6	7…9	10…12
Testing	PGS		COX	
WM Levels	L1	L2	L3	
Test samples	13…15	16…18	19,20	21,22

Table 3.6. Training and test samples for Stools for experiment 4

Message size	1/40 of image size	1/10 of image size
Training samples	1…6	7…12
Test samples	13…17	18…22

Table 3.7. Training and test samples for Jsteg for experiment 4

Message size	1/100 of image size
Training samples	1…12
Test samples	13…22

Table 3.8. Training and test samples for Steganos for experiment 4. (Note: In certain images the Steganos did not let the messages to be embedded no matter what their size)

Message size	1/40 of image size	1/10 of image size
Training samples	2,4,8	10,11,13
Test samples	15,17	19,20,21

Table 3.9. Training and test samples for pooled steganography algorithms for experiment 5

	Steganos		Stools		Jsteg
Message size	1/40	1/10	1/40	1/10	1/100
Training samples	2,4	8,10	1,3	5,6	7,9,11,12
Test samples	13,15	17,19	14,16	18,20	21,22

Table 3.10. Training and test samples for experiment 6: train on Steganos and Stools, test on Jsteg

Training	Steganos		Stools	
Msg. Size	1/40	1/10	1/40	1/10
Training samples	2,4,8	10,11	1,3,5,6	7,9,12
Testing	Jsteg			
Msg. Size	1/100			
Test samples	13…22			

Table 3.11. Training and test samples for pooled watermarking and steganography algorithms for experiment 7

	Digimarc		PGS		COX	Steganos		Stools		Jsteg
Level or msg size	L2	L3	L2	L3	1000 cof	1/40	1/10	1/40	1/10	1/100
Training samples	7	8	9	10	11,12	2	4	1	3	5,6
Test samples	18	19	20	21	22	13	15	14	16	17

Table 3.12. Training and test samples for experiment 8: train on Steganos, Stools and Digimarc, test on Jsteg and Cox

Training	Digimarc		PGS		COX	Steganos		Stools		Jsteg
Level, msg. size	L2	L3	L2	L3	1000 cof	1/40	1/10	1/40	1/10	1/100
Train samples	7	9	11	12		2,4	8,10	1,3	5,6	
Testing	Digimarc		PGS		COX	Steganos		Stools		Jsteg
Level, msg. size	L2	L3	L2	L3	1000 cof	1/40	1/10	1/40	1/10	1/100
Test samples					13…17					18..22

Table 3.13. Performance of the Steganalyzer for All the Experiments

Experiment	False Alarm Rate	Miss Rate	Correct Detection
1. a. Digimarc	2/10	2/10	16/20
1. b. PGS	2/10	1/10	17/20
1. c. Cox	4/10	2/10	14/20
2. Pooled Watermarking	3/10	3/10	14/20
3 Train on Digimarc, Test on PGS and Cox	5/10	2/10	13/20
4. a. Steganos	2/5	1/5	7/10
4. b. Stools	4/10	1/10	15/20
4. c. Jsteg	3/10	3/10	14/20
5 Pooled Steganography	5/10	0/10	15/20
6 Train on Steganos and Stools, Test on Jsteg	3/10	3/10	14/20
7 Pooled Watermarking and Steganography	5/10	1/10	14/20
8 Train on Digimarc, PGS, Steganos, Stools Test on Cox and Jsteg	4/10	3/10	13/20

The performance of the steganalyzer is given in Table 3.13. Simulation results indicate that the steganalyzer is well enough to do a classification of marked and non-marked images. The classifier is still able to do a classification when the tested images come from an embedding technique unknown to it, indicating that it has a generalizing capability of capturing the intrinsic characteristics of watermarking and steganographic techniques. As we have noted, the features were regressed to output labels –1 and +1 in every experiment. The false alarms, especially in experiments 3, 5 and 7, can be fixed to desired rate by choosing the output regression labels asymmetrically.

3.5. Conclusions

In this section, we have addressed the problem of steganalysis of watermarked and stegonagraphically marked images. That is, we develop techniques for discriminating between cover-images and stego-images. Our approach is based on the hypothesis that a particular message embedding scheme leaves statistical evidence or structure that can be exploited for detection with the aid of proper selection of image features and multivariate regression analysis. We showed that the distance between an unmarked image and its

filtered version is different than the distance between a marked image and its filtered version. We used image quality metrics as the feature set to distinguish between marked and non-marked images. To identify specific quality measures, which provide the best discriminative power, we used ANOVA technique. We have also pointed out the image features that should be taken more seriously into account in the design of more successful watermarking or steganographic techniques to eschew detection. After selecting an appropriate feature set, we used multivariate regression techniques to get an optimal classifier using an image and its filtered version. Simulation results with well known and commercially available watermarking and steganographic techniques indicate that the proposed technique is successful in classification of marked and non-marked images.

4. LOSSLESS AND NEAR-LOSSLESS IMAGE COMPRESSION WITH SUCCESSIVE REFINEMENT

4.1. Introduction

Lossless or reversible compression refers to compression approaches in which the reconstructed data exactly matches the original. Near-lossless compression denotes compression methods, which give quantitative guarantees on the nature of the loss that is introduced. Typically, most of the near-lossless compression techniques proposed in the literature provide a guarantee that no pixel difference between the original and the compressed image is above a given value [81]. Near-lossless compression is potentially useful in remote sensing, medical imaging, space imaging and image archiving applications, where the huge data size could require lossy compression for efficient storage or transmission. However, the need to preserve the validity of subsequent image analysis performed on the data set to derive information of scientific or clinical value puts strict constraints on the error between compressed image pixel values and their originals. In such cases, near-lossless compression can be used as it yields significantly higher compression ratios compared to lossless compression and at the same time, the quantitative guarantees it provides on the nature of loss introduced by the compression process are more desirable compared to the uncertainties that are faced when using lossy compression.

Another way to deal with the lossy-lossless dilemma faced in applications such as medical imaging and remote sensing is to use a successively refinable compression technique that provides a bitstream that leads to a progressive reconstruction of the image. The increasingly popular wavelet based image compression techniques, for example, provide an embedded bit stream from which various levels of rate and distortion can be obtained. With reversible integer wavelets, one gets a progressive transmission capability all the way to lossless reconstruction. Hence such techniques have been widely cited for potential use in applications like tele-radiology where a physician can request portions of an image at increased quality (including lossless reconstruction) while accepting unimportant portions at much lower quality, thereby reducing the overall bandwidth

required for transmitting an image [82, 83]. Indeed, the new still image compression standard, JPEG 2000, provides such features in its extended forms [84].

Although reversible integer wavelet based image compression techniques provide integration of lossless and lossy compression in one single framework, the compression performance they provide is typically inferior to state-of-the-art non-embedded and DPCM based techniques like CALIC [85]. In addition, although lossless compression is possible by receiving the entire bit stream (corresponding to a block or the entire image), the lossy reconstruction at intermediate stages provides no guarantees on the nature of the distortion that may be present. Near-lossless compression in such a framework is only possible either by an appropriate pre-quantization of the wavelet coefficients and lossless transmission of the resulting bit stream, or by truncation of the bit stream at an appropriate point followed by transmission of a residual layer to provide the near-lossless bound. Both these approaches have been shown to provide inferior compression as compared to near-lossless compression in conjunction with DPCM coding [81].

We propose a technique that unifies the above two approaches. The proposed technique produces a bitstream that results in progressive reconstruction of the image just like what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after each layer of the successively refinable bitstream is decoded (note, however that these bounds need to be pre-decided at compression time and cannot be changed during decompression). Furthermore, the compression performance provided by the proposed technique is superior or comparable to the best-known lossless and near-lossless techniques proposed in the literature [86, 87, 88].

This section is organized as follows: We review the concepts of successive refinement, density estimation and the data model in Section 4.2. The compression method is described in Section 4.3. In Section 4.4. we give experimental results and conlusions are drawn in Section 4.5.

4.2. Problem Formulation

The key problem in lossless compression involves estimating the p.m.f. (probability mass function) of the current pixel based on previously known pixels (or previously received information). With this in mind, the problem of successive refinement can then be viewed as the process of obtaining improved estimates of the p.m.f.'s with each pass of the image. If we also restrict the "support" of the p.m.f. to a given length we then integrate near-lossless compression and successive refinement with lossless compression in one single framework. That is we obtain a bitstream, which gives us near-lossless reconstruction after each pass in the sense that each pixel is within δ counts of its original value. The length of the interval, $\lambda = 2\delta + 1$, in which the pixel is, decreases with successive passes and in the final pass we have lossless reconstruction, $\lambda = 1$ and $\delta = 0$. In order to design a compression technique with these properties we consider image data compression as asking the optimal question to determine the exact value or the interval of the pixel depending on whether we are interested in lossless or near-lossless compression, respectively. Our aim is to find the minimum description length of every pixel based on the knowledge we have about its neighbors. We know from the Kraft Inequality that a code length is just another way to express a probability distribution. Massey [89] observed that the average number of guesses to determine the value of a random variable is minimized by a strategy that guesses the possible values of the random variable in decreasing order of probability. Our strategy is to estimate the probability density of the current pixel using previous information, and based on this density to determine the interval of the pixel by questioning the most probable interval where the pixel lies.

