The Report committee for Zeina Sinno certifies that this is the Approved version of the following report:

A Closed-Form Correlation Model of Oriented Bandpass Natural Images Beyond Adjacent Responses

APPROVED BY

SUPERVISING COMMITTEE:

Alan C. Bovik, Supervisor

Joydeep Ghosh

A Closed-Form Correlation Model of Oriented Bandpass Natural Images Beyond Adjacent Responses

by

Zeina Sinno, B.E.

REPORT

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN ENGINEERING

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2015

Dedicated to my parents, Hala and Zaki.

Acknowledgments

First, I would like to express my gratitude to my super supervisor Professor Alan Bovik; for his exceptional teaching style that inspired me to pursue my research in this area, guidance, support, valuable ideas, constructive comments... and most importantly his infinite encouragement.

I would also like to acknowledge Professor Joydeep Ghosh as the second reader of this report, Professor Brian Evans for his assistance in my first year, Dr. Che Chun Su for his helping hand in this project and fellow LIVEr Todd Goodall for his valuable feedback for this report.

I must express my very profound gratitude to my family; my mom, my dad, my brother Mounir, my sister Farah and my brother-in-law Maher for their transcontinental affection, unfailing support, and big faith in me. This accomplishment would not have been possible without them.

Least but not last I would like to thank my friends for listening, offering me advice and supporting me through the entire process and fellow LIVErs for creating an agreeable lab environment.

A Closed-Form Correlation Model of Oriented Bandpass Natural Images Beyond Adjacent Responses

Zeina Sinno, M.S.E.
The University of Texas at Austin, 2015

Supervisor: Alan C. Bovik

Building natural scene statistical models is crucial for a large set of applications starting from the design of faithful image and video quality metrics to image enhancing techniques. Most predominant statistical models of natural images characterize univariate distributions of divisively normalized bandpass responses or wavelet-like decomposition of them. Previous models focusing on these bandpass natural responses offer optimized solutions to numerous problems in image processing however, these models have not focused on finding a closed-form quantative model capturing the bivariate natural statistics. Towards the efforts for filling this gap, Su *et. al.* [1] recently modeled spatially horizontally neighboring bandpass image responses on multiple scales; however, the latter work did not cover the response of distant bandpass image responses with various spatial orientations. This work builds on Su. et al 's model and extends the closed-form correlation model to cover distant bandpass image responses, up to a distance of 10 pixels; with multi-

ple spatial orientations, encompassing all the discrete spatial angles for the lastly-mentioned distances on multiple scales.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The objective of this chapter is to emphasize the relationship between Natural Scene Statistics (NSS) and the Human Visual System (HVS) and to present relevant work that exploited a part of this interlacing relationship across neighboring bandpass image responses.

1.1 An Interlaced Relationship between NSS and the HVS

Understanding the functionalities of the different components of the Human Visual System (HVS) and modeling the Natural Scene Statistics (NSS) of the perceived images and videos are the building blocks of many reliable image and video processing algorithms. These algorithms span a wide range of applications starting from image/video quality assessment metrics [2–4] to state of the art image enhancement techniques such as: image denoising [5], image defocus [6], and image super-resolution [7]. Tremendous effort has been made in an attempt to fathom these two interlacing components and to uncover the links between them.

The HVS is able of processing highly demanding tasks in a relatively short amount of time such as fully visually recognizing the objects of an image within 40ms [8]. Processing this enormous amount of information is enabled by the parallel and concurrent processing capability found in the primary visual cortex [9]. The different cells of the primary visual cortex have a very well defined map in the spatial information in the vision [10]. Thus, the connection between a given location in primary visual cortex and in the subjective visual field is very precise. This area can be modeled using a Gabor filter bank, capable of decorrelating the received signal over multiple scales and orientations [11]. The resulting signal after decorrelation follows a well behaved Gaussian model distribution.

Ruderman [12], showed that a local mean subtraction operation followed by division by local variance, known as divisive normalization reproduces this decorrelation effect. Therefore, the resulting signal after these two operations is Gaussian, for which the signal coefficients follow a robust statistical model [12]. This concept was deeply exploited in the establishment of first order statistical models; and in attempts to understand the models across scales and orientations which is not completely uncovered. A summary of these attempts will be presented next.

1.2 Relevant Observations and Models

Before diving into relevant models to our problem, some notable observations from the literature will be briefly presented.

1.2.1 Relevant Observations

Previous work led to a few prominent observations concerning wavelet coefficients computed from natural images. First, Simoncelli *et. al.* [13] observed that the coefficients of orthonormal wavelet decomposition of natural images are decorrelated but not independent. Secondly, Liu *et. al.* [14] noted the presence of inter and intra-scale dependencies between wavelet coefficients. Thirdly, Sendur. *et. al.* [15] used a circularly symmetric bivariate distribution to model the dependencies between image wavelet coefficients and their parents (at coarser scale locations).

1.2.2 Relevant Models

Inspired by the work accomplished by Geman *et. al.* [16] , which makes use of a Markov random field to implement a image restoration technique at low signal-to-noise ratios, Portilla *et. al.* [17, 18] targeted the problem of texture modeling of the images. The latter method uses a set of parametric constraints on pairs of complex wavelet coefficients at adjacent spatial locations, orientations and scales in addition to incorporating Markov Random Field and non-Gaussian Statistics. The major issue with this method is in the choice of statistical constraints, obtained by applying a form of reverse-engineering

of the early HVS. The selection process of the parameters is achieved by observing failures to synthesize particular types of texture. As a result, this selection process does not guarantee the optimality of the solution. Given the under-determined nature of the solution, additional constraints are required. Subsequently, this method does not guarantee success in modeling the texture.

On the other hand, Po *et. al.* [19] modeled natural images using a hidden Markov tree, Gaussian mixture model, and a two dimensional contourlet to capture interlocation, interscale, and interdirection dependencies. The contourlet transform is an extension of the wavelet transform using multiscale and directional filter banks. This model is not limited to texture retrieval but it could be used for denoising applications and offers a valuable tool in image processing.

Mumford *et. al.* [20] proposed an infinitely divisible model of generic image statistics. This model assumes that the environment may be subdivided into objects cast against an ergodic field while also containing regions with very little information (e.g: blue sky). This model seemed to fit a small portion of the data only; but is short of capturing all the basic qualitative properties of many images.

Among the previously presented models characterizing the bivariate behavior, none offers a closed form model. The presence of such a model however

would provide an optimal solution to various image processing applications. Su *et. al.* [1] proposed a closed form correlation model of oriented bandpass natural images for horizontally adjacent pixels across sub-band bandpass orientations across multiple scales. This model is valid for all natural images, correlating highly with the HVS, proving to be very useful in a wide variety of image processing applications including stereoscopic image quality metrics with depth information [21]. Su. *et. al.* designed a 3D image quality predictor based in part on the new model. This predictor outperforms state-of-the-art full- and no-reference 3D IQA algorithms on both symmetrically and asymmetrically distorted stereoscopic image pairs. An additional application of this model is a depth estimator based on single luminance images [22]. Consequently, extending the above mentioned model will be very propitious.

This work will provide an extension to Su *et. al.*'s model. The latter work did not cover the response of distant bandpass image responses with various spatial orientations. This work will fill in this gap by covering distant bandpass image responses, up to a distance of 10 pixels; with multiple spatial orientations, targeting all the discrete spatial angles for the lastly-mentioned distances on multiple scales. In chapter 2, we will present the proposed model and how it relates to the human visual system. In chapter 3, we will present the extended model and validate it. In chapter 4, we will compare our model to the truncated fourier series expansion. In chapter 5, we will open the door for future extensions of our model.

Chapter 2

The Steps Followed

In this chapter we present the steps to obtain our model out of the input images. The steps consist of a steerable pyramid decomposition accompanied by a divisive normalization for preprocessing followed by modeling the resulting coefficients with a bivariate joint distribution model. The motivation behind each step will be presented with similarity aspects to the HVS. In order to build our model, the LIVE IQA database [3] was used.

2.1 Color Space Transformation and Steerable Pyramid Decomposition Steps

First, the obtained image is transformed to the CIELAB color space. This color space mirrors that of the human color perception [23]. The lightness component, L, of this color space matches closely the human perception of lightness. Only the L component will be fed into the next stages. Second, the steerable pyramid is a linear multi-scale, multi-orientation image decomposition. It has been deployed in a wide variety of applications in image processing as it overcomes some limitations of orthogonal separable wavelet decompositions. Its properties include independent representations of scale and orientation, translation and rotation invariance and most importantly similarity to

the bandpass responses of simple cells in the primary visual cortex [24–26].

Simple cells compute a linearly weighted sum of the input over space and time (usually a Gabor-like function). This simple cells response is then normalized by the responses of neighboring neurons and passed through a pointwise nonlinearity. They are characterized by their oriented receptive field and their responses are space-time separable. The steerable pyramid obtains decorrelated representations over scale and orientation which resembles these seperable responses found in the HVS. Specifically, the steerable pyramid models the orientation selective behavior of individual simple cell responses found in the primary visual cortex. These simple cells are combined to cover all orientations and receptive field sizes [27]. Lastly, receptive field size is modeled by the multiple scales in the steerable pyramid.

Note that in our case, we used a steerable pyramid of 4 scales where scale 1 represents the finest and scale 4 represents the coarsest scale and fifteen frequency tuning orientations of values $[0, \pi/15, 2\pi/15, ..., \pi]$.

2.2 Divisive Normalization Step

Divisive normalization on all the wavelet coefficients is obtained from the steerable pyramid decomposition output. Applying this step is analogous to the normalization taking place in the area V1 of the visual cortex [28]. The divisive normalization is defined as:

$$u(x_i, y_i) = \frac{w(x_i, y_i)}{\sqrt{s + \boldsymbol{w_g^T w_g}}} = \frac{w(x_i, y_i)}{\sqrt{s + \sum_j g(x_j, y_j) w(x_j, y_j)^2}} \tag{2.1}$$

where (x_i, y_i) are the spatial coordinates, w are the wavelet coefficients, u represent the coefficients obtained after the divisive normalization, s is the semi saturation constant (in our case it is set to 0.0001). The weighted sum is computed over a spatial neighborhood of pixels in the same sub-band index by j (in our case we considered a window of dimensions 3×3 hence $j = 9$). The Gaussian weighting function, $g(x_i, y_i)$, is circularly symetric and unit volume.

2.3 Modeling the Bivariate Joint Distribution

The bivariate joint distribution targets pixels with a gap in the distance between 1 and 10, and covering all possible discrete angles in an image in the divisive normalization result, or u . This is the extension to Su *et. al.*'s work [1], where the bivariate joint distribution targeted adjacent pixels horizontally and vertically only.

Inspired by the fact that the univariate generalized Gaussian distribution was used to model univariate natural scene statistics [29, 30], the multivariate generalized Gaussian distribution (MGGD) will be used to model the bivariate joint histogram of the two target pixels. This choice is also justified by the fact that MGGD was an accurate tool for modeling multi-dimensional image histograms [31]. The probability density function of the MGGD used is:

$$p(\boldsymbol{x}; \boldsymbol{M}, \alpha, \beta) = \frac{1}{|\boldsymbol{M}|^{\frac{1}{2}}} g_{\alpha,\beta}(\boldsymbol{x}^{\boldsymbol{T}} \boldsymbol{M}^{-\mathbf{1}} \boldsymbol{x}) \tag{2.2}$$

where $\boldsymbol{x} \in \mathbb{R}^N$, $\boldsymbol{M}$ is an $N \times N$ scatter matrix α and β are the scale and shape parameters respectively and $g_{\alpha,\beta}(.)$ is the density generator defined by:

$$g_{\alpha,\beta}(y) = \frac{\beta \Gamma(\frac{N}{2})}{(2^{\frac{1}{\beta}} \pi \alpha)^{\frac{N}{2}} \Gamma(\frac{N}{2\beta})} e^{-\frac{1}{2}(\frac{y}{\alpha})^{\beta}} \tag{2.3}$$

where $y \in \mathbb{R}^+$. We note that if $\beta = 0.5$ then equation (2.2) yields to the multivariate Laplacian distribution and when $\beta = 1$ then equation (2.2) corresponds to the multivariate Gaussian distribution.

The bivariate empirical histograms of the target sub-band coefficients in natural images is modeled using a bivariate generalized Gaussian distribution (BGGD), by setting $N = 2$. Similarly to Su *et. al.* [1] the parameters of the BGDD were estimated using Maximum Likelihood Estimation.

Chapter 3

The Generalized Model

Before describing the extended model and validating it, we will present some observations that directed Su *et. al.* towards the closed form model for the horizontally adjacent bandpass responses.

3.1 Su et. al's Model

The BGGD model allows us to observe that the shape and height of the bivariate distributions vary with the tuning orientation of the sub-band responses. Su *et. al.* [1] observed the case of horizontally spatially adjacent responses. When the spatial relationship between bandpass samples matches the sub-band tuning orientation, the joint distribution becomes peaky and extremely elliptical, implying that the bandpass responses are highly correlated, as depicted in Figure 3.1. Conversely, when the spatial relationship and the sub-band tuning orientation are orthogonal, the joint distribution becomes nearly a circular Gaussian indicating that we have uncorrelated sub-band responses.

In order to more deeply understand the effect of the relative orientation, Su *et. al.* [1] modeled the correlation coefficients as a function of the

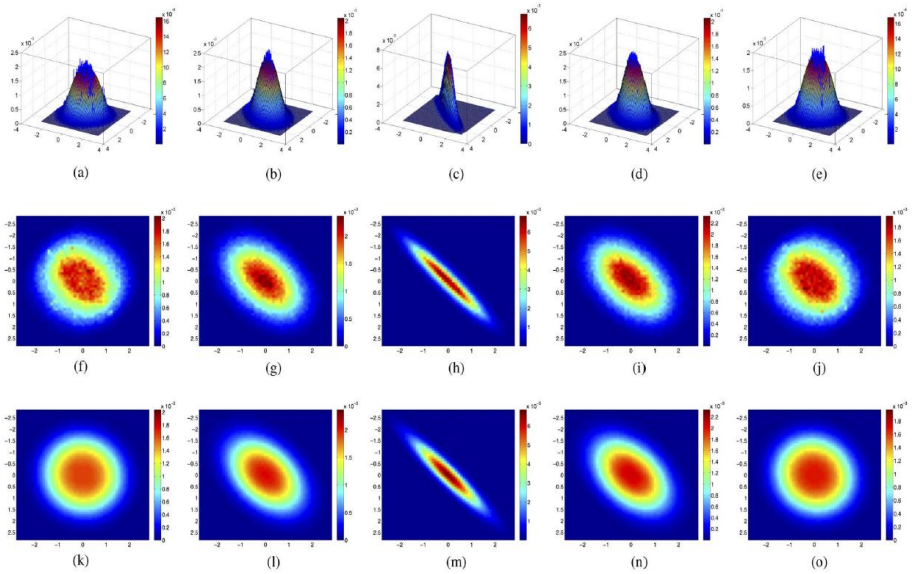

Figure 3.1: Joint histograms of horizontally adjacent bandpass coefficients from a pristine image and the corresponding BGGD fits at the finest scale with different orientations. From left column to right column: 0 (rad), $\frac{\pi}{4}$, $\frac{\pi}{2}$, $\frac{3\pi}{4}$, and $\frac{11\pi}{12}$. Top row: 3D illustration of bivariate histogram and BGGD fit, middle row: 2D iso-probability contour plot of histogram, and bottom row: 2D iso-probability contour plot of BGGD fit (a) 0(rad) (b) $\frac{\pi}{4}$ (c) $\frac{\pi}{2}$(d) $\frac{3\pi}{4}$(e) $\frac{11\pi}{12}$ (f) 0(rad) (g) $\frac{\pi}{4}$ (h) $\frac{\pi}{2}$ (i) $\frac{3\pi}{4}$(j) $\frac{11\pi}{12}$(k) 0(rad) (l) $\frac{\pi}{4}$(m) $\frac{\pi}{2}$(n) $\frac{3\pi}{4}$(o) $\frac{11\pi}{12}$, source [1]