In the first pass, we assume that the data in use at a coarse level is stationary and Gaussian in a small neighborhood and we hence use linear prediction. We fit a Gaussian density for the current pixel, with the linear prediction value taken as the optimal estimate of its mean, and linear prediction error as its variance. We divide the support of the current pixel's p.m.f., [0,255], into equal $2\delta + 1$ length intervals, δ being integer. The intervals are sorted with respect to their probability mass. If the pixel is found to lie in the interval with highest probability mass the probability mass outside the interval is zeroed out and the event 1 is fed to the entropy coder; otherwise the next question is asked whether it lies in the next highest probability interval. Every time one receives a negative answer, the

probability mass within the given interval is zeroed out and the event 0 is fed to the entropy coder till the right interval is found. At the end of the first pass, we have a "crude" approximation of the image but the maximum error in reconstructed image $\|e\|_\infty$ is δ when the midvalue of the interval is selected as the reconstructed pixel value.

In the remaining passes we then refine the p.m.f. for each pixel by narrowing the size of the interval in which it is now known to lie. The key problem here is how to refine the p.m.f. of each pixel based on p.m.f.'s of its neighboring pixels. Note that the causal pixels already have a refined p.m.f. but the non-causal pixels do not. The non-causal pixels now give us vital information (like the presence of edges and texture patterns), which can be used to get better estimates of the refined p.m.f.. However care must be taken, as the available information is less redundant than in the first pass with respect to the current p.m.f. estimation. That is we know in which interval the current pixel is, we have more precise information of causal pixels and less precise information of non-causal pixels. We need to estimate/refine the current p.m.f. within the constraint of its support. The refinement of the current p.m.f. should take all these into account. The p.m.f. estimation method for second and remaining passes, outlined in the next section, which is simply a causal and non-causal p.m.f. summation over the current p.m.f.'s support takes successfully all the information into account. Once the p.m.f. is estimated/refined for the current pixel the same strategy, guessing the correct interval or the value depending on their probability, is applied to constrain the pixel to narrower intervals or to their exact values.

In the following sub-sections we review some key concepts and results from known literature and show how we propose to use these in order to develop the proposed technique for successively refinable lossless and near-lossless image compression.

4.2.1. Successive Refinement

Successive refinement of information consists of first approximating data using a few bits of information, and then iteratively improving the approximation as more and more information is supplied. In their paper on successive refinement of information [90] Equitz and Cover state that rate distortion problem is successively refinable if and only if the

individual solutions of the rate distortion problems can be written as Markov chain. Then they give examples of signals along with distortion measures for which successive refinement is possible, i.e. if the source is Gaussian and MSE (Mean Square Error) is used as the distortion measure the source is successively refinable. Massey [89] considered the problem of guessing the value of a realization of a random variable X by asking the questions of the form "Is X equal to x" until the answer is "Yes". It is observed that the average number of guesses is minimized by a guessing strategy that guesses the possible values of X in decreasing order of probability. In near-lossless compression we are interested in intervals where the pixel lies rather than in their exact values, so the optimal strategy for minimizing the average number of guesses is to guess the interval in decreasing order of probability masses contained in the intervals. In either case, we first need to construct probability mass estimates in order to use this strategy. In what follows, we describe probability mass estimation for different passes.

4.2.2. P.m.f. Estimation in the First Pass

The Gaussian Model: Natural images in general do not satisfy Gaussianity or stationarity assumptions. But at a coarse level, in a reasonable size neighborhood, the statistics can be assumed not to differ from the above assumptions and the results of Gauss-Markov property can be used. We use linear prediction in the first pass assuming the data in a small neighborhood as stationary and Gaussian. We fit a Gaussian density for the current pixel, with the linear prediction value taken as the optimal estimate of its mean.

We use causal pixel to predict the current pixel via normal linear regression model. Suppose $X_{i-1}, X_{i-2}, \ldots, X_{i-N}$ are random variables representing the causal neighbors of the current pixel X_i, shown in Figure 4.1. Let $x_{(i-1)k}, x_{(i-2)k}, \ldots, x_{(i-N)k}$, k = 1,…,K denote their realizations. We assume a discrete-time scalar-valued random process $\{X_i\}$ that satisfies the Nth-order linear prediction equation

$$X_i = \sum_{j=1}^{N} \beta_j \left(x_{i-j}\right) + \nu_i \tag{4.1}$$

where $\{\beta_j\}_{j=1}^{N}$ are real-valued linear prediction coefficients of the process, and $\{\nu_i\}$ is a sequence that consists of i.i.d. random variables having a Gaussian density with zero mean and variance σ^2. Optimal MMSE (Minimum MSE) linear prediction for an N^{th} order stationary Gauss-Markov process $\{X_i\}$ can be formulated as:

$$E\left[X_i \middle| X_{i-1}, X_{i-2}, \ldots, X_{i-N}\right] = \sum_{j=1}^{N} \beta_j \left(x_{i-j}\right). \tag{4.2}$$

For this standard linear model, according to Gauss-Markov theorem, the minimum variance linear unbiased estimator $\boldsymbol{\beta} = [\beta_1 \ldots \beta_N]$ is the least square solution of (4.2) and is given by [91, 54]

$$\boldsymbol{\beta} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \left(\mathbf{X}^T \mathbf{y}\right) \tag{4.3}$$

where $\mathbf{y} = [X_{i-1}, X_{i-2}, \ldots, X_{i-K}]$ denote the K context pixels given in Figure 4.2., while the data matrix $\mathbf{X}$

$$\mathbf{X} = \begin{bmatrix} X_{i-1-1} & \cdots & X_{i-1-N} \\ \vdots & \ddots & \vdots \\ X_{i-K-1} & \cdots & X_{i-K-N} \end{bmatrix}$$

consists of the prediction neighbors of $\mathbf{y}$. The expected value of X_i is given by (4.2) and an unbiased estimator of prediction error variance σ^2 can be obtained [54] as

$$\sigma^2 = \frac{1}{K - N - 1} \left(\mathbf{y}^T \mathbf{y} - \boldsymbol{\beta}^T \mathbf{X}^T \mathbf{y}\right). \tag{4.4}$$

Based on the principle that the mean-square prediction for a normal random variable is its mean value, then the density of X_i conditioned on causal neighbors is given by

$$f\left(x_i \middle| x_{i-1}, x_{i-2}, \ldots, x_{i-N}\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\left(\frac{1}{2\sigma^2} \right) \left[x_i - \sum_{j=1}^{N} \beta_j \left(x_{i-j} \right) \right]^2 \right), \qquad (4.5)$$

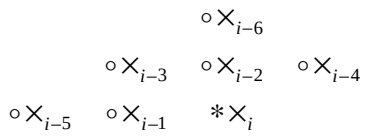


Figure 4.1. Ordering of the causal prediction neighbors of the current pixel $\times_i$, N=6.

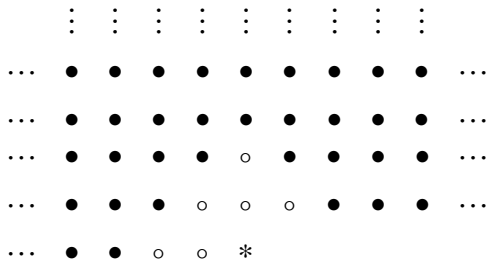


Figure 4.2. The context pixels, denoted by $\bullet$ and $\circ$, used in the covariance estimation of the current pixel $*$. The number of context pixels is K=40.

4.2.3. P.m.f. Estimations in the Second Pass

L_2 Norm Minimizing Probability Estimation: In finer quantal resolutions after the first pass we have to leave aside the gaussian assumption since the image data at finer resolutions behaves more randomly lacking correlation. We thus assume data is independent and update the estimate of the current pixel density by using neighboring densities, that is by minimizing the L_2 norm of causal and non-causal densities. This stems from the fact that in the first pass, most of the time the interval is guessed correctly in one question, leading to Gaussian distributions which fit well to pixels at low resolution (large λ). In later passes the data becomes more independent of each other as more of the redundancy is removed after each pass, resulting in decreased correlation. At this stage we can use the non-causal densities as well, which are densities from the non-causal neighborhood of the pixel from the previous pass. Several probability mass update methods are presented for the second and higher passes. The prediction neighbors used in probability mass

estimation in the second and higher passes are given in Figure 4.3. Note that we have the chance to use the non-causal pixels for prediction.

$$\begin{array}{ccc} \circ\times_{i-3} & \circ\times_{i-2} & \circ\times_{i-4} \\ \circ\times_{i-1} & *\times_{i} & \bullet\times_{i-5} \\ \bullet\times_{i-7} & \bullet\times_{i-6} & \bullet\times_{i-8} \end{array}$$

Figure 4.3. Causal, $\circ$, and non-causal, $\bullet$, neighbors of the current pixel , $*$, used for probability mass estimation in the second and higher passes.

Let $p_i(x)$ denote the probability mass function to be estimated given the causal and non-causal distributions $\{p_{i-j}\}_{j=1}^{N}$. Minimizing $\sum_j \left(p_i(x) - p_{i-j}(x)\right)^2$ subject to the constraint $\sum p_i(x) = 1$ and using Lagrange multipliers we have

$$J(p_i) = \sum_{1 \le j \le N} \left(p_i(x) - p_{i-j}(x)\right)^2 + \lambda\left(\sum p_i(x)\right). \tag{4.6}$$

Using the variational derivative with respect to $p_i(x)$ one finds the distribution to be of the form

$$p_i^{*}(x) = \frac{1}{N}\left(\sum_{j=1}^{N} p_{i-j}(x)\right). \tag{4.7}$$

The method has some desirable properties. If the neighboring interval censored p.m.f.s do not overlap with the current one then they have no negative effect on the estimate. If there exist some overlapping, then an evidence gathering from the causal and non-causal neighbors for the indices of the current interval occurs as they give rise to higher accumulated probabilities for some indices in the interval. Notice this method of summing neighboring densities gives automatically more importance to more precise information residing in the causal neighbor p.m.f.'s concentrated in narrower intervals than to the less precise information in the non-causal ones.

Turnbull Probability Estimator: An interval-censored observation of X is of the form $(X_L, X_R]$ with $X_L < X_R$, where the actual value of X lies in $(X_L < X \le X_R]$. Suppose a sample of N i.i.d observations $(X_L^i, X_R^i]$, for $i = 1,...,N$, is given. Define the indicator variables $\alpha_s^i = I\{x_s \in (X_L^i, X_R^i]\}$, $s = 1,...,255$. With this notation, a self-consistent density estimator p, adopted from Turnbull [92], is given by

$$p[s] = \frac{1}{N} \sum_{1 \le i \le N} \frac{\alpha_s^i p[s]}{\sum_{1 \le l \le 255} \alpha_l^i p[l]} \tag{4.8}$$

Turnbull also gives an iterative algorithm to compute (4.8). A heuristic justification of this method can be made by multiplying both sides by *N*. Then the left hand side of (4.6) is the expected number of individuals in the neighborhood having $X = s$, while the right-hand side is the conditional expected number of individuals in the neighborhood having $X = s$ given the observed interval data.

Note that Turnbull's estimator accumulates the likelihood for the current index *s* the same way L_2 minimizing estimator does, that is the probability of the indices *s* for which i.i.d. intervals overlap is higher than the other indices for which the intervals do not overlap. This seems intuitively reasonable thing to do when the samples are i.i.d.

Hellinger Norm Minimizing Probability Estimation: The relative entropy $D(p\|q)$ is a measure of distance between two distributions. It is a measure of the inefficiency of assuming that the distribution is q when the true distribution is p. For example if we knew true distribution of the random variable, then we could construct a code with average length $H(p)$. If instead, we used the code for distribution q, we would need $H(p) + D(p\|q)$ bits on the average to describe the random variable. The squared Hellinger norm between distributions with densities p and q is defined as

$$H^2(p,q) = \int \left(\sqrt{p(x)} - \sqrt{q(x)}\right)^2 dx \tag{4.9}$$

Many, if not all, smooth function classes satisfy the equivalence $D(p\|\ q)\approx H^2(p,q)$. The advantage of H^2 is that it satisfies triangle inequality while D does not. However D brings in clean information theoretic identities [93] such as minimum description length principle, stochastic complexity, etc. Taking advantage of the equivalence between D and H^2 we can use one for the other in the derivation of the optimal $p_i^*(x)$.

When we have a class of candidate densities $\{p_{i-j} : j=1,\ldots,N\}$ and want to find the $p_i^*(x)$, which minimizes the inefficiency of assuming the distribution was p_{i-j}, we can minimize the total extra bits to obtain the shortest description length on the average:

$$J(p_i)=\sum_{1\le j\le N}\int\left(\sqrt{p_i(x)}-\sqrt{p_{i-j}(x)}\right)^2 dx+\lambda\int p_i(x)dx \tag{4.10}$$

where λ is the Lagrange multiplier. Again finding the variational derivative with respect to $p_i(x)$ and setting it equal to zero, we get

$$p_i^*(x)=T\left(\sum_{1\le j\le N}\sqrt{p_{i-j}(x)}\right)^2 \tag{4.11}$$

where *T* is the normalizing constant. In general the relative entropy or Kullback-Leibler distance has a close connection with more traditional statistical estimation measures such as L_2 norm (MSE) and Hellinger norm when the distributions are bounded away from zero, and is equivalent to MSE when both p and q are Gaussian distributions with the same covariance structure [93]. Like Turnbull's method this method is not used in p.m.f. estimation because it is similar in performance to (4.5) but computationally more expensive.

4.2.4. Multi-hypothesis Testing

We can treat the problem of estimating the interval where the current pixel value is in, within the framework of multi-hypotheses testing [94]. Let the $H_1,...,H_M$ denote the M hypotheses where every hypothesis m is associated with an interval $(L_m, R_m]$ that has a length of $2\delta+1$. The random variable X has a probability mass under each hypothesis $H=m$ and denote this probability mass by

$$p_{X|H}(x|m)=\sum_{i\in(L_m,R_m]} p_X(i) \tag{4.12}$$

When each hypothesis has an a priori probability, $p_m=\Pr\{H=m\}$, the cumulative probability mass of $H=m$ and $X=x$ is then $p_{X|H}(x|m)p_m$. The a posteriori probability that $H=m$ conditional on $X=x$ is

$$p_{H|X}(m|x)=\frac{p_{X|H}(x|m)p_m}{\sum_{m=1}^{M} p_{X|H}(x|m)p_m}. \tag{4.13}$$

The rule for maximizing the probability of being correct, so as to minimize the number of guesses in finding the correct interval, is to choose that m for which $p_{H|X}(m|x)$ is maximized. This is denoted by

$$\hat{H}=\arg\max_m\left[p_m\, p_{H|X}(m|x)\right] \tag{4.14}$$

and known as maximum a posteriori (MAP) rule. For equal probabilities, this becomes the maximum likelihood (ML) rule where we simply choose the hypothesis with the largest likelihood

$$\hat{H}=\arg\max_m\left[p_{X|H}(x|m)\right]. \tag{4.15}$$

We use ML rule in the first pass, while MAP rule is used in the following passes since we have the a priori probability mass from the previous passes. Defining the indicator function as

$$\chi_x^m = I\{x \in (L_m, R_m]\} \tag{4.16}$$

where after hypothesis test the $(L_m, R_m]$ is the correct interval for the current pixel with highest probability mass in it, the entropy coder is fed with one or zero, respectively, if the indicator function is one or zero.