relative orientation. The authors observed a periodic behavior with a period π. The author was able to obtain a significant fit of the correlation coefficients for horizontally adjacent bandpass responses using:

$$\rho = Acos(\theta_2 - \theta_1)^{2\gamma} + c \tag{3.1}$$

where θ_1 and θ_2 represent the sub-band and spatial tuning orientations,

A represents the amplitude, γ is the exponent and c is the offset. The relative orientation is represented by $\theta_2 - \theta_1$. Consequently, the authors were able to obtain a closed form model by fitting the correlation coefficients between horizontally adjacent bandpass responses as a function of the sub-band tuning orientations using the images of the LIVE IQA database. The exponentiated sine model for horizontally adjacent bandpass responses of Su et. al is presented in Figure 3.2.

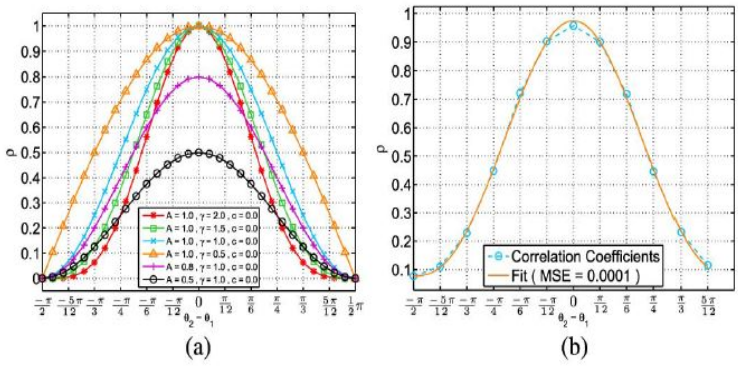

Figure 3.2: The exponentiated cosine function and its fit to correlation coefficients as a function of relative orientation (a) Exponentiated cosine function (b) Fit to correlation coefficients., source [1]

3.2 Extending the Model

Our goal is to be able to generalize Su *et. al.* 's model across more points in the space. We redefined the relative distance to be equal to $\theta_1 + \theta_2$. θ_2 is defined to be equal to $\arctan(\frac{\delta_y}{\delta_x})$ where δ_x and δ_y represent the relative row and column differences between the correlated pixels in the lightness image

obtained after the divisive normalization. The reason behind redefining the relative orientation is that the new convention allowed us to align the maximum correlation coefficients at 0 and multiples of π. On the other hand, if we follow the convention of the previous model, the maximum correlation will occur at $2(\theta_2 - \theta_1)$, hence if we consider points not horizontally or vertically adjacent the peak will not occur at 0 or multiples of π .

To examine the effect of the distance on the observed correlation, we considered different points with the same digital distance. To do so, we used the midpoint circle algorithm [32] to obtain the digital circles of radius varying between 1 to 10, and considered all the points covered by these circles. These different points are shows in Figure 3.3. Our model considers 158 points where θ_2 lies in $[-\pi/2, \pi/2)$ in each scale. Since we have four different scales due to the use of the steerable pyramid, we obtained a total of 632 cases. Since the model is π periodic in the space in terms of θ_2, it accounts for a total of 1264 points.

To compute the correlation model, we consider an origin pixel at the top leftmost corner of the image obtained after the divisive normalization, then we move a sliding window across all the image and store the values of these two pixels in two vector. Next, we correlate the two vectors using the Pearson Correlation. At a fixed (δ_x, δ_y) corner to the window, the spatial orientation θ_2 is fixed. Hence we repeat the process across all the 15 sub-band orientations $0, \frac{\pi}{15}, \frac{2\pi}{15}, ..., , \frac{14\pi}{15}$ rad and across the four scales obtained from the steerable pyramid. This process is also repeated across all the images of the LIVE IQA

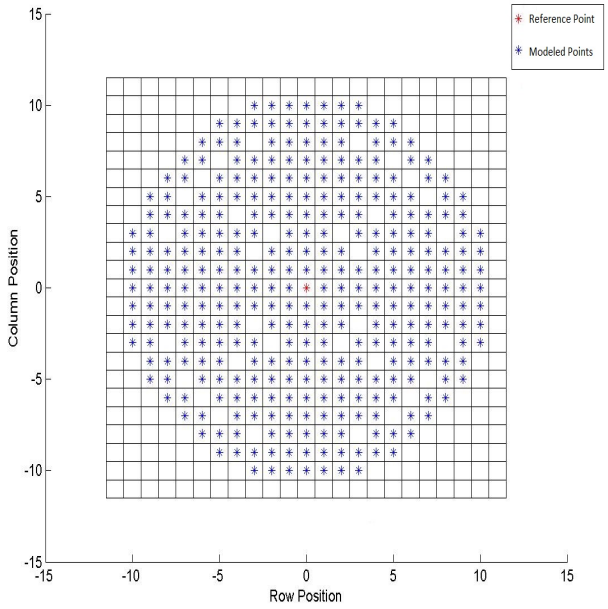

Figure 3.3: The covered points of the extended model

database. For fitting our model, we considered the average correlation across all the images of this database.

The first observation that could be drawn by looking at the average correlation is that the maximal correlation is obtained when $\theta_1 + \theta_2$ is equal to 0 and at multiples of π. This observation is always true when the average correlation is high enough. The maximal correlation drops as the relative distance between the origin and target increases. This could be seen in Figure 3.4 where we present the maximal correlation as a function of δ_x and δ_y for the case of scale 1 (the finest scale obtained using the steerable pyramid decomposition). Similar trends are observed across the other more coarse scales.