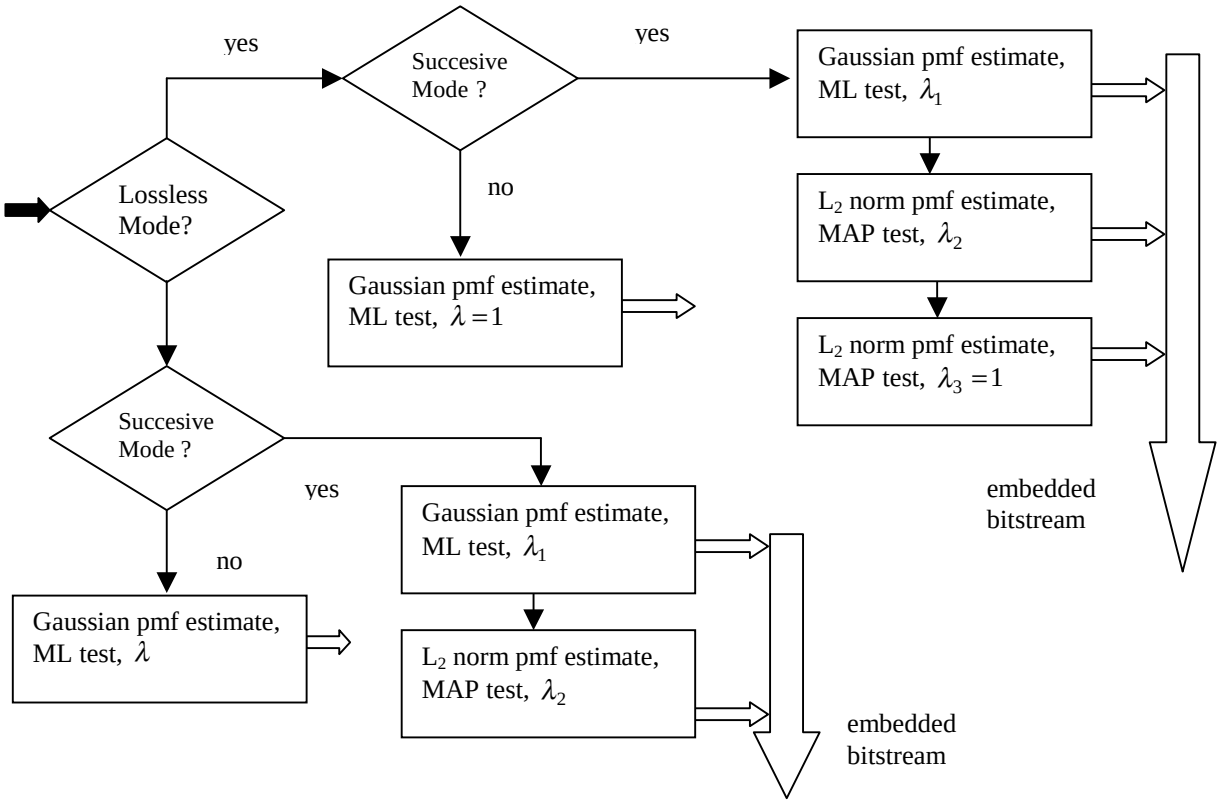


Figure 4.4. Schematic description of the overall compression scheme.

4.3. The Compression Method

The schematic description of our algorithm is given in Figure 4.4. We have integrated near-lossless compression and successive refinement with lossless compression

in one single framework. Lossless or near-lossless compression can be achieved either with successive refinement or in one step, called the direct method.

In successive mode lossless compression, the support of the p.m.f. is successively refined and narrowed down to one. In the first pass, for every pixel the p.m.f. is estimated with (4.5) using linear prediction and multi-hypothesis testing (4.15) to constrain the support (or length) of the current p.m.f. to λ_1. The details of the p.m.f. support constraining are given in Figure 4.5.

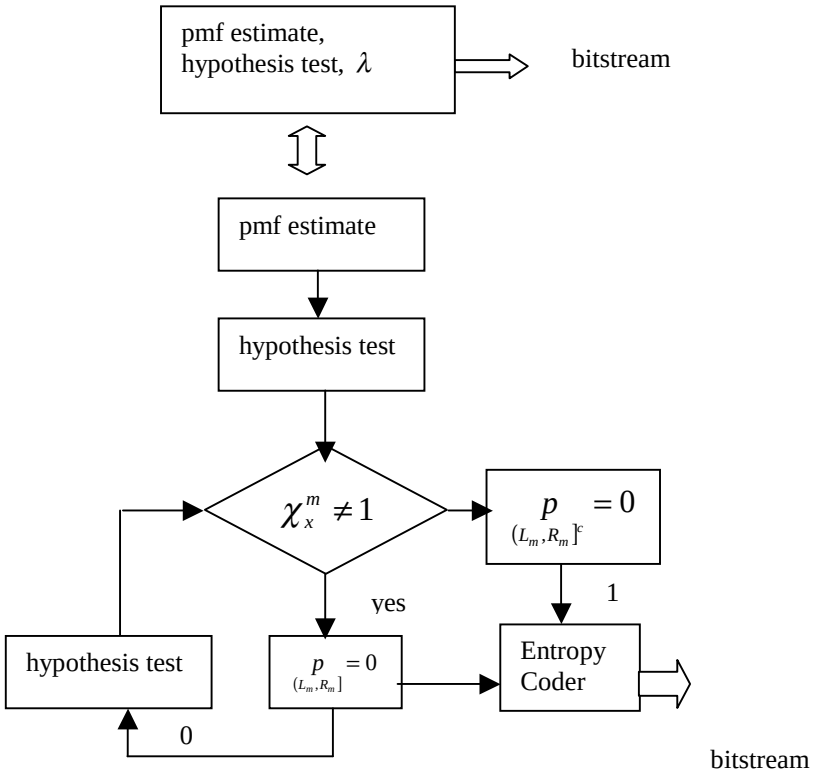


Figure 4.5 Details of the encoder block used in Figure 4.4. Here λ is the length of the interval $(L_m, R_m]$.

The probability mass within an interval is zeroed out every time we fail in making the correct guess and feed an arithmetic encoder with zero until we find the correct interval. The whole probability mass is within a fixed length interval when we proceed to the next pixel.

Successive near-lossless compression is similarly performed in several passes. The first pass is the same as the direct near-lossless case, but in the following passes, for every pixel the p.m.f.s are successively refined and the lengths of the supports for the pixels are narrowed down to the desired precision. P.m.f.s are estimated with method (4.7) and the intervals where the pixel lies are determined by MAP test (4.14) as we have the a priori probability mass for the current pixel from the previous pass.

In successive mode, for both lossless and near-lossless compression, the interval of the current pixel at second or the following passes can be narrowed in two ways: One way is to split it into two intervals and to perform binary hypothesis test. The other way is to split the current interval into more than two equal length intervals and to perform multiple hypothesis test.

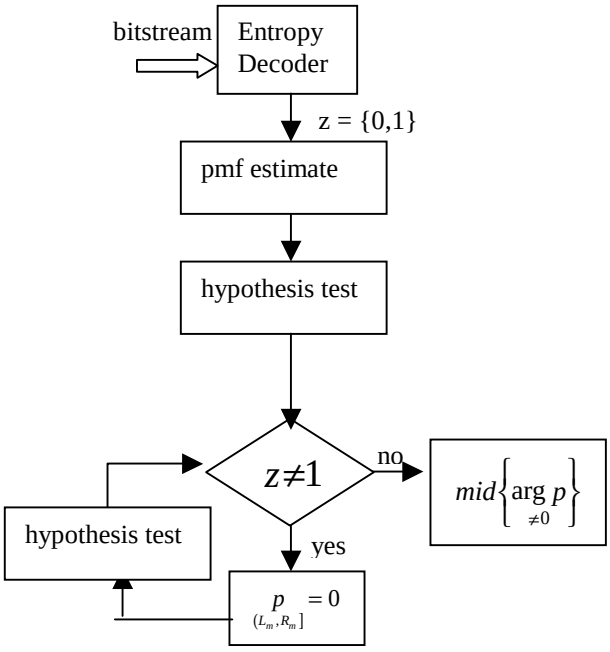


Figure 4. 6. The decoder is a replica of the encoder.

The decoder given in Figure 4.6. is similar to the encoder. Based on the estimated p.m.f. and decoded sequences of successes and failures, the p.m.f. support for the current pixel is found. Within the support of the estimated p.m.f. the intervals are sorted with respect to their probability mass and the correct interval is found using decoded failure and success events. Recall that the decoded events are the outcomes of the tests whether the

current pixel lies in that interval or not. Whenever the decoded event *z* is not success for the given interval, this interval is discarded, until the correct interval is tested with success. The reconstructed value of the current pixel is taken as the midvalue of the successfully tested interval, which guarantees that the error in reconstructed value is not more than half of the interval length minus one, that is $\|e\|_\infty = (\lambda - 1)/2$, where λ is the length of the interval.

4.4. Experimental Results

Successive and direct mode lossless compression results, which are obtained with three passes are given in Table 4.1. The interval lengths of the pixel values are taken 8 in the first pass, 4 in the second pass and 1 in the final third pass in the successive mode. One-pass results are also given in the same table. Near-lossless results obtained with one pass for $\delta = 1$, $\delta = 3$ and $\delta = 7$, are given in Tables 4.2., 4.3., and 4.4. respectively.

Table 4.1. Comparison of lossless compression results: proposed method versus CALIC.

	Proposed		CALIC [85]
$\delta = 0$	3-pass	1-pass	
BARB	4.18	4.21	4.45
LENA	4.07	4.07	4.11
ZELDA	3.81	3.79	3.86
GOLDHILL	4.65	4.69	4.63
BOAT	4.06	4.08	4.15
MANDRILL	5.89	5.96	5.88
PEPPERS	4.41	4.37	4.38
MEAN	4.44	4.45	4.49

Table 4.2. Comparison of 4 different methods of near-lossless compression $(\delta = 1)$

Image	SPIHT [52]			CALIC [85]			CALIC [95]			Proposed		
	Bpp	PSNR	$\|e\|_\infty$	bpp	PSNR	$\|e\|_\infty$	bpp	PSNR	$\|e\|_\infty$	Bpp	PSNR	$\|e\|_\infty$
Hotel	2.70	76.93	6	2.76	49.95	1	2.70	50.07	1	2.84	49.92	1
Ultrasound	1.60	46.33	8	1.99	52.36	1	1.60	51.76	1	1.70	49.84	1
Café	3.22	47.09	7	3.30	49.97	1	3.22	50.09	1	3.44	49.85	1
Barbara				2.94	49.91	1				2.65	49.88	1
Finger	3.82	47.27	5	3.90	49.89	1	3.82	49.98	1	3.80	49.89	1

Table 4.3. Comparison of 4 different methods of near-lossless compression ($\delta = 3$)

Image	SPIHT [52]			CALIC [85]			CALIC [95]			Proposed		
	Bpp	PSNR	$\|e\|_\infty$	Bpp	PSNR	$\|e\|_\infty$	bpp	PSNR	$\|e\|_\infty$	Bpp	PSNR	$\|e\|_\infty$
Hotel	1.69	41.77	11	1.74	42.31	3	1.67	42.37	3	1.75	42.30	3
Ultrasound	1.09	41.42	16	1.51	45.04	3	1.09	44.86	3	1.30	42.22	3
Café	2.19	41.42	16	2.27	42.41	3	2.19	42.49	3	2.31	42.21	3
Barbara				1.92	42.23	3				1.61	42.27	3
Finger	2.67	40.71	15	2.73	42.11	3	2.67	42.18	3	2.63	42.10	3

Table 4.4. Comparison of 4 different methods of near-lossless compression ($\delta = 7$)

Image	SPIHT [52]			CALIC [85]			CALIC [95]			Proposed		
	Bpp	PSNR	$\|e\|_\infty$	Bpp	PSNR	$\|e\|_\infty$	bpp	PSNR	$\|e\|_\infty$	Bpp	PSNR	$\|e\|_\infty$
Hotel	0.95	37.76	20	0.97	36.50	7	0.95	36.75	7	0.96	36.42	7
Ultrasound	0.72	37.17	25	0.99	37.19	7	0.72	38.55	7	0.78	36.48	7
Café	1.43	36.49	30	1.50	36.31	7	1.43	36.45	7	1.48	35.80	7
Barbara				1.19	36.21	7				0.91	36.27	7
Finger	1.77	35.90	21	1.80	35.43	7	1.77	35.59	7	1.73	35.43	7

Table 4.5. Comparison of bit/pixel efficiency and peak signal to noise ratio in dB of the proposed algorithm versus the CALIC [85] algorithm.

	bit/pixel gain	PSNR gain
$\delta = 1$	0.05	-0.01
$\delta = 2$	0.07	0.01
$\delta = 3$	0.08	0.03
$\delta = 4$	0.09	0.05
$\delta = 5$	0.10	0.07
$\delta = 6$	0.11	0.07
$\delta = 7$	0.10	0.08

Notice that the lossless compression results in bits per pixel obtained with one pass and three passes are almost the same, contrary to one's expectation that more than one pass would yield better performance as the non-causal information is available in the following passes. This is because the non-causal information is less precise than the causal information.

4.5. Conclusion

In this section, we have presented a technique that unifies progressive transmission and near-lossless compression in one single bit stream. The proposed technique produces a bitstream that results in progressive reconstruction of the image just like what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after each layer of the successively refinable bitstream is decoded. Furthermore, the compression performance provided by the proposed technique is superior or comparable to the best-known lossless and near-lossless techniques proposed in the literature [81, 85, 95, 52].

The originality of the method consists in looking at the image data compression as one of asking the optimal questions to determine the interval in which the current pixel lies. With successive passes of the image, the length of this interval is gradually narrowed until it becomes of length one, in which case we have lossless reconstruction. Stopping the process in any intermediate stage gives near-lossless compression. Although our experimental results show that the proposed method brings in only modest gains in dB measure or bit per pixel efficiency, we believe that there is room for improvement. Our future work will explore different avenues for improving upon the results given in this paper. For example, we have no mechanism for learning global or repeated patterns in the image. Context based techniques like CALIC, keep a history of past events in a suitably quantized form and use these to better model subsequent events. We believe such mechanisms when incorporated within our framework will give additional improvements.

The proposed techniques provides a flexible framework and many variations of the basic method are possible. For example, the quality of reconstruction as defined by the near-lossless parameter k can be made to vary from region to region or even from pixel to pixel based on image content or other requirements. Given this fact, different regions in the image can be refined to different desired precision using HVS properties. To this effect, flat regions where compression artifact visibility is higher can be refined more accurately, thus achieving perceptually lossless compression. In addition, it would be interesting to extend the technique to multispectral images.

5. CONCLUSIONS AND FUTURE PERSPECTIVES

Firstly, we have presented collectively a set of image quality measures in their multispectral version and categorized them. Statistical investigation of 26 different measures using a ANOVA analyses has revealed that local phase-magnitude measures (*S2 or S5)*, HVS-filtered *L1*, *L2* norms, edge stability measure are most sensitive to coding, blur and artifacts, while the mean square error (*D1*) remains as the measure for additive noise. One can state that combined spectral phase-magnitude measures and HVS filtered error norms should be paid more attention in the design of coding algorithms and sensor evaluation. On the other hand the pixel-difference based measures remain still to be the measures responsive to distortions and least affected by image variety.

The Kohonen map of the measures has been useful in depicting similar ones, and identifying the ones that are sensitive possibly to different distortion artifacts in compressed images. The correlation between various measures has been depicted via Kohonen's Self-Organizing Map. The placement of measures in the two-dimensional map has been in agreement with one's intuitive grouping.

Future work will address subjective experiments and prediction of subjective image quality using the above salient measures identified. Another possible avenue is to combine various "fundamental" metrics for better performance prediction.

Secondly, we have addressed the problem of steganalysis of watermarked and steganographically marked images. That is, we developed techniques for discriminating between cover-images and stego-images. Our approach is based on the hypothesis that a particular message embedding scheme leaves statistical evidence or structure that can be exploited for detection with the aid of proper selection of image features and multivariate regression analysis. We showed that the distance between an unmarked image and its filtered version is different than the distance between a marked image and its filtered version. We used image quality metrics as the feature set to distinguish between marked and non-marked images. To identify specific quality measures, which provide the best discriminative power, we used ANOVA technique. We have also pointed out the image

features that should be taken more seriously into account in the design of more successful watermarking or steganographic techniques to eschew detection. After selecting an appropriate feature set, we used multivariate regression techniques to get an optimal classifier using an image and its filtered version. Simulation results with well known and commercially available watermarking and steganographic techniques indicate that the selected IQMs form a multidimensional feature space whose points cluster well enough to do a classification of marked and non-marked images. The classifier is still able to do a classification when the tested images come from an embedding technique unknown to it, indicating that it has a generalizing capability of capturing the general intrinsic characteristics of watermarking and steganographic techniques.

Proposed steganalysis technique is its infancy. The measures that are incorporated in the watermarking and steganographic system designs should be investigated for steganalysis. We have reduced the problem to classification, so the wealth of tools in pattern recognition i.e., feature selection methods, nonlinear or neural classifiers can be used for a better classification performance, since the real world problems are non-linear. In fact, one of the consequences of this work is that current watermarking and steganographic techniques are so weak for steganographic applications that they lend themselves easily to linear solutions.

Thirdly, we have presented a technique that unifies progressive transmission and near-lossless compression in one single bit stream. The proposed technique produces a bitstream that results in progressive reconstruction of the image just like what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after each layer of the successively refinable bitstream is decoded. Furthermore, the compression performance provided by the proposed technique is superior or comparable to the best-known lossless and near-lossless techniques proposed in the literature.

The originality of the method consists in looking at the image data compression as one of asking the optimal questions to determine the interval in which the current pixel lies. With successive passes of the image, the length of this interval is gradually narrowed until it becomes of length one, in which case we have lossless reconstruction. Stopping the process in any intermediate stage gives near-lossless compression. Although our

experimental results show that the proposed method brings in only modest gains in dB measure or bit per pixel efficiency, we believe that there is room for improvement. Our future work will explore different avenues for improving upon the results given in this thesis. For example, we have no mechanism for learning global or repeated patterns in the image. Context based techniques like CALIC, keep a history of past events in a suitably quantized form and use these to better model subsequent events. We believe such mechanisms when incorporated within our framework will give additional improvements.