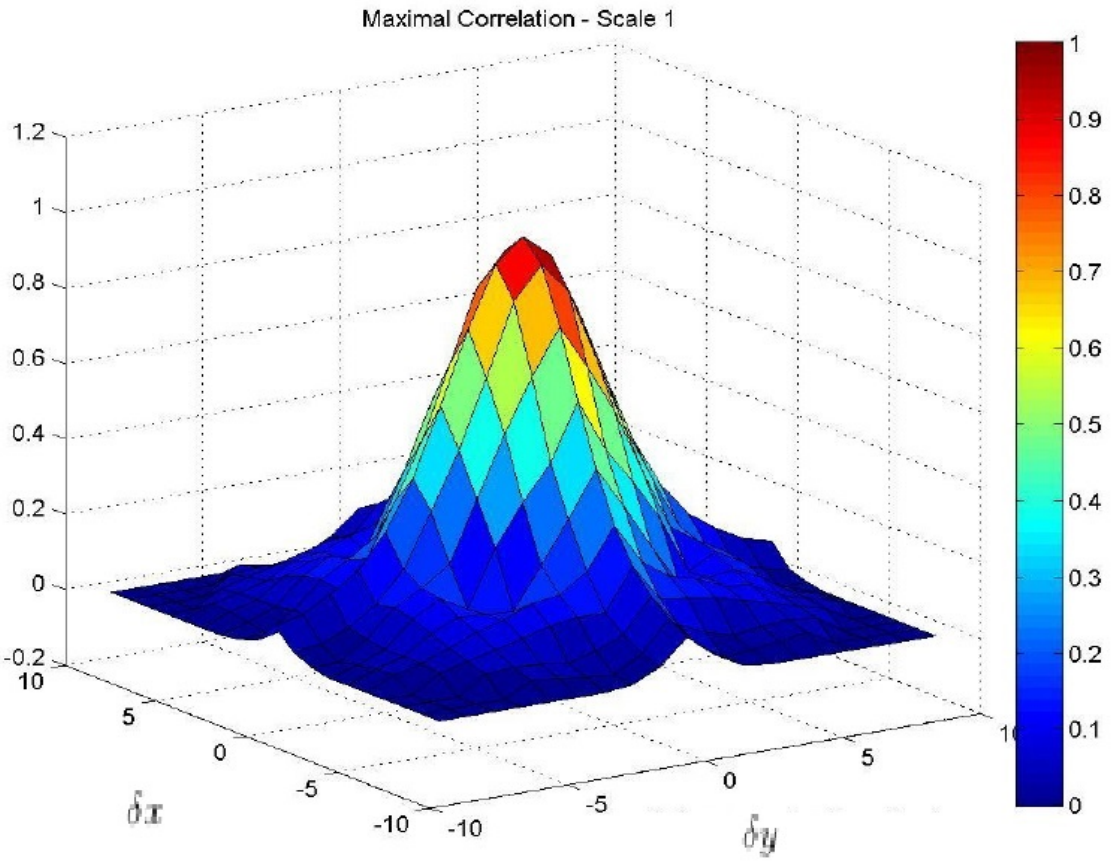

Figure 3.4: Maximal correlation as function of δ_x and δ_y

To study the impact of θ_2, we fixed the distance and observed the obtained correlations. The spatial angle $\theta_2 = \pi/2$ rad seems to achieve the highest maximal correlation followed by $\theta_2 = 0$ rad. The reason behind this could be explained by the fact that the world contains many vertical and horizontal structures. Hence these two orientations contain most of the energy

[33]. The maximal correlation drops at other values of θ_2.

We observed that at small distances the correlation has a Gaussian shape. As the distance increases, we start to observe lobes on the right and left sides of the origin. As the distance increases further, we remark that the lobes increase in magnitude while the maximal correlations continues to drop. Although for high distances the correlation was extremely low (in the order of less than 0.05), our model was still able to perform well.

In order to fit our obtained correlations and account for the different observed lobes we considered the following new model:

$$\rho' = A_1 cos(\frac{2(\theta_1+\theta_2)}{1}) \quad + \quad A_2 cos(\frac{2(\theta_1+\theta_2)}{2}) + A_3 cos(\frac{2(\theta_1+\theta_2)}{3}) \quad + \quad A_4 cos(\frac{2(\theta_1+\theta_2)}{4}) \quad + c \quad (3.2)$$

One advantage of this model compared to Su. et al's model is that we were able to get rid of the exponent term. Some of the correlations and fitted models are presented in Figures 3.5 to Figure 3.9. All these correlations are at a fixed $\theta_2 = \pi/2$ with a varying relative distance. Scale 1 represents the finest scale and scale 4 the coarsest. We present the cases when the relative distance is equal to 1, 3, 5, 7 and 9. As we can observe that for a given spatial distance, the same trend holds across the different scales.

For the displayed cases in Figure 3.5 to Figure 3.9, the obtained parameters using a non-constrained non-linear regression, are shown in Table 3.1.

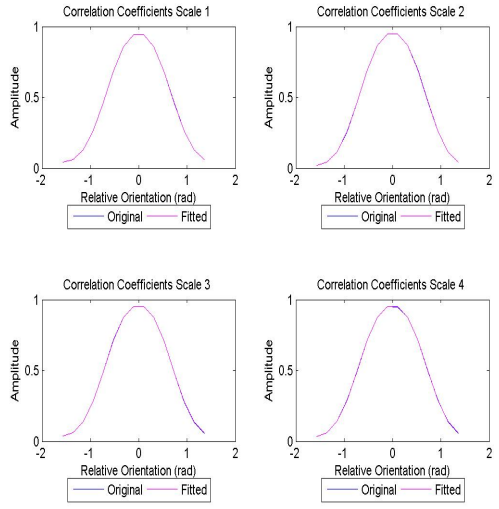

Figure 3.5: Correlation coefficients and fitted model for $\delta_x = 0, \delta_y = 1$

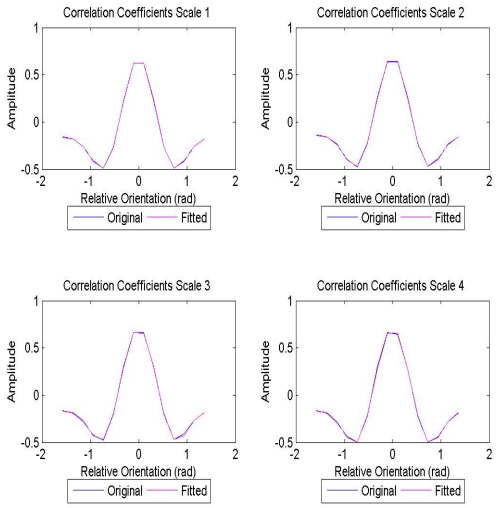

Figure 3.6: Correlation coefficients and fitted model for $\delta_x = 0, \delta_y = 3$

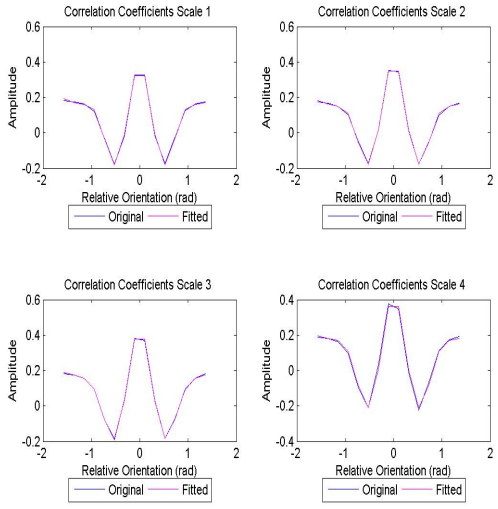

Figure 3.7: Correlation coefficients and fitted model for $\delta_x = 0, \delta_y = 5$

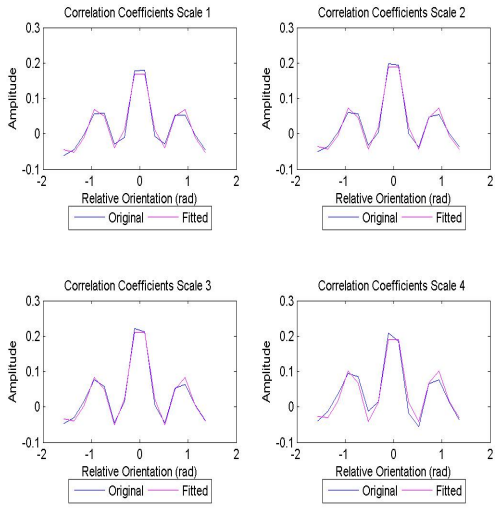

Figure 3.8: Correlation coefficients and fitted model for $\delta_x = 0, \delta_y = 7$

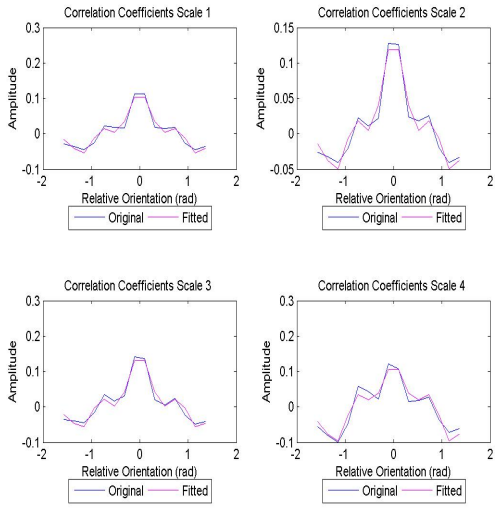

Figure 3.9: Correlation coefficients and fitted model for $\delta_x = 0, \delta_y = 9$

Distance	Scale	δ_x	δ_y	A1	A2	A3	A4	c
1	1	0	1	0.470409	0.043307	-0.01347	-0.0023	0.456836
1	2	0	1	0.484666	0.037684	-0.01676	-0.00317	0.453898
1	3	0	1	0.47898	0.031725	-0.01758	-0.00447	0.470419
1	4	0	1	0.477693	0.027308	-0.01586	-0.00352	0.47322
3	1	0	3	0.305609	0.377222	0.117863	-0.0035	-0.11247
3	2	0	3	0.310583	0.375779	0.111905	-0.01086	-0.08856
3	3	0	3	0.346922	0.378423	0.099425	-0.01558	-0.08922
3	4	0	3	0.34467	0.389593	0.099488	-0.01561	-0.10387
5	1	0	5	-0.04864	0.133167	0.144128	0.065612	0.088165
5	2	0	5	-0.03018	0.148893	0.141853	0.059077	0.085984
5	3	0	5	-0.02539	0.169892	0.149169	0.053445	0.089213
5	4	0	5	-0.04393	0.169591	0.160285	0.056721	0.083716
7	1	0	7	0.029407	0.048917	0.012009	0.056447	0.010242
7	2	0	7	0.034771	0.051545	0.021183	0.059737	0.008786
7	3	0	7	0.041484	0.053756	0.025799	0.069	0.007483
7	4	0	7	0.047003	0.042343	0.016912	0.070032	-0.00752
9	1	0	9	0.009523	0.057285	0.01457	0.008165	0.012609
9	2	0	9	0.014562	0.059066	0.017154	0.012592	0.015227
9	3	0	9	0.015124	0.067409	0.018245	0.015866	0.010489
9	4	0	9	0.002443	0.080375	0.003974	-0.00058	0.014513

Table 3.1: The parameters obtained for some samples points

3.3 General trend of the parameters in the Generalized Model

We studied how the coefficients vary as a function of the distance for a given spatial orientation. We studied in total 12 cases of θ_2: -1.57 rad, 0 rad,-1.24 rad, 1.24 rad, -1.10 rad, 1.10 rad, -0.785 rad, 0.785 rad, -0.463 rad, 0463 rad, -0.321 rad and 0.321 rad.

It was interesting to observe that the parameters on the different scales behave similarly for the 12 cases. For small distances, the first cosine in the model contributes the most. As the distance increases, its contribution decreases, and the other cosines with higher frequency start to take over.

Also we observed some form of symmetry in the behavior of the parameters for θ_2 around 0 rad. That is, we remarked a similar behavior of the parameters at -0.321 rad and 0.321 rad, -0.463 rad and 0.463 rad... Also we observed a symmetric behavior between 1.57 rad and 0 rad. Figures 3.10 and 3.11 provide an example of the latter case.

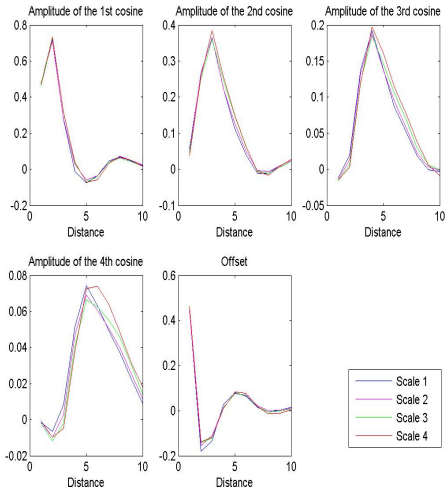

Figure 3.10: General trend in the parameters of the model as a function of the distance for $\theta_2 = 0$ rad

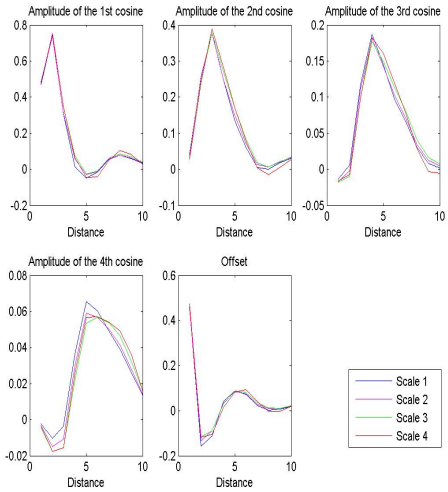

Figure 3.11: General trend in the parameters of the model as a function of the distance for $\theta_2 = 1.57$ rad

3.4 Validation of the Generalized Model

We computed the Mean Squared Error (MSE) and χ^2 test across the images of the LIVE IQA database [3] to validate our model. The χ^2 test is computed as:

$$\chi^2 = \sum_{i=1}^{N} \sum_{j=1}^{S} \frac{(\rho_{ij} - \rho'_j)^2}{\rho'_j} \tag{3.3}$$

where $\{\rho'_j\} = \boldsymbol{\rho'} \in \mathbb{R}^D$is the model, $\{\rho_{ij}\} = \boldsymbol{\rho_i} \in \mathbb{R}^D$ are the correlation coefficients as a function of the $i-th$ pristine image, and N is the number of pristine images. $N = 29$ for the LIVE IQA database.The MSE and χ^2 test of the samples previously presented are shown in Table 3.2.

Additionally, we observed the values of the parameters estimated individually on each each image of the LIVE IQA database [3]. We visualized these parameters using boxplots (Figure 3.12 to Figure 3.16). It is notable to observe how the parameters obtained by estimating the average correlation closely match the parameters estimated on the images individually.

Distance	Scale	δ_x	δ_y	MSE	χ^2 for the LIVE IQA Database
1	1	0	1	3.94E-07	3.761252571
1	2	0	1	1.44E-06	5.646435045
1	3	0	1	2.69E-06	8.925839408
1	4	0	1	4.46E-06	15.00274839
3	1	0	3	2.32E-05	-0.980905718
3	2	0	3	5.01E-05	-1.690234587
3	3	0	3	9.26E-05	-5.078580082
3	4	0	3	9.84E-05	-11.37689956
5	1	0	5	4.47E-05	-98.50399357
5	2	0	5	2.09E-05	57.87348179
5	3	0	5	1.98E-05	38.62702038
5	4	0	5	0.000155	345.8776777
7	1	0	7	0.000149	41.37185147
7	2	0	7	0.000116	24.19320404
7	3	0	7	9.34E-05	62.36532159
7	4	0	7	0.00028	106.7090611
9	1	0	9	0.000124	21.39735822
9	2	0	9	0.00012	27.68330054
9	3	0	9	0.000138	40.69777252
9	4	0	9	0.000261	64.6202476