The proposed techniques provides a flexible framework and many variations of the basic method are possible. For example, the quality of reconstruction as defined by the near-lossless parameter k can be made to vary from region to region or even from pixel to pixel based on image content or other requirements. Given this fact, different regions in the image can be refined to different desired precision using HVS properties. To this effect, flat regions where compression artifact visibility is higher can be refined more accurately, thus achieving perceptually lossless compression. It would be interesting to extend the technique to multispectral images.

Finally, multidimensional pmf estimation and optimal question approach may result in several birds with one shot.

REFERENCES

1. Perlmutter, S.M., P.C. Cosman, R.M. Gray, R.A. Olshen, D. Ikeda, C.N. Adams, B.J. Betts, M.B. Williams, K.O. Perlmutter, J. Li, A. Aiyer, L. Fajardo, R. Birdwell and B.L. Daniel, "Image Quality in Lossy Compressed Digital Mammograms", *Signal Processing*, Vol. 59, pp. 189-210, 1997.

2. Lambrecht, C. B., Ed., "Special Issue on Image and Video Quality Metrics", *Signal Processing*, Vol. 70, 1998.

3. Lehmann, T., A. Sovakar, W. Schmitt and R. Repges, "A comparison of Similarity Measures for Digital Subtraction Radiography", *Computational Biological Medicine*, Vol. 27, No. 2, pp. 151-167, 1997.

4. Eskicioğlu, A. M., "Application of Multidimensional Quality Measures to Reconstructed Medical Images", *Optical Engineering* Vol. 35, No. 3, pp. 778-785, 1996.

5. Eskicioğlu, A. M. and P. S. Fisher, "Image Quality Measures and Their Performance", *IEEE Transactions on Communications*, 43(12), 2959-2965 (1995).

6. Ridder, H., "Minkowsky Metrics as a Combination Rule for Digital Image Coding Impairments", in *Proceedings SPIE 1666: Human Vision, Visual Processing, and Digital Display III*, pp. 17-27 1992.

7. Watson, A. B. (Ed.), *Digital Images and Human Vision,* Cambridge, MA, MIT Press, 1993.

8. Girod, B., "What's Wrong with Mean-squared Error", in A. B. Watson (Ed.), *Digital Images and Human Vision,* Chapter 15, Cambridge, MA, MIT Press 1993.

9. Miyahara, M., K. Kotani and V. R. Algazi, "Objective Picture Quality Scale (PQS) for Image Coding", *IEEE Transactions On Communications*, Vol. 46, No. 9, pp. 1213-1226, 1998.

10. Nill, N. B. and B. H. Bouzas, "Objective Image Quality Measure Derived From Digital Image Power Spectra", *Optical Engineering*, Vol. 31, No. 4, pp. 813-825, 1992.

11. Franti, P., "Blockwise Distortion Measure for Statistical and Structural Errors in Digital Images" *Signal Processing: Image Communication,* 13, 89-98 (1998).

12. Winkler, S., "A perceptual distortion metric for digital color images." in *Proceedings of the Fifth International Conference on Image Processing*, vol. 3, pp. 399-403, Chicago, Illinois, October 4-7, 1998.

13. Daly, S., "The visible differences predictor: An algorithm for the assessment of image fidelity", in A. B.Watson (Ed.), *Digital Images and Human Vision,* Cambridge, MA, MIT Press, pp. 179-205, 1993.

14. Frese, T., C. A. Bouman and J. P. Allebach, "Methodology for Designing Image Similarity Metrics Based on Human Visual System Models", *Proceedings of SPIE/IS&T Conference on Human Vision and Electronic Imaging II*, Vol. 3016, pp. 472-483, San Jose, CA, 1997.

15. CCIR, Rec. 500-2, Method for the Subjective Assessment of the Quality of Television Pictures, 1986.

16. Van Dijk, M. and J. B. Martens, "Subjective Quality Assessment of Compressed Images", *Signal Processing*, Vol. 58, pp. 235-252, 1997.

17. Rohaly, A.M., P. Corriveau, J. Libert, A. Webster, V. Baroncini, J. Beerends, J.L Blin, L. Contin, T. Hamada, D. Harrison, A. Hekstra, J. Lubin, Y. Nishida, R. Nishihara, J. Pearson, A. F. Pessoa, N. Pickford, A. Schertz, M. Visca, A. B. Watson and S.

Winkler: "Video Quality Experts Group: Current results and future directions." *Proceedings SPIE Visual Communications and Image Processing*, Vol. 4067, Perth, Australia, June 21-23, 2000.

18. Corriveau, P. and A. Webster, "VQEG Evaluation of Objective Methods of Video Quality Assessment", SMPTE Journal, Vol. 108, pp. 645-648, 1999.

19. Kanugo, T. and R. M. Haralick, "A Methodology for Quantitative Performance Evaluation of Detection Algorithms", *IEEE Transactions on Image Processing*, Vol. 4, No. 12, pp. 1667-1673, 1995.

20. Matrik, R., M. Petrou and J. Kittler, "Error-Sensitivity Assessment of Vision Algorithms", *IEE Proceedings-Vis. Image Signal Proces*sing, Vol. 145, No. 2, pp. 124-130, 1998.

21. Grim, M. and H. Szu, "Video Compression Quality Metrics Correlation with Aided Target Recognition (ATR) Applications", *Journal of Electronic Imaging*, Vol. 7, No. 4, pp. 740-745, 1998.

22. Barrett, H. H., "Objective Assessment of Image Quality: Effects of Quantum Noise and Object Variability", *Journal of Optical Society of America*, Vol. A, No. 7, pp. 1261-1278, 1990.

23. Barrett, H. H., J. L. Denny, R. F. Wagner and K. J. Myers, "Objective Assessment of Image Quality II: Fisher Information, Fourier-Crosstalk, and Figures of Merit for Task Performance", *Journal of Optical Society of America*, Vol. A, No. 12, pp. 834-852, 1995.

24. Halford, C.E., K.A. Krapels, R.G. Driggers and E.E. Burroughs, "Developing Operational Performance Metrics Using Image Comparison Metrics and the Concept of Degradation Space", *Optical Engineering*, Vol. 38, No. 5, pp. 836-844, 1999.

25. Dubuisson, M. P. and A. K. Jain, "A Modified Hausdorff Distance for Object Matching", *International Conference on Pattern Recognition*, Vol. A, pp. 566-569, Jerusalem, 1994.

26. International Commission of Illumination (CIE), *Recommendations on Uniform Color Spaces, Color Difference Equations, Psychometric Color Terms*, Publication CIE 15 (E.-1.3.1), Supplement No. 2, Bureau Central de la CIE, Vienna, 1971.

27. Jain, A. K., *Fundamentals of Digital Image Processing*, Prentice Hall, New Jersey, 1989.

28. DiGesu, V., and V. V. Staravoitov, "Distance-based Functions for Image Comparison", *Pattern Recognition Letters*, Vol. 20, No. 2, pp. 207-213, 1999.

29. Starovoitov, V. V., C. Köse and B. Sankur, "Generalized Distance Based Matching of Nonbinary Images", *IEEE International Conference on Image Processing*, Chicago, 1998.

30. Juffs, P., E. Beggs and F. Deravi, "A Multiresolution Distance Measure for Images", *IEEE Signal Processing Letters*, Vol. 5, No. 6, pp.138-140, 1998.

31. Andreutos, D., K. N. Plataniotis and A. N. Venetsanopoulos, "Distance Measures for Color Image Retrieval", *IEEE International Conference On Image Processing,* Chicago, 1998.

32. Pratt, W. K., *Digital Image Processing*, New York, Wiley, 1978.

33. Trahanias, P. E., D. Karakos, A. N. Venetsanopoulos, "Directional Processing of Color Images: Theory and Experimental Results", *IEEE Transactions Image Processing*, Vol. 5, No. 6, pp. 868-880, 1996.

34. Zetsche, C., E. Barth and B. Wegmann, "The Importance of Intrinsically Two-Dimensional Image Features in Biological Vision and Picture Coding," in A. B.

Watson (Ed.), *Digital Images and Human Vision,* Cambridge, MA, MIT Press, pp. 109-138, 1993.

35. Rajan, P. K. and J. M. Davidson, "Evaluation of Corner Detection Algorithms", *Proceedings of Twenty-First Southeastern Symposium on System Theory*, pp. 29-33, 1989.