Table 3.2: The MSE Error and χ^2 of the LIVE IQA database for some sample points

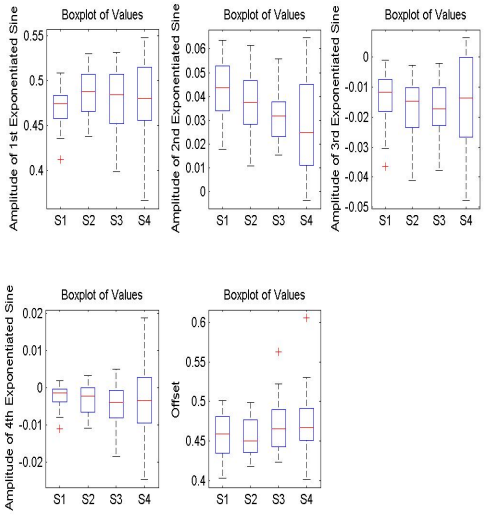

Figure 3.12: Boxplot of the parameters for $\delta_x = 0, \delta_y = 1$

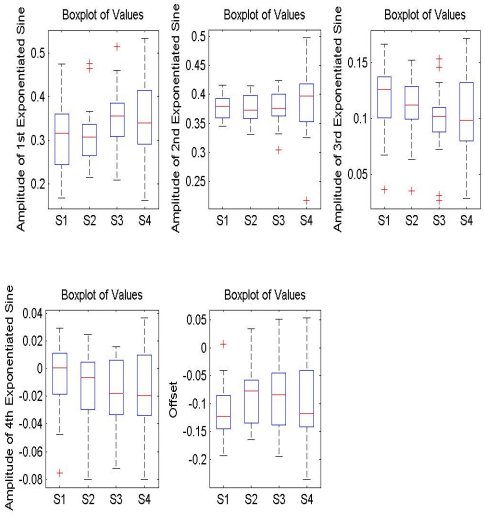

Figure 3.13: Boxplot of the parameters for $\delta_x = 0, \delta_y = 3$

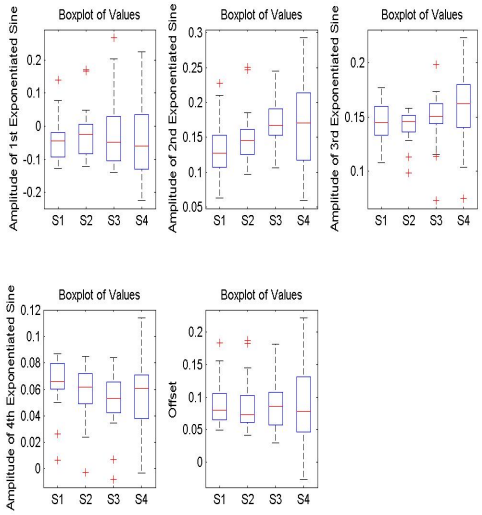

Figure 3.14: Boxplot of the parameters for $\delta_x = 0, \delta_y = 5$

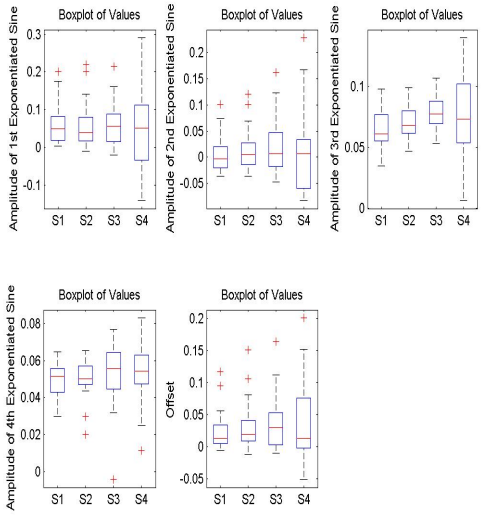

Figure 3.15: Boxplot of the parameters for $\delta_x = 0, \delta_y = 7$

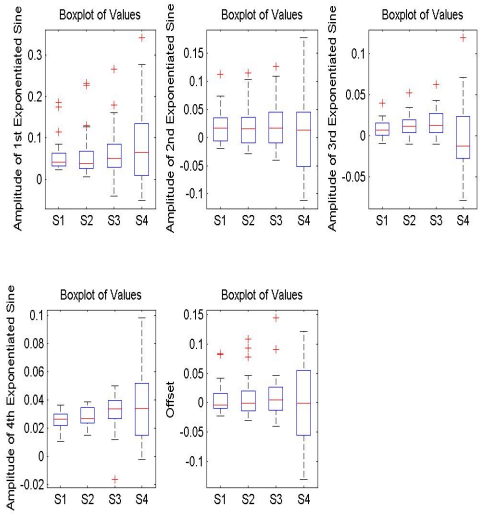

Figure 3.16: Boxplot of the parameters for $\delta_x = 0, \delta_y = 9$

Chapter 4

Comparison to the Short Fourier Series Expansion

In this section we compare our model against the truncated fourier series expansion.

4.1 Motivation behind the Fourier Series Expansion

The motivation behind comparing our model against the fourier series expansion is that any periodic signal $x(t)$ with period T_0 and fundamental frequency $\omega_0 = 2\pi/T_0$ can be represented by a sum of scaled sines and cosines at multiples of the fundamental frequency. The series can also be expressed as sums of scaled complex exponentials at multiples of the fundamental frequency [34]. That is:

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n cos(n\omega_0 t) + b_n cos(n\omega_0 t) \tag{4.1}$$

We computed the truncated Fourier Series of order 4 using least-square fits. This method makes use of the fact that Fourier coefficients give the best least-squares fit when a function is expanded in a set of orthonormal

functions[35]. Sampling the function:

$$x(tn) = x_n, tn = \frac{nT}{N}, n = 1, 2, ...N \tag{4.2}$$

We obtain:

$$x_n = a_0 + \sum_{m=1}^{M} a_m cos(\frac{2\pi mn}{N}) + b_m sin(\frac{2\pi mn}{N}) \tag{4.3}$$

As a result, the problem can be reformulated as; given a column vector $x = [x_1 x_2 ... x_N]^T$, find the set of coefficients $coef = [a_0 a_1 ... a_n b_1 ... b_n]^T$ that fits the expansion above the best. Define:

$$A = \begin{bmatrix} 1 & cos(w_0.t_1) & ... & cos(nw_0.t_1) & sin(w_0.t_1) & ... & sin(nw_0.t_1) \\ 1 & cos(w_0.t_2) & ... & cos(nw_0.t_2) & sin(w_0.t_2) & ... & sin(nw_0.t_2) \\ ... & ... & ... & ... & ... & ... & \\ 1 & cos(w_0.t_N) & ... & cos(nw_0.t_N) & sin(w_0.t_N) & ... & sin(nw_0.t_N) \end{bmatrix} \tag{4.4}$$

And use ordinary least-squares to find the coefficient matrix $coef$ such that:

$$x = A.coef \tag{4.5}$$

implying:

$$coef = A^{-1}.x \tag{4.6}$$

4.2 Results of the Short Fourier Series Expansion

For a fair comparison with the model presented in the previous chapter, we computed the fourier series expansion truncated at $m = 4$:

$$x_n = a_0 + \sum_{m=1}^{4} a_m cos(\frac{2\pi mn}{N}) + b_m sin(\frac{2\pi mn}{N}) \tag{4.7}$$

The obtained plots are presented in Figures 4.1 to 4.5. The resulting coefficients, MSE and χ^2 test can be observed in Table 4.1 to Table 4.3.