36. Canny, J., "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679-698, 1986 .

37. Carevic, D. and T. Caelli, "Region Based Coding of Color Images Using KLT", *Graphical Models and Image Processing,* Vol. 59, No. 1, pp. 27-38, 1997.

38. Tao, H. and T. Huang, "Color Image Edge Detection using Cluster Analysis", *IEEE International Conference On Image Processing*, pp. 834-836*,* California, 1997.

39. Trahanias, P. E. and A. N. Venetsanopoulos, "Vector Order Statistics Operators as Color Edge Detectors", *IEEE Transactions on System Man and Cybernetics*, Vol. 26, No. 1, pp. 135-143, 1996.

40. Lipschutz, M. M., *Theory and Problems of Differential Geometry*, McGraw-Hill Inc., 1969.

41. McIvor, M. and R. J. Valkenburg, "A Comparison of Local Surface Geometry Estimation Methods", *Machine Vision and Applications*, Vol. 10, pp.17-26, 1997.

42. Avcıbaş, İ., B. Sankur and K. Sayood, "New Distortion Measures for Color Image Vector Quantization", *ISAS'98: World Multiconference on Systemics, Cybernetics and Informatics II,* pp. 496-503, July 12-16, Orlando, USA, 1998.

43. Duda, R. O. and P. E. Hart, *Pattern Recognition and Scene Analysis*, New-York, Wiley, 1973.

44. Popat, K. and R. Picard, "Cluster Based Probability Model and It's Application to Image and Texture Processing", *IEEE Transactions on Image Processing*, Vol. 6, No. 2, pp.268-284, 1997.

45. Basseville, M., "Distance Measures for Signal Processing and Pattern Recognition", *Signal Processing* Vol. 18, pp. 349-369, 1989.

46. Nill, N. B., "A Visual Model Weighted Cosine Transform for Image Compression and Quality Assessment", *IEEE Transactions on Communications*, Vol. 33, No. 6, pp. 551-557, 1985.

47. Avcıbaş, İ. and B. Sankur, "Statistical Analysis of Image Quality Measures", in *Proceedings EUSIPCO'2000: 10th European Signal Processing Conference*, pp. 2181-2184, 5-8 September, Tampere, Finland, 2000.

48. Avcıbaş, İ. and B. Sankur, "Çokbantli Imgelerde Nitelik Ölçülerinin Istatistiksel Davranisi", *SIU'99: IEEE 1999 Sinyal İşleme ve Uygulamaları Kurultayı*, Bilkent, Ankara, 1999.

49. Avcıbaş, İ. and B. Sankur, "Kodlama Algoritmalarinin Bilgi Içerigini Koruma Basarimlari", *SIU'98: 6. Sinyal Isleme ve Uygulamalari Kurultayi*, pp. 248-253, Kizilcahamam, Ankara, 1998.

50. Frese, T., C. A. Bouman and J. P. Allebach, *A Methodology for Designing Image Similarity Metrics Based on Human Visual System Models*, Tech. Rep. TR-ECE 97-2, Purdue University, West Lafayette, IN, 1997.

51. Wallace, G. K., "The JPEG Still Picture Compression Standard", *IEEE Transactions on Consumer Electronics*, Vol. 38, No. 1, pp. 18-34, 1992.

52. Said, A. and W. A. Pearlman, *"*A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees*"*, *IEEE Transactions Circuits and Systems: Video Technology*, Vol. 6, No. 3, pp. 243-250, 1996.

53. Martinez, A.M. and R. Benavente, *The AR Face Database*, CVC Technical Report No. 24, June 1998.

54. Rencher, A. C., *Methods of Multivariate Analysis*, John Wiley, New York, 1995.

55. Kohonen, T., *Self-Organizing Maps,* Springer-Verlag, Heidelberg, 1995.

56. Watson, A. B., "DCTune: A Technique for Visual Optimization of DCT Quantization Matrices for Individual Images", *Society for Information Display Digest of Technical Papers*, Vol. 24, pp. 946-949, 1993.

57. Simmons, G. J., "The prisoners' problem and the subliminal channel," in *Proceedings IEEE Workshop Communications Security CRYPTO'8*3, Santa Barbara, CA, 1983, pp. 51–67.

58. Simmons, G. J., "The history of subliminal channels," *IEEE Journal of Selected Areas on Communications*, vol. 16, pp. 452–462, May 1998.

59. Petitcolas, F. A. P., R. J. Anderson and M. G. Kuhn, "Information Hiding—A Survey", *Proceedings of the IEEE*, Vol. 87, pp.1062-1078, July 1999.

60. Johnson, N. F. and S. Katzenbeisser, "A Survey of steganographic techniques", in S. Katzenbeisser and F. Petitcolas (Eds.): *Information Hiding*, pp. 43-78. Artech House, Norwood, MA, 2000.

61. Dugelay, J-L. and S. Roche, "A Survey of current watermarking techniques", in S. Katzenbeisser and F. Petitcolas (Eds.): *Information Hiding*, pp. 43-78. Artech House, Norwood, MA, 2000.

62. Langelaar, G. C., I. Setyawan and R. L. Lagendijk, "Watermarking Digital Image and Video Data, A State of-the-Art Overview", *IEEE Signal Processing Magazine*, pp. 20-46, September 2000.

63. Hartung, F. and M. Kutter, "Multimedia Watermarking Techniques", *Proceedings of the IEEE*, Vol. 87, pp.1079-1107, July 1999.

64. Westfield, A. and A. Pfitzmann, "Attacks on Steganographic Systems", in A. Pfitzmann (Ed.): *Information Hiding*, LNCS 1768, pp. 61-76, Springer-Verlag Berlin Heidelberg, 1999.

65. Johnson, N. F. and S. Jajodia, "Steganalysis: The investigation of Hidden Information", *IEEE Information Technology Conference*, Syracuse, NY, USA, 1-3 September 1998.

66. Johnson, N. F. and S. Jajodia, "Steganalysis of Images created using current steganography software", in D. Aucsmith (Ed.): *Information Hiding*, LNCS 1525, pp. 32-47. Springer-Verlag Berlin Heidelberg 1998.

67. Fridrich, J., R. Du and M. Long, "Steganalysis of LSB Encoding in Color Images", *ICME 2000*, New York City, July 31-August 2, New York, USA, 2000.

68. Avcıbaş, İ., N. Memon and B. Sankur, "Steganalysis Using Image Quality Metrics", (under review), *IEEE Transactions on Image Processing*, 2001.

69. Avcıbaş, İ., N. Memon and B. Sankur, "Automatic Detection of the Presence of Stego-signals and Watermarks in Images", to be presented, *ECVP'2001: European Conference on Visual Perception*, Antalya, Turkey, 2001.

70. Avcıbaş, İ., N. Memon and B. Sankur, "Steganalysis Based On Image Quality Metrics", to be presented, *MMSP'2001: IEEE Workshop on Multimedia Signal Processing*, Cannes, France, 2001.

71. Avcıbaş, İ., N. Memon and B. Sankur, "Staganalysis of Watermarking Techniques Using Image Quality Metrics", *SPIE Conference 4314: Security and Watermarking of Multimedia Contents III*, San Jose, USA, 2001.

72. Avcıbaş, İ., B. Sankur and N. Memon, "İmge Nitelik Ölçütlerine Dayanarak Damgalama Sezimi", *SIU'2001: IEEE 2001 Sinyal İşleme ve Uygulamaları Kurultayı*, Lefkoşa, KKTC, 2001.

73. Avcıbaş, İ., B. Sankur and K. Sayood, "Statistical Evaluation of Image Quality Measures", (under review), *Journal of Electronic Imaging*, 2001.

74. Kutter, M., S. Voloshynovskiy and A. Herrigel, "The Watermark Copy Attack", *SPIE Conf. on Security and Watermarking of Multimedia Contents II*, pp. 371-380, San Jose, USA, 2000.

75. PictureMarc, Embed Watermark, v.1.00.45, Copyright © 1996, Digimarc Corporation.

76. Kutter, M. and F. Jordan, *JK-PGS (Pretty Good Signature).* Signal Processing Laboratory at Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, 1998, http://ltswww.epfl.ch/~kutter/watermarking/JK_PGS.html.

77. Cox, J., J. Kilian, T. Leighton and T. Shamoon, "Secure Spread Spectrum Watermarking For Multimedia*", IEEE Transactions Image Processing*, Vol. 6, pp. 1673-1686, 1997.

78. Steganos II, 2001, http://www.steganos.com/english/steganos/download.htm.

79. Brown, A., S-Tools 4.0, 1996, http://members.tripod.com/steganography/stego/s-tools4.html.

80. Korejwa, J., Jsteg Shell 2.0, 2001, http://www.tiac.net/users/korejwa/steg.htm.

81. Ansari, R., N. Memon and E. Ceran, "Near-lossless Image Compression Techniques", *Journal of Electronic Imaging,* Vol. 7, No. 3, pp. 486-494, July 1998.

82. Ebrahimi, T. and M. Kunt, "Visual data compression for multimedia applications", *Proceedings of the IEEE*, vol. 86, pp.1109-1125, May 1999.

83. Egger, O., P. Fleury, T. Ebrahimi, and M. Kunt, "High-Performance Compression of Visual Information—A Tutorial Review—Part I: Still Pictures", *Proceedings of the IEEE*, Vol. 87, No. 6, June 1999.

84. Marcellin, M. W., M. J. Gormish, A. Bilgin and M. P. Boliek, "An overview of JPEG-2000", *Proceedings 2000 Data Compression Conference*, Snowbird, Utah, March 2000.

85. Wu, X. and N. Memon, "Context-Based, Adaptive, Lossless Image Coding", *IEEE Transactions On Communications*, pp. 437-444, April 1997.

86. Avcıbaş, İ., N. Memon and B. Sankur, "Lossless and Near-Lossless Image Compression with Successive Refinement", (under review), *IEEE Signal Processing Letters*, 2001.

87. Avcıbaş, İ., N. Memon, B. Sankur and K. Sayood, "Lossless and Near-Lossless Compression with Successive Refinement", *SPIE Conference 4310: Visual Communications and Image Processing*, San Jose, USA, 2001.

88. Avcıbaş, İ., B. Sankur and N. Memon, "Yitimsiz ve Sınırlı -Yitimli Kodlamada Bağlam Olasılıklarının Kullanımı", *SIU'2001: IEEE 2001 Sinyal İşleme ve Uygulamaları Kurultayı*, Lefkoşa, KKTC, 2001.

89. Massey, J. L., "Guessing and Entropy", *in Proceedings 1994 IEEE International Symposium on Information Theory,* p. 204, Trondheim, Norway, 1994.

90. Equitz, H. R. and T. Cover, "Successive Refinement of Information", *IEEE Transactions on Information Theory*, Vol. 37, No. 2, pp. 269-275, March 1991.

91. Jayant, N. S. and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, 1984.

92. Turnbull, B. W., "The Empirical Distribution Function with Arbitrary Grouped, Censored and Truncated Data", *Journal Royal Statist. Soc. Ser. B*, Vol. 38, pp. 290-295, 1976.

93. Barron, A., J. Rissanen and B. Yu, "The Minimum Description Length Principle in Coding and Modeling", *IEEE Transactions on Information Theory*, Vol. 44, No. 6, pp. 2743-2760, October 1998.

94. Van Trees, H. L., *Detection Estimation and Modulation Theory Part I*, Wiley, 1968.

95. Wu, X. and P. Bao, "L_∞ Constrained High-Fidelity Image Compression via Adaptive Context Modeling", *IEEE Transactions on Image Processing*, pp.536-542, Apr. 2000.

REFERENCES NOT CITED

Arikan, E., "An Inequality on Guessing and its Application to Sequential Decoding", *IEEE Transactions On Information Theory*, Vol. 42, No. 1, pp. 99-105, January 1991.

Avcıbaş, İ., B. Sankur, K. Sayood and N. Memon, "Component Ratio Preserving Compression for Remote Sensing Applications", *Proceedings SPIE Conf. 3974: Image and Video Communications and Processing*, pp. 700-708, San Jose, USA, 2000.

Barni, M., F. Bartolini, and A. Piva, "Digital Waermarking Of Visual Data: State of the Art and New Trends", *EUSIPCO'2000: 10th European Signal Processing Conference*, 5-8 September, Tampere, Finland, 2000.

Barron, A. R. and T. Cover, "Minimum Complexity Density estimation", *IEEE Transactions on Information Theory*, Vol. 37, No. 4, pp. 1034-1054, July 1991.

Berger, T. and J. D. Gibson, "Lossy Source Coding", *IEEE Transactions On Information Theory*, Vol. 44, No. 6, pp. 2693-2723, October 1998.

Brocket, P. L., A. Charnes and K. H. Paick, "Information Theoretic Approach to Unimodal Density Estimation", *IEEE Transactions on Information Theory*, Vol. 41, No. 3, pp. 824-829, January 1991.

Cover, T. M. and J. A. Thomas, *Elements of Information Theory*, New York, Wiley, 1991.

Derin, H., and P. A. Kely, "Discrete-Index Markov Type Random Processes", *Proceedings of the IEEE*, Vol. 77. No. 10, pp. 1485-1510, October 1989.

Gubner, J. A., "Distributed Estimation and Quantization", *IEEE Transactions on Information Theory*, Vol. 39, No. 4, pp. 1456-1459, July 1993.

Hoffman, R. L. and A. K. Jain, "Segmentation and Classification of Range Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 9, No. 5, pp. 608-620, September 1987.

Hoover, "An Experimental Comparison of Range Image Segmentation Algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 7, July 1996.

Howard, P. G. and J. S. Vitter, "Arithmetic Coding for Data Compression", *Proceedings of the IEEE*, Vol. 82, No. 6, pp. 857-865, June 1994.

Rissanen, J. J., "A universal data compression system", *IEEE Transactions on Information Theory*, Vol. 29, No. 5, pp. 656-664, 1983.

Jayant, N., J. Johnston and R. Safranek, "Signal Compression based on models of human perception", *Proceedings of the IEEE*, Vol. 81. No. 10., pp. 1385-1422, 1993.

Johnson, N. and S. Jajodia, "Exploring Steganography: Seeing the Unseen", *IEEE Computer*, pp. 26-34, February 1998.

Karray, L., P. Duhamel and O. Rioul, "Image Coding with an L-Infinity Norm and Confidence Interval Criteria", *IEEE Transactions on Image Processing* Vol.7, pp. 621-631, 1998.

Kohonen, T., E. Oja, O. Simula, A. Sisa and J. Kangas, "Engineering Applications of the Self-Organizing Map", *Proceedings of the IEEE*, Vol. 84, No. 10, pp. 1358-1378, October 1996.

Linde, Y., A. Buzo and R. M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Transactions On Communications*, Vol. COM-28, pp. 84-95, 1980.

Lohmann, A. W., D. Mendelovic and G. Shabtay, "Significance of Phase and Amplitude in the Fourier Domain", *JOSA*, Vol. 14, pp. 2901–2904, 1997.

Memon, N. and X. Wu, "Recent Developments in Context-Based Predictive Techniques for Lossless Image Compression", *The Computer Journal*, Vol. 40, No. 2/3, pp. 127-136, 1997.

Olives, J. L., B. Lamiscarre, M. Gazalet, "Optimization of Electro-Optical System with an Image Quality Measure", *SPIE* Vol. 3025, 1997.

Ortega, A. and K. Ramchandran,"Rate Distortion Methods for Image and Video Compression", *IEEE Signal Processing Magazine*, pp. 23-50, November 1998.

Popat, K. and R. Pickard, "Exaggerated Consensus in Lossless Image Compression", *in Proceedings IEEE ICIP-1994*, Austin, Texas, November 1994.

Rissanen, J., T. P. Speed and B. Yu, "Density Estimation by Stochastic Copmplexity", *IEEE Transactions on Information Theory*, Vol. 38, No. 2, pp. 315-323, March 1992.

Sayood, K., *Introduction to Data Compression*, Morgan Kaufman, Publishers, Inc., 1996.

Schafer, R., G. Heising, A. Smolic, "Improving Image Compression- Is It Worth the Effort?", *EUSIPCO'2000: 10th European Signal Processing Conference*, 5-8 September, Tampere, Finland, 2000.

Shapiro, J. M., "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE Transactions on Signal Processing*, Vol. 41, No. 12, pp. 3445-3462, December 1993.

Fischer, T. R., J. D. Gibson, and B. Koo, "Estimation and noisy source coding", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 23-34, Jan. 1990.

Torres, L. and E. Delp, "New Trends in Image and Video Compression", *EUSIPCO'2000: 10th European Signal Processing Conference*, 5-8 September, Tampere, Finland, 2000.

Weinberger, M. J., J. J. Rissanen, and R. B. Arps, "Applications of Universal Context Modeling to Lossless Compression of Gray-Scale Images", *IEEE Transactions on Image Processing*, Vol. 5, No. 4, pp. 575-586, April 1996.

Weinberger, M. J., G. Seroussi and G. Sapiro, "The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS", *IEEE Transactions on Image Processing*, Vol. 9, No. 8, pp. 1309-1324, August 2000.