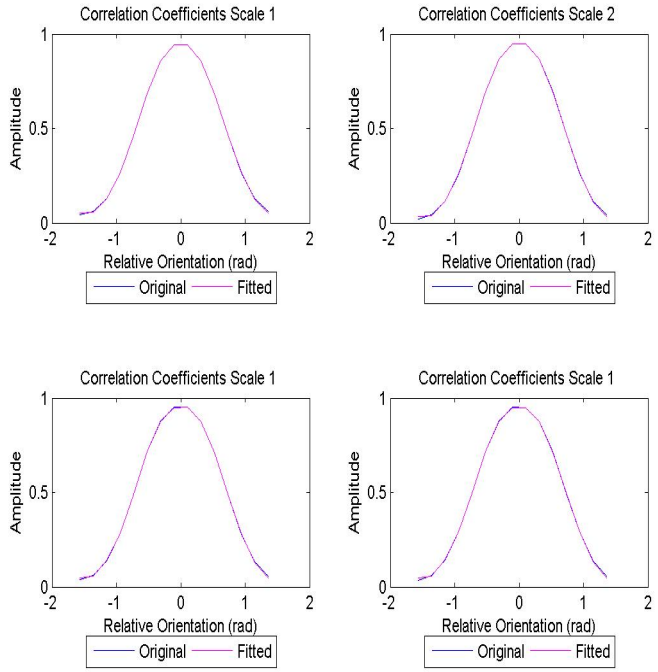

Figure 4.1: Original plot and fitted model using truncated fourier series for $\delta_x = 0, \delta_y = 1$

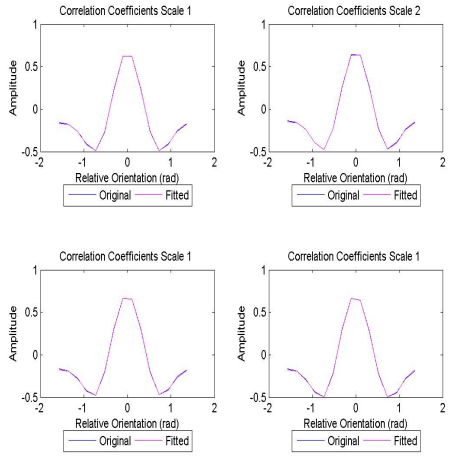

Figure 4.2: Original plot and fitted model using truncated fourier series for $\delta_x = 0, \delta_y = 3$

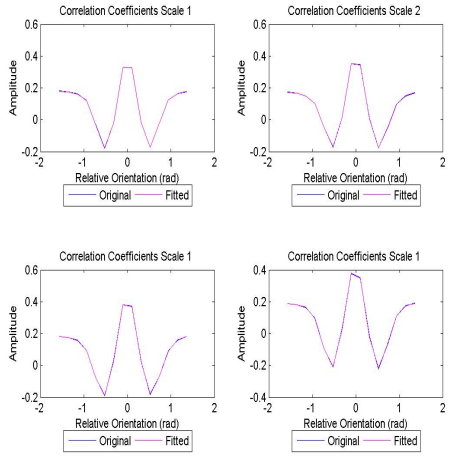

Figure 4.3: Original plot and fitted model using truncated fourier series for $\delta_x = 0, \delta_y = 5$

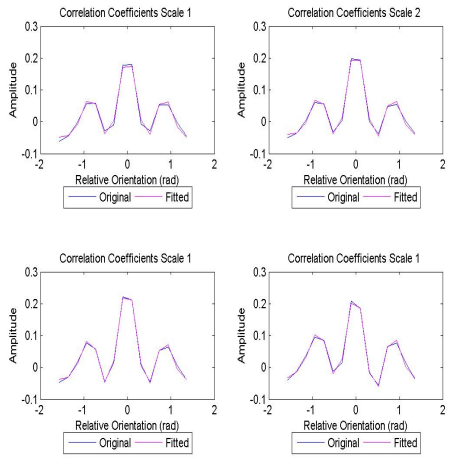

Figure 4.4: Original plot and fitted model using truncated fourier series for $\delta_x = 0, \delta_y = 7$

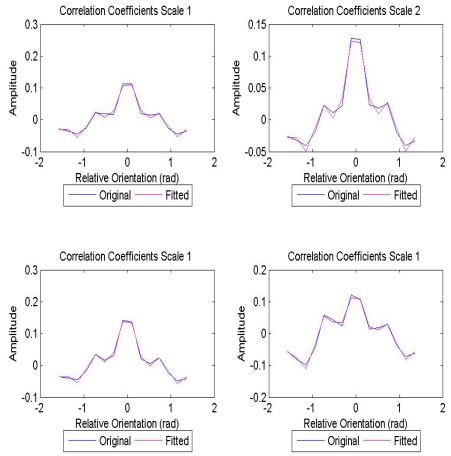

Figure 4.5: Original plot and fitted model using truncated fourier series for $\delta_x = 0, \delta_y = 9$

Distance	Scale	δ_x	δ_y	a0	a1	a2	a3	a4
1	1	0	1	0.485679	0.450819	0.015282	-0.00593	-0.00303
1	2	0	1	0.483929	0.463436	0.00935	-0.00725	-0.00363
1	3	0	1	0.500467	0.457618	0.004113	-0.00765	-0.00486
1	4	0	1	0.5035	0.455344	0.000706	-0.00519	-0.00475
3	1	0	3	-0.10832	0.337741	0.33756	0.05891	-0.0032
3	2	0	3	-0.08406	0.34229	0.336292	0.054526	-0.00697
3	3	0	3	-0.0829	0.377731	0.33698	0.044175	-0.00997
3	4	0	3	-0.0984	0.378793	0.347147	0.044887	-0.0056
5	1	0	5	0.08168	-0.03486	0.140418	0.10823	0.030228
5	2	0	5	0.079919	-0.01417	0.154338	0.104424	0.026893
5	3	0	5	0.082668	-0.00769	0.173369	0.108296	0.025399
5	4	0	5	0.076321	-0.02602	0.181452	0.122393	0.025846
7	1	0	7	0.016187	0.011575	-0.04947	0.018159	0.015918
7	2	0	7	0.015723	0.004995	-0.04432	0.021445	0.015858
7	3	0	7	0.017437	0.00242	-0.04868	0.02481	0.017102
7	4	0	7	0.025644	0.000524	-0.05819	0.030337	0.020398
9	1	0	9	0.008712	0.058906	0.012949	0.009786	0.017019
9	2	0	9	0.013671	0.060848	0.015371	0.014375	0.018715
9	3	0	9	0.014152	0.069352	0.016302	0.01781	0.020406
9	4	0	9	0.001281	0.082698	0.00165	0.00174	0.0244

Table 4.1: Coefficients of the truncated fourier series expansion part 1

Distance	Scale	δ_x	δ_y	b1	b2	b3	b4
1	1	0	1	0.103337	0.0084627	-0.00628	-0.00166
1	2	0	1	0.106717	0.0053964	-0.00858	-0.00221
1	3	0	1	0.102436	0.0038912	-0.00935	-0.00256
1	4	0	1	0.102763	0.0019906	-0.00855	-0.00263
3	1	0	3	0.07862	0.1635395	0.049134	-0.00749
3	2	0	3	0.080063	0.1609028	0.043031	-0.01339
3	3	0	3	0.089716	0.1555794	0.036318	-0.01573
3	4	0	3	0.084351	0.1592687	0.033488	-0.01798
5	1	0	5	-0.00604	0.0681758	0.087381	0.036753
5	2	0	5	-0.00521	0.0722994	0.084045	0.03154
5	3	0	5	-0.00122	0.0836137	0.084577	0.02489
5	4	0	5	-0.00588	0.0731232	0.084762	0.025887
7	1	0	7	-0.0192	-0.017494	0.025865	0.042851
7	2	0	7	-0.01712	-0.015946	0.027585	0.041445
7	3	0	7	-0.02238	-0.022718	0.026215	0.043071
7	4	0	7	-0.0136	-0.017117	0.032936	0.049285
9	1	0	9	0.012609	0.0060478	0.009385	0.021825
9	2	0	9	0.015227	0.0081978	0.010616	0.021787
9	3	0	9	0.010489	0.0056245	0.013824	0.025046
9	4	0	9	0.014513	-0.011261	0.010888	0.029166

Table 4.2: Coefficients of the truncated fourier series expansion part 2

Distance	Scale	δ_x	δ_y	MSE	χ^2 for the LIVE IQA
1	1	0	1	3.94E-07	3.786108
1	2	0	1	1.44E-06	5.302242
1	3	0	1	2.69E-06	8.599795
1	4	0	1	4.46E-06	14.30204
3	1	0	3	2.32E-05	-1.05237
3	2	0	3	5.01E-05	-1.74872
3	3	0	3	9.26E-05	-5.09735
3	4	0	3	9.84E-05	-11.4321
5	1	0	5	4.47E-05	-54.5984
5	2	0	5	2.09E-05	96.3525
5	3	0	5	1.98E-05	44.03968
5	4	0	5	0.000155	0.239492
7	1	0	7	0.000149	150.4983
7	2	0	7	0.000116	49.64248
7	3	0	7	9.34E-05	49.57876
7	4	0	7	0.00028	109.6757
9	1	0	9	0.000124	21.72763
9	2	0	9	0.00012	33.19539
9	3	0	9	0.000138	-29.6458
9	4	0	9	0.000261	87.47455

Table 4.3: MSE and χ^2 test results for the truncated fourier series expansion

4.3 The Generalized Model versus The Truncated Fourier Series

In terms of performance, the generalized model and the truncated fourier series method perform as well. However the truncated fourier series model requires more coefficients; we need to use 9 coefficients instead of 5 (case of the generalized model). So the chance of overfitting our model is smaller. As a result, using the proposed extended model is more favorable as it is less complex. In fact, because our correlation coefficients are symmetric around the origin, except for high distances where the correlation is low using cosines to model the correlation coefficient should be sufficient. Furthermore, we remark a similarity between of the values of the coefficients between the truncated fourier series and our model. In fact, a_0, a_1, a_2, a_3, and a_4 in the truncated fourier series model take values close to c, A_1, A_2, $A3$ and $A4$ repectively in the generalized model.

Chapter 5

Conclusion

In this report, we proposed a new closed-form correlation model of oriented bandpass natural images covering distant points in the space for up to a distance of 10 pixels and over four different scales. We modeled in total 632 points in the space, accounting in total for 1264 data points. We were able to validate our model statistically. Furthermore, we compared our model to the truncated fourier series expansion, and showed that both models achieve a similar performance; however, our generalized model requires less coefficients 5 only versus 9 for the truncated fourier series expansion. For this reason the extended model is more favorable. For future work, we plan to extend the generalized model further by studying the impact of the sub-band orientation and modeling the correlation coefficients accordingly.

Bibliography

[1] C.-C. Su, L. Cormack, and A. Bovik, "Closed-form correlation model of oriented bandpass natural images," *Signal Processing Letters, IEEE*, vol. 22, pp. 21–25, Jan 2015.

[2] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600–612, April 2004.

[3] H. Sheikh, M. Sabir, and A. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *Image Processing, IEEE Transactions on*, vol. 15, pp. 3440–3451, Nov 2006.

[4] Z. Wang and A. Bovik, "A universal image quality index," *Signal Processing Letters, IEEE*, vol. 9, pp. 81–84, March 2002.

[5] W. S. G. Jeffrey S. Perry, "Natural scene statistics for image denoising," tech. rep., Center for Perceptual Systems, The University of Texas at Austin, 2013.

[6] J. Burge and W. S. Geisler, "Optimal defocus estimation in individual natural images," *Proceedings of the National Academy of Sciences*, vol. 108, no. 40, pp. 16849–16854, 2011.

[7] A. D. D'Antona, J. S. Perry, and W. S. Geisler, "Humans make efficient use of natural image statistics when performing spatial interpolation," *Journal of vision*, vol. 13, no. 14, p. 11, 2013.

[8] K. Grill-Spector and R. Sayres, "Object recognition insights from advances in fmri methods," *Current Directions in Psychological Science*, vol. 17, no. 2, pp. 73–79, 2008.

[9] J. J. Nassi and E. M. Callaway, "Parallel processing strategies of the primate visual system," *Nature Reviews Neuroscience*, vol. 10, no. 5, pp. 360–372, 2009.

[10] A. W. Roe, S. L. Pallas, J.-O. Hahm, and M. Sur, "A map of visual space induced in primary auditory cortex," *Science*, vol. 250, no. 4982, pp. 818–820, 1990.

[11] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex," *Journal of neurophysiology*, vol. 58, no. 6, pp. 1233–1258, 1987.

[12] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Physical review letters*, vol. 73, no. 6, p. 814, 1994.

[13] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, pp. 188–195, International Society for Optics and Photonics, 1999.

[14] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *Image Processing, IEEE Transactions on*, vol. 10, no. 11, pp. 1647–1658, 2001.

[15] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *Signal Processing, IEEE Transactions on*, vol. 50, no. 11, pp. 2744–2756, 2002.

[16] S. Geman and C. Graffigne, "Markov random field image models and their applications to computer vision," in *Proceedings of the International Congress of Mathematicians*, vol. 1, p. 2, 1986.

[17] J. Portilla and E. Simoncelli, "Texture modeling and synthesis using joint statistics of complex wavelet coefficients," vol. 12, 1999.

[18] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–70, 2000.

[19] D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," *Image Processing, IEEE Transactions on*, vol. 15, no. 6, pp. 1610–1620, 2006.

[20] D. Mumford and B. Gidas, "Stochastic models for generic images," *Quarterly of applied mathematics*, vol. 59, no. 1, pp. 85–112, 2001.

[21] C.-C. Su, L. Cormack, and A. Bovik, "Oriented correlation models of distorted natural images with application to natural stereopair quality

evaluation," *Image Processing, IEEE Transactions on*, vol. 24, pp. 1685–1699, May 2015.

[22] A. C. B. Che-Chun Su, Lawrence K. Cormack, "Bayesian depth estimation from monocular natural images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, January 2015.

[23] U. Rajashekar, Z. Wang, and E. P. Simoncelli, "Perceptual quality assessment of color images using adaptive signal representation," in *IS&T/SPIE Electronic Imaging*, pp. 75271L–75271L, International Society for Optics and Photonics, 2010.

[24] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA A*, vol. 4, no. 12, pp. 2379–2394, 1987.

[25] M. Clark and A. C. Bovik, "Experiments in segmenting texton patterns using localized spatial filters," *Pattern Recognition*, vol. 22, no. 6, pp. 707–717, 1989.

[26] B. A. Olshausen and D. J. Field, "How close are we to understanding v1?," *Neural computation*, vol. 17, no. 8, pp. 1665–1699, 2005.

[27] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Image Processing, International Conference on*, vol. 3, pp. 3444–3444, IEEE Computer Society, 1995.

[28] M. Carandini, D. J. Heeger, and J. A. Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *The Journal of Neuroscience*, vol. 17, no. 21, pp. 8621–8644, 1997.

[29] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *Image Processing, IEEE Transactions on*, vol. 20, no. 12, pp. 3350–3364, 2011.

[30] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *Image Processing, IEEE Transactions on*, vol. 21, no. 12, pp. 4695–4708, 2012.

[31] F. Pascal, L. Bombrun, J.-Y. Tourneret, and Y. Berthoumieu, "Parameter estimation for multivariate generalized gaussian distributions," *Signal Processing, IEEE Transactions on*, vol. 61, no. 23, pp. 5960–5971, 2013.

[32] J. R. Van Aken, "An efficient ellipse-drawing algorithm," *Computer Graphics and Applications, IEEE*, vol. 4, no. 9, pp. 24–35, 1984.

[33] P. J. Hancock, R. J. Baddeley, and L. S. Smith, "The principal components of natural images," *Network: computation in neural systems*, vol. 3, no. 1, pp. 61–70, 1992.

[34] S. A. Tretter, *Notes on Fourier Series*. University of Maryland.

[35] L. Tan and J. Jiang, *Digital signal processing: fundamentals and applications*. Academic Press, 2013